Machine learning-based clustering and classification of mouse behaviors via respiratory patterns

Highlights
Respiration varies widely across mouse behaviors using multiple recording methods

Intranasal pressure recordings provide understudied respiratory features

Clustering reveals distinct respiratory features across behaviors

A classifier separates mouse immobility from differing contexts via respiration
SUMMARY
Breathing is dynamically modulated by metabolic needs as well as by emotional states. Even though rodents are invaluable models for investigating the neural control of respiration, current literature lacks systematic characterization of breathing dynamics across a broad spectrum of rodent behaviors. Here we uncover a wide diversity in breathing patterns across spontaneous, attractive odor-, stress-, and fear-induced behaviors in mice. Direct recordings of intranasal pressure afford more detailed respiratory information than more traditional whole-body plethysmography. K-means clustering groups 11 well-defined behavioral states into four clusters with distinct key respiratory features. Furthermore, we implement RUSBoost (random undersampling boost) classification, a supervised machine learning model, and find that breathing patterns can separate these behaviors with an accuracy of 80%. Taken together, our findings highlight the tight relationship between breathing and behavior and the potential use of breathing patterns to aid in distinguishing similar behaviors and inform about their internal states.

INTRODUCTION
As a vital sign, breathing is constantly modulated by physical and mental states. Rodents are incredibly instrumental models for understanding the neural control of respiration; however, the potential of using highly dynamic breathing patterns to reliably separate behaviors and even probe associated internal states has not been fully explored in rodents.

Breathing must be dynamic in order to maintain oxygen supply and expel carbon dioxide throughout the changing metabolic needs across different behaviors. Many studies have established the stimulating effects of hypoxia, hypercapnia, and exercise on breathing1,2 and the potential neural circuits that drive these changes.3,4 Furthermore, coordination of breathing also exists with more nuanced orofacial behaviors such as swallowing, chewing, licking, and whisking.5-9 Breathing also facilitates specific behaviors including odor sampling10,11 and vocalization.12-14 Collectively, breathing patterns likely vary drastically across behaviors and potentially serve as a physiological signal that helps to distinguish gross behaviors.

Beyond modulation by physical movement, breathing can also be influenced by emotional states, which coincide with autonomic responses.15 Evidence for a bidirectional relationship between breathing and emotion exists in both humans and rodents.16-18 The autonomic nervous system influences breathing through sympathetic and parasympathetic activation.19,20 In rodents, commonly used fear- and stress-based tasks induce stereotypical defensive behaviors, including both active (e.g., escape and struggling) and passive phases (e.g., freezing and immobility). These behaviors serve as proxies for internal fear and reflect diverging coping strategies. Within fear tasks, sympathetic tone dominates during the active response while parasympathetic tone dominates during the passive response.21 Recently, differences in breathing rate have been observed across passive fear behavior (freezing) elicited by distinct stimuli of auditory fear, contextual fear, and innate fear.22 These data suggest that passive behaviors evoked across fear paradigms appear behaviorally similar but may encompass different breathing patterns that potentially reflect underlying affective states.
An increasing number of studies have begun implementing breathing recordings (e.g.,\textsuperscript{12,14,22–27}), especially considering respiration-entrained rhythmic activity is detected in widespread brain regions in rodents and in humans.\textsuperscript{28–31} Most studies have limited their analysis to a small number of respiratory features such as breathing rate, peak inspiratory flow (or depth), and volume. However, a single breath consists of as many as four phases: an inhale, an inhale pause, an exhale, and a following exhale pause prior to the initiation of the next inhale. The dynamics of these individual breath phases not only impact overall frequency but also influence measurements of duty cycle and variability of breathing patterns.\textsuperscript{31} Here, we implement BreathMetrics, a MATLAB toolbox, which first reliably labels these four phases within respiration and then uses the labeling to generate a comprehensive quantitative summary that includes over 20 respiratory parameters.\textsuperscript{31} The BreathMetrics summary importantly evaluates previously understudied parameters such as inhale and exhale pauses and measurements of variability, which may inform physical and mental states in distinct ways. Multiple techniques, such as movement, muscle, temperature, and airflow/pressure recordings, have been used to detect respiration in freely behaving rodents.\textsuperscript{32,33} In particular, airflow/pressure measurements, especially the intranasal pressure recording, can provide robust measurements of respiratory features including the inhale/exhale pauses in both humans and rodents.\textsuperscript{31}

Machine learning has recently been effectively applied in rodents to recognize specific body movements,\textsuperscript{34} behavioral syllables,\textsuperscript{35} facial expressions in response to emotionally salient stimuli,\textsuperscript{36} pain responses,\textsuperscript{37} and ultrasonic vocalizations.\textsuperscript{38} Likewise, human respiration data have been used to detect emotions with some success.\textsuperscript{39} Machine learning may prove useful in objectively identifying nuanced differences in breathing patterns of rodents across emotionally salient behavioral contexts.

Here we sought to probe the wide spectrum of rodent breathing patterns and the use of breathing to distinguish behaviors. We utilized two airflow-based methods, whole-body plethysmography (WBP) and intranasal pressure recording, to monitor breathing during spontaneous behaviors as well as in response to attractive odor, fearful stimuli, and stress in freely behaving mice. We assessed the diversity in breathing patterns across a range of well-defined behaviors using principal component analysis (PCA) and unbiased k-means clustering. Through the application of RUSBoost classification to intranasal pressure recordings, we classified these behaviors with an accuracy of nearly 80%. This study highlights the strong interaction between breathing and rodent behavior and its potential to aid in separating behaviorally similar states.

### RESULTS

**Breathing patterns across behaviors measured by whole-body plethysmography**

WBP is a widely used, noninvasive method for measuring respiratory function in awake animals. Using this method, we initially evaluated breathing patterns during olfactory fear conditioning trials from five different behaviors: investigation, grooming, quiescence, and shock-related arousal and freezing (Figures 1A–1F, Table S1). To standardize respiration data across animals, the recordings were the first baseline corrected and z-scored using the BreathMetrics toolbox.\textsuperscript{31} BreathMetrics was then employed to identify specific respiratory features to create a quantitative summary of breathing from each defined behavior (Figure 1G and Table 1). Reliably detected respiratory features such as breathing rate, average peak inspiratory flow, coefficient of variation (CV) of breathing volume, CV breathing rate, and both duty cycle of the inhale and exhale significantly differed across the behaviors we evaluated (Figures 2A–2G). Given that we aimed to understand whether the combination of these features may potentially generate unique breathing patterns for each behavior, we compiled 7 defining features of WBP respiration into PCA (Figure 2H and Table 1).

We observed that, in PC space, breathing from these behaviors was predominantly divided into a high breathing rate group (investigation, post-shock arousal) versus a low breathing rate group (grooming, freezing, quiescence) (Figures 2A and 2H). We suspected that within the high or low breathing rate group, these behavioral states are associated with varying levels of emotional salience or arousal, and breathing patterns may reflect this. While post-shock arousal and investigation exhibited differences in peak inspiratory flow and CV breathing rate (Figures 2E and 2F), these distinctions were not large enough in magnitude to separate these breathing patterns in PC space. Similarly, quiescence has a minor decrease in breathing rate and CV breathing volume in comparison to freezing breathing patterns. However, these behaviors remain largely interspersed in breathing PC space when considering all features comprehensively. To classify these behaviors using breathing, we implemented random undersampling boost algorithm (RUSBoost), a supervised machine learning method. This method works well with imbalanced sample sizes of breathing across the behaviors in the training data and operates off raw respiratory parameters from BreathMetrics.\textsuperscript{31} To optimize classification, we created a
series of models to evaluate performance across 20 different combinations of respiratory parameters from BreathMetrics (Figure S1A). In testing the series of models, the removal of each respiratory parameter was intended to preserve variability in the data by removing highly correlated variables first. Each model was first created on 80% of the overall breathing data shown in Figure 2 as a training dataset and then hypertuned using 10-fold cross-validation on the training data (see STAR methods for details). Prediction accuracy was determined by evaluating the classification of unlabeled testing data (remaining 20% of the overall data in Figure 2). The RUSBoost classifier with highest performance resulted in a prediction accuracy of 68.3% across the five behaviors (Figures S1B and S1C). Across behaviors, investigation and post-shock arousal showed reasonable prediction accuracies, 69.1% and 65.0% respectively, but were commonly mislabeled as each other seen in Figure S1D. The lowest prediction accuracy occurred for freezing (34.1%) and arose from predominantly the mislabeling of freezing to quiescence behavior (Figure S1D). While quiescence demonstrated a higher prediction accuracy of 81.3%, quiescence was also mislabeled as freezing. The increased CV of breathing volume and rate during grooming may sufficiently differentiate this breathing pattern from other low-frequency patterns, allowing for the highest prediction accuracy of 92.0% (Figures 2B, 2E, and S1C). Overall, the RUSBoost classifier based on the WBP data is quite successful in distinguishing certain behaviors (e.g., grooming) but not others (e.g., freezing).

Breathing patterns across behaviors measured via intranasal pressure

Given the inability to distinguish emotionally salient freezing from other behaviors using WBP, we sought an additional technique to record respiration. In contrast to WBP, intranasal pressure recordings

---

Table 1. Principal component analysis of whole-body plethysmography data using 7 BreathMetrics parameters

| Respiratory Parameter | PC1   | PC2   | PC3   |
|-----------------------|-------|-------|-------|
| Breathing Rate (Hz)   | 0.543 | -0.131| -0.320|
| Duty Cycle Inhale     | 0.377 | -0.207| 0.643 |
| CV Breathing Rate     | 0.324 | 0.581 | 0.113 |
| CV Breathing Volume   | 0.034 | 0.769 | 0.011 |
| Peak Inspiratory Flow| 0.028 | -0.030| -0.072|
| Avg Inhale Duration (s)| -0.460| 0.065 | 0.591 |
| Duty Cycle Exhale     | -0.496| 0.070 | -0.340|
| Variance Explained (%)| 46.692| 23.705| 17.787|

Loading coefficients for 7 respiratory parameters across 3 PCs. Bolded loading coefficients for each PC have a magnitude >0.4 and strongly contribute to that PC. PC, principal component.
We first examined respiration across 11 well-defined behaviors in freely behaving mice using intranasal pressure recordings (Figure 3A). These 11 behaviors included attractive odor (peanut oil)-induced investigation (Figure 3B), grooming (Figure 3C), non-rapid eye movement (NREM) sleep (Figure 3D; see STAR methods for brain state determination), and four emotionally salient tests (Innate Fear-2MT, Conditioned Fear Retrieval, Tail Suspension Test (TST), and Restraint Stress) each of which had an active and an inactive state (Figures 3E–3H). These emotionally salient tests are frequently employed in rodent literature to investigate fear responses and motivation in response to stress. Innate fear was induced via a brief exposure to 2MT (2-methyl-2-thiazoline), a synthetic predator odor analog. 2MT induced robust freezing, a stereotypical fear behavior, with intermittent locomotion (Figure 3E). We also evaluated conditioned fear via a two-day protocol. Day 1 consisted of four pairings of an auditory cue with a mild 1.0 s footshock; on day 2, mice underwent fear retrieval through re-exposure to the auditory cue four times in a novel environment. Similar to innate fear, mice alternated between locomotion and freezing during the retrieval session (Figure 3F). Tail suspension test (e.g.,24), consisting of suspending each mouse by the tail for 6 min, induced alternating bouts of struggling and immobility (Figure 3G). Lastly, restraint stress consisted of 30 min constraint in a clear tube with an opening for ventilation. Mice also responded to restraint stress with alternating bouts of struggling and immobility (Figure 3H).
We used BreathMetrics to quantify respiratory parameters of breathing during these behaviors following baseline correction and z-scoring. Compared to WBP (Figure 1G), intranasal pressure data allow consistent inhale and exhale pause detection (note a few inhale pauses in Figure 4A). To minimize potential noise that may arise from redundant parameters in PC space, we reduced the initial 20 quantitative BreathMetrics parameters to the first critical 13 respiratory parameters shown in Figure 5A. While we focused on these 13 parameters, we observed consistent clustering across a range of parameter combinations, including the comprehensive 20 parameters. For clustering, the 7 removed parameters were either correlated with or derived from the 13 included parameters (Figure S2). For instance, volume estimations are calculated from corresponding peak flow and inhale/exhale duration parameters. Furthermore, some parameters exhibit strong relationships with other included parameters, such as peak expiratory flow’s anti-correlation with inspiratory flow and CV inhale duty cycle’s correlation with CV breathing rate. Given these natural underlying correlations between parameters, it is unsurprising that clustering remains consistent while evaluating various combinations of parameters. Therefore, we compiled these 13 parameter summaries for a total of 938 bouts of behavior (Table S2) across 24 mice into PCA (Table 2). BreathMetrics parameters quantifying inhale and exhale pauses have high loading coefficients within PC1 and PC3, respectively, indicating the importance of these measures in breathing variability (Table 2).

**Similarities in breathing across behaviors revealed by k-means clustering**

To understand the underlying associations between breathing and behavioral contexts, we used PCA for dimension reduction and k-means clustering to test whether the breathing patterns fall into separate clusters. Based on an elbow plot analysis (Figure S3), we determined that four clusters adequately captured the grouping of breathing patterns (color coded in Figure 4B). Each of the four respiratory clusters represents a unique combination of respiratory features (Figures 4C–4O). Cluster 1 (orange; termed High Rate) was...
Figure 4. K-means clustering reveals defining features for clusters of intranasal pressure breathing recordings

(A) BreathMetrics labeling of intra-nasal pressure respiratory features. (B) PCA was performed using 13 respiratory parameters, shown in (C–O), for each well-defined behavior bout (n = 938). Four clusters, represented by different colors, were defined by k-means clustering. Orange cluster (n = 379 bouts), Blue cluster (n = 249 bouts), Purple cluster (n = 199 bouts), Green cluster (n = 111 bouts).

(D and H) High breathing rate and high inspiratory flow breathing comprise the orange cluster. (E and I) Low CV of breathing rate and increased duty cycle of the exhale define the blue cluster. (F and J) Increased percentage of breaths with inhale pauses and longer duration inhale pauses exist in the purple cluster. (G and K) Breathing bouts with higher CV of breathing volume and increased percentage of breaths with exhale pauses comprise the green cluster.

(C) The green cluster also included an increased CV exhale duty cycle that may stem from the increased exhale pause percentage (K) and duration (O). Kruskal-Wallis test followed by post-hoc Dunn’s multiple comparisons test in (C–O). Data are displayed as mean ± SEM. For clarity, nonsignificant post-hoc comparisons are not shown. Six outlier data points are not shown: 1 in CV exhale duty cycle (C), 1 in inspiratory flow (H), 1 in inhale pause duration (J), 2 in inhale duration (L), 1 in exhale duration (N). PC, principal component. CV, coefficient of variation. *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001. See also Figures S2–S5.
Characterized by higher breathing rate and higher peak inspiratory flow than the remaining three clusters (Figures 4D and 4H). Cluster 2 (blue; termed Steady Rate) contained steady, even breathing patterns, as demonstrated by lower CV breathing rate and an even exhale-to-inhale length (duty cycle exhale close to 0.5) (Figures 4E and 4I). Cluster 3 (purple; termed Inhale Pauses) consisted of a higher percentage of breaths with longer inhale pauses (Figures 4F and 4J). Lastly, Cluster 4 (green; termed Exhale Pauses) contained breathing patterns with higher CV breathing volume and an increased percentage of breaths with exhale pauses (Figures 4G and 4K). These four respiratory clusters remain highly stable in their key features when the total dataset is split into two from different subsets of mice (Figure S4).

Each respiratory cluster captured breathing patterns for a different group of behavioral contexts (Table 3). The High Rate cluster was predominantly comprised of respiration during odor investigation, 2MT locomotion, retrieval locomotion, and struggling during restraint. Two respiratory clusters, Steady Rate and Inhale Pauses, largely contained breathing patterns during inactive behaviors. The Steady Rate cluster was composed of respiration from sleep, immobility during restraint stress, freezing from fear retrieval, and immobility during TST. The Inhale Pauses cluster consisted of breathing from 2MT freezing, immobility during TST, and retrieval freezing. The final Exhale Pauses cluster was composed of respiration from grooming, struggling during TST, 2MT freezing, and struggling during restraint (Table 3).

Through tracking whether respiratory bouts from the same behavior were consistently assigned to the same cluster, we assessed how robust respiration patterns are for each behavior. Respiratory bouts within 7 of 11 behavioral classes were consistently assigned to the same cluster (>80% of total behavioral bouts).
These results highlight the robust nature of breathing patterns regardless of potential individual differences. Behaviors with consistent breathing included odor investigation, 2MT locomotion, retrieval locomotion, struggling during restraint, sleep, immobility during restraint, and 2MT freezing. Surprisingly, breathing patterns associated with retrieval freezing, both states during TST, and grooming were distributed across multiple clusters, indicating higher variability of breathing patterns during these behaviors. To understand this diversity in clustering, we investigated how bout duration of each behavior varies across the clusters. We found no relationship suggesting that respiratory features are independent of the duration of retrieval freezing and both behavioral phases of TST (Figures S5A–S5D). Another potential explanation for the spread across clusters could be sampling timepoint from each behavioral test, particularly in tasks like fear retrieval and TST. For instance, by the end of TST, mice potentially experience physical fatigue, and thus, breathing at the beginning of the task may comprise different attributes from breathing at the end of the task. While no obvious trend exists for retrieval freezing or TST struggling, interestingly, initiation

Table 2. Principal component analysis of intranasal pressure data using 13 BreathMetrics parameters

| Respiratory Parameter                              | PC1       | PC2       | PC3       |
|---------------------------------------------------|-----------|-----------|-----------|
| Avg Inhale Pause Duration (s)                      | 0.430     | −0.118    | −0.242    |
| % Breaths Inhale Pause                            | 0.380     | −0.283    | −0.207    |
| Breathing Rate (Hz)                               | −0.376    | −0.359    | 0.178     |
| Duty Cycle Inhale                                 | −0.360    | −0.098    | 0.124     |
| Avg Exhale Duration (s)                           | 0.081     | 0.594     | −0.113    |
| Duty Cycle Exhale                                 | −0.341    | 0.413     | 0.095     |
| Avg Exhale Pause Duration (s)                      | 0.167     | 0.177     | 0.482     |
| % Breaths Exhale Pause                            | 0.151     | 0.183     | 0.445     |
| CV Breathing Volume                               | 0.154     | −0.111    | 0.411     |
| CV Exhale Duty Cycle                              | 0.290     | 0.088     | 0.315     |
| CV Breathing Rate                                 | 0.265     | −0.211    | 0.295     |
| Avg Inhale Duration (s)                           | 0.206     | 0.268     | −0.170    |
| Peak Inspiratory                                  | 0.078     | 0.205     | −0.121    |
| Variance Explained (%)                            | 28.782    | 19.638    | 15.445    |

Loading coefficients for 13 respiratory parameters across 3 PCs. Bolded high magnitude (>0.3) loading coefficients exhibit which parameters have a higher contribution to each PC. Note the importance of inhale pause parameters for PC1 and exhale pause parameters for PC3. PC, principal component.

Table 3. The distribution of behaviors across clusters from intranasal pressure data shows the consistency of breathing patterns during most behaviors

| Consistency                                      | High Rate (orange) | Steady Rate (blue) | Inhale Pauses (purple) | Exhale Pauses (green) |
|--------------------------------------------------|---------------------|--------------------|------------------------|-----------------------|
| Odor Investigation                               | 100                 | –                  | –                      | 100                   |
| 2MT Locomotion                                   | 86.25               | –                  | 12.50                  | 1.25                  | 100                   |
| Retrieval Locomotion                             | 97.10               | –                  | 2.90                   | –                     | 100                   |
| Restraint Struggling                             | 88.42               | 1.06               | 2.11                   | 8.42                  | 100                   |
| TST Struggling                                   | 42.31               | 3.85               | 25.64                  | 28.21                 | 100                   |
| Grooming                                         | 37.74               | –                  | 1.89                   | 60.38                 | 100                   |
| Sleep                                            | –                   | 87.80              | –                      | 12.20                 | 100                   |
| Restraint Immobility                             | 2.36                | 82.68              | 4.72                   | 10.24                 | 100                   |
| Retrieval Freezing                               | 30.43               | 38.04              | 27.17                  | 4.35                  | 100                   |
| 2MT Freezing                                      | 4.17                | 2.50               | 81.67                  | 11.67                 | 100                   |
| TST Immobility                                   | 7.69                | 38.46              | 44.87                  | 8.97                  | 100                   |

Consistency measures the percentage of each behavior that is assigned to each cluster. Numbers in bold denote >80%.
timepoint of the behavior may play a role in the distribution of breathing from TST immobility across the two inactive clusters (Figures S5E–S5G). Finally, for these behaviors, individual mice showed variability in the clustering of their breathing patterns (Figures S5H–S5J) excluding the possibility that individual differences solely explain the spread in clustering.

**Prediction of behaviors from breathing patterns via random undersampling boost classification**

Given the consistent clustering of breathing for multiple behaviors, we next investigated how accurately respiration distinguishes behaviors using supervised classification. We first began with a parameter optimization study by assessing model performance across a series of 20 parameter combinations (Figure 5A). For each combination of parameters, an RUSBoost model was built and trained on 80% of the overall intranasal pressure data. Each trained model was hypertuned using 10-fold cross-validation (see STAR methods for details). Prediction accuracy was then assessed on the unlabeled testing data, the remaining 20% of the intranasal pressure data. Given that 2MT locomotion and retrieval locomotion were very similar to odor investigation, we collapsed these behavioral states into a single category (exploration) in the RUSBoost model. Model performance stayed in a relatively tight range of 67%–78.8% when the number of parameters increased from 3 to 19. This robust performance likely relates to the natural correlations between parameters (Figure S2). Classification accuracy deteriorated when only using the common breathing measures of breathing rate and peak inspiratory flow or amplitude. The overall highest prediction accuracy was 78.8% for the 9 behavioral states with some variations between classes (Figure 5C). The confusion matrix shown in Figure 5D depicts the cases of inaccurate classification for each behavioral label. We further demonstrate stable classification accuracy across individual mice for stress- and fear-induced behaviors (Figures S5K–S5N). Overall, active behaviors are more likely to be mislabeled as other active behaviors whereas inactive behaviors are more likely to be mislabeled as other inactive behaviors. Given this trend, we attempted to further distinguish behaviors by creating a separate “active only” classifier vs an “inactive only” classifier. We found no significant changes in prediction accuracy across behaviors when utilizing two separate models in comparison to the comprehensive model shown in Figure 5C (data not shown).

**Identification of similar breathing patterns for novel behavioral states**

Lastly, we aimed to demonstrate the ease for others to reference novel breathing patterns to the previous clustering respiratory space (Figures 6A–6D). Using BreathMetrics, we summarized breathing patterns for bouts of novel cage rearing, post-shock arousal and post-shock freezing from fear conditioning trials, and home cage quiescence. For each of these breathing bouts from additional behaviors, we first assessed the distance to pre-existing cluster centers (Figures 6E–6H). Short distances along with the visualization of each novel behavior in the clustering space (Figures S6A–S6D) suggest that each behavior does not constitute a novel breathing cluster previously unidentified. Rearing, an investigative rodent behavior, most commonly fell in the High Rate cluster and had the shortest distances to the exploration and restraint struggling categories (Figures 6E and 6I), as did breathing during post-shock arousal (Figures 6F and 6J). Interestingly, breathing during post-shock arousal did not display significant similarity to active stress response behaviors such as struggling during restraint or TST. Both rearing and post-shock arousal are statistically further away from most inactive behaviors in comparison to chance (Figures S6E and S6F).

Given that we observed a separation in clustering from a conditioned freezing response and an innate freezing response, we evaluated the breathing from an additional odor-independent innate freezing response through post-shock freezing. Similar to retrieval freezing, breathing from post-shock freezing had similar distances to three of the four clusters (Figure 6G). Notably, breathing from post-shock freezing showed shortest distance to the retrieval freezing category rather than 2MT freezing (Figure 6K) and was statistically closer to retrieval, 2MT freezing, and TST immobility in comparison to chance (Figure S6G). This overall suggests that the differences in breathing during 2MT freezing in the clustering may originate from the sensory modality of 2MT and not its innate fear quality. Finally, our clustering of breathing incorporates multiple inactive behavioral states with varying levels of arousal. We, therefore, evaluated behavioral quiescence which encompasses wakefulness from the brain state, no movement, and low arousal, as it typically occurred prior to sleep. Quiescence bouts showed a larger spread in the distance to cluster centers but were closest to the Steady Rate cluster (Figure 6H). In evaluating proximity to the 9 individual behaviors, quiescence is closest in distance to sleep in comparison to other forms of inactive behaviors evoked by emotionally salient tasks (Figures 6L and S6H). These results suggest that the clustering space can potentially be used to evaluate respiration patterns in new behavioral states not included in the current study.
DISCUSSION

In this study, we examined a wide array of breathing patterns that manifest during rodent behavior, including emotionally salient states. Compared to WBP, intranasal pressure recordings allow the extraction of more breathing features using BreathMetrics. Through PCA and k-means clustering, we identified four clusters of breathing data that are comprised of distinct features to create unique breathing patterns. Applying supervised machine learning, we then demonstrated that breathing distinguishes these behaviors from each other with acceptable accuracy. Notably, several behaviorally similar states (e.g., immobility and freezing) are differentiated based on breathing patterns, highlighting the tight relationship between breathing and rodent behavior. Importantly, these understudied nuances in breathing, such as measurements in variability or number of inhale/exhale pauses, serve as strong future candidates for investigating potential differences in rodent internal states during emotionally salient behavioral tasks.

Multiple methods have been adapted to record respiration in behaving animals, and each has its own strengths and weaknesses.\textsuperscript{32} The WBP breathing data separated behaviors largely by breathing rate as seen in PCA (Figures 1 and 2; Table 1). Intranasal pressure recordings permitted more refined breathing feature extraction and further separation of rodent behavior (Figures 3, 4, and 5; Tables 2 and 3). Practically, WBP restricts a mouse to a small chamber which may influence overall physiology and stress levels. While WBP is susceptible to large motion artifacts,\textsuperscript{44} intranasal pressure recording requires tethering with tubing.
and may be impacted by nasal clogging.\textsuperscript{45} Despite the differences in these two methods, our study identified some similarities in breathing parameters across behaviors. We found that investigation and post-shock arousal share similar breathing features and that grooming respiration contains higher variability using both methods (Figures 1, 2, 3, and 4). More importantly, using intranasal pressure measurement, we were able to obtain breathing patterns from more behaviors including emotionally salient tasks that would not be feasible with WBP. However, this difference in total behaviors evaluated within WBP and intranasal pressure recordings limits the ability to directly compare the prediction accuracy of certain behaviors (i.e., grooming) between the methods. We acknowledge that prediction accuracy can be influenced by which behaviors are included in the model.

K-means clustering based on intranasal pressure measurement reveals which behaviors share similar breathing patterns. The High Rate cluster (characterized by high peak inspiratory flow and high rate breathing) depicts the shared features between odor investigation, locomotion during 2MT exposure and conditioned fear retrieval, and struggling during restraint (Figure 4 and Table 3). Given the observed movement during the active phase of 2MT and retrieval tests was predominantly low-intensity locomotion, our results highlight that the stimulating effect of locomotion on breathing resembles odor investigation in these cases. In contrast, struggling during TST did not map to the High Rate cluster as consistently as the other active behaviors (Table 3). While the average peak inspiratory flow was similar for struggling during TST versus other active behaviors, the average breathing rate was lower during TST struggling (Figure S7). Some potential reasons for the additional diversity in breathing from struggling during TST relate to the physical context of this test or the intensity of physical exertion. Inversion of the mouse during TST likely increases the force necessary to contract the diaphragm due to gravity, and the interaction between abdominal viscera and lungs may promote coupling of breathing to activity.\textsuperscript{46} The clustering analysis also reveals which behaviors have highly consistent respiration patterns (Table 3). For these behavioral states (e.g., sleep), a significant deviation in the breathing pattern may predict underlying pathological conditions as shown in a recent study on patients with Parkinson’s disease.\textsuperscript{47}

Notably, incorporating BreathMetrics features surrounding all phases of breathing (e.g., inhale and exhale pauses) facilitates the separation of behavioral states by breathing. Previously, the simultaneous implementation of neck electromyogram and electroencephalogram has enabled the distinction of sleep from freezing.\textsuperscript{48} In this analysis, two respiratory clusters (Steady Rate and Inhale Pauses) emerge for inactive behaviors, and a diverging feature between them is percentage and duration of inhale pauses (Figure 4). The Inhale Pauses cluster primarily contains breathing from 2MT freezing, immobility during TST, and retrieval freezing. Increased inhale pauses during freezing may be a result of parasympathetic dominance, typically measured through bradycardia,\textsuperscript{19–31} and the accompanying heightened sensory arousal. Interestingly, restraint stress breathing in both the active and inactive behaviors lacks inhale pauses while TST breathing contains a similar level across active and inactive states (Figures S7 and S8). The lack of inhale pauses during restraint stress potentially reflects enduring dominance of sympathetic activation across both states during restraint.\textsuperscript{52,53} The Steady Rate cluster shows a strong overlap in breathing features between sleep and immobility during restraint (Table 3), which do not have these inhale pauses (Figure S8). During sleep (relaxed situation), lack of inhale pauses may be due to passive expiration, while during restraint immobility (stressed situation), this may be potentially caused by hyperventilation without breath holds. Interestingly retrieval freezing breathing patterns showed variability across the clusters in comparison to the recent reports on the relatively steady respiration-entrained rhythmic activity (centered around 4 Hz) found in the fear circuits.\textsuperscript{22,25,26} The Exhale Pauses cluster, characterized by increased amount of exhale pauses and higher variability, both in the rate and volume, mainly contains breathing from grooming and struggling during TST (Table 3 and Figure 4). Grooming contains its own nuanced patterns of stereotyped motor behavior that can be categorized into several separate phases.\textsuperscript{54} Given the stereotypical nature of grooming, in the future determining the potential phase coupling between grooming strokes and respiration would be of interest.

Implementing RUSBoost classification distinguishes 9 behavioral states from each other using breathing with reasonable accuracy (Figure 5). We combined the locomotion of 2MT and conditioned fear retrieval with odor investigation into one label (Exploration), given the overall similarity of these behaviors. As sniffing serves to sample the external environment,\textsuperscript{1} a natural component of low-speed locomotion, the commonalities between breathing during locomotion and odor investigation are logical.
Most behaviors, including those spread across the clusters (e.g., retrieval freezing), exhibit reasonable prediction accuracies (Figure 5). Furthermore, our study shows the importance of investigating detailed features of breathing across behaviors in addition to the standard components of breathing rate and peak inspiratory flow. We observe robust performance of the model across a range of parameter combinations which deteriorates when performance is based solely on breathing rate and peak inspiratory flow.

The strong modulation of breathing through 2MT exposure in our study aligns with prior work demonstrating a reduction in breathing rate associated with 2MT.22,55 Here intranasal pressure recordings highlight that the reduction in breathing frequency manifests from an increase in longer inhale pauses. Given this effect on pauses is most dramatic during 2MT freezing, an innate fear, in comparison to freezing from conditioned fear during retrieval, we investigated breathing during an additional odor-independent innate fear, foot shock. Using our respiratory space to assess the similarity of breathing during foot shock-induced freezing to current breathing patterns, we found that innate freezing from foot shock has a more similar breathing pattern to fear retrieval freezing than 2MT freezing (Figure 6K). This implies that the odorant quality of 2MT may intensify the effect of inhale pauses during freezing. Noxious vapors evoke a protective reflex of inhale pauses or apnea during the post-inspiratory phase of breathing through the trigeminal system.56 Notably, recent literature supports this potential mechanism for 2MT effects on breathing as it exerts partial action through trigeminal transient receptor potential (TRP) A1 receptors57 and increases c-fos expression in the spinal trigeminal nucleus.55

For some behaviors, similarity in breathing pattern likely reflects similar emotional states while for other behaviors, convergent breathing patterns may arise from disparate neural circuits and emotional states. In assessing the similarity of rearing to breathing patterns in our respiratory space, we found this behavior overlaps with exploration (Figures 6A, 6E, and 6I). Serving as a proof of concept, this investigative rodent behavior unsurprisingly encompasses sniffing, an exploratory breathing pattern. Post-shock arousal breathing also overlaps with exploration and does not significantly overlap with breathing during highly active struggling from restraint or TST (Figures 6B, 6F, and 6J). This result implies that breathing from the struggling states may predominantly reflect the physical activity over the stress response itself, particularly for struggling during TST. Post-shock arousal behavior typically encompasses minor orienting movements and not the magnitude of gross body movements observed in struggling. Our finding that post-shock arousal shares high similarity with exploration, which contains odor investigation, relates to work showing that habituation to odor through multiple presentations is associated with a reduction in high-frequency sniffing to that odor.23 In addition, sniffing frequency may reflect reward anticipation as it remains high prior to reward receipt in a go/no go task, with higher frequency sniffing tracking learning of the reward paired odor.94 Finally, quiescence breathing displays similarities to breathing during sleep (Figures 6D, 6H, and 6L). As quiescence represents an awake inactive behavior without emotional salience, it is noteworthy that quiescence resembles sleep and is not statistically closer to other immobile behaviors evoked by emotionally salient tests (Figure 6H).

This work presents a current respiratory “space” defined by BreathMetics features including spontaneous, stress-induced, fear-induced, and positive-odor-induced behaviors. Through monitoring additional breathing parameters outside the common breathing rate and peak inspiratory flow, we enhance the classification accuracy and separation of behaviors by breathing. Of particular note are the understudied breathing parameters (inhale and exhale pauses, measurements of variability, duty cycle) that vary significantly across emotionally salient behaviors which appear overall the same judging by eye. While in this work, breathing appears to be somewhat continuous, with the overlap of behaviors within clusters rather than distinct clusters for each particular behavior, the analysis provides a future framework for separating breathing patterns and their potential underlying internal state and physiology. For instance, our study highlights the previously unidentified variability in the number and duration of inhale pauses across forms of behavioral immobility. The neural underpinnings of passive defensive responses such as freezing and immobility have been consistently linked to the ventrolateral periaqueductal gray (vPAG).59-61 While passive behaviors studied here display some variability in breathing responses (c.f., retrieval freezing versus TST immobility), it would be of interest to investigate whether neural spiking dynamics across various cell types in the vPAG further separates these forms of breathing into individual clusters. An alternative is that these differences in passive breathing patterns arise from different combinations of upstream activation independent of vPAG activation or via additional physiological measurements of autonomic output. Finally, we suspect that the unique breathing pattern seen
during 2MT freezing represents the concomitant activation of innate freezing pathways and the trigeminal reflex. Probing the neural pathways associated with these breathing patterns will aid in clarifying how these breathing outputs relate to rodent internal state.

**Limitations of the study**
It is technically challenging and time-consuming to obtain a large amount of respiratory data via intranasal pressure recordings from behaving mice. This hinders our capability of applying more advanced machine learning algorithms (e.g., deep learning via convolutional neural networks) that require larger datasets. These models can learn from the rich information embedded in the raw respiration waveform and may enhance the prediction accuracy of behavior.

**STAR METHODS**
Detailed methods are provided in the online version of this paper and include the following:

- **KEY RESOURCES TABLE**
- **RESOURCE AVAILABILITY**
  - Lead contact
  - Materials availability
  - Data and code availability
- **EXPERIMENTAL MODEL AND SUBJECT DETAILS**
  - Animals
- **METHOD DETAILS**
  - Surgical implantation
  - Whole-body plethysmography recordings
  - Intranasal pressure recordings
  - Behavioral tests
  - Behavioral scoring
  - Data Analysis
- **QUANTIFICATION AND STATISTICAL ANALYSIS**

**SUPPLEMENTAL INFORMATION**
Supplemental information can be found online at https://doi.org/10.1016/j.isci.2022.105625.

**ACKNOWLEDGMENTS**
We thank Anamaria Cotelo-Larrea for her assistance in behavior scoring. This work was supported by the National Institutes of Health R01DC006213 to MM, R01NS117061, R01DA049545, and R01DA049449 to MM and DWW, R21DC019193 to JPB, and F31MH124372 and the Hearst Foundation Fellowship to EJ.

**AUTHOR CONTRIBUTIONS**
Conceptualization, EJ and MM; Methodology, all authors; Investigation, EJ, MZ, SER, JPB, MRS, AHM; Formal Analysis, Data Curation, and Visualization, EJ; Writing-Original Draft, EJ and MM; Writing-Review & Editing, WL, LD, and DWW; Supervision, DWW and MM; Funding Acquisition, EJ, JPB, DWW, and MM.

**DECLARATION OF INTERESTS**
The authors declare no competing interests.

**INCLUSION AND DIVERSITY**
We worked to ensure sex balance in the selection of non-human subjects.

Received: May 3, 2022
Revised: October 10, 2022
Accepted: November 16, 2022
Published: December 22, 2022
REFERENCES

1. Tipton, M. J., Harper, A., Paton, J. F. P., and Costello, J. T. (2017). The human ventilatory response to stress: rate or depth? J. Physiol. 595, 5729–5752. https://doi.org/10.1113/JPHYSIOL.2016.030339.

2. Mitchell, G. S., and Johnson, S. M. (2003). Neuroplasticity in respiratory motor control. J. Appl. Physiol. 94, 358–374. https://doi.org/10.1152/japplphysiol.00523.2002.

3. Del Negro, C. A., Funk, G. D., and Feldman, J. L. (2018). Breathing matters. Nat. Rev. Neurosci. 19, 5729–5752. https://doi.org/10.1038/s41583-018-0003-6.

4. Guyenet, P. G., and Bayliss, D. A. (2015). Olfactory inputs modulate respiration-ultrasonic vocalizations-behavior acquisition setup for assessing emotional memory in rats. Front. Behav. Neurosci. 9, 25. https://doi.org/10.3389/fnbeh.2011.00025.

5. Anderson, D. J., and Adolphs, R. (2014). A framework for studying emotions across species. Cell 157, 187–200. https://doi.org/10.1016/j.cell.2014.03.003.

6. Ashhad, S., Kam, K., Del Negro, C. A., and Feldman, J. L. (2022). Breathing rhythm and pattern and their influence on emotion. Annu. Rev. Neurosci. 45, 223–247. https://doi.org/10.1146/annurev-neuro-091012-014424.

7. Welker, W. I. (1964). Analysis of sniffing of the albino rat. Beyond Behav. Neurosci. 13, 251–367. https://doi.org/10.1038/jn.1964.251.

8. Bautista, T. G., Sun, Q. J., and Pilowsky, P. M. (2018). Breathing waveforms using BreathMetrics: a minimally invasive highly precise monitoring system. Meas. Sci. Rev. 18, 573–583. https://doi.org/10.1016/j.jsp.2019.02.009.

9. Mathis, A., Mamicchina, 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. https://doi.org/10.1038/s41593-018-0209-y.

10. Wiltshin, A. B. O., Johnson, M. J., Gurli, G. P., Peterson, R. E., Katon, J. M., Pashkovski, L. S., Abraira, V. E., Adams, R. P., and Datta, S. R. (2015). Mapping sub-second structure in mouse behavior. Neuroimage 88, 1121–1135. https://doi.org/10.1016/j.neuroimage.2015.11.031.

11. Dolensek, N., Gehlach, D. A., Klein, A. S., and Gogolla, N. (2020). Facial expressions of emotion states and their neuronal correlates.
in mice. Science 368, 89–94. https://doi.org/10.1126/science.aaz9468.

37. Abdus-Saboor, I., Fried, N.T., Lay, M., Burdge, J., Swanson, K., Fischer, R., Jones, J., Dong, P., Cai, W., Guo, X., et al. (2019). Development of a mouse pain scale using sub-second behavioral mapping and statistical modeling. Cell Rep. 28, 1623–1634.e4. https://doi.org/10.1016/j.celrep.2019.07.017.

38. Fonseca, A.H., Santana, G.M., Bosque Ortiz, G.M., Bampi, S., and Dietrich, M.O. (2021). Analysis of ultrasonic vocalizations from mice using computer vision and machine learning. Elife 10, e59161. https://doi.org/10.7554/ eLife.59161.

39. Siddiqui, H.U.R., Shahzad, H.F., Saleem, A.A., and Napolitano, A. (2008). RUSBoost: improving classification performance when training data is skewed. In 19th International Conference on Pattern Recognition, pp. 1–4. https://doi.org/10.1109/ICPR.2008.4761297.

40. Seiffert, C., Khoshgoftaar, T.M., Van Hulse, J., and Dudley, S. (2021). Respiration-based non-invasive approach for emotion assessment of Parkinson’s disease using nocturnal breathing signals. Nat. Med. 28, 2207–2215. https://doi.org/10.1038/s41591-022-01932-x.

41. Verhagen, J.V., Wesson, D.W., Netoff, T.I., White, J.A., and Wachowiak, M. (2007). Sniffing controls an adaptive filter of sensory input to the olfactory bulb. Nat. Neurosci. 10, 631–639. https://doi.org/10.1038/nn1892.

42. Shusterman, R., Smear, M.C., Koulakov, A.A., and Rinberg, D. (2011). Precise olfactory responses tile the sniff cycle. Nat. Neurosci. 14, 1039–1044. https://doi.org/10.1038/nn.2877.

43. Isosaka, T., Matsuo, T., Yamaguchi, T., Funabiki, K., Nakashima, S., Kobayakawa, R., and Kobayakawa, K. (2015). Htr2a-Expressing cells in the central amygdala control the hierarchy between innate and learned fear. Cell 163, 1153–1164. https://doi.org/10.1016/j.cell.2015.10.047.

44. Hérent, C., Diem, S., Fortin, G., and Bouvier, J. (2020). Absent phosphorylation of respiratory and locomotor rhythms in running mice. Elife 9, e61919. https://doi.org/10.7554/eLife.61919.

45. Reiser, J., Golden, G.J., Matsumura, K., Smear, M., Rinberg, D., and Gelperin, A. (2014). Comparing thoracic and intra-nasal pressure transients to monitor active odor sampling during odor-guided decision making in the mouse. J. Neurosci. Methods 221, 8–14. https://doi.org/10.1016/j. jneumeth.2013.09.006.

46. Bramble, D.M., and Jenkins, F.A., Jr. (1993). Mammalian locomotor-respiratory integration: implications for diaphragmatic and pulmonary design. Science 262, 235–240. https://doi.org/10.1126/science.2811141.

47. Yang, Y., Yuan, Y., Zhang, G., Wang, H., Chen, Y.C., Liu, Y., Taroll, C.G., Crepeau, D., Bukarty, J., Junna, M.R., et al. (2022). Artificial intelligence-enabled detection and assessment of Parkinson’s disease using nocturnal breathing signals. Nat. Med. 28, 2207–2215. https://doi.org/10.1038/s41591-022-01932-x.

48. Steenland, H.W., and Zhuo, M. (2009). Neck electromyography is an effective measure of fear behavior. J. Neurosci. Methods 177, 355–360. https://doi.org/10.1016/j.jneumeth.2008.10.020.

49. Koba, S., Inoue, R., and Watanabe, T. (2016). Role played by periaqueductal gray neurons in parasympathetically mediated fear bradycardia in conscious rats. Physiol. Rep. 4, e12831. https://doi.org/10.14814/phy2.12831.

50. Walker, P., and Carrive, P. (2003). Role of ventrolateral periaqueductal gray neurons in the behavioral and cardiovascular responses to contextual conditioned fear and poststress recovery. Neuroscience 116, 897–912. https://doi.org/10.1016/S0306-4522(02)00744-3.

51. Vianna, D.M.L., and Carrive, P. (2005). Changes in cutaneous and body temperature during and after conditioned fear to context. Eur. J. Neurosci. 206–212. https://doi.org/10.1038/nature21047.

52. Johnson, S.B., Emmons, E.B., Lingg, R.T., Anderson, R.M., Romig-Martin, S.A., Murphy, D.L., and Tuohimaa, P. (2007). Comparing thoracic and intra-nasal accelerometer data of mice running on the actiGraph: a marker for locomotor activity. Physiol. Rep. 14, 102957. https://doi.org/10.1016/j.autneu.2018.11421.8.2018.

53. Yakumichi, N., Kobayakawa, K., and Kobayakawa, K. (2021). Artificial hibernation/life-protective state induced by thiazoline-related innate fear odors. Commun. Biol. 4, 101. https://doi.org/10.1038/s42003-020-01629-2.

54. Dutschmann, M., Jones, S.E., Subramanian, H.H., Stanic, D., and Bautista, T.G. (2014). The physiological significance of postinspiration in respiratory control. Prog. Brain Res. 212, 113–130. https://doi.org/10.1016/B978-0-444-34888-7.00007-0.

55. Wang, Y., Cao, L., Lee, C.Y., Matsuo, T., Wu, K., Asher, G., Tang, L., Saitoh, T., Russell, J., Kleve-Nebenius, D., et al. (2018). Large-scale forward genetics screening identifies Trp1 as a chemosensor for predator odor-evoked innate fear responses. Nat. Commun. 9, 2041. https://doi.org/10.1038/s41467-018-04324-3.

56. Liu, P., Cao, T., Xu, J., Mao, X., Wang, D., and Li, A. (2020). Plasticity of sniffing pattern and neural activity in the olfactory bulb of behaving mice during odor sampling, anticipation, and reward. Neurosci. Bull. 36, 598–610. https://doi.org/10.1007/s12264-019-00463-9.

57. Fukud, J.P., Krabbe, S., Markovic, M., Courtin, J., Xu, C., Massi, L., Botta, P., Bylund, K., Muller, C., Kovacevic, A., et al. (2017). A competitive inhibitory circuit for selection of active and passive fear responses. Nature 542, 96–100. https://doi.org/10.1038/nature21047.

58. Tovote, P., Espósito, M.S., Botta, P., Chaudun, F., Fukud, J.P., Markovic, M., Wolff, S.B.E., Ramakrishnan, C., Fenno, L., Desseroth, K., et al. (2016). Midbrain circuits for defensive behaviour. Nature 534, 206–212. https://doi.org/10.1038/nature17996.

59. Johnson, S.B., Emmons, E.B., Liu, G.H., Bautista, T.G., LaLumiere, R.T., Narayanan, N.S., Vau, V., and Radley, J.J. (2019). Prefrontal-bad nucleus circuit modulation of a passive coping response set. J. Neurosci. 39, 1405–1419. https://doi.org/10.1523/JNEUROSCI.1421-18.2018.

60. Schreck, M.R., Zhuang, L., Janke, E., Mobley, A.H., Bhattarai, J.P., Gottfried, J.A., Wesson, D.W., and Ma, M. (2022). State-dependent olfactory processing in freely behaving mice. Cell Rep. 38, 110450. https://doi.org/10.1016/j.celrep.2022.110450.
STAR METHODS

KEY RESOURCES TABLE

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---------------------|--------|------------|
| Deposited data      |        |            |
| WBP BreathMetrics data | This paper | https://doi.org/10.17632/h8vzwynysb.1 |
| Intranasal pressure BreathMetrics data | This paper | https://doi.org/10.17632/n29x529y6g.1 |
| Experimental models: Organisms/strains | | |
| Mouse: C57BL/6J | Jackson Laboratory | RRID:IMSR_JAX:000664 |
| Software and algorithms | | |
| MATLAB (R2019b) | Mathworks | https://www.mathworks.com/products/new_products/release2019b.html (RRID:SCR_001622) |
| BreathMetrics | Noto et al., 2018 | https://github.com/zelanolab/breathmetrics |
| Chronux Version 2.11 | Open source originally through the Mitra Lab in Cold Spring Harbor | http://chronux.org/ (RRID:SCR_005547) |
| Respiratory feature extraction, clustering, and classification of WBP and intranasal pressure data | This paper | https://doi.org/10.5281/zenodo.7261849 |
| Synapse | Tucker-Davis Technologies | https://www.tdt.com/support/downloads/ |
| Synapse Lite | Tucker-Davis Technologies | https://www.tdt.com/support/downloads/ |
| AxoScope within pCLAMP | Molecular Devices | https://www.moleculardevices.com/products/axon-patch-clamp-system/acquisition-and-analysis-software/pclamp-software-suite |

RESOURCE AVAILABILITY

Lead contact
Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Minghong Ma (minghong@pennmedicine.upenn.edu).

Materials availability
This study did not generate new unique reagents.

Data and code availability
- BreathMetrics summary statistics of whole-body plethysmography and intranasal pressure recordings during behavior have been deposited at Mendeley data and are publicly available as of the date of publication. The DOIs are listed in the key resources table.
- All original code has been deposited at Zenodo and is publicly available as of the date of publication. The DOI is in the key resources table.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Animals
All experiments were conducted in C57BL/6J mice from Jackson Laboratory (RRID:IMSR_JAX:000664). Both male and female mice (2–6 months) were used and evaluated across all behavior conditions. Since no sex differences were evident, data from male and female mice were pooled together. Mice were group-housed prior to surgery and then were singly housed following surgical implantation. Mice were maintained on a 12–12 hr light-dark cycle with ad libitum food and water in a temperature and...
humidity-controlled animal facility. All experimental procedures conform to the relevant institutional and national guidelines and were approved by the Institutional Animal Care and Use Committees of the University of Pennsylvania and the University of Florida.

METHOD DETAILS

Surgical implantation

Anesthesia was induced by exposing mice to 3% isoflurane in oxygen. Mice were then moved to a stereotaxic system (Model 940, David Kopf Instruments), and anesthesia was maintained at a reduced level of 1.5–2% isoflurane for the remainder of the surgery. A temperature control system (TC-1000, CWE Inc.) with a heating pad regulated mouse body temperature at 37°C. A local anesthetic, bupivacaine (2mg/kg), was administered at the incision site along the nasal bone. A hole was drilled in the nasal bone 7.0 mm caudal from the tip of the nose on either the right or left nasal plate. The olfactory epithelium was punctured to allow for the cannula (C311G/SPC, with corresponding dummy cannula C311DC, P1 Technologies) to have access to the nasal cavity, and two bilateral skull screws (Fine Science Tools) were inserted for stabilization. The implant was then sealed with 3M Vetbond Adhesive (World Precision Instruments) and dental cement. For sleep and quiescence classification, an additional 8-pin board (Pinnacle Technology, 8415-SM) was mounted on the skull for electroencephalogram (EEG) and electromyogram (EMG) recordings. Stainless steel wires connected the board to skull screws above the hippocampus (~1.5 mm AP, 1.5 mm ML) for EEG. Then two stainless steel wires were implanted into the neck muscle for EMG. Reference and grounding screws were placed bilaterally above the cerebellum. Mice recovered for at least 7 days prior to any respiratory recordings.

Whole-body plethysmography recordings

Animals were habituated to the plethysmograph chamber with no stimulation 24 h before olfactory fear conditioning. A plethysmograph chamber (Data Sciences International) was adapted for infusion of a neutral odor, isopentyl acetate (1 torr), through a custom air-dilution olfactometer at a flow rate of 1 L/min (20 s) followed by a foot shock (0.5 mA for 1 s). Each fear conditioning session consisted of 10 trials of odor pairing with foot shock and 3 min inter-trial intervals. Respiration was detected using a Data Sciences flow transducer and digitized (0.1–20 Hz) at 300 Hz in Synapse Lite (Tucker Davis Technologies) following a gain amplification of 100X (Cygnus Technology Inc). Positive pressure of room air was applied to the chamber using a stable-output air pump (Tetra Whisper). Animal behaviors were recorded through the entire task with two cameras (Microsoft) from both sides of the plethysmograph chamber.

Intranasal pressure recordings

Respiration recordings require the nasal cannula be attached to a well-sealed tube (P1 Technologies, C315CT) which connects to a pressure sensor (Honeywell, CPXL04GF). Respiration was amplified through DP-301 single-channel differential amplifier (Warner Instruments) and filtered from 0.1 Hz to 100 Hz. For sleep classification, EEG and EMG recordings were pre-amplified through a four-channel mouse preamplifier (Pinnacle Technology) prior to amplification through DP-304 Differential amplifier (Warner Instruments). EMG was filtered between 0.1 Hz and 100 Hz. Respiration, and when applicable, EEG, and EMG were simultaneously acquired at 1017 Hz through RZ5P Processor (Tucker-Davis Technologies). Signals and behavior videos were visualized in Synapse software. An additional data acquisition system was used in earlier experiments; pre-amplified respiration as previously stated was digitized at 2000 Hz through Digidata1320A (Axon Instruments) and visualized in AxoScope software. Sleep was characterized by high-amplitude, low-frequency EEG oscillations in the delta range (0.5–4 Hz) and low EMG activity as previously described.62

Behavioral tests

Odor exposures were conducted in a clean, novel cage. Either peanut oil or diluted 2-methyl-2-thiazoline (1:250 in neutral mineral oil, Sigma Aldrich CAS 2346-00-1) was pipetted into a small petri dish in a randomized corner of the cage. Odor exposure occurred for 5–10 min before petri dish removal. Tail suspension test (TST) consisted of placing a 0.5 inch hollow tube over the base of the mouse’s tail to prevent tail climbing during suspension. A 15 cm piece of tape was then attached approximately 3 mm from the tip of the tail, and the mouse was suspended approximately 18 inch in the air for a total of 6 min. Restraint stress consisted of placing each mouse in a custom transparent tube (1.5 in diameter) for 30 min. This tube has a slit cut down the midline of the tube to allow for tethering of respiration tubing. Auditory fear conditioning
consisted of 4 conditioning trials where a 10 s 5kHz tone was followed by a 1s 0.5 mA shock (Med Associates chamber) at randomized intervals between 2-3 min. Auditory fear retrieval was performed 24 h following fear conditioning. Mice were placed in a novel environment and re-exposed to the tone for 4 trials at randomized intervals between 2-3 min. Sleep recordings were performed in the home cage in the morning.

Behavioral scoring
Behavior was manually scored from video. For WBP, investigation was characterized by whisking and exploratory head movements. Freezing consisted of rigid immobility while quiescence consisted of relaxed, hunched immobility. Post shock arousal consisted of the 2 s following foot shock. Intranasal pressure, odor investigation was defined as the mouse approaching a 2 inch zone from the petri dish containing peanut oil. Grooming and rearing bouts were isolated from baseline and recovery periods following peanut oil exposure. NREM sleep and quiescence were classified using EEG and EMG signal (see data analysis for extended details on sleep classification). These states were further verified with video referencing. For both TST and restraint, full body movements were scored as struggling, and no movement, except for breathing motions, was scored as immobility. Similarly, in auditory fear conditioning, fear retrieval, and 2MT exposure, freezing was scored as the complete absence of movement except for breathing. Locomotion during fear retrieval and 2MT exposure was scored as active phase in these conditions. Post shock arousal was scored as the 4 s following shock offset during fear conditioning. For clarity, all behavior bouts included in the respiratory analysis were 2s in duration or longer.

Data Analysis
Feature extraction
Acquired data were analyzed using BreathMetrics toolbox and custom MATLAB (MathWorks; RRID:SCR_001622) scripts. For WBP, respiration was smoothed using MATLAB’s built in smooth function with 15 datapoint bins. Intranasal pressure did not require smoothing given the accurate labeling of raw respiration by BreathMetrics. Next, respiration was baseline corrected and z-scored using BreathMetrics, and then secondary respiratory features were extracted for each clearly defined bout of behavior from that recording. Duration of the breathing bout analyzed reflected the natural duration of the behavior bout from video scoring. Importantly, for WBP, outward deflection represents inhalation whereas in intranasal pressure recordings, outward deflection represents exhalation. In BreathMetrics airflow analysis, outward deflection represents inhalation. Thus, for secondary feature extraction of intranasal pressure which has the opposite orientation of BreathMetrics, inspiratory parameters extracted must be appropriately renamed as expiratory parameters and vice versa. All respiratory features presented for intranasal pressure in this manuscript reflect the accurate orientation of intranasal pressure recordings in our setup, not the inverse orientation by BreathMetrics. Technical limitations of intranasal cannula implants and study design precluded each mouse have respiratory data from all behavioral states. In the following analysis, each mouse typically contributed respiratory data to at least 4 of the 11 behavioral states.

Sleep classification was performed using a combination of EEG and EMG signals. EEG spectrograms of the 0–15 Hz range were determined using multitaper methods from the Chronux toolbox with a moving window of 0.5 s. First NREM sleep was characterized by high-amplitude, low-frequency EEG oscillations in the delta range (0.5–4 Hz) and low EMG activity, whereas REM sleep was defined by EEG oscillations in the theta range (6–8 Hz) and even further reduced EMG. Wake was characterized by low-amplitude, fast EEG oscillations and high EMG activity. Quiescence was then defined during wake by thresholding the EMG for periods with no movement and these periods in the home cage typically occurred prior to sleep onset.

Clustering
Secondary respiratory features from BreathMetrics were compiled across behaviors, and individual parameters were z-scored. Principal component analysis was then conducted. To increase the signal to noise ratio, factor analysis and iterative PCA were performed to reduce the total number of respiratory features included in the clustering analysis. Volume estimations, which are calculated from breath duration and peak amplitude, were omitted. Peak expiratory flow and CV inhale duty cycle were also omitted due to correlations with additional respiratory features included in the analysis. K-means clustering was then conducted on the respiratory data in PC space. An elbow plot of summed variation determined the optimal number of clusters to fit to the data.
**RUSBoost classification**

For behavior classification using WBP or intranasal pressure recordings, the random undersampling boost (RUSBoost) algorithm was used. Training data consisted of 80% of the overall data, equally distributed across behaviors and individual mice. The remaining 20% of overall data served as test cases. Considering fewer grooming bouts were obtained using WBP, longer bouts, >20 s, were segmented into multiple bouts (each no shorter than 10 s) to increase testing data available. Each model was hypertuned on training data using 10-fold cross validation in MATLAB’s classification learner app. The hypertuning method consisted of a Bayesian optimization approach through 50 iterations using an acquisition function of expected improvement per second plus. Hyperparameters searched were: 10–500 learners; 0.001–1 learning rate; 1 to (n-1) max number of splits per learner where n is the number of data points in the training set. 10 iterations were run for each model, randomizing the training and testing dataset used. Shuffled RUSBoost performance for WBP and intranasal pressure data was evaluated by testing the identical 10 iterations of data while shuffling labels within the training dataset. Euclidean distances were calculated to show similarity of novel behaviors to existing clusters.

**QUANTIFICATION AND STATISTICAL ANALYSIS**

Statistical tests were performed in GraphPad Prism. Each figure legend contains the corresponding statistical tests implemented within each panel as well as the sample identity (e.g., bouts, replicates) and sample size. All data are displayed as mean ± SEM. Normality assumptions were determined via Shapiro-Wilk test and through evaluating Normal QQ plots. Kruskal-Wallis test was employed followed by post-hoc Dunn’s multiple comparisons test for non-parametric conditions. For paired samples, Wilcoxon tests were used in non-parametric comparisons while paired t-tests were used for parametric comparisons. Mann-Whitney tests were used in non-parametric conditions in addition to parametric two sample t-tests for statistical testing from independent samples. *p < 0.05; **p < 0.01, ***p < 0.001, ****p < 0.0001. Specifically in Figure S6, #p < 0.0001.