

# Delay and trace fear conditioning in C57BL/6 and DBA/2 mice: issues of measurement and performance

Megan E. Tipps,<sup>1</sup> Jonathan D. Raybuck,<sup>1</sup> Kari J. Buck,<sup>1,2</sup> and K. Matthew Lattal<sup>1</sup>

<sup>1</sup>Department of Behavioral Neuroscience, Oregon Health & Science University, Portland, Oregon 97239, USA; <sup>2</sup>Portland Alcohol Research Center, Portland VA Medical Center, Portland, Oregon 97239, USA

Strain comparison studies have been critical to the identification of novel genetic and molecular mechanisms in learning and memory. However, even within a single learning paradigm, the behavioral data for the same strain can vary greatly, making it difficult to form meaningful conclusions at both the behavioral and cellular level. In fear conditioning, there is a high level of variability across reports, especially regarding responses to the conditioned stimulus (CS). Here, we compare C57BL/6 and DBA/2 mice using delay fear conditioning, trace fear conditioning, and a nonassociative condition. Our data highlight both the significant strain differences apparent in these fear conditioning paradigms and the significant differences in conditioning type within each strain. We then compare our data to an extensive literature review of delay and trace fear conditioning in these two strains. Finally, we apply a number of commonly used baseline normalization approaches to compare how they alter the reported differences. Our findings highlight three major sources of variability in the fear conditioning literature: CS duration, number of CS presentations, and data normalization to baseline measures.

[Supplemental material is available for this article.]

For decades, inbred mouse strains have been used to probe the genetic and molecular underpinnings of behavior. By comparing genetically distinct lines, researchers have been able to identify chromosomal regions and individual genes that contribute to numerous behaviors, including learning and memory (Wehner et al. 1997; Wilson et al. 2011; Parker et al. 2012). Molecular pathways and the importance of specific brain regions for these behaviors have also been identified (Wehner et al. 1990; Lu and Wehner 1997; Nguyen and Gerlai 2002). Thus, the results from inbred strain studies are often used as the basis for explorations into the specific neural mechanisms that drive learning and memory.

The strain comparison approach has been widely applied to fear conditioning, a learning paradigm in which a neutral conditioned stimulus (CS) is paired with an aversive unconditioned stimulus (US) to generate a conditioned fear response to both the CS and the context in which the US was presented (Blanchard and Blanchard 1969; Fanselow 1980). The C57BL/6 (B6) and DBA/2 (D2) strains, which show significant differences in multiple learning paradigms (Paylor et al. 1996; Dellu et al. 2000; Schimanski and Nguyen 2004; Restivo et al. 2006; Graybeal et al. 2014), have been used by several groups to identify genetic and molecular contributions to fear conditioned learning, including long-standing hypotheses regarding the role of the hippocampus and amygdala (Schimanski and Nguyen 2004; Yang et al. 2008; Andre et al. 2012). However, the behavioral findings for fear conditioning in these strains are highly variable across labs, particularly with regard to differences in CS responses (e.g., Ammassari-Teule et al. 2001; Balogh et al. 2002; Gale et al. 2009). The lack of consensus regarding strain differences at the behavioral level makes drawing conclusions about the underlying neural mechanisms difficult.

Much of the variability between labs may be due to methodological differences in how these behaviors are established, tested, and analyzed. Although most would agree that changes to the training or testing parameters affect the outcome of any behavioral assay, little has been done to establish field-wide comparisons of the reported variations in fear conditioning procedures or the experimental design differences that influence them. This confounds future work, as many existing fear conditioning studies cannot be directly compared, making it difficult to draw meaningful conclusions about either the implications of the behavioral differences in learning and memory or the underlying neural mechanisms. Thus, there is a need to assess both the work done in these strains to date and to address the primary sources of variation in these comparison studies.

Here, we compare B6 and D2 mice using two forms of fear conditioning: delay fear conditioning and trace fear conditioning. The CS associations in these paradigms are thought to rely on distinct neural systems (Raybuck and Lattal 2011), allowing for comparisons across multiple learning pathways using a common endpoint. We compare the results for contextual and cued conditioning across strains and training paradigms and look at the data relative to a nonassociative training condition. In addition, we compare the utility of alternative behavioral measures to the standard measure of freezing as an index of learning. Finally, we use our data to compare the effects of several commonly used baseline normalization approaches. We then discuss our findings in the context of a comprehensive literature review of the fear conditioning data available for these strains. We highlight numerous sources of variability within the existing literature and propose approaches to address these sources of variation in future work.

© 2014 Tipps et al. This article is distributed exclusively by Cold Spring Harbor Laboratory Press for the first 12 months after the full-issue publication date (see <http://learnmem.cshlp.org/site/misc/terms.xhtml>). After 12 months, it is available under a Creative Commons License (Attribution-Non-Commercial 4.0 International), as described at <http://creativecommons.org/licenses/by-nc/4.0/>.

Corresponding author: [metipps@gmail.com](mailto:metipps@gmail.com)

Article is online at <http://www.learnmem.org/cgi/doi/10.1101/lm.035261.114>.

## Results

Data were collected for freezing, grooming, and rearing in B6 and D2 mice following both delay and trace training. The training period, context test, and cue test were all assessed. Average minute-by-minute data for each of these measures can be found in the Supplemental Material.

### Training

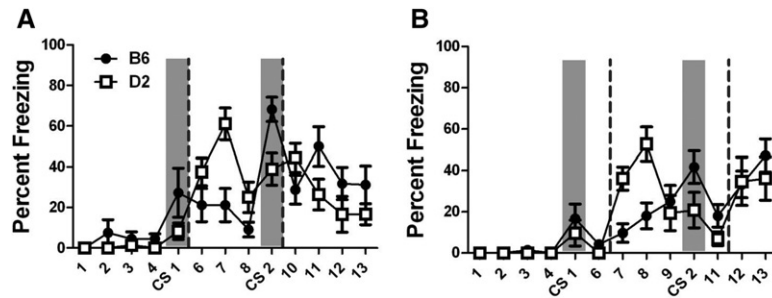
B6 and D2 mice were trained with two CS-US pairings using either delay (Fig. 1A) or trace (Fig. 1B) fear conditioning. The chambers, CS, and US were identical across these paradigms, with only the timing between the US and CS differing. As shown in Figure 1, both strains showed low novel context freezing (bins 1–4) and a mild freezing response to the first CS presentation (CS 1). A repeated measures ANOVA (strain  $\times$  training  $\times$  CS) found a significant effect of strain ( $F_{(1,44)} = 9.091$ ,  $P = 0.004$ ) and training ( $F_{(1,44)} = 4.430$ ,  $P = 0.041$ ). There was also a significant effect of CS ( $F_{(1,44)} = 49.212$ ,  $P < 0.001$ ) with a significant CS  $\times$  training interaction ( $F_{(1,44)} = 5.748$ ,  $P = 0.021$ ). For both strains, freezing to the second CS was significantly higher than freezing to the first CS in the delay paradigm (B6,  $P = 0.001$ ; D2,  $P = 0.001$ ), but in the trace paradigm, this was only significant for B6 mice (B6,  $P = 0.022$ ; D2,  $P = 0.28$ ). B6 mice froze significantly more to the second CS presentation than D2 mice in delay training with a trend for the same difference in the trace training (Delay,  $P = 0.004$ ; Trace,  $P = 0.075$ ). Finally, although there were no differences in responses to the first CS presentation between delay- and trace-trained mice of either strain (B6,  $P = 0.519$ ; D2,  $P = 0.846$ ), freezing to the second CS was higher in delay-trained B6 mice than in trace-trained B6 mice ( $P = 0.008$ ). There was no significant effect of training type in D2 mice, but there was a trend for higher responses in delay-trained mice than in trace-trained mice ( $P = 0.11$ ).

An interesting pattern of freezing also emerged during the inter-CS period for both delay and trace training (bins 6–8 for delay and 7–9 for trace). Following the first US presentation (dashed lines in Fig. 1), D2 mice showed higher levels of freezing than B6 mice (Delay,  $P = 0.001$ ; Trace,  $P = 0.014$ ), but the same difference is not observed following the second US presentation. The basis for this early increased freezing in D2 mice is unclear, but has been observed by other groups (Gerlai 1998; Nie and Abel 2001; Fitch et al. 2002; Wilson et al. 2011), and is often used as evidence that reactivity to the US is not responsible for low D2 freezing during the context and CS tests. In agreement with this, we found a similar footshock threshold between these strains (B6,  $0.163 \pm 0.008$  mA; D2,  $0.169 \pm 0.009$  mA;  $P = 0.618$ ), which is consistent with previous reports (Lu and Wehner 1997; Liu et al. 2003).

These data illustrate three important points: (1) the initial CS response (CS 1) does not differ between B6 and D2 mice; (2) both strains respond after the US with increased freezing behavior; and (3) even at the early stages, the conditioned freezing response evoked by the CS following pairing with the US is stronger for delay training than for trace training in both B6 and D2 mice.

### Context test

Twenty-four hours after training, the mice were returned to the training room and placed into the training chambers for a 12-min context test. As shown in Figure 2, A and B (Supplemental



**Figure 1.** Delay and trace training. Percent freezing (mean  $\pm$  SEM) during training for the (A) delay and (B) trace fear conditioning paradigms. The gray bars represent the timing of the 30-sec CS, and the dashed lines represent the timing of the 2-sec US. Data are shown in 30-sec bins to match the CS and trace interval durations. Significant fear differences between the strains and training types are described in the text.

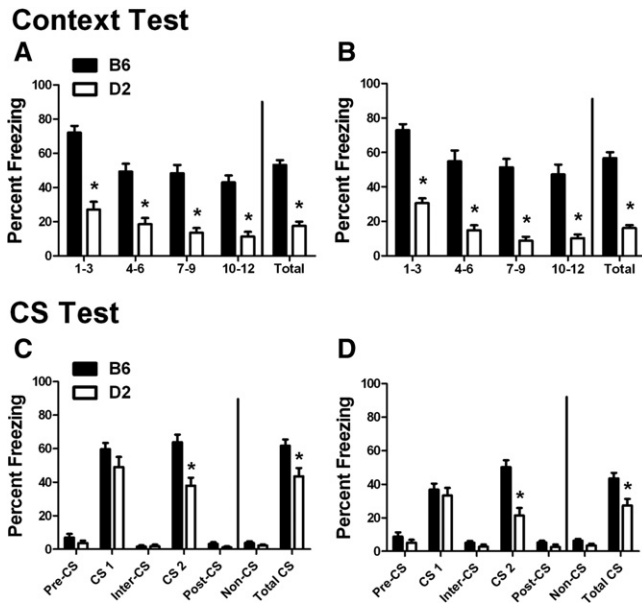
Fig. 1), we observed a robust and consistent difference between these strains. A repeated measures ANOVA (strain  $\times$  training  $\times$  time bin) found a significant effect of strain ( $F_{(1,44)} = 215.48$ ,  $P < 0.001$ ) and time bin ( $F_{(3,132)} = 37.729$ ,  $P < 0.001$ ). For both delay- and trace-trained mice, D2 mice froze significantly less than B6 mice ( $P < 0.001$  for all time bins). Both B6 and D2 mice froze more during the first time bin (min 1–3) than the fourth (min 10–12) under both training conditions, illustrating a general decrease in freezing as the test progressed ( $P < 0.01$  for all comparisons) (see Supplemental Fig. 1 for a minute-by-minute analysis). Interestingly, there was no effect of training, suggesting that contextual learning for either strain was not affected by training type.

### CS test

Twenty-four hours after the context test (48 h after training), the mice were moved to a second testing room and placed in novel chambers where they were exposed to a 15-min CS test that included two 3-min presentations of the CS (Fig. 2C,D; Supplemental Fig. 1). Our analysis of the two CS presentations (strain  $\times$  training  $\times$  CS) found a significant effect of strain ( $F_{(1,44)} = 20.593$ ,  $P < 0.001$ ) and training ( $F_{(1,44)} = 20.593$ ,  $P < 0.001$ ) with a significant interaction between CS and strain ( $F_{(1,44)} = 24.662$ ,  $P < 0.001$ ). Similar to our results for the context test, D2 mice froze significantly less to the CS when the two CS periods were combined (Total CS freezing) than B6 mice under both training paradigms (Delay,  $P = 0.004$ ; Trace,  $P = 0.003$ ). Interestingly, when the CS presentations were analyzed separately, this strain difference was only apparent during the second CS presentation ( $P < 0.001$  for both Delay and Trace), but was robust enough to make the total CS freezing levels significantly different. There were no significant strain differences during the non-CS (Pre-CS, Inter-CS, Post-CS) periods.

### Alternative behaviors

A number of groups have suggested that the differences in freezing between strains may be reflective, not of learning differences, but of differences in the behavioral expression of learning (Griebel et al. 1997; Stiedl et al. 1999). Specific to the B6/D2 comparisons, while several groups see deficits in passive avoidance measures for D2 mice (such as freezing in fear conditioning), these mice actually perform better on active avoidance tasks (Weinberger et al. 1992; Heyser et al. 1999), suggesting that freezing may not be the best measure of their associative abilities. The various behaviors, including freezing, observed during different phases of fear conditioning have been shown to have complex correlations, suggesting that some of the behaviors associated with fear may



**Figure 2.** Context and CS tests. Percent freezing (mean ± SEM) during the context and CS tests. (A,B) Freezing to the context is presented in 3-min bins and as the total for the 12-min test in (A) delay- and (B) trace-trained mice. (C,D) Freezing to the CS is presented in 3-min bins corresponding to the individual periods of the test and as the total freezing for the non-CS (pre-CS, inter-CS, and post-CS) and CS (CS 1 and CS 2) periods for (C) delay- and (D) trace-trained mice. (\*)  $P < 0.05$  for B6 versus D2.

be independent of freezing under certain conditions (Fitch et al. 2002; Holahan and White 2002). Thus, quantifying other behavioral elements could provide a better description of the overall behavioral responses to fear conditioning across differently behaving strains. A handful of alternative behaviors have been investigated, including grooming, rearing, and boli (e.g., Gerlai 1998; Fitch et al. 2002; Wilson et al. 2011; Choy et al. 2012). We chose the two most commonly observed behaviors, grooming and rearing, and compared them across our strains in delay and trace fear conditioning.

**Grooming**

Grooming is a repetitive, highly stereotyped behavior observed across a number of species. For mice, this behavior has been linked to changes in emotional state, particularly measures of anxiety and stress (Kalueff and Tuohimaa 2005b; Estanislau et al. 2013; O’Leary et al. 2013), both of which are related to fear conditioning responses (Ponder et al. 2007; Ryoke et al. 2014). Thus, grooming behavior could be a pathway for fear learning expression in high grooming strains, such as D2.

Prior to the first CS presentation during training (Supplemental Fig. 2A,B), we found no differences in grooming behavior between strains (Delay,  $P = 0.102$ ; Trace,  $P = 0.09$ ). In the post-CS period, D2 mice groomed significantly more than B6 mice in the delay condition ( $P = 0.001$ ), but no strain differences were observed for the trace condition ( $P = 0.106$ ). During the context test (Supplemental Fig. 2C,D), a repeated measures ANOVA (strain × training × time) found a significant effect of strain ( $F_{(1,44)} = 92.592$ ,  $P < 0.001$ ) and time ( $F_{(11,484)} = 4.752$ ,  $P < 0.001$ ) with a significant time × strain interaction ( $F_{(11,484)} = 1.903$ ,  $P = 0.037$ ). As shown in Figure 3, A and B, grooming behavior in B6 mice is significantly lower than that of D2 mice during the

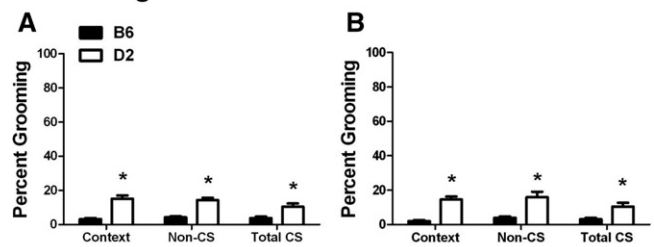
context test and the CS and non-CS periods of the CS test ( $P < 0.005$  for both Delay and Trace). Similar to freezing, there was no effect of training type on grooming during the context test for either strain. However, unlike the freezing measure, there was also no effect of training type on grooming responses to the CS in D2 mice. A repeated measures ANOVA (strain × training × CS) found a significant effect of strain ( $F_{(1,44)} = 21.238$ ,  $P < 0.001$ ) with D2 mice grooming more than B6 mice during both CS presentations for both training paradigms ( $P < 0.05$  for all comparisons) (Supplemental Fig. 2).

**Rearing**

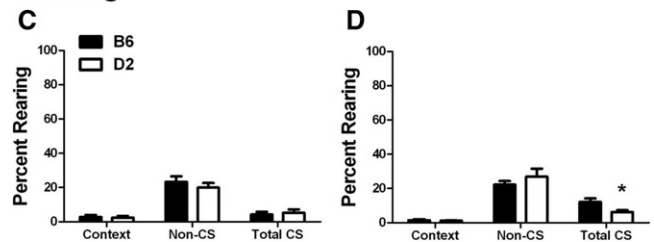
Rearing behavior, in which both front feet are lifted off the ground and the body is extended, has been categorized as both a general exploratory behavior, which should be decreased in the presence of a fearful context or stimulus, and an escape/avoidance behavior, which would be increased under the same conditions (Perez-Saad and Bures 1983; Crusio 2001). Strain differences in rearing have been observed, but these differences are often dependent upon the context in which rearing is assessed (e.g., homecage rearing vs. open field rearing) (Tang et al. 2002; Bothe et al. 2005). Given that D2 mice show enhanced performance on active avoidance tasks compared to B6 (Weinberger et al. 1992; Heyser et al. 1999), it is possible that rearing could be part of their behavioral response to fearful stimuli during fear conditioning tests.

There were no differences in rearing behavior between B6 and D2 mice during the pre-CS or post-CS period of either trace or delay training (Supplemental Fig. 3A,B). For the context test (Fig. 3C,D; Supplemental Fig. 3C,D), we did not detect an effect of strain, training, or time. However, in the CS test, a repeated measures ANOVA (strain × training × CS) found a significant effect of training ( $F_{(1,44)} = 6.854$ ,  $P = 0.012$ ) and a training × strain interaction ( $F_{(1,44)} = 4.242$ ,  $P = 0.045$ ). Although there was no effect of CS alone, there were significant interactions for CS × strain

**Grooming**



**Rearing**



**Figure 3.** Grooming and rearing. Percent grooming and rearing (mean ± SEM) during the context and CS test. (A,B) Grooming behavior for the 12-min context test and the non-CS and CS periods of the CS test for (A) delay- and (B) trace-trained mice. (C,D) Rearing behavior for the 12-min context test and the non-CS and CS periods of the CS test for (C) delay- and (D) trace-trained mice. An extended analysis for both behaviors can be found in Supplemental Figures 2 and 3. (\*)  $P < 0.05$  for B6 versus D2.

( $F_{(1,44)} = 8.246$ ,  $P = 0.006$ ) and CS  $\times$  strain  $\times$  training ( $F_{(1,44)} = 4.465$ ,  $P = 0.04$ ). These effects were driven by the trace-trained group, in which B6 mice reared significantly more than D2 mice during the first CS presentation ( $P = 0.003$ ) (Supplemental Fig. 3E,F), and this difference was robust enough to make the combined CS rearing higher for trace-trained B6 mice ( $P = 0.024$ ) (Fig. 3D). No other strain or training differences were observed.

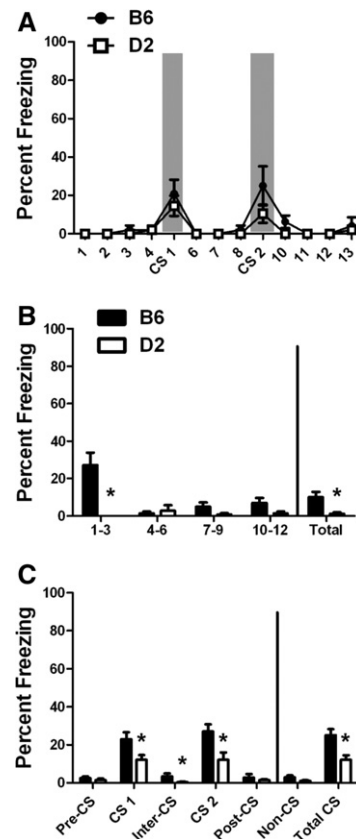
### Nonassociative effects

A major issue with any strain comparison study is how to assess and account for inherent differences in “baseline” behavior that might affect the interpretation of the outcome. B6 and D2 strains differ in their locomotor and exploratory behaviors (Trullas and Skolnick 1993; Podhorna and Brown 2002) as well as general auditory startle responses (Paylor and Crawley 1997), which could lead to biased responses following fear conditioning to an auditory cue. Over the years, a number of control measures have been proposed to account for these differences. Ideally, the CS and US would be presented in a completely random fashion, with the probability of the US in the presence and the absence of the CS being the same (e.g., Rescorla 1967). However, this is challenging in studies of fear conditioning in which no more than a few US presentations are given. Although not ideal, the most commonly used method is to present “unpaired” sets of the CS and US in which the CS–US interval is very long or the US is given in the absence of the CS (Paylor et al. 1994; Gerlai 1998; Hwang et al. 2010). The duration of this interval is critical to these assumptions, given that trace fear conditioning between the CS and US can form across a wide range of inter-stimulus intervals (Misane et al. 2005). Alternatively, under these conditions the animal could learn that the CS signals the explicit absence of a US. Thus, it is difficult to argue that absolutely no excitatory or inhibitory association is formed between the CS and US when the US and novel CS both occur in a fixed environment. In addition, this approach cannot be used to assess contextual association strength, as the context and US are still paired.

Rather than attempt to break apart the associative processes, we used a conditioning procedure designed to determine whether simply removing the US from the training process entirely would result in behaviors that were significantly different than those observed in either the delay- or trace-trained animals. We used a nonassociative method in which animals were exposed to the training and testing process in the absence of a US (no footshock) to assess the baseline freezing to the training context and CS used in our paradigm (Fig. 4; Supplemental Fig. 4). The grooming and rearing behavior observed for this condition can be found in Supplemental Figure 5.

During No-US training (Fig. 4A), both B6 and D2 mice had freezing responses to the CS with no significant differences between strains, suggesting that baseline differences in CS response are not responsible for later behavioral differences. Interestingly, there were also no differences in freezing between the period prior to the first CS exposure (bins 1–4) and the period following the last CS exposure (bins 10–13), suggesting that repeated exposure to a startling CS did not alter post-CS freezing levels. This contrasts with the significant difference observed between the pretraining and post-training freezing levels observed in both the delay- (B6,  $P < 0.001$ ; D2,  $P = 0.007$ ) and trace-trained mice (B6,  $P < 0.001$ ; D2,  $P = 0.001$ ) (see Fig. 1).

In the context test (Fig. 4B), there was a significant effect of strain ( $F_{(1,14)} = 9.966$ ,  $P = 0.007$ ) and time ( $F_{(3,42)} = 11.478$ ,  $P < 0.001$ ), with a significant strain  $\times$  time interaction ( $F_{(3,42)} = 15.785$ ,  $P < 0.001$ ). While freezing in both strains was much lower than that observed for either the delay or trace groups, B6 mice showed a moderate level of freezing during the first time bin, while



**Figure 4.** No-US training, context test, and CS test. Percent freezing (mean  $\pm$  SEM) for mice trained in the absence of a US. (A) Training: freezing during the No-US training procedure is shown in 30-sec bins to match the CS duration. The CS was presented twice with the same timing as the delay and trace training paradigms (gray bars). (B) Context test: freezing is presented in 3-min bins along with the total freezing for the entire 12-min period. (C) CS test: data are shown in 3-min bins to match the intervals of the CS test. Total freezing during the non-CS and CS intervals is also shown. (\*)  $P < 0.05$  for B6 versus D2.

D2 mice showed little to no freezing across the entire test, resulting in significantly more freezing in B6 mice ( $P = 0.007$ ).

In the CS test (Fig. 4C), we observed a low level of freezing during the CS presentations in both strains. There was a significant effect of strain ( $F_{(1,14)} = 11.103$ ,  $P = 0.005$ ), with B6 mice freezing more than D2 mice during both CS presentations (CS 1,  $P = 0.02$ ; CS 2,  $P = 0.009$ ; Total,  $P = 0.004$ ). Interestingly, unlike the delay and trace conditions (see Fig. 2), D2 mice showed no change in freezing between the first and second CS presentation, suggesting that acclimation to the CS itself is not responsible for changes in freezing during the second CS in delay and trace fear conditioning.

The presence of mild freezing following training to the CS alone suggests that, even in the absence of a US, the CS used here (85-dB white noise) may be mildly aversive or startling on its own. Differences in sensitivity to the CS itself could account for the freezing differences observed here, given that B6 mice show increased acoustic startle responses compared to D2 mice (Paylor and Crawley 1997; Liu et al. 2003). Similar strain differences in mice trained with the CS alone have been reported by some groups (Stiedl et al. 1999; Balogh et al. 2002), while other groups using a CS only condition have reported no freezing responses to either the context or the tone (Amassari-Teule et al. 2001; Bolivar et al. 2001). However, the studies in which no freezing

was observed used very short CS presentations (<10 sec), making them difficult to compare to the current study. It is possible that lowering the intensity or duration of the CS could remove these effects in B6 mice; however, given the propensity of D2 mice to show hearing impairments (Zheng et al. 1999), the louder, more complex white noise is preferred for this strain. Additionally, Bolivar et al. (2001) used a louder CS than the one used here and saw no freezing, suggesting that the CS dB level alone is not the cause of these effects. Studies have also found that B6 mice show equivalent locomotor profiles during the CS test if they were trained with a CS-US pairing or the US alone (Gerlai 1998; Bolivar et al. 2001), suggesting that this strain may show generalized startle responses under a variety of conditions that are not indicative of CS-US associative learning.

### Training comparisons

Although low levels of freezing to both the context and cue were observed in the nonassociative condition, training with either delay or trace significantly increased freezing in both tests for both strains (Supplemental Fig. 5). To analyze these differences, we compared total freezing, grooming, and rearing to both the context and CS across our three training paradigms (Fig. 5).

#### Freezing

For the total context freezing (Fig. 5A), there was a significant effect of strain ( $F_{(1,58)} = 179.932, P < 0.001$ ) and training ( $F_{(2,58)} = 79.927, P < 0.001$ ) with a significant strain  $\times$  training interaction ( $F_{(2,58)} = 19.122, P < 0.001$ ). In all training types, B6 mice froze

more than D2 mice (No-US,  $P = 0.007$ ; Delay,  $P < 0.001$ ; Trace,  $P < 0.001$ ), and both strains froze more following delay and trace training compared to the No-US group ( $P < 0.001$  for all comparisons).

For CS freezing, there was a significant effect of strain ( $F_{(1,58)} = 26.416, P < 0.001$ ) and training ( $F_{(2,58)} = 39.021, P < 0.001$ ). Similar to the total context freezing, B6 mice froze significantly more than D2 mice in all three paradigms (No-US,  $P = 0.004$ ; Delay,  $P = 0.005$ ; Trace,  $P = 0.003$ ), and for both strains, freezing for the delay- and trace-trained groups was significantly higher than that for the No-US group ( $P < 0.01$  for all comparisons). Thus, we can conclude that both the delay and trace training paradigms result in a freezing response above and beyond that which occurs after simple exposure to the CS in B6 and D2 mice. The further increase after delay compared to trace conditioning suggests that temporal pairing of the CS and US contributes to the freezing response.

#### Grooming

For total context grooming (Fig. 5B), we found a significant effect of strain ( $F_{(1,58)} = 71.017, P < 0.001$ ), with D2 mice grooming more than B6 mice for all training types (No-US,  $P = 0.037$ ; Delay,  $P < 0.001$ ; Trace,  $P < 0.001$ ). Following delay or trace training, B6 mice show a clear reduction in grooming behavior (Delay,  $P = 0.007$ ; Trace,  $P < 0.001$ ) compared to the No-US condition, suggesting that in this strain, grooming decreases as freezing increases. Interestingly, D2 mice do not show a decrease in grooming following delay or trace training. In fact, there was no effect of training type on grooming in this strain under any condition tested.

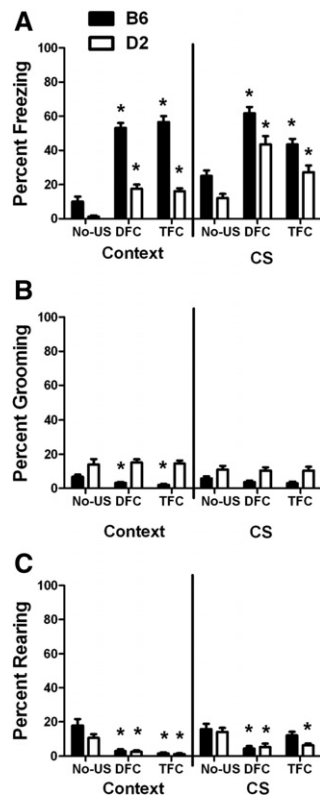
For the CS test (Fig. 5B), we found a significant effect of strain ( $F_{(1,58)} = 632.808, P < 0.001$ ) but not of training. D2 mice showed more grooming behavior than B6 mice under all training types (No-US,  $P = 0.03$ ; Delay,  $P = 0.003$ ; Trace,  $P = 0.004$ ). For the CS responses, there was no significant effect of training in either strain, although there were nonsignificant trends for a reduction in grooming in the B6 groups (Delay,  $P = 0.1$ ; Trace,  $P = 0.06$ ).

These data illustrate three points: (1) in B6 mice, grooming decreases as freezing increases, suggesting a simple suppression effect; (2) grooming behavior is not reflective of D2 learning in this paradigm, as training type has no effect; (3) grooming in D2 mice is not altered by the increase in freezing, indicating that the grooming process in these mice is an important behavior, even in the presence of fearful stimuli. Grooming behavior is commonly used in the assessment of anxiety and has been used to evaluate the effectiveness of anxiolytic drugs (Kalueff and Tuohimaa 2005a). Given that total grooming was not altered by training type, assessing grooming in fear conditioning paradigms does not seem to be a useful indicator of learning; however, it is possible that this behavior could be an important indicator of underlying anxiety levels in D2 mice, making it a useful control measure for assessing drugs or other manipulations targeted specifically to learning or anxiety in this paradigm.

#### Rearing

Rearing behavior in the context (Fig. 5C) showed a significant effect of strain ( $F_{(1,58)} = 4.903, P = 0.031$ ) and training ( $F_{(2,58)} = 40.934, P < 0.001$ ) with a significant strain  $\times$  training interaction ( $F_{(2,58)} = 3.247, P = 0.046$ ). For both B6 and D2 mice, rearing was significantly reduced by delay and trace training compared to No-US training ( $P < 0.001$  for all comparisons).

In the CS test, we found a significant effect of training ( $F_{(2,58)} = 12.944, P < 0.001$ ) but not strain. Here, rearing was reduced in all groups except the trace-trained B6 mice ( $P < 0.01$  for all comparisons, except B6 Trace,  $P = 0.29$ ). These results are



**Figure 5.** Training type comparisons. Percent (A) freezing, (B) grooming, and (C) rearing (mean  $\pm$  SEM) for B6 and D2 mice trained under No-US, delay (DFC) or trace (TFC) conditions. Total behaviors for the 12-min context test and the total CS freezing during the CS test are shown for each behavior. (\*)  $P < 0.05$  compared to the No-US group.

consistent with the hypothesis that rearing in this assay is an exploratory behavior that is reduced by the presence of fearful stimuli. The idea that rearing may be an active avoidance or escape behavior for D2 mice in this assay is not supported by our data.

It should be noted that for both rearing and grooming it is possible that it is specific sub-behaviors (rearing vs. leaning) or the microstructure of these behaviors, rather than the gross total measure, that are reflective of important learning or motivational differences (Fitch et al. 2002; Kalueff et al. 2005a). However, none of the analysis performed here suggest that either grooming or rearing can be used as an alternative measure of learning in these strains. Finally, although neither rearing nor grooming served as reliable indicators of learning here, these behaviors should be considered when selecting a scoring method for fear conditioning experiments. A number of automated scoring approaches, including beambreak, motion detectors, and tracking software, have been used to assess fear conditioned freezing (e.g., Bolivar et al. 2001; Fitch et al. 2002; Gale et al. 2009; Lattal and Maughan 2012). Given that each uses slightly different criteria for assessing movement, it is important to consider the ability of any scoring system to distinguish between freezing and other nonambulatory behaviors (such as grooming or rearing). For example, if the grooming and freezing data from the current experiments were combined, we would still observe robust strain differences in the context test, but, due to the high grooming levels in the D2 mice, the B6 and D2 measures in the CS test would no longer be significantly different (Delay,  $P = 0.06$ ; Trace,  $P = 0.07$ ). Thus, scoring approaches that give accurate assessments of strains with low levels of nonambulatory behaviors may not be optimal when applied to strains with higher grooming or rearing behaviors.

## Discussion

Although a number of groups have studied B6 and D2 mice using fear conditioning, this is the first report to compare differences both between strains and within each strain using delay and trace fear conditioning. While the detailed analysis presented here highlights several interesting differences between these strains and conditioning types, it is necessary to discuss these findings in the context of the existing literature. Primarily, it is important to address the existence of conflicting reports within the field and how they might be resolved. A list of the published works comparing B6 and D2 mice in contextual (no CS), delay, and trace fear conditioning paradigms can be found in Table 1. This table includes both the reported results and the experimental parameters used for each of the 30 studies reviewed here.

Although we focus on several variables that can affect the outcome of fear conditioning studies, it should be noted that our analysis of the experimental factors that contribute to observed strain differences is limited by certain aspects of the reported data. There are numerous variables that could also contribute to the variability in the literature that are not commonly manipulated or examined across studies. For example, the role of sex differences, testing during different stages of the light–dark cycle, and the order of context and CS testing could all affect the detection of strain differences. Many of these variables are left unmanipulated in an attempt to maintain continuity with previous work (e.g., testing during the light phase, testing context before CS), resulting in a dearth of information on these effects. For others, measures vary so greatly from study to study that a comprehensive analysis of their effects across the literature is difficult. For example, 11 of the studies reviewed here used both male and female subjects, but the level of analysis for sex differences varied greatly across studies, and the results include a mixture of no sex effects and sex effects that manifest

in one strain but not the other. The variability in these reports highlights the need for systematic investigations into these variables to determine their potential effects on fear conditioning comparisons.

## Context freezing

Consistent with our data (Fig. 2A,B), the majority of studies (25 of 29) on B6 and D2 mice find that D2 mice show reduced contextual freezing compared to B6 mice when they are trained using delay (23), trace (one), or contextual (five) fear conditioning (see Context Results in Table 1). This strain difference is very robust, persisting in spite of variations in the number of US exposures (Lattal and Maughan 2012), the duration of the context test (e.g., Ammassari-Teule et al. 2001; Lattal and Maughan 2012), and whether the context was processed as the foreground or background stimulus (e.g., Gerlai 1998; Andre et al. 2012). This difference has become one of the primary pieces of behavioral evidence for the hypothesis that contextual learning relies on the hippocampus, a structure that differs in both size and several functional parameters between these two strains (Nguyen et al. 2000; Balogh et al. 2002; Andre et al. 2012). Although it is clear that B6 and D2 mice have differing hippocampal processing, the role those differences play in the contextual conditioning behaviors observed in fear conditioning has been questioned (Gewirtz et al. 2000; Gerlai 2002). There is conflicting evidence regarding the necessity of the hippocampus in contextual conditioning (Matus-Amat et al. 2004; Wiltgen et al. 2006), and the underlying differences in hippocampal function may mean that contextual information gets processed very differently in these two strains (Rossi-Arnaud and Ammassari-Teule 1998; Ammassari-Teule et al. 2001; Sung et al. 2008). However, it is clear that regardless of the underlying mechanisms, contextual learning in these strains differs for both delay and trace fear conditioning.

It should be noted that while the majority of studies examined here report a difference in contextual responses in these strains, the effect is not universally observed. Six of the reviewed studies report that these strains perform identically during contextual tests (Owen et al. 1997b; Liu et al. 2003; Gale et al. 2009; Choy et al. 2012). A comparison of these reports does not yield a clear source for these differences. Two studies (Gale et al. 2009; Choy et al. 2012) use a high number of training shocks (three and four, respectively) and an automated scoring system, while the other two (Owen et al. 1997b; Liu et al. 2003) hand-score the data following two or one CS–US pairings, respectively. The length of the context test varies across these studies; however, our data, which report context freezing over a 12-min period, show that the strain differences in this test are robust throughout the testing period (Supplemental Fig. 1), making it unlikely that a very short or very long testing period would obscure the results. Two of these studies (Owen et al. 1997b; Choy et al. 2012) pre-exposed the mice to the training context prior to testing. Pre-exposure has been shown to alter later freezing responses to the context (Paylor et al. 1994), and interestingly, pre-exposure prior to training is one of the only differences between Owen et al. (1997b), in which no context difference was observed, and Owen et al. (1997a), in which a significant difference was reported. Choy et al. (2012) gave saline injections to their D2 mice, but not their B6 mice, suggesting that the stress of injection could have increased the level of D2 freezing. Interestingly, another paper comparing saline control groups in these strains found that the strain differences in both context and cued freezing varied depending on how the saline was delivered (i.p. injection vs. infusion pump) (Portugal et al. 2012). However, no single variable or clear combination of variables seems to explain the lack of contextual differences reported across these four studies.

**Table 1.** Reviewed literature

Author (date)	Animal information	Training	CS	US	Context test	CS test	Scoring method	Context results	CS results	Data presentation
Ammassari-Teule et al. (2001)	Charles River Not given	DFC, two pairings Pre-exposure	70-dB tone 2 × 8 sec	0.7 mA 2 sec	2 min	Five presentations 8 + 8-sec CS	Automated	B6 > D2	D2 > B6	Total seconds freezing
Andre et al. (2012)	Jackson Male	DFC and CTX, two pairings	85-dB white noise 30 sec	0.57 mA 2 sec	5 min	N/A	Handscored 10-sec intervals	B6 > D2	N/A	Percent observations
Balogh et al. (2002)	In house Male/Female	DFC, two pairings	80-dB clicker 30 sec	0.7 mA 2 sec	5 min	One presentation 3-min CS	Handscored 10-sec intervals	B6 > D2	B6 = D2	Percent observations freezing
Balogh and Wehner (2003)	In house Male	DFC, two pairings	80-dB clicker 30 sec	0.7 mA 2 sec	5 min	One presentation 3-min CS	Handscored 10-sec intervals	B6 > D2	B6 > D2	Percent observations freezing
Bolivar et al. (2001)	Jackson Male/Female	DFC, three pairings	90-dB tone 7 sec	0.5 mA 2 sec	5 min	Three presentations 5-sec CS	Photobeam breaks	B6 > D2	Variable	Percent of baseline activity
Bothe et al. (2005)	Jackson (B6) In house (D2) Male/Female	DFC, three pairings	90-dB tone 5 sec	0.5 mA 0.5 sec	5 min	Three presentations 5-sec CS	Photobeam breaks	B6 > D2	B6 > D2	Percent of baseline activity
Brigman et al. (2009) <sup>a</sup>	In house Male/Female	DFC, three pairings	80 dB tone 30 sec	0.6 mA 2 sec	5 min	One presentation 3-min CS	Handscored 5-sec intervals	B6 > D2	B6 > D2	Percent observations
Choy et al. (2012)	Not given Male Saline injections (D2)	CTX, four pairings Three pre-exposures	N/A	1.0 mA 3.5 sec	5 min	N/A	Automated (1-sec threshold)	B6 = D2	N/A	Percent time freezing
Fitch et al. (2002)	Not given Male	DFC, three pairings One pre-exposure	80-dB tone 20 sec	0.7 mA 1 sec	6 min	Three presentations 20-sec CS	Automated and handscored 10-sec intervals	B6 > D2	Variable	Percent time freezing
Fordyce et al. (1995)	In house Male Saline injections	DFC, one pairing	80-dB <sup>b</sup> clicker 30 sec	0.7 mA <sup>b</sup> 2 sec	5 min	One presentation 3-min CS	Handscored 10-sec intervals	B6 > D2	D2 > B6	Percent observations (adjusted)
Gale et al. (2009)	Jackson Male	DFC, three pairings	80-dB tone 15 sec	0.6 mA 1 sec	8 min	10 presentations 20-sec CS	Automated 20-sec intervals	B6 = D2	B6 > D2	Percent immobility/min
Gerlai (1998)	Not given Male	DFC, three pairings	80-dB tone 20 sec	0.7 mA 1 sec	6 min	Three presentations 20-sec CS	Automated	B6 > D2	B6 > D2	Percent freezing/min
Holmes et al. (2002)	Jackson Male/Female	TFC, four pairings, 2.5-sec interval DFC, four pairings	80-dB tone (TFC) or white noise (DFC/TFC) 30 sec	0.5 mA 2-sec	5 min	One presentation 3-min CS	Handscored 10-sec intervals	B6 > D2	B6 > D2	Percent observations
Hwang et al. (2010)	Charles River Male	TFC, six pairings, 18-sec interval One pre-exposure	80-dB tone 20 sec	0.5 mA 2 sec	N/A	Four presentations 20-sec CS	Handscored 1-sec intervals	N/A	Variable	Percent observations

*Continued*

**Table 1.** *Continued*

Author (date)	Animal information	Training	CS	US	Context test	CS test	Scoring method	Context results	CS results	Data presentation
Lattal and Maughan (2012)	In house Male/Female	CTX, one, four, and eight pairings	N/A	0.35 and 1.75 mA 2 sec	24 min	N/A	Activity monitor	B6 > D2	N/A	Percent freezing
Liu et al. (2003)	Jackson Male	DFC, one pairing	85-dB <sup>c</sup> tone 30 sec	0.70 mA 2 sec	Not given	One presentation 3-min CS	Handscored 5-sec intervals	B6 = D2	B6 = D2	Percent observations
Lu and Wehrner (1997)	In house Male Vehicle injections	DFC, two pairings	80-dB clicker 30 sec	0.35 mA 2 sec	5 min	One presentation 3-min CS	Handscored 10-sec intervals	B6 > D2	B6 = D2	Percent observations
Metz et al. (2006)	In house Male/Female Backcross littermates	DFC, one pairing <sup>b</sup>	80-dB clicker 30 sec <sup>b</sup>	0.6 mA 2 sec <sup>b</sup>	3 min	N/A	Handscored 10-sec intervals	B6 > D2	N/A	Percent observations
Nguyen et al. (2000)	Jackson Male	DFC, one pairing	85-dB tone 30 sec	0.7 mA 2 sec	5 min	One presentation 3-min CS	Handscored 5-sec intervals	B6 > D2	B6 > D2	Percent observations
Nie and Abel (2001)	Jackson Male	DFC, one pairing	83-dB white noise 30 sec	0.7 mA 2 sec	5 min	One presentation 3-min CS	Handscored 5-sec intervals	B6 > D2	B6 > D2	Percent observations
Owen et al. (1997a)	In house Male/Female	DFC, one pairing	80-dB clicker 30 sec	0.35 mA 2 sec	5 min	One presentation 3-min CS	Handscored 10-sec intervals	B6 > D2	B6 = D2	Percent observations
Owen et al. (1997b)	Jackson Male/Female	DFC, two pairings One pre-exposure	80-dB clicker 30 sec	0.35 mA 2 sec	5 min	One presentation 3-min CS	Handscored 10-sec intervals	B6 = D2	B6 = D2	Percent observations
Paylor et al. (1994)	In house Male	DFC, one pairing CTX, one pairing	80-dB clicker 30 sec	0.7 mA 2 sec	5 min	One presentation 3-min CS	Handscored	B6 > D2	B6 = D2	Percent observations (adjusted)
Portugal et al. (2012)	Jackson Male Saline injections and infusions	DFC, two pairings	85-dB white noise 30 sec	0.57 mA 2 sec	5 min	One presentation 3-min CS	Handscored 10-sec intervals	Variable	Variable	Percent observations
Stiedl et al. (1999)	Harlan-Winkelmann (B6) Bomholtgård (D2) Male	DFC, one pairing	75-dB tone 30 sec	0.7 mA 2 sec	3 min	One presentation 3-min CS	Beambreak and handscored 10-sec interval	B6 > D2	B6 > D2	Activity (cm/sec) Percent observations
Valentinuzzi et al. (1998)	Jackson Male	DFC, one pairing	White noise (dB unknown) 30 sec	0.6 mA 2 sec	5 min	One presentation 3-min CS	Beambreak 5-sec and handscored 5- and 10-sec intervals	Variable	B6 = D2	Percent observations Adjusted beam break scores

*Continued*



**Table 1.** *Continued*

Author (date)	Animal information	Training	CS	US	Context test	CS test	Scoring method	Context results	CS results	Data presentation
Voikar et al. (2005)	Harlan Male	DFC, two pairings	75-dB tone 30 sec	0.7 mA 2 sec	Not given	One presentation 2-min CS	Beambreak 3-sec threshold	B6 > D2	B6 > D2	Percent time with no beambreaks
Wehner et al. (1997)	In house Male/Female	DFC, two pairings	80-dB clicker 30 sec	0.35 mA 2 sec	5 min	One presentation 3-min CS	Handscored 10-sec intervals	B6 > D2	B6 > D2	Percent observations
Wilson et al. (2011)	In house Male/Female	CTX, one pairing	N/A	1.0 mA 3 sec	3 min	N/A	Handscored	B6 > D2	N/A	Percent observations
Yamada 2010	JCL Inc. Male	DFC, three pairings	70-dB white noise 15 sec	0.3 mA 2 sec	N/A	One presentation 2-min CS	Automated	N/A	B6 > D2	Percent time freezing
Young et al. (2000)	In house Male	DFC, two pairings	80-dB clicker 30 sec	0.35 mA 2 sec	5 min	One presentation 3-min CS	Handscored	B6 > D2	B6 = D2	Percent observations

The experimental parameters and results for published B6/D2 comparisons: Animal information: source of animals, sexes tested, and any manipulations; training: type and number of CS-US pairings (CS-US interval length is given for TFC); CS: type, duration, and dB level; US: magnitude and duration; Context Test: length of context exposure; CS Test: number of CS presentations and CS duration; Scoring Method: method used and parameters applied; Data Presentation: data format.

(N/A) not applicable, (DFC) delay fear conditioning, (TFC) trace fear conditioning, (CTX) context fear conditioning (no CS).

<sup>a</sup>Experimental parameters not listed in text: reported in Yang et al. (2008).

<sup>b</sup>Not listed in text; reported in Paylor et al. (1994).

<sup>c</sup>Reported as 8.5 dB.

## CS freezing

### *Delay fear conditioning*

The bulk of the fear conditioning literature for B6 and D2 mice uses delay training to compare CS responses, thus we will focus on that literature first. Unlike context freezing, there is a high level of variability in reports regarding CS responses between these strains. In our review of the literature, 13 groups report that B6 freeze more than D2, three groups report that B6 freeze less than D2, and 12 groups report no difference between the strains (see CS Results in Table 1). This level of variability suggests that CS responses may be more sensitive to variations in procedure than the more consistent contextual responses. Here, we highlight three major sources of variance in the existing literature for CS responses: (1) number of CS presentations during testing; (2) the duration of CS presentations; (3) baseline measures and data normalization.

### *Number of CS presentations*

In the present data (see Fig. 2C), D2 mice show reduced total CS freezing following delay training; however, the analysis of freezing behavior for each CS presentation shows that these differences are only observed during the second CS presentation, while freezing to the first CS presentation is equivalent across strains. Other reports in which multiple CS presentations were analyzed separately also report differences in responding between the CS presentations (Gerlai 1998; Fitch et al. 2002). The basis for this CS to CS difference is unclear. One possibility is that D2 mice exhibit inter-trial extinction between the two CS presentations. D2 mice show more rapid extinction than B6 mice (Balogh et al. 2002; Lattal and Maughan 2012), and it is possible that this extinction could occur within a single testing period. Alternatively, the reduced freezing to the second CS presentation could be an expression of weaker CS-US associations in D2 mice; however, it is very difficult to distinguish between a weaker initial association and an enhanced extinction process when the genetic background differs during both conditioning and extinction (Isiegas et al. 2006; Lattal and Maughan 2012). It is unlikely that this difference is due to simple CS acclimation, given that a similar decrease between CS presentations was not observed in our No-US group (see Fig. 4C).

This observation suggests that the number of CS presentations used during the CS test could account for some of the variability across reports. Our data indicate that studies in which the CS is only presented once are likely to find no difference between the strains. In agreement with this hypothesis, for nine of the 19 references that report equivalent CS freezing, only a single CS presentation was used during the CS test. However, not all reports using a single CS presentation reported equivalent freezing. Ten papers using a single CS presentation found differential freezing between strains. Interestingly, of these, eight used a high level of footshock (0.6–0.7 mA) (Paylor et al. 1994; Fordyce et al. 1995; Stiedl et al. 1999; Nguyen et al. 2000; Nie and Abel 2001; Balogh and Wehner 2003; Voikar et al. 2005; Brigman et al. 2009) and three used a high number of CS-US pairings (three or four pairings) (Holmes et al. 2002; Brigman et al. 2009; Yamada 2010), suggesting that increasing the training intensity could affect the detection of differences for a single CS presentation. Another paper that used a single CS and reported strain differences was a genetic screen that used a high number of mice per group (86 B6 and 79 D2) (Wehner et al. 1997), which might make it possible to pull out small strain differences that are not apparent at the more conventionally used animal numbers. This could mean that strain differences do exist in the initial CS responses, but that the standard 8–12 animals per group is too underpowered to detect

them. Taken together, the variations in these reports suggest that at lower training levels, a single CS presentation during testing is insufficient to detect strain differences. In addition, the ability to find strain differences in a single pairing using either increased training intensity or very high animal numbers suggests that these strain differences exist, albeit in a milder form, across all CS responses, and are somehow enhanced during later CS presentations.

### *Duration of CS presentation*

In addition to the number of CS presentations, the duration of the CS presentations may also affect the reported results. The minute-by-minute analysis of our data (Supplemental Fig. 1) shows that during the first minute of our CS presentations, B6 and D2 freezing is very similar, followed by a divergence between the groups as the CS progresses. The initial response to any auditory stimulus, whether trained or untrained, is likely a mixture of orienting behavior and startle response, especially when the stimulus is a loud noise. Thus, it is possible that freezing to the initial seconds of a CS is not reflective of learning, but rather of auditory startle/orienting. Six of the papers reviewed here used a short (<30 sec) CS presentation (Gerlai 1998; Ammassari-Teule et al. 2001; Bolivar et al. 2001; Fitch et al. 2002; Bothe et al. 2005; Gale et al. 2009). Of these, one report showed increased freezing in D2 mice compared to B6 (Ammassari-Teule et al. 2001), while three show decreased D2 freezing (Gerlai 1998; Bothe et al. 2005; Gale et al. 2009). The remaining two papers (Bolivar et al. 2001; Fitch et al. 2002) show variable CS differences depending on how the data were acquired (Fitch et al. 2002) and processed (Bolivar et al. 2001).

### *Trace fear conditioning*

To date, only two studies have compared B6 and D2 CS responses using trace fear conditioning (Holmes et al. 2002; Hwang et al. 2010). Consistent with the present data (Fig. 2D), both report that D2 mice freeze less than B6 mice during the CS test; however, both raise interesting parametric issues.

Holmes et al. (2002) found no difference between delay and trace fear conditioned D2 mice, while our data show a robust effect of training on CS response levels (Fig. 5A; Supplemental Fig. 5). A common finding across the trace fear conditioning literature is that the trace CS-US association is weaker than the one formed during delay conditioning, thus most groups find lower freezing levels following trace conditioning (Raybuck and Lattal 2014). The lack of training type effect in the Holmes et al. (2002) paper may be due to the very short trace interval used (2.5 sec). Trace fear conditioning is thought to activate the hippocampus and prefrontal regions due to the necessity of maintaining a CS-trace memory during the trace interval (Raybuck and Lattal 2014). As a result, the length of the interval is directly correlated with the recruitment of these mechanisms and the eventual response. The 2.5-sec interval used by Holmes et al. (2002) has been shown to be insufficient to recruit hippocampal processing (Misane et al. 2005), which may indicate that there are temporal boundaries to the durational effects on trace conditioning that will influence the detection of differences between trace and delay conditioned animals. Investigations into the relevance of the trace duration are important for a second reason: the use of “unpaired” CS-US presentations as a control group. As discussed previously, the use of unpaired control groups has several issues, one of which is the duration of the interval between the CS and US. Given the broad spectrum of CS-US intervals in which trace processing can occur (Misane et al. 2005) and the limited number of CS-US pairings typically used to generate fear conditioning, the potential to

form weak trace fear conditioning associations rather than an unpaired control group is a serious consideration for this approach. This is particularly true when one considers that the efficacy of a particular interval to act as a trace interval changes depending on the overall temporal context in which it is placed (e.g., Balsam 1984).

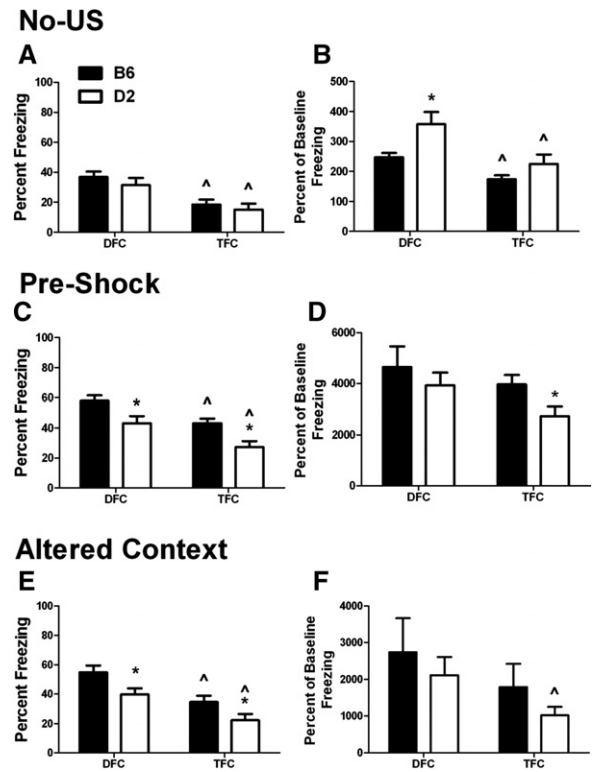
Hwang et al. (2010) also looked at TFC in B6 and D2 mice. Interestingly, their results for CS strain differences were variable. The raw data showed a decrease in D2 freezing compared to B6; however, when they included a control group exposed to the US only into their analyses, these differences were no longer significant. This leads to the final major source of variation across the published fear conditioning literature: the use of control groups and the normalization of data to baseline measures.

#### Baseline measures and data normalization approaches

A major concern within the field is how to address the presence of baseline behavioral differences when comparing across strains. B6 and D2 mice differ on a number of behavioral measures prior to training in any paradigm (Logue et al. 1997; Podhorna and Brown 2002), which may or may not have an effect on the learned behavior. In three of the papers reviewed here (Valentinuzzi et al. 1998; Bolivar et al. 2001; Hwang et al. 2010), adjusting the raw data to a baseline or control group changed the significance of the strain comparison. This leaves the question of which is truly representative of learning differences in these strains, the raw data or the adjusted measures? Thus, the question is two-fold: (1) should we use a baseline measure to normalize data across groups with inherent differences; (2) how do we determine the appropriate baseline measure and subsequent mathematical manipulation to use? When two groups differ in baseline performance, it becomes very difficult to conclude that there are differences in post-conditioning performance because all baseline normalization transformations necessarily include assumptions about how learning is reflected in performance at different parts of the behavioral scale. For example, two of the most commonly used transformations to assess changes in performance before and after conditioning (absolute difference or percent change from baseline) can lead to very different conclusions about learning processes (see Lattal 1999). However, the presence of significant differences between strains at pretraining or in control groups has been observed by several groups, leading to the presentation of adjusted measures in fear conditioning reports (Jacobs et al. 2010).

For those who chose to control for baseline, there is a variety of both the measure used to represent “baseline” behavior and the mathematical manipulations employed in the normalization, both of which can significantly alter the reported effects. To illustrate the potential variability introduced by different baseline adjustments, we transformed our trace and delay CS data using three possible baseline measures: (1) our No-US responses, (2) altered context responses (e.g., Paylor et al. 1994), and (3) pretraining responses (e.g., Andre et al. 2012). For all three baselines, we used two common mathematical normalization approaches: (1) subtraction of baseline (Total CS freezing–baseline freezing) and (2) percentage of baseline freezing (Total CS freezing/baseline freezing  $\times$  100) (Jacobs et al. 2010). The results of each transformation are shown in Figure 6 (note: for baseline measures of 0, an adjusted measure of 1 was used in the percentage method to avoid mathematical errors).

In Figure 6A, the average total CS response of the No-US group was subtracted from the total CS responses of each animal in the delay and trace fear conditioned groups. This manipulation yields a significant effect of training ( $F_{(1,44)} = 20.593, P < 0.001$ ), but no effect of strain. While the trace groups still show reduced freezing compared to the delay groups (B,  $P < 0.001$ ; D2,  $P =$



**Figure 6.** Baseline normalization. Total CS freezing for delay (DFC) and trace (TFC) fear conditioning are shown normalized to three baseline measures using either a subtraction (A,C,E) or a percentage (B,D,F) method. (A,B) Normalization to No-US CS responses using (A) subtraction and (B) percentage; (C,D) normalization to the preshock period of training using (C) subtraction and (D) percentage; (E,F) normalization to altered context freezing during the CS test using (E) subtraction and (F) percentage. (\*)  $P < 0.05$  for B6 versus D2; (^)  $P < 0.05$  for DFC versus TFC.

0.013), the B6 and D2 differences for each training type are no longer significant. When the CS response is presented as a percentage of the average No-US CS response (Fig. 6B), we retain both the significant effect of strain ( $F_{(1,44)} = 9.38, P = 0.004$ ) and training ( $F_{(1,44)} = 15.321, P < 0.001$ ); however, significant strain differences are only detected for the delay-trained mice ( $P = 0.013$ ), and the D2 mice now show increased freezing compared to B6.

In Figure 6, C and D, we use the same mathematical manipulations but this time use each animal’s preshock freezing from the training period as the baseline. For the subtraction condition (Fig. 6C), we see a significant effect of both strain ( $F_{(1,44)} = 17.135, P < 0.001$ ) and training ( $F_{(1,44)} = 17.135, P < 0.001$ ), with a significant strain difference for both delay- ( $P = 0.016$ ) and trace-trained mice ( $P = 0.003$ ), and both strains showing reduced freezing following trace fear conditioning compared to delay fear conditioning (B6,  $P = 0.003$ ; D2,  $P = 0.013$ ). However, when the data are normalized to preshock freezing using the percentage method (Fig. 6D), there is no significant effect of strain or training ( $P = 0.069$  and  $P = 0.077$ , respectively).

Finally, we normalized to freezing in the altered context (first 3 min of the CS test) (Fig. 6E,F). Using the subtraction approach, there was a significant effect of both strain ( $F_{(1,44)} = 11.058, P = 0.002$ ) and training ( $F_{(1,44)} = 20.747, P < 0.001$ ) with D2 freezing less than B6 and trace conditioned animals freezing less than delay conditioned animals. However, the percentage manipulation for this baseline measure shows no effect of either (Strain,  $P = 0.251$ ; Training,  $P = 0.094$ ).

In general, the percent of baseline method had the greatest effect on the significance of our findings, primarily due to the large amount of variability this manipulation introduced to the data. For many of our subjects, freezing during the preshock and altered context periods was at zero, making those percentage measures very high. The low levels of preshock and altered context freezing also explain why the subtraction methods for these baselines had very little effect on the overall outcome. However, these baseline measures are not reliably low across different reports. Altered context freezing, sometimes referred to as “generalized freezing,” has been reported to differ in B6 and D2 mice (e.g., Young et al. 2000; Balogh et al. 2002; Balogh and Wehner 2003; Hwang et al. 2010), but these differences are not consistently found (e.g., Lu and Wehner 1997; Nie and Abel 2001) and were absent in our data (Fig. 2B,C; Supplemental Fig. 1). Moreover, it is unclear how the altered context measure actually relates to the expression of CS–US or context–US associations, as it has been shown to correlate with both CS freezing and context freezing (Wood and Anagnostaras 2011). Preshock freezing during training, on the other hand, is correlated with context freezing, but not CS freezing (Wood and Anagnostaras 2011), suggesting that it is, in fact, a poor baseline for CS response normalization.

The nonassociative responses presented in this paper have the advantage of allowing for the direct comparison of CS responses, rather than relying on some other measure that may or may not be reflective of CS freezing. However, the major limitation for this measure is the inability to compare behavior within a single subject, as the No-US, delay, and trace groups were all separate. The most interesting outcome for manipulations using the No-US as a baseline is the effect on strain differences. The subtraction method (Fig. 6A) shows that the magnitude in differences between the strains for the three training paradigms is very similar, which could suggest that the strain differences observed during the CS test are actually reflective of nonassociative differences in CS responding, rather than learning differences.

Most importantly, these differences between outcomes for different normalization approaches highlight the need to report raw data. For reports that give only adjusted measures, it is unclear if the reported effects were also detected in the raw data, making it difficult not only to interpret the findings, but also to compare across papers with differing approaches to baseline normalization.

## Conclusions

Here, we show a detailed analysis of B6 and D2 behavior following delay and trace fear conditioning. We found both a significant effect of strain on both types of fear conditioning and a significant difference between training types within each strain. While our data are in agreement with a number of previously published studies, our analysis and literature review highlight a number of important experimental parameters that contribute to the variability in previously reported results. Specifically, the number and duration of CS presentations during testing and the normalization of data seem to be the primary sources of differences in findings across labs. Most importantly, these differences highlight the need to include raw measurements for all reports in order to facilitate future comparisons within this paradigm.

## Materials and Methods

### Animals

Male C57BL/6 and DBA/2 mice (Jackson Laboratory, Bar Harbor, ME) between 8 and 12 wk of age were used for all studies. The delay and trace experiments had 12 mice per group, while the No-US condition had eight mice per group. The mice were housed four

per cage (28 cm × 17.5 cm × 11 cm) under standard housing conditions (12-h light–dark cycle) with access to food (50 LOD Irradiated Rodent Diet, LabDiet) and water ad libitum. Mice were allowed to acclimate to the colony room for 1 wk prior to any experimentation. Experiments were performed during the light portion of the light–dark cycle. All studies were approved by the Institutional Animal Care and Use Committee at Oregon Health & Science University and conducted in accordance with the guidelines provided by the National Institutes of Health.

### Apparatus

Training and contextual testing were conducted in circular Plexiglas chambers (diameter 14.5 cm) mounted on a rod floor as described in Stafford and Lattal (2009). Each chamber was housed in an individual sound-attenuating box (Med-Associates). An 85-dB white noise CS was administered through a sound generator (Coulbourn Instruments), and a 0.35-mA footshock US was administered through the rod floor via a shock scrambler/generator (Coulbourn). To provide a distinct olfactory cue, the apparatus was cleaned with 0.1% acetic acid prior to training and testing. The training and context testing sessions were controlled by an IBM-PC running Graphic State software (Coulbourn). Testing for cued fear conditioning was conducted in rectangular conditioning chambers with Plexiglas floors (Med-Associates). Each chamber was housed in an individual sound-attenuating box and located in a room that was different from the training/contextual testing room. In these chambers, the same CS used during training was generated with ANL-926 (Med-Associates) and administered through speakers mounted on the left wall of the chambers. The CS presentations were controlled by an IBM-PC running MED-PC 4 (Med-Associates). The cue testing chambers were cleaned with 90% ethanol.

### Fear conditioning

Animals were acclimated to the training room and handling for 2 d prior to training. No pre-exposure to the chamber or CS was used. Each day (including training and testing days) mice were moved into the room 1 h prior to any procedures and left in the room for 1 h after all procedures were completed. The delay and trace training methods follow those described by Raybuck and Lattal (2011).

### Training

Each session began with the activation of a house light. After 2 min, a 30-sec CS was activated. For delay fear conditioning, the 30-sec CS co-terminated with a 2-sec 0.35-mA footshock (US). Under the trace fear conditioning paradigm, the 30-sec CS was followed by a 30-sec trace interval and then the 2-sec US. For both training paradigms, animals received two CS–US pairings in a 6.5-min session. In the no-US condition, the same CS presentations and timing were used, but the US was omitted.

### Testing

To evaluate contextual learning, mice were returned to the training chambers 24 h after training. The houselight was activated, and behavior was assessed for 12 min. To test cued learning, mice were placed in the novel cue testing apparatus 48 h after training. The animals were assessed for responses to an altered context for 3 min, followed by two 3-min CS presentations separated by a 3-min inter-trial interval and followed by a 3-min post-CS period, for a total test time of 15 min.

### Scoring

All training sessions, contextual tests, and cue tests were hand-scored by two experimenters. Animals were scored at 10-sec intervals for freezing, rearing, and grooming, and the scores were averaged across the two experimenters. Freezing was defined as the absence of all movement except respiration (Blanchard and Blanchard 1969). Grooming behavior included licking or chewing

the paws or body and rubbing/scratching the head and body. Rearing was defined as both front paws lifted off the floor with an extended body and included leaning (front paws resting on the chamber wall) and free rearing (rearing with no wall support). Data are presented as the percentage of observations for each behavior (six observations per minute).

### Footshock threshold

Mice from the no-US group were tested for footshock threshold 1 wk after testing for fear conditioning. The mice were placed in the chambers used for cue testing with the Plexiglass floors removed and allowed 30 sec to acclimate to the chamber. To test the sensitivity to the US, we started at a subthreshold shock intensity (0.05 mA) that generated no response in the animals. Subsequent shocks at increasing intensities were administered every 30 sec. The shock intensity was increased by 0.05 mA until animals responded via either changes in locomotor behavior or vocalization, and the threshold was defined as the intensity at which the first response was observed.

### Analysis

The data were analyzed using Excel (Microsoft) and Systat 13 (Cranes Software International Ltd) programs, and the Prism 5 (GraphPad Software) program was used to generate all graphs. The data are reported as the mean  $\pm$  standard error of the mean (SEM) of the percent observations. The training data are shown in 30-sec bins to correspond with the CS duration. The context and CS test data are shown in 1-min bins (Supplemental Material) and 3-min bins along with total context, non-CS, and CS averages. All comparisons were made using repeated or two-way ANOVAs followed by Student's *t*-test for comparisons between individual groups. A *P*-value  $<0.05$  was considered significant.

### Literature search

The reports on fear conditioning in male B6 and D2 mice were collected through the PubMed.gov website. Only reports published in the last 20 years that looked at male mice in both strains using either contextual (no CS), delay, or trace fear conditioning were included.

### Acknowledgments

This work was funded by National Institutes of Health 1F32 AA022011 (to M.E.T.); National Institutes of Health F32DA031537 (to J.D.R.); National Institutes of Health R01 DA005228, P60 AA10760, R01 AA011114, and R24 AA020245 (to K.J.B.); National Institutes of Health DA02592 and US Department of the Army/DOD-TATRC:W81XWH-12-2-0048 (to K.M.L.); and National Institutes of Health T32 DA07262 and T32 AA07468 and the Portland VA Medical Center.

### References

Amassari-Teule M, Restivo L, Pietteur V, Passino E. 2001. Learning about the context in genetically-defined mice. *Behav Brain Res* **125**: 195–204.

Andre JM, Cordero KA, Gould TJ. 2012. Comparison of the performance of DBA/2 and C57BL/6 mice in transitive inference and foreground and background contextual fear conditioning. *Behav Neurosci* **126**: 249–257.

Balogh SA, Wehner JM. 2003. Inbred mouse strain differences in the establishment of long-term fear memory. *Behav Brain Res* **140**: 97–106.

Balogh SA, Radcliffe RA, Logue SF, Wehner JM. 2002. Contextual and cued fear conditioning in C57BL/6J and DBA/2J mice: context discrimination and the effects of retention interval. *Behav Neurosci* **116**: 947–957.

Balsam P. 1984. Relative time in trace conditioning. *Ann N Y Acad Sci* **423**: 211–227.

Blanchard RJ, Blanchard DC. 1969. Crouching as an index of fear. *J Comp Physiol Psychol* **67**: 370–375.

Bolivar VJ, Pooler O, Flaherty L. 2001. Inbred strain variation in contextual and cued fear conditioning behavior. *Mamm Genome* **12**: 651–656.

Bothe GW, Bolivar VJ, Vedder MJ, Geistfeld JG. 2005. Behavioral differences among fourteen inbred mouse strains commonly used as disease models. *Comp Med* **55**: 326–334.

Brigman JL, Poonam M, Lu L, Williams RW, Holmes A. 2009. Genetic relationship between anxiety- and fear-related behaviors in BXD recombinant inbred mice. *Behav Pharmacol* **20**: 204–209.

Choy KH, Yu J, Hawkes D, Mayorov DN. 2012. Analysis of vigilant scanning behavior in mice using two-point digital video tracking. *Psychopharmacology (Berl)* **221**: 649–657.

Crusio WE. 2001. Genetic dissection of mouse exploratory behaviour. *Behav Brain Res* **125**: 127–132.

Dellu F, Contarino A, Simon H, Koob GF, Gold LH. 2000. Genetic differences in response to novelty and spatial memory using a two-trial recognition task in mice. *Neurobiol Learn Mem* **73**: 31–48.

Estanislau C, Diaz-Moran S, Canete T, Balzquez G, Tobena A, Fernandez-Teruel A. 2013. Context-dependent differences in grooming behavior among the NIH heterogeneous stock and the Roman high- and low-avoidance rats. *Neurosci Res* **77**: 187–201.

Fanselow MS. 1980. Conditioned and unconditional components of post-shock freezing. *Pavlov J Biol Sci* **15**: 177–182.

Fitch T, Adams B, Chaney S, Gerlai R. 2002. Force transducer-based movement detection in fear conditioning in mice: a comparative analysis. *Hippocampus* **12**: 4–17.

Fordyce DE, Clark VJ, Paylor R, Wehner JM. 1995. Enhancement of hippocampally-mediated learning and protein kinase C activity by oxiracetam in learning-impaired DBA/2 mice. *Brain Res* **672**: 170–176.

Gale GD, Yazdi RD, Khan AH, Lulusi AJ, Davis RC, Smith DJ. 2009. A genome-wide panel of congenic mice reveals widespread epistasis of behavior quantitative trait loci. *Mol Psychiatry* **14**: 631–645.

Gerlai R. 1998. Contextual learning and cue association in fear conditioning in mice: a strain comparison and a lesion study. *Behav Brain Res* **95**: 191–203.

Gerlai R. 2002. Hippocampal LTP and memory in mouse strains: Is there evidence for a causal relationship? *Hippocampus* **12**: 657–666.

Gewirtz JC, McNish KA, Davis M. 2000. Is the hippocampus necessary for contextual fear conditioning? *Behav Brain Res* **110**: 83–95.

Graybeal C, Bachu M, Mozhui K, Saksida LM, Bussey TJ, Sagalyn E, Williams RW, Holmes A. 2014. Strains and stressors: an analysis of touchscreen learning in genetically diverse mouse strains. *PLoS One* **9**: e87745.

Griebel G, Sanger DJ, Perrault G. 1997. Genetic differences in the mouse defense test battery. *Aggress Behav* **23**: 19–31.

Heyser CJ, McDonald JS, Polis IY, Gold LH. 1999. Strain distribution of mice in discriminated Y-maze avoidance learning: genetic and procedural differences. *Behav Neurosci* **113**: 91–102.

Holahan MR, White NM. 2002. Conditioned memory modulation, freezing, and avoidance as measures of amygdala-mediated conditioned fear. *Neurobiol Learn Mem* **77**: 250–275.

Holmes A, Wrenn CC, Harris AP, Thayer KE, Crawley JN. 2002. Behavioral profiles of inbred strains on novel olfactory, spatial and emotional tests for reference memory in mice. *Genes Brain Behav* **1**: 55–69.

Hwang YK, Song JC, Han SH, Cho J, Smith DR, Gallagher M, Han JS. 2010. Differences in hippocampal CREB phosphorylation in trace fear conditioning of two inbred mouse strains. *Brain Res* **1345**: 156–163.

Isiegas C, Park A, Kandel ER, Abel T, Lattal KM. 2006. Transgenic inhibition of neuronal protein kinase A activity facilitates fear extinction. *J Neurosci* **26**: 12700–12707.

Jacobs NS, Cushman JD, Fanselow MS. 2010. The accurate measurement of fear memory in Pavlovian conditioning: resolving the baseline issue. *J Neurosci Methods* **190**: 235–239.

Kalueff AV, Tuohimaa P. 2005a. Mouse grooming microstructure is a reliable anxiety marker bidirectionally sensitive to GABAergic drugs. *Eur J Pharmacol* **508**: 147–153.

Kalueff AV, Tuohimaa P. 2005b. Contrasting grooming phenotypes in three mouse strains markedly different in anxiety and activity (129S1, BALB/c and NMRI). *Behav Brain Res* **160**: 1–10.

Lattal KM. 1999. Trial and intertrial durations in Pavlovian conditioning: issues of learning and performance. *J Exp Psychol Anim Behav Process* **25**: 433–450.

Lattal KM, Maughan DK. 2012. A parametric analysis of factors affecting acquisition and extinction of contextual fear in C57BL/6 and DBA/2 mice. *Behav Processes* **90**: 49–57.

Liu D, Singh RP, Khan AH, Bhavsar K, Lulusi AJ, Davis RC, Smith DJ. 2003. Identifying loci for behavioral traits using genome-tagged mice. *J Neurosci Res* **74**: 562–569.

Logue SF, Owen EH, Rasmussen DL, Wehner JM. 1997. Assessment of locomotor activity, acoustic and tactile startle, and prepulse inhibition of startle in inbred mouse strains and F1 hybrids: implications of genetic background for single gene and quantitative trait loci analyses. *Neuroscience* **80**: 1075–1086.

Lu Y, Wehner JM. 1997. Enhancement of contextual fear-conditioning by putative (+/-)-alpha-amino-3-hydroxy-5-methylisoxazole-4-

- propionic acid (AMPA) receptor modulators and N-methyl-D-aspartate (NMDA) receptor antagonists in DBA/2J mice. *Brain Res* **768**: 197–207.
- Matus-Amat P, Higgins EA, Barrientos RM, Rudy JW. 2004. The role of the dorsal hippocampus in the acquisition and retrieval of context memory representations. *J Neurosci* **24**: 2431–2439.
- Metz AV, Chynoweth J, Allan AM. 2006. Influence of genetic background on alcohol drinking and behavioral phenotypes of 5-HT3 receptor over-expressing mice. *Pharmacol Biochem Behav* **84**: 120–127.
- Misane I, Tovote P, Meyer M, Spiess J, Ogren SO, Stiedl O. 2005. Time-dependent involvement of the dorsal hippocampus in trace fear conditioning in mice. *Hippocampus* **15**: 418–426.
- Nguyen PV, Gerlai R. 2002. Behavioural and physiological characterization of inbred mouse strains: prospects for elucidating the molecular mechanisms of mammalian learning and memory. *Genes Brain Behav* **1**: 72–81.
- Nguyen PV, Abel T, Kandel ER, Bourtochouladze R. 2000. Strain-dependent differences in LTP and hippocampus-dependent memory in inbred mice. *Learn Mem* **7**: 170–179.
- Nie T, Abel T. 2001. Fear conditioning in inbred mouse strains: an analysis of the time course of memory. *Behav Neurosci* **115**: 951–956.
- O'Leary TP, Gunn RK, Brown RE. 2013. What are we measuring when we test strain differences in anxiety in mice? *Behav Genet* **43**: 34–50.
- Owen EH, Christensen SC, Paylor R, Wehner JM. 1997a. Identification of quantitative trait loci involved in contextual and auditory-cued fear conditioning in BxD recombinant inbred strains. *Behav Neurosci* **111**: 292–300.
- Owen EH, Logue SF, Rasmussen DL, Wehner JM. 1997b. Assessment of learning by the Morris water task and fear conditioning in inbred mouse strains and F1 hybrids: implications of genetic background for single gene mutations and quantitative trait loci analyses. *Neuroscience* **80**: 1087–1099.
- Parker CC, Sokoloff G, Cheng R, Palmer AA. 2012. Genome-wide association for fear conditioning in an advanced intercross mouse line. *Behav Genet* **42**: 437–448.
- Paylor R, Crawley JN. 1997. Inbred strain differences in prepulse inhibition of the mouse startle response. *Psychopharmacology (Berl)* **132**: 169–180.
- Paylor R, Tracy R, Wehner J, Rudy JW. 1994. DBA/2 and C57BL/6 mice differ in contextual fear but not auditory fear conditioning. *Behav Neurosci* **108**: 810–817.
- Paylor R, Baskall-Baldini L, Yuva L, Wehner JM. 1996. Developmental differences in place-learning performance between C57BL/6 and DBA/2 mice parallel the ontogeny of hippocampal protein kinase C. *Behav Neurosci* **110**: 1415–1425.
- Perez-Saad H, Bures J. 1983. Cortical spreading depression blocks naloxone-induced escape behaviour in morphine pretreated mice. *Pharmacol Biochem Behav* **18**: 145–147.
- Podhorna J, Brown RE. 2002. Strain differences in activity and emotionality do not account for differences in learning and memory performance between C57BL/6 and DBA/2 mice. *Genes Brain Behav* **1**: 96–110.
- Ponder CA, Kliethermes CL, Drew MR, Muller J, Das K, Risbrough VB, Crabbe JC, Gilliam TC, Palmer AA. 2007. Selection for contextual fear conditioning affects anxiety-like behaviors and gene expression. *Genes Brain Behav* **6**: 736–749.
- Portugal GS, Wilkinson DS, Kenny JW, Sullivan C, Gould TJ. 2012. Strain dependent effects of acute, chronic, and withdrawal from chronic nicotine on fear conditioning. *Behav Genet* **42**: 133–150.
- Raybuck JD, Lattal KM. 2011. Double dissociation of amygdala and hippocampal contributions to trace and delay fear conditioning. *PLoS One* **6**: e15982.
- Raybuck JD, Lattal KM. 2014. Bridging the interval: theory and neurobiology of trace conditioning. *Behav Processes* **101C**: 103–111.
- Rescorla RA. 1967. Pavlovian conditioning and its proper control procedures. *Psychol Rev* **74**: 71–80.
- Restivo L, Chaillan FA, Ammassari-Teule M, Roman FS, Marchetti E. 2006. Strain differences in rewarded discrimination learning using the olfactory tubing maze. *Behav Genet* **36**: 923–934.
- Rossi-Arnaud C, Ammassari-Teule M. 1998. What do comparative studies of inbred mice add to current investigations on the neural basis of spatial behaviors? *Exp Brain Res* **123**: 36–44.
- Ryoke R, Yamada K, Ichitani Y. 2014. Long-term effects of traumatic stress on subsequent contextual fear conditioning in rats. *Physiol Behav* **129C**: 30–35.
- Schimanski LA, Nguyen PV. 2004. Multidisciplinary approaches for investigating the mechanisms of hippocampus-dependent memory: a focus on inbred mouse strains. *Neurosci Biobehav Rev* **28**: 463–483.
- Stafford JM, Lattal KM. 2009. Direct comparisons of the size and persistence of anisomycin-induced consolidation and reconsolidation deficits. *Learn Mem* **16**: 494–503.
- Stiedl O, Radulovic J, Lohmann R, Birkenfeld K, Palve M, Kammermeier J, Sananbenesi F, Spiess J. 1999. Strain and substrain differences in context- and tone-dependent fear conditioning of inbred mice. *Behav Brain Res* **104**: 1–12.
- Sung JY, Goo JS, Lee DE, Jin DQ, Bizon JL, Gallagher M, Han JS. 2008. Learning strategy selection in the water maze and hippocampal CREB phosphorylation differ in two inbred strains of mice. *Learn Mem* **15**: 183–188.
- Tang X, Orchard SM, Sanford LD. 2002. Home cage activity and behavioral performance in inbred and hybrid mice. *Behav Brain Res* **136**: 555–569.
- Trullas R, Skolnick P. 1993. Differences in fear motivated behaviors among inbred mouse strains. *Psychopharmacology (Berl)* **111**: 323–331.
- Valentinuzzi VS, Kolker DE, Vitaterna MH, Shimomura K, Whiteley A, Low-Zeddies S, Turek FW, Ferrari EA, Paylor R, Takahashi JS. 1998. Automated measurement of mouse freezing behavior and its use for quantitative trait locus analysis of contextual fear conditioning in (BALB/CJ × C57BL/6J)F2 mice. *Learn Mem* **5**: 391–403.
- Voikar V, Polus A, Vasar E, Rauvala H. 2005. Long-term individual housing in C57BL/6J and DBA/2 mice: assessment of behavioral consequences. *Genes Brain Behav* **4**: 240–252.
- Wehner JM, Sleight S, Upchurch M. 1990. Hippocampal protein kinase C activity is reduced in poor spatial learners. *Brain Res* **523**: 181–187.
- Wehner JM, Radcliffe RA, Rosmann ST, Christensen SC, Rasmussen DL, Fulker DW, Wiles M. 1997. Quantitative trait locus analysis of contextual fear conditioning in mice. *Nat Genet* **17**: 331–334.
- Weinberger SB, Koob GF, Martinez JL Jr. 1992. Differences in one-way active avoidance learning in mice of three inbred strains. *Behav Genet* **22**: 177–188.
- Wilson YM, Brodnicki TC, Lawrence AJ, Murphy M. 2011. Congenic mouse strains enable discrimination of genetic determinants contributing to fear and fear memory. *Behav Genet* **41**: 278–287.
- Wiltgen BJ, Sanders MJ, Anagnostaras SG, Sage JR, Fanselow MS. 2006. Context fear learning in the absence of the hippocampus. *J Neurosci* **26**: 5484–5491.
- Wood SC, Anagnostaras SG. 2011. Interdependence of measures in Pavlovian conditioned freezing. *Neurosci Lett* **505**: 134–139.
- Yamada K. 2010. Strain differences of selective attention in mice: effect of Kamin blocking on classical fear conditioning. *Behav Brain Res* **213**: 126–129.
- Yang RJ, Mzhui K, Karlsson RM, Cameron HA, Williams RW, Holmes A. 2008. Variation in mouse basolateral amygdala volume is associated with differences in stress reactivity and fear learning. *Neuropsychopharmacology* **33**: 2595–2604.
- Young EA, Owen EH, Meiri KF, Wehner JM. 2000. Alterations in hippocampal GAP-43 phosphorylation and protein level following contextual fear conditioning. *Brain Res* **860**: 95–103.
- Zheng QY, Johnson KR, Erway LC. 1999. Assessment of hearing in 80 inbred strains of mice by ABR threshold analyses. *Hear Res* **130**: 94–107.

Received April 1, 2014; accepted in revised form June 10, 2014.