Mapping Sub-Second Structure in Mouse Behavior

Alexander B. Wiltschko[1,2], **Matthew J. Johnson**[1,2], **Giuliano Iurilli**[1], **Ralph E. Peterson**[1], **Jesse M. Katon**[1], **Stan L. Pashkovski**[1], **Victoria E. Abraira**[1], **Ryan P. Adams**[2], and **Sandeep Robert Datta**[1,*]

[1]Department of Neurobiology, Harvard Medical School, Boston, Massachusetts, USA 02115

[2]School of Engineering and Applied Sciences, Harvard University, Cambridge, Massachusetts, USA 02138

Summary

Complex animal behaviors are likely built from simpler modules, but their systematic identification in mammals remains a significant challenge. Here we use depth imaging to show that three-dimensional (3D) mouse pose dynamics are structured at the sub-second timescale. Computational modeling of these fast dynamics effectively describes mouse behavior as a series of reused and stereotyped modules with defined transition probabilities. We demonstrate this combined 3D imaging and machine learning method can be used to unmask potential strategies employed by the brain to adapt to the environment, to capture both predicted and previously-hidden phenotypes caused by genetic or neural manipulations, and to systematically expose the global structure of behavior within an experiment. This work reveals that mouse body language is built from identifiable components and is organized in a predictable fashion; deciphering this language establishes an objective framework for characterizing the influence of environmental cues, genes and neural activity on behavior.

Introduction

Innate behaviors are sculpted by evolution into stereotyped forms that enable animals to accomplish particular goals (such as exploring or avoiding a predator). Ultimately understanding how neural circuits create these patterned behaviors requires a clear framework for characterizing how behavior is organized and evolves over time. One conceptual approach to addressing this challenge arises from ethology, which proposes that the brain builds coherent behaviors by expressing stereotyped modules of simpler action in specific sequences (Tinbergen, 1951).

Although behavioral modules have traditionally been identified, one at a time, through careful human observation, recent technical advances have enabled more comprehensive characterization of the components of behavior. For example, in invertebrates, behavioral modules and their associated transition probabilities can now be discovered systematically through automated machine vision, clustering and classification algorithms (Berman et al., 2014; Croll, 1975; Garrity et al., 2010; Stephens et al., 2008, 2010; Vogelstein et al., 2014).

*Correspondence: srdatta@hms.harvard.edu.

Furthermore, identifying behavioral modules and transition probabilities has uncovered context-specific strategies used by invertebrate brains to adapt behavior to changes in the environment, which include both the emission of new behavioral modules (such as when the animal switches from "exploring" to "mating"), and the generation of new behaviors through re-sequencing existing modules. In *C. elegans*, for example, neural circuits that respond to appetitive cues alter transition probabilities between a core set of locomotor-related behavioral modules, thereby creating new behavioral sequences that enable taxis towards attractive odorants (Gray et al., 2005; Pierce-Shimomura et al., 1999). Similar observations have been made for sensory-driven behaviors in fly larvae (Garrity et al., 2010).

Comparable systematic approaches to discovering behavioral modules have not yet been implemented in mice. Instead, traditional behavioral classification approaches have been instantiated *in silico*, enabling machine vision algorithms to replace tedious and unreliable human scoring of videotaped behavior (de Chaumont et al., 2012; Jhuang et al., 2010; Kabra et al., 2013; Weissbrod et al., 2013). These approaches are powerful, and have been successfully used to quantify components of innate exploratory, grooming, approach, aggressive and reproductive behaviors. However, because they depend upon the prior specification, by human observers, of what constitutes a meaningful behavioral module, the insight from these methods is bounded by human perception and intuition. Currently-available approaches therefore focus on identifying a small number of pre-specified modules within a given experiment, rather than on discovering new behavioral modules (which potentially encapsulate novel patterns of action), describing the global structure of behavior, or predicting future actions based upon those in the past.

Systematically describing the structure of behavior in mice — and understanding how the brain alters that structure to enable adaptation — requires overcoming three challenges. First, it is not clear which features are important to measure when identifying candidate behavioral modules. Mice interact with the world by expressing complex three-dimensional (3D) pose dynamics, but because these are difficult to capture, most current methods track two-dimensional parameters such as the position, velocity, or 2D contour of the mouse (de Chaumont et al., 2012; Gomez-Marin et al., 2012; Jhuang et al., 2010; Kabra et al., 2013; Spink et al., 2001; but see Ou-Yang et al., 2011). Second, given that behavior evolves on many timescales in parallel, it is not clear how to objectively identify the relevant spatiotemporal scales at which to modularize behavior. Finally, even stereotyped modules of behavior exhibit moment-to-moment and animal-to-animal variability (Colgan, 1978). This variability raises significant challenges for identifying the number and content of behavioral modules, or with associating observed actions with specific behavioral modules.

Here we describe a novel method, based upon recent advances in machine vision and learning, that identifies behavioral modules and their associated transition probabilities without human supervision. This approach uses 3D imaging to capture the pose dynamics of mice as they freely behave in a variety of experimental contexts; these data reveal a surprising regularity that appears to divide mouse behavior into recognizable behavioral motifs that are organized at the sub-second timescale. A computational model then takes advantage of the observed fast temporal structure to describe mouse behavior as a series of

reused modules, each a brief and stereotyped 3D trajectory through pose space that is connected in time to other modules through predictable transitions. We use this combined 3D imaging/modeling approach to explore how the underlying structure of behavior is altered after distinct environmental, genetic or neural manipulations, and show that this method can detect both predicted changes in action and new features of behavior that had not been previously described. This work reveals that defining behavioral modules based upon structure in the 3D behavioral data itself — rather than using *a priori* definitions for what should constitute a measurable unit of action — can yield key information about the elements of behavior, offer insight into adaptive behavioral strategies used by mice, and enable discovery of subtle alterations in patterned action.

Results

3D Imaging Captures Inherent Structure in Mouse Pose Dynamics

We wished to develop a method that would allow unsupervised phenotyping of mice based upon patterns of 3D movement. However, it is not clear whether spontaneous behaviors exhibited by mice have a definable underlying structure that can be used to characterize action as it evolves over time. To ask whether such a structure might exist, we measured how the shape of a mouse's body changes as it freely explores a circular open field (Experimental Procedures). We used a single depth camera placed above the arena to capture these 3D pose dynamics, and then extracted the image of the mouse from the arena, corrected imaging artifacts due to parallax effects, centered and aligned the mouse along the inferred axis of its spine, and then quantified how the mouse's pose changed over time (Fig. 1A, Fig. S1, Movie S1).

Plotting these 3D data over time revealed that mouse behavior is characterized by periods during which pose dynamics evolve slowly, punctuated by fast transitions that separate these periods; this pattern appears to divide the behavioral imaging data into blocks typically lasting 200–900 ms (Figs. 1B, S2A). This temporal structure is apparent in the raw imaging pixels (Fig. 1B, top), the inferred shape of the mouse's spine (Fig. 1B, middle), and dimensionally-reduced data that randomly samples from the depth images on the sensor (Fig. 1B, bottom, see Supplemental Experimental Procedures). This structure is absent in data acquired from anesthetized or dead mice, but is present for the entire duration of each experiment in the pose dynamics data of mice exploring behavioral arenas of different shapes and after exposure to a wide variety of sensory cues, suggesting that it is a fundamental and ubiquitous feature of mouse behavior (data not shown, Fig. S2).

To characterize this fast temporal structure we performed three separate quantitative analyses. First, approximate boundaries between blocks in the behavioral imaging data were identified by a changepoints algorithm, which revealed that the mean block duration was about 350 ms, roughly matching the timescale of the blocks apparent upon visual inspection (Figs. 1B, 1C, Supplemental Experimental Procedures). Second, we performed temporal autocorrelation analysis on the pose dynamics data, which demonstrated that autocorrelation in the mouse's pose largely dissipated after 400 ms (tau = 340 ± 58 ms, Figs. 1C, S2B). This pattern of autocorrelation reflects specific behavioral dynamics organized at sub-second timescales, as it was destroyed by shuffling the behavioral data at timescales of 500 ms or

less, and was not observed in synthetic mouse behavioral data designed to evolve with either random walk or Levy flight characteristics (Figs. 1C, S2C). Third, we used a Wiener filter analysis to compare power spectral densities taken from live and dead mice; this approach identifies frequencies that must be changed in imaging data taken from a dead mouse to match the power spectrum of a live mouse. Nearly all of the frequency content differentiating behaving from dead mice was concentrated between 1 and 6 Hz (measured by spectrum ratio, mean $3.75 \pm .56$ Hz, Fig. 1C, S2B). Taken together, the qualitative appearance of block structure in the pose dynamics data, along with the convergent results obtained with these three quantitative analyses, demonstrate that mouse pose dynamics exhibit structure at the sub-second timescale.

The observed temporal structure within the pose dynamics data suggests a timescale at which continuous behavior may be naturally segmented into meaningful components, as visual inspection of 3D movies revealed that each of the sub-second blocks of behavior appears to encode a recognizable action (e.g., a dart, a pause, the first half of a rear, see Movie S2). To explore the possibility that these sub-second actions are stereotyped (and therefore reproducibly performed at different times during an experiment), the 3D mouse imaging data was subjected to wavelet decomposition followed by principal component analysis (PCA), which transformed each block in the pose dynamics data into a continuous trajectory through principal component (PC) space (Figs. 1D). By scanning the behavioral data using a template matching method (using Euclidean distance among the first 10 PCs, which explain 88% of the data variance, Fig. S2D), additional instances were identified in which each template action was reused (Fig. 1D, Supplemental Experimental Procedures, see Fig. S3 for additional examples). These anecdotal observations suggest that mice create complex behaviors through the serial expression of stereotyped and reused behavioral modules (Tinbergen, 1951).

Mouse Behavior Can Be Described and Predicted with Modules and Transitions

Although our analysis suggests a timescale at which behavioral modules might exist, and qualitative inspection of 3D video and PCA trajectories is consistent with the possibility that sub-second blocks of behavior correspond to reused modules, current methods do not allow for the systematic identification of candidate behavioral modules in mice. Indeed, available approaches neither reveal whether dividing behavior into modules actually helps to explain the overall structure of behavior, nor identify the most likely number of modules expressed within any given dataset, or the content and durations of those modules.

To address these issues we built a family of computational models, each of which proposes a unique underlying structure for behavior, and asked which of these models best predicts the pose dynamics of freely behaving mice (Figs. 2A, S4); we reasoned that the model that most closely fit behavioral data (to which the model had not been exposed) would reveal key features of the underlying organization of behavior, and could be used to characterize its components. After pre-processing the imaging data, the top 10 PCs of the data (Fig. S4E) were used to fit each model; this use of PCs (which directly reflect the pixel data) as a basis for modeling minimized potential biases from feature engineering. Models were fit using Bayesian nonparametric and Markov Chain Monte Carlo techniques that can automatically

identify structure within large datasets, including the optimal state number for a given dataset and model (see Supplemental Experimental Procedures, Fig. S4). Each model was trained on one set of pose dynamics data, and then tested for its ability to explain a separate set of held-out data, a metric canonically used to compare unsupervised learning models (Hastie et al., 2009).

Our simplest model proposed that each behavioral module is a single fixed pose with no defined transition probabilities between the modules. We iteratively added structure to build increasingly complex models, which incorporated modules with more elaborate internal structures (ranging from mixtures of poses to smooth pose trajectories), allowed predictable transitions between specific modules (by embedding the modules within a Markov model), or both (Fig. S4). Where relevant, the fitting procedures were explicitly focused to search for behavioral modules at the sub-second timescale matching the temporal structure identified using our model-free methods; this approach provided an important — and previously unavailable — constraint, given the multiple possible timescales upon which behavior evolves simultaneously.

The model that best fit previously-unseen behavioral data describes mouse behavior as a sequence of modules (each capturing a brief motif of 3D body motion) that switch from one to another at the sub-second timescale identified by our model-free analysis of pose dynamics (Fig. 2A, S4D, Supplemental Experimental Procedures). We refer to this model as an AR-HMM, as each behavioral module was modeled as a vector autoregressive (AR) process capturing a stereotyped trajectory through pose space, and the switching dynamics between different modules were modeled using a hidden Markov model (HMM). In other words, the model is a hierarchical description of behavior, with the "internals" of each module reflecting the mouse's pose dynamics over short timescales, and the longer-timescale relationships between behavioral modules (i.e., the possible module sequences) governed by the transition probabilities specified by an HMM (Fig. S4B, S4C). The observation that the AR-HMM outperforms alternative models (Fig. 2A) demonstrates that modularity and transition structure at fast timescales are critical for describing mouse behavior, a key prediction from ethology.

Model-Identified Behavioral Modules Are Stereotyped and Distinct

The AR-HMM systematically identifies modules and their transition probabilities from behavioral data without human supervision; this suggests that the AR-HMM can be used to identify behavioral modules and their associated transition probabilities — and thereby expose the underlying structure of behavior — in a wide variety of experimental contexts. We therefore performed a series of control analyses to establish whether the AR-HMM can indeed reliably identify behavioral modules encoding repeatedly-used and stereotyped motifs of distinguishable behavior that are organized at sub-second timescales.

Although the AR-HMM is tuned to identify modules at a particular timescale, it is possible that after training the model could fail to effectively capture temporal structure in behavior. However, the AR-HMM successfully identified modules at the fast behavioral timescale defined by the model-free methods, as the distribution of module durations was similar to the duration distribution for changepoints-identified blocks (Fig. S5A). Importantly, the

ability of the AR-HMM to identify behavioral modules depended upon the inherent sub-second organization of mouse pose data, as shuffling the behavioral data in small chunks (i.e., < 500 milliseconds) substantially degraded model performance, whereas shuffling the data in bigger chunks had little effect (Fig. S5B).

We then asked whether model-identified modules encode repeatedly-used and stereotyped motifs of behavior. The pose trajectories associated with a specific model-identified behavioral module took similar paths through PC space, and visual inspection of the 3D movies associated with multiple instances of this module revealed they all encoded a stereotyped rearing behavior (Fig. 2B, 2C, Movie S3). In contrast, data instances drawn from different behavioral modules traced distinct (and stereotyped) paths through PC space (Fig. 2B, see Fig. S5C for additional examples). Furthermore, visual inspection of the 3D movies assigned to different modules reveals that each encoded a coherent pattern of 3D motion that *post hoc* can be distinguished and labeled with descriptors (see Movie S4 for "walk," "pause," and "low rear" modules depicted in Fig. 2C, as well as additional examples).

The modules identified by the AR-HMM are distinct from each other, as a cross-likelihood analysis demonstrated that the imaging data associated with a given module are best assigned to that module, and not to any of the other behavioral modules (Fig. 2D, Supplemental Experimental Procedures). In contrast, the AR-HMM failed to identify any well-separated modules in a synthetic mouse behavioral dataset that lacks modularity but otherwise matches all multidimensional and intertemporal correlations of the real data, demonstrating that the AR-HMM does not discover modularity where none exists (Fig. 2D).

Furthermore, rerunning the AR-HMM training process from random starting points generated highly similar behavioral modules (R^2 = .94 ± 0.03, n = 15 restarts, Supplemental Experimental Procedures); in comparison, models with lower held-out likelihood scores have lower consistency (R^2 < .4), suggesting that these alternatives fail to reliably identify underlying structure in behavior. The AR-HMM output was robust to the specific training data used, as models created from subsets of a larger dataset representing a single experiment were highly similar to each other (R^2 > .9). These findings demonstrate that the AR-HMM converges on a consistent set of behavioral modules regardless of the specific training data (within a given experiment) or how the model is initialized.

Finally, the modules of behavior and the associated transition probabilities identified by the AR-HMM appear to fully capture the richness of mouse behavior, as a 3D movie of a behaving mouse generated by a trained AR-HMM was qualitatively difficult to distinguish from a 3D movie of behavior exhibited by a real animal (Movie S5, Fig. S4E); in contrast, movies generated by more poorly performing models appeared discontinuous and were easily distinguished from real animals (data not shown). Taken together with our model-free and model-based analyses and controls, the observation that after training our model can synthesize a convincing replica of 3D mouse behavior from learned modules and transition probabilities is consistent with the hypothesis that mouse behavior is organized into distinct sub-second modules that are combined to create coherent patterns of action.

Using the AR-HMM to Characterize Baseline Patterns of Behavior

The AR-HMM identifies two key features of mouse behavior (from the perspective of 3D pose dynamics): which behavioral modules are expressed during behavior, and how those modules transition into each other over time to create action. The AR-HMM identified ~60 reliably-used behavioral modules from a circular open field dataset, which is representative of normal mouse exploratory behavior in the laboratory (51 modules explained 95 percent of imaging frames, and 65 modules explained 99 percent of imaging frames, Figs. 3A, S5D, S5E; subjective categorization of the 51 most-used modules is shown in Fig. S6A). These modules were connected to each other over time in a highly non-uniform manner, with each module preferentially linked to some modules and not others (Fig. 3B, Fig. S5F; average node degree after thresholding transitions that occur with < 5% probability, 4.08 ± .10). This specific transition pattern among modules constrained the module sequences observed in the dataset (~17,000/~75,000 possible sequences of three modules ("trigrams") given the total data size) demonstrating that certain module sequences were favored over others, and that mouse behavior is therefore predictable (per frame entropy rate without self-transitions 3.78 ± .03 bits, with self-transitions .72 ± .01 bits, entropy rate in a uniform matrix 6.022 bits; average mutual information without self-transitions 1.92 ± .02 bits, with self-transitions 4.84 bits ± .03 bits; see Movie S6 for multiple examples of the sequence of three modules depicted in Fig. 2B). Note that while estimating coarse changes in specific trigram frequencies is possible, accurately estimating higher-order k-gram transition statistics is difficult as the amount of data required grows exponentially with k. We therefore focus our analysis on lower-order statistics such as module usage frequencies and temporal interactions between pairs of modules.

Using the AR-HMM to Characterize the Nature of Behavioral Change

We tested whether the AR-HMM could effectively capture changes in behavior (both predicted and unpredicted) elicited by a range of simple experimental interventions designed to probe the influence of the environment, genes, or neural circuit activity on behavior. We first asked how mice behavior adapts to changes in apparatus shape. We imaged mice within a small square box and then co-trained our model with both the circular open field and square box data, thereby enabling direct comparisons of modules and transitions under both conditions; the modules identified by this co-training procedure did not erroneously lump together data that would otherwise be distinguishable, as each behavioral module's mean pose trajectory was stereotyped across the different experimental arenas (data not shown). Although mice tended to explore the corners of the square box and the walls of the circular open field, the overall usage of most modules was similar between these apparatuses, consistent with exploratory behavior sharing many common features across arenas (Fig. 3C).

However, the AR-HMM also identified a small number of behavioral modules that were deployed selectively in just one context, consistent with the idea that different physical environments drive expression of new behavioral modules (Fig. 3C). For example, one circular arena-specific module encoded a behavior in which mice walk near the arena wall with a body posture that matches the curvature of the wall, while within the square box mice expressed a context-specific module that encodes a dart out of the center of the square (Figs. 3D and E). Several behavioral modules were also differentially enriched (but not exclusively

expressed) in one context or the other. In the circular arena, for example, mice preferentially executed a rear characterized by pointing outwards while pausing near the center of the open field, whereas in the smaller square box mice preferentially expressed a high rear in the corners of the box (Fig. 3D, data not shown).

These results demonstrate that the AR-HMM can effectively capture predictable changes in behavior resulting from altering the physical environment (like walking along a curved wall or rearing in a corner). Importantly, these results also demonstrate that the AR-HMM can unmask arena-specific patterns of behavior that are expressed in the center of both arenas, away from the physically-constraining walls (like the darting across the center of the square box and the outward pointing behavior expressed in the circle); this surprisingly suggests that arena shape influences mouse behavior in a manner that extends significantly beyond the predictable changes in action at the walls themselves. Taken together, these experiments reveal that the AR-HMM can suggest strategies used by the mouse brain to adapt to new physical environments: in the case of a change in environmental geometry, this strategy includes the recruitment of a limited set of context-specific behavioral modules into baseline patterns of action, and a broad rewriting of where in space modules are expressed with respect to the arena boundaries.

The small number of behavioral modules distinguishing the circular and square arenas suggests only modest differences in the global pattern of behavior in these two experiments. To ask how the AR-HMM captures changes in the underlying structure of behavior after an overt change in behavioral state, we exposed mice to an ethologically-relevant olfactory cue, the aversive fox odor trimethylthiazoline (TMT), which was delivered to one quadrant of the square box via an olfactometer. This odorant profoundly changes mouse behavior, inducing odor investigation, escape, and freezing behaviors that are accompanied by increases in corticosteroid and endogenous opioid levels (Fendt et al., 2005; Wallace and Rosen, 2000). Consistent with these known effects, mice sniffed the odor-containing quadrant and then avoided the quadrant containing the predator cue, displaying prolonged periods of immobility traditionally described as freezing behavior (Fig. 4A, Fig. S6B).

Given that TMT-induced behaviors are dramatically different than those observed at baseline, one might predict that TMT should induce new behavioral modules that underlie the generation of these new actions. Surprisingly, the AR-HMM revealed that the TMT-induced suite of new behaviors was best explained by the same set of behavioral modules that were expressed during normal exploration; several modules were up- or down-regulated after TMT exposure, but new modules were not introduced or eliminated relative to control (Fig. 4B).

We therefore asked whether the changes in observed behavior were the consequence of altered connections between behavioral modules. Plotting the module transitions altered after exposure to TMT defined two neighborhoods within the behavioral statemap; the first included an expansive set of transitions that was modestly downregulated by TMT, and the second included a focused set of transitions that was upregulated by TMT (Fig. 4C). Many of these newly-interconnected modules encoded different forms of freezing behavior (average velocity .14 ± .54 mm/sec, for other modules 34.7 ± 53 mm/sec, Movie S8, Fig.

S6B). In addition, the TMT-initiated modulation of transition probabilities altered the expression of specific behavioral sequences; for example, the most commonly-observed sequence of three freezing modules was expressed 716 times after TMT exposure (in 300 minutes of imaging), as opposed to just 17 times under control conditions (in 480 minutes of imaging). The stimulus-evoked rewriting of transition probabilities was accompanied by an increase in the overall predictability of mouse behavior (per frame entropy rate fell from $3.92 \pm .02$ bits to $3.66 \pm .08$ bits without self-transitions, and from $.82 \pm .01$ bits to $.64 \pm .02$ bits with self-transitions) consistent with the mouse enacting an avoidance strategy that was more deterministic in nature than locomotor exploration.

Proximity to the odor source also governed the usage of specific behavioral modules (Figs. 4D, 4E). For example, a set of freezing-related modules tended to be expressed in the quadrant most distant from the odor source, while the expression of an investigatory rearing module (whose overall usage was not altered by TMT) was specifically enriched within the odor quadrant (Figs. 4D, 4E, Movie S9).

Although TMT is known to induce dramatic changes in behavior, it has not been possible to systematically identify those specific behavioral features altered in response to this odorant, or to place those altered features in context with normal patterns of exploration. Analysis by the AR-HMM suggests that the strategy used by the mouse brain to adapt to the presence of TMT overlaps with — and yet is distinct from — that used to accommodate physical changes in the environment. As was true for a changed physical environment, exposure to TMT alters the spatial deployment of modules and sequences to support particular patterns of action; in contrast, the complete cohort of behaviors elicited by TMT, including seemingly "new" behaviors such as freezing, are the consequence of altered transition structure between individual modules. Behavioral modules are not, therefore, simply reused over time, but instead act as flexibly interlinked components whose specific sequencing and deployment in space has profound consequences for the generation of adaptive behavior.

Sub-Second Architecture of Behavior Reflects the Influence of Genes and Neural Activity

As described above, the AR-HMM shows the fine-timescale structure of behavior to be sensitive to persistent changes in the physical or sensory environment. However, manipulation of individual genes or neural circuits influences behavior across a range of spatiotemporal scales and with variable penetrance and reliability; these changes in behavior may or may not be effectively captured by a classification method designed to characterize the sub-second structure of action. We therefore directly asked whether the AR-HMM could systematically reveal the behavioral consequences of manipulating the mouse genome or activity within the nervous system.

To explore this possibility, we used the AR-HHM to characterize the phenotype of mice mutant for the retinoid-related orphan receptor 1β (*Ror1*β) gene, which is expressed in neurons in the brain and spinal cord. This mouse was selected for analysis because adult homozygous mutant animals permanently exhibit abnormal gait, which would be expected to be observed during a brief open-field experiment (André et al., 1998; Eppig et al., 2015; Liu et al., 2013; Masana et al., 2007). Analysis with the AR-HMM revealed that littermate control mice are nearly indistinguishable from fully inbred C57/Bl6 mice, whereas

homozygous mutant mice express a unique behavioral module encoding a waddling gait (Fig. 5A, 5C, Movie S10, see Supplemental Experimental Procedures for statistical details, data were co-trained with both circular open field and TMT datasets to facilitate comparisons with C57 animals). Conversely, the expression of five behavioral modules encoding normal forward locomotion in wild-type and C57 mice was downregulated in *Ror1β* mutants (Fig. 5A, average during-module velocity = 114.6 ± 76.3 mm/sec).

Previously unobserved phenotypes in the *Ror1β* mutant mice were also uncovered by the AR-HMM, as the expression of a set of four modules that encoded brief pauses and headbobs was upregulated in mutant mice (Fig. 5A, average during-module velocity = 8.8 ± 25.3 mm/sec); this pausing phenotype had not been previously reported in the literature. Furthermore, heterozygous mice – which have no reported phenotype (André et al., 1998; Eppig et al., 2015; Liu et al., 2013; Masana et al., 2007), exhibit wild-type running wheel behavior (Masana et al., 2007), and appear normal by eye — were also found to express a fully-penetrant mutant phenotype: they overexpressed the same set of pausing modules that were upregulated in the full *Ror1β* mutants, while failing to express the more dramatic waddling phenotype (Fig. 5A). Differences between wild-type animals and both heterozygotes and mutants were also observed in transition probabilities associated with these pausing modules (Fig. 5B).

The AR-HMM therefore describes the pathological behavior of *Ror1β* mice as the combination of a single neomorphic waddling module, decreased expression of normal locomotion modules, and increased expression of a small group of physiological modules encoding pausing behaviors; heterozygous mice express a defined subset of these behavioral abnormalities, whose penetrance is not intermediate but equals that observed in the mutant. These results suggest that the sensitivity of the AR-HMM allows fractionation of severe and subtle behavioral abnormalities within the same litter of animals, facilitates comparisons amongst genotypes and enables discovery of new phenotypes.

In the case of the *Ror1β* animals, a permanent alteration in DNA sequence is translated into an ongoing change in the overall sub-second statistical structure of behavior, one that is expressed continuously over the lifetime of the animal. We also wished to characterize how transient changes in activity in specific neural circuits influence the moment-to-moment structure of behavior; furthermore, given that the relationship between neural circuit activity and behavior can vary on a trial-to-trial basis, we wanted to probe the ability of the AR-HMM to afford insight into the probabilistic relationships between neural circuit activity and behavior.

To address these questions we unilaterally expressed the light-gated ion channel Channelrhodopsin-2 in a subset of layer 5 corticostriatal neurons in the right hemisphere and assessed behavioral responses before, during, and after two seconds of light-mediated activation of motor cortex (Glickfeld et al., 2013). At negligible power levels no light-induced changes in behavior were observed, whereas at the highest power levels the AR-HMM identified two behavioral modules whose expression was reliably induced by the light, as on nearly every trial either one or the other module was expressed (Fig. 6A). As would be expected, both of these modules encode forms of spinning-to-the-left behavior,

and neither of these modules was expressed during normal mouse locomotion (Fig. 6B, Movie S11). In addition we noted that approximately 40 percent of the time the overall pattern of behavior did not return to baseline for several seconds after the end of optogenetic stimulation. This deviation from baseline was not due to continued expression of the same spinning modules that were triggered at light onset; instead, mice often paradoxically expressed a pausing module at light offset (average during-module velocity = $.8 \pm 7$ mm/sec, Fig. 6A, see cross).

The behavioral changes induced by high intensity optogenetic stimulation were reliable, as on nearly every trial the animal emitted one of the two spinning modules. To ask whether the AR-HMM could characterize the pattern of behavior observed when behavioral modules are generated more probabilistically, the levels of light stimulation were reduced; we identified conditions under which one of the two spinning modules was no longer detected, and the other was expressed in only 25 percent of trials. Under these conditions the AR-HMM detected the upregulation of a distinct set of behavioral modules, each of which was expressed in a fraction of trials (Fig. 6A, see stars). Unlike the spinning modules triggered by high intensity stimulation, these modules were not neomorphic; rather, these modules were normally expressed during physiological exploration, and encoded distinct forms of forward locomotion behavior (data not shown). Interestingly, although each of these individual light-regulated modules was emitted probabilistically on any given trial, low intensity neural activation reliably influenced behavior across all trials when the behavioral modules were considered in aggregate (Fig. 6A).

Thus both low and high intensity optogenetic stimulation dramatically alters behavior, but the identity of these behaviors and their relative probabilities of expression vary across light levels. This effect would not have been apparent without the ability of the AR-HMM to distinguish new from previously-expressed behavioral modules, and to identify the specific behavioral module induced at light onset on each trial; indeed under low light stimulation the behavioral phenotype apparent by eye on any given trial often appears to be an extension of normal mouse exploratory behavior. Taken together, these data demonstrate that the AR-HMM can identify and characterize both obvious and subtle optogenetically-induced phenotypes, distinguish "new" optogenetically-induced behaviors from upregulated expression of "old" behaviors, and reveal the trial-by-trial relationships between neural activity and action.

Discussion

It has been long hypothesized that innate behaviors are composed of stereotyped modules, and that specific sequences of these modules encode coherent and adaptive patterns of action (Bizzi et al., 2000; Brown, 1911; Drai et al., 2000; Lashley, 1967; Sherrington, 1907; Tinbergen, 1951). However, most efforts to explore the underlying structure of mouse behavior have relied on *ad hoc* definitions of what constitutes a behavioral module, and have focused on specific behaviors rather than systematically considering behavior as a whole. As a consequence, we lack insight into the global organization of mouse behavior, into the relationships between currently expressed actions and past or future behaviors, and into the strategies used by the brain to generate behavioral change. Furthermore, we lack a

comprehensive framework for characterizing the influence of individual genes or neural circuits on behavior.

Here we use 3D imaging to identify a sub-second spatiotemporal scale at which mouse behavior may be organized. Using this finding as a constraint, we then built a family of computational models, each of which represents a different hypothesis for the potential structure of behavior, and compared the ability of these models to predict and explain mouse behavior. The best performing model (the AR-HMM) specifically searches for modularity at sub-second timescales similar to those observed in the pose dynamics data, and quantitatively describes behavior as a series of sub-second modules with defined transition probabilities. This combined 3D imaging/modeling approach can be used to automatically identify the behavioral modules expressed during a variety of experiments and to systematically discover how the architecture of behavior is altered as the mouse adapts to a changing world.

The AR-HMM Automatically and Systematically Captures Known and New Phenotypes

Our experiments reveal that the AR-HMM identifies both predicted changes in action and new features of behavior that had not been previously described. The high sensitivity of the AR-HMM — illustrated by the identification of a previously-undescribed phenotype in heterozygous Ror1β mice — raises the possibility that the AR-HMM could be useful for extracting subtle phenotypes from mouse models, including those in which single gene alleles are mutated; such patterns of mutation are common in human disease, but rarely explored in mouse models. Furthermore, the ability of the AR-HMM to place changes in behavior into context — illustrated by the ability of the AR-HMM to identify the modules expressed after modest optogenetic stimulation on a trial-by-trial basis — suggests that the AR-HMM may be useful for relating unreliable or noisy patterns of neural activity to the probabilistic expression of specific actions. The AR-HMM (or similar approaches based upon unsupervised machine learning) may therefore be useful for discovering the behavioral consequences of a wide variety of experimental manipulations, particularly those enabled by the ever-growing toolbox of gene editing and optogenetic techniques.

The AR-HMM Identifies Possible Mechanisms for Behavioral Change

The AR-HMM suggests three regulatory strategies that may be used by the brain to implement behavioral adaptation. First, behavioral modules and their transitions appear to be selectively — and to some extent independently — vulnerable to alteration. None of the physiological and pathological deviations from baseline described here, from freezing to a waddling gait, caused global changes in the underlying structure of action; instead, new behaviors were well-described as changes in a small number of specific modules or transitions. This suggests that the brain can focally manipulate individual modules or transitions to generate new behaviors, and furthermore that mice can accommodate pathological actions (such as waddling) without catastrophic alterations in behavioral patterning.

Second, dramatically new behaviors can be created by altering the transitions statistics between modules alone — without invoking new behavioral modules — as was observed in

animals exposed to TMT. This strategy has been shown to underlie sensory-driven negative and positive taxis behaviors in bacteria, worms and flies (Berg and Brown, 1972; Garrity et al., 2010; Gray et al., 2005; Pierce-Shimomura et al., 1999); here we show that this strategy is conserved in mice, and used by vertebrate nervous systems to create complex patterns of action in response to an external cue. In the specific case of the TMT response, the induced behaviors extend beyond taxis to adaptations like freezing, suggesting that the restructuring of transition probabilities may be a general mechanism for creating new patterns of action.

Third, we find that modulation of where in space behavioral modules are expressed supports the generation of specific adaptive behaviors. For example, the rearing module used by mice to investigate TMT is not significantly upregulated or resequenced relative to control, and yet its spatial pattern of expression in the quadrant containing TMT facilitates detection (and therefore avoidance) of the aversive odorant. Characterizing where in space behavioral modules are expressed also reveals that changes in physical context — such as the difference between a square and circular arena — elicit "state"-like changes in mouse behavior that extend beyond predictable changes in action at apparatus boundaries. Because the training data for the AR-HMM does not include any explicit allocentric parameters (such as the mouse's spatial position within the apparatus), the ability of the model to uncover meaningful relationships between allocentric space and egocentric pose dynamics is an important validation that its segmentation of behavior is informative.

Potential Neural Underpinnings for Modules and Transitions

The observations generated by the AR-HMM lead to several predictions about the neural control of behavior. One such prediction is that specific behavioral modules and their associated transition probabilities will have explicit neural correlates, whose pattern of activity should reflect the ~2–5 Hz timescale at which modules are expressed; it is tempting to speculate that neural correlates representing transition probabilities between behavioral modules will be encoded in higher-order neural circuits tasked with behavioral sequencing, while neural correlates for the behavioral modules themselves might be encoded in central pattern generators or related circuit motifs in the brainstem or spinal cord. The relevant neural circuits may include both evolutionarily-ancient regions of the brain involved in releasing innate behaviors, such as the amygdala, hypothalamus and brainstem, as well as other areas, such as the striatum, that regulate fine-timescale behavioral sequencing (Aldridge et al., 2004; Swanson, 2000). Testing these predictions will require simultaneous characterization of neural activity and assessment of behavior; future embellishments of the modeling approach described here may allow for inference of joint structure between dense neural and behavioral data, and therefore be useful for revealing mechanistic relationships between the dynamics of neural activity and action.

Unsupervised Behavioral Characterization via Modeling of 3D Data: Strengths, Caveats and Future Directions

Although there are multiple possible approaches to acquiring 3D pose dynamics data, we chose to implement our imaging system using a single standard depth camera because such cameras are widely available, adaptable to a variety of different experimental circumstances, and can characterize behavior under most lighting conditions and in animals with any coat

color (due to the use of infrared light as an illumination source). The "unsupervised" modeling approach taken here is also transparent, insofar as the assumptions of the model are explicitly stated (with no constraints supplied by the researcher other than the structure of the model, and the specification of a single parameter that acts as a tunable "lens" to focus the model on behavior at a particular timescale). This makes clear the precise bounds of human influence on the output of the AR-HMM, and insulates key aspects of that output from the vagaries of human perception and intuition. Given that inter-observer reliability in scoring even single mouse behaviors can be low (from 50 to 70 percent, with reliability falling as the number of scored behaviors increases), developing methods free from observer bias is essential for informatively characterizing mouse behavior (Garcia et al., 2010).

In addition, the generative modeling and inferential fitting methods described herein offer several practical advantages over the approaches typically used to analyze mouse behavior (Crawley, 2003), including: the explicit time-series modeling of behavioral data (as opposed to simple clustering of dimensionally-reduced data); the ability to directly inspect and explore each behavioral module; the flexible discovery of previously-unobserved behavioral modules (rather than characterization of behavior from the perspective of "known" phenotypes); and the ability to generate synthetic behavioral data, thereby allowing quantitation of how well a given model predicts the structure of behavior. This quantitative framework for comparing the performance of different methods for dividing up and measuring behavior is critical for advancing behavioral neuroscience, as it enables objective evaluation of alternative models for behavior, and assessment of future extensions that incorporate inevitable improvements in camera resolution, model structure, and fitting procedures.

On the other hand, our conclusions regarding the underlying structure of behavior (including its timescale) are limited by the simplicity of our experimental manipulations, which were designed to expose differences in motor outputs. Furthermore "mouse behavior" as described by the model is restricted to the imaged pose dynamics of the mouse at a particular spatiotemporal resolution and within a controlled laboratory experiment. Because the AR-HMM directly models the pixel data (after pre-processing), comparisons can only be made between mice of roughly similar size and shape. Furthermore, there are clearly important physical features of mouse behavior (operating at a variety of spatiotemporal scales) not captured in the pose data and therefore not modeled — these range from individual joint dynamics and paw position to sniffing, whisking and breathing. In the future, complementary datastreams that capture different facets of behavior could be integrated with 3D pose data to generate more comprehensive behavioral models.

In addition, the modeling approach itself has several important limitations. The AR-HMM uses Bayesian nonparametric approaches to identify the most likely number of modules that describe behavior at a particular temporal scale. However, this insight also comes at a cost: as the amount of data fed to the algorithm increases, the number of discovered modules necessarily rises. This challenge parallels the well-described phenomenon in ethology in which the number of discovered behaviors increases in subjects that have been observed either more frequently or for longer (Colgan, 1978). This monotonic (although sublinear) relationship between data size and the number of discovered modules limits comparisons of

behavior across experiments without co-training models, as was done here. One potential approach to address this challenge could be the incorporation of a "canonical" behavioral dataset against which additional data could be compared; such a framework (in which the "canonical" modules are either frozen or flexible) may enable analysis of new behavioral experiments within a fixed frame of reference.

Finally, the AR-HMM cannot explicitly disambiguate those features of behavior that are the consequence of the biomechanics of the mouse — for example, transitions between specific modules that are impossible due to physical constraints — from those that are the consequence of the action of the nervous system. Given that the nervous system and the body in which it is embedded co-evolved to facilitate action, ultimately disentangling the relative contributions of each to the organization of behavior may be difficult (for discussion of this issue, see for example (Tresch and Jarc, 2009)). However, changes in the structure of behavior that are induced by experimental intervention arise principally from the action of the nervous system; the observed context-dependent flexibility of the transition statistics between modules, taken with the ability of the mouse to emit new behavioral modules in response to internal or external cues, together suggest that — at the spatiotemporal scale captured by our methods — the nervous system plays a key role in regulating the overall structure of behavior.

Mouse Body Language: Syllables and Grammar

Candidate behavioral modules have been recognized in a variety of different contexts and on a wide range of spatiotemporal scales, and accordingly researchers have given them a diverse set of names, including motor primitives, behavioral motifs, motor synergies, prototypes and movemes (Anderson and Perona, 2014; Flash and Hochner, 2005; Tresch et al., 1999). The behavioral modules identified by the AR-HMM here find their origin in switching dynamics that are expressed on timescales of 100s of milliseconds. Mouse behavior is clearly also organized at the varied and interdigitated timescales at which internal state (e.g. neural activity, endocrine function and development) and external state (e.g. daily, monthly, seasonal and annual variation in behavior) unfold. The behavioral modules we have characterized therefore likely exist at an intermediate hierarchical level within the overall structure of behavior, albeit one that captures many of the behavioral changes induced by experimental manipulations carried out at both short and long timescales.

The modules identified by the AR-HMM do not exist in isolation; instead they are given behavioral meaning through a transition structure that governs their sequencing. The observation of both modularity and transition structure within the pose dynamics of mice suggests strong analogies to birdsong, which is also hierarchically organized and composed of identifiable modules whose sequence is governed by definable transition statistics; importantly, birdsong has also been well described using generative modeling and low-order Markov processes (Berwick et al., 2011; Markowitz et al., 2013; Wohlgemuth et al., 2010). By analogy to birdsong we therefore propose to refer to the modules we have identified as behavioral "syllables," and the statistical interconnections between these syllables as behavioral "grammar." Such a grammatical structure has been previously proposed for

Author Manuscript

restricted subsets of mouse behavior (such as grooming) in which behavioral modules were defined on an *ad hoc* basis (Berridge et al., 1987; Fentress and Stilwell, 1973); through unsupervised identification of behavioral syllables here we show that the notion of a regulatory grammar is general, and can be used to explain a wide variety of behavioral phenotypes. As is true for birdsong, the grammar we describe is highly restricted in nature (as only low-order interactions are modeled) and lacks the richness and flexibility of context-dependent grammars that have been explored in human language (Berwick et al., 2011). Despite this limitation, the experiments described herein expose an underlying structure for mouse body language organized at the sub-second timescale; this structure encapsulates mouse behavior (as detected by a depth sensor) within a given experiment, and reveals a balance between stochasticity and determinism that is dynamically modulated as the mouse varies its pattern of action to adapt to challenges in the environment.

Experimental Procedures

All errors bars indicated in the paper are SEM as determined by bootstrap analysis unless noted otherwise. For complete details on methods used, please consult Supplemental Experimental Procedures.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

Aldridge JW, Berridge KC, Rosen AR. Basal ganglia neural mechanisms of natural movement sequences. Canadian Journal of Physiology and Pharmacology. 2004; 82:732–739. [PubMed: 15523530]

Anderson DJ, Perona P. Toward a science of computational ethology. Neuron. 2014; 84:18–31. [PubMed: 25277452]

André E, Conquet F, Steinmayr M, Stratton SC, Porciatti V, Becker-André M. Disruption of retinoid-related orphan receptor beta changes circadian behavior, causes retinal degeneration and leads to vacillans phenotype in mice. The EMBO journal. 1998; 17:3867–3877. [PubMed: 9670004]

Berg HC, Brown DA. Chemotaxis in Escherichia coli analysed by three-dimensional tracking. Nature. 1972; 239:500–504. [PubMed: 4563019]

Berman GJ, Choi DM, Bialek W, Shaevitz JW. Mapping the structure of drosophilid behavior. Journal of the Royal Society, Interface / the Royal Society. 2014; 11:20140672.

Berridge KC, Fentress JC, Parr H. Natural syntax rules control action sequence of rats. Behavioural brain research. 1987; 23:59–68. [PubMed: 3828046]

Berwick RC, Okanoya K, Beckers GJL, Bolhuis JJ. Songs to syntax: the linguistics of birdsong. Trends in cognitive sciences. 2011; 15:113–121. [PubMed: 21296608]

Bizzi E, Tresch MC, Saltiel P, d'Avella A. New perspectives on spinal motor systems. Nature Reviews Neuroscience. 2000; 1:101–108. [PubMed: 11252772]

Brown, TG. The intrinsic factors in the act of progression in the mammal; Proceedings of the Royal Society of London Series B; 1911.

Colgan, PW. Quantitative ethology. John Wiley & Sons; 1978.

Crawley JN. Behavioral phenotyping of rodents. Comparative medicine. 2003; 53:140–146. [PubMed: 12784847]

Croll NA. Components and patterns in the behaviour of the nematode Caenorhabditis elegans. Journal of zoology. 1975; 176:159–176.

de Chaumont F, Coura RD-S, Serreau P, Cressant A, Chabout J, Granon S, Olivo-Marin J-C. Computerized video analysis of social interactions in mice. Nature Methods. 2012; 9:410–417. [PubMed: 22388289]

Drai D, Benjamini Y, Golani I. Statistical discrimination of natural modes of motion in rat exploratory behavior. Journal of neuroscience methods. 2000; 96:119–131. [PubMed: 10720676]

Eppig JT, Blake JA, Bult CJ, Kadin JA, Richardson JE, Group MGD. The Mouse Genome Database (MGD): facilitating mouse as a model for human biology and disease. Nucleic Acids Research. 2015; 43:D726–D736. [PubMed: 25348401]

Fendt M, Endres T, Lowry CA, Apfelbach R, McGregor IS. TMT-induced autonomic and behavioral changes and the neural basis of its processing. Neurosci Biobehav Rev. 2005; 29:1145–1156. [PubMed: 16099043]

Fentress JC, Stilwell FP. Letter: Grammar of a movement sequence in inbred mice. Nature. 1973; 244:52–53. [PubMed: 4582491]

Flash T, Hochner B. Motor primitives in vertebrates and invertebrates. Current opinion in neurobiology. 2005; 15:660–666. [PubMed: 16275056]

Garcia VA, Crispim Junior CF, Marino-Neto J. Assessment of observers' stability and reliability - a tool for evaluation of intra- and inter-concordance in animal behavioral recordings. Conference proceedings : Annual International Conference of the IEEE Engineering in Medicine and Biology Society IEEE Engineering in Medicine and Biology Society Conference. 2010; 2010:6603–6606.

Garrity PA, Goodman MB, Samuel AD, Sengupta P. Running hot and cold: behavioral strategies, neural circuits, and the molecular machinery for thermotaxis in C. elegans and Drosophila. Genes & Development. 2010; 24:2365–2382. [PubMed: 21041406]

Glickfeld LL, Andermann ML, Bonin V, Reid RC. Cortico-cortical projections in mouse visual cortex are functionally target specific. Nature Neuroscience. 2013; 16:219–226. [PubMed: 23292681]

Gomez-Marin A, Partoune N, Stephens GJ, Louis M, Brembs B. Automated tracking of animal posture and movement during exploration and sensory orientation behaviors. PLoS ONE. 2012; 7:e41642. [PubMed: 22912674]

Gray JM, Hill JJ, Bargmann CI. A circuit for navigation in Caenorhabditis elegans. Proceedings of the National Academy of Sciences of the United States of America. 2005; 102:3184–3191. [PubMed: 15689400]

Hastie, T.; Tibshirani, R.; Friedman, J. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Second. Springer; 2009.

Jhuang H, Garrote E, Yu X, Khilnani V, Poggio T, Steele AD, Serre T. Automated home-cage behavioural phenotyping of mice. Nature Communications. 2010; 1:68.

Kabra M, Robie AA, Rivera-Alba M, Branson S, Branson K. JAABA: interactive machine learning for automatic annotation of animal behavior. Nature Methods. 2013; 10:64–67. [PubMed: 23202433]

Lashley, KS. The problem of serial order in behavior. Jeffress, LA., editor. Psycholinguistics: A book of readings; 1967.

Liu H, Kim S-Y, Fu Y, Wu X, Ng L, Swaroop A, Forrest D. An isoform of retinoid-related orphan receptor β directs differentiation of retinal amacrine and horizontal interneurons. Nature Communications. 2013; 4:1813.

Markowitz JE, Ivie E, Kligler L, Gardner TJ. Long-range Order in Canary Song. PLoS computational biology. 2013; 9:e1003052. [PubMed: 23658509]

Masana MI, Sumaya IC, Becker-Andre M, Dubocovich ML. Behavioral characterization and modulation of circadian rhythms by light and melatonin in C3H/HeN mice homozygous for the RORbeta knockout. American journal of physiology Regulatory, integrative and comparative physiology. 2007; 292:R2357–R2367.

Ou-Yang T-H, Tsai M-L, Yen C-T, Lin T-T. An infrared range camera-based approach for three-dimensional locomotion tracking and pose reconstruction in a rodent. Journal of neuroscience methods. 2011; 201:116–123. [PubMed: 21835202]

Pierce-Shimomura JT, Morse TM, Lockery SR. The fundamental role of pirouettes in Caenorhabditis elegans chemotaxis. Journal of Neuroscience. 1999; 19:9557–9569. [PubMed: 10531458]

Sherrington C. The Integrative Action of the Nervous System. The. Journal of Nervous and Mental Disease. 1907

Spink AJ, Tegelenbosch RA, Buma MO, Noldus LP. The EthoVision video tracking system--a tool for behavioral phenotyping of transgenic mice. Physiology & behavior. 2001; 73:731–744. [PubMed: 11566207]

Stephens GJ, Johnson-Kerner B, Bialek W, Ryu WS. Dimensionality and Dynamics in the Behavior of C. elegans. PLoS computational biology. 2008; 4:e1000028. [PubMed: 18389066]

Stephens GJ, Johnson-Kerner B, Bialek W, Ryu WS. From Modes to Movement in the Behavior of Caenorhabditis elegans. PLoS ONE. 2010; 5:e13914. [PubMed: 21103370]

Swanson LW. Cerebral hemisphere regulation of motivated behavior. Brain research. 2000; 886:113–164. [PubMed: 11119693]

Tinbergen, N. The study of instinct. Oxford: Clarendon Press; 1951.

Tresch MC, Jarc A. The case for and against muscle synergies. Current opinion in neurobiology. 2009; 19:601–607. [PubMed: 19828310]

Tresch MC, Saltiel P, Bizzi E. The construction of movement by the spinal cord. Nature Neuroscience. 1999; 2:162–167. [PubMed: 10195201]

Vogelstein JT, Park Y, Ohyama T, Kerr RA, Truman JW, Priebe CE, Zlatic M. Discovery of brainwide neural-behavioral maps via multiscale unsupervised structure learning. Science (New York, NY). 2014; 344:386–392.

Wallace KJ, Rosen JB. Predator odor as an unconditioned fear stimulus in rats: elicitation of freezing by trimethylthiazoline, a component of fox feces. Behav Neurosci. 2000; 114:912–922. [PubMed: 11085605]

Weissbrod A, Shapiro A, Vasserman G, Edry L, Dayan M, Yitzhaky A, Hertzberg L, Feinerman O, Kimchi T. Automated long-term tracking and social behavioural phenotyping of animal colonies within a semi-natural environment. Nature Communications. 2013; 4:2018.

Wohlgemuth MJ, Sober SJ, Brainard MS. Linked control of syllable sequence and phonology in birdsong. Journal of Neuroscience. 2010; 30:12936–12949. [PubMed: 20881112]
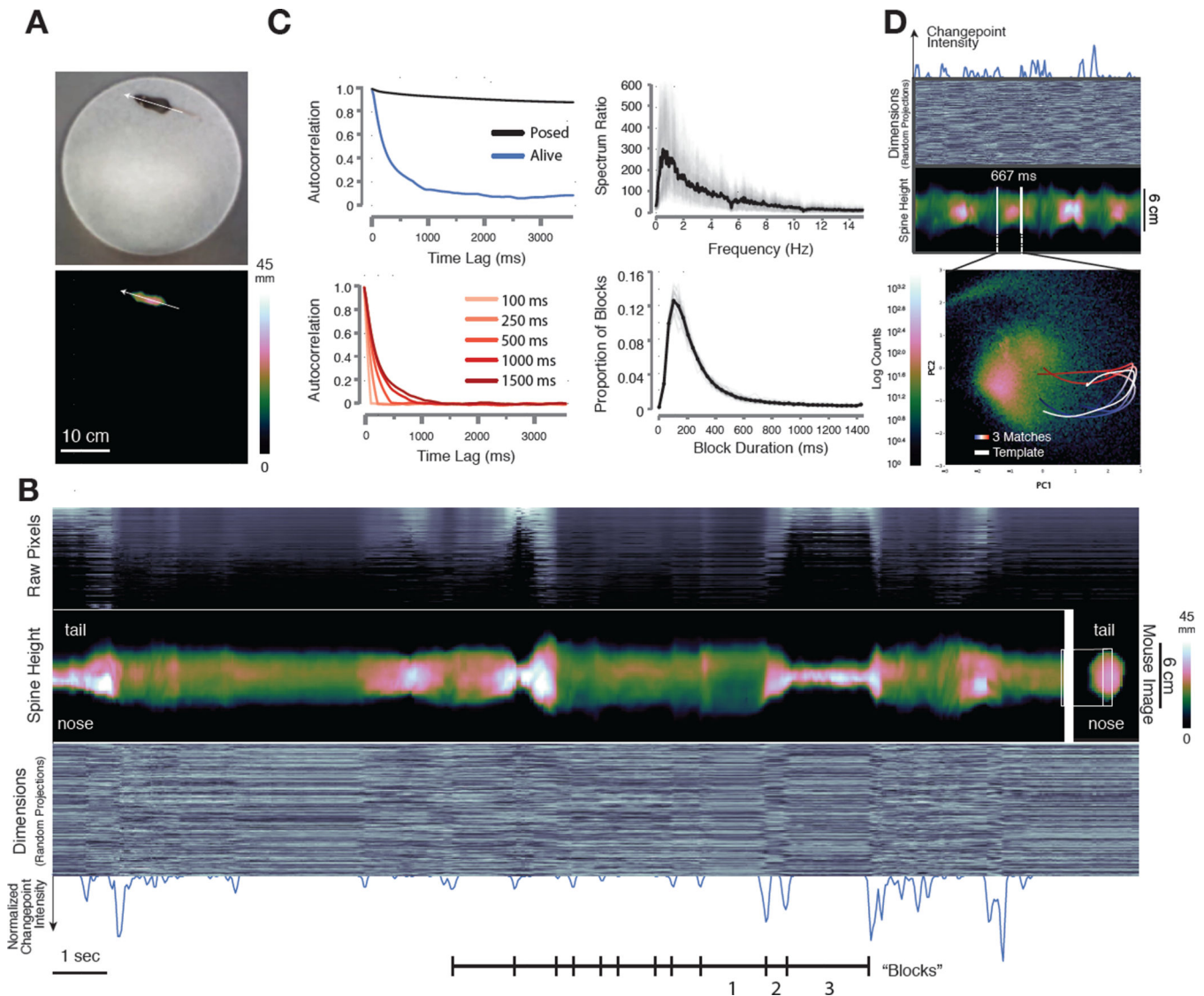
Fig. 1. Depth imaging reveals block structure in 3D mouse pose dynamics data
A. Mouse imaged in the circular open field with a standard RGB (top) and 3D depth camera
(bottom, mm = mm above floor). Arrow indicates the inferred axis of the animal's spine. **B**.
Raw pixels of the extracted and aligned 3D mouse image (**top panel**, sorted by mean
height), compressed data (**bottom panel**, 300 dimensions compressed via random
projections arrayed on the Y axis, pixel brightness proportional to its value), and height at
each inferred position of the mouse's spine (**middle panel**, "spine" data extracted from the
mouse as indicated on the right, mm = mm above floor) each reveal sporadic, sharp
transitions in the pose data over time. Note that the cross-sectional profile of the spine with
respect to the camera varies depending upon the morphology of the mouse; when reared this
profile becomes smaller, and when on all fours it becomes larger. Changepoints analysis
(**bottom panel**, blue trace = normalized changepoint probability) identifies approximate
boundaries between blocks. Blocks encode a variety of behaviors including locomotion (1),
a pointing rear with the mouse's body elongated with respect to the sensor (2), and a true

rear (3). **C, upper left**. Autocorrelation analysis performed on the top 10 dimensions of principal components data reveals that temporal correlation in the mouse's pose stabilizes after about 400 milliseconds (tau = 340 ± 58 ms). **C, lower left**. Shuffling behavioral data in blocks of 500 milliseconds or shorter destroys autocorrelation structure (shuffle block size indicated). **C, upper right**. Spectral power ratio between a behaving and dead mouse (mean plotted in black, individual mice plotted in grey) reveals most frequency content is represented between 1 and 6 Hz (mean = 3.75 ± .56 Hz). **C, lower right**. Changepoints-identified block duration distribution (mean = 358 ms, SD 495 ms, mean plotted in black, individual mice in gray, n=25 mice, 500 total minutes imaging,). **D**. Projecting mouse pose data (**top panels**, random projections and spine data depicted as in **B**) into Principal Component (PC) space (**bottom**) reveals that blocks of pose data encode reused trajectories (density of all recorded poses colormapped behind trajectories). Tracing out the path associated with a block highlighted by changepoint detection (**top**) identifies a trajectory through PC space (white). Similar trajectories identified through template matching (time indicated as progression from blue to red), are superimposed. Note that this procedure uses the first 10 PCs to identify matched trajectories, although only the first two PCs are depicted here.
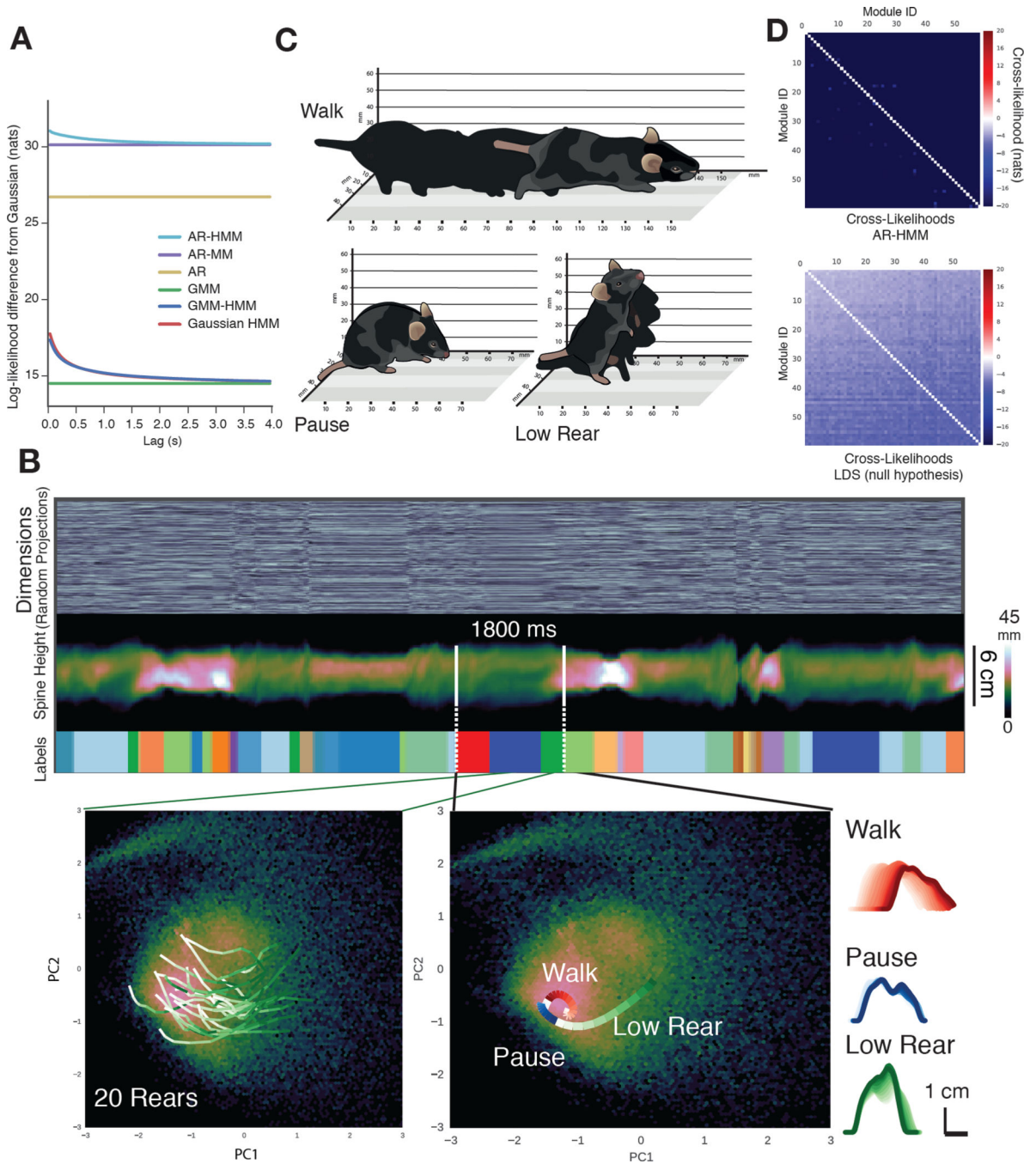
Fig. 2. Reused behavioral modules within mouse pose dynamics data
A. Predictive performance comparison of computational models describing possible
structures for behavior (details of each model and the comparison metric provided in
Supplemental Experimental Procedures). Models range from a Gaussian model (which
proposes that mouse behavior is built from modules, each a single Gaussian in pose space)
to an AR-HMM (which proposes that mouse behavior is built from modules, each of which
encodes an autoregressive trajectory through pose space, and which transition from one to
another with definable transition statistics; AR-MM = autoregressive mixture model, AR =

autoregressive model, GMM = Gaussian mixture model, GMM-HMM = GMM hidden Markov model, Gaussian HMM = Gaussian hidden Markov model). Model performance plotted as the log likelihood (Y-axis) ascribed to held-out test data at some time lag (X-axis) into the future (after subtracting Gaussian model performance). **B**. The AR-HMM parses behavioral data into identifiable modules (**top panels** – marked "labels", each module is uniquely color coded). Multiple data instances associated with a single behavioral module (between green lines, encoding a rear) each take a stereotyped trajectory through PC space (**bottom left**, trajectories progress from white to green over time, see Movie S4); multiple trajectories define behavioral sequences (**bottom center**, see Movie S6). Each trajectory within a sequence encodes a different elemental action (side-on view of the mouse calculated from depth data, **bottom right**, time indicated as increasingly darker lines, from module start to end). **C**. Isometric-view illustrations of the 3D imaging data associated with walk, pause and low rear modules (also see Movie S4). **D**. Cross-likelihood analysis depicting the likelihood that a data instance assigned to a particular module is well-modeled by another module. Cross-likelihoods were computed for the open field dataset (**above**, see Supplemental Experimental Procedures, units are nats, where e^{nats} is the likelihood ratio) and for module-free synthetic data whose autocorrelation structure matches actual mouse data (**below**).
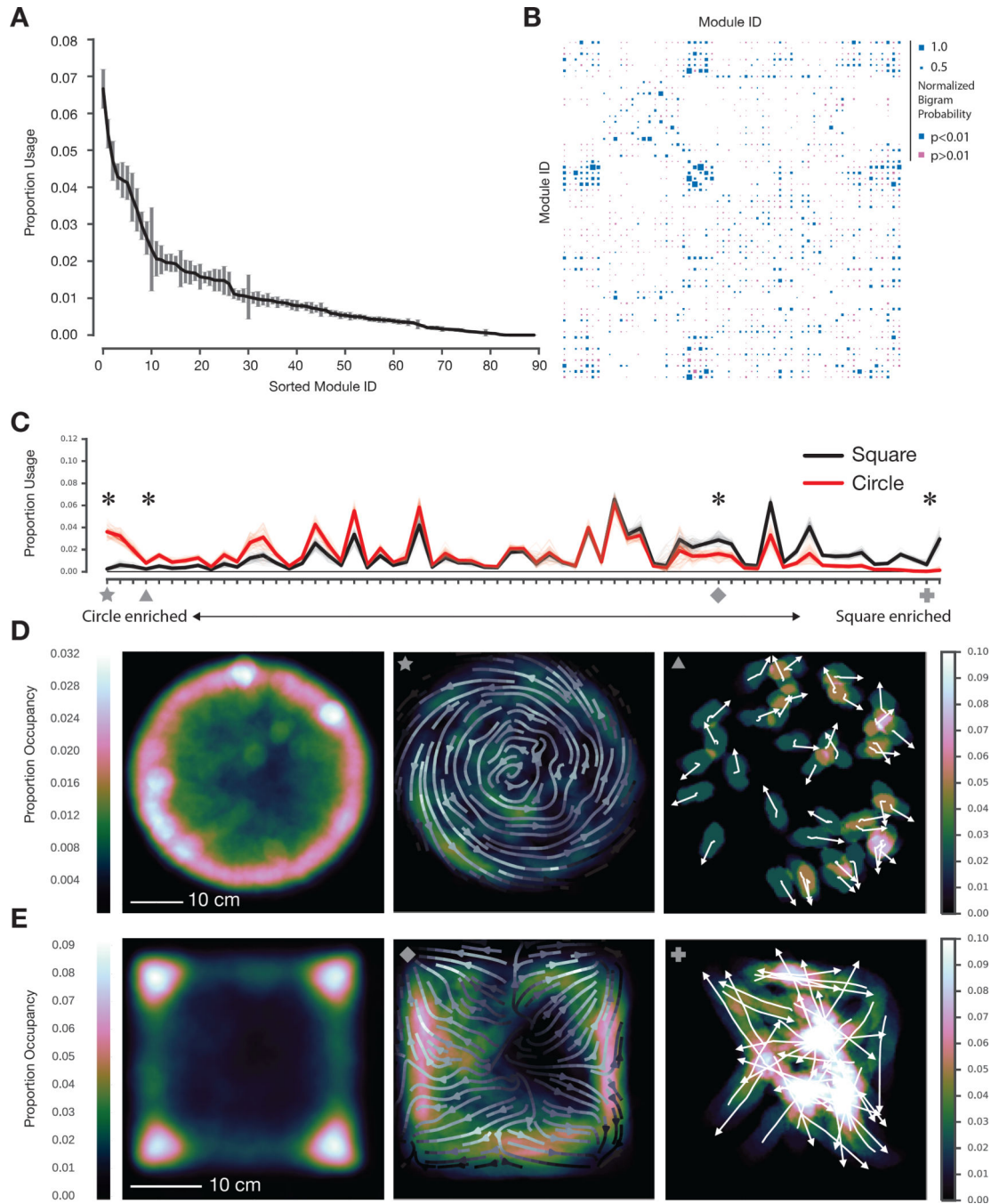
Fig. 3. The physical environment influences module usage and spatial pattern of expression
A. Modules identified by the AR-HMM, sorted by usage (n = 25 mice, 20 minutes per
mouse, data from circular open field, error bars are SEs calculated using bootstrap
estimation, n=100 bootstrap estimates, see Fig. S5E for Bayesian credible intervals). **B**.
Hinton diagram depicting the probability that any pair of modules is observed as an ordered
pair (p-values calculated via bootstrap estimation and color coded); modules were sorted by
spectral clustering to emphasize neighborhood structure. **C**. Module usage, sorted by context
(with "circleness" on left, overall usages differ significantly, $p < 10^{-15}$, Hotelling two-

sample t-squared test, see Supplemental Experimental Procedures for sorting details). Mean usages across animals depicted with dark lines, with bootstrap estimates depicted in fainter lines (n=100). Marked modules discussed in main text and shown in panel **D**: star = circular wall-hugging locomotion ("thigmotaxis"), triangle = outward-facing rears, diamond = square thigmotaxis, cross = square dart, see Movie S7. Usage for all marked modules significantly modulated by context (indicated by asterisk, Wald Test, Holm-Bonferroni adjusted p < 0.006 for square dart, otherwise p < 10^{-5}). **D**. Occupancy plot of mice in circular open field (**left**, n=25, 500 minutes total) indicating cumulative spatial positions across all experiments. Occupancy plot depicting deployment of circular thigmotaxis module (**middle**, average orientation across the experiment indicated as arrow field) and circle-enriched outward-facing rear module (**right**, orientation of individual animals indicated with arrows). **E**. Occupancy plot of mice in square box (**left**, n=15, 300 minutes total) indicating cumulative spatial positions across all experiments. Occupancy plot depicting a square-enriched thigmotaxis module (**middle**, average orientation across the experiment indicated as arrow field), and square-specific darting module (**right**, orientation of individual animals indicated with arrows).
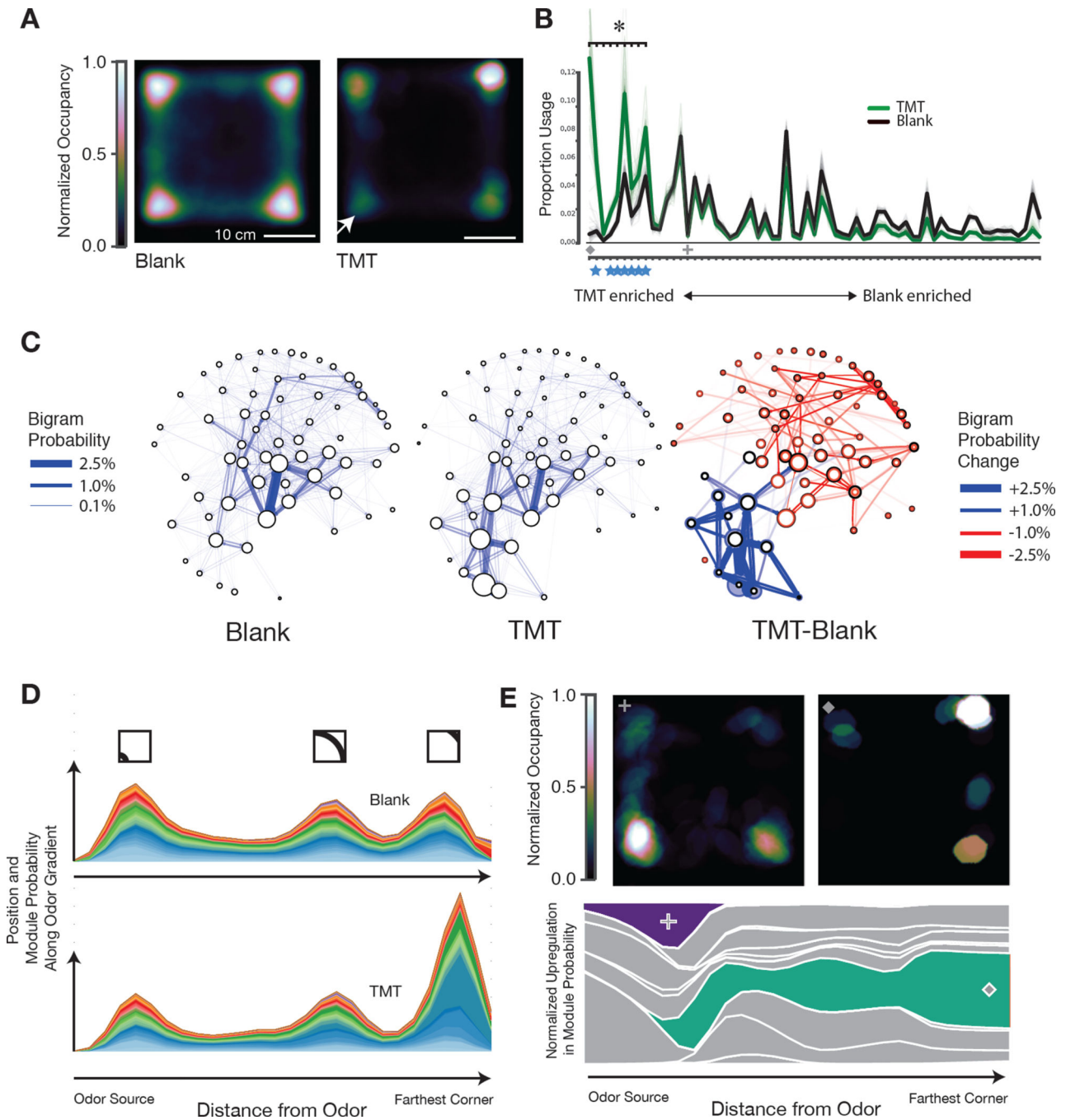
Fig. 4. Odor-driven innate avoidance alters transition probabilities
A. Occupancy plot of mice under control conditions (n=24, 480 minutes total) and after exposure to the fox-derived odorant trimethylthiazoline (TMT, 5% dilution, n=15, 300 minutes total, model co-trained on both conditions) in the lower left quadrant (arrow). Plots normalized such that maximum occupancy = 1. **B**. Module usage plot sorted by "TMT-ness" (Dark lines depict mean usages, bootstrap estimates depicted in fainter lines, sorting as in Fig. 3). Marked modules discussed in main text and panel **E**: cross = sniff in TMT quadrant, diamond = freeze away from TMT, see Movies S8 and S9. Blue stars indicate freezing

modules. Asterisk indicates statistically-significant regulation (Wald test, Holm-Bonferroni corrected, $p < 10^{-4}$). **C, left and middle**. Behavioral state maps for mice exploring a square box before and after TMT exposure, with modules depicted as nodes (usage proportional to the diameter of each node), and bigram transition probabilities depicted as directional edges. Graph layout minimizes the length of edges and is seeded by spectral clustering to emphasize local structure. **C, right**. Statemap depiction of the difference between blank and TMT. Usage differences are indicated by the newly-sized colored circles (upregulation indicated in blue, downregulation indicated in red, previous usages in control conditions indicated in black). Altered bigram transition probabilities are indicated in the same color code; only those significant transition probabilities (p<0.01) are depicted. **D**. Mountain plot depicting the joint probability of module expression and spatial position, plotted with respect to the TMT corner (X-axis); note that the "bump" two-thirds of the way across the graph occurs due to the two corners equidistant from the odor source (see inset for approximate position in square box, modules are color coded). **E**. Occupancy plot (**upper**) indicating spatial position at which mice after TMT exposure emit an investigatory sniffing module (**left**) or a pausing module (**right**, see Movie S8). Mountain plot (**lower**) indicating the differential deployment of these two modules (purple, green; other modules in grey) with respect to distance from the odor source.
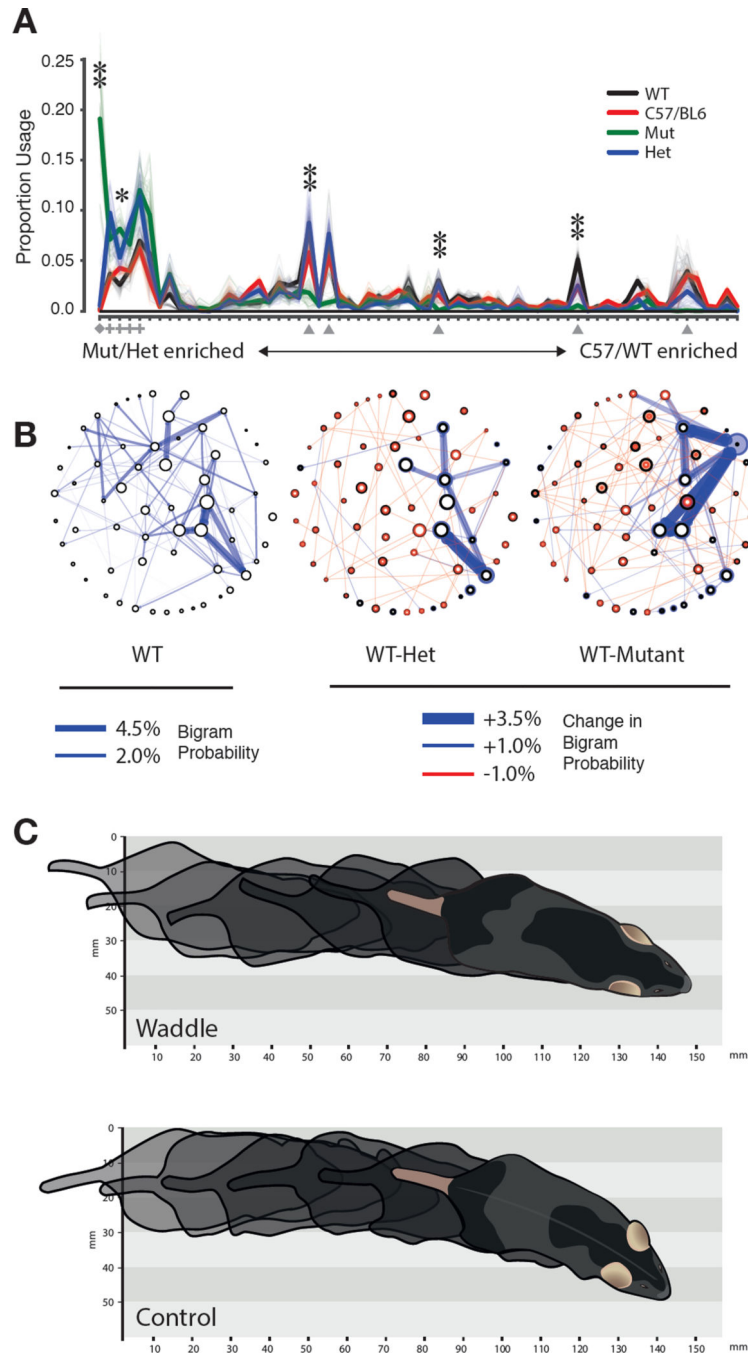
Fig. 5. The AR-HMM disambiguates wild-type, heterozygous and homozygous *Ror1*β mice
A. Usage plot of modules exhibited by littermate *Ror1*β mice (n=25 C57/BL6, n = 3 +/+, n = 6 +/−, n = 4 −/−, open field assay, 20 minute trials), sorted by "mutant-ness" (sorting and depiction as in Fig. 3). Wild-type module usage is not statistically different from C57, but differs significantly from homozygote and heterozygote (Hotelling two-sample t-squared test, $p < 10^{-15}$). Marked modules described in main text, diamond = waddle, triangle = normal locomotion, cross = pause. Single asterisk indicates significant usage difference between mutant and wildtype, $p < 0.05$, double asterisk indicates $p < 0.01$ under Wald test,

Holm-Bonferroni corrected. **B**. State map depiction of baseline OFA behavior for +/+ animals as in Fig. 4C (**left**); difference state maps as in Fig. 4C between the +/+ and +/− genotype (**middle**), and +/+ and −/− genotype (**right**); all depicted transitions that distinguish genotypes are statistically significant, p<.01. **C**. Illustration of the "waddle" module in which the hind limbs of the animal are elevated during walking (see Movie S10).
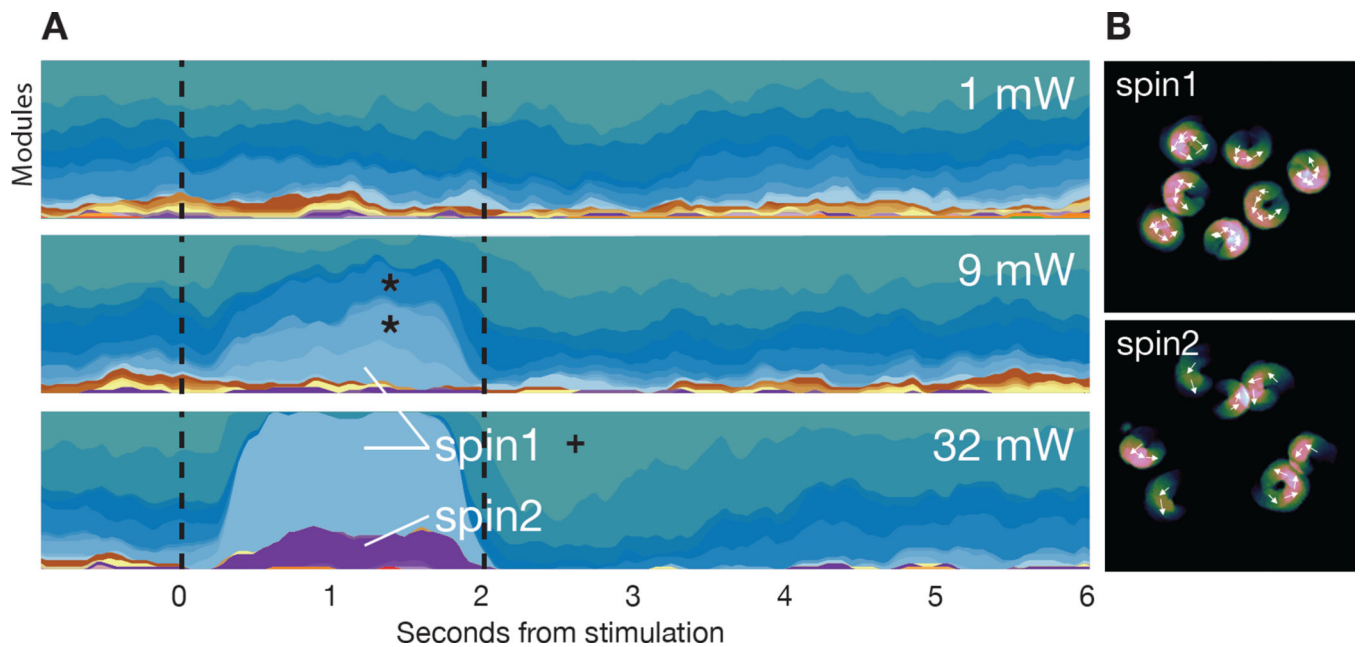
Fig. 6. Optogenetic perturbation of motor cortex yields both neomorphic and physiological modules

A. Mountain plot depicting the probability of expression of each behavioral module (assigned a unique color on the Y axis) over time (X axis), with light stimulation indicated by dashed vertical lines (each plot is the average of 50 trials). Note that modest variations in the baseline pattern of behavior, due to trial structure, are captured before light onset. Stars indicate two modules expressed during baseline conditions that are also upregulated at intermediate powers (11 mW) but not high powers (32 mW, Wald test, Holm-Bonferroni adjusted $p < 10^{-5}$); cross indicates pausing module upregulated at light offset (Wald test, Holm-Bonferroni adjusted, $p < 10^{-5}$). **B**. Average position of example mice (with arrows indicating orientation over time) of the two modules induced under the highest stimulation conditions (see Movie S11). Note that **A** and **B** are generated from one animal and that the observed modulations are representative of the complete dataset (n= 4 mice, model was trained separately from previous experiments).