이학박사학위논문

편도체 및 변연계아래피질 내 공포학습관련 신경활성

Neural correlates of fear learning in the amygdala and the infralimbic cortex

2014년 7월

서울대학교 대학원
협동과정 뇌과학 전공
안 보 배

Abstract

Neural correlates of fear learning in the amygdala and the infralimbic cortex

Bobae An

Interdisciplinary Program of Neuroscience

The Graduate School

Seoul National University

Fear is one of the most intensely studied fields in emotion, due to its simple and well-known animal model, the Pavlovian fear conditioning. Numerous studies have reported that the amygdala and its surrounding brain network are critically involved in fear conditioning and extinction. However, the long-term effects of fear learning have remained largely unknown since most of the previous studies used behavioral paradigms in which memory

retrieval was tested only in the short-term. Therefore, I employed a fear learning paradigm that consists of fear conditioning and extensive extinction that spans several days.

In the first chapter, I examined how neurons in the lateral amygdala (LA), a key brain structure of fear associative learning, represents fear memory during fear conditioning and subsequent extensive extinction, reconditioning. I found that the ensemble activity of LA neurons correlated tightly with conditioned fear responses of rats in the reconditioning paradigm. Further analysis revealed that among the LA neurons that displayed increased responses to the CS after fear conditioning, some exhibited weakened responses after extinction (extinction-sensitive), whereas others remained potentiated (extinction-resistant) after extinction. These results suggest the existence of distinct neuronal populations that encode various facets of fear memory and provide insights into the neuronal mechanisms underlying fear memory modulation.

In the second chapter, I questioned whether the inhibitory network, which consists of the infralimbic cortex (IL) and the intercalated amygdala cells (ITC), is crucial for fear extinction, represents long-term correlates of fear learning that consisted of fear conditioning and extensive extinction. Single unit recordings and biochemical lesion techniques were employed to

investigate the long-term effects of fear learning. I found that the CS-responses of IL neurons which emerged after single extinction dissipated with additional extinction. In keeping with this, I also found that ITC lesions that impaired the retrieval of extinction caused no deficit if lesions were made after multiple extinction sessions. These results suggest that single and extensive extinction involves different neural mechanisms.

In summary, I investigated the long-term neural correlates of fear learning involving extensive extinction and reconditioning. First, LA neuronal population represented dynamic changes in CS-US association, while distinct sub-populations encoding various aspects of fear learning existed. Second, IL neurons and ITC activities were critical for single extinction, but not for extensive extinction. Together, these findings provide insights into the neural mechanisms underlying fear memory modulation and the treatment of fear-related mental disorders.

Key words: Lateral amygdala, Infralimbic cortex, Intercalated amygdala neurons, fear conditioning, fear extinction

Student number*:* 2009-20445

Contents

Page

Chapter 2. Neural correlates of extensive extinction learning in the infralimbic cortex and the amygdala intercalated neurons

List of Figures

Page

Chapter 1.

Chapter 2.

Background and Purpose

1. Background

1.1. Pavlovian fear conditioning

1.1.1. Characteristics of Pavlovian fear conditioning

Fear is one of the most vigorously and extensively studied fields in emotion, due to the presence of a well-verified animal model, the Pavlovian fear conditioning. When a neutral stimulus (Conditioned stimulus, CS), often a tone, is repeatedly presented with a noxious stimulus (Unconditioned stimulus, US), such as a foot shock, animals quickly learn that the CS is a predictive signal of an aversive event (Fig. 1). As a result, CS elicits defensive behavior, freezing and physiological alterations in heart rate, blood pressure and hormones, controlled by the autonomic nervous system or the endocrine system (Kapp et al., 1979; Davis, 1992; LeDoux, 2014).

Pavlovian fear conditioning has been a useful tool for studying the underlying mechanisms of fear-related mental disorders, such as post-traumatic stress disorders (PTSD) and phobias (Davis, 1992; LeDoux, 2000; Davis and Whalen, 2001). The model can be utilized across a wide range of animals, from vertebrates to invertebrates (Carew et al., 1981; LeDoux, 2000; Lau et al., 2013). It is readily and rapidly acquired, even with one CS presentation paired with a noxious stimulus (Fanselow, 1994). Once

established, fear memory is firm and long-lasting, often persists throughout life (Gale et al., 2004).

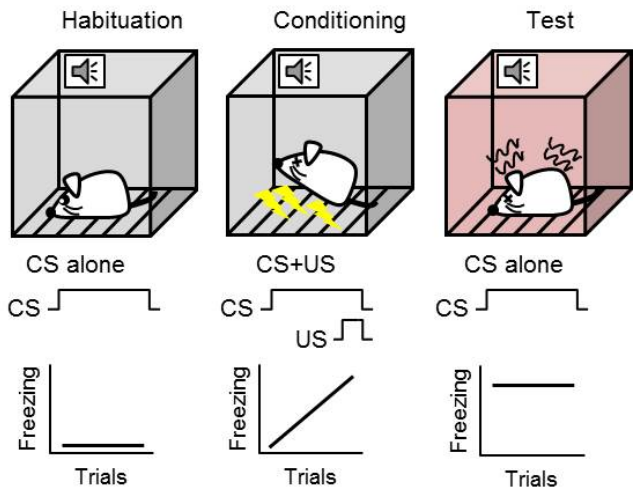

Figure 1. Pavlovian fear conditioning in rodents. Rats do not show fear responses to a neutral tone (CS) during the habituation session. After fear conditioning is performed by presenting the CS with a noxious foot shock (US) repeatedly, rats show fear responses to the tone, even if CS is presented without a shock.

1.1.2. Neural mechanisms underlying fear conditioning

Amygdala. A large body of evidence suggests the amygdala as the locus of fear memory storage and modulation (Davis, 1992; LeDoux, 2000; Pare and Duvarci, 2012), especially in the case of auditory fear conditioning (Fig. 2). Both experimentally amygdala-lesioned animals and human patients whose amygdala is damaged show deficits in acquiring the CS-US association (Phillips and LeDoux, 1992; LaBar et al., 1995; Phelps and LeDoux, 2005). Auditory thalamus and cortical inputs to the amygdala are potentiated after fear conditioning (McKernan and Shinnick-Gallagher, 1997; Quirk et al., 1997), resulting increased output signal to the downstream so as to evoke aversive behavior (Davis and Whalen, 2001). Accordingly, it has been reported that tone-evoked neural activity in the amygdala increases after fear conditioning, and decreases after closely following extinction (Quirk et al., 1997; Rogan et al., 1997), correlates well with the behavioral fear responses.

The rodent amygdala consists of distinct sub-regions (Pitkanen et al., 1997). Particularly, the lateral, basal and central part of the amygdala has been critically involved in fear and anxiety. The lateral amygdala (LA) is the main target of sensory afferents from the thalamus and cortex. Accordingly, LA neurons respond to auditory and somatosensory stimuli with short

latencies, as fast as 10 ms (Bordi et al., 1993; Quirk et al., 1995). LA has been regarded as the locus where CS-US association occurs since auditory and somatosensory information converges in the region (Bordi et al., 1993; Romanski et al., 1993). Auditory fear conditioning increases CS-responses of LA neurons (Quirk et al., 1995; Repa et al., 2001; An et al., 2012). The central amygdala (CeA) is the main output region of the amygdala. It receives inputs from the LA and the basal amygdala and sends outputs to the brainstem and the hypothalamus to control autonomic and behavioral responses (Maren and Fanselow, 1996; Pitkanen et al., 1997). Recently, it has been reported that CeA neurons respond to auditory CS and fear learning modulate CS-responses of CeA (Haubensak et al., 2010). The basal amygdala (BA) is believed to modulate CS-US association since it is reciprocally connected with various sub-cortical and cortical regions. Particularly, it receives inputs from the medial prefrontal cortex (mPFC) and the hippocampus, which are the regions involved fear extinction and contextual information processing, respectively (Maren and Fanselow, 1996; Maren and Quirk, 2004; Herry et al., 2008). The BA also interacts with neuromodulatory system, such as noradrenergic and cholinergic system, and influences on fear memory consolidation (McGaugh, 2000).

Other cortical areas. There has been accumulating evidence that

other cortical areas, such as the hippocampus, medial prefrontal cortex (mPFC) and sensory cortices also participate in fear conditioning.

The hippocampus is critical for learning the association between a neutral context and a fearful event. Hippocampal lesioned animals show deficits in contextual fear conditioning, where a neutral context is associated with a noxious foot shock, while no deficit in auditory cued fear learning (Phillips and LeDoux, 1992). It is believed that hippocampus provides more complicated CS information, which is not processed in the level of sensory thalamus, to the BA (Fanselow, 2000). Increased theta synchronization between the hippocampus and the LA during the retrieval of fear memory has also been reported, suggesting that the functional connectivity between the hippocampus and the amygdala is important for the storage and the expression of fear memory (Seidenbecher et al., 2003).

The dorsomedial part of mPFC, prelimbic cortex (PL) has also been implicated in the expression of fear memory, whereas its ventral part, infralimbic cortex (IL) is involved in fear extinction (Sotres-Bayon and Quirk, 2010). The two sub-regions of the mPFC are believed to modulate fear responses bidirectionally through their divergent projections to the amygdala. PL supports the expression of fear memory via its excitatory connection to the BA (Milad and Quirk, 2012). PL inactivation impairs fear

learning (Corcoran and Quirk, 2007; Laurent and Westbrook, 2009) and CS activates PL neurons after fear conditioning (Santini et al., 2008; Burgos-Robles et al., 2009). PL neurons show sustained increased activity that mirrors the time course of freezing responses, lasting tens of seconds (Burgos-Robles et al., 2009). Secondary sensory cortices also have been critically involved in the storage of remote fear memory (Sacco and Sacchetti, 2010). Secondary sensory cortices lesions abolish one-month-old memory, whereas recently formed memories remain intact.

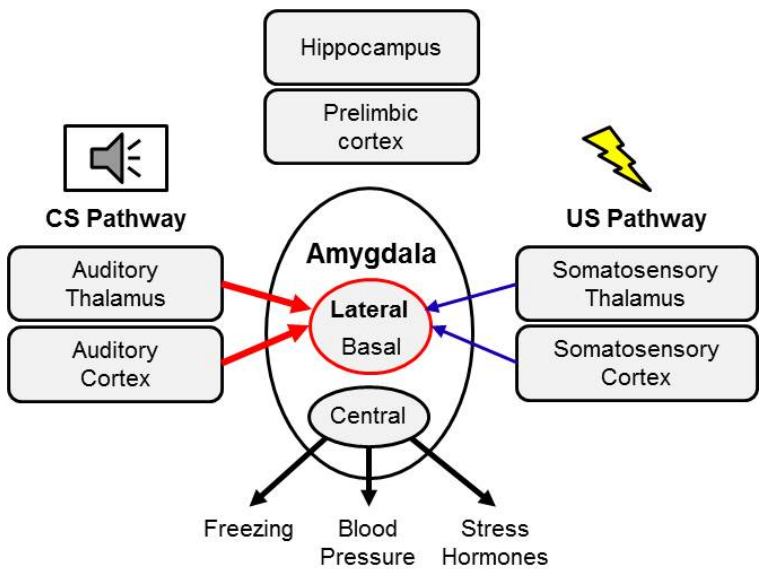

Figure 2. Brain regions involved in fear conditioning. The lateral and the basal amygdala receive sensory information of CS and US from thalamic and cortical areas. The central amygdala sends outputs to brainstem to control behavioral and autonomic responses to the CS.

1.2. Fear extinction

1.2.1. Characteristics of fear extinction

Repeated presentations of the CS in the absence of harmful stimuli, foot shocks, lead to a weakening of conditioned fear response, eventually to the point where fearful responses disappear (Fig. 3). This phenomenon is termed as fear extinction and has been a useful animal model of the exposure therapy, the most common and useful treatment for aberrant fear memory-related disorders, such as PTSD and phobia (Quirk et al., 2006; Maren, 2011).

Fear extinction is gradually acquired, unlike fear conditioning, requiring numerous CS presentations without noxious stimuli (Myers and Davis, 2007). Extinction memory is formed in a highly context-dependent manner, thus it is retrieved only in the same context where extinction learning has occurred (Bouton, 2002; Maren and Quirk, 2004). In another context, however, conditioned fear responses reappear even after extensive extinction, a phenomenon termed fear renewal, suggesting substantial remnants of the originally learned fear survive even after extensive extinction (Bouton, 2002; Chang et al., 2009). Moreover, extinction memory is less stable than fear memory, thus fear responses spontaneously

reappeared weeks after extinction training. It also supports the notion that original fear memory is not erased during fear extinction, rather inhibited temporarily (Maren and Quirk, 2004). The remnants of the original fear memory also support relearning which occurs more rapidly and with a lower threshold, compared to the initial fear learning.

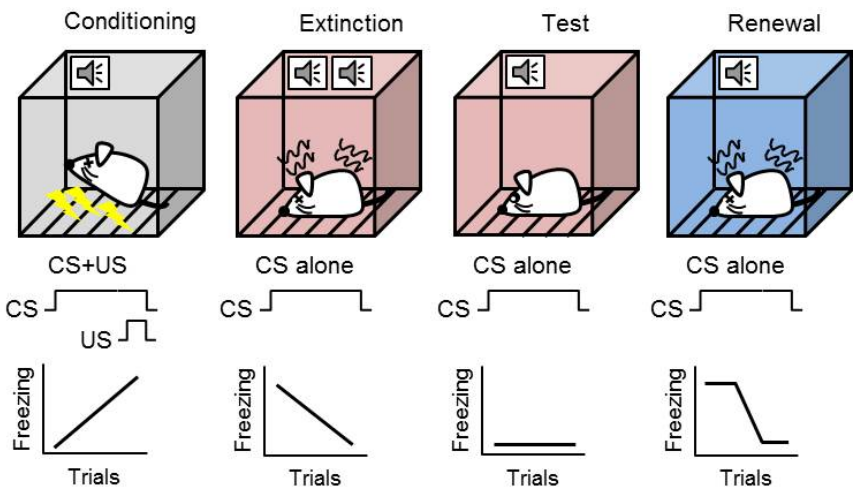

Figure 3. Fear extinction in rodents. Numerous presentations of CS alone, in fear extinction, decrease fear responses to the CS. However, conditioned fear responses reappear in various circumstances. For example, fear responses can be renewed when the rats are exposed in another context, different from the context where extinction learning has occurred.

1.2.2. Neural mechanisms underlying fear extinction

Prefrontal cortex. The ventromedial part of the medial prefrontal cortex, infralimbic cortex (IL) has been considered as a critical regulator of fear extinction, which inhibits conditioned fear behavior after extinction (Quirk et al., 2006; Sotres-Bayon and Quirk, 2010) (Fig. 4). Thalamic and hippocampal inputs to the IL are potentiated after fear extinction (Herry and Garcia, 2003), which are relayed to the medial subnuclei of the central amygdala (CeM) via the BA and amygdala-intercalated neurons to inhibit conditioned fear behavior (Maren and Quirk, 2004; Haubensak et al., 2010; Pape and Pare, 2010; Amir et al., 2011). NMDA receptor blockers infused into the IL immediately following extinction impair the retrieval of extinction, suggesting that neuronal plasticity in the IL is critical for the consolidation of extinction memory (Miserendino et al., 1990; Falls et al., 1992; Sotres-Bayon et al., 2007). Accordingly, IL neuronal activities are potentiated in animals that successfully retrieved with extinction (Milad and Quirk, 2002; Knapska and Maren, 2009) and stimulation of IL facilitates extinction (Milad and Quirk, 2002).

Amygdala. The amygdala is also critical in fear extinction. NMDA receptor blockers infused into the amygdala impair both fear conditioning

and extinction, suggesting that neuronal plasticity in the amygdala is crucial for both events (Miserendino et al., 1990; Falls et al., 1992; Sotres-Bayon et al., 2007). Similar to fear conditioning, sub-divisions of the amygdala also represent various aspects of fear extinction. LA neurons show decreased responses to the CS after extinction (Quirk et al., 1995), same as the CeA neurons (McEchron et al., 1995). However, some LA neurons retain CS-responses after fear extinction, representing the original fear memory (Repa et al., 2001; An et al., 2012). A neuronal population in the BA starts to signal the CS after extinction, named extinction neurons, suggesting BA plays a unique role in fear extinction (Herry et al., 2008). Extinction also induces depotentiation at LA input synapses (Kim et al., 2007; Dalton et al., 2008; Hong et al., 2009), and enhances local inhibitory signals (Chhatwal et al., 2005; Lin et al., 2009), all leading to decreased fear-related behavior.

Importantly, intercalated amygdala neurons (ITC), a probable mediator of prefrontal inhibition over the amygdala (Royer et al., 1999; Pape and Pare, 2010; Pare and Duvarci, 2012) are critically involved in fear extinction. ITCs are densely packed clusters of cells, mostly GABAergic neurons that surround the BLA. ITC clusters that are located between the BLA and the CeA have been implicated in fear extinction and thus described further, whereas the involvement of ITC clusters which lie between the BLA

and the cerebral cortex is elusive (Pare and Duvarci, 2012). ITC clusters at BLA-CeA border receive a dense projection from the IL and the BA (Sesack et al., 1989; McDonald et al., 1996; Freedman et al., 2000) and send its inhibitory outputs to the CeM (Pare and Smith, 1993a, b). Fear extinction potentiates BA inputs to the ITC cells that project to the CeM and this requires IL neuronal activities (Amano et al., 2010). ITC lesions impair the recall of extinction and activation of ITC cells facilitates extinction learning (Jungling et al., 2008; Likhtik et al., 2008).

Hippocampus. The hippocampus has been implicated in contextual modulation of fear extinction. Context-dependency of fear extinction is impaired if the hippocampus is inactivated before extinction training (Corcoran and Maren, 2001; Corcoran et al., 2005). Hippocampal inactivation also disrupts the context-dependent reappearance of fear after extinction, fear renewal (Corcoran and Maren, 2001; Hobin et al., 2003; Corcoran et al., 2005; Hobin et al., 2006).

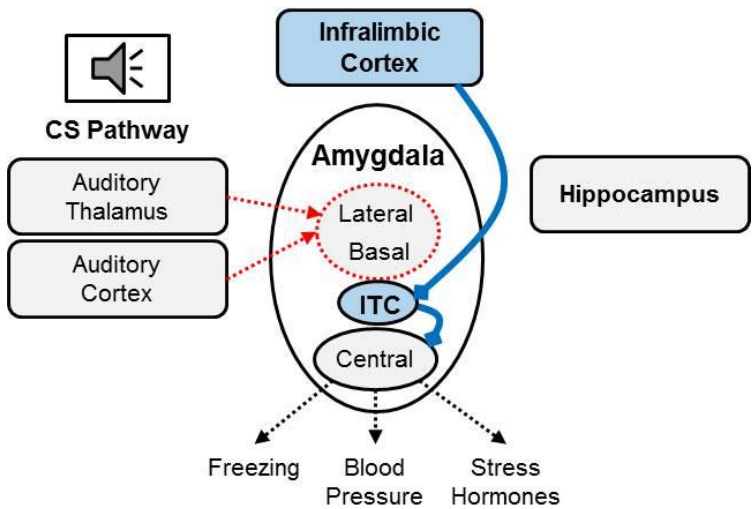

Figure 4. Brain regions involved in fear extinction. The infralimbic cortex sends its inhibitory controls over the amygdala via intercalated amygdala neurons (ITC) in the amygdala to suppress conditioned fear responses. Synaptic inputs to the lateral amygdala are also weakened by extinction. Hippocampus is implicated in contextual modulation of fear extinction.

2. Purpose

Fear is one of the most intensely studied fields in emotion, due to its simple and well-known animal model, the Pavlovian fear conditioning. Numerous studies have reported that the amygdala and its inputs and outputs are critically involved in fear conditioning and extinction. However, the long-term effects of fear learning have remained largely unknown since most of the previous studies employed a short behavioral paradigm that consists of fear conditioning and single extinction session. Therefore, I employed a fear conditioning paradigm that consists of fear conditioning and extensive extinction, spanning several days.

In the first chapter, I examined how neurons in the LA, a key brain structure where CS-US association takes place, represent the long-term correlates of fear learning which consists of fear conditioning, extinction and reconditioning. In the second chapter, I questioned whether the inhibitory network which is critically involved in fear extinction, including the prefrontal cortex and intercalated amygdala cell masses, represents the long-term correlates of fear learning encompassing fear conditioning and extensive extinction. To investigate the long-term effects of fear learning, single unit recordings and biochemical lesion techniques were employed. Together, these questions and answers could provide insights into the neural

mechanisms underlying fear memory modulation and the treatment of fear-related mental disorders.

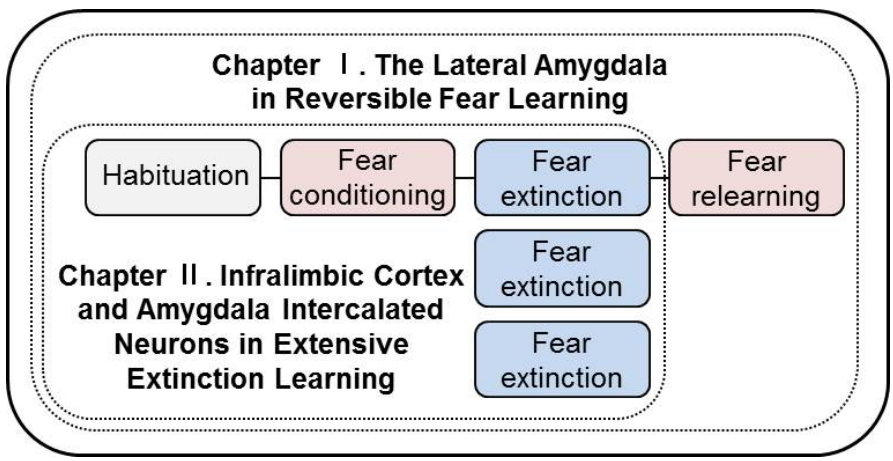

Figure 5. Schematic diagram of thesis. In chapter 1, neuronal activity of the lateral amygdala was examined in fear learning that consisted of fear conditioning, extensive extinction and reconditioning. In chapter 2, activities of the infralimbic cortex and the amygdala intercalated neurons were examined in fear conditioning and extensive extinction.

Chapter 1.

Long-term neural correlates of reversible fear learning in the lateral amygdala

Abstract

The lateral amygdala (LA) is a primary locus of auditory cued fear memory storage. LA neuronal responses to conditioned stimuli (CS) increase after fear conditioning and decrease during closely following extinction. However, the long-term effects of repeated fear conditioning and extinction on firing patterns of LA neurons have not been fully explored. Here I show, using single unit recording techniques, that the ensemble activity of LA neurons correlates tightly with behavioral fear responses of rats in fear conditioning, extensive extinction and reconditioning. The CS-evoked LA ensemble activity increased after fear conditioning, decreased after extinction, and was re-potentiated after reconditioning. Further analysis revealed that among the LA neurons that displayed increased CS-responses after fear conditioning, some showed weakened responses after extinction (extinction-sensitive), whereas others remained potentiated (extinction-resistant) after extensive extinction. The majority of extinction-sensitive neurons exhibited strong potentiation after reconditioning, suggesting that this distinct sub-population (‘reversible fear neurons’) dynamically encodes updated CS-US association strength. Interestingly, these reversible fear

neurons displayed more rapid potentiation during reconditioning compared to the initial fear conditioning, providing a neural correlate of ‘savings’ after extinction. In contrast, the extinction-resistant fear neurons did not show further increases after reconditioning, suggesting that this sub-population encodes persistent fear memory representing the original CS-US association. These results constitute the first longitudinal report on LA neuronal activity during reversible fear learning and provide insight into the neuronal mechanisms underlying fear memory modulation.

Key words: Lateral amygdala, fear conditioning, fear extinction

Introduction

Fear conditioning is the association between a neutral CS and an aversive unconditioned stimulus (US), which leads to fear responses to CS-alone presentations (LeDoux, 2000). After fear memory consolidation, which requires > 4~6 hours (McGaugh, 2000; Schafe et al., 2000), fear memory becomes remarkably resistant to perturbation, giving way only to numerous unreinforced CS presentations which leads to the extinction of conditioned fear responses. However, substantial remnants of the originally learned fear survive even after extensive extinction and cause the re-appearance of behavioral fear responses in a variety of circumstances, such as fear renewal and facilitated re-acquisition (Bouton, 2002). These observations suggest that extinction does not lead to complete reversal of fear learning, but rather a unique state in which the original fear memory traces are inhibited temporarily. The mechanisms of subsequent relearning are largely unknown, although it is well known that relearning occurs both more rapidly and with a lower threshold (i.e.'savings'; Napier et al., 1992).

The LA is essential in the acquisition and consolidation of auditory cued fear conditioning (Davis, 1992; Blair et al., 2001). Fear conditioning potentiates thalamic and cortical auditory inputs to the LA (McKernan and

Shinnick-Gallagher, 1997; Quirk et al., 1997; Tsvetkov et al., 2002), which are relayed to the basal and central amygdala to evoke aversive behavior (LeDoux, 2000; Davis and Whalen, 2001). Fear extinction recruits the infralimbic (IL) cortex to exert inhibitory influence on the medial subnuclei of the central amygdala (CeM) via the basal amygdala (BA) and amygdala-intercalated neurons (Maren and Quirk, 2004; Haubensak et al., 2010; Pape and Pare, 2010; Amir et al., 2011). Extinction also induces depotentiation at LA input synapses (Kim et al., 2007; Dalton et al., 2008; Hong et al., 2009), and enhances local inhibition (Chhatwal et al., 2005; Lin et al., 2009), all leading to decreased fear-related responses. Interestingly, NMDA receptor blockers infused into the LA impair both fear conditioning and extinction, suggesting that neuronal plasticity in the LA is critical for both events (Miserendino et al., 1990; Falls et al., 1992; Sotres-Bayon et al., 2007). Reconditioning has been less well explored, and although savings has been regarded as proof of the persistence of fear memory after extinction, the neural substrates which support the rapid relearning are largely unknown.

Previous LA unit recording studies have demonstrated that LA neurons increase their response to fear-conditioned stimuli and decrease when the stimuli become less fearful (Quirk et al., 1995; Collins and Pare, 2000; Repa et al., 2001; Goosens et al., 2003). Most of these reports

employed behavioral paradigms in which memory retrieval was tested only in the short-term, thus falling short of demonstrating the long-term modulation of fear memory involving extensive extinction and subsequent relearning. I thereby used high signal-to-noise ratio single unit recordings to track longitudinal changes in neuronal firing during fear conditioning, extinction and reconditioning. My results reveal distinct sub-populations in the LA which persistently represent the original CS-US association or dynamically encode updated CS-US association throughout the course of reversible fear learning.

Materials and Methods

Animals. Male Sprague-Dawley rats (n=45, 8 weeks old) were individually housed for 4~5 days before all experiments under an inverted 12 hours light/dark cycle (lights off at 09:00) and provided with food and water ad libitum. Behavioral training was done in the dark portion of the cycle. All procedures were approved by the Institute of Laboratory Animal Resources of Seoul National University.

Surgery and recording. Rats were anesthetized with sodium pentobarbital (50 mg/kg, i.p.) and maintained with isoflurane (1~1.5%) in O2. Rats were secured in a stereotaxic frame and bilaterally implanted with fixed-wire electrodes targeted to the LA: 2.8 mm posterior to bregma; 5.2 mm lateral to midline; and 6.3 mm to 6.9 mm deep from the cortical surface. The electrodes consisted of 8 individually insulated nichrome microwires (50 μm outer diameter, impedance 0.5~1 MΩ at 1 kHz; California Fine Wire) contained in a 21 gauge stainless steel guide cannula. The implant was secured using dental cement (Vertex). After surgery, analgesia (Metacam, Boehringer) and antibiotics were applied and rats were allowed to recover for 6~7 days. Neural activity was acquired and analyzed using a Plexon

MAP system, as previously described (Herry et al., 2008).

Behavioral procedures. Fear conditioning and extinction took place in two different contexts (context A and B) to minimize the influence of contextual associations. Reconditioning was conducted in the same context as extinction to avoid renewal and to observe savings. Context A was a rectangular Plexiglas box with a metal grid floor connected to an electrical current source (Coulbourn Instruments) which was set in a sound attenuating chamber. The chamber was illuminated with white light and was cleaned with a 70% ethanol solution. Context B was a cylindrical Plexiglas chamber with a metal grid floor which was illuminated with a red light and was cleaned with 1% acetic acid. In the retention test for the second unpairing (Post-UP2), a different context (context C) was used to avoid contextual fear. Context C was a trapezoid black opaque box with a flat black Formica floor illuminated with a red light that was cleaned with scented soap. All of the training sessions were videotaped and conditioned freezing was quantified by trained observers. The animals were considered to be freezing when there was no movement except for respiratory activity for 2 s during the 30 s CS presentation. The total freezing time was normalized to the duration of the CS presentation (Kim et al., 2010). On day

1, rats were habituated to the context and the CS in context A, in which they were placed in the recording chamber twice for 10 min, first without any cue and later with 4 presentations of the CS. The CS was a 29.089 s series of twenty-seven 2.8 kHz pure tone pips (200 ms duration repeated at 0.9 Hz, 85 dB sound pressure level) which has been used previously to enhance the signal-to-noise ratio for neural recordings (Rogan et al., 1997; Repa et al., 2001; Herry et al., 2008). On day 2, rats were given 4 presentations of the CS to determine basal LA neural responses to the CS (Hab). An hour later, fear conditioning was conducted by pairing the CS with a mild electric foot shock (0.5 mA, 1 s, 7 CS/US pairings; inter-trial interval: 80~120 s) co-initiating with the onset of the last tone pip. Extinction training took place 8 hours after fear conditioning in context B, in which rats were presented with 20 non-reinforced CS presentations (Post-FC). Two additional extinction sessions were conducted on the next day. On day 4, the behavioral and neuronal outcome of three extinction sessions was observed in a short 4 CS test session (Post-EX), followed 1 hour later by the reconditioning session in a manner similar to the initial fear learning. Eight hours after reconditioning, a retention test session was conducted (Post-REFC). To control for non-associative effects of conditioning, a separate group of rats (unpaired group, n=13) was exposed to explicitly unpaired CS and US

presentations during the conditioning and reconditioning sessions, with all the other procedures applied identically.

Single-unit spike sorting and analysis. Unit discrimination was performed using Offline Sorter (OFS, Plexon). All waveforms were plotted in a principal component space and clusters consisting of similar waveforms were first defined automatically and then verified manually. A cluster of waveforms distinct from other clusters in principal component space and showing a clear refractory period (>1 ms) was considered to be generated from a single neuron. At most, two distinct units were identified per channel, and single channel recordings proved sufficient to discern single unit responses, due to the low neuronal density of the LA (Quirk et al., 1997; Pare et al., 2004). Single unit isolation was graded using two statistic parameters, J3 and the Davies-Bouldin validity metric (DB), and neurons with a low grade were discarded. J3 reflects the ratio of between-cluster separation to within-cluster density calculated in a principal component space, and the DB is the ratio between the sum of within-cluster density to the degree of separation between clusters, and thus a high J3 and low DB value indicates a compact, well-separated unit cluster (Nicolelis et al., 2003). The long-term stability of a single-unit isolation was first determined using

Wavetracker (Plexon), in which the principal component space-cylinders of a unit recorded from different sessions were plotted (Herry et al., 2008; Tseng et al., 2011). A straight cylinder suggests that the clusters of a unit have a similar principal component composition, and that the same set of single units was recorded during the entire training session. Next I calculated the linear correlation values (r) between the template waveforms obtained over the entire set of behavioral sessions (Jackson and Fetz, 2007) to evaluate the similarity of waveform shape. Only stable units ($r > 0.93$) were considered for further analysis.

To investigate the effects of training on the LA cells, CS-evoked neural activities were normalized using a standard z-score transformation (bin size, 10ms). I adopted a moving average baseline (Pare and Gaudreau, 1996; Oyama et al., 2010) to exclude possible errors arising from extremely low spontaneous firing rates of the LA, and to reflect the in-session changes of basal firing rate. Unit responses were normalized to the firing rates of 500 ms preceding tone pip-onset for three consecutive CS (81 individual tone pips in total), except for units that did not exhibit any firing within this interval, which were normalized to the basal firing rates calculated from all pre-pip intervals of the session. Z-score peri-event time histograms (PETHs) of averaged CS-responses were constructed for each neuron and each pip

and then averaged for every CS (27 tone pips). A unit was regarded as being CS-onset or -offset responsive if the firing in 2 consecutive bins within 100 ms following CS-onset or -offset was significantly different from the baseline (500 ms preceding the CS) in an averaged PETH of all training sessions ($p < 0.05$, one-tailed t test) (Quirk et al., 1995). The onset latency of the CS-evoked responses was defined as the first bin to become significantly different from the baseline, and the bin which displayed the greatest firing within the 100-ms interval provided the peak response latency. To investigate the effects of behavioral training on the entire LA neuronal population, the population z-score PETH of all recorded neurons was calculated for each CS consisting of 27 tone pips and the mean z-values of 0~100 ms following CS-onset and -offset from the first 4 CSs of each session were compared throughout the course of behavioral training. The mean z-values in the two conditioning sessions were calculated using the first 25 tone pips of the CS to avoid foot shock-induced artifacts in the last pips.

Cell-by-cell analysis was further conducted to explore the effects of reversible fear learning on individual LA neurons. Analysis was restricted to neurons that were responsive to CS-onset. To determine responsiveness in each session, the CS-responses PETHs of 4 CSs (108 individual tone pips in

total) were averaged and the maximum z-score of the 0~100 ms interval after CS-onset was calculated for each neuron and compared to the significant z-score, 1.65 ($p < 0.05$, one-tailed t test) (Herry et al., 2008). A neuron was determined to be a 'fear neuron' if it exhibited significant CS-evoked responses in fear memory recall sessions (Post-FC or Post-REFC) and increased responses relative to the preceding sessions (Hab or Post-EX). I also sought 'extinction neurons', defined as neurons displaying strong CS-responses only after the extinction session (Post-EX), but found only one, and thus the characteristics of the fear neurons were compared to all of the other CS-responsive neurons.

Histology. At the end of each experiment, rats were anesthetized with urethane (1 g/kg, i.p.) and electrolytic lesions were made by passing a current (10 μA, 5~20 s) through recording microwires from which discrete units were identified. The duration of current injection was varied among recording microwires to identify the exact region where each unit was located. Longer current injections produced larger and more visible lesions. Animals were then transcardially perfused with 0.9% saline solution and 10% buffered formalin. Brains were removed and post-fixated overnight. Coronal sections (90 μm thick) were obtained using a vibroslicer (NVSL; World

Precision Instruments) and stained with cresyl violet. The placement of the recording microwires was examined under a light microscope.

Statistical analysis. To compare the behavioral results among behavioral sessions, averaged data points were analyzed using repeated-measures ANOVA with subsequent Newman-Keuls post hoc comparison. The CS-responsiveness of LA units was determined using unpaired t tests. For the analysis of CS-responses of LA sub-populations, the Friedman test (non-parametric one-way ANOVA for repeated measurements) and subsequent Dunn's post-hoc tests were used (Duclos et al., 2008). To detect changes in the CS-responses of the entire LA ensemble average activity (including both CS-responsive and non-responsive units), the linear sum of all CS-evoked activity was computed and the tone-to-tone variation was used for statistical deduction with parametric one-way ANOVA and Newman-Keuls post-hoc tests. Correlation between neuronal firings and behavioral responses were calculated using Pearson's correlation test. A probability value of $p < 0.05$ was considered indicative of statistical significance.

Results

Reversible fear learning dynamically regulates defensive behavior

A total of 32 rats underwent a reconditioning paradigm as described (see Methods) (Fig. 6A) and their fear-related behavior to the CS were examined. The CS was a series of twenty-seven 2.8 kHz pure tone pips (200 ms duration repeated at 0.9 Hz). Eight hours after the initial fear learning, rats displayed robust freezing when they were exposed to the CS in a different context ($F(3,93) = 781.70$, $p < 0.0001$, repeated-measures ANOVA; Hab vs. Post-FC, $p < 0.05$, Newman-Keuls posttest) (Fig. 6B) and the conditioned fear behavior diminished progressively over three extinction sessions (Fig. 6C). Reconditioning was conducted after CS-evoked fear returned to pre-conditioning levels with extinction training (Hab vs. Post-EX, $p > 0.05$) and resulted in stronger fear responses compared to the initial fear learning (Post-FC vs. Post-REFC, $p < 0.05$). In contrast, the 13 rats that received unpaired CS-US presentations showed no evidence of CS-induced fear, except immediately after shock delivery ($F(3,36) = 0.83$, $p > 0.5$, repeated-measures ANOVA; $p > 0.05$ for all pairs, Newman-Keuls posttest).

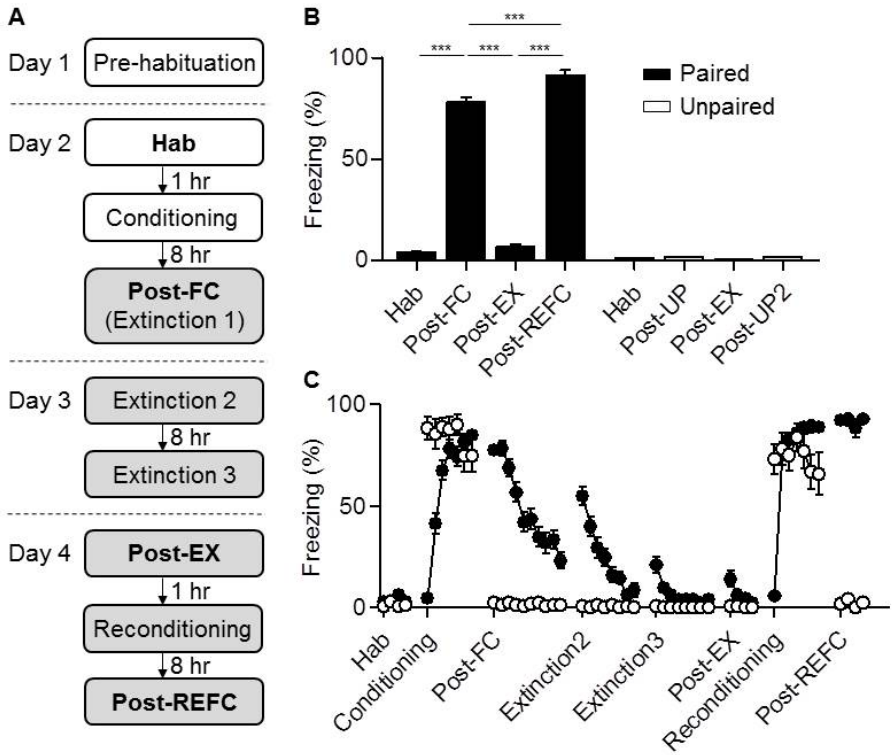

Figure 6. Behavioral procedures and results. A, The behavioral procedure used in the experiment. The white and gray shades represent different contexts. **B,** Averaged freezing responses during the first 4 CS presentations of the retention test sessions (bold characters in A) in each group (paired group, n=32 rats; unpaired controls, n=13 rats). **C,** The learning curves of the entire behavioral session (paired group, filled circle; unpaired controls, open circle). Error bars indicate SEM. Abbreviations: Hab, habituation; Post-FC, post-conditioning; Post-EX, post-extinction; Post-REFC, post-reconditioning.

Electrophysiological characteristics of the LA neurons

Only stable, high signal-to-noise ratio LA neurons verified by principal component comparisons and correlation analysis were included in the data analysis (Fig. 7). In total 188 LA neurons were analyzed, 114 from the fear-conditioned group and 74 from the unpaired controls. Histological analysis revealed that recorded cells were located within the dorsal and ventral LA (Fig. 8). Consistent with previous reports, the LA neurons displayed low spontaneous firing rates (Quirk et al., 1995; Pare and Collins, 2000; Repa et al., 2001). The average firing rate was 0.68 Hz, ranging from 0.01 to 13 Hz, and the averaged spike width (the time between the maximum and minimum peak) was 0.43 ms, ranging from 0.12 to 0.75 ms. In accordance with previous results (Quirk et al., 1995), most of the recorded LA cells showed wide spike widths and low firing rates and the waveform width and firing rate were inversely correlated ($r = -0.48$, $p < 0.0001$, Pearson's correlation test), consistent with the pyramidal-type projection neurons which are prevalent in the LA (McDonald, 1982; Davis et al., 1994; Medina et al., 2002). The average basal firing rates were not different among the behavioral sessions ($F(5,565) = 1.64$, $p > 0.1$, repeated-measures ANOVA).

Forty five of 114 (39%) neurons in the fear-conditioned group and

22 of 74 (30%) neurons in the unpaired controls were determined as CS-responsive based on the averaged CS-evoked neural activities in all of the training sessions. These neurons displayed phasic responses to tone within 40 ms following pip-onset (Fig. 9A), with an average onset response latency of 26.3 ± 1.9 ms (paired group, 25.3 ± 2.5 ms; unpaired group, 29.1 ± 2.7 ms; $p > 0.1$, unpaired t test). The pip-evoked excitation appeared reliably throughout the individual CS presentations of 27 individual pips, thus the pip-evoked responses were averaged to enhance signal-to-noise ratio of CS-responses as shown in previous studies (Rogan et al., 1997; Repa et al., 2001; Herry et al., 2008). The number of CS-responsive neurons in each separate session was not largely changed throughout the course of reversible fear learning, while repeated unpairing resulted in fewer neurons being responsive (Table 1). Histological analysis revealed that LAd neurons responded to the CS with shorter response latencies than LAv neurons (LAd, 24.3 ± 2.1 ms; LAv, 31.6 ± 3.8 ms; $p < 0.05$, unpaired t test) (Bordi et al., 1993). Interestingly, 43% of the CS-onset responsive neurons (n=29) also displayed CS-offset responses (Fig. 9B), while 20 neurons were only responsive to CS-offset with an average latency of 26.5 ± 3.2 ms.

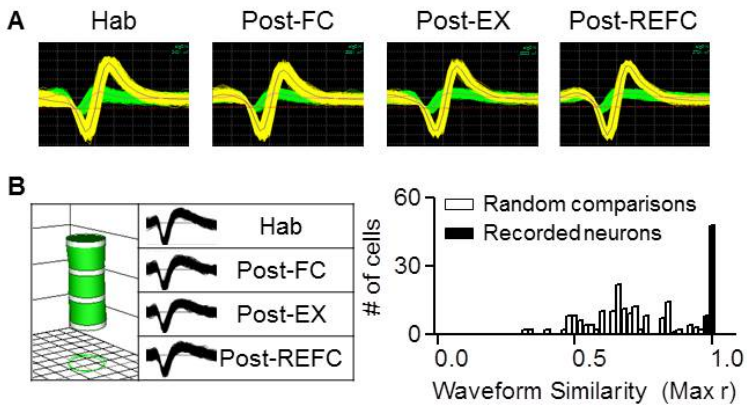

Figure 7. Long-term single unit recordings in the LA. A, Representative waveforms of two neurons recorded from a single electrode and were stably observed throughout the behavioral training period. Grid: 55μV, 100μs. **B,** Verification of long-term stable single unit recordings using principal component space cylinders (Left). A straight cylinder suggests that the same set of single units was recorded in different behavioral sessions. Quantitative evaluation of waveform similarity from units recorded on different days (Right). Randomly selected waveforms were used as a control.

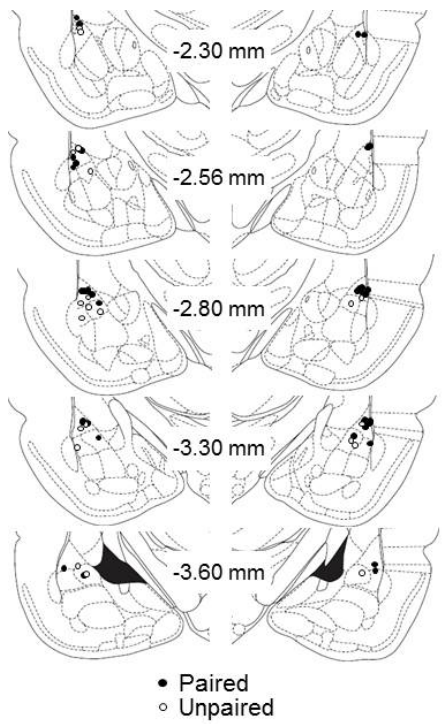

Figure 8. Histological verification of the electrode placements. The electrode placements were found within the LA, varied in dorsal-ventral and anterior-posterior axes. The paired group is indicated with filled circle and the unpaired controls with open circle.

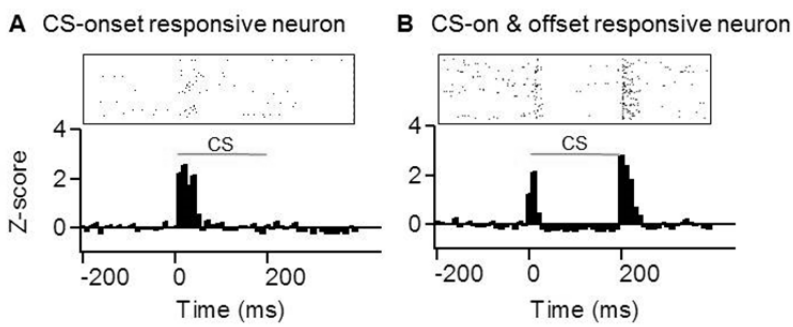

Figure 9. LA neuronal responses to the CS. A, A representative unit showing phasic responses to CS-onset. **B,** Both CS-onset and CS-offset induced strong excitation in a representative LA neuron.

	All sessions averaged			HAB			Post-FC			Post-EX			Post-REFC		
	Firing rates (Hz)	% of responsive cells		Firing rates (Hz)	% of responsive cells		Firing rates (Hz)	% of responsive cells (Newly emerged)		Firing rates (Hz)	% of responsive cells (Newly emerged)		Firing rates (Hz)	% of responsive cells (Newly emerged)	
		Onset	Offset		Onset	Offset		Onset	Offset		Onset	Offset		Onset	Offset
Paired	0.79 ± 0.19	39.47	27.19	0.89 ± 0.21	28.95	20.18	0.82 ± 0.24	34.21 (15.79)	28.95 (18.42)	0.85 ± 0.22	25.44 (10.53)	14.04 (10.53)	0.68 ± 0.18	29.83 (13.16)	21.05 (15.79)
Unpaired	0.52 ± 0.21	29.73	24.32	0.55 ± 0.22	25.68	12.16	0.55 ± 0.22	22.97 (6.76)	25.68 (20.27)	0.42 ± 0.16	18.92 (8.11)	6.76 (4.05)	0.50 ± 0.22	12.16 (6.76)	8.11 (8.11)

Table 1. Basal firing rates and CS-response properties of the recorded LA neurons in the paired group (n=114) and the unpaired controls (n=74) throughout the reversible fear learning.

LA ensemble activity represents updated CS-US association strength in reversible fear learning

It has been reported that the CS-evoked responses of LA neurons increase after fear conditioning, and that closely following extinction results in decreased tone responses of LA neurons in vivo (Quirk et al., 1995; Collins and Pare, 2000; Repa et al., 2001; Goosens et al., 2003). However, neural representations of fear memory involving extensive extinction and subsequent reconditioning have remained elusive because most previous studies have used behavioral paradigms in which memory retrieval was tested only in the short-term. Therefore, I investigated LA responses to the CS in reversible fear learning comprising extensive extinction and reconditioning. Fear conditioning-induced changes in tone-evoked firings were examined eight hours after the initial fear conditioning, a time at which fear memory is fully consolidated (Schafe et al., 2000; Schafe and LeDoux, 2000).

I constructed a population z-score PETH throughout the reversible fear learning and found that LA neurons showed potent excitation in response to CS-onset and their activity was dynamically modulated in the reversible fear learning, corresponding to the CS-US association strength. Fear conditioning resulted in a strong CS-evoked excitation of LA neurons,

while this excitation was weakened during extensive extinction, and reconditioning reinstated a strong CS-response (Fig. 10). In the unpaired controls, however, CS-evoked responses were largely unchanged by the initial unpairing, and were weakened by the second unpairing.

The average CS-evoked responses of LA neuronal population were quantified as a mean z-value of 0~100 ms following CS-onset and compared across retention test sessions of reversible fear learning. Fear conditioning significantly increased the averaged CS-response compared to habituation ($F(3,12) = 14.03$, $p < 0.001$, one-way ANOVA; Hab vs. Post-FC, $p < 0.05$, Newman-Keuls posttest), whereas unpairing did not alter LA neuronal responses ($F(3,12) = 3.52$, $p < 0.05$, one-way ANOVA; Hab vs. Post-UP, $p > 0.05$, Newman-Keuls posttest) (Fig. 11A). Three CS-alone extinction sessions resulted in decreased LA responses indiscernible with habituation (Hab vs. Post-EX, $p > 0.05$). These results are consistent with previous reports, which demonstrated the short-term effects of fear conditioning and extinction on LA neurons (Quirk et al., 1995; Repa et al., 2001) and further suggest that the updating of CS-US association strength that takes place during the reversible fear learning is dynamically represented in the LA even after memory consolidation. Consistently, reconditioning again increased CS-evoked responses of the LA compared to both the preceding

extinction retrieval session and the habituation session (Post-EX vs. Post-REFC, $p < 0.05$; Hab vs. Post-REFC, $p < 0.05$). In the unpaired controls, LA neuronal responses to CS-onset slightly decreased after the second unpairing, possibly due to safety learning (Lolordo, 1969; Rogan et al., 2005), but not to statistically significant levels (Post-EX vs. Post-UP2, $p > 0.05$) (Fig. 11A). The averaged LA population activity was positively correlated with the freezing behavior in the paired group ($r = 0.55$, $p < 0.001$, Pearson's correlation test), but not in the unpaired control ($r = 0.08$, $p > 0.1$, Pearson's correlation test) (Fig. 11B).

Importantly, CS-evoked response latencies were also reversibly altered; the CS-evoked response arose and peaked more rapidly following the initial fear conditioning and reconditioning compared to the preceding sessions (onset response latencies, Hab vs. Post-FC, Post-FC vs. Post-EX, Post-EX vs. Post-REFC, $p < 0.05$, paired t test; peak response latencies, $p < 0.05$ for the same pairs, paired t test) (Fig. 11C). Again, unpaired controls did not show significant changes ($p > 0.1$ for the same pairs, paired t test) (data not shown). Faster response latencies are consistent with strengthened influences from the short-latency thalamic pathway (McKernan and Shinnick-Gallagher, 1997; Quirk et al., 1997). These intricate, dynamic changes in the CS-response profile further support the involvement of

specific plastic mechanisms reversibly recruited in my learning paradigm.

Additionally, I checked whether CS-offset responses were altered following reversible fear learning, because a considerable number of LA neurons were responsive to CS-offset. Fear conditioning, however, did not significantly alter the CS-offset responses of the LA neurons and the responses disappeared following extensive extinction (Fig. 12). Collectively, these results suggest that the average LA ensemble activity represents updated CS-US association strength in the reversible fear learning and maintains this representation beyond memory consolidation, consistent with previous reports (Maren, 2000; Goosens et al., 2003; Hong et al., 2011).

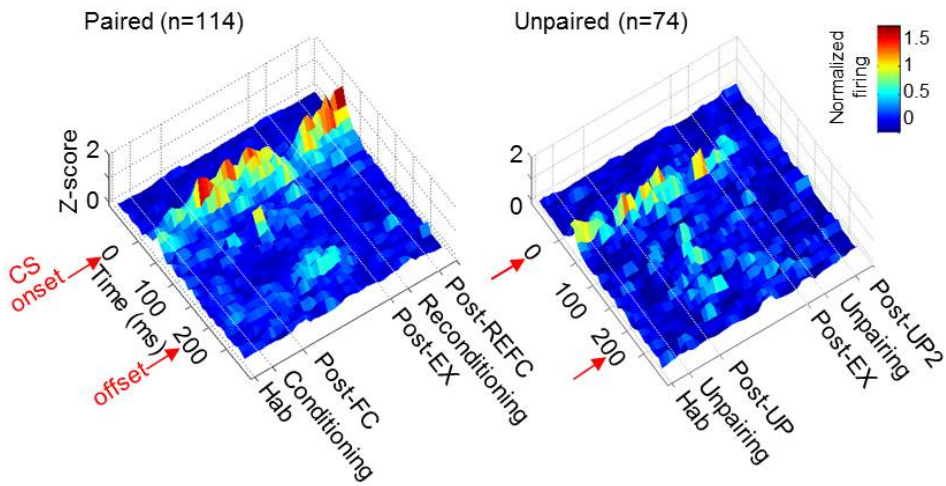

Figure 10. LA ensemble activity during reversible fear learning. Population z-score PETH throughout the behavioral training in the paired group (n=114, left) and the unpaired controls (n=74, right). The surface plot of the normalized firing rate was calculated and was smoothed for ±1 trials.

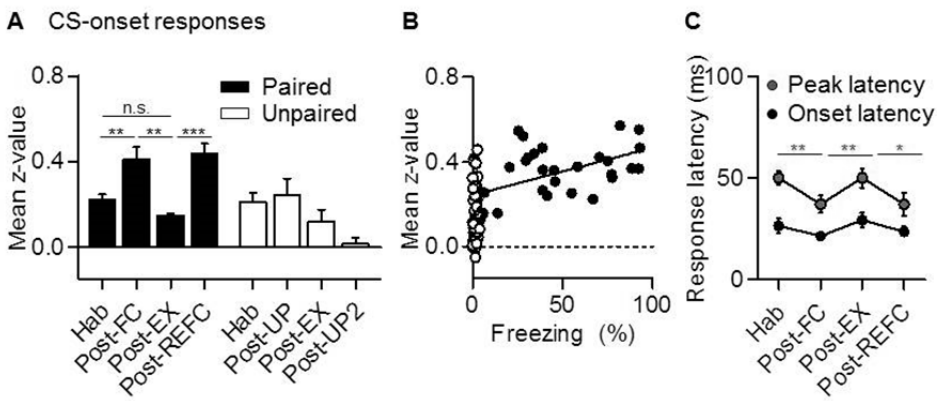

Figure 11. Quantification of LA ensemble activity to CS-onset. A, Comparisons of mean z-values calculated in a period of 0~100 ms following CS-onset. The paired group displayed reversible CS-evoked responses in contrast to the unpaired controls. **B,** Correlation analysis between neural responses and freezing behavior. A significant correlation was observed only in the conditioned group (r = 0.55; filled circle), not in the unpaired controls (r = 0.08; empty circle). **C,** Comparison of the onset onset and peak response latency across the retention test sessions. Conditioning resulted in a more rapid onset and peak response latency compared to the preceding sessions. Error bars indicate SEM.

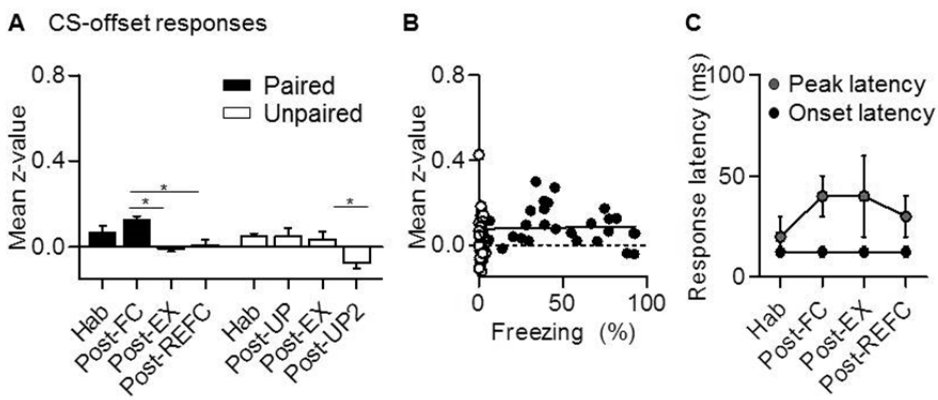

Figure 12. Quantification of LA ensemble activity to CS-offset. A–C, The same quantification as CS-onset responses was performed for the CS-offset responses. Fear conditioning did not significantly alter the CS-offset responses of LA neurons. Error bars indicate SEM.

Distinct sub-populations of LA fear neurons represent the updated and original CS-US association strength in reversible fear learning

It has been demonstrated that fear conditioning results in a strong potentiation of CS-evoked LA field potentials (Rogan et al., 1997), while only 10~30% of LA neurons display increased CS-evoked responses after fear conditioning and this subset of neurons exhibits various types of learning-induced plasticity, such as transient or persistent potentiation by fear conditioning (Quirk et al., 1995; Repa et al., 2001). I thus further analyzed the data on a cell-by-cell basis to identify distinct LA neuronal sub-populations that encode the various facets of reversible fear learning. I focused on CS-onset responsive neurons, since the LA population displayed stronger excitation in response to CS-onset and this response was dynamically modulated during reversible fear learning.

I first identified neurons which displayed significant and increased responses to CS-onset after either of the two fear conditioning sessions (Post-FC or Post-REFC) compared to the preceding sessions (Hab or Post-EX), and these neurons were defined as ‘fear neurons’ (n=25, 56% of CS-onset responsive units) (Fig. 13). I also sought for 'extinction neurons' displaying increased CS-responses only after extinction and found only one,

consistent with previous results showing that they reside mostly in the BA (Herry et al., 2008). 68% of the fear neurons increased their responses to CS after the initial fear conditioning ('conditioning-potentiated fear neurons', n=17) (Fig. 14A) and a larger number of neurons exhibited potentiated responses following reconditioning ('reconditioning-potentiated fear neurons', n=21, 84% of fear neurons) (Fig. 14B). Both conditioning- and reconditioning-potentiated fear neurons displayed reversible changes of CS-evoked firing patterns throughout the course of reversible fear learning, while small and relatively constant responses were observed in the other CS-responsive neurons that were categorized as non-fear-encoding neurons ('other neurons', n=20, 44% of CS-onset responsive units) (Fig. 14C). The basal firing rates and spike duration of fear neurons were not different from the other CS-responsive neurons ($p > 0.1$, unpaired t test) (Fig. 15A). However, fear neurons responded to the CS with a shorter response latency compared to the other neurons (fear neurons, 24.0 ± 1.6 ms; other neurons, 32.5 ± 5.2 ms; $p < 0.05$, unpaired t test) (Fig. 15B) and were frequently found in the dorsal part of the LA, with a few in the ventral LA (Fig. 15C), suggesting potent innervation by short-latency thalamic inputs. Interestingly, I found that there was a large overlap between neurons that were potentiated after the original fear conditioning and reconditioning; 76% of the

conditioning-potentiated fear neurons was re-potentiated by reconditioning (n=13) (Fig. 13), suggesting that traces of the initial fear learning remained even after extensive extinction, which allowed neurons to be readily recruited by the subsequent relearning.

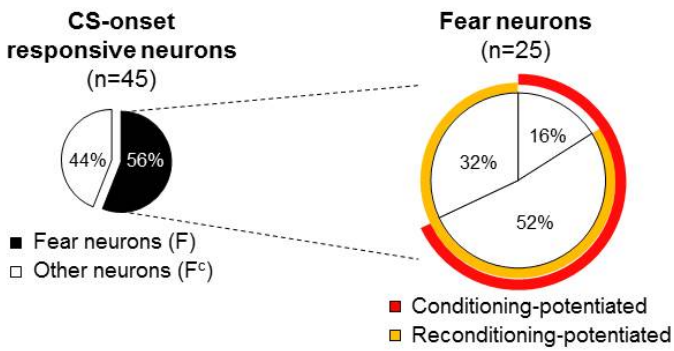

Figure 13. Fear-encoding neurons in the LA. Pie chart shows the percentage of fear neurons among the CS-onset responsive neurons (left, n=45 cells) and the subcategories of fear neurons (right, n=25 cells). A large overlap between the conditioning-potentiated fear neurons (n=17 cells) and reconditioning-potentiated fear neurons (n=21 cells) was observed.

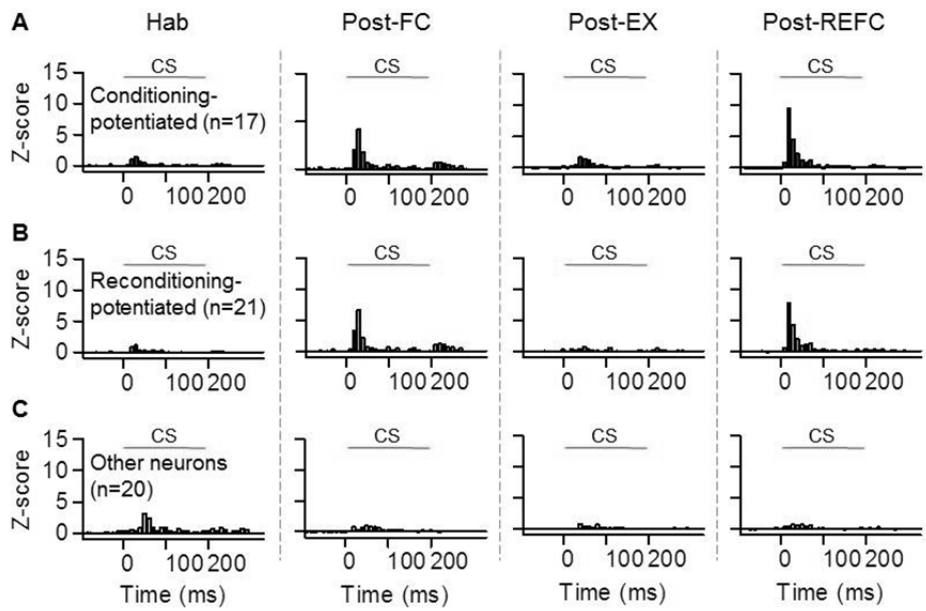

Figure 14. CS-responses of fear-encoding LA neurons. A, Z-score PETH of conditioning-potentiated fear neurons (n=17, 68% of fear neurons). **B,** Z-score PETH of reconditioning-potentiated fear neurons (n=21, 84% of fear neurons). **C,** Z-score PETH of CS-onset responsive, but not fear-encoding neurons (other neurons, n=20, 44% of CS-onset responsive units).

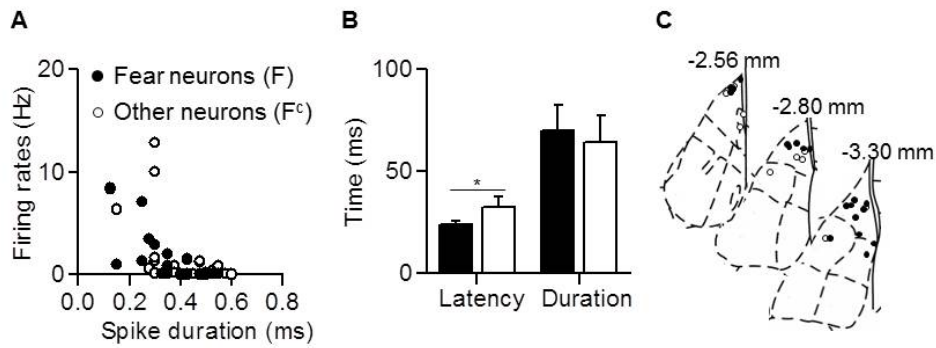

Figure 15. Characteristics of fear-encoding LA neurons. A, The basal firing rates and spike duration of fear neurons were not different from the other neurons. **B,** A comparison of onset latency and response duration. Fear neurons responded with a shorter latency to CS-onset compared to the other neurons. Error bars indicate SEM. **C,** Histological analysis revealed that fear neurons were preferentially found in the LAd.

To identify distinct LA neuronal sub-populations that represent various facets of reversible fear learning, I tracked the changes in CS-evoked responses of neurons that were potentiated following the initial fear conditioning ('conditioning-potentiated fear neurons') in subsequent extinction and reconditioning. Although the averaged CS-evoked responses of the conditioning-potentiated fear neurons appeared to be reversibly modulated (Fig. 14A), a cell-by-cell analysis revealed that this population was not homogeneous; two distinct classes of neurons were identified based on their responses to extinction (Fig. 16). Half of the conditioning-potentiated neurons exhibited significantly decreased CS-evoked responses after extinction ('extinction-sensitive fear neurons', n=9, 53% of conditioning-potentiated fear neurons) (Fig. 17A), while the other half retained increased CS-responses even after extensive extinction ('extinction-resistant fear neurons', n=8, 47% of conditioning-potentiated fear neurons) (Fig. 17B). These results are consistent with a previous study which reported similar neuronal populations within a single extinction session conducted 1 hour after fear conditioning (Repa et al., 2001). Interestingly, the extinction-sensitive fear neurons exhibited typical phasic and strong responses to CS-onset corresponding to short-latency sensory inputs, whereas extinction-resistant fear neurons exhibited smaller but more

sustained responses to the tone (over 100 ms). The onset latencies were not different between these two populations (extinction-sensitive fear neurons, 20.0 ± 2.9 ms; extinction-resistant fear neurons, 22.5 ± 3.1 ms; $p > 0.1$, unpaired t test) (Fig. 18A) and histological analysis confirmed that both neuronal populations were located in the dorsal part of the LA (Fig. 18C). However, the CS-evoked responses of extinction-resistant fear neurons lasted much longer (extinction-sensitive fear neurons, 45.6 ± 16.1 ms; extinction-resistant fear neurons, 111.3 ± 21.9 ms; $p < 0.05$, unpaired t test) (Fig. 18A), and were weaker (mean z-value, extinction-sensitive fear neurons, 9.9 ± 2.1; extinction-resistant fear neurons, 3.5 ± 0.4; $p < 0.005$, unpaired t test) (data not shown), suggesting distinct connectivity. The longer, persistent responses in the extinction-resistant fear neurons may involve multi-synaptic local sensory inputs and/or innervations from cortical regions (Repa et al., 2001), and may represent the persistence of the original fear memory after extinction.

Importantly, extinction-sensitive and -resistant neurons were also distinguished by their CS-evoked activities after reconditioning. The average CS-evoked responses of extinction-sensitive fear neurons were strongly potentiated after reconditioning, resembling LA ensemble activity (Fig. 17A), whereas extinction-resistant fear neurons did not show further

increases after reconditioning (Fig. 17B). Intriguingly, a cell-by-cell analysis revealed that all of the extinction-sensitive fear neurons but for a single exception showed increased and significant responses after reconditioning, and thus comprise a sub-population encoding dynamic changes in CS-US association strength during reversible fear learning ('reversible fear neurons', n=8, 89% of extinction-sensitive fear neurons, and 47% of conditioning-potentiated fear neurons) (Fig. 16). In contrast, all of the other CS-responsive neurons ('other CS-responsive neurons', n=37) (Fig. 17C) displayed weak, constant CS-evoked responses. I compared the mean z-values of the reversible fear neurons across sessions and found that their responses were reversibly altered in a manner similar to LA population ensemble activity, but to a greater extent ($p < 0.001$, Friedman test; Hab vs. Post-FC, Post-FC vs. Post-EX, Post-EX vs. Post-REFC, $p < 0.05$, Dunn's posttest). In contrast, the mean z-values of the other CS-responsive neurons remained relatively constant ($p > 0.05$, Friedman test; $p > 0.05$ for the same pairs, Dunn's posttest) (Fig. 18B), suggesting the reversible fear neurons lead the LA neuronal ensemble activity in reversible fear learning. Reversible fear neurons displayed a shorter responses latency compared to the other CS-responsive neurons (reversible fear neurons, 18.8 ± 3.0 ms; other CS-responsive neurons, 30.3 ± 3.1 ms; $p < 0.05$, unpaired t test), but

with a similar response duration (reversible fear neurons, 47.5 ± 18.1 ms; other CS-responsive neurons, 71.9 ± 10.0 ms; $p > 0.1$, unpaired t test) (Fig. 18A). Consistent with these electrophysiological characteristics, histological analysis revealed that reversible fear neurons were preferentially located in the dorsal part of the LAd (Fig. 18C), which is known to receive dense thalamic short-latency innervations (LeDoux et al., 1990; Quirk et al., 1997). Together, these results suggest there are two distinct sub-populations of fear-encoding neurons in the LA; one is dynamically regulated by fear conditioning and extinction while the other represents persistence of the original fear memory.

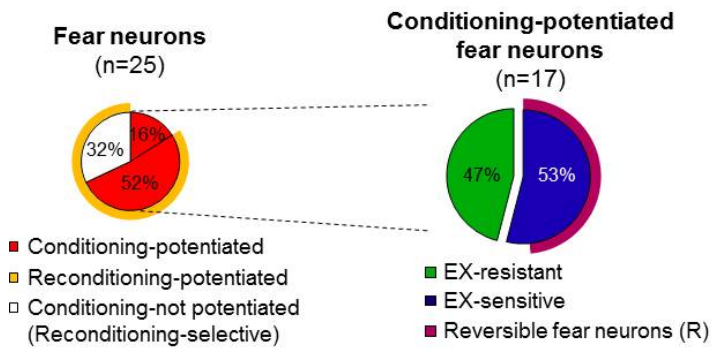

Figure 16. Fear-encoding sub-populations in the LA. Pie chart summarizes how the subcategories of conditioning-potentiated fear neurons responded to subsequent extinction and reconditioning. Conditioning-potentiated fear neurons were categorized into extinction-resistant fear neurons (n=8 cells) and extinction-sensitive fear neurons (n=9 cells). The left pie chart represents identical fear neurons as those in Figure 13.

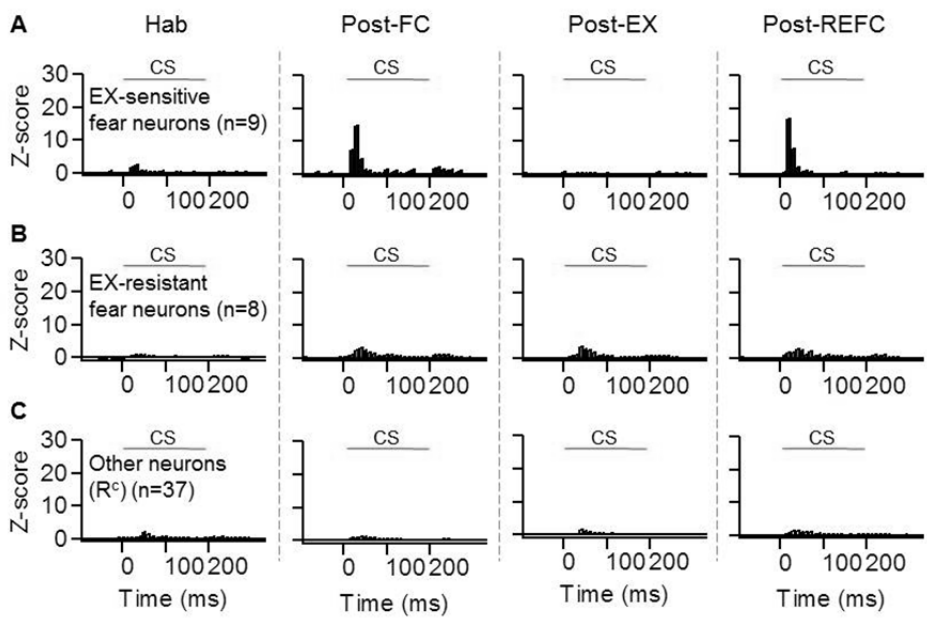

Figure 17. CS-responses of fear-encoding sub-populations. A, Z-score PETH of extinction-sensitive fear neurons (n=9, 53% of conditioning-potentiated fear neurons). **B,** Z-score PETH of extinction-resistant fear neurons (n=8, 47% of conditioning-potentiated fear neurons), which retained increased CS responses after extensive extinction**. C,** Z-score PETH of other CS-responsive neurons (n=37) that were not categorized as reversible fear neurons.

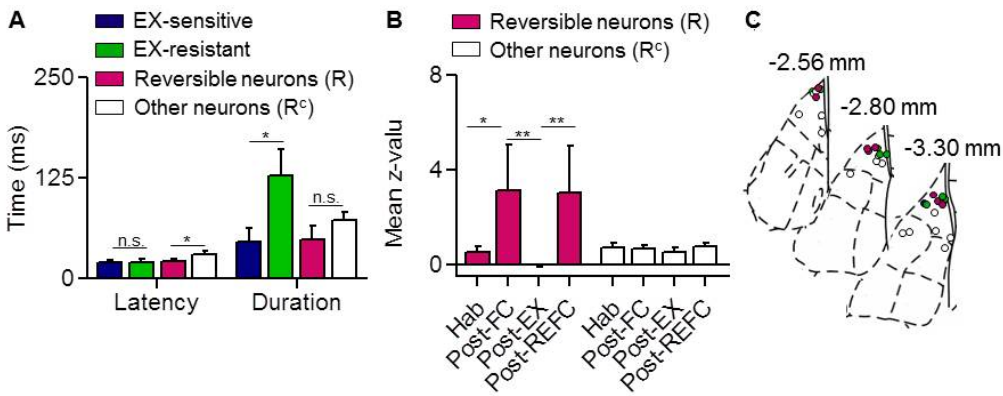

Figure 18. Characteristics of fear-encoding sub-populations. A, Comparison of onset response latency and response duration. Extinction-resistant fear neurons displayed sustained responses compared with extinction-sensitive fear neurons. The response latency of the reversible fear neurons was shorter than the other CS-responsive neurons. **B,** The mean z-value comparisons of reversible fear neurons and the other CS-responsive neurons. Error bars indicate SEM. **C,** Histological analysis confirmed that conditioning-potentiated fear neurons, including reversible fear neurons, were preferentially located in the dorsal part of the LA.

Reversible fear neurons represent savings effect after extinction

The relearning of fear occurs much faster than original fear learning even after extensive extinction, and this phenomenon is known as the 'savings' (Kehoe, 1988; Rescorla, 2001). Although savings has been widely suggested as empirical evidence of memory persistence after extinction (Bouton, 2002), the neural correlates of savings have not been identified.

In accordance with previous reports (Rescorla, 2001), I found that the freezing responses progressively increased during the initial fear conditioning, but increased more rapidly during reconditioning. CS-evoked freezing was indistinguishable between pre-conditioning sessions, Hab and Post-EX ($p > 0.05$, paired t test), and at the first pairing of the two conditioning sessions ($p > 0.1$, paired t test). However, the discrepancy between the learning curves of fear conditioning and reconditioning was significant at the second CS-US pairing ($p < 0.0001$, paired t test), the third pairing ($p < 0.005$, paired t test) and the fifth pairing ($p < 0.005$, paired t test) (Fig. 19A). Although the difference in conditioned freezing disappeared by the end of the conditioning sessions ($p > 0.1$, paired t test), stronger freezing was also observed in the retention test of reconditioning ($p < 0.0001$, paired t test) compared to the initial fear conditioning.

Interestingly, the CS-evoked responses of the reversible fear neurons increased more rapidly during reconditioning, in tight correlation with the behavioral results. The mean z-values in the two conditioning sessions diverged at the second CS-US pairing ($p < 0.05$, paired t test) (Fig. 19B), while the CS-responses in the pre-conditioning sessions and at the first pairing were not significantly different. The statistical difference disappeared at the third pairing ($p > 0.1$, paired t test), suggesting that the potentiation of the neural responses reached a ceiling faster than the behavioral responses. The rapid increases of LA neuronal responses during the reconditioning session were further confirmed by comparison of the slope of CS-response increase between the first and second CS-US pairings ($p < 0.05$, paired t test) (Fig. 19C). These results suggest that 'reversible fear neurons' not only integrate the reversible changes in CS-US association strength, but also are primed by prior learning-induced changes so as to detect a given CS-US association more rapidly during subsequent relearning. The persistently potentiated CS-responses of extinction-resistant fear neurons may also trigger/support this rapid re-potentiation of the CS-responses observed in reversible fear neurons. In addition to the more rapid in-session learning upon reconditioning, stronger freezing was also observed in the retention test of reconditioning ($p < 0.0001$, paired t test) compared to

the initial conditioning, which is likely to be supported by the larger number of neurons recruited by reconditioning (Fig. 13) compared to the initial fear conditioning.

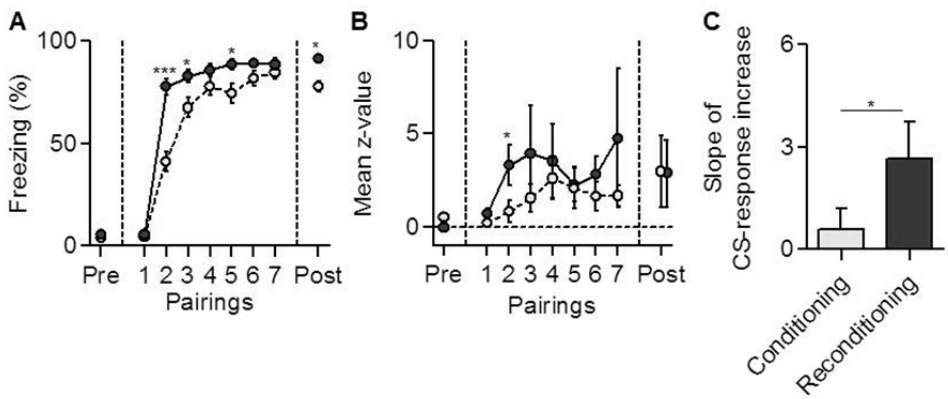

Figure 19. The neural correlate of savings after extinction. A, Behavioral evidence of savings. Reconditioning occurred faster than the initial fear learning. **B,** The mean z-values of reversible fear neurons. CS-evoked responses were larger and more rapidly increased during reconditioning (conditioning, dark gray circle; reconditioning, light gray circle). **C**, Comparison of the slope of CS-response increases between the first and second pairings. Error bars indicate SEM.

Discussion

LA neurons were found to reversibly encode updated CS-US association strength throughout the course of sequential fear learning. The LA neuronal population displayed increased average CS-evoked firing after conditioning, decreased responses after extinction and re-potentiated responses after reconditioning, in tight correlation with the changes in conditioned freezing responses. Cell-by-cell analysis revealed the two distinct sub-populations of fear-encoding neurons in the LA; one showed reversible encoding of fear learning that corresponded to the LA population activity (‘reversible fear neurons’), whereas the other was resistant to change during extinction and reconditioning (‘extinction-resistant fear neurons’), likely supporting the persistence of fear memory. Interestingly, reversible fear neurons exhibited both a stronger and more rapid acquisition of CS-US association during reconditioning relative to the initial fear conditioning, providing a neural correlate of the savings effect during reconditioning.

The ‘reversible fear neurons’ observed in the present study exhibit remarkably similar characteristics to distinct BA neurons that are responsive to fear conditioning, extinction and renewal in a reversible manner and also

a subset of LA neurons encoding the renewal of extinguished fear (Hobin et al., 2003; Herry et al., 2008). Since LA excitatory neurons are known to drive the activation of the central amygdala and fear expression via BA excitatory neurons (LeDoux, 2000; Pape and Pare, 2010; Amir et al., 2011), it is possible that the subset of LA neurons that responds to renewal (Hobin et al., 2003) largely overlaps with the 'reversible fear neurons' identified here and that both preferentially innervate 'fear neurons' in the BA (Herry et al., 2008), thus controlling central amygdala activity and contributing to reversible fear expression. Alternatively, reversible LA neuronal firing may alter activity of the amygdala-intercalated neurons and inhibitory central amygdala neurons (Pare et al., 2004; Amano et al., 2010; Haubensak et al., 2010). The extraordinary plasticity of these reversible fear neurons suggests that LA neural circuits can be dynamically modified even after memory consolidation.

The 'extinction-resistant fear neurons' found in my study provide a neural substrate for the persistent fear memory trace which had been predicted earlier (Pearce and Hall, 1980; Bouton and King, 1983). These neurons displayed CS-responses of longer duration (Fig. 18A), suggesting the influence of cortical regions where traces of persistent fear have also been identified (Corcoran and Quirk, 2007; Burgos-Robles et al., 2009;

Sacco and Sacchetti, 2010; Sotres-Bayon and Quirk, 2010). The persistent potentiated firing of the 'extinction-resistant fear neurons' may contribute to the renewal or spontaneous recovery of fear even after extensive extinction. In spite of the persistent fear-encoding in these neurons, after extinction, the expression of fearful responses is likely to be inhibited downstream of the LA (Ehrlich et al., 2009; Pape and Pare, 2010; Maren, 2011). Well-known inhibitory mechanisms involving the prefrontal cortex (Milad and Quirk, 2002; Rosenkranz et al., 2003; Likhtik et al., 2005; Sotres-Bayon et al., 2006; Quirk and Mueller, 2008) and amygdala ITC neurons (Chhatwal et al., 2005; Likhtik et al., 2008; Ehrlich et al., 2009) may provide inhibition at the BA or CeM leading to the suppression of fear responses. The context-dependent disinhibition of these subnuclei and the LA are believed to underlie the renewal of fear (Hobin et al., 2003; Likhtik et al., 2008; Ehrlich et al., 2009).

Extinction is thought to involve both inhibition and unlearning of original associations (Bouton, 2002). The relative contribution of new learning and unlearning in the behavioral extinction of many forms of associative memory has been a key issue in memory research (Medina et al., 2002; Barad, 2006). In previous studies involving different learning paradigms, the immediate reversal of CS-US contingencies resulted in the

reversal of neural responses in a subset of amygdala neurons (Schoenbaum et al., 1999; Paton et al., 2006). Consistent with these findings, my results in auditory cued-fear conditioning demonstrate that the CS-responses of some LA neurons are suppressed after extinction and exhibit savings during relearning, but there are other neurons which exhibit persistent potentiation after extinction, suggesting that unlearning and new learning are both integrated at the level of the LA neurons. Consistent with previous reports (Repa et al., 2001), 'extinction-resistant' fear neurons retained potentiated CS-responses even after extensive extinction, while 'extinction-sensitive' fear neurons showed a clear decrease in CS-responses (Fig. 17); Together, this resulted in a net reduction of the LA ensemble activity after extensive extinction. Although the net CS-response after extinction was indiscernible from pre-training levels, individual neurons displayed different responses, suggesting that network changes in LA connectivity upon fear conditioning persist after extinction. Because early- and late-extinction (within and beyond 6 hours post-conditioning, respectively) involves different mechanisms and leads to different neural changes (Myers et al., 2006; Chang et al., 2009), and most previous recordings were limited to early-extinction paradigms, my results constitute important evidence for the mechanisms underlying late-extinction.

Reconditioning after extinction has been less well explored, although the rapid re-acquisition of fear has been regarded as proof of the persistence of memory after extinction (Bouton, 2002). My findings show that whereas extinction does not return the network changes in LA connectivity to the pre-conditioning state, reconditioning appears to return the system to the pre-extinction state. Reconditioning resulted in an increase of the LA ensemble activity, which had decreased to baseline levels after extinction (Fig. 11), suggesting that LA neurons are able to adaptively represent updated CS-US association strength throughout the course of reversible fear learning. This re-potentiation was supported by a majority of the conditioning-potentiated fear neurons, demonstrating a significant overlap of fear-encoding neurons. This overlap is accounted for the extinction-induced inhibitory mechanisms that temporarily suppress fear conditioned responses. Interestingly, the CS-responses of reversible fear neurons appeared to be more readily potentiated upon reconditioning compared to the initial fear conditioning (Fig. 19), supporting the hypothesis that reconditioning reverses extinction-induced network changes. Together, these results suggest the conditioning-induced plasticity was temporarily inhibited by extinction and reconditioning eliminated this inhibition (Bouton and King, 1983; Quirk et al., 2006; Myers and Davis, 2007).

The strong reversible encoding of CS-US association strength in 'reversible fear neurons' (Fig. 16) dominates the LA population coding (shown in Fig. 10), suggesting that it is the plasticity of these neurons which is detected using field potential (Rogan et al., 1997) or immediate-early gene methods (Hall et al., 2001; Han et al., 2007; Reijmers et al., 2007). These fear neurons amount to only 10~30% of all the LA neurons, suggesting a rather sparse and restricted encoding of CS-US associations (Quirk et al., 1995; Repa et al., 2001; Han et al., 2007). In contrast, fear learning-induced synaptic potentiation has been observed in the general population of LA neurons (McKernan and Shinnick-Gallagher, 1997; Kim et al., 2007; Zhou et al., 2009), leading to the previously suggested possibility that a majority of LA neurons are strongly inhibited by GABAergic interneurons (Pare and Gaudreau, 1996) and are thus virtually undetectable by either in vivo recordings or immediate-early gene staining methods. Interestingly, a previous report demonstrated that targeted ablation of the roughly ~15% of LA neurons that preferentially participated in learning can significantly impair auditory fear memory, whereas ablating a similarly sized random population had no effect (Han et al., 2009). It is tempting to hypothesize the similarly sized 'reversible fear neuron' population in my recordings largely overlaps with the population targeted in the previous

studies.

Traces of persistent fear memory have been suggested to reside in cortical regions (Corcoran and Quirk, 2007; Burgos-Robles et al., 2009; Sacco and Sacchetti, 2010; Sotres-Bayon and Quirk, 2010), but how they may interact with the LA and support later savings or memory relapse has been largely unknown. My findings show a strong neural correlate of savings in fear-encoding LA neurons, which may be innervated and influenced by memory-preserving cortical regions to allow the more rapid detection of changes in CS-US association. Metaplastic mechanisms that enable more rapid synaptic plasticity at input synapses may also support the enhanced potentiation of CS-responses in these neurons (Abraham, 2008; Lee et al., 2013). Extinction-resistant fear neurons, which were potentiated after the initial fear learning and retained the potentiation even after extensive extinction, may also play an important role in the persistence of fear memory and relapse after extinction.

Fear conditioning and extinction have served as primary models for the treatment of PTSD and anxiety disorders. Although most PTSD research aimed at thwarting the renewal of fear memory has focused on the dysfunctions or manipulations of the prefrontal cortex (Quirk et al., 2006; Sotres-Bayon et al., 2006), my research suggests that a component of

persistent fear memory lies within the LA, thus providing an alternative target for clinical treatments.

Chapter 2.

Neural correlates of extensive extinction learning in the infralimbic cortex and the amygdala intercalated neurons

Abstract

Repeated presentations of the conditioned stimuli (CS) in the absence of aversive outcomes lead to a weakening of the conditioned fear responses, a process known to extinction. It has been believed that fear extinction recruits inhibitory network involving the infralimbic cortex (IL) and the amygdala-intercalated neurons (ITC), leading to the suppression of fear responses. Accordingly, CS-evoked responses in the IL and ITC cell activities develop after extinction. However, the long-term effects of extensive extinction learning on the inhibitory network have not been explored. Here I show that the CS-responses of IL neurons which emerged after single extinction dissipated with additional extinction sessions. The CS-evoked responses of IL neurons appeared in rats that showed less freezing in the recall of the first extinction session, but not in rats with high freezing. Surprisingly, the CS-evoked responses of IL neurons observed in the recall of the initial extinction disappeared with additional CS presentations in the same session and the CS-responses of IL never emerged in the subsequent extinction and recall sessions. In keeping with this, I also showed that ITC lesions resulted in marked deficits in the expression of

extinction caused no deficit if lesions were made after multiple extinction sessions. This first longitudinal report on the inhibitory network activity during extensive extinction learning suggests that single and extensive extinction involve different neural mechanisms and provides insight into the treatments of aberrant fear memory-related disorders.

Key words: Infralimbic cortex, Intercalated amygdala neurons, fear extinction

Introduction

Repeated presentations of the conditioned stimuli (CS) in the absence of the unconditioned stimuli (US) leads to a weakening of the conditioned response (CR), eventually to the point where the CR disappears. This phenomenon is termed as extinction and has been used as a useful animal model for the treatment of aberrant fear memory-related disorders (Maren and Quirk, 2004; Barad, 2005; Myers and Davis, 2007). However, substantial remnants of the originally learned fear survive even after extensive extinction and cause the re-appearance of fear-related behavior in a variety of circumstances, such as fear renewal and spontaneous recovery (Bouton, 2002; Myers and Davis, 2007). These observations suggest that extinction does not lead to complete reversal of original fear learning, but rather a unique state in which original traces are inhibited temporarily.

The infralimbic cortex (IL), the ventromedial part of the prefrontal cortex, has been considered as a negative regulator of aversive conditioning (Sotres-Bayon and Quirk, 2010). IL neuronal activities are potentiated in animals that successfully retrieved with extinction (Milad and Quirk, 2002; Knapska and Maren, 2009) and stimulation of IL facilitates extinction (Milad and Quirk, 2002). Moreover, NMDA receptor blockers infused into

the IL immediately following extinction impair the retrieval of extinction, suggesting that neuronal plasticity in the IL is crucial for the consolidation of extinction memory (Falls et al., 1992; Burgos-Robles et al., 2007; Sotres-Bayon et al., 2009).

Intercalated amygdala neurons (ITC), a probable mediator of prefrontal inhibition over the amygdala (Royer et al., 1999; Pape and Pare, 2010; Pare and Duvarci, 2012) receives a dense projection from the IL (Sesack et al., 1989; McDonald et al., 1996; Freedman et al., 2000) and the basolateral amygdala (BLA) and sends its inhibitory outputs to the medial subnuclei of the central amygdala (CeM) (Pare and Smith, 1993b, a), the main output nucleus of the amygdala for conditioned fear responses (Davis and Whalen, 2001). Fear extinction potentiates BLA inputs to the ITC cells that project to the CeM, which requires IL activity (Amano et al., 2010). ITC lesions impaired the recall of extinction and activation of ITC cells facilitated extinction (Jungling et al., 2008; Likhtik et al., 2008).

Although accumulating evidence indicates that the inhibitory network consisting of the prefrontal cortex and inhibitory neurons in the amygdala is crucial for fear extinction, most previous studies employed short behavioral procedures consisted of single extinction, thus falling short of demonstrating the long-term modulation of fear memory involving

extensive extinction. I thereby used high signal-to-noise ratio single unit recordings and biochemical lesions to track longitudinal changes in inhibitory network during three extinction sessions. My results revealed that CS-responses of IL neurons which emerged after single extinction session dissipated with additional extinction sessions. Moreover, ITC lesions which impaired the expression of single extinction caused no deficit if lesions were made after three extinction sessions, suggesting that different neural mechanisms underlie in single and extensive extinction.

Materials and Methods

Animals. Male Sprague-Dawley rats (n=101, 8 weeks old) were individually housed for 4~5 days before all experiments under an inverted 12 hours light/dark cycle (lights off at 09:00) and provided with food and water ad libitum. Behavioral training was done in the dark portion of the cycle (An et al., 2012). All procedures were approved by the Institute of Laboratory Animal Resources of Seoul National University.

Surgery. Rats were anesthetized with sodium pentobarbital (50 mg/kg, i.p.) and maintained with isoflurane (1~1.5%) in O2. For the IL recordings, rats were secured in a stereotaxic frame and bilaterally implanted with angled fixed-wire electrodes targeted to the IL: 2.85 mm anterior to bregma, 1.2 to 1.5 mm lateral to midline, and 4.2 to 4.6 mm deep from the cortical surface. The electrodes consisted of 8 individually insulated nichrome microwires (50 μm outer diameter, impedance 0.5~1 MΩ at 1 kHz; California Fine Wire) contained in a 21 gauge stainless steel guide cannula. The implant was secured using dental cement (Vertex). After surgery, analgesia (Metacam, Boehringer) and antibiotics were applied and rats were allowed to recover for 6~7 days. For the ITC lesion study, rats with ≤15% freezing at the end of the first extinction session were secured in a stereotaxic frame. Either D-Sap

(3 pmol/0.3 μl/hemisphere) or the same concentration and volume of a scrambled peptide conjugated to saporin (B-Sap; Advanced Targeting systems) were bilaterally infused through a micro-syringe (30 gauge) targeted to the ITC: 2.65 mm posterior to bregma, 4.75 mm lateral to midline, and 8.65 mm deep from the cortical surface. The micro-syringe was removed ten minutes after the end of the infusion to minimize diffusion along the needle tract.

Apparatus. In all experiments, fear conditioning and extinction took place in two different contexts (context A and B) to minimize the influence of contextual associations. Context A was a rectangular Plexiglas box with a metal grid floor connected to an electrical current source (Coulbourn Instruments) which was set in a sound attenuating chamber. The chamber was illuminated with white light and was cleaned with a 70% ethanol solution. Context B was a cylindrical Plexiglas chamber, with a metal grid floor which was illuminated with a red light for IL unit recordings (An et al., 2012) and a flat black Formica floor with the light off for ITC lesions (Kim et al., 2010) and the both were cleaned with 1% acetic acid. All of the training sessions were videotaped and conditioned freezing was quantified by trained observers.

Behavioral procedures. For IL unit recordings, rats were first habituated to the context and the CS in context A, in which they were placed in the recording chamber twice for 10 min, first without any cue and later with one CS presentation (Pre-habituation). The CS was a 30 s 4 kHz pure tone (85 dB sound pressure level) (Milad and Quirk, 2002). On day 2, rats were given 5 presentations of the CS to determine basal IL neural responses to the CS (Hab). Fear conditioning was conducted by pairing the CS with a mild electric foot shock (0.5 mA, 0.5 s, 5 CS/US pairings; inter-trial interval: 80~120 s) co-terminating with the CS. Extinction training took place 8 hours after fear conditioning in context B, in which rats were presented with 20 non-reinforced CS presentations (Post-Cond). Two additional extinction sessions were conducted on the next day. On day 4, the behavioral and neuronal outcome of three extinction sessions was observed in a short 5 CS test session (Post-Ext3).

For ITC lesions, rats were first habituated to the context A, in which they were placed in the recording chamber for 20 min (habituation). On day 2, fear conditioning was conducted by pairing the CS with a mild electric foot shock (0.4 mA, 1 s, 4 CS/US pairings; inter-trial interval: 80~120 s) co-terminating with the CS (Likhtik et al., 2008). The CS was a 30 s 4 kHz

pure tone (85 dB sound pressure level). On the next day, extinction training took place in context B. Two additional extinction sessions were conducted to investigate the effects of extensive extinction on the inhibitory network involving ITC. The animals were considered to be freezing when there was no movement except for respiratory activity for 2 s during the 30 s CS presentation. The total freezing time was normalized to the duration of the CS presentation (Kim et al., 2010).

Single-unit spike sorting and analysis. Neural activity was acquired and analyzed using a Plexon MAP system, as previously described (Herry et al., 2008). Unit discrimination was performed using Offline Sorter (OFS, Plexon) as previously described (An et al., 2012). Briefly, all waveforms were plotted in a principal component space and clusters consisting of similar waveforms were defined automatically and manually. Single unit isolation was graded using two statistic parameters, J3 and the Davies-Bouldin validity metric (DB). A high J3 and low DB value indicates a compact, well-separated unit cluster (Nicolelis et al., 2003), and neurons with a low grade were discarded. The long-term stability of a single-unit isolation was determined using Wavetracker (Plexon), in which the principal components of a unit recorded from different sessions were compared, and

the linear correlation values (r) between the template waveforms obtained over the entire set of behavioral sessions (Jackson and Fetz, 2007). Only stable units ($r > 0.97$) were considered for further analysis.

To investigate the effects of extinction training on the IL cells, CS-evoked neural activities were normalized using a standard z-score transformation (bin size, 100ms). Unit responses were normalized to the firing rates of four pre-tone bins. Z-score peri-event time histograms (PETHs) of averaged CS-responses were constructed for each neuron and then averaged for every CS. The mean z-values of 0~400 ms following CS-onset from the first 5 CSs of each session were compared throughout the course of behavioral training.

Histology. To identify location of recording microwires, rats were anesthetized with urethane (1 g/kg, i.p.) and electrolytic lesions were made by passing a current (10 μA, 5~20 s) through recording microwires from which discrete units were identified at the end of experiments (An et al., 2012). Animals were then transcardially perfused with 0.9% saline solution and 10% buffered formalin. Brains were removed and post-fixated overnight. Coronal sections (90 μm thick) were obtained using a vibroslicer (NVSL; World Precision Instruments) and stained with cresyl violet. The placement

of the recording microwires was examined under a light microscope.

To reveal μOR immunoreactivity, rats were anesthetized with urethane and transcardially perfused. Brains were removed and post-fixated overnight. The amygdala-containing sections (60 μm thick) were obtained from 2.0~3.0 posterior to bregma using a vibroslicer (NVSL; World Precision Instruments) and stored in PBS. The sections were incubated in 1% sodium borohydride for 30 min and pre-incubated in a blocking solution (10% goat serum, 1% BSA, 0.3% Triton-X100). Then, sections were incubated in the primary antibody solution containing μOR (ImmunoStar, 1:2000) and NeuN antibody (ImmunoStar, 1:2000) in 1% normal goat serum, 1% BSA, and 0.3% Triton-X100 in PBS for 1 hr, followed by incubation in the cocktail of the fluorescent secondary antibodies (Merck, 1:500) for 2 hrs. Cell counting was conducted as previously described (Likhtik et al., 2008), but slightly modified. Contour areas that are stained for μOR and located between the BLA complex and the CeA were defined as ITC regions. In 1-in-4 series of sections, the regions of interest (ROI) were systematically sampled (ITC counting frame, 25 X 25 μm; grid size, 45 X 43 μm; CEA, counting frame, 35 X 35 μm; grid size, 115 X 115 μm) and NeuN-positive cells in the ROI areas were counted. The optical dissector height was 10 μm.

Statistical analysis. To compare the behavioral and neural responses among behavioral sessions, averaged data points were analyzed using repeated-measures ANOVA with subsequent Newman-Keuls post hoc comparison. A probability value of $p < 0.05$ was considered indicative of statistical significance.

Results

IL neuronal activities represent CS-US dissociation after single extinction, but not after extensive extinction

It has been reported that responses of IL neurons to the CS, which emerged in the retrieval phase of extinction in fear extinguished rats, were inversely correlated with freezing at the retrieval test (Milad and Quirk, 2002). However, neural representations of extinction memory involving multiple extinction sessions have remained obscure because previous study has employed short behavioral procedures. Therefore, I investigated IL responses to the CS during multiple extinction sessions.

To investigate the effects of extensive extinction on IL neuronal activity, I employed an extensive extinction paradigm consisting of fear conditioning and subsequent three extinction sessions and IL neuronal activities were recorded throughout the behavioral training. A total of 19 rats underwent an extensive extinction paradigm as described (see Methods) (Fig. 20A) and their fear levels to the CS were examined. Eight hours after the initial fear learning, rats displayed robust freezing when they were exposed to the CS in a different context ($F(4,94) = 110.1$, $p < 0.0001$,

repeated-measures ANOVA; Hab vs. Post-Cond, $p < 0.05$, Newman-Keuls posttest) (Fig. 20C). The conditioned fear progressively diminished over three extinction sessions (Fig. 20B) and freezing levels of the rats in the last test session became undistinguishable from the pre-conditioning levels (Hab vs. Post-Ext3, $p > 0.05$, Newman-Keuls posttest).

A total of 72 cells were recorded from the IL across three days. Histological analysis revealed that recorded cells were located within the anterior part of the IL (Fig. 20D). IL neurons displayed low spontaneous firing rates, averaged firing rate of 0.98 Hz. The average basal firing rates were not different among the behavioral sessions ($F(4,349) = 1.64$, $p > 0.1$, repeated-measures ANOVA). Only stable, high signal-to-noise ratio IL neurons verified by principal component comparisons and correlation analysis were included in the data analysis (Fig. 21).

I constructed a population z-score PETH throughout the behavioral training to investigate the effects of extensive extinction learning on the neural responses of IL to the auditory CS. Since responses of IL neurons to the CS have been shown to be inversely correlated with freezing at the retrieval test, rats were divided into two groups; one with ≤ 50% recovery of freezing (n=14) and the other with > 50% recovery of freezing (n=5) in the early part of the second extinction (Milad and Quirk, 2002) (Fig. 23). In

accordance with previous results, IL neurons signaled extinguished CS in the retrieval session of fear extinction, while they were unresponsive to the CS during the first extinction session (Fig. 22B). The CS-evoked excitation of IL neurons emerged after extinction training was found only in rats with low recovery of freezing (Fig. 22B), suggesting IL neuronal responses is important for the retrieval of extinction memory. Surprisingly, subsequent extinction abolished the CS-evoked excitation of IL neurons and IL neurons remained silent during the additional extinction session and the test session on the next day (Fig. 22B). In rats with high recovery of freezing, however, CS-evoked responses of IL neurons were largely unchanged throughout the course of fear learning involving extensive extinction.

The CS-evoked responses of IL neurons were quantified as a mean z-value of 0~400 ms following the first 5 CSs and compared throughout the behavioral training. Single extinction significantly increased the averaged CS-response of IL neurons compared to the preceding two sessions ($F(4,279) = 3.35$, $p < 0.01$, one-way ANOVA; Post-Ext1 vs. Hab, Post-Ext1 vs. Post-Cond, $p < 0.05$, Newman-Keuls posttest) in rats with low recovery of freezing (Fig. 23B), whereas IL neuronal responses were not altered in rats with high recovery of freezing ($F(4,79) = 3.52$, $p > 0.5$, one-way ANOVA) (Fig. 23B). Intriguingly, CS-responses of IL neurons in rats with low fear

recovery decreased to the habituation level in the following extinction session (Post-Ext2 vs. Post-Ext1, $p < 0.05$, Post-Ext2 vs. Hab, $p > 0.05$, Newman-Keuls posttest), although rats still successfully retrieved with extinction memory (Fig. 20B). Moreover, CS-evoked responses of IL neurons were not found in all rats during the test session conducted on day 4 (Post-Ext3 vs. Hab, $p > 0.05$, Newman-Keuls posttest for the low fear recovery group), suggesting IL neuronal activity is not required for the expression of extinction memory after extensive extinction. I further analyzed IL neuronal activity in the early and the late part of each extinction session to see if the CS-responses of IL neurons alter within the extinction sessions. It was found that CS-responses of IL neurons which emerged at the start of the second extinction disappeared at the end of the same session ($F(7,385) = 3.12$, $p < 0.005$, one-way ANOVA; Post-Ext1 early vs. late, $p < 0.05$, Post-Ext1 late vs. Hab, $p > 0.05$, Newman-Keuls posttest) (Fig. 23C). Collectively, these results suggest that IL neuronal activity is differently involved in single and extensive extinction learning.

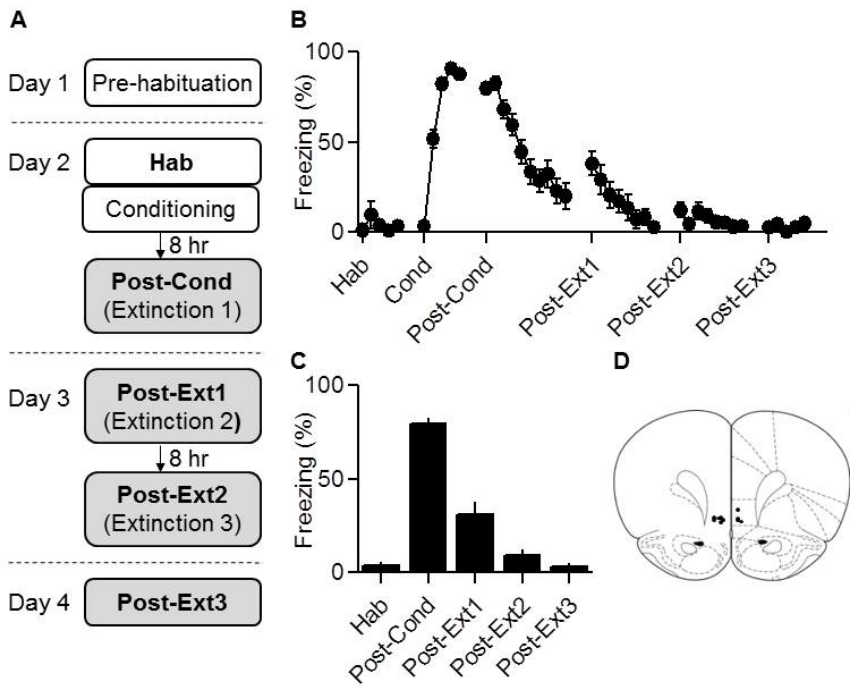

Figure 20. Behavioral procedures and results. A, The behavioral procedure used in the experiment. The white and gray shades represent different contexts. **B,** The learning curves of the entire behavioral session. **C,** Averaged freezing responses during the first five CS presentations of the retention test sessions (bold characters in A) in all rats (n=19). Error bars indicate SEM. Abbreviations: Hab, habituation; Post-Cond, post-conditioning; Post-Ext, post-extinction. **D,** Histological verification of the electrodes placements.

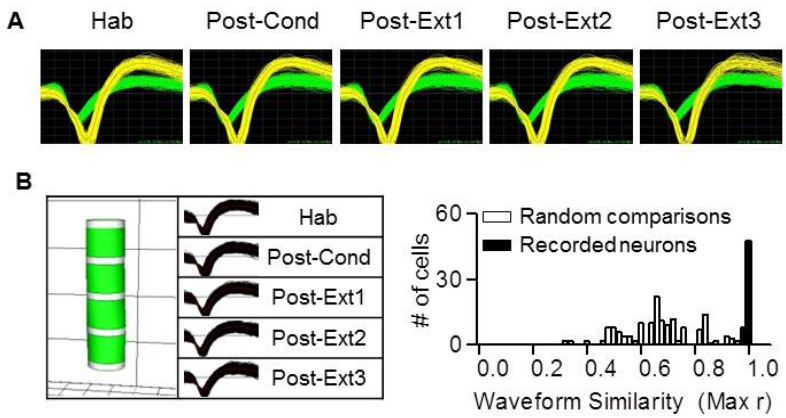

Figure 21. Long-term single unit recordings in the IL. A, Representative waveforms of two neurons recorded from a single electrode and stably observed throughout the behavioral training period. Grid: 55μV, 100μs. **B**, Verification of long-term stable single unit recordings using principal component space cylinders (Left). A straight cylinder suggests that the same set of single units was recorded in different behavioral sessions. Quantitative evaluation of waveform similarity from units recorded on different days. Randomly selected waveforms were used as a control (Right).

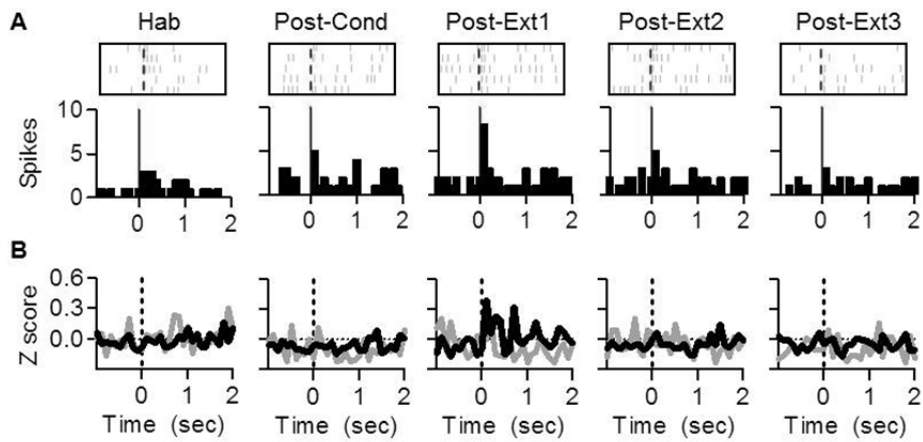

Figure 22. IL neuronal responses to the CS during fear learning. IL neurons represent extinguished CS after single extinction, but not after multiple extinction sessions. **A,** Representative neurons displaying CS-evoked responses after the first extinction. Responses decreased during subsequent extinction and test. **B,** Averaged responses of IL neurons in rats with a good recall of extinction memory (n=14, Black line) and rats with higher freezing (n=5, Gray line).

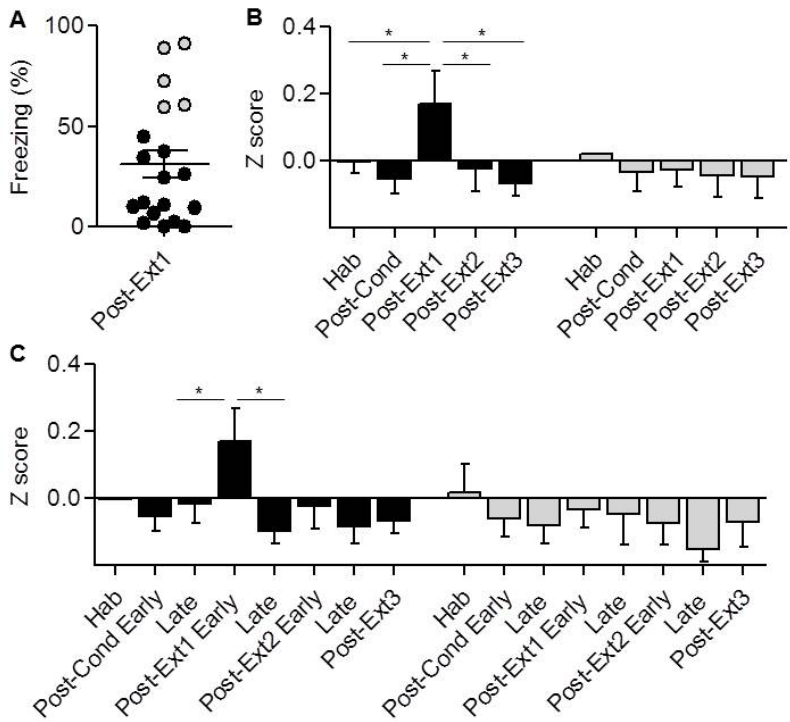

Figure 23. Quantification of IL responses to the CS. A, Rats were divided into two groups, according to their freezing levels in the second extinction session (Post-Ext1). **B,** Comparison of mean z-values calculated in a period of 0~400 ms following CS-onset. The low recovery of fear group displayed CS-responses in the second extinction session, retrieving the initial extinction memory. **C**, The CS-responses of IL neurons emerged in the early part of the second extinction and disappeared in the late part of the same session. Error bars indicate SEM.

Amygdala intercalated neurons are required for the expression of single extinction, but not extensive extinction

Thus far, I have demonstrated that IL neurons signal extinguished CS only after single extinction, but not after extensive extinction learning. I next tested whether ITC, which is the most probable mediator of prefrontal inhibition over the amygdala, is also involved in single and extensive extinction differently. To address this, I employed selective ITC lesions with a ribosome inactivating toxin (D-Sap) that was conjugated to an agonist with a high selectivity and affinity for μ-opioid receptors (μORs), dermorphin (Pare and Smith, 1993a). It has been reported that μORs are more abundantly expressed among ITC neurons, compared to adjacent BA or CeA cells (Likhtik et al., 2008). As a control, a scrambled peptide conjugated to toxin (B-Sap) was utilized.

I first tested the effects of selective ITC lesions obtained by the toxin on single extinction. Rats underwent a single extinction paradigm as described (Fig. 24) and their fear levels to the CS were examined. Either D-Sap or the same concentration and volume of a control peptide was bilaterally infused to the ITC the day after extinction session. After 7 days of recovery, the retrieval of extinction memory was tested and freezing levels to the CS were quantified in a blind manner. Only rats with syringe

tips located at the BLA-CeA border were included. In consistent with previous study (Likhtik et al., 2008), D-Sap infusions resulted in a marked reduction in μOR staining restricted to the region adjacent to infusion site, whereas more distant ITC clusters at the external capsule were not affected (Fig. 25A). μOR expression was not altered in B-Sap treated rats (Fig. 25B).

To evaluate the selective ITC lesions obtained by D-Sap infusions, I performed unbiased stereological estimates of the number of NeuN positive cells. Compared to B-Sap treated rats, the number of ITC neurons were significantly reduced in rats that received D-Sap infusions into the ITC (Fig. 25C; B-Sap, 136.6 ± 16.3, n=10; D-Sap, 62.3 ± 7.5, n=12; $p < 0.001$, unpaired t test). In contrast, the number of CeA neurons were identical in the two groups (Fig. 25C; B-Sap, 717.3 ± 23.4, n=6; D-Sap, 671.3 ± 45.5, n=6; $p > 0.1$, unpaired t test). Consistent with the previous report which showed inverse correlation between freezing levels during extinction recall and the number of survived ITC cells (Likhtik et al., 2008), D-Sap infused rats displayed impaired expression of extinction memory, whereas rats with B-Sap infusions successfully retrieved with extinction (Fig. 26; B-Sap, 29.1 ± 5.2; D-Sap, 60.0 ± 7.8; $p < 0.05$, unpaired t test). These results suggest IL neuronal activity is required for the expression of extinction memory after single extinction.

Having established the effects of selective ITC lesion on single extinction, I next examined the effects of ITC lesions on extensive extinction by employing two additional extinction sessions. D-Sap or B-Sap infusions were conducted the day after the last extinction session. Consistent with the single extinction experiment, the number of ITC neurons were significantly decreased in rats that received D-Sap infusions in the ITC compared to B-Sap treated rats (Fig. 27C; B-Sap, 161.6 ± 13.2, n=10; D-Sap, 57.93 ± 9.4; $p < 0.0001$, unpaired t test). The number of CeA neurons was identical in the two behavioral groups (B-Sap, 700.8 ± 28.1, n=6; D-Sap, 689.0 ± 27.8; $p > 0.5$, unpaired t test). Surprisingly, freezing levels of D-Sap infused rats in the recall test were not different from those of B-Sap treated rats (Fig. 27B; B-Sap, 20.8 ± 6.7; D-Sap, 14.1 ± 3.8; $p > 0.1$, unpaired t test), although toxin-mediated selective ITC lesions were effective as much as shown in the single extinction experiment. I further confirmed that fear renewal, which is one of the behavioral characteristics of fear extinction besides spontaneous recovery and savings, was normally induced after single (Fig. 28A) and extensive extinction (Fig. 28B) paradigm, suggesting that both single and extensive extinction did not erase the original fear memory. Rats displayed strong freezing when they were exposed to the context where fear conditioning had occurred, whereas no fear responses

were observed when rats were exposed to the extinction context, no matter how many extinction sessions they had experienced (Fig. 28A; ABA, 21.7 ± 1.3, ABB, 9.2 ± 2.3, $p < 0.0001$, unpaired t test for single extinction group) (Fig. 28B; ABA, 21.1 ± 2.2, ABB, 6.1 ± 2.7, $p < 0.005$, unpaired t test for extensive extinction group). Collectively, these results suggest that ITC neuronal activity is not required for the maintenance and the expression of extinction memory in extensive extinction learning consisting of three extinction sessions.

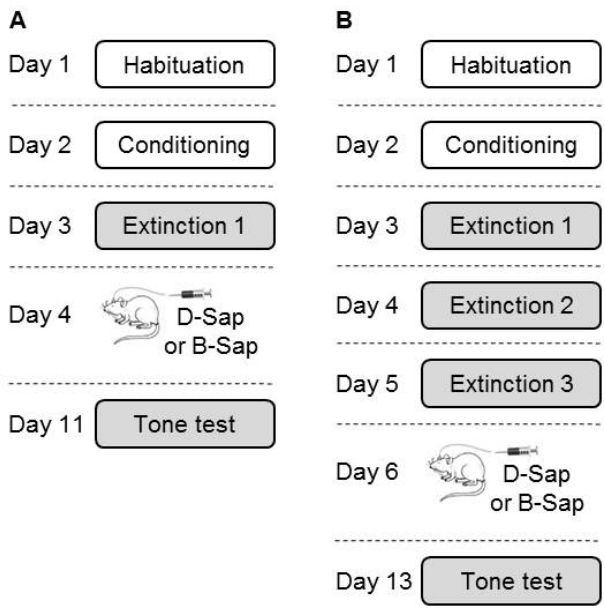

Figure 24. Experimental designs. Behavioral training and toxin infusions in **A,** single extinction and **B,** extensive extinction paradigm. Either toxin (D-Sap) or control toxin (B-Sap) was infused the next day of the last extinction session. Extinction recall was tested after 7 days of recovery.

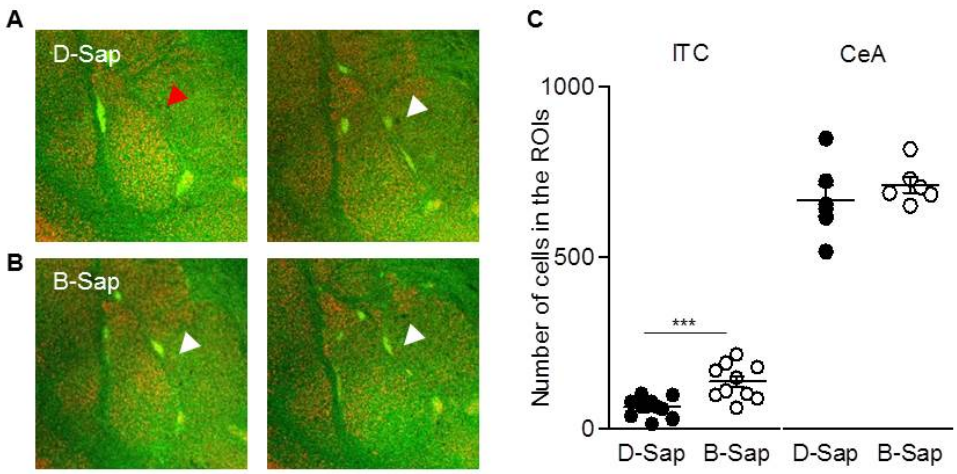

Figure 25. Selective ITC lesions. A, μOR staining in rats infused with D-Sap. μOR staining is reduced adjacent to infusion site (Red arrow), whereas distant ITC clusters were not affected (White arrow). **B**, μOR staining was not decreased by B-Sap infusion. **C**, Number of NeuN-positive cells in the ITC and the CeA. The number of ITC neurons is decreased in D-Sap treated rats, compared to the B-Sap infused rats. CeA neurons were not affected by D-Sap or B-Sap infusion. Error bars indicate SEM. Abbreviations: ITC, intercalated cells; CeA, central amygdala.

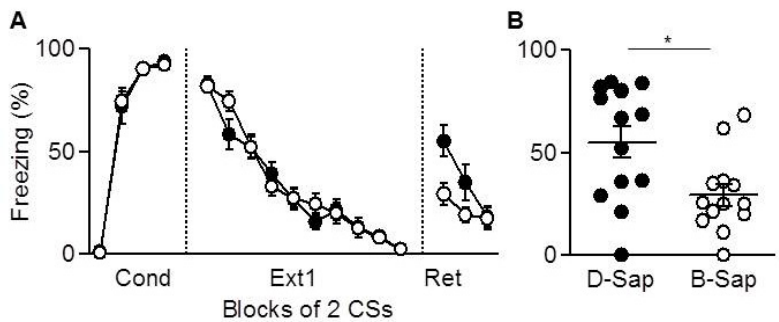

Figure 26. The effects of ITC lesions on single extinction. A, The learning curves of the entire behavioral session. On the next day of extinction, either D-Sap (Black circle) or B-Sap (White circle) was infused aimed to the ITC. Extinction memory was tested after 7 days of recovery. **B,** D-Sap treated rats displayed higher freezing in the test session, compared to the B-Sap treated rats. Error bars indicate SEM. Abbreviations: Cond, fear conditioning; Ext, extinction; Ret, retrieval session.

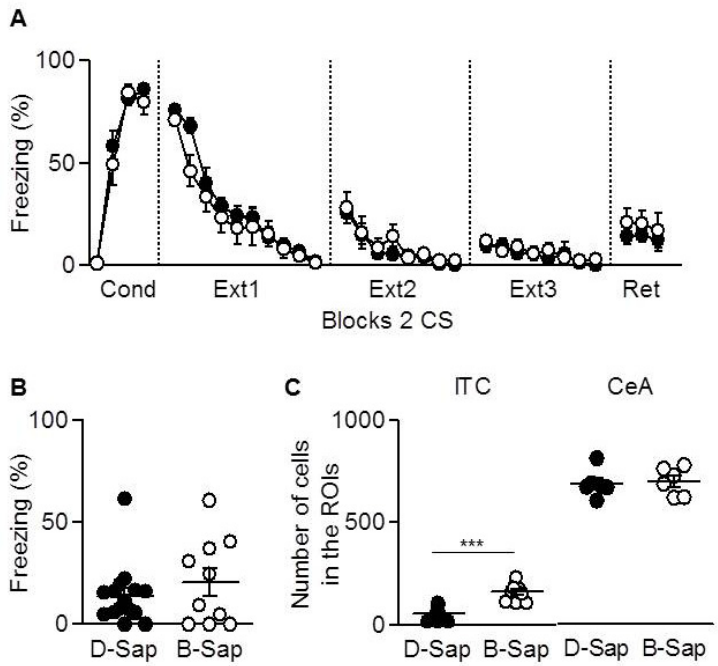

Figure 27. The effects of ITC lesions on extensive extinction. A, The learning curves of the entire behavioral session. On the next day of the last extinction, either D-Sap (Black circle) or B-Sap (White circle) was infused aimed to the ITC. Extinction memory was tested after 7 days of recovery. **B**, D-Sap treated rats displayed low freezing responses in the test session, similar to the B-Sap treated rats. **C,** Number of NeuN-positive cells in the ITC and the CeA. The number of ITC neurons is decreased in D-Sap treated rats, compared to the B-Sap infused rats. CeA neurons were not affected by D-Sap or B-Sap infusion. Error bars indicate SEM.

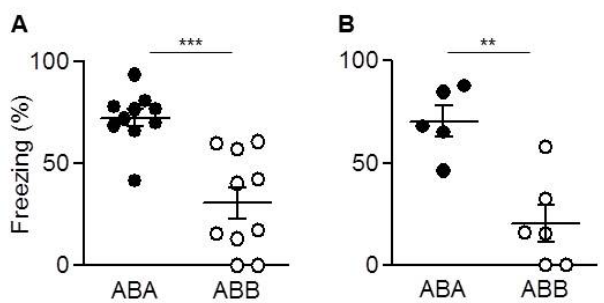

Figure 28. Renewal of fear in single and extensive extinction. Fear responses were examined in the same context where the extinction (ABB retention, white circle) or the fear conditioning (ABA renewal, black circle) took place. When rats were exposed to the conditioning context, renewal of fear was observed in rats that underwent **A,** single extinction and **B,** multiple extinction sessions. Error bars indicate SEM.

Discussion

It has been believed that the inhibitory network, including the prefrontal cortex and the inhibitory neurons in the amygdala, is critical in the acquisition and the expression of extinction memory (Maren and Quirk, 2004; Pape and Pare, 2010; Sotres-Bayon and Quirk, 2010). However, I found that the IL and the ITC, the essential brain regions constituting the inhibitory network, were crucial for single extinction, but not for extensive extinction. Consistent with previous report (Milad and Quirk, 2002), IL neurons only in rats which showed successful recall of extinction memory displayed increased CS-evoked firing in the retrieval session after the first extinction session. However, CS-responses of IL neurons decreased to pre-training level during the same session and never emerged in subsequent extinction sessions. In keeping with these results, I also showed that ITC lesions which resulted in a marked deficit in the expression of single extinction caused no deficit if lesions were made after multiple extinction sessions. Together, these results suggest that the inhibitory network is crucial for single extinction training, however, a different neural network is recruited with additional extinction sessions.

IL has long been considered as a critical regulator of aversive

conditioning (Maren and Quirk, 2004; Quirk et al., 2006; Sotres-Bayon and Quirk, 2010). NMDA receptor blockers infused into the IL immediately following extinction impair the retrieval of extinction, suggesting that neuronal plasticity in the IL is crucial for the consolidation of extinction memory (Falls et al., 1992; Burgos-Robles et al., 2007; Sotres-Bayon et al., 2009). Consistent with previous reports (Milad and Quirk, 2002; Knapska and Maren, 2009), I have observed CS-evoked excitation of IL neurons emerged after extinction in rats that successfully retrieved with extinction (Fig. 22). These potentiated CS-responses of IL neurons after extinction have been considered to mediate the consolidation and the expression of extinction memory. IL is reciprocally connected with the BLA in which a neuronal population representing extinguished CS has been reported (Herry et al., 2008). NMDA receptor blockers and protein kinase inhibitors infused into the BLA impair fear extinction, suggesting that neuronal plasticity in the BLA is crucial for the extinction of conditioned fear (Falls et al., 1992; LeDoux, 2000). IL also sends robust projections to the ITC (Sesack et al., 1989; McDonald et al., 1996) which in turn strongly inhibit output from the central nucleus of the amygdala (Royer et al., 1999), leading to the suppression of fear conditioned responses after extinction. Recently, it was reported that theta synchronization between the prefrontal cortex and the

BLA increase in response to safe cues that are not associated with noxious shocks (Likhtik et al., 2014), suggesting the IL might generally represent learned safety. Importantly, I found that CS no longer elicited excitatory responses in the IL when rats underwent additional extinction sessions (Fig. 22), suggesting IL neuronal responses is not required for the expression of extinction memory in extensive extinction. It has been reported that the cortical areas represent salient events and the saliency-related cortical activities rapidly disappear with repeated exposures to the events (Ranganath and Rainer, 2003). It is possible that IL responses to the extinguished CS might represent saliency of the CS which has been dissociated from the US. Thus, IL responses would decrease with repetitive CS presentations, since CS-US dissociation became firm and thus less salient.

ITC is one of probable mediators of prefrontal inhibition over the amygdala after extinction (Royer et al., 1999; Pape and Pare, 2010; Pare and Duvarci, 2012). Fear extinction potentiates BLA inputs to the ITC cells that project to the CeM and synaptic potentiation between the BLA and the ITC is impaired by IL inactivation (Amano et al., 2010). Consistent with previous reports (Jungling et al., 2008; Likhtik et al., 2008), I found that ITC lesions following single extinction impaired the retrieval of extinction

memory (Fig. 26), suggesting ITC is critical for single extinction. ITC receives a dense projection from the IL (Sesack et al., 1989; McDonald et al., 1996; Freedman et al., 2000) and the BLA and sends its inhibitory outputs to the CeM (Pare and Smith, 1993b, a), the main output nucleus of the amygdala for conditioned fear responses (Davis and Whalen, 2001), so as to inhibit conditioned fear behavior after extinction. However, I found that ITC lesions no longer affect the expression of extinction memory when the lesions were made after three extinction sessions (Fig. 27), suggesting ITC neuronal activity is not required for the inhibition of conditioned fear behavior after extensive extinction. Extinction recall after extensive extinction is likely to be mediated by decreased LA inputs to the CeM. LA synaptic inputs are depotentiated after extinction learning (Kim et al., 2007) and I also observed that LA ensemble activity to the CS decreased after extensive extinction (An et al., 2012).

Traces of persistent fear memory have been suggested to reside in cortical regions (Corcoran and Quirk, 2007; Burgos-Robles et al., 2009; Sacco and Sacchetti, 2010; Sotres-Bayon and Quirk, 2010). In the previous chapter, I found a subset of LA neurons also represents the original CS-US association even after extensive extinction (‘extinction-resistant fear neurons’). It has been believed that well-known inhibitory mechanisms

involving the prefrontal cortex (Milad and Quirk, 2002; Rosenkranz et al., 2003; Likhtik et al., 2005; Sotres-Bayon et al., 2006; Quirk and Mueller, 2008) and the ITC neurons (Chhatwal et al., 2005; Likhtik et al., 2008; Ehrlich et al., 2009) may provide inhibition at the BA or the CeM leading to the suppression of fear responses. Accordingly, re-appearance of fear memory after extinction has been regarded to be mediated by the context-dependent disinhibition of the inhibitory network over the amygdala (Hobin et al., 2003; Likhtik et al., 2008; Ehrlich et al., 2009). However, my results indicate that the essential brain regions constituting the inhibitory network, the prefrontal cortex and the amygdala ITC neurons play minor roles in extensive extinction, although the renewal of fear is normally observed. It is possible that the inhibitory network supports the expression of fear extinction in the beginning and additional extinction trainings recruit other brain network. Further researches are required to understand how the LA and other brain network support later savings or memory relapse after extensive extinction when the inhibitory influences of the prefrontal cortex disappeared. Metaplastic mechanisms that enable more rapid synaptic plasticity at input synapses may also support the enhanced potentiation of CS-responses in these neurons (Abraham, 2008; Lee et al., 2013). LA neurons representing the original fear memory after extensive extinction

may also play an important role in the persistence of fear memory and relapse after extinction (An et al., 2012).

Fear conditioning and extinction have served as primary models for the treatment of PTSD and other anxiety disorders. Although most PTSD research aimed at preventing the relapse of fear memory has focused on the dysfunctions or manipulations of the prefrontal cortex (Quirk et al., 2006; Sotres-Bayon et al., 2006), my results suggests that the inhibitory influences of the prefrontal cortex over the amygdala is no longer critical for the maintenance and the expression of extensive extinction. It is consistent with clinical studies which showed the connectivity between the prefrontal cortex and the amygdala progressively decreased with repetitive presentations of the traumatic script (Gilboa et al., 2004; Rauch et al., 2006). Further researches will be required to find appropriate targets for clinical treatment of fear-related mental disorders using the extensive extinction paradigm.

References

Abraham WC (2008) Metaplasticity: tuning synapses and networks for plasticity. Nature Reviews Neuroscience 9:387.

Amano T, Unal CT, Pare D (2010) Synaptic correlates of fear extinction in the amygdala. Nature Neuroscience 13:489-494.

Amir A, Amano T, Pare D (2011) Physiological identification and infralimbic responsiveness of rat intercalated amygdala neurons. Journal of Neurophysiology 105:3054-3066.

An B, Hong I, Choi S (2012) Long-term neural correlates of reversible fear learning in the lateral amygdala. The Journal of Neuroscience 32:16845-16856.

Barad M (2005) Fear extinction in rodents: basic insight to clinical promise. Current Opinion in Neurobiology 15:710-715.

Barad M (2006) Is extinction of fear erasure or inhibition? Why both, of course. Learning & Memory 13:108-109.

Blair HT, Schafe GE, Bauer EP, Rodrigues SM, LeDoux JE (2001) Synaptic plasticity in the lateral amygdala: a cellular hypothesis of fear conditioning. Learning & Memory 8:229-242.

Bordi F, LeDoux J, Clugnet MC, Pavlides C (1993) Single-unit activity in the lateral nucleus of the amygdala and overlying areas of the striatum in freely behaving rats: rates, discharge patterns, and

responses to acoustic stimuli. Behavioral Neuroscience 107:757-769.

Bouton ME (2002) Context, ambiguity, and unlearning: sources of relapse after behavioral extinction. Biological Psychiatry 52:976-986.

Bouton ME, King DA (1983) Contextual control of the extinction of conditioned fear: tests for the associative value of the context. Journal of Experimental Psychology 9:248-265.

Burgos-Robles A, Vidal-Gonzalez I, Quirk GJ (2009) Sustained conditioned responses in prelimbic prefrontal neurons are correlated with fear expression and extinction failure. The Journal of Neuroscience 29:8474-8482.

Burgos-Robles A, Vidal-Gonzalez I, Santini E, Quirk GJ (2007) Consolidation of fear extinction requires NMDA receptor-dependent bursting in the ventromedial prefrontal cortex. Neuron 53:871-880.

Carew TJ, Walters ET, Kandel ER (1981) Associative learning in Aplysia: cellular correlates supporting a conditioned fear hypothesis. Science 211:501-504.

Chang CH, Knapska E, Orsini CA, Rabinak CA, Zimmerman JM, Maren S (2009) Fear extinction in rodents. Current Protocols in Neuroscience 8:8-23.

Chhatwal JP, Myers KM, Ressler KJ, Davis M (2005) Regulation of

gephyrin and GABAA receptor binding within the amygdala after fear acquisition and extinction. The Journal of Neuroscience 25:502-506.

Collins DR, Pare D (2000) Differential fear conditioning induces reciprocal changes in the sensory responses of lateral amygdala neurons to the CS(+) and CS(-). Learning & Memory 7:97-103.

Corcoran KA, Maren S (2001) Hippocampal inactivation disrupts contextual retrieval of fear memory after extinction. The Journal of Neuroscience 21:1720-1726.

Corcoran KA, Quirk GJ (2007) Activity in prelimbic cortex is necessary for the expression of learned, but not innate, fears. The Journal of Neuroscience 27:840-844.

Corcoran KA, Desmond TJ, Frey KA, Maren S (2005) Hippocampal inactivation disrupts the acquisition and contextual encoding of fear extinction. The Journal of Neuroscience 25:8978-8987.

Dalton GL, Wang YT, Floresco SB, Phillips AG (2008) Disruption of AMPA receptor endocytosis impairs the extinction, but not acquisition of learned fear. Neuropsychopharmacology 33:2416-2426.

Davis M (1992) The role of the amygdala in fear and anxiety. Annual Review of Neuroscience 15:353-375.

Davis M, Whalen PJ (2001) The amygdala: vigilance and emotion. Molecular Psychiatry 6:13-34.

Davis M, Rainnie D, Cassell M (1994) Neurotransmission in the rat amygdala related to fear and anxiety. Trends in Neurosciences 17:208-214.

Duclos Y, Schmied A, Burle B, Burnet H, Rossi-Durand C (2008) Anticipatory changes in human motoneuron discharge patterns during motor preparation. The Journal of Physiology 586:1017-1028.

Ehrlich I, Humeau Y, Grenier F, Ciocchi S, Herry C, Luthi A (2009) Amygdala inhibitory circuits and the control of fear memory. Neuron 62:757-771.

Falls WA, Miserendino MJ, Davis M (1992) Extinction of fear-potentiated startle: blockade by infusion of an NMDA antagonist into the amygdala. The Journal of Neuroscience 12:854-863.

Fanselow MS (1994) Neural organization of the defensive behavior system responsible for fear. Psychonomic Bulletin & Review 1:429-438.

Fanselow MS (2000) Contextual fear, gestalt memories, and the hippocampus. Behavioural Brain Research 110:73-81.

Freedman LJ, Insel TR, Smith Y (2000) Subcortical projections of area 25 (subgenual cortex) of the macaque monkey. Journal of Comparative

Neurology 421:172-188.

Gale GD, Anagnostaras SG, Godsil BP, Mitchell S, Nozawa T, Sage JR, Wiltgen B, Fanselow MS (2004) Role of the basolateral amygdala in the storage of fear memories across the adult lifetime of rats. The Journal of Neuroscience 24:3810-3815.

Gilboa A, Shalev AY, Laor L, Lester H, Louzoun Y, Chisin R, Bonne O (2004) Functional connectivity of the prefrontal cortex and the amygdala in posttraumatic stress disorder. Biological Psychiatry 55:263-272.

Goosens KA, Hobin JA, Maren S (2003) Auditory-evoked spike firing in the lateral amygdala and Pavlovian fear conditioning: mnemonic code or fear bias? Neuron 40:1013-1022.

Hall J, Thomas KL, Everitt BJ (2001) Cellular imaging of zif268 expression in the hippocampus and amygdala during contextual and cued fear memory retrieval: selective activation of hippocampal CA1 neurons during the recall of contextual memories. The Journal of Neuroscience 21:2186-2193.

Han JH, Kushner SA, Yiu AP, Cole CJ, Matynia A, Brown RA, Neve RL, Guzowski JF, Silva AJ, Josselyn SA (2007) Neuronal competition and selection during memory formation. Science 316:457-460.

Han JH, Kushner SA, Yiu AP, Hsiang HL, Buch T, Waisman A, Bontempi B,

Neve RL, Frankland PW, Josselyn SA (2009) Selective erasure of a fear memory. Science 323:1492-1496.

Haubensak W, Kunwar PS, Cai H, Ciocchi S, Wall NR, Ponnusamy R, Biag J, Dong HW, Deisseroth K, Callaway EM, Fanselow MS, Luthi A, Anderson DJ (2010) Genetic dissection of an amygdala microcircuit that gates conditioned fear. Nature 468:270-276.

Herry C, Garcia R (2003) Behavioral and paired-pulse facilitation analyses of long-lasting depression at excitatory synapses in the medial prefrontal cortex in mice. Behavioural Brain Research 146:89-96.

Herry C, Ciocchi S, Senn V, Demmou L, Muller C, Luthi A (2008) Switching on and off fear by distinct neuronal circuits. Nature 454:600-606.

Hobin JA, Goosens KA, Maren S (2003) Context-dependent neuronal activity in the lateral amygdala represents fear memories after extinction. The Journal of Neuroscience 23:8410-8416.

Hobin JA, Ji J, Maren S (2006) Ventral hippocampal muscimol disrupts context-specific fear memory retrieval after extinction in rats. Hippocampus 16:174-182.

Hong I, Song B, Lee S, Kim J, Kim J, Choi S (2009) Extinction of cued fear memory involves a distinct form of depotentiation at cortical input synapses onto the lateral amygdala. The European Journal of

Neuroscience 30:2089-2099.

Hong I, Kim J, Lee J, Park S, Song B, Kim J, An B, Park K, Lee HW, Lee S, Kim H, Park SH, Eom KD, Lee S, Choi S (2011) Reversible plasticity of fear memory-encoding amygdala synaptic circuits even after fear memory consolidation. PloS One 6:e24260.

Jackson A, Fetz EE (2007) Compact movable microwire array for long-term chronic unit recording in cerebral cortex of primates. Journal of Neurophysiology 98:3109-3118.

Jungling K, Seidenbecher T, Sosulina L, Lesting J, Sangha S, Clark SD, Okamura N, Duangdao DM, Xu YL, Reinscheid RK, Pape HC (2008) Neuropeptide S-mediated control of fear expression and extinction: role of intercalated GABAergic neurons in the amygdala. Neuron 59:298-310.

Kapp BS, Frysinger RC, Gallagher M, Haselton JR (1979) Amygdala central nucleus lesions: effect on heart rate conditioning in the rabbit. Physiology & Behavior 23:1109-1117.

Kehoe EJ (1988) A layered network model of associative learning: learning to learn and configuration. Psychological Review 95:411-433.

Kim J, Song B, Hong I, Kim J, Lee J, Park S, Eom JY, Lee CJ, Lee S, Choi S (2010) Reactivation of fear memory renders consolidated amygdala synapses labile. The Journal of Neuroscience 30:9631-

9640.

Kim J, Lee S, Park K, Hong I, Song B, Son G, Park H, Kim WR, Park E, Choe HK, Kim H, Lee C, Sun W, Kim K, Shin KS, Choi S (2007) Amygdala depotentiation and fear extinction. Proceedings of the National Academy of Sciences of the United States of America 104:20955-20960.

Knapska E, Maren S (2009) Reciprocal patterns of c-Fos expression in the medial prefrontal cortex and amygdala after extinction and renewal of conditioned fear. Learning & Memory 16:486-493.

LaBar KS, LeDoux JE, Spencer DD, Phelps EA (1995) Impaired fear conditioning following unilateral temporal lobectomy in humans. The Journal of Neuroscience 15:6846-6855.

Lau HL, Timbers TA, Mahmoud R, Rankin CH (2013) Genetic dissection of memory for associative and non-associative learning in Caenorhabditis elegans. Genes, Brain, and Behavior 12:210-223.

Laurent V, Westbrook RF (2009) Inactivation of the infralimbic but not the prelimbic cortex impairs consolidation and retrieval of fear extinction. Learning & Memory 16:520-529.

LeDoux JE (2000) Emotion circuits in the brain. Annual Review of Neuroscience 23:155-184.

LeDoux JE (2014) Coming to terms with fear. Proceedings of the National Academy of Sciences of the United States of America 111:2871-2878.

LeDoux JE, Farb C, Ruggiero DA (1990) Topographic organization of neurons in the acoustic thalamus that project to the amygdala. The Journal of Neuroscience 10:1043-1054.

Lee S, Song B, Kim J, Park K, Hong I, An B, Song S, Lee J, Park S, Kim J, Park D, Lee CJ, Kim K, Shin KS, Tsien RW, Choi S (2013) GluA1 phosphorylation at serine 831 in the lateral amygdala is required for fear renewal. Nature Neuroscience 16:1436-1444.

Likhtik E, Pelletier JG, Paz R, Pare D (2005) Prefrontal control of the amygdala. The Journal of Neuroscience 25:7429-7437.

Likhtik E, Popa D, Apergis-Schoute J, Fidacaro GA, Pare D (2008) Amygdala intercalated neurons are required for expression of fear extinction. Nature 454:642-645.

Likhtik E, Stujenske JM, Topiwala MA, Harris AZ, Gordon JA (2014) Prefrontal entrainment of amygdala activity signals safety in learned fear and innate anxiety. Nature Neuroscience 17:106-113.

Lin HC, Mao SC, Gean PW (2009) Block of gamma-aminobutyric acid-A receptor insertion in the amygdala impairs extinction of conditioned fear. Biological Psychiatry 66:665-673.

Lolordo VM (1969) Positive conditioned reinforcement from aversive situations. Psychological Bulletin 72:193-203.

Maren S (2000) Auditory fear conditioning increases CS-elicited spike firing in lateral amygdala neurons even after extensive overtraining. The European Journal of Neuroscience 12:4047-4054.

Maren S (2011) Seeking a spotless mind: extinction, deconsolidation, and erasure of fear memory. Neuron 70:830-845.

Maren S, Fanselow MS (1996) The amygdala and fear conditioning: has the nut been cracked? Neuron 16:237-240.

Maren S, Quirk GJ (2004) Neuronal signalling of fear memory. Nature Reviews Neuroscience 5:844-852.

McDonald AJ (1982) Neurons of the lateral and basolateral amygdaloid nuclei: a Golgi study in the rat. The Journal of Comparative Neurology 212:293-312.

McDonald AJ, Mascagni F, Guo L (1996) Projections of the medial and lateral prefrontal cortices to the amygdala: a Phaseolus vulgaris leucoagglutinin study in the rat. Neuroscience 71:55-75.

McEchron MD, McCabe PM, Green EJ, Llabre MM, Schneiderman N (1995) Simultaneous single unit recording in the medial nucleus of the medial geniculate nucleus and amygdaloid central nucleus

throughout habituation, acquisition, and extinction of the rabbit's classically conditioned heart rate. Brain Research 682:157-166.

McGaugh JL (2000) Memory--a century of consolidation. Science 287:248-251.

McKernan MG, Shinnick-Gallagher P (1997) Fear conditioning induces a lasting potentiation of synaptic currents in vitro. Nature 390:607-611.

Medina JF, Repa JC, Mauk MD, LeDoux JE (2002) Parallels between cerebellum- and amygdala-dependent conditioning. Nature Reviews Neuroscience 3:122-131.

Milad MR, Quirk GJ (2002) Neurons in medial prefrontal cortex signal memory for fear extinction. Nature 420:70-74.

Milad MR, Quirk GJ (2012) Fear extinction as a model for translational neuroscience: ten years of progress. Annual Review of Psychology 63:129-151.

Miserendino MJ, Sananes CB, Melia KR, Davis M (1990) Blocking of acquisition but not expression of conditioned fear-potentiated startle by NMDA antagonists in the amygdala. Nature 345:716-718.

Myers KM, Davis M (2007) Mechanisms of fear extinction. Molecular Psychiatry 12:120-150.

Myers KM, Ressler KJ, Davis M (2006) Different mechanisms of fear

extinction dependent on length of time since fear acquisition. Learning & Memory 13:216-223.

Napier RM, Macrae M, Kehoe EJ (1992) Rapid reaquisition in conditioning of the rabbit's nictitating membrane response. Journal of Experimental Psychology 18:182-192.

Nicolelis MA, Dimitrov D, Carmena JM, Crist R, Lehew G, Kralik JD, Wise SP (2003) Chronic, multisite, multielectrode recordings in macaque monkeys. Proceedings of the National Academy of Sciences of the United States of America 100:11041-11046.

Oyama K, Hernadi I, Iijima T, Tsutsui K (2010) Reward prediction error coding in dorsal striatal neurons. The Journal of Neuroscience 30:11447-11457.

Pape HC, Pare D (2010) Plastic synaptic networks of the amygdala for the acquisition, expression, and extinction of conditioned fear. Physiological Reviews 90:419-463.

Pare D, Smith Y (1993a) Distribution of GABA immunoreactivity in the amygdaloid complex of the cat. Neuroscience 57:1061-1076.

Pare D, Smith Y (1993b) The intercalated cell masses project to the central and medial nuclei of the amygdala in cats. Neuroscience 57:1077-1090.

Pare D, Gaudreau H (1996) Projection cells and interneurons of the lateral and basolateral amygdala: distinct firing patterns and differential relation to theta and delta rhythms in conscious cats. The Journal of Neuroscience 16:3334-3350.

Pare D, Collins DR (2000) Neuronal correlates of fear in the lateral amygdala: multiple extracellular recordings in conscious cats. The Journal of Neuroscience 20:2701-2710.

Pare D, Duvarci S (2012) Amygdala microcircuits mediating fear expression and extinction. Current Opinion in Neurobiology 22:717-723.

Pare D, Quirk GJ, Ledoux JE (2004) New vistas on amygdala networks in conditioned fear. Journal of Neurophysiology 92:1-9.

Paton JJ, Belova MA, Morrison SE, Salzman CD (2006) The primate amygdala represents the positive and negative value of visual stimuli during learning. Nature 439:865-870.

Pearce JM, Hall G (1980) A model for Pavlovian learning: variations in the effectiveness of conditioned but not of unconditioned stimuli. Psychological Review 87:532-552.

Phelps EA, LeDoux JE (2005) Contributions of the amygdala to emotion processing: from animal models to human behavior. Neuron 48:175-187.

Phillips RG, LeDoux JE (1992) Differential contribution of amygdala and hippocampus to cued and contextual fear conditioning. Behavioral Neuroscience 106:274-285.

Pitkanen A, Savander V, LeDoux JE (1997) Organization of intra-amygdaloid circuitries in the rat: an emerging framework for understanding functions of the amygdala. Trends in Neurosciences 20:517-523.

Quirk GJ, Mueller D (2008) Neural mechanisms of extinction learning and retrieval. Neuropsychopharmacology 33:56-72.

Quirk GJ, Repa C, LeDoux JE (1995) Fear conditioning enhances short-latency auditory responses of lateral amygdala neurons: parallel recordings in the freely behaving rat. Neuron 15:1029-1039.

Quirk GJ, Armony JL, LeDoux JE (1997) Fear conditioning enhances different temporal components of tone-evoked spike trains in auditory cortex and lateral amygdala. Neuron 19:613-624.

Quirk GJ, Garcia R, Gonzalez-Lima F (2006) Prefrontal mechanisms in extinction of conditioned fear. Biological Psychiatry 60:337-343.

Ranganath C, Rainer G (2003) Neural mechanisms for detecting and remembering novel events. Nature Reviews Neuroscience 4:193-202.

Rauch SL, Shin LM, Phelps EA (2006) Neurocircuitry models of

posttraumatic stress disorder and extinction: human neuroimaging research--past, present, and future. Biological Psychiatry 60:376-382.

Reijmers LG, Perkins BL, Matsuo N, Mayford M (2007) Localization of a stable neural correlate of associative memory. Science 317:1230-1233.

Repa JC, Muller J, Apergis J, Desrochers TM, Zhou Y, LeDoux JE (2001) Two different lateral amygdala cell populations contribute to the initiation and storage of memory. Nature Neuroscience 4:724-731.

Rescorla RA (2001) Retraining of extinguished Pavlovian stimuli. Journal of Experimental Psychology 27:115-124.

Rogan MT, Staubli UV, LeDoux JE (1997) Fear conditioning induces associative long-term potentiation in the amygdala. Nature 390:604-607.

Rogan MT, Leon KS, Perez DL, Kandel ER (2005) Distinct neural signatures for safety and danger in the amygdala and striatum of the mouse. Neuron 46:309-320.

Romanski LM, Clugnet MC, Bordi F, LeDoux JE (1993) Somatosensory and auditory convergence in the lateral nucleus of the amygdala. Behavioral Neuroscience 107:444-450.

Rosenkranz JA, Moore H, Grace AA (2003) The prefrontal cortex regulates

lateral amygdala neuronal plasticity and responses to previously conditioned stimuli. The Journal of Neuroscience 23:11054-11064.

Royer S, Martina M, Pare D (1999) An inhibitory interface gates impulse traffic between the input and output stations of the amygdala. The Journal of Neuroscience 19:10575-10583.

Sacco T, Sacchetti B (2010) Role of secondary sensory cortices in emotional memory storage and retrieval in rats. Science 329:649-656.

Santini E, Quirk GJ, Porter JT (2008) Fear conditioning and extinction differentially modify the intrinsic excitability of infralimbic neurons. The Journal of Neuroscience 28:4028-4036.

Schafe GE, LeDoux JE (2000) Memory consolidation of auditory pavlovian fear conditioning requires protein synthesis and protein kinase A in the amygdala. The Journal of Neuroscience 20:RC96.

Schafe GE, Atkins CM, Swank MW, Bauer EP, Sweatt JD, LeDoux JE (2000) Activation of ERK/MAP kinase in the amygdala is required for memory consolidation of pavlovian fear conditioning. The Journal of Neuroscience 20:8177-8187.

Schoenbaum G, Chiba AA, Gallagher M (1999) Neural encoding in orbitofrontal cortex and basolateral amygdala during olfactory discrimination learning. The Journal of Neuroscience 19:1876-1884.

Seidenbecher T, Laxmi TR, Stork O, Pape HC (2003) Amygdalar and hippocampal theta rhythm synchronization during fear memory retrieval. Science 301:846-850.

Sesack SR, Deutch AY, Roth RH, Bunney BS (1989) Topographical organization of the efferent projections of the medial prefrontal cortex in the rat: an anterograde tract-tracing study with Phaseolus vulgaris leucoagglutinin. The Journal of Comparative Neurology 290:213-242.

Sotres-Bayon F, Quirk GJ (2010) Prefrontal control of fear: more than just extinction. Current Opinion in Neurobiology 20:231-235.

Sotres-Bayon F, Cain CK, LeDoux JE (2006) Brain mechanisms of fear extinction: historical perspectives on the contribution of prefrontal cortex. Biological Psychiatry 60:329-336.

Sotres-Bayon F, Bush DE, LeDoux JE (2007) Acquisition of fear extinction requires activation of NR2B-containing NMDA receptors in the lateral amygdala. Neuropsychopharmacology 32:1929-1940.

Sotres-Bayon F, Diaz-Mataix L, Bush DE, LeDoux JE (2009) Dissociable roles for the ventromedial prefrontal cortex and amygdala in fear extinction: NR2B contribution. Cerebral Cortex 19:474-482.

Tseng WT, Yen CT, Tsai ML (2011) A bundled microwire array for long-term chronic single-unit recording in deep brain regions of behaving

rats. Journal of Neuroscience Methods 201:368-376.

Tsvetkov E, Carlezon WA, Benes FM, Kandel ER, Bolshakov VY (2002) Fear conditioning occludes LTP-induced presynaptic enhancement of synaptic transmission in the cortical pathway to the lateral amygdala. Neuron 34:289-300.

Zhou Y, Won J, Karlsson MG, Zhou M, Rogerson T, Balaji J, Neve R, Poirazi P, Silva AJ (2009) CREB regulates excitability and the allocation of memory to subsets of neurons in the amygdala. Nature Neuroscience 12:1438-1443.

국문초록

공포학습 시 편도체 및 변연계아래피질의 신경활성

서울대학교 대학원
협동과정 뇌과학 전공
안 보 배

중성적 자극과 유해한 자극을 반복적으로 제시하여 이를 연합하는 공포 조건화 학습 방법은 외상 후 스트레스 장애 등 공포관련 질환의 동물 모델로 유용하게 사용되어 왔다. 과거 수많은 연구자들은 공포 조건화 학습 모델을 이용하여 편도체 및 그와 연결된 신경네트워크가 공포 학습 및 소거에 필수적임을 제안하였다.

그러나 이전 연구들은 단기 공포 학습 모델을 이용함으로써, 공포 학습 및 소거가 편도체 및 신경네트워크에 미치는 장기적 영향에 대해서는 밝히지 못하였다. 그러므로 본 연구에서는 장기 공포 학습 및 반복 소거 학습이 편도체 및 신경네트워크에 미치는 영향을 살펴보고자 하였다. 제 1장에서는 장기 공포학습 및 소거, 재학습 동안 공포 연합 학습의 중추로 알려진 등쪽 편도체 내 신경세포의 활성을 관찰하였다. 일련의 실험을 통하여 등쪽 편도체 내 신경세포들이 역동적으로 변화하는 공포 연합 기억을 표상함을 발견하였다. 나아가, 등쪽 편도체 내 공포 소거 학습 기억을 표상하는 집단(공포 소거 순응 신경세포)과 공포 소거 학습 기억을 표상하지 않는 집단 (공포 소거 저항 신경세포)이 있음을 발견하였다. 이러한 결과는 등쪽 편도체가 공포 조건화 학습의 다양한 측면을 표상함을 의미한다.

제 2장에서는 장기 공포학습 및 반복 소거 학습 동안 공포 소거 학습의 중추로 알려진 편도체 및 변연계아래피질의 활성을 관찰하였다. 일련의 실험을 통하여 변연계아래피질 신경세포들이 단일 공포 소거 기억은 표상하지만, 반복 공포 소거 기억은 표상하지 않음을 발견하였다. 또한 편도체 내 억제 신경세포의 활성이 반복 소거 학습 시 필요하지 않음을 발견하였다. 이러한 결과는 단일 및 반복 공포 소거 학습이 다른 신경학적 기전에 의해 매개됨을 의미한다.

요약적으로, 본 연구는 장기 공포 학습이 편도체 및 신경네

트워크에 미치는 영향을 살펴보았다. 먼저, 등쪽 편도체 신경세포가 공포 연합 기억의 다양한 측면을 역동적으로 표상함을 관찰하였다. 다음으로, 편도체와 변연계아래피질의 신경 활성이 단일 공포 소거 학습에는 중요하지만, 반복 소거 학습에는 필요하지 않음을 발견하였다. 이러한 결과들은 공포 기억이 조절되는 신경학적 기반에 대한 이해를 도모하고, 나아가 공포 관련 정신 질환 치료의 기반을 제시한다.

핵심어: 편도체, 변연계아래피질, 공포 조건화 학습, 공포 소거 학습

학번: 2009-20445

이학박사학위논문

편도체 및 변연계아래피질 내 공포학습관련 신경활성

Neural correlates of fear learning in the amygdala and the infralimbic cortex

2014년 7월

서울대학교 대학원
협동과정 뇌과학 전공
안 보 배

Abstract

Neural correlates of fear learning in the amygdala and the infralimbic cortex

Bobae An

Interdisciplinary Program of Neuroscience

The Graduate School

Seoul National University

Fear is one of the most intensely studied fields in emotion, due to its simple and well-known animal model, the Pavlovian fear conditioning. Numerous studies have reported that the amygdala and its surrounding brain network are critically involved in fear conditioning and extinction. However, the long-term effects of fear learning have remained largely unknown since most of the previous studies used behavioral paradigms in which memory

retrieval was tested only in the short-term. Therefore, I employed a fear learning paradigm that consists of fear conditioning and extensive extinction that spans several days.

In the first chapter, I examined how neurons in the lateral amygdala (LA), a key brain structure of fear associative learning, represents fear memory during fear conditioning and subsequent extensive extinction, reconditioning. I found that the ensemble activity of LA neurons correlated tightly with conditioned fear responses of rats in the reconditioning paradigm. Further analysis revealed that among the LA neurons that displayed increased responses to the CS after fear conditioning, some exhibited weakened responses after extinction (extinction-sensitive), whereas others remained potentiated (extinction-resistant) after extinction. These results suggest the existence of distinct neuronal populations that encode various facets of fear memory and provide insights into the neuronal mechanisms underlying fear memory modulation.

In the second chapter, I questioned whether the inhibitory network, which consists of the infralimbic cortex (IL) and the intercalated amygdala cells (ITC), is crucial for fear extinction, represents long-term correlates of fear learning that consisted of fear conditioning and extensive extinction. Single unit recordings and biochemical lesion techniques were employed to

investigate the long-term effects of fear learning. I found that the CS-responses of IL neurons which emerged after single extinction dissipated with additional extinction. In keeping with this, I also found that ITC lesions that impaired the retrieval of extinction caused no deficit if lesions were made after multiple extinction sessions. These results suggest that single and extensive extinction involves different neural mechanisms.

In summary, I investigated the long-term neural correlates of fear learning involving extensive extinction and reconditioning. First, LA neuronal population represented dynamic changes in CS-US association, while distinct sub-populations encoding various aspects of fear learning existed. Second, IL neurons and ITC activities were critical for single extinction, but not for extensive extinction. Together, these findings provide insights into the neural mechanisms underlying fear memory modulation and the treatment of fear-related mental disorders.

Key words: Lateral amygdala, Infralimbic cortex, Intercalated amygdala neurons, fear conditioning, fear extinction

Student number*:* 2009-20445

Contents

Page

Chapter 2. Neural correlates of extensive extinction learning in the infralimbic cortex and the amygdala intercalated neurons

List of Figures

Page

Chapter 1.

Chapter 2.

Background and Purpose

1. Background

1.1. Pavlovian fear conditioning

1.1.1. Characteristics of Pavlovian fear conditioning

Fear is one of the most vigorously and extensively studied fields in emotion, due to the presence of a well-verified animal model, the Pavlovian fear conditioning. When a neutral stimulus (Conditioned stimulus, CS), often a tone, is repeatedly presented with a noxious stimulus (Unconditioned stimulus, US), such as a foot shock, animals quickly learn that the CS is a predictive signal of an aversive event (Fig. 1). As a result, CS elicits defensive behavior, freezing and physiological alterations in heart rate, blood pressure and hormones, controlled by the autonomic nervous system or the endocrine system (Kapp et al., 1979; Davis, 1992; LeDoux, 2014).

Pavlovian fear conditioning has been a useful tool for studying the underlying mechanisms of fear-related mental disorders, such as post-traumatic stress disorders (PTSD) and phobias (Davis, 1992; LeDoux, 2000; Davis and Whalen, 2001). The model can be utilized across a wide range of animals, from vertebrates to invertebrates (Carew et al., 1981; LeDoux, 2000; Lau et al., 2013). It is readily and rapidly acquired, even with one CS presentation paired with a noxious stimulus (Fanselow, 1994). Once

established, fear memory is firm and long-lasting, often persists throughout life (Gale et al., 2004).

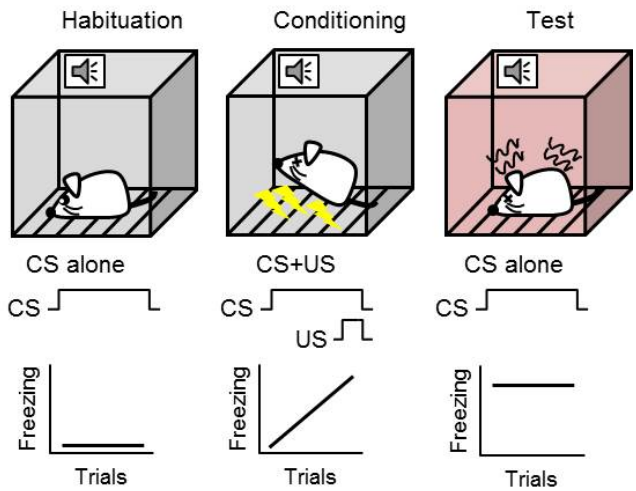

Figure 1. Pavlovian fear conditioning in rodents. Rats do not show fear responses to a neutral tone (CS) during the habituation session. After fear conditioning is performed by presenting the CS with a noxious foot shock (US) repeatedly, rats show fear responses to the tone, even if CS is presented without a shock.

1.1.2. Neural mechanisms underlying fear conditioning

Amygdala. A large body of evidence suggests the amygdala as the locus of fear memory storage and modulation (Davis, 1992; LeDoux, 2000; Pare and Duvarci, 2012), especially in the case of auditory fear conditioning (Fig. 2). Both experimentally amygdala-lesioned animals and human patients whose amygdala is damaged show deficits in acquiring the CS-US association (Phillips and LeDoux, 1992; LaBar et al., 1995; Phelps and LeDoux, 2005). Auditory thalamus and cortical inputs to the amygdala are potentiated after fear conditioning (McKernan and Shinnick-Gallagher, 1997; Quirk et al., 1997), resulting increased output signal to the downstream so as to evoke aversive behavior (Davis and Whalen, 2001). Accordingly, it has been reported that tone-evoked neural activity in the amygdala increases after fear conditioning, and decreases after closely following extinction (Quirk et al., 1997; Rogan et al., 1997), correlates well with the behavioral fear responses.

The rodent amygdala consists of distinct sub-regions (Pitkanen et al., 1997). Particularly, the lateral, basal and central part of the amygdala has been critically involved in fear and anxiety. The lateral amygdala (LA) is the main target of sensory afferents from the thalamus and cortex. Accordingly, LA neurons respond to auditory and somatosensory stimuli with short

latencies, as fast as 10 ms (Bordi et al., 1993; Quirk et al., 1995). LA has been regarded as the locus where CS-US association occurs since auditory and somatosensory information converges in the region (Bordi et al., 1993; Romanski et al., 1993). Auditory fear conditioning increases CS-responses of LA neurons (Quirk et al., 1995; Repa et al., 2001; An et al., 2012). The central amygdala (CeA) is the main output region of the amygdala. It receives inputs from the LA and the basal amygdala and sends outputs to the brainstem and the hypothalamus to control autonomic and behavioral responses (Maren and Fanselow, 1996; Pitkanen et al., 1997). Recently, it has been reported that CeA neurons respond to auditory CS and fear learning modulate CS-responses of CeA (Haubensak et al., 2010). The basal amygdala (BA) is believed to modulate CS-US association since it is reciprocally connected with various sub-cortical and cortical regions. Particularly, it receives inputs from the medial prefrontal cortex (mPFC) and the hippocampus, which are the regions involved fear extinction and contextual information processing, respectively (Maren and Fanselow, 1996; Maren and Quirk, 2004; Herry et al., 2008). The BA also interacts with neuromodulatory system, such as noradrenergic and cholinergic system, and influences on fear memory consolidation (McGaugh, 2000).

Other cortical areas. There has been accumulating evidence that

other cortical areas, such as the hippocampus, medial prefrontal cortex (mPFC) and sensory cortices also participate in fear conditioning.

The hippocampus is critical for learning the association between a neutral context and a fearful event. Hippocampal lesioned animals show deficits in contextual fear conditioning, where a neutral context is associated with a noxious foot shock, while no deficit in auditory cued fear learning (Phillips and LeDoux, 1992). It is believed that hippocampus provides more complicated CS information, which is not processed in the level of sensory thalamus, to the BA (Fanselow, 2000). Increased theta synchronization between the hippocampus and the LA during the retrieval of fear memory has also been reported, suggesting that the functional connectivity between the hippocampus and the amygdala is important for the storage and the expression of fear memory (Seidenbecher et al., 2003).

The dorsomedial part of mPFC, prelimbic cortex (PL) has also been implicated in the expression of fear memory, whereas its ventral part, infralimbic cortex (IL) is involved in fear extinction (Sotres-Bayon and Quirk, 2010). The two sub-regions of the mPFC are believed to modulate fear responses bidirectionally through their divergent projections to the amygdala. PL supports the expression of fear memory via its excitatory connection to the BA (Milad and Quirk, 2012). PL inactivation impairs fear

learning (Corcoran and Quirk, 2007; Laurent and Westbrook, 2009) and CS activates PL neurons after fear conditioning (Santini et al., 2008; Burgos-Robles et al., 2009). PL neurons show sustained increased activity that mirrors the time course of freezing responses, lasting tens of seconds (Burgos-Robles et al., 2009). Secondary sensory cortices also have been critically involved in the storage of remote fear memory (Sacco and Sacchetti, 2010). Secondary sensory cortices lesions abolish one-month-old memory, whereas recently formed memories remain intact.

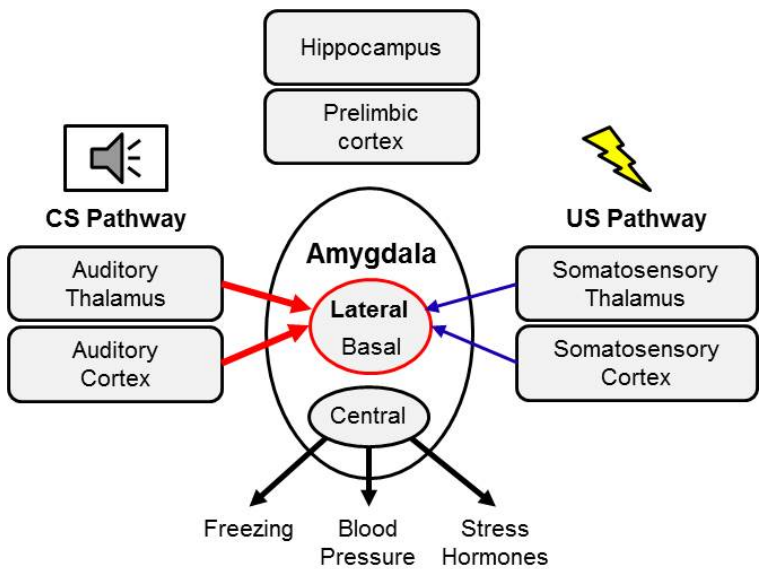

Figure 2. Brain regions involved in fear conditioning. The lateral and the basal amygdala receive sensory information of CS and US from thalamic and cortical areas. The central amygdala sends outputs to brainstem to control behavioral and autonomic responses to the CS.

1.2. Fear extinction

1.2.1. Characteristics of fear extinction

Repeated presentations of the CS in the absence of harmful stimuli, foot shocks, lead to a weakening of conditioned fear response, eventually to the point where fearful responses disappear (Fig. 3). This phenomenon is termed as fear extinction and has been a useful animal model of the exposure therapy, the most common and useful treatment for aberrant fear memory-related disorders, such as PTSD and phobia (Quirk et al., 2006; Maren, 2011).

Fear extinction is gradually acquired, unlike fear conditioning, requiring numerous CS presentations without noxious stimuli (Myers and Davis, 2007). Extinction memory is formed in a highly context-dependent manner, thus it is retrieved only in the same context where extinction learning has occurred (Bouton, 2002; Maren and Quirk, 2004). In another context, however, conditioned fear responses reappear even after extensive extinction, a phenomenon termed fear renewal, suggesting substantial remnants of the originally learned fear survive even after extensive extinction (Bouton, 2002; Chang et al., 2009). Moreover, extinction memory is less stable than fear memory, thus fear responses spontaneously

reappeared weeks after extinction training. It also supports the notion that original fear memory is not erased during fear extinction, rather inhibited temporarily (Maren and Quirk, 2004). The remnants of the original fear memory also support relearning which occurs more rapidly and with a lower threshold, compared to the initial fear learning.

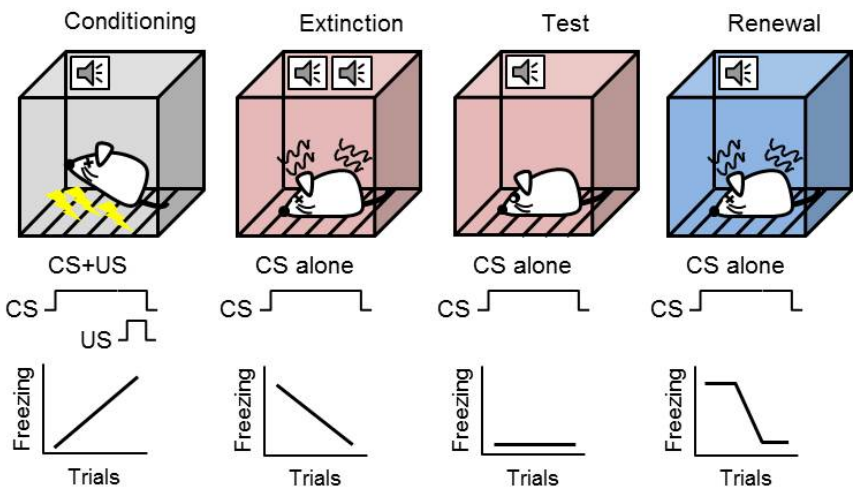

Figure 3. Fear extinction in rodents. Numerous presentations of CS alone, in fear extinction, decrease fear responses to the CS. However, conditioned fear responses reappear in various circumstances. For example, fear responses can be renewed when the rats are exposed in another context, different from the context where extinction learning has occurred.

1.2.2. Neural mechanisms underlying fear extinction

Prefrontal cortex. The ventromedial part of the medial prefrontal cortex, infralimbic cortex (IL) has been considered as a critical regulator of fear extinction, which inhibits conditioned fear behavior after extinction (Quirk et al., 2006; Sotres-Bayon and Quirk, 2010) (Fig. 4). Thalamic and hippocampal inputs to the IL are potentiated after fear extinction (Herry and Garcia, 2003), which are relayed to the medial subnuclei of the central amygdala (CeM) via the BA and amygdala-intercalated neurons to inhibit conditioned fear behavior (Maren and Quirk, 2004; Haubensak et al., 2010; Pape and Pare, 2010; Amir et al., 2011). NMDA receptor blockers infused into the IL immediately following extinction impair the retrieval of extinction, suggesting that neuronal plasticity in the IL is critical for the consolidation of extinction memory (Miserendino et al., 1990; Falls et al., 1992; Sotres-Bayon et al., 2007). Accordingly, IL neuronal activities are potentiated in animals that successfully retrieved with extinction (Milad and Quirk, 2002; Knapska and Maren, 2009) and stimulation of IL facilitates extinction (Milad and Quirk, 2002).

Amygdala. The amygdala is also critical in fear extinction. NMDA receptor blockers infused into the amygdala impair both fear conditioning

and extinction, suggesting that neuronal plasticity in the amygdala is crucial for both events (Miserendino et al., 1990; Falls et al., 1992; Sotres-Bayon et al., 2007). Similar to fear conditioning, sub-divisions of the amygdala also represent various aspects of fear extinction. LA neurons show decreased responses to the CS after extinction (Quirk et al., 1995), same as the CeA neurons (McEchron et al., 1995). However, some LA neurons retain CS-responses after fear extinction, representing the original fear memory (Repa et al., 2001; An et al., 2012). A neuronal population in the BA starts to signal the CS after extinction, named extinction neurons, suggesting BA plays a unique role in fear extinction (Herry et al., 2008). Extinction also induces depotentiation at LA input synapses (Kim et al., 2007; Dalton et al., 2008; Hong et al., 2009), and enhances local inhibitory signals (Chhatwal et al., 2005; Lin et al., 2009), all leading to decreased fear-related behavior.

Importantly, intercalated amygdala neurons (ITC), a probable mediator of prefrontal inhibition over the amygdala (Royer et al., 1999; Pape and Pare, 2010; Pare and Duvarci, 2012) are critically involved in fear extinction. ITCs are densely packed clusters of cells, mostly GABAergic neurons that surround the BLA. ITC clusters that are located between the BLA and the CeA have been implicated in fear extinction and thus described further, whereas the involvement of ITC clusters which lie between the BLA

and the cerebral cortex is elusive (Pare and Duvarci, 2012). ITC clusters at BLA-CeA border receive a dense projection from the IL and the BA (Sesack et al., 1989; McDonald et al., 1996; Freedman et al., 2000) and send its inhibitory outputs to the CeM (Pare and Smith, 1993a, b). Fear extinction potentiates BA inputs to the ITC cells that project to the CeM and this requires IL neuronal activities (Amano et al., 2010). ITC lesions impair the recall of extinction and activation of ITC cells facilitates extinction learning (Jungling et al., 2008; Likhtik et al., 2008).

Hippocampus. The hippocampus has been implicated in contextual modulation of fear extinction. Context-dependency of fear extinction is impaired if the hippocampus is inactivated before extinction training (Corcoran and Maren, 2001; Corcoran et al., 2005). Hippocampal inactivation also disrupts the context-dependent reappearance of fear after extinction, fear renewal (Corcoran and Maren, 2001; Hobin et al., 2003; Corcoran et al., 2005; Hobin et al., 2006).

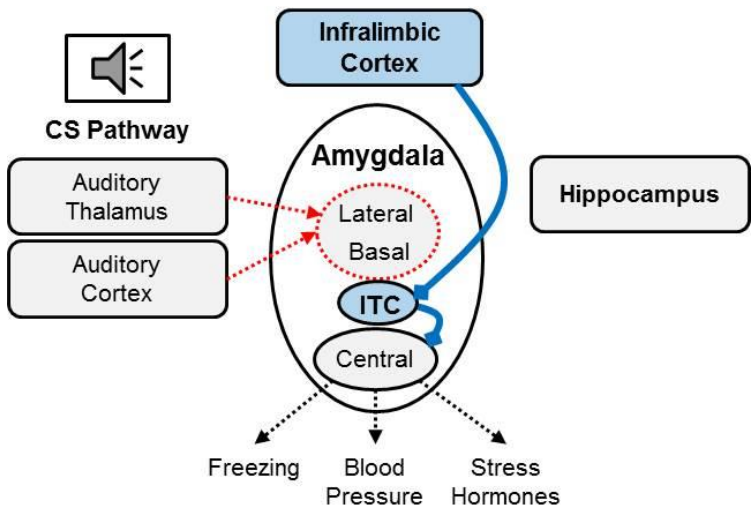

Figure 4. Brain regions involved in fear extinction. The infralimbic cortex sends its inhibitory controls over the amygdala via intercalated amygdala neurons (ITC) in the amygdala to suppress conditioned fear responses. Synaptic inputs to the lateral amygdala are also weakened by extinction. Hippocampus is implicated in contextual modulation of fear extinction.

2. Purpose

Fear is one of the most intensely studied fields in emotion, due to its simple and well-known animal model, the Pavlovian fear conditioning. Numerous studies have reported that the amygdala and its inputs and outputs are critically involved in fear conditioning and extinction. However, the long-term effects of fear learning have remained largely unknown since most of the previous studies employed a short behavioral paradigm that consists of fear conditioning and single extinction session. Therefore, I employed a fear conditioning paradigm that consists of fear conditioning and extensive extinction, spanning several days.

In the first chapter, I examined how neurons in the LA, a key brain structure where CS-US association takes place, represent the long-term correlates of fear learning which consists of fear conditioning, extinction and reconditioning. In the second chapter, I questioned whether the inhibitory network which is critically involved in fear extinction, including the prefrontal cortex and intercalated amygdala cell masses, represents the long-term correlates of fear learning encompassing fear conditioning and extensive extinction. To investigate the long-term effects of fear learning, single unit recordings and biochemical lesion techniques were employed. Together, these questions and answers could provide insights into the neural

mechanisms underlying fear memory modulation and the treatment of fear-related mental disorders.

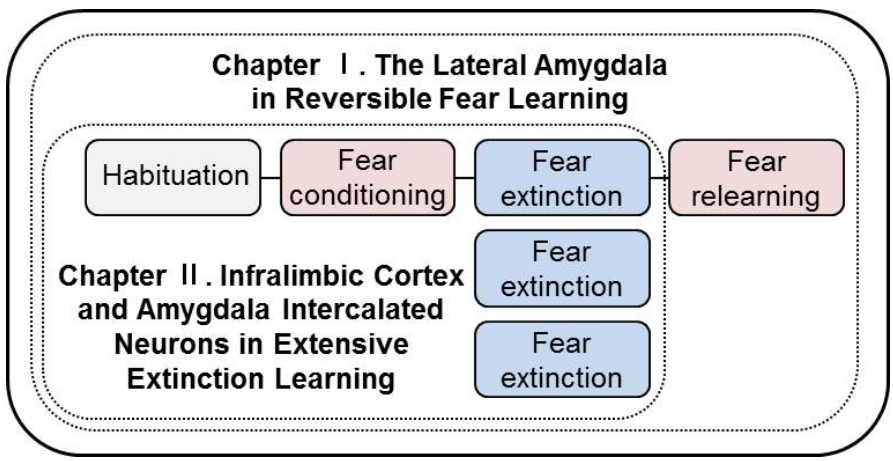

Figure 5. Schematic diagram of thesis. In chapter 1, neuronal activity of the lateral amygdala was examined in fear learning that consisted of fear conditioning, extensive extinction and reconditioning. In chapter 2, activities of the infralimbic cortex and the amygdala intercalated neurons were examined in fear conditioning and extensive extinction.

Chapter 1.

Long-term neural correlates of reversible fear learning in the lateral amygdala

Abstract

The lateral amygdala (LA) is a primary locus of auditory cued fear memory storage. LA neuronal responses to conditioned stimuli (CS) increase after fear conditioning and decrease during closely following extinction. However, the long-term effects of repeated fear conditioning and extinction on firing patterns of LA neurons have not been fully explored. Here I show, using single unit recording techniques, that the ensemble activity of LA neurons correlates tightly with behavioral fear responses of rats in fear conditioning, extensive extinction and reconditioning. The CS-evoked LA ensemble activity increased after fear conditioning, decreased after extinction, and was re-potentiated after reconditioning. Further analysis revealed that among the LA neurons that displayed increased CS-responses after fear conditioning, some showed weakened responses after extinction (extinction-sensitive), whereas others remained potentiated (extinction-resistant) after extensive extinction. The majority of extinction-sensitive neurons exhibited strong potentiation after reconditioning, suggesting that this distinct sub-population (‘reversible fear neurons’) dynamically encodes updated CS-US association strength. Interestingly, these reversible fear

neurons displayed more rapid potentiation during reconditioning compared to the initial fear conditioning, providing a neural correlate of ‘savings’ after extinction. In contrast, the extinction-resistant fear neurons did not show further increases after reconditioning, suggesting that this sub-population encodes persistent fear memory representing the original CS-US association. These results constitute the first longitudinal report on LA neuronal activity during reversible fear learning and provide insight into the neuronal mechanisms underlying fear memory modulation.

Key words: Lateral amygdala, fear conditioning, fear extinction

Introduction

Fear conditioning is the association between a neutral CS and an aversive unconditioned stimulus (US), which leads to fear responses to CS-alone presentations (LeDoux, 2000). After fear memory consolidation, which requires > 4~6 hours (McGaugh, 2000; Schafe et al., 2000), fear memory becomes remarkably resistant to perturbation, giving way only to numerous unreinforced CS presentations which leads to the extinction of conditioned fear responses. However, substantial remnants of the originally learned fear survive even after extensive extinction and cause the re-appearance of behavioral fear responses in a variety of circumstances, such as fear renewal and facilitated re-acquisition (Bouton, 2002). These observations suggest that extinction does not lead to complete reversal of fear learning, but rather a unique state in which the original fear memory traces are inhibited temporarily. The mechanisms of subsequent relearning are largely unknown, although it is well known that relearning occurs both more rapidly and with a lower threshold (i.e.'savings'; Napier et al., 1992).

The LA is essential in the acquisition and consolidation of auditory cued fear conditioning (Davis, 1992; Blair et al., 2001). Fear conditioning potentiates thalamic and cortical auditory inputs to the LA (McKernan and

Shinnick-Gallagher, 1997; Quirk et al., 1997; Tsvetkov et al., 2002), which are relayed to the basal and central amygdala to evoke aversive behavior (LeDoux, 2000; Davis and Whalen, 2001). Fear extinction recruits the infralimbic (IL) cortex to exert inhibitory influence on the medial subnuclei of the central amygdala (CeM) via the basal amygdala (BA) and amygdala-intercalated neurons (Maren and Quirk, 2004; Haubensak et al., 2010; Pape and Pare, 2010; Amir et al., 2011). Extinction also induces depotentiation at LA input synapses (Kim et al., 2007; Dalton et al., 2008; Hong et al., 2009), and enhances local inhibition (Chhatwal et al., 2005; Lin et al., 2009), all leading to decreased fear-related responses. Interestingly, NMDA receptor blockers infused into the LA impair both fear conditioning and extinction, suggesting that neuronal plasticity in the LA is critical for both events (Miserendino et al., 1990; Falls et al., 1992; Sotres-Bayon et al., 2007). Reconditioning has been less well explored, and although savings has been regarded as proof of the persistence of fear memory after extinction, the neural substrates which support the rapid relearning are largely unknown.

Previous LA unit recording studies have demonstrated that LA neurons increase their response to fear-conditioned stimuli and decrease when the stimuli become less fearful (Quirk et al., 1995; Collins and Pare, 2000; Repa et al., 2001; Goosens et al., 2003). Most of these reports

employed behavioral paradigms in which memory retrieval was tested only in the short-term, thus falling short of demonstrating the long-term modulation of fear memory involving extensive extinction and subsequent relearning. I thereby used high signal-to-noise ratio single unit recordings to track longitudinal changes in neuronal firing during fear conditioning, extinction and reconditioning. My results reveal distinct sub-populations in the LA which persistently represent the original CS-US association or dynamically encode updated CS-US association throughout the course of reversible fear learning.

Materials and Methods

Animals. Male Sprague-Dawley rats (n=45, 8 weeks old) were individually housed for 4~5 days before all experiments under an inverted 12 hours light/dark cycle (lights off at 09:00) and provided with food and water ad libitum. Behavioral training was done in the dark portion of the cycle. All procedures were approved by the Institute of Laboratory Animal Resources of Seoul National University.

Surgery and recording. Rats were anesthetized with sodium pentobarbital (50 mg/kg, i.p.) and maintained with isoflurane (1~1.5%) in O2. Rats were secured in a stereotaxic frame and bilaterally implanted with fixed-wire electrodes targeted to the LA: 2.8 mm posterior to bregma; 5.2 mm lateral to midline; and 6.3 mm to 6.9 mm deep from the cortical surface. The electrodes consisted of 8 individually insulated nichrome microwires (50 μm outer diameter, impedance 0.5~1 MΩ at 1 kHz; California Fine Wire) contained in a 21 gauge stainless steel guide cannula. The implant was secured using dental cement (Vertex). After surgery, analgesia (Metacam, Boehringer) and antibiotics were applied and rats were allowed to recover for 6~7 days. Neural activity was acquired and analyzed using a Plexon

MAP system, as previously described (Herry et al., 2008).

Behavioral procedures. Fear conditioning and extinction took place in two different contexts (context A and B) to minimize the influence of contextual associations. Reconditioning was conducted in the same context as extinction to avoid renewal and to observe savings. Context A was a rectangular Plexiglas box with a metal grid floor connected to an electrical current source (Coulbourn Instruments) which was set in a sound attenuating chamber. The chamber was illuminated with white light and was cleaned with a 70% ethanol solution. Context B was a cylindrical Plexiglas chamber with a metal grid floor which was illuminated with a red light and was cleaned with 1% acetic acid. In the retention test for the second unpairing (Post-UP2), a different context (context C) was used to avoid contextual fear. Context C was a trapezoid black opaque box with a flat black Formica floor illuminated with a red light that was cleaned with scented soap. All of the training sessions were videotaped and conditioned freezing was quantified by trained observers. The animals were considered to be freezing when there was no movement except for respiratory activity for 2 s during the 30 s CS presentation. The total freezing time was normalized to the duration of the CS presentation (Kim et al., 2010). On day

1, rats were habituated to the context and the CS in context A, in which they were placed in the recording chamber twice for 10 min, first without any cue and later with 4 presentations of the CS. The CS was a 29.089 s series of twenty-seven 2.8 kHz pure tone pips (200 ms duration repeated at 0.9 Hz, 85 dB sound pressure level) which has been used previously to enhance the signal-to-noise ratio for neural recordings (Rogan et al., 1997; Repa et al., 2001; Herry et al., 2008). On day 2, rats were given 4 presentations of the CS to determine basal LA neural responses to the CS (Hab). An hour later, fear conditioning was conducted by pairing the CS with a mild electric foot shock (0.5 mA, 1 s, 7 CS/US pairings; inter-trial interval: 80~120 s) co-initiating with the onset of the last tone pip. Extinction training took place 8 hours after fear conditioning in context B, in which rats were presented with 20 non-reinforced CS presentations (Post-FC). Two additional extinction sessions were conducted on the next day. On day 4, the behavioral and neuronal outcome of three extinction sessions was observed in a short 4 CS test session (Post-EX), followed 1 hour later by the reconditioning session in a manner similar to the initial fear learning. Eight hours after reconditioning, a retention test session was conducted (Post-REFC). To control for non-associative effects of conditioning, a separate group of rats (unpaired group, n=13) was exposed to explicitly unpaired CS and US

presentations during the conditioning and reconditioning sessions, with all the other procedures applied identically.

Single-unit spike sorting and analysis. Unit discrimination was performed using Offline Sorter (OFS, Plexon). All waveforms were plotted in a principal component space and clusters consisting of similar waveforms were first defined automatically and then verified manually. A cluster of waveforms distinct from other clusters in principal component space and showing a clear refractory period (>1 ms) was considered to be generated from a single neuron. At most, two distinct units were identified per channel, and single channel recordings proved sufficient to discern single unit responses, due to the low neuronal density of the LA (Quirk et al., 1997; Pare et al., 2004). Single unit isolation was graded using two statistic parameters, J3 and the Davies-Bouldin validity metric (DB), and neurons with a low grade were discarded. J3 reflects the ratio of between-cluster separation to within-cluster density calculated in a principal component space, and the DB is the ratio between the sum of within-cluster density to the degree of separation between clusters, and thus a high J3 and low DB value indicates a compact, well-separated unit cluster (Nicolelis et al., 2003). The long-term stability of a single-unit isolation was first determined using

Wavetracker (Plexon), in which the principal component space-cylinders of a unit recorded from different sessions were plotted (Herry et al., 2008; Tseng et al., 2011). A straight cylinder suggests that the clusters of a unit have a similar principal component composition, and that the same set of single units was recorded during the entire training session. Next I calculated the linear correlation values (r) between the template waveforms obtained over the entire set of behavioral sessions (Jackson and Fetz, 2007) to evaluate the similarity of waveform shape. Only stable units ($r > 0.93$) were considered for further analysis.

To investigate the effects of training on the LA cells, CS-evoked neural activities were normalized using a standard z-score transformation (bin size, 10ms). I adopted a moving average baseline (Pare and Gaudreau, 1996; Oyama et al., 2010) to exclude possible errors arising from extremely low spontaneous firing rates of the LA, and to reflect the in-session changes of basal firing rate. Unit responses were normalized to the firing rates of 500 ms preceding tone pip-onset for three consecutive CS (81 individual tone pips in total), except for units that did not exhibit any firing within this interval, which were normalized to the basal firing rates calculated from all pre-pip intervals of the session. Z-score peri-event time histograms (PETHs) of averaged CS-responses were constructed for each neuron and each pip

and then averaged for every CS (27 tone pips). A unit was regarded as being CS-onset or -offset responsive if the firing in 2 consecutive bins within 100 ms following CS-onset or -offset was significantly different from the baseline (500 ms preceding the CS) in an averaged PETH of all training sessions ($p < 0.05$, one-tailed t test) (Quirk et al., 1995). The onset latency of the CS-evoked responses was defined as the first bin to become significantly different from the baseline, and the bin which displayed the greatest firing within the 100-ms interval provided the peak response latency. To investigate the effects of behavioral training on the entire LA neuronal population, the population z-score PETH of all recorded neurons was calculated for each CS consisting of 27 tone pips and the mean z-values of 0~100 ms following CS-onset and -offset from the first 4 CSs of each session were compared throughout the course of behavioral training. The mean z-values in the two conditioning sessions were calculated using the first 25 tone pips of the CS to avoid foot shock-induced artifacts in the last pips.

Cell-by-cell analysis was further conducted to explore the effects of reversible fear learning on individual LA neurons. Analysis was restricted to neurons that were responsive to CS-onset. To determine responsiveness in each session, the CS-responses PETHs of 4 CSs (108 individual tone pips in

total) were averaged and the maximum z-score of the 0~100 ms interval after CS-onset was calculated for each neuron and compared to the significant z-score, 1.65 ($p < 0.05$, one-tailed t test) (Herry et al., 2008). A neuron was determined to be a 'fear neuron' if it exhibited significant CS-evoked responses in fear memory recall sessions (Post-FC or Post-REFC) and increased responses relative to the preceding sessions (Hab or Post-EX). I also sought 'extinction neurons', defined as neurons displaying strong CS-responses only after the extinction session (Post-EX), but found only one, and thus the characteristics of the fear neurons were compared to all of the other CS-responsive neurons.

Histology. At the end of each experiment, rats were anesthetized with urethane (1 g/kg, i.p.) and electrolytic lesions were made by passing a current (10 μA, 5~20 s) through recording microwires from which discrete units were identified. The duration of current injection was varied among recording microwires to identify the exact region where each unit was located. Longer current injections produced larger and more visible lesions. Animals were then transcardially perfused with 0.9% saline solution and 10% buffered formalin. Brains were removed and post-fixated overnight. Coronal sections (90 μm thick) were obtained using a vibroslicer (NVSL; World

Precision Instruments) and stained with cresyl violet. The placement of the recording microwires was examined under a light microscope.

Statistical analysis. To compare the behavioral results among behavioral sessions, averaged data points were analyzed using repeated-measures ANOVA with subsequent Newman-Keuls post hoc comparison. The CS-responsiveness of LA units was determined using unpaired t tests. For the analysis of CS-responses of LA sub-populations, the Friedman test (non-parametric one-way ANOVA for repeated measurements) and subsequent Dunn's post-hoc tests were used (Duclos et al., 2008). To detect changes in the CS-responses of the entire LA ensemble average activity (including both CS-responsive and non-responsive units), the linear sum of all CS-evoked activity was computed and the tone-to-tone variation was used for statistical deduction with parametric one-way ANOVA and Newman-Keuls post-hoc tests. Correlation between neuronal firings and behavioral responses were calculated using Pearson's correlation test. A probability value of $p < 0.05$ was considered indicative of statistical significance.

Results

Reversible fear learning dynamically regulates defensive behavior

A total of 32 rats underwent a reconditioning paradigm as described (see Methods) (Fig. 6A) and their fear-related behavior to the CS were examined. The CS was a series of twenty-seven 2.8 kHz pure tone pips (200 ms duration repeated at 0.9 Hz). Eight hours after the initial fear learning, rats displayed robust freezing when they were exposed to the CS in a different context ($F(3,93) = 781.70$, $p < 0.0001$, repeated-measures ANOVA; Hab vs. Post-FC, $p < 0.05$, Newman-Keuls posttest) (Fig. 6B) and the conditioned fear behavior diminished progressively over three extinction sessions (Fig. 6C). Reconditioning was conducted after CS-evoked fear returned to pre-conditioning levels with extinction training (Hab vs. Post-EX, $p > 0.05$) and resulted in stronger fear responses compared to the initial fear learning (Post-FC vs. Post-REFC, $p < 0.05$). In contrast, the 13 rats that received unpaired CS-US presentations showed no evidence of CS-induced fear, except immediately after shock delivery ($F(3,36) = 0.83$, $p > 0.5$, repeated-measures ANOVA; $p > 0.05$ for all pairs, Newman-Keuls posttest).

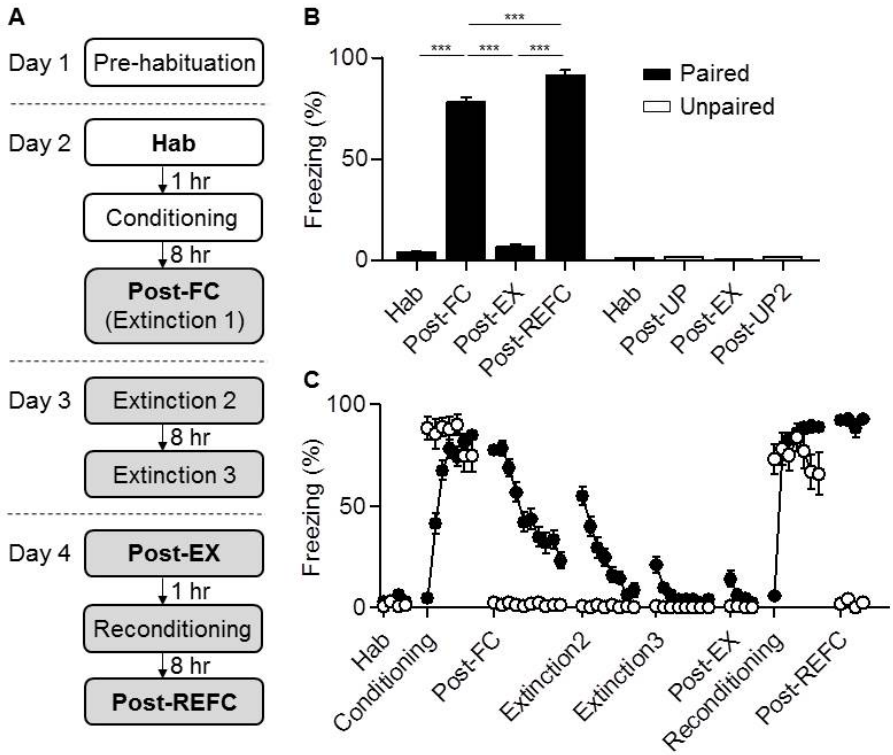

Figure 6. Behavioral procedures and results. A, The behavioral procedure used in the experiment. The white and gray shades represent different contexts. **B,** Averaged freezing responses during the first 4 CS presentations of the retention test sessions (bold characters in A) in each group (paired group, n=32 rats; unpaired controls, n=13 rats). **C,** The learning curves of the entire behavioral session (paired group, filled circle; unpaired controls, open circle). Error bars indicate SEM. Abbreviations: Hab, habituation; Post-FC, post-conditioning; Post-EX, post-extinction; Post-REFC, post-reconditioning.

Electrophysiological characteristics of the LA neurons

Only stable, high signal-to-noise ratio LA neurons verified by principal component comparisons and correlation analysis were included in the data analysis (Fig. 7). In total 188 LA neurons were analyzed, 114 from the fear-conditioned group and 74 from the unpaired controls. Histological analysis revealed that recorded cells were located within the dorsal and ventral LA (Fig. 8). Consistent with previous reports, the LA neurons displayed low spontaneous firing rates (Quirk et al., 1995; Pare and Collins, 2000; Repa et al., 2001). The average firing rate was 0.68 Hz, ranging from 0.01 to 13 Hz, and the averaged spike width (the time between the maximum and minimum peak) was 0.43 ms, ranging from 0.12 to 0.75 ms. In accordance with previous results (Quirk et al., 1995), most of the recorded LA cells showed wide spike widths and low firing rates and the waveform width and firing rate were inversely correlated ($r = -0.48$, $p < 0.0001$, Pearson's correlation test), consistent with the pyramidal-type projection neurons which are prevalent in the LA (McDonald, 1982; Davis et al., 1994; Medina et al., 2002). The average basal firing rates were not different among the behavioral sessions ($F(5,565) = 1.64$, $p > 0.1$, repeated-measures ANOVA).

Forty five of 114 (39%) neurons in the fear-conditioned group and

22 of 74 (30%) neurons in the unpaired controls were determined as CS-responsive based on the averaged CS-evoked neural activities in all of the training sessions. These neurons displayed phasic responses to tone within 40 ms following pip-onset (Fig. 9A), with an average onset response latency of 26.3 ± 1.9 ms (paired group, 25.3 ± 2.5 ms; unpaired group, 29.1 ± 2.7 ms; $p > 0.1$, unpaired t test). The pip-evoked excitation appeared reliably throughout the individual CS presentations of 27 individual pips, thus the pip-evoked responses were averaged to enhance signal-to-noise ratio of CS-responses as shown in previous studies (Rogan et al., 1997; Repa et al., 2001; Herry et al., 2008). The number of CS-responsive neurons in each separate session was not largely changed throughout the course of reversible fear learning, while repeated unpairing resulted in fewer neurons being responsive (Table 1). Histological analysis revealed that LAd neurons responded to the CS with shorter response latencies than LAv neurons (LAd, 24.3 ± 2.1 ms; LAv, 31.6 ± 3.8 ms; $p < 0.05$, unpaired t test) (Bordi et al., 1993). Interestingly, 43% of the CS-onset responsive neurons (n=29) also displayed CS-offset responses (Fig. 9B), while 20 neurons were only responsive to CS-offset with an average latency of 26.5 ± 3.2 ms.

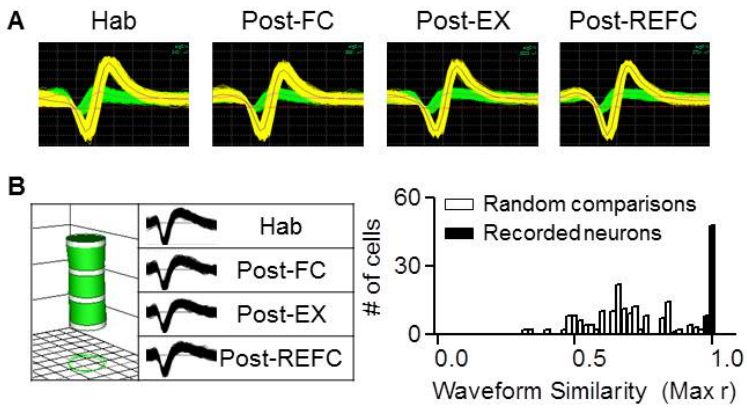

Figure 7. Long-term single unit recordings in the LA. A, Representative waveforms of two neurons recorded from a single electrode and were stably observed throughout the behavioral training period. Grid: 55μV, 100μs. **B,** Verification of long-term stable single unit recordings using principal component space cylinders (Left). A straight cylinder suggests that the same set of single units was recorded in different behavioral sessions. Quantitative evaluation of waveform similarity from units recorded on different days (Right). Randomly selected waveforms were used as a control.

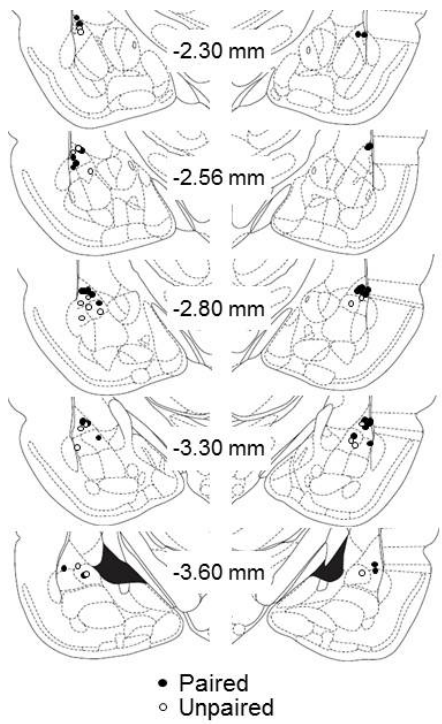

Figure 8. Histological verification of the electrode placements. The electrode placements were found within the LA, varied in dorsal-ventral and anterior-posterior axes. The paired group is indicated with filled circle and the unpaired controls with open circle.

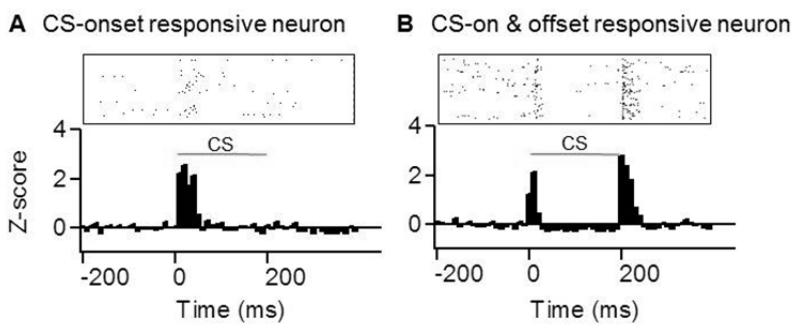

Figure 9. LA neuronal responses to the CS. A, A representative unit showing phasic responses to CS-onset. **B,** Both CS-onset and CS-offset induced strong excitation in a representative LA neuron.

	All sessions averaged			HAB			Post-FC			Post-EX			Post-REFC		
	Firing rates (Hz)	% of responsive cells		Firing rates (Hz)	% of responsive cells		Firing rates (Hz)	% of responsive cells (Newly emerged)		Firing rates (Hz)	% of responsive cells (Newly emerged)		Firing rates (Hz)	% of responsive cells (Newly emerged)	
		Onset	Offset		Onset	Offset		Onset	Offset		Onset	Offset		Onset	Offset
Paired	0.79 ± 0.19	39.47	27.19	0.89 ± 0.21	28.95	20.18	0.82 ± 0.24	34.21 (15.79)	28.95 (18.42)	0.85 ± 0.22	25.44 (10.53)	14.04 (10.53)	0.68 ± 0.18	29.83 (13.16)	21.05 (15.79)
Unpaired	0.52 ± 0.21	29.73	24.32	0.55 ± 0.22	25.68	12.16	0.55 ± 0.22	22.97 (6.76)	25.68 (20.27)	0.42 ± 0.16	18.92 (8.11)	6.76 (4.05)	0.50 ± 0.22	12.16 (6.76)	8.11 (8.11)

Table 1. Basal firing rates and CS-response properties of the recorded LA neurons in the paired group (n=114) and the unpaired controls (n=74) throughout the reversible fear learning.

LA ensemble activity represents updated CS-US association strength in reversible fear learning

It has been reported that the CS-evoked responses of LA neurons increase after fear conditioning, and that closely following extinction results in decreased tone responses of LA neurons in vivo (Quirk et al., 1995; Collins and Pare, 2000; Repa et al., 2001; Goosens et al., 2003). However, neural representations of fear memory involving extensive extinction and subsequent reconditioning have remained elusive because most previous studies have used behavioral paradigms in which memory retrieval was tested only in the short-term. Therefore, I investigated LA responses to the CS in reversible fear learning comprising extensive extinction and reconditioning. Fear conditioning-induced changes in tone-evoked firings were examined eight hours after the initial fear conditioning, a time at which fear memory is fully consolidated (Schafe et al., 2000; Schafe and LeDoux, 2000).

I constructed a population z-score PETH throughout the reversible fear learning and found that LA neurons showed potent excitation in response to CS-onset and their activity was dynamically modulated in the reversible fear learning, corresponding to the CS-US association strength. Fear conditioning resulted in a strong CS-evoked excitation of LA neurons,

while this excitation was weakened during extensive extinction, and reconditioning reinstated a strong CS-response (Fig. 10). In the unpaired controls, however, CS-evoked responses were largely unchanged by the initial unpairing, and were weakened by the second unpairing.

The average CS-evoked responses of LA neuronal population were quantified as a mean z-value of 0~100 ms following CS-onset and compared across retention test sessions of reversible fear learning. Fear conditioning significantly increased the averaged CS-response compared to habituation ($F(3,12) = 14.03$, $p < 0.001$, one-way ANOVA; Hab vs. Post-FC, $p < 0.05$, Newman-Keuls posttest), whereas unpairing did not alter LA neuronal responses ($F(3,12) = 3.52$, $p < 0.05$, one-way ANOVA; Hab vs. Post-UP, $p > 0.05$, Newman-Keuls posttest) (Fig. 11A). Three CS-alone extinction sessions resulted in decreased LA responses indiscernible with habituation (Hab vs. Post-EX, $p > 0.05$). These results are consistent with previous reports, which demonstrated the short-term effects of fear conditioning and extinction on LA neurons (Quirk et al., 1995; Repa et al., 2001) and further suggest that the updating of CS-US association strength that takes place during the reversible fear learning is dynamically represented in the LA even after memory consolidation. Consistently, reconditioning again increased CS-evoked responses of the LA compared to both the preceding

extinction retrieval session and the habituation session (Post-EX vs. Post-REFC, $p < 0.05$; Hab vs. Post-REFC, $p < 0.05$). In the unpaired controls, LA neuronal responses to CS-onset slightly decreased after the second unpairing, possibly due to safety learning (Lolordo, 1969; Rogan et al., 2005), but not to statistically significant levels (Post-EX vs. Post-UP2, $p > 0.05$) (Fig. 11A). The averaged LA population activity was positively correlated with the freezing behavior in the paired group ($r = 0.55$, $p < 0.001$, Pearson's correlation test), but not in the unpaired control ($r = 0.08$, $p > 0.1$, Pearson's correlation test) (Fig. 11B).

Importantly, CS-evoked response latencies were also reversibly altered; the CS-evoked response arose and peaked more rapidly following the initial fear conditioning and reconditioning compared to the preceding sessions (onset response latencies, Hab vs. Post-FC, Post-FC vs. Post-EX, Post-EX vs. Post-REFC, $p < 0.05$, paired t test; peak response latencies, $p < 0.05$ for the same pairs, paired t test) (Fig. 11C). Again, unpaired controls did not show significant changes ($p > 0.1$ for the same pairs, paired t test) (data not shown). Faster response latencies are consistent with strengthened influences from the short-latency thalamic pathway (McKernan and Shinnick-Gallagher, 1997; Quirk et al., 1997). These intricate, dynamic changes in the CS-response profile further support the involvement of

specific plastic mechanisms reversibly recruited in my learning paradigm.

Additionally, I checked whether CS-offset responses were altered following reversible fear learning, because a considerable number of LA neurons were responsive to CS-offset. Fear conditioning, however, did not significantly alter the CS-offset responses of the LA neurons and the responses disappeared following extensive extinction (Fig. 12). Collectively, these results suggest that the average LA ensemble activity represents updated CS-US association strength in the reversible fear learning and maintains this representation beyond memory consolidation, consistent with previous reports (Maren, 2000; Goosens et al., 2003; Hong et al., 2011).

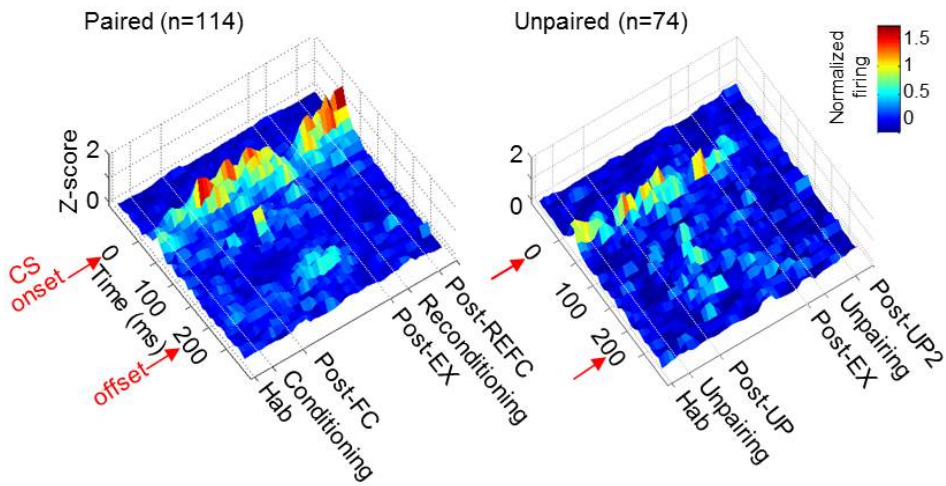

Figure 10. LA ensemble activity during reversible fear learning. Population z-score PETH throughout the behavioral training in the paired group (n=114, left) and the unpaired controls (n=74, right). The surface plot of the normalized firing rate was calculated and was smoothed for ±1 trials.

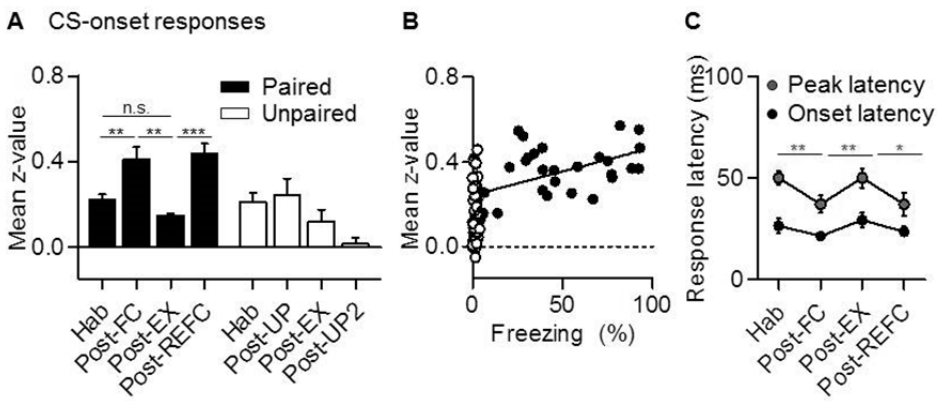

Figure 11. Quantification of LA ensemble activity to CS-onset. A, Comparisons of mean z-values calculated in a period of 0~100 ms following CS-onset. The paired group displayed reversible CS-evoked responses in contrast to the unpaired controls. **B,** Correlation analysis between neural responses and freezing behavior. A significant correlation was observed only in the conditioned group (r = 0.55; filled circle), not in the unpaired controls (r = 0.08; empty circle). **C,** Comparison of the onset onset and peak response latency across the retention test sessions. Conditioning resulted in a more rapid onset and peak response latency compared to the preceding sessions. Error bars indicate SEM.

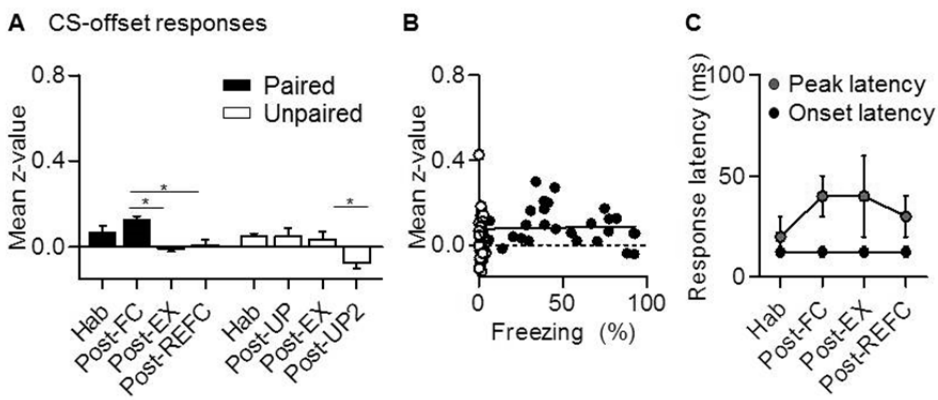

Figure 12. Quantification of LA ensemble activity to CS-offset. A–C, The same quantification as CS-onset responses was performed for the CS-offset responses. Fear conditioning did not significantly alter the CS-offset responses of LA neurons. Error bars indicate SEM.

Distinct sub-populations of LA fear neurons represent the updated and original CS-US association strength in reversible fear learning

It has been demonstrated that fear conditioning results in a strong potentiation of CS-evoked LA field potentials (Rogan et al., 1997), while only 10~30% of LA neurons display increased CS-evoked responses after fear conditioning and this subset of neurons exhibits various types of learning-induced plasticity, such as transient or persistent potentiation by fear conditioning (Quirk et al., 1995; Repa et al., 2001). I thus further analyzed the data on a cell-by-cell basis to identify distinct LA neuronal sub-populations that encode the various facets of reversible fear learning. I focused on CS-onset responsive neurons, since the LA population displayed stronger excitation in response to CS-onset and this response was dynamically modulated during reversible fear learning.

I first identified neurons which displayed significant and increased responses to CS-onset after either of the two fear conditioning sessions (Post-FC or Post-REFC) compared to the preceding sessions (Hab or Post-EX), and these neurons were defined as ‘fear neurons’ (n=25, 56% of CS-onset responsive units) (Fig. 13). I also sought for 'extinction neurons' displaying increased CS-responses only after extinction and found only one,

consistent with previous results showing that they reside mostly in the BA (Herry et al., 2008). 68% of the fear neurons increased their responses to CS after the initial fear conditioning ('conditioning-potentiated fear neurons', n=17) (Fig. 14A) and a larger number of neurons exhibited potentiated responses following reconditioning ('reconditioning-potentiated fear neurons', n=21, 84% of fear neurons) (Fig. 14B). Both conditioning- and reconditioning-potentiated fear neurons displayed reversible changes of CS-evoked firing patterns throughout the course of reversible fear learning, while small and relatively constant responses were observed in the other CS-responsive neurons that were categorized as non-fear-encoding neurons ('other neurons', n=20, 44% of CS-onset responsive units) (Fig. 14C). The basal firing rates and spike duration of fear neurons were not different from the other CS-responsive neurons ($p > 0.1$, unpaired t test) (Fig. 15A). However, fear neurons responded to the CS with a shorter response latency compared to the other neurons (fear neurons, 24.0 ± 1.6 ms; other neurons, 32.5 ± 5.2 ms; $p < 0.05$, unpaired t test) (Fig. 15B) and were frequently found in the dorsal part of the LA, with a few in the ventral LA (Fig. 15C), suggesting potent innervation by short-latency thalamic inputs. Interestingly, I found that there was a large overlap between neurons that were potentiated after the original fear conditioning and reconditioning; 76% of the

conditioning-potentiated fear neurons was re-potentiated by reconditioning (n=13) (Fig. 13), suggesting that traces of the initial fear learning remained even after extensive extinction, which allowed neurons to be readily recruited by the subsequent relearning.

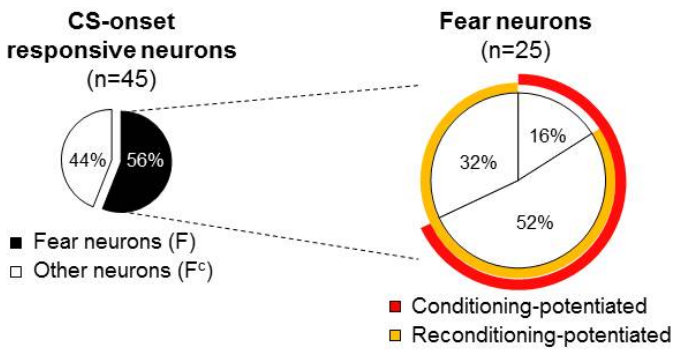

Figure 13. Fear-encoding neurons in the LA. Pie chart shows the percentage of fear neurons among the CS-onset responsive neurons (left, n=45 cells) and the subcategories of fear neurons (right, n=25 cells). A large overlap between the conditioning-potentiated fear neurons (n=17 cells) and reconditioning-potentiated fear neurons (n=21 cells) was observed.

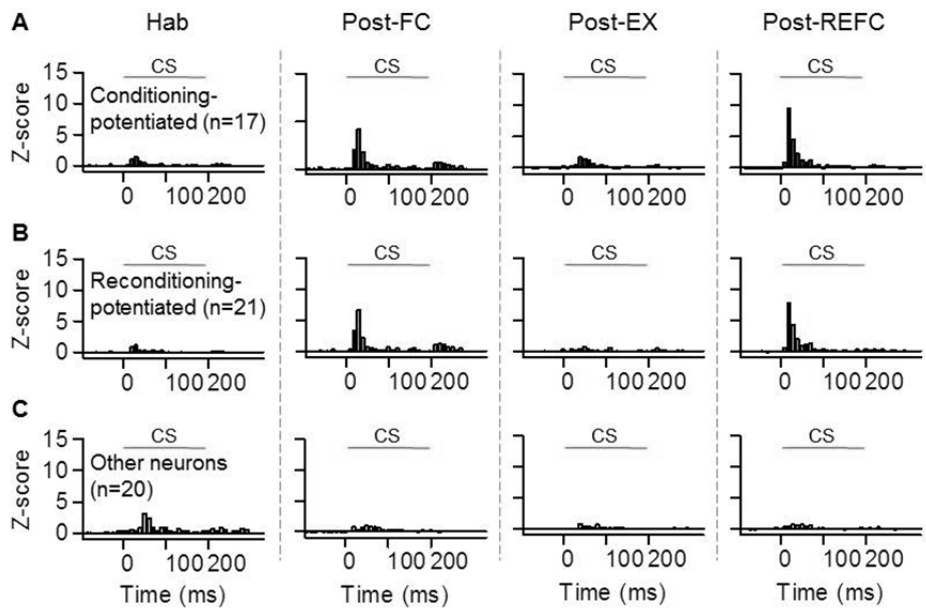

Figure 14. CS-responses of fear-encoding LA neurons. A, Z-score PETH of conditioning-potentiated fear neurons (n=17, 68% of fear neurons). **B,** Z-score PETH of reconditioning-potentiated fear neurons (n=21, 84% of fear neurons). **C,** Z-score PETH of CS-onset responsive, but not fear-encoding neurons (other neurons, n=20, 44% of CS-onset responsive units).

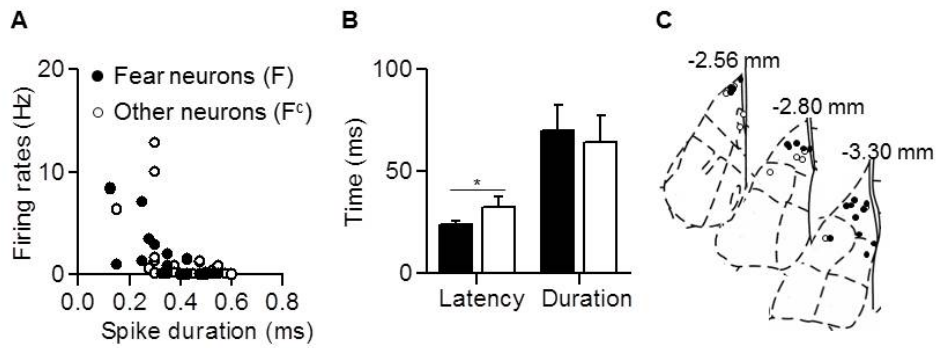

Figure 15. Characteristics of fear-encoding LA neurons. A, The basal firing rates and spike duration of fear neurons were not different from the other neurons. **B,** A comparison of onset latency and response duration. Fear neurons responded with a shorter latency to CS-onset compared to the other neurons. Error bars indicate SEM. **C,** Histological analysis revealed that fear neurons were preferentially found in the LAd.

To identify distinct LA neuronal sub-populations that represent various facets of reversible fear learning, I tracked the changes in CS-evoked responses of neurons that were potentiated following the initial fear conditioning ('conditioning-potentiated fear neurons') in subsequent extinction and reconditioning. Although the averaged CS-evoked responses of the conditioning-potentiated fear neurons appeared to be reversibly modulated (Fig. 14A), a cell-by-cell analysis revealed that this population was not homogeneous; two distinct classes of neurons were identified based on their responses to extinction (Fig. 16). Half of the conditioning-potentiated neurons exhibited significantly decreased CS-evoked responses after extinction ('extinction-sensitive fear neurons', n=9, 53% of conditioning-potentiated fear neurons) (Fig. 17A), while the other half retained increased CS-responses even after extensive extinction ('extinction-resistant fear neurons', n=8, 47% of conditioning-potentiated fear neurons) (Fig. 17B). These results are consistent with a previous study which reported similar neuronal populations within a single extinction session conducted 1 hour after fear conditioning (Repa et al., 2001). Interestingly, the extinction-sensitive fear neurons exhibited typical phasic and strong responses to CS-onset corresponding to short-latency sensory inputs, whereas extinction-resistant fear neurons exhibited smaller but more

sustained responses to the tone (over 100 ms). The onset latencies were not different between these two populations (extinction-sensitive fear neurons, 20.0 ± 2.9 ms; extinction-resistant fear neurons, 22.5 ± 3.1 ms; $p > 0.1$, unpaired t test) (Fig. 18A) and histological analysis confirmed that both neuronal populations were located in the dorsal part of the LA (Fig. 18C). However, the CS-evoked responses of extinction-resistant fear neurons lasted much longer (extinction-sensitive fear neurons, 45.6 ± 16.1 ms; extinction-resistant fear neurons, 111.3 ± 21.9 ms; $p < 0.05$, unpaired t test) (Fig. 18A), and were weaker (mean z-value, extinction-sensitive fear neurons, 9.9 ± 2.1; extinction-resistant fear neurons, 3.5 ± 0.4; $p < 0.005$, unpaired t test) (data not shown), suggesting distinct connectivity. The longer, persistent responses in the extinction-resistant fear neurons may involve multi-synaptic local sensory inputs and/or innervations from cortical regions (Repa et al., 2001), and may represent the persistence of the original fear memory after extinction.

Importantly, extinction-sensitive and -resistant neurons were also distinguished by their CS-evoked activities after reconditioning. The average CS-evoked responses of extinction-sensitive fear neurons were strongly potentiated after reconditioning, resembling LA ensemble activity (Fig. 17A), whereas extinction-resistant fear neurons did not show further

increases after reconditioning (Fig. 17B). Intriguingly, a cell-by-cell analysis revealed that all of the extinction-sensitive fear neurons but for a single exception showed increased and significant responses after reconditioning, and thus comprise a sub-population encoding dynamic changes in CS-US association strength during reversible fear learning ('reversible fear neurons', n=8, 89% of extinction-sensitive fear neurons, and 47% of conditioning-potentiated fear neurons) (Fig. 16). In contrast, all of the other CS-responsive neurons ('other CS-responsive neurons', n=37) (Fig. 17C) displayed weak, constant CS-evoked responses. I compared the mean z-values of the reversible fear neurons across sessions and found that their responses were reversibly altered in a manner similar to LA population ensemble activity, but to a greater extent ($p < 0.001$, Friedman test; Hab vs. Post-FC, Post-FC vs. Post-EX, Post-EX vs. Post-REFC, $p < 0.05$, Dunn's posttest). In contrast, the mean z-values of the other CS-responsive neurons remained relatively constant ($p > 0.05$, Friedman test; $p > 0.05$ for the same pairs, Dunn's posttest) (Fig. 18B), suggesting the reversible fear neurons lead the LA neuronal ensemble activity in reversible fear learning. Reversible fear neurons displayed a shorter responses latency compared to the other CS-responsive neurons (reversible fear neurons, 18.8 ± 3.0 ms; other CS-responsive neurons, 30.3 ± 3.1 ms; $p < 0.05$, unpaired t test), but

with a similar response duration (reversible fear neurons, 47.5 ± 18.1 ms; other CS-responsive neurons, 71.9 ± 10.0 ms; $p > 0.1$, unpaired t test) (Fig. 18A). Consistent with these electrophysiological characteristics, histological analysis revealed that reversible fear neurons were preferentially located in the dorsal part of the LAd (Fig. 18C), which is known to receive dense thalamic short-latency innervations (LeDoux et al., 1990; Quirk et al., 1997). Together, these results suggest there are two distinct sub-populations of fear-encoding neurons in the LA; one is dynamically regulated by fear conditioning and extinction while the other represents persistence of the original fear memory.

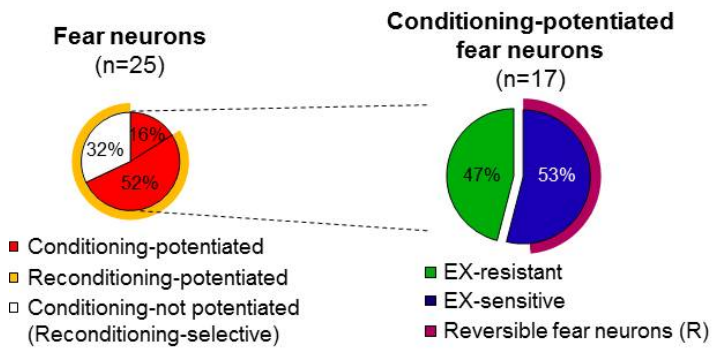

Figure 16. Fear-encoding sub-populations in the LA. Pie chart summarizes how the subcategories of conditioning-potentiated fear neurons responded to subsequent extinction and reconditioning. Conditioning-potentiated fear neurons were categorized into extinction-resistant fear neurons (n=8 cells) and extinction-sensitive fear neurons (n=9 cells). The left pie chart represents identical fear neurons as those in Figure 13.

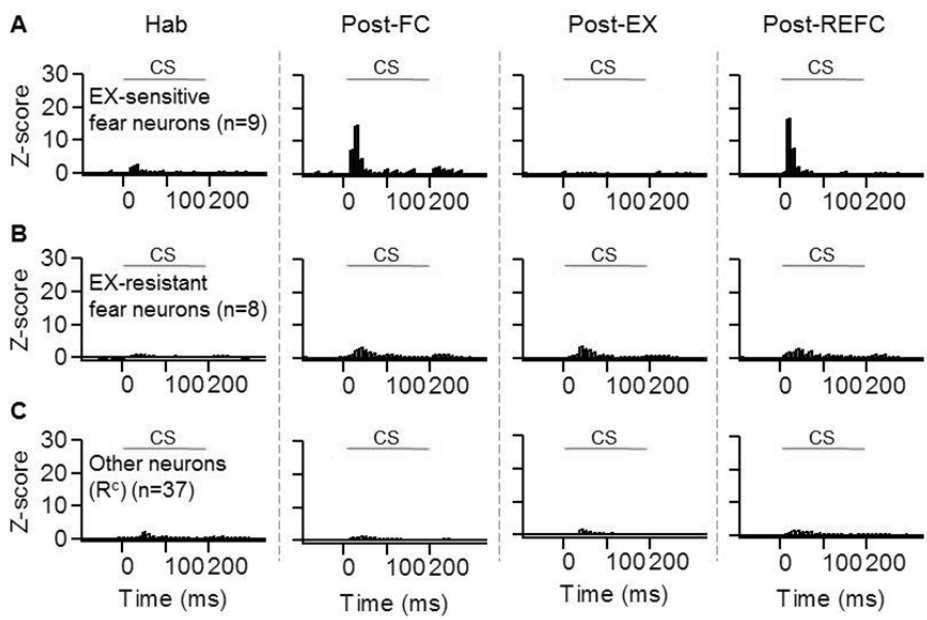

Figure 17. CS-responses of fear-encoding sub-populations. A, Z-score PETH of extinction-sensitive fear neurons (n=9, 53% of conditioning-potentiated fear neurons). **B,** Z-score PETH of extinction-resistant fear neurons (n=8, 47% of conditioning-potentiated fear neurons), which retained increased CS responses after extensive extinction**. C,** Z-score PETH of other CS-responsive neurons (n=37) that were not categorized as reversible fear neurons.

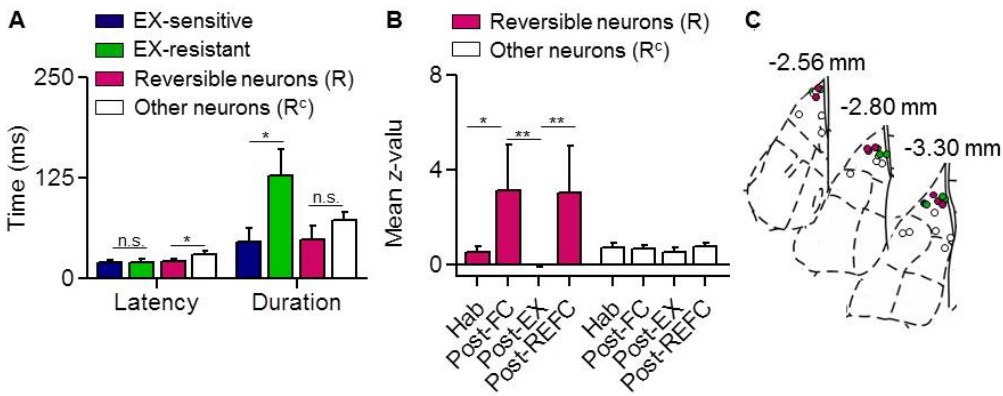

Figure 18. Characteristics of fear-encoding sub-populations. A, Comparison of onset response latency and response duration. Extinction-resistant fear neurons displayed sustained responses compared with extinction-sensitive fear neurons. The response latency of the reversible fear neurons was shorter than the other CS-responsive neurons. **B,** The mean z-value comparisons of reversible fear neurons and the other CS-responsive neurons. Error bars indicate SEM. **C,** Histological analysis confirmed that conditioning-potentiated fear neurons, including reversible fear neurons, were preferentially located in the dorsal part of the LA.

Reversible fear neurons represent savings effect after extinction

The relearning of fear occurs much faster than original fear learning even after extensive extinction, and this phenomenon is known as the 'savings' (Kehoe, 1988; Rescorla, 2001). Although savings has been widely suggested as empirical evidence of memory persistence after extinction (Bouton, 2002), the neural correlates of savings have not been identified.

In accordance with previous reports (Rescorla, 2001), I found that the freezing responses progressively increased during the initial fear conditioning, but increased more rapidly during reconditioning. CS-evoked freezing was indistinguishable between pre-conditioning sessions, Hab and Post-EX ($p > 0.05$, paired t test), and at the first pairing of the two conditioning sessions ($p > 0.1$, paired t test). However, the discrepancy between the learning curves of fear conditioning and reconditioning was significant at the second CS-US pairing ($p < 0.0001$, paired t test), the third pairing ($p < 0.005$, paired t test) and the fifth pairing ($p < 0.005$, paired t test) (Fig. 19A). Although the difference in conditioned freezing disappeared by the end of the conditioning sessions ($p > 0.1$, paired t test), stronger freezing was also observed in the retention test of reconditioning ($p < 0.0001$, paired t test) compared to the initial fear conditioning.

Interestingly, the CS-evoked responses of the reversible fear neurons increased more rapidly during reconditioning, in tight correlation with the behavioral results. The mean z-values in the two conditioning sessions diverged at the second CS-US pairing ($p < 0.05$, paired t test) (Fig. 19B), while the CS-responses in the pre-conditioning sessions and at the first pairing were not significantly different. The statistical difference disappeared at the third pairing ($p > 0.1$, paired t test), suggesting that the potentiation of the neural responses reached a ceiling faster than the behavioral responses. The rapid increases of LA neuronal responses during the reconditioning session were further confirmed by comparison of the slope of CS-response increase between the first and second CS-US pairings ($p < 0.05$, paired t test) (Fig. 19C). These results suggest that 'reversible fear neurons' not only integrate the reversible changes in CS-US association strength, but also are primed by prior learning-induced changes so as to detect a given CS-US association more rapidly during subsequent relearning. The persistently potentiated CS-responses of extinction-resistant fear neurons may also trigger/support this rapid re-potentiation of the CS-responses observed in reversible fear neurons. In addition to the more rapid in-session learning upon reconditioning, stronger freezing was also observed in the retention test of reconditioning ($p < 0.0001$, paired t test) compared to

the initial conditioning, which is likely to be supported by the larger number of neurons recruited by reconditioning (Fig. 13) compared to the initial fear conditioning.

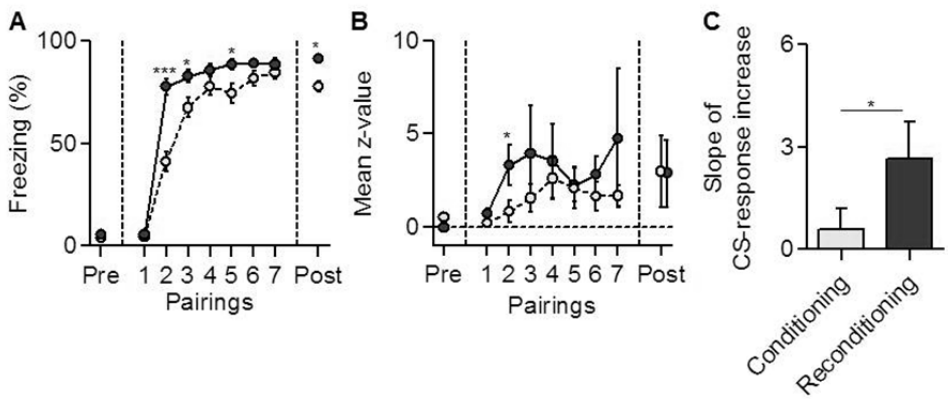

Figure 19. The neural correlate of savings after extinction. A, Behavioral evidence of savings. Reconditioning occurred faster than the initial fear learning. **B,** The mean z-values of reversible fear neurons. CS-evoked responses were larger and more rapidly increased during reconditioning (conditioning, dark gray circle; reconditioning, light gray circle). **C**, Comparison of the slope of CS-response increases between the first and second pairings. Error bars indicate SEM.

Discussion

LA neurons were found to reversibly encode updated CS-US association strength throughout the course of sequential fear learning. The LA neuronal population displayed increased average CS-evoked firing after conditioning, decreased responses after extinction and re-potentiated responses after reconditioning, in tight correlation with the changes in conditioned freezing responses. Cell-by-cell analysis revealed the two distinct sub-populations of fear-encoding neurons in the LA; one showed reversible encoding of fear learning that corresponded to the LA population activity (‘reversible fear neurons’), whereas the other was resistant to change during extinction and reconditioning (‘extinction-resistant fear neurons’), likely supporting the persistence of fear memory. Interestingly, reversible fear neurons exhibited both a stronger and more rapid acquisition of CS-US association during reconditioning relative to the initial fear conditioning, providing a neural correlate of the savings effect during reconditioning.

The ‘reversible fear neurons’ observed in the present study exhibit remarkably similar characteristics to distinct BA neurons that are responsive to fear conditioning, extinction and renewal in a reversible manner and also

a subset of LA neurons encoding the renewal of extinguished fear (Hobin et al., 2003; Herry et al., 2008). Since LA excitatory neurons are known to drive the activation of the central amygdala and fear expression via BA excitatory neurons (LeDoux, 2000; Pape and Pare, 2010; Amir et al., 2011), it is possible that the subset of LA neurons that responds to renewal (Hobin et al., 2003) largely overlaps with the 'reversible fear neurons' identified here and that both preferentially innervate 'fear neurons' in the BA (Herry et al., 2008), thus controlling central amygdala activity and contributing to reversible fear expression. Alternatively, reversible LA neuronal firing may alter activity of the amygdala-intercalated neurons and inhibitory central amygdala neurons (Pare et al., 2004; Amano et al., 2010; Haubensak et al., 2010). The extraordinary plasticity of these reversible fear neurons suggests that LA neural circuits can be dynamically modified even after memory consolidation.

The 'extinction-resistant fear neurons' found in my study provide a neural substrate for the persistent fear memory trace which had been predicted earlier (Pearce and Hall, 1980; Bouton and King, 1983). These neurons displayed CS-responses of longer duration (Fig. 18A), suggesting the influence of cortical regions where traces of persistent fear have also been identified (Corcoran and Quirk, 2007; Burgos-Robles et al., 2009;

Sacco and Sacchetti, 2010; Sotres-Bayon and Quirk, 2010). The persistent potentiated firing of the 'extinction-resistant fear neurons' may contribute to the renewal or spontaneous recovery of fear even after extensive extinction. In spite of the persistent fear-encoding in these neurons, after extinction, the expression of fearful responses is likely to be inhibited downstream of the LA (Ehrlich et al., 2009; Pape and Pare, 2010; Maren, 2011). Well-known inhibitory mechanisms involving the prefrontal cortex (Milad and Quirk, 2002; Rosenkranz et al., 2003; Likhtik et al., 2005; Sotres-Bayon et al., 2006; Quirk and Mueller, 2008) and amygdala ITC neurons (Chhatwal et al., 2005; Likhtik et al., 2008; Ehrlich et al., 2009) may provide inhibition at the BA or CeM leading to the suppression of fear responses. The context-dependent disinhibition of these subnuclei and the LA are believed to underlie the renewal of fear (Hobin et al., 2003; Likhtik et al., 2008; Ehrlich et al., 2009).

Extinction is thought to involve both inhibition and unlearning of original associations (Bouton, 2002). The relative contribution of new learning and unlearning in the behavioral extinction of many forms of associative memory has been a key issue in memory research (Medina et al., 2002; Barad, 2006). In previous studies involving different learning paradigms, the immediate reversal of CS-US contingencies resulted in the

reversal of neural responses in a subset of amygdala neurons (Schoenbaum et al., 1999; Paton et al., 2006). Consistent with these findings, my results in auditory cued-fear conditioning demonstrate that the CS-responses of some LA neurons are suppressed after extinction and exhibit savings during relearning, but there are other neurons which exhibit persistent potentiation after extinction, suggesting that unlearning and new learning are both integrated at the level of the LA neurons. Consistent with previous reports (Repa et al., 2001), 'extinction-resistant' fear neurons retained potentiated CS-responses even after extensive extinction, while 'extinction-sensitive' fear neurons showed a clear decrease in CS-responses (Fig. 17); Together, this resulted in a net reduction of the LA ensemble activity after extensive extinction. Although the net CS-response after extinction was indiscernible from pre-training levels, individual neurons displayed different responses, suggesting that network changes in LA connectivity upon fear conditioning persist after extinction. Because early- and late-extinction (within and beyond 6 hours post-conditioning, respectively) involves different mechanisms and leads to different neural changes (Myers et al., 2006; Chang et al., 2009), and most previous recordings were limited to early-extinction paradigms, my results constitute important evidence for the mechanisms underlying late-extinction.

Reconditioning after extinction has been less well explored, although the rapid re-acquisition of fear has been regarded as proof of the persistence of memory after extinction (Bouton, 2002). My findings show that whereas extinction does not return the network changes in LA connectivity to the pre-conditioning state, reconditioning appears to return the system to the pre-extinction state. Reconditioning resulted in an increase of the LA ensemble activity, which had decreased to baseline levels after extinction (Fig. 11), suggesting that LA neurons are able to adaptively represent updated CS-US association strength throughout the course of reversible fear learning. This re-potentiation was supported by a majority of the conditioning-potentiated fear neurons, demonstrating a significant overlap of fear-encoding neurons. This overlap is accounted for the extinction-induced inhibitory mechanisms that temporarily suppress fear conditioned responses. Interestingly, the CS-responses of reversible fear neurons appeared to be more readily potentiated upon reconditioning compared to the initial fear conditioning (Fig. 19), supporting the hypothesis that reconditioning reverses extinction-induced network changes. Together, these results suggest the conditioning-induced plasticity was temporarily inhibited by extinction and reconditioning eliminated this inhibition (Bouton and King, 1983; Quirk et al., 2006; Myers and Davis, 2007).

The strong reversible encoding of CS-US association strength in 'reversible fear neurons' (Fig. 16) dominates the LA population coding (shown in Fig. 10), suggesting that it is the plasticity of these neurons which is detected using field potential (Rogan et al., 1997) or immediate-early gene methods (Hall et al., 2001; Han et al., 2007; Reijmers et al., 2007). These fear neurons amount to only 10~30% of all the LA neurons, suggesting a rather sparse and restricted encoding of CS-US associations (Quirk et al., 1995; Repa et al., 2001; Han et al., 2007). In contrast, fear learning-induced synaptic potentiation has been observed in the general population of LA neurons (McKernan and Shinnick-Gallagher, 1997; Kim et al., 2007; Zhou et al., 2009), leading to the previously suggested possibility that a majority of LA neurons are strongly inhibited by GABAergic interneurons (Pare and Gaudreau, 1996) and are thus virtually undetectable by either in vivo recordings or immediate-early gene staining methods. Interestingly, a previous report demonstrated that targeted ablation of the roughly ~15% of LA neurons that preferentially participated in learning can significantly impair auditory fear memory, whereas ablating a similarly sized random population had no effect (Han et al., 2009). It is tempting to hypothesize the similarly sized 'reversible fear neuron' population in my recordings largely overlaps with the population targeted in the previous

studies.

Traces of persistent fear memory have been suggested to reside in cortical regions (Corcoran and Quirk, 2007; Burgos-Robles et al., 2009; Sacco and Sacchetti, 2010; Sotres-Bayon and Quirk, 2010), but how they may interact with the LA and support later savings or memory relapse has been largely unknown. My findings show a strong neural correlate of savings in fear-encoding LA neurons, which may be innervated and influenced by memory-preserving cortical regions to allow the more rapid detection of changes in CS-US association. Metaplastic mechanisms that enable more rapid synaptic plasticity at input synapses may also support the enhanced potentiation of CS-responses in these neurons (Abraham, 2008; Lee et al., 2013). Extinction-resistant fear neurons, which were potentiated after the initial fear learning and retained the potentiation even after extensive extinction, may also play an important role in the persistence of fear memory and relapse after extinction.

Fear conditioning and extinction have served as primary models for the treatment of PTSD and anxiety disorders. Although most PTSD research aimed at thwarting the renewal of fear memory has focused on the dysfunctions or manipulations of the prefrontal cortex (Quirk et al., 2006; Sotres-Bayon et al., 2006), my research suggests that a component of

persistent fear memory lies within the LA, thus providing an alternative target for clinical treatments.

Chapter 2.

Neural correlates of extensive extinction learning in the infralimbic cortex and the amygdala intercalated neurons

Abstract

Repeated presentations of the conditioned stimuli (CS) in the absence of aversive outcomes lead to a weakening of the conditioned fear responses, a process known to extinction. It has been believed that fear extinction recruits inhibitory network involving the infralimbic cortex (IL) and the amygdala-intercalated neurons (ITC), leading to the suppression of fear responses. Accordingly, CS-evoked responses in the IL and ITC cell activities develop after extinction. However, the long-term effects of extensive extinction learning on the inhibitory network have not been explored. Here I show that the CS-responses of IL neurons which emerged after single extinction dissipated with additional extinction sessions. The CS-evoked responses of IL neurons appeared in rats that showed less freezing in the recall of the first extinction session, but not in rats with high freezing. Surprisingly, the CS-evoked responses of IL neurons observed in the recall of the initial extinction disappeared with additional CS presentations in the same session and the CS-responses of IL never emerged in the subsequent extinction and recall sessions. In keeping with this, I also showed that ITC lesions resulted in marked deficits in the expression of

extinction caused no deficit if lesions were made after multiple extinction sessions. This first longitudinal report on the inhibitory network activity during extensive extinction learning suggests that single and extensive extinction involve different neural mechanisms and provides insight into the treatments of aberrant fear memory-related disorders.

Key words: Infralimbic cortex, Intercalated amygdala neurons, fear extinction

Introduction

Repeated presentations of the conditioned stimuli (CS) in the absence of the unconditioned stimuli (US) leads to a weakening of the conditioned response (CR), eventually to the point where the CR disappears. This phenomenon is termed as extinction and has been used as a useful animal model for the treatment of aberrant fear memory-related disorders (Maren and Quirk, 2004; Barad, 2005; Myers and Davis, 2007). However, substantial remnants of the originally learned fear survive even after extensive extinction and cause the re-appearance of fear-related behavior in a variety of circumstances, such as fear renewal and spontaneous recovery (Bouton, 2002; Myers and Davis, 2007). These observations suggest that extinction does not lead to complete reversal of original fear learning, but rather a unique state in which original traces are inhibited temporarily.

The infralimbic cortex (IL), the ventromedial part of the prefrontal cortex, has been considered as a negative regulator of aversive conditioning (Sotres-Bayon and Quirk, 2010). IL neuronal activities are potentiated in animals that successfully retrieved with extinction (Milad and Quirk, 2002; Knapska and Maren, 2009) and stimulation of IL facilitates extinction (Milad and Quirk, 2002). Moreover, NMDA receptor blockers infused into

the IL immediately following extinction impair the retrieval of extinction, suggesting that neuronal plasticity in the IL is crucial for the consolidation of extinction memory (Falls et al., 1992; Burgos-Robles et al., 2007; Sotres-Bayon et al., 2009).

Intercalated amygdala neurons (ITC), a probable mediator of prefrontal inhibition over the amygdala (Royer et al., 1999; Pape and Pare, 2010; Pare and Duvarci, 2012) receives a dense projection from the IL (Sesack et al., 1989; McDonald et al., 1996; Freedman et al., 2000) and the basolateral amygdala (BLA) and sends its inhibitory outputs to the medial subnuclei of the central amygdala (CeM) (Pare and Smith, 1993b, a), the main output nucleus of the amygdala for conditioned fear responses (Davis and Whalen, 2001). Fear extinction potentiates BLA inputs to the ITC cells that project to the CeM, which requires IL activity (Amano et al., 2010). ITC lesions impaired the recall of extinction and activation of ITC cells facilitated extinction (Jungling et al., 2008; Likhtik et al., 2008).

Although accumulating evidence indicates that the inhibitory network consisting of the prefrontal cortex and inhibitory neurons in the amygdala is crucial for fear extinction, most previous studies employed short behavioral procedures consisted of single extinction, thus falling short of demonstrating the long-term modulation of fear memory involving

extensive extinction. I thereby used high signal-to-noise ratio single unit recordings and biochemical lesions to track longitudinal changes in inhibitory network during three extinction sessions. My results revealed that CS-responses of IL neurons which emerged after single extinction session dissipated with additional extinction sessions. Moreover, ITC lesions which impaired the expression of single extinction caused no deficit if lesions were made after three extinction sessions, suggesting that different neural mechanisms underlie in single and extensive extinction.

Materials and Methods

Animals. Male Sprague-Dawley rats (n=101, 8 weeks old) were individually housed for 4~5 days before all experiments under an inverted 12 hours light/dark cycle (lights off at 09:00) and provided with food and water ad libitum. Behavioral training was done in the dark portion of the cycle (An et al., 2012). All procedures were approved by the Institute of Laboratory Animal Resources of Seoul National University.

Surgery. Rats were anesthetized with sodium pentobarbital (50 mg/kg, i.p.) and maintained with isoflurane (1~1.5%) in O2. For the IL recordings, rats were secured in a stereotaxic frame and bilaterally implanted with angled fixed-wire electrodes targeted to the IL: 2.85 mm anterior to bregma, 1.2 to 1.5 mm lateral to midline, and 4.2 to 4.6 mm deep from the cortical surface. The electrodes consisted of 8 individually insulated nichrome microwires (50 μm outer diameter, impedance 0.5~1 MΩ at 1 kHz; California Fine Wire) contained in a 21 gauge stainless steel guide cannula. The implant was secured using dental cement (Vertex). After surgery, analgesia (Metacam, Boehringer) and antibiotics were applied and rats were allowed to recover for 6~7 days. For the ITC lesion study, rats with ≤15% freezing at the end of the first extinction session were secured in a stereotaxic frame. Either D-Sap

(3 pmol/0.3 μl/hemisphere) or the same concentration and volume of a scrambled peptide conjugated to saporin (B-Sap; Advanced Targeting systems) were bilaterally infused through a micro-syringe (30 gauge) targeted to the ITC: 2.65 mm posterior to bregma, 4.75 mm lateral to midline, and 8.65 mm deep from the cortical surface. The micro-syringe was removed ten minutes after the end of the infusion to minimize diffusion along the needle tract.

Apparatus. In all experiments, fear conditioning and extinction took place in two different contexts (context A and B) to minimize the influence of contextual associations. Context A was a rectangular Plexiglas box with a metal grid floor connected to an electrical current source (Coulbourn Instruments) which was set in a sound attenuating chamber. The chamber was illuminated with white light and was cleaned with a 70% ethanol solution. Context B was a cylindrical Plexiglas chamber, with a metal grid floor which was illuminated with a red light for IL unit recordings (An et al., 2012) and a flat black Formica floor with the light off for ITC lesions (Kim et al., 2010) and the both were cleaned with 1% acetic acid. All of the training sessions were videotaped and conditioned freezing was quantified by trained observers.

Behavioral procedures. For IL unit recordings, rats were first habituated to the context and the CS in context A, in which they were placed in the recording chamber twice for 10 min, first without any cue and later with one CS presentation (Pre-habituation). The CS was a 30 s 4 kHz pure tone (85 dB sound pressure level) (Milad and Quirk, 2002). On day 2, rats were given 5 presentations of the CS to determine basal IL neural responses to the CS (Hab). Fear conditioning was conducted by pairing the CS with a mild electric foot shock (0.5 mA, 0.5 s, 5 CS/US pairings; inter-trial interval: 80~120 s) co-terminating with the CS. Extinction training took place 8 hours after fear conditioning in context B, in which rats were presented with 20 non-reinforced CS presentations (Post-Cond). Two additional extinction sessions were conducted on the next day. On day 4, the behavioral and neuronal outcome of three extinction sessions was observed in a short 5 CS test session (Post-Ext3).

For ITC lesions, rats were first habituated to the context A, in which they were placed in the recording chamber for 20 min (habituation). On day 2, fear conditioning was conducted by pairing the CS with a mild electric foot shock (0.4 mA, 1 s, 4 CS/US pairings; inter-trial interval: 80~120 s) co-terminating with the CS (Likhtik et al., 2008). The CS was a 30 s 4 kHz

pure tone (85 dB sound pressure level). On the next day, extinction training took place in context B. Two additional extinction sessions were conducted to investigate the effects of extensive extinction on the inhibitory network involving ITC. The animals were considered to be freezing when there was no movement except for respiratory activity for 2 s during the 30 s CS presentation. The total freezing time was normalized to the duration of the CS presentation (Kim et al., 2010).

Single-unit spike sorting and analysis. Neural activity was acquired and analyzed using a Plexon MAP system, as previously described (Herry et al., 2008). Unit discrimination was performed using Offline Sorter (OFS, Plexon) as previously described (An et al., 2012). Briefly, all waveforms were plotted in a principal component space and clusters consisting of similar waveforms were defined automatically and manually. Single unit isolation was graded using two statistic parameters, J3 and the Davies-Bouldin validity metric (DB). A high J3 and low DB value indicates a compact, well-separated unit cluster (Nicolelis et al., 2003), and neurons with a low grade were discarded. The long-term stability of a single-unit isolation was determined using Wavetracker (Plexon), in which the principal components of a unit recorded from different sessions were compared, and

the linear correlation values (r) between the template waveforms obtained over the entire set of behavioral sessions (Jackson and Fetz, 2007). Only stable units ($r > 0.97$) were considered for further analysis.

To investigate the effects of extinction training on the IL cells, CS-evoked neural activities were normalized using a standard z-score transformation (bin size, 100ms). Unit responses were normalized to the firing rates of four pre-tone bins. Z-score peri-event time histograms (PETHs) of averaged CS-responses were constructed for each neuron and then averaged for every CS. The mean z-values of 0~400 ms following CS-onset from the first 5 CSs of each session were compared throughout the course of behavioral training.

Histology. To identify location of recording microwires, rats were anesthetized with urethane (1 g/kg, i.p.) and electrolytic lesions were made by passing a current (10 μA, 5~20 s) through recording microwires from which discrete units were identified at the end of experiments (An et al., 2012). Animals were then transcardially perfused with 0.9% saline solution and 10% buffered formalin. Brains were removed and post-fixated overnight. Coronal sections (90 μm thick) were obtained using a vibroslicer (NVSL; World Precision Instruments) and stained with cresyl violet. The placement

of the recording microwires was examined under a light microscope.

To reveal μOR immunoreactivity, rats were anesthetized with urethane and transcardially perfused. Brains were removed and post-fixated overnight. The amygdala-containing sections (60 μm thick) were obtained from 2.0~3.0 posterior to bregma using a vibroslicer (NVSL; World Precision Instruments) and stored in PBS. The sections were incubated in 1% sodium borohydride for 30 min and pre-incubated in a blocking solution (10% goat serum, 1% BSA, 0.3% Triton-X100). Then, sections were incubated in the primary antibody solution containing μOR (ImmunoStar, 1:2000) and NeuN antibody (ImmunoStar, 1:2000) in 1% normal goat serum, 1% BSA, and 0.3% Triton-X100 in PBS for 1 hr, followed by incubation in the cocktail of the fluorescent secondary antibodies (Merck, 1:500) for 2 hrs. Cell counting was conducted as previously described (Likhtik et al., 2008), but slightly modified. Contour areas that are stained for μOR and located between the BLA complex and the CeA were defined as ITC regions. In 1-in-4 series of sections, the regions of interest (ROI) were systematically sampled (ITC counting frame, 25 X 25 μm; grid size, 45 X 43 μm; CEA, counting frame, 35 X 35 μm; grid size, 115 X 115 μm) and NeuN-positive cells in the ROI areas were counted. The optical dissector height was 10 μm.

Statistical analysis. To compare the behavioral and neural responses among behavioral sessions, averaged data points were analyzed using repeated-measures ANOVA with subsequent Newman-Keuls post hoc comparison. A probability value of $p < 0.05$ was considered indicative of statistical significance.

Results

IL neuronal activities represent CS-US dissociation after single extinction, but not after extensive extinction

It has been reported that responses of IL neurons to the CS, which emerged in the retrieval phase of extinction in fear extinguished rats, were inversely correlated with freezing at the retrieval test (Milad and Quirk, 2002). However, neural representations of extinction memory involving multiple extinction sessions have remained obscure because previous study has employed short behavioral procedures. Therefore, I investigated IL responses to the CS during multiple extinction sessions.

To investigate the effects of extensive extinction on IL neuronal activity, I employed an extensive extinction paradigm consisting of fear conditioning and subsequent three extinction sessions and IL neuronal activities were recorded throughout the behavioral training. A total of 19 rats underwent an extensive extinction paradigm as described (see Methods) (Fig. 20A) and their fear levels to the CS were examined. Eight hours after the initial fear learning, rats displayed robust freezing when they were exposed to the CS in a different context ($F(4,94) = 110.1$, $p < 0.0001$,

repeated-measures ANOVA; Hab vs. Post-Cond, $p < 0.05$, Newman-Keuls posttest) (Fig. 20C). The conditioned fear progressively diminished over three extinction sessions (Fig. 20B) and freezing levels of the rats in the last test session became undistinguishable from the pre-conditioning levels (Hab vs. Post-Ext3, $p > 0.05$, Newman-Keuls posttest).

A total of 72 cells were recorded from the IL across three days. Histological analysis revealed that recorded cells were located within the anterior part of the IL (Fig. 20D). IL neurons displayed low spontaneous firing rates, averaged firing rate of 0.98 Hz. The average basal firing rates were not different among the behavioral sessions ($F(4,349) = 1.64$, $p > 0.1$, repeated-measures ANOVA). Only stable, high signal-to-noise ratio IL neurons verified by principal component comparisons and correlation analysis were included in the data analysis (Fig. 21).

I constructed a population z-score PETH throughout the behavioral training to investigate the effects of extensive extinction learning on the neural responses of IL to the auditory CS. Since responses of IL neurons to the CS have been shown to be inversely correlated with freezing at the retrieval test, rats were divided into two groups; one with ≤ 50% recovery of freezing (n=14) and the other with > 50% recovery of freezing (n=5) in the early part of the second extinction (Milad and Quirk, 2002) (Fig. 23). In

accordance with previous results, IL neurons signaled extinguished CS in the retrieval session of fear extinction, while they were unresponsive to the CS during the first extinction session (Fig. 22B). The CS-evoked excitation of IL neurons emerged after extinction training was found only in rats with low recovery of freezing (Fig. 22B), suggesting IL neuronal responses is important for the retrieval of extinction memory. Surprisingly, subsequent extinction abolished the CS-evoked excitation of IL neurons and IL neurons remained silent during the additional extinction session and the test session on the next day (Fig. 22B). In rats with high recovery of freezing, however, CS-evoked responses of IL neurons were largely unchanged throughout the course of fear learning involving extensive extinction.

The CS-evoked responses of IL neurons were quantified as a mean z-value of 0~400 ms following the first 5 CSs and compared throughout the behavioral training. Single extinction significantly increased the averaged CS-response of IL neurons compared to the preceding two sessions ($F(4,279) = 3.35$, $p < 0.01$, one-way ANOVA; Post-Ext1 vs. Hab, Post-Ext1 vs. Post-Cond, $p < 0.05$, Newman-Keuls posttest) in rats with low recovery of freezing (Fig. 23B), whereas IL neuronal responses were not altered in rats with high recovery of freezing ($F(4,79) = 3.52$, $p > 0.5$, one-way ANOVA) (Fig. 23B). Intriguingly, CS-responses of IL neurons in rats with low fear

recovery decreased to the habituation level in the following extinction session (Post-Ext2 vs. Post-Ext1, $p < 0.05$, Post-Ext2 vs. Hab, $p > 0.05$, Newman-Keuls posttest), although rats still successfully retrieved with extinction memory (Fig. 20B). Moreover, CS-evoked responses of IL neurons were not found in all rats during the test session conducted on day 4 (Post-Ext3 vs. Hab, $p > 0.05$, Newman-Keuls posttest for the low fear recovery group), suggesting IL neuronal activity is not required for the expression of extinction memory after extensive extinction. I further analyzed IL neuronal activity in the early and the late part of each extinction session to see if the CS-responses of IL neurons alter within the extinction sessions. It was found that CS-responses of IL neurons which emerged at the start of the second extinction disappeared at the end of the same session ($F(7,385) = 3.12$, $p < 0.005$, one-way ANOVA; Post-Ext1 early vs. late, $p < 0.05$, Post-Ext1 late vs. Hab, $p > 0.05$, Newman-Keuls posttest) (Fig. 23C). Collectively, these results suggest that IL neuronal activity is differently involved in single and extensive extinction learning.

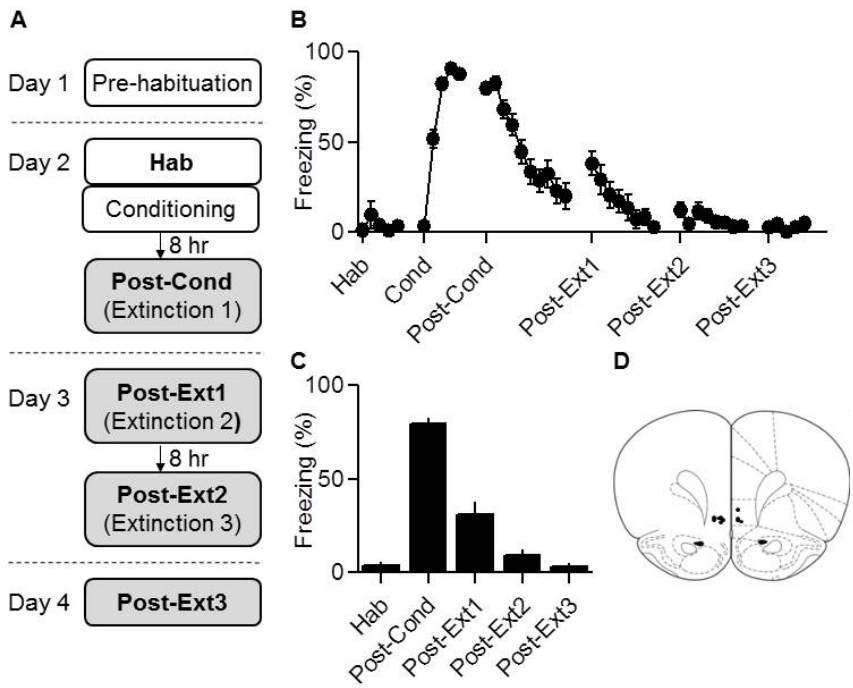

Figure 20. Behavioral procedures and results. A, The behavioral procedure used in the experiment. The white and gray shades represent different contexts. **B,** The learning curves of the entire behavioral session. **C,** Averaged freezing responses during the first five CS presentations of the retention test sessions (bold characters in A) in all rats (n=19). Error bars indicate SEM. Abbreviations: Hab, habituation; Post-Cond, post-conditioning; Post-Ext, post-extinction. **D,** Histological verification of the electrodes placements.

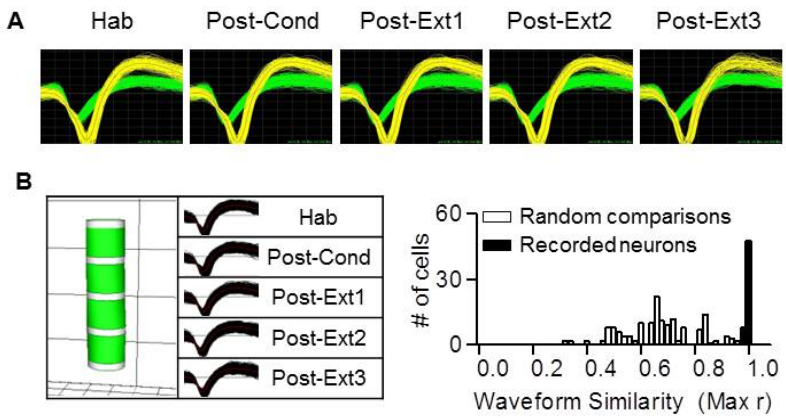

Figure 21. Long-term single unit recordings in the IL. A, Representative waveforms of two neurons recorded from a single electrode and stably observed throughout the behavioral training period. Grid: 55μV, 100μs. **B**, Verification of long-term stable single unit recordings using principal component space cylinders (Left). A straight cylinder suggests that the same set of single units was recorded in different behavioral sessions. Quantitative evaluation of waveform similarity from units recorded on different days. Randomly selected waveforms were used as a control (Right).

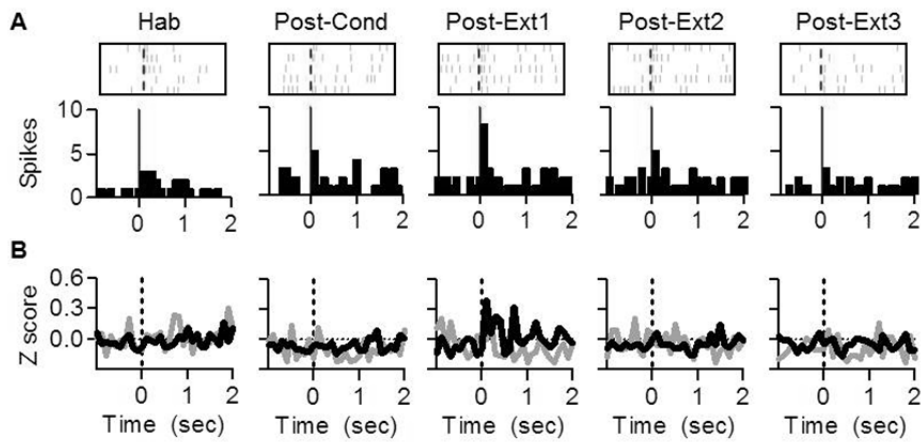

Figure 22. IL neuronal responses to the CS during fear learning. IL neurons represent extinguished CS after single extinction, but not after multiple extinction sessions. **A,** Representative neurons displaying CS-evoked responses after the first extinction. Responses decreased during subsequent extinction and test. **B,** Averaged responses of IL neurons in rats with a good recall of extinction memory (n=14, Black line) and rats with higher freezing (n=5, Gray line).

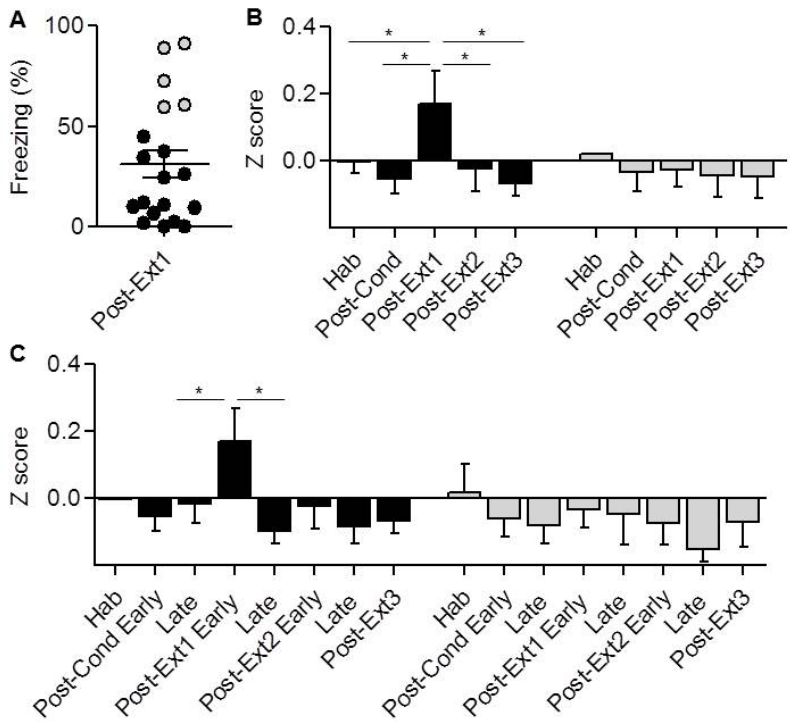

Figure 23. Quantification of IL responses to the CS. A, Rats were divided into two groups, according to their freezing levels in the second extinction session (Post-Ext1). **B,** Comparison of mean z-values calculated in a period of 0~400 ms following CS-onset. The low recovery of fear group displayed CS-responses in the second extinction session, retrieving the initial extinction memory. **C**, The CS-responses of IL neurons emerged in the early part of the second extinction and disappeared in the late part of the same session. Error bars indicate SEM.

Amygdala intercalated neurons are required for the expression of single extinction, but not extensive extinction

Thus far, I have demonstrated that IL neurons signal extinguished CS only after single extinction, but not after extensive extinction learning. I next tested whether ITC, which is the most probable mediator of prefrontal inhibition over the amygdala, is also involved in single and extensive extinction differently. To address this, I employed selective ITC lesions with a ribosome inactivating toxin (D-Sap) that was conjugated to an agonist with a high selectivity and affinity for μ-opioid receptors (μORs), dermorphin (Pare and Smith, 1993a). It has been reported that μORs are more abundantly expressed among ITC neurons, compared to adjacent BA or CeA cells (Likhtik et al., 2008). As a control, a scrambled peptide conjugated to toxin (B-Sap) was utilized.

I first tested the effects of selective ITC lesions obtained by the toxin on single extinction. Rats underwent a single extinction paradigm as described (Fig. 24) and their fear levels to the CS were examined. Either D-Sap or the same concentration and volume of a control peptide was bilaterally infused to the ITC the day after extinction session. After 7 days of recovery, the retrieval of extinction memory was tested and freezing levels to the CS were quantified in a blind manner. Only rats with syringe

tips located at the BLA-CeA border were included. In consistent with previous study (Likhtik et al., 2008), D-Sap infusions resulted in a marked reduction in μOR staining restricted to the region adjacent to infusion site, whereas more distant ITC clusters at the external capsule were not affected (Fig. 25A). μOR expression was not altered in B-Sap treated rats (Fig. 25B).

To evaluate the selective ITC lesions obtained by D-Sap infusions, I performed unbiased stereological estimates of the number of NeuN positive cells. Compared to B-Sap treated rats, the number of ITC neurons were significantly reduced in rats that received D-Sap infusions into the ITC (Fig. 25C; B-Sap, 136.6 ± 16.3, n=10; D-Sap, 62.3 ± 7.5, n=12; $p < 0.001$, unpaired t test). In contrast, the number of CeA neurons were identical in the two groups (Fig. 25C; B-Sap, 717.3 ± 23.4, n=6; D-Sap, 671.3 ± 45.5, n=6; $p > 0.1$, unpaired t test). Consistent with the previous report which showed inverse correlation between freezing levels during extinction recall and the number of survived ITC cells (Likhtik et al., 2008), D-Sap infused rats displayed impaired expression of extinction memory, whereas rats with B-Sap infusions successfully retrieved with extinction (Fig. 26; B-Sap, 29.1 ± 5.2; D-Sap, 60.0 ± 7.8; $p < 0.05$, unpaired t test). These results suggest IL neuronal activity is required for the expression of extinction memory after single extinction.

Having established the effects of selective ITC lesion on single extinction, I next examined the effects of ITC lesions on extensive extinction by employing two additional extinction sessions. D-Sap or B-Sap infusions were conducted the day after the last extinction session. Consistent with the single extinction experiment, the number of ITC neurons were significantly decreased in rats that received D-Sap infusions in the ITC compared to B-Sap treated rats (Fig. 27C; B-Sap, 161.6 ± 13.2, n=10; D-Sap, 57.93 ± 9.4; $p < 0.0001$, unpaired t test). The number of CeA neurons was identical in the two behavioral groups (B-Sap, 700.8 ± 28.1, n=6; D-Sap, 689.0 ± 27.8; $p > 0.5$, unpaired t test). Surprisingly, freezing levels of D-Sap infused rats in the recall test were not different from those of B-Sap treated rats (Fig. 27B; B-Sap, 20.8 ± 6.7; D-Sap, 14.1 ± 3.8; $p > 0.1$, unpaired t test), although toxin-mediated selective ITC lesions were effective as much as shown in the single extinction experiment. I further confirmed that fear renewal, which is one of the behavioral characteristics of fear extinction besides spontaneous recovery and savings, was normally induced after single (Fig. 28A) and extensive extinction (Fig. 28B) paradigm, suggesting that both single and extensive extinction did not erase the original fear memory. Rats displayed strong freezing when they were exposed to the context where fear conditioning had occurred, whereas no fear responses

were observed when rats were exposed to the extinction context, no matter how many extinction sessions they had experienced (Fig. 28A; ABA, 21.7 ± 1.3, ABB, 9.2 ± 2.3, $p < 0.0001$, unpaired t test for single extinction group) (Fig. 28B; ABA, 21.1 ± 2.2, ABB, 6.1 ± 2.7, $p < 0.005$, unpaired t test for extensive extinction group). Collectively, these results suggest that ITC neuronal activity is not required for the maintenance and the expression of extinction memory in extensive extinction learning consisting of three extinction sessions.

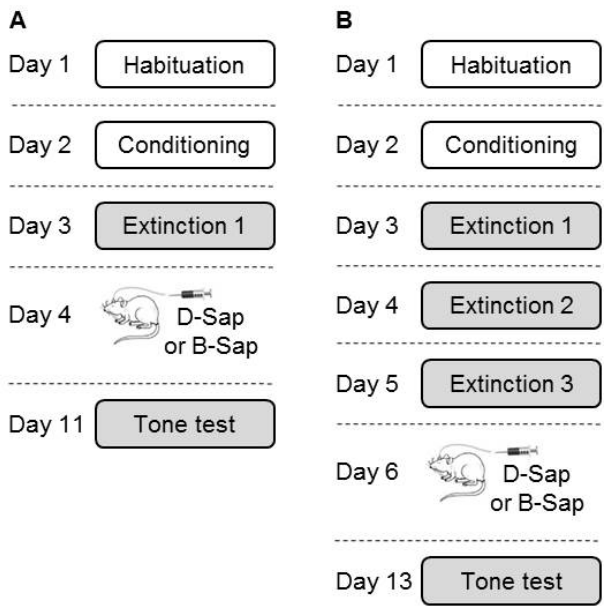

Figure 24. Experimental designs. Behavioral training and toxin infusions in **A,** single extinction and **B,** extensive extinction paradigm. Either toxin (D-Sap) or control toxin (B-Sap) was infused the next day of the last extinction session. Extinction recall was tested after 7 days of recovery.

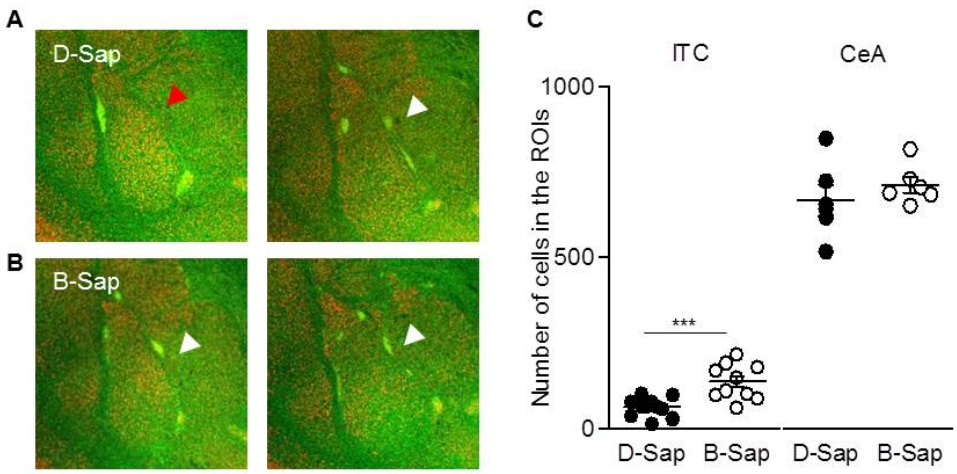

Figure 25. Selective ITC lesions. A, μOR staining in rats infused with D-Sap. μOR staining is reduced adjacent to infusion site (Red arrow), whereas distant ITC clusters were not affected (White arrow). **B**, μOR staining was not decreased by B-Sap infusion. **C**, Number of NeuN-positive cells in the ITC and the CeA. The number of ITC neurons is decreased in D-Sap treated rats, compared to the B-Sap infused rats. CeA neurons were not affected by D-Sap or B-Sap infusion. Error bars indicate SEM. Abbreviations: ITC, intercalated cells; CeA, central amygdala.

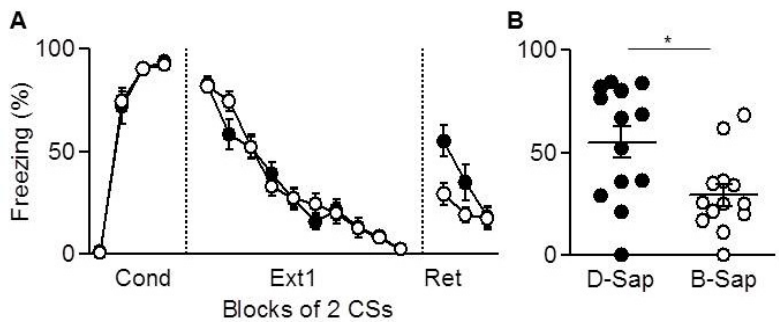

Figure 26. The effects of ITC lesions on single extinction. A, The learning curves of the entire behavioral session. On the next day of extinction, either D-Sap (Black circle) or B-Sap (White circle) was infused aimed to the ITC. Extinction memory was tested after 7 days of recovery. **B,** D-Sap treated rats displayed higher freezing in the test session, compared to the B-Sap treated rats. Error bars indicate SEM. Abbreviations: Cond, fear conditioning; Ext, extinction; Ret, retrieval session.

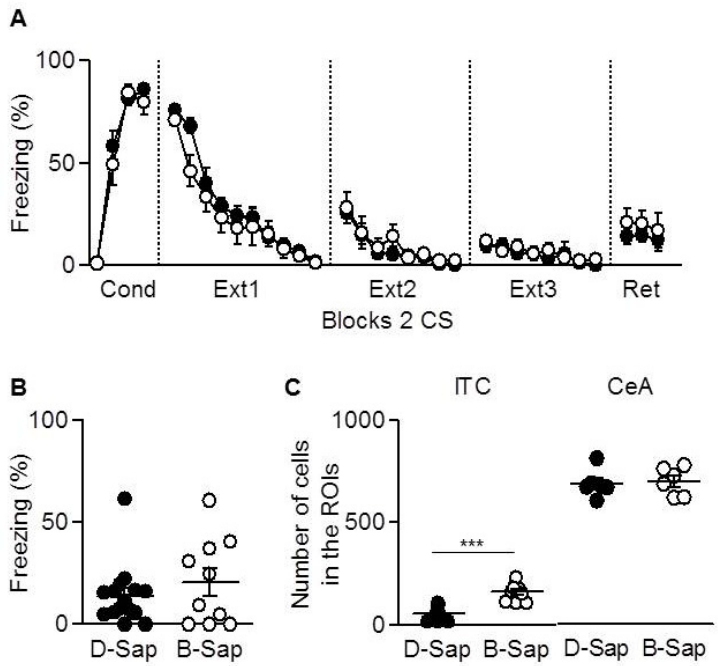

Figure 27. The effects of ITC lesions on extensive extinction. A, The learning curves of the entire behavioral session. On the next day of the last extinction, either D-Sap (Black circle) or B-Sap (White circle) was infused aimed to the ITC. Extinction memory was tested after 7 days of recovery. **B**, D-Sap treated rats displayed low freezing responses in the test session, similar to the B-Sap treated rats. **C,** Number of NeuN-positive cells in the ITC and the CeA. The number of ITC neurons is decreased in D-Sap treated rats, compared to the B-Sap infused rats. CeA neurons were not affected by D-Sap or B-Sap infusion. Error bars indicate SEM.

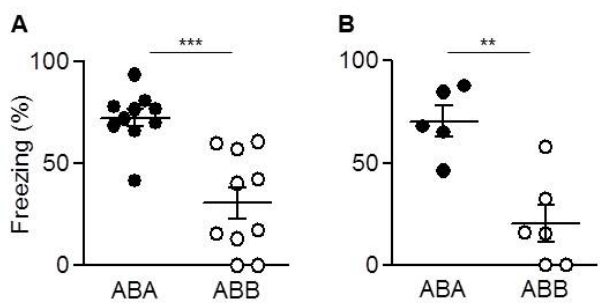

Figure 28. Renewal of fear in single and extensive extinction. Fear responses were examined in the same context where the extinction (ABB retention, white circle) or the fear conditioning (ABA renewal, black circle) took place. When rats were exposed to the conditioning context, renewal of fear was observed in rats that underwent **A,** single extinction and **B,** multiple extinction sessions. Error bars indicate SEM.

Discussion

It has been believed that the inhibitory network, including the prefrontal cortex and the inhibitory neurons in the amygdala, is critical in the acquisition and the expression of extinction memory (Maren and Quirk, 2004; Pape and Pare, 2010; Sotres-Bayon and Quirk, 2010). However, I found that the IL and the ITC, the essential brain regions constituting the inhibitory network, were crucial for single extinction, but not for extensive extinction. Consistent with previous report (Milad and Quirk, 2002), IL neurons only in rats which showed successful recall of extinction memory displayed increased CS-evoked firing in the retrieval session after the first extinction session. However, CS-responses of IL neurons decreased to pre-training level during the same session and never emerged in subsequent extinction sessions. In keeping with these results, I also showed that ITC lesions which resulted in a marked deficit in the expression of single extinction caused no deficit if lesions were made after multiple extinction sessions. Together, these results suggest that the inhibitory network is crucial for single extinction training, however, a different neural network is recruited with additional extinction sessions.

IL has long been considered as a critical regulator of aversive

conditioning (Maren and Quirk, 2004; Quirk et al., 2006; Sotres-Bayon and Quirk, 2010). NMDA receptor blockers infused into the IL immediately following extinction impair the retrieval of extinction, suggesting that neuronal plasticity in the IL is crucial for the consolidation of extinction memory (Falls et al., 1992; Burgos-Robles et al., 2007; Sotres-Bayon et al., 2009). Consistent with previous reports (Milad and Quirk, 2002; Knapska and Maren, 2009), I have observed CS-evoked excitation of IL neurons emerged after extinction in rats that successfully retrieved with extinction (Fig. 22). These potentiated CS-responses of IL neurons after extinction have been considered to mediate the consolidation and the expression of extinction memory. IL is reciprocally connected with the BLA in which a neuronal population representing extinguished CS has been reported (Herry et al., 2008). NMDA receptor blockers and protein kinase inhibitors infused into the BLA impair fear extinction, suggesting that neuronal plasticity in the BLA is crucial for the extinction of conditioned fear (Falls et al., 1992; LeDoux, 2000). IL also sends robust projections to the ITC (Sesack et al., 1989; McDonald et al., 1996) which in turn strongly inhibit output from the central nucleus of the amygdala (Royer et al., 1999), leading to the suppression of fear conditioned responses after extinction. Recently, it was reported that theta synchronization between the prefrontal cortex and the

BLA increase in response to safe cues that are not associated with noxious shocks (Likhtik et al., 2014), suggesting the IL might generally represent learned safety. Importantly, I found that CS no longer elicited excitatory responses in the IL when rats underwent additional extinction sessions (Fig. 22), suggesting IL neuronal responses is not required for the expression of extinction memory in extensive extinction. It has been reported that the cortical areas represent salient events and the saliency-related cortical activities rapidly disappear with repeated exposures to the events (Ranganath and Rainer, 2003). It is possible that IL responses to the extinguished CS might represent saliency of the CS which has been dissociated from the US. Thus, IL responses would decrease with repetitive CS presentations, since CS-US dissociation became firm and thus less salient.

ITC is one of probable mediators of prefrontal inhibition over the amygdala after extinction (Royer et al., 1999; Pape and Pare, 2010; Pare and Duvarci, 2012). Fear extinction potentiates BLA inputs to the ITC cells that project to the CeM and synaptic potentiation between the BLA and the ITC is impaired by IL inactivation (Amano et al., 2010). Consistent with previous reports (Jungling et al., 2008; Likhtik et al., 2008), I found that ITC lesions following single extinction impaired the retrieval of extinction

memory (Fig. 26), suggesting ITC is critical for single extinction. ITC receives a dense projection from the IL (Sesack et al., 1989; McDonald et al., 1996; Freedman et al., 2000) and the BLA and sends its inhibitory outputs to the CeM (Pare and Smith, 1993b, a), the main output nucleus of the amygdala for conditioned fear responses (Davis and Whalen, 2001), so as to inhibit conditioned fear behavior after extinction. However, I found that ITC lesions no longer affect the expression of extinction memory when the lesions were made after three extinction sessions (Fig. 27), suggesting ITC neuronal activity is not required for the inhibition of conditioned fear behavior after extensive extinction. Extinction recall after extensive extinction is likely to be mediated by decreased LA inputs to the CeM. LA synaptic inputs are depotentiated after extinction learning (Kim et al., 2007) and I also observed that LA ensemble activity to the CS decreased after extensive extinction (An et al., 2012).

Traces of persistent fear memory have been suggested to reside in cortical regions (Corcoran and Quirk, 2007; Burgos-Robles et al., 2009; Sacco and Sacchetti, 2010; Sotres-Bayon and Quirk, 2010). In the previous chapter, I found a subset of LA neurons also represents the original CS-US association even after extensive extinction (‘extinction-resistant fear neurons’). It has been believed that well-known inhibitory mechanisms

involving the prefrontal cortex (Milad and Quirk, 2002; Rosenkranz et al., 2003; Likhtik et al., 2005; Sotres-Bayon et al., 2006; Quirk and Mueller, 2008) and the ITC neurons (Chhatwal et al., 2005; Likhtik et al., 2008; Ehrlich et al., 2009) may provide inhibition at the BA or the CeM leading to the suppression of fear responses. Accordingly, re-appearance of fear memory after extinction has been regarded to be mediated by the context-dependent disinhibition of the inhibitory network over the amygdala (Hobin et al., 2003; Likhtik et al., 2008; Ehrlich et al., 2009). However, my results indicate that the essential brain regions constituting the inhibitory network, the prefrontal cortex and the amygdala ITC neurons play minor roles in extensive extinction, although the renewal of fear is normally observed. It is possible that the inhibitory network supports the expression of fear extinction in the beginning and additional extinction trainings recruit other brain network. Further researches are required to understand how the LA and other brain network support later savings or memory relapse after extensive extinction when the inhibitory influences of the prefrontal cortex disappeared. Metaplastic mechanisms that enable more rapid synaptic plasticity at input synapses may also support the enhanced potentiation of CS-responses in these neurons (Abraham, 2008; Lee et al., 2013). LA neurons representing the original fear memory after extensive extinction

may also play an important role in the persistence of fear memory and relapse after extinction (An et al., 2012).

Fear conditioning and extinction have served as primary models for the treatment of PTSD and other anxiety disorders. Although most PTSD research aimed at preventing the relapse of fear memory has focused on the dysfunctions or manipulations of the prefrontal cortex (Quirk et al., 2006; Sotres-Bayon et al., 2006), my results suggests that the inhibitory influences of the prefrontal cortex over the amygdala is no longer critical for the maintenance and the expression of extensive extinction. It is consistent with clinical studies which showed the connectivity between the prefrontal cortex and the amygdala progressively decreased with repetitive presentations of the traumatic script (Gilboa et al., 2004; Rauch et al., 2006). Further researches will be required to find appropriate targets for clinical treatment of fear-related mental disorders using the extensive extinction paradigm.

References

Abraham WC (2008) Metaplasticity: tuning synapses and networks for plasticity. Nature Reviews Neuroscience 9:387.

Amano T, Unal CT, Pare D (2010) Synaptic correlates of fear extinction in the amygdala. Nature Neuroscience 13:489-494.

Amir A, Amano T, Pare D (2011) Physiological identification and infralimbic responsiveness of rat intercalated amygdala neurons. Journal of Neurophysiology 105:3054-3066.

An B, Hong I, Choi S (2012) Long-term neural correlates of reversible fear learning in the lateral amygdala. The Journal of Neuroscience 32:16845-16856.

Barad M (2005) Fear extinction in rodents: basic insight to clinical promise. Current Opinion in Neurobiology 15:710-715.

Barad M (2006) Is extinction of fear erasure or inhibition? Why both, of course. Learning & Memory 13:108-109.

Blair HT, Schafe GE, Bauer EP, Rodrigues SM, LeDoux JE (2001) Synaptic plasticity in the lateral amygdala: a cellular hypothesis of fear conditioning. Learning & Memory 8:229-242.

Bordi F, LeDoux J, Clugnet MC, Pavlides C (1993) Single-unit activity in the lateral nucleus of the amygdala and overlying areas of the striatum in freely behaving rats: rates, discharge patterns, and

responses to acoustic stimuli. Behavioral Neuroscience 107:757-769.

Bouton ME (2002) Context, ambiguity, and unlearning: sources of relapse after behavioral extinction. Biological Psychiatry 52:976-986.

Bouton ME, King DA (1983) Contextual control of the extinction of conditioned fear: tests for the associative value of the context. Journal of Experimental Psychology 9:248-265.

Burgos-Robles A, Vidal-Gonzalez I, Quirk GJ (2009) Sustained conditioned responses in prelimbic prefrontal neurons are correlated with fear expression and extinction failure. The Journal of Neuroscience 29:8474-8482.

Burgos-Robles A, Vidal-Gonzalez I, Santini E, Quirk GJ (2007) Consolidation of fear extinction requires NMDA receptor-dependent bursting in the ventromedial prefrontal cortex. Neuron 53:871-880.

Carew TJ, Walters ET, Kandel ER (1981) Associative learning in Aplysia: cellular correlates supporting a conditioned fear hypothesis. Science 211:501-504.

Chang CH, Knapska E, Orsini CA, Rabinak CA, Zimmerman JM, Maren S (2009) Fear extinction in rodents. Current Protocols in Neuroscience 8:8-23.

Chhatwal JP, Myers KM, Ressler KJ, Davis M (2005) Regulation of

gephyrin and GABAA receptor binding within the amygdala after fear acquisition and extinction. The Journal of Neuroscience 25:502-506.

Collins DR, Pare D (2000) Differential fear conditioning induces reciprocal changes in the sensory responses of lateral amygdala neurons to the CS(+) and CS(-). Learning & Memory 7:97-103.

Corcoran KA, Maren S (2001) Hippocampal inactivation disrupts contextual retrieval of fear memory after extinction. The Journal of Neuroscience 21:1720-1726.

Corcoran KA, Quirk GJ (2007) Activity in prelimbic cortex is necessary for the expression of learned, but not innate, fears. The Journal of Neuroscience 27:840-844.

Corcoran KA, Desmond TJ, Frey KA, Maren S (2005) Hippocampal inactivation disrupts the acquisition and contextual encoding of fear extinction. The Journal of Neuroscience 25:8978-8987.

Dalton GL, Wang YT, Floresco SB, Phillips AG (2008) Disruption of AMPA receptor endocytosis impairs the extinction, but not acquisition of learned fear. Neuropsychopharmacology 33:2416-2426.

Davis M (1992) The role of the amygdala in fear and anxiety. Annual Review of Neuroscience 15:353-375.

Davis M, Whalen PJ (2001) The amygdala: vigilance and emotion. Molecular Psychiatry 6:13-34.

Davis M, Rainnie D, Cassell M (1994) Neurotransmission in the rat amygdala related to fear and anxiety. Trends in Neurosciences 17:208-214.

Duclos Y, Schmied A, Burle B, Burnet H, Rossi-Durand C (2008) Anticipatory changes in human motoneuron discharge patterns during motor preparation. The Journal of Physiology 586:1017-1028.

Ehrlich I, Humeau Y, Grenier F, Ciocchi S, Herry C, Luthi A (2009) Amygdala inhibitory circuits and the control of fear memory. Neuron 62:757-771.

Falls WA, Miserendino MJ, Davis M (1992) Extinction of fear-potentiated startle: blockade by infusion of an NMDA antagonist into the amygdala. The Journal of Neuroscience 12:854-863.

Fanselow MS (1994) Neural organization of the defensive behavior system responsible for fear. Psychonomic Bulletin & Review 1:429-438.

Fanselow MS (2000) Contextual fear, gestalt memories, and the hippocampus. Behavioural Brain Research 110:73-81.

Freedman LJ, Insel TR, Smith Y (2000) Subcortical projections of area 25 (subgenual cortex) of the macaque monkey. Journal of Comparative

Neurology 421:172-188.

Gale GD, Anagnostaras SG, Godsil BP, Mitchell S, Nozawa T, Sage JR, Wiltgen B, Fanselow MS (2004) Role of the basolateral amygdala in the storage of fear memories across the adult lifetime of rats. The Journal of Neuroscience 24:3810-3815.

Gilboa A, Shalev AY, Laor L, Lester H, Louzoun Y, Chisin R, Bonne O (2004) Functional connectivity of the prefrontal cortex and the amygdala in posttraumatic stress disorder. Biological Psychiatry 55:263-272.

Goosens KA, Hobin JA, Maren S (2003) Auditory-evoked spike firing in the lateral amygdala and Pavlovian fear conditioning: mnemonic code or fear bias? Neuron 40:1013-1022.

Hall J, Thomas KL, Everitt BJ (2001) Cellular imaging of zif268 expression in the hippocampus and amygdala during contextual and cued fear memory retrieval: selective activation of hippocampal CA1 neurons during the recall of contextual memories. The Journal of Neuroscience 21:2186-2193.

Han JH, Kushner SA, Yiu AP, Cole CJ, Matynia A, Brown RA, Neve RL, Guzowski JF, Silva AJ, Josselyn SA (2007) Neuronal competition and selection during memory formation. Science 316:457-460.

Han JH, Kushner SA, Yiu AP, Hsiang HL, Buch T, Waisman A, Bontempi B,

Neve RL, Frankland PW, Josselyn SA (2009) Selective erasure of a fear memory. Science 323:1492-1496.

Haubensak W, Kunwar PS, Cai H, Ciocchi S, Wall NR, Ponnusamy R, Biag J, Dong HW, Deisseroth K, Callaway EM, Fanselow MS, Luthi A, Anderson DJ (2010) Genetic dissection of an amygdala microcircuit that gates conditioned fear. Nature 468:270-276.

Herry C, Garcia R (2003) Behavioral and paired-pulse facilitation analyses of long-lasting depression at excitatory synapses in the medial prefrontal cortex in mice. Behavioural Brain Research 146:89-96.

Herry C, Ciocchi S, Senn V, Demmou L, Muller C, Luthi A (2008) Switching on and off fear by distinct neuronal circuits. Nature 454:600-606.

Hobin JA, Goosens KA, Maren S (2003) Context-dependent neuronal activity in the lateral amygdala represents fear memories after extinction. The Journal of Neuroscience 23:8410-8416.

Hobin JA, Ji J, Maren S (2006) Ventral hippocampal muscimol disrupts context-specific fear memory retrieval after extinction in rats. Hippocampus 16:174-182.

Hong I, Song B, Lee S, Kim J, Kim J, Choi S (2009) Extinction of cued fear memory involves a distinct form of depotentiation at cortical input synapses onto the lateral amygdala. The European Journal of

Neuroscience 30:2089-2099.

Hong I, Kim J, Lee J, Park S, Song B, Kim J, An B, Park K, Lee HW, Lee S, Kim H, Park SH, Eom KD, Lee S, Choi S (2011) Reversible plasticity of fear memory-encoding amygdala synaptic circuits even after fear memory consolidation. PloS One 6:e24260.

Jackson A, Fetz EE (2007) Compact movable microwire array for long-term chronic unit recording in cerebral cortex of primates. Journal of Neurophysiology 98:3109-3118.

Jungling K, Seidenbecher T, Sosulina L, Lesting J, Sangha S, Clark SD, Okamura N, Duangdao DM, Xu YL, Reinscheid RK, Pape HC (2008) Neuropeptide S-mediated control of fear expression and extinction: role of intercalated GABAergic neurons in the amygdala. Neuron 59:298-310.

Kapp BS, Frysinger RC, Gallagher M, Haselton JR (1979) Amygdala central nucleus lesions: effect on heart rate conditioning in the rabbit. Physiology & Behavior 23:1109-1117.

Kehoe EJ (1988) A layered network model of associative learning: learning to learn and configuration. Psychological Review 95:411-433.

Kim J, Song B, Hong I, Kim J, Lee J, Park S, Eom JY, Lee CJ, Lee S, Choi S (2010) Reactivation of fear memory renders consolidated amygdala synapses labile. The Journal of Neuroscience 30:9631-

9640.

Kim J, Lee S, Park K, Hong I, Song B, Son G, Park H, Kim WR, Park E, Choe HK, Kim H, Lee C, Sun W, Kim K, Shin KS, Choi S (2007) Amygdala depotentiation and fear extinction. Proceedings of the National Academy of Sciences of the United States of America 104:20955-20960.

Knapska E, Maren S (2009) Reciprocal patterns of c-Fos expression in the medial prefrontal cortex and amygdala after extinction and renewal of conditioned fear. Learning & Memory 16:486-493.

LaBar KS, LeDoux JE, Spencer DD, Phelps EA (1995) Impaired fear conditioning following unilateral temporal lobectomy in humans. The Journal of Neuroscience 15:6846-6855.

Lau HL, Timbers TA, Mahmoud R, Rankin CH (2013) Genetic dissection of memory for associative and non-associative learning in Caenorhabditis elegans. Genes, Brain, and Behavior 12:210-223.

Laurent V, Westbrook RF (2009) Inactivation of the infralimbic but not the prelimbic cortex impairs consolidation and retrieval of fear extinction. Learning & Memory 16:520-529.

LeDoux JE (2000) Emotion circuits in the brain. Annual Review of Neuroscience 23:155-184.

LeDoux JE (2014) Coming to terms with fear. Proceedings of the National Academy of Sciences of the United States of America 111:2871-2878.

LeDoux JE, Farb C, Ruggiero DA (1990) Topographic organization of neurons in the acoustic thalamus that project to the amygdala. The Journal of Neuroscience 10:1043-1054.

Lee S, Song B, Kim J, Park K, Hong I, An B, Song S, Lee J, Park S, Kim J, Park D, Lee CJ, Kim K, Shin KS, Tsien RW, Choi S (2013) GluA1 phosphorylation at serine 831 in the lateral amygdala is required for fear renewal. Nature Neuroscience 16:1436-1444.

Likhtik E, Pelletier JG, Paz R, Pare D (2005) Prefrontal control of the amygdala. The Journal of Neuroscience 25:7429-7437.

Likhtik E, Popa D, Apergis-Schoute J, Fidacaro GA, Pare D (2008) Amygdala intercalated neurons are required for expression of fear extinction. Nature 454:642-645.

Likhtik E, Stujenske JM, Topiwala MA, Harris AZ, Gordon JA (2014) Prefrontal entrainment of amygdala activity signals safety in learned fear and innate anxiety. Nature Neuroscience 17:106-113.

Lin HC, Mao SC, Gean PW (2009) Block of gamma-aminobutyric acid-A receptor insertion in the amygdala impairs extinction of conditioned fear. Biological Psychiatry 66:665-673.

Lolordo VM (1969) Positive conditioned reinforcement from aversive situations. Psychological Bulletin 72:193-203.

Maren S (2000) Auditory fear conditioning increases CS-elicited spike firing in lateral amygdala neurons even after extensive overtraining. The European Journal of Neuroscience 12:4047-4054.

Maren S (2011) Seeking a spotless mind: extinction, deconsolidation, and erasure of fear memory. Neuron 70:830-845.

Maren S, Fanselow MS (1996) The amygdala and fear conditioning: has the nut been cracked? Neuron 16:237-240.

Maren S, Quirk GJ (2004) Neuronal signalling of fear memory. Nature Reviews Neuroscience 5:844-852.

McDonald AJ (1982) Neurons of the lateral and basolateral amygdaloid nuclei: a Golgi study in the rat. The Journal of Comparative Neurology 212:293-312.

McDonald AJ, Mascagni F, Guo L (1996) Projections of the medial and lateral prefrontal cortices to the amygdala: a Phaseolus vulgaris leucoagglutinin study in the rat. Neuroscience 71:55-75.

McEchron MD, McCabe PM, Green EJ, Llabre MM, Schneiderman N (1995) Simultaneous single unit recording in the medial nucleus of the medial geniculate nucleus and amygdaloid central nucleus

throughout habituation, acquisition, and extinction of the rabbit's classically conditioned heart rate. Brain Research 682:157-166.

McGaugh JL (2000) Memory--a century of consolidation. Science 287:248-251.

McKernan MG, Shinnick-Gallagher P (1997) Fear conditioning induces a lasting potentiation of synaptic currents in vitro. Nature 390:607-611.

Medina JF, Repa JC, Mauk MD, LeDoux JE (2002) Parallels between cerebellum- and amygdala-dependent conditioning. Nature Reviews Neuroscience 3:122-131.

Milad MR, Quirk GJ (2002) Neurons in medial prefrontal cortex signal memory for fear extinction. Nature 420:70-74.

Milad MR, Quirk GJ (2012) Fear extinction as a model for translational neuroscience: ten years of progress. Annual Review of Psychology 63:129-151.

Miserendino MJ, Sananes CB, Melia KR, Davis M (1990) Blocking of acquisition but not expression of conditioned fear-potentiated startle by NMDA antagonists in the amygdala. Nature 345:716-718.

Myers KM, Davis M (2007) Mechanisms of fear extinction. Molecular Psychiatry 12:120-150.

Myers KM, Ressler KJ, Davis M (2006) Different mechanisms of fear

extinction dependent on length of time since fear acquisition. Learning & Memory 13:216-223.

Napier RM, Macrae M, Kehoe EJ (1992) Rapid reaquisition in conditioning of the rabbit's nictitating membrane response. Journal of Experimental Psychology 18:182-192.

Nicolelis MA, Dimitrov D, Carmena JM, Crist R, Lehew G, Kralik JD, Wise SP (2003) Chronic, multisite, multielectrode recordings in macaque monkeys. Proceedings of the National Academy of Sciences of the United States of America 100:11041-11046.

Oyama K, Hernadi I, Iijima T, Tsutsui K (2010) Reward prediction error coding in dorsal striatal neurons. The Journal of Neuroscience 30:11447-11457.

Pape HC, Pare D (2010) Plastic synaptic networks of the amygdala for the acquisition, expression, and extinction of conditioned fear. Physiological Reviews 90:419-463.

Pare D, Smith Y (1993a) Distribution of GABA immunoreactivity in the amygdaloid complex of the cat. Neuroscience 57:1061-1076.

Pare D, Smith Y (1993b) The intercalated cell masses project to the central and medial nuclei of the amygdala in cats. Neuroscience 57:1077-1090.

Pare D, Gaudreau H (1996) Projection cells and interneurons of the lateral and basolateral amygdala: distinct firing patterns and differential relation to theta and delta rhythms in conscious cats. The Journal of Neuroscience 16:3334-3350.

Pare D, Collins DR (2000) Neuronal correlates of fear in the lateral amygdala: multiple extracellular recordings in conscious cats. The Journal of Neuroscience 20:2701-2710.

Pare D, Duvarci S (2012) Amygdala microcircuits mediating fear expression and extinction. Current Opinion in Neurobiology 22:717-723.

Pare D, Quirk GJ, Ledoux JE (2004) New vistas on amygdala networks in conditioned fear. Journal of Neurophysiology 92:1-9.

Paton JJ, Belova MA, Morrison SE, Salzman CD (2006) The primate amygdala represents the positive and negative value of visual stimuli during learning. Nature 439:865-870.

Pearce JM, Hall G (1980) A model for Pavlovian learning: variations in the effectiveness of conditioned but not of unconditioned stimuli. Psychological Review 87:532-552.

Phelps EA, LeDoux JE (2005) Contributions of the amygdala to emotion processing: from animal models to human behavior. Neuron 48:175-187.

Phillips RG, LeDoux JE (1992) Differential contribution of amygdala and hippocampus to cued and contextual fear conditioning. Behavioral Neuroscience 106:274-285.

Pitkanen A, Savander V, LeDoux JE (1997) Organization of intra-amygdaloid circuitries in the rat: an emerging framework for understanding functions of the amygdala. Trends in Neurosciences 20:517-523.

Quirk GJ, Mueller D (2008) Neural mechanisms of extinction learning and retrieval. Neuropsychopharmacology 33:56-72.

Quirk GJ, Repa C, LeDoux JE (1995) Fear conditioning enhances short-latency auditory responses of lateral amygdala neurons: parallel recordings in the freely behaving rat. Neuron 15:1029-1039.

Quirk GJ, Armony JL, LeDoux JE (1997) Fear conditioning enhances different temporal components of tone-evoked spike trains in auditory cortex and lateral amygdala. Neuron 19:613-624.

Quirk GJ, Garcia R, Gonzalez-Lima F (2006) Prefrontal mechanisms in extinction of conditioned fear. Biological Psychiatry 60:337-343.

Ranganath C, Rainer G (2003) Neural mechanisms for detecting and remembering novel events. Nature Reviews Neuroscience 4:193-202.

Rauch SL, Shin LM, Phelps EA (2006) Neurocircuitry models of

posttraumatic stress disorder and extinction: human neuroimaging research--past, present, and future. Biological Psychiatry 60:376-382.

Reijmers LG, Perkins BL, Matsuo N, Mayford M (2007) Localization of a stable neural correlate of associative memory. Science 317:1230-1233.

Repa JC, Muller J, Apergis J, Desrochers TM, Zhou Y, LeDoux JE (2001) Two different lateral amygdala cell populations contribute to the initiation and storage of memory. Nature Neuroscience 4:724-731.

Rescorla RA (2001) Retraining of extinguished Pavlovian stimuli. Journal of Experimental Psychology 27:115-124.

Rogan MT, Staubli UV, LeDoux JE (1997) Fear conditioning induces associative long-term potentiation in the amygdala. Nature 390:604-607.

Rogan MT, Leon KS, Perez DL, Kandel ER (2005) Distinct neural signatures for safety and danger in the amygdala and striatum of the mouse. Neuron 46:309-320.

Romanski LM, Clugnet MC, Bordi F, LeDoux JE (1993) Somatosensory and auditory convergence in the lateral nucleus of the amygdala. Behavioral Neuroscience 107:444-450.

Rosenkranz JA, Moore H, Grace AA (2003) The prefrontal cortex regulates

lateral amygdala neuronal plasticity and responses to previously conditioned stimuli. The Journal of Neuroscience 23:11054-11064.

Royer S, Martina M, Pare D (1999) An inhibitory interface gates impulse traffic between the input and output stations of the amygdala. The Journal of Neuroscience 19:10575-10583.

Sacco T, Sacchetti B (2010) Role of secondary sensory cortices in emotional memory storage and retrieval in rats. Science 329:649-656.

Santini E, Quirk GJ, Porter JT (2008) Fear conditioning and extinction differentially modify the intrinsic excitability of infralimbic neurons. The Journal of Neuroscience 28:4028-4036.

Schafe GE, LeDoux JE (2000) Memory consolidation of auditory pavlovian fear conditioning requires protein synthesis and protein kinase A in the amygdala. The Journal of Neuroscience 20:RC96.

Schafe GE, Atkins CM, Swank MW, Bauer EP, Sweatt JD, LeDoux JE (2000) Activation of ERK/MAP kinase in the amygdala is required for memory consolidation of pavlovian fear conditioning. The Journal of Neuroscience 20:8177-8187.

Schoenbaum G, Chiba AA, Gallagher M (1999) Neural encoding in orbitofrontal cortex and basolateral amygdala during olfactory discrimination learning. The Journal of Neuroscience 19:1876-1884.

Seidenbecher T, Laxmi TR, Stork O, Pape HC (2003) Amygdalar and hippocampal theta rhythm synchronization during fear memory retrieval. Science 301:846-850.

Sesack SR, Deutch AY, Roth RH, Bunney BS (1989) Topographical organization of the efferent projections of the medial prefrontal cortex in the rat: an anterograde tract-tracing study with Phaseolus vulgaris leucoagglutinin. The Journal of Comparative Neurology 290:213-242.

Sotres-Bayon F, Quirk GJ (2010) Prefrontal control of fear: more than just extinction. Current Opinion in Neurobiology 20:231-235.

Sotres-Bayon F, Cain CK, LeDoux JE (2006) Brain mechanisms of fear extinction: historical perspectives on the contribution of prefrontal cortex. Biological Psychiatry 60:329-336.

Sotres-Bayon F, Bush DE, LeDoux JE (2007) Acquisition of fear extinction requires activation of NR2B-containing NMDA receptors in the lateral amygdala. Neuropsychopharmacology 32:1929-1940.

Sotres-Bayon F, Diaz-Mataix L, Bush DE, LeDoux JE (2009) Dissociable roles for the ventromedial prefrontal cortex and amygdala in fear extinction: NR2B contribution. Cerebral Cortex 19:474-482.

Tseng WT, Yen CT, Tsai ML (2011) A bundled microwire array for long-term chronic single-unit recording in deep brain regions of behaving

rats. Journal of Neuroscience Methods 201:368-376.

Tsvetkov E, Carlezon WA, Benes FM, Kandel ER, Bolshakov VY (2002) Fear conditioning occludes LTP-induced presynaptic enhancement of synaptic transmission in the cortical pathway to the lateral amygdala. Neuron 34:289-300.

Zhou Y, Won J, Karlsson MG, Zhou M, Rogerson T, Balaji J, Neve R, Poirazi P, Silva AJ (2009) CREB regulates excitability and the allocation of memory to subsets of neurons in the amygdala. Nature Neuroscience 12:1438-1443.

국문초록

공포학습 시 편도체 및 변연계아래피질의 신경활성

서울대학교 대학원
협동과정 뇌과학 전공
안 보 배

중성적 자극과 유해한 자극을 반복적으로 제시하여 이를 연합하는 공포 조건화 학습 방법은 외상 후 스트레스 장애 등 공포관련 질환의 동물 모델로 유용하게 사용되어 왔다. 과거 수많은 연구자들은 공포 조건화 학습 모델을 이용하여 편도체 및 그와 연결된 신경네트워크가 공포 학습 및 소거에 필수적임을 제안하였다.

그러나 이전 연구들은 단기 공포 학습 모델을 이용함으로써, 공포 학습 및 소거가 편도체 및 신경네트워크에 미치는 장기적 영향에 대해서는 밝히지 못하였다. 그러므로 본 연구에서는 장기 공포 학습 및 반복 소거 학습이 편도체 및 신경네트워크에 미치는 영향을 살펴보고자 하였다. 제 1장에서는 장기 공포학습 및 소거, 재학습 동안 공포 연합 학습의 중추로 알려진 등쪽 편도체 내 신경세포의 활성을 관찰하였다. 일련의 실험을 통하여 등쪽 편도체 내 신경세포들이 역동적으로 변화하는 공포 연합 기억을 표상함을 발견하였다. 나아가, 등쪽 편도체 내 공포 소거 학습 기억을 표상하는 집단(공포 소거 순응 신경세포)과 공포 소거 학습 기억을 표상하지 않는 집단 (공포 소거 저항 신경세포)이 있음을 발견하였다. 이러한 결과는 등쪽 편도체가 공포 조건화 학습의 다양한 측면을 표상함을 의미한다.

제 2장에서는 장기 공포학습 및 반복 소거 학습 동안 공포 소거 학습의 중추로 알려진 편도체 및 변연계아래피질의 활성을 관찰하였다. 일련의 실험을 통하여 변연계아래피질 신경세포들이 단일 공포 소거 기억은 표상하지만, 반복 공포 소거 기억은 표상하지 않음을 발견하였다. 또한 편도체 내 억제 신경세포의 활성이 반복 소거 학습 시 필요하지 않음을 발견하였다. 이러한 결과는 단일 및 반복 공포 소거 학습이 다른 신경학적 기전에 의해 매개됨을 의미한다.

요약적으로, 본 연구는 장기 공포 학습이 편도체 및 신경네

트워크에 미치는 영향을 살펴보았다. 먼저, 등쪽 편도체 신경세포가 공포 연합 기억의 다양한 측면을 역동적으로 표상함을 관찰하였다. 다음으로, 편도체와 변연계아래피질의 신경 활성이 단일 공포 소거 학습에는 중요하지만, 반복 소거 학습에는 필요하지 않음을 발견하였다. 이러한 결과들은 공포 기억이 조절되는 신경학적 기반에 대한 이해를 도모하고, 나아가 공포 관련 정신 질환 치료의 기반을 제시한다.

핵심어: 편도체, 변연계아래피질, 공포 조건화 학습, 공포 소거 학습

학번: 2009-20445