Systems/Circuits

# Infralimbic Projections to the Substantia Innominata–Ventral Pallidum Constrain Defensive Behavior during Extinction Learning

Carolina Fernandes-Henriques,[1,2]* Yuval Guetta,[3]* Mia G. Sclar,[2]* Rebecca Zhang,[2] Yuka Miura,[1,2] Katherine R. Surrence,[2] Allyson K. Friedman,[1,2] and Ekaterina Likhtik[1,2]

[1]Biology Program, The Graduate Center, CUNY, New York 10016 and Departments of, [2]Biological Sciences, and [3]Psychology, Hunter College, CUNY, New York 10065

Fear extinction is critical for decreasing fear responses to a stimulus that is no longer threatening. While it is known that the infralimbic (IL) region of the medial prefrontal cortex mediates retrieval of an extinction memory through projections to the basolateral amygdala (BLA), IL pathways contributing to extinction learning are not well understood. Given the dense projection from the IL to the substantia innominata–ventral pallidum (SI/VP), an area that processes aversive and appetitive cues, we compared how the IL→SI/VP functions in extinction compared with the IL→BLA pathway in male mice. Using retrograde tracing, we demonstrate that IL projections to the SI/VP originate from superficial [Layer (L)2/3] and deep cortical layers (L5) and that they are denser than IL projections to the BLA. Next, combining retrograde tracing with labeling for the immediate early gene cFos, we show increased activity of L5 IL→SI/VP output during extinction learning and increased activity of L2/3 IL→BLA output during extinction retrieval. Then, using in vitro recordings, we demonstrate that neurons in the IL→SI/VP pathway are more excitable during extinction learning than retrieval. Finally, using optogenetics, we inactivate the IL→SI/VP pathway and show that this increases defensive freezing during extinction learning and re-extinction, without affecting memory. Taken together, we demonstrate that the IL→SI/VP pathway is active during extinction learning, when it constrains the defensive freezing response. We propose that the IL acts as a switchboard operator, increasing IL L5 communication with the SI/VP during extinction learning and IL L2/3 communication with the BLA during extinction retrieval.

*Key words:* amygdala; fear extinction; infralimbic; learning; substantia innominata; ventral pallidum

---

## Significance Statement

Fear extinction is a widely used behavioral approach to decrease conditioned fear, and projections from the infralimbic cortex to the amygdala are known to mediate extinction memory retrieval. However, less is known about the role of infralimbic pathways in extinction learning. We use neuroanatomical tracing, behavior, slice recordings, and circuit manipulation to show that infralimbic output to the substantia innominata–ventral pallidum (SI/VP), a region that processes aversive and appetitive stimuli, is denser than to the amygdala and is more active during extinction learning than retrieval, when it acts to constrain the defensive freezing response. Thus, we posit that during extinction, the infralimbic uses several lines of communication, one with the SI/VP during learning one with the amygdala during retrieval.

---

## Introduction

The ventral portion of the medial prefrontal cortex (mPFC), known as the infralimbic (IL) region in rodents, is critical for fear extinction whereas the more dorsal prelimbic (PL) is more associated with fear and discrimination behavior (Milad and Quirk, 2002; Phelps et al., 2004; Milad et al., 2007; Sierra-Mercado et al., 2011; Meyer and Bucci, 2014). IL neurons fire during extinction consolidation and retrieval (Milad and Quirk, 2002; Burgos-Robles et al., 2007), and inhibition of the IL impedes extinction memory consolidation and retrieval, whereas IL stimulation accelerates extinction acquisition and retrieval (Vidal-Gonzalez et al., 2006; Burgos-Robles et al.,

2007; Laurent and Westbrook, 2009; Sotres-Bayon et al., 2009; Do-Monte et al., 2015; Kim et al., 2016; Szeska et al., 2022; Bayer et al., 2024). However, the pathways through which IL integrates various components of learning are not well known, complicating translatability to humans (Roberts and Clarke, 2019). For example, overactivation of the mPFC in rodents and nonhuman primates drives movement-based behavioral and cardiovascular responses to threat (Halladay and Blair, 2017; Alexander et al., 2020). Likewise, in rats, IL neural firing decreases during defensive freezing (Giustino et al., 2016), suggesting that this region may bias behavior away from defensive freezing and toward more movement, which is observed both during extinction and when processing future threats (Wallis et al., 2017; Grunfeld and Likhtik, 2018; Alexander et al., 2020). Thus, it is critical to understand how IL interactions with target structures affect extinction behavior.

Prefrontal projections to the basolateral complex of the amygdala (BLA) have been identified as a key pathway that undergoes synaptic plasticity after extinction learning (Amano et al., 2010; Cho et al., 2013) and is critical for extinction memory retrieval (Likhtik et al., 2008; Amir et al., 2011; Bukalo et al., 2015; Bloodgood et al., 2018). Accordingly, inhibiting IL projections to the BLA during extinction learning does not impair extinction acquisition but hampers extinction retrieval, indicating that IL inputs to the BLA mediate extinction memory consolidation (Laurent and Westbrook, 2009; Bukalo et al., 2015; Do-Monte et al., 2015; Bloodgood et al., 2018). However, little is known regarding the contribution of other IL pathways to extinction.

Using rats, the IL was shown to prominently project to the substantia innominata (SI) and the ventral pallidum (VP) of the basal forebrain (BF), the center of attentional processing (Room et al., 1985; Hurley et al., 1991; Zaborszky et al., 1997; Vertes, 2004). The SI/VP contains a mix of cholinergic, GABAergic, and glutamatergic neurons, with the cholinergic cells shown to modulate cued fear learning via projections to the amygdala and cortex (Henny and Jones, 2008; McDonald et al., 2011; Unal et al., 2015; Xu et al., 2015; Crimmins et al., 2023; Bratsch-Prince et al., 2024; Rajebhosale et al., 2024). Furthermore, GABAergic and glutamatergic VP neurons differentially modulate behavior in response to appetitive and aversive cues (Lin and Nicolelis, 2008; Stephenson-Jones et al., 2020; Moaddab et al., 2021; Hegedüs et al., 2024). Thus, given the importance of IL communication for extinction and the prominence of the IL→SI/VP projection in rats, we asked if this pathway is involved in extinction. To this end, we first use retrograde tracing to compare the density and distribution of prefrontal pathways emanating to the SI/VP versus the BLA and show that prefrontal output to the SI/VP is overall denser than to the BLA. Then, using the immediate early gene cFos, we show that IL→SI/VP Layer (L)5 projections are more active during extinction learning, whereas IL→BLA L2/3 projections are more active when defensive freezing is lower during extinction retrieval. Then, using in vitro recordings, we demonstrate that L5 IL→SI/VP projectors are more excitable during extinction learning than retrieval. Finally, we demonstrate that inhibiting the IL→SI/VP, but not the PL→SI/VP pathway during extinction learning, impairs the within-session decrement in freezing without affecting memory. Taken together, we show that L5 IL→SI/VP pathway constrains defensive fear expression during active extinction learning.

## Materials and Methods

*Animals.* Adult male C57BL/6J mice (Jackson Laboratory), aged 9–11 weeks, were group housed (2–4 per cage) under a 12 h light/dark cycle (lights on 8A.M. to 8P.M.) with *ad libitum* access to food and water.

All procedures were conducted under the regulation of the Hunter College Institutional Animal Care and Use Committee.

*Surgeries—microinjections and optic fiber implantations.* For all surgeries, mice were anesthetized with 2% isoflurane in oxygen, placed in a stereotaxic frame (Kopf Instruments) and maintained on 1.5% isoflurane throughout surgery (oxygen at a flow rate of 1 L/min). Temperature was maintained at 37C ± 1C with a feedback-regulated heating pad. Mice received dexamethasone (1 mg/ml, s.c.) and bupivacaine under the scalp (5 mg/ml, s.c.) prior to incision.

For tracing/cFos experiments, two craniotomies were performed using a drill with a burr attachment. Mice were injected unilaterally in the right hemisphere with the 0.3 μl of the retrograde tracers cholera toxin subunit B (CTB) and Alexa Fluor 488 and 647 (Invitrogen) into the SI/VP (−0.15 mm AP, +1.5 mm ML, −4.5 mm DV from the brain surface) and BLA (−1.6 mm AP, +3.15 mm ML, −4.1 mm DV from the brain surface), counterbalanced by fluorophore and region, at a rate of 0.08 μl/min with 10 μl Hamilton syringes (QSW Stereotax Injector, Stoelting). The craniotomies were then closed with bone wax, and the skin was closed with poly(glycolide-coacaprolactone) monofilament absorbable sutures (Covetrus). Postoperatively, all mice received carpofen for pain relief (1 mg/ml, i.p.) and were group housed in cages warmed by a heating pad until recovery. Mice were allowed to recover from surgery for at least 1 week prior to handling.

For the optogenetic manipulation experiments, mice (n = 5–9/grp) were injected bilaterally with 0.15 μl of the inhibitory anterograde virus rAAV5-hSyn-eArch3.0-EYFP or its matched control rAAV5-hSyn-EYFP (10 × 1,012 vg/ml; UNC Vector Core) into the IL (+1.6 mm AP, ±0.4 mm ML, −2.0 mm DV from the brain surface) or the PL (+1.6 mm AP, ±0.4 mm ML, −1.6 mm DV from the brain surface) at a rate of 0.08 μl/min with 10 μl Hamilton syringes (QSI Stereotax Injector, Stoelting). Custom-order ferrules with attached optic fibers (exposed fiber length, 6 mm, Newdoon) were then implanted over the BF (−0.15 mm AP, ±1.5 mm ML, −4.4 mm DV from the brain surface) and cemented onto the skull using both opaque C&B Metabond (Parkell) and an additional layer of different-colored dental cement (Teets, Lang Dental) for animal identification. Mice were allowed to recover on a heating pad, and then the virus was left to express for 4–5 weeks prior to handling.

For the in vitro recordings, n = 16 mice were injected bilaterally with 0.3 μl of the retrograde virus rAAV-retro-hSyn-eYFP (1 × 1012 vg/ml; UNC Vector Core) into the SI/VP (−0.15 mm AP, ±1.5 mm ML, −4.4 mm DV from the brain surface), and the skin was closed with poly(glycolide-coacaprolactone) monofilament absorbable sutures (Covetrus). Mice were allowed to recover for 2 weeks.

*Behavioral experiments.* Context A: Animals underwent fear conditioning in a plexiglass chamber with aluminum walls and a stainless-steel rod floor capable of delivering scrambled footshock (Med Associates). Overhead lamps maintained light levels at ∼40 lux, and the conditioning box was cleaned with ethanol between animals.

*Context B*: Animals underwent extinction in a custom-made gray wood box (45 cm length × 13 cm width × 20 cm height), with a smooth paper floor that was changed between animals. Light levels were maintained at ∼70 lux.

Auditory cues were delivered via an audio speaker (ENV-224AM, Med Associates) located in the wall of the chamber (Context A) or above the enclosure (Context B) at approximately the same height as in Context A. Mice were presented with a 2 kHz pure tone throughout the protocol, except for one group that also received 8 kHz pure tones, all tones delivered at 100 dB. Behavior was recorded using an infrared OptiTrack camera and Neuromotive software running in conjunction with the Central software (Blackrock Neurotech). Timestamped video data were analyzed off-line.

*Behavioral protocols.* Handling and habituation: Mice were brought to the behavioral-adjacent room and allowed to acclimate 1 h before the experiments started each day. Mice were first handled by the experimenter for 5 min. The next day, mice were first handled for 5 min

and, at least 1 h later, they underwent habituation to Context B for ~8 min without any cues or with laser light-on habituation trials in the case of optogenetic experiments. At least 1 h later, they were habituated to Context A, where they were exposed to five trials of the 30-s-long conditioned stimulus (CS), a 2 kHz tone (ITI, 60–120 s). Each CS consisted of 50 ms pips (amplitude modulated with 25 ms linear increase followed by 25 ms linear decrease), delivered once per second for 30 s, as reported previously (Stujenske et al., 2022).

For the experiments assessing IL→SI/VP pathway activity during extinction via cFos, on Day 1, mice were allocated to one of three groups: a tone control group that received the same numbers of tone-alone trials as the other groups but did not undergo associative learning, an extinction learning group, and an extinction retrieval group.

*Tone control*: Day 1, 5 trials of a 2 kHz tone CS in Context A; Day 2, 20 trials of a 2 kHz tone CS in Context B; Day 3, 10 trials of a 2 KHz CS tone in Context B.

*Extinction learning*: Day 1: 5 trials of a 2 kHz tone CS coterminating with a 1 s unconditioned stimulus (US; 0.7 mA scrambled electric footshock) in Context A; Day 2, 20 trials of a novel 8 kHz tone in Context B; Day 3, 5 trials of the 8 kHz tone followed by 5 extinction trials of the fear conditioned 2 kHz CS tone in Context B.

*Extinction retrieval*: Day 1, 5 trials of a 2 kHz tone CS coterminating with a 1 s US (0.7 mA scrambled electric footshock) in Context A; Day 2, 20 extinction trials of the 2 kHz tone CS in Context B; Day 3, 10 extinction retrieval trials of the 2 kHz tone CS in Context B.

All mice were killed and perfused 90 min after the sixth tone was delivered on Day 3 of the protocol to quantify expression of the immediate early gene cFos in IL projectors to the SI/VP or BLA.

For the experiments that were testing IL→SI/VP pathway excitability in vitro, following handling and habituation as described above, mice were divided into one of the four groups.

*Tone control*: Day 1, five trials of 2 kHz tone-alone exposure; Day 2, two 2 kHz tone-alone trials, perfusion 10 min later.

*Early extinction*: Day 1, five trials of 2 kHz CS paired with US (0.7 mA scrambled electric footshock); Day 2, two 2 kHz tone-alone trials of extinction learning, perfusion 10 min later.

*Late extinction*: Day 1, five trials of 2 kHz CS paired with US (0.7 mA scrambled electric footshock); Day 2, 20 2 kHz tone-alone trials of extinction learning, perfusion 10 min later.

*Extinction retrieval*: Day 1, five trials of 2 kHz CS paired with US (0.7 mA scrambled electric footshock); Day 2, 20 2 kHz tone-alone trials of extinction learning; Day 3, two 2 kHz tone-alone trials of extinction retrieval, perfusion 10 min later.

For experiments with optogenetic inhibition of IL inputs to the SI/VP, handling and habituation were as described above, with the exception that during habituation to Context B, all mice were also exposed to five trials of 35 s laser light followed by habituation to the tone CS in Context A. Then, mice underwent fear conditioning (context A), with five CS–US paired trials where the 2 kHz CS delivery coterminated with a 1 s, 0.7 mA footshock US. The next day, during extinction acquisition (Context B), mice received 10 trials of the 2 kHz CS alone, coupled with a green laser (532 nm, continuous stimulation, ~10 mW per hemisphere, ramp-modulated during onset and offset; Laserglow Technologies). The next day, during extinction retrieval (Context B), animals were exposed to 10 trials of the 2 kHz tone-alone.

*Scoring behavior*: For all behavioral analyses, time spent showing defensive freezing was manually quantified by an experimenter blind to group. The scoring consisted of measuring the amount of time spent freezing during the 30 s prior to Trial 1 (baseline) and during the 30 s of all CS presentations. Freezing was defined as total immobility for longer than 1 s, apart from breathing. Periods of immobility when the animal was curled up asleep or not orienting to CS onset or offset more than five trials of moving around during the CS were not counted as freezing.

*Tissue collection and immunostaining*. For the neuroanatomical tracing and cFos immunostaining experiments, mice were deeply anesthetized with a mixture of ketamine (100 mg/kg, i.p.) and xylazine (2 mg/kg, i.p.) and transcardially perfused with cold phosphate-buffered saline (PBS), followed by 4% paraformaldehyde (PFA) in PBS 90 min after the onset of CS #6 on Day 3. Brains were extracted and postfixed

in 4% PFA overnight. After cryoprotection in 30% sucrose in PBS, 40 µ histological sections were prepared on a cryostat (Cryostar) to evaluate (1) SI/VP and BLA injection sites and (2) cFos+ and CTB + mPFC cell bodies. Mounted SI/VP and BLA sections were coverslipped with ProLong Gold plus DAPI antifade mounting medium (Thermo Fisher Scientific) and imaged with a fluorescent episcope (Olympus BX53) to identify CTB injection placements. Tracing and cFos analyses were only carried out in mice with correct placements in SI/VP or BLA and visible CTB+ cells in the mPFC (Fig. 1). Immunohistochemistry was performed on mPFC sections at three defined points: bregma AP +2.0; +1.8; +1.6 mm.

Free-floating sections were subsequently washed in 1× PBS (three times, 5 min), blocked in 5% normal donkey serum (Jackson ImmunoResearch Laboratories) in PBS Triton X-100 1% (Sigma-Aldrich) at room temperature (1 h), and then incubated overnight in rabbit anti c-Fos antibody (1:2,000, Abcam, #ab190289) and mouse anti-myelin basic protein (1:1,000, BioLegend, #808401) in blocking solution at 4C, staining the corpus callosum to visualize the width of the mPFC. The next day, sections were washed in PBS (three times, 5 min) and incubated in donkey anti-rabbit Alexa Fluor 594 secondary antibody (1:500, Life Technologies #A21207) and donkey anti-mouse Alexa Fluor 405 (Invitrogen, #a48257) in blocking solution for 2 h at room temperature. Sections were subsequently washed in 1× PBS (three times, 5 min) before being mounted and coverslipped with ProLong Gold antifade mounting medium (Thermo Fisher Scientific).

*Cell counting*. The three mPFC slices were imaged on a confocal microscope (40×, Leica SP8) with a Z-stack of six optical slices (3–4 µm/slice). Images were analyzed in ImageJ (NIH), with each $z$-stack analyzed as a maximum intensity projection, with regions of interest (ROI) for IL and PL extracted for each slice, based on stereotaxic coordinates (Franklin and Paxinos, 2013). For each ROI, fluorescing cells were counted using the Multipoint tool, and the accompanying $x$ and $y$ coordinates for each cell were saved. The coordinates were analyzed with custom-written scripts (MATLAB) that used the $x$ and $y$ coordinates of each cell to locate it in the mPFC. The $x$-axis was binned (25 µm bins), and densities of cells per bin were calculated for each cell type. To calculate layer-specific parameters, bins encompassing 175–375 µm from the surface were averaged into L2/3, and bins spanning 376–550 µm from cortical surface were averaged as L5 for each animal (Little and Carter, 2012).

*Cell recordings*. Viral injections were targeted to the IL and optrodes, consisting of a 28-channel stereotrode bundle with an optic fiber, and were inserted in the SI/VP. Recordings took place during laser habituation trials in Context B, when light was ramp-modulated to turn on for 35 s. Recordings were made using the Cerebus Neural Signal Processor (Blackrock Neurotech), with filters open, and data sampled at 30 kHz. Recorded channels were then processed using Kilosort2 (post patch) with bandpass filtering between 300 and 6,000 Hz and then manually curated using Phy2. Given that the same units can be seen across some stereotrodes, data were clustered by inputting stereotrode wires as a linear probe, such that 28 channels were considered per each spike. Wires within a stereotrode were adjacent, but otherwise ordering was random. Single-unit and multiunit responses were analyzed in nonoverlapping 500 ms bins in the 35 s before light onset and during the 35 s of light. To compare firing rates, we computed the average firing rate for 35 s prior to light onset and compared it with the 35 s in the light-on period to the firing rate during the 35 s of the light-on period. To compute the percentage change in the firing rate, we took the difference in the firing rate during the 35 s during light relative to the 35 s prelight.

*Slice preparation*. Slice preparation was performed as previously described (Friedman et al., 2016). Artificial cerebrospinal fluid (aCSF) was prepared in the following ion concentration (in mM): 128 NaCl; 10 D-glucose; 1.25 $NaH_2PO_4$; 25 $NaHCO_3$; 2 $MgCl_2$; 3 KCl; and 2 CaCl2. aCSF was ice-cold and oxygenated with 95% oxygen and 5% carbon dioxide. Ten minutes following the behavioral timepoint of interest, mice were anesthetized with isoflurane (1-chloro-2,2,2-trifluoroethyl-

difluoromethylether). After confirming that the mouse was deeply anesthetized, an incision was made on the chest. Ice-cold oxygenated aCSF was trascardially perfused prior to rapid decapitation. After harvesting, the brain was blocked into mPFC-containing and BF-containing block. The mPFC-containing block was fixed on the buffer tray of a Microslicer (Microslicer DTK- 1000, Dosaka EM) and the BF-containing block was soaked in 4% PFA in 1× PBS for placement evaluation. Acute brain slices containing mPFC neurons were cut at 250 μm thick in cold oxygenated sucrose aCSF (in mM: 227 sucrose; 10 D-glucose; 1.25 $NaH_2PO_4$; 24 $NaHCO_3$; 2 $MgCl_2$; 3 KCl; 2 $CaCl_2$) using the microslicer. These slices were then transferred to a recovery chamber with oxygenated aCSF for 1 h at 36°C. The recovery chamber was then moved to room temperature with continuous oxygenation, and slices were used for recording for up to a 4 h period.

*Whole-cell patch–clamp recordings.* Recordings were performed at 37°C using an inline solution heater (SH-27B) and temperature controller (TC-324C, Warner Instruments). Slices were transferred to a recording chamber that was continually perfused with oxygenated aCSF at a flow rate of 3.0 ml/min. Recording pipettes were made from thick-walled borosilicate glass (BF150-86-10, Sutter Instrument). Glass pipets were pulled by P-97 Flaming/Brown micropipette puller (Sutter Instrument).
Patch pipet for whole-cell voltage clamp and current clamp (3–8 mΩ) was filled with internal solution (in mM: 115 K-gluconate; 20 KCl; 1.5 MgCl2; 10 phosphocreatine; 2 K-ATP; 0.5 Na-GTP; 10 HEPES; 284 mOsm), pH 7.4. IL was identified by anatomical location guided by anterior commissure spanning both hemispheres and lateral ventricles and visualized with 4× objective lens (PLN 4×, Olympus). IL neurons were visualized under infrared light with 40× objective (LUMPLFLN, Olympus) immersed in aCSF. eYFP-labeled neurons were visualized with a fluorescent lamp (X-Cite 120Q, Lumen Dynamics) with a 470 nm filter, and recordings were made from eYFP-labeled neurons. Neurons of interest were identified by the presence of fluorescence in the soma. Neurons without fluorescence were not recorded. After the creation of a giga-ohm seal, the cell membrane was ruptured by small suction to create a whole-cell configuration. Different measurements of excitability were taken ∼2 min following the establishment of the whole-cell configuration in the current-clamp mode ($I = 0$): (1) resting membrane potential (RMP); (2) rheobase, i.e., as the minimum amount of current required to fire an action potential (AP) using a current ramp; and (3) the relationship between increasing steps of current and the number of APs fired. For the latter, voltage responses to depolarizing current were recorded from 0 to +200 pA for 200 ms in increments of 10 pA. To control for differences in RMP, current-injection protocols were performed at both RMP and −70 mV. Signals were digitized using a Multiclamp 700 B amplifier (Molecular Devices), and data were acquired using Axon Digidata 1550 B (Molecular Devices). The number of spikes during depolarizing current injection was counted by event detection (Clampfit). Rheobase was counted as the minimum current injection needed to elicit an AP during the excitability protocol.

*Experimental design and statistical analysis.* All analyses were performed with the Prism 10 software except for the mixed-effect model that was constructed in SPSS (version 25). For the anatomy analysis of mPFC projector distributions, we ran a mixed model with fixed effects of projector (SI/VP, BLA), mPFC subregion (PL, IL), and cortical layer (L2/3, L5), and a random intercept for subject. For the cFos analysis, two-way repeated–measure (rm)ANOVAs were used to study differences across groups for the two layers (repeated measures) for each pathway and two-way ANOVAs to study differences between SI/VP and BLA projectors across groups for L2/3 and L5. Due to variability in CTB tracer expression between cohorts, outlier analyses were run on the anatomical tracing experiment, using GrpahPad (Prism 10), which resulted in removal of two animals with abnormally high labeling through layers. For the behavioral analyses, a two-way rmANOVA was performed with "group" and "trial" as factors and Tukey's or Sidák's as the post hoc test. For comparisons between specific behavioral trial bins, RMP, and rheboase across groups, unpaired *t* tests were performed in normal distributions or Mann–Whitney tests if the distributions failed to pass

the normality test. Differences in the number of APs between groups were investigated using a two-way rmANOVA with "current step" and "group" as factors. All data are expressed as mean ± SEM, and significance was defined as $p < 0.05$.

## Results

The mPFC has a strong reciprocal connection from L2/3 with the BLA, which partakes in fear and extinction (Quirk et al., 2003; Little and Carter, 2012; Arruda-Carvalho and Clem, 2014; Burgos-Robles et al., 2017; Klavir et al., 2017; Bloodgood et al., 2018), but mPFC connectivity with the SI/VP and whether it plays a role in extinction are less well understood. Thus, we aimed to compare the distribution of mPFC projections to the SI/VP of the BF with those of the BLA and to determine whether these two pathways have similar or different activity profiles during extinction. Previous work in rats, using the anterograde tracer *Phaseolus vulgaris*-leucoaggultinin (Room et al., 1985; Hurley et al., 1991; Vertes, 2004), showed that the IL has a stronger projection to the SI/VP (BF) than the PL. Thus, we were also interested in understanding whether this projection pattern is retained in mice. To answer these questions, we injected the two different CTB retrograde tracers, one in the SI/VP (BF) and one BLA (Fig. 1A–D), and assigned mice to one of three behavioral conditions, tone-control, extinction learning, or extinction retrieval. We first mapped the distribution of CTB-labeled mPFC soma projecting to the SI/VP and the BLA (Fig. 1E), to evaluate the anatomical distribution of these pathways across mPFC layers (Fig. 1F–K). We then analyzed the overlap of CTB and cFos expression during behavior to investigate pathway-specific activity related to each behavioral condition (Figs. 2, 3).

### mPFC has denser projections to the SI/VP than to the BLA

To compare the distribution of mPFC pathways to the SI/VP and BLA, we injected C57B/6J male mice with the retrograde tracer CTB (0.3 μl) in the posterior SI/VP region of the BF (−0.15 mm AP, +1.5 mm ML, −4.5 mm DV from the brain surface) and in the BLA (−1.6 mm AP, +3.15 mm ML, −4.1 mm DV from the brain surface; Fig. 1A,B). The CTB was counterbalanced for fluorophores (CTB-488 or CTB-647) at each injection site. All sites were checked for correct targeting, and those with incorrect injection placements were removed from analysis, with the resulting $n = 16$ in the SI/VP of the BF and $n = 11$ BLA (Fig. 1C,D). The density and distribution of CTB-expressing soma in the mPFC were analyzed and averaged across three anterior–posterior locations (bregma +2.0 mm, +1.8 mm, +1.6 mm). At each location, the $x$ coordinates of CTB-expressing soma were binned (25 μm bins) and mapped along the mediolateral axis of the mPFC, spanning from the pial surface to the corpus callosum, and then averaged across superficial (L2/3, 175–374 μm from midline) and deep layers (L5, 375–550 μm from midline; Fig. 1E).
Using a mixed-model analysis to account for mPFC cortical region (PL vs IL), subcortical projection target (SI/VP vs BLA), and layer (L2/3 vs L5), we first identified a significant effect of subcortical projection target ($F_{(1,129)} = 54.36$; $p < 0.001$), with a denser mPFC projection to the SI/VP than to the BLA (BF, 75.1 cells/mm2 ± 8.8, vs BLA, 7.6 cells/mm2 ± 10; $p < 0.001$), with both the PL and the IL showing the same pattern (cortical regions × subcortical target; $F_{(1,110)} = 0.64$; $p = 0.43$; Fig. 1F–H). Next, we looked at the density of subcortical innervation from different cortical layers of the mPFC and saw a significant difference in density of SI/VP versus BLA innervation across layers (subcortical target × layer; $F_{(1,109)} = 14.79$; $p < 0.001$). A pairwise comparison showed that mPFC (PL and IL) projections to the SI/VP are denser
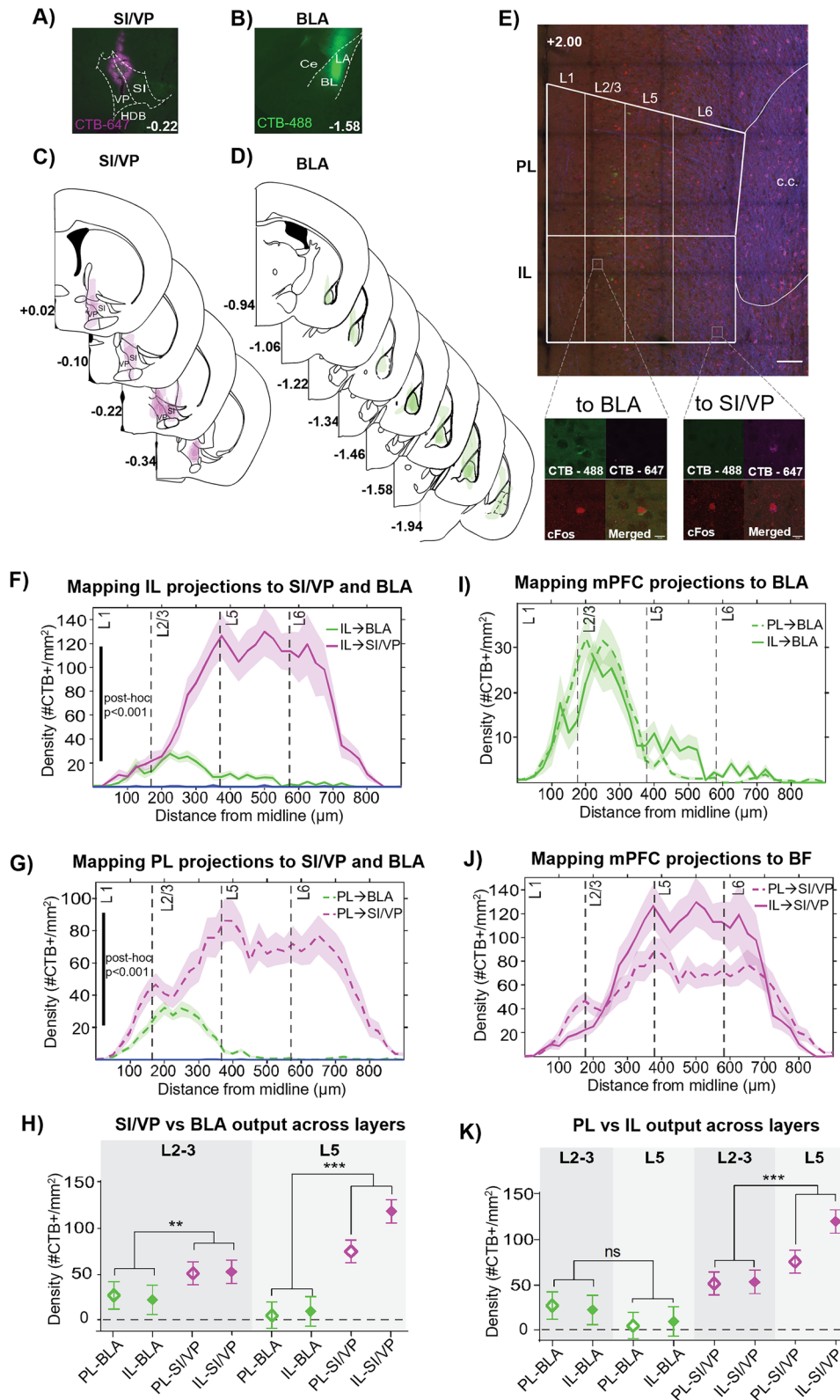
**Figure 1.** Comparative distribution of mPFC projections to the SI/VP and BLA across cortical layers. ***A,B***, Examples of CTB injections in the SI/VP and BLA. ***C,D***, Mapping of the full extent of CTB injections in the SI/VP and BLA. ***E***, Example of CTB and cFos labeling in the mPFC. Scale bar, 100 μm. Insets, Left, Example of CTB-488 labeling an IL→BLA projector, which was also cFos+. Right, An example of CTB-647 labeling of an IL→SI/VP projector, which was also cFos+. Scale bar, 10 μm. ***F***, Comparative density mapping shows a significantly denser IL output to the SI/VP than BLA (post hoc comparison of cortical region × subcortical target; $p < 0.001$). ***G***, Comparative density mapping shows a significantly larger PL output to the SI/VP than BLA (post hoc comparison of cortical region × subcortical target; $p < 0.001$). ***H***, Mixed-model comparing density of SI/VP and BLA output shows denser PL and IL projections to the SI/VP from both L2/3 and L5. ***I***, Comparative density mapping shows that PL and IL projections to the BLA are similarly dense, with BLA projections from both peaking in L2/3 but spreading to deeper layers as well. ***J***, Comparative density mapping showing IL and PL output to the SI/VP. Although L5 IL→SI/VP projections appear to be more numerous than L5 PL→SI/VP projections, the three-way comparison of cortical region × subcortical target × layer was not significant ($F_{(1,109)} = 1.007$; $p > 0.05$). ***K***, Mixed-model comparing densities of PL and IL output across layers. Post hoc comparisons

than to the BLA out of superficial (L2/3, $p = 0.004$) and deep layers (L5, $p < 0.001$; Fig. 1H). However, overall, L5 of the mPFC sends a denser projection to the SI/VP than L2/3 ($p < 0.001$). On the other hand, although the mPFC-to-BLA projection is dense from L2/3, there was enough variability in BLA output throughout the layers, such that there were no differences in mPFC→BLA output between L2/3 and L5 ($p = 0.17$; Fig. 1K, green).

Next, we were interested in whether mice show a denser projection from the IL than the PL to the SI/VP of the BF, as was previously observed in rats (Vertes, 2004). The mixed-model analysis showed a threshold effect of cortical region (IL vs PL; $F_{(1,110)} = 3.92$; $p = 0.05$), suggesting denser overall projections from the IL than PL to subcortical targets (IL, 49.5 cells/mm2 ± 9; PL, 33.2 cells/mm2 ± 9). Note that although L5 IL→SI/VP projections were seemingly more numerous than L5 PL→SI/VP projections (Fig. 1J,K, magenta), the three-way comparison of cortical region × subcortical target × layer was not significant ($F_{(1,109)} = 1.007$; $p > 0.05$).

Finally, we were interested in whether the mPFC→BLA and mPFC→SI/VP pathways collateralize or are independent. Interestingly, the number of SI/VP and BLA coprojectors was very sparse in both the PL and IL, suggesting that mPFC projections to the BLA and the SI/VP are largely noncollateralizing (data not shown). Thus, overall, our anatomical analyses show that (1) the mPFC sends a denser projection to the SI/VP than to the BLA from both superficial and deep layers, (2) the mPFC-to-SI/VP projection is overall denser from deeper than superficial layers, (3) projections to subcortical targets from the IL are slightly denser than from the PL, and (4) mPFC projections to the BLA and SI/VP have minimal collateralization.

### During behavior, overall IL activity is similar across layers, whereas overall PL activity is higher in superficial than deep layers

Having established a differential distribution of mPFC pathways to the SI/VP and the BLA, we were interested in identifying whether these pathways have a similar activity profile during extinction. To do so, the animals that were injected with CTB were divided into three groups, tone control ($n = 9$), extinction learning ($n = 9$), and extinction retrieval ($n = 10$), based on their behavioral condition during cFos capture (Fig. 2A). Mice first underwent handling and habituation (see Materials and Methods) and then were exposed to a 3 d paradigm, where on Day 1 they were either exposed to five 2 kHz tone-alone trials (tone control) or fear conditioned with five paired 2 kHz CS–US trials (extinction learning and retrieval groups). On Day 2, mice were either exposed to 20 2 kHz tone-alone control trials (tone control group), 20 trials of a novel neutral tone (8 kHz) to control for tone exposure (extinction learning group), or 20 2 kHz tone-alone extinction trials (extinction retrieval group). Then, on Day 3 we used cFos to identify mPFC activity during behavior. Animals in the tone control group were exposed to 10 2 kHz tone-alone trials. Mice in the extinction learning group were first exposed to five trials of the same neutral 8 kHz tone as the day before, to control for total tone exposure,

followed by five trials of the fear conditioned 2 kHz CS to capture extinction learning. Mice in the extinction retrieval group were exposed to 10 trials of the previously extinguished 2 kHz CS to capture extinction retrieval, when fear was similarly low to the controls. Animals were perfused 90 min after Trial 6, targeting cFos expression to the start of acquisition in the extinction learning group and compare it to low freezing due to good retrieval in the extinction retrieval group and to low freezing due to the absence of fear learning in the tone controls (Fig. 2A,B).

Fear conditioning (Fig. 2B) resulted in significant differences in freezing across groups ($F_{(2,125)} = 57.99$; $p < 0.0001$) and a trial by group interaction ($F_{(10,125)} = 29.09$; $p < 0.0001$), with the two fear conditioned groups freezing more than the tone control group by the end of the session (Trial 5, control 0% vs extinction learning group 64.1% ± 6; extinction retrieval group 68.7% ± 5.6; both $p < 0.001$). Importantly, the two fear conditioned groups showed similar levels of fear by the end of training (Trial 5, extinction learning 64.1% ± 6 vs retrieval group 68.7% ± 5.7; $p = 0.84$). Then, on Day 2 (Fig. 2B), there were significant differences between groups ($F_{(2,25)} = 7.25$; $p = 0.003$) with a trial × group interaction ($F_{(40,500)} = 3.88$; $p < 0.0001$). Post hoc comparisons showed that the controls did not freeze to the tone, whereas mice that underwent 20 extinction trials significantly decreased freezing during the session (Trial 1 vs 20; $p = 0.02$). Notably, animals that were exposed to 20 trials of a new 8 kHz tone on Day 2 showed some freezing at the beginning of the session, which was likely due to novelty or fear generalization to this novel tone (Trials 1–2 controls vs extinction learning; $p = 0.04$), but this response was gone by Trial 3, when their freezing was the same as controls ($p = 0.13$), staying low for the rest of the session. On Day 3, the tone control group continued to show low freezing throughout the session, whereas the extinction learning group also showed low freezing during the 8 kHz neutral tone presentations (Trials 1–5) but then increased freezing when exposed to the fear conditioned CS starting with Trial 6, reflecting fear retrieval (Fig. 2B, orange), when they froze significantly more than both the control (extinction learning 67.1% ± 9.7 vs tone controls 9.5% ± 3.1; $p = 0.006$) and the extinction retrieval group (extinction learning 67.1% ± 9.7 vs retrieval 21.2% ± 5; $p = 0.014$), whereas the control and extinction retrieval groups showed similarly low freezing by Trial 6 ($p = 0.29$). The extinction learning group then decreased freezing over the next five trials, such that it was no longer different from the extinction retrieval or control groups by Trial 10 ($p > 0.05$ for both). Thus, our cFos timing (Trial 6) aimed to capture the extinction learning group in a relatively higher fear state than both the control and extinction retrieval groups, which were in a similarly low fear state by comparison (Fig. 2B, gray box).

Turning to neural activity, we first asked whether overall, the PL or IL is differentially active during fear extinction learning or when fear is suppressed during retrieval. We first compared the density of cFos+ cells between superficial and deep mPFC layers in all behavioral groups (two-way rmANOVA layer × behavioral group) and found no main effect of the group in the PL or IL (PL, $F_{(2,20)} = 0.304$; $p = 0.74$; IL, $F_{(2,20)} = 0.018$; $p = 0.98$). However,

←

show that there are no differences in PL versus IL outputs to the BLA across layers. However, there is a significantly denser projection from L5 than L2/3 PL and IL to the SI/VP (post hoc, $p < 0.001$). Note that mPFC-BLA projections are illustrated with green (CTB-488), and mPFC→SI/VP projections are illustrated with magenta (CTB-647) for visualization purposes only. During experiments, CTB-488 and CTB-647 injections were counterbalanced between SI/VP and BLA. Main effects, *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$. All data are shown as mean ± SEM. Abbreviations: BL, basolateral nucleus; BLA, basolateral amygdala; Ce, central nucleus; HDB, horizontal limb of the diagonal band of Broca; IL, infralimbic cortex; LA, lateral nucleus; PL, prelimbic cortex; SI, substantia innominata; VP, ventral pallidum.
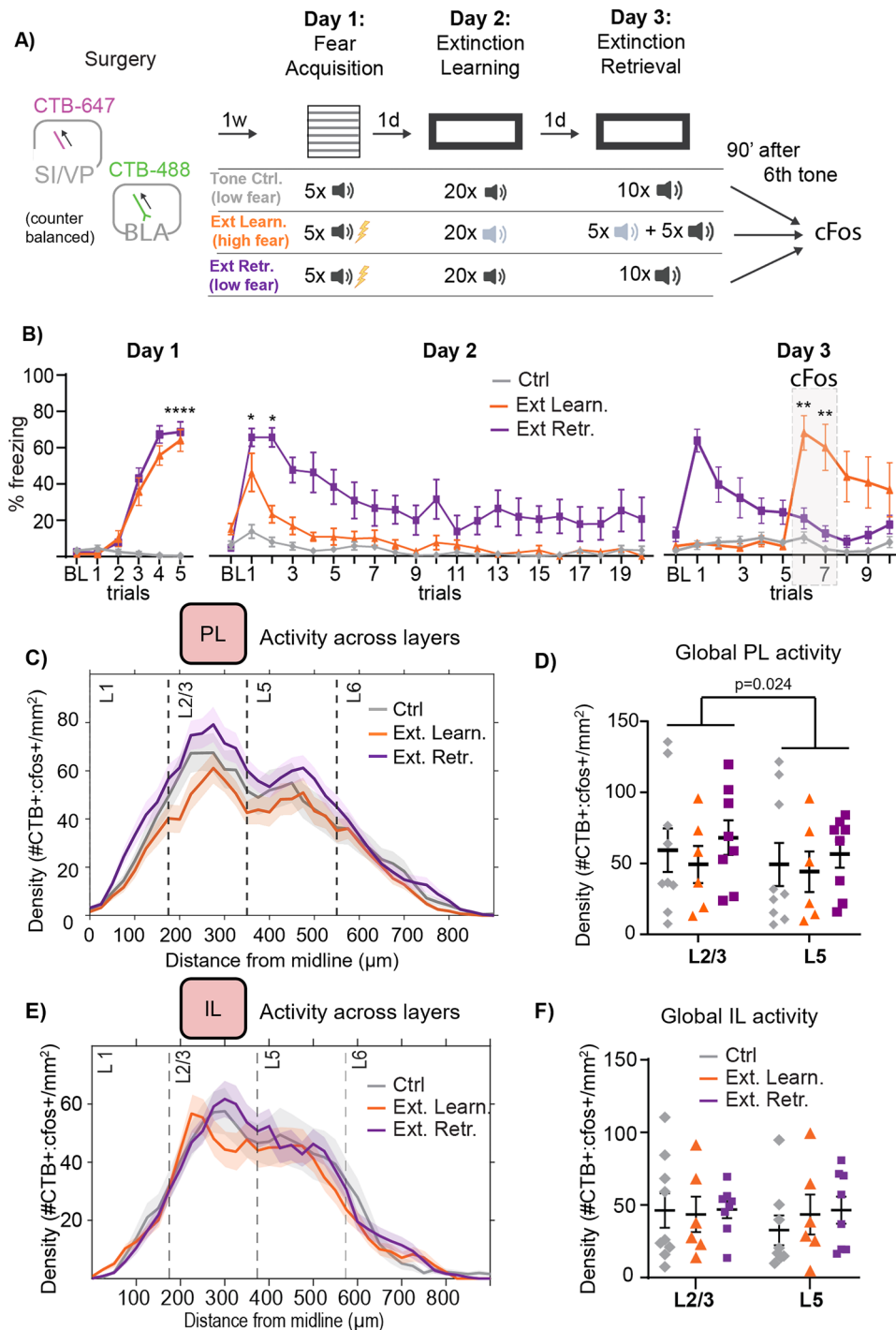
**Figure 2.**    Global PL activity is higher in L2/3 than L5, whereas global IL activity is similar across layers during behavior. ***A***, Timeline of CTB injection surgeries and behavioral paradigm tailored for cFos expression. One week after CTB injections in the VP/SI and BLA, mice were randomly assigned to one of the three groups. The tone control group [Tone Ctrl. (low fear), gray] was exposed to unpaired tones across Days 1–3 and thus was in a low fear state when sampled for cFos. The extinction learning group [Ext Learn. (high fear), orange] was fear conditioned with five CS–US pairings on Day 1. Then, to control for tone exposure, this group received 20 trials of a new, unpaired 8 kHz tone on Day 2 and then, on Day 3, another five trials of the unpaired 8 kHz tone, followed by five trials of the fear-conditioned CS for extinction learning. This group was in a relatively high fear state when sampled for cFos. The extinction retrieval group [Ext Ret. (low fear), purple] was fear conditioned with five CS–US pairings on Day 1, extinguished with 20 CS trials on Day 2 and, on Day 3, underwent extinction retrieval with 10 CS trials. This group was in a relatively low fear state when sampled for cFos. All animals were perfused 90 min after the sixth tone on Day 3. ***B***, Percentage defensive freezing in all groups throughout Days 1–3 of the behavioral paradigm. Day 3, Gray box highlights the trials for timing cFos capture, when the extinction learning group freezing was significantly higher than both in controls and extinction retrieval groups. ***C***, Density mapping of PL cFos+ cells across all layers in all behavioral groups. ***D***, The average number of PL cFos+ cells was higher in L2/3 than L5 for all behavioral groups. ***E***, Density mapping of IL cFos+ cells across all layers in all behavioral groups. ***F***, There were no differences in the average number of IL cFos+ cells across layers in all groups. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$. All data are shown as mean ± SEM.

in the PL, there was an overall main effect of the layer, with higher overall activity in L2/3 than L5 in all behavioral groups ($F_{(1,20)} = 5.945$; $p = 0.024$; Fig. 2C,D), but no main effect of the layer in the IL ($F_{(1,20)} = 0.0004$; $p = 0.984$; Fig. 2E,F). Thus, whereas the PL is overall more active in L2/3 than L5, the IL shows similar levels of activity in superficial and deep layers.
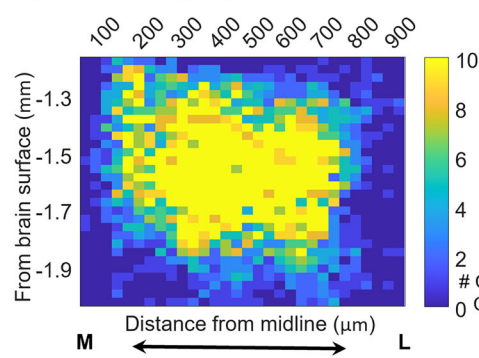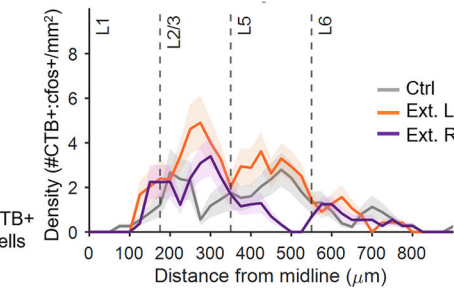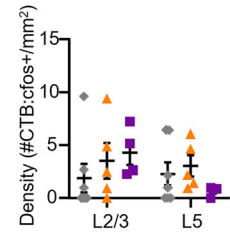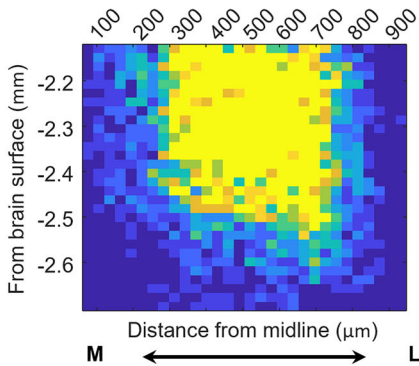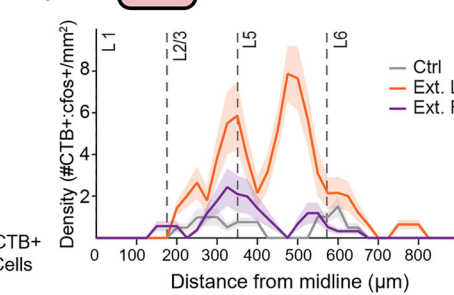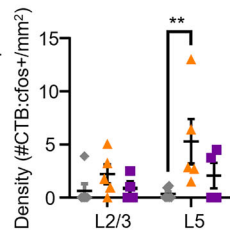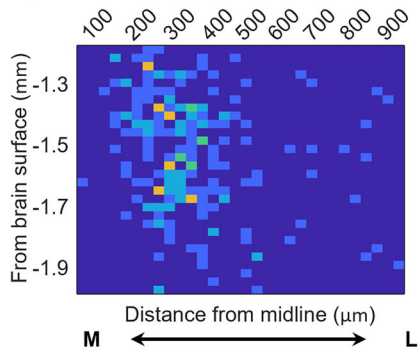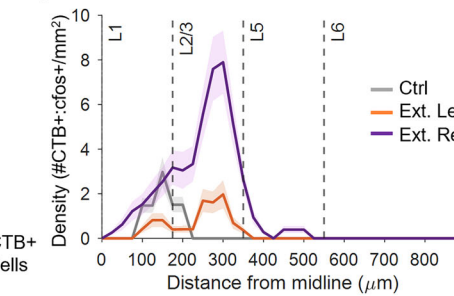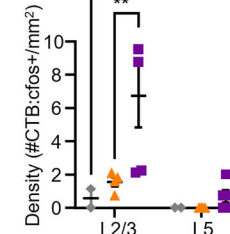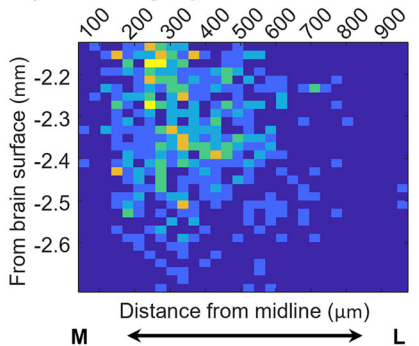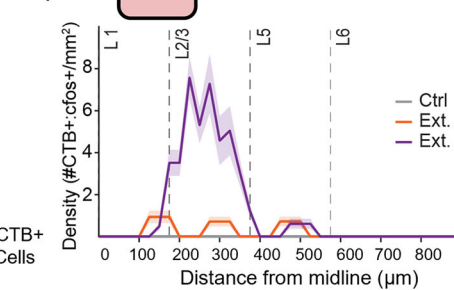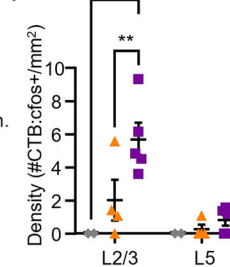
**Figure 3.** The L5 IL→SI/VP pathway is more active during extinction learning, whereas the L2/3 PL→BLA and IL→BLA pathways are more active during extinction retrieval. *A*, Heatmap showing the anatomical distribution of PL→SI/VP projectors along the mediolateral and dorsoventral axes. *B1*, A density map of cFos+ cells in the PL→SI/VP pathway across cortical layers in all behavioral groups. *B2*, Quantification of PL→SI/VP pathway activity (cFos+ cells) in superficial and deep layers in all behavioral groups. The PL→SI/VP pathway does not show any differences in activity between groups. *C*, Heatmap showing the anatomical distribution of IL→SI/VP projectors along the mediolateral and dorsoventral axes. *D1*, A density map of cFos+ cells in the IL→SI/VP pathway across cortical layers in all behavioral groups. *D2*, Quantification of the IL→SI/VP pathway activity in superficial and deep layers in all behavioral groups. There are significantly more active L5 IL→SI/VP projectors in the extinction learning group than in controls. *E*, Heatmap showing the anatomical distribution of PL→BLA projectors along the mediolateral and dorsoventral axes. *F1*, A density map of cFos+ cells in the PL→BLA pathway across cortical layers in all behavioral groups. *F2*, Quantification of the PL→BLA pathway activity in superficial and

## L5 IL projections to the SI/VP are more active during extinction learning; L2/3 PL and IL projections to the BLA are more active during low fear in extinction retrieval

Next, we took advantage of our retrograde tracers in the SI/VP and BLA to identify pathway-specific activity in the PL and IL during extinction. Given the large mPFC projection to the SI/VP peaking in L5 (Fig. 3A), we were interested in whether there was evidence of the PL or IL output to the SI/VP being active at any phase of extinction. A group × layer rmANOVA showed no significant differences in the activity across layers of PL→SI/VP pathway during different phases of behavior ($F_{(2,12)} = 2.73$; $p = 0.11$; Fig. 3B1,2). However, the same analysis in the IL→SI/VP pathway (Fig. 3C) showed a significant effect of behavior ($F_{(2,12)} = 4.98$; $p = 0.02$; Fig. 3D1,2), with Tukey's multiple comparisons revealing that the density of active L5 IL→SI/VP projectors was significantly higher in the extinction learning than the tone control group ($p < 0.01$), whereas the extinction retrieval group did not differ in activity from tone controls (Fig. 3D1,2; $p = 0.62$). This increase in IL→SI/VP activity was only observed in L5, without any significant differences in L2/3 IL→SI/VP activity across behavioral groups (all $p > 0.05$; Fig. 3D1,2). These findings indicate that L5 IL-SI/VP projectors increase activity during fear retrieval and extinction learning compared with tone-alone controls, whereas the activity of L5 PL→SI/VP projectors does not.

Turning to the mPFC→BLA pathway, the overall anatomy showed that mPFC output to the BLA peaked in L2/3 but also distributed to deeper layers and was overall less dense than the projection to the SI/VP of the BF (Figs. 1F–J, 3E,G). To investigate the behavior-dependent activity in PL and IL output to the BLA, we used a two-way rmANOVA layer × group analysis and found that both the PL→BLA and the IL→BLA pathways showed a significant main effect of layer (PL, $F_{(1,8)} = 9.91$; $p = 0.01$; IL, $F_{(1,7)} = 12.05$; $p = 0.01$), behavioral group (PL, $F_{(2,8)} = 4.52$; $p = 0.049$; IL, $F_{(2,8)} = 5.95$; $p = 0.027$), and a significant layer by group interaction in the PL→BLA pathway ($F_{(2,8)} = 4.54$; $p = 0.048$) and a borderline significant layer by group interaction in the IL→BLA pathway ($F_{(2,7)} = 4.58$; $p = 0.054$). Interestingly, Tukey's post hoc tests revealed that the density of active L2/3 PL→BLA and L2/3 IL→BLA projectors was higher in the extinction retrieval group compared with tone controls (PL, $p = 0.01$; IL, $p < 0.01$) and compared with the extinction learning group (PL, $p < 0.01$; Fig. 3E,F1,2; IL, $p = 0.01$; Fig. 3G,H1,2). However, activity in L5 PL→BLA and L5 IL→BLA projectors were equally low in both groups ($p > 0.05$; Fig. 3F,H), indicating that the deep layer mPFC projections to the BLA are less involved in extinction retrieval when fear is low. These findings are in line with previous work showing that mPFC output to the BLA is active during extinction retrieval and fear discrimination (Bukalo et al., 2015; Bloodgood et al., 2018; Stujenske et al., 2022).

## The IL→SI/VP pathway gains excitability with extinction learning and becomes less excitable at extinction retrieval

Given our finding that L5 IL output to the SI/VP was more active during a period that conflated fear retrieval and fear extinction learning, we were interested in pinpointing which of these processes involved this pathway more. Despite timing perfusions for cFos assessment to the beginning of extinction learning in our experiments (Figs. 2, 3), cFos transcription has a relatively low temporal resolution, making it difficult to know whether this pathway is more likely to be excited early on in extinction learning, which is also driven by fear retrieval, or later in extinction acquisition. To address this question in more detail, we injected the mice with the retrograde virus AAV2-hSyn-eYFP (UNC Vector Core) in the SI/VP (Fig. 4A), and then we obtained in vitro patch recordings from identified IL→SI/VP projectors at RMP and when the membrane was held at −70 mV, in mice that were either exposed to the tone alone, after two trials of extinction, when fear retrieval is high, late in extinction learning which has less fear retrieval, or in extinction retrieval (Fig. 4A). To do so, after 2 weeks of viral expression, mice were handled and habituated to the training and extinction contexts, followed by five trials of tone-alone exposure (tone controls) or five trials of paired CS–US fear conditioning in the training Context A. The next day, in vitro recordings were obtained from IL→SI/VP projecting cells located in deeper IL layers, either after two 2 kHz tone-alone trials (tone controls, $n = 14$ cells; $n = 3$ mice; Fig. 4C, "gray arrow"), two 2 kHz tone-alone extinction learning trials (early extinction, $n = 26$ cells; $n = 5$ mice; Fig. 4C, "orange arrow"), or 20 2 kHz tone-alone extinction learning trials (late extinction, $n = 16$ cells; $n = 3$ mice; Fig. 4C, "red arrow"). A fourth group of animals went through 20 2 kHz tone-alone trials of extinction learning, and then the next day was exposed to 2 2 kHz tone-alone extinction retrieval trials prior to in vitro patch recordings (extinction retrieval, $n = 28$ cells; $n = 5$ mice; Fig. 4C, "purple arrow").

Freezing behavior on conditioning day was analyzed with a rmANOVA that tested freezing over trials in the four groups. This analysis showed a main effect of trial ($F_{(3.5,42.2)} = 46.06$; $p < 0.0001$), group ($F_{(3,12)} = 11.19$; $p < 0.001$), and a trial by group interaction ($F_{(15,60)} = 5.76$; $p < 0.001$). Subsequent multiple comparisons showed that by trial 5 of fear conditioning, the fear conditioned groups had higher freezing than the tone control group ($p < 0.05$ for each conditioned group; Fig. 4C), whereas the fear conditioned groups had similar amounts of defensive freezing ($p > 0.05$), indicating that they had learned the CS–US association similarly.

On extinction learning day, there was a significant effect of the group (one-way ANOVA; $F_{(3,12)} = 4.78$; $p = 0.02$) and trial ($F_{(3.2,24.2)} = 15.2$; $p < 0.0001$), indicating that animals learned to extinguish fear. A one-way ANOVA evaluating defensive freezing in all groups during the two trials preceding in vitro recordings showed a significant effect of the group ($p = 0.002$), with post hoc comparisons showing that the early extinction group froze significantly more during the first two trials of extinction learning/fear retrieval than the tone control and late extinction groups ($p < 0.01$; Fig. 4C,D). Patch recordings of IL→SI/VP projector excitability at RMP showed a significant effect of the group ($F_{(3,80)} = 3$; $p = 0.03$), current step ($F_{(3,256)} = 351$; $p < 0.0001$), and a current step × group interaction ($F_{(60,1600)} = 1.96$; $p < 0.0001$). Interestingly, post hoc comparisons showed that despite mice being in high versus low fear states for the early extinction versus tone control groups (Fig. 4D), respectively, the excitability

---

deep layers in all behavioral groups. The L2/3 PL→BLA pathway is more active during extinction retrieval than control and extinction learning groups. **G,** Heatmap showing the anatomical distribution of IL→BLA projectors along the mediolateral and dorsoventral axes. **H1,** A density map of cFos+ cells in the IL→BLA pathway across cortical layers in all behavioral groups. **H2,** Quantification of the IL→BLA pathway activity in superficial and deep layers in all behavioral groups. The L2/3 IL→BLA pathway is more active during extinction retrieval than control and extinction learning groups. Mean and SEM are shown throughout. The color bar indicates the average number of CTB+ cells. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.
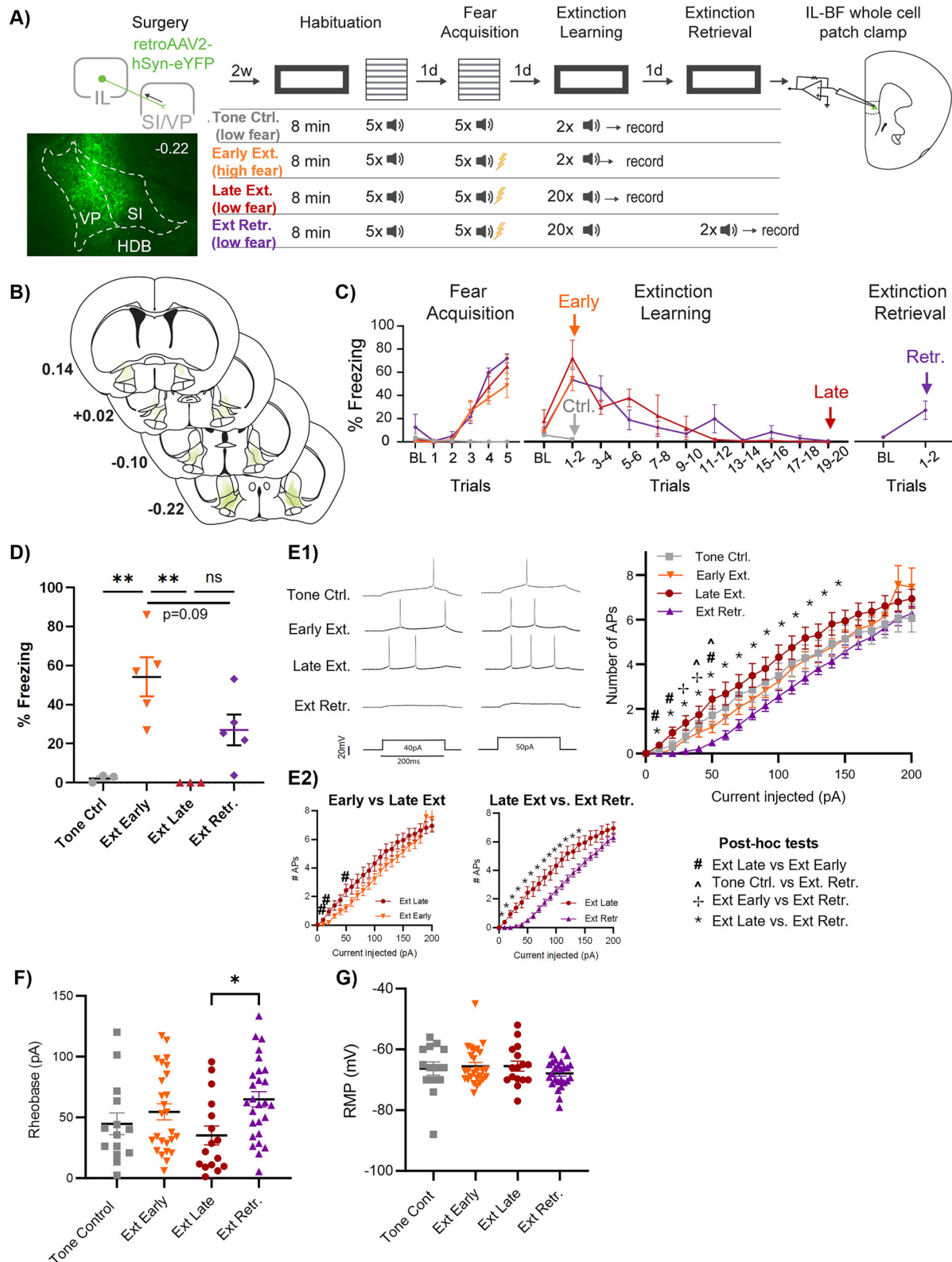
**Figure 4.** Deep layer IL→SI/VP projectors are more excitable during extinction learning than retrieval. ***A***, Timeline of injection surgeries and behavioral paradigm prior to in vitro recordings of IL projectors to the SI/VP. Example injection of rgAAV2-hSyn-eYFP in the SI/VP to identify IL→SI/VP projectors. After 2 weeks of expression, animals were randomly assigned to one of the four groups to probe IL→SI/VP projector excitability during early extinction when fear was high [Early Ext. (high fear), orange], late extinction when fear was low [Late Ext. (low fear), red], extinction retrieval [Ext. Retr. (low fear), purple], or tone control when fear was low, and no learning had occurred [Tone Ctrl. (low fear), gray]. Mice were killed for in vitro recordings 10 min after the last behavioral trial. ***B***, Anatomical mapping of all viral injections in the SI/VP. ***C***, The percentage defensive freezing for all behavioral groups. Vertical arrows mark the last two trials of behavior for each group, after which the mPFC was sliced for in vitro recordings. ***D***, Average percentage freezing for each group in the last two trials before perfusion. ***E1***, Left, Example traces showing IL→SI/VP projector excitability in each group at 40 and 50 pA current injection steps. Right, Excitability curves at RMP shown by the number of APs in response to increasing injections of current for each behavioral group. Significant post hoc tests are marked with their respective symbols shown in the key. ***E2***, Replotting of excitability curves in subsets of groups for clarity. All significance

of IL→SI/VP cells at RMP in these groups was not different (Fig. 4E; tone controls, $n = 14$ cells; $N = 3$ mice, gray; early extinction, $n = 26$ cells; $N = 5$ mice, orange; $p > 0.05$), indicating that early in extinction learning, when fear retrieval is high, the IL→SI/VP pathway did not differ in excitability from the tone-alone control condition. Later in extinction learning, freezing for the late extinction group decreased, such that by Trials 19–20, defensive freezing was significantly lower than during Trials 1–2 ($p < 0.001$) and was lower than Trials 1–2 for the early extinction group ($p < 0.01$; Fig. 4C,D). Interestingly, Tukey's multiple comparisons revealed that in the late extinction group, IL→SI/VP excitability at RMP increased relative to the early extinction group at several steps in the more physiological range (Fig. 4E1,2; 10 mA $p = 0.08$; 30 mA $p = 0.05$; 50 mA; $p = 0.08$).

The next day, the extinction retrieval group showed a trend toward decreased freezing from early extinction ($p = 0.09$) and no difference in freezing from the late extinction group ($p > 0.05$; Fig. 4D). Interestingly, during extinction retrieval, IL→SI/VP excitability at RMP was significantly lower than in the late extinction group (Fig. 4E1,2; 10–140 mA pulses; all $p < 0.05$). IL→SI/VP excitability during extinction retrieval was also lower than during early extinction and lower than tone control at several current-level steps in the lower, more physiological stimulation range (Fig. 4E1, 10–50 mA pulses). We then tested IL→SI/VP projector rheobase or the current needed to drive a cell to spike at RMP, in all behavioral groups. This analysis showed that rheobase was significantly different between behavioral groups (Kruskal–Wallis, $p < 0.05$), with multiple comparisons revealing that less current was needed to drive an IL→SI/VP cell to fire in the late extinction group than in the extinction retrieval group ($p < 0.05$; Fig. 4F). These findings confirm that the IL→SI/VP pathway is more excitable during extinction learning and then becomes less excitable during extinction retrieval, when fear is already suppressed. Notably, changes in excitability were not accompanied by any change in RMP (Fig. 4G; one-way ANOVA; $F_{(3,78)} = 0.81$; $p = 0.49$), and there were no changes in excitability observed when IL→SI/VP projectors were held at −70 mV (data not shown), indicating that IL→SI/VP projectors are likely to be more synaptically driven by the network during extinction learning rather than retrieval, without intrinsically changing these cells.

## IL→SI/VP but not PL→SI/VP pathway constrains freezing during extinction learning

Given that during extinction learning, L5 IL→SI/VP neurons upregulate their activity and become more excitable, we next wanted to know if this pathway functionally contributes to extinction. To this end, we injected the inhibitory opsin AAV$_5$-hsyn-eArch3.0-eYFP ($n = 9$) or its control AAV$_5$-hsyn-eYFP ($n = 9$) in the IL and bilaterally implanted optic fibers over the SI/VP (Fig. 5A,B). After ~4 weeks of viral expression, animals were habituated to tone CS in the fear conditioning context (Context A) and to 35 s exposure of 532 nm laser in the extinction context (Context B, 3 s on and 2 s off ramps; Mahn et al., 2016), mice then underwent tone fear conditioning in Context A 24 h later. The next day, animals underwent extinction learning with inhibition of IL terminals in the SI/VP

during each CS in Context B. The next day, mice underwent a second extinction session in the absence of laser in Context B (Fig. 5C).To check whether light at IL terminals affected neural activity (Mahn et al., 2016), we recorded SI/VP single units ($n = 3$) and multiunit activity (MUA; $n = 6$) in two mice ($n = 1$ eYFP and $n = 1$ eArch) during light habituation trials in Context B (Fig. 5D–H). Comparisons of firing rates before and after light onset (Fig. 5E,F) showed no change in cell firing. Single units with different baseline firing rates showed no effect of light on their firing rates compared with the 35 s preceding light onset (paired $t$ tests; $p > 0.05$; Fig. 5G). Furthermore, the percentage change in the overall firing rate with light onset was not different between the eYFP and eArch animals (unpaired $t$ test; $p > 0.05$; Fig. 5H).

Mice in both groups did not freeze during the habituation to the tone CS (Context A; eYFP 2.8%; eArch 2.2%; two-way rmANOVA; trial × group; $F_{(5,90)} = 0.45$; $p = 0.81$) or during habituation to the laser (Context B, eYFP, 0.5%; eArch, 1.1%; two-way rmANOVA; trial × group; $F_{(5,90)} = 1.25$; $p = 0.29$; Fig. 5I). The next day, during fear conditioning, both groups acquired the CS–US association similarly (Fig. 5J; two-way rmANOVA; trial; $F_{(2.72, 43.05)} = 69.46$; $p < 0.0001$; group $F_{(1,16)} = 0.64$; $p = 0.44$; trial × group; $F_{(5,79)} = 0.75$; $p = 0.59$), without any differences in average freezing during the conditioning session (Fig. 5K; unpaired $t$ test; $p > 0.05$). Then, during an extinction learning session, the laser was administered during each CS, which resulted in a significant difference in tone-evoked freezing between groups (Fig. 5J; two-way rmANOVA; $F_{(1,16)} = 6.45$; $p = 0.02$) as well as a group by trial interaction (Fig. 5J; $F_{(5,80)} = 3.99$; $p = 0.003$). Average freezing during the session was higher in the eArch than in the eYFP group (Fig. 5L, left; $t$ test; $p = 0.019$). A comparison of freezing at the beginning versus the end of the session (Fig. 5L, right) using a two-way rmANOVA of the group by trial freezing showed a significant effect of the group ($F_{(1,16)} = 5.6$; $p = 0.03$), trial ($F_{(1.75, 28)} = 38.24$; $p < 0.0001$), and a group × trial interaction ($F_{(2,32)} = 4.96$; $p = 0.01$), with the eArch group showing significantly higher freezing than eYFP controls at the beginning (Trials 1–2, eArch, 74.8%; eYFP, 50%) and at the end of extinction learning (Trials 9–10, eArch, 44%; eYFP, 15%). When we took all trials into account, Sidak's multiple comparisons indicated that a significant difference in freezing emerged on Trials 3–4 (Fig. 5J, extinction learning). Thus, activity in the IL→SI/VP pathway constrains the defensive freezing response during extinction learning. Interestingly, when tested without any light the next day during Extinction 2, probing a mixture of extinction retrieval and re-extinction, there were no differences in freezing between groups [Fig. 5J, Extinction 2; group; $F_{(1,16)} = 0.04$; $p > 0.84$], indicating that the IL→SI/VP pathway is important for constraining the expression of defensive freezing during active fear decrement and does not affect memory or behavior the next day. Notably, turning the light on during extinction (before the CS started) did not alter single unit activity ($n = 3$), but when the CS and light were on, there was an overall change in firing from prelight (Extended Data Fig. 5-1), indicating that inhibition of the IL→SI/VP pathway during extinction learning affected SI/VP activity.

To test whether this was a pathway-specific effect, we repeated this experiment in a small cohort of mice, this time inhibiting PL

←

testing was done on four groups. Left, Early versus late extinction excitability curves; post hoc tests that reached significance are marked with #, indicating increased excitability in late versus early extinction in IL→SI/VP projectors. Right, IL→SI/VP projectors in late extinction are significantly more excitable than in extinction retrieval in a wide range of testing conditions. *F*, Rheobase (pA), or the lowest level of current to evoke an AP, across groups. IL→SI/VP projectors show significantly lower rheobase in late extinction than extinction retrieval. *G*, IL→SI/VP RMP (mV) shows no difference across groups.
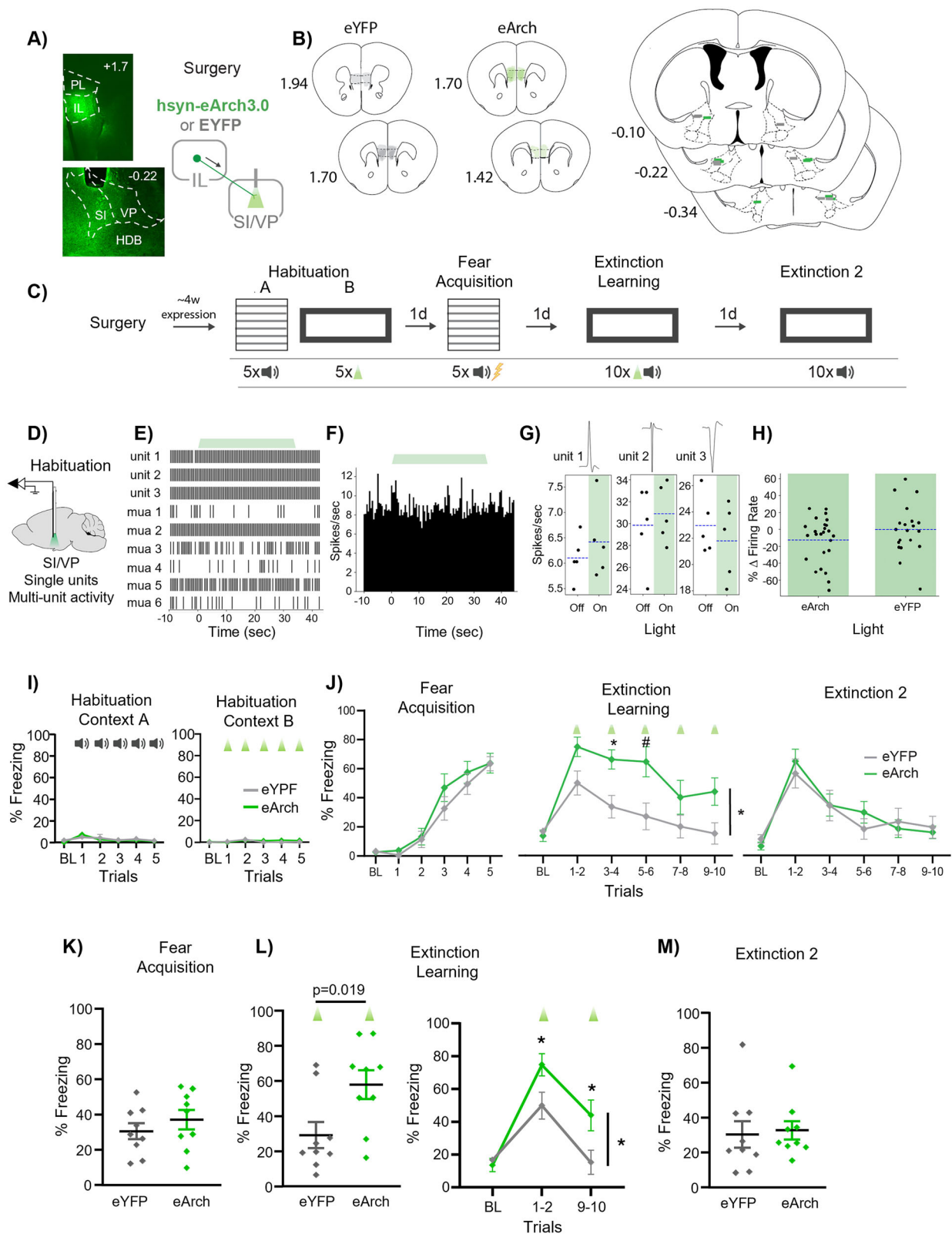
**Figure 5.** IL projections to the SI/VP constrain defensive freezing during extinction learning. **A**, Example of viral injection in the IL and fiber placement in the SI/VP and schematic showing the injections and fiber placement surgeries. **B**, Mapping of the viral spread of the virus in the IL (left) and fiber placements in the SI/VP (right). Gray, eYFP; green, eArch. **C**, Schematic of the behavioral paradigm. Mice were first habituated to the tone in fear conditioning Context A and to the laser in extinction Context B. The next day, mice were fear conditioned with five CS–US pairings. The next day, mice underwent fear extinction for 10 trials with laser light delivery during the tone, inhibiting IL inputs to the VP/SI. One day later, mice were re-exposed to the extinction context during a second 10-trial session of extinction, testing a mixture of extinction retrieval and re-extinction. **D**, Virus validation: recording setup during optogenetic manipulation. After IL injections of the AAV5-hsyn-eArch3.0-eYFP (*n* = 1) or the AAV5-hsyn-eYFP (*n* = 1) virus, an optrode, consisting of the optic fiber and stereotrode bundle, was placed in the SI/VP to record single units and MUA during a laser light-habituation session in Context B. Additional recordings showing the effects of IL→SI/VP inhibition on cell activity during extinction learning are shown in Extended Data Figure 5-1. **E**, A raster plot showing an example of single units and MUA firing for 10 s before light onset, during a 35 s light-on trial (green bar, ramp-on), and for 10 s after light

terminals in the SI/VP during extinction learning (Extended Data Fig. 5-2A–C). As before, exposure to tones in Context A or light during habituation in Context B did not affect defensive freezing in either group (Extended Data Fig. 5-2D; tone in Context A, eYFP 1.3%; eArch 1.1%; two-way rmANOVA; trial × group; $F_{(5,20)} = 1.5$; $p = 0.23$; light in Context B, eYFP, 1.5%; eArch, 2.6%; two-way rmANOVA; trial × group; $F_{(5,20)} = 1.2$; $p = 0.33$). During fear conditioning, both groups acquired the defensive freezing response similarly (Extended Data Fig. 5-2E; two-way rmANOVA; trial, $F_{(1.55,6.2)} = 39.2$; $p < 0.001$; group $F_{(1,4)} = 0.38$; $p = 0.57$; trial × group; $F_{(5,20)} = 0.81$; $p = 0.55$) and showed the same levels of average freezing during the conditioning session (Extended Data Fig. 5-2F; unpaired $t$ test; $p > 0.05$). During extinction learning, inhibition of PL terminals in the SI/VP had no effect on behavior, with both groups displaying similar decrement in defensive freezing as the session progressed (Extended Data Fig. 5-2E; two-way rmANOVA; trial, $F_{(1.7,6.6)} = 6.2$; $p = 0.03$; group, $F_{(1,4)} = 0.26$; $p = 0.64$; trial × group; $F_{(5,20)} = 0.26$; $p = 0.93$) and no group differences in average freezing during the extinction learning session (Extended Data Fig. 5-2G; unpaired $t$ test; $p > 0.05$). Likewise, during Extinction 2 the next day, both groups showed similar levels of memory retrieval and re-extinction across the session (Extended Data Fig. 5-2E; two-way rmANOVA; trial, $F_{(1.72,6.9)} = 9.3$; $p = 0.01$; group, $F_{(1,4)} = 0.16$; $p = 0.71$; trial × group, $F_{(5,20)} = 0.29$; $p = 0.91$), without any differences in overall freezing (Extended Data Fig. 5-2H; unpaired $t$ test; $p > 0.05$).

Taken together, these findings indicate that during extinction learning, IL inputs to the SI/VP constrain the freezing response as animals learn, whereas PL inputs to the SI/VP do not.

## IL inputs to the SI/VP constrain freezing during active fear decrement

The shorter, 10-trial extinction session used in the optogenetic experiments resulted in good within-session extinction, but the next day, both groups showed relatively high freezing at the beginning of Extinction 2, which quickly decreases to baseline levels (Fig. 5J; Extended Data Fig. 5-2E), suggesting re-extinction. To test whether IL→SI/VP pathway affects re-extinction, we used optogenetic inhibition of IL-SI/VP terminals during Extinction 2 (Fig. 6A,B). As previously, animals were injected with the inhibitory opsin AAV5-hsyn-eArch3.0-eYFP ($n = 7$) or its control virus, AAV5-hsyn-eYFP ($n = 5$), in the IL and received bilateral implantations of optic fibers over the SI/VP (Fig. 6A). They then underwent behavior as previously described, but with IL→SI/VP terminal inhibition occurring only during tone presentations on Extinction 2 (Fig. 6B). During habituation to tone in Context A or light in Context B, there was little defensive freezing in both groups (Fig. 6C; tone in Context A, eYFP, 1.9%; eArch, 2.7%; two-way rmANOVA; trial × group; $F_{(5,50)} = 0.31$; $p = 0.91$; light in Context B, eYFP, 4%; eArch, 2.7%; two-way rmANOVA; trial × group; $F_{(5,50)} = 1.74$; $p = 0.14$). During fear acquisition, animals in both groups showed similar increases in defensive freezing (Fig. 6D; two-

way rmANOVA; group, $F_{(1,10)} = 0.13$; $p = 0.73$; trial, $F_{(2.23, 22.91)} = 73.31$; $p < 0.0001$; group × trial, $F_{(5,50)} = 0.11$; $p = 0.99$), and there were no differences in average freezing during the fear acquisition session (Fig. 6E; $t$ test; $p = 0.71$). Both groups also showed similar decrements in freezing during extinction learning (Fig. 6D; two-way rmANOVA; group, $F_{(1,10)} = 0.07$; $p = 0.80$; trial, $F_{(3,30.05)} = 34.52$; $p < 0.0001$; group × trial, $F_{(5,50)} = 0.94$; $p = 0.46$), and no differences in average freezing during extinction learning (Fig. 6F; $t$ test; $p = 0.6$). Interestingly, IL→SI/VP inhibition during Extinction 2 did not impact the decrement in freezing seen across trials (Fig. 6D; two-way rmANOVA; group, $F_{(1,10)} = 1.38$; $p = 0.27$; trial, $F_{(2.94, 29.42)} = 16.26$; $p < 0.0001$; group × trial, $F_{(5,50)} = 1.76$; $p = 0.14$), and average freezing during Extinction 2 was the same in both groups (Fig. 6G, left; $t$ test; $p = 0.21$), suggesting no overall effects of IL→SI/VP pathway inhibition during this session. However, a planned two-way rmANOVA (as in Fig. 5L) comparing the first two and last two trials showed a group by trial interaction (Fig. 6G, right; group, $F_{(1,10)} = 0.67$; $p = 0.43$; trial, $F_{(2,20)} = 44.08$; $p < 0.0001$; group × trial, $F_{(2,20)} = 3.64$; $p = 0.045$), with post hoc comparisons indicating significantly higher freezing in the eArch than eYFP controls at the beginning of Extinction 2 ($p = 0.027$) and no differences between groups by the end of Extinction 2 ($p = 0.61$).

These findings suggest that the IL→SI/VP pathway during Extinction 2 constrained freezing only during the early trials of Extinction 2, when animals underwent re-extinction, without an effect on behavior by the end of the session. However, given that the control group in this experiment showed good extinction retrieval (paired $t$ test; eYFP extinction learning freezing during Trials 1–2, 78.3 ± 9.4%, vs eYFP Extinction 2 freezing during Trials 1–2, 49.7 ± 10.6%; $p < 0.01$), the most parsimonious interpretation of these data is that the IL→SI/VP pathway helps constrain freezing during active fear decrement that occurred during the early trials of this session, which is likely to be driven by a mix of extinction retrieval and re-extinction.

## Discussion

The IL is a critical region for extinction (Milad et al., 2005; Giustino and Maren, 2015; Kim et al., 2016), prompting us to investigate whether the IL→SI/VP pathway (Room et al., 1985; Hurley et al., 1991; Zaborszky et al., 1997; Vertes, 2004) plays a role in this behavior. We show that IL→ and PL→SI/VP pathways originate from L2/3 and L5, but only L5 IL→SI/VP output increases activity during extinction learning, whereas L2/3 IL→BLA output increases activity during extinction retrieval. We also demonstrate that IL→SI/VP projectors are more excitable during extinction learning than retrieval. Furthermore, we demonstrate that inhibiting the IL→ but not the PL→SI/VP pathway during extinction learning increases defensive freezing, without affecting memory. Collectively, we show that during extinction learning, the IL→SI/VP pathway increases activity and constrains expression of defensive freezing.

←

offset (ramp-off). *F*, A peristimulus time histogram showing the average firing rate of all units shown individually in panel *E*. *G*, An average firing rate (spikes/sec) of three single units during five habituation trials during light-off (white) and light-on (green) periods. The average waveform of each unit is shown. Unit 1 has a firing rate of 6.1 spikes/sec, whereas Units 2 and 3 have higher firing rates (30 and 23 spikes/sec, respectively). None of the units change their firing rates during light-on periods. *H*, Average percentage change in firing (single units and MUA) during the light shows no change. Blue line, mean. *I*, Average percentage defensive freezing on each trial during habituation to tone presentations in Context A and optogenetic light stimulation in Context B in both groups. *J*, Average percentage defensive freezing on each trial in eYFP and eArch groups across fear conditioning, extinction learning, and Extinction 2. *K*, Average percentage defensive freezing during the fear acquisition session is similar in both groups. *L*, During extinction learning, left, average percentage defensive freezing is higher in the eArch than the eYFP group (unpaired $t$ test; $p = 0.019$). Right, During extinction learning, average defensive freezing in Trials 1–2 and Trials 9–10 is higher in the eArch than eYFP group. *M*, Average percentage freezing during Extinction 2 is similar in both groups. All data, unless otherwise specified, are shown as mean ± SEM; *$p < 0.05$; #$p < 0.07$. In Extended Data Figure 5-2, we show that inhibition of the PL→SI/VP pathway during extinction learning does not affect behavior. Abbreviations: IL, infralimbic; MUA, multiunit activity; SI, substantia innominata; VP, ventral pallidum.
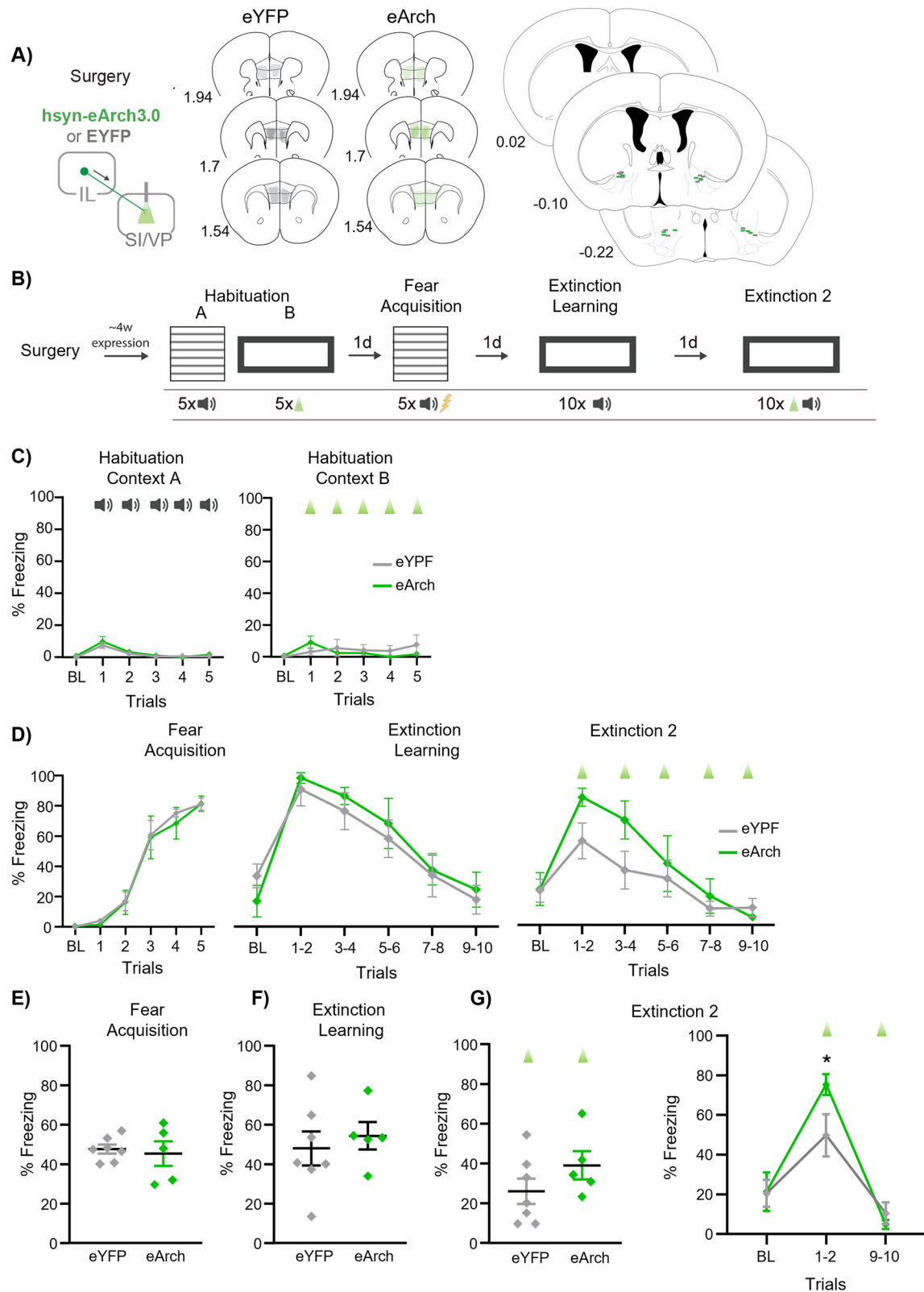
**Figure 6.** IL inputs to the SI/VP constrain freezing during active fear decrement during Extinction 2. *A*, Schematic showing the IL viral injections and SI/VP fiber placements and the mapping of the viral spread of the virus in the IL (left) and fiber placements in the SI/VP (right). Gray, eYFP; green, eArch. *B*, Schematic of behavioral paradigm. Mice were first habituated to the fear conditioning Context A and the tone, as well as to the extinction Context B and the laser. The next day, mice were fear conditioned with five CS–US pairings in Context A. The next day, mice underwent fear extinction learning for 10 trials in Context B. One day later, during Extinction 2, mice were re-exposed to the extinction Context B during a second 10-trial session of extinction with laser light delivery during the tones, inhibiting IL inputs to the SI/VP. *C*, Average percentage defensive freezing on each trial during habituation to tone presentations in Context A and optogenetic light stimulation in Context B in both groups. *D*, Average percentage defensive freezing on each trial in eYFP and eArch groups across fear acquisition, extinction learning, and Extinction 2. There were no significant differences in behavior when taking into account all trials in all conditions. *E*, Average percentage

The IL projects to cortical, subcortical, and brainstem targets, which collectively influence cognitive, motor, and cardiovascular responses (Room et al., 1985; Hurley et al., 1991; Vertes, 2004). Interestingly, optogenetic inactivation of the IL during extinction was reported not to impair extinction memory (Do-Monte et al., 2015), suggesting that the IL facilitates extinction via its target regions, such as the BLA (Quirk et al., 2003; Amano et al., 2010; Cho et al., 2013; Strobel et al., 2015). Notably, inhibiting the IL→BLA pathway during learning does not impair within-session extinction but instead dampens extinction retrieval (Bukalo et al., 2015; Bloodgood et al., 2018), indicating that IL inputs drive longer-term plasticity in the amygdala. Accordingly, IL in vitro recordings, including IL→BLA projectors, show that they become more excitable after extinction (Santini et al., 2008; Cruz et al., 2014; Bloodgood et al., 2018). Likewise, recent work has uncovered the paraventricular thalamus (PVT) as another important IL target for mediating extinction retrieval (Tao et al., 2021). Taken together, these findings indicate that cellular and molecular consolidation occur in IL→BLA projectors postextinction learning (Santini et al., 2004; Burgos-Robles et al., 2007) and shape BLA processing during retrieval (Royer and Pare, 2002; Quirk et al., 2003; Likhtik et al., 2008; Amano et al., 2010; Amir et al., 2011; Cho et al., 2013; Strobel et al., 2015; Hagihara et al., 2021).

Less is known about the role of IL output to the SI/VP, which processes threatening and rewarding stimuli (Lin and Nicolelis, 2008; Stephenson-Jones et al., 2020; Moaddab et al., 2021; Hegedüs et al., 2024), prompting us to use cFos to assess its activity during extinction learning and retrieval. This approach showed that L5 IL→SI/VP output was upregulated during extinction learning more than in control conditions. Conversely, as shown previously (Bukalo et al., 2015; Bloodgood et al., 2018), the L2/3 IL→BLA output was upregulated during extinction retrieval. Interestingly, in the extinction retrieval group, activity also increased in the L2/3 PL→BLA pathway, suggesting that PL communication with the BLA may be active during fear and extinction (Marek et al., 2018; Watanabe et al., 2021). One important caveat to our approach is the low temporal resolution of cFos expression, whereby the extinction learning group was timed to the beginning of extinction, conflating fear retrieval and extinction learning. For extinction retrieval, cFos was timed for equally low freezing in the retrieval and control groups to isolate the effects of learning, thereby conflating retrieval and re-extinction. Adding another behavioral readout to this experiment, such as lever-pressing for food, could help better parse these phases. However, overall, this experiment showed that the L5 IL→SI/VP pathway is upregulated during active fear decrement, whereas the L2/3 IL→BLA pathway is upregulated when fear is suppressed.

For a more granular temporal analysis, we used in vitro recordings of IL→SI/VP projector excitability, timed to different trials of extinction. We show that IL→SI/VP projectors become more excitable later in extinction learning, a period less contaminated by fear retrieval. Furthermore, after two trials of extinction retrieval, IL→SI/VP projectors were less excitable than in late extinction, confirming that the IL→SI/VP pathway is more active during extinction learning than retrieval. Notably, IL excitability increased without changes in the RMP, suggesting that it is due to
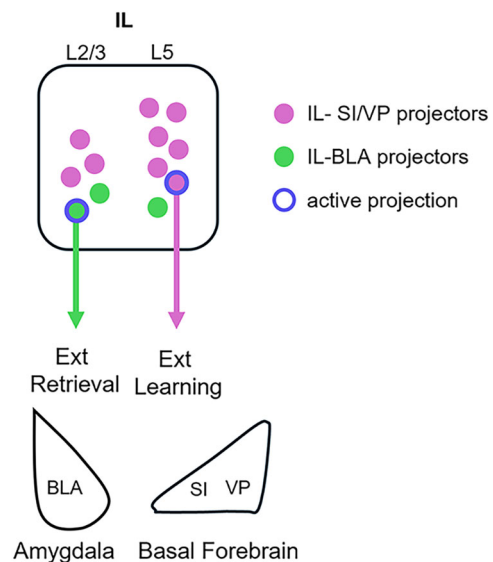


**Figure 7.** A summary model. Schematic of the neuroanatomical distribution of IL projectors to the BLA (green circles) and to the SI/VP (pink circles). L5 IL→SI/VP projectors are more active (blue rim) during extinction learning, when fear is high and decreasing, whereas L2/3 IL→BLA projectors are more active (blue rim) during extinction retrieval when fear is suppressed.

network effects. Increased input could come from several IL afferents, including the ventral hippocampus (Kim and Cho, 2017; Brockway et al., 2023), the PVT (Russo and Parsons, 2022), the BLA (Manoocheri and Carter, 2022), and the PL (Marek et al., 2018; Watanabe et al., 2021).

To probe the functional contribution of the IL→SI/VP pathway, we used inhibitory optogenetics during a shorter extinction protocol. We discovered that inhibiting the IL→ (but not PL→) SI/VP pathway during learning resulted in higher within-session freezing, without affecting next-day behavior. Thus, the IL→SI/VP pathway constrains freezing expression during learning but does not affect memory. Interestingly, when we inhibited this pathway during a second extinction session, we found that freezing was again higher in the inhibited group but only during the early trials, suggesting that the manipulation was affecting re-extinction. Given that the control group showed good extinction retrieval in this experiment, it does not completely rule out the possibility that IL→SI/VP contributes to fear suppression early in extinction retrieval and in re-extinction. In summary, these findings complement the cFos and in vitro experiments, indicating that the IL→SI/VP pathway constrains freezing only during active fear decrement associated with extinction. Thus emerges a picture of brain-wide IL communication patterns, whereby IL efferents downregulate fear expression during extinction learning via the SI/VP and promote memory for extinction retrieval via the BLA (Fig. 7).

Accumulating evidence shows that the VP engages in adaptive behavioral responses during changing circumstances (Stephenson-Jones et al., 2016, 2020; Saga et al., 2017; Hegedüs et al., 2024), as is the case during extinction learning. Indeed, the VP was shown to modulate appetitive and aversive responses while accounting for motivational state by engaging different cell

←

defensive freezing during the fear acquisition session is similar in both groups. *F*, Average percentage defensive freezing is also similar for both groups during extinction learning. *G*, Left, During Extinction 2, average percentage defensive freezing is similar across groups. Right, Average defensive freezing in Trials 1–2 is higher in the eArch than eYFP group, whereas both groups are at similarly low levels of freezing by Trials 9–10. All data are shown as mean ± SEM, *$p < 0.05$. Abbreviations: IL, infralimbic; SI, substantia innominata; VP, ventral pallidum.

types and firing patterns (Lin and Nicolelis, 2008; Richard et al., 2016; Stephenson-Jones et al., 2020; Moaddab et al., 2021). For example, VP GABAergic firing was shown to increase reward seeking, whereas glutamatergic firing increased threat avoidance (Stephenson-Jones et al., 2020), suggesting potential means for regulation of defensive behavior in extinction. In addition, SI/VP cholinergic cells innervate regions in the fear discrimination circuit such as the amygdala and prefrontal cortex (Zaborszky et al., 2015; Gielow and Zaborszky, 2017), where BF cholinergic signaling drives attentional cue processing (Parikh et al., 2007; Gritton et al., 2016). Lesioning or inhibiting the cholinergic BF can impair within-session cued extinction (Knox, 2016; Knox and Keller, 2016) and slow contextual extinction (Zelikowsky et al., 2013), indicating that cholinergic activity partakes in extinction acquisition. However, BF cholinergic signaling is also upregulated during fear learning, when acetylcholine is released in the BLA (Rajebhosale et al., 2024). This can counteract fear extinction (Jiang et al., 2016; Crimmins et al., 2023), increase BLA firing (Unal et al., 2015, Knox, 2016), and entrain threat-associated BLA theta oscillations (Lee et al., 2005; Aitta-Aho et al., 2018; Bratsch-Prince et al., 2024; Cattani et al., 2024). Thus, given the importance of cholinergic signaling for both fear conditioning and extinction, it is likely that cholinergic tone modulates attention to the CS during multiple kinds of learning (Likhtik and Johansen, 2019; Crouse et al., 2020). However, too much cholinergic signaling is associated with anxiety and depression, necessitating a balance that facilitates focus during learning but prevents overactivation (Vythilingam et al., 2007; Mineur and Picciotto, 2021). Our finding that IL projections to the SI/VP constrain fear expression during extinction learning suggests that this pathway could modulate learning rates via its interactions with different cell types in the SI/VP. One important avenue going forward is to understand which SI/VP cell types are postsynaptic to the IL and how the IL interacts with local SI/VP circuit dynamics during extinction.

Extinction is a form of prediction-error–based learning (Rescorla and Wagner, 1972), when US omission violates the predicted CS–US association that was established during conditioning, facilitating new learning about the CS (Bouton et al., 2021). Prediction-error models highlight the importance of attention to the CS for learning the changing CS–US association (Mackintosh, 1975; Pearce and Hall, 1980; Dunsmoor et al., 2015), as seen in increased attentional network activity in humans during extinction learning (Wen et al., 2021). Interestingly, acetylcholine signaling increases when a known CS becomes unreliable (Yu and Dayan, 2005), suggesting the highest cholinergic activity is at the beginning of extinction, when the prediction error is high due to initial violations of the previously established CS–US contingency. Notably, the cholinergic SI/VP sends a projection to the IL (Henny and Jones, 2008; Bloem et al., 2014; Zaborszky et al., 2015), where extinction consolidation is regulated via cholinergic receptors (Santini et al., 2012; Wilson and Fadel, 2017). Thus, there may exist an attention driven SI/VP→IL→SI/VP loop (Sarter et al., 2005), whereby SI/VP cholinergic communication with the mPFC and BLA is part of bottom-up signal about the CS. Then, IL→SI/VP top–down signaling would downregulate cholinergic (and possibly glutamatergic) activity, decreasing within-session fear expression. Indeed, prefrontal interactions with the attention network in emotion regulation (Sharpe and Killcross, 2014; Sharpe and Killcross, 2015; Wen et al., 2021) suggest that this is an important loop for cognitive control during learning.

# References

Aitta-Aho T, Hay YA, Phillips BU, Saksida LM, Bussey TJ, Paulsen O, Apergis- Schoute J (2018) Basal forebrain and brainstem cholinergic neurons differentially impact amygdala circuits and learning-related behavior. Curr Biol 28:2557–2569 e2554.

Alexander L, Wood CM, Gaskin PLR, Sawiak SJ, Fryer TD, Hong YT, McIver L, Clarke HF, Roberts AC (2020) Over-activation of primate subgenual cingulate cortex enhances the cardiovascular, behavioral and neural responses to threat. Nat Commun 11:5386.

Amano T, Unal CT, Pare D (2010) Synaptic correlates of fear extinction in the amygdala. Nat Neurosci 13:489–494.

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

Arruda-Carvalho M, Clem RL (2014) Pathway-selective adjustment of prefrontal- amygdala transmission during fear encoding. J Neurosci 34: 15601–15609.

Bayer H, Hassell JE Jr, Oleksiak CR, Garcia GM, Vaughan HL, Juliano VAL, Maren S (2024) Pharmacological stimulation of infralimbic cortex after fear conditioning facilitates subsequent fear extinction. Neuropsychopharmacology 49:1951–1957.

Bloem B, Schoppink L, Rotaru DC, Faiz A, Hendriks P, Mansvelder HD, van de Berg WD, Wouterlood FG (2014) Topographic mapping between basal forebrain cholinergic neurons and the medial prefrontal cortex in mice. J Neurosci 34:16234–16246.

Bloodgood DW, Sugam JA, Holmes A, Kash TL (2018) Fear extinction requires infralimbic cortex projections to the basolateral amygdala. Transl Psychiatry 8:60.

Bouton ME, Maren S, McNally GP (2021) Behavioral and neurobiological mechanisms of pavlovian and instrumental extinction learning. Physiol Rev 101:611–681.

Bratsch-Prince JX, Warren JW, Jones GC, McDonald AJ, Mott DD (2024) Acetylcholine engages distinct amygdala microcircuits to gate internal theta rhythm. J Neurosci 44:e1568232024.

Brockway ET, Simon S, Drew MR (2023) Ventral hippocampal projections to infralimbic cortex and basolateral amygdala are differentially activated by contextual fear and extinction recall. Neurobiol Learn Mem 205: 107832.

Bukalo O, et al. (2015) Prefrontal inputs to the amygdala instruct fear extinction memory formation. Sci Adv 1:e1500251.

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.

Burgos-Robles A, et al. (2017) Amygdala inputs to prefrontal cortex guide behavior amid conflicting cues of reward and punishment. Nat Neurosci 20:824–835.

Cattani A, Arnold DB, McCarthy M, Kopell N (2024) Basolateral amygdala oscillations enable fear learning in a biophysical model. Elife 12:RP89519.

Cho JH, Deisseroth K, Bolshakov VY (2013) Synaptic encoding of fear extinction in mPFC-amygdala circuits. Neuron 80:1491–1507.

Crimmins BE, Lingawi NW, Chieng BC, Leung BK, Maren S, Laurent V (2023) Basal forebrain cholinergic signaling in the basolateral amygdala promotes strength and durability of fear memories. Neuropsychopharmacology 48: 605–614.

Crouse RB, et al. (2020) Acetylcholine is released in the basolateral amygdala in response to predictors of reward and enhances the learning of cue-reward contingency. Elife 9:e57335.

Cruz E, Lopez AV, Porter JT (2014) Spontaneous recovery of fear reverses extinction- induced excitability of infralimbic neurons. PLoS One 9: e103596.

Do-Monte FH, Manzano-Nieves G, Quinones-Laracuente K, Ramos-Medina L, Quirk GJ (2015) Revisiting the role of infralimbic cortex in fear extinction with optogenetics. J Neurosci 35:3607–3615.

Dunsmoor JE, Niv Y, Daw N, Phelps EA (2015) Rethinking extinction. Neuron 88:47–63.

Franklin KBJ, Paxinos G (2013) *Paxinos and Franklin's the mouse brain in stereotaxic coordinates*. Amsterdam: Academic Press. an imprint of Elsevier.

Friedman AK, et al. (2016) KCNQ channel openers reverse depressive symptoms via an active resilience mechanism. Nat Commun 7:11671.

Gielow MR, Zaborszky L (2017) The input-output relationship of the cholinergic basal forebrain. Cell Rep 18:1817–1830.

Giustino TF, Maren S (2015) The role of the medial prefrontal cortex in the conditioning and extinction of fear. Front Behav Neurosci 9:298.

Giustino TF, Fitzgerald PJ, Maren S (2016) Fear expression suppresses medial prefrontal cortical firing in rats. PLoS One 11:e0165256.

Gritton HJ, Howe WM, Mallory CS, Hetrick VL, Berke JD, Sarter M (2016) Cortical cholinergic signaling controls the detection of cues. Proc Natl Acad Sci U S A 113:E1089–1097.

Grunfeld IS, Likhtik E (2018) Mixed selectivity encoding and action selection in the prefrontal cortex during threat assessment. Curr Opin Neurobiol 49:108–115.

Hagihara KM, et al. (2021) Intercalated amygdala clusters orchestrate a switch in fear state. Nature 594:403–407.

Halladay LR, Blair HT (2017) Prefrontal infralimbic cortex mediates competition between excitation and inhibition of body movements during pavlovian fear conditioning. J Neurosci Res 95:853–862.

Hegedüs P, Király B, Schlingloff D, Lyakhova V, Velencei A, Szabó Í, Mayer MI, Zelenak Z, Nyiri G, Hangya B (2024) Parvalbumin-expressing basal forebrain neurons mediate learning from negative experience. Nat Commun 15:4768.

Henny P, Jones BE (2008) Projections from basal forebrain to prefrontal cortex comprise cholinergic, GABAergic and glutamatergic inputs to pyramidal cells or interneurons. Eur J Neurosci 27:654–670.

Hurley KM, Herbert H, Moga MM, Saper CB (1991) Efferent projections of the infralimbic cortex of the rat. J Comp Neurol 308:249–276.

Jiang L, Kundu S, Lederman JD, Lopez-Hernandez GY, Ballinger EC, Wang S, Talmage DA, Role LW (2016) Cholinergic signaling controls conditioned fear behaviors and enhances plasticity of cortical-amygdala circuits. Neuron 90:1057–1070.

Kim WB, Cho JH (2017) Synaptic targeting of double-projecting ventral CA1 hippocampal neurons to the medial prefrontal cortex and basal amygdala. J Neurosci 37:4868–4882.

Kim HS, Cho HY, Augustine GJ, Han JH (2016) Selective control of fear expression by optogenetic manipulation of infralimbic cortex after extinction. Neuropsychopharmacology 41:1261–1273.

Klavir O, Prigge M, Sarel A, Paz R, Yizhar O (2017) Manipulating fear associations via optogenetic modulation of amygdala inputs to prefrontal cortex. Nat Neurosci 20:836–844.

Knox D (2016) The role of basal forebrain cholinergic neurons in fear and extinction memory. Neurobiol Learn Mem 133:39–52.

Knox D, Keller SM (2016) Cholinergic neuronal lesions in the medial septum and vertical limb of the diagonal bands of Broca induce contextual fear memory generalization and impair acquisition of fear extinction. Hippocampus 26:718–726.

Laurent V, Westbrook RF (2009) Inactivation of the infralimbic but not the prelimbic cortex impairs consolidation and retrieval of fear extinction. Learn Mem 16:520–529.

Lee MG, Hassani OK, Alonso A, Jones BE (2005) Cholinergic basal forebrain neurons burst with theta during waking and paradoxical sleep. J Neurosci 25:4365–4369.

Likhtik E, Johansen JP (2019) Neuromodulation in circuits of aversive emotional learning. Nat Neurosci 22:1586–1597.

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.

Lin SC, Nicolelis MA (2008) Neuronal ensemble bursting in the basal forebrain encodes salience irrespective of valence. Neuron 59:138–149.

Little JP, Carter AG (2012) Subcellular synaptic connectivity of layer 2 pyramidal neurons in the medial prefrontal cortex. J Neurosci 32:12808–12819.

Mackintosh NJ (1975) A theory of attention: variations in the associability of stimuli with reinforcement. Psychol Rev 82:276–298.

Mahn M, Prigge M, Ron S, Levy R, Yizhar O (2016) Biophysical constraints of optogenetic inhibition at presynaptic terminals. Nat Neurosci 19:554–556.

Manoocheri K, Carter AG (2022) Rostral and caudal basolateral amygdala engage distinct circuits in the prelimbic and infralimbic prefrontal cortex. Elife 11:e82688.

Marek R, Xu L, Sullivan RKP, Sah P (2018) Excitatory connections between the prelimbic and infralimbic medial prefrontal cortex show a role for the prelimbic cortex in fear extinction. Nat Neurosci 21:654–658.

McDonald AJ, Muller JF, Mascagni F (2011) Postsynaptic targets of GABAergic basal forebrain projections to the basolateral amygdala. Neuroscience 183:144–159.

Meyer HC, Bucci DJ (2014) The contribution of medial prefrontal cortical regions to conditioned inhibition. Behav Neurosci 128:644–653.

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

Milad MR, Quinn BT, Pitman RK, Orr SP, Fischl B, Rauch SL (2005) Thickness of ventromedial prefrontal cortex in humans is correlated with extinction memory. Proc Natl Acad Sci U S A 102:10706–10711.

Milad MR, Wright CI, Orr SP, Pitman RK, Quirk GJ, Rauch SL (2007) Recall of fear extinction in humans activates the ventromedial prefrontal cortex and hippocampus in concert. Biol Psychiatry 62:446–454.

Mineur YS, Picciotto MR (2021) The role of acetylcholine in negative encoding bias: too much of a good thing? Eur J Neurosci 53:114–125.

Moaddab M, Ray MH, McDannald MA (2021) Ventral pallidum neurons dynamically signal relative threat. Commun Biol 4: 43.

Parikh V, Kozak R, Martinez V, Sarter M (2007) Prefrontal acetylcholine release controls cue detection on multiple timescales. Neuron 56:141–154.

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

Phelps EA, Delgado MR, Nearing KI, LeDoux JE (2004) Extinction learning in humans: role of the amygdala and vmPFC. Neuron 43:897–905.

Quirk GJ, Likhtik E, Pelletier JG, Paré D (2003) Stimulation of medial prefrontal cortex decreases the responsiveness of central amygdala output neurons. J Neurosci 23:8800–8807.

Rajebhosale P, et al. (2024) Functionally refined encoding of threat memory by distinct populations of basal forebrain cholinergic projection neurons. Elife 13:e86581.

Rescorla RA, Wagner AR (1972) A theory of pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. In: Classical conditioning II: current research and theory (Black AH, Prokasy WF, eds), pp 64–99. New York: Appleton Century Crofts.

Richard JM, Ambroggi F, Janak PH, Fields HL (2016) Ventral Pallidum neurons encode incentive value and promote cue-elicited instrumental actions. Neuron 90:1165–1173.

Roberts AC, Clarke HF (2019) Why we need nonhuman primates to study the role of ventromedial prefrontal cortex in the regulation of threat- and reward-elicited responses. Proc Natl Acad Sci U S A 116:26297–26304.

Room P, Russchen FT, Groenewegen HJ, Lohman AH (1985) Efferent connections of the prelimbic (area 32) and the infralimbic (area 25) cortices: an anterograde tracing study in the cat. J Comp Neurol 242: 40–55.

Royer S, Pare D (2002) Bidirectional synaptic plasticity in intercalated amygdala neurons and the extinction of conditioned fear responses. Neuroscience G 115:455–462.

Russo AS, Parsons RG (2022) Neural activity in afferent projections to the infralimbic cortex is associated with individual differences in the recall of fear extinction. Sci Rep 12:13703.

Saga Y, Richard A, Sgambato-Faure V, Hoshi E, Tobler PN, Tremblay L (2017) Ventral pallidum encodes contextual information and controls aversive behaviors. Cereb Cortex 27:2528–2543.

Santini E, Ge H, Ren K, Pena de Ortiz S, Quirk GJ (2004) Consolidation of fear extinction requires protein synthesis in the medial prefrontal cortex. J Neurosci 24:5704–5710.

Santini E, Quirk GJ, Porter JT (2008) Fear conditioning and extinction differentially modify the intrinsic excitability of infralimbic neurons. J Neurosci 28:4028–4036.

Santini E, Sepulveda-Orengo M, Porter JT (2012) Muscarinic receptors modulate the intrinsic excitability of infralimbic neurons and consolidation of fear extinction. Neuropsychopharmacology 37:2047–2056.

Sarter M, Hasselmo ME, Bruno JP, Givens B (2005) Unraveling the attentional functions of cortical cholinergic inputs: interactions between signal-driven and cognitive modulation of signal detection. Brain Res Brain Res Rev 48:98–111.

Sharpe MJ, Killcross S (2014) The prelimbic cortex uses higher-order cues to modulate both the acquisition and expression of conditioned fear. Front Syst Neurosci 8:235.

Sharpe MJ, Killcross S (2015) The prelimbic cortex directs attention toward predictive cues during fear learning. Learn Mem 22:289–293.

Sierra-Mercado D, Padilla-Coreano N, Quirk GJ (2011) Dissociable roles of prelimbic and infralimbic cortices, ventral hippocampus, and basolateral amygdala in the expression and extinction of conditioned fear. Neuropsychopharmacology 36:529–538.

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. Cereb Cortex 19:474–482.

Stephenson-Jones M, Bravo-Rivera C, Ahrens S, Furlan A, Xiao X, Fernandes-Henriques C, Li B (2020) Opposing contributions of GABAergic and glutamatergic ventral pallidal neurons to motivational behaviors. Neuron 105:921–933.

Stephenson-Jones M, Yu K, Ahrens S, Tucciarone JM, van Huijstee AN, Mejia LA, Penzo MA, Tai LH, Wilbrecht L, Li B (2016) A basal ganglia circuit for evaluating action outcomes. Nature 539:289–293.

Strobel C, Marek R, Gooch HM, Sullivan RK, Sah P (2015) Prefrontal and auditory input to intercalated neurons of the amygdala. Cell Rep 10:1435–1442.

Stujenske JM, et al. (2022) Prelimbic cortex drives discrimination of non-aversion via amygdala somatostatin interneurons. Neuron 110:2258–2267.e11.

Szeska C, Punjer H, Riemann S, Meinzer M, Hamm AO (2022) Stimulation of the ventromedial prefrontal cortex blocks the return of subcortically mediated fear responses. Transl Psychiatry 12:394.

Tao Y, Cai CY, Xian JY, Kou XL, Lin YH, Qin C, Wu HY, Chang L, LuoCX, Zhu DY (2021) Projections from infralimbic Cortex to paraventricular thalamus mediate fear extinction retrieval. Neurosci Bull 37:229–241.

Unal CT, Pare D, Zaborszky L (2015) Impact of basal forebrain cholinergic inputs on basolateral amygdala neurons. J Neurosci 35:853–863.

Vertes RP (2004) Differential projections of the infralimbic and prelimbic cortex in the rat. Synapse 51:32–58.

Vidal-Gonzalez I, Vidal-Gonzalez B, Rauch SL, Quirk GJ (2006) Microstimulation reveals opposing influences of prelimbic and infralimbic cortex on the expression of conditioned fear. Learn Mem 13:728–733.

Vythilingam M, et al. (2007) Biased emotional attention in post-traumatic stress disorder: a help as well as a hindrance? Psychol Med 37:1445–1455.

Wallis CU, Cardinal RN, Alexander L, Roberts AC, Clarke HF (2017) Opposing roles of primate areas 25 and 32 and their putative rodent homologs in the regulation of negative emotion. Proc Natl Acad Sci U S A 114:E4075–E4084.

Watanabe M, Uematsu A, Johansen JP (2021) Enhanced synchronization between prelimbic and infralimbic cortices during fear extinction learning. Mol Brain 14:175.

Wen Z, Chen ZS, Milad MR (2021) Fear extinction learning modulates large-scale brain connectivity. Neuroimage 238:118261.

Wilson MA, Fadel JR (2017) Cholinergic regulation of fear learning and extinction. J Neurosci Res 95:836–852.

Xu M, et al. (2015) Basal forebrain circuit for sleep-wake control. Nat Neurosci 18:1641–1647.

Yu AJ, Dayan P (2005) Uncertainty, neuromodulation, and attention. Neuron 46:681–692.

Zaborszky L, Gaykema RP, Swanson DJ, Cullinan WE (1997) Cortical input to the basal forebrain. Neuroscience 79:1051–1078.

Zaborszky L, Csordas A, Mosca K, Kim J, Gielow MR, Vadasz C, Nadasdy Z (2015) Neurons in the basal forebrain project to the cortex in a complex topographic organization that reflects corticocortical connectivity patterns: an experimental study based on retrograde tracing and 3D reconstruction. Cereb Cortex 25:118–137.

Zelikowsky M, Hast TA, Bennett RZ, Merjanian M, Nocera NA, Ponnusamy R, Fanselow MS (2013) Cholinergic blockade frees fear extinction from its contextual dependency.