2022

Sensory and Neuromodulatory Entrainment of Foraging States in C. Elegans

Elias Scheer

SENSORY AND NEUROMODULATORY ENTRAINMENT OF FORAGING STATES IN *C. ELEGANS*

A Thesis Presented to the Faculty of
The Rockefeller University
in Partial Fulfillment of the Requirements for
the degree of Doctor of Philosophy

by
Elias Scheer
June 2022

SENSORY AND NEUROMODULATORY ENTRAINMENT OF FORAGING STATES IN *C. ELEGANS*

Elias Scheer, Ph.D.
The Rockefeller University 2022

Foraging animals dynamically adjust their behavioral patterns over time to optimize feeding in their environment. In the nematode *Caenorhabditis. elegans,* foraging animals alternate between states of locomotory arousal. I observed that animals in low activity states remain nearby food patches, whereas highly aroused animals often exit patches to explore the external environment. Using different methods of behavioral state classification reveals different aspects of foraging behavior: previously described high arousal roaming and low arousal dwelling states reflect long-term changes in locomotory patterns, whereas an Autoregressive Hidden Markov model that classifies behavior on shorter time scales effectively segments food leaving behaviors.

Adaptive foraging strategies rely on animals' ability to integrate internally sensed physiological needs and external sensory cues with ongoing internal state. Consistent with this, I find that animals acutely unable to eat food drastically increase high arousal behavioral states and food leaving behavior, motivated by internal sensing of the absence of food. I also find that mutants lacking sensory transduction or neuromodulatory signaling each show dysregulated behavioral states surrounding food leaving behavior.

In *C. elegans,* animals spontaneously alternate between roaming and dwelling states in environments that contain a uniform food density. However, I find that animals encountering environments with non-uniform food distributions entrain their behavioral state transitions to locally experienced changes in food density. Experience of dwindling food resources over several minutes triggers increased locomotion speed and stimulates roaming behavior, whereas entry into denser food patches suppresses speed and causes animals to dwell. In mutants with deficient chemosensory transduction, this sensory entrainment is lost. I also find that animals lacking a functional PDF receptor (PDFR-1) show both reduced roaming and attenuated sensory coupling of locomotory behavior. Restoring *pdfr-1* expression in interneurons stimulates roaming but fails to rescue sensory coupling, suggesting that other sites of *pdfr-1* expression may drive this.

I conclude that internal states integrate sensory stimuli and physiological needs to facilitate effective foraging strategies on both short and long time scales.

For Kiwi, who helps me find the light.

ACKNOWLEDGMENTS

There are many people I have to thank for helping me to finish this work.

First, I have the deepest gratitude to my advisor and mentor Cori Bargmann for supporting this work by creating a lab environment of openness and curiosity, making it fun to do science. Thank you for showing me how to connect my lofty ideas about neuroscience to the concrete tasks of careful data collection, observation, and analysis. But more than anything, thank you for being a kind, steadying hand on this work through a turbulent time in the world that coincided with my time in graduate school.

I also have to thank the many wonderful lab mates and colleagues I have worked with during my time in graduate school. In particular, I am thankful for my close friendship with former graduate student Alejandro López-Cruz, which has been very influential in my personal growth and growth as a scientist. I would also like to thank former lab member Sagi Levy for his generous help with setting up behavioral analysis code and current lab member Phil Kidd for numerous discussions about quantitative methods for behavioral analysis. Thanks also to Aylesse Sordillo for always having my back in lab when I have a favor to ask or need advice (also for just being dope!). Nothing in the Bargmann lab is possible without the support from Priscilla, Hernán, and Manoush, and they have my deepest gratitude for running the lab so well. Also, special shout out to Manoush for plying me with delicious Za'atar bread.

I would also like to thank my committee members Vanessa Ruta and Winrich Freiwald for providing crucial support to me when the path forward for me was unclear. Your critiques and encouragement have prodded me on to do better science.

A big thank you to everyone in the Dean's Office for making the process of graduate school so streamlined and easy to navigate. Special thanks to Dean Strickland for helping me strategize how to finish this PhD and plan for what's next. And thank you to Cris for always being so kind and helpful in planning the Thesis Defense!

For all of my friends inside and outside of Rockefeller, thank you for being there for me during this long time in graduate school, and a special shout out to the "Cool nerds", an iconic group chat of dear friends who have cheered me to take this thesis over the finish line.

Last but far from least, thank you to my family who show me the unconditional love and encouragement to pursue my work to the highest and best it can be. I love you all.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 2

CHAPTER 3

CHAPTER 4

CHAPTER 5

LIST OF TABLES

CHAPTER 1: INTRODUCTION

In the wild, most animals live in environments where food is unevenly distributed. This creates challenges: most simply, how to find a patch of food, and once found, how long to stay there before moving on to search for others. Foraging behavior is an expression of an animal's response to these questions. Optimal foraging theory predicts that the ideal time for an animal to leave a food patch in search of others occurs when its marginal rate of food intake matches its average rate of food intake across the environment (Charnov, 1976). There has been considerable debate concerning whether animals forage optimally, but in any case, effective foraging requires animal locomotion patterns to be coherently organized through time toward the goal of finding and exploiting food resources.

Foraging behavior can be subdivided into behavioral states that persist on time scales from seconds to minutes or even days to months. In the field of ethology, behavioral states can be conceptualized using an ethogram, which segments behavior into discrete sequential modules that serve to accomplish adaptive goals (Tinbergen, 1951). Ultimately, behavioral states are emergent phenomena of nervous systems, which process sensory information and internal states and enact motor responses. Understanding the genetic and neuronal mechanisms that generate internal states and interpret physiological and external information to direct behavior is an important goal of neuroscience.

Behavioral states have been used to describe foraging behavior in diverse animals and environments. GPS measurements of wild elk in Ontario, Canada revealed movement patterns that can be described using a high activity "exploratory" state and a low activity "encamped" state; the latter is used more often in open habitats (Morales et al., 2004). Similarly, the movements of Red-cockaded woodpeckers in Florida, USA can be divided into high activity "foraging" and low activity "resting" states, whose frequency and duration depend on group size and habitat quality across seasonal changes (McKellar et al., 2015). Given the widespread use of behavioral states across animals, could there be common genetic or neural mechanisms that generate and regulate these states? By studying the foraging patterns of a simple organism in well-controlled laboratory environments, can we discover organizing principles that animate behavioral states? In my thesis work, I have used the nematode *Caenorhabditis elegans* to improve the description of behavioral states used during foraging while deepening understanding of the molecular and neural mechanisms that govern these states.

Approaches to classifying behavioral states across organisms

Classical ethologists manually segmented observations of animal behavior into states using subjective criteria (Baerends, 1976; Tinbergen, 1951). While this method yielded insights that have been highly influential, it is now possible to automatically define behavioral states based on estimates of animal positions through time in video data.

However, the process of deriving a set of discrete behavioral states from continuous recordings requires thought. First, researchers must decide which metrics of animal behavior they wish to measure for later use in generating behavioral state classifications. For many low-level descriptions of animal behavior, it is sufficient to measure an animal's centroid position through time. Centroid positions can be used to derive speed and turning rate, which are coarse but highly informative descriptors of locomotory behavior. Most animals have anterior/posterior axis, which can be used to define the direction an animal is moving relative to its own coordinate frame: forward, backward, or sideways. Finer analysis of animal movements can involve pose estimation of specific body parts or the entire body posture through time, and these will vary based on the biomechanical properties of the animal. For example, in *C. elegans*, body posture can be summarized by fitting a smooth sinusoidal spline to the body shape, generating measurements of a set of points that undulate with animal movement (Husson et al., 2012). In animals with rigid body structures, independent appendages must be tracked separately to generate pose estimates. In recent years, machine learning pose estimation techniques like DeepLabCut and LEAP have been used to track a set of user-defined key points on an animal's body through time, facilitating this process (Mathis et al., 2018; Pereira et al., 2019).

Once body part positions have been tracked, these raw data are transformed into usable dimensions for behavioral state classification. Practically, the temporal resolution of behavior can vary, so the relevant time scale on which to analyze behavior must be chosen carefully. For example, discrete movements in mice can be studied on the order of milliseconds, or over diurnal sleep/wake cycles, yielding different behavioral state classification schemes (Adamah-Biassi et al., 2013; Wiltschko et al., 2015). After coarse-graining data to the time scale of choice, dimensionality reduction techniques like principal components analysis (PCA) can be used to reduce the number of features. In *C. elegans*, PCA allows the full set of tracked points along the body spline (typically ~40-100 points) to be efficiently represented by just four numbers that are weights on a new basis set of "eigenworm" postures, explaining 95% of the variance in the original posture space (Stephens et al., 2008). For behaviors that have a cyclic nature, such as walking in fruit flies or web weaving in spiders, movements of tracked appendages can be decomposed into the frequency domain, which permits temporal clustering of similar actions (Berman et al., 2014; Corver et al., 2021).

After the desired time scales and feature representations have been defined, there are several choices of statistical methods for classifying behavioral states. One approach that has been used in flies and other organisms involves manual annotation of a training set of video frames containing a behavior of interest and then allowing a supervised machine learning algorithm to identify other video frames that share that pattern (Kabra et al., 2013). Alternatively, behavioral states can be identified using unsupervised machine learning approaches. Among these, Hidden Markov Models (HMMs) are especially prevalent and powerful.

Hidden Markov Models in neuroscience

Hidden Markov Models assert that the values of measured time-series feature dimensions are probabilistically generated from hidden states according to "emission probabilities" (Baum and Petrie, 1966; Rabiner, 1989). On each successive time step *t*, a hidden state is chosen according to a "transition probability" distribution conditioned on the state used at *t-1.* Due to the inherent noisiness of time-varying data collected across neuroscience, HMMs have been widely adopted to infer the presence of latent processes that generate the noisy observations.

One useful application of HMMs is to infer the presence of latent states that generate neural activity. The collective activity of neurons in rat gustatory cortex shows diverse and noisy responses to presentation of tastant stimuli when data is aligned to the onset of trials. However, when activity is aligned to the onset of HMM-decoded hidden states across the ensemble neural activity, responses become coherent and reliable (Jones et al., 2007). This shows that sensory responses of cortical neurons encode systems-level dynamic processes present across distributed neuronal ensembles. On a smaller scale, HMMs can also be used to infer whether an ion channel is in an open or closed state from noisy cell-attached voltage recordings (Becker et al., 1994; Chung et al., 1990).

In the context of behavioral neuroscience, the hidden states can be thought of as the behavioral states we seek to measure. Training HMMs with different numbers of hidden states generates models that can decode the same behavioral time series in different ways. By definition, models with more hidden states tend to split behavioral sequences into shorter-duration states, whereas models with fewer states collapse more data into longer states. For example, *C. elegans* behavior on bacterial food has been classified into two states (discussed more below, Chapter 4) or up to nine states depending on the HMM used (Ben Arous et al., 2009; Cermak et al., 2020; Flavell et al., 2013; Fujiwara et al., 2002). Comparing the occupancy of states from the nine-state model against the two-state model reveals that eight of the nine states in the former model are shorter sub-states of one the states in the two-state model.

The basic structure of the HMM can be elaborated to admit different parametrizations of transition and emission probability distributions (Linderman, 2016). One such HMM, in which observation emissions are modeled according to autoregressive processes (AR-HMM), has been used to classify mouse behavior "syllables" (Wiltschko et al., 2015). This model posits a hierarchy of behavioral organization whereby discrete hidden states on longer time scales govern sets of linear dynamics that generate continuous observations of movement features on shorter timescales. Empirically, AR-HMMs produce more likely state predictions on data not used during the model training procedure than HMMs without autoregressive time dynamics, indicating that state descriptions that encapsulate both discrete and continuous dynamics may generalize better across conditions.

Internal states probabilistically drive behaviors

Behavioral states are thought to be the outward manifestation of persistent changes in the internal state of the brain's neural activity. Internal states, in turn, represent the evolving goals of animal behavior motivated by physiological needs and instinctual drives (Anderson and Adolphs, 2014).

Discrete motor actions (or "decisions") may sit in a nested hierarchy of internal states and behavioral states. The courtship behavior of male *Drosophila* illustrates this. Males engaged in courtship have a persistent internal state that drives a suite of behaviors toward a female (Clowney et al., 2015; Inagaki et al., 2014). These behaviors can be segmented into different behavioral states based on both the male's own movements and his spatial relationship to the female (Calhoun et al., 2019). These behavioral states, in turn, change the male's probability of producing two discrete types of courtship songs with his wing, sine and pulse song. When the male is further from the female, he is more likely in the "Chase" state, chasing her at high speed while favoring production of pulse song. At close range, he more often in the "Close" state, in which he moves slowly and favors production of sine song. Neither the internal state nor the behavioral state deterministically decides the type of song type performed; rather it biases the decision in favor of one song type or the other probabilistically.

Internal hunger state drives foraging behavior

Foraging behaviors are regulated by a critical internal state: hunger. The physiological need for nutrients creates the internal state of hunger, which drives behavioral states that comprise foraging behavior. In mammals, the hormone ghrelin is released from the enteroendocrine cells of the gut when nutrients are low to signal a hunger state (Ariyasu et al., 2001). Ghrelin activates the AgRP neurons in the arcuate nucleus of the hypothalamus (Cowley et al., 2003), which release neuropeptide Y that drives feeding and food-seeking behavior (Clark et al., 1984; Ollmann et al., 1997). In the presence of food, animals with activated AgRP neurons eat voraciously (Aponte et al., 2011). However, if food is absent, AgRP activity stimulates exploratory food-seeking behaviors reminiscent of foraging in natural environments (Dietrich et al., 2015). Once food has been found but before feeding commences, AgRP neuronal activity decreases, indicating that the internal state of food-seeking is distinct from feeding itself (Chen et al., 2015).

Internal hunger state gates sensorimotor transformations that guide foraging

One way that internal hunger states affect behavior is by potentiating sensory responses to food-associated cues. For example, in *Drosophila*, starvation causes release of short neuropeptide F (sNPF), which induces presynaptic facilitation of olfactory neurons, thereby increasing odor-evoked neurotransmitter release to drive food search behaviors (Root et al., 2011). Once the flies are sated, insulin-like peptides suppress transcription of the sNPF receptor, thereby decreasing food search behavior (Root et al., 2011). Starvation also enhances the proboscis extension response to sucrose (Meunier et al., 2007; Scheiner et al., 2004). Increased dopamine release on to primary gustatory neurons enhances this appetitive behavioral response (Inagaki et al., 2012). Together, these results show that starvation-induced neuromodulatory signaling enhances multiple sensory modalities to drive food-seeking behaviors.

Learned association of sensory stimuli with water and food produces flexible behavioral responses to otherwise neutral cues. For example, mice that are trained to associate an odor cue with water delivery lick vigorously at a water spout following odor presentation when thirsty but otherwise do not. The internal state of thirst affects firing patterns of neurons broadly distributed throughout the brain, galvanizing animals to produce adaptive behaviors that sate them, like drinking in response to a paired cue (Allen et al., 2019).

The molecular and circuit mechanisms that allow internal states to persist through time while integrating sensory inputs and orchestrating behavioral outputs are topics of active investigation.

Molecular, cellular, and circuit mechanisms that generate circuit states

Internal states are generated by patterns of activity in neural circuits that persist through time. These patterns of activity, in turn, emerge from the connections among cells in the circuit and their constituent electrophysiological properties. Studies in the stomatogastric ganglion (STG) of the crab *Cancer borealis* illustrate these mechanisms elegantly. The STG contains neurons whose electrical activity partakes in two rhythmic motor patterns: the pyloric rhythm and the gastric mill rhythm (other rhythms exist but are excluded for simplicity). The pyloric rhythm is an oscillating triphasic signal driven by a set of coupled electrical pacemakers, whereas the gastric mill rhythm is episodically active and depends on input from sensory neurons (Nusbaum and Beenhakker, 2002). Neurons in the STG are densely interconnected by inhibitory chemical synapses and electrical synapses that couple activity patterns between cells, allowing individual neurons to dynamically synchronize their activity to either one rhythm or the other (Marder and Bucher, 2007; Nusbaum and Beenhakker, 2002). Neurons in the STG also exhibit varying electrophysiological properties—some are tonically spiking, others fire bursts of spikes followed by periods of low activity. Both the intrinsic biophysical properties and the weighting of synaptic connections between cells are under

neuromodulatory control. Neuromodulators, including both the biogenic amines like dopamine and serotonin as well as neuropeptides, can dynamically drive stable participation in one rhythm or the other over time (Marder, 2012).

C. elegans as a model for studying the neuromodulatory and sensory control of behavioral states in foraging

C. elegans is a powerful model system for interrogating mechanisms underlying behavioral states in foraging. It has a compact nervous system comprising 302 neurons whose synaptic connections have all been mapped (Cook et al., 2019; White et al., 1986). It has long been studied as a genetic model system and in addition, neuronal cell types can be precisely manipulated by genetic tools, permitting causal experimentation of neuronal function. Furthermore, its optical transparency permits direct measurements of neural activity using fluorescent indicators. Finally, *C. elegans*' inexpensive culture and fast generation time facilitate experiments on large numbers of animals to generate robust statistics.

C. elegans behavioral states are defined based on characterization of a limited repertoire of locomotory maneuvers including forward runs, reversals, pauses, turns, fine scale head movements, and changes in posture (Schwarz et al., 2015; Von Stetina et al., 2006; Steuer Costa et al., 2019). Forward and reverse movements are produced by alternating contractions of body wall muscles that generate sinusoidal waves that propagate along the body. During forward locomotion, the wave begins at head and propagates backward, whereas reversals initiate waves from the tail that propagate forward; both forward and reversal speeds vary continuously from 0-0.3 mm/sec. During foraging, movements are modified according to the presence or absence of food. Animals off food move at higher speeds on average than animals on food, and slow down when encountering a bacterial lawn (Sawin et al., 2000). Reversals also vary in their speed and duration according to the presence of food: on food, *C. elegans* exhibits mostly short reversals, lasting 2-3 seconds, whereas off food reversals of 6 seconds or longer are also observed (Chalasani et al., 2007; Gray et al., 2005; Sordillo and Bargmann, 2021).

On food, *C. elegans* adopts two major behavioral states termed roaming and dwelling. Roaming is characterized by high speed and low angular speed, whereas dwelling is defined by low speed and high angular speed (see Chapter 4) (Ben Arous et al., 2009; Flavell et al., 2013; Fujiwara et al., 2002). The fraction of time an animal spends roaming is regulated by its satiety state and sensory stimuli: animals roam less when food quality is low or aversive chemicals are present in bacterial food (Ben Arous et al., 2009; Shtonda, 2006), but the bistability of roaming and dwelling states is governed by neuromodulatory signaling: the neuropeptide Pigment Dispersing Factor (PDF-1) promotes roaming by signaling through its cognate receptor PDFR-1, while the biogenic amine serotonin promotes dwelling by activating the serotonin-gated chloride channel MOD-1 (Flavell et al., 2013).

Upon removal from food, *C. elegans* exhibits a different pair of behavioral states called local search and global search. Immediately after removal from food, animals begin local search, during which they maintain a high reversal rate that contains their locomotion to a small area, as if seeking food they just left behind. After approximately 15 minutes, animals transition into the global search state by suppressing their reversal rates, sending them on long forward runs that disperse towards new food sources (Calhoun et al., 2015; Gray et al., 2005; López-Cruz et al., 2019). Prolonged glutamate release from mechanosensory and chemosensory neurons maintains the local search state by activating the inhibitory G-protein coupled metabotropic glutamate receptor MGL-1 (López-Cruz et al., 2019). MGL-1 inhibits interneurons that normally promote forward movement, thereby suppressing forward movements in favor of reversals. Over time off food, glutamatergic release lessens, disinhibiting the forward-promoting interneurons that initiate global search.

Across both roaming/dwelling and local search/global search states, an animal's sensory and metabolic history influences its behavioral state usage. Animals that are starved for one hour show half as much roaming when re-introduced to food than animals with continuous access to food (Ben Arous et al., 2009). There is evidence that animals form a memory of both the abundance and variability of food distributed in their environment over time that guides subsequent local search states. Animals removed from lawns of higher food density do more vigorous local search than animals removed from less dense lawns upon removal from food (López-Cruz et al., 2019). Furthermore, animals removed from lawns where food density is more spatially variable show reduced local search compared to animals removed from uniformly dense food lawns (Calhoun et al., 2015). Analysis of single animals' local search behavior following exposure to varying food environments reveals that sensory experience up to thirty minutes in the past affects behavior in the present. In addition, early life food experiences affect behavioral state usage in adulthood. Animals that are starved during their first larval stage enter an alternative life stage of developmental arrest called dauer (Golden and Riddle, 1982, 1984). Re-exposure to food induces dauers to resume development into adults. Although basic features of post-dauer adult locomotion are similar to that of animals who were never starved, post-dauers persistently roam less and do fewer exploratory bouts away from food (Pradhan et al., 2019). These results demonstrate long-lasting effects of early-life stress mediated through developmental plasticity on adult behavioral states.

Just as behavioral states result from the differential usage of a common set of locomotory maneuvers, the neurons and neuromodulators often contribute to multiple behavioral states. Some neurons produce the same locomotory behavior regardless of which behavioral state an animal is using. For example, optogenetically activating the RIM interneurons promotes reversals across both local search and global search states (López-Cruz et al., 2019). Other neurons promote alternate behavioral states depending on the current state of the animal. Optogenetic activation of the AIA interneurons,

whose activity integrates sensory signals from several sensory neurons, causes roaming animals to dwell and dwelling animals to roam (Dobosiewicz et al., 2019; Ji et al., 2021). While some of the mechanisms that produce behavioral states and link them to sensory experience have been dissected, a holistic understanding of the neural organization of behavioral states remains to be uncovered.

Thesis Overview

Spontaneous internal states have been examined in uniform environments. Acute sensory responses have been examined across different internal states. How does a more heterogeneous natural environment shape the properties of internal states and their effects on behavior? The goal of this thesis is to investigate the genetic and neuronal mechanisms that generate behavioral states and link them to changing sensory experience during foraging.

In Chapter 2, I delineate new methods for classifying *C. elegans* behavioral states on small lawns of bacteria, where food density is not uniformly distributed. I compare my newly derived behavioral states with those from previously described work. I investigate how behavioral states that persist on various time scales affect the prevalence of momentary locomotory patterns used by animals at the borders of bacterial lawns. These discrete maneuvers include lawn exits, whereby animals completely leave bacterial food to explore their environment, and head pokes, in which an animal briefly pokes its head outside the lawn before retracting it, preventing a lawn exit. I find that lawn exits but not head pokes are enriched in behavioral states that correspond to exceptionally high locomotory arousal.

In Chapter 3, I investigate sensory and neuromodulatory mechanisms that generate behavioral states and lawn leaving behavior. I find that an animal's probability of lawn leaving scales inversely with bacterial food quality and the animal's ability to ingest food. I investigate changes in lawn leaving and behavioral states across a set of strains in which genes or neurons that regulate sensory and neuromodulatory processes are inactivated. A mutant with loss of function in the neuropeptide Y-like gene *npr-1* leads to the largest increase in lawn leaving and roaming behavior, so I study it further. I find that the *npr-1(lf)* high lawn leaving and roaming phenotypes act through synaptic output from RMG neurons, the core of the well-characterized circuit in which NPR-1 signals. They also require PDFR-1 signaling on a set of downstream interneurons including RIB. By chemogenetic silencing, I find that RIB neuronal activity modulates behavioral states and lawn leaving.

In Chapter 4, I compare the behavior of animals on uniform lawns and small lawns of bacteria containing a dense edge region. I find that animals on small lawns roam less than animals on uniform lawns and show that the probability of roaming is inversely related to increasing bacterial density of the lawn. I then analyze the same set of strains used in Chapter 3, in which sensory and neuromodulatory genes and neurons are inactivated, for differences in animals' experienced bacterial density and distance to the lawn boundary. I find that mutants in *tax-4*, a critical component of chemosensory signaling in several sensory neurons, as well as mutants in the *pdfr-1* signaling pathway roam less than wild type and are found consistently further inside bacterial lawns, where bacterial density is lower. My results suggest that *pdfr-1* signaling is required both for the generation of roaming states and their entrainment to bacterial density, and that

these two outcomes require different sites of *pdfr-1* expression. I identify the RIB neurons as an important site of *pdfr-1* expression that generates roam states.

I conclude that *C. elegans* behavioral states are sculpted from an interaction between internal circuit mechanisms that promote locomotory changes and sensory information that regulates when states are deployed. My discussion will focus on possible mechanisms that combine these two streams of neural activity to produce adaptive behavioral states in foraging.

CHAPTER 2: QUANTIFYING FORAGING BEHAVIOR IN HETEROGENEOUS FOOD ENVIRONMENTS

Introduction

Measurement of movement phenotypes in *C. elegans* has been crucial in dissecting the genetic and neural mechanisms that underpin behavior. In particular, quantifying animal locomotion on lawns of bacterial food OP50 has yielded major insights into the structure of *C. elegans* foraging behaviors. Over the years, methods for measuring behavior have become more precise, but simple movement descriptors have yielded major insights into the structure of animal behavior. Simply by counting the number of body bends per time interval, researchers can detect large changes in speed between animals on or off food (Sawin et al., 2000). Similarly, counting the number of squares entered on a grid overlayed on the food lawn illuminated the structure of behavioral states (Flavell et al., 2013). Automated behavioral recording techniques allow researchers to examine behavioral differences across genotypes and conditions in greater detail. Even low spatial resolution video recordings allow reliable discernment of animal speed and direction of movement or timing of discrete events like egg laying (De Bono and Bargmann, 1998; Waggoner et al., 1998).

Higher resolution video recording techniques show the worm in enough pixels to derive the shape and posture of individual animals using computer vision (Geng et al., 2004). This posture data inspires new ways of describing locomotion patterns. Using principal components analysis, most of the variance in *C. elegans* behavior can be parametrized by decomposing posture vectors into a set of weights on eigenmodes (“eigenworms”) that correspond to animal shapes used over time during locomotion (Hums et al., 2016; Stephens et al., 2008). Measured postures can also be compared to a set of reference postures that capture the full range of shapes a worm can assume. Animals on bacterial food use some postures more frequently than off food and also alter the sequence or “syntax” of postures used through time (Schwarz et al., 2015).

In addition to instantaneous movement changes, neuroscientists wish to discern when animal behavior undergoes a longer-lasting change. It is commonly assumed that persistent changes in locomotion patterns correspond to transitions in behavioral state. In order to rigorously demarcate the boundary of one behavioral state and the next, statistical methods are needed. Hidden Markov Models (HMMs) provide one such abstraction for inferring the presence and duration of behavioral states from measured input dimensions (Baum and Petrie, 1966; Rabiner, 1989). A two-state HMM has been used to categorize *C. elegans* behavior on food into either roaming (high speed, low angular speed) or dwelling (low speed, high angular speed) (Flavell et al., 2013). Other researchers have suggested HMMs with different numbers of states. In addition to roaming and dwelling, one study proposed a third very low speed state called quiescence but also suggested that animals may adopt “mixed states” that exist in a continuous feature space (Gallagher et al., 2013). More recently, an HMM parametrized

by posture measurements found nine states, eight of which subdivide the dwelling state into shorter sub-states (Cermak et al., 2020).

While foraging in natural environments, both an animal's behavioral patterns and its sensory experience vary through time. Therefore, to simplify the study of behavioral states, a good first step is to measure behavior in a "featureless" environment. This allows researchers to analyze the self-generated aspects of behavior but omits important relationships between behavior and external sensory cues. In this thesis, I interrogate the effect of uneven food distribution on the generation of behavioral states by studying *C. elegans* behavior on small lawns of bacterial food. Because of the capillary action underlying the "coffee-stain effect", bacteria applied to an agar surface in liquid droplets is deposited most densely at the outer edge of the lawn as it dries (Deegan et al., 1997). In small lawns, the ratio of circumference to surface area is greater, resulting in an overall more variable food environment than in larger lawns (Calhoun et al., 2015). Animals detect these changes in bacterial density through two pairs of chemosensory neurons, ASK and ASI, both of which signal to downstream interneurons.

In this chapter, I describe the methods I used to extract locomotory and sensory features of an animal's environment on small lawns of bacteria using video recording and computer vision techniques. I then discuss a new class of behaviors generated from interactions with the boundary of a food lawn. This leads into the derivation of Hidden Markov Models used to segment behavioral states on small lawns and how these behavioral states correlate with lawn boundary interaction behaviors. I show that lawn leaving behavior emerges particularly from high arousal behavioral states, whereas head poke reversals that return an animal to the lawn are evenly distributed across arousal states. Using a Generalized Linear Model, I demonstrate that midbody and head speed are highly correlated and predictive of lawn leaving events.

Results

Recording *C. elegans* behavior at high spatial and temporal resolution on bacterial food

In order to discern how *C. elegans* navigates its environment, I recorded and identified its locomotory maneuvers with high precision. I adapted and expanded an imaging system to record behavior of individual animals on bacterial lawns at 3 frames per second and spatial resolution of 10μm (Stern et al., 2017). Since adult *C. elegans* are approximately 1mm long and 30-50μm wide, this system shows individual animals in high detail. Each camera is aimed at a custom-made laser-cut multi-well plate such that up to 6 animals can be recorded simultaneously per camera (Figure 2-1 A). Using 12 cameras, up to 72 single animals can be recorded in one experiment lasting 1 hour, totaling 10,800 movie frames per animal. I developed custom tracking codes in MATLAB to automatically detect and track individual animals from background-

subtracted videoframes. In these images, I discern the animal's shape and posture after computationally applying a smooth spline. Using this spline, I unambiguously identified the head, tail and centroid positions in 98% of frames. Head annotation is done by looking for the end of the spline that shows higher overall displacement throughout the video. Using the spline and spatial landmarks, the script then determines when animals are moving forward, reversing, pausing, and additionally defines other behavioral parameters including speed, angular speed, and relative angles between various body parts through time (for a full list of extracted features see Methods).

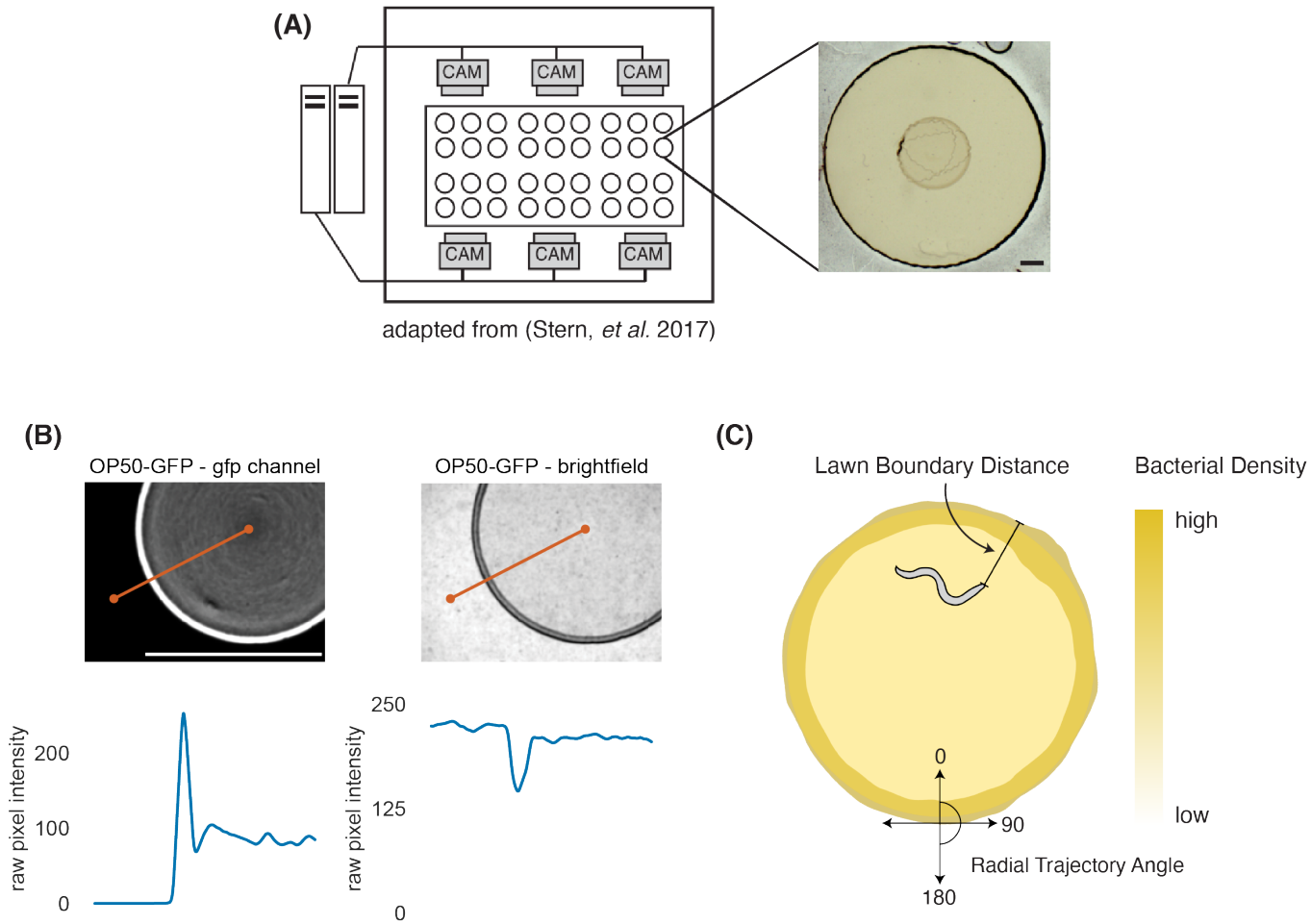
Many previous studies have investigated *C. elegans* behavioral repertoire on its standard bacterial food, *E. coli* OP50. Here, I sought to understand how animals behave at the boundary of bacterial lawns – the border between food and no food. To do this, I grew small lawns of OP50 (diameter = ~4 mm) to a defined density in single wells of multi-well agar assay plates (see Methods). To evaluate the food abundance distribution in such lawns, I grew sample lawns using a strain of OP50 that expresses GFP (OP50-GFP) and measured the lawn fluorescence as a proxy for bacterial abundance using an epifluorescence microscope (Labrousse et al., 2000). Small lawns of OP50-GFP are more than twice as bright at the edge than the center, indicating at least a twofold increase in lawn density in the border region (Figure 2-1 B). This thick border is 100-300μm wide, thus representing about 16% of the area of the bacterial lawn traversed by the animal. Higher food density at the border is also measurable as darker pixels in the brightfield image. When tracking animal movements in assay videos, I estimate the relative local bacterial density an animal experiences as it navigates a food lawn by measuring the grayscale pixel values of the background image at the coordinates of the nose tip (see Methods). This parameter is called "Bacterial Density". Additionally, the tracking code measures the "Lawn Boundary Distance" as the distance from the animal's nose tip to the closest point on the lawn boundary in every frame. By convention, Lawn Boundary Distance is defined as a positive number inside the lawn, 0 at the border, and negative outside. The "Radial Trajectory Angle" is the angle of an animal's trajectory relative to the radius of the lawn – 0 means pointing toward the lawn center, 180 means pointing away, and 90 follows the tangent trajectory along the border.

In this chapter, I analyze data from a cohort (n = 978) of wild type (PD1074) animals on small bacterial lawns that have passed a variety of quality control tests (see Methods).

Across this dataset, the distribution of animals' measured lawn boundary distance and bacterial density reveals that animals spend 60% of time in the narrow border region, where bacterial density is the highest. Since the lawn border accounts for 16% of the lawn area, this amounts to a ~4-fold enrichment in animal residence in the border region (Table 2-1). The distribution of Radial Trajectory Angle (RTA) peaks at 90 degrees, representing animals that circumnavigate the boundary.

Figure 2-1. Quantifying locomotory and environmental variables on small lawns of bacterial food.
(A) Multi-camera imaging system for tracking locomotion of single animals at high spatiotemporal resolution. Inset shows a single animal on a small lawn of *E. coli* OP50 grown on NGM agar inside a laser-cut well, scale bar in lower right is 1 mm.
(B) A small bacterial lawn expressing green fluorescent protein (GFP) shows higher concentrations on the edge compared to the center. Top row: left, an image of GFP intensity across the lawn. Right, an image of the same lawn in brightfield. Bottom row: pixel intensity profile is plotted across the orange transect line plotted in each of the above images. White scale bar in GFP image is 1 mm.
(C) Schematic illustrating Lawn Boundary Distance, Bacterial Density, and Radial Trajectory Angle (see Methods).
(D) Empirical distribution of Lawn Boundary Distance.
(E) Scatter plot of Bacterial Density versus Lawn Boundary Distance for a random sample of 10,000 points.
(F) Empirical distribution of Radial Trajectory Angle.
n = 978 wild type animals on small lawns.

Figure 2-1


(A)
CAM
CAM
CAM
CAM
CAM
CAM
adapted from (Stern, *et al.* 2017)
(B)
OP50-GFP - gfp channel
OP50-GFP - brightfield
raw pixel intensity
200
100
0
250
125
0
(C)
Lawn Boundary Distance
Bacterial Density
high
low
0
90
180
Radial Trajectory Angle
(D)
normalized frequency
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
−0.5
0.0
0.5
1.0
1.5
Lawn Boundary Distance (mm)
(E)
Bacterial Density (0-1)
1.0
0.8
0.6
0.4
0.2
0.0
−0.5
0.0
0.5
1.0
1.5
Lawn Boundary Distance (mm)
(F)
normalized frequency
0.16
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
0
50
100
150
Radial Trajectory Angle (deg)

Classifying lawn boundary interaction behaviors

Freely moving animals showed several kinds of behavioral interactions with the lawn boundary. When an animal moved its head onto the bacteria-free agar, I designated this behavior a “head poke”. A head poke followed by a reversal that returned the head to the lawn interior was called a “head poke reversal” (Figure 2-2 A). Alternately, a head poke followed by curved forward movement that returned the entire animal was called a “head poke forward” (Figure 2-2 B). When an animal’s movement ceased for following a head poke before resuming forward motion, it was called a “head poke pause” (Figure 2-2 C). Finally, when animals leave the lawn entirely to explore the agar surface, it was called a “lawn exit” (Figure 2-2 D).

To see how these types of lawn boundary interaction coincide with locomotion metrics like speed and angular speed as well as lawn boundary distance and experienced bacterial density, I plotted these metrics together for an individual animal across the entire 40-minute assay (Figure 2-3).

Head poke types

Different interactions with the lawn boundary occur with different frequency. 73% of head pokes are head poke reversals, which occur approximately once per minute (Figure 2-4). Head poke forward and head poke reversals account for roughly 14% and 13% of head pokes, respectively and occur once every ~5 minutes. Lawn exits are the rarest behavior at the lawn boundary, only occurring only in 30% of animals with a rate of one lawn exit every ~90 minutes (see Table 2-1 for all lawn boundary interaction statistics).

The initiation of a head poke is associated with an increase in midbody speed and radial trajectory angle detectable ~3 seconds before the measured event (Figure 2-4 C, D). The resolution of the head poke into a reversal, pause, or continued forward movement was completed within 5 seconds after its initiation. Notably, the resolution of the head poke was predictable at least one minute before its initiation: the average midbody speed was highest for animals that performed the head poke forward maneuver, followed by head poke reversal, and lowest for animals in which head pokes were followed by a pause. This difference in speed persisted for at least one minute after the head poke, suggesting that the behavioral resolution of a head poke is influenced by a sustained behavioral state reflected in locomotion speed.

The resolution of head pokes was also predictable at least one minute before its initiation based on environmental parameters (Figure 2-4 E, F). The head poke pause behavior was preceded by a sustained interval at higher bacterial density, closer to the lawn boundary, than head poke behaviors that were resolved in reversals or continued forward movement. Like locomotion speed, these variations persisted for at least one minute after the head poke.

At a more subtle level, head poke reversals were characterized by the largest change in radial trajectory (Figure 2-4 D) and in the distance that the head covered outside the lawn (Figure 2-4 E) compared to other events.

In summary, head pokes represent behaviors whose resolution is influenced by both ongoing behavioral state of the animal, represented by speed, and by environment, represented by bacterial density and lawn boundary distance.

Lawn Exits

In a lawn exit event, the entire body of the animal translocates outside the lawn before its return. Analyzing lawn exit events using the parameters described above for head pokes revealed similarities and differences in the preceding behavior (Figure 2-4). Before lawn exits, midbody speed and lawn boundary distance are greater than the corresponding metrics are before all types of head pokes. Similarly, animals are found at persistently lower bacterial density before lawn exits than before head pokes. These distinctions identify lawn exits as a distinct behavioral decision and will be considered further below. The exploratory bouts off food following lawn exit varied in length from several seconds to several minutes, when animals re-entered the lawn ("lawn entry") (Figure 2-5). In the lower quartile of bout duration, animals typically executed a large reversal and reorientation, returning to the lawn within 30 seconds of lawn exit. In the upper quartile of bout duration, animals did not reverse immediately after lawn exit, and lingered outside the lawn for several minutes before returning.

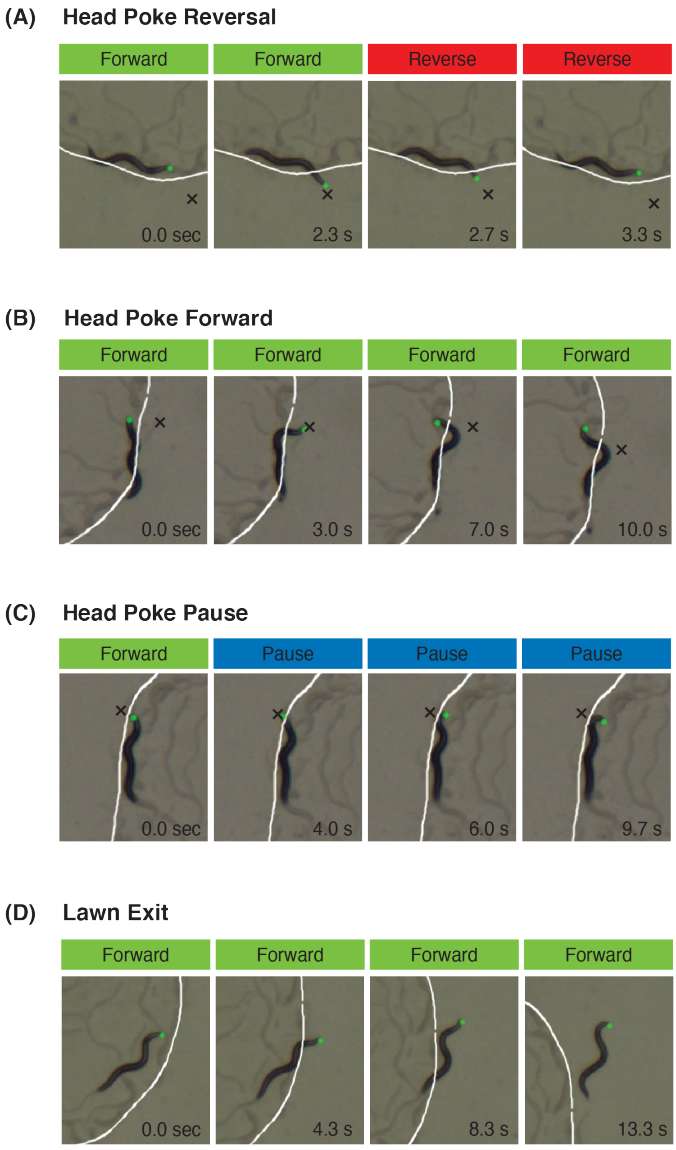


Figure 2-2. Classifying types of lawn boundary interactions.
(A) Four example images from a behavioral sequence generating a head poke reversal. First animal pokes head outside lawn while moving forward, then reverses to retract its head on to bacterial lawn.
(B) Same as (A) for head poke forward. The animal's head pokes outside the lawn during forward movement, followed by a head swing that continues forward movement inside the lawn.
(C) Same as (A) for head poke pause. The animal's head pokes outside and retracts before pausing.
(D) Same as (A) for lawn exit. The animal approaches and fully crosses through the lawn boundary to explore the bacteria-free agar outside of the lawn.
White lines in each image indicate the lawn boundary. Black x in (A-C) depicts the position of maximum head displacement outside the lawn during head poke.

Figure 2-3

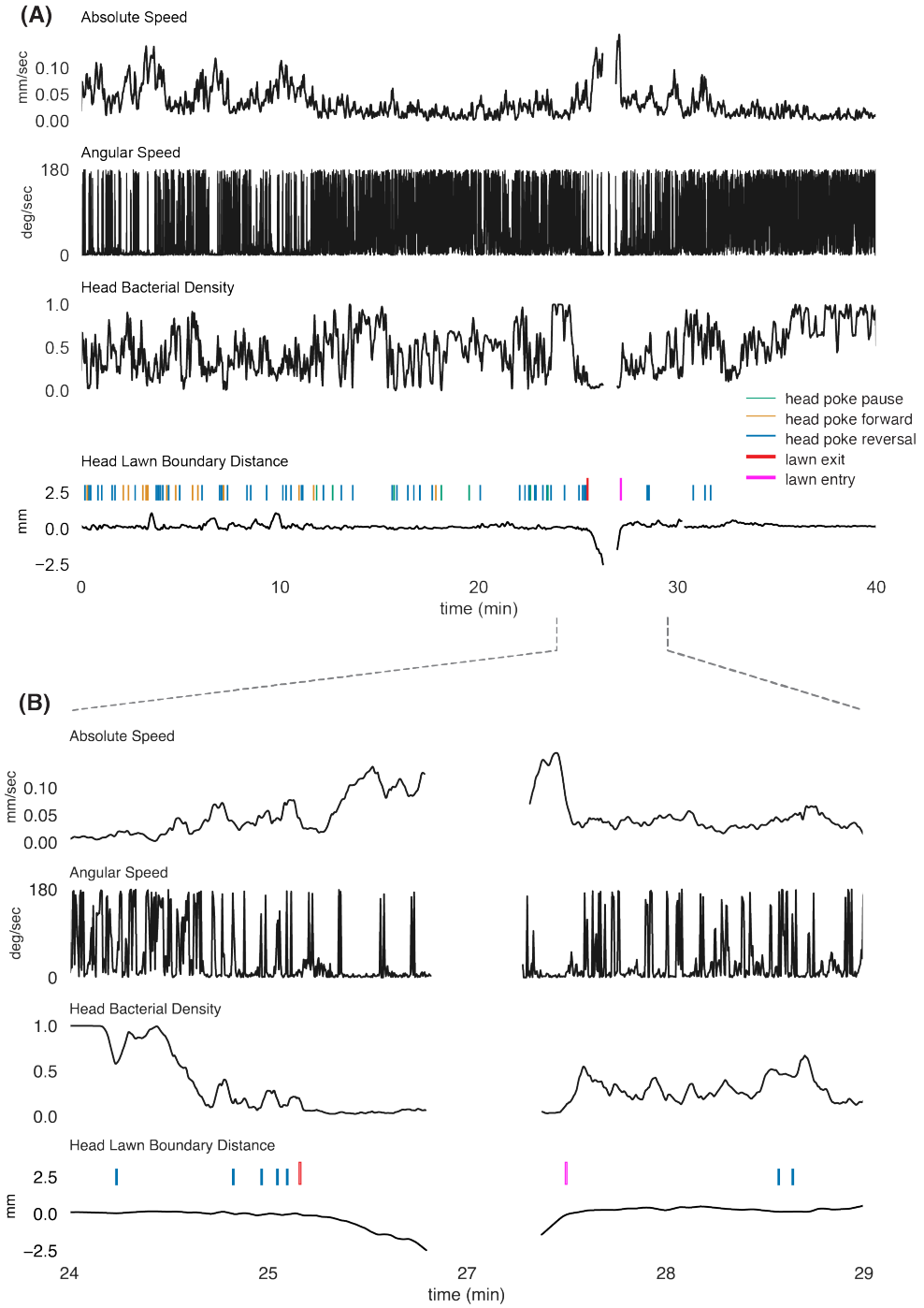


Figure 2-3. Example behavioral trace with annotated locomotion and environmental variables.
(A) Absolute speed, Angular speed, Bacterial Density, and Lawn Boundary Distance for an entire behavioral trajectory on a small lawn for a single animal. Lawn boundary interactions are annotated colored tick marks above the Lawn Boundary Distance trace. **(B)** A 5-minute chunk of the behavioral trajectory displayed in (A) centered on a lawn exit.

Figure 2-4. Head poke types correlate with different patterns in locomotory and environmental variables.
(A) Empirical distributions of number of lawn boundary interaction events per animal per 40-mintue assay.
(B) Mean frequency of different types of lawn boundary interactions in events per minute. Each dot represents one animal. Boxes indicate median and interquartile range.
(C) Midbody speed aligned to different types of lawn boundary interactions. Dark line represents the mean and shaded region represents the standard error in (C-F)
(D) Same as (C) for Radial Trajectory Angle.
(E) Same as (C) for Lawn Boundary Distance.
(F) Same as (C) for Bacterial Density.
See Table 2-1 for mean frequencies of all lawn boundary interaction behaviors.

Figure 2-4

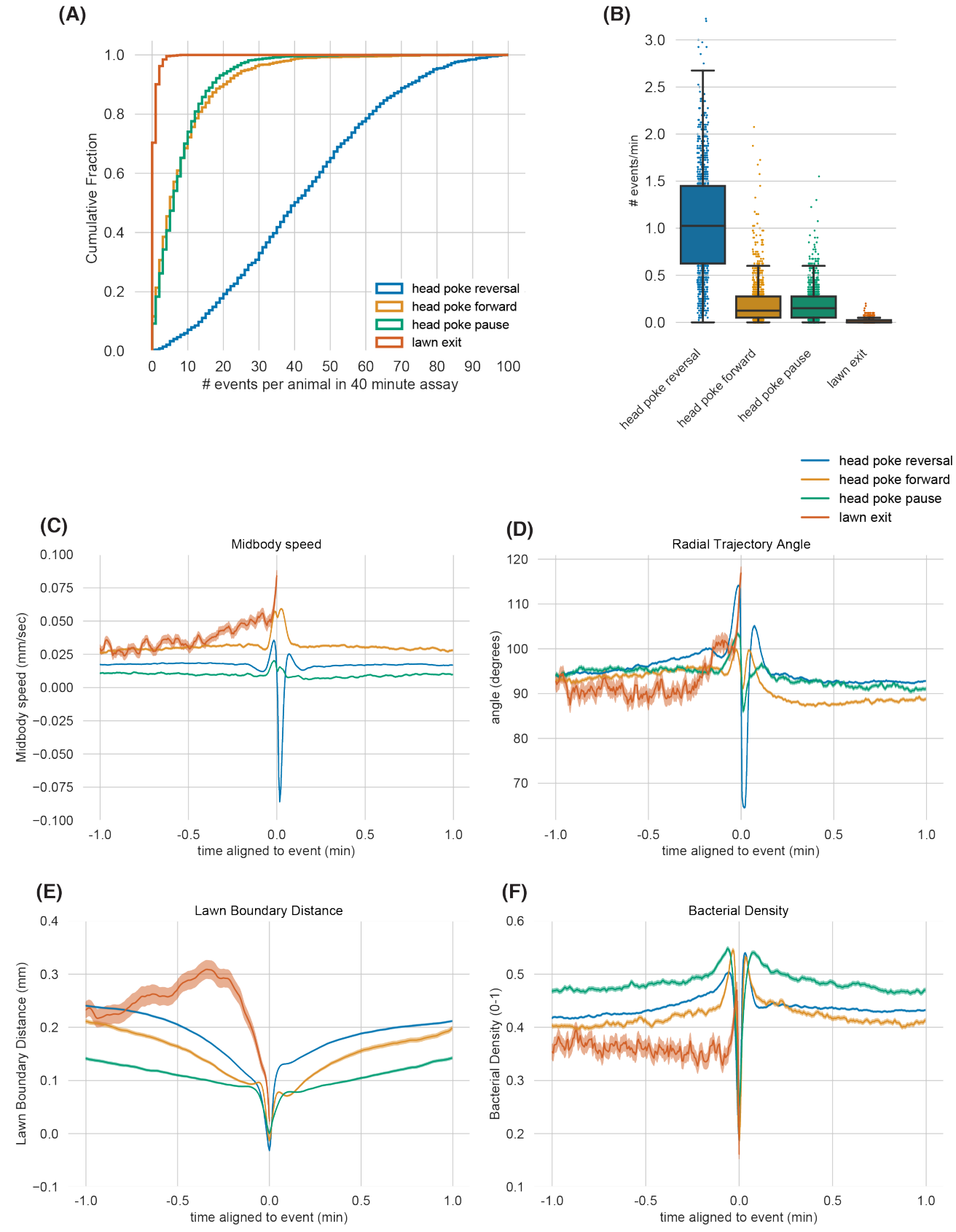

(A)
Cumulative Fraction
events per animal in 40 minute assay
head poke reversal
head poke forward
head poke pause
lawn exit
(B)
events/min
head poke reversal
head poke forward
head poke pause
lawn exit
(C)
Midbody speed
Midbody speed (mm/sec)
time aligned to event (min)
(D)
Radial Trajectory Angle
angle (degrees)
time aligned to event (min)
(E)
Lawn Boundary Distance
Lawn Boundary Distance (mm)
time aligned to event (min)
(F)
Bacterial Density
Bacterial Density (0-1)
time aligned to event (min)

Table 2-1. Quantification of behaviors throughout small bacterial lawns.
Top, mean values of locomotion features on different regions of small bacterial lawns: at the lawn border, lawn interior, or during off-lawn bouts. LBD = lawn boundary distance.
Middle, frequency of different lawn boundary interaction behaviors.
Bottom, relative frequency and duration of different head poke types before resuming forward movement.
Relates to data in Figures 2-2 A-D and 2-4 A and B.

Locomotion features throughout the lawn

	at border (LBD < 0.3 mm)	*lawn interior* (LBD >= 0.3 mm)	*Off-lawn bouts*	fold enrichment at the border
percentage of time spent	60	38	2	3.7
absolute speed (mm/sec)	0.03	0.03	0.09	
angular speed (deg/sec)	75.2	69.1	43.1	
bacterial density	0.52	0.25	0.21	

Lawn Boundary Interaction Frequencies

	fraction of animals that do one	*# events/min*	*# minutes per event*	*# events per 40-minute assay*
lawn exit	0.30	0.01	87.1	0.5
head poke reversal	1.00	1.07	0.9	42.6
head poke forward	0.88	0.21	4.8	8.3
head poke pause	0.91	0.19	5.2	7.7

	fraction of total head pokes in each category	time before resuming forward movement (sec)
head poke reversal	0.73	2.8
head poke forward	0.14	n/a
head poke pause	0.13	3.4

Figure 2-5

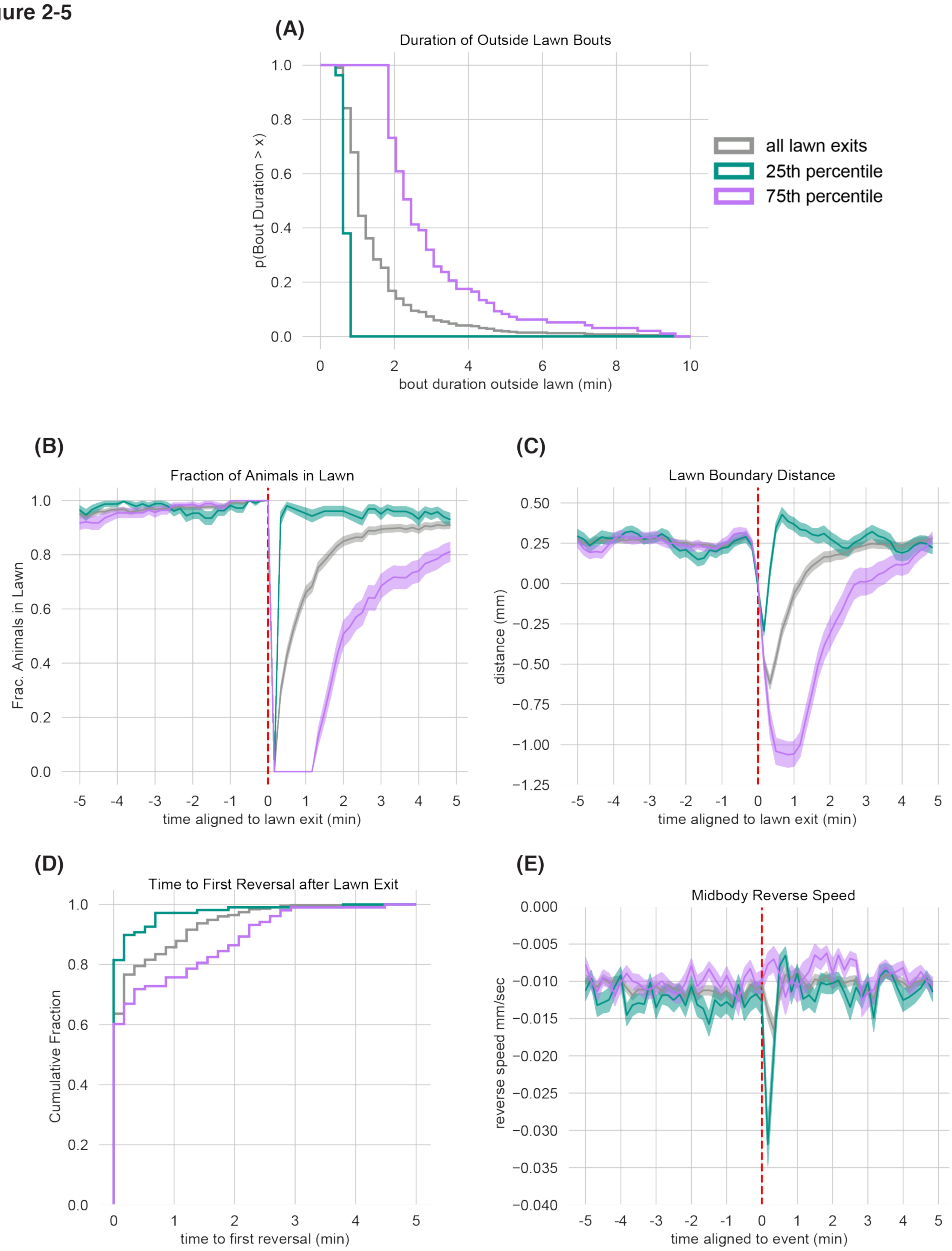


Figure 2-5. Lawn exit bout duration correlates with behavior off food.
(A) Complementary cumulative distribution function (ccdf) for bout duration off food following lawn exits. Data from all bouts, 25th, and 75th percentile bout durations are plotted in all panels.
(B) Fraction of animals in the lawn before lawn exits.
(C) Lawn Boundary Distance in 10 second bins before lawn exit.
(D) Cumulative distribution of the time in minutes to first reversal after lawn exit.
(E) Midbody Reverse Speed in 10 second bins before lawn exit.
In panels B,C,E, dark line is mean computed across behavior aligned to events, shaded area is standard error.

Segmenting Behavioral States on Bacterial Food

The results described above suggest that behavioral states that last one minute or more influence instantaneous decisions at the lawn edge. In previous work studying *C. elegans* behavior on uniform lawns of bacteria, long lasting high and low arousal states called roaming and dwelling have been rigorously defined using Hidden Markov Models (HMM) (Cermak et al., 2020; Flavell et al., 2013). To ask whether similar behavioral states were generated on the small non-uniform lawns that I used, I trained a two-state HMM using the same parameters as the roaming/dwelling HMM, absolute speed and angular speed, as feature dimensions. I observed a similar separation between low speed, high angular speed dwelling states and high speed, low angular speed roaming states (Figure 2-6 A). Using established parameters to separate these behaviors, I derived transition probabilities and state durations for Roaming and Dwelling and found that animals spent 17% of the time roaming and 83% dwelling (Figure 2-6 B-D, Table 2-2). The differences in behavior between animals on small lawns and uniform lawns will be considered in greater detail in Chapter 4.

In order to capture a fuller picture of *C. elegans* behavior on small food lawns, I expanded the measured behavioral features beyond speed and angular speed. Data was segmented data into consecutive 10 second bins (30 frames each at 3 frames per second). 10 second bins were chosen based on prior work (Flavell et al., 2013) and on the duration of lawn interaction behaviors described above. The following features were derived (Figure 2-7):

1) fraction of time moving forward per bin
2) fraction of time reversing per bin
3) fraction of time paused per bin
4) mean head speed per bin (mm/sec)
5) mean head angular speed per bin (deg/sec)
6) mean midbody forward speed per bin (mm/sec)
7) mean midbody reverse speed per bin (mm/sec)
8) mean midbody angular speed per bin (deg/sec)
9) mean tail speed per bin (mm/sec)
10) mean tail angular speed per bin (deg/sec)
11) mean head angular velocity relative to midbody per bin (deg/sec)
12) mean head radial velocity relative to midbody per bin (mm/sec)
13) mean quirkiness per bin. $Q = \sqrt{1 - \frac{a^2}{A^2}}$, where a is the minor and A is the major axis of a bounding box surrounding the animal, respectively. Values closer to 1 means the animal's shape is more elongated and thinner; closer to 0 indicates a more rounded shape.

Head, midbody and tail anchor points were derived following (Javer et al., 2018): using a skeletonized spline comprised of 49 points along the worm's body, the head, midbody and tail were defined as the centroid positions of the first eight body points, body points

17-33, and body points 45-49, respectively. Speed and angular speed were calculated for head, midbody and tail (see Methods for more details).

Many of these metrics are correlated since they measure body parts that have correlated movement. To remove these correlations and ascertain the dimensionality of these features, I performed Principal Components Analysis (PCA). This method finds a set of orthogonal bases that comprise the original dataset such that each principal component maximizes more variance than the next. Performing this analysis revealed that the first 3 principal components (PCs) explained ~85% of the observed overall variance in the dataset (Figure 2-8). Inspecting the loadings (eigenvalues) across PCs shows that PC1 (61% of variance explained) correlates to forward movement and speed. By contrast, PC2 (16% of variance explained) is correlated with reversals and reverse speed. PC3 (8% of variance explained) is primarily anticorrelated with the animal's quirkiness, which represents sharp turns and coiled animal shapes. Histograms of PC1 and PC2 across the entire wildtype dataset yielded a gaussian shape for PC1 with a long positive tail and a bimodal shape in PC2. Plotting the bivariate histogram of PC1 vs. PC2 yielded two clusters, corresponding to 10 second bins that did or did not include a reversal (Figure 2-8 B-D). The bins without reversals (lower cluster) showed more skew towards high values of PC1, corresponding to high forward speed uninterrupted by reversals, the characteristic behavior of roaming animals.

Figure 2-6

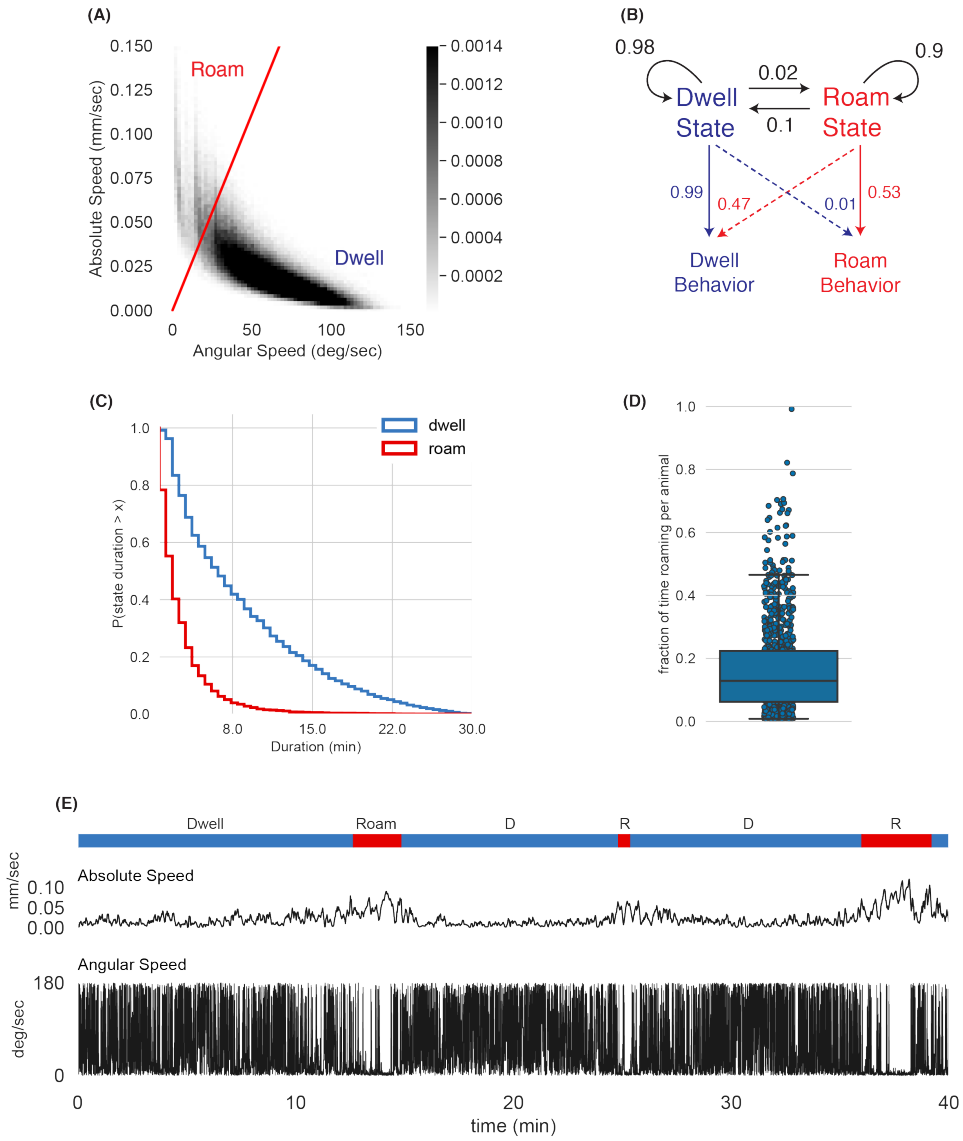
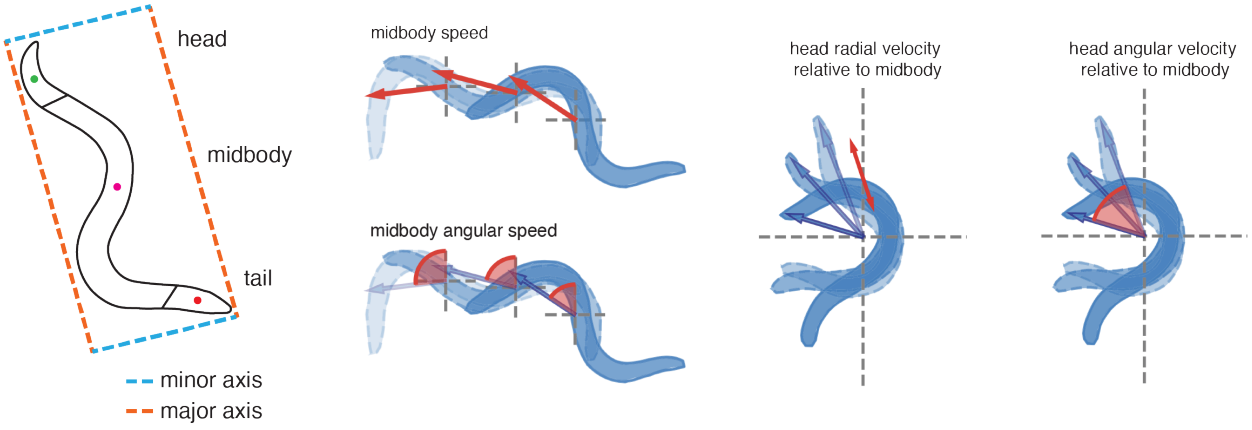


Figure 2-6. Roaming and Dwelling states in non-uniform food environments.
(A) Scatter plot of average absolute speed and angular speed in 10 s intervals for animals on small lawns (n = 978, 234,720 bins). Two clusters of data points correspond to: high speed/low angular speed roaming and low speed/high angular speed dwelling.
(B) Two-state Hidden Markov Model trained on wild-type animals. Roaming Behavior and Dwelling Behavior correspond to 10 sec intervals either above or below the separating line in (A), respectively. Roaming and dwelling states are inferred from the Hidden Markov Model. Black arrows indicate transition probabilities between states and colored arrows represent emission probabilities.
(C) Ccdfs for roam and dwell state durations in minutes.
(D) Fraction of time spent roaming per animal. Boxes indicate median and interquartile range.
(E) An example behavioral trajectory decoded by roaming/dwelling HMM.

Figure 2-7

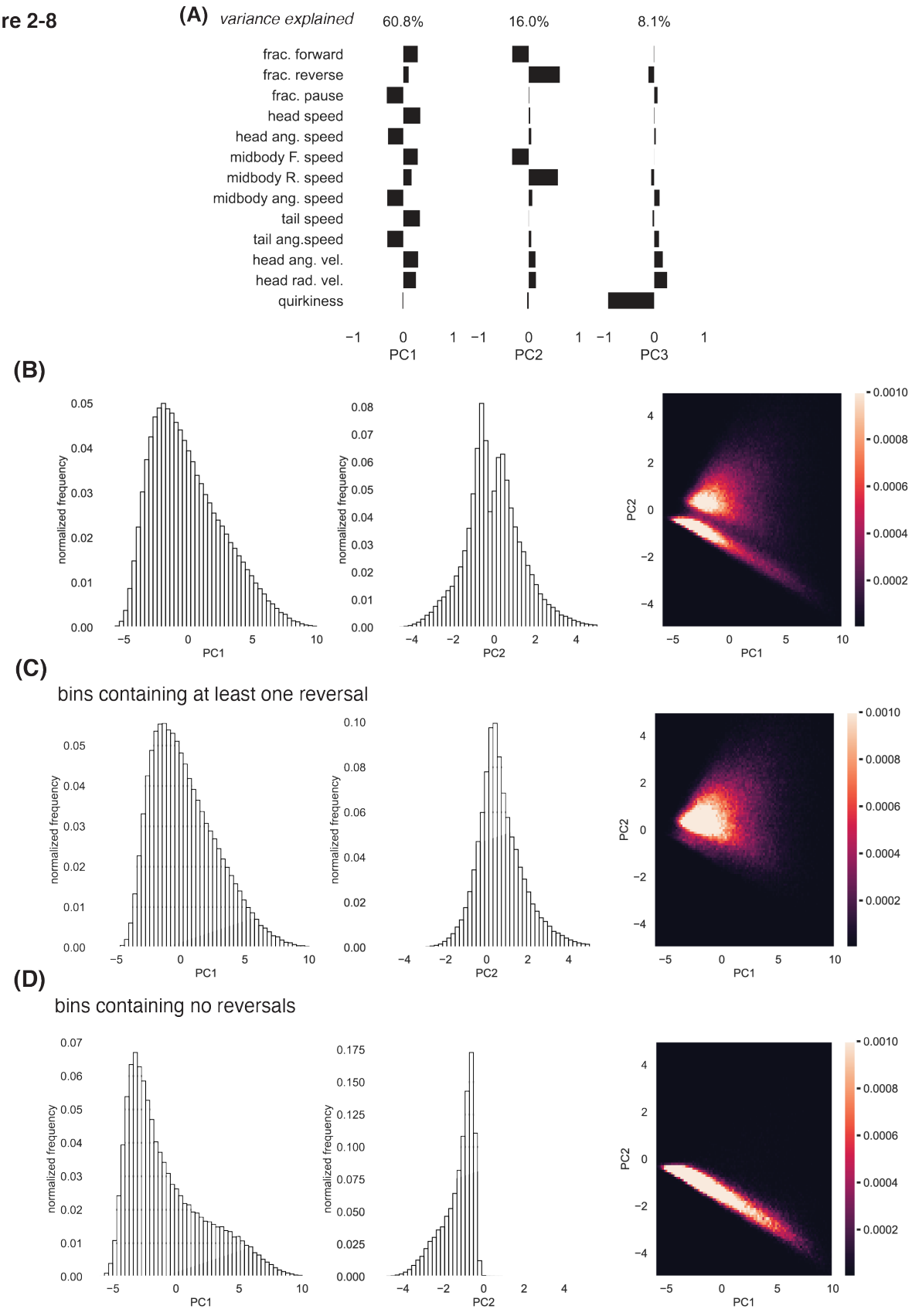


images adapted from Javer, et. al. (2018)

Figure 2-7. Deriving locomotion features.
(A) Schematic of animal body shape with key body points indicated. The minor and major axes of the bounding box are indicated.
(B) Schematics illustrating derived behavioral features. *Adapted from Javer, et al (2018)*.

Figure 2-8. Principal components analysis reveals behavioral data bins with and without reversals.
(A) Principal component weights and signs across 13 locomotion features.
(B) Empirical distributions of PC1, PC2 and joint distribution of PC1 and PC2 for 10 s intervals for animals on small lawns (n = 978, 234,720 bins).
(C) Same as (B) but only bins containing at least one reversal are shown.
(D) Same as (B) but only bins containing no reversals are shown.

Figure 2-8


(A)
variance explained
60.8%
16.0%
8.1%
frac. forward
frac. reverse
frac. pause
head speed
head ang. speed
midbody F. speed
midbody R. speed
midbody ang. speed
tail speed
tail ang.speed
head ang. vel.
head rad. vel.
quirkiness
PC1
PC2
PC3
(B)
normalized frequency
PC1
PC2
(C)
bins containing at least one reversal
(D)
bins containing no reversals

Modeling behavioral states using Hidden Markov Models with Gaussian emissions and Principal Component Observations

I decided to model the distribution of behaviors using an HMM with Gaussian emission probabilities $p(x|z) = \prod_{k=1}^{K} \mathcal{N}(x|\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)^{z_k}$, where *x* is the vector of features (PC1 and PC2), and z_k are the hidden states that are each parametrized by a 2-dimensionsal gaussian distribution with mean and covariance matrix given by $\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k$ respectively. Observations drawn from this joint distribution would yield data in clusters that tile the parameter space (Figure 2-9). The performance of trained HMMs was evaluated by comparing the likelihood of a given model to that of a baseline model consisting of a simple multivariate Gaussian over the input dimensions $x \sim \mathcal{N}(\mu, \Sigma)$, where μ, Σ are the maximum likelihood estimates of the mean and covariance.

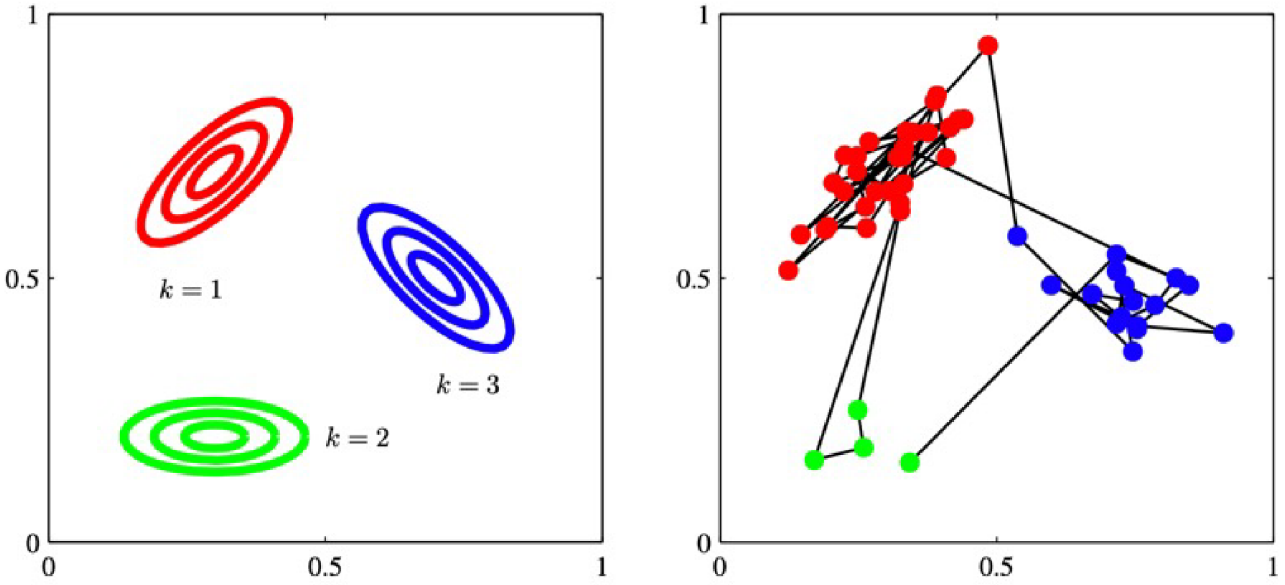


Figure 13.8 Illustration of sampling from a hidden Markov model having a 3-state latent variable $\mathbf{z}$ and a Gaussian emission model $p(\mathbf{x}|\mathbf{z})$ where $\mathbf{x}$ is 2-dimensional. (a) Contours of constant probability density for the emission distributions corresponding to each of the three states of the latent variable. (b) A sample of 50 points drawn from the hidden Markov model, colour coded according to the component that generated them and with lines connecting the successive observations. Here the transition matrix was fixed so that in any state there is a 5% probability of making a transition to each of the other states, and consequently a 90% probability of remaining in the same state.

Adapted from Bishop (2006)

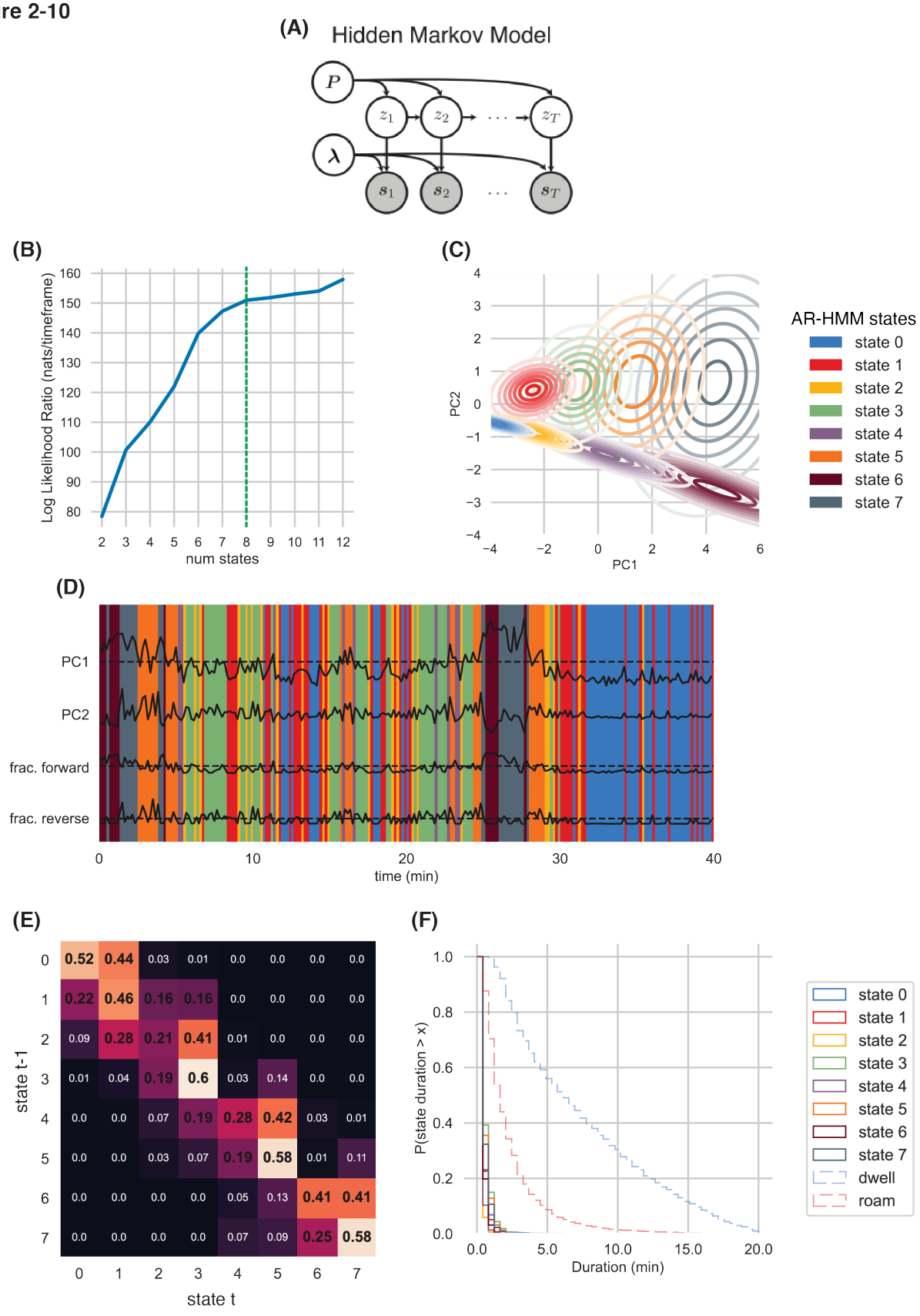
Figure 2-9. Illustration of sampling from a Hidden Markov model.
Adapted from Bishop, 2006.

To choose the optimal number of hidden states *k* for modeling data in the PC1, PC2 feature space, I trained HMMs over a range of hidden states on a subset of the total wild type dataset, then evaluated the log likelihood of these trained models on a held-out test data set. This “predictive log likelihood” is measured in nats per timeframe improvement over the baseline Gaussian model (one nat is the information content of an event whose probability is 1/*e*), and measures how well a set of model parameters generalizes to unseen data. The resulting HMMs yielded monotonically increasing predictive log likelihood that plateaued after k = 8 hidden states over the range of k tested (Figure 2-10 B). As is often the case, adding more hidden states subdivides the feature space into smaller and smaller regions, leading to diminishing returns in overall log-likelihood (Buchanan et al., 2017). Therefore, we chose the eight-state model for further analysis.

The emission distributions for the eight-state model were found exclusively either in high PC2 or low PC2 clusters: states 1, 3, 5, and 7 each contained reversals whereas states 0, 2, 4, and 6 did not (Figure 2-10 C). Inspecting example traces of inferred hidden states generated by this model showed that many states frequently changed as animals switched back and forth between forward and reverse locomotion (Figure 2-10 D). To further define the state metrics, we observed the transitions between the eight states. The transition matrix reveals a block structure along the diagonal where transitions to adjacent states (state i-1 or state i+1 for a given state i) were approximately as likely as self-transitions (Figure 2-10 E). A straightforward interpretation of these results is that the eight states can be collapsed into approximately four longer-lasting states according to their average PC1 levels, each including bins with and without a reversal (0 merged with 1, 2 merged with 3, 4 merged with 5, and 6 merged 7) (Figure 2-10 E).

Figure 2-10. Modeling behavioral states across principal component feature space using a Gaussian Hidden Markov Model.
(A) A graphical model for a Hidden Markov Model. Nodes represent random variables, shaded nodes are observed variables, unshaded nodes are hidden variables. Edges represent conditional dependencies. Hidden z_t states show Markovian conditional dependency on the previous hidden state. P and λ, represent the transition and emission probabilities, respectively. Here, λ are Gaussian-distributed over PC1 and PC2 features. *Adapted from Linderman, 2016.*
(B) Predictive log likelihood ratio (computed on held out test data) with respect to a multivariate normal distribution of Gaussian HMMs with increasing number of states. Green dotted line indicates chosen 8 state model.
(C) λ Gaussian emission distribution for each hidden state. Contours of constant probability density for emission distributions correspond to each of the 8 states of the latent variable.
(D) Gaussian HMM state classifications for an example 40-minute behavioral trace. PC1, PC2, and Z-scored fraction of time moving forward and fraction of time reversing per 10 sec bin are shown. Dotted lines are at 0 for each dimension.
(E) Transition matrix for 8-state Gaussian HMM. Rows index state at *t-1*, columns index state at *t.*
(F) Empirical ccdf for state durations according to 8-state Gaussian HMM compared to state durations from roaming/dwelling model (n = 978 wild type animals on small lawns).

Figure 2-10


(A)
Hidden Markov Model
P
z_1
z_2
z_T
λ
s_1
s_2
s_T
(B)
Log Likelihood Ratio (nats/timeframe)
num states
(C)
PC2
PC1
AR-HMM states
state 0
state 1
state 2
state 3
state 4
state 5
state 6
state 7
(D)
PC1
PC2
frac. forward
frac. reverse
time (min)
(E)
state t-1
state t
(F)
P(state duration > x)
Duration (min)
state 0
state 1
state 2
state 3
state 4
state 5
state 6
state 7
dwell
roam

Modeling behavioral states using Autoregressive Hidden Markov Models and 5 feature dimensions

For the next analysis, I de-emphasized reversals and prioritized the variance in speed and persistence of forward locomotion. This analysis included the following five features (all in 10 second bins) without PCA dimensionality reduction:

1) fraction of time moving forward per bin
2) mean midbody forward speed per bin (mm/sec)
3) mean midbody angular speed per bin (deg/sec)
4) mean head angular velocity relative to midbody per bin (deg/sec)
5) mean head radial velocity relative to midbody per bin (mm/sec)

To capture longer time scales of behavioral correlations, I changed the HMM architecture to an Autoregressive Hidden Markov Model (AR-HMM), which has been used to model behavioral states in worms and other animals (Buchanan et al., 2017; Wiltschko et al., 2015). In addition to discrete hidden states, the AR-HMM builds in a Markovian conditional dependence of each observation vector x_t on x_{t-1}, generating a smooth prediction of behavior on short time scales based on the current hidden state (Figure 2-11). Each hidden state or "behavioral mode" has its own set of autoregressive dynamics that must be learned by the fitting procedure.

The notation for AR-HMMs that follows is borrowed from (Buchanan et al., 2017).

In order to build in the desired time duration of inferred behavioral modes, we add a "sticky" prior on the rows of the transition matrix, $\pi_k \sim \mathrm{Dir}(\alpha 1_K + \kappa e_k)$, where 1_K is a length-K vector of ones and e_k is a length K unit vector with a 1 in the k-th position. $\alpha > 0 \text{ and } \kappa > 0$ parameters relate to the "concentration" and "stickiness," respectively. Large values of κ push the transition matrix closer to the identity matrix, thus increasing the probability of self-transition at time *t* and leading to longer behavioral state durations. The value of α has been shown to have little effect on predictive log likelihood during model fitting, so we fixed its value to 1 throughout this work (Buchanan et al., 2017).

AR-HMMs were trained over a range of values of k (number of states) and κ (stickiness parameter) and the predictive log likelihood was evaluated on held out data as described above (Figure 2-11 B). As before, the log likelihood of a given model was compared to the log likelihood of a Gaussian multivariate model by taking a likelihood ratio. Like the Gaussian HMMs discussed above, predictive log likelihood increased with increasing k. Increased κ increases self-transition weights across states, in essence reducing model sensitivity to small changes in the input dimensions and yields longer behavioral state durations. For values of κ = 10,000 and above, the predictive log likelihood began to plateau at k = 4 states. Models at k = 4, κ = 10,000, 25,000 and 100,000 yielded lengthening state durations as κ increases; a model with κ = 25,000 and k = 4 yielded state durations of the same order of magnitude as roaming or dwelling

states (Figure 2-11 C). All modeling work requires tradeoffs among hyperparameters but ultimately proves its worth by illuminating biological interpretation. In my estimation, this model captures important aspects of our datasets, as described below.

As expected, and consistent with the earlier HMM, behavioral state emission distributions tiled the observation dimensions in order of increasing speed and persistent forward locomotion (Figure 2-11 D). Each state is scaled such that such that state 0 and state 3 have the lowest and highest mean and variance, respectively. State 1 was the most commonly observed (~40%), followed by state 2 (~30%), state 0 (~20%) and state 3 (~10%) (Figure 2-11 E). The transition matrix reveals large values of self-transition probability, as expected from our sticky prior (Figure 2-11 F). The blocked structure of off-diagonal transition probabilities reveals that animals change states sequentially without skipping over intermediates, likely reflecting the continuous nature of forward speed.

To gain a better understanding of the meaning of each AR-HMM state, I returned to the full set of behavioral parameters and examined them in each behavioral state. Incrementing from state 0 to state 3 increases mean speed, fraction of time moving forward, head angular velocity, and head radial velocity (Figure 2-12 A), all facets of the principal components defined in Figure 2-8. Incrementing state is also associated with faster midbody reversal speed, midbody angular speed and fraction of time pausing per bin. The fraction of time reversing is lowest in state 0, the paused state, and similar in states 1, 2, and 3. The most dramatic difference between states 2 and 3 is the near absence of pauses in state 3 (Figure 2-12 A).

To build intuition about these behavior states, I examined their overlap with roaming and dwelling states (Figure 2-12 B). Dwelling is composed of state 0, 1 and 2, with state 1 representing the most often co-occurring state. Roaming is composed of states 2 and 3. Conversely, states 0 and 1 uniformly correspond to dwelling states and states 2 and 3 preferentially align with dwelling and roaming, respectively, but overlap. Thus, different behavioral classification schemes yield related but not identical behavioral states.

Figure 2-11. Modeling behavioral states across locomotory feature dimensions using an Autoregressive Hidden Markov Model.
(A) A graphical model for an Autoregressive Hidden Markov Model. The hidden (discrete states) follow Markovian dynamics and the observed features x_t (continuous state) depends only on the preceding observed features and the current hidden state. The global parameters θ include the discrete state transition matrix and the autoregressive dynamics parameters (see Methods)*. Adapted from Buchanan, et al 2017.*
(B) Predictive log likelihood ratio (computed on held out test data) with respect to a multivariate normal distribution for Gaussian HMMs trained across increasing values of *kappa* and increasing number of states, *k.* Green dotted line indicates chosen AR-HMM model at *kappa* = 25,000, k = 4 states.
(C) Empirical ccdf for state durations according to 4-state AR-HMMs across 3 values of *kappa* = 10000, 25000, and 100000, compared to state durations from roaming/dwelling model. Green dotted line indicates chosen AR-HMM model at *kappa* = 25,000, k = 4 states (n = 978 wild type animals on small lawns).
(D) Gaussian emission distributions of 4-state AR-HMM model. Contours of constant probability density are plotted among pairs of z-scored feature dimensions.
(E) Empirical mean fraction of time spent in each AR-HMM state per animal in small lawn dataset (n = 978). See Table 2-2 for the mean values per state.
(F) Transition matrix for 4-state AR-HMM. Rows index state at *t-1*, columns index state at *t.*

Figure 2-11

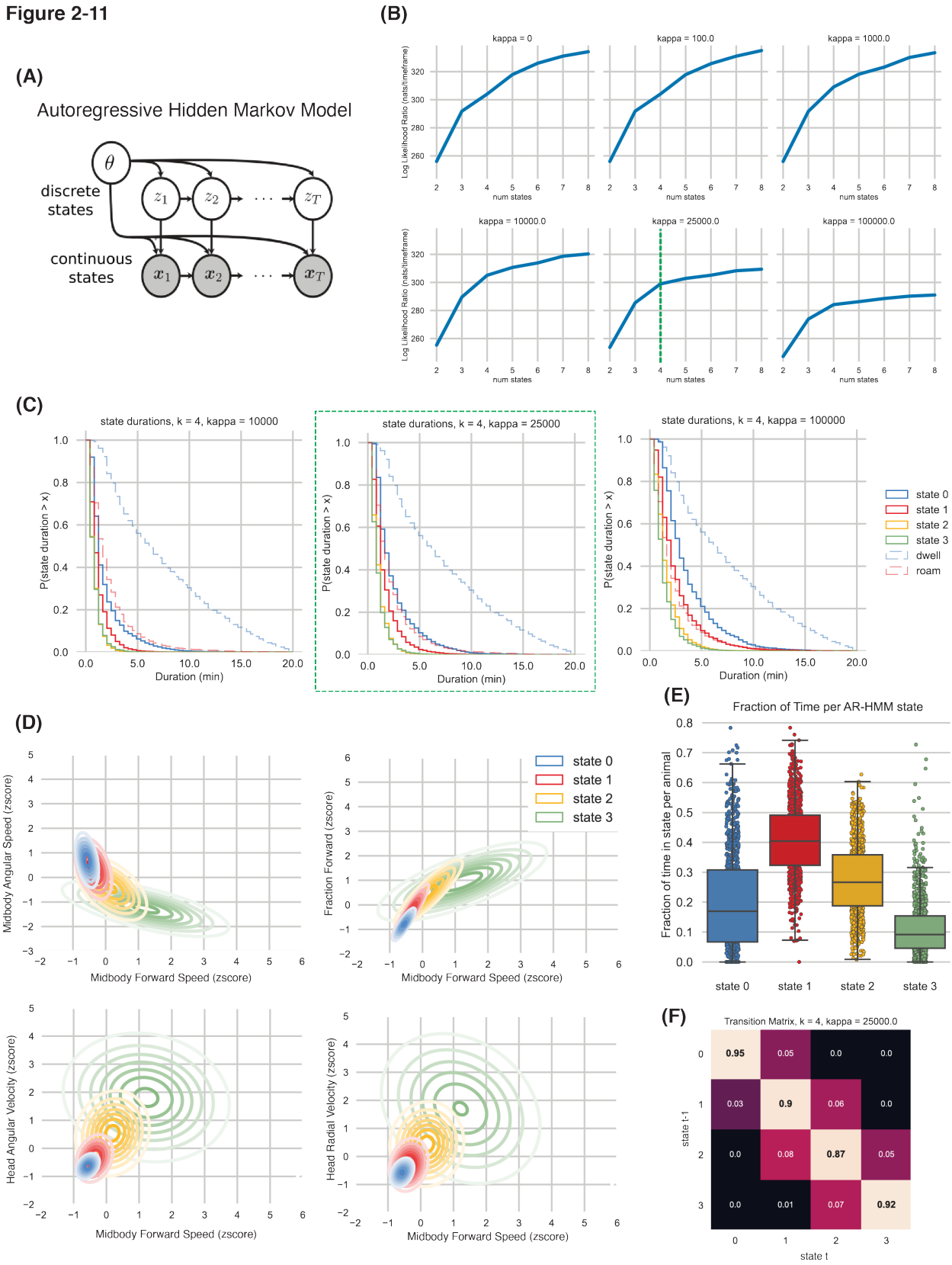
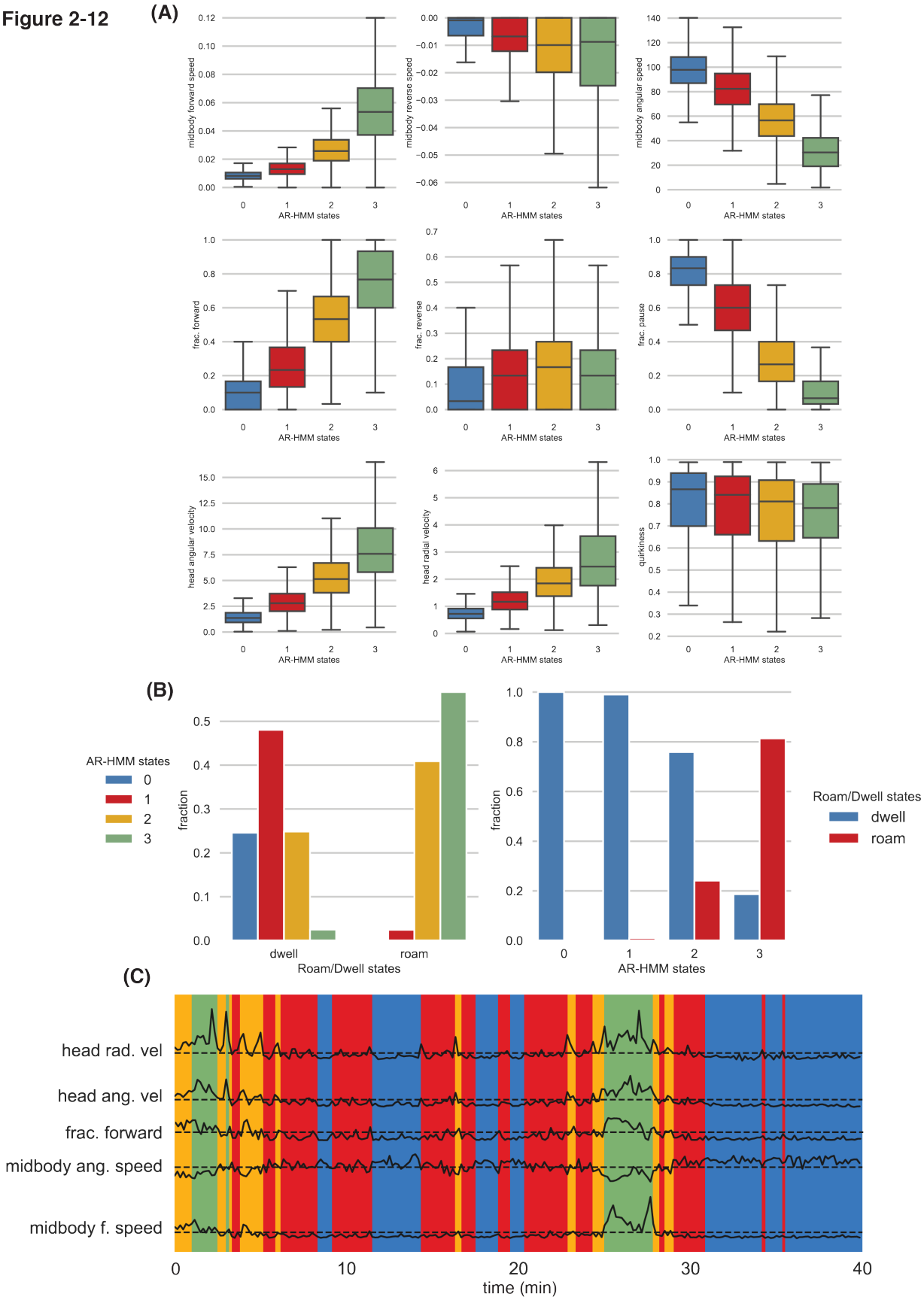

(A)
Autoregressive Hidden Markov Model
discrete states
continuous states
(B)
kappa = 0
kappa = 100.0
kappa = 1000.0
kappa = 10000.0
kappa = 25000.0
kappa = 100000.0
Log Likelihood Ratio (nats/timeframe)
num states
(C)
state durations, k = 4, kappa = 10000
state durations, k = 4, kappa = 25000
state durations, k = 4, kappa = 100000
P(state duration > x)
Duration (min)
state 0
state 1
state 2
state 3
dwell
roam
(D)
Midbody Angular Speed (zscore)
Fraction Forward (zscore)
Head Angular Velocity (zscore)
Head Radial Velocity (zscore)
Midbody Forward Speed (zscore)
(E)
Fraction of Time per AR-HMM state
Fraction of time in state per animal
(F)
Transition Matrix, k = 4, kappa = 25000.0
state t-1
state t
0.95
0.05
0.0
0.0
0.03
0.9
0.06
0.0
0.0
0.08
0.87
0.05
0.0
0.01
0.07
0.92

Figure 2-12. AR-HMM behavioral states correlate with increasing levels of locomotory arousal.
(A) Empirical distributions of feature dimensions classified among 4 AR-HMM states. Boxes indicate median and interquartile range.
(B) Co-occupancy of roaming and dwelling with 4 AR-HMM states. Left, fraction of AR-HMM states found in roaming and dwelling states. Right, fraction of roaming and dwelling states found in each AR-HMM state.
(C) AR-HMM state classifications for an example 40-minute behavioral trace. Z-scored versions of each of the 5 observation dimensions are plotted vertically, dotted lines are at 0 for each dimension. Top to bottom: head rad. vel = mean head radial velocity relative to midbody, head ang. vel = mean head angular velocity relative to midbody, frac. forward = fraction of time moving forward per 10 sec bin, midbody ang. speed = mean midbody angular speed, midbody f. speed = mean midbody forward speed. All features are mean per 10 sec bin.

Figure 2-12


(A)
midbody forward speed
midbody reverse speed
midbody angular speed
frac. forward
frac. reverse
frac. pause
head angular velocity
head radial velocity
quirkiness
AR-HMM states
(B)
AR-HMM states
0
1
2
3
fraction
dwell
roam
Roam/Dwell states
Roam/Dwell states
(C)
head rad. vel
head ang. vel
frac. forward
midbody ang. speed
midbody f. speed
0
10
20
30
40
time (min)

Lawn exits occur in high arousal states.
Informed by these behavioral analyses, I returned to the lawn-interaction behaviors of head pokes and lawn exits. 80% of lawn exits occurred during roaming states, whereas 20% occurred during dwelling. As dwelling is more common than roaming, fewer than 0.1% of dwelling state bins, but close to 1% of roaming state bins, contained lawn exits. Together, there was a 5-fold enrichment in lawn leaving from roaming states versus dwelling states (Figure 2-13). Among AR-HMM states, a ~7-fold enrichment of lawn exit events from state 3 was observed. By contrast, head poke reversals showed only a 1.6-fold enrichment from roaming states and a 1.7-fold enrichment from AR-HMM state 3 (Figure 2-14).
Consistent with this, the probability of roaming showed roughly two-fold enrichment over baseline rates five minutes before lawn leaving (35% versus 17% at baseline) and began ramping two minutes before leaving to finally reach a 5-fold enrichment (81% of animals roaming) at the time of lawn exit (Figure 2-13, Table 2-2). Five minutes before leaving, AR-HMM state 3 was two-fold enriched over baseline (21% versus 12% at baseline), and rapidly ramped up one minute before leaving to reach 7-fold enrichment by lawn exit. Aligning state usage to head poke reversals yielded a strikingly different pattern (Figure 2-14, Table 2-2). The fraction of animals roaming and dwelling five minutes before each event corresponded to the baseline with a gradual 60% rise in roaming before head pokes. Similarly, AR-HMM states were enriched for states 2 and 3 at the expense of 1 and 0 before head pokes by ~50-70%, a much smaller enrichment than occurred before lawn exits.

Next, we aligned the full set of behavioral parameters to lawn exits and head poke reversals (Figure 2-15). Approximately three minutes before lawn exit, midbody angular speed and fraction of time moving forward began to increase; an additional increase in midbody forward speed and head radial and angular velocity occurred one minute before lawn exit, along with a small decrease in reversals. Head poke reversals were preceded by smaller changes in locomotory features. I remained curious about the remaining 20% of lawn exits that left from dwelling or non- AR-HMM state 3 states. Do these lawn exits correlate with a different set of behavioral of sensory patterns prior to or after leaving? I compared aligned features to lawn exits that occurred during roaming (“roam-leaving”) or dwelling (“dwell-leaving”) (Figure 2-16). Roam-leaving animals exhibit the characteristic ramping in roaming prior to leaving. However, dwell-leaving animals remain dwelling throughout the five minutes before leaving. They showed no speed ramping prior to leaving and remained at lower speeds than roam-leavers even off food. Upon re-entry to the lawn, both roam-leaving and dwell-leaving animals returned to their pre-leaving baselines within two minutes.

Figure 2-13. Lawn exits occur in high arousal states.
(A) Co-occupancy of lawn exits with roaming and dwelling. Left, percent of time bins classified as roaming or dwelling that contain a lawn exit. Middle, percent of time bins containing a lawn exit that are found in roaming or dwelling states. Right, fold enrichment of lawn exits contained by roaming or dwelling bins scaled by the frequency of roaming or dwelling state bins.
(B) Roaming and dwelling aligned to lawn exits. Left, heatmap of roaming/dwelling state classifications aligned to lawn exit. Right, probability of roaming or dwelling aligned to lawn exit. n = 431 lawn exits from n = 978 animals.
(C) Co-occupancy of lawn exits with AR-HMM states. Left, percent of time bins classified as each AR-HMM state that contain a lawn exit. Middle, percent of time bins containing a lawn exit that are found AR-HMM states. Right, fold enrichment of lawn exits contained by AR-HMM states scaled by the frequency of each AR-HMM state.
(D) AR-HMM states aligned to lawn exits. Left, heatmap of AR-HMM state classifications aligned to lawn exit. Right, probability of AR-HMM states aligned to lawn exit. n = 431 lawn exits from n = 978 animals. All data derived from 10 second bins.

Figure 2-13

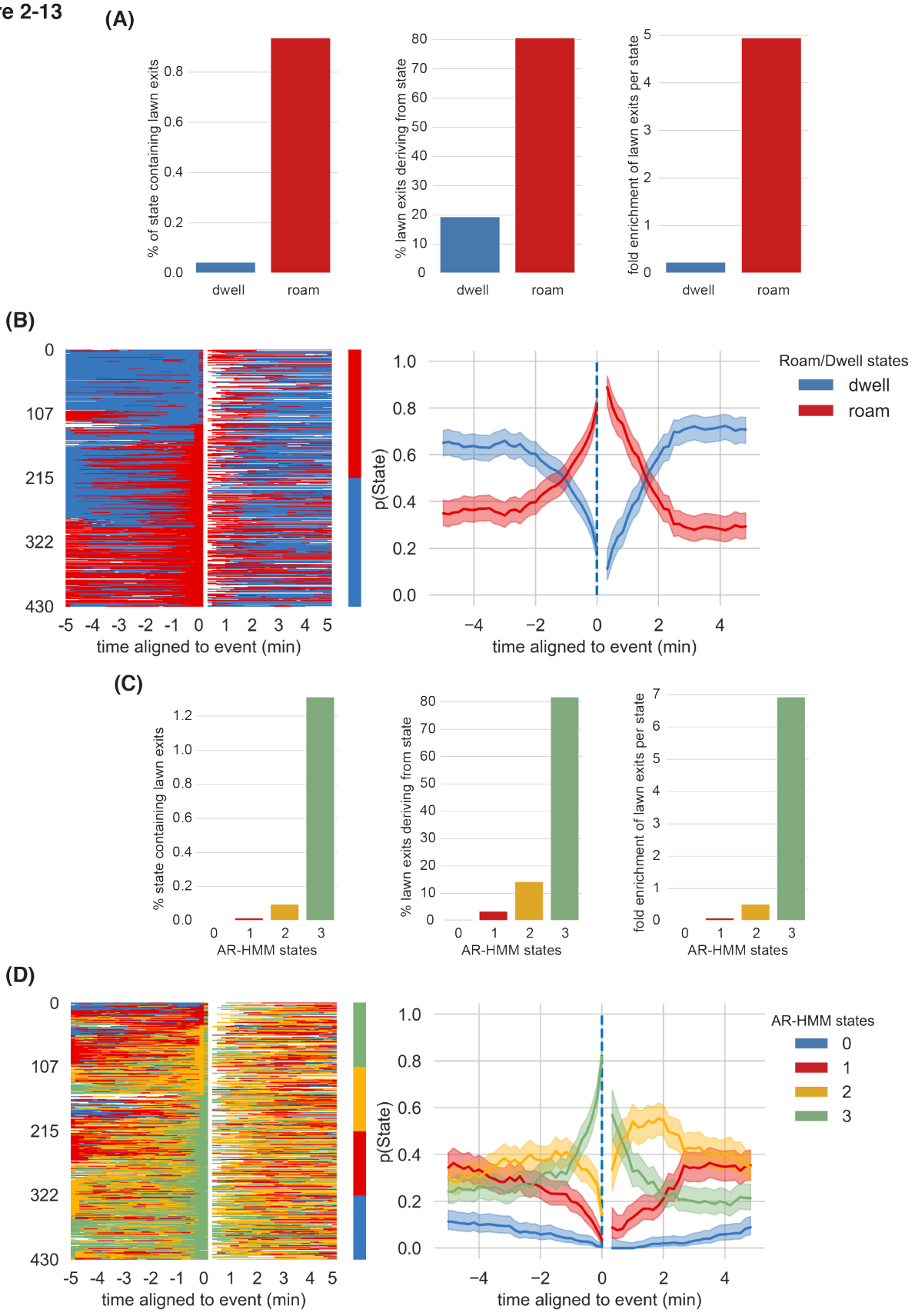

(A)
% of state containing lawn exits
0.0
0.2
0.4
0.6
0.8
dwell
roam
% lawn exits deriving from state
0
10
20
30
40
50
60
70
80
dwell
roam
fold enrichment of lawn exits per state
0
1
2
3
4
5
dwell
roam
(B)
0
107
215
322
430
-5 -4 -3 -2 -1 0 1 2 3 4 5
time aligned to event (min)
p(State)
0.0
0.2
0.4
0.6
0.8
1.0
−4 −2 0 2 4
time aligned to event (min)
Roam/Dwell states
dwell
roam
(C)
% state containing lawn exits
0.0
0.2
0.4
0.6
0.8
1.0
1.2
0 1 2 3
AR-HMM states
% lawn exits deriving from state
0
10
20
30
40
50
60
70
80
0 1 2 3
AR-HMM states
fold enrichment of lawn exits per state
0
1
2
3
4
5
6
7
0 1 2 3
AR-HMM states
(D)
0
107
215
322
430
-5 -4 -3 -2 -1 0 1 2 3 4 5
time aligned to event (min)
p(State)
0.0
0.2
0.4
0.6
0.8
1.0
−4 −2 0 2 4
time aligned to event (min)
AR-HMM states
0
1
2
3

Figure 2-14. Head poke reversals are not highly enriched by behavioral states.
(A) Co-occupancy of head poke reversals with roaming and dwelling. Left, percent of 10 sec bins classified as roaming or dwelling that contain a head poke reversal. Middle, percent of 10 sec bins containing a head poke reversal that are found in roaming or dwelling states. Right, fold enrichment of head poke reversals contained by roaming or dwelling bins scaled by the frequency of roaming or dwelling state bins.
(B) Roaming and dwelling aligned to head poke reversals. Left, heatmap of roaming/dwelling state classifications aligned to head poke reversal. Right, probability of roaming or dwelling aligned to head poke reversal. n = 37,379 head poke reversals from n = 978 animals.
(C) Co-occupancy of head poke reversals with AR-HMM states. Left, percent of 10 sec bins classified as each AR-HMM state that contain a head poke reversal. Middle, percent of 10 sec bins containing a head poke reversal that are found AR-HMM states. Right, fold enrichment of head poke reversals contained by AR-HMM states scaled by the frequency of each AR-HMM state.
(D) AR-HMM states aligned to head poke reversals. Left, heatmap of AR-HMM state classifications aligned to head poke reversal. Right, probability of AR-HMM states aligned to head poke reversal. n = 37,379 head poke reversals from n = 978 animals.

Figure 2-14

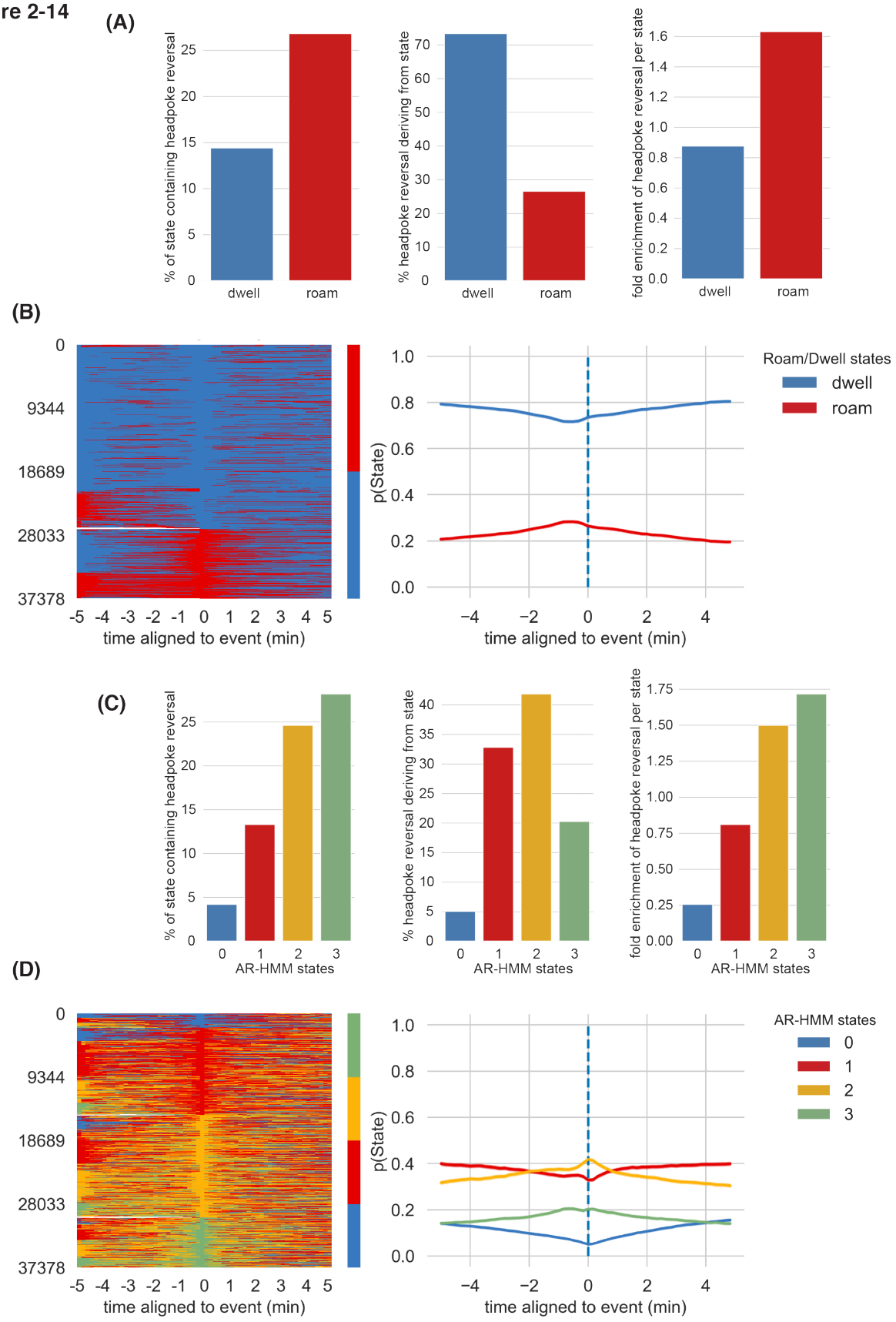
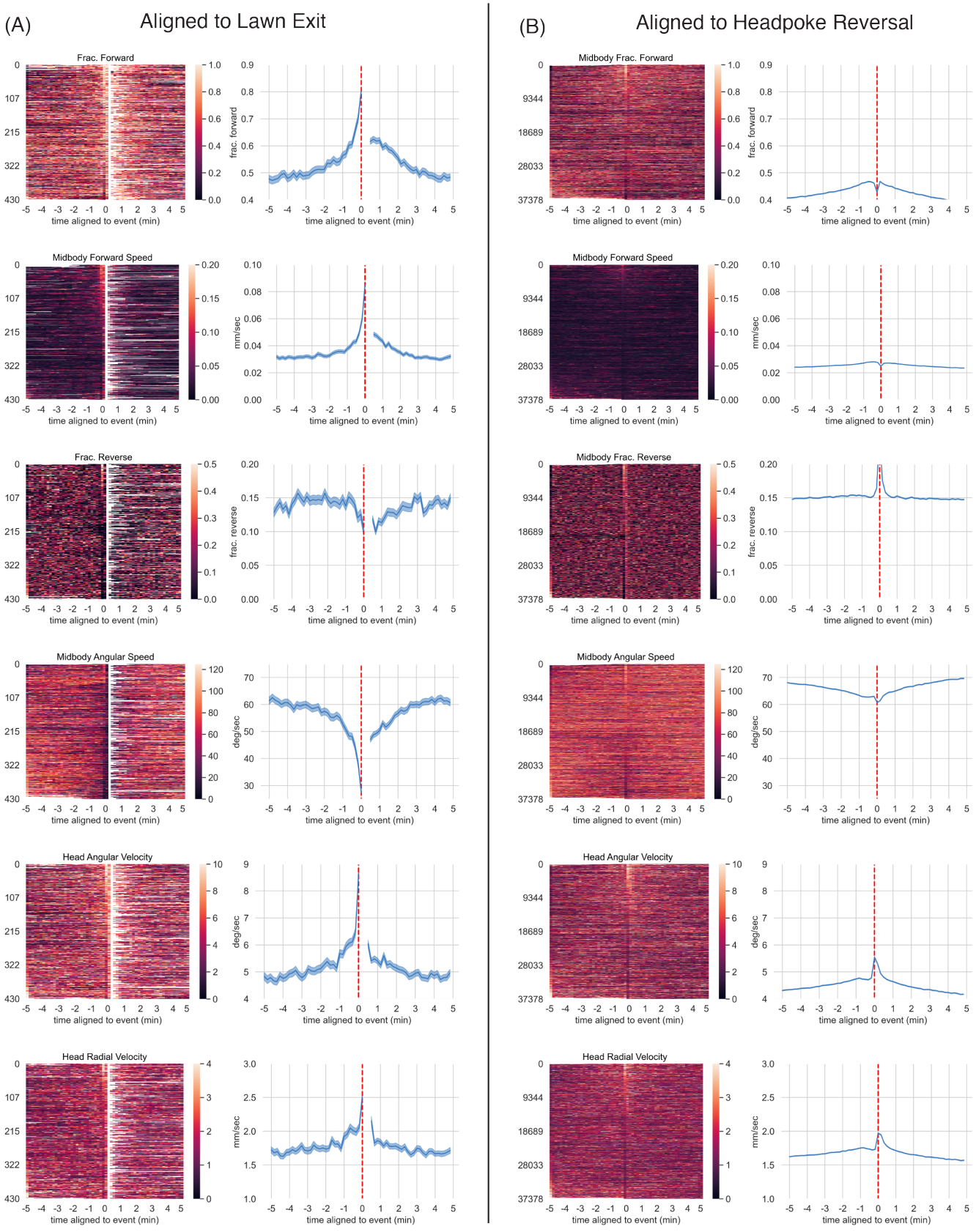

(A)
% of state containing headpoke reversal
dwell
roam
% headpoke reversal deriving from state
fold enrichment of headpoke reversal per state
(B)
time aligned to event (min)
p(State)
Roam/Dwell states
dwell
roam
(C)
AR-HMM states
(D)
AR-HMM states
0
1
2
3

Table 2-2. Quantification of state usages and durations.
Top, percentage of time animals use each state (baseline rate) and mean state durations.
Middle, probability of each state 5 minutes before lawn exit, at time of lawn exit, and enrichment of p(State) at lawn exit over baseline rate.
Bottom, probability of each state 5 minutes before head poke reversal, at time of head poke reversal, and enrichment of p(State) at head poke reversal over baseline rate.
Relates to data in Figures 2-6 D, 2-1 E, 2-13 A and C, 2-14 A and C.

State usage and durations			
state	*percent of time in state*	*mean state duration (min)*	
dwell	83	9.1	
roam	17	2.3	
AR-HMM state 0	20	2.2	
AR-HMM state 1	40	1.4	
AR-HMM state 2	28	0.9	
AR-HMM state 3	12	0.8	
Probability of states aligned to lawn exits			
state	*p(State) 5 min before lawn exit*	*p(State) at lawn exit*	*p(State) enrichment at lawn exit over baseline*
dwell	0.65	0.19	0.2
roam	0.35	0.81	4.7
AR-HMM state 0	0.10	0.00	0.0
AR-HMM state 1	0.30	0.03	0.1
AR-HMM state 2	0.26	0.14	0.5
AR-HMM state 3	0.21	0.82	6.9
Probability of states aligned to head poke reversals			
state	*p(State) 5 min before head poke reversal*	*p(State) at head poke reversal*	*p(State) enrichment at head poke reversal over baseline*
dwell	0.79	0.74	0.9
roam	0.21	0.26	1.6
AR-HMM state 0	0.12	0.05	0.3
AR-HMM state 1	0.34	0.33	0.8
AR-HMM state 2	0.27	0.42	1.5
AR-HMM state 3	0.12	0.20	1.7

Figure 2-15. Behavioral patterns preceding lawn exits differ from those preceding head poke reversals.
(A) Top to bottom, 10-second binned behavioral features aligned to lawn exits. Left, heatmap of individual traces (aligned to n = 431 lawn exits from n = 978 animals), Right, average computed across traces in the heatmap.
(B) Top to bottom, 10-second binned behavioral features aligned to head poke reversals. Left, heatmap of individual traces (aligned to n = 37,379 head poke reversals from n = 978 animals), Right, average computed across traces in the heatmap.

Figure 2-15


(A)
Aligned to Lawn Exit
(B)
Aligned to Headpoke Reversal
Frac. Forward
Midbody Frac. Forward
Midbody Forward Speed
Frac. Reverse
Midbody Frac. Reverse
Midbody Angular Speed
Head Angular Velocity
Head Radial Velocity
frac. forward
frac. reverse
mm/sec
deg/sec
time aligned to event (min)

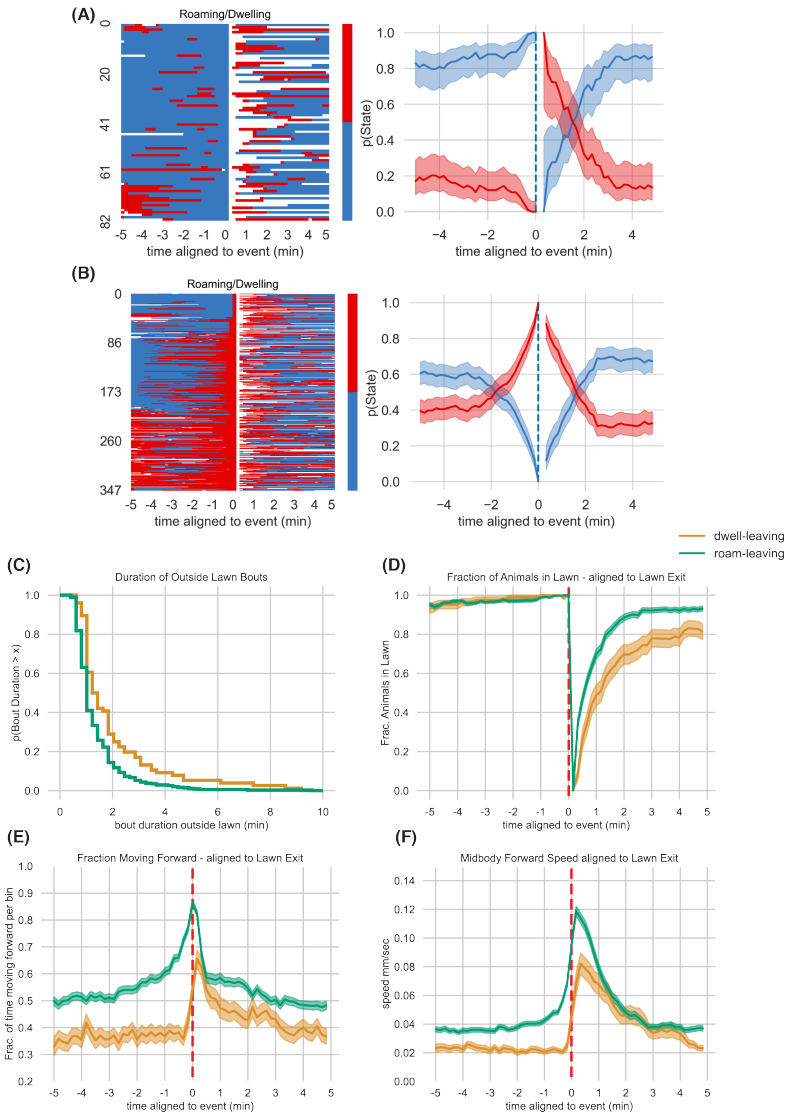


Figure 2-16. Roaming and Dwelling animals show persistent behavioral differences before and after lawn exit.

(A) Roaming and dwelling traces for animals that left lawns from the dwelling state. Left, heatmap of roaming/dwelling state classifications aligned to lawn exit. Right, probability of roaming or dwelling aligned to lawn exit. n = 83 lawn exits from 79 animals.

(B) Roaming and dwelling traces for animals that left lawns from the roaming state. Left, heatmap of roaming/dwelling state classifications aligned to lawn exit. Right, probability of roaming or dwelling aligned to lawn exit. n = 348 lawn exits from 228 animals.

(C) Ccdf of bout duration following lawn exit for animals that left while roaming or dwelling.

(D) Fraction of animals in lawn aligned to lawn exit for animals that left while roaming or dwelling.

(E) Fraction of time spent moving forward per 10-second bin for animals that left while roaming or dwelling.

(F) Mean midbody forward speed for animals that left while roaming or dwelling.

Prediction of lawn leaving using generalized linear models

Finally, given the robust correlations of locomotion features with lawn exits, I asked how well instantaneous patterns predicted lawn leaving, building and testing a set of Generalized Linear Models (GLM) across locomotion features to regress against lawn exits. As the name suggests, GLMs are a generalized version of linear regression that allows the response variable (in this case lawn exits) to have a non-gaussian error distribution via a nonlinear link function (see Methods).
GLMs were trained on a dataset of the following predictors regressed against lawn exits:

1) midbody forward speed
2) midbody reverse speed
3) midbody angular speed
4) head speed
5) head angular speed
6) head angular velocity
7) head radial velocity
8) quirkiness
9) bacterial density
10) lawn boundary distance
11) AR-HMM states

I also added a bias term, composed of a vector of ones to allow the fitting procedure to learn a weight that scales to the overall likelihood of observing the event. Frequent events earned a higher weight on the bias term and vice versa. After training, I evaluated the performance of this model by predicting lawn exits on a held-out test dataset. A given vector of 12 inputs (including the bias) generated a probability of observing a lawn exit, which I thresholded at various values to predict whether the event occurred or not. By varying this threshold over a range from 0 to 1, I calculated a true positive rate and a false positive rate for each threshold value, yielding the Receiver Operating Characteristic (ROC) curve. The area under the curve (AUC) is a common metric for evaluating the goodness of model predictions, with values closer to 1 optimal and a value of 0.5 indicating a model that has no discriminative capacity. This GLM yielded an AUC value of 0.87 (Figure 2-17).

The model had a large negative weight on the bias term, indicating low overall probability of lawn exits, strong positive weights on midbody forward speed and head speed, and a small negative weight on midbody reverse speed and lawn boundary distance. All other facets had a weight near 0. This model agrees with others in its general observation that lawn exits emanate from high speed, low reversal states at the lawn boundary. To ask if single behavioral variables could decode lawn leaving events as well as the ensemble, I trained a series of logit GLMs using only a bias term and a single input feature (Figure 2-18). As expected, the standard errors derived on these weights were smaller than in the multi-input GLM since only one variable was

regressed. Among trained models, those using midbody forward speed, midbody angular speed, head speed, and head angular speed all had significant discriminatory power in predicting lawn leaving events, as judged by their AUC metrics. The GLM derived from midbody speed and head speed attained AUCs of 0.79 and 0.83, respectively on held-out test data. The ROC curve for AR-HMM states alone yielded an AUC of 0.82. Thus, speed metrics and AR-HMM behavioral states both show a high discriminative power for predicting lawn exit events.

Figure 2-17

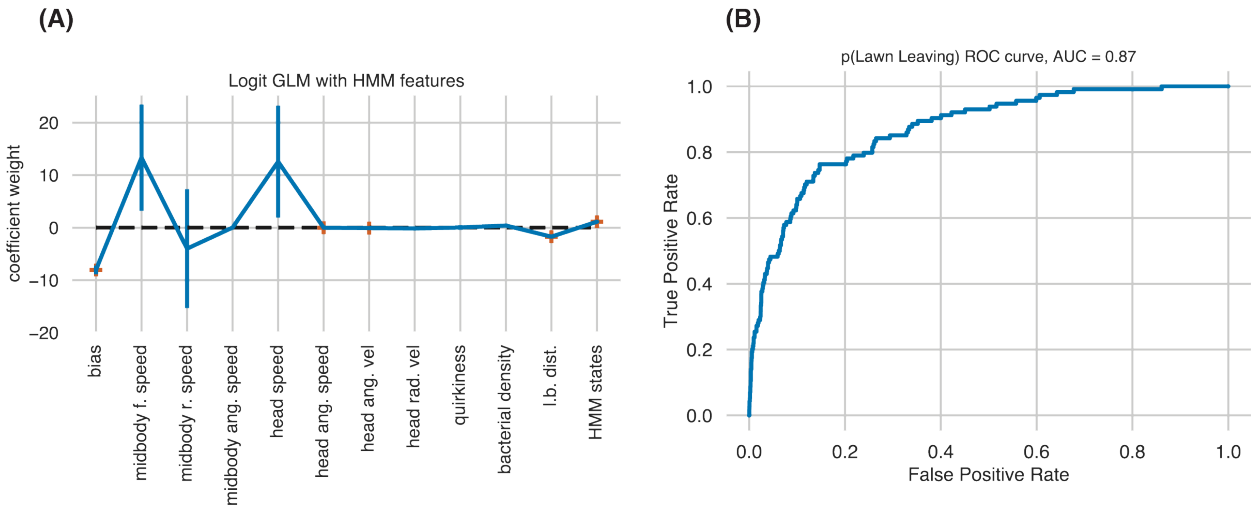
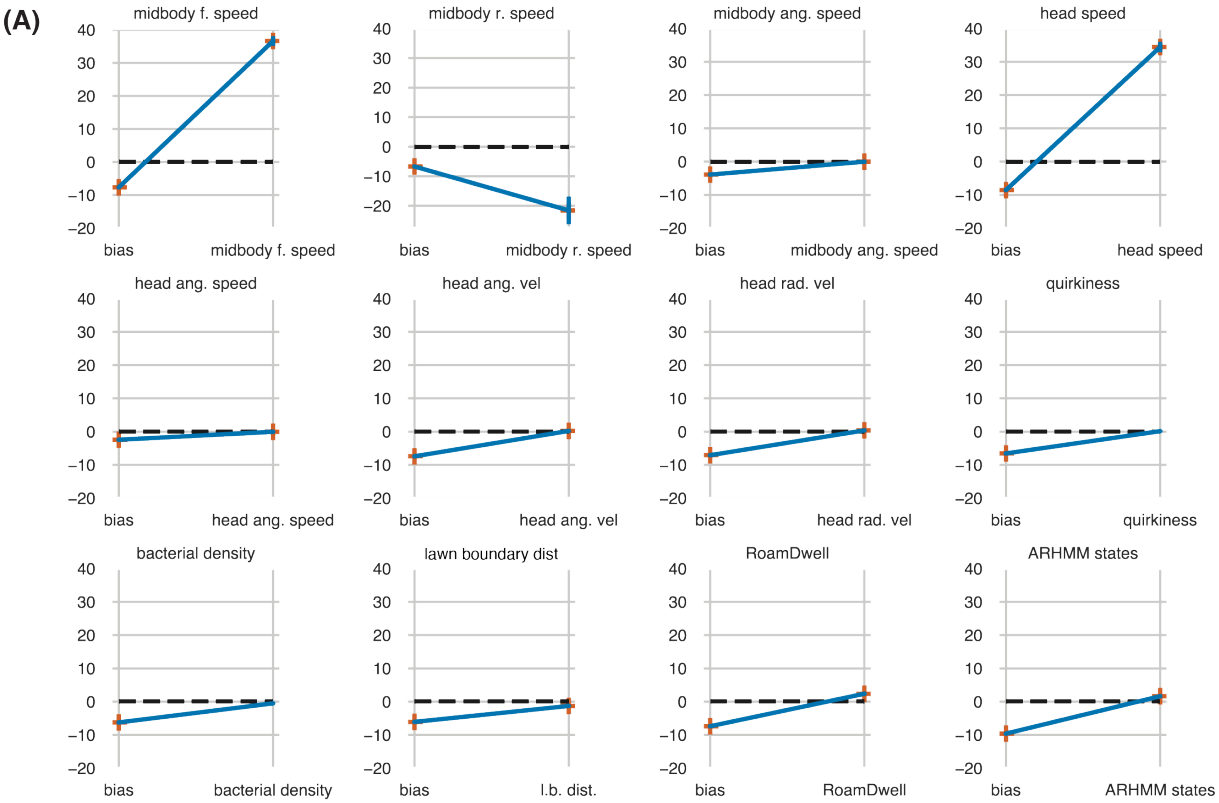
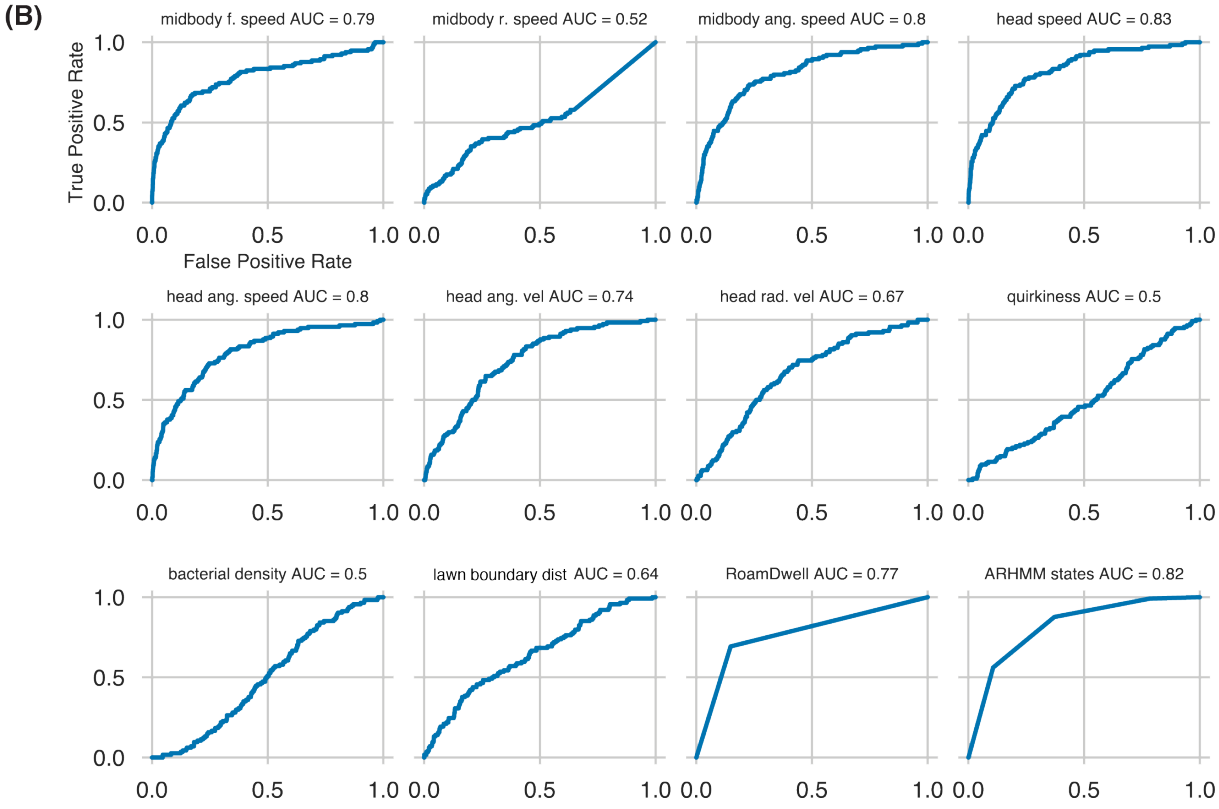


Figure 2-17. Generalized linear model based on locomotory and environmental variables predicts lawn exits with high accuracy.
(A) Generalized linear model (GLM) weight coefficients on behavioral features and HMM states for predicting lawn exits.
(B) Receiver operating characteristic (ROC) curve for GLM trained on behavioral features in (A) evaluated on a test dataset not used for training. Area under the curve (AUC) is 0.87 for predicting lawn leaving in this test dataset.

Figure 2-18. Generalized linear models based on speed-relate features predict lawn exits with high accuracy.
(A) Weight coefficients for GLMs trained with single behavioral features and a bias term to predict lawn exits.
(B) ROC curves for GLMs trained on single behavioral features with a bias term for predicting lawn leaving in a test dataset not used for training. AUC values are reported on each plot.

Figure 2-18


(A)
midbody f. speed
midbody r. speed
midbody ang. speed
head speed
head ang. speed
head ang. vel
head rad. vel
quirkiness
bacterial density
lawn boundary dist
RoamDwell
ARHMM states
bias
midbody f. speed
midbody r. speed
midbody ang. speed
head speed
head ang. speed
head ang. vel
head rad. vel
quirkiness
bacterial density
l.b. dist.
RoamDwell
ARHMM states


(B)
midbody f. speed AUC = 0.79
midbody r. speed AUC = 0.52
midbody ang. speed AUC = 0.8
head speed AUC = 0.83
head ang. speed AUC = 0.8
head ang. vel AUC = 0.74
head rad. vel AUC = 0.67
quirkiness AUC = 0.5
bacterial density AUC = 0.5
lawn boundary dist AUC = 0.64
RoamDwell AUC = 0.77
ARHMM states AUC = 0.82
True Positive Rate
False Positive Rate

Summary

Using high resolution video tracking across a large dataset of wild type animals, I measured *C. elegans* behavior on small lawns of bacterial food. These lawns show an uneven distribution of bacteria, with highest density found at the lawn boundary. Animals interacting with the lawn boundary exhibited several different behaviors including head pokes, in which the animal returns to the lawn, and lawn exits, in which the entire animal leaves the lawn. Using Hidden Markov Models to infer the boundaries between behavioral states, I found that lawn leaving is closely allied with highly aroused locomotion, whereas head pokes are found across all behavioral states. This correlation was strong enough that I could train a Generalized Linear Model to predict lawn leaving behavior on unseen data using movement speed alone as a predictor.

Lawn exit behavior appears to be a discrete foraging event that is a feature of locomotory arousal states. To understand more about how lawn exits come about, I asked how arousal states are generated, which I will explore in more detail in Chapter 4.

CHAPTER 3: DISSECTING THE PHYSIOLOGICAL DETERMINANTS OF LAWN LEAVING BEHAVIOR

Introduction

In environments where food is distributed in scattered patches, foraging animals continually face the question of how long to exploit local resources in one patch rather than explore elsewhere. There are many reasons for an animal to leave or remain in a food patch, including external appetitive and aversive cues, internal states of arousal, and nutritional status.

In *C. elegans,* patch leaving behavior is studied by recording behavior on small lawns of bacterial food, as described in Chapter 2. Even in these controlled environments, there are a range of salient internal and external stimuli that allow an animal to evaluate the quality of a food patch, thereby influencing its decision to leave the lawn. To a worm, the quality of a food lawn has several dimensions: the bacterial density or abundance, the ease of ingesting it, its nutritive value to the animal, and the presence or absence of appetitive or aversive compounds. These dimensions all influence *C. elegans* growth rate on the lawn, which constitute one summary metric of food quality (Avery and Shtonda, 2003).

Sensory information about food quality influences lawn leaving behavior acutely and on longer time scales. Difficulty eating food due to large bacterial cell size or mutations that cause the animal's pharynx (mouth parts) to lose coordination increases lawn leaving behavior (Shtonda, 2006). Animals also quickly leave lawns of bacteria spiked with aversive compounds (Pradel et al., 2007). By contrast, animals slowly increase their lawn leaving rates as food in the lawn is depleted (Milward et al., 2011), an effect that becomes more pronounced as new larvae hatch on the lawn and secrete pheromones that signal adult hermaphrodites to leave (Scott et al., 2017). Additionally, conditioning animals on high quality food and then transferring them to lower quality food elevates lawn leaving for several hours before a return to baseline rates, indicating that animals form a long-lasting memory of food quality that guides foraging (Shtonda, 2006).

Certain bacteria are pathogenic to *C. elegans*, necessitating different behavioral responses to those bacterial lawns. For example, worms are initially attracted to lawns of the opportunistic pathogen *Pseudomonas aeruginosa* PA14, but gradually leave the lawn over the course of hours as they become infected (Laws et al., 2006; Singh and Aballay, 2019). Lawn leaving on this long time scale involves neuroendocrine signaling from infected non-neuronal tissues to the brain to effect behavioral avoidance (McEwan et al., 2012; Melo and Ruvkun, 2012).

Although there are many ultimate reasons for an animal to leave food while foraging, the proximate causes often involve behavioral state changes driven by neuropeptide signaling. One example comes from studying lawn leaving behavior in *C. elegans*

males. Well-fed *C. elegans* males show increased probability of lawn leaving to search for hermaphrodite mates if none are present on the lawn. Loss of function mutations in the gene encoding the neuropeptide Pigment Dispersing Factor 1 (PDF-1) and its receptor PDFR-1 suppress lawn leaving in search of mates (Barrios et al., 2012). *pdf-1* and *pdfr-1* are expressed in many neurons, but only a subset of these neurons are required for mate search behavior: AIM interneurons that release PDF-1 and URY, PQR, and PHA sensory neurons that express the PDFR-1 receptor.

Another example of peptidergic control of lawn leaving comes from studying laboratory adaptation and natural trait variation of *C. elegans.* The laboratory strain N2 and the Hawaiian isolate CB4856 have different lawn leaving rates associated with two X-linked quantitative trait loci, corresponding to the genes encoding the neuropeptide Y-like receptor NPR-1 and the tyramine/octopamine receptor TYRA-3 (Bendesky et al., 2011). Strains that harbor low-activity alleles of *npr-1* and *tyra-3* exhibit high locomotion speed and high lawn leaving rates. *tyra-3* expression in ASK and BAG sensory neurons suppresses lawn leaving in the high-leaving Hawaiian wild isolate; NPR-1 signaling is discussed below.

NPR-1 signaling and behavioral arousal

The *npr-1* gene was first identified through a loss of function mutation in the N2 laboratory strain that caused aggregation into feeding groups, accumulation at the border of a bacterial lawn, and high speed on food. These behavioral differences arose from inactivation of the *npr-1* gene, which encodes a G protein coupled receptor. Natural isolates of *C. elegans* share these behavioral phenotypes with *npr-1(lf)* mutants, and remarkably bear a single amino acid difference from the laboratory strain (De Bono and Bargmann, 1998; McGrath et al., 2009).

The TAX-4 cyclic GMP-gated channel, which is essential for transduction in many sensory neurons, is required for the aggregation and high speed of *npr-1(lf)* animals, indicating that sensory activity is important in generating these behaviors (Coates and De Bono, 2002). The most important among the *tax-4* expressing neurons are PQR, AQR, and URX (Coates and De Bono, 2002). These neurons are sensors of high environmental oxygen, which they detect using a family of soluble guanylate cyclases including GCY-35 and GCY-36 (Cheung et al., 2005; Gray et al., 2004). In agreement with the hypothesis that oxygen drives aggregation and other *npr-1*-dependent behaviors, lowering oxygen concentrations or inactivating *gcy-35* also abolishes *npr-1(lf)* aggregation phenotypes. Accumulation at the lawn border appears to be a result of aerotaxis behavior: animals are attracted to thickest part of the bacterial lawn, where oxygen concentration is lowest due to bacterial metabolism. Aggregation also involves inputs from nociceptive neurons and pheromone sensing neurons (De Bono et al., 2002; Macosko et al., 2009).

The two RMG interneurons are the most important site of NPR-1 action (Macosko et al.,

2009). RMG neurons promote oxygen-dependent aggregation and high locomotion speed, and their activity is inhibited by NPR-1. The neuroanatomical wiring diagram shows that RMG interneurons form many gap junctions with oxygen-sensing neurons and other sensory neurons. Reducing the function of RMG gap junctions results in suppression of oxygen-regulated behavior (Jang et al., 2017).

Inhibiting RMG synaptic release with tetanus toxin also suppresses oxygen-related behaviors (Macosko et al., 2009). RMG is highly peptidergic so one hypothesis is that NPR-1 regulates release of neuropeptides affecting locomotory circuits. Indeed, NPR-1 signaling is required to promote behavioral quiescence during the larval molt between the last larval stage and adult—an effect which relies upon functional neuropeptide processing through the EGL-3 proprotein convertase enzyme (Choi et al., 2013). A likely mediator of this effect is the arousal neuropeptide PDF-1, whose secretion is upregulated in *npr-1(lf)* animals. Loss of function mutations in *pdf-1* and its receptor *pdfr-1* suppress *npr-1(lf)* hyperactivity during lethargus.

Results

Food availability controls locomotory arousal states and lawn leaving behavior

Given the well-established connection between food availability and behavioral state changes in *C. elegans*, I hypothesized that decreasing food abundance would lead to increased lawn leaving. To test this, I seeded bacterial lawns from bacterial cultures grown to OD1 or concentrated 2x or 4x to generate lawns designated OD1, OD2 and OD4 (OD2 is the standard bacterial density used throughout this work unless otherwise noted, See Methods for details). After acclimating animals to the experimental lawn density for 45 minutes, I measured animal locomotory behavior and lawn leaving at each of these lawn densities. Increasing lawn density significantly decreased the fraction of animals leaving the lawn during a 40-minute assay, as well as the number of lawn exits per animal (Figure 3-1 A,B). Similarly, using the analysis described in Chapter 2 showed that the fraction of time roaming and the mean midbody forward speed were also decreased with increasing food density (Figure 3-1 C,D).

Next, I asked whether acute inhibition of animals' food intake on bacterial lawns affected lawn leaving. For the experiment, I used a strain in which optogenetic activation of the pharyngeal muscle paralyzed it and prevented food intake (Rhoades et al., 2019). This strain transgenically expresses the red-shifted channelrhodopsin ReaChR (Lin et al., 2013) in the pharyngeal muscle used to grind and pump in food. Animals were acclimated to small lawns for 20 minutes before recording their behavior during alternating 10-minute intervals of light-induced pharyngeal paralysis (Figure 3-2 B). Feeding inhibition induced a dramatic increase in fraction of time spent roaming, 10% to 51% (Figure 3-2 C), as well as increased occupancy of AR-HMM state 3 and decreased occupancy of AR-HMM state 0 and state 1 (see Chapter 2) (Figure 3-2 D). Within one minute of turning the light on, animals accelerated from a baseline of 0.02 mm/sec to 0.06 mm/sec (Figure 3-2 E). Finally, the optogenetic stimulus resulted in a robust increase from 3% to 54% of animals leaving the lawn within 10 minutes (Figure 3-2 F, G). In contrast with the rapid increase in speed within two minutes of light onset, lawn leaving was distributed across the 10-minute light ON interval (Figure 3-2 G).

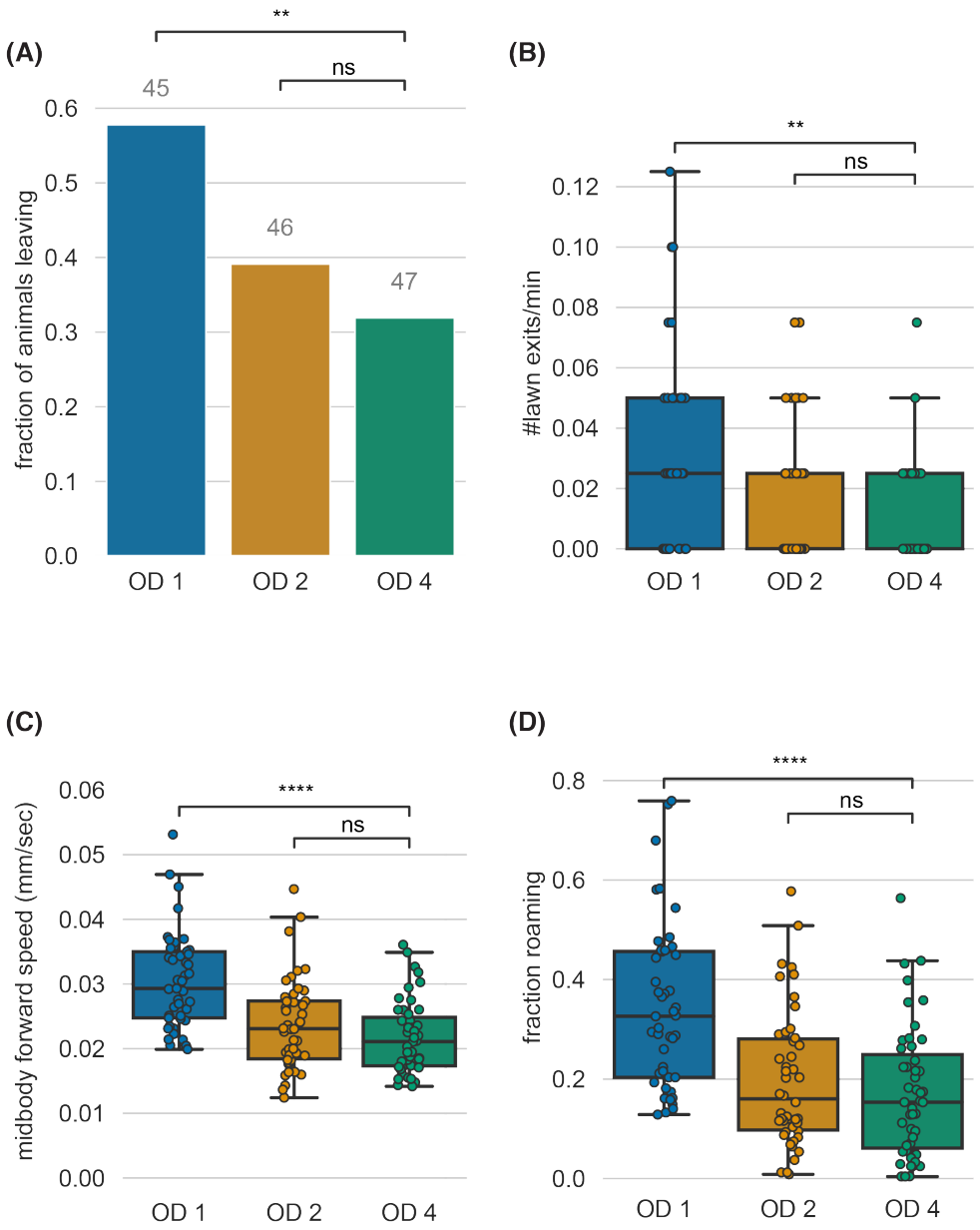


Figure 3-1. Increasing bacterial density suppresses locomotion speed and lawn leaving.

(A) Fraction of animals executing at least one lawn exit per 40-minute assay. Optical density of bacteria seeded for each assay is listed as OD1, OD2, OD4. Number of animals per condition shown in gray above bars applies to all figure panels.

(B) Lawn leaving rate measured in number of lawn exits per minute per animal.

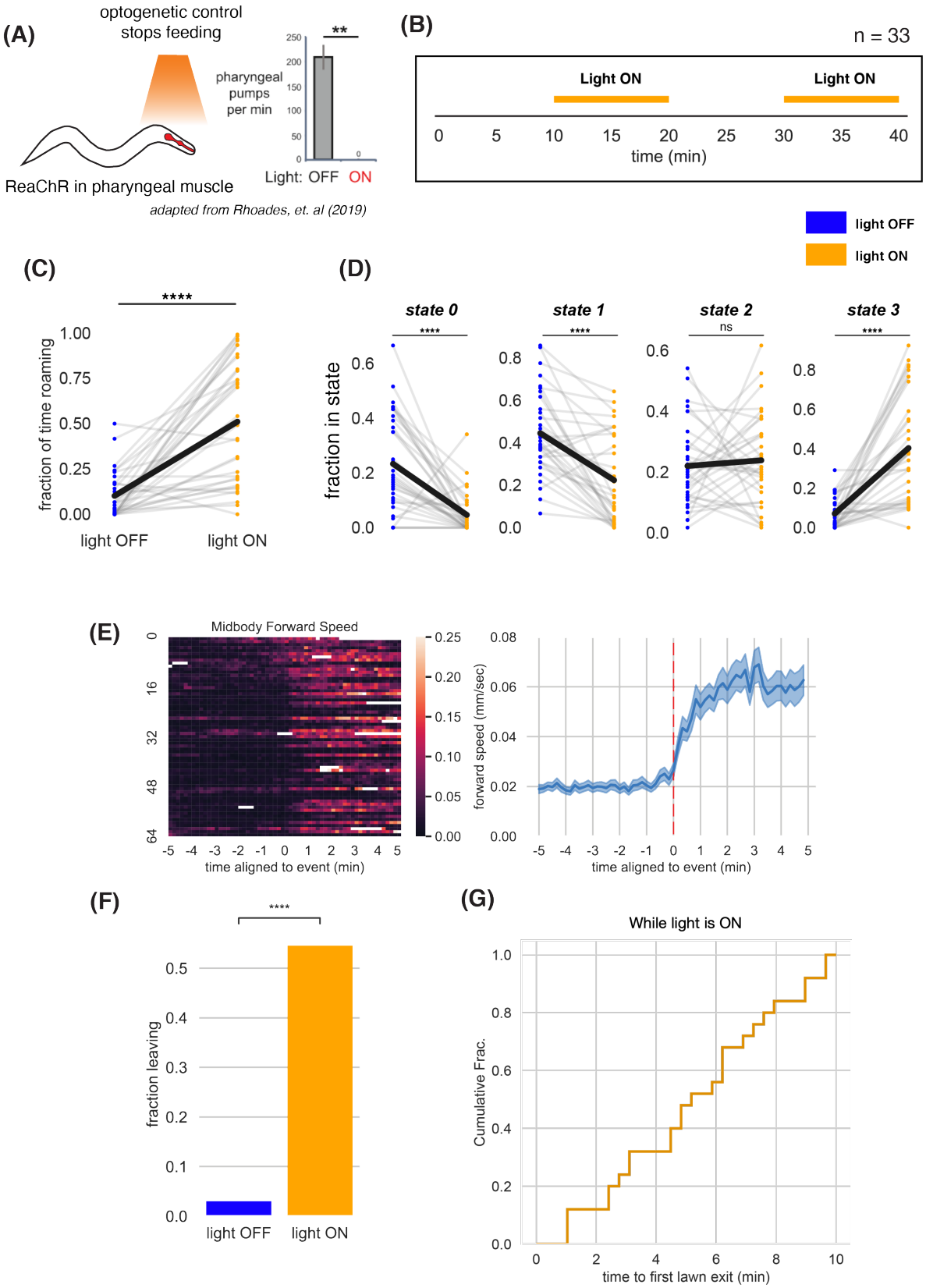
(C) Mean midbody forward speed per animal.

(D) Fraction of time spent roaming per animal.

For panels B-D, each dot represents one animal. Boxes indicate median and interquartile range. ** $p < 0.01$, **** $p < 10^{-4}$. In (A) Proportions z-test was used; in B-D Mann-Whitney two-sided test in B-D. All p values are Bonferroni corrected.

Figure 3-2. Acute feeding inhibition stimulates high speed states and lawn leaving.
(A) Optogenetic strategy to stop feeding using transgenic animals expressing ReaChR in pharyngeal muscle (*myo-2::ReaChR)*. In the presence of red-shifted light, pharyngeal pumping is suppressed. *Adapted and modified from Rhoades, et. al (2019).*
(B) Schematic for delivery of light pulses for two 10-minute intervals during a 40-minute assay.
(C) Paired comparison of fraction of time roaming between light OFF and light ON conditions. Each gray line represents fraction of roaming for a single animal. The average across all animals is plotted in black. **** $p < 10^{-4}$ by Wilcoxon signed rank test.
(D) Same as (C) but for each AR-HMM state in 4-state model derived in Chapter 2.
(E) Midbody forward speed aligned to light ON transitions at minute 10 and minute 30. Left, heatmap of individual traces. Right, average midbody forward speed is plotted in dark blue and standard error is shaded in light blue.
(F) Fraction of animals that execute at least one lawn leaving event over the two light OFF or light ON intervals. **** $p < 10^{-4}$ by proportions z-test.
(G) Cumulative distribution of time until the first lawn exit during light ON intervals.

Figure 3-2


(A)
optogenetic control
stops feeding
pharyngeal
pumps
per min
**
250
200
150
100
50
0
Light: OFF ON
ReaChR in pharyngeal muscle
adapted from Rhoades, et. al (2019)
(B)
n = 33
Light ON
Light ON
0 5 10 15 20 25 30 35 40
time (min)
light OFF
light ON
(C)
****
fraction of time roaming
1.00
0.75
0.50
0.25
0.00
light OFF
light ON
(D)
state 0
state 1
state 2
state 3
****
****
ns
****
fraction in state
0.6
0.4
0.2
0.0
0.8
0.6
0.4
0.2
0.0
0.6
0.4
0.2
0.0
0.8
0.6
0.4
0.2
0.0
(E)
Midbody Forward Speed
0
16
32
48
64
-5 -4 -3 -2 -1 0 1 2 3 4 5
time aligned to event (min)
0.25
0.20
0.15
0.10
0.05
0.00
forward speed (mm/sec)
0.08
0.06
0.04
0.02
0.00
-5 -4 -3 -2 -1 0 1 2 3 4 5
time aligned to event (min)
(F)
****
fraction leaving
0.5
0.4
0.3
0.2
0.1
0.0
light OFF
light ON
(G)
While light is ON
Cumulative Frac.
1.0
0.8
0.6
0.4
0.2
0.0
0 2 4 6 8 10
time to first lawn exit (min)

Genetics and neuronal functions that affect lawn leaving behavior

To investigate the genetic and neuronal mechanisms that generate arousal and lawn leaving behavior, I recorded behavior of 19 strains with mutations in sensory molecules or neuromodulators, or transgenic strains that ablated specific sensory neurons (Figure 3-3). As the values varied across days in wild type controls, all mutants were compared to wild type animals tested on the same day (Table 3-1, 3-2). Among the sensory mutants tested, *tax-4* and a strain in which the OLL neurons were genetically ablated both showed a significant increase in the fraction of animals leaving the lawn. These strains differed in their basal arousal state: *tax-4* mutants had a reduced fraction of time spent roaming, whereas OLL-ablated animals had a slightly higher roaming fraction (Figure 3-3 C).

Wild type animals increase their speed and change their behavioral state two minutes before lawn leaving (Chapter 2). *tax-4(lf)* mutant animals also increased their forward speed, but the speed ramp began less than one minute before lawn leaving (Figure 3-4 A). Similarly, *tax-4(lf)* mutants entered a roaming state for approximately one minute before lawn exit, compared to the 3 minute ramp in wild type animals (Figure 3-4 B, 2-15). In general, individual animals left the lawn no more than once per 40 minute assay and remained outside the lawn significantly longer than wild type animals before returning to the lawn (Figure 3-4 C,D). Together, these results suggest that *tax-4* expressing neurons affect transitions into the high arousal states associated with lawn leaving but preserve the local relationship between arousal and lawn leaving.

Among neuromodulatory mutants tested, mutants that roamed more tended also to leave lawns more readily (and vice versa) (Figure 3-3 B,D). However, after multiple hypothesis test correction, only *npr-1(ad609)* reached statistical significance for increased fraction leaving (Figure 3-3 B). *npr-1(lf)* mutants also left the lawn more times within a 40 minute assay than wild type (Figure 3-5 C). The time duration outside the lawn in each bout was of similar duration to wild type (Figure 3-5 D). *npr-1(lf)* animals had a higher basal forward speed and fraction of time spent roaming, and prior to lawn leaving events, speed and occupancy of the roaming state increased further (Figure 3-5 A,B). These results confirm the strong effects of *npr-1* on overall arousal and are consistent with an effect on lawn leaving that is caused by the high arousal state. By contrast, *tax-4* mutants leave more despite an overall decrease in arousal, consistent with an underlying sensory defect in assessing the quality or quantity of food.

Figure 3-3. Candidate genetic and neuronal inactivation reveals sensory and neuromodulatory control of roaming and lawn leaving.
(A) Fraction of animals that execute at least one lawn exit for different experimental groups targeting sensory neurons alongside wild type control data collected on the same day.
(B) Same as (A) for neuromodulatory mutants and their paired wild type controls.
(C) Fraction of time roaming per animal plotted for different experimental groups targeting sensory neurons alongside wild type control data collected on the same day.
(D) Same as (C) for neuromodulatory mutants and paired wild type controls.
Boxes indicate median and interquartile range. Asterisks indicate significant difference between 2 conditions using Mann-Whitney two-sided test before multiple hypothesis test correction. ns = no significant difference, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 10^{-4}$. Experimental groups shaded red and blue are significantly increased or decreased relative to wild type after Bonferroni multiple hypothesis test correction.

Figure 3-3

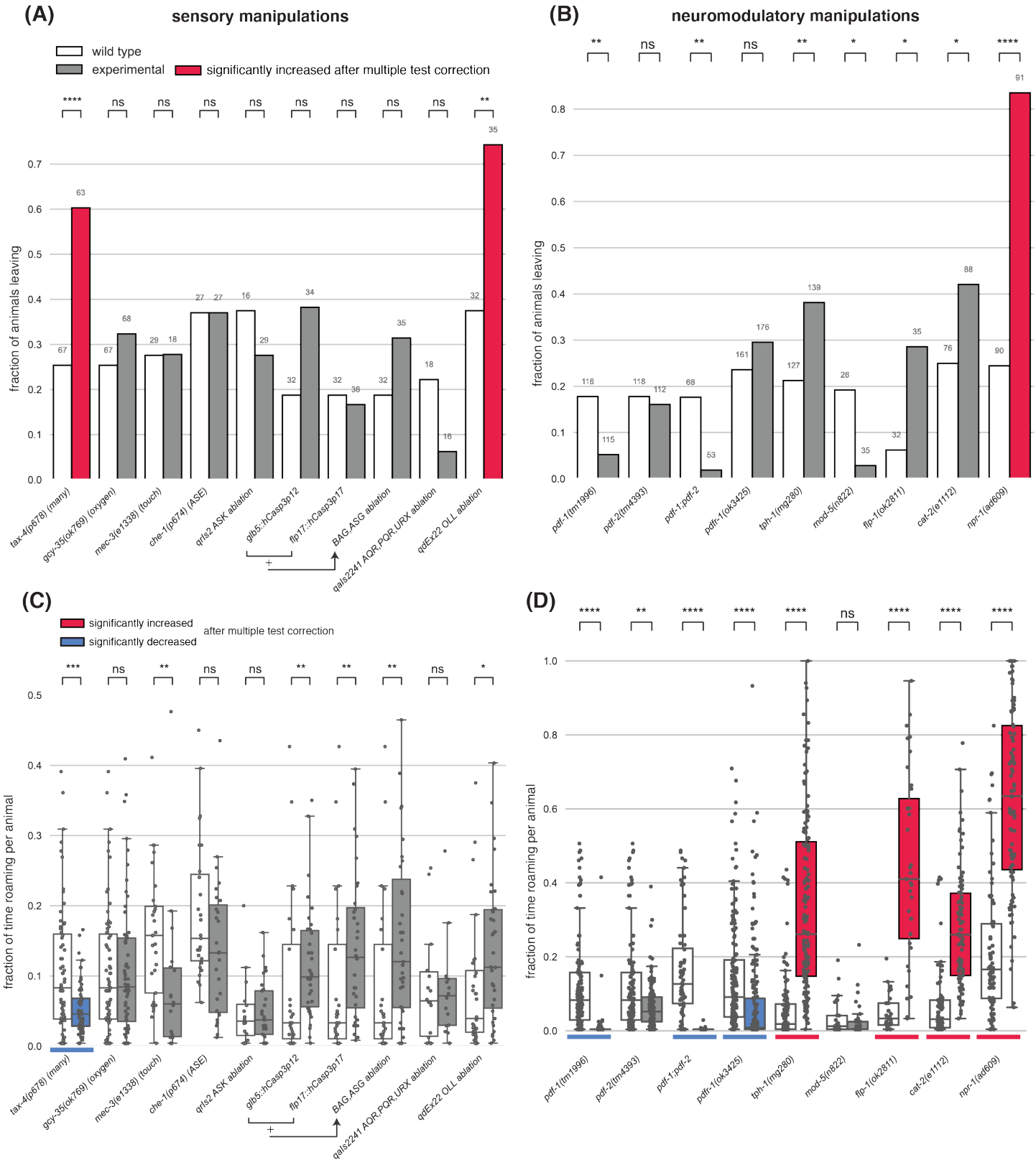

(A)
sensory manipulations
wild type
experimental
significantly increased after multiple test correction
fraction of animals leaving
tax-4(p678) (many)
gcy-35(ok769) (oxygen)
mec-3(e1338) (touch)
che-1(p674) (ASE)
qrfs2 ASK ablation
glb5::hCasp3p12
flp17::hCasp3p17
BAG,ASG ablation
qaIs2241 AQR,PQR,URX ablation
qdEx22 OLL ablation
(B)
neuromodulatory manipulations
fraction of animals leaving
pdf-1(tm1996)
pdf-2(tm4393)
pdf-1;pdf-2
pdfr-1(ok3425)
tph-1(mg280)
mod-5(n822)
flp-1(ok2811)
cat-2(e1112)
npr-1(ad609)
(C)
significantly increased
significantly decreased
after multiple test correction
fraction of time roaming per animal
(D)
fraction of time roaming per animal

Table 3-1. Quantification of Fraction of Animals Leaving.
Fraction of animals that execute at least one lawn exit for different experimental groups targeting sensory neurons and neuromodulators alongside wild type control data collected on the same day. Data from Figure 3-3 A-B. p values by Mann-Whitney two-sided test. Red indicates significant p values after Bonferroni multiple hypothesis test correction.

genotype	n	fraction of animals leaving	pvalue	adjusted pvalue
pdf-1(tm1996)	115	0.05	0.003	0.051
wild type	118	0.18		
pdf-2(tm4393)	112	0.16	0.727	1
wild type	118	0.18		
pdf-1;pdf-2	53	0.02	0.005	0.104
wild type	68	0.18		
pdfr-1(ok3425)	176	0.30	0.218	1
wild type	161	0.24		
tph-1(mg280)	139	0.38	0.003	0.052
wild type	127	0.21		
mod-5(n822)	35	0.03	0.034	0.640
wild type	26	0.19		
flp-1(ok2811)	35	0.29	0.017	0.329
wild type	32	0.06		
cat-2(e1112)	88	0.42	0.022	0.412
wild type	76	0.25		
npr-1(ad609)	91	0.84	1.53E-15	2.90E-14
wild type	90	0.24		
tax-4(p678) (many)	63	0.60	5.57E-05	0.001
wild type	67	0.25		
gcy-35(ok769) (oxygen)	68	0.32	0.371	1
wild type	67	0.25		
mec-3(e1338) (touch)	18	0.28	0.989	1
wild type	29	0.28		
che-1(p674) (ASE)	27	0.37	1	1
wild type	27	0.37		
qrls2 ASK ablation	29	0.28	0.492	1
wild type	16	0.38		
glb5::hCasp3p12	34	0.38	0.081	1
wild type	32	0.19		
flp17::hCasp3p17	36	0.17	0.822	1
wild type	32	0.19		
BAG,ASG ablation	35	0.31	0.234	1
wild type	32	0.19		
qaIs2241 AQR,PQR,URX ablation	16	0.06	0.189	1
wild type	18	0.22		
qdEx22 OLL ablation	35	0.74	0.002	0.046
wild type	32	0.38		

Table 3-2. Quantification of fraction of time roaming.
Mean fraction of time spent roaming for different experimental groups targeting sensory neurons and neuromodulators alongside wild type control data collected on the same day. Data from Figure 3-3 C-D. p values by Mann-Whitney two-sided test. Red, blue indicates significant increased and decreased p values after Bonferroni multiple hypothesis test correction.

genotype	n	fraction of time roaming	pvalue	adjusted pvalue
pdf-1(tm1996)	115	0.01	9.16E-33	1.74E-31
wild type	118	0.12		
pdf-2(tm4393)	112	0.07	0.001	0.025
wild type	118	0.12		
pdf-1;pdf-2	53	0.00	6.30601E-20	1.19814E-18
wild type	68	0.17		
pdfr-1(ok3425)	176	0.08	8.8902E-15	1.68914E-13
wild type	161	0.14		
tph-1(mg280)	139	0.34	1.88E-31	3.58E-30
wild type	127	0.06		
mod-5(n822)	35	0.03	0.117	1
wild type	26	0.04		
flp-1(ok2811)	35	0.44	5.20E-11	9.88E-10
wild type	32	0.05		
cat-2(e1112)	88	0.28	2.93E-19	5.56E-18
wild type	76	0.07		
npr-1(ad609)	91	0.62	2.34E-21	4.45E-20
wild type	90	0.21		
tax-4(p678) (many)	63	0.05	0.0002	0.0045
wild type	67	0.11		
gcy-35(ok769) (oxygen)	68	0.11	0.444	1
wild type	67	0.11		
mec-3(e1338) (touch)	18	0.09	0.002	0.042
wild type	29	0.16		
che-1(p674) (ASE)	27	0.14	0.030	0.575
wild type	27	0.19		
qrls2 ASK ablation	29	0.05	0.462	1
wild type	16	0.05		
glb5::hCasp3p12	34	0.12	0.004	0.077
wild type	32	0.12		
flp17::hCasp3p17	36	0.15	0.003	0.050
wild type	32	0.12		
BAG,ASG ablation	35	0.17	0.004	0.068
wild type	32	0.12		
qaIs2241 AQR,PQR,URX ablation	16	0.11	0.365	1
wild type	18	0.11		
qdEx22 OLL ablation	35	0.15	0.005	0.101
wild type	32	0.10		

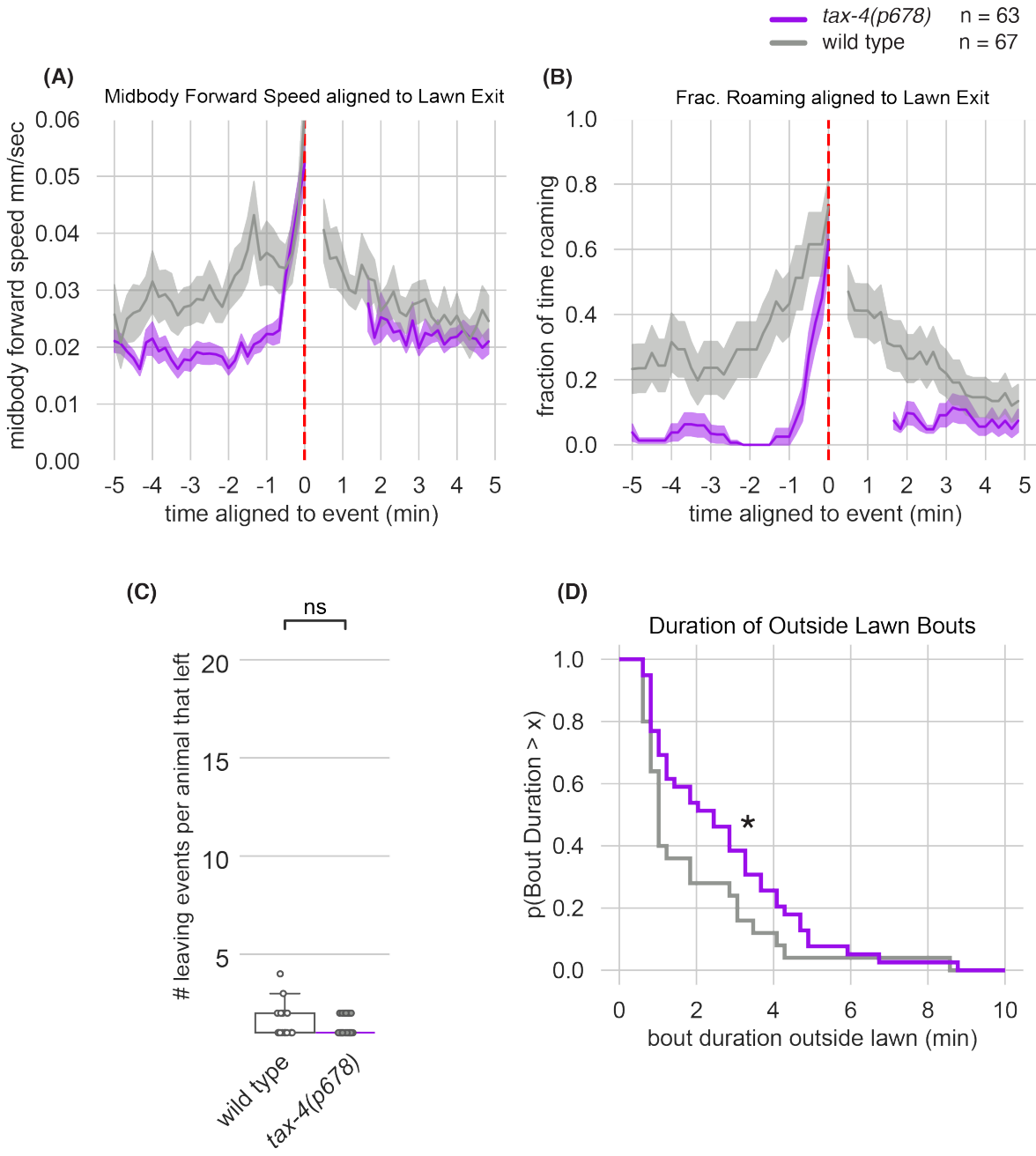


Figure 3-4. Loss of *tax-4* signaling increases lawn leaving and duration outside lawns.
(A) Midbody forward speed of wild type and *tax-4* mutants aligned to lawn exit. Dark line is mean across animals, shaded area is standard error.
(B) Fraction of time roaming per 10 second bin aligned to lawn exit. Dark line is mean across animals, shaded error is standard error.
(C) Number of lawn exits per animal that every left the lawn.
(D) Complementary cumulative distribution function (ccdf) for bout duration off food following lawn exits. * $p < 0.05$ by Kolmogorov-Smirnov two-sided test.

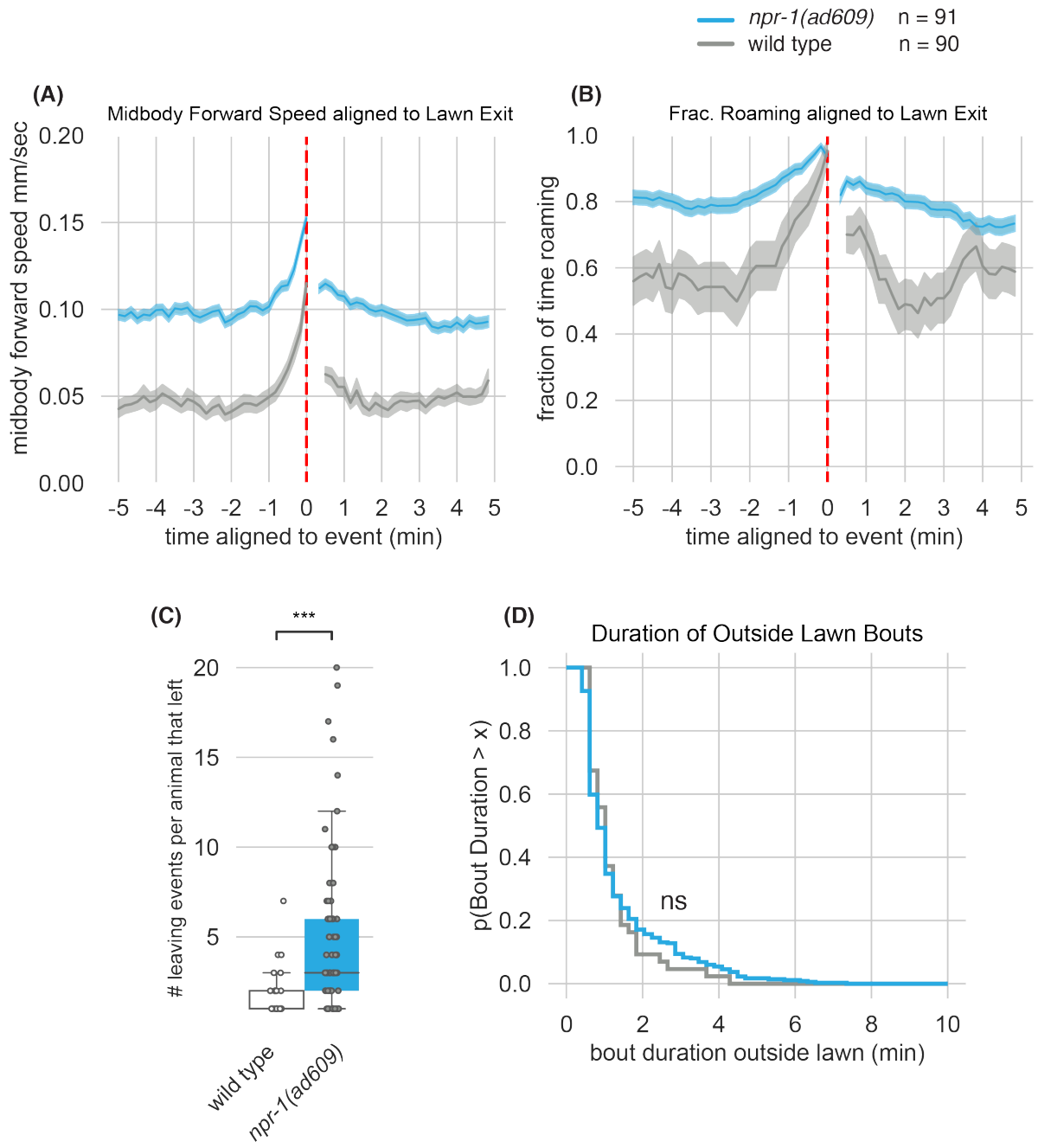


Figure 3-5. Loss of *npr-1* signaling increases roaming and lawn leaving rate.
(A) Midbody forward speed of wild type and *npr-1* mutants aligned to lawn exit. Dark line is mean across animals, shaded area is standard error.
(B) Fraction of time roaming per 10 second bin aligned to lawn exit. Dark line is mean across animals, shaded error is standard error.
(C) Number of lawn exits per animal that every left the lawn. *** $p < 0.001$ by Mann-Whitney two-sided test.
(D) Complementary cumulative distribution function (ccdf) for bout duration off food following lawn exits. ns = not significant by Kolmogorov-Smirnov two-sided test.

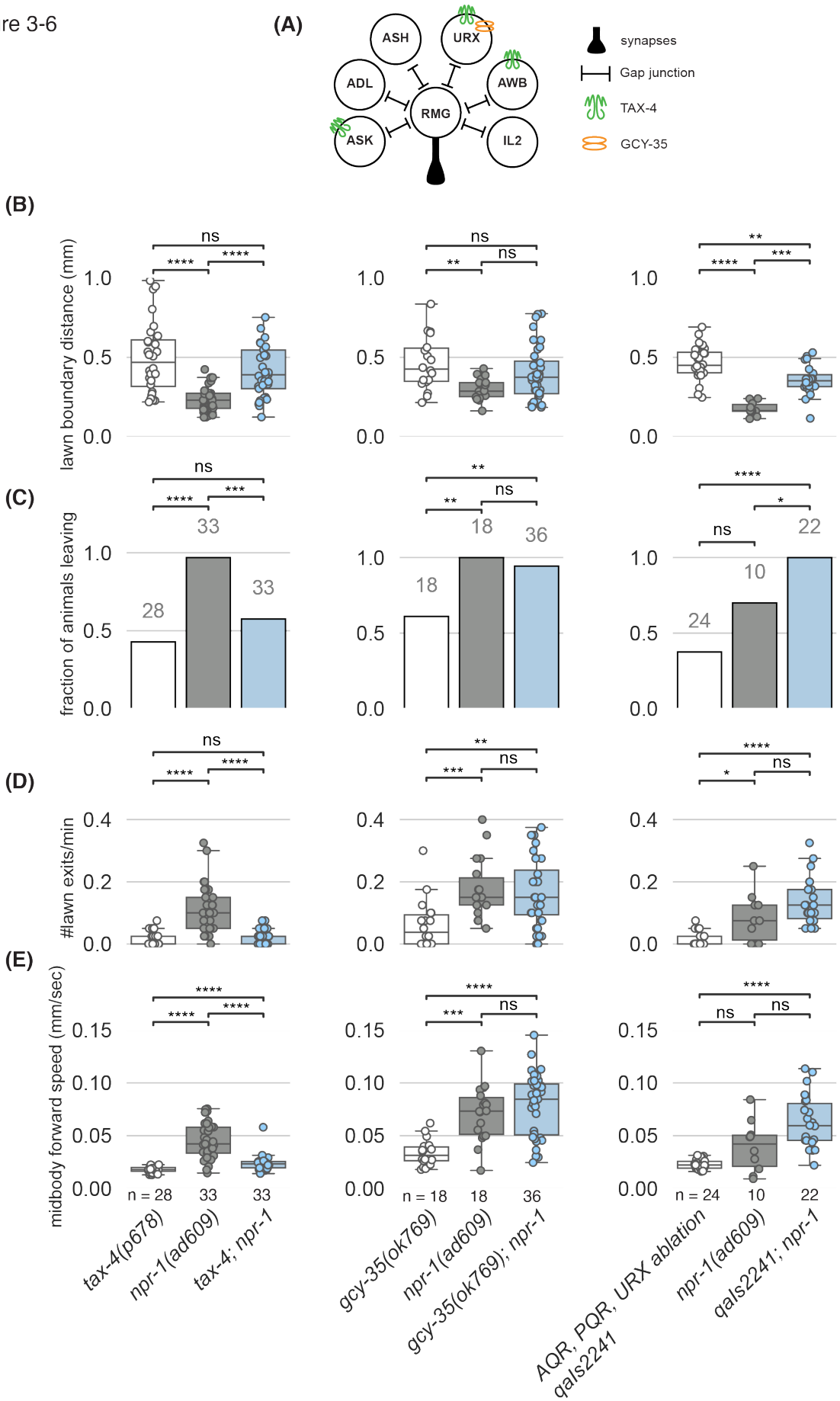
TAX-4 signaling is required for *npr-1(lf)* high speed and lawn leaving.

The activity of oxygen-sensing neurons is required for *npr-1(lf)* aggregation and bordering behaviors, as is the *tax-4* sensory transduction channel (see Introduction). Consistent with this, animals lacking functional TAX-4, GCY-35, or with oxygen-sensing neurons AQR, PQR, and URX genetically ablated, all exhibited more dispersal toward the lawn interior relative to *npr-1(lf)* animals as measured by increased lawn boundary distance. The lawn boundary distance of *tax-4*; *npr-1* mutants was indistinguishable from that of *tax-4* alone, whereas *gcy-35; npr-1* and oxygen neuron-ablations in the *npr-1(lf)* background yielded lawn boundary distances intermediate between single mutants (Figure 3-6 B). These results suggest that oxygen sensation and TAX-4 sensory transduction promotes the bordering of *npr-1(lf)* animals even in single animals on small lawns of bacteria.

I next asked whether inactivating *tax-4* or oxygen-sensing neurons would have a similar suppressive effect on arousal and lawn-leaving in *npr-1(lf)* animals. Although a higher fraction of *tax-4(lf)* mutant animals left lawns than wild type (Figure 3-3 A), inactivating *tax-4* in *npr-1(lf)* animals reduced the fraction of animals leaving and lawn leaving rates to *tax-4(lf)* levels (Figure 3-6 C,D). Midbody forward speed of *tax-4; npr-1* animals was intermediate between the speeds of single mutants (Figure 3-6 E). By comparison, inactivating oxygen-sensing either by loss of function of *gcy-35* or by genetic ablation of oxygen-sensing neurons AQR, PQR, and URX in *npr-1(lf)* animals did not suppress its high lawn leaving nor its high midbody forward speed (Figure 3-6 C-E). Together, these results suggest that TAX-4 signaling is required for *npr-1*'s high lawn leaving behavior but oxygen-sensation is not.

Figure 3-6. TAX-4 signaling but not hyperoxia-sensing is required for *npr-1(lf)* high speed and lawn leaving.
(A) Circuit diagram of neurons with gap junctions to RMG neurons. Neurons expressing *tax-4* and *gcy-35* are indicated.
(B-E) Three datasets, each comparing genotypes collected on the same day, are plotted left to right. Left column: *tax-4, npr-1, tax-4; npr-1*. Middle column: *gcy-35, npr-1, gcy-35; npr-1*. Right column: *qaIs2241* oxygen-sensing neuron AQR, PQR, URX genetic ablation, *npr-1, qaIs2241; npr-1*.
(B) Mean lawn boundary distance per animal. p value by Mann-Whitney 2-sided test.
(C) Fraction of animals that execute at least one lawn exit in each genotype. Gray numbers indicate number of animals in each group. p value by proportions z test.
(D) Lawn exit rate is shown in lawn exits per minute. p value by Mann-Whitney 2-sided test.
(E) Mean midbody forward speed per animal. p value by Mann-Whitney 2-sided test. Each dot represents one animal. Boxes indicate median and interquartile range.
* $p <= 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 10^{-4}$, ns = not significant. All p values are Bonferroni corrected.

Figure 3-6


(A)
ASH
URX
ADL
AWB
RMG
ASK
IL2
synapses
Gap junction
TAX-4
GCY-35
(B)
lawn boundary distance (mm)
(C)
fraction of animals leaving
(D)
#lawn exits/min
(E)
midbody forward speed (mm/sec)
n = 28
33
33
n = 18
18
36
n = 24
10
22
tax-4(p678)
npr-1(ad609)
tax-4; npr-1
gcy-35(ok769)
npr-1(ad609)
gcy-35(ok769); npr-1
AQR, PQR, URX ablation
qaIs2241
npr-1(ad609)
qaIs2241; npr-1

NPR-1 acts in RMG neurons to modulate lawn leaving

The aggregation and arousal phenotypes of *npr-1(lf)* animals require high activity in the RMG interneurons. Using a previously used intersectional approach, I reconstituted NPR-1 cDNA expression solely in RMG interneurons in the *npr-1(ad609)* background (Macosko et al., 2009), and observed a reduction in locomotion speed, fraction of time roaming, and fraction of animals leaving the lawn (Figure 3-7). Blocking RMG synaptic output by expressing tetanus toxin light chain in RMG also resulted in a reduction in locomotion speed, fraction of time roaming, and fraction of animals leaving, suggesting that release of a neurotransmitter or neuropeptide from RMG stimulates lawn leaving (Figure 3-8). RMG expresses many neuropeptides but is not known to express classical neurotransmitters, so I asked whether *npr-1(lf)* mutants release neuropeptides that promote lawn leaving. Neuropeptides are derived from precursor molecules that undergo endoproteolytic cleavage by proprotein convertase enzymes (Figure 3-9 A) (Kass et al., 2001; Thacker and Rose, 2000). Of the four convertase enzymes in *C. elegans,* RMG expresses only *egl-3* at significant levels (Figure 3-9 B) (Taylor et al., 2021). To generate an RMG-specific knockout of *egl-3*, I used an intersectional floxed *egl-3* locus (a gift of D. Colón-Ramos, unpublished) and the promoters of the RMG-selective genes *flp-5* and *nlp-56* to drive Cre recombinase (Figure 3-9 C) (Jang et al., 2017; Taylor et al., 2021). If *egl-3*-dependent peptide were the functional output of RMG, I expected a reduction in forward speed or leaving by the RMG-selective knockouts. This did not happen; instead, there was a slight by statistically significant increase in lawn-leaving and locomotion speed upon *egl-3* knockout in RMG (Figure 3-9 D,E).

A second step in the processing of some neuropeptides is mediated by EGL-21, a carboxypeptidase E enzyme that removes basic residues from pro-peptides (Figure 3-9 A) (Jacob and Kaplan, 2003). *egl-21(lf)* mutations suppress *npr-1(lf)* aggregation behavior, and RMG-specific *egl-21* knockdown reduces locomotory responses to changes in oxygen concentration (Laurent et al., 2015; Macosko, 2009). A floxed *egl-21* transgene was combined with expression of Cre recombinase in the RMG neurons to generate a conditional knockout of *egl-21* (Figure 3-10 A). In these experiments, I compared lawn leaving and arousal states of the single copy *egl-21* insertion (*kySi61)* to its genetic background *egl-21; npr-1* as well as 2 strains generated by crossing in an RMG-specific Cre recombinase. *kySi61* significantly increased speed and trended toward higher lawn exit rate over *egl-21; npr-1* levels, although the increase in lawn exit rate was not statistically significant. Specific inactivation of *egl-21* in RMG neurons did not alter lawn leaving and locomotion speed in the strains bearing the *egl-21* single copy insertion (Figure 3-10 B,C). These results suggest that EGL-3-dependent neuropeptides from RMG are not required for lawn leaving and high speed of *npr-1(lf)* mutants. To rationalize the different results with RMG tetanus toxin and RMG *egl-3* knockout, I suggest two hypotheses: First, some peptides do not require EGL-3 for processing; one or more of these peptides could be the RMG transmitter affected by tetanus toxin (for example, an insulin-like peptide). Second, EGL-3 protein could persist after deletion of the gene in RMG if low levels are sufficient for activity.

Figure 3-7. Restoring *npr-1* expression in RMG neurons rescues *npr-1(lf)* hyper-leaving phenotype.
(A) Intersectional Cre/Lox strategy to express *npr-1* specifically in RMG. *Adapted from Macosko, et al (2009).*
(B) Circuit diagram of neurons with gap junctions to RMG neurons. Neurons expressing *npr-1* are indicated.
(C) Fraction of *npr-1(ad609)* animals carrying *ncs-1::nCre* and/or *flp-21::LoxStopLox::npr-1* transgenes that execute at least one lawn exit. Gray numbers indicate number of animals in each group. p value by proportions z test.
(D) Same as (C) for lawn exit rate.
(E) Same as (C) for mean midbody forward speed per animal.
(F) Same as (C) for fraction of time roaming per animal.
Each dot represents one animal. Boxes indicate median and interquartile range.
For panels D-F, Kruskal-Wallis with Dunn's multiple comparison correction was performed. * $p <= 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 10^{-4}$, ns = not significant.

Figure 3-7

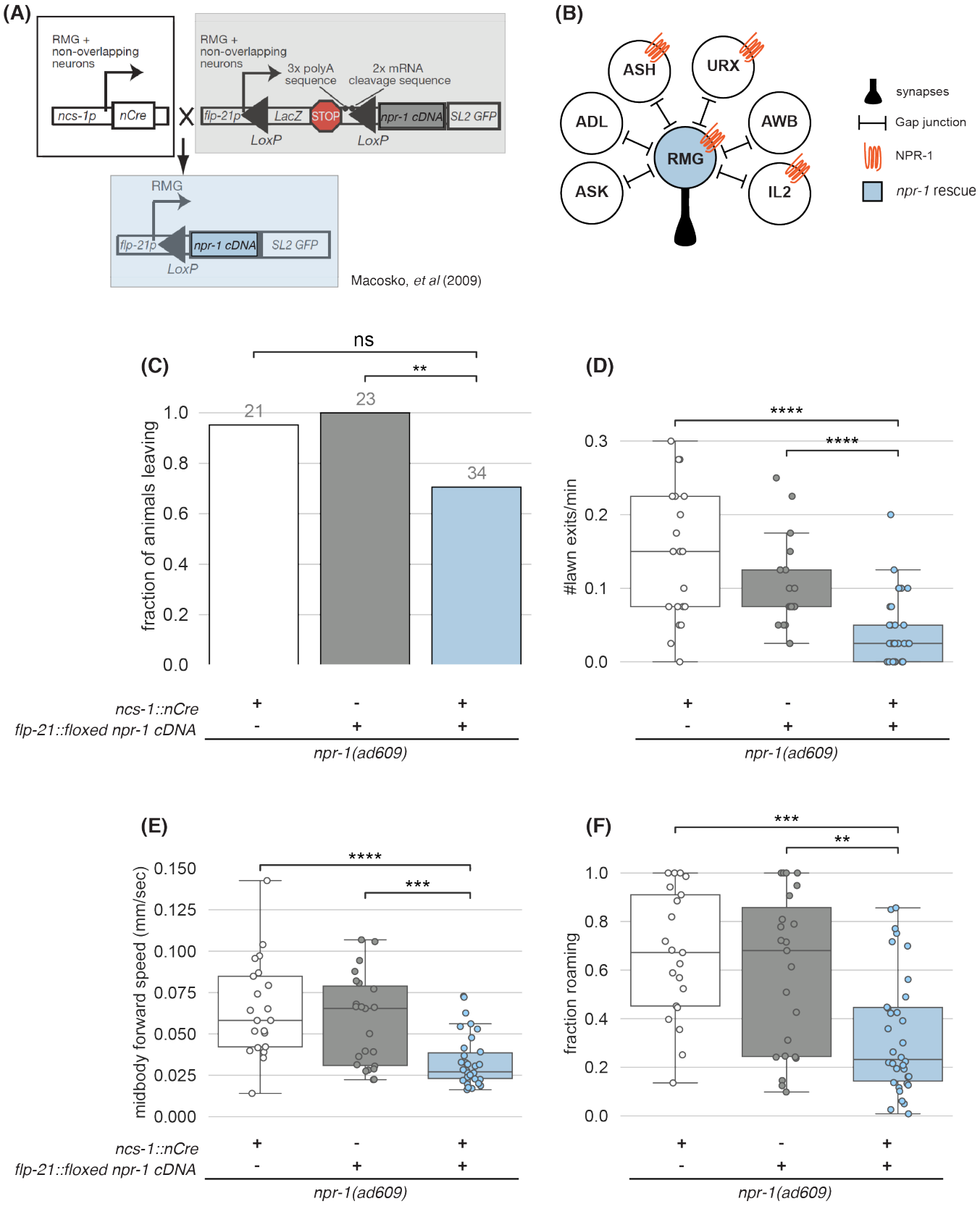

(A)
RMG + non-overlapping neurons
ncs-1p
nCre
RMG + non-overlapping neurons
3x polyA sequence
2x mRNA cleavage sequence
flp-21p
LacZ
STOP
npr-1 cDNA
SL2 GFP
LoxP
LoxP
RMG
flp-21p
npr-1 cDNA
SL2 GFP
LoxP
Macosko, et al (2009)
(B)
ASH
URX
ADL
AWB
RMG
ASK
IL2
synapses
Gap junction
NPR-1
npr-1 rescue
(C)
ns
**
21
23
34
fraction of animals leaving
1.0
0.8
0.6
0.4
0.2
0.0
ncs-1::nCre
flp-21::floxed npr-1 cDNA
+
-
+
-
+
+
npr-1(ad609)
(D)
****
****
#lawn exits/min
0.3
0.2
0.1
0.0
+
-
+
-
+
+
npr-1(ad609)
(E)
****
***
midbody forward speed (mm/sec)
0.150
0.125
0.100
0.075
0.050
0.025
0.000
ncs-1::nCre
flp-21::floxed npr-1 cDNA
+
-
+
-
+
+
npr-1(ad609)
(F)
***
**
fraction roaming
1.0
0.8
0.6
0.4
0.2
0.0
+
-
+
-
+
+
npr-1(ad609)

Figure 3-8. RMG synaptic release is required for *npr-1(lf)* hyper-leaving.
(A) Intersectional Cre/Lox strategy to express tetanus toxin light chain (TeTx) specifically in RMG. *Adapted from Macosko, et al (2009).*
(B) Circuit diagram of neurons with gap junctions to RMG neurons. Neurons expressing *npr-1* are indicated.
(C) Fraction of *npr-1(ad609)* animals carrying *ncs-1::nCre* and/or *flp-21::LoxStopLox::TeTx* transgenes that execute at least one lawn exit. Gray numbers indicate number of animals in each group. p value by proportions z test.
(D) Same as (C) for lawn exit rate.
(E) Same as (C) for mean midbody forward speed per animal.
(F) Same as (C) for fraction of time roaming per animal.
Each dot represents one animal. Boxes indicate median and interquartile range.
For panels D-F, Kruskal-Wallis with Dunn's multiple comparison correction was performed. * $p <= 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 10^{-4}$, ns = not significant.

Figure 3-8

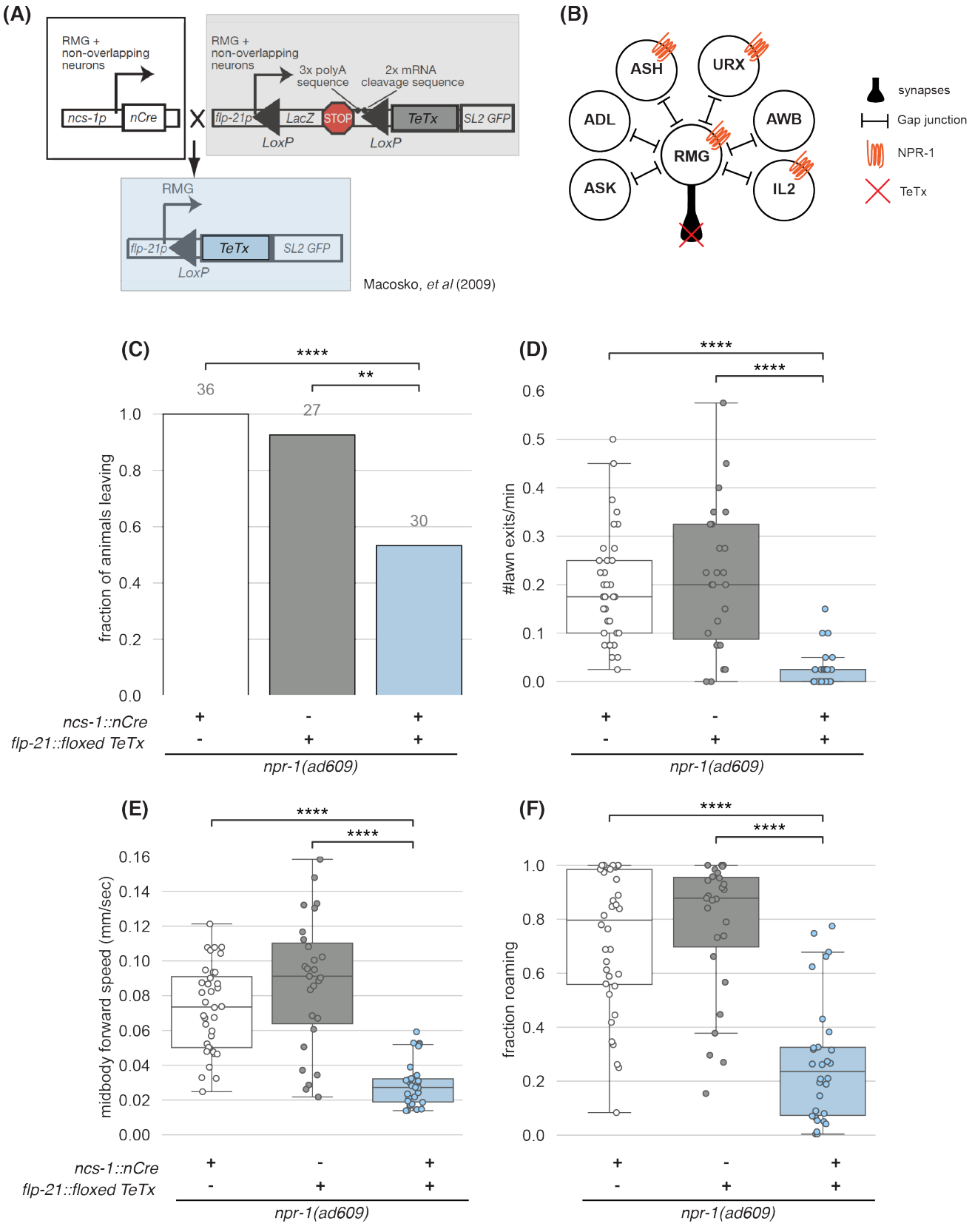
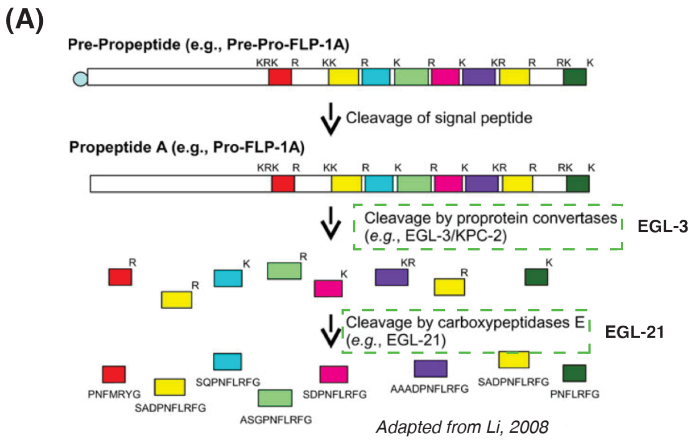

(A)
RMG + non-overlapping neurons
ncs-1p
nCre
RMG + non-overlapping neurons
3x polyA sequence
2x mRNA cleavage sequence
flp-21p
LacZ
STOP
TeTx
SL2 GFP
LoxP
LoxP
RMG
flp-21p
TeTx
SL2 GFP
LoxP
Macosko, et al (2009)
(B)
ASH
URX
ADL
AWB
RMG
ASK
IL2
synapses
Gap junction
NPR-1
TeTx
(C)
fraction of animals leaving
36
27
30
****
**
ncs-1::nCre
flp-21::floxed TeTx
npr-1(ad609)
(D)
#lawn exits/min
****
****
(E)
midbody forward speed (mm/sec)
****
****
(F)
fraction roaming
****
****

Figure 3-9. EGL-3-dependent neuropeptides are not required in RMG for *npr-1(lf)* hyper leaving.
(A) "Processing of a neuropeptide gene product: *flp-1* as an example. After translation of the flp-1A transcript, pre-pro-FLP-1A is cleaved by signal peptidase in the endoplasmic reticulum to release the signal peptide. The propeptide pro-FLP-1A is further cleaved C-terminal to mono-, di-, or tribasic residues (indicated by K and R) by proprotein convertases, such as EGL-3/KPC-2. The basic amino acids are removed by carboxypeptidases E, such as EGL-21, to yield the basic neuropeptides." *Quoted from* (*Li, 2008).* EGL-3 and EGL-21 steps highlighted with green dotted line by me.
(B) Single-cell expression data in transcripts per million, TPM, for neuropeptide processing genes in RMG neurons. Data from (Taylor et al., 2021).
(C) Genetic strategy for conditional knockout of *egl-3*. CRISPR-edited endogenous *egl-3* gene locus contains loxP sites flanking the ORF. Transgenic expression of Cre recombinase excises the ORF and turns on mCherry.
(D) Lawn exit rate across *npr-1(ad609)* animals bearing the floxed *egl-3* locus and two transgenes expressing Cre under RMG-selective promoters.
(E) same as (D) for mean midbody speed.
Each dot represents one animal. Boxes indicate median and interquartile range.
For panels D-E, Kruskal-Wallis with Dunn's multiple comparison correction was performed. * $p <= 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 10^{-4}$, ns = not significant.

Figure 3-9

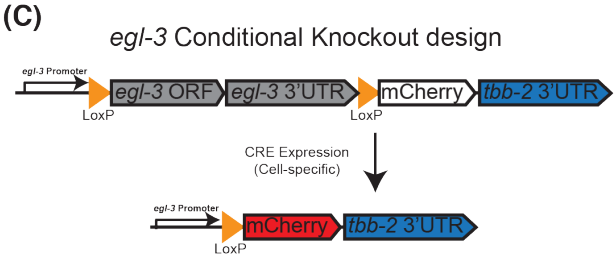


Adapted from Li, 2008

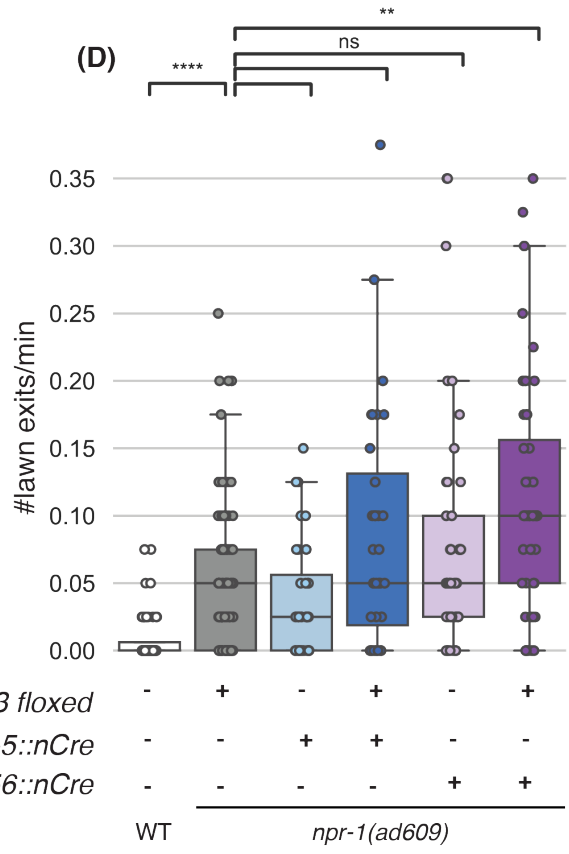
(B)

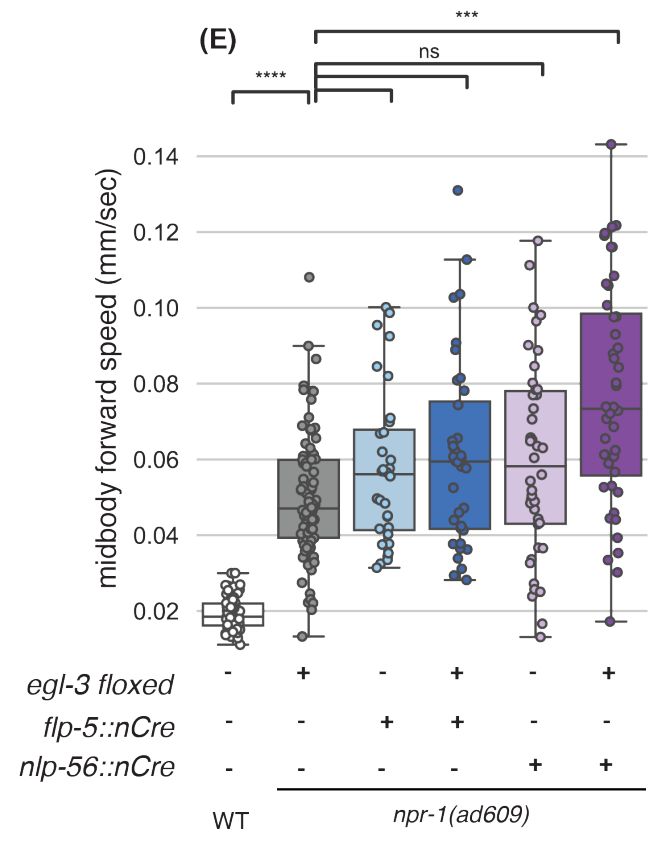
type	Gene name	RMG Expression
proprotein convertase	*egl-3*	1894.423
proprotein convertase	*kpc-1*	0
proprotein convertase	*aex-5*	0
proprotein convertase	*bli-4*	0
carboxypeptidase	*egl-21*	5412.994

Data from Taylor, et. al. (2021)



reproduced with permission from J. Marquina





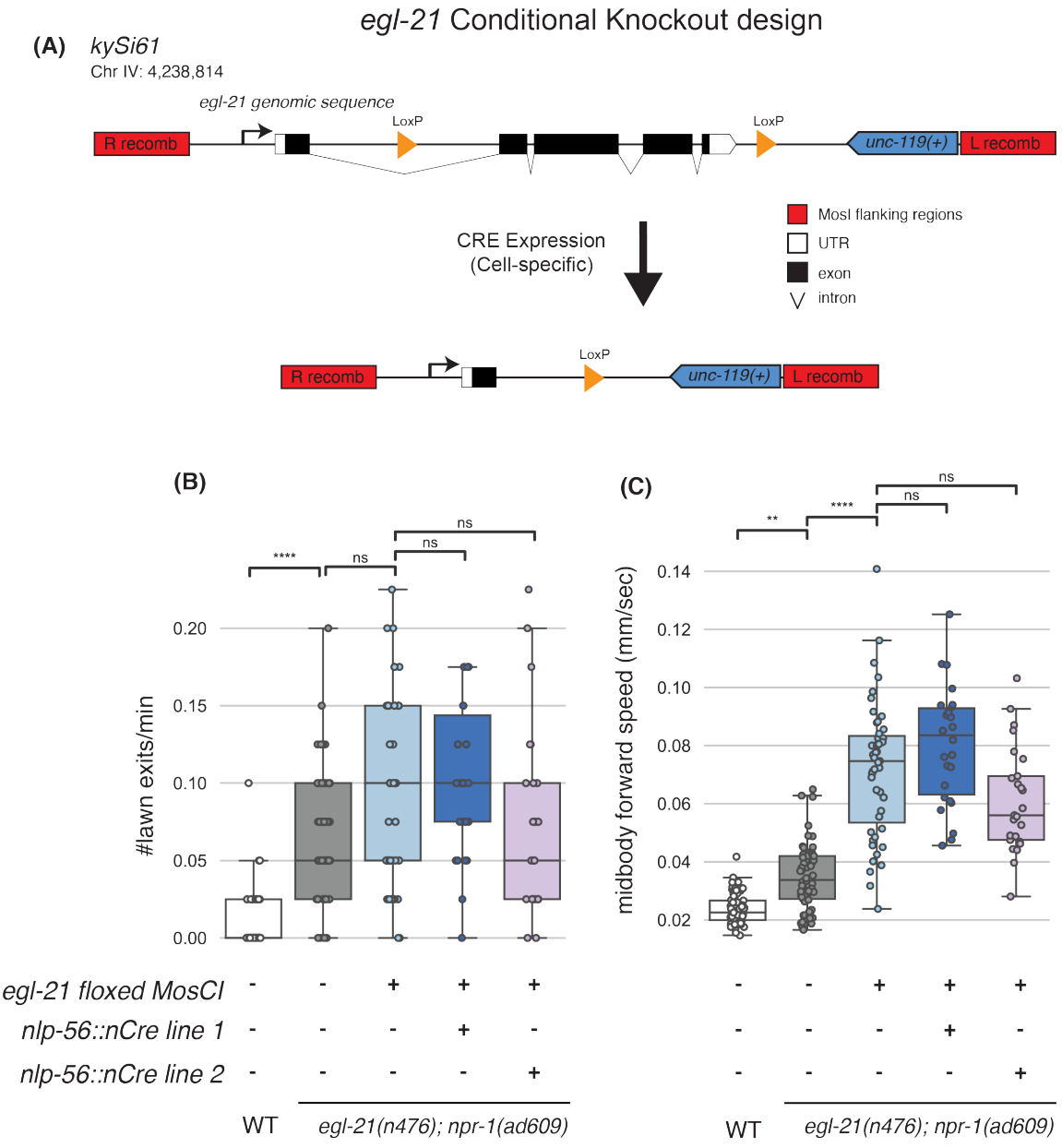


Figure 3-10. EGL-21-dependent neuropeptides are not required in RMG for *npr-1(lf)* hyper-leaving.

(A) Genetic strategy for conditional knockout of *egl-21*. *egl-21* gene locus flanked by loxP sites is integrated as a single-copy insertion on Chr IV. Transgenic expression of Cre recombinase excises the ORF.

(B) Lawn exit rate in *egl-21; npr-1* animals carrying *egl-21* MosCI insertion and/or transgenes expressing Cre in RMG-specific promoters.

(C) same as (B) for mean midbody speed.

Each dot represents one animal. Boxes indicate median and interquartile range. For panels B-C, Kruskal-Wallis with Dunn's multiple comparison correction was performed. * $p <= 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 10^{-4}$, ns = not significant.

PDFR-1 signaling in multiple neurons can increase forward speed and lawn leaving

Mutations of the neuropeptide *pdf-1* suppress *npr-1(lf)* hyper-arousal during lethargus (Choi et al., 2013). Similarly, introducing a *pdf-1* null mutation into *npr-1(lf)* animals significantly reduced the fraction of animals leaving, the lawn exit rate, forward speed, roaming and AR-HMM state 3. However, in no case were these metrics fully suppressed to the level of *pdf-1(lf)* alone (Figure 3-11). The single characterized receptor for *pdf-1* is *pdfr-1*. Introducing a *pdfr-1* null mutant into *npr-1(lf)* animals reduced lawn exit rate, speed, and roaming compared to *npr-1(lf)*, but the fraction of animals leaving and occupancy of AR-HMM state 3 were not decreased (Figure 3-12). *pdf-1* is expressed in many neurons, including the RMG neurons (Choi et al., 2013; Taylor et al., 2021). A transgene driving RMG-specific *pdf-1* expression in *pdf-1; npr-1* mutants partially rescued speed, roaming and lawn leaving rate in one out of two lines tested (Figure 3-13). Thus, RMG may act as a source of PDF-1 to drive *npr-1(lf)* lawn leaving and other arousal behaviors, but other neurons may also produce PDF-1 to modify these behaviors.

To define the sites of *pdfr-1* expression required for *npr-1(lf)* lawn leaving and high locomotion speed, I used a previously characterized intersectional system that restores *pdfr-1* expression in targeted groups of cells in *pdfr-1; npr-1* mutant animals (Figure 3-14 A) (Flavell et al., 2013). This method uses two transgenes: one that expresses inverted, floxed *pdfr-1* cDNA under the *pdfr-1* distal promoter, and another to drive Cre using a promoter with characterized cell-specific expression. Since *pdfr-1* is expressed in many neurons and in muscle, I first activated it in large groups of neurons. Restoring *pdfr-1* expression under a pan-neuronal promoter increased lawn leaving rate and forward speed beyond the basal *npr-1(lf)* levels (Figure 3-14 B). While this result is promising, it suggests that the transgene does not perfectly reflect endogenous *pdfr-1* function. Possible explanations for this include overexpression from the transgene, the use of a single *pdfr-1* cDNA when the endogenous gene has multiple splice forms, or a partial expression pattern as the transgene does not include all regulatory sequences. Nonetheless, I continued with the strain because it had robust function in my assay.

I next examined subsets of neurons for their rescuing activity. In *pdfr-1; npr-1* double mutants, *pdfr-1* expression in *pdfr-1*-expressing glutamatergic neurons yielded partial rescue of fraction of animals leaving, lawn leaving rate, and speed (Figure 3-14 C). In *pdfr-1; npr-1* double mutants, expressing *pdfr-1* in *pdfr-1*-expressing cholinergic neurons partially rescued speed, but did not rescue fraction of animals leaving or leaving rate (Figure 3-14 D). This suggests that multiple *pdfr-1*-expressing cells contribute to speed and lawn leaving in *npr-1(lf)* animals. Intersectional expression of *pdfr-1* using the *tdc-1* promoter, which reconstitutes expression in the RIM neurons, increased lawn leaving fraction and lawn leaving rate, but not locomotion speed, the reciprocal result to cholinergic expression (Figure 3-15 A). Intersectional expression of *pdfr-1* using the *mod-1* promoter, which is expressed in AIY and RIB, significantly

increased fraction leaving, leaving rate, and speed (Figure 3-15 B). Intersectional expression of *pdfr-1* using the *sto-3* promoter, which directs expression exclusively to RIB neurons (Figure 3-16), robustly increased speed with only a small increase in lawn leaving rate (Figure 3-15 C). Intersectional expression of *pdfr-1* in sensory neurons did not restore lawn leaving, speed, roaming of AR-HMM state 3 (data not shown).

In summary, these results suggest that *pdfr-1* can act in multiple neurons—glutamatergic, and cholinergic, at least—to increase locomotion speed and lawn leaving. One neuron of particular interest is RIB, because it drives behavioral effects on speed on its own.

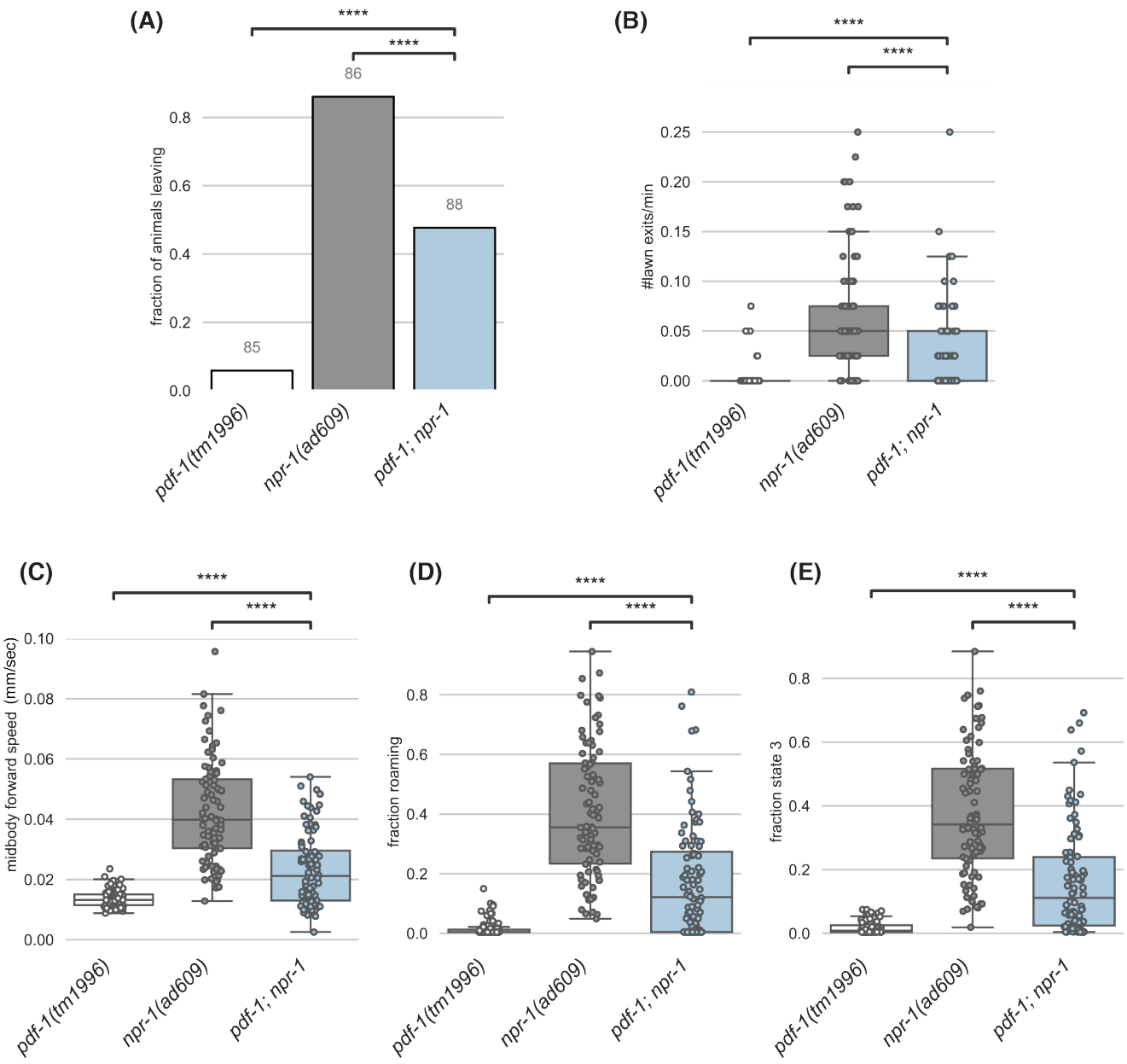


Figure 3-11. PDF-1 signaling is required for *npr-1(lf)* hyper-leaving.
(A) Fraction of *pdf-1, npr-1,* and *pdf-1; npr-1* mutant animals that execute at least one lawn exit in each genotype. Gray numbers indicate number of animals in each group. p value by proportions z test.
(B) Same as (A) for lawn exit rate.
(C) Same as (A) for mean midbody forward speed per animal.
(D) Same as (A) for fraction of time roaming per animal.
(E) Same as (A) for fraction of time in AR-HMM state 3 per animal.
Each dot represents one animal. Boxes indicate median and interquartile range. For panels B-E, Kruskal-Wallis with Dunn's multiple comparison correction was performed. * $p <= 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 10^{-4}$, ns = not significant.

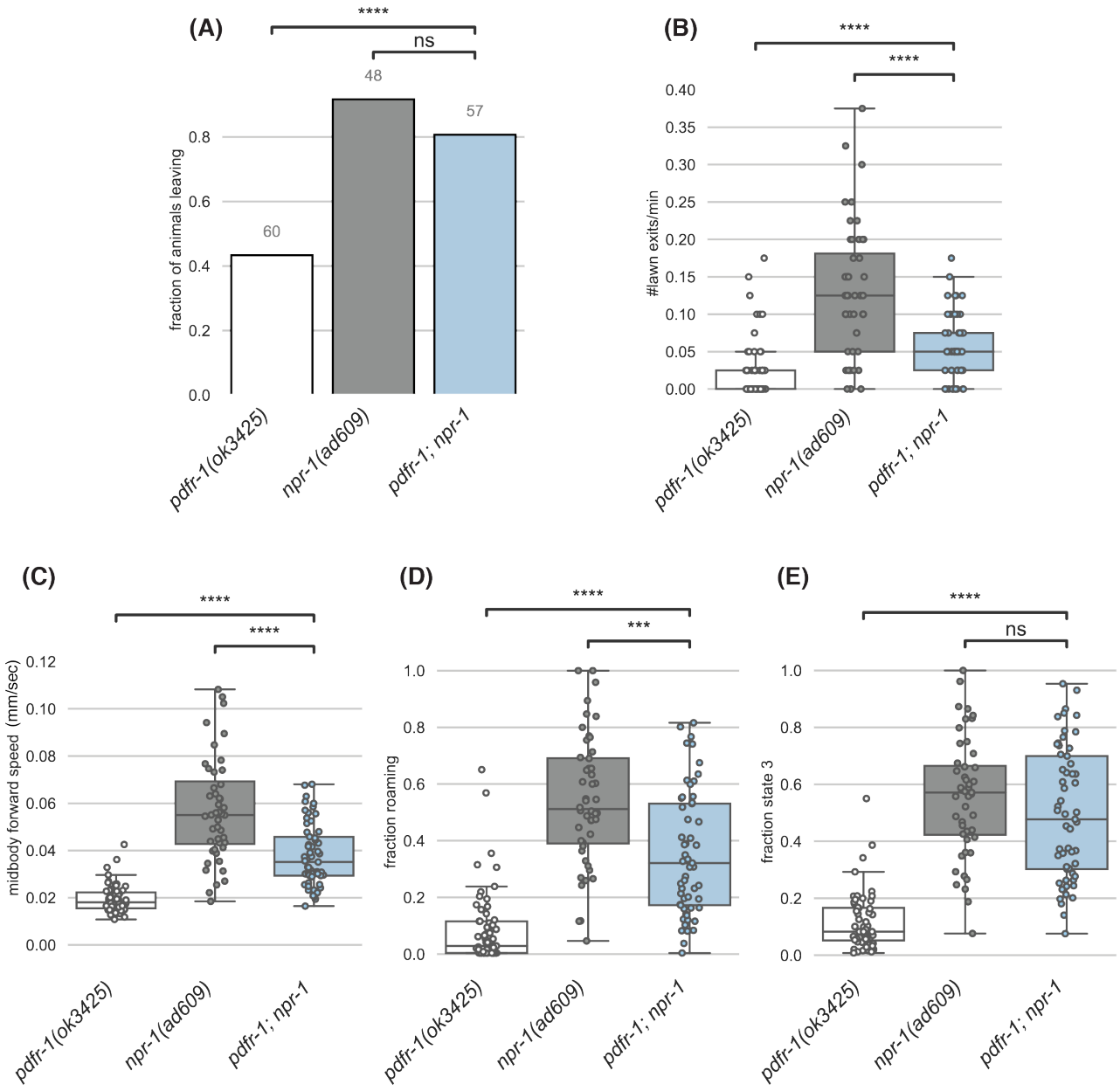


Figure 3-12. PDFR-1 signaling is required for *npr-1(lf)* hyper-leaving.
(A) Fraction of *pdfr-1, npr-1,* and *pdfr-1; npr-1* mutant animals that execute at least one lawn exit in each genotype. Gray numbers indicate number of animals in each group. p value by proportions z test.
(B) Same as (A) for lawn exit rate.
(C) Same as (A) for mean midbody forward speed per animal.
(D) Same as (A) for fraction of time roaming per animal.
(E) Same as (A) for fraction of time in AR-HMM state 3 per animal.
Each dot represents one animal. Boxes indicate median and interquartile range.
For panels B-E, Kruskal-Wallis with Dunn's multiple comparison correction was performed. * $p <= 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 10^{-4}$, ns = not significant.

Figure 3-13. RMG may be one site of PDF-1 release in *npr-1(lf)* animals.
(A) Circuit diagram of neurons with gap junctions to RMG neurons. Neurons expressing *npr-1* and *pdf-1* are indicated.
(B) Fraction of animals that execute at least one lawn exit in each genotype for two strains of *pdf-1; npr-1* animals carrying a transgene that expresses *pdf-1* specifically in RMG. Gray numbers indicate number of animals in each group. p value by proportions z test.
(C) same as (B) for lawn exit rate.
(D) same as (B) for mean midbody forward speed.
(E) same as (B) for fraction of time roaming.
Each dot represents one animal. Boxes indicate median and interquartile range.
For panels B-E, Kruskal-Wallis with Dunn's multiple comparison correction was performed. * $p <= 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 10^{-4}$, ns = not significant.

Figure 3-13

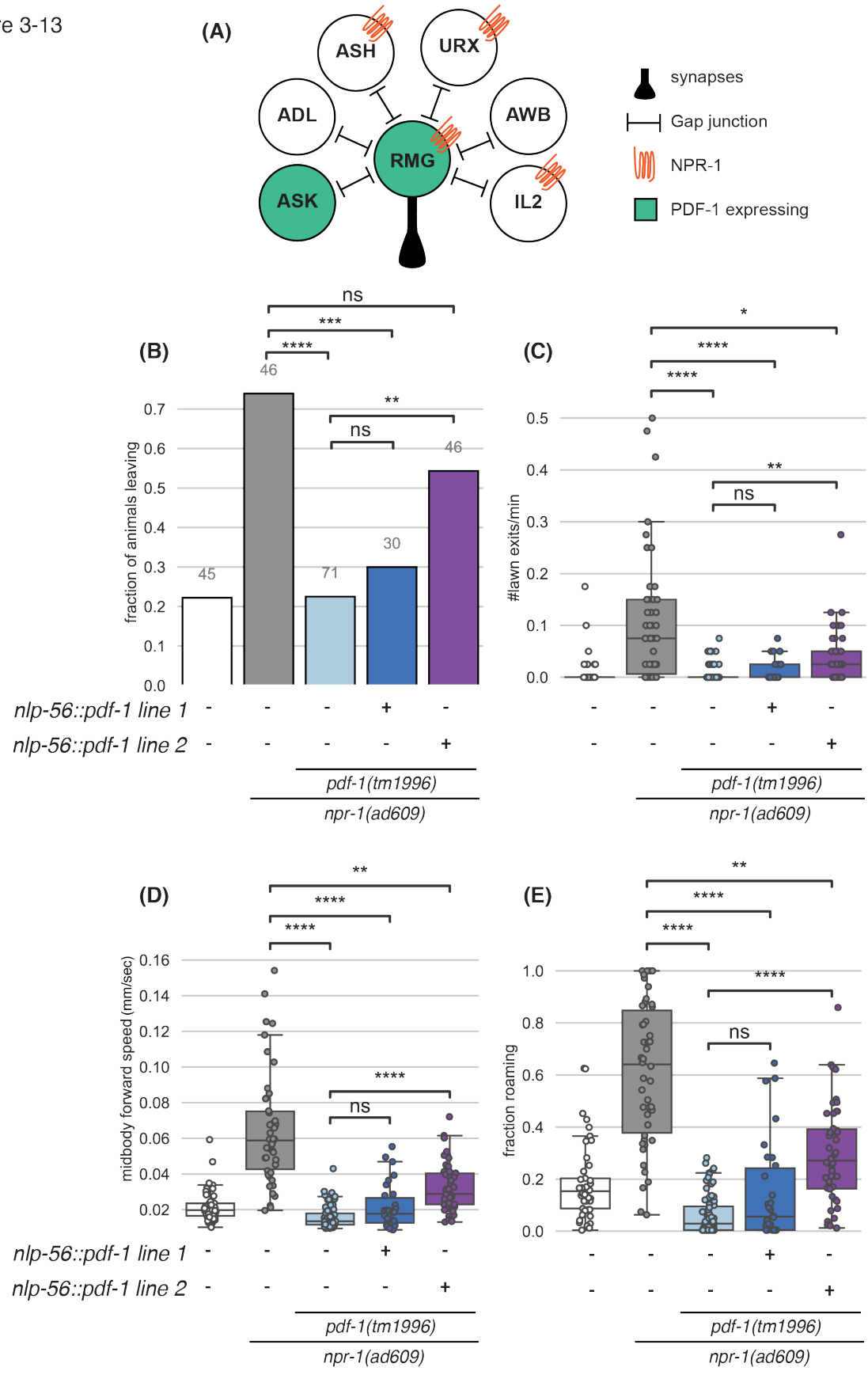

(A)
ASH
URX
ADL
AWB
RMG
ASK
IL2
synapses
Gap junction
NPR-1
PDF-1 expressing
(B)
fraction of animals leaving
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
ns
***
****
**
ns
46
45
71
30
46
(C)
#lawn exits/min
0.0
0.1
0.2
0.3
0.4
0.5
*
****
****
**
ns
(D)
midbody forward speed (mm/sec)
0.02
0.04
0.06
0.08
0.10
0.12
0.14
0.16
**
****
****
****
ns
(E)
fraction roaming
0.0
0.2
0.4
0.6
0.8
1.0
**
****
****
****
ns
nlp-56::pdf-1 line 1
nlp-56::pdf-1 line 2
pdf-1(tm1996)
npr-1(ad609)

Figure 3-14. *pdfr-1* expression in glutamatergic and cholinergic neurons controls different aspects of *npr-1(lf)* on-food phenotypes.
(A) Intersectional cell-specific rescue of *pdfr-1* using a Cre-Lox strategy. One transgene carries an inverted floxed *pdfr-1* cDNA ("*inv[pdfr-1]*"). A second transgene drives Cre in specific cell types to activate *pdfr-1* expression. *Adapted from (Flavell, et al 2013).*
(B-D) Fraction of animals that leave lawns, lawn exit rate, and midbody forward speed are compared across *pdfr-1* rescue in three different groups of cells. Each dataset compares genotypes collected on the same day.
(B) *pdfr-1; npr-1* animals carrying *inv[pdfr-1]* with and without pan-neuronal Cre are compared to *npr-1(lf)* alone.
(C) glutamatergic *pdfr-1* expression compared to *inv[pdfr-1]*, pan-neuronal *pdfr-1* expression in *pdfr-1; npr-1* animals.
(D) cholinergic *pdfr-1* expression compared to *inv[pdfr-1]*, pan-neuronal *pdfr-1* expression in *pdfr-1; npr-1* animals.
Each dot represents one animal. Boxes indicate median and interquartile range.
For panels B-D, Proportions z-test with Bonferroni correction was performed on fraction of animals leaving, Kruskal-Wallis with Dunn's multiple comparison correction was performed for lawn exit rate and midbody forward speed. * $p <= 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 10^{-4}$, ns = not significant.

Figure 3-14

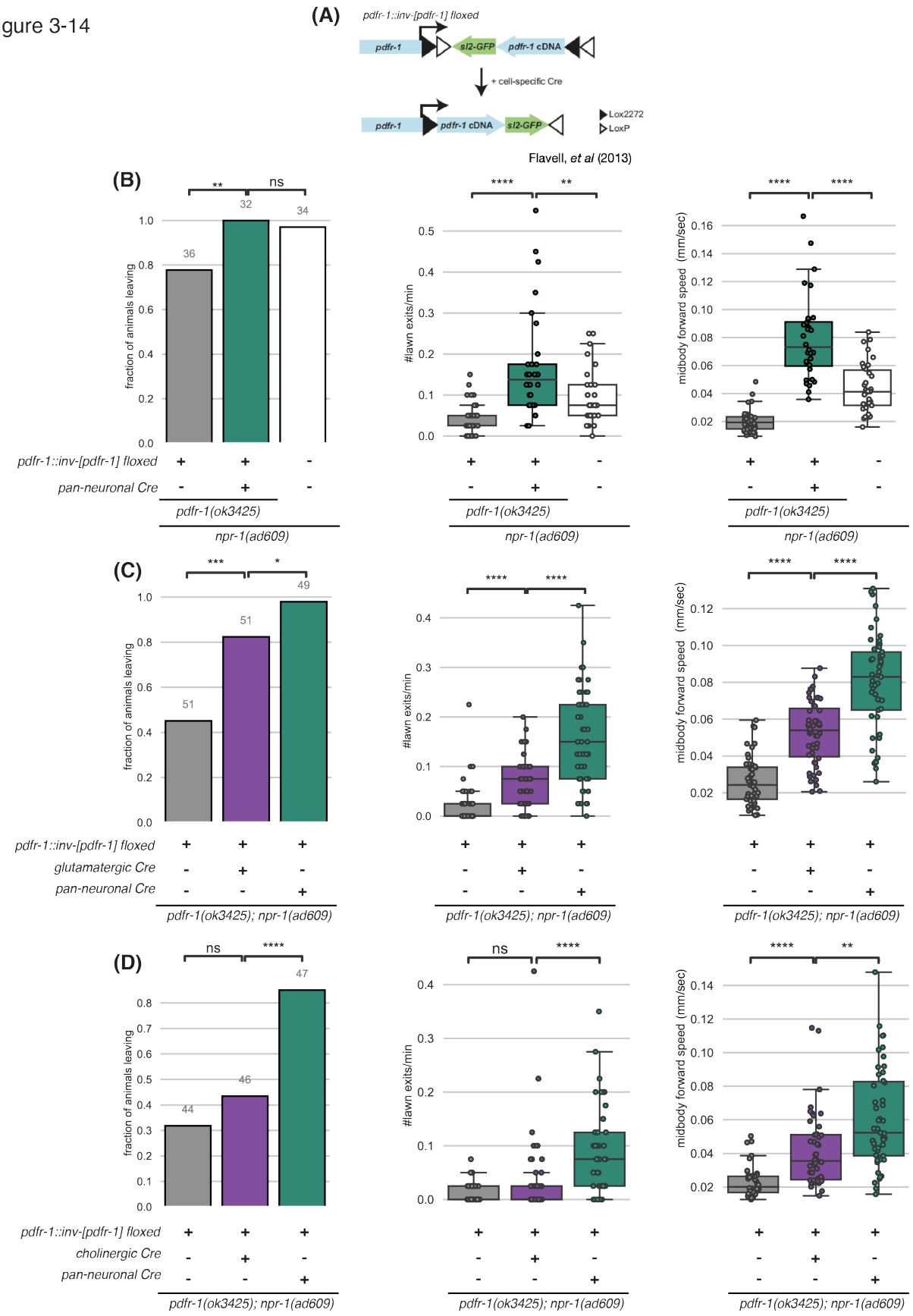
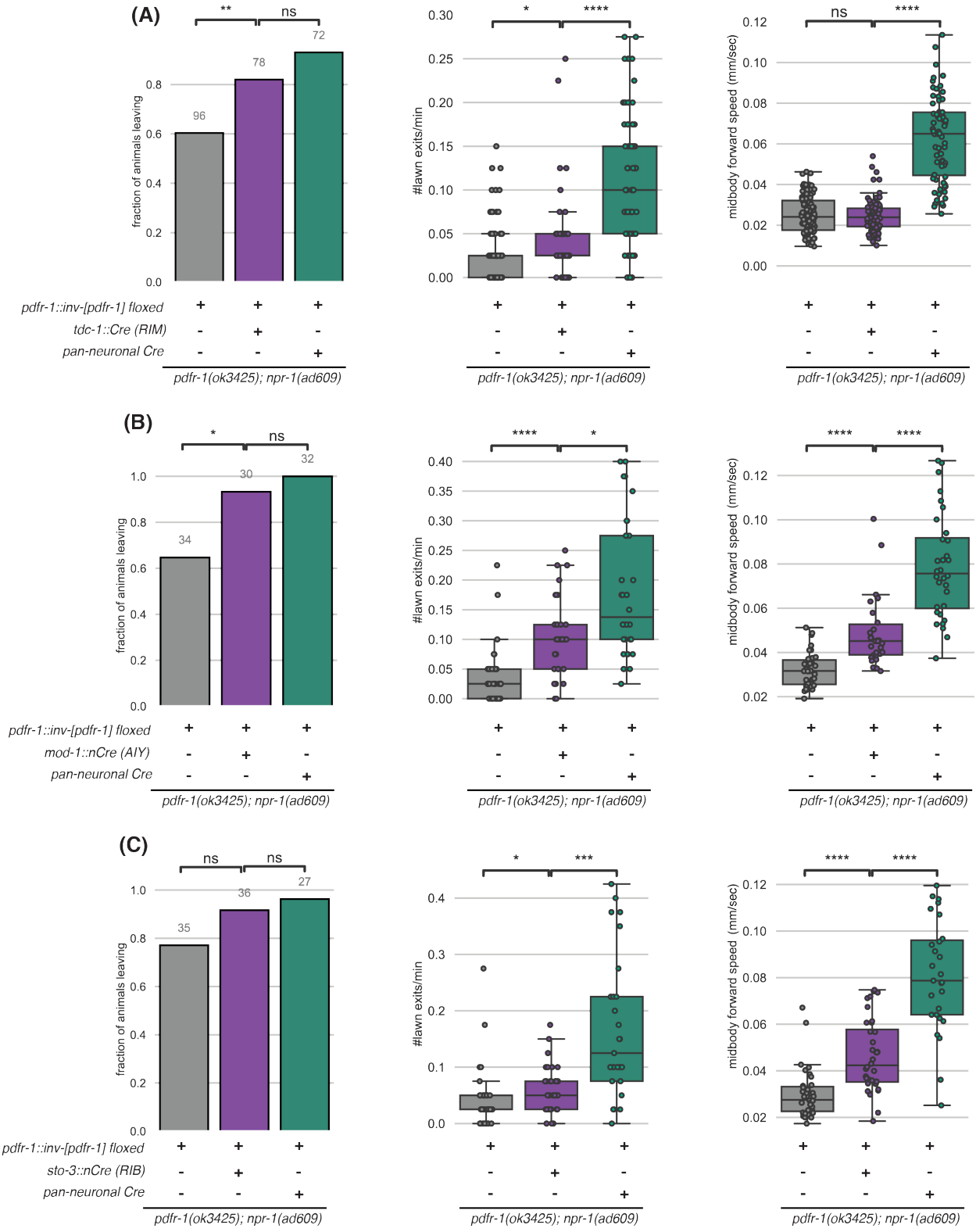

(A)
pdfr-1::inv-[pdfr-1] floxed
pdfr-1
sl2-GFP
pdfr-1 cDNA
+ cell-specific Cre
pdfr-1
pdfr-1 cDNA
sl2-GFP
Lox2272
LoxP
Flavell, et al (2013)
(B)
fraction of animals leaving
#lawn exits/min
midbody forward speed (mm/sec)
pdfr-1::inv-[pdfr-1] floxed
pan-neuronal Cre
pdfr-1(ok3425)
npr-1(ad609)
(C)
fraction of animals leaving
#lawn exits/min
midbody forward speed (mm/sec)
pdfr-1::inv-[pdfr-1] floxed
glutamatergic Cre
pan-neuronal Cre
pdfr-1(ok3425); npr-1(ad609)
(D)
fraction of animals leaving
#lawn exits/min
midbody forward speed (mm/sec)
pdfr-1::inv-[pdfr-1] floxed
cholinergic Cre
pan-neuronal Cre
pdfr-1(ok3425); npr-1(ad609)

Figure 3-15. *pdfr-1* expression in RIM and RIB neurons controls different aspects of *npr-1(lf)* on-food phenotypes.
(A-C) Fraction of animals that leave lawns, lawn exit rate, and midbody forward speed are compared across *pdfr-1* rescue in three different groups of cells. Each dataset compares genotypes collected on the same day.
(A) *pdfr-1* expression in RIM neurons compared to *inv[pdfr-1]*, pan-neuronal *pdfr-1* expression in *pdfr-1; npr-1* animals.
(B) *pdfr-1* expression in AIY neurons compared to *inv[pdfr-1]*, pan-neuronal *pdfr-1* expression in *pdfr-1; npr-1* animals.
(C) *pdfr-1* expression in RIB neurons compared to *inv[pdfr-1]*, pan-neuronal *pdfr-1* expression in *pdfr-1; npr-1* animals.
Each dot represents one animal. Boxes indicate median and interquartile range.
For panels A-C, Proportions z-test with Bonferroni correction was performed on fraction of animals leaving, Kruskal-Wallis with Dunn's multiple comparison correction was performed for lawn exit rate and midbody forward speed. * $p <= 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 10^{-4}$, ns = not significant.

Figure 3-15


(A)
**
ns
72
78
96
fraction of animals leaving
0.0
0.2
0.4
0.6
0.8
pdfr-1::inv-[pdfr-1] floxed
tdc-1::Cre (RIM)
pan-neuronal Cre
pdfr-1(ok3425); npr-1(ad609)
*
****
#lawn exits/min
0.00
0.05
0.10
0.15
0.20
0.25
0.30
ns
****
midbody forward speed (mm/sec)
0.00
0.02
0.04
0.06
0.08
0.10
0.12
(B)
*
ns
32
30
34
mod-1::nCre (AIY)
****
*
****
****
(C)
ns
ns
27
36
35
sto-3::nCre (RIB)
*
***
****
****

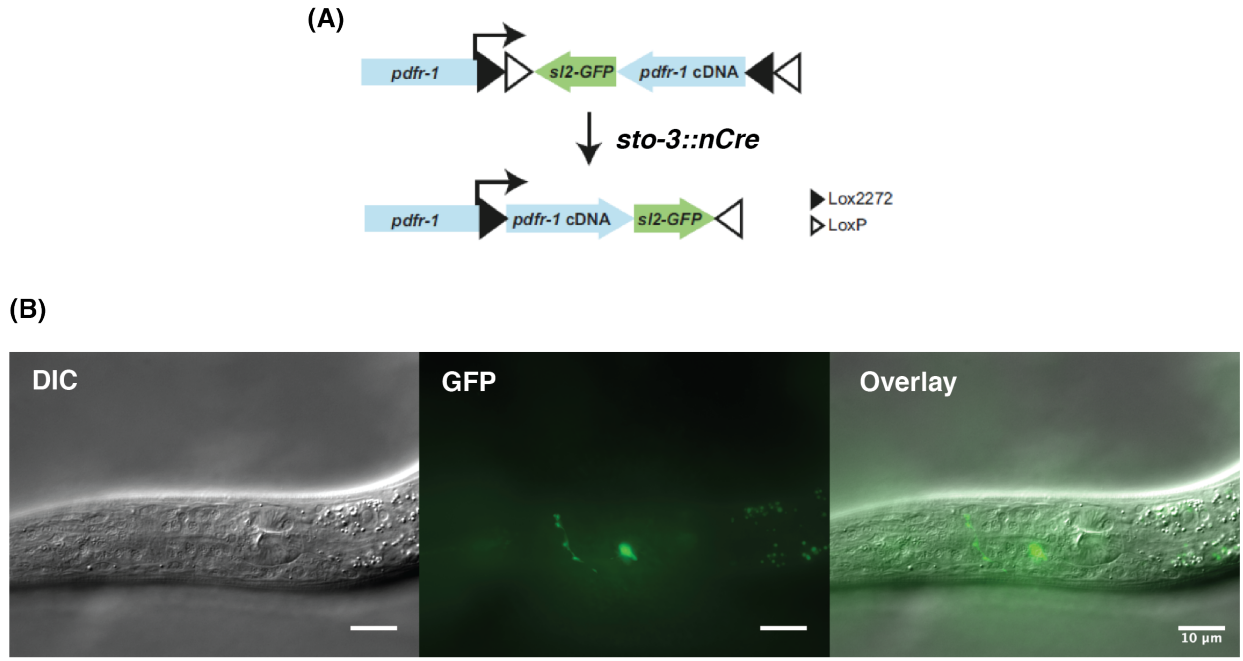


Figure 3-16. Intersectional expression patterns drive *pdfr-1* expression in RIB interneurons.
(A) Intersectional cell-specific rescue of *pdfr-1* in RIB interneurons using a Cre-Lox strategy *Adapted from (Flavell, et al 2013).*
(B) Validation of *pdfr-1* recombination in RIB interneurons. Differential interference contrast (DIC) (left), GFP fluorescence (middle), and merged (right) images of animals expressing *pdfr-1* in RIB. Scale bar is 10 μM.

Inactivation of RIB decreases locomotion speed and lawn leaving in *npr-1(lf)* animals

PDFR-1 is a secretin-receptor family G-protein coupled receptor (GPCR) that signals through Gαs to increase cyclic AMP (cAMP) and through Gαq to increase PIP2 turnover (Meelkop et al., 2012). In many neurons, cAMP and PIP2 mobilization increase excitability and synaptic transmission (Steuer Costa et al., 2017). Therefore, rescuing PDFR-1 signaling in RIB neurons may potentiate RIB activity or synaptic output. To test the hypothesis that RIB activity contributes to lawn leaving and high speed of *npr-1(lf)* animals, I used a chemogenetic approach to hyperpolarize RIB neurons. I generated two transgenic strains that express the histamine-gated chloride channel HisCl1 in RIB neurons and exposed them to exogenous histamine. *C. elegans* does not use endogenous histamine as a neurotransmitter, but can express exogenous histamine-gated chloride channels that drive neuronal hyperpolarization in response to histamine (Pokala et al., 2014). In both strains, silencing RIB dramatically reduced forward speed of *npr-1(lf)* animals (Figue 3-17 A), and decreased roaming (Figure 3-17 B). Although RIB silencing did not change the fraction of animals leaving the lawn (Figure 3-17 C), it did decrease the number of leaving events per animal (Figure 3-17 D), suggesting RIB can affect lawn leaving to some degree.

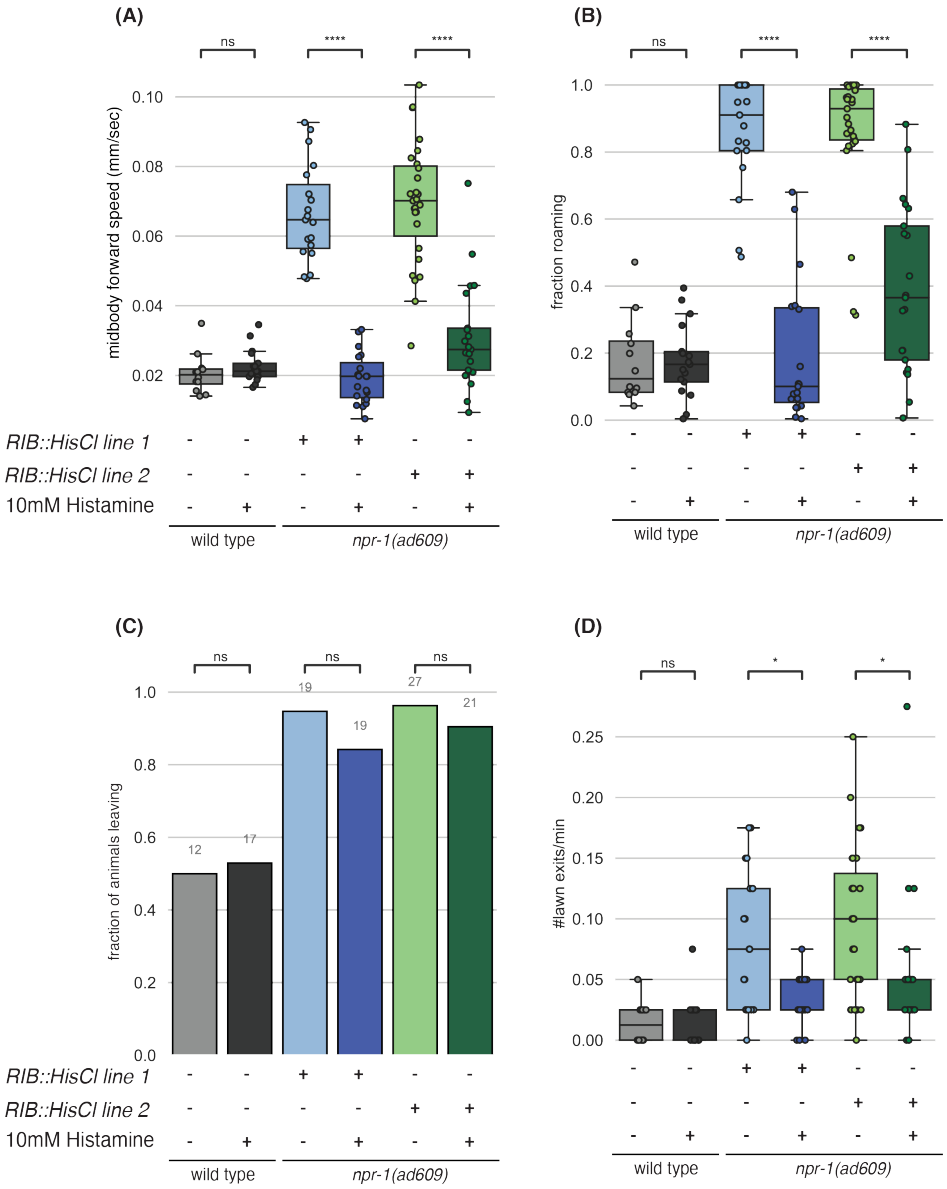


Figure 3-17. RIB activity is required for *npr-1(lf)* high speed on food and lawn leaving.

(A) Midbody forward speed of *npr-1(ad609)* animals carrying a transgene expressing the histamine gated chloride channel HisCl1 in RIB neurons with and without exogenous histamine.

(B) same as (A) for fraction of time spent roaming.

(C) same as (A) for fraction of animals that leave lawns.

(D) same as (A) for lawn exit rate.

Each dot represents one animal. Boxes indicate median and interquartile range.

For panel C, Proportions z-test with Bonferroni correction was performed.

For panels A,B,D, Kruskal-Wallis with Dunn's multiple comparison correction was performed. * $p <= 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 10^{-4}$, ns = not significant.

Discussion and Future Directions

In this chapter, I investigated environmental and internal triggers of lawn leaving behavior in *C. elegans.* Animals move faster in less dense lawns of bacteria, and a higher fraction of animals leave the lawn. Differences in speed on lawns of different densities may be responses to the absolute abundance of food experienced in the present or may reflect a memory of these animals' previous growth on denser or less dense foods in the past. Further experiments can distinguish between these possibilities by varying the density of food during the acclimation phase before measuring speed and lawn leaving in small lawn assays.

When food intake is acutely inhibited by optogenetic paralysis of the pharynx, wild type animals immediately increase their speed. In data not presented here, I also found that animals lacking serotonin (*tph-1* mutants) and animals lacking *pdfr-1* signaling were still able to induce high speed roaming-like behavior upon pharyngeal paralysis. Given that these signaling pathways are required for dwelling and roaming, respectively, this raises the question of how the endogenous molecular mechanisms that generate roaming and dwelling states interact with signaling pathways induced by pharyngeal paralysis. Since the serotonergic pharyngeal neuron NSM is known to surveil food intake in the alimentary tract through its minor neurite (Rhoades et al., 2019), one hypothesis is that acute loss of serotonin signaling from NSM can induce roaming. However, since *tph-1* animals still maintain their ability to roam upon pharyngeal inhibition, this is not likely to be the only mechanism. One possibility is that NSM also secretes additional factors that communicate the presence of food to the rest of the nervous system. Another possibility is that additional neurons or non-neuronal tissues in the pharynx secrete signals that stimulate roaming-like behaviors. To test this, I could express the tetanus toxin light chain in the NSM neurons during pharyngeal paralysis to see if the synaptic output of NSM is required for the roaming-like behaviors generated.

Although animals rapidly increase movement speed following optogenetic pharyngeal paralysis, they do not uniformly leave the bacterial lawn when the light turns on. Instead, lawn leaving events occur uniformly over a ten-minute interval of paralysis. I speculate that increased speed is a prerequisite for lawn leaving, but a second trigger or probabilistic event may drive the actual lawn leaving event.

Based on my analysis of *tax-4* mutants, the activity of sensory neurons that detect bacterial density may prevent animals from leaving food lawns. *tax-4(lf)* animals roam less than wild type but show a higher fraction of animals that leave a food lawn. Before leaving, *tax-4(lf)* have a sudden burst of speed. These observations suggest that *tax-4(lf)* animals are unlikely to enter high arousal states, but once aroused, are more likely to leave the lawn, perhaps because of a loss of sensory signals from food. To test this hypothesis in a future experiment, I could inhibit pharyngeal pumping in *tax-4(lf)* animals. If TAX-4-related sensory activity normally prevents animals from leaving during aroused states, I expect these animals to leave food lawns rapidly upon acute feeding

deprivation.

TAX-4 affects many sensory neurons, and different aspects of its functions are evident in *npr-1(lf)* animals. In *npr-1(lf)* animals, arousal driven by oxygen-sensing neurons drives aggregation, bordering, and high speed on food (Coates and De Bono, 2002, and this work). All of these behaviors depend upon high RMG activity through gap junctions with multiple sensory neurons, including several that require TAX-4 for their activity (Jang et al., 2017; Macosko et al., 2009). One possibility is that sensory neurons lacking TAX-4 are effectively hyperpolarized and can inhibit RMG through gap junctions, thereby suppressing high locomotion speed, aggregation and lawn leaving. The idea of inhibition spreading through gap junctions has a precedent in recent work on behavioral outputs of the RIM interneurons (Sordillo and Bargmann, 2021).

In contrast with NPR-1-dependent aggregation and bordering, I found that the oxygen-sensing neurons affected by *gcy-35(lf)* were not essential for regulating lawn leaving behavior. Similarly, killing AQR, PQR and URX neurons did not suppress locomotion speed or lawn leaving. Since URX expresses *tax-4,* I surmise that loss of URX activity is not essential for the *tax-4* mutant phenotype. Future rescue experiments can be used to determine which additional *tax-4*-expressing neurons contribute to arousal and lawn leaving behavior. AWC and ASI sensory neurons, which express *tax-4*, are good candidates, given their roles in promoting roaming behavior on uniform food lawns (Ben Arous et al., 2009).

OLL neurons, which do not express *tax-4*, inhibit lawn leaving. OLL neurons are mechanosensory neurons that inhibit leaving of pathogenic bacterial lawns (Chang et al., 2011). I speculate that these neurons may play a general mechanosensory role that causes animals to stay on bacterial lawns of all kinds.

In agreement with previous work on the NPR-1 hub-and-spoke circuit, I found that RMG neurons are the crucial site of *npr-1* expression. The suppression of lawn leaving by TeTx in RMG suggests that regulated release of a neurotransmitter or neuropeptide from RMG mediates its downstream effects. However, I failed to elicit a suppression of *npr-1(lf)* hyper leaving by RMG-specific knockout of neuropeptide processing enzymes EGL-3 or EGL-21. RMG is not thought to secrete classical neurotransmitter: Among the classes of neurotransmitters that do not rely on neuropeptide processing machinery are the insulin-like peptides and non-canonical neurotransmitters such as betaine (Chen et al., 2013; Hardege et al., 2021), which could be secreted by RMG to promote arousal.

Finally, I found that the lawn leaving behaviors of *npr-1* mutants, and other arousal behaviors, can be suppressed by mutations in the neuropeptide PDF-1 and its receptor PDFR-1. Interestingly, multiple classes of neurons can contribute to PDFR-1's function in lawn leaving, just as multiple classes contribute to its role in roaming and mate search in males (Barrios et al., 2012; Flavell et al., 2013). The neurons that express *pdfr-1* and promote lawn leaving include both cholinergic and glutamatergic neurons.

Among these classes, PDFR-1 signaling in cholinergic RIB neurons or glutamatergic RIM neurons contributed to speed and lawn leaving of *npr-1(lf)*. In future studies, I would also like to examine the effects of RIM silencing in the *npr-1(lf)* background.

CHAPTER 4: INVESTIGATING ENTRAINMENT OF BEHAVIORAL STATES TO FOOD ABUNDANCE

Introduction

On bacterial food, *C. elegans* locomotion can be categorized into two behavioral states: Roam states that are characterized by high speed forward movement, and Dwell states defined by low speed and frequent reversals (Ben Arous et al., 2009; Flavell et al., 2013; Fujiwara et al., 2002). The relative fraction of roaming and dwelling behavior is strongly influenced by sensory information about the food environment. Dense nutritious bacterial lawns promote prolonged Dwell states, while sparse food promotes Roam states (Ben Arous et al., 2009; Shtonda, 2006). Animals with loss-of-function mutations in chemosensory signaling genes like *che-2* and *tax-4* show increased dwelling as do mutants defective in AWC sensory neuron function (Ben Arous et al., 2009; Fujiwara et al., 2002). On the other hand, a mutation in the cyclic GMP-dependent protein kinase *egl-4*, which promotes sensory adaptation, leads to increased roaming (McCloskey et al., 2017). These results suggest that sensory activity that relays information about food quality stimulates roaming, while its absence causes dwelling. A recent study showed that animals on food extend Roam state durations to chemotax toward food odors, while triggering dwelling if odor concentration decreases (Ji et al., 2021).

An additional input that controls roaming and dwelling comes from an animal's continuous monitoring of its food intake. Animals acutely unable to ingest bacterial food due to pharyngeal paralysis increase roaming (Chapter 3). Similarly, treating *E. coli* with a drug that makes bacteria too large to ingest causes animals to roam (Ben Arous et al., 2009). *C. elegans* can monitor its own feeding through the pharyngeal neuron NSM that senses the presence of food in the alimentary tract and whose activity leads to dwelling through serotonin release (Rhoades et al., 2019).

Roam and Dwell states persist even in environments with uniform bacterial density and without any obvious sensory gradients. How does the internal state of the nervous system orchestrate selection of behavioral states? The answer may be that the underlying circuit architecture that controls Roam and Dwell states has "winner-take-all" dynamics. Previous work has shown that two sets of neurons antagonistically promote either dwelling, via serotonin released by NSM neurons, or roaming, via secretion of the neuropeptide PDF-1 by several neurons (Flavell et al., 2013; Ji et al., 2021). Large-scale calcium imaging in freely moving animals shows that in wild type animals, these two neuronal groups have mutually exclusive activity patterns (Ji et al., 2021). However, mutations in the PDF receptor gene *pdfr-1* appear to abolish this mutual exclusivity, leading to disorganized and abortive Roam states (Ji et al., 2021). Both PDFR-1 and its ligand PDF-1 are expressed in many neurons and act at multiple sites to support roaming behavior (Flavell et al., 2013).

In this chapter, I ask how sensory information from experienced changes in bacterial density influences the transition between Roam and Dwell states. My results suggest that PDFR-1 may play different roles in different neurons to regulate arousal and lawn leaving. In particular, I found that expression of *pdfr-1* in RIB neurons increases the fraction of time spent roaming but did not entrain Roam states to experienced bacterial density.

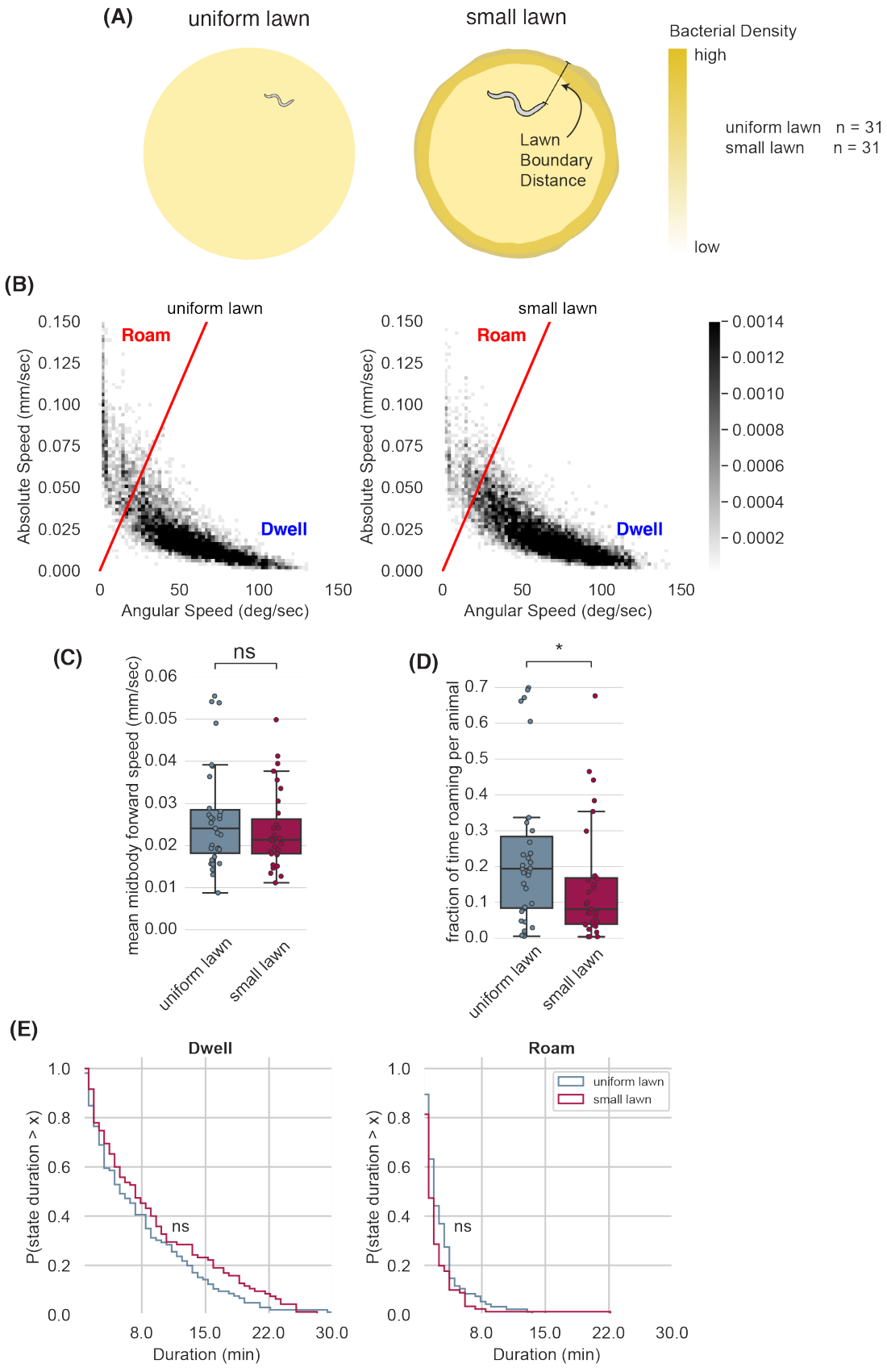
Results

Roaming is decreased on small lawns

To investigate the effect of unevenly distributed food abundance on *C. elegans* behavioral states, I compared animal behavior on small bacterial lawns (as described in Chapter 2), with behavior on lawns of bacteria grown to even density ("uniform lawns", see Methods). Although animals move with similar speeds in these environments, animals on uniform lawns roamed 23% of the time while animals on small lawns roamed only 14% of the time (Figure 4-1). Nonetheless, the duration of Roam and Dwell states did not differ greatly between these contexts, suggesting that basic regulatory features are maintained. These results suggest that prevalence of Roam states is affected by the spatial distribution of bacterial food in the environment.

Figure 4-1. Animals roam less on small lawns than uniform lawns.
(A) Schematic of animal of an animal on either a uniform lawn or small lawn, illustrating Lawn Boundary Distance and Bacterial Density (see Methods). Note that uniform lawns are much larger than small lawns, so worm appears smaller.
(B) Scatter plot of average absolute speed and angular speed in 10 s intervals for animals on uniform lawns (n = 31, 7440 bins) and animals on small lawns (n = 31, 7440 bins). Two clusters of data points correspond to: high speed/low angular speed roaming and low speed/high angular speed dwelling.
(C) Mean midbody forward speed per animal compared between animals on uniform and small lawns. Each dot represents one animal. Boxes indicate median and interquartile range. ns = no significant difference between conditions using Mann-Whitney two-sided test.
(D) Fraction of time spent roaming per animal on uniform and small lawns. Each dot represents one animal. Boxes indicate median and interquartile range. * $p < 0.01$ by Mann-Whitney two-sided test; ns = not significant.
(E) Complementary cumulative distribution function (ccdf) for Dwell and Roam state durations in wild-type animals on uniform and small lawns. ns = no significant difference between conditions using Kolmogorov-Smirnov two-sided test.

Figure 4-1

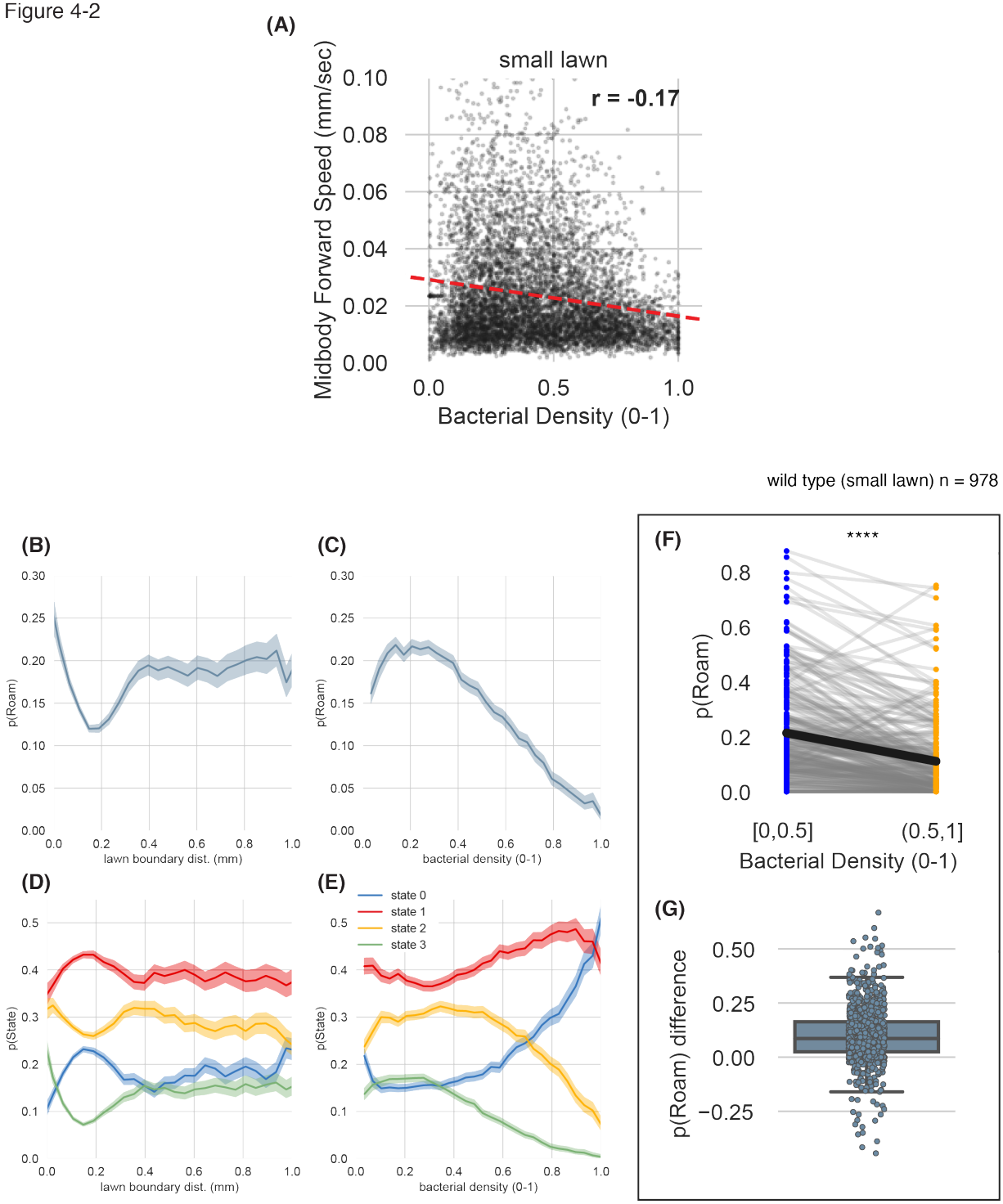

(A)
uniform lawn
small lawn
Bacterial Density
high
low
uniform lawn n = 31
small lawn n = 31
Lawn
Boundary
Distance
(B)
uniform lawn
small lawn
Roam
Dwell
Absolute Speed (mm/sec)
Angular Speed (deg/sec)
(C)
ns
mean midbody forward speed (mm/sec)
uniform lawn
small lawn
(D)
*
fraction of time roaming per animal
(E)
Dwell
Roam
P(state duration > x)
Duration (min)
ns

Bacterial density is inversely correlated with speed and roaming on small lawns

I hypothesized that the presence of a dense edge region in small lawns causes animals to dwell more. In small lawns, I observed a negative correlation between bacterial density and forward speed, such that speed is suppressed at high bacterial density and vice versa (Figure 4-2 A). Next, I investigated the relationship between bacterial density and lawn boundary distance on the probability of roaming “p(Roam)” on small lawns. p(Roam) is suppressed within 0.3 mm of the lawn boundary (Figure 4-2 B), where food is densest and visible thickening can be seen (Figure 2-1 B). Similarly, p(Roam) is constant below bacterial density of 0.4 and decreases monotonically as bacterial density increases (Figure 4-2 C). Looking at p(Roam) in just two intervals of bacterial density, [0,0.5] and (0.5,1] shows reliable suppression in the high-density interval with a mean difference in p(Roam) of 0.1 across the two intervals. A similar phenomenon was observed in AR-HMM states: probability of low-speed AR-HMM states 0 and 1 showed enrichment within 0.3 mm of the lawn boundary, while high speed AR-HMM states 2 and 3 were depleted (Figure 4-2 D). Among AR-HMM states, state 0, which corresponds to the lowest speeds and pausing, was the most strongly modulated by varying bacterial density: at bacterial density = 0, p(state 0) = 0.2 and at bacterial density = 1, p(state 0) = 0.5. This enrichment likely indicates that animals pause more at high bacterial density.

Figure 4-2. Bacterial density suppresses forward speed and probability of roaming in small lawns food environments.
(A) Scatter plot of Midbody Forward Speed and Bacterial Density in 10 s intervals for animals on small lawns (n = 31, 7440 bins). Linear least squares regression line is plotted as a red dashed line. r is the Pearson correlation coefficient.
(B) Probability of roaming "p(Roam)" calculated over 30 contiguous bins across Lawn Boundary Distance ranging between 0 to 1 mm.
(C) Probability of roaming "p(Roam)" calculated over 30 contiguous bins across Bacterial Density ranging between 0 to 1.
(D) Probability of each AR-HMM state calculated over the same bins as in (B).
(E) Probability of each AR-HMM state calculated over the same bins as in (C).
(F) Paired comparison of p(Roam) between low Bacterial Density [0,0.5] and high Bacterial Density (0.5,1]. Each gray line represents values of p(Roam) per animal. The average across all animals is plotted in black. **** $p < 10^{-4}$ by Wilcoxon signed rank test.
(G) Difference in p(Roam) between low Bacterial Density [0,0.5] and high Bacterial Density (0.5,1]. Each gray dot represents data from one animal. Box indicates median and interquartile range.
For panels B-G, n = 978 wild type animals on small lawns.
For panels B-E, dark line is mean, shaded area is standard error.

Figure 4-2


(A)
small lawn
r = -0.17
Midbody Forward Speed (mm/sec)
Bacterial Density (0-1)
wild type (small lawn) n = 978
(B)
p(Roam)
lawn boundary dist. (mm)
(C)
p(Roam)
bacterial density (0-1)
(D)
p(State)
lawn boundary dist. (mm)
(E)
state 0
state 1
state 2
state 3
p(State)
bacterial density (0-1)
(F)
****
p(Roam)
[0,0.5]
(0.5,1]
Bacterial Density (0-1)
(G)
p(Roam) difference

Changes in experienced bacterial density precede behavioral state changes

In order to discern how bacterial density affects behavioral state changes, I aligned midbody forward speed, bacterial density, and lawn boundary distance to Roam and Dwell state transitions. As expected, switching from Dwell to Roam or Roam to Dwell states corresponded to stable increases and decreases in mean forward speed, respectively. Interestingly, these state transitions also aligned well with changes in lawn boundary distance and bacterial density. Animals switching to Roam from Dwell states transiently deflected their heads towards higher lawn boundary distance (further inside the lawn). This corresponded to a transient dip in bacterial density as the food is less dense in the lawn center (Figure 4-3 A-C). The opposite was true of transitions from Roam to Dwell states (Figure 4-3 D-F).

These observations suggest that bacterial density changes are associated with behavioral state transitions. To explore these relationships further, I performed auto-correlation and cross-correlation analysis. Cross correlation is a measure of how similar two time-series are to each other at different time lags, with a maximum value of 1 for complete correlation or -1 for complete anticorrelation. Autocorrelation is the cross correlation of a time series with itself and is a useful measure of the duration of temporal patterns in the data. I found positive autocorrelation of midbody forward speed that extended to lags up to +/-5 minutes (Figure 4-4 A). Thus, for a given time *t*, temporal patterns in midbody forward speed persist within a 10-minute window centered on *t*; in this context, the timescales match well with roaming duration (Figure 4-1 E), which has a survival time of roughly 9 minutes. The cross-correlation of bacterial density with midbody forward speed (bacterial density $\star$ midbody forward speed) shows a negative correlation over approximately the same duration, suggesting that periods of rapid locomotion coincide with occupancy of low bacterial density, and slow locomotion with occupancy at high bacterial density.

Lawn boundary distance $\star$ midbody forward speed shows a positive correlation, indicating that animals move at slower speed at the lawn border. Interestingly, a peak correlation at -40 second +/- 5 seconds suggests that transitions between the lawn border to the lawn center precede, and might predict, transitions between low and high speed locomotion. Because temporal structure in autocorrelation functions can create spurious structure in cross-correlations, I also calculated the autocorrelations and cross-correlations of a shuffled version of these datasets. This yielded identical autocorrelations but abolished temporal structure in cross-correlation functions, indicating that cross-correlation results are not artifacts of the autocorrelation functions. The magnitudes of the cross-correlations indicate that bacterial density is more reliably anticorrelated with speed than lawn boundary distance is correlated with speed. Accordingly, in subsequent cross-correlation analyses, I focused on the bacterial density-speed correlation.

Figure 4-3. Changes in bacterial density precede Roam/Dwell state transitions.
(A) Heatmap of Midbody Forward Speed aligned to Dwell to Roam state transitions (left). Mean across all traces on left is plotted as dark blue with standard error of the mean plotted in light blue (right). n = 2648 Dwell to Roam transitions for (A-C)
(B) same as (A) for Lawn Boundary Distance.
(C) same as (A) for Bacterial Density.
(D) Heatmap of Midbody Forward Speed aligned to Roam to Dwell state transitions (left). Mean across all traces on left is plotted as dark blue with standard error of the mean plotted in light blue (right). n = 2765 Roam to Dwell transitions for (D-F).
(E) same as (D) for Lawn Boundary Distance.
(F) same as (D) for Bacterial Density.
For panels, A-F, n = 978 wild type animals on small lawns.

Figure 4-3

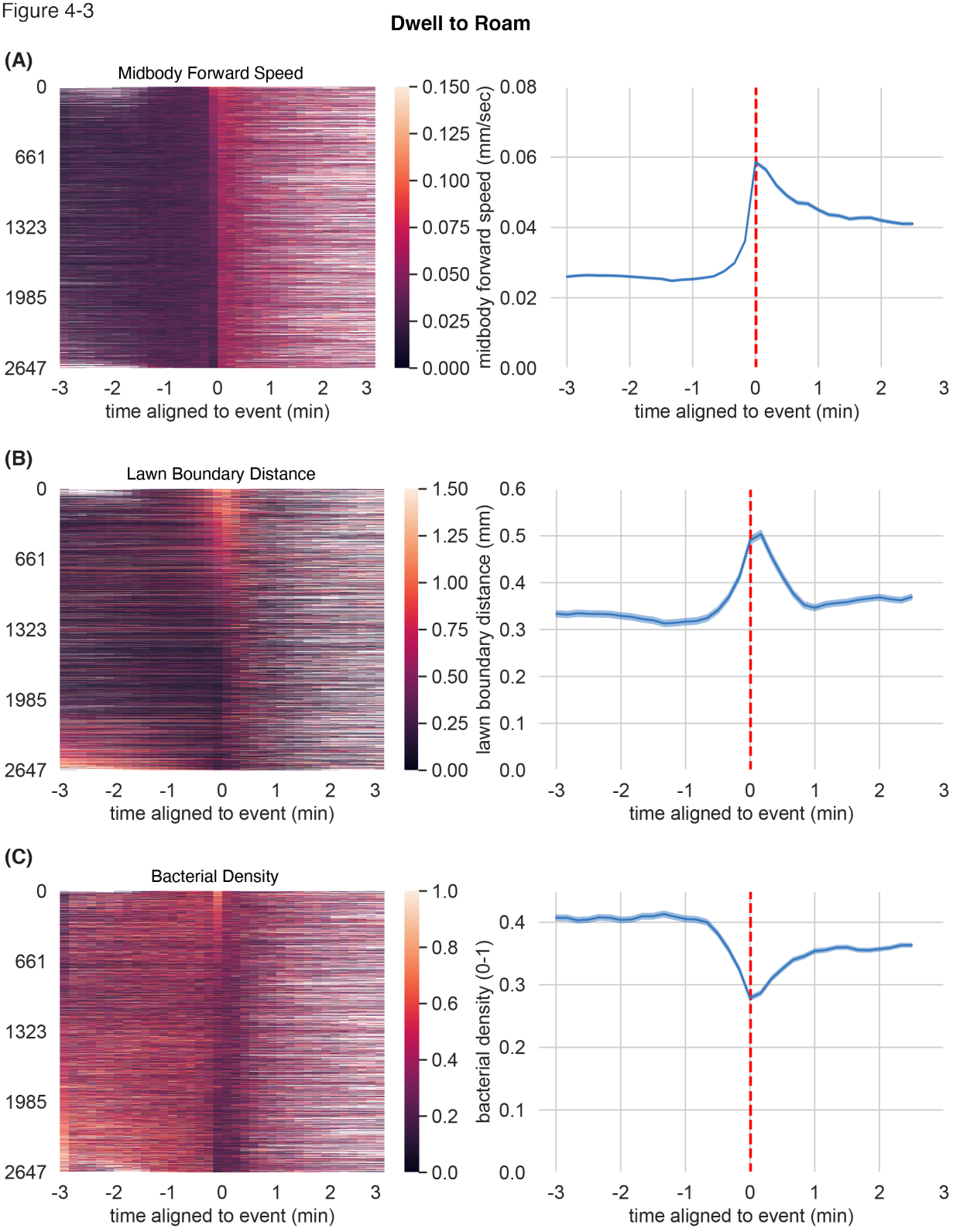
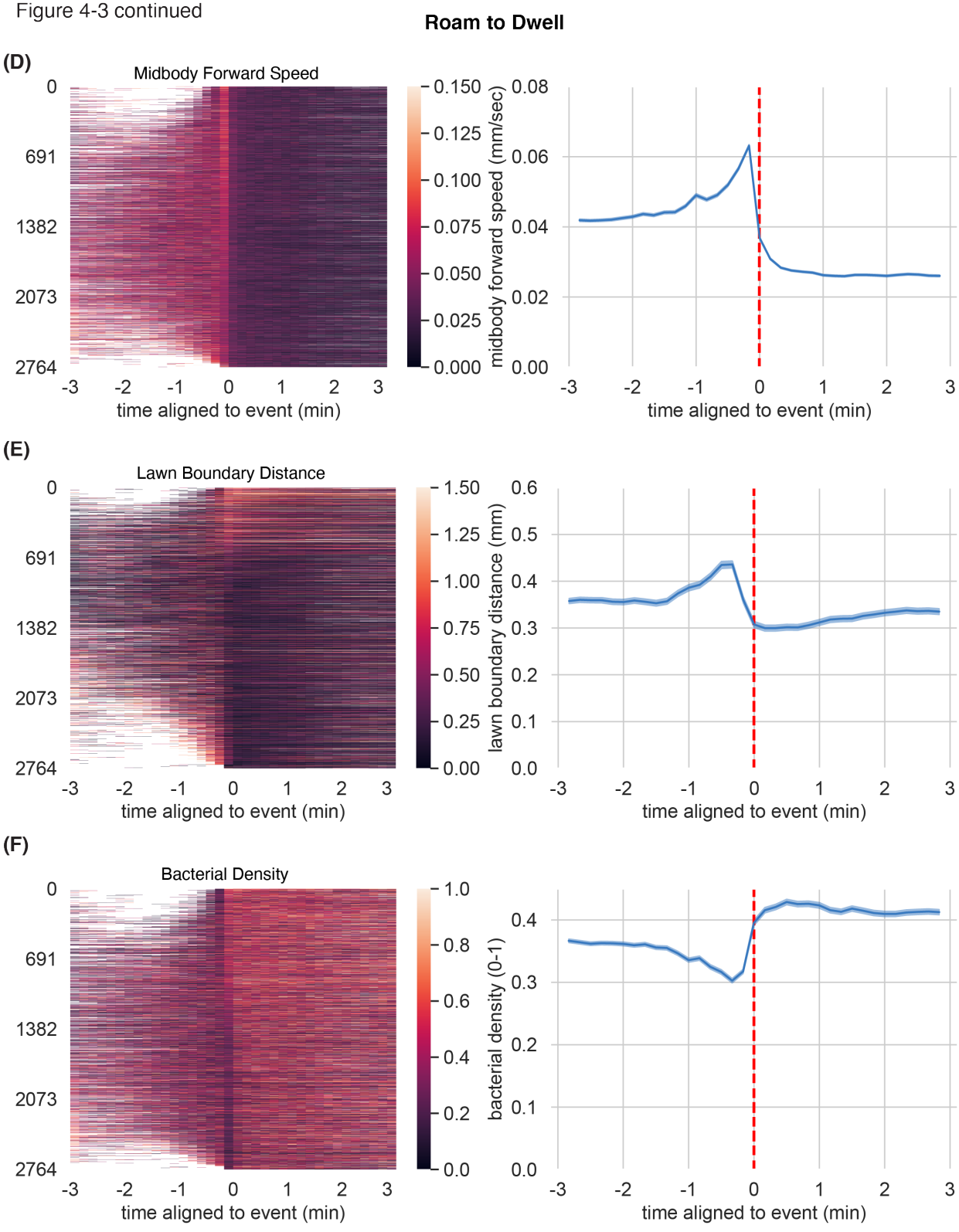

Dwell to Roam
(A)
Midbody Forward Speed
0
661
1323
1985
2647
-3 -2 -1 0 1 2 3
time aligned to event (min)
0.150
0.125
0.100
0.075
0.050
0.025
0.000
midbody forward speed (mm/sec)
0.08
0.06
0.04
0.02
0.00
-3 -2 -1 0 1 2 3
time aligned to event (min)
(B)
Lawn Boundary Distance
0
661
1323
1985
2647
-3 -2 -1 0 1 2 3
time aligned to event (min)
1.50
1.25
1.00
0.75
0.50
0.25
0.00
lawn boundary distance (mm)
0.6
0.5
0.4
0.3
0.2
0.1
0.0
-3 -2 -1 0 1 2 3
time aligned to event (min)
(C)
Bacterial Density
0
661
1323
1985
2647
-3 -2 -1 0 1 2 3
time aligned to event (min)
1.0
0.8
0.6
0.4
0.2
0.0
bacterial density (0-1)
0.4
0.3
0.2
0.1
0.0
-3 -2 -1 0 1 2 3
time aligned to event (min)

Figure 4-3 continued



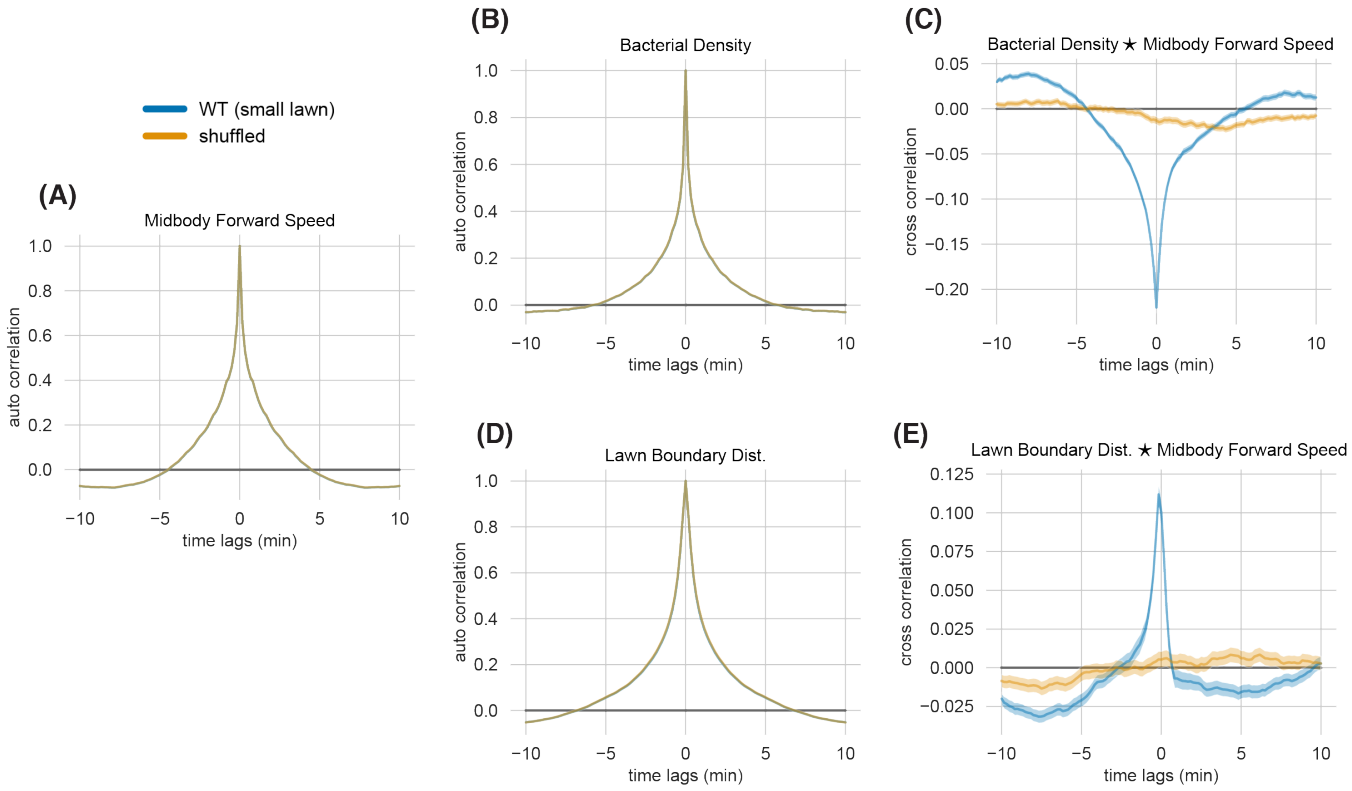


Figure 4-4. Changing bacterial density and lawn boundary is anticorrelated with changes in forward speed.
(A) Mean autocorrelation function for Midbody Forward Speed is plotted as dark blue with standard error of the mean plotted in light blue. Shuffled autocorrelations are shown in yellow.
(B) Same as (A) for Bacterial Density.
(C) Mean zero-normalized cross-correlation for Bacterial Density ⋆ Midbody Forward Speed is plotted as dark blue with standard error of the mean plotted in light blue. Shuffled cross-correlations are plotted in yellow.
(D) Same as (A) for Lawn Boundary Distance.
(E) Mean zero-normalized cross-correlation for Lawn Boundary Distance ⋆ Midbody Forward Speed is plotted as dark blue with standard error of the mean plotted in light blue. Shuffled cross-correlations are plotted in yellow.

Genetics and neuronal functions that affect responses to bacterial density

To investigate the genetic and neuronal mechanisms that impact *C. elegans*' responses to food distribution, I measured the average lawn boundary distance and bacterial density across the same 19 strains that were assayed for differences in lawn leaving and roaming (see Chapter 3). Due to the variation in control data across days, all mutants were compared to wild type animals tested on the same day (Table 4-1). Among sensory manipulations, two strains with genetic ablation of sensory neurons showed significant increases in lawn boundary distance after multiple hypothesis test correction: one that lacks the oxygen-sensing neurons AQR, PQR and URX (*qaIs2241*) and one that lacks the mechanosensory OLL neurons (*qdEx22*). When comparing the average bacterial density, a different set of strains showed significant decreases compared to wild type: *tax-4(lf),* a strain in which the BAG and ASG sensory neurons are genetically ablated, and the OLL-ablated strain (Figure 4-5). In all assays, several additional strains showed significant differences from same-day controls that did not survive a multiple hypothesis test correction (Figure 4-5, gray bars with an asterisk, Table 4-1, 4-2). Among neuromodulatory manipulations, mutations in genes involved in PDF-1 signaling (*pdf-1, pdf-1; pdf-2,* and *pdfr-1*) all showed significant increases in lawn boundary distance and *pdfr-1* showed significant decreases in bacterial density.

Figure 4-5. Candidate genetic and neuronal inactivation reveals sensory and neuromodulatory control of lawn positioning and bacterial density experience.
(A) Lawn boundary distance plotted for different experimental groups targeting sensory neurons alongside wild type control data collected on the same day.
(B) Same as (A) for neuromodulatory mutants and their paired wild type controls.
(C) Bacterial density plotted for different experimental groups targeting sensory neurons alongside wild type control data collected on the same day.
(D) Same as (C) for neuromodulatory mutants and paired wild type controls.
Boxes indicate median and interquartile range. Asterisks indicate significant difference between 2 conditions using Mann-Whitney two-sided test before multiple hypothesis test correction. ns = no significant difference, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 10^{-4}$. Experimental groups shaded red and blue are significantly increased or decreased relative to wild type after Bonferroni multiple hypothesis test correction.

Figure 4-5

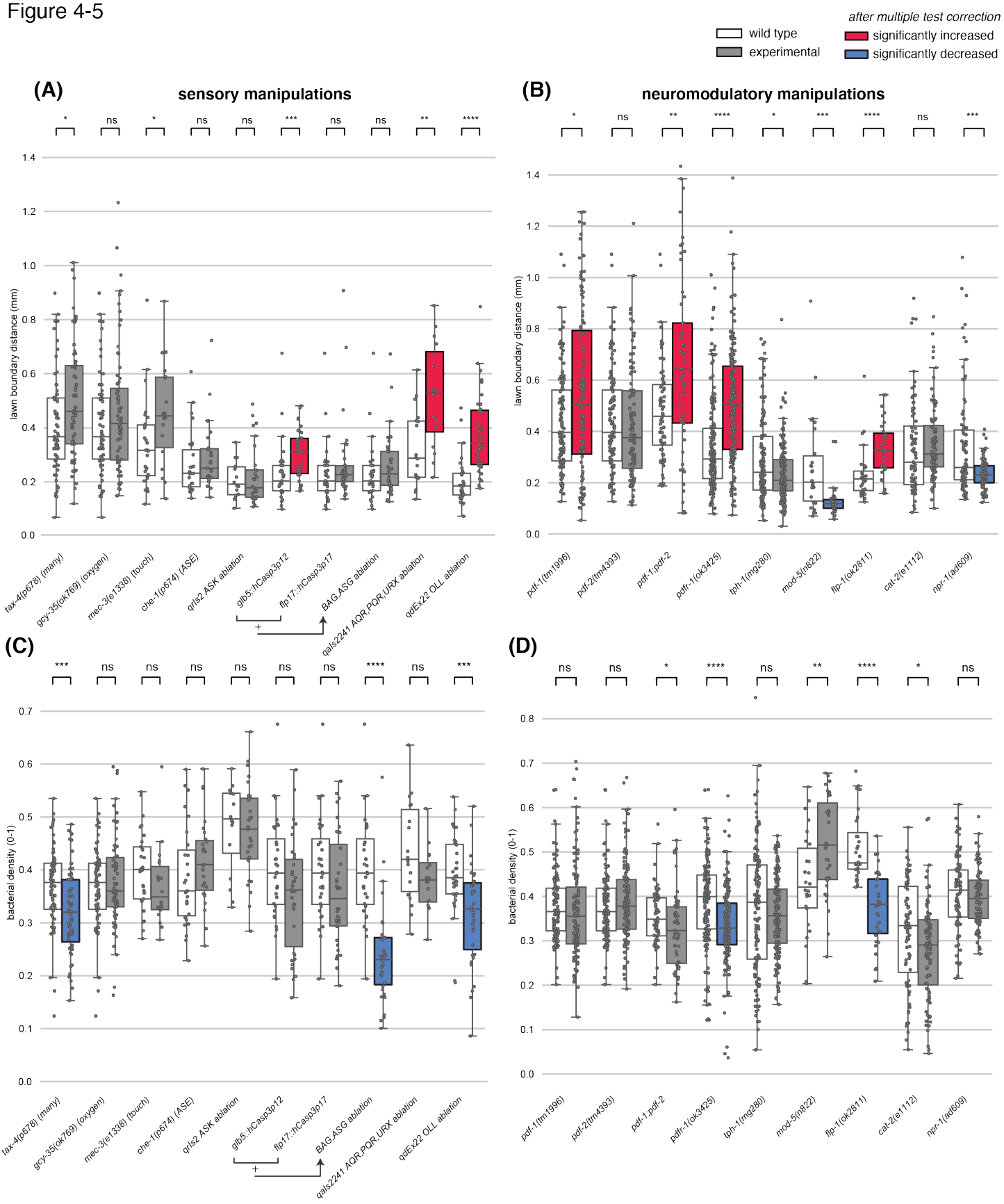

after multiple test correction
wild type
experimental
significantly increased
significantly decreased
(A)
sensory manipulations
(B)
neuromodulatory manipulations
(C)
(D)
lawn boundary distance (mm)
bacterial density (0-1)
tax-4(p678) (many)
gcy-35(ok769) (oxygen)
mec-3(e1338) (touch)
che-1(p674) (ASE)
qrIs2 ASK ablation
glb5::hCasp3p12
flp17::hCasp3p17
BAG,ASG ablation
qaIs2241 AQR,PQR,URX ablation
qdEx22 OLL ablation
pdf-1(tm1996)
pdf-2(tm4393)
pdf-1;pdf-2
pdfr-1(ok3425)
tph-1(mg280)
mod-5(n822)
flp-1(ok2811)
cat-2(e1112)
npr-1(ad609)

Table 4-1. Quantification of lawn boundary distance.
Mean lawn boundary distance (mm) for different experimental groups targeting sensory neurons and neuromodulators alongside wild type control data collected on the same day. Data from Figure 4-5 A-B. p values by Mann-Whitney two-sided test. Red, blue indicates significant increased and decreased p values after Bonferroni multiple hypothesis test correction.

genotype	n	lawn boundary dist	pvalue	adjusted pvalue
pdf-1(tm1996)	115	0.55	0.01	0.24
wild type	118	0.44		
pdf-2(tm4393)	112	0.43	0.23	1
wild type	118	0.44		
pdf-1;pdf-2	53	0.66	0.00	0.02
wild type	68	0.48		
pdfr-1(ok3425)	176	0.51	4.56E-14	8.66E-13
wild type	161	0.34		
tph-1(mg280)	139	0.24	0.01	0.19
wild type	127	0.30		
mod-5(n822)	35	0.13	0.00	0.00
wild type	26	0.26		
flp-1(ok2811)	35	0.33	3.19E-05	0.00
wild type	32	0.23		
cat-2(e1112)	88	0.36	0.04	0.73
wild type	76	0.33		
npr-1(ad609)	91	0.23	0.00	0.00
wild type	90	0.34		
tax-4(p678) (many)	63	0.50	0.01	0.17
wild type	67	0.41		
gcy-35(ok769) (oxygen)	68	0.46	0.15	1
wild type	67	0.41		
mec-3(e1338) (touch)	18	0.45	0.01	0.25
wild type	29	0.34		
che-1(p674) (ASE)	27	0.29	0.15	1
wild type	27	0.27		
qrls2 ASK ablation	29	0.22	0.37	1
wild type	16	0.21		
glb5::hCasp3p12	34	0.30	0.00	0.01
wild type	32	0.23		
flp17::hCasp3p17	36	0.27	0.08	1
wild type	32	0.23		
BAG,ASG ablation	35	0.26	0.09	1
wild type	32	0.23		
qaIs2241 AQR,PQR,URX ablation	16	0.52	0.00	0.02
wild type	18	0.32		
qdEx22 OLL ablation	35	0.38	3.50E-08	6.65E-07
wild type	32	0.20		

Table 4-2. Quantification of bacterial density.
Mean bacterial density for different experimental groups targeting sensory neurons and neuromodulators alongside wild type control data collected on the same day. Data from Figure 4-5 C-D. p values by Mann-Whitney two-sided test. Blue indicates significantly decreased p values after Bonferroni multiple hypothesis test correction.

genotype	n	bacterial density	pvalue	adjusted pvalue
pdf-1(tm1996)	115	0.37	0.04	0.76
wild type	118	0.38		
pdf-2(tm4393)	112	0.39	0.20	1
wild type	118	0.38		
pdf-1;pdf-2	53	0.33	0.01	0.26
wild type	68	0.35		
pdfr-1(ok3425)	176	0.34	1.24E-09	2.35E-08
wild type	161	0.39		
tph-1(mg280)	139	0.36	0.20	1.00
wild type	127	0.37		
mod-5(n822)	35	0.51	0.00	0.07
wild type	26	0.43		
flp-1(ok2811)	35	0.37	1.04E-08	1.98E-07
wild type	32	0.51		
cat-2(e1112)	88	0.28	0.01	0.16
wild type	76	0.32		
npr-1(ad609)	91	0.40	0.13	1.00
wild type	90	0.41		
tax-4(p678) (many)	63	0.32	1.89E-04	3.60E-03
wild type	67	0.37		
gcy-35(ok769) (oxygen)	68	0.38	0.37	1
wild type	67	0.37		
mec-3(e1338) (touch)	18	0.37	0.04	0.85
wild type	29	0.40		
che-1(p674) (ASE)	27	0.42	0.07	1
wild type	27	0.39		
qrIs2 ASK ablation	29	0.47	0.28	1
wild type	16	0.48		
glb5::hCasp3p12	34	0.35	0.06	1.00
wild type	32	0.40		
flp17::hCasp3p17	36	0.36	0.04	0.843664228
wild type	32	0.40		
BAG,ASG ablation	35	0.24	5.02E-09	9.53E-08
wild type	32	0.40		
qaIs2241 AQR,PQR,URX ablation	16	0.38	0.05	0.90
wild type	18	0.43		
qdEx22 OLL ablation	35	0.32	3.12E-04	5.94E-03
wild type	32	0.39		

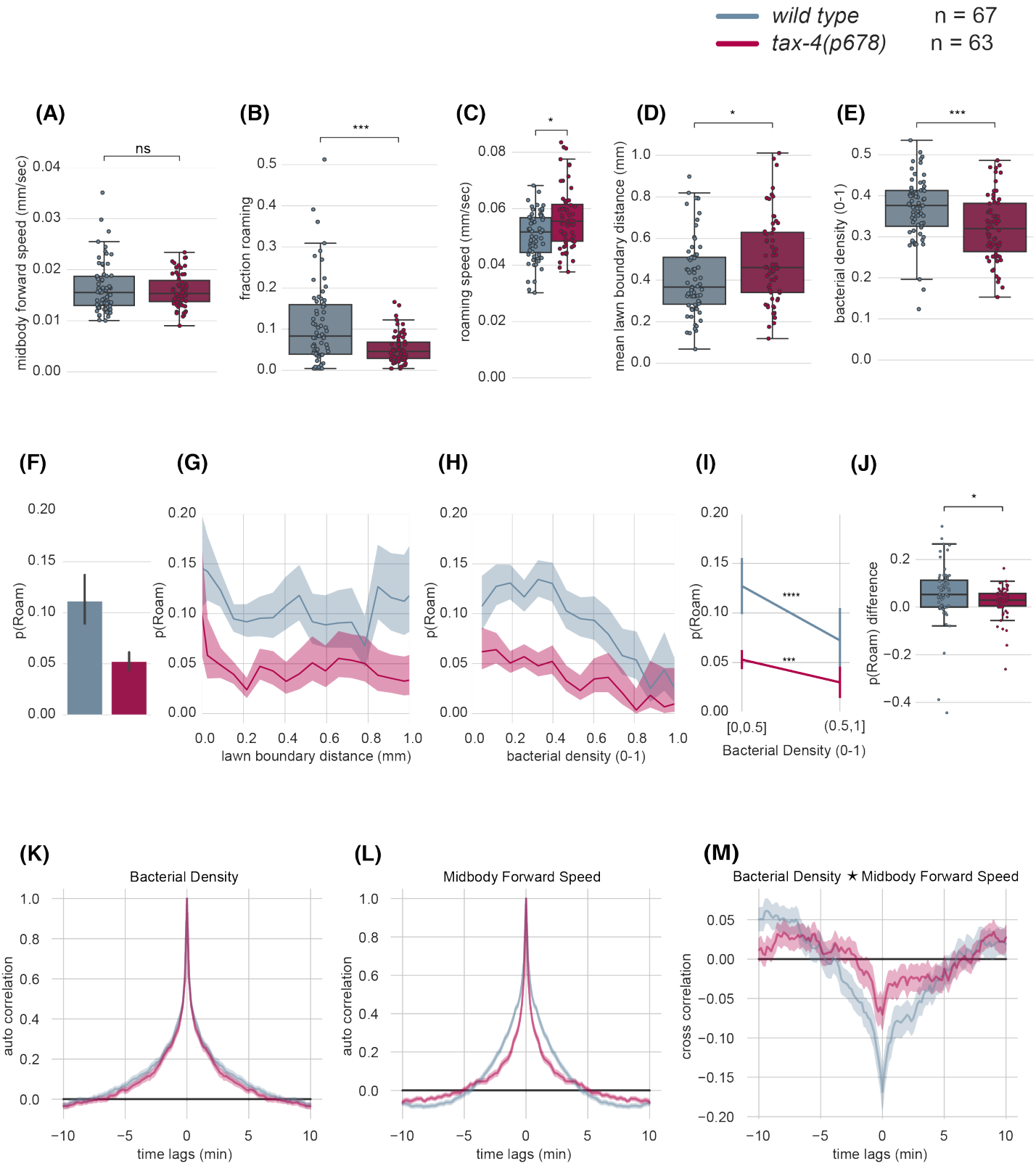
TAX-4 is required for entrainment of Roam states to bacterial density

As previously reported, *tax-4(lf)* animals roam less but have a similar average speed compared to wild type (Figure 4-6 A-B) (Coates and De Bono, 2002; Fujiwara et al., 2002). *tax-4(lf)* animals also have a slightly faster roaming speed than wild type (Figure 4-6 C), and resided at a higher mean lawn boundary distance and correspondingly lower bacterial density than wild type (Figure 4-6 D,E). Across intervals of increasing bacterial density, they showed attenuated suppression of roaming probability (Figure 4-6 H-J).

tax-4(lf) midbody forward speed showed a narrower autocorrelation function than wild type (this remains to be quantified), indicating that the speed of these animals lacks long term temporal structure observed in wild type (Figure 4-6 L). Autocorrelation of bacterial density was similar to wild type (Figure 4-6 K). The lower magnitude of cross-correlation between bacterial density and forward speed suggests that *tax-4(lf)* is less influenced by experienced changes in bacterial density than wild type (Figure 4-6 M). These results suggest that TAX-4-dependent sensory activity promotes the entrainment of animal locomotion speed to bacterial density.

Figure 4-6. Loss of *tax-4* signaling attenuates entrainment of Roam states to experienced bacterial density.
(A) Mean midbody forward speed per animal compared between wild type and *tax-4(lf)* animals.
(B) Same as (A) for fraction of time spent roaming per animal.
(C) Same as (A) for midbody forward speed of animals during roaming.
(D) Same as (A) for mean lawn boundary distance.
(E) Same as (A) for bacterial density.
(F) Overall probability of roaming "p(Roam)" is equivalent to fraction of time spent roaming across all animals.
(G) p(Roam) calculated in 30 contiguous bins across Lawn Boundary Distance ranging between 0 to 1 mm.
(H) p(Roam) calculated in 30 contiguous bins across Bacterial Density ranging between 0 to 1.
(I) Paired comparison of average p(Roam) between low Bacterial Density [0,0.5] and high Bacterial Density (0.5,1]. Vertical lines indicate 95% confidence intervals. Statistical test by Wilcoxon signed rank test.
(J) Difference in p(Roam) between low Bacterial Density [0,0.5] and high Bacterial Density (0.5,1]. Statistical test by Wilcoxon signed rank test.
(K) Mean autocorrelation functions for Bacterial Density.
(L) Mean autocorrelation functions for Midbody Forward Speed.
(M) Mean zero-normalized cross-correlation for Bacterial Density $\star$ Midbody Forward Speed.
For all boxplots: each dot represents one animal. Boxes indicate median and interquartile range. In panels A-E, p values are from Mann-Whitney two-sided test. ns = no significant difference, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 10^{-4}$.

Figure 4-6


wild type n = 67
tax-4(p678) n = 63
(A)
midbody forward speed (mm/sec)
ns
(B)
fraction roaming
***
(C)
roaming speed (mm/sec)
*
(D)
mean lawn boundary distance (mm)
*
(E)
bacterial density (0-1)
***
(F)
p(Roam)
(G)
p(Roam)
lawn boundary distance (mm)
(H)
p(Roam)
bacterial density (0-1)
(I)
p(Roam)
****
***
[0,0.5]
(0.5,1]
Bacterial Density (0-1)
(J)
p(Roam) difference
*
(K)
Bacterial Density
auto correlation
time lags (min)
(L)
Midbody Forward Speed
auto correlation
time lags (min)
(M)
Bacterial Density ★ Midbody Forward Speed
cross correlation
time lags (min)

PDF / PDFR-1 signaling is required for entrainment of Roam states to bacterial density

Consistent with previous results on uniform lawns (Flavell et al., 2013), *pdfr-1(lf)* animals showed a small but significant reduction in midbody forward speed and a strong suppression in time roaming (Figure 4-7 A-B). Unlike *tax-4(lf)* animals, *pdfr-1(lf)* animals also show reduced roaming speed (Figure 4-7 C). Compared to wild type animals, *pdfr-1(lf)* animals are found further inside the lawn (Figure 4-7 D). Correspondingly, these mutants also show a small but significant decrease in average bacterial density (Figure 4-7 E), as well as a reduced effect of bacterial density and lawn boundary distance on roaming probability (Figure 4-7 G-J).

Like *tax-4(lf)*, *pdfr-1(lf)* midbody forward speed autocorrelation was much narrower than wild type, indicating lost temporal structure, whereas bacterial density autocorrelation was similar to wild type (Figure 4-7 L). Strikingly, *pdfr-1(lf)* animals show an almost complete loss of cross-correlation between bacterial density and midbody forward speed (Figure 4-7 M), suggesting that the speed of these animals is relatively insensitive to changes in bacterial density. PDFR-1 signaling may promote both the generation of roaming states and also the entrainment of locomotion speed to bacterial density.

The PDFR-1 receptor in *C. elegans* has two known ligands, encoded by *pdf-1* and *pdf-2* (Janssen et al., 2008, 2009). Consistent with previous reports, I observed a robust reduction in midbody forward speed and fraction of time roaming in *pdf-1(lf)* mutants (Figure 4-8 A-B) (Flavell et al., 2013; Meelkop et al., 2012). In contrast to previous results, I also found that inactivation of *pdf-2* reduces speed and fraction of time roaming, although to a lesser extent than *pdf-1(lf)* (O'Donnell et al., 2018). *pdf-1; pdf-2* double mutant animals moved more slowly than *pdf-1(lf)* or *pdf-2(lf)* animals alone, suggesting an additive effect of PDF-1 and PDF-2 peptides on locomotion speed. Both *pdf-1(lf)* and *pdf-1; pdf-2* animals showed an almost complete absence of roaming.

pdf-1; pdf-2 double mutants were found at higher lawn boundary distance and lower bacterial density than wild type animals. Most of this effect was due to loss of function of *pdf-1*. *pdf-2(lf)* showed a reduced roaming probability but was still suppressed by high bacterial density and the lawn boundary (Figure 4-8 F-I). The low roaming probability in *pdf-1(lf)* and *pdf-1; pdf-2* animals precluded analysis of effects of lawn boundary distance and bacterial density on roaming in these animals. Together, these results emphasize the primary role of PDF-1 in roaming behaviors with a secondary role of PDF-2.

Moving from genetic analysis to site of action, the *pdfr-1* gene has at least two distinct promoters, one proximal to the translational start site that drives expression primarily in sensory neurons (OLL, URY, PHA, and others) and one located upstream of the first exon ("distal promoter") that drives expression mostly in interneurons (AIY, RIM, RIA

and others) (Barrios et al., 2012; Flavell et al., 2013) (Figure 4-9 A). Transgenic rescue of *pdfr-1(lf)* mutants with a 10kb genomic fragment that includes the proximal but not distal promoter ("PCR fragment #1") failed to rescue midbody forward speed, fraction roaming, or bacterial effects on roaming probability (Figure 4-9 B-E). By contrast, restoring expression using a 24kb genomic fragment spanning both distal and proximal promoters ("PCR fragment #2") fully rescued roaming probability and resulted in abnormally rapid midbody forward speed compared to either wild type or *pdfr-1(lf)* mutants. It is likely that this is a gain of function overexpression phenotype. Regulation of roaming probability by bacterial density was also restored, albeit to a higher baseline level than wild type (Figure 4-9 F-J). These results indicated that generation of Roam states and entrainment of roaming to bacterial density requires *pdfr-1* expression either from the distal promoter alone or a combination of the proximal and distal promoters.

To further define the necessary sites of *pdfr-1* expression for generation and bacterial entrainment of Roam states, I used the same intersectional strategy to rescue *pdfr-1* as was used in Chapter 3 (this time all experiments were performed in the *pdfr-1(lf)* background without *npr-1(lf))*. In these experiments, the presence of Cre activates transcription of *pdfr-1* cDNA from the *pdfr-1* distal promoter. Using a pan-neuronal promoter to drive Cre fully rescued midbody forward speed and fraction of time roaming but did not rescue modulation of *pdfr-1(lf)* roaming probability by lawn boundary distance and bacterial density (Figure 4-10 B-E). Following the results from *pdfr-1* rescue in the *npr-1(lf)* background, I found that restoring *pdfr-1* expression in RIB neurons resembled pan-neuronal rescue with the transgene, with normal forward speed and roaming fraction, but minimal modulation by bacterial density (Figure 4-10 F-G). These results indicate *pdfr-1* expression from its distal promoter promotes generation but not density-dependence of Roam states and highlights RIB as a prominent site of activity.

Figure 4-7. PDFR-1 signaling is required for density-dependent suppression of Roam states.
(A) Mean midbody forward speed per animal compared between wild type and *pdfr-1(lf)* animals.
(B) Same as (A) for fraction of time spent roaming per animal.
(C) Same as (A) for midbody forward speed of animals during roaming.
(D) Same as (A) for mean lawn boundary distance.
(E) Same as (A) for bacterial density.
(F) Overall probability of roaming "p(Roam)" is equivalent to fraction of time spent roaming across all animals.
(G) p(Roam) calculated in 30 contiguous bins across Lawn Boundary Distance ranging between 0 to 1 mm.
(H) p(Roam) calculated in 30 contiguous bins across Bacterial Density ranging between 0 to 1.
(I) Paired comparison of average p(Roam) between low Bacterial Density [0,0.5] and high Bacterial Density (0.5,1]. Vertical lines indicate 95% confidence intervals. Statistical test by Wilcoxon signed rank test.
(J) Difference in p(Roam) between low Bacterial Density [0,0.5] and high Bacterial Density (0.5,1]. Statistical test by Wilcoxon signed rank test.
(K) Mean autocorrelation functions for Bacterial Density.
(L) Mean autocorrelation functions for Midbody Forward Speed.
(M) Mean zero-normalized cross-correlation for Bacterial Density $\star$ Midbody Forward Speed.
For all boxplots: each dot represents one animal. Boxes indicate median and interquartile range. In panels A-E, p values are from Mann-Whitney two-sided test. ns = no significant difference, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 10^{-4}$.

Figure 4-7

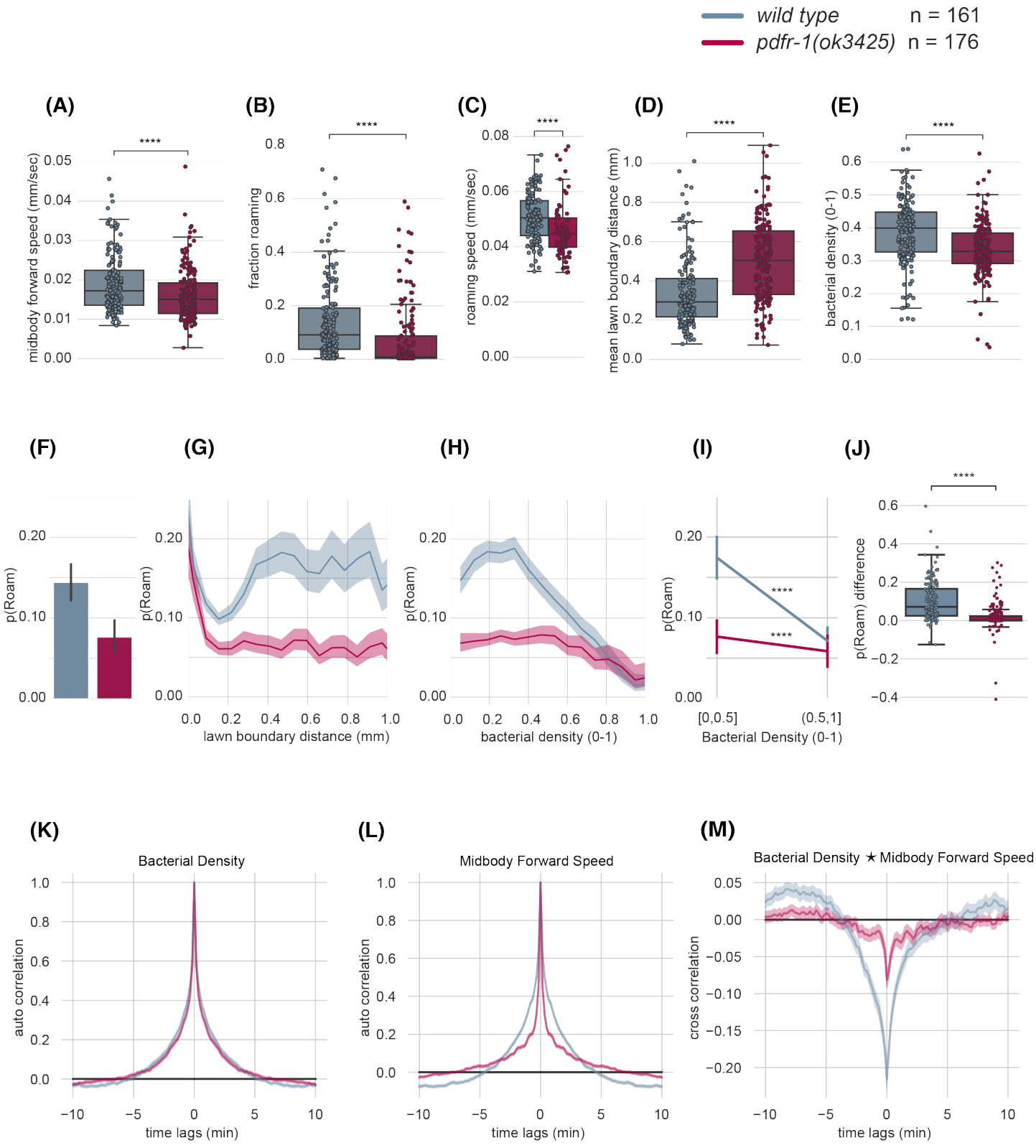

wild type n = 161
pdfr-1(ok3425) n = 176
(A)
midbody forward speed (mm/sec)
(B)
fraction roaming
(C)
roaming speed (mm/sec)
(D)
mean lawn boundary distance (mm)
(E)
bacterial density (0-1)
****
(F)
p(Roam)
(G)
lawn boundary distance (mm)
(H)
bacterial density (0-1)
(I)
Bacterial Density (0-1)
[0,0.5]
(0.5,1]
(J)
p(Roam) difference
(K)
Bacterial Density
auto correlation
time lags (min)
(L)
Midbody Forward Speed
(M)
Bacterial Density ★ Midbody Forward Speed
cross correlation

Figure 4-8. PDF-1 and PDF-2 act additively to generate Roam states.
(A) Mean midbody forward speed per wild type, *pdf-1*, *pdf-2*, and *pdf-1; pdf-2* animals.
(B) Same as (A) for fraction roaming per animal.
(C) Same as (A) for mean lawn boundary distance per animal.
(D) Same as (A) for mean bacterial density per animal.
(E) p(Roam) per genotype.
(F) p(Roam) calculated over 30 contiguous bins across Lawn Boundary Distance ranging between 0 to 1 mm.
(G) p(Roam) calculated over 30 contiguous bins across Bacterial Density ranging between 0 to 1.
(H) Paired comparison of average p(Roam) between low Bacterial Density [0,0.5] and high Bacterial Density (0.5,1]. Vertical lines indicate 95% confidence intervals. Statistical test by Wilcoxon signed rank test.
(I) Difference in p(Roam) between low Bacterial Density [0,0.5] and high Bacterial Density (0.5,1]. Statistical test using Kruskal-Wallis test with Dunn's multiple comparisons test.
For panels A-D Kruskal-Wallis test with Dunn's multiple comparisons test was used.
For all boxplots: each dot represents one animal. Boxes indicate median and interquartile range. ns = no significant difference, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 10^{-4}$.

Figure 4-8

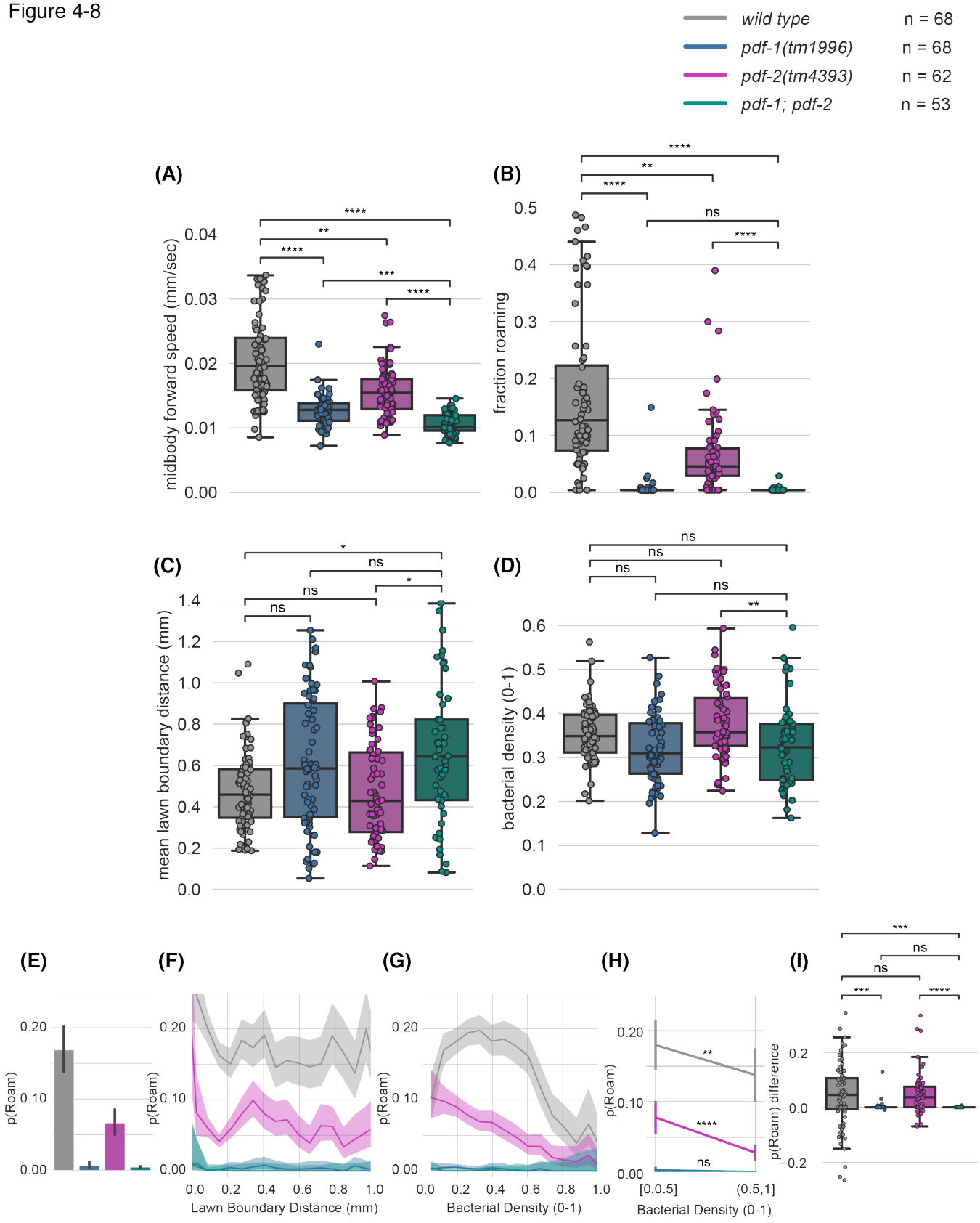

wild type n = 68
pdf-1(tm1996) n = 68
pdf-2(tm4393) n = 62
pdf-1; pdf-2 n = 53
(A)
midbody forward speed (mm/sec)
(B)
fraction roaming
(C)
mean lawn boundary distance (mm)
(D)
bacterial density (0-1)
(E)
p(Roam)
(F)
Lawn Boundary Distance (mm)
(G)
Bacterial Density (0-1)
(H)
[0,0.5]
(0.5,1]
Bacterial Density (0-1)
(I)
p(Roam) difference

Figure 4-9. *pdfr-1* expression from both proximal and distal promoters is required for generation and density-dependent entrainment of Roam states.
(A) Genomic region surrounding *pdfr-1* (thick boxes are exons; thin boxes are non-coding exons), region deleted in *ok3425* shown in blue, PCR fragments used for transgenic rescue and *pdfr-1* distal and proximal promoters. *Adapted from Flavell, et al (2013)*
(B) Mean midbody forward speed per wild type, *pdfr-1(ok3425)* and transgenic rescues with *pdfr-1* genomic fragments.
(C) Same as (B) for fraction roaming per animal.
(D) Same as (B) for mean lawn boundary distance per animal.
(E) Same as (B) for bacterial density per animal.
(F) p(Roam) per genotype.
(G) p(Roam) calculated over 30 contiguous bins across Lawn Boundary Distance ranging between 0 to 1 mm.
(H) p(Roam) calculated over 30 contiguous bins across Bacterial Density ranging between 0 to 1.
(I) Paired comparison of average p(Roam) between low Bacterial Density [0,0.5] and high Bacterial Density (0.5,1]. Vertical lines indicate 95% confidence intervals. Statistical test by Wilcoxon signed rank test.
(J) Difference in p(Roam) between low Bacterial Density [0,0.5] and high Bacterial Density (0.5,1]. Statistical test using Kruskal-Wallis test with Dunn's multiple comparisons test.
(B-E) p values tested by Mann-Whitney two-sided test.
For panels B-E, Kruskal-Wallis test with Dunn's multiple comparisons test was used.
For all boxplots: each dot represents one animal. Boxes indicate median and interquartile range. ns = no significant difference, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 10^{-4}$.

Figure 4-9

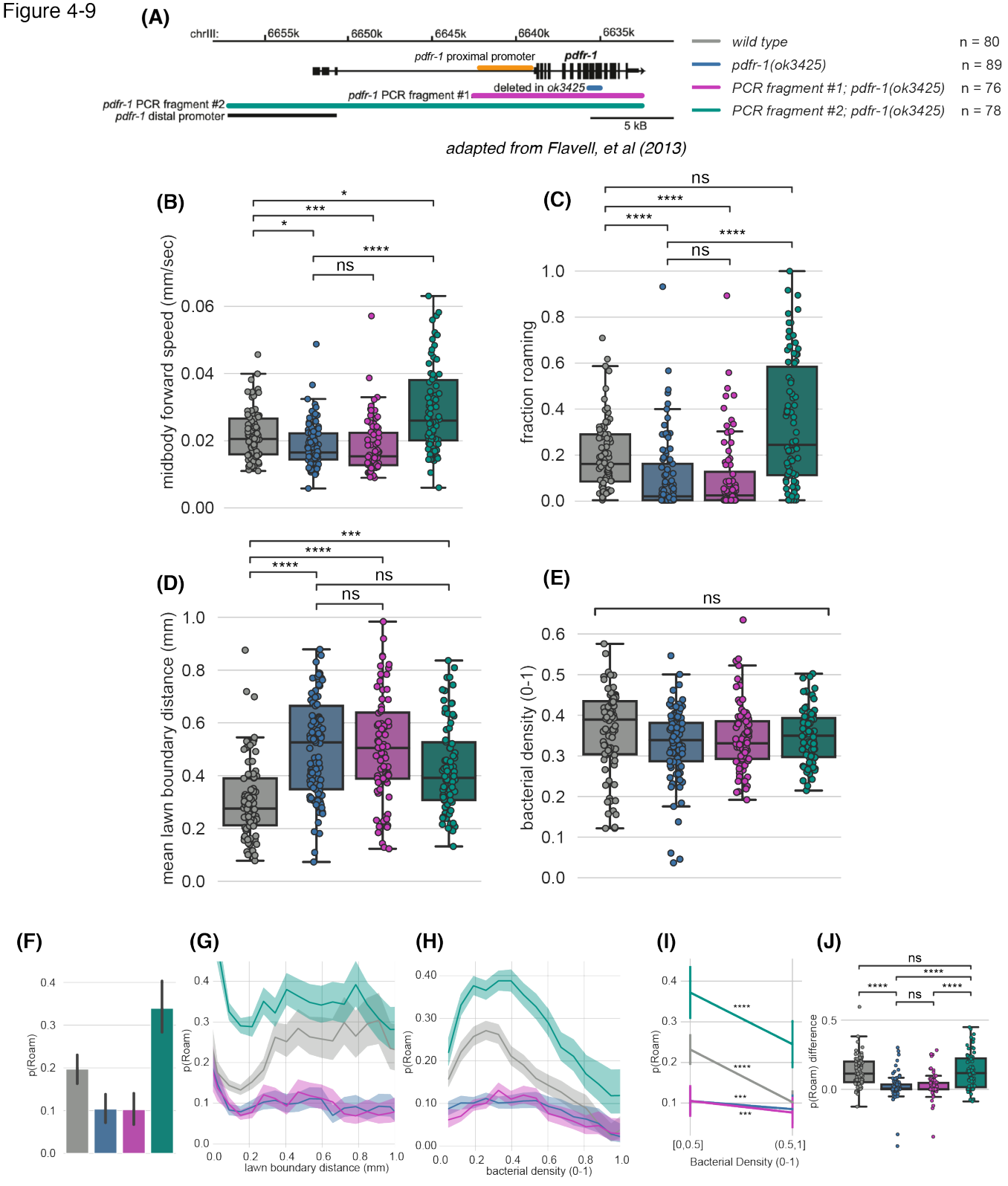
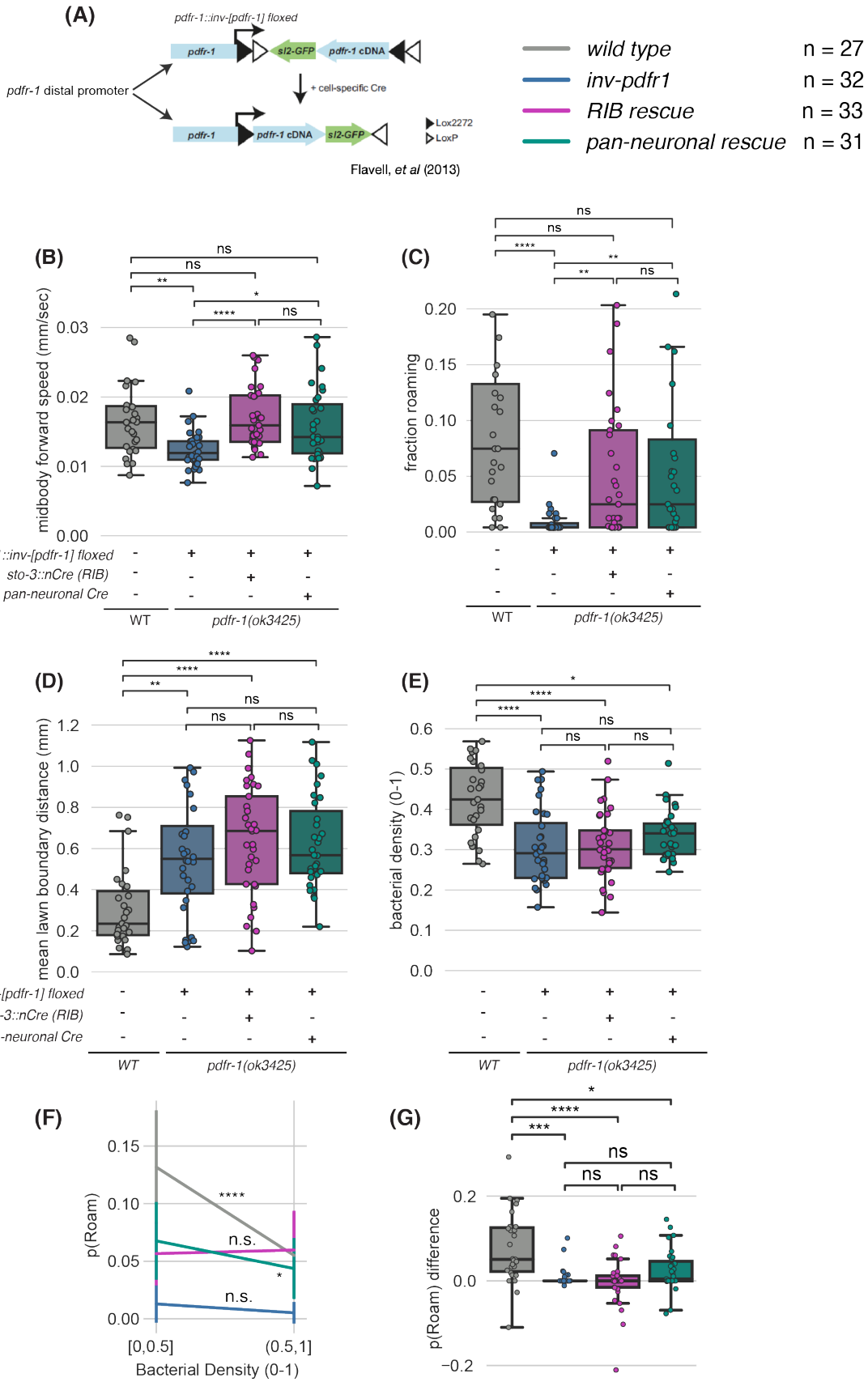

(A)
chrIII: 6655k 6650k 6645k 6640k 6635k
pdfr-1 proximal promoter
pdfr-1
deleted in ok3425
pdfr-1 PCR fragment #1
pdfr-1 PCR fragment #2
pdfr-1 distal promoter
5 kB
wild type n = 80
pdfr-1(ok3425) n = 89
PCR fragment #1; pdfr-1(ok3425) n = 76
PCR fragment #2; pdfr-1(ok3425) n = 78
adapted from Flavell, et al (2013)
(B)
midbody forward speed (mm/sec)
(C)
fraction roaming
(D)
mean lawn boundary distance (mm)
(E)
bacterial density (0-1)
(F)
p(Roam)
(G)
p(Roam)
lawn boundary distance (mm)
(H)
p(Roam)
bacterial density (0-1)
(I)
p(Roam)
[0,0.5]
(0.5,1]
Bacterial Density (0-1)
(J)
p(Roam) difference

Figure 4-10. *pdfr-1* expression in RIB interneurons is sufficient to generate Roam states but not to suppress them at high bacterial density.
(A) Intersectional cell-specific rescue of *pdfr-1* using a Cre-Lox strategy. One transgene carries an inverted floxed *pdfr-1* cDNA ("*inv[pdfr-1]*") driven by the *pdfr-1* distal promoter. A second transgene drives Cre in specific cell types to activate *pdfr-1* expression. *Adapted from (Flavell, et al 2013).*
(B) Mean midbody forward speed in *pdfr-1(ok3425)* animals carrying *inv[pdfr-1]* and/or pan-neuronal Cre or RIB-specific Cre.
(C) Same as (B) for fraction roaming per animal.
(D) Same as (B) for mean lawn boundary distance per animal.
(E) Same as (B) for mean bacterial density per animal.
(F) Paired comparison of average p(Roam) between low Bacterial Density [0,0.5] and high Bacterial Density (0.5,1]. Vertical lines indicate 95% confidence intervals. Statistical test by Wilcoxon signed rank test.
(G) Difference in p(Roam) between low Bacterial Density [0,0.5] and high Bacterial Density (0.5,1]. Statistical test using Kruskal-Wallis test with Dunn's multiple comparisons test.
For panels B-E, Kruskal-Wallis test with Dunn's multiple comparisons test was used.
For all boxplots: each dot represents one animal. Boxes indicate median and interquartile range. ns = no significant difference, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 10^{-4}$.

Figure 4-10


(A)
pdfr-1::inv-[pdfr-1] floxed
pdfr-1
sl2-GFP
pdfr-1 cDNA
pdfr-1 distal promoter
+ cell-specific Cre
pdfr-1
pdfr-1 cDNA
sl2-GFP
Lox2272
LoxP
Flavell, et al (2013)
wild type n = 27
inv-pdfr1 n = 32
RIB rescue n = 33
pan-neuronal rescue n = 31
(B)
midbody forward speed (mm/sec)
pdfr-1::inv-[pdfr-1] floxed
sto-3::nCre (RIB)
pan-neuronal Cre
WT
pdfr-1(ok3425)
(C)
fraction roaming
(D)
mean lawn boundary distance (mm)
(E)
bacterial density (0-1)
(F)
p(Roam)
n.s.
[0,0.5]
(0.5,1]
Bacterial Density (0-1)
(G)
p(Roam) difference

Discussion and Future Directions

In this chapter, I investigated *C. elegans'* behavioral responses to unevenly distributed food. Previous work has shown that animals spontaneously alternate between Roam and Dwell states on uniform bacterial lawns. The overall food concentration scales the propensity of animals to roam: higher concentration leads to less roaming and vice versa (Ben Arous et al., 2009). Here, I demonstrated that Roam and Dwell states can also be regulated by an animal's dynamically changing experience of bacterial density in environments where food distribution is non-uniform. Intriguingly, I observed that decreasing bacterial density reliably preceded the onset of Roam states, whereas increasing bacterial density preceded Dwell states.

Due to the radial symmetry of the small lawns used in these assays, some animals that initiate Roam states after experiencing low bacterial density moved through the center of the lawn and quickly reached high bacterial density on the opposite side lawn boundary. This triggered them to slow down and begin dwelling again. One question left unanswered by this work is how reliably changes in bacterial density elicit behavioral changes. Once roaming commences, is there a minimum amount of time animals are committed to that state or can they be immediately switched back to dwelling whenever they encounter a denser food patch? In subsequent experiments, I could measure behavior across environments with different spatial distributions of bacterial food to see if Roam state durations are always determined by changing sensory experience of bacterial density or if roaming states have internally generated characteristic durations.

Across genetic and neuronal manipulations, I found no consistent pattern relating changes in lawn leaving or fraction of time roaming with lawn boundary distance or bacterial density. While some mutants like *tax-4(lf)* and *pdfr-1(lf)* roamed less and attained lower bacterial density than wild type, others like *flp-1(lf)* and OLL-ablated animals roamed more and attained lower bacterial density than wild type (Figure 3-3, Figure 4-5). This suggests that genetic and neuronal mechanisms for sensing and responding to bacterial density are overlapping but not identical to those that generate locomotion speed and arousal states.

The observation that *tax-4(lf)* animals roam less often but move faster while roaming is reminiscent of their abrupt speed increase that precedes lawn leaving (Figure 3-4). The loss of sensory transduction in a large group of sensory neurons in *tax-4(lf)* mutants seems to dampen overall locomotory arousal while perhaps also releasing inhibition of sudden bursts of speed. In future experiments, I could perform cell-specific rescues of *tax-4* to identify important neuronal sources of sensory activity for bacterial density-dependent entrainment of Roam states and forward speed.

Previous work has highlighted the importance of PDFR-1 signaling in generating Roam states in uniform bacterial lawns (Flavell et al., 2013). Here I showed that PDFR-1 signaling may be required both for the generation of Roam states and their entrainment

to experienced changes in bacterial density. Although PDFR-1 ligands PDF-1 and PDF-2 both contribute to the generation of Roam states, PDF-1 plays a larger role in bacterial density-dependent entrainment. In future work, I would like to identify the relevant sources of PDF-1 and PDF-2 that promote this effect by cell-specific knockouts of *pdf-1* and *pdf-2.* Single-cell sequencing datasets reveal that *pdf-1* has a broader expression pattern than *pdf-2* although both genes are expressed in an overlapping subset of neurons (Taylor et al., 2021).

One important unresolved question is how PDF-1 and PDF-2 signaling impacts propagation of sensory information to affect behavior. In one model, these peptides could be tonically released by sensory neurons but suppressed when animals encounter dense food in order to promote dwelling. Another possibility is that PDF-1 and PDF-2 could be released by interneurons or motoneurons to potentiate sensory responses to dense food, thereby suppressing Roaming when dense food is encountered. Others have found that PDF-1/PDFR-1 signaling can affect gene expression in the ASJ sensory neurons, lending credence to the idea that centrally released PDF-1 can affect sensory neuron signaling (Hilbert and Kim, 2018).

Based on my results, expression of *pdfr-1* from a distal promoter is sufficient to rescue Roam states but leaves them untethered to changes in bacterial density. Previous work found that *pdfr-1* expression in AIY, RIM and RIA interneurons provided partial rescue of Roam states on uniform lawns (Flavell et al., 2013). Here, I found that restoring *pdfr-1* expression in RIB interneurons provided rescue commensurate to pan-neuronal rescue from the distal promoter. This result is consistent with RIB's known role in promoting forward locomotion and the suppression of RIB calcium activity before the onset of dwelling (Ji et al., 2021; Kato et al., 2015; Wang et al., 2020). In my experiments, *pdfr-1* expression in RIB did not rescue density-dependent roaming suppression, highlighting the possibility that PDFR-1 signaling may orchestrate the generation of Roam states and promote bacterial density-dependent entrainment of Roam states in separate sets of neurons perhaps driven by distinct *pdfr-1* promoters.

Expression of *pdfr-1* from the proximal promoter drives expression in a different subset of cells than the distal promoter. Since many of the cells driven by the proximal promoter are sensory neurons, it is plausible that *pdfr-1* in one or more of them could provide the bacterial density-entrainment effect (Barrios et al., 2012). Given the requirement of TAX-4 signaling in generating density-dependent Roam states, one hypothesis is that cells that express both *pdfr-1* and *tax-4* play a role. In future studies, I plan to test this hypothesis by restoring *pdfr-1* expression in both RIB and under the *tax-4* promoter to try to rescue density-dependent roaming. Another possibility is that different isoforms of PDFR-1 may be required for modulating Roaming by bacterial density. In this case, transgenes expressing these other isoforms may be required.

CHAPTER 5: SUMMARY AND FUTURE DIRECTIONS

Summary

In this thesis, I have used behavioral analyses, genetic manipulations, and perturbations of neural activity to investigate the behavioral states *C. elegans* adopts in an environment where food is unevenly distributed. Specifically, I have used Hidden Markov Models to infer the presence of behavioral states including the previously described roaming and dwelling states and an alternative set of states corresponding to different levels of locomotion speed on food. I found that animals in roaming or high-speed states show enhanced probability of leaving a bacterial food lawn and that lawn leaving events are predictable based on a stereotyped behavioral pattern, most notably identifiable by ramping locomotion speed. Animals in environments with lower food density or unable to ingest bacterial food increase roaming and lawn leaving, suggesting that integration of sensory and satiety signals generates behavioral arousal that leads to lawn leaving.

Across a set of genetic manipulations, I found that several mutants that inactivate sensory and neuromodulatory processes alter roaming and lawn leaving behavior. NPR-1 signaling in RMG neurons inhibits roaming and lawn leaving. In the *npr-1(lf)* background, animals roam more and leave food more readily; TAX-4-dependent signal transduction is required in sensory neurons for this effect. These results suggest that neural activity in sensory neurons promotes high arousal states in the *npr-1(lf)* background. Increased roaming and leaving is also dependent on PDF-1 neuropeptide signaling sensed through its receptor PDFR-1 in multiple interneurons, including RIB. Silencing RIB in *npr-1(lf)* animals decreases roaming and lawn leaving, consistent with its role in promoting forward locomotion and roaming.

I investigated how behavioral states were affected by changing sensory experience. I found that animals decreased roaming in regions of the environment where bacterial food was denser; in fact, roaming to dwelling transitions aligned well with increasing food abundance and vice versa. Again, I found that TAX-4 and PDFR-1 signaling were required for generating roaming behavior and modulating animal locomotion by bacterial density. I showed that PDFR-1 signaling in RIB interneurons contributes to the generation of roaming behavior but not bacterial density modulation, suggesting there may be other cells where PDFR-1 signaling entrains roaming to food abundance.

Model

To unify my results, I propose a model that relates bacterial density, Roam states and lawn leaving decisions (Figure 5-1). As shown in Chapter 2, Roam states give rise to high-speed locomotion on food and increase the probability of lawn leaving although they do not determine the exact timing of lawn leaving events. Lawn leaving events are a discrete decision behavior commonly associated with the fastest speeds achieved during roaming defined by AR-HMM state 3. My results in Chapter 3 indicate that TAX-4 signaling in sensory neurons suppresses lawn leaving and may determine the acute timing of lawn leaving events as animals interact with the lawn boundary.

Bacterial density is a complex sensory stimulus that can either promote or inhibit roam states depending on the context. Chemosensory cues from bacteria including odors and local concentrations of oxygen can be sensed by sensory neurons that use TAX-4 to transduce chemical signals to electrical activity. Since animals lacking functional *tax-4* show reduced roaming, TAX-4 signaling in these neurons normally promotes roaming. Noting the role of *pdfr-1* in coupling roaming to sensed changes in bacterial density and that *pdfr-1* is expressed in most *tax-4(+)* sensory neurons (Taylor et al., 2021), one hypothesis is that PDFR-1 signaling potentiates sensory activity in *tax-4(+)* cell(s) that promote roaming when food odors decrease, thereby coupling Roam states to bacterial density. As established in Chapter 4, PDFR-1 signaling also promotes roaming irrespective of experienced changes in bacterial density through its action in RIB interneurons. As established in Chapter 3, The presence of bacterial density can also suppress Roam states through internal sensing of food intake in the pharynx.

This model shows illuminates the way that *C. elegans* organizes its foraging into behavioral states according to levels of arousal, which can be coupled to changing sensory context by sensory and neuromodulatory signaling.

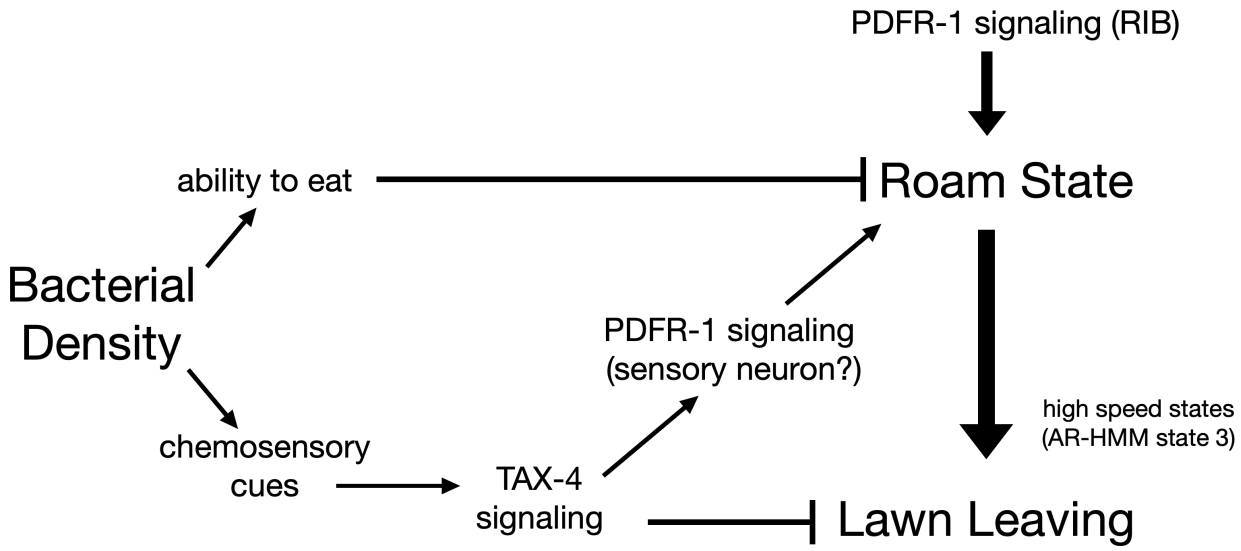


Figure 5-1. Model
Schematic depicting the relationships between bacterial density, Roaming and lawn leaving, connected by the proposed neuronal and molecular mechanisms.

Future direction: How does neural activity promote and maintain roaming in dynamic sensory contexts?

It remains unclear how animals can generate behavioral states that persist through time that are also sensitive to changing sensory experience. The internal states that give rise to behavioral states often correspond to neural activity patterns distributed across multiple neurons or brain areas. In *C. elegans*, the internal state of roaming is generated by release of PDF-1 and PDF-2 neuropeptide ligands by multiple cells that activate PDFR-1 receptors on another group of cells (Flavell et al., 2013). Among these cells, which ones contain the neural activity or biochemical processes that generate the persistence of roaming over minutes? Is it the same or a different set of cells and processes that entrain roaming states to sensory experience of bacterial density? How are these patterns integrated into a coherent internal state that generates behavioral output?

Transition into roaming behavior changes the activity of many neurons due to correlations with locomotion speed, but only a subset of these neurons are likely to actively generate the internal state of roaming (Hallinen et al., 2021; Ji et al., 2021; Kato et al., 2015). One possibility is that neurons that release PDF neuropeptide drive roaming state persistence. According to this model, the internal roaming state is active as long as these molecules are secreted. Lending credence to this hypothesis, excessive release of PDF-1 mediates the hyperactivity of *npr-1(lf)* animals during the larval molt before adulthood (Choi et al., 2013). Another model, which does not preclude the last, is that PDFR-1-expressing neurons generate persistent activity that outlasts initial activation by PDF peptide. This hypothesis is supported by the observation that brief optogenetic activation cAMP production in PDFR-1-expressing cells causes animals to roam for minutes after the stimulus ends (Flavell et al., 2013). However, the calcium activities of neurons that express *pdf-1* or *pdfr-1* remain correlated with roam states throughout their duration, making it difficult to discern if these models are correct. Nevertheless, decreasing activity in RIB neurons is a particularly good predictor of transitions into dwelling states, suggesting that its activity may drive roam state persistence (Ji et al., 2021).

In future experiments, I can test which neurons in the roaming circuit promote roaming state persistence by brief hyperpolarization of *pdf-1* or *pdfr-1*- expressing cells during roaming. Using a light-gated anion channel like GtACR1 (Yamanashi et al., 2019), I can deliver a brief inhibitory pulse to single neurons (RIB, for example) in a roaming animal to test whether acute silencing alters only locomotory behavior during the light pulse, or whether roam states are curtailed. If animals start to dwell, I can conclude that neural activity is required for roaming state persistence; if behavior is altered during the light pulse but does not lead to behavioral state changes, I will conclude that other neurons' activity is also required for roaming state persistence.

Once established, what are the neural circuit mechanisms that modulate roam states by experienced food density? Since TAX-4 signaling is required, the activity of one or more of the *tax-4* expressing sensory neurons must influence the core roaming circuitry that expresses *pdf-1* or *pdfr-1*. The neural activity that promotes roaming and dwelling is normally mutually exclusive, enforced by “winner take all” dynamics. In animals lacking *pdfr-1*, the activity of roaming- and dwelling-promoting neurons instead becomes coactive, degenerating the internal roam state (Ji et al., 2021). My results indicate that *pdfr-1* animals are unable to entrain their locomotion speed to changes in bacterial density, highlighting the possibility that the influence of food density on speed is gated by the internal roaming state. In this model, sensory information about food density can only influence behavior when the roaming state is intact. *pdfr-1* expression in RIB neurons was insufficient to entrain roaming to bacterial density although it was able to generate roaming behavior, suggesting that other sites of *pdfr-1* expression may be required to generate a functional roam state.

In future experiments, I will identify the *tax-4* neurons, whose activity modulates roaming by bacterial density by performing genetic rescue experiments. Among *tax-4*-expressing neurons, the activity of ASI and AWC sensory neurons have been shown to promote roaming (Ben Arous et al., 2009), suggesting they may also be required for bacterial density modulation. The *tax-4* cells that entrain roaming to bacterial density may be the same as the missing *pdfr-1*-expressing cells required for bacterial density-dependent roaming modulation—for example, AWC and ASI also express *pdfr-1*. If this is true, it would suggest an interesting model, in which PDF-1 promotes roaming behavior activating RIB interneurons, while simultaneously potentiating responses of sensory neurons that sense bacterial density. This hypothetical mechanism would provide one way to generate a persistent roaming state that remains flexible to dynamic changes in sensory context.

Methods

Nematode Culture

Animals were grown at room temperature (21-22°C on the laboratory bench) or in a 20°C incubator on nematode growth media (NGM; 51.3 mM NaCl, 1.7% agar, 0.25% peptone, 1 mM CaCl2, 12.9 µM cholesterol, 1 mM MgSO4, 25 mM KPO4, pH 6) plates seeded with 200 μL of an E. coli OP50 bacterial liquid culture (Brenner, 1974). This culture was grown at room temperature for 48 hours or overnight at 37ºC (without shaking) from a single colony of OP50 in 100 mL of sterile LB. This liquid culture was stored at 4ºC. The Bristol N2 strain was used as wild type throughout all experiments.

Strains

Transgenic animals were generated by microinjection of DNA containing the genetic construct of interest, a fluorescent co-injection marker (*myo-2p*::mCherry, *myo-3p*::mCherry, *elt-2p*::nGFP, *elt-2p*::mCherry), and empty pSM vector to reach a final DNA concentration of 100 ng/μL. Unless noted all mutant strains and strains containing integrated transgenes were backcrossed into wild-type to reduce background mutations.

Transgenic and mutant strains were always compared to control animals assayed in on the same day. For the candidate mutant screen (Figure 3-3, Figure 4-5), at least 16 animals were tested in each genotype, hypothesis tests were corrected for multiple comparisons using the Benjamini-Hochberg method.

See Appendix A for strain details.

Small lawn foraging assay

For all assays, *E. coli* OP50 was grown overnight in a shaking LB liquid culture from a single colony at 37ºC. On the morning of the assay, 400 μL of saturated liquid culture was diluted into 5 mL of LB and allowed to grow to OD1 at 37 ºC (takes ~1.5 hours), as measured by spectrophotometer. The liquid culture was then spun down and resuspended in M9 buffer (3 g KH_2PO_4, 6 g Na_2HPO_4, 5 g NaCl, 1 ml 1 M $MgSO_4$, H_2O to 1 liter) then concentrated to a density of OD2 (or 1x for OD1, 4x for OD4, as used in Figure 3-1). To generate the test lawns, 2 μL of this concentrated bacterial resuspension was seeded onto NGM agar within each well of a custom-made laser-cut multi-well plate (Stern et al., 2017). 50 μL of bacterial resuspension was seeded onto a separate NGM agar plate to be used as a food density acclimation plate. Lawns were grown at 20-22ºC for 2 hours. Adult hermaphrodites picked as L4s 16-20 hours before were then transferred to acclimation plates and left there for 45 minutes – 1.5 hours. Animals were then transferred to an unseeded NGM, cleaned of *E. coli*, and then transferred singly into each well of the assay plates. Animals were placed away on the bacteria-free agar and allowed to find the small food lawn on their own. Assays were recorded on to 4 computers for 1 hour at 3 frames per second using 12 8.8 MP USB3 cameras (Pointgrey, Flea3) and 35 mm high-resolution objectives (Edmund optics). LED backlights (Metaphase Technologies) provided uniform illumination of the assay plates. Commercial software (Flycapture, Pointgrey) was used to record the movies.

Uniform lawn assay

Assays testing worm behavior on uniform bacterial lawns were performed exactly as in the small lawn foraging assay with the exception that instead of 2 μL of bacteria seeded in the center of the well, 15 μL of OD2 bacterial suspension is spread evenly throughout the well. Bacteria grown for 2 hours before acclimation and starting the assay.

Behavioral tracking and lawn feature detection

Because each video recorded the behavior of up to 6 individual animals, videos were manually cropped so each surrounded just a single animal using FFmpeg software (Tomar, 2006). To extract animal positions and postures, captured movies were analyzed by custom made scripts in MATLAB (Mathworks, version 2021a) using the image processing toolbox, and computer vision toolbox. In each frame of the movie, the worm is segmented by background subtraction and its XY position is tracked over time using a Kalman filter. From the background-subtracted worm image, a smooth spline of 49 points is computationally applied and features relating to the movement of points along the body are derived following (Javer et al., 2018). Disambiguation of the head versus tail was determined by assigning the head as the end of the spline that had greater cumulative displacement over the video assay. This facilitated determination of times when the animal moved forward and backward.

The outline of the bacterial lawn was determined by edge detectors applied to the background averaged across movie frames. Across every frame, the closest boundary point to the animal's head was determined and used to calculate the "Lawn Boundary Distance." The lawn center point was calculated by averaging the X and Y positions of points on the lawn boundary and then used to compute the Radial Trajectory Angle, the angle formed between an animal's midbody trajectory through time and the lawn center point (0 = pointing toward lawn center, 180 = pointing directly outside the lawn).

The "Bacterial Density" experienced at the nose tip of an animal throughout the video was calculated as follows: 1) the minimum and maximum allowable grayscale values in the average background image of the video were calculated from pixels in the dense lawn boundary region and regions off food, respectively. 2) the raw grayscale value of pixels at the nose tip was extracted from the time-averaged background image. 3) these raw grayscale values are truncated to the range between the minimum and maximum grayscale values and scaled between 0 and 1.

Lawn Boundary Distance and movement direction were used to classify a set of lawn interaction behaviors. Head pokes were classified based on an excursion of the head lawn boundary distance that peaks outside the lawn before returning to the lawn interior. In the period following maximal displacement outside the lawn and before resuming locomotion inside the lawn ("recovery interval"), 3 types of head pokes were categorized: head poke forward, in which at least half of the recovery intervals is spent moving forward, head poke reversal, in which the animal executes a reversal during the recovery interval, or head poke pauses, in which an animal spends at least half of the recovery interval with speed less than 0.02 mm/sec. Lawn Exits were categorized when an animal's entire body translocated outside the lawn before returning.

Quality control for including animals in subsequent analyses

Behavior and lawn features were detected and tracked over the 1-hour assay but only the latter 40 minutes of data was retained for analysis in order to minimize the effects of mechanosensory stimulation from picking animals on recorded behavior. Data from single animals was only retained if in subsequent analyses if the following conditions were met: 1) the worm was visible in the video for at least half of the time the worm was recorded (cumulatively 30 minutes), 2) the worm was inside the bacterial lawn for at least one minute within the first 20 minutes of the video (before usable data was collected), 3) the plate was not bumped during the recording.

Quantitative Locomotion Analysis and HMM analyses

All quantitative analyses of locomotion and Hidden Markov Model building were performed in Python. For all analyses of animal locomotion and model-building, behavioral data from animals outside the food lawn was censored.

To classify roaming and dwelling states, speed and angular speed of animal centroid position was averaged into contiguous 10 second intervals. Roaming and dwelling states were classified as in (Flavell et al., 2013). Briefly, two classes of intervals corresponding to high angular speed / low speed and low angular speed / high speed were identified and separated by a line drawn at y = x/450. Behavior can then be instantaneously classified into roaming intervals when values are above the line, or dwelling intervals, when values are below the line. A categorical Hidden Markov Model was then trained on these behavioral sequences to generate roaming and dwelling hidden states using the SSM package (Linderman et al., 2020).

As described in Chapter 2, principal components analysis (PCA) was performed across 13 behavioral features to generate a 2-dimensional feature representation of animal locomotion using the scikit-learn package (Pedregosa et al., 2011). HMM analysis was performed on these principal component dimensions using the Gaussian observation model in the SSM package (Linderman et al., 2020).

An Autoregressive Hidden Markov Model (AR-HMM) was also trained to segment animal behavioral states on food using a different set of behavioral features: fraction of time moving forward per 10-sec bin, midbody forward speed, midbody angular speed, head angular velocity relative to midbody, head radial velocity relative to midbody (Chapter 2). Formally, at each time step, hidden states z_t follows Markovian dynamics, $z_{t+1} \mid z_t, \{\pi_k\}_{k=1}^K \sim \pi_{z_t}$, where $\{\pi_k\}_{k=1}^K$ is the transition matrix and $\pi_k \in [0,1]^K$ corresponds to the *k*-th row. Given hidden states z_t, the resulting feature dynamics are given by $x_t \mid x_{t-1}, z_t \sim \mathcal{N}\left(A_{z_t} x_{t-1} + b_{z_t}, Q_{z_t}\right)$, where $A_k, Q_k \in \mathbb{R}^{5x5}$ are linear dynamics and covariance matrices, respectively, and $b_k \in \mathbb{R}^5$ is the bias. The linear dynamics matrix *A* defines a set of equations that define a vector field in the feature dimensions. The bias term *b* can drive the dynamics in a particular direction. In the case where *A* is all zeros, the system has no dynamics and this amounts to a Hidden Markov Model with Gaussian emissions. AR-HMMs were also trained using the SSM package (Linderman et al., 2020). AR-HMM performance was evaluated by calculating the ratio of the log-

likelihood of held-out test data set using the AR-HMM same data under the AR-HMM and a multivariate Gaussian model $x_t \sim \mathcal{N}(\mu, \Sigma)$.

Generalized Linear Models

Generalized Linear Models (GLMs) can be expressed as $E(\mathbf{Y}|\mathbf{X}) = \mu = g^{-1}(\mathbf{X}\beta)$, where X are the features (also called predictors), Y is the response variable, *g* is the link function, and β are the coefficients or weights on the input features. Lawn exits were modeled as Bernoulli observations (0 = no exit, 1 = exit) and the link function we use is logit, $g(p) = ln(\frac{p}{1-p})$, where *p* is the probability of lawn exit. *g(p)* is therefore the log odds of lawn exit to no exit. GLMs were trained on a dataset using the statsmodels package (Seabold and Perktold, 2010) and evaluated for their prediction accuracy on a held-out test dataset to generate ROC curves.

Behavioral Optogenetic Manipulations

Pharyngeal pumping was inhibited using the the red-shifted channelrhodopsin ReaChR (Lin et al., 2013). ReaChR was cloned into the pSM vector and used the *myo-2* promoter to drive expression in pharyngeal muscle. Animals were stimulated while navigating small food lawns described above. Experimental animals were treated with 50 μM all-trans retinal overnight before assays. Control animals were placed on overnight lawns made in parallel that did not contain retinal. A 590 nm Precision LED with Uniform Illumination (Mightex) controlled with custom MATLAB software to deliver optogenetic stimuli was used. Animals were acclimated to small lawns for 20 minutes before being exposed to alternating 10-minute intervals of light OFF and light ON using 590 nm light at 4.5 mW/mm^2 (Figure 3-2). Recording hardware and software is identical to that of off-food foraging assays without optogenetic stimulation except that the optogenetic set up is only parallelized to accommodate 2 recordings at once.

Cell-specific knockouts of neuropeptide processing enzymes

To generate an RMG-specific knockout of *egl-3*, a strain containing a CRISPR-edited *egl-3* gene locus such that loxP sites flank the ORF (gift of D. Colon-Ramos, unpublished). To generate an RMG-specific knockout of *egl-21*, the loxP-flanked *egl-21* gene and its associated regulatory sequences was integrated by single copy Mos1 transposon-mediated insertion (MosSCI) into an *egl-21; npr-1* double null mutant (Frøkjær-Jensen et al., 2008).

pdfr-1 Rescue

The same rescue strains were tested for *pdfr-1* rescue using PCR'd genomic fragments (Figure 4-9) as were used in (Flavell et al., 2013). For cell-selective *pdfr-1* cDNA expression under the *pdfr-1* distal promoter using the floxed rescue construct (Figure 3-14, Figure 3-15, Figure 4-10), Cre expression was driven by a pan-neuronal promoter (*tag-168*), in glutamatergic cells (*eat-4*), in cholinergic cells (*cho-1*), in RIM (*tdc-1*), in AIY and others (*mod-1*), and in RIB (*sto-3*).

Acute histamine-mediated neuronal silencing

The *Drosophila* histamine-gated chloride channel HisCl1 was expressed under the *sto-3* promoter. Animals were treated with histamine following (Pokala et al., 2014). Histamine dihydrochloride (Sigma-Aldrich H7250) was dissolved in Milli-Q purified water, filtered, and stored at –20°C. Histamine solution was added to cooled NGM (45–50°C) for a final concentration of 10 mM to make assay plates. Animals were habituated on homogeneous OP50 lawns on histamine-free NGM plates, transferred to histamine-free NGM plates for cleaning, and then recorded on 10 mM histamine assay plates one hour during the small lawn foraging assay.

Appendix A: List of Strains Used

Strain name	Transgene, genotype, etc.	Figures	Notes
CX0001	wild type (N2) – lab strain	*	
PD1074	wild type (N2) – sequenced strain	* all figures	
CX16279	myo-2::ReaChR::sl2-GFP (1ng/uL), myo-3::mCherry (5ng/uL)	3-2	
CX14246	pdf-1(tm1996) LSC27 outcrossed to N2 2x	3-11, 4-8 *	
CX14484	pdf-2(tm4393)	4-8 *	
CX14893	pdf-1; pdf-2	4-8 *	
CX14295	pdfr-1(ok3425) VC2609 outcrossed to N2 5x	3-12, 4-7, 4-9 *	
MT15434	tph-1(mg280)II Backcrossed	*	
CX13630	mod-5 (n822) - MT8944 outcrossed to N2 3x	*	
CX17981	flp-1 (ok2811) 6x backrossed to N2	*	
CX11078	cat-2(e1112) II outcrossed 6x	*	
CX13663	npr-1(ad609)	3-5, 3-6, 3-11, 3-12 *	
CX13078	tax-4 (p678)	3-4,3-5, 3-6, 4-6 *	
CX6448	gcy-35(ok769)	3-6 *	
CX8767	mec-3(e1338) IV	*	
CX17674	che-1 (p674) backcrossed 6x to N2 (crossed out npr-1 215F)	*	
CX15305	qrIs2 ASKkill	*	
CX11693	kyIs538 (glb5::hCasp3 p12)	*	
CX11691	kyIs536 (BAG::hCasp3 p17)	*	
CX11697	kyIs536; kyIs538 (BAG and ASG caspase kill integrated)	*	
CX7102	lin-15(n765); qaIs2241 X	3-6 *	
ZD653	qdEx22 (OLL cell kill)	*	
CX4819	tax4 (p678); npr-1(ad609)	3-6	
CX7157	gcy-35(ok769); npr-1(ad609)	3-6	
CX7158	qaIs2241; npr-1(ad609)	3-6	
CX17802	pdf-1(tm1996); npr-1(ad609)	3-11	
CX15947	pdfr-1(ok3425); npr-1(ad609)	3-12	
CX13186	ncs-::nCre; npr-1(ad609), ofm-1::dsRed	3-7, 3-8	
CX13507	flp-21::LoxSTOPLox::npr-1; npr-1(ad609), elt-2::mcherry (1ng/uL)	3-7	
CX13186 x CX13507	RMG *npr-1* rescue; npr-1(ad609)	3-7	
CX13505	flp-21::LoxSTOPLox::TeTx; npr-1(ad609), elt-2::mcherry (1ng/uL)	3-8	
CX13186 x CX13505	RMG *TeTx*; npr-1(ad609)	3-8	
CX18134	flp-5p::nCre (20 ng/uL); elt-2p::nls-GFP (2 ng/uL) ; npr-1(ad609)	3-9	

CX18136	flp-5p::nCre (20 ng/uL) ; egl-3(nu1711); npr-1(ad609)	3-9	
N/A	egl-3(nu1711); npr-1(ad609)	3-9	CX18141 array negative animals used
CX18142	nlp-56::nCre (20ng/uL); elt-2p::nls-GFP (2 ng/uL); npr-1(ad609)	3-9	
CX18141	nlp-56::nCre (20ng/uL); elt-2p::nls-GFP (2 ng/uL); egl-3(nu1711); npr-1(ad609)	3-9	
CX18177	egl-21(n476); npr-1(ad609)	3-10	
CX18198	kySi61; egl-21(n476); npr-1(ad609)	3-10	
CX18201	kySi61 = mosSCI (Chr IV) of loxP-egl-21genomic-loxP; nlp-56::nCre (20 ng/ul), myo-3::mCherry (5 ng/ul); egl-21(n476); npr-1(ad609) (line 1)	3-10	
CX18199	kySi61 = mosSCI (Chr IV) of loxP-egl-21genomic-loxP; nlp-56::nCre (20 ng/ul), myo-3::mCherry (5 ng/ul); egl-21(n476); npr-1(ad609) (line 2)	3-10	
CX18158	nlp-56::pdf-1.1::sl2-GFP (1ng/uL); myo-3::mCherry (5ng/uL); pdf-1(tm1996); npr-1(ad609) (line 1)	3-13	
CX18156	nlp-56::pdf-1.1::sl2-GFP (1ng/uL); myo-3::mCherry (5ng/uL); pdf-1(tm1996); npr-1(ad609) (line 2)	3-13	
CX17754	pdfr-1(distal)::inv[pdfr-1::sl2GFP] (20ng/uL); myo-2::mCherry (1ng/uL); pdfr-1(ok3425); npr-1(ad609)	3-14, 3-15	
CX17755	h20::nCre (15ng/uL); myo-3::mCherry(5ng/uL); pdfr-1(distal)::inv[pdfr-1::sl2GFP] (20ng/uL); myo-2::mCherry (1ng/uL); pdfr-1(ok3425); npr-1(ad609)	3-14, 3-15	
CX17816	eat-4::nCre (5ng/uL); myo-3::mCherry(5ng/uL); pdfr-1(distal)::inv[pdfr-1::sl2GFP] (20ng/uL); myo-2::mCherry (1ng/uL); pdfr-1(ok3425); npr-1(ad609)	3-14	
CX18005	cho-1::nCre (25ng/uL); myo-3::mCherry (5ng/uL); pdfr-1(distal)::inv[pdfr-1::sl2GFP] (20ng/uL); myo-2::mCherry (1ng/uL); pdfr-1(ok3425); npr-1(ad609)	3-14	
CX18048	mod-1::nCre (18ng/uL), myo-3::mCherry (5ng/uL); pdfr-1(distal)::inv[pdfr-1::sl2GFP] (20ng/uL); myo-2::mCherry (1ng/uL); pdfr-1(ok3425); npr-1(ad609)	3-15	
CX17975	tdc-1::nCre (10ng/uL), myo-3::mCherry (5ng/uL); pdfr-1(distal)::inv[pdfr-1::sl2GFP] (20ng/uL); myo-2::mCherry (1ng/uL); pdfr-1(ok3425); npr-1(ad609)	3-15	
N/A	sto-3::nCre (40ng/uL), myo-3::mCherry (5ng/uL); pdfr-1(distal)::inv[pdfr-1::sl2GFP] (20ng/uL); myo-2::mCherry (1ng/uL); pdfr-1(ok3425); npr-1(ad609)	3-15, 3-16	strain lost

CX18340	sto-3::HisCl-sl2-mCherry (40 ng/uL), elt-2::mCherry (2 ng/uL); npr-1(ad609) (line 1)	3-17	
CX18341	sto-3::HisCl-sl2-mCherry (40 ng/uL), elt-2::mCherry (2 ng/uL); npr-1(ad609) (line 2)	3-17	
CX14378	PCR fragment #1 10kB genomic PCR of pdfr-1 (2.5 ng/uL); myo-3::GFP (25ng/uL); pdfr-1(ok3425)	4-9	
CX14383	PCR fragment #2 24kB genomic PCR of pdfr-1 (2.5 ng/uL); myo-3::GFP (25ng/uL); pdfr-1(ok3425)	4-9	
CX14485	pdfr-1(upstream)::inv[pdfr-1::sl2GFP] (20ng/uL); myo-2::mCherry (1ng/uL); pdfr-1(ok3425)	4-10	
CX18302	sto-3::nCre (40ng/uL), myo-3::mCherry (5ng/uL); pdfr-1(distal)::inv[pdfr-1::sl2GFP] (20ng/uL); myo-2::mCherry (1ng/uL); pdfr-1(ok3425); npr-1(ad609)	4-10	
CX14488	h20::nCre (15ng/uL); myo-3::mCherry(5ng/uL); pdfr-1(distal)::inv[pdfr-1::sl2GFP] (20ng/uL); myo-2::mCherry (1ng/uL); pdfr-1(ok3425); npr-1(ad609)	4-10	

* data is also found in: Figure 3-3, Figure 4-5, Table 3-1, Table 3-2, Table 4-1, Table 4-2.
The matched wild type control data in these figures was either CX0001 or PD1074 as follows:

genotype	**wild type used for comparison**
pdf-1(tm1996)	PD1074
pdf-2(tm4393)	PD1074
pdf-1;pdf-2	PD1074
pdfr-1(ok3425)	PD1074
tph-1(mg280)	CX0001
mod-5(n822)	CX0001
flp-1(ok2811)	CX0001
cat-2(e1112)	CX0001
npr-1(ad609)	PD1074
tax-4(p678) (many)	PD1074
gcy-35(ok769) (oxygen)	PD1074
mec-3(e1338) (touch)	CX0001
che-1(p674) (ASE)	CX0001
qrIs2 ASK ablation	CX0001
glb5::hCasp3p12	CX0001
flp17::hCasp3p17	CX0001
BAG,ASG ablation	CX0001
qaIs2241 AQR,PQR,URX ablation	CX0001
qdEx22 OLL ablation	CX0001

References

Adamah-Biassi, E.B., Stepien, I., Hudson, R.L., and Dubocovich, M.L. (2013). Automated video analysis system reveals distinct diurnal behaviors in C57BL/6 and C3H/HeN mice. Behav. Brain Res. *243*, 306–312.

Allen, W.E., Chen, M.Z., Pichamoorthy, N., Tien, R.H., Pachitariu, M., Luo, L., and Deisseroth, K. (2019). Thirst regulates motivated behavior through modulation of brainwide neural population dynamics. Science *364*, 253.

Anderson, D.J., and Adolphs, R. (2014). A framework for studying emotions across species. Cell *157*, 187–200.

Aponte, Y., Atasoy, D., and Sternson, S.M. (2011). AGRP neurons are sufficient to orchestrate feeding behavior rapidly and without training. Nat. Neurosci. *14*, 351–355.

Ariyasu, H., Takaya, K., Tagami, T., Ogawa, Y., Hosoda, K., Akamizu, T., Suda, M., Koh, T., Natsui, K., Toyooka, S., et al. (2001). Stomach Is a Major Source of Circulating Ghrelin, and Feeding State Determines Plasma Ghrelin-Like Immunoreactivity Levels in Humans. J. Clin. Endocrinol. Metab. *86*, 4753–4758.

Ben Arous, J., Laffont, S., and Chatenay, D. (2009). Molecular and Sensory Basis of a Food Related Two-State Behavior in C. elegans. PLoS One *4*, e7584.

Avery, L., and Shtonda, B.B. (2003). Food transport in the C. elegans pharynx. J. Exp. Biol. *206*, 2441–2457.

Baerends, G.P. (1976). The functional organization of behaviour. Anim. Behav. *24*, 726–738.

Barrios, A., Ghosh, R., Fang, C., Emmons, S.W., and Barr, M.M. (2012). PDF-1 neuropeptide signaling modulates a neural circuit for mate-searching behavior in C. elegans. Nat. Neurosci. *15*, 1675–1682.

Baum, L.E., and Petrie, T. (1966). Statistical Inference for Probabilistic Functions of Finite State Markov Chains. Ann. Math. Stat. *37*, 1554–1563.

Becker, J.D., Honerkamp, J., Hirsch, J., Frobe, U., Schlatter, E., and Greger, R. (1994). Analysing ion channels with hidden Markov models. Pflügers Arch. Eur. J. Physiol. *426*, 328–332.

Bendesky, A., Tsunozaki, M., Rockman, M. V., Kruglyak, L., and Bargmann, C.I. (2011). Catecholamine receptor polymorphisms affect decision-making in C. elegans. Nature *472*, 313–318.

Berman, G.J., Choi, D.M., Bialek, W., and Shaevitz, J.W. (2014). Mapping the

stereotyped behaviour of freely moving fruit flies. J. R. Soc. Interface *11*.

De Bono, M., and Bargmann, C.I. (1998). Natural variation in a neuropeptide Y receptor homolog modifies social behavior and food response in C. elegans. Cell *94*, 679–689.

De Bono, M., Tobin, D.M., Davis, M.W., Avery, L., and Bargmann, C.I. (2002). Social feeding in Caenorhabditis elegans is induced by neurons that detect aversive stimuli. Nature *419*, 899–903.

Brenner, S. (1974). The genetics of Caenorhabditis elegans. Genetics *77*, 71–94.

Buchanan, E.K., Linderman, S., and Paninski, L. (2017). Quantifying the behavioral dynamics of C . elegans with autoregressive hidden Markov models. 1–5.

Calhoun, A.J., Tong, A., Pokala, N., Fitzpatrick, J.A.J., Sharpee, T.O., and Chalasani, S.H. (2015). Neural mechanisms for evaluating environmental variability in caenorhabditis elegans. Neuron *86*, 428–441.

Calhoun, A.J., Pillow, J.W., and Murthy, M. (2019). Unsupervised identification of the internal states that shape natural behavior. Nat. Neurosci. *22*, 2040–2049.

Cermak, N., Yu, S.K., Clark, R., Huang, Y.C., Baskoylu, S.N., and Flavell, S.W. (2020). Whole-organism behavioral profiling reveals a role for dopamine in statedependent motor program coupling in C. Elegans. Elife *9*, 1–34.

Chalasani, S.H., Chronis, N., Tsunozaki, M., Gray, J.M., Ramot, D., Goodman, M.B., and Bargmann, C.I. (2007). Dissecting a circuit for olfactory behaviour in Caenorhabditis elegans. Nature *450*, 63–70.

Chang, H.C., Paek, J., and Kim, D.H. (2011). Natural polymorphisms in C. elegans HECW-1 E3 ligase affect pathogen avoidance behaviour. Nature *480*, 525–529.

Charnov, E.L. (1976). Optimal foraging, the marginal value theorem. Theor. Popul. Biol. *9*, 129–136.

Chen, Y., Lin, Y.C., Kuo, T.W., and Knight, Z.A. (2015). Sensory Detection of Food Rapidly Modulates Arcuate Feeding Circuits. Cell *160*, 829–841.

Chen, Z., Hendricks, M., Cornils, A., Maier, W., Alcedo, J., and Zhang, Y. (2013). Two Insulin-like Peptides Antagonistically Regulate Aversive Olfactory Learning in C. elegans. Neuron *77*, 572–585.

Cheung, B.H.H., Cohen, M., Rogers, C., Albayram, O., and De Bono, M. (2005). Experience-dependent modulation of C. elegans behavior by ambient oxygen. Curr. Biol. *15*, 905–917.

Choi, S., Chatzigeorgiou, M., Taylor, K.P., Schafer, W.R., and Kaplan, J.M. (2013).

Analysis of NPR-1 reveals a circuit mechanism for behavioral quiescence in C.elegans. Neuron *78*, 869–880.

Chung, S., Moore, J., Xia, L., Premkumar, L., and Gage, P. (1990). Characterization of single channel currents using digital signal processing techniques based on Hidden Markov Models. Philos. Trans. R. Soc. London. Ser. B Biol. Sci. *329*, 265–285.

Clark, J.T., Kalra, P.S., Crowley, W.R., and Kalra, S.P. (1984). Neuropeptide Y and human pancreatic polypeptide stimulate feeding behavior in rats. Endocrinology *115*, 427–429.

Clowney, E.J., Iguchi, S., Bussell, J.J., Scheer, E., and Ruta, V. (2015). Multimodal Chemosensory Circuits Controlling Male Courtship in Drosophila. Neuron *87*, 1036–1049.

Coates, J.C., and De Bono, M. (2002). Antagonistic pathways in neurons exposed to body fluid regulate social feeding in Caenorhabditis elegans. Nature *419*, 925–929.

Cook, S.J., Jarrell, T.A., Brittin, C.A., Wang, Y., Bloniarz, A.E., Yakovlev, M.A., Nguyen, K.C.Q., Tang, L.T.-H., Bayer, E.A., Duerr, J.S., et al. (2019). Whole-animal connectomes of both Caenorhabditis elegans sexes. Nature *571*, 63–71.

Corver, A., Wilkerson, N., Miller, J., and Gordus, A. (2021). Distinct movement patterns generate stages of spider web building. Curr. Biol. *31*, 4983-4997.e5.

Cowley, M.A., Smith, R.G., Diano, S., Tschöp, M., Pronchuk, N., Grove, K.L., Strasburger, C.J., Bidlingmaier, M., Esterman, M., Heiman, M.L., et al. (2003). The distribution and mechanism of action of ghrelin in the CNS demonstrates a novel hypothalamic circuit regulating energy homeostasis. Neuron *37*, 649–661.

Deegan, R.D., Bakajin, O., Dupont, T.F., Huber, G., Nagel, S.R., and Witten, T.A. (1997). Capillary flow as the cause of ring stains from dried liquid drops. Nature *389*, 827–829.

Dietrich, M.O., Zimmer, M.R., Bober, J., and Horvath, T.L. (2015). Hypothalamic Agrp neurons drive stereotypic behaviors beyond feeding. Cell *160*, 1222–1232.

Dobosiewicz, M., Liu, Q., and Bargmann, C.I. (2019). Reliability of an interneuron response depends on an integrated sensory state. Elife *8*, 1–27.

Flavell, S.W., Pokala, N., Macosko, E.Z., Albrecht, D.R., Larsch, J., and Bargmann, C.I. (2013). Serotonin and the neuropeptide PDF initiate and extend opposing behavioral states in C. Elegans. Cell *154*, 1023–1035.

Frøkjær-Jensen, C., Wayne Davis, M., Hopkins, C.E., Newman, B.J., Thummel, J.M., Olesen, S.P., Grunnet, M., and Jorgensen, E.M. (2008). Single-copy insertion of

transgenes in Caenorhabditis elegans. Nat. Genet. *40*, 1375–1383.

Fujiwara, M., Sengupta, P., and McIntire, S.L. (2002). Regulation of body size and behavioral state of C. elegans by sensory perception and the egl-4 cGMP-dependent protein kinase. Neuron *36*, 1091–1102.

Gallagher, T., Bjorness, T., Greene, R., You, Y.-J., and Avery, L. (2013). The Geometry of Locomotive Behavioral States in C. elegans. PLoS One *8*, e59865.

Geng, W., Cosman, P., Berry, C.C., Feng, Z., and Schafer, W.R. (2004). Automatic tracking, feature extraction and classification of C. elegans phenotypes. IEEE Trans. Biomed. Eng. *51*, 1811–1820.

Golden, J.W., and Riddle, D.L. (1982). A Pheromone Influences Larval Development in the Nematode Caenorhabditis elegans. Science (80-.). *218*, 578–580.

Golden, J.W., and Riddle, D.L. (1984). The Caenorhabditis elegans dauer larva: Developmental effects of pheromone, food, and temperature. Dev. Biol. *102*, 368–378.

Gray, J.M., Karow, D.S., Lu, H., Chang, A.J., Chang, J.S., Ellis, R.E., Marietta, M.A., and Bargmann, C.I. (2004). Oxygen sensation and social feeding mediated by a C. elegans guanylate cyclase homologue. Nature *430*, 317–322.

Gray, J.M., Hill, J.J., and Bargmann, C.I. (2005). A circuit for navigation in Caenorhabditis elegans. Proc. Natl. Acad. Sci. *102*, 3184–3191.

Hallinen, K.M., Dempsey, R., Scholz, M., Yu, X., Linder, A., Randi, F., Sharma, A., Shaevitz, J.W., and Leifer, A.M. (2021). Decoding locomotion from population neural activity in moving C. Elegans. Elife *10*, 1–28.

Hardege, I., Morud, J., Yu, J., Wilson, T., Schroeder, F., and Schafer, W. (2021). Neuronally Produced Betaine Acts via a Novel Ligand Gated Ion Channel to Control Behavioural States.

Hilbert, Z.A., and Kim, D.H. (2018). PDF-1 neuropeptide signaling regulates sexually dimorphic gene expression in shared sensory neurons of C. Elegans. Elife *7*.

Hums, I., Riedl, J., Mende, F., Kato, S., Kaplan, H.S., Latham, R., Sonntag, M., Traunmüller, L., and Zimmer, M. (2016). Regulation of two motor patterns enables the gradual adjustment of locomotion strategy in Caenorhabditis elegans. Elife *5*, 1–36.

Husson, S.J., Costa, W.S., Schmitt, C., and Gottschalk, A. (2012). Keeping track of worm trackers. WormBook 1–17.

Inagaki, H.K., Ben-Tabou De-Leon, S., Wong, A.M., Jagadish, S., Ishimoto, H., Barnea, G., Kitamoto, T., Axel, R., and Anderson, D.J. (2012). Visualizing neuromodulation in vivo: TANGO-mapping of dopamine signaling reveals appetite control of sugar sensing.

Cell *148*, 583–595.

Inagaki, H.K., Jung, Y., Hoopfer, E.D., Wong, A.M., Mishra, N., Lin, J.Y., Tsien, R.Y., and Anderson, D.J. (2014). Optogenetic control of Drosophila using a red-shifted channelrhodopsin reveals experience-dependent influences on courtship. Nat. Methods *11*, 325–332.

Jacob, T.C., and Kaplan, J.M.. (2003). The EGL-21 Carboxypeptidase E Facilitates Acetylcholine Release at Caenorhabditis elegans Neuromuscular Junctions. J. Neurosci. *23*, 2122–2130.

Jang, H., Levy, S., Flavell, S.W., Mende, F., Latham, R., Zimmer, M., and Bargmann, C.I. (2017). Dissection of neuronal gap junction circuits that regulate social behavior in Caenorhabditis elegans. Proc. Natl. Acad. Sci. *114*, E1263–E1272.

Janssen, T., Husson, S.J., Lindemans, M., Mertens, I., Rademakers, S., Donck, K. Ver, Geysen, J., Jansen, G., and Schoofs, L. (2008). Functional Characterization of Three G Protein-coupled Receptors for Pigment Dispersing Factors in Caenorhabditis elegans.

Janssen, T., Husson, S.J., Meelkop, E., Temmerman, L., Lindemans, M., Verstraelen, K., Rademakers, S., Mertens, I., Nitabach, M., Jansen, G., et al. (2009). Discovery and characterization of a conserved pigment dispersing factor-like neuropeptide pathway in Caenorhabditis elegans. J. Neurochem. *111*, 228–241.

Javer, A., Ripoll-Sánchez, L., and Brown, A.E.X. (2018). Powerful and interpretable behavioural features for quantitative phenotyping of Caenorhabditis elegans. Philos. Trans. R. Soc. B Biol. Sci.

Ji, N., Madan, G.K., Fabre, G.I., Dayan, A., Baker, C.M., Kramer, T.S., Nwabudike, I., and Flavell, S.W. (2021). A neural circuit for flexible control of persistent behavioral states. Elife *10*.

Jones, L.M., Fontanini, A., Sadacca, B.F., Miller, P., and Katz, D.B. (2007). Natural stimuli evoke dynamic sequences of states in sensory cortical ensembles. Proc. Natl. Acad. Sci. U. S. A. *104*, 18772–18777.

Kabra, M., Robie, A.A., Rivera-Alba, M., Branson, S., and Branson, K. (2013). JAABA: Interactive machine learning for automatic annotation of animal behavior. Nat. Methods *10*, 64–67.

Kass, J., Jacob, T.C., Kim, P., and Kaplan, J.M. (2001). The EGL-3 proprotein convertase regulates mechanosensory responses of Caenorhabditis elegans. J. Neurosci. *21*, 9265–9272.

Kato, S., Kaplan, H.S., Schrödel, T., Skora, S., Lindsay, T.H., Yemini, E., Lockery, S., and Zimmer, M. (2015). Global Brain Dynamics Embed the Motor Command Sequence

of Caenorhabditis elegans. Cell *163*, 656–669.

Labrousse, A., Chauvet, S., Couillault, C., Léopold Kurz, C., and Ewbank, J.J. (2000). Caenorhabditis elegans is a model host for Salmonella typhimurium. Curr. Biol. *10*, 1543–1545.

Laurent, P., Soltesz, Z., Nelson, G.M., Chen, C., Arellano-Carbajal, F., Levy, E., and De Bono, M. (2015). Decoding a neural circuit controlling global animal state in C. elegans. Elife *4*.

Laws, T.R., Atkins, H.S., Atkins, T.P., and Titball, R.W. (2006). The pathogen Pseudomonas aeruginosa negatively affects the attraction response of the nematode Caenorhabditis elegans to bacteria. Microb. Pathog. *40*, 293–297.

Li, C. (2008). Neuropeptides. WormBook 1–36.

Lin, J.Y., Knutsen, P.M., Muller, A., Kleinfeld, D., and Tsien, R.Y. (2013). ReaChR: A red-shifted variant of channelrhodopsin enables deep transcranial optogenetic excitation. Nat. Neurosci. *16*, 1499–1508.

Linderman, S.W. (2016). Bayesian Methods for Discovering Structure in Neural Spike Trains.

Linderman, S., Antin, B., Zoltowski, D., and Glaser, J. (2020). SSM: Bayesian Learning and Inference for State Space Models.

López-Cruz, A., Sordillo, A., Pokala, N., Liu, Q., McGrath, P.T., and Bargmann, C.I. (2019). Parallel Multimodal Circuits Control an Innate Foraging Behavior. Neuron *102*, 407-419.e8.

Macosko, E.Z. (2009). The Neural Circuitry Of Social Behavior In C . elegans. The Rockefeller University, New York.

Macosko, E.Z., Pokala, N., Feinberg, E.H., Chalasani, S.H., Butcher, R.A., Clardy, J., and Bargmann, C.I. (2009). A hub-and-spoke circuit drives pheromone attraction and social behaviour in C. elegans. Nature *458*, 1171–1175.

Marder, E. (2012). Neuromodulation of Neuronal Circuits: Back to the Future. Neuron *76*, 1–11.

Marder, E., and Bucher, D. (2007). Understanding circuit dynamics using the stomatogastric nervous system of lobsters and crabs. Annu. Rev. Physiol. *69*, 291–316.

Mathis, A., Mamidanna, P., Cury, K.M., Abe, T., Murthy, V.N., Mathis, M.W., and Bethge, M. (2018). DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. Nat. Neurosci. *21*, 1281–1289.

McCloskey, R.J., Fouad, A.D., Churgin, M.A., and Fang-Yen, C. (2017). Food responsiveness regulates episodic behavioral states in Caenorhabditis elegans. J Neurophysiol *117*, 1911–1934.

McEwan, D.L., Kirienko, N. V., and Ausubel, F.M. (2012). Host translational inhibition by Pseudomonas aeruginosa exotoxin A triggers an immune response in Caenorhabditis elegans. Cell Host Microbe *11*, 364–374.

McGrath, P.T., Rockman, M. V., Zimmer, M., Jang, H., Macosko, E.Z., Kruglyak, L., and Bargmann, C.I. (2009). Quantitative Mapping of a Digenic Behavioral Trait Implicates Globin Variation in C. elegans Sensory Behaviors. Neuron *61*, 692–699.

McKellar, A.E., Langrock, R., Walters, J.R., and Kesler, D.C. (2015). Using mixed hidden Markov models to examine behavioral states in a cooperatively breeding bird. Behav. Ecol. *26*, 148–157.

Meelkop, E., Temmerman, L., Janssen, T., Suetens, N., Beets, I., Van Rompay, L., Shanmugam, N., Husson, S.J., and Schoofs, L. (2012). PDF receptor signaling in Caenorhabditis elegans modulates locomotion and egg-laying. Mol. Cell. Endocrinol. *361*, 232–240.

Melo, J.A., and Ruvkun, G. (2012). Inactivation of conserved C. elegans genes engages pathogen- and xenobiotic-associated defenses. Cell *149*, 452–466.

Meunier, N., Belgacem, Y.H., and Martin, J.-R. (2007). Regulation of feeding behaviour and locomotor activity by takeout in Drosophila. J. Exp. Biol. *210*, 1424–1434.

Milward, K., Busch, K.E., Murphy, R.J., De Bono, M., and Olofsson, B. (2011). Neuronal and molecular substrates for optimal foraging in Caenorhabditis elegans. Proc. Natl. Acad. Sci. U. S. A. *108*, 20672–20677.

Morales, J.M., Haydon, D.T., Frair, J., Holsinger, K.E., and Fryxell, J.M. (2004). Extracting more out of relocation data: building movement models as mixtures of random walks. Ecology *85*, 2436–2445.

Nusbaum, M.P., and Beenhakker, M.P. (2002). A small-systems approach to motor pattern generation. Nature *417*, 343–350.

O'Donnell, M.P., Chao, P.H., Kammenga, J.E., and Sengupta, P. (2018). Rictor/TORC2 mediates gut-to-brain signaling in the regulation of phenotypic plasticity in C. elegans. PLoS Genet. *14*, 1–27.

Ollmann, M.M., Wilson, B.D., Yang, Y.-K., Kerns, J.A., Chen, Y., Gantz, I., and Barsh, G.S. (1997). Antagonism of Central Melanocortin Receptors in Vitro and in Vivo by Agouti-Related Protein. Science (80-.). *278*, 135–138.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., et al. (2011). Scikit-learn: Machine Learning in {P}ython. J. Mach. Learn. Res. *12*, 2825–2830.

Pereira, T.D., Aldarondo, D.E., Willmore, L., Kislin, M., Wang, S.S.H., Murthy, M., and Shaevitz, J.W. (2019). Fast animal pose estimation using deep neural networks. Nat. Methods *16*, 117–125.

Pokala, N., Liu, Q., Gordus, A., and Bargmann, C.I. (2014). Inducible and titratable silencing of Caenorhabditis elegans neurons in vivo with histamine-gated chloride channels. Proc. Natl. Acad. Sci. U. S. A. *111*, 2770–2775.

Pradel, E., Zhang, Y., Pujol, N., Matsuyama, T., Bargmann, C.I., and Ewbank, J.J. (2007). Detection and avoidance of a natural product from the pathogenic bacterium Serratia marcescens by Caenorhabditis elegans. Proc. Natl. Acad. Sci. U. S. A. *104*, 2295–2300.

Pradhan, S., Quilez, S., Homer, K., and Hendricks, M. (2019). Environmental Programming of Adult Foraging Behavior in C. elegans. Curr. Biol. *29*, 2867-2879.e4.

Rabiner, L.R. (1989). Tutorial on Hmm and Applications.Pdf. Proc. IEEE *77*, 257–286.

Rhoades, J.L., Nelson, J.C., Nwabudike, I., Yu, S.K., McLachlan, I.G., Madan, G.K., Abebe, E., Powers, J.R., Colón-Ramos, D.A., and Flavell, S.W. (2019). ASICs Mediate Food Responses in an Enteric Serotonergic Neuron that Controls Foraging Behaviors. Cell *176*, 85-97.e14.

Root, C.M., Ko, K.I., Jafari, A., and Wang, J.W. (2011). Presynaptic facilitation by neuropeptide signaling mediates odor-driven food search. Cell *145*, 133–144.

Sawin, E.R., Ranganathan, R., and Horvitz, H.R. (2000). C. elegans locomotory rate is modulated by the environment through a dopaminergic pathway and by experience through a serotonergic pathway. Neuron *26*, 619–631.

Scheiner, R., Sokolowski, M.B., and Erber, J. (2004). Activity of cGMP-Dependent Protein Kinase (PKG) Affects Sucrose Responsiveness and Habituation in Drosophila melanogaster. Learn. Mem. *11*, 303–311.

Schwarz, R.F., Branicky, R., Grundy, L.J., Schafer, W.R., and Brown, A.E.X. (2015). Changes in Postural Syntax Characterize Sensory Modulation and Natural Variation of C. elegans Locomotion. PLoS Comput. Biol. *11*, 1–16.

Scott, E., Hudson, A., Feist, E., Calahorro, F., Dillon, J., De, R., Wand, M., Schoofs, L., Connor, V.O., and Holden-dye, L. (2017). An oxytocin-dependent social interaction between larvae and adult C . elegans. 1–13.

Seabold, S., and Perktold, J. (2010). statsmodels: Econometric and statistical modeling with python. In 9th Python in Science Conference, p.

Shtonda, B.B. (2006). Dietary choice behavior in Caenorhabditis elegans. J. Exp. Biol. *209*, 89–102.

Singh, J., and Aballay, A. (2019). Intestinal infection regulates behavior and learning via neuroendocrine signaling. Elife *8*, 1–22.

Sordillo, A., and Bargmann, C.I. (2021). Behavioral control by depolarized and hyperpolarized states of an integrating neuron. Elife *10*, 2021.02.19.431690.

Stephens, G.J., Johnson-Kerner, B., Bialek, W., and Ryu, W.S. (2008). Dimensionality and dynamics in the behavior of C. elegans. PLoS Comput. Biol. *4*.

Stern, S., Kirst, C., and Bargmann, C.I. (2017). Neuromodulatory Control of Long-Term Behavioral Patterns and Individuality across Development. Cell *171*, 1649-1662.e10.

Von Stetina, S.E., Treinin, M., and Miller, D.M. (2006). The Motor Circuit. Int. Rev. Neurobiol. *69*, 125–167.

Steuer Costa, W., Yu, S. chieh, Liewald, J.F., and Gottschalk, A. (2017). Fast cAMP Modulation of Neurotransmission via Neuropeptide Signals and Vesicle Loading. Curr. Biol. *27*, 495–507.

Steuer Costa, W., Van der Auwera, P., Glock, C., Liewald, J.F., Bach, M., Schüler, C., Wabnig, S., Oranth, A., Masurat, F., Bringmann, H., et al. (2019). A GABAergic and peptidergic sleep neuron as a locomotion stop neuron with compartmentalized Ca2+ dynamics. Nat. Commun. *10*, 4095.

Taylor, S.R., Santpere, G., Weinreb, A., Barrett, A., Reilly, M.B., Xu, C., Varol, E., Oikonomou, P., Glenwinkel, L., McWhirter, R., et al. (2021). Molecular topography of an entire nervous system. Cell *184*, 4329-4347.e23.

Thacker, C., and Rose, A.M. (2000). A look at the Caenorhabditis elegans Kex2/Subtilisin-like proprotein convertase family. BioEssays *22*, 545–553.

Tinbergen, N. (1951). The study of instinct (United Kingdom: Oxford, Clarendon Press).

Tomar, S. (2006). Converting video formats with FFmpeg. Linux J. *2006*, 10.

Waggoner, L.E., Zhou, G.T., Schafer, R.W., and Schafer, W.R. (1998). Control of alternative behavioral states by serotonin in Caenorhabditis elegans. Neuron *21*, 203–214.

Wang, Y., Zhang, X., Xin, Q., Hung, W., Florman, J., Huo, J., Xu, T., Xie, Y., Alkema, M.J., Zhen, M., et al. (2020). Flexible motor sequence generation during stereotyped

escape responses. Elife.

White, J.G., Southgate, E., Thomson, J.N., and Brenner, S. (1986). The structure of the nervous system of the nematode Caenorhabditis elegans. Philos. Trans. R. Soc. London. B, Biol. Sci. *314*, 1–340.

Wiltschko, A.B., Johnson, M.J., Iurilli, G., Peterson, R.E., Katon, J.M., Pashkovski, S.L., Abraira, V.E., Adams, R.P., and Datta, S.R. (2015). Mapping Sub-Second Structure in Mouse Behavior. Neuron *88*, 1121–1135.

Yamanashi, T., Maki, M., Kojima, K., Shibukawa, A., Tsukamoto, T., Chowdhury, S., Yamanaka, A., Takagi, S., and Sudo, Y. (2019). Quantitation of the neural silencing activity of anion channelrhodopsins in Caenorhabditis elegans and their applicability for long-term illumination. Sci. Rep. *9*, 1–11.