Published in final edited form as:

Nat Neurosci. 2019 May; 22(5): 753-761. doi:10.1038/s41593-019-0361-z.

# Distinct hippocampal engrams control extinction and relapse of fear memory

Anthony F. Lacagnina<sup>1</sup>, Emma T. Brockway<sup>1</sup>, Chelsea R. Crovetti<sup>1</sup>, Francis Shue<sup>1</sup>, Meredith J. McCarty<sup>1</sup>, Kevin P. Sattler<sup>1</sup>, Sean C. Lim<sup>2,3</sup>, Sofia Leal Santos<sup>2,3,4,5</sup>, Christine A. Denny<sup>2,3</sup>, Michael R. Drew<sup>1,\*</sup>

<sup>1</sup>Center for Learning and Memory, Department of Neuroscience, University of Texas at Austin, Austin, TX 78712, USA

<sup>2</sup>Department of Psychiatry, Columbia University, New York, NY 10032, USA

<sup>3</sup>Division of Systems Neuroscience, New York State Psychiatric Institute (NYSPI)/Research Foundation for Mental Hygiene, Inc. (RFMH), New York, NY 10032, USA

<sup>4</sup>Life and Health Sciences Research Institute (ICVS), School of Medicine, University of Minho, Braga, Portugal

<sup>5</sup>ICVS/3B's, PT Government Associate Laboratory, Braga/Guimarães, Portugal

#### **Abstract**

Learned fear often relapses after extinction, suggesting that extinction training generates a new memory that coexists with the original fear memory; however, the mechanisms governing expression of competing fear and extinction memories remain unclear. We used activity-dependent neural tagging to investigate representations of fear and extinction memories in the dentate gyrus (DG). We demonstrate that extinction training suppresses reactivation of context fear engram cells, while activating a second ensemble, a putative extinction engram. Optogenetic inhibition of neurons that were active during extinction training increased fear after extinction training, whereas silencing neurons that were active during fear training reduced spontaneous recovery of fear. Optogenetic stimulation of fear acquisition neurons increased fear, while stimulation of extinction neurons suppressed fear and prevented spontaneous recovery. Our results indicate the hippocampus generates a fear extinction representation and that interactions between hippocampal fear and extinction representations govern suppression and relapse of fear after extinction.

AUTHOR CONTRIBUTIONS

A.F.L., C.A.D. and M.R.D. conceived the project and designed experiments. A.F.L., S.C.L. and S.L.S. performed the cell reactivation experiments. A.F.L. performed the optogenetic behavioral experiments. A.F.L., E.T.B., C.R.C., F.S., M.J.M., K.P.S., S.C.L., and S.L.S. quantified cells and analyzed data. C.A.D. generated the  $ArcCreER^{T2}$  mice. A.F.L. and M.R.D. wrote the manuscript. M.R.D. supervised the project.

COMPETING INTERESTS STATEMENT

The authors declare no competing interests.

Users may view, print, copy, and download text and data-mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use:http://www.nature.com/authors/editorial\_policies/license.html#terms

<sup>\*</sup>Correspondence: mdrew@utexas.edu.

## INTRODUCTION

Maladaptive learned fear is commonly treated using therapies based on extinction—repeated exposure to a fear-evoking stimulus in the absence of threat<sup>1,2</sup>. Extinction does not permanently eliminate the fear memory. Fear can, for instance, relapse with the passage of time, a phenomenon known as spontaneous recovery<sup>3,4</sup>. The transience of extinction suggests that learned fear is not erased by extinction training but, instead, establishes a new memory that suppresses the fear memory or competes with it for expression<sup>5,6</sup>. The acquisition of an extinction memory depends critically on prefrontal cortex (PFC)-to-amygdala projections<sup>7,8</sup>, but the mechanisms governing expression of extinction are not well understood. Based on evidence that inhibiting hippocampal activity interferes with extinction learning<sup>9</sup> and the context-dependency of extinction retrieval<sup>10–12</sup>, we investigated how activity of hippocampal ensembles regulates the expression of an extinction memory.

The hippocampal DG plays a critical role in acquisition of context fear memory—fear of a place where an aversive experience occurred. Context fear acquisition activates a sparse ensemble of DG granule cells, sometimes termed "fear engram cells," whose reactivation is necessary and sufficient 14–17 for expression of context fear. Fear engram cells retain their ability to evoke fear when optogenetically stimulated at remote time points 18 and after amnestic treatments 16,19, suggesting that they constitute a stable neural representation of the contextual fear memory. However, it is unknown how extinction training affects the activity of fear engram cells. We recently discovered that activity in the DG is necessary for extinction of context fear which led us to hypothesize that extinction training might suppress fear by altering the activity of fear engram cells in DG.

We used an activity-dependent neuronal-tagging transgenic mouse line <sup>13</sup> to indelibly label and manipulate DG granule cells active during either contextual fear acquisition or extinction. We report that extinction training suppresses reactivation of the neurons that were active during fear training (fear acquisition neurons), while causing a different population—putative extinction neurons—to become active. Silencing neurons tagged during fear acquisition decreased fear, whereas silencing neurons tagged during extinction training increased fear after extinction. Conversely, optogenetic stimulation of neurons tagged during fear acquisition increased fear, whereas stimulation of neurons tagged during extinction decreased fear. Our data lead us to hypothesize that fear acquisition and fear extinction are represented by unique ensembles of DG granule cells and that suppression and relapse of fear after extinction are controlled by the activity of these ensembles.

#### **RESULTS**

# Extinction suppresses reactivation of fear acquisition neurons

We used the ArcCreER<sup>T2</sup> transgenic mouse line to tag and manipulate neurons active during either contextual fear conditioning (CFC) acquisition or fear extinction. In these mice, activity of the immediate early gene (IEG) *Arc* drives expression of tamoxifen-dependent CreER<sup>T2</sup> recombinase. An injection of 4-hydroxytamoxifen (4-OHT) transiently activates recombinase activity, thereby permanently tagging Arc-expressing neurons with a reporter (Figure 1A). Previous work demonstrates that (1) ArcCreER<sup>T2</sup> reporter expression closely

mirrors that of endogenous Arc protein, (2) the system detects unique hippocampal ensembles active in different contexts, (3) CFC retrieval is associated with reactivation of neurons tagged during fear acquisition, and (4) silencing neurons tagged during fear acquisition impairs behavioral expression of fear 13,20,21.

We first sought to determine if extinction training affects reactivation of DG fear acquisition neurons. ArcCreER<sup>T2</sup>::channelrhodopsin (ChR2)-eYFP<sup>flx</sup> mice were injected with 4-OHT prior to CFC training to tag active neurons with eYFP. Five days later, one group of mice (Group Ext, n = 8) began extinction training, consisting of ten 3-min exposures to the conditioned context without shock. The other groups remained in the homecage during this time. Mice were then exposed to the conditioned context (No Ext, n = 8) or an alternate context (Alt Ctx, n = 6) to test for conditioned fear (Figure 1B). Freezing to the conditioned context decreased in Group Ext across the 10 extinction sessions (Figure 1C; RM one-way ANOVA,  $F_{(9,63)} = 45.52$ , p < 0.0001). During the test session 2 d after the final day of extinction (and 16 d after training), freezing of Group No Ext exceeded that of Groups Ext and Alt Ctx ( $F_{(2,19)} = 78.99$ , p < 0.0001). Mice were perfused 90 min after the retrieval session. Immunohistochemistry against Arc and eYFP was used to assess reactivation of the eYFP<sup>+</sup> neurons that had been tagged during fear acquisition (Figure 1D). Groups Ext, No Ext, and Alt Ctx displayed similar numbers of eYFP<sup>+</sup> cells (Figure 1E;  $F_{(2,19)} = 1.53$ , p =0.243) and Arc<sup>+</sup> cells (Figure 1F;  $F_{(2,19)} = 1.15$ , p = 0.339) as a percentage of DAPI<sup>+</sup> cells. Reactivation of eYFP+ cells was assessed by dividing the percentage of eYFP+/Arc+ colabeled cells among DAPI<sup>+</sup> cells by the chance percentage ((eYFP<sup>+</sup>/DAPI<sup>+</sup>) \* (Arc<sup>+</sup>/DAPI<sup>+</sup>) \* 100)<sup>22–24</sup>. Reactivation was reduced in Groups Ext and Alt Ctx as compared to Group No Ext (Figure 1G, S1, Table S1;  $F_{(2,19)} = 7.55$ , p = 0.004), demonstrating that extinction suppresses reactivation of DG fear acquisition neurons to a similar extent as exposure to a neutral, alternate context. Furthermore, because the number of Arc+ neurons did not differ among the groups, we infer that retrieval of fear and retrieval of extinction similarly engage the DG but activate distinct neuronal ensembles.

#### Fear retrieval and extinction retrieval reactivate distinct ensembles

To test the hypothesis that retrieval of fear and retrieval of extinction recruit unique DG ensembles, we performed an experiment in which neurons active during fear training or fear extinction were tagged in separate cohorts of mice. ArcCreER<sup>T2</sup>::halorhodopsin (Halo)-eYFP<sup>flx</sup> mice were subjected to CFC and extinction. 4-OHT was injected either immediately after fear acquisition (Acq-Tag; n = 16) or immediately after the  $10^{th}$  day of extinction (Ext-Tag; n = 15). 4-OHT injected immediately after CFC effectively tags fear engram neurons<sup>20</sup> and was performed to prevent the injection from becoming a conditioned stimulus or retrieval cue. Mice were returned to the context either 5 d later for an extinction retrieval test or 28 d later for a spontaneous recovery test (Figure 2A). Freezing to the conditioned context decreased across the extinction sessions in both groups (Figure 2B; RM two-way ANOVA, effect of Session [ $F_{(9,261)} = 69.4$ , p < 0.0001], no effect of Group [F < 1] or Group X Session interaction [F < 1]). Freezing remained low during the extinction retrieval test but was elevated during the spontaneous recovery test (two-way ANOVA, effect of Test [ $F_{(1,27)} = 10.04$ , p = 0.004]). All mice were perfused 90 min after their test session to evaluate the reactivation of fear- and extinction-tagged neurons during low fear (extinction retrieval) and

high fear (spontaneous recovery) states (Figure 2C, 2D). We predicted that extinction-tagged neurons would be reactivated during the 5 d test, whereas fear acquisition-tagged neurons would be reactivated during the 28 d test.

Acq-Tag and Ext-Tag groups displayed similar numbers of eYFP<sup>+</sup> cells (Figure 2E; no effect of Test [F< 1], Group [F(1,27) = 1.87, p = 0.183], or Test X Group interaction [F(1,27) = 1.76, p = 0.196]), and Arc<sup>+</sup> cells (Figure 2F; no effect of Test [F(1,27) = 3.32, p = 0.080], Group [F< 1], or Test X Group interaction [F(1,27) = 1.68, p = 0.206]). During the 5 d test, Ext-Tag mice displayed an increased amount of reactivated eYFP<sup>+</sup>/Arc<sup>+</sup> neurons compared to Acq-Tag mice (Figure 2G, S1; significant Test X Group interaction [F(1,27) = 41.47, p< 0.0001]). The pattern reversed during the 28 d test, with Acq-Tag mice displaying a higher amount of reactivated eYFP<sup>+</sup>/Arc<sup>+</sup> neurons than Ext-Tag mice. These data demonstrate that fear acquisition and extinction activate different populations of DG neurons. Expression of extinction is associated with reactivation of neurons that were active during extinction training, whereas expression of fear during spontaneous recovery is associated with reactivation of neurons that were active during fear acquisition.

### Silencing extinction-tagged neurons impairs extinction retrieval

Next, we evaluated whether the reactivation of extinction-tagged neurons is necessary for expression of extinction. ArcCreER<sup>T2</sup>(+)::Halo-eYFP<sup>flx</sup> and their ArcCreER<sup>T2</sup>(-)::Halo-eYFP<sup>flx</sup> littermates (n = 9 per group) were implanted bilaterally with optical fibers targeting the dorsal DG (Figure 3A, 3B). Two to four weeks later, all mice received CFC followed by extinction training. 4-OHT was injected immediately after the 10<sup>th</sup> day of extinction, thereby expressing Halo in the Arc<sup>+</sup> cells activated during extinction (Figure 3C).

Freezing declined across the extinction sessions and did not differ between ArcCreER<sup>T2</sup>(+) and ArcCreER<sup>T2</sup>(-) mice (Figure 3D; effect of Session  $[F_{(9.135)} = 35.43, p < 0.0001]$ , no effect of Genotype [F < 1] or Genotype X Session interaction [F < 1]). Five days later, mice were returned to the context for a 12-min extinction retrieval test. Extinction-tagged neurons were silenced bilaterally with green light (9-12 mW, continuous) during minutes 0-3 and 6-9 of the test. Silencing extinction-tagged neurons increased freezing during the first light ON epoch, and freezing returned to control levels during the light OFF epoch (Figure 3E, 3F; significant Genotype X Epoch X Light interaction  $[F_{1,14}] = 5.40$ , p = 0.036]). Interestingly, inhibiting extinction-tagged neurons had no effect on freezing during the second light ON epoch (Figure 3E, S2), suggesting that although these cells contribute to the initiation of extinction retrieval, they are not necessary for maintaining retrieval after it has occurred. Silencing extinction-tagged neurons had no effect on freezing in a neutral, alternate context (Figure 3G, 3H; no effect of Genotype  $[F_{(1.14)} = 2.82, p = 0.12]$  or Interactions [p's > 0.20]). To rule out an order effect, one day after the alternate context test, mice were given a second exposure to the conditioning context with light presentations. Silencing extinction-tagged neurons again increased freezing during the first light ON epoch but not during the second epoch (Figure S3).

When mice were returned to the conditioned context 28 d after the alternate context test, silencing extinction-tagged neurons had no effect on spontaneous recovery (Figure 3I, 3J; no effect of Genotype [F < 1] or interactions [F < 1]). These results demonstrate that DG

neurons active during fear extinction are necessary for expression of the extinction memory. Activity of extinction-tagged neurons is not necessary during spontaneous recovery of fear, possibly because spontaneous recovery reflects the failure to reactivate the extinction memory.

# Silencing fear acquisition-tagged neurons reduces spontaneous recovery of fear

Reactivation of hippocampal fear acquisition neurons is necessary for expression of contextual fear  $^{13,25}$ . It is unknown if activity of these neurons is also required for expression of fear after extinction, such as during spontaneous recovery. To address this question, ArcCreER $^{T2}$ (+)::Halo-eYFP $^{flx}$  (n = 7) and ArcCreER $^{T2}$ (-)::Halo-eYFP $^{flx}$  littermates (n = 8) were implanted bilaterally with optical fibers targeting the dorsal DG (Figure 4A, 4B). Two to four weeks later, mice received CFC training followed immediately by an injection of 4-OHT, thereby expressing Halo in fear acquisition neurons, and then a course of extinction training (Figure 4C).

Context fear decreased across the 10 d of extinction, and freezing did not differ between genotypes (Figure 4D; effect of Session  $[F_{(9,117)} = 27.07, p < 0.0001]$ , no effect of Genotype  $[F_{(1,13)} = 1.25, p = 0.284]$  or Genotype X Session interaction  $[F_{(9,117)} = 1.48, p = 0.165]$ ). Five days after the final extinction session, mice were returned to the context for an extinction retrieval test with green light delivered during minutes 0-3 and 6-9. Silencing fear acquisition-tagged neurons had no effect on freezing during light ON epochs of this test (Figure 4E, 4F). Although the Genotype X Epoch X Light interaction reached significance  $(F_{(1,11)} = 5.42, p = .040)$ , the pairwise between-group tests were not significant. The next day, light was delivered during exposure to an alternate context. Silencing fear acquisitiontagged neurons had no effect on freezing behavior (Figure 4G, 4H; no effect of Genotype  $[F_{(1,11)} = 1.60, p = .232]$  or interactions [p's > .07]). In contrast, when mice were returned to conditioning context for a spontaneous recovery test 28 d later, silencing fear acquisitiontagged neurons reduced context fear (Figure 4I, 4J; significant Genotype X Epoch X Light interaction [ $F_{(1,11)} = 10.01$ , p = 0.009]). Freezing in ArcCreER<sup>T2</sup>(+) mice increased during the first light OFF epoch but was reduced again during the second light ON epoch. These results suggest that the neurons active during fear acquisition are not required for extinction retrieval but are necessary for spontaneous recovery of fear.

# Stimulating fear acquisition-tagged neurons potentiates fear whereas stimulating extinction-tagged neurons suppresses fear.

Our findings suggest that reactivation of neurons active during extinction training suppresses fear and the failure of these neurons to reactivate contributes to fear relapse after extinction. To determine whether the activity of extinction neurons is sufficient to suppress fear and attenuate fear relapse, we generated mice expressing the optogenetic neural activator ChR2 in fear acquisition- or extinction-tagged neurons. ArcCreER<sup>T2</sup>(+)::ChR2-eYFP<sup>flx</sup> and their ArcCreER<sup>T2</sup>(-)::ChR2-eYFP<sup>flx</sup> littermates were implanted bilaterally with optical fibers targeting the dorsal DG (Figure 5A, 5B). Two to four weeks later, all mice received CFC and extinction training (Figure 5C). 4-OHT was injected either immediately after fear acquisition (Acq-Tag(+): n = 4; Acq-Tag(-): n = 6) or immediately after the  $10^{th}$  day of extinction (Ext-Tag(+): n = 4; Ext-Tag(-): n = 6), thereby expressing ChR2 in fear acquisition- or

extinction-tagged neurons. We confirmed that 3 min of unilateral blue light (10 Hz, 20 ms pulses, 0.75-0.9 mW) delivered in an alternate context increased the percentage of Arc<sup>+</sup> cells among eYFP<sup>+</sup> cells in Acq-Tag(+) and Ext-Tag(+) mice (Figure 5D, 5E, S4;  $t_{(6)}$  = 4.86, p = 0.003). Groups not expressing ChR2 (Acq-Tag(-) and Ext-Tag(-)) performed similarly in all test sessions and were combined into a single control group.

Five days after fear acquisition and one day before extinction training, mice were tested for 9 min in an alternate context with light delivered during minutes 3-6. This test was performed to confirm that stimulation of fear acquisition-tagged neurons potentiates fear expression. Indeed, stimulating fear acquisition-tagged neurons increased freezing during the light ON epoch (Figure 5G, 5H; significant Group X Light interaction [ $F_{(2,17)} = 3.65$ , p = 0.048]). Because this test occurred prior to 4-OHT administration in Ext-Tag(+) mice, light stimulation failed to alter behavior in these mice.

Context fear decreased across the 10 d of extinction, and freezing did not differ between groups (Figure 5F; effect of Session [ $F_{(9,153)} = 41.55$ , p < 0.0001], no effect of Group [ $F_{(2,17)} = 3.34$ , p = 0.060] or Genotype X Session interaction [F < 1]). Five days after the final extinction session, mice were returned to the context for an extinction retrieval test with light delivered during minutes 3-6. Stimulating extinction-tagged neurons reduced freezing, and stimulating fear acquisition-tagged neurons had no effect on freezing, (Figure 5I, 5J; significant Group X Light interaction [ $F_{(2,17)} = 6.78$ , p = 0.007]). In an open field (OF) test, stimulation of fear acquisition- or extinction-tagged neurons failed to affect center time, center distance, or total distance traveled (Figure S5).

The mice were returned to the context 28 d later for a spontaneous recovery test with light delivered during minutes 0-3 and 6-9. In contrast to previous test sessions, light stimulation was delivered at the beginning of the session to determine if stimulation of extinction-tagged neurons could prevent spontaneous recovery of fear, which is typically strongest at the start of the test session. Stimulating extinction-tagged neurons blocked spontaneous recovery, an effect that persisted even after light termination (Figure 5K, 5L; significant Group X Light interaction [ $F_{(2,17)} = 4.24$ , p = 0.032]). For consistency with the previous light stimulation sessions, mice were given a second exposure to the context on the following day with light delivered in the middle of the session (Figure S6). Stimulating extinction-tagged neurons again reduced freezing, and there was no effect of stimulating fear acquisition-tagged neurons. These results demonstrate that neurons active during fear acquisition and fear extinction have opposing effects on fear behavior. Whereas activity of fear acquisition neurons is sufficient to induce fear, activity of extinction neurons is sufficient to suppress fear.

#### DISCUSSION

We used activity-dependent neuronal tagging to investigate how extinction training influences context fear representations in the DG. Previous research demonstrates that recall of context fear requires reactivation of the DG neurons that were active during fear acquisition 13. Our results demonstrate that extinction training suppresses reactivation of the fear acquisition ensemble and recruits a different ensemble of DG neurons putatively

representing the extinction memory. These fear extinction neurons were more active than fear acquisition neurons during a test session 5 d after extinction, when fear expression was low. In contrast, during a spontaneous recovery test 28 d after extinction, fear acquisition neurons were more active than extinction neurons. Optogenetic manipulations revealed that reactivation of fear extinction neurons is necessary and sufficient for expression of extinction, whereas reactivation of fear acquisition neurons is necessary for spontaneous recovery. Our data suggest that fear acquisition and extinction training have unique ensemble representations in the DG, and competition between these ensembles controls expression of fear after extinction.

The role of the hippocampus in CFC is thought to involve generating context representations that acquire the ability to activate expression of fear through hippocampal interactions with the amygdala<sup>26–28</sup>. Behavioral evidence suggests that changes in context valence are associative in nature, in that they reflect changes in the ability of a stable context representation to evoke fear, rather the generation of new context representations<sup>29</sup>. This view is supported by evidence that rodents can be fear conditioned to a remembered context representation<sup>30</sup> and by evidence that pairing artificial reactivation of a context representation with shock is sufficient to produce fear of a previously neutral context<sup>15,16</sup>. Other evidence, however, suggests that changes in context valence might alter context representations in the hippocampus. In hippocampal CA1, the place fields of some neurons remap during fear conditioning and/or extinction<sup>31</sup>. Fear acquisition and extinction also induce IEG expression in different ensembles of CA1 pyramidal cells<sup>32</sup>. Our results demonstrate that unique ensemble representations of fear and extinction are present upstream in the DG and that the activity of these DG ensembles plays a causal role in expression of fear and extinction. The maintenance of separate fear and extinction representations in the hippocampus appears to provide a mechanism for fear relapse after extinction.

Whereas lesions and nonspecific silencing of DG fail to impair expression of context fear and other learned responses<sup>33–35</sup>, paradoxically, we find that silencing sparse populations of fear acquisition- or extinction-tagged neurons alters fear expression. A recent study using computational simulations of hippocampal function provides a potential resolution to this paradox<sup>9</sup>. The simulations suggest that when only engram cells are silenced, the remaining activity in DG is contextually inappropriate and interferes with recall. Large-scale DG silencing minimizes this inappropriate activity, allowing accurate recall through cortical inputs to CA3 (bypassing DG), which are sufficient for correct recall under some circumstances. This model leads us to speculate that silencing extinction cells could lead to increased activity of fear acquisition neurons, while silencing fear acquisition neurons could lead to increased activity of extinction neurons. Testing this hypothesis will require new methods that enable tagging of multiple ensembles in the same mice.

Although our conclusions are different, our results are not inconsistent with a recent study showing that reactivation of the DG neurons active during recall of a remote fear memory contributes to fear extinction<sup>36</sup>. In that study, neurons were tagged during a shock-free test session, which presumably evoked both fear recall and extinction. Our findings suggest that such a procedure will yield a heterogeneous tagged population. It is also possible that the

hippocampal mechanisms of extinction vary depending on whether the memory being extinguished is recent (our study) or remote<sup>36,37</sup>.

The existence of distinct populations of fear acquisition and extinction neurons in the hippocampus parallels what has been observed in the PFC and amygdala. The basolateral amygdala (BLA) contains unique populations of excitatory neurons that respond to extinguished or non-extinguished fear cues and project to different targets<sup>38,39</sup>. In the PFC, neurons in the prelimbic subdivision preferentially respond to fear-associated cues, while those in the infralimbic subdivision respond to extinguished cues<sup>40–42</sup>. Whereas the opposing cell types in the BLA and PFC can be distinguished based on molecular profiles and connectivity<sup>43–45</sup>, it remains to be determined whether the populations in the DG are similarly differentiated. The fear and extinction-tagged neurons we observed in DG were intermingled and could not be distinguished from each other based on morphology or location. Nevertheless, we speculate that activity of DG fear acquisition neurons induces fear expression through interactions with ventral CA1 projections to the BLA<sup>46</sup>. The activity of extinction neurons might reduce fear by interfering with recall of hippocampal fear representations or by activating fear-suppressive circuitry in the amygdala, PFC, or another region through polysynaptic mechanisms<sup>8,47,48</sup>. It is also possible that fear and extinction ensembles are associated with different oscillation frequencies, similar to what has been shown in PFC-amygdala circuits<sup>49</sup>.

In conclusion, we demonstrated that fear and extinction memories have distinct ensemble representations in the DG. We hypothesize that competition between these ensemble representations in the hippocampus determines whether fear is expressed or suppressed after extinction training. Because fear relapse represents a significant challenge for the treatment of anxiety and fear disorders<sup>50</sup>, interventions that potentiate activity of the hippocampal extinction ensemble or suppress reactivation of the fear acquisition ensemble may be of therapeutic value.

#### **METHODS**

#### **Subjects**

Adult male and female ArcCreER<sup>T2</sup>::Halo-eYFP<sup>flx</sup> and ArcCreER<sup>T2</sup>::ChR2-eYFP<sup>flx</sup> mice aged two to seven months were used for all experiments (Table S2). ArcCreER<sup>T2</sup>::Halo-eYFP<sup>flx</sup> and ArcCreER<sup>T2</sup>::ChR2-eYFP<sup>flx</sup> mice were generated by breeding heterozygous (+/-) ArcCreER<sup>T2</sup> mice<sup>13</sup> with homozygous (+/+) Rosa26-CAG-stop<sup>flx</sup>-eNpHR3.0-eYFP (Ai39) or Rosa26-CAG-stop<sup>flx</sup>-ChR2(H134R)-eYFP (Ai32) mice<sup>51</sup>, originally purchased from The Jackson Laboratories. These crosses generated ArcCreER<sup>T2</sup>(+) or (-) mice heterozygous for either the Halo-eYFP<sup>flx</sup> or ChR2-eYFP<sup>flx</sup> allele. Mice were housed with their littermates in groups of one to five in plastic cages with woodchip bedding and maintained on a 12 h light/dark cycle (7:00-19:00) in a temperature- and humidity-controlled vivarium. Food and water were provided *ad libitum*. Experiments were conducted during the light phase. Mice were randomly assigned to groups before the start of each experiment. As sex by genotype statistical interactions were not present, male and female data were aggregated. A cohort of 6 mice shipped from Columbia and potentially exposed to extreme heat did not extinguish and was removed from further analysis. Five mice were

found to have no eYFP expression suggestive of a genotyping error and were removed from further analysis. No statistical methods were used to predetermine sample size, but the sample sizes were based on those in previously published studies <sup>13,15</sup>. All procedures were approved by the University of Texas at Austin Institutional Animal Care and Use Committee.

#### Surgery

For optogenetic experiments, mice received stereotaxic surgery 2-4 weeks prior to behavioral experimentation. Mice were anesthetized with 3% isoflurane (1.5 L/min) vaporized in pure oxygen, head-fixed in a stereotaxic surgical frame, and maintained under anesthesia with 1.5% isoflurane concentration (0.75 L/min). Ophthalmic ointment was applied to prevent the eyes from drying. A skin incision was made to expose the skull. The skull was scoured with an acidic gel etchant and a layer of OptiBond epoxy (Kerr Dental) was applied. Fiber optic cannulas were bilaterally implanted at a 20° angle targeting the dorsal DG (from bregma: A/P: –2.0 mm; M/L:  $\pm 1.3$  mm; D/V: –1.5 mm). Optic fiber implants were constructed using previously published methods<sup>52</sup>, using 1.25 mm ceramic ferrules (Kientec Systems) and 200  $\mu$ m core, 0.39 numerical aperture multimodal fiber (ThorLabs). The fibers were secured in place with a layer of dental cement (Bosworth). Following the surgery, mice were injected subcutaneously with ketoprofen (5 mg/kg) in sterile saline (0.9%) to provide analgesia.

#### **Neuronal Tagging**

Recombination was induced with 4-hydroxytamoxifen (4-OHT; Sigma). 4-OHT was dissolved by sonication in 10% EtOH / 90% sunflower seed oil at 10 mg/mL. Mice were transported from the vivarium to an adjacent holding room at least 3 hours prior to 4-OHT injections to minimize transportation-induced IEG activity. Activity-dependent neuronal tagging was induced by a single intraperitoneal injection of 4-OHT (55 mg/kg) administered either immediately following the CFC session for fear acquisition-tagged mice, or immediately following the 10<sup>th</sup> extinction session for extinction-tagged mice. Following the 4-OHT injection, mice were moved to an isolated room and dark housed for 3 d, which helps minimize non-specific labeling 13. Mice were then returned to the vivarium with a regular 12h light/dark cycle for the remainder of the experiment.

#### **Optogenetics**

Cranial implants were bilaterally connected via fiber optic patch cables to a light source interfaced with a FC/PC rotary joint (Doric Lenses). Photoinhibition was performed bilaterally with a 140 mW, 532 nm laser (Shanghai Dream Lasers Technology) delivered continuously during light ON epochs with an intensity of 9-12 mW at the end of the fiber optic implant. ChR2 photostimulation was performed unilaterally with a 17.2 mW, 470 nm LED (ThorLabs) delivered in 20 ms pulses at 10 Hz with an intensity of 0.75-0.9 mW at the end of the fiber optic implant.

#### **Contextual Fear Conditioning**

All mice were handled for 1-2 min per day for 4-5 d prior to behavioral testing. During handling sessions, mice were habituated to fiber optic cables by allowing them to explore a novel cage while attached to patch cables. Subjects were transported from the vivarium to a holding room adjacent to the test room at least one hour before experimentation. Mice were moved individually to and from the conditioning room in an opaque container. The transport containers were cleaned with a 70% EtOH solution between uses. Mice were connected to fiber optic patch cables for every behavioral session.

Behavioral testing occurred in  $30.5 \times 24 \times 21$  cm conditioning chambers (Med Associates), with two aluminum side walls, a Plexiglas door and ceiling, and a white vinyl back wall. Chambers were contained within a larger, sound-attenuating chamber equipped with a fan to provide ~65 dB ambient noise. An overhead white light illuminated the chamber continuously throughout the procedures.

Two distinct contexts were used – a conditioning context and an alternate context. The conditioning context contained a straight stainless-steel rod floor (36 rods, spaced 8mm from center to center), was cleaned with a 70% EtOH solution between uses and was scented with 1% acetic acid solution in the waste tray below the floor. The alternate context occurred in the same chambers and contained a flat plastic floor covered in woodchip bedding and a curved plastic insert along the back wall. The chamber was cleaned and scented with Clorox disinfectant wipes.

For fear conditioning, three 2-s 0.75 mA scrambled foot shocks were delivered through the floor 180, 240, and 300 s after the mice were placed in the context. Mice were removed 30 s after the final footshock and returned to their homecage.

All behavioral sessions were video recorded at 30 frames/s using a near-infrared camera mounted to the interior door of the chamber. The videos were manually scored for freezing behavior by an investigator blind to the experimental conditions. Freezing was defined as the absence of movement, with the exception of those related to breathing.

#### Extinction

Extinction sessions consisted typically of 5 min unreinforced exposures to the conditioning context once per day for 10 d (exception: 3 min sessions for Figure 1B–G).

#### **Tests and Optogenetic Manipulation**

For test sessions in cell reactivation experiments, mice were placed in the conditioning or alternate context for 5 min. Tests for silencing experiments took place in either context for 12 min. Light ON epochs occurred during minutes 0-3 and 6-9. Tests for ChR2 stimulation experiments took place in either context for 9 min. The light ON epoch occurred during minutes 3-6 for the extinction retrieval and alternate context tests. Light ON epochs occurred during minutes 0-3 and 6-9 for the spontaneous recovery test. Spontaneous recovery tests occurred 27-28 d after the last test session. The photostimulation-induced IEG experiment took place in the alternate context for 9 min. For mice that received photostimulation, light was delivered during minutes 3-6.

#### Open Field

The OF was a  $40 \times 40$  cm arena with opaque plastic walls 35 cm high. Overhead lights provided 80 lux illumination measured in the center of the arena. Mice were connected to fiber optic patch cables and placed into the arena for 15 min. Photostimulation consisted of a 5 min epoch from minutes 5-10. Light parameters were identical to those reported for tests. The center was defined as an  $18.5 \times 18.5$  cm zone in the center of the arena. The behavior was recorded using a ceiling-mounted digital camera and analyzed with video-tracing software (ANY-Maze).

#### Tissue Preparation and Immunohistochemistry

Ninety minutes after the CFC test, mice were deeply anesthetized with ketamine/xylazine (150/15 mg/kg) and transcardially perfused with 1× phosphate buffered saline (PBS), followed by 4% paraformaldehyde (PFA) in 1× PBS. Brains were extracted and post-fixed overnight at 4°C in 4% PFA and then transferred to a 30% sucrose in 1× PBS at 4°C for two days. Thirty-five  $\mu$ m coronal sections were collected on a cryostat and stored in cryoprotectant at -20°C.

For immunohistochemistry, sections were washed in  $1 \times PBS$  and blocked at room temperature (RT) for 2 h in 10% normal donkey serum (NDS) in  $1 \times PBS$  with 0.5% Triton-X (PBS-T). Sections were incubated with primary antibodies (1:2,000 rabbit anti-Arc (Synaptic Systems); 1:500 chicken anti-GFP (Abcam); 1:1000 rabbit anti-c-Fos (Millipore)) diluted in 5% NDS in  $1 \times PBS$ -T overnight at 4°C. Sections were rinsed in  $1 \times PBS$ -T and incubated in secondary antibodies (Jackson ImmunoResearch; 1:500 donkey anti-rabbit Cy3; 1:500 biotinylated donkey anti-chicken) in  $1 \times PBS$ -T for 2 h at RT. Sections were rinsed in  $1 \times PBS$ -T and incubated in tertiary antibody (1:250 avidin Cy-2 (Jackson ImmunoResearch)) and 1:1000 DAPI for 1 h at RT. Sections were washed in  $1 \times PBS$ , mounted onto slides, and coverslipped with ProLong Gold (Invitrogen). See the Life Sciences Reporting Summary for additional information.

#### **Imaging and Quantification**

For Figures 1–2, immunoreactive Arc<sup>+</sup> cells in the granule cell layer of the DG from every 6<sup>th</sup> section throughout the entire bilateral rostrocaudal axis of the hippocampus were counted exhaustively under fluorescent illumination (Zeiss Axio Imager M2) with a 40× objective. Every Arc<sup>+</sup> cell was evaluated for eYFP<sup>+</sup> immunoreactivity to obtain eYFP<sup>+</sup>/Arc<sup>+</sup> colabeled cell counts. For eYFP<sup>+</sup> cell estimates, fluorescent confocal images were obtained (Leica DM6000 CFS) with a 40x objective from 4 locations sampled throughout the rostrocaudal axis of the DG. Each image stack was acquired at 1 µm optical sections, and eYFP<sup>+</sup> cells were quantified from the z-stack images using Image-J. To obtain DAPI<sup>+</sup> estimates, the DG volume from each mouse was measured using Stereo Investigator. Next, a DAPI<sup>+</sup> density estimate was obtained by counting DAPI<sup>+</sup> cells in the DG from 3 image stacks obtained with a 63x objective of 4 ArcCreER<sup>T2</sup>::ChR2-eYFP<sup>flx</sup> mice. Finally, the DAPI<sup>+</sup> density estimate was multiplied by the DG volume. For the photostimulation-induced IEG experiment, eYFP<sup>+</sup> and Arc<sup>+</sup> or c-Fos<sup>+</sup> cells in the dorsal DG were quantified from 3 image stacks obtained with a 63× objective from a section near the site of the fiber optic implant. For CA3 counts, c-Fos<sup>+</sup> and eYFP<sup>+</sup> cells were counted exhaustively from every 6<sup>th</sup>

section throughout the dorsal hippocampus. An investigator blind to the treatment status performed all cell counts. The amount of reactivation was normalized for chance overlap by dividing the percentage of co-labeled<sup>+</sup> among DAPI<sup>+</sup> by chance ((eYFP<sup>+</sup>/DAPI<sup>+</sup>) \*  $(Arc^+/DAPI^+)$  \*  $(100)^{22-24}$ .

#### **Statistical Analysis**

Data were analyzed using two-sided *t*-tests, the Mann-Whitney U test, or ANOVA, using repeated measures when appropriate. Significant ANOVAs were followed *post hoc* Sidak's test for multiple comparisons. Data distribution was assumed to be normal but this was not formally tested. Data analysis was performed on Prism 6 (GraphPad Software) or JMP (SAS Institute). The  $\alpha$  value was set at 0.05 for all analyses. All data are presented as mean  $\pm$  1 s.e.m.

#### **Data Availability**

All relevant data supporting the findings of this study are available from the corresponding author upon reasonable request.

# **Supplementary Material**

Refer to Web version on PubMed Central for supplementary material.

#### ACKNOWLEDGMENTS

We thank J. Dunsmoor for comments on the manuscript. A.F.L. was supported by NIH F31 MH111243 and NIH T32 MH106454. S.L.S. was supported by PD/BD/128076/2016 from the Portuguese Foundation for Science and Technology. Research supported by NIH DP5 OD017908 and New York Stem Cell Science (NYSTEM) C-029157 to C.A.D., NIH R01 MH102595 and NIH R21 EY026446 to M.R.D.

#### REFERENCES

- Powers MB, Halpern JM, Ferenschak MP, Gillihan SJ & Foa EB A meta-analytic review of prolonged exposure for posttraumatic stress disorder. Clin. Psychol. Rev 30, 635–641 (2010). [PubMed: 20546985]
- 2. Milad MR & Quirk GJ Fear extinction as a model for translational neuroscience: Ten years of progress. Annu. Rev. Psychol 63, 129–151 (2012). [PubMed: 22129456]
- 3. Rescorla RA Spontaneous recovery. Learn. Mem 11, 501–509 (2004). [PubMed: 15466300]
- 4. Pavlov IP Conditioned Reflexes. (Oxford University Press, 1927).
- 5. Bouton ME, Westbrook RF, Corcoran KA & Maren S Contextual and temporal modulation of extinction: Behavioral and biological mechanisms. Biol. Psychiatry 60, 352–360 (2006). [PubMed: 16616731]
- 6. Myers KM & Davis M Mechanisms of fear extinction. Mol. Psychiatry 12, 120–150 (2007). [PubMed: 17160066]
- 7. Tovote P, Fadok JP & Lüthi A Neuronal circuits for fear and anxiety. Nat. Rev. Neurosci 16, 317–331 (2015). [PubMed: 25991441]
- 8. Trouche S, Sasaki JM, Tu T & Reijmers LG Fear extinction causes target-specific remodeling of perisomatic inhibitory synapses. Neuron 80, 1054–1065 (2013). [PubMed: 24183705]
- 9. Bernier BE et al. Dentate gyrus contributes to retrieval as well as encoding: Evidence from context fear conditioning, recall, and extinction. J. Neurosci 37, 6359–6371 (2017). [PubMed: 28546308]
- Corcoran KA & Maren S Hippocampal inactivation disrupts contextual retrieval of fear memory after extinction. J. Neurosci 21, 1720–1726 (2001). [PubMed: 11222661]

11. Corcoran KA & Maren S Factors regulating the effects of hippocampal inactivation on renewal of conditional fear after extinction. Learn. Mem 11, 598–603 (2004). [PubMed: 15466314]

- 12. Hobin JA, Ji J & Maren S Ventral hippocampal muscimol disrupts context-specific fear memory retrieval after extinction in rats. Hippocampus 16, 174–182 (2006). [PubMed: 16358312]
- 13. Denny CA et al. Hippocampal memory traces are differentially modulated by experience, time, and adult neurogenesis. Neuron 83, 189–201 (2014). [PubMed: 24991962]
- 14. Liu X et al. Optogenetic stimulation of a hippocampal engram activates fear memory recall. Nature 484, 381–385 (2012). [PubMed: 22441246]
- 15. Ramirez S et al. Creating a false memory in the hippocampus. Science 341, 387–391 (2013). [PubMed: 23888038]
- 16. Ryan TJ, Roy DS, Pignatelli M, Arons A & Tonegawa S Engram cells retain memory under retrograde amnesia. Science 348, 1007–1013 (2015). [PubMed: 26023136]
- 17. Redondo RL et al. Bidirectional switch of the valence associated with a hippocampal contextual memory engram. Nature 513, 426–430 (2014). [PubMed: 25162525]
- 18. Kitamura T et al. Engrams and circuits crucial for systems consolidation of a memory. Science 356, 73–78 (2017). [PubMed: 28386011]
- 19. Roy DS et al. Memory retrieval by activating engram cells in mouse models of early Alzheimer's disease. Nature 531, 508–512 (2016). [PubMed: 26982728]
- 20. Cazzulino AS, Martinez R, Tomm NK & Denny CA Improved specificity of hippocampal memory trace labeling. Hippocampus 26, 752–762 (2016). [PubMed: 26662713]
- 21. Sun Q et al. Proximodistal heterogeneity of hippocampal CA3 pyramidal neuron intrinsic properties, connectivity, and reactivation during memory recall. Neuron 95, 656–672.e3 (2017). [PubMed: 28772124]
- Tayler KK, Tanaka KZ, Reijmers LG & Wiltgen BJ Reactivation of neural ensembles during the retrieval of recent and remote memory. Curr. Biol 23, 99–106 (2013). [PubMed: 23246402]
- 23. Cai DJ et al. A shared neural ensemble links distinct contextual memories encoded close in time. Nature 534, 115–118 (2016). [PubMed: 27251287]
- Reijmers LG, Perkins BL, Matsuo N & Mayford M Localization of a stable neural correlate of associative memory. Science 317, 1230–1233 (2007). [PubMed: 17761885]
- 25. Tanaka KZ et al. Cortical representations are reinstated by the hippocampus during memory retrieval. Neuron 84, 347–354 (2014). [PubMed: 25308331]
- Fanselow MS Contextual fear, gestalt memories, and the hippocampus. Behav. Brain Res 110, 73– 81 (2000). [PubMed: 10802305]
- 27. Rudy JW & O'Reilly RC Conjunctive representations, the hippocampus, and contextual fear conditioning. Cogn. Affect. Behav. Neurosci 1, 66–82 (2001). [PubMed: 12467104]
- 28. O'Reilly RC & Rudy JW Conjunctive representations in learning and memory: principles of cortical and hippocampal function. Psychol. Rev 108, 311–345 (2001). [PubMed: 11381832]
- 29. Rudy JW, Huff NC & Matus-Amat P Understanding contextual fear conditioning: insights from a two-process model. Neurosci. Biobehav. Rev 28, 675–685 (2004). [PubMed: 15555677]
- 30. Rudy JW, Barrientos RM & O'Reilly RC Hippocampal formation supports conditioning to memory of a context. Behav. Neurosci 116, 530–538 (2002). [PubMed: 12148921]
- Wang ME, Yuan RK, Keinath AT, Ramos Alvarez MM & Muzzio IA Extinction of learned fear induces hippocampal place cell remapping. J. Neurosci 35, 9122–9136 (2015). [PubMed: 26085635]
- 32. Tronson NC et al. Segregated populations of hippocampal principal CA1 neurons mediating conditioning and extinction of contextual fear. J. Neurosci 29, 3387–3394 (2009). [PubMed: 19295145]
- 33. Kheirbek MA et al. Differential control of learning and anxiety along the dorsoventral axis of the dentate gyrus. Neuron 77, 955–968 (2013). [PubMed: 23473324]
- 34. Lassalle J-M, Bataille T & Halley H Reversible inactivation of the hippocampal mossy fiber synapses in mice impairs spatial learning, but neither consolidation nor memory retrieval, in the Morris navigation task. Neurobiol. Learn. Mem 73, 243–257 (2000). [PubMed: 10775494]

35. Madroñal N et al. Rapid erasure of hippocampal memory following inhibition of dentate gyrus granule cells. Nat. Commun 7, 10923 (2016). [PubMed: 26988806]

- 36. Khalaf O et al. Reactivation of recall-induced neurons contributes to remote fear memory attenuation. Science 360, 1239–1242 (2018). [PubMed: 29903974]
- 37. Corcoran KA, Leaderbrand K & Radulovic J Extinction of remotely acquired fear depends on an inhibitory NR2B/PKA pathway in the retrosplenial cortex. J. Neurosci 33, 19492–19498 (2013). [PubMed: 24336715]
- 38. Herry C et al. Switching on and off fear by distinct neuronal circuits. Nature 454, 600–606 (2008). [PubMed: 18615015]
- 39. Grewe BF et al. Neural ensemble dynamics underlying a long-term associative memory. Nature 543, 670–675 (2017). [PubMed: 28329757]
- 40. Milad MR & Quirk GJ Neurons in medial prefrontal cortex signal memory for fear extinction. Nature 420, 70–74 (2002). [PubMed: 12422216]
- 41. Sotres-Bayon F, Sierra-Mercado D, Pardilla-Delgado E & Quirk GJ Gating of fear in prelimbic cortex by hippocampal and amygdala inputs. Neuron 76, 804–812 (2012). [PubMed: 23177964]
- Burgos-Robles A, Vidal-Gonzalez I & Quirk GJ Sustained conditioned responses in prelimbic prefrontal neurons are correlated with fear expression and extinction failure. J. Neurosci 29, 8474– 8482 (2009). [PubMed: 19571138]
- 43. Namburi P et al. A circuit mechanism for differentiating positive and negative associations. Nature 520, 675–678 (2015). [PubMed: 25925480]
- 44. Kim J, Pignatelli M, Xu S, Itohara S & Tonegawa S Antagonistic negative and positive neurons of the basolateral amygdala. Nat. Neurosci 19, 1636–1646 (2016). [PubMed: 27749826]
- 45. Senn V et al. Long-range connectivity defines behavioral specificity of amygdala neurons. Neuron 81, 428–437 (2014). [PubMed: 24462103]
- 46. Xu C et al. Distinct hippocampal pathways mediate dissociable roles of context in memory retrieval. Cell 167, 961–972.e16 (2016). [PubMed: 27773481]
- 47. Marek R et al. Hippocampus-driven feed-forward inhibition of the prefrontal cortex mediates relapse of extinguished fear. Nat. Neurosci 463, 36 (2018).
- 48. Klavir O, Prigge M, Sarel A, Paz R & Yizhar O Manipulating fear associations via optogenetic modulation of amygdala inputs to prefrontal cortex. Nat. Neurosci 20, 836–844 (2017). [PubMed: 28288126]
- 49. Davis P, Zaki Y, Maguire J & Reijmers LG Cellular and oscillatory substrates of fear extinction learning. Nat. Neurosci 20, 1624–1633 (2017). [PubMed: 28967909]
- 50. Vervliet B, Craske MG & Hermans D Fear extinction and relapse: State of the art. Annu. Rev. Clin. Psychol 9, 215–248 (2013). [PubMed: 23537484]
- 51. Madisen L et al. A toolbox of Cre-dependent optogenetic transgenic mice for light-induced activation and silencing. Nat. Neurosci 15, 793–802 (2012). [PubMed: 22446880]
- 52. Sparta DR et al. Construction of implantable optical fibers for long-term optogenetic manipulation of neural circuits. Nat. Protoc 7, 12–23 (2011). [PubMed: 22157972]

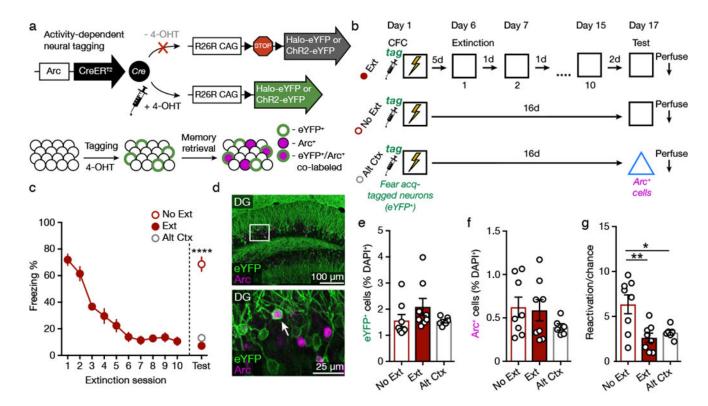


Figure 1. Extinction Suppresses Reactivation of Fear Acquisition-Tagged Neurons.

- (A) Genetic design. Administration of 4-OHT to ArcCreER<sup>T2</sup> mice activates permanent expression of a reporter in neurons active around the time of the injection.
- (B) Experimental design.
- (C) Freezing behavior declined across the 3 min extinction sessions. Group No Ext had significantly higher freezer than Groups Ext and Alt Ctx during a retrieval test (one-way ANOVA with Sidak's post hoc test,  $F_{(2,19)} = 78.99$ , p < 0.0001.
- (D) Representative image of eYFP<sup>+</sup> and Arc<sup>+</sup> immunofluorescence in the DG. White box indicates area of magnification (bottom). White arrowhead denotes co-labeled eYFP<sup>+</sup>/Arc<sup>+</sup> cell.
- (E) Groups Ext, No Ext, and Alt Ctx displayed a similar percentage of eYFP<sup>+</sup> cells among DAPI<sup>+</sup> cells.
- (F) Fear retrieval, extinction retrieval, and alternate context exposure activated a similar percentage of Arc<sup>+</sup> cells among DAPI<sup>+</sup> cells.
- (G) The number of eYFP<sup>+</sup>/Arc<sup>+</sup> reactivated cells was significantly higher in the Group No Ext as compared to Groups Ext and Alt Ctx. Reactivation was calculated as the percentage of eYFP<sup>+</sup>/Arc<sup>+</sup> cells divided by the chance percentage (one-way ANOVA with Sidak's post hoc test,  $F_{(2.19)} = 7.55$ , p = 0.004).

No Ext: n = 8 mice; Ext: n = 8 mice; Alt Ctx: n = 6 mice. Data are means  $\pm$  s.e.m. \* p < 0.05, \*\* p < 0.01, \*\*\*\* p < 0.0001.

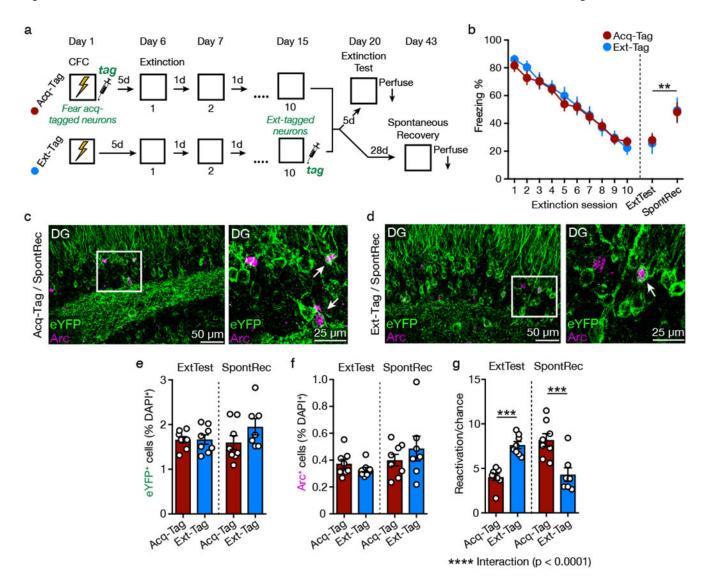


Figure 2. Fear Retrieval and Extinction Retrieval Reactivate Distinct Neural Ensembles in the DG.

- (A) Experimental design.
- (B) Freezing behavior declined across the 5 min extinction sessions. Freezing remained low in mice tested 5 d after extinction (extinction retrieval) but was increased in mice tested 28 d after extinction (spontaneous recovery) (two-way ANOVA, effect of Test,  $F_{(1,27)} = 10.04$ , p = 0.004).
- (C and D) Representative images of eYFP $^+$  and Arc $^+$  immunofluorescence in the DG of Acq-Tag (C) and Ext-Tag (D) mice during a spontaneous recovery test. White box indicates area of magnification (right). White arrowhead denotes co-labeled eYFP $^+$ /Arc $^+$  cells.
- (E) Acq-Tag and Ext-Tag groups displayed a similar percentage of eYFP<sup>+</sup> cells among DAPI <sup>+</sup> cells.
- (F) Extinction retrieval and spontaneous recovery activated a similar percentage of Arc<sup>+</sup> cells among DAPI<sup>+</sup> cells.

(G) Ext-Tag mice displayed a significantly greater percentage of reactivated eYFP<sup>+</sup>/Arc<sup>+</sup> cells (relative to chance) than Acq-Tag mice during an extinction retrieval test. The pattern reversed during the spontaneous recovery test (two-way ANOVA with Sidak's post hoc test, significant Test X Group interaction,  $F_{(1,27)} = 41.47$ , p < 0.0001).

Acq-Tag/ExtTest: n = 8 mice; Ext-Tag/ExtTest: n = 8 mice; Acq-Tag/SpontRec: n = 8 mice; Ext-Tag/SpontRec: n = 7 mice. Data are means  $\pm$  s.e.m. \*\* p < 0.01, \*\*\* p < 0.001, \*\*\*\* p < 0.0001.

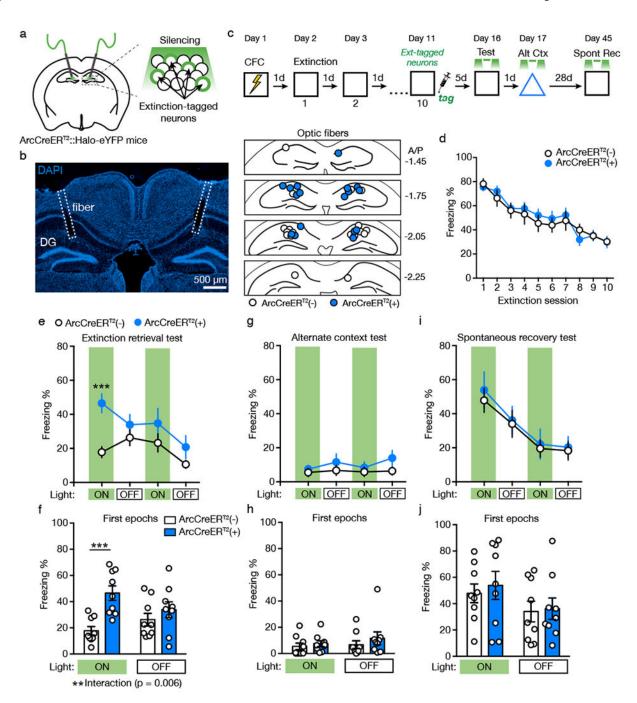


Figure 3. Silencing Extinction-tagged Neurons Impairs Extinction Retrieval.

- (A) Schematic for silencing extinction-tagged neurons in the DG of ArcCreER<sup>T2</sup>::Halo-eYFP<sup>flx</sup> mice.
- (B) Left: Example of optic fiber placement. Right: Coronal figures showing fiber implantation sites in the dorsal DG.
- (C) Experimental design. Light was delivered during context tests at minutes 0-3 and 6-9.
- (D) Freezing declined across the 5 min extinction sessions and did not differ between  $ArcCreER^{T2}(+)$  and  $ArcCreER^{T2}(-)$  mice.

(E) Freezing behavior during the extinction retrieval test (three-way RM ANOVA, significant Genotype X Epoch X Light interaction,  $F_{(1,14)} = 5.40$ , p = 0.036).

- (F) In the extinction retrieval test, silencing extinction-tagged neurons increased freezing in ArcCreER<sup>T2</sup>(+) mice during the first light ON epoch compared to ArcCreER<sup>T2</sup>(-) mice (two-way RM ANOVA with Sidak's post hoc test, significant Genotype X Light interaction,  $F_{(1.16)} = 10.24$ , p = 0.006).
- (G) Freezing behavior during the alternate context test.
- (H) In the alternate context test, silencing extinction-tagged neurons had no effect on freezing behavior during the first light ON and OFF epochs.
- (I) Freezing behavior during the spontaneous recovery test.
- (J) In the spontaneous recovery test, silencing extinction-tagged neurons had no effect on freezing behavior during the first light ON and OFF epochs.

ArcCreER<sup>T2</sup>(-): n = 9 mice; ArcCreER<sup>T2</sup>(+): n = 9 mice. Data are means  $\pm$  s.e.m. \*\* p < 0.01, \*\*\* p < 0.001.

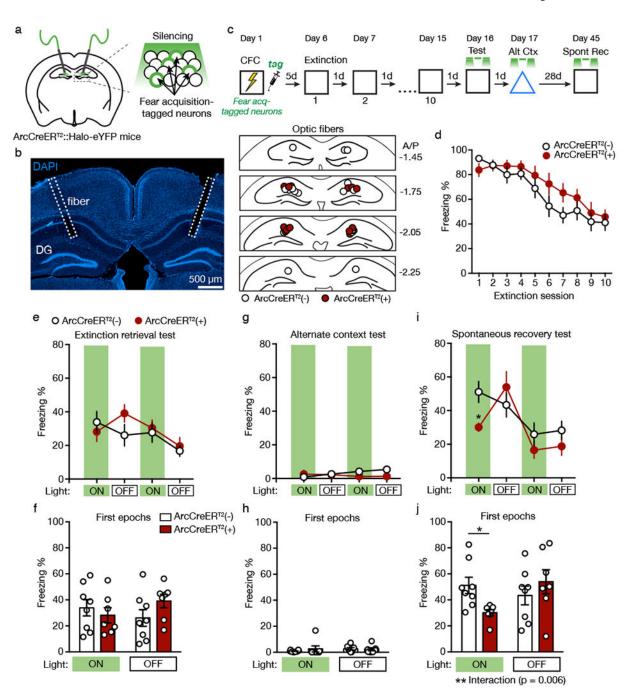


Figure 4. Silencing Fear Acquisition-tagged Neurons Reduces Spontaneous Recovery of Fear.

- (A) Schematic for silencing fear acquisition-tagged neurons in the DG of  $ArcCreER^{T2}$ ::Halo-eYFP<sup>flx</sup> mice.
- (B) Left: Example of optic fiber placement. Right: Coronal figures showing fiber implantation sites in the dorsal DG.
- (C) Experimental design. Light was delivered during context tests at minutes 0-3 and 6-9.
- (D) Freezing declined across the 5 min extinction sessions and did not differ between  $ArcCreER^{T2}(+)$  and  $ArcCreER^{T2}(-)$  mice.

- (E) Freezing behavior during the extinction retrieval test.
- (F) In the extinction retrieval test, silencing fear acquisition-tagged neurons had no effect on freezing behavior during the first light ON and OFF epochs.
- (G) Freezing behavior during the alternate context test.
- (H) In the alternate context test, silencing fear acquisition-tagged neurons had no effect on freezing behavior during the first light ON and OFF epochs.
- (I) Freezing behavior during the spontaneous recovery test (three-way RM ANOVA, significant Genotype X Epoch X Light interaction,  $F_{(1,11)} = 10.01$ , p = 0.009).
- (J) In the spontaneous recovery test, silencing fear acquisition-tagged neurons reduced freezing behavior in  $ArcCreER^{T2}(+)$  mice during the first light ON epoch compared to  $ArcCreER^{T2}(-)$  mice (two-way RM ANOVA with Sidak's post hoc test,  $F_{(1,13)} = 10.68$ , p = 0.006).

ArcCreER<sup>T2</sup>(-): n = 8 mice; ArcCreER<sup>T2</sup>(+): n = 7 mice. Data are means  $\pm$  s.e.m. \* p < 0.05, \*\* p < 0.01

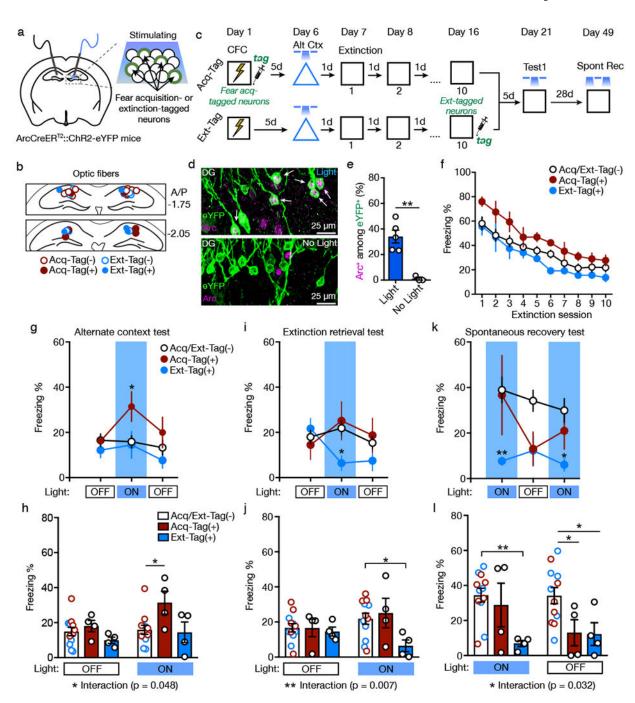


Figure 5. Stimulating Fear Acquisition-tagged Neurons Potentiates Fear whereas Stimulating Extinction-tagged Neurons Suppresses Fear.

- (A) Schematic for stimulating fear acquisition- or extinction-tagged neurons in the DG of  $ArcCreER^{T2}$ :: $ChR2-eYFP^{flx}$  mice.
- (B) Coronal figures showing fiber implantation sites in the dorsal DG.
- (C) Experimental design. Light was delivered during alternate context and extinction tests during at minutes 3-6. For the spontaneous recovery test, light was delivered during minutes 0-3 and 6-9.

(D) Representative images of eYFP $^+$  and Arc $^+$  immunofluorescence in the DG of Acq-Tag( $^+$ ) mice presented with light (top) or without light (bottom). White arrowheads denote co-labeled eYFP $^+$ /Arc $^+$  cells.

- (E) Light increased the percentage of Arc<sup>+</sup> cells among eYFP<sup>+</sup> cells in Acq-Tag(+) and Ext-Tag(+) mice (two-sided *t*-test,  $t_{(6)} = 4.86$ , p = 0.003). White arrowheads denote co-labeled eYFP<sup>+</sup>/Arc<sup>+</sup> cells. Light: n = 5 mice; No Light: n = 3 mice.
- (F) Freezing declined across the 5 min extinction sessions and did not differ between groups.
- (G) Freezing behavior during the alternate context test (two-way RM ANOVA with Sidak's post hoc test, effect of Light,  $F_{(2.34)} = 3.31$ , p = 0.048).
- (H) In the alternate context test, stimulating fear acquisition-tagged neurons increased freezing during the light ON epoch in Acq-Tag(+) mice compared to Acq/Ext-Tag(-) mice. Light OFF epochs averaged (two-way RM ANOVA with Sidak's post hoc test, significant Group X Light interaction,  $F_{(2,17)} = 3.65$ , p = 0.048).
- (I) Freezing behavior during the extinction retrieval test (two-way RM ANOVA with Sidak's post hoc test, significant Group X Light interaction,  $F_{(4.34)} = 3.09$ , p = 0.029).
- (J) In the extinction retrieval test, stimulating extinction-tagged neurons decreased freezing during the light ON epoch in Ext-Tag(+) mice compared to Acq/Ext-Tag(-) mice. Light OFF epochs averaged (two-way RM ANOVA with Sidak's post hoc test, significant Group X Light interaction,  $F_{(2,17)} = 6.78$ , p = 0.007).
- (K) Freezing behavior during the spontaneous recovery test (two-way RM ANOVA with Sidak's post hoc test, effect of Group,  $F_{(2.17)} = 4.92$ , p = 0.021).
- (L) In the spontaneous recovery test, stimulating extinction-tagged neurons reduced freezing during the light ON epochs in Ext-Tag(+) mice compared to Acq/Ext-Tag(-) mice. Light ON epochs averaged (two-way RM ANOVA with Sidak's post hoc test, significant Group X Light interaction,  $F_{(2,17)} = 4.24$ , p = 0.032).

Acq-Tag(-): n = 6 mice; Acq-Tag(+): n = 4 mice; Ext-Tag(-): n = 6 mice; Ext-Tag(+): n = 4 mice. Data are means  $\pm$  s.e.m. \* p < 0.05, \*\* p < 0.01.