

Prefrontal somatostatin interneurons encode fear memory

Kirstie A. Cummings and Roger L. Clem *

Theories stipulate that memories are encoded within networks of cortical projection neurons. Conversely, GABAergic interneurons are thought to function primarily to inhibit projection neurons and thereby impose network gain control, an important but purely modulatory role. Here we show in male mice that associative fear learning potentiates synaptic transmission and cue-specific activity of medial prefrontal cortex somatostatin (SST) interneurons and that activation of these cells controls both memory encoding and expression. Furthermore, the synaptic organization of SST and parvalbumin interneurons provides a potential circuit basis for SST interneuron-evoked disinhibition of medial prefrontal cortex output neurons and recruitment of remote brain regions associated with defensive behavior. These data suggest that, rather than constrain mnemonic processing, potentiation of SST interneuron activity represents an important causal mechanism for conditioned fear.

Associative memory is a critical function of cortical brain networks, which are primarily populated by excitatory projection neurons (PNs) and inhibitory interneurons. The most abundant of these cell types, PNs are a key substrate for interregional brain signaling that is critical for memory expression^{1,2}. Accordingly, retrieval cues activate subsets of PNs that are hypothesized to encode stimulus associations through persistent changes in excitatory synapse strength and density³. In contrast, GABAergic interneurons are generally thought to inhibit PNs^{4–8}, which has been suggested to play a role in optimizing the dynamic range of PN firing to indirectly modulate the strength and specificity of learning. However, while several studies credit interneurons with such ‘supporting roles’, it remains unclear whether they can directly mediate the encoding of cue associations through their own functional plasticity.

Fear conditioning is a powerful model of such learning where an animal acquires survival-based defensive reactions to a conditioned stimulus (CS) that predicts imminent threat. The expression of fear memory in rodents requires neural activity in the prelimbic subregion of the medial prefrontal cortex (mPFC)⁹, where both PNs and interneurons sampled by extracellular recordings exhibit CS-evoked changes in firing rate after conditioning⁴. However, whether learning induces synaptic plasticity in prelimbic circuits and, if so, whether interneuron activity is modulated by these changes in conjunction with memory encoding is unknown. In this study, we address these questions in mice with a combination of synaptic electrophysiology, calcium imaging, optogenetic manipulation and brain activity mapping of prelimbic interneurons and associated circuitry. We demonstrate that SST interneurons exhibit properties indicative of a memory storage substrate, including (1) learning-dependent potentiation of synaptic transmission, (2) cue-specific activation during memory retrieval and (3) bidirectional modulation of memory expression. Moreover, prelimbic SST interneurons exert potent disinhibitory control over a fear-related brain network, suggesting a fundamental role for these cells in orchestrating conditioned fear responses.

Results

Cued fear learning potentiates SST interneuron excitatory input.

While experience-dependent plasticity is considered to be the most probable mechanism for cortical information storage^{2,3}, the extent

to which learning is associated with plasticity of cortical inhibitory circuits remains poorly understood. Therefore, to determine whether fear learning alters the synaptic properties of prefrontal interneurons, we obtained *ex vivo* electrophysiological recordings from parvalbumin (PV)- and SST-expressing cells, which together comprise the majority of cortical GABAergic interneurons¹⁰. To identify these cell types in acute brain slices we generated interneuron-specific expression of tdTomato by crossing the Ai9 reporter line to PV- or SST-Cre driver mice, which exhibit highly selective recombination in the prelimbic cortex (Fig. 1a) (ref. ¹¹). These animals then underwent behavioral training entailing either paired or unpaired presentations of an auditory CS and footshock (unconditioned stimulus (US); Fig. 1b and Supplementary Fig. 1a,f). Because only paired animals acquire CS-evoked defensive freezing, unpaired mice served as a control for the nonassociative effects of stimulus exposure^{12,13}. At 24 h after training, spontaneous excitatory postsynaptic currents (sEPSCs) and inhibitory postsynaptic currents were recorded and analyzed as a proxy for potential synaptic plasticity in prelimbic interneurons (Fig. 1b,c and Supplementary Fig. 1). Mice that received CS–US pairing displayed higher sEPSC frequency in SST but not PV interneurons residing in layer 2/3 (L2/3), compared to naïve and unpaired controls (Fig. 1c). No other differences in sEPSC or spontaneous inhibitory postsynaptic current properties were associated specifically with CS–US pairing (Supplementary Fig. 1). Because sEPSC frequency can reflect differences in presynaptic efficacy, we next measured the response of L2/3 SST interneurons to local paired-pulse stimulation, a well-established assay for neurotransmitter release probability (Fig. 1d) (ref. ¹⁴). Consistent with increased glutamate release onto SST interneurons after fear conditioning, evoked EPSCs exhibited a higher paired-pulse ratio in animals that received CS–US stimulus pairing but not unpaired training. These results confirm that cued fear learning is associated with potentiation of excitatory synapses onto SST interneurons.

SST interneurons signal the CS after learning. Given the observed learning-related potentiation of their excitatory input (Fig. 1), we hypothesized that SST interneurons in conditioned mice might exhibit increased CS-related activity and thereby participate in memory processing. To test this possibility, we utilized fiber

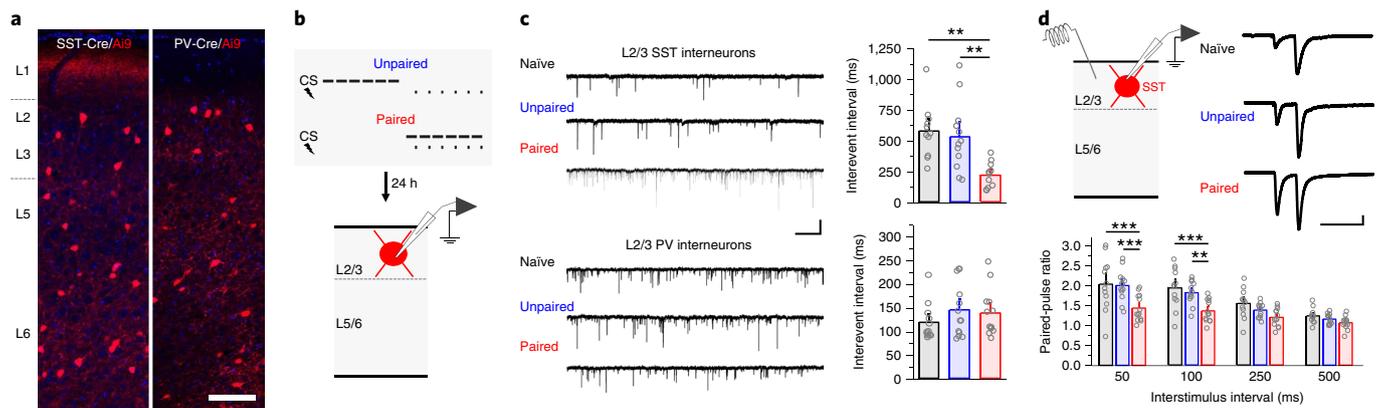


Fig. 1 | Potentiation of synaptic transmission in prefrontal SST interneurons after cued fear learning. **a**, Prelimbic SST interneurons and PV interneurons were identified by tdTomato expression in SST-IRES-Cre/Ai9 or PV-IRES-Cre/Ai9 mice. Scale bar, 200 μ m. **b**, Animals were either cage-experienced (naïve) or trained using paired or unpaired presentations of CS (2 kHz, 80 dB, 20 s) and US (1 mA, 2 s), followed 24 h later by whole-cell recording in acute brain slices. **c**, Recordings of sEPSCs were obtained from L2/3 SST interneurons (top) and PV interneurons (bottom). Example raw traces and mean interevent intervals are depicted for experimental and control groups. SST interneurons: $F_{2,31} = 9.03$, $P = 8.11 \times 10^{-4}$, one-way ANOVA; naïve, $n = 11$ cells (7 slices from 3 mice); unpaired $n = 12$ cells (9 slices from 3 mice); paired $n = 11$ cells (12 slices from 4 mice). PV interneurons: $F_{2,45} = 0.52$, $P = 0.60$, one-way ANOVA; naïve, $n = 13$ cells (9 slices from 3 mice); unpaired, $n = 13$ cells (8 slices from 3 mice); paired, $n = 13$ cells (11 slices from 4 mice). Scale bars, 20 pA \times 1 s. **d**, EPSC recordings in L2/3 SST interneurons during paired-pulse stimulation. Paired-pulse ratio: $F_{6,66} = 3.46$, $P = 0.0049$, interaction between ratio and training, two-way repeated measures ANOVA; naïve, $n = 13$ cells (7 slices from 3 mice); unpaired, $n = 13$ cells (6 slices from 3 mice); paired, $n = 13$ cells (6 slices from 3 mice). Scale bars, 100 pA \times 100 ms. ** $P < 0.01$, *** $P < 0.001$, Tukey's post hoc test. The bar graphs depict the mean \pm s.e.m.

photometry to monitor *in vivo* activity of SST interneurons after conditional viral expression of GCaMP6f (AAV-DIO-GCaMP6f-eYFP) in SST-Cre mice (Fig. 2 and Supplementary Fig. 2). Mice were unilaterally implanted with an optic fiber (400 μ m core diameter) directed at the prelimbic cortex and imaged under freely behaving conditions during CS exposure before, during and after auditory fear conditioning (Fig. 2b and Supplementary Fig. 3). During CS–US pairing, we observed an increase in CS-related calcium signals during trials 5 and 6 compared to the initial conditioning trial (Fig. 2b,c), which mirrored a trial-dependent increase in CS-evoked freezing (Supplementary Fig. 3). One day after conditioning, CS-related calcium signals during reexposure to the CS were markedly increased compared to a preconditioning test (Fig. 2b,d). Because CS presentation leads to defensive freezing, the recruitment of SST interneurons during memory retrieval could be a consequence of fear expression rather than CS modulation per se. Therefore, we additionally analyzed calcium signals during intertrial freezing bouts, during which fear-related SST interneuron activity can be dissociated from CS processing. Unlike the CS trials, transitions from movement to freezing during the intertrial intervals were associated with negligible fluorescence changes (Fig. 2d).

To further test whether fearful states modulate SST interneuron activity independent of the CS, we next performed SST interneuron-specific fiber photometry in mice that underwent unpaired CS–US training (Supplementary Fig. 4a–c). Importantly, although unpaired training results in context conditioned fear, it does not induce synaptic plasticity in SST interneurons (Fig. 1c,d). During unpaired conditioning, US trials were associated with large calcium signals, confirming that changes in SST interneuron activity could be readily detected (Supplementary Fig. 4d,f). However, regardless of whether imaging was conducted in a new context (context B) or the original training arena (context A), no changes in calcium signals were associated with spontaneous freezing bouts (Supplementary Fig. 4e,f). Moreover, there were no overall differences in the frequency of calcium transients (peaks) during exposure to contexts A and B relative to a preconditioning baseline in context A (Supplementary Fig. 4g), despite robust differences in freezing between these tests (Supplementary Fig. 4b). These data indicate that SST interneurons

do not generally signal a high fear state, but are instead specifically activated by the threat-associated cues.

To independently confirm the activation of SST interneurons in response to memory retrieval, we performed immunohistochemical labeling for the activity reporter *c-Fos* (Supplementary Fig. 5), which is a marker of neurons strongly activated by mnemonic cues³. To elicit memory retrieval in SST-Cre/Ai9 reporter mice, four CS trials were presented at 24 h after fear conditioning in a context distinct from the training arena; as control conditions, we examined mice where either conditioning or memory retrieval were omitted. Retrieval CS trials triggered a robust increase in freezing only in mice that had been previously conditioned (Supplementary Fig. 5b–d). Following behavioral testing, substantially more SST interneurons exhibited *c-Fos* immunoreactivity in these animals compared to those in either control group (Supplementary Fig. 5a,e,f). Increased SST interneuron activation in conditioned mice could also be observed specifically in L2/3, consistent with a causal role for lamina-specific SST interneuron plasticity (Fig. 1). In contrast to these results, PV interneurons did not exhibit a detectable increase in *c-Fos* immunoreactivity under the same conditions in PV-Cre/Ai9 mice (Supplementary Fig. 6).

SST interneuron activity controls memory expression. CS activation of SST interneurons suggests that recruitment of these cells could be important for memory retrieval. On the other hand, previous studies suggested that SST interneuron activity might also function to constrain fear expression through the inhibition of cue-responsive PNs^{4–8}. To establish the behavioral impact of SST interneuron recruitment, we utilized optogenetics to modulate SST interneuron activity in conjunction with CS-evoked memory retrieval. We microinjected into the prelimbic cortex viral vectors encoding archaerhodopsin (Arch; AAV-FLEX-Arch3.0-GFP), channelrhodopsin-2 (ChR2; AAV-DIO-ChR2-eYFP) or an opsin-negative control vector (enhanced yellow fluorescent protein (eYFP); AAV-DIO-eYFP). Importantly, optic illumination was sufficient to hyperpolarize or depolarize Arch- or channelrhodopsin-expressing SST interneurons, respectively, and thereby reliably control the firing of these cells (Supplementary Fig. 7a,b). Following

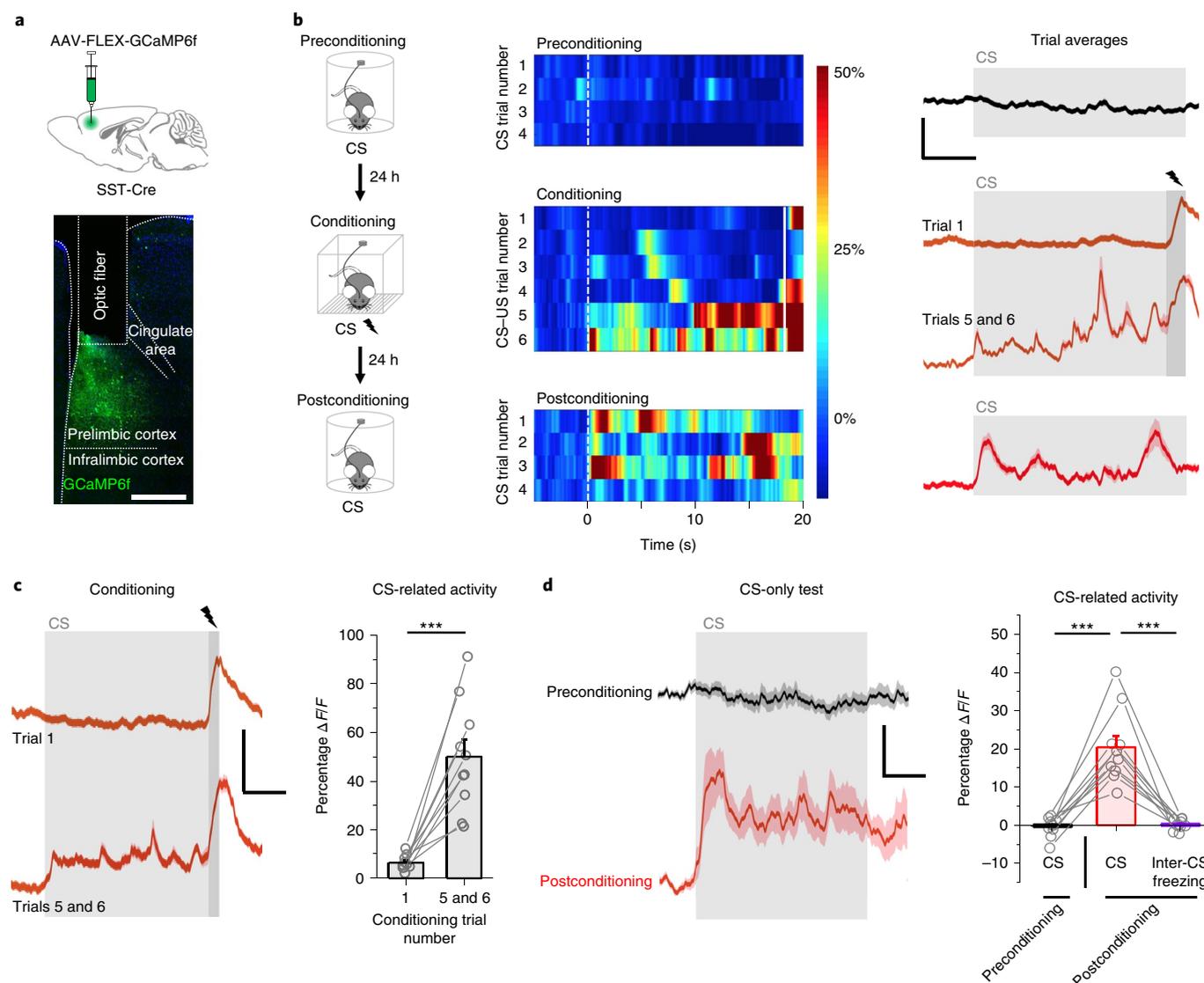


Fig. 2 | CS-related SST interneuron activity increases in tandem with memory acquisition. **a**, For Ca^{2+} -based imaging of SST interneuron activity, SST-IRES-Cre mice ($n = 10$) received injections of conditional vector encoding GCaMP6f and were implanted with a single optic fiber (400 μm core diameter) directed at the prelimbic cortex. Scale bar, 500 μm . **b**, After surgical recovery, CS-related freezing and the Ca^{2+} -dependent fluorescence signal were monitored before, during and after CS-US stimulus pairing, which entailed six coterminating trials of CS (2 kHz, 20 s, 80 dB) and US (0.7 mA footshock, 2 s). Pre- and postconditioning CS tests consisted of four CS trials in a context distinct from the training arena. Heat maps (middle) and mean Ca^{2+} traces (right) \pm s.e.m. (shaded area) depict SST interneuron responses to all conditioned and unconditioned trials for a representative animal. Scale bars, 30% $\Delta F/F \times 5$ s. **c**, Group-averaged Ca^{2+} traces (left) \pm s.e.m. (shaded area) for CS-US trials and mean change in CS-associated peak fluorescence (percentage $\Delta F/F$) for trial 1 and trials 5 and 6. Effect of trial: $n = 10$ mice, $t_9 = 5.95$, $P = 2.15 \times 10^{-4}$, two-sided paired t -test. Scale bars, 50% $\Delta F/F \times 5$ s. **d**, Group-averaged Ca^{2+} traces (left) \pm s.e.m. (shaded area). Right: comparison of percentage $\Delta F/F$ between pre- and postconditioning CS presentations, as well as freezing-related epochs that occurred independently of the CS during the inter-trial intervals of the postconditioning test. CS-related activity: $n = 10$ mice, $F_{2,18} = 42.82$, $P = 1.44 \times 10^{-7}$, one-way repeated-measures ANOVA. The experiment was performed in three independent cohorts, and mice were pooled together for analysis. The traces represent the average of four CS trials. Scale bars, 10% $\Delta F/F \times 5$ s. *** $P < 0.001$ by two-sided paired t -test (**c**) or Tukey's post hoc test (**d**). The bar graphs depict the mean \pm s.e.m.

virus infusion, mice were implanted with optic fibers directed at the prelimbic cortex (Fig. 3a and Supplementary Fig. 7c,d). One week after surgery, these mice underwent CS-US pairing in the absence of photostimulation.

At 24 h after training, a memory retrieval test was conducted where the independent and combined effects of light and CS were examined in a context distinct from the training arena. Compared to CS-only trials, a marked reduction in freezing was observed in Arch-expressing mice during light (532 nm, 20 s, constant) + CS trials (Fig. 4b,d), indicating that SST interneuron activity is required

for cued memory expression. No reduction in freezing was observed in Arch-expressing animals during light-only trials compared to baseline, suggesting that SST interneuron activity is not required for generalized context fear. Conversely, in Chr2-expressing mice, photoexcitation of SST interneurons (473 nm, 10 ms pulses, 20 Hz), even in the absence of CS presentation, was sufficient to increase freezing over baseline levels and thereby mimic a conditioned response (Fig. 3c,d). Combined CS + light presentation in these animals did not increase freezing beyond that observed during interleaved CS-only trials, suggesting the possibility of a ceiling effect.

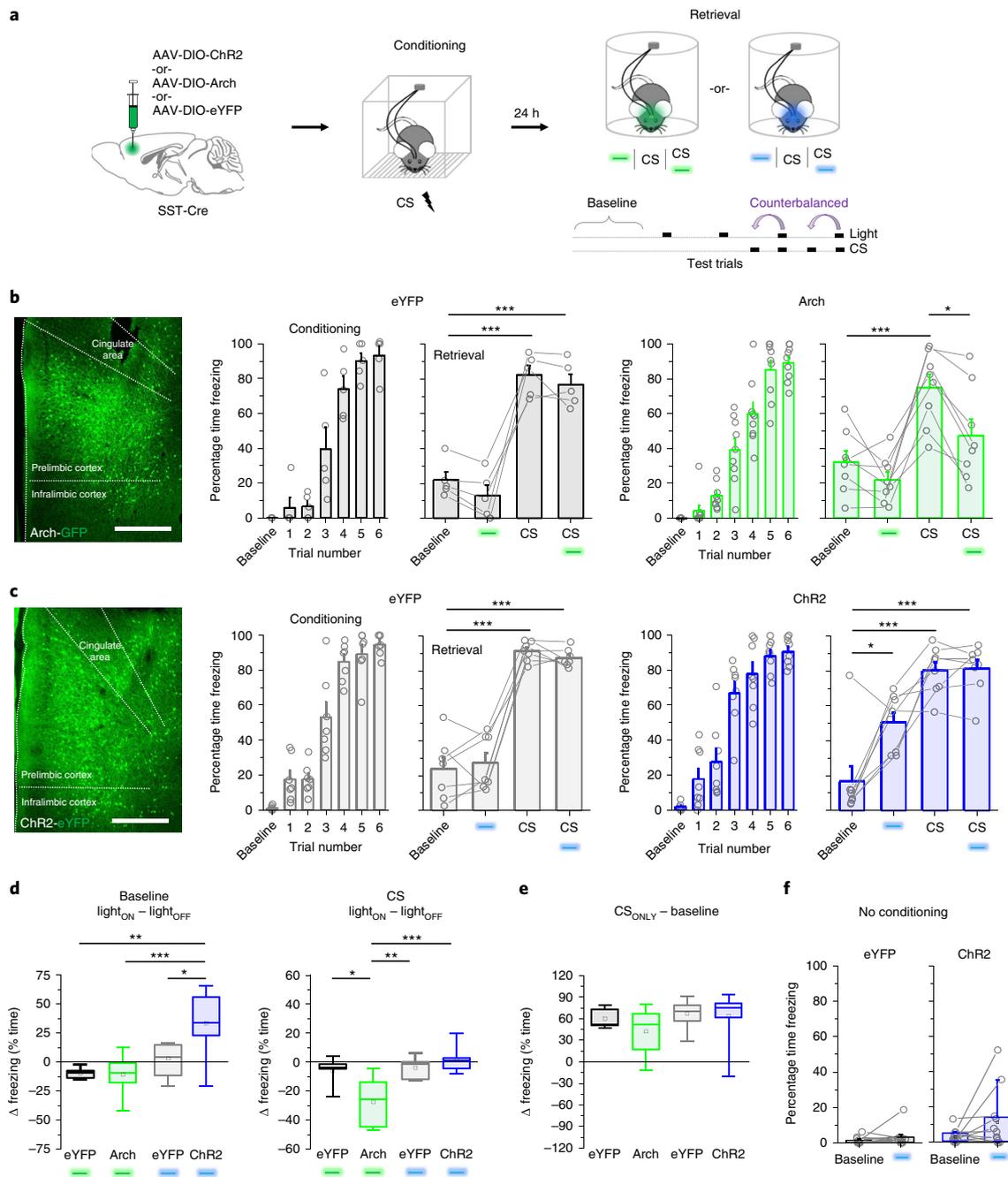


Fig. 3 | SST interneuron activity is necessary for memory expression and induces freezing in conditioned mice. a, For in vivo activity manipulation, SST-IRES-Cre mice received injections of conditional Arch, ChR2 or eYFP control vectors and were implanted with optic ferrules directed at the prelimbic cortex. After surgical recovery, all animals underwent auditory fear conditioning in the absence of optic illumination. Freezing was quantified 24 h later during independent and simultaneous presentation of CS and photostimulation in a distinct context. **b**, Modulation of freezing by CS and light (532 nm; constant, 20 s epochs) in Arch (green) or eYFP control mice (black). Arch: $F_{3,21} = 11.3$, $P = 1.29 \times 10^{-4}$, one-way repeated-measures ANOVA, $n = 8$ mice; eYFP: $F_{3,12} = 68.6$, $P = 8.03 \times 10^{-8}$, repeated-measures ANOVA, $n = 5$ mice. Scale bar, 500 μm . The experiment was performed in three independent cohorts, and mice were pooled together for analysis. **c**, Modulation of freezing by CS and light (473 nm, 10 ms pulses at 20 Hz) in ChR2 (blue) or eYFP control mice (gray). ChR2: $F_{3,21} = 19.3$, $P = 3.06 \times 10^{-6}$, one-way repeated-measures ANOVA; $n = 9$ mice. eYFP: $F_{3,18} = 87.7$, $P = 6.20 \times 10^{-11}$, one-way repeated-measures ANOVA; $n = 7$ mice. Scale bar, 500 μm . The experiment was performed in four independent cohorts, and mice were pooled together for analysis. **d**, Mean change in freezing induced by photostimulation during baseline and CS periods, as compared across all light and opsin combinations used in **b** and **c**. Photostimulation effect on baseline freezing (light_{ON} - light_{OFF}): $F_{3,24} = 8.88$, $P = 3.87 \times 10^{-4}$, one-way ANOVA. Photostimulation effect on CS-evoked freezing (light_{ON} - light_{OFF}): $F_{3,24} = 9.61$, $P = 2.37 \times 10^{-4}$, one-way ANOVA. **e**, Mean change in freezing induced by CS presentation compared across all experimental groups used in **b** and **c**. CS-evoked change in freezing (CS_{ONLY} at baseline): $F_{3,24} = 1.12$, $P = 0.36$, one-way ANOVA. **f**, Freezing behavior during light-only trials (473 nm, 10 ms pulses, 20 Hz) in naïve (unconditioned) mice expressing either the ChR2 or eYFP control vectors. ChR2: $W = 17$, $P = 0.088$, two-sided Wilcoxon signed-rank test; $n = 12$ mice. eYFP: $W = 12$, $P = 0.46$, two-sided Wilcoxon signed-rank test; $n = 11$ mice. The experiment was performed in four independent cohorts, and mice were pooled together for analysis. CS (2 kHz tone, 80 dB, 20 s duration). Blue and green boxes, laser illumination. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$ Tukey's post hoc test (**b,c,d**). The bar graphs depict the mean \pm s.e.m. The box plots depict the median (center line), mean (black box), quartiles and 10–90% range (whiskers).

Importantly, light-evoked freezing cannot be explained by a non-specific motor deficit because photoexcitation of ChR2-expressing SST interneurons did not alter the standard metrics of locomotion in the open field test (Supplementary Fig. 8). Finally, no light effects were observed in eYFP control groups that were stimulated with the same parameters used in Arch- (Fig. 3b,d) or ChR2-expressing mice (Fig. 3c,d); all groups displayed a CS-evoked increase in freezing of similar magnitude and therefore did not differ in memory strength (Fig. 3e).

To establish whether fear-promoting properties are intrinsic to SST interneurons or acquired through learning, we performed photoexcitation of SST interneurons without prior fear conditioning. In contrast to fear-conditioned mice (Fig. 3c), no consistent effect of photoexcitation was observed in naïve animals (Fig. 3f).

SST interneuron activation and plasticity mediates memory acquisition. Given the critical role of SST interneurons in memory expression, we next sought to utilize photoinhibition to determine whether SST interneuron activity at the time of CS–US pairing is required for memory acquisition as well as associated plasticity of SST interneurons. After prelimbic infusion of AAV-FLEX-Arch3.0-GFP or eYFP control vectors, SST-Cre mice underwent CS–US pairing, during which photoinhibition was timed to coincide with each of the six CS–US trials (Fig. 4a and Supplementary Fig. 9). On the following day, memory retrieval was examined in the absence of photoinhibition during presentation of four CS retrieval trials. Remarkably, while CS presentations triggered an increase in freezing in both Arch-expressing mice and eYFP controls, the magnitude of CS-evoked increase in freezing was dramatically lower in Arch-expressing mice (Fig. 4b,c). This effect cannot be attributed solely to Arch expression because a retrieval deficit was not observed in Arch-expressing mice that were conditioned without photoinhibition (Fig. 3e). Following memory retrieval, acute brain slices were obtained to compare excitatory synaptic transmission in prelimbic SST interneurons of Arch versus eYFP animals, focusing on L2/3 cells located within 1 mm below the optic fiber track. Compared to eYFP controls, Arch-expressing SST interneurons exhibited lower sEPSC frequency and higher paired-pulse ratios of evoked EPSCs (Fig. 4d,e).

While these results suggest that prefrontal SST interneurons participate in memory encoding, previous studies have established that

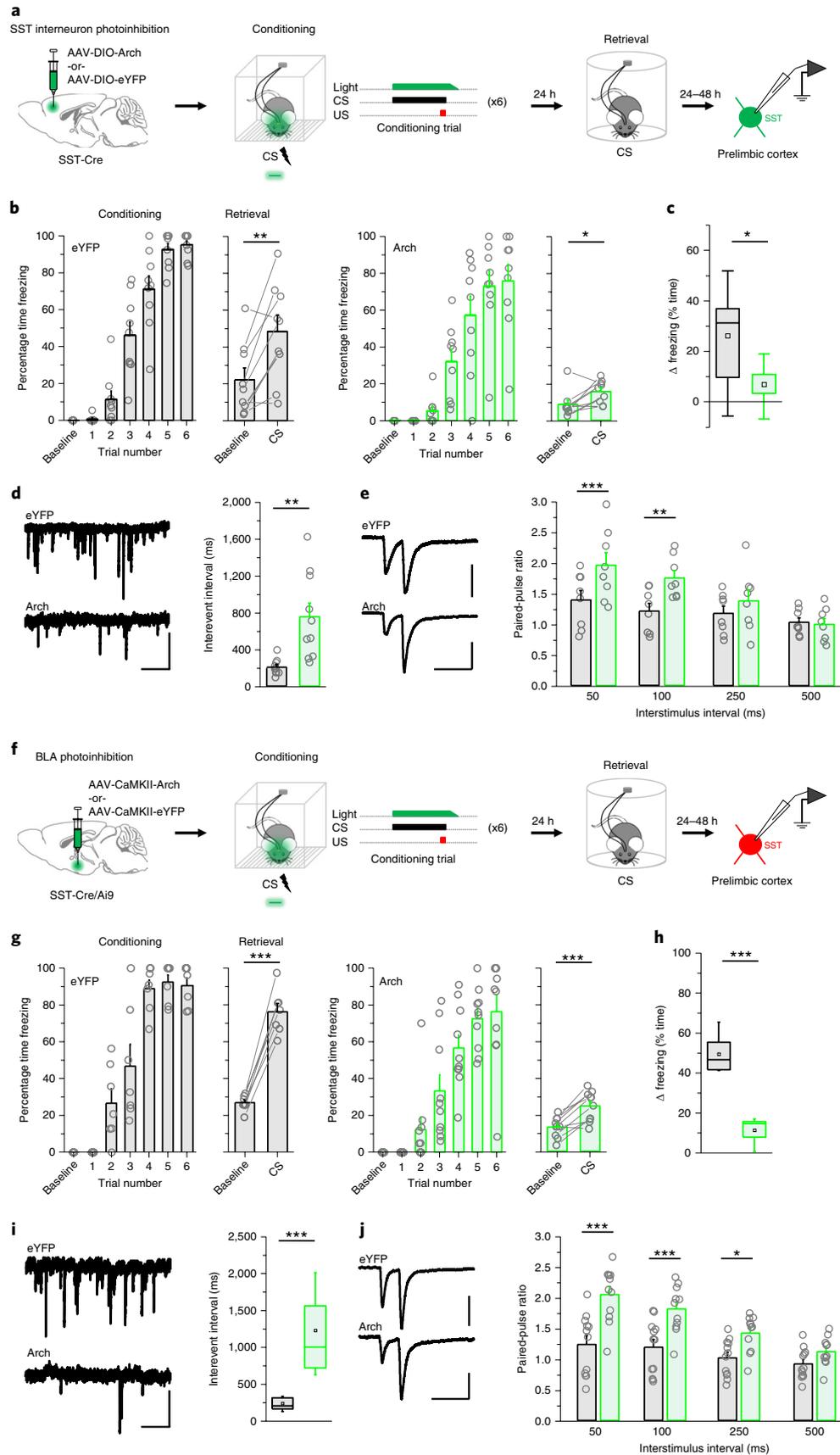
fear conditioning requires basolateral amygdala (BLA) activity¹⁵ and memory storage is widely believed to be mediated by plasticity of synaptic connections within the lateral and basal nuclei^{12,13,16–22}. Therefore, because activation of prelimbic SST interneurons evokes a fear response (Fig. 3c), we wondered whether deficits in learning after amygdala silencing can be explained in part by reduced prefrontal SST interneuron transmission. To test this hypothesis, we performed photoinhibition of BLA excitatory neurons during CS–US pairing. One week before conditioning, SST-Cre/Ai9 mice received BLA injections of Arch or eYFP control vectors under the control of a Ca²⁺/calmodulin-dependent protein kinase II (CaMKII) promoter (AAV-CaMKII-ArchT-GFP or AAV-CaMKII-eYFP) and were implanted with optic ferrules directed at the BLA (Fig. 4f and Supplementary Fig. 10). Fear conditioning and retrieval tests were conducted as described for prelimbic SST interneuron photoinhibition (Fig. 4f). While freezing occurred in both groups during conditioning (Fig. 4g), Arch mice exhibited a deficit in CS-evoked responses relative to eYFP mice during the memory retrieval test (Fig. 4h). Following retrieval, whole-cell recordings revealed that SST interneurons exhibited lower spontaneous EPSC frequency as well as increased paired-pulse ratios of evoked EPSCs (Fig. 4i,j) in Arch-expressing mice, compared to eYFP controls. These data independently confirm the relationship between memory encoding and increased SST interneuron transmission, and suggest that SST interneuron plasticity depends at least in part on BLA activity.

Microcircuit organization and opposing behavioral roles of SST and PV interneurons. An increase in freezing on SST interneuron photoexcitation (Fig. 3c) is surprising given that SST interneurons probably form inhibitory contacts onto excitatory PNs that control memory expression¹⁵. However, we considered the possibility that another class of interneurons also receives input from SST interneurons and, when inhibited via these connections, might be responsible for disinhibition of PNs²³. In particular, previous work indicates that mPFC PV interneurons exhibit firing rate inhibition during CS exposure, and that photoinhibition of PV interneurons elicits freezing⁴. Therefore, we utilized optogenetics to establish whether SST interneurons form functional connections with PV interneurons, and compared the strength of such connections with those formed onto PNs from the same brain slices. In addition, we asked whether CS–US pairing affects the balance of transmission from SST or PV

Fig. 4 | SST interneuron activation and plasticity mediates memory formation. **a**, For in vivo manipulation of prelimbic SST interneurons, SST-IRES-Cre mice received injections of conditional Arch or eYFP control vectors and were implanted with optic ferrules directed at the prelimbic cortex. After surgical recovery, all animals underwent auditory fear conditioning, during which light stimulation (532 nm, constant, 20 s epoch, ramp offset) coincided with each of 6 CS–US trials. Freezing was quantified 24 h later during presentation of a 4 CS trial in a context distinct from the training arena. During the 24–48 h subsequent to the retrieval test, ex vivo recordings were obtained from eYFP⁺ neurons (SST interneurons) located not more than 1 mm below the tip of the optic fiber in L2/3 of acute brain slices. **b**, Modulation of freezing by CS retrieval in Arch (green) or eYFP control mice (black). Arch retrieval: $t_8 = 2.71$, $P = 0.026$, two-sided paired t -test; $n = 9$ mice. eYFP retrieval: $t_8 = 3.78$, $P = 0.0054$, paired t -test; $n = 9$ mice. **c**, Mean change in freezing induced by CS presentation during memory retrieval for the experimental groups in **b**. Arch versus eYFP: $t_{16} = 2.60$, $P = 0.019$, two-sided unpaired t -test. **d**, Example raw traces and interevent intervals for sEPSCs in Arch versus eYFP mice. Interevent interval: $t_{16} = 2.60$, $P = 0.019$, two-sided unpaired t -test; Arch, $n = 10$ cells (5 slices from 3 mice); eYFP, $n = 10$ cells (5 slices from 3 mice). Scale bars, 10 pA \times 0.5 s. **e**, EPSC recordings in SST interneurons during local paired-pulse electrical stimulation. Example traces collected at 50 ms interstimulus intervals. Paired-pulse ratio: $F_{3,21} = 5.55$, $P = 0.0058$, interaction between ratio and training, two-way repeated-measures ANOVA; Arch, $n = 8$ cells (4 slices from 3 mice); eYFP, $n = 8$ cells (4 slices from 3 mice). Scale bars, 40 pA \times 100 ms. **f**, Experimental design for in vivo manipulation of BLA PNs was the same as in **a**, except that vector injections and light stimulation were delivered to the BLA, and ex vivo recordings were obtained from Tomato⁺ SST interneurons in the prelimbic cortex of SST-IRES-Cre/Ai9 mice. **g**, Modulation of freezing by CS retrieval in Arch (green) or eYFP control mice (black). Arch retrieval: $t_9 = 5.69$, $P = 2.98 \times 10^{-4}$, two-sided paired t -test; $n = 10$ mice. eYFP retrieval: $t_6 = 14.74$, 6.14×10^{-6} , paired t -test; $n = 7$ mice. **h**, Mean change in freezing induced by CS presentation during memory retrieval for the experimental groups in **b**. Arch versus eYFP: $U = 0$, $P = 1.03 \times 10^{-4}$, two-sided Mann-Whitney U -test. **i**, Example raw traces and interevent intervals for sEPSCs in Arch versus eYFP mice. Interevent interval: $U = 0$, $P = 9.97 \times 10^{-4}$, two-sided Mann-Whitney U -test; Arch, $n = 14$ cells (6 slices from 3 mice); eYFP, $n = 13$ cells (8 slices from 4 mice). Scale, 10 pA \times 0.5 s. **j**, EPSC recordings in SST interneurons during local paired-pulse electrical stimulation. Example traces collected at the 50 ms interstimulus interval. Paired-pulse ratio: $F_{3,30} = 11.18$, $P = 4.31 \times 10^{-5}$, interaction between ratio and training, two-way repeated-measures ANOVA; Arch, $n = 12$ cells (6 slices from 3 mice); eYFP, $n = 12$ cells (6 slices from 3 mice). Scale bars, 50 pA \times 100 ms. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$ by two-sided paired t -test (**b,g**), two-sided unpaired t -test (**c,d**), two-sided Mann-Whitney U -test (**h,i**) or Tukey's post hoc test (**j**). The bar graphs depict the mean \pm s.e.m. The box plots depict the median (center line), mean (black box), quartiles and 10–90% range (whiskers).

interneurons onto inhibitory versus excitatory populations, which might modulate whether SST or PV interneuron recruitment leads predominantly to inhibition or disinhibition of PNs.

First, to enable selective interrogation of synaptic outputs from SST interneurons, we microinjected an INTRSECT ChR2 vector (AAV-Cre_{OFF}/Flp_{ON}-ChR2-eYFP)²⁴ into the prelimbic cortex



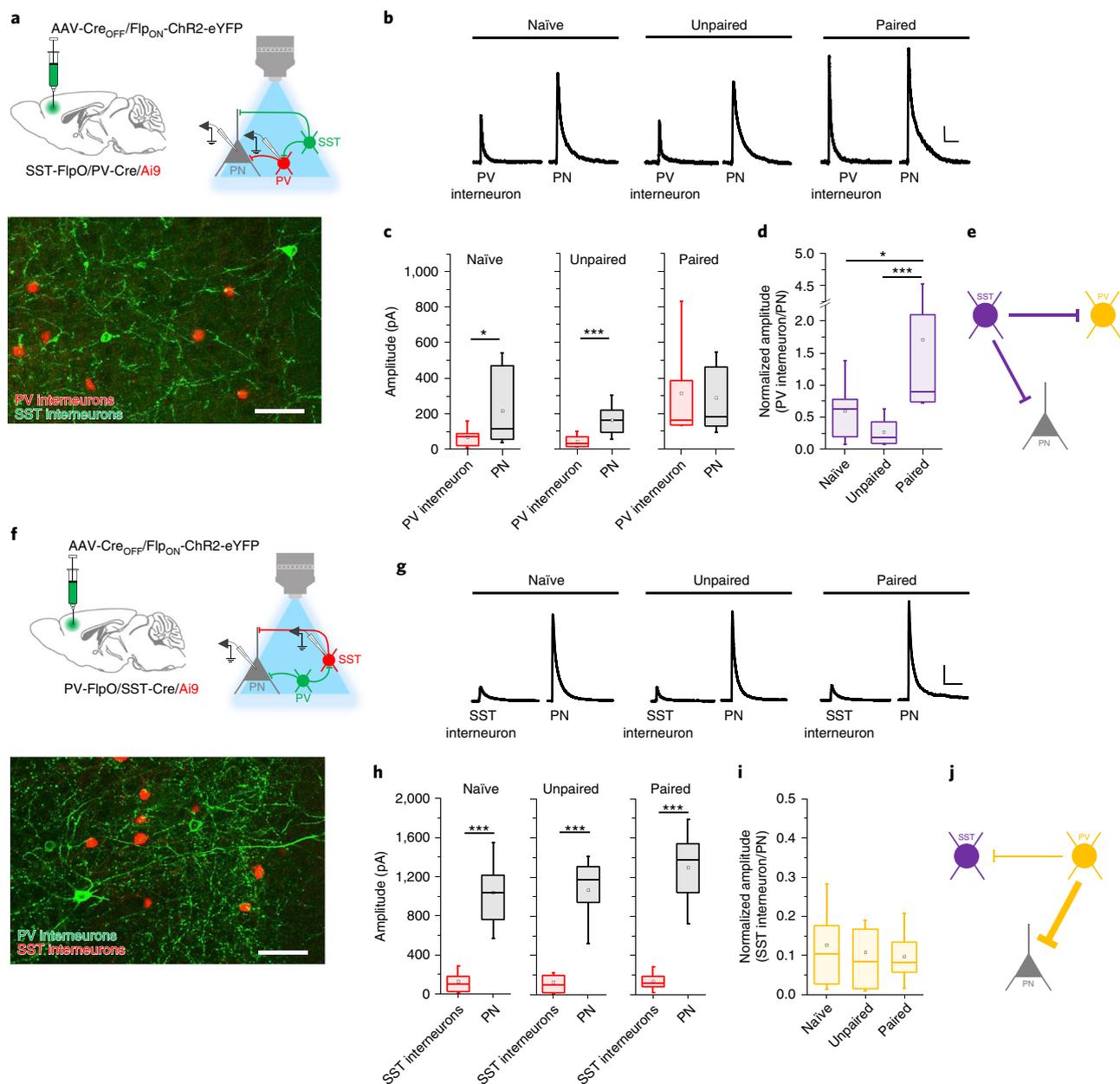


Fig. 5 | SST interneurons elicit relatively potent inhibition of PV interneurons, while PV interneurons preferentially inhibit PNs. a, To isolate monosynaptic responses to SST interneuron photostimulation in PV interneurons, we first infused a Flp_{ON}, Cre_{OFF} INTRASECT Chr2 vector into SST-IRES-FlpO/PV-IRES-Cre/Ai9 triple-transgenic mice. At 24 h after CS-US pairing, recordings were obtained from Tomato⁺ PV interneurons as well as surrounding PNs. Scale bar, 50 μ m. **b**, Example IPSC traces. Scale bars, 200 pA \times 1 s. **c**, Amplitude of IPSCs resulting from SST interneuron photoexcitation (460 nm, 1 ms pulse, 0.1 Hz) in slices from naïve mice (8 slices from 4 mice), as well as those that received unpaired (6 slices from 3 mice) or paired training (8 slices from 4 mice). Effect of cell type (naïve): $U = 46.5$, $P = 0.017$, two-sided Mann-Whitney U -test; PV interneuron, $n = 14$ cells; PN, $n = 14$ cells. Effect of cell type (unpaired): $U = 17$, $P = 8.22 \times 10^{-4}$, two-sided Mann-Whitney U -test; PV interneuron, $n = 12$ cells; PN, $n = 12$ cells. Effect of cell type (paired): $U = 108.5$, $P = 0.88$, two-sided Mann-Whitney U -test; PV interneuron, $n = 15$ cells; PN, $n = 15$ cells. **d**, Amplitude of IPSCs in PV interneurons normalized to median values from PNs in the same slices. Effect of training: $\chi^2 = 20.63$ (2), $P = 3.31 \times 10^{-5}$, Kruskal-Wallis ANOVA. **e**, Relative strength of SST interneuron transmission onto PV interneurons versus PNs in conditioned mice. **f**, Interrogation of monosynaptic connections from PV interneurons onto SST interneurons and PNs. Flp_{ON}, Cre_{OFF} INTRASECT Chr2 vector into PV-IRES-FlpO/SST-IRES-Cre/Ai9 triple-transgenic mice. **g**, Example IPSC traces. Scale bars, 200 pA \times 1 s. **h**, Amplitude of IPSCs resulting from PV interneuron photoexcitation (460 nm, 1 ms pulse, 0.1 Hz) in slices from naïve mice (6 slices from 3 mice), as well as those that received unpaired (6 slices from 3 mice) or paired training (6 slices from 3 mice). Effect of cell type (naïve): $U = 6$, $P = 1.15 \times 10^{-5}$, two-sided Mann-Whitney U -test; SST interneuron, $n = 14$ cells; PN, $n = 15$ cells. Effect of cell type (unpaired): $U = 3$, $P = 1.04 \times 10^{-5}$, two-sided Mann-Whitney U -test; PV interneuron, $n = 12$ cells; PN, $n = 11$ cells. Effect of cell type (paired): $U = 6$, $P = 1.15 \times 10^{-5}$, two-sided Mann-Whitney U -test; PV interneuron, $n = 12$ cells; PN, $n = 13$ cells. **i**, Amplitude of IPSCs in PV interneurons normalized to median values from PNs in the same slices. Effect of training: $\chi^2 = 0.33$ (2), $P = 0.85$, Kruskal-Wallis ANOVA. **j**, Relative strength of PV interneuron transmission onto SST interneurons versus PNs in conditioned mice. * $P < 0.05$, *** $P < 0.001$ by Mann-Whitney U -test (**c**, **h**) or Dunn's post hoc test (**d**). The box plots depict the median (center line), mean (black box), quartiles and 10–90% range (whiskers). During optic stimulation, 100% of postsynaptic cells that were sampled in these analyses exhibited synaptic responses.

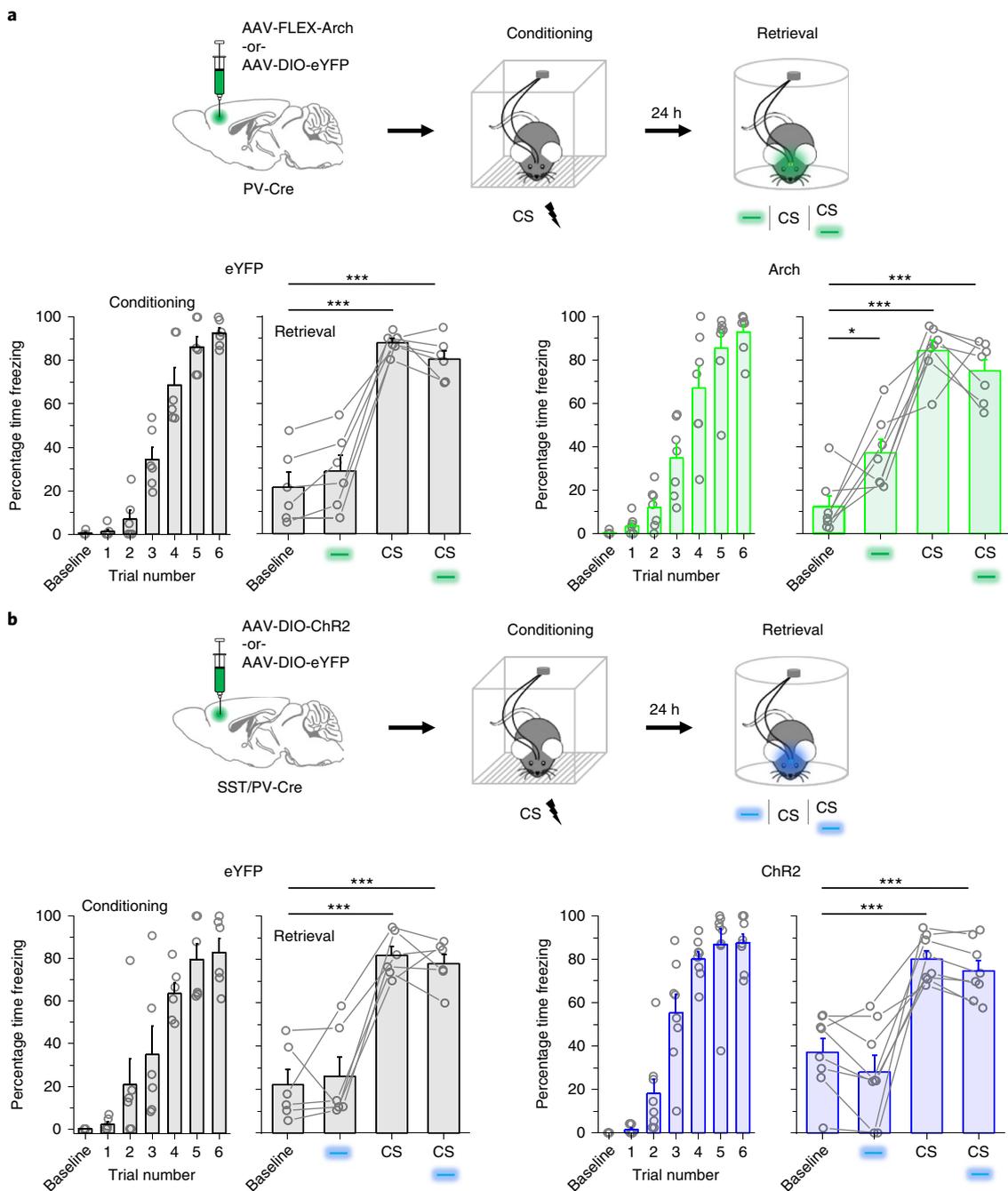


Fig. 6 | SST interneuron-evoked freezing requires suppression of PV interneurons. a,b, In vivo dissection of PV interneuron contributions to stimulus-dependent freezing by CS and light (532 nm, constant, 20 s) at 24 h after CS-US pairing in PV-IRES-Cre mice expressing conditional Arch (green) or eYFP control vectors (black). Arch: $F_{3,18} = 33.2$, $P = 1.52 \times 10^{-7}$, one-way repeated-measures ANOVA; $n = 6$ mice. eYFP: $F_{3,15} = 50.1$, $P = 4.74 \times 10^{-8}$, one-way repeated-measures ANOVA; $n = 7$ mice. The experiment was performed in two independent cohorts, and mice were pooled together for analysis. **b**, Modulation of freezing by CS and light (473 nm, 10 ms pulses, 20 Hz) 24 h after fear conditioning in SST-IRES-Cre/PV-IRES-Cre double-transgenic mice expressing conditional ChR2 (blue) or eYFP (black) control vectors. ChR2: $F_{3,21} = 23.3$, $P = 6.94 \times 10^{-7}$, one-way repeated-measures ANOVA; $n = 8$ mice. eYFP: $F_{3,15} = 41.3$, $P = 1.74 \times 10^{-7}$, one-way repeated-measures ANOVA; $n = 6$ mice. The experiment was performed in two independent cohorts and mice were pooled together for analysis. * $P < 0.05$, *** $P < 0.001$ by Tukey's post hoc test. The bar graphs depict the mean \pm s.e.m.

of SST-IRES-FlpO/PV-IRES-Cre/Ai9 triple transgenic mice, permitting independent tagging of SST (ChR2) and PV interneurons (tdTomato) (Fig. 5a). We then examined the inhibitory responses to SST interneuron photoexcitation in PV interneurons as well as PNs from the same brain slices to control for group differences attributable to viral expression. In both naïve and unpaired control mice, SST interneurons elicited monosynaptic inhibitory PSCs (IPSCs) in

PV interneurons that were less potent than those recorded in surrounding PNs (Fig. 5b,c and Supplementary Fig. 11). In contrast, after CS-US pairing, responses to SST interneuron photoexcitation in these cell types were similar in amplitude. Comparison of SST interneuron-evoked responses in PV interneurons normalized to those in PNs confirmed that CS-US pairing increases the relative strength of SST interneuron \rightarrow PV interneuron transmission

(Fig. 5d). This suggests that learning shifts the balance of SST interneuron output to favor inhibition of PV interneurons, which may increase the likelihood of SST interneuron-evoked disinhibition.

Next, we utilized a similar genetic approach to interrogate PV interneuron transmission onto SST interneurons and PNs in pre- limbic L2/3 (Fig. 5f). Strikingly, this revealed that regardless of the training condition, PV interneurons elicit IPSCs that are approximately tenfold larger in amplitude in PNs compared to surrounding SST interneurons (Fig. 5g,h). Comparison of SST interneuron responses normalized to those in PNs revealed no effect of training on the balance of transmission (Fig. 5i). These results indicate the presence of a strong bias in PV interneuron output that potentially favors preferential control of PN over SST interneuron firing (Fig. 5j). In contrast, SST interneurons exhibit a much weaker bias (approximately twofold) for PNs over PV interneurons and this bias is completely eliminated by conditioning (Fig. 5e). This implies that SST interneurons in the pre- limbic cortex could have a unique capacity to evoke PN disinhibition.

Prelimbic circuit organization suggests that SST interneurons might interact directly with PV interneurons to mediate fear expression through PN disinhibition. As an *in vivo* test of this model, we first sought to confirm that photoinhibition of PV interneurons elicits freezing, as reported previously⁴. Indeed, after fear conditioning, PV interneuron photoinhibition resulted in an increase in freezing during light-only trials (Fig. 6a and Supplementary Fig. 12). We then tested whether inhibition of PV interneurons is required specifically for SST interneuron-evoked freezing. In contrast to SST interneuron-specific manipulations (Fig. 3c), concurrent photoactivation of SST and PV interneurons in SST-Cre/PV-Cre double transgenic mice negated the fear-promoting effect of SST interneuron activity (Fig. 6b). These results imply that interaction between SST and PV interneurons is important for processing the behavioral output of SST interneuron activity.

BLA afferent connectivity of pre- limbic interneurons. Having established that SST interneurons directly inhibit PV interneurons, we next considered whether these interneuron populations are engaged by long-range inputs to the pre- limbic cortex. In addition to modulating plasticity of pre- limbic SST interneurons (Fig. 4), BLA contains pre- limbic cortex PNs that exhibit increased firing during memory retrieval^{25,26} and regulate fear memory expression^{27,28}. Projections from these cells primarily target L2/3 and are thus well positioned to recruit potentiated SST interneurons²⁹. To test whether pre- limbic interneurons are directly modulated by these projections, we infused a CaMKII-driven ChR2 vector (AAV-CaMKII-hChR2-

eYFP; CaMKII-hChR2) into BLA (Fig. 7a), leading to axonal ChR2 accumulation in pre- limbic L2/3 of SST- (Fig. 7b) and PV-IRES-Cre/Ai9 mice (Fig. 7g). Optic stimulation of these projections in naïve mice elicited compound EPSCs and feedforward IPSCs in both SST and PV interneurons, as well as surrounding PNs (Supplementary Fig. 13). Because much of this transmission occurred at long latencies after stimulation, recurrent activity within pre- limbic circuits is probably responsible for its generation. Interestingly, compared to PNs from the same brain slices, SST but not PV interneurons exhibited a higher ratio of excitatory-to-inhibitory charge during these events (Supplementary Fig. 13c). This could be attributed to less potent network inhibition of SST interneurons, since inhibitory charge in SST interneurons was lower than in surrounding PNs (Supplementary Fig. 13b).

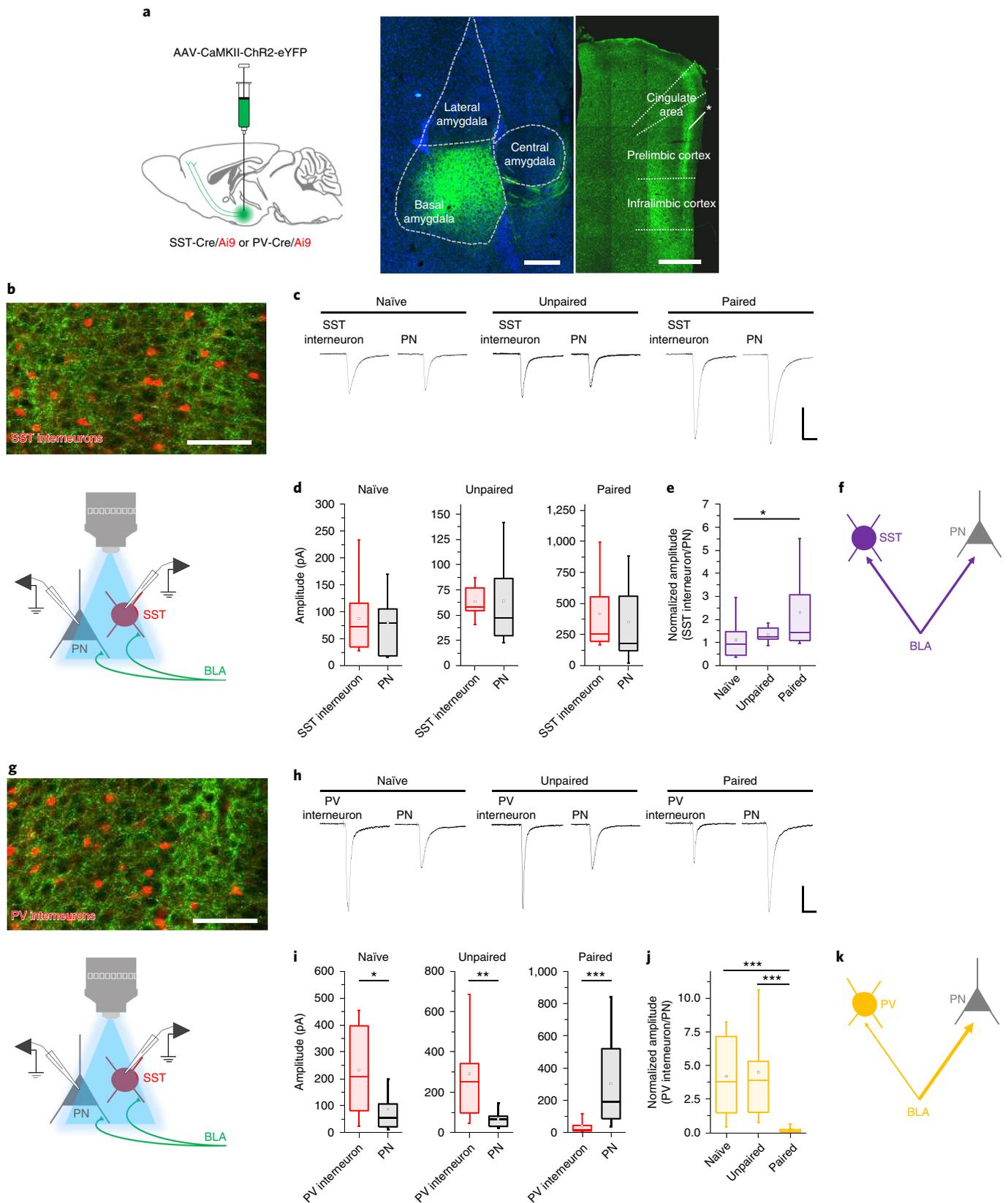
While these results are intriguing, the complexity of BLA-evoked activity prevented the analysis of monosynaptic transmission at connections between BLA axons and pre- limbic interneurons. Therefore, we applied a pharmacological cocktail to eliminate action potential propagation and prevent polysynaptic transmission^{12,13,30}. When using this approach in SST-Cre/Ai9 mice, we found that regardless of training condition, BLA afferents evoke responses of similar amplitude in SST interneurons compared to PNs (Fig. 7c,d and Supplementary Fig. 14). However, when SST interneuron responses were normalized to those of PNs, this revealed a slightly higher ratio of SST interneuron/PN transmission in paired mice compared to naïve controls (Fig. 7e). In contrast to these results, BLA terminal stimulation in naïve and unpaired PV-Cre/Ai9 mice evoked EPSCs that were larger in amplitude in PV interneurons compared to PNs (Fig. 7h–j). This is in agreement with similar experiments that examined the potency of BLA transmission onto PV interneurons in the infralimbic cortex³¹. However, in animals that received CS–US pairing, the relative strength of BLA transmission in PV interneurons and PNs was effectively reversed (Fig. 7j). These data collectively imply that, following conditioning, SST interneurons are probably strongly activated by BLA afferents and that circuit plasticity may favor their recruitment over PV interneurons (Fig. 7f,k).

Network disinhibition underlies SST interneuron-evoked fear expression. Because the balance of ongoing excitatory and inhibitory transmission determines the firing rate of excitatory PNs, the activity of pre- limbic output neurons could be modulated solely through the relief of PV interneuron-mediated inhibition. Therefore, to reveal the extent to which SST interneuron activity disinhibits pre- limbic networks, we conducted immunolabeling for the activity marker *c-Fos* in pre- limbic cortex and potential

Fig. 7 | Relative strength of BLA transmission onto pre- limbic interneurons and PNs is modulated by learning. **a**, To examine pre- limbic neuronal responses to BLA afferent stimulation, a CaMKII promoter-dependent ChR2 vector was infused into the BLA (left) of SST- or PV-IRES-Cre/Ai9 mice, leading to projection-specific expression of ChR2-eYFP in pre- limbic L2/3 (right). The asterisk indicates pre- limbic ChR2 accumulation. Scale bars, 250 μ m (left) and 500 μ m (right). At 24 h after CS–US pairing, recordings were obtained from Tomato⁺ interneurons and surrounding PNs during BLA afferent photoexcitation (460 nm, 1 ms pulse, 0.1 Hz). **b**, Axonal ChR2 accumulation in the L2/3 of SST-IRES-Cre/Ai9 mice and recording configuration. **c**, Example EPSC traces. Scale bars, 100 pA \times 50 ms. **d**, Amplitude of BLA-evoked SST interneuron and PN EPSCs in slices from naïve mice (8 slices from 4 mice), as well as those that received unpaired (6 slices from 3 mice) or paired training (8 slices from 4 mice). Effect of cell type (naïve): $U = 134$, $P = 0.51$, two-sided Mann–Whitney U -test; SST interneuron, $n = 18$ cells; PN, $n = 13$ cells. Effect of cell type (unpaired): $U = 106$, $P = 0.49$, two-sided Mann–Whitney U -test; SST interneuron, $n = 13$ cells, PN, $n = 14$ cells. Effect of cell type (paired): $U = 107$, $P = 0.37$, two-sided Mann–Whitney U -test; PV interneuron, $n = 16$ cells; PN, $n = 11$ cells. **e**, Amplitude of EPSCs in SST interneurons normalized to median values from PNs in the same slices. Effect of training: $\chi^2 = 7.2$ (2), $P = 0.027$, Kruskal–Wallis ANOVA. **f**, Relative strength of BLA transmission in conditioned mice. **g**, Axonal ChR2 accumulation in L2/3 of PV-IRES-Cre/Ai9 mice. **h**, Example EPSC traces. Scale bars, 100 pA \times 50 ms. **i**, Amplitude of BLA-evoked PV interneuron and PN EPSCs in slices from naïve mice (6 slices from 3 mice), as well as those that received unpaired (6 slices from 3 mice) or paired training (8 slices from 4 mice). Effect of cell type (naïve): $U = 118$, $P = 0.025$, two-sided Mann–Whitney U -test; PV interneuron, $n = 14$ cells; PN, $n = 11$ cells. Effect of cell type (unpaired): $U = 85$, $P = 0.0066$, two-sided Mann–Whitney U -test; PV interneuron, $n = 10$ cells, PN, $n = 10$ cells. Effect of cell type (paired): $U = 19$, $P = 7.56 \times 10^{-4}$, two-sided Mann–Whitney U -test; PV interneuron, $n = 13$ cells; PN, $n = 12$ cells. **j**, Amplitude of EPSCs in PV interneurons normalized to median values from PNs in the same slices. Effect of training: $\chi^2 = 21.6$ (2), $P = 2.04 \times 10^{-5}$, Kruskal–Wallis ANOVA. **k**, Relative strength of SST interneuron transmission in conditioned mice. $*P < 0.05$, $**P < 0.01$, $***P < 0.001$ by two-sided Mann–Whitney U -test (**i**) or Dunn’s post hoc test (**e,j**). The box plots depict the median (center line), mean (black box), quartiles and 10–90% range (whiskers).

downstream brain regions following SST interneuron photoexcitation in the absence of any CS exposure 24 h after conditioning (Fig. 8a and Supplementary Fig. 15). Similar to a previous experiment (Fig. 3), ChR2-expressing mice but not eYFP controls exhibited an increase in freezing in response to photoexcitation (Supplementary

Fig. 15). After behavioral testing, ChR2-expressing mice exhibited higher c-Fos labeling of SST interneurons compared to eYFP controls, as well as higher c-Fos labeling of surrounding eYFP⁻ cells, consistent with disinhibition of other prelimbic cell types (Fig. 8b,d). When quantification was extended to downstream targets of the



prelimbic cortex, higher numbers of c-Fos⁺ cells were also detected in the BLA, paraventricular thalamus, lateral habenula, ventrolateral periaqueductal gray and dorsomedial hypothalamus (Fig. 8c,e). However, several other regions including the nucleus accumbens, caudate putamen, ventral hippocampus (area CA1), dentate gyrus and mediodorsal thalamus were unaffected by photostimulation (Fig. 8e and Supplementary Fig. 15).

To test whether regional c-Fos induction by SST interneuron photoexcitation occurs independently of fear expression, we next quantified c-Fos expression in a randomly selected subset of naïve mice that received optogenetic manipulation of prelimbic SST interneurons without prior fear conditioning (Fig. 3f). Consistent with the larger group, these mice exhibited no increase in freezing over baseline levels during photostimulation (Fig. 8f and Supplementary Fig. 16). Examination of stimulated prelimbic tissue confirmed that consistent with photoactivation, higher c-Fos labeling was present in eYFP⁺ cells of Chr2-expressing relative to eYFP-expressing mice (Fig. 8g and Supplementary Fig. 16). However, there was no group difference in c-Fos expression in surrounding eYFP⁻ cells, indicating that in contrast to animals that received CS-US pairing (Fig. 8b,d), SST interneuron photoexcitation in naïve mice does not activate surrounding prelimbic neurons to a notable degree. In addition, remote brain regions that were modulated by photoexcitation in conditioned mice did not exhibit any group differences in the number of c-Fos⁺ neurons (Fig. 8h and Supplementary Fig. 16). Thus, acquisition of SST interneuron-evoked freezing correlates with a change in SST interneuron recruitment of a specific brain network, including prelimbic neurons indirectly activated by SST interneuron photoexcitation, presumably via disinhibition.

Finally, to test whether the network-level effects of memory retrieval resemble those evoked by SST interneuron photoexcitation in conditioned mice, we performed c-Fos analysis following CS exposure (Fig. 8i). Presentation of 4 CS trials elicited increased freezing in mice that received CS-US pairing 24 h before the memory retrieval

test, but not in nonconditioned controls (Supplementary Fig. 17). After CS exposure, a higher number of c-Fos⁺ cells was observed in conditioned relative to nonconditioned mice in the majority (5 out of 6) of brain regions that were modulated by SST interneuron photoexcitation (Fig. 8j). The remaining region (ventrolateral periaqueductal gray (vIPAG)) showed a trend toward higher c-Fos labeling in conditioned mice ($P=0.082$). Conversely, areas where c-Fos immunoreactivity was unaffected by SST interneuron photoexcitation also exhibited no differences in c-Fos⁺ cells following CS-evoked memory retrieval. Together these data argue against the notion that network c-Fos induction by photostimulation results from nonphysiological activity patterns and suggest that CS recruitment of SST interneurons mediates disinhibition of prelimbic outputs to remote brain regions underlying memory expression.

Discussion

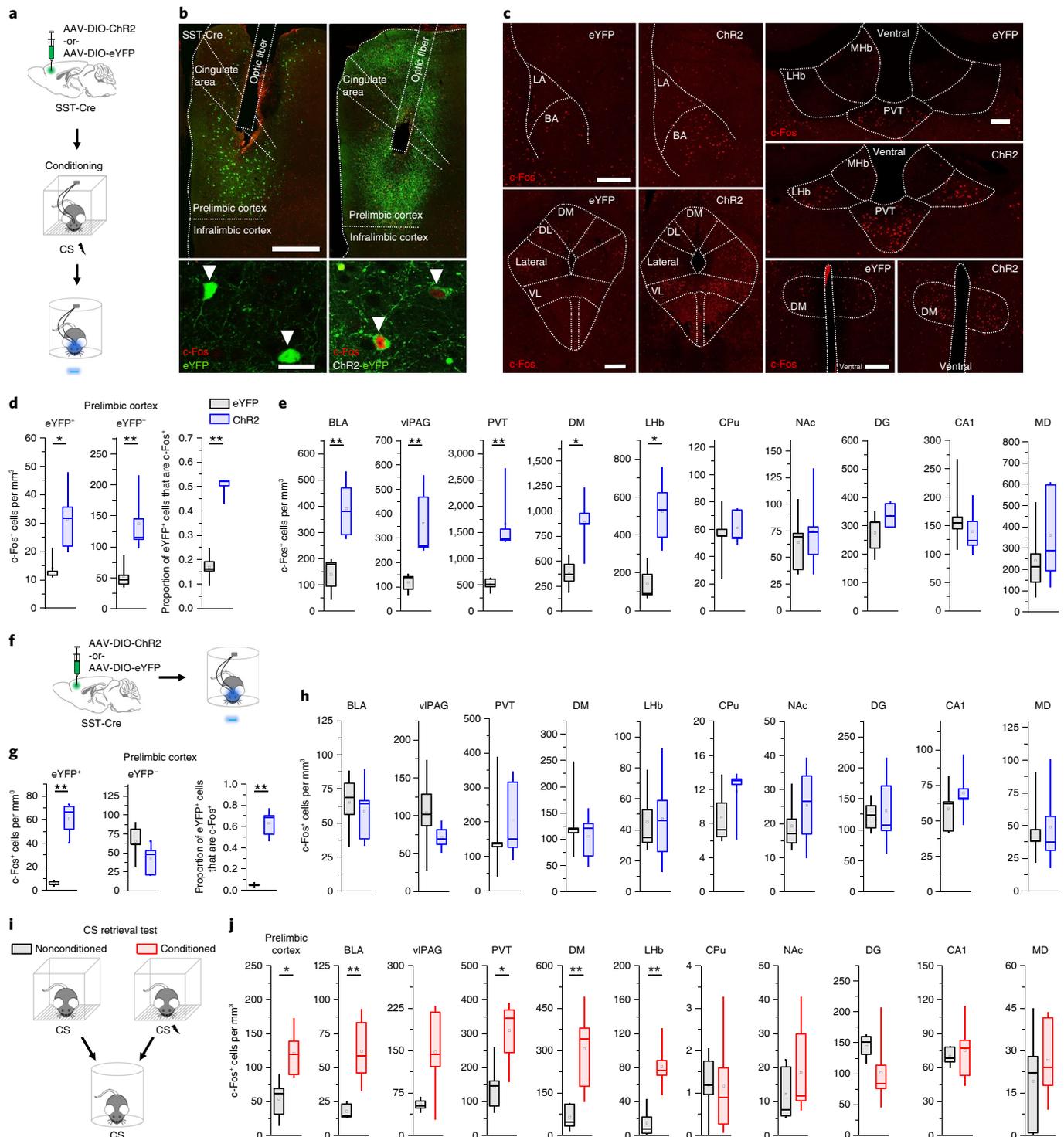
In this study, we demonstrate that associative fear conditioning potentiates the function of prefrontal SST interneurons at the level of synaptic transmission and in vivo activity. Correlated with memory acquisition, SST interneurons exhibit an experience-dependent increase in CS-evoked firing and specifically signal the learned CS rather than the general fear state of the animal. Activation of SST interneurons in turn plays a causal role in both the expression and initial acquisition of cued fear responses. The paradoxical role of SST interneurons in memory expression can be explained by their encoding of cue-related disinhibition of prelimbic PNs, which are in turn responsible for the recruitment of a distributed brain network for defensive responding. Although our results do not exclude the possibility of a similar role in context-evoked fear, SST interneuron activity was not modulated by the conditioning context (Supplementary Fig. 4) and there was no baseline effect of SST interneuron photoinhibition on generalized context freezing (Fig. 3), suggesting that SST interneurons preferentially mediate the encoding of cue associations.

Fig. 8 | SST interneuron activation recruits a specific brain network in conditioned mice. **a**, Preparation of Chr2 ($n=5$) and eYFP mice ($n=5$) used for c-Fos activity mapping after SST interneuron photoexcitation. All mice underwent CS-US pairing 24 h before stimulation. **b**, Induction of c-Fos by photoexcitation (6 trials, 473 nm, 10 ms pulses, 20 Hz) in the prelimbic cortex. The insets (bottom) depict eYFP/c-Fos double-labeled SST interneurons. The arrowheads denote SST interneurons. Scale bars, 500 μm and 50 μm , respectively. **c**, Elevated c-Fos expression in selected brain regions after optic stimulation of conditioned mice. BA, basal amygdala; DL, dorsolateral; DM, dorsomedial (hypothalamus); LA, lateral amygdala; MHB, medial habenula; LHb, lateral habenula; PVT, paraventricular thalamus; VL, ventrolateral (periaqueductal gray). Scale bar, 200 μm . **d**, Comparison of c-Fos⁺ cell counts in the prelimbic cortex of stimulated Chr2-eYFP (blue; $n=5$ mice) versus eYFP-only (black; $n=5$ mice) mice. In both groups, eYFP⁺ cells represent transfected SST interneurons, while eYFP⁻ cells represent nonexpressing neighbors. Group differences established by two-sided Mann-Whitney *U*-test and controlled for false discovery rate (type 1 error) by the Benjamini-Hochberg method. eYFP⁺ cells: $U=1$, $P=0.016$. eYFP⁻ cells: $U=0$, $P=0.0079$. Proportion of eYFP⁺ cells that are c-Fos⁺: $U=0$, $P=0.0079$. **e**, Comparison of c-Fos⁺ cell counts for all brain regions ($n=5$ mice per eYFP or Chr2 group, 2 slices per brain region) using the same statistical procedures (two-sided Mann-Whitney *U*-test) as in **c**. BLA: $U=0$, $P=0.0079$. ventrolateral periaqueductal gray (vIPAG): $U=0$, $P=0.0079$. PVT: $U=0$, $P=0.0079$. DM: $U=1$, $P=0.0022$. LHb: $U=0$, $P=0.0012$. Caudate putamen (CPu): $U=14$, $P=0.84$. Nucleus accumbens (NAc): $U=10$, $P=0.69$. Dentate gyrus (DG) of the dorsal hippocampus: $U=7$, $P=0.29$. Cornu ammonis area 1 (CA1) of the ventral hippocampus: $U=16$, $P=0.55$. Mediodorsal thalamus (MD): $U=8$, $P=0.42$. The experiment was conducted in two independent cohorts and the results were pooled together. c-Fos counts for each mouse are reported as the average of two representative brain slices per mouse. Tissues quantified in **b-e** are from the same mice. **f**, Preparation of Chr2 ($n=5$) and eYFP mice ($n=5$ mice) used for c-Fos activity mapping after SST interneuron photoexcitation in naïve (unconditioned) mice. **g**, Comparison of c-Fos⁺ cell counts in the prelimbic cortex of stimulated Chr2-eYFP (blue, $n=5$ mice) versus eYFP-only (black, $n=5$ mice) naïve mice (two-sided Mann-Whitney *U*-test). eYFP⁺ cells: $U=0$, $P=0.0079$. eYFP⁻ cells: $U=20$, $P=0.15$. Proportion eYFP⁺ cells that are c-Fos⁺: $U=0$, $P=0.0079$. **h**, Comparison of c-Fos⁺ cell counts in selected brain regions of stimulated Chr2-eYFP (blue, $n=5$ mice) and eYFP-only (black, $n=5$ mice) naïve mice; two-sided Mann-Whitney *U*-test. BLA: $U=14$, $P=0.84$. vIPAG: $U=19$, $P=0.22$. PVT: $U=11$, $P=0.84$. DM: $U=13$, $P=1$. LHb: $U=13$, $P=1$. CPu: $U=7$, $P=0.31$. NAc: $U=9$, $P=0.55$. DG: $U=12$, $P=1$. CA1: $U=6$, $P=0.22$. MD: $U=14$, $P=0.84$. The experiment was conducted in two independent cohorts, and the results were pooled together. c-Fos counts for each mouse are reported as the average of two representative brain slices per mouse. The tissues quantified in **g-h** are from the same mice. **i**, Preparation of SST-IRES-Cre mice used for c-Fos activity mapping after CS exposure (4 trials, 2 kHz, 80 dB, 20 s) in mice that were conditioned 24 h prior ($n=6$) or in those that did not receive conditioning ($n=5$). **j**, Comparison of c-Fos⁺ cell counts in selected brain regions of conditioned ($n=6$ mice) versus nonconditioned ($n=5$ mice) mice, as conducted for photostimulated mice (**e,h**). Prelimbic cortex: $U=2$, $P=0.0017$. BLA: $U=0$, $P=0.0043$. vIPAG: $U=5$, $P=0.082$. PVT: $U=3$, $P=0.030$. DM: $U=0$, $P=0.0043$. LHb: $U=0$, $P=0.0043$. CPu: $U=17$, $P=0.79$. NAc: $U=9$, $P=0.33$. DG: $U=25$, $P=0.082$. CA1: $U=12$, $P=0.66$. MD: $U=12$, $P=0.66$. The experiment was conducted in two independent cohorts and the results were pooled together. c-Fos counts for each mouse are reported as the average of two representative brain slices per mouse. The tissues quantified in **i-j** are from the same mice. * $P < 0.05$, ** $P < 0.01$ by two-sided Mann-Whitney *U*-test. The box plots depict the median (center line), mean (black box), quartiles and 10–90% range (whiskers).

In network disinhibition, prelimbic circuits formed by SST and PV interneurons exhibit functional differences that may facilitate their complementary roles. In particular, output from PV interneurons is heavily biased toward PNs over SST interneurons (Fig. 5), implying that they are specialized for the suppression of PN firing. In contrast, there is a greater potential for PN disinhibition during SST interneuron activity owing to their relatively strong inhibition of PV interneurons. Accordingly, photoactivation of SST interneurons indirectly activates not only surrounding prelimbic cells but also remote brain regions that receive excitatory connections from

the mPFC (Fig. 8). A learning-dependent shift in SST interneuron output could in part explain why these effects are specific to conditioned mice. Alternatively, plasticity in other circuits could prime this network for SST interneuron-evoked disinhibition.

Although all prelimbic cell types we examined receive direct input from BLA projections, it is notable that these afferents evoke complex synaptic activity with an overall higher ratio of excitatory:inhibitory transmission in SST interneurons (Supplementary Fig. 13). In conditioned mice, PV interneurons but not SST interneurons exhibited a lower strength of monosynaptic BLA transmission compared to



neighboring PNs (Fig. 7). Given these findings, it seems possible that, after learning, BLA afferent activity favors the recruitment of SST interneurons over PV interneurons, and that the resulting PN disinhibition could underlie the freezing behavior elicited by BLA afferent stimulation in conditioned mice²⁷. In addition, BLA projections could also mediate the effect of amygdala activity on learning-related SST interneuron plasticity (Fig. 4). Further interrogation of synaptic and functional plasticity throughout this circuitry is necessary to establish the degree to which other cell types, and projection-specific populations, cooperate with SST interneurons in memory encoding and retrieval.

An important implication of SST interneuron plasticity is that phasic disinhibition, which has been established primarily as a mechanism for cortical processing of behavioral feedback signals^{23,32–35}, can be encoded by potentiation of GABAergic transmission to mediate novel cue associations. This means that both activated (for example, SST interneurons) and inhibited interneuron populations (for example, PV interneurons) can play a role in memory expression and exhibit cue-evoked activity patterns that are causally linked. In this way, the modification of a sparse interneuron population can extensively reorganize stimulus processing in a broader local network. Indeed, previous findings confirm that firing rate inhibition of PV interneurons is a potent regulator of PN recruitment, synchrony and synaptic plasticity^{4,8,23,32,35}. Nevertheless, it is unclear why SST interneuron-evoked disinhibition triggers such selective changes in behavior and circuit activity. One possibility is that different subnetworks of prelimbic PNs, for example, populations with discrete projection patterns, vary in the level of inhibition they receive from SST versus PV interneurons. The balance of ongoing activity from these inhibitory sources could in part determine whether a given cell is predominantly activated, inhibited or unaffected by cue-responsive SST interneurons.

Several brain regions activated by prelimbic SST interneuron stimulation, including the BLA¹⁵, paraventricular thalamus^{36,37}, lateral habenula³⁸, ventrolateral periaqueductal gray³⁹ and dorsomedial hypothalamus⁴⁰, have established roles in mediating defensive behavioral responses to CS and US. It is possible that PNs that give rise to direct projections to these regions from the prelimbic cortex are preferentially disinhibited during SST interneuron firing. Alternatively, this pattern of network activity may be completed by relay circuits from other downstream effectors. Intriguingly, however, neither the striatum nor the mediodorsal thalamus, which are among the areas receiving the highest density of input from the prelimbic cortex^{41,42}, are activated during the expression of SST interneuron-evoked freezing. This implies a high level of specificity in the computational logic of prelimbic microcircuits.

While our *in vivo* manipulations were targeted to SST interneurons as a whole, SST expression demarcates a cell population with heterogeneous protein expression, firing patterns and morphology⁴³, whose function is additionally determined by their laminar location⁴⁴. Our results imply that SST interneurons that encode fear-related disinhibition reside in the superficial layers of the prelimbic cortex, but further parcellation of this cell class could reveal functionally discrete subpopulations. For example, previous work has identified a subset of SST interneurons that express the oxytocin receptor and contribute to the regulation of anxiety and social behavior in a sex-dependent manner^{45,46}. At present, it is unclear what features of an aversive experience might be encoded by SST interneurons or how this information is represented by specific SST interneuron populations. Future studies could resolve the function of individual neurons through the use of genetic markers, cellular tags and *in vivo* imaging. In addition, although circuit disinhibition can be explained by fast GABAergic transmission, it is also important to consider that SST is not just a marker of interneurons but also a peptide transmitter that can influence memory acquisition⁴⁷ and recall⁴⁸. The release of this peptide from

dense-core vesicles may contribute to the memory function of potentiated SST interneurons.

In conclusion, our results outline an important causal role for inhibitory signaling in associative memory. In pursuit of memory engrams, it is therefore critical to consider the contributions of interneurons that, despite their inhibitory output, have the capacity to encode and reactivate a specific pattern of excitatory neuronal activity. Interrogation of these cells and their associated circuitry could reveal important computational principles for the storage, consolidation and retrieval of information. Indeed, during the preparation of our revised manuscript, a new report demonstrated an important causal role for prefrontal SST interneuron activity in learned social fear⁴⁹.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information, details of author contributions and competing interests, and statements of data and code availability are available at <https://doi.org/10.1038/s41593-019-0552-7>.

Received: 7 November 2018; Accepted: 31 October 2019;

Published online: 16 December 2019

References

- Holtmaat, A. & Caroni, P. Functional and structural underpinnings of neuronal assembly formation in learning. *Nat. Neurosci.* **19**, 1553–1562 (2016).
- Kandel, E. R., Dudai, Y. & Mayford, M. R. The molecular and systems biology of memory. *Cell* **157**, 163–186 (2014).
- Josselyn, S. A., Köhler, S. & Frankland, P. W. Finding the engram. *Nat. Rev. Neurosci.* **16**, 521–534 (2015).
- Courtin, J. et al. Prefrontal parvalbumin interneurons shape neuronal activity to drive fear expression. *Nature* **505**, 92–96 (2014).
- Rashid, A. J. et al. Competition between engrams influences fear memory formation and recall. *Science* **353**, 383–387 (2016).
- Siwani, S. et al. OLMα2 cells bidirectionally modulate learning. *Neuron* **99**, 404–412.e3 (2018).
- Stefanelli, T., Bertolini, C., Lüscher, C., Müller, D. & Mendez, P. Hippocampal somatostatin interneurons control the size of neuronal memory ensembles. *Neuron* **89**, 1074–1085 (2016).
- Wolff, S. B. et al. Amygdala interneuron subtypes control fear learning through disinhibition. *Nature* **509**, 453–458 (2014).
- Corcoran, K. A. & Quirk, G. J. Activity in prelimbic cortex is necessary for the expression of learned, but not innate, fears. *J. Neurosci.* **27**, 840–844 (2007).
- Rudy, B., Fishell, G., Lee, S. & Hjerling-Lefler, J. Three groups of interneurons account for nearly 100% of neocortical GABAergic neurons. *Dev. Neurobiol.* **71**, 45–61 (2011).
- Kim, D. et al. Distinct roles of parvalbumin- and somatostatin-expressing interneurons in working memory. *Neuron* **92**, 902–915 (2016).
- Arruda-Carvalho, M. & Clem, R. L. Pathway-selective adjustment of prefrontal-amygdala transmission during fear encoding. *J. Neurosci.* **34**, 15601–15609 (2014).
- Lucas, E. K., Jegarl, A. M., Morishita, H. & Clem, R. L. Multimodal and site-specific plasticity of amygdala parvalbumin interneurons after fear learning. *Neuron* **91**, 629–643 (2016).
- Graziane, N. & Dong, Y. *Electrophysiological Analysis of Synaptic Transmission* (Humana Press, 2016).
- Herry, C. & Johansen, J. P. Encoding of fear learning and memory in distributed neuronal circuits. *Nat. Neurosci.* **17**, 1644–1654 (2014).
- Amano, T., Unal, C. T. & Paré, D. Synaptic correlates of fear extinction in the amygdala. *Nat. Neurosci.* **13**, 489–494 (2010).
- Asede, D., Bosch, D., Lüthi, A., Ferraguti, F. & Ehrlich, I. Sensory inputs to intercalated cells provide fear-learning modulated inhibition to the basolateral amygdala. *Neuron* **86**, 541–554 (2015).
- Namburi, P. et al. A circuit mechanism for differentiating positive and negative associations. *Nature* **520**, 675–678 (2015).
- Pattwell, S. S., Bath, K. G., Casey, B. J., Ninan, I. & Lee, F. S. Selective early-acquired fear memories undergo temporary suppression during adolescence. *Proc. Natl Acad. Sci. USA* **108**, 1182–1187 (2011).
- Tsvetkov, E., Carlezon, W. A., Benes, F. M., Kandel, E. R. & Bolshakov, V. Y. Fear conditioning occludes LTP-induced presynaptic enhancement of synaptic transmission in the cortical pathway to the lateral amygdala. *Neuron* **34**, 289–300 (2002).

21. Zhou, Y. et al. CREB regulates excitability and the allocation of memory to subsets of neurons in the amygdala. *Nat. Neurosci.* **12**, 1438–1443 (2009).
22. Clem, R. L. & Huganir, R. L. Calcium-permeable AMPA receptor dynamics mediate fear memory erasure. *Science* **330**, 1108–1112 (2010).
23. Pi, H. J. et al. Cortical interneurons that specialize in disinhibitory control. *Nature* **503**, 521–524 (2013).
24. Fenno, L. E. et al. Targeting cells with single vectors using multiple-feature Boolean logic. *Nat. Methods* **11**, 763–772 (2014).
25. Senn, V. et al. Long-range connectivity defines behavioral specificity of amygdala neurons. *Neuron* **81**, 428–437 (2014).
26. Herry, C. et al. Switching on and off fear by distinct neuronal circuits. *Nature* **454**, 600–606 (2008).
27. Burgos-Robles, A. et al. Amygdala inputs to prefrontal cortex guide behavior amid conflicting cues of reward and punishment. *Nat. Neurosci.* **20**, 824–835 (2017).
28. Klavir, O., Prigge, M., Sarel, A., Paz, R. & Yizhar, O. Manipulating fear associations via optogenetic modulation of amygdala inputs to prefrontal cortex. *Nat. Neurosci.* **20**, 836–844 (2017).
29. Little, J. P. & Carter, A. G. Synaptic mechanisms underlying strong reciprocal connectivity between the medial prefrontal cortex and basolateral amygdala. *J. Neurosci.* **33**, 15333–15342 (2013).
30. Cruikshank, S. J., Urabe, H., Nurmikko, A. V. & Connors, B. W. Pathway-specific feedforward circuits between thalamus and neocortex revealed by selective optical stimulation of axons. *Neuron* **65**, 230–245 (2010).
31. McGarry, L. M. & Carter, A. G. Inhibitory gating of basolateral amygdala inputs to the prefrontal cortex. *J. Neurosci.* **36**, 9391–9406 (2016).
32. Lee, S., Kruglikov, I., Huang, Z. J., Fishell, G. & Rudy, B. A disinhibitory circuit mediates motor integration in the somatosensory cortex. *Nat. Neurosci.* **16**, 1662–1670 (2013).
33. Garcia-Junco-Clemente, P. et al. An inhibitory pull–push circuit in frontal cortex. *Nat. Neurosci.* **20**, 389–392 (2017).
34. Ibrahim, L. A. et al. Cross-modality sharpening of visual cortical processing through layer-1-mediated inhibition and disinhibition. *Neuron* **89**, 1031–1045 (2016).
35. Letzkus, J. J. et al. A disinhibitory microcircuit for associative fear learning in the auditory cortex. *Nature* **480**, 331–335 (2011).
36. Beas, B. S. et al. The locus coeruleus drives disinhibition in the midline thalamus via a dopaminergic mechanism. *Nat. Neurosci.* **21**, 963–973 (2018).
37. Do-Monte, F. H., Quiñones-Laracuent, K. & Quirk, G. J. A temporal shift in the circuits mediating retrieval of fear memory. *Nature* **519**, 460–463 (2015).
38. Stamatakis, A. M. & Stuber, G. D. Activation of lateral habenula inputs to the ventral midbrain promotes behavioral avoidance. *Nat. Neurosci.* **15**, 1105–1107 (2012).
39. Tovote, P. et al. Midbrain circuits for defensive behaviour. *Nature* **534**, 206–212 (2016).
40. Johnson, P. L. & Shekhar, A. Panic-prone state induced in rats with GABA dysfunction in the dorsomedial hypothalamus is mediated by NMDA receptors. *J. Neurosci.* **26**, 7093–7104 (2006).
41. Arruda-Carvalho, M., Wu, W. C., Cummings, K. A. & Clem, R. L. Optogenetic examination of prefrontal-amygdala synaptic development. *J. Neurosci.* **37**, 2976–2985 (2017).
42. Xu, W. & Südhof, T. C. A neural circuit for memory specificity and generalization. *Science* **339**, 1290–1295 (2013).
43. Yavorska, I. & Wehr, M. Somatostatin-expressing inhibitory interneurons in cortical circuits. *Front. Neural Circuits* **10**, 76 (2016).
44. Muñoz, W., Tremblay, R., Levenstein, D. & Rudy, B. Layer-specific modulation of neocortical dendritic inhibition during active wakefulness. *Science* **355**, 954–959 (2017).
45. Li, K., Nakajima, M., Ibañez-Tallon, I. & Heintz, N. A cortical circuit for sexually dimorphic oxytocin-dependent anxiety behaviors. *Cell* **167**, 60–72.e11 (2016).
46. Nakajima, M., Görlich, A. & Heintz, N. Oxytocin modulates female sociosexual behavior through a specific class of prefrontal cortical interneurons. *Cell* **159**, 295–305 (2014).
47. Kluge, C., Stoppel, C., Szinyei, C., Stork, O. & Pape, H. C. Role of the somatostatin system in contextual fear memory and hippocampal synaptic plasticity. *Learn. Mem.* **15**, 252–260 (2008).
48. Einstein, E. B. et al. Somatostatin signaling in neuronal cilia is critical for object recognition memory. *J. Neurosci.* **30**, 4306–4314 (2010).
49. Xu, H. et al. A disinhibitory microcircuit mediates conditioned social fear in the prefrontal cortex. *Neuron* **102**, 668–682.e5 (2019).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature America, Inc. 2019

Methods

Animals. All experimental procedures were approved by the Institutional Animal Care and Use Committee at the Icahn School of Medicine at Mount Sinai. Experiments were performed on male mice aged postnatal day 42 (P42)–60. Mice were acquired from The Jackson Laboratory and maintained in the C57Bl/6J background. The following genotypes were used: SST-IRES-Cre (stock no. 028864), PV-Cre (stock no. 017320), SST-IRES-Flp (stock no. 028579), PV-FlpO (stock no. 022730) and Ai9 (stock no. 007909). Mice were housed 2–5 per cage in a 12 h light–dark cycle with access to food and water *ad libitum*.

Stereotaxic vector infusion and optic fiber implantation for *in vivo* optogenetics.

Viral constructs were purchased from the University of Pennsylvania Vector Core or Addgene and included AAV1-EF1a-DIO-hChR2(H134R)-eYFP-WPRE (plasmid no. 20298; Addgene), AAV1-CBA-FLEX-Arch-GFP (plasmid no. 22222), AAV1-EF1a-DIO-eYFP-WPRE (plasmid no. 27056), AAV1-CaMKII α -ArchT-GFP (plasmid no. 99039; Addgene) and AAV1-CaMKII α -eYFP-WPRE (plasmid no. 105622). Optic fibers (200 μ m diameter; Thorlabs) were fixed inside ferrules (1.25 mm outer diameter, 230 μ m inner diameter; Precision Fiber Products) using heat-cured epoxy, cut to 2.5 mm length for mPFC or 5.0 mm length for BLA and polished using aluminum oxide lapping paper with incremental decreases in graininess (Thorlabs). Light transmittance was measured using a light intensity meter (Thorlabs) at the wavelengths used in each respective experiment. Stereotaxic surgeries were conducted at P45–49. Anesthesia was induced using 5% inhaled isoflurane vaporized in oxygen delivered at a rate of 1–1.5 min^{-1} . Mice were then mounted in stereotaxic frames and maintained at 1–1.5% isoflurane (Stoelting or Kopf). Viral constructs were injected bilaterally into the prelimbic cortex (300 nl; anteroposterior (AP): +1.9; dorsoventral (DV): –2.0; mediolateral (ML): ± 0.9 at a 10° angle) or BLA (250 nl; AP: –1.4; DV: –5.1; ML: ± 3.3) using motorized injectors (Stoelting and World Precision Instruments) at a rate of 100 nl min^{-1} . Following infusion, the needle was left in place for an additional 10 min before slow removal at a rate of 0.03 mm s^{-1} . For *in vivo* optogenetics, following bilateral injections, optic ferrules were bilaterally implanted and directed toward the prelimbic cortex (AP: +1.9; DV: –1.6; ML: ± 0.9 at a 10° angle) or BLA (AP: –1.4; DV: –4.8; ML: ± 3.3). Ferrules were fixed in place using C&B Metabond luting cement (Parkell) and dental cement. Postsurgical analgesia was achieved with banamine (2.5 mg kg^{-1}). Mice recovered in home cages for at least 1 week before experimental manipulation.

Stereotaxic vector infusion for electrophysiology. AAV1-CaMKII α -hChR2(H134R)-eYFP-WPRE (plasmid no. 26969; Addgene) was purchased from the University of Pennsylvania Vector Core. rAAVDJ/nEF-Cre_{OFF}/Flp_{ON}-hChR2(H134R)-eYFP was purchased from the University of North Carolina Gene Therapy Center Vector Core. Stereotaxic surgeries were done at P42–45 for AAV1-CaMKII α -hChR2(H134R)-eYFP-WPRE and from P25 to P28 for rAAVDJ-Cre_{OFF}/Flp_{ON}-hChR2(H134R)-eYFP. AAV1-CaMKII α -hChR2(H134R)-eYFP-WPRE and rAAVDJ-Cre_{OFF}/Flp_{ON}-hChR2(H134R)-eYFP were bilaterally infused into the BLA (250 nl; AP: –1.3; DV: –5.1; ML ± 3.3) or prelimbic cortex (400 nl; AP: +1.9; DV: –1.6; ML ± 0.2), respectively. Mice recovered for 10 d (AAV1-CaMKII α -eYFP-WPRE) or 25–27 d (rAAVDJ/nEF-Cre_{OFF}/Flp_{ON}-hChR2(H134R)-eYFP) before electrophysiological recordings, which we determined to be the incubation time required for effective expression of ChR2 by these viral serotypes.

Optogenetic behavioral manipulations. Mice were acclimated to handling and patch cord tethering for three consecutive days before behavioral testing. Handling consisted of 10 min handling sessions, followed by 15 min habituation to patch cords in a fresh cage, followed by 5 min of additional handling. Fear conditioning and retrieval tests were conducted in sound-attenuating chambers with automated stimulus delivery software (Med Associates). Each session began with a 200 s baseline period after which various stimuli were presented with a fixed 80 s interstimulus interval. Auditory fear conditioning entailed 6 pairings of an auditory tone (CS; 2 kHz, 80 dB, 20 s) with a coterminating footshock (US; 0.7 mA, 2 s). For experiments where optogenetic manipulations were performed during memory retrieval, animals were conditioned while connected to patch cords but no light stimuli were delivered; 24 h after conditioning, modulation of freezing behavior by CS and light stimuli were examined in a context distinct from the conditioning arena (context B). In the retrieval test, mice underwent two laser-only trials, followed by four CS presentations alternating with and without concurrent laser stimulation. The CS-only or CS + light presentation order was counterbalanced. The two laser-only, two CS-only and two CS + light trials were each averaged and used for subsequent analysis. In the experiments involving naïve mice, animals underwent two trials of laser stimulation in context B without prior fear conditioning. For trials including optogenetic stimulation, a final transmitted intensity of 7–9 mW for 473 nm (for ChR2; 20 Hz, 10 ms for 20 s epochs) or 563 nm (for Arch; constant light, 20 s) laser-generated light (Opto Engine) was used. Parameters for photoexcitation are consistent with other recent studies of cortical SST interneurons^{50,51}; the 20 Hz firing rate observed during *ex vivo* stimulation of SST interneurons (Supplementary Fig. 7) is within the endogenous range of activity in prelimbic SST interneurons⁵². In experiments where neurons were silenced during learning, light was presented simultaneously with each CS and extended 3 s after CS/US termination, after which it was manually ramped

down to 0 mW over a period of 3 s. Twenty-four hours later, mice were tethered to patch cords in context B and underwent four CS presentations in the absence of light stimulation. Mice were then used for electrophysiology recordings 24 h after retrieval. Behavior was recorded by video and scored off-line by a trained experimenter blind to the identity of the animal. Scoring by a second blinded experimenter was used to validate results. Fiber tip placement and viral expression were confirmed by fluorescence microscopy. Mice having misplaced fiber tips or lacking viral expression encompassing approximately 50% of the target structure, in either hemisphere, were excluded from the analysis by an investigator blinded to the experimental results.

Open field experiments. Naïve SST-IRES-Cre mice expressing ChR2 or eYFP in prefrontal SST interneurons were used for the open field experiments. Mice were habituated to the room for 30 min before the start of the experiment. Mice were tethered to patch cords and placed into the center of a 42 cm (long) \times 42 cm (wide) \times 30 cm (high) square arena equipped with 15 infrared beams/detectors on each side. Tests were 20 min long and consisted of alternating light on (473 nm, 7–9 mW delivered at 20 Hz, 10 ms pulses) and light off periods lasting 5 min each. Light on/off epochs were counterbalanced in both groups. Open field arenas were connected to a PC running the Fusion v.5.6 SuperFlex software, which was used to analyze infrared beam breaks and quantify locomotor parameters. Locomotor metrics are reported as an average of each of the two light on or two light off periods.

Fiber photometry surgery and calcium imaging. SST-IRES-Cre mice received unilateral infusion (400 nl; AP: +1.9; DV: –1.6; ML: ± 0.2) of AAV1-hSyn-FLEX-GCaMP6f-WPRE (plasmid no. 100833). Following viral infusion, imaging fibers manufactured from 400 μ m core 0.48 numerical aperture optic fiber fixed in a 2.5 mm diameter metal ferrule (2.0 mm length; Doric Lenses) were chronically implanted above the injection site in the prelimbic cortex (AP: +1.9; DV: –1.5; ML: ± 0.2). Fibers were fixed to the skull using luting and dental cement. Mice were returned to their home cage for 4 weeks before the start of imaging, which was determined to be the amount of time required to achieve maximal GCaMP6f expression in prefrontal SST interneurons.

Mice were habituated to handling (10 min) and patch cord tethering (15 min) for 3 consecutive days before the start of the imaging experiments. Additionally, calcium signals were streamed during handling and the habituation sessions to ensure robust and reproducible signals before the start of experiments. Mice were connected to an imaging patch cord (400 μ m core; 0.48 numerical aperture), which was connected to a 6-port fluorescence minicube (Doric Lenses). Blue (465 nm for GCaMP6f excitation) and violet (405 nm for control artifact fluorescence) light was transmitted into the brain at 20–80 μ W and kept constant across experimental sessions. Emitted light was passed through a dichroic mirror and a 500–540 nm filter before detection by a visible Femtowatt Photoreceiver (model 2151; Newport). Analog signals were recorded using an RZ5 processor and a PC equipped with the Synapse software (Tucker-Davis Technologies, version 90, build 90-39092P). All conditioning was performed in Med Associates' operant chambers. For paired fear conditioning experiments, mice were tethered to a patch cord and exposed to 4 CS tones (2 kHz, 80 dB, 20 s) during two test sessions at 24 h pre- and postconditioning. Paired conditioning consisted of 6 pairings of CS with a coterminating US (0.7 mA, 2 s).

For unpaired conditioning, mice were tethered to the patch cord for both CS and US presentations but were untethered and returned to the home cage for 15 min between sessions. After 24 and 48 h, mice were tethered to the patch cord and exposed to context B and context A, respectively. Calcium signals were sampled at 6 kHz and were continuously recorded throughout all tests, including conditioning, CS and context memory retrieval experiments. All imaging sessions started with a 2 min baseline period during which calcium signals were recorded before the start of the experiment. The start and end of each experimental session as well as the CS and US onset and offset timestamps were generated using transistor–transistor logic (TTL) signals triggered by the Med Associates software to enable precise temporal analysis of calcium signals.

Fiber photometry data analysis. Extraction and analysis of fiber photometry signals was performed using custom code in MATLAB R2018b (MathWorks; original code available on the Tucker-Davis Technologies website: <https://www.tdt.com/support/matlab-sdk/>), following conventional practices reported elsewhere^{36,53,54}. Demodulated 465 nm and 405 nm signals were digitally filtered and scaled. To account for artifacts resulting from movement, photobleaching and autofluorescence, which are reflected by fluctuations in the control 405 nm trace, the control signal was subtracted from the GCaMP6f signal (465 nm). The resulting fitted trace was then normalized to the 405 nm trace ($\Delta F/F = (465 \text{ nm signal} - \text{fitted } 405 \text{ nm signal}) / \text{fitted } 405 \text{ nm signal}$) and used to calculate changes in fluorescence during behavior. Changes in peak fluorescence in response to a 20-s CS or 2-s US presentation were calculated by normalization to a baseline period of equal duration occurring immediately before CS or US onset, respectively. To analyze the inter-CS freezing and contextual freezing epochs, behavior videos were scored to extract freezing onset and offset. Calcium traces associated with the freezing epochs were then aligned and averaged. Changes in fluorescence during the first 2 s of freezing, which was the minimum duration of freezing bouts, were compared to a 2 s prefreezing baseline period immediately preceding the freezing epoch. Percentage change in signal ($\% \Delta F/F$) was calculated by

subtracting the peak fluorescence signal in the test (CS, US or freezing epoch) period from the peak fluorescence signal in the respective baseline period and dividing the resulting signal by the peak fluorescence signal in the baseline period ($\% \Delta F/F = (F_{\text{peakTEST}} - F_{\text{peakBASELINE}}) / F_{\text{peakBASELINE}}$). For frequency analysis of calcium-related events, the median absolute deviation of the fitted signal trace was calculated and all transients that exceeded 2.91 s.d. were manually counted.

Fear conditioning for electrophysiology. Cued auditory fear conditioning entailed 6 pairings of an auditory tone (CS; 2 kHz, 80 dB, 20 s) with a coterminating footshock (US; 1 mA, 2 s). Control and experimental mice were selected from the same litter and recordings from these groups were interleaved. All mice were randomly assigned to behavior groups. Naïve controls received only cage experience, while unpaired controls underwent 6 CS presentations followed 15 min later by 6 US presentations. Mice were killed for electrophysiological analyses 24 h after conditioning.

Slice electrophysiology. Mice were deeply anesthetized using isoflurane inhalation before decapitation. Acute brain slices were prepared from the mPFC at 350 μm thickness on a VT1200S Vibratome (Leica Microsystems) in a low sodium sucrose solution bubbled with carbogen (95% O₂, 5% CO₂) and consisting of 210 mM sucrose, 26.2 mM NaHCO₃, 11 mM glucose, 2.5 mM KCl, 1 mM NaH₂PO₄, 0.5 mM ascorbate, 4 mM MgCl₂ and 0.5 mM CaCl₂, and chilled to -3 – 4 °C. Slices were transferred to a recovery chamber continuously bubbled with carbogen and containing normal artificial cerebrospinal fluid consisting of 119 mM NaCl, 26.2 mM NaHCO₃, 11 mM glucose, 2.5 mM KCl, 1 mM NaH₂PO₄, 2 mM MgCl₂ and 2 mM CaCl₂, and warmed to 34 °C for 45 min. Following recovery, slices were maintained at room temperature until recordings started. Whole-cell electrodes were pulled from borosilicate glass and filled with a low-chloride solution (for voltage clamp recordings) consisting of 120 mM cesium methanesulfonate, 10 mM HEPES, 10 mM sodium phosphocreatine, 8 mM NaCl, 1 mM QX-314, 0.5 mM EGTA, 4 mM Mg-ATP and 0.4 mM Na-GTP or a potassium-based solution (for current clamp recordings) consisting of 127.5 mM potassium methanesulfonate, 10 mM HEPES, 5 mM KCl, 5 mM sodium phosphocreatine, 2 mM MgCl₂, 0.6 mM EGTA, 2 mM Mg-ATP and 0.3 mM Na-GTP. Internal solutions were adjusted to pH 7.25 and 290–300 mOsm. Slices were visualized on an upright differential interference contrast microscope and LED-coupled (Prizmatix) 40 \times objectives were used to identify fluorescently tagged cells as well as optogenetic stimulation. SST and PV interneurons were identified based on tdTomato or eYFP fluorescence, with high membrane resistance (>100 M Ω) or low capacitance (<90 pF) as confirmation of interneuron identity. Principal excitatory PNs were identified based on morphology (large pyramidal soma and prominent apical dendrite), with low membrane resistance (<75 M Ω) or high capacitance (>100 pF) as additional criteria.

EPSC and IPSC postsynaptic currents were isolated by clamping neurons at -60 or 0 mV, respectively, in low-chloride internal solution. A total trace duration of at least 5 min was sampled at each potential for spontaneous currents. For paired-pulse measurements, EPSCs were evoked with a bipolar stimulating electrode placed in L2 of the prelimbic cortex adjacent to the targeted postsynaptic neuron. To stimulate light-evoked transmission, we used transistor-transistor logic-pulsed microscope objective-coupled, light-emitting diodes (460 nm, 20 mW mm⁻²; Prizmatix). This intensity evoked maximal response amplitudes at a pulse duration of 1 ms. Spontaneous postsynaptic currents (Fig. 1 and Supplementary Fig. 1), paired-pulse analysis (Fig. 1) and light-evoked compound currents (Supplementary Fig. 13) were conducted in standard artificial cerebrospinal fluid. For more stringent isolation of monosynaptic currents during optical stimulation (Figs. 5 and 7), recordings were conducted in the presence of 1 μM tetrodotoxin (Abcam) and 100 μM 4-aminopyrimidine (Abcam), which results in complete elimination of polysynaptic activity^{12,13,30}.

Data were low-pass-filtered at 3 (evoked) or 10 kHz (spontaneous) and acquired at 10 kHz using Multiclamp 700B Microelectrode Amplifier (Molecular Devices) and pClamp 10 software (Molecular Devices, version 10.3.1). Mice were randomly assigned to behavior groups and the experimenter was blinded to cell type and experimental condition during the analysis of evoked (Clampfit 10; Molecular Devices) and spontaneous currents (MiniAnalysis version 6.03; Synaptosoft). Neurons were excluded before analysis if they did not meet the cell type criteria for passive membrane properties (as detailed earlier) or yielded unstable current or voltage traces. These rejected neurons accounted for a very small number of recordings.

c-Fos immunofluorescence. For analysis of CS-evoked c-Fos (Supplementary Figs. 5, 6 and 17 and Fig. 8i,j), mice underwent 4 CS presentations in context B. To analyze c-Fos induction by SST interneuron photoexcitation (Fig. 8a–h), Chr2- or eYFP-expressing mice underwent 6 bouts of 20 s stimulation (473 nm, 10 ms pulse, 20 Hz). Ninety minutes after CS or light stimuli, mice were deeply anesthetized via isoflurane inhalation and transcardially perfused with PBS followed by 4% paraformaldehyde in PBS (pH 7.45). Brains were postfixed in paraformaldehyde for 14–16 h postperfusion and sectioned in 50 μm thick slices on the coronal plane on a VT1000S Vibratome (Leica Microsystems). Immunofluorescence staining against c-Fos was conducted on floating sections using a rabbit anti-c-Fos primary antibody (1:1,000; catalog no. ABE457; Merck Millipore)⁵⁵. Fluorescence-conjugated secondary antibodies included goat anti-rabbit conjugated to fluorescein isothiocyanate (1:500; catalog no. 111-095-003; Jackson ImmunoResearch) and goat

anti-rabbit conjugated to Alexa Fluor 647 (1:500; catalog no. 111-605-003; Jackson ImmunoResearch). Slices were blocked in 2% goat serum in 0.3% Tween 20 PBS for 1 h at room temperature. The primary antibody was incubated overnight at 4 °C in 2% goat serum in 0.3% Tween 20 PBS. Following PBS washes, slices were incubated with the secondary antibody in 2% goat serum in 0.3% Tween 20 for 2 h at room temperature. Following additional PBS washes, slices were incubated for 7 min at room temperature with filtered 1 mg ml⁻¹ 4,6-diamidino-2-phenylindole solution, followed by thorough PBS washes. Slices were then mounted with ProLong Gold Antifading Mountant (Thermo Fisher Scientific) and imaged on a ZEISS confocal microscope operating the ZEISS Zen software version 8.0.0.273. tdTomato⁺ and eYFP⁺ SST interneurons, as well as c-Fos⁺ nuclei, were manually quantified using the Cell Counter plugin in ImageJ version 1.47 (National Institutes of Health) while blinded to experimental condition.

Statistical analysis. Before parametric statistical analysis, the Shapiro–Wilk and Levene's tests were used to establish normality of data and homogeneity of variance, respectively. Failing these assumptions or in cases where group sizes might be too small to establish normality, we utilized nonparametric statistical comparisons. All figures utilizing bar graphs contain individual sample data, means and s.e.m. bars. The parameters used for the box plots are included in the corresponding figure legends. For reported effects, statistical power exceeded a minimum of 0.8 and in most cases 0.9. *P* values obtained from the network analysis of c-Fos⁺ cells were corrected for a false discovery rate of 10% using the Benjamini–Hochberg method. Power analyses were conducted at the outset of the project to estimate sample sizes for electrophysiological recordings and freezing behavior based on previous experiments published by our laboratory. Statistical analysis and graphing were conducted in Prism 6 (GraphPad Software) and OriginPro 2016 (OriginLab).

All additional information regarding experimental animals, statistical parameters, software and code, study design, materials and methods can be found in the associated Nature Research Reporting Summary document.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

- Keum, S. et al. A missense variant at the *Nrxn3* locus enhances empathy fear in the mouse. *Neuron* **98**, 588–601.e5 (2018).
- Yaeger, C. E., Ringach, D. L. & Trachtenberg, J. T. Neuromodulatory control of localized dendritic spiking in critical period cortex. *Nature* **567**, 100–104 (2019).
- Kvitsiani, D. et al. Distinct behavioural and network correlates of two interneuron types in prefrontal cortex. *Nature* **498**, 363–366 (2013).
- Calipari, E. S. et al. In vivo imaging identifies temporal signature of D1 and D2 medium spiny neurons in cocaine reward. *Proc. Natl Acad. Sci. USA* **113**, 2726–2731 (2016).
- Gunaydin, L. A. et al. Natural neural projection dynamics underlying social behavior. *Cell* **157**, 1535–1551 (2014).
- Lagerlöf, O. et al. The nutrient sensor OGT in PVN neurons regulates feeding. *Science* **351**, 1293–1296 (2016).

Acknowledgements

We thank S. Bayshtok and E. Beckett for their expert technical assistance, behavioral scoring and assistance with experimenter blinding; C. Román-Ortiz for help with the optogenetic behavioral manipulations; and D. Cai for comments on the manuscript. These experiments were supported by funds from National Institute of Mental Health grant nos. RO1 MH105414, RO1 MH116145 and R21 MH114170 to R.L.C. and grant no. F32 MH115688 to K.A.C.

Author contributions

K.A.C. and R.L.C. initiated the project. R.L.C. supervised the research. K.A.C. and R.L.C. designed the experiments. K.A.C. performed the research and data analysis. R.L.C. and K.A.C. wrote the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41593-019-0552-7>.

Correspondence and requests for materials should be addressed to R.L.C.

Peer review information *Nature Neuroscience* thanks Mario Penzo for his contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection

Molecular Devices Clampex and Clampfit 10.3.1.
Zeiss ZEN 2012 SP2
MedAssociates Video Freeze Version 2.5.5.0
Tucker-Davis Technologies Synapse software, Fiber Photometry gizmo
pClamp10, Version 10.6.2.2

Data analysis

Synaptosoft MiniAnalysis 6.03
NIH ImageJ 1.47
Adobe Photoshop 12.1
OriginPro 2016
GraphPad Prism 5
Clampfit10 version 10.6.2.2
Omnitech Electronics Fusion v5.6 SuperFlex
MATLAB R2018b

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The authors declare that all data supporting the findings of this study are available within the paper and its supplementary information files.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

- Life sciences Behavioural & social sciences Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Power analyses were conducted at the outset of the project (during NIH funding application) to estimate sample sizes for electrophysiological recordings and freezing behavior based on previous experiments published by the laboratory. For cFos analysis, power analysis was conducted on samples from preliminary studies in the amygdala as well as prefrontal cortex tissue. 80% power was considered sufficient.
Data exclusions	For electrophysiological recordings, neurons were excluded prior to analysis if they did not meet cell-type criteria for passive membrane properties (as detailed in Materials and Methods) or yielded unstable current or voltage traces. These rejected neurons accounted for a very small number of recordings. All brains receiving viral injections and/or optic ferrule implantation were sectioned and mounted to confirm targeting. Mistargeted animals were excluded blind to experimental condition or analysis results, if available. Neither animals nor recordings were excluded on the basis of analysis results, for example because they were statistical outliers. All exclusion criteria were pre-established.
Replication	There have been no failures to replicate any findings reported. Independent replications of in vivo optogenetics experiments are described in the paper (e.g. ChR2 manipulations of SST-INs) and have also been conducted by graduate trainees in the laboratory in unreported experiments. cFos analysis of SST-INs after CS exposure as well as ChR2-mediated stimulation was independently replicated in unreported triple antibody labeling experiments. Electrophysiological analysis of BLA-->mPFC synaptic plasticity was in principle replicated in a standard ACSF recording solution for both PV-Cre and SST-Cre mice.
Randomization	Subjects were randomly assigned to experimental groups. All experiments were conducted with interleaved control groups selected from the same mouse litters.
Blinding	Investigator was in all cases blinded to group allocation during both data exclusion and analysis (including electrophysiological, histological and behavioral data). For some electrophysiological experiments in Fig. 1 and 5 the investigator was also blind to group allocation during data collection.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
<input type="checkbox"/>	<input checked="" type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology
<input type="checkbox"/>	<input checked="" type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Human research participants
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data

Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

Antibodies

Antibodies used

We used rabbit anti-cFos (1:1000, Millipore, ABE457, Lot # 2905394), goat anti-rabbit conjugated to FITC (1:500, Jackson

Immuno Research, 111-095-003, lot # 118332) and goat anti-rabbit conjugated to Alexa 647 (1:500, Jackson Immuno Research, 111-605-003, lot # 128500).

Validation

Rabbit polyclonal anti-c-Fos antibody is used to tag c-Fos for detection and quantitation by immunocytochemical and immunohistochemical (IHC) techniques. The manufacturer cites 37 publications for this application. This antibody was validated in fixed floating mouse brain slices (O. Lagerlöf et al., The nutrient sensor OGT in PVN neurons regulates feeding. Science 351, 6279 (Mar 2016)).

Animals and other organisms

Policy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research

Laboratory animals

Mice were acquired from Jackson Laboratories (Bar Harbor, ME, USA) and maintained in the C57Bl/6J background. The following genotypes were used: SST-IRES-Cre (Stock No. 028864), PV-IRES-Cre (Stock No. 017320), SST-IRES-FlpO (Stock No. 028579), PV-T2A-FlpO (022730), and Ai9 (Stock No. 007909). Male mice at P42-60 were used in this study.

Wild animals

This study did not include wild animals.

Field-collected samples

This study did not include field-collected samples.

Ethics oversight

All experimental procedures were approved by the Institutional Animal Care and Use Committee at the Icahn School of Medicine at Mount Sinai.

Note that full information on the approval of the study protocol must also be provided in the manuscript.