



Supplementary Materials for

Engrams and circuits crucial for systems consolidation of a memory

Takashi Kitamura*, Sachie K. Ogawa*, Dheeraj S. Roy*, Teruhiro Okuyama, Mark D. Morrissey, Lillian M. Smith, Roger L. Redondo & Susumu Tonegawa

Correspondence to: tonegawa@mit.edu

This PDF file includes:

Materials and Methods
Figs. S1 to S12
Statistics and Sample Sizes

Materials and Methods

Subjects

All procedures relating to mouse care and treatment conformed to institutional and NIH guidelines. Animals were individually housed in a 12 hour (7am-7pm) light/dark cycle, with food and water ad libitum. We used 15-25 weeks-old wild-type male C57BL/6J mice (WT: Jackson Laboratory), c-fos:tTA transgenic mice (Jackson Laboratory; strain B6.Cg-Tg(Fos-tTA,Fos-EGFP*)1Mmay/J; stock number 018306) (16) and TRE:H2B-GFP transgenic mice (Jackson Laboratory: strain Tg(tetO-HIST1H2BJ/GFP)47Efu/J; stock number 005104) (29, 34).

Preparation of Adeno-Associated Viruses

The pAAV-TRE:ChR2-mCherry, pAAV-TRE:ArchT-eGFP, pAAV-TRE:HTG plasmids (Addgene plasmid # 27437) and pAAV-c-fos:tTA were constructed as previously reported (26, 35-38). The pAAV-TRE:TeTX was constructed by replacing the ChR2-eYFP fusion gene in the pAAV-TRE:ChR2-eYFP plasmid from Liu et al. (17) with a gene of TeTX-LC from pGEMTEZ-TeTXLC (a gift from Richard Axel & Joseph Gogos & C. Ron Yu (Addgene plasmid # 32640) (39). These plasmids were used to generate by the Gene Therapy Center and Vector Core at the University of Massachusetts Medical School or ViGENE Biosciences. Viral titrations were 8.0×10^{12} genome copy per ml for AAV₉-TRE:ChR2-mCherry, 1.0×10^{13} genome copy per ml for AAV₉-TRE:ArchT-eGFP, 1.4×10^{13} genome copy per ml for AAV₅-TRE:HTG, 1.5×10^{13} genome copy per ml for AAV₉-c-fos:tTA and 3.7×10^{14} genome copy per ml for AAV₉-TRE:TeTX. For the long-term labeling of DG engram cells, we employed split-Cre AAV constructs, in which the Cre coding sequence is divided into N-terminus (NCre) and C-terminus (CCre) parts and their dimerization via a leucine zipper re-forms the functional enzyme (40). TRE3G split Cre AAV (TRE3G:NCre and TRE3G:CCre) did not have detectable leak expression in 293T cells or in the brain (41). AAV_{1,5,8,9}-TRE3G:NCre (5.7×10^{13} genome copy/ml) and AAV_{1,5,8,9}-TRE3G:CCre (1.6×10^{13} genome copy/ml) were generated with the a cocktail of 4 serotypes (2/1, 2/5, 2/8, 2/9) by Yasuyuki Shima and Sacha B Nelson (Brandeis University) in house (41). The AAV₅-CaMKII α :eYFP (5.2×10^{12} genome copy/ml), AAV₅-CaMKII α :eArchT-eYFP (2.5×10^{12} genome copy/ml), AAV₈-CaMKII α :eYFP (3.5×10^{12} genome copy/ml) and AAV₈-CaMKII α :eArchT-eYFP (2.7×10^{12} genome copy/ml) and AAV₉-

CaMKII α :eArchT-eYFP (2.0×10^{13} genome copy/ml) were generated by and acquired from the University of North Carolina at Chapel Hill (UNC) Vector Core. The AAV₅-Syn:GCaMP6f (1.5×10^{13} genome copy/ml) and AAV₅-EF1 α :DIO-ChR2-mCherry (2.0×10^{13} genome copy/ml) was generated by and acquired from the University of Pennsylvania Vector Core.

Preparation of Rabies Virus

Deletion-mutant rabies virus was produced by replacing the eGFP gene in cSPBN-4GFP with the gene encoding mCherry. Packaging with the ASLV-A envelope protein was made as previously described (42, 43).

Stereotactic Injection and Fiber Optic Implants

Stereotactic viral injections, microendoscope implantations, and optic fiber implantations were all performed in accordance with MIT's CAC guidelines. Mice were anaesthetized using 500 mg/kg avertin. Viruses were injected using a glass micropipette attached to a 10 μ l Hamilton microsyringe through a microelectrode holder filled with mineral oil. A microsyringe pump and its controller were used to control the speed of the injection. The needle was slowly lowered to the target site and remained for 10 minutes after the injection. For the terminal inhibition of MEC-Va projection at PFC, cACC (caudal part of ACC), RSC and BLA, bilateral viral delivery into MEC of WT male mice was aimed at these coordinates relative to Bregma: AP: -4.85 mm, ML, ± 3.45 mm, DV, -3.30 mm. Mice were bilaterally injected with 200 nl of AAV₈-CaMKII α :eArchT-eYFP or AAV₈-CaMKII α :eYFP as control group, and were bilaterally implanted with optical fibers into the PFC (AP: +1.7 mm, ML, ± 0.35 mm, DV, -1.6 mm), cACC (AP: ± 0.0 mm, ML, ± 0.35 mm, DV, -1.2 mm, posterior part of ACC), RSC (AP: -1.94 mm, ML, ± 0.35 mm, DV, -0.75 mm) or BLA (AP -1.5 mm, ML +3.3 mm, DV -4.3 mm). For the terminal inhibition of BLA cells at PFC, bilateral viral delivery into BLA of WT male mice was aimed at these coordinates relative to Bregma: AP: -1.5 mm, ML, ± 3.3 mm, DV, -4.68 mm. Mice were bilaterally injected with 150 nl of AAV₉-CaMKII α :eArchT-eYFP or AAV₉-CaMKII α :eYFP as control group, and were bilaterally implanted with optical fiber into the PFC (AP: +1.7 mm, ML, ± 0.35 mm, DV, -1.6 mm). For the cell body activation/inhibition of PFC engram cells, c-fos:tTA mice were bilaterally injected with 150 nl of AAV₉-TRE:ArchT-eGFP or AAV₉-TRE:ChR2-mCherry into the PFC (AP: +1.7 mm, ML, ± 0.35 mm, DV, -1.7 mm), and were

bilaterally implanted with optical fiber into the PFC (AP: +1.7 mm, ML, \pm 0.35 mm, DV, -1.6 mm) under the ON-Dox condition. For the cell body activation of DG engram cells, c-fos:tTA mice were bilaterally injected with 200 nl of AAV₉-TRE:ChR2-mCherry into the DG (AP: -2.0 mm, ML, \pm 1.3 mm, DV, -2.05 mm), and were bilaterally implanted with optical fiber into the DG (AP: -2.0 mm, ML, \pm 1.3 mm, DV, -1.5 mm) under the ON-Dox condition. The top part of an Eppendorf tube was inserted to protect the implant and the incision was closed with sutures⁽¹⁷⁾¹⁰. 100 nl of CTB-Alexa555 (0.5 % wt/vol) was unilaterally injected into PFC (AP: +1.7 mm, ML, +0.35 mm, DV, -1.7 mm), caudal ACC (cACC, AP: \pm 0.0 mm, ML, +0.35 mm, DV, -1.2 mm), RSC (AP: -1.94 mm, ML, +0.35 mm, DV, -0.75 mm) or BLA (AP -1.5 mm, ML +3.3 mm, DV -4.68 mm). Five to seven days after the CTB injection. Mice were perfused for brain sampling. For the GCaMP6f activity monitoring from the PFC, bilateral viral delivery into the PFC of the WT male mice was aimed at these coordinates relative to Bregma: AP: -1.78 mm, ML, \pm 1.3 mm, DV, -1.7 mm. WT mice were bilaterally injected with 150 nl of AAV2/5-Syn:GCaMP6f. One month after AAV injection, we implanted a microendoscope into the right side of PFC (AP: -1.78 mm, ML, +0.35 mm, DV, -1.6 mm) of the WT mice. One month after implantation of the microendoscope, the baseplate for a miniaturized microscope camera (22, 27) was attached above the implanted microendoscope in the mice. After the baseplate surgery, animals were habituated to the attachment of the microscope camera for two weeks.

Histology

Mice were transcardially perfused with 4 % paraformaldehyde (PFA) in phosphate buffered saline (PBS). Brains were post-fixed with the same solution for 24 hours, and then sectioned using a vibratome by sagittal sections or coronal sections depends on purpose of experiments. For immunohistochemistry, sliced tissue sections were incubated in 0.3 % Triton-X PBS with 5 % normal goat serum (NGS) for 1 hour. Primary antibodies were then added to a 5 % NGS 0.3 % triton-X in PBS solution and incubated overnight at 4°C. Primary antibodies: chicken anti-GFP antibody (Thermo Fisher Scientific, A-10262, 1:1000), rabbit anti-PCP4 antibody (Sigma, HPA005792-100UL, 1:200), mouse anti-NeuN antibody (Millipore, MAB377, 1:1000), rabbit anti-c-Fos antibody (Santa Cruz Biotechnology, SC-52, 1/500), rabbit anti-RFP antibody (Rockland, 600-401-379, 1/1000), and rabbit anti-VAMP2 antibody (Synaptic Systems, 104 202, 1/250). After rinsing with PBS 3 times for 15 min each, sliced tissue sections were subsequently

incubated with AlexaFluor405, AlexaFluor488, AlexaFluor546, or AlexaFluor633 conjugated secondary antibodies (Thermo Fisher Scientific, 1:250). Sliced tissue sections were then washed in PBS 3 times for 15 min and mounted in VECTASHIELD antifade mounting medium for Fluorescence on glass slides. Some sections were stained by DAPI or Nissl (1:500). Fluorescence images were taken by confocal microscopy using 10X, 20X, 40X, 63X objectives and by fluorescent microscopy using 10X and 20X objectives. Z-projected confocal images were generated by Zenblack.

Contextual Fear Conditioning and Recall Testing

Fear conditioning was performed on male mice aged between 15 and 25 weeks, in the animal facility during the light cycle with minor modifications to the method described previously (19, 27). Before fear conditioning, all animals were habituated to human experimenters for 3 days. For contextual fear conditioning, we used two different contexts (Context-A and Context-B). Context-A was a chamber (29 cm (W) x 23 cm (D) x 20 cm (H)) with distinct visual cues and a grid floor which consisted of 36 stainless steel rods (Context-A). Context-B was a chamber (29 cm (W) x 23 cm (D) x 20 cm (H)) with different visual cues, compared to Context-A, and has a white plastic floor (Context-B). These chambers were cleaned with 70% ethanol prior to the introduction of each individual mouse. When we examined the terminal inhibition during conditioning (fig, S1E), on Day-1, the optical fiber implants was bilaterally connected to a 561 nm laser controlled by a function generator and green light stimulation (15 mW, each hemisphere) was illuminated during the entire training period (300 s). Subject mice were kept in the conditioning chamber in Context-A for total 300 s. Foot-shock (2 s, 0.75mA) was delivered at 120 s, 180 s and 240 s. Mice remained in the conditioning chamber for a total of 300 s. On Day-2, Day-8, Day-15, Day-22, mice were placed into Context-A (or Context-B) and allowed to explore for 180 s to monitor their freezing behavior. When we examined the terminal inhibition during recall testing (fig, S1F), on Day-1, subject mice were kept in the conditioning chamber in Context-A for total 300 s. Foot-shock (2 s, 0.75mA) was delivered at 120 s, 180 s and 240 s. Mice remained in the conditioning chamber for a total of 300 s. On Day-2, Day-8, Day-15 and Day-22, the optical fiber implants was bilaterally connected to a 561 nm laser controlled by a function generator and green light stimulation (15 mW, each hemisphere) was illuminated during the entire testing period (180 s) in Context-A (or Context-B).

Activity-Dependent Cell Labeling

For activity-dependent cell labeling, we employed c-fos:tTA transgenic mice combined with AAV-TRE system or TRE transgenic mice (16) (17, 29, 37, 38), a) AAV₉-TRE:ChR2-mCherry, AAV₉-TRE:ArchT-eGFP or AAV₅-TRE:HTG injection into c-fos:tTA mice, and b) c-fos:tTA mice crossed with TRE:H2B-GFP transgenic mice. These mice were maintained under the food condition containing 40 mg/kg Dox from their birth. Mice were single housed post-surgery and throughout the rest of the experiments. Before labeling cells, all animals were habituated to human experimenters for 3 days. In order to open a window of activity-dependent labeling for CFC in Context-A, the mice were then taken off Dox for 24 hours. Animals were exposed to Context-A for 300 s for CFC, and DOX diets were resumed immediately after CFC. When we did not open a window for activity-dependent labeling during CFC, there was no cell labeling in PFC, DG, nor BLA (Fig. 1Q, fig. S6). Moreover, there was no reduction of number of H2BGFP positive cells in PFC, DG, and BLA through the 12 days after CFC. In the DG engram cells labeling with ChR2, we have a slight reduction of the expression of ChR2 in DG engram cells. So, To avoid the problem, we employed the TRE-dependent a split-Cre virus combined with c-fos:tTA system (Fig. 3M-N). Although we had a small leak expression under ON-Dox condition when we conducted the long-term labeling in the DG with a split-Cre system (Fig. 3M-N), the leak expression was much smaller than homecage labeling and we clearly observed the CFC-induced increment of cell labeling with ChR2-mCherry in the DG.

Cell Counting

Sixty minutes after the exposure to Context-A or CFC, animals were perfused and their brains were fixed with 4% PFA to detect the c-Fos immunoreactivities in the DG, CA3 and PFC with anti-c-Fos antibody (Santa Cruz Biotechnology, SC-52, 1/1000). Figure 1M-O (PFC) and Figure 2Q (PFC), a sampling of c-Fos⁺ cells was conducted throughout PFC (from AP: +1.94 mm to AP: +1.34 mm) in coronal sections for the quantification analysis of the number of c-Fos⁺ cells in dorsal medial PFC in each group of mice. Five coronal PFC sections were used to count the number of c-Fos⁺ cells and DAPI⁺ cells in the PFC regions and calculated the percentage of c-Fos⁺/DAPI⁺ in each groups. In Figure 3D, a sampling of c-Fos⁺ cells was conducted throughout dorsal CA3 in sagittal sections for the quantification analysis of the number of c-Fos⁺ cells in

CA3 in each group of mice. Ten sagittal hippocampal sections were used to count the number of c-Fos⁺ cells and DAPI⁺ cells in CA3 regions and calculated the percentage of c-Fos⁺/DAPI⁺ in each groups. In Figure 4F-K, a sampling of c-Fos⁺ cells was conducted in anterior BLA (from AP: -1.0 mm to AP: -1.6 mm) in coronal sections for the quantification analysis of the number of c-Fos⁺ cells in BLA in each group of mice, since fear-related neurons are localized more anterior part of BLA (44, 45). Five coronal BLA sections were used to count the number of c-Fos⁺ cells and DAPI⁺ cells in BLA regions. All counting was performed blind as to the group and condition. When we evaluated the reactivation of engram cells by natural recall cues (or non-conditioned cues) in PFC (Fig. 2A-C, Fig. 3E-F), DG (Fig. 3I-L), and BLA (Fig. 4F-K) (we referred in text as overlap experiment), the numbers of H2B-GFP⁺ cells, c-Fos⁺ cells and DAPI⁺ cells in the region of interest (ROI) obtained by the confocal microscopy were counted, and the ratios of c-Fos⁺ H2B-GFP⁺ / H2B-GFP⁺ and the ratio of c-Fos⁺ H2B-GFP⁻ / H2B-GFP⁻ were calculated. Reactivation of hippocampal engram cells by natural cues were previously conducted by several different groups (20, 29, 32, 46, 47). Total number of DAPI⁺ cells studied were 3,984 (Fig. 2C), 7,894 (Fig. 3L), 2,091 (Fig. 3F), and 870 (Fig. 4H, K) cells. Chance level for overlap in BLA (Fig. 4H, K) was calculated as (c-Fos⁺/DAPI⁺) x (H2B- GFP⁺/DAPI⁺) as previously reported (48).

Dendritic Spines Density in PFC and DG Engram Cells

Dendritic spine counting was conducted as previously described (24, 26) with minor changes. Engram cells in the PFC and the DG were labelled using c-Fos-tTA-driven synthesis of Chr2-mCherry. The mCherry signal was amplified using immunohistochemistry procedures with anti-RFP antibody (Rockland, 600-401-379, 1/1000), after which fluorescence z-stacks were taken by confocal microscopy using a × 40 objective. Maximum intensity projections were generated using ZEN Black software. Three mice per experimental group were analyzed for dendritic spines. For each mouse, 20–30 dendritic fragments of 10 μm length were quantified (n = 70–110 fragments per group, detail is shown in *Statistics and Sample Sizes*). To measure spine density of DG engram cells, dendritic fragments in the middle molecular layer were selected. For PFC engram cells, apical and basal dendritic fragments were selected. To compute spine density, the number of spines counted on each fragment was normalized by the cylindrical approximation of the surface of the specific fragment. Experiments and analyses were conducted blind to

experimental group. Averages of spines density in the PFC cells and DG cells were comparable with previous reports (14, 24, 26, 49).

Optogenetic activation of PFC and DG Engram Cells

In this study, we tested whether activating PFC and DG engram cells labelled by ChR2–eYFP during CFC is sufficient for memory recall at recent and remote time points. For light-induced freezing behavior, a context distinct from the CFC training chamber (Context-B) was used. These were 30 × 25 × 33 cm chambers with white plastic floors, square ceilings, white lighting, and scented with 0.25 % benzaldehyde. Chamber ceilings were customized to hold a rotary joint connected to two 0.32-m patch cords. All mice had patch cords fitted to the optic fiber implant before testing. ChR2 was stimulated at 4 Hz (15 ms pulse width) or 20 Hz (15 ms pulse width) using a 473 nm laser (10–15 mW), for the designated epochs. Testing sessions were 12 min in duration, consisting of four 3 min epochs, with the first and third as light-OFF epochs, and the second and fourth as light-ON epochs. At the end of 12 min, the mouse was detached and returned to its home cage. Floors of chambers were cleaned with 70 % ethanol before test. Freezing counting experiments were conducted double blind to experimental group. Experiments that resulted in significant behavioral effects were replicated three times in the laboratory. Following behavioral protocols, brain sections were prepared to confirm efficient viral labelling in target areas. Animals lacking adequate labelling were excluded before behavior quantification.

TeTX Expression in HPC Engram Cells

To inhibit the synaptic transmission from HPC engram cells after CFC, AAV₉-TRE:TeTX was targeted to bilaterally inject into the hippocampal dentate gyrus of *c-fos*:tTA mice and the virus was often distributed to not only DG but also CA3 region. To avoid DOX-independent TeTX expression (leak expression) in the HPC, we injected the 100-fold diluted AAV solution with PBS. To examine the effect of TeTX expression in DG engram cells on the reactivation of PFC engram cells at remote timepoint (Fig. 3E-F), we injected the AAV₅-TRE:H2B-GFP into the PFC and AAV₉-TRE:TeTX into the HPC of *c-fos*:tTA mice. To examine the effect of TeTX expression in DG engram cells on the spine density of the PFC engram cells at remote timepoint (Fig. 3G), we injected the AAV₉-TRE:ChR2-mCherry into the PFC and AAV₉-TRE:TeTX into the HPC of *c-fos*:tTA mice. To examine the effect of TeTX expression in DG engram cells on

the GCaMP activity of PFC cells at recent and remote timepoint (Fig. 3H), we injected the AAV₅-Syn:GCaMP6f into the PFC and AAV₉-TRE:TeTX into the HPC of c-fos:tTA mice. The c-fos:tTA mice were maintained under the food condition containing 40 mg/kg Dox from their birth.

Monosynaptic Tracing with Rabies Virus

For the monosynaptic retrograde tracing experiment (fig. S5), 150 nl of AAV₅-TRE:HTG (H2BGFP-TVA-G) was stereotaxically delivered into the PFC of c-fos:tTA mice (AP: +1.7 mm, ML, ±0.35 mm, DV, -1.7 mm). Experiments were all conducted in 15–25-week-old male and female mice. The c-fos:tTA mice were maintained under the food condition containing 40 mg/kg Dox from their birth. Mice were single housed post-surgery and throughout the rest of the experiments. In order to open a window of activity-dependent labeling for CFC in Context-A, the mice were then taken off Dox for 24 hours. Animals were exposed to Context-A for 300 s for CFC, and DOX diets were resumed immediately after CFC. One day after CFC, 50 nl of rabies virus, RVΔG-mCherry(EnvA), was injected at the same coordinates, and mice were perfused 3-5 days later.

Calcium Imaging

Calcium imaging from PFC was performed on WT mice in the animal facility during the light cycle. Mice were habituated to human experimenters as well as the experimental room two weeks following AAV injection, microendoscope implant, and baseplate surgeries. Imaging from PFC via microendoscope was previously reported (50). Experiments were run in several different pairs of contexts (Fig. 2J). On Day-1, mice were first exposed to Context-B followed by CFC in Context-A. Mice were then re-exposed to both contexts in the same order one day later (Day-2) and 14 days later (Day-15). Ca²⁺ signals from GCaMP6f were imaged during entire exposure periods. The chambers were cleaned with 70 % ethanol between sessions. Immediately before and after imaging sessions, the mouse rested on a pedestal next to the experimental room. Ca²⁺ signals from GCaMP6f were captured at 20 Hz on a miniature microscope. The movie of Ca²⁺ signals was then motion corrected using Inscopix Mosaic software (correction type: translation and rotation; reference region with spatial mean (r = 20 pixels) subtracted, inverted, and spatial mean applied (r = 5 pixels)). Finally, it was processed by ImageJ (dividing each image, pixel by

pixel, by a low-passed ($r = 20$ pixels) filtered version), and the $\Delta F/F$ signal was calculated by Inscopix Mosaic software as previously described (27, 51). Cell locations were semi-automatically identified from the stacked image, and selected as small regions of interest (ROIs) ($\sim 1/3$ of cell body size) at the center of the cell bodies. This was done in ImageJ for each processed movie. Approximately 40-80 cells were selected per animal, and their $\Delta F/F$ signals were isolated. Calcium traces were calculated at these ROIs for each processed movie, in ImageJ. Calcium events were detected by thresholding (> 3 Standard Deviations from the $\Delta F/F$ signal,) at the local maxima of the $\Delta F/F$ signal. In Fig 2G-O, total 220 cells in the PFC are monitored from 3 mice. In Fig 3H, total 89 cells in the PFC are monitored from 2 mice. A rate difference index (defined below) was calculated for all cells that had > 10 Ca^{2+} events during exposure to a context. To define the shock-responding cells in PFC, we calculated the averaged frequency of each PFC cell during Context-A (5 min), Context-B (5 min, excluding shock periods), and Shock periods in Context-B (2 min) on Day-1. If the Averaged frequency in shock-period was 4 times higher than that in Context-A and Context-B (before shock), the cells are identified as shock responding (SR) cells.

Rate Difference Index of Ca^{2+} Events

We used the following score to determine the score of “context preference” in each single cell:

$$\text{Rate Difference Index} = \frac{n_{\text{events},X} - n_{\text{events},Y}}{n_{\text{events},X} + n_{\text{events},Y}}$$

Where $n_{\text{events},X}$ is the number of Ca^{2+} events in context X . Similar definition for $n_{\text{events},Y}$.

The rate difference score was calculated for every cell in each animal for Context-A versus Context-B on Day-1, Day-2 and Day-15 to evaluate the preference of Ca^{2+} activity in different contexts. In Figure 2, we displayed the cumulative probability of the rate difference indices for different contexts (A vs B) on Day-1, Day-2 and Day-15. When the distribution shifts to right, this indicates a larger change in the number of Ca^{2+} events and higher median rate difference index.

Statistical Analysis

GraphPad Prism version 6.01 for Windows, GraphPad Software, La Jolla California USA was used for statistical analysis. All data are presented as mean \pm SEM. N indicates number of animals. n indicates number of cells or spines. Comparisons between two-group data were analyzed by two-tailed unpaired t-test or two-tailed paired t-test. Comparisons of distribution data between two-groups were analyzed by Kolmogorov–Smirnov (KS) test. Multiple group comparisons were assessed using a one-way, two-way (ANOVA), followed by the post-hoc Tukey-Kramer test when significant main effects or interactions were detected. The null hypothesis was rejected at the $P < 0.05$ level. Data met assumptions of statistical tests. Sample sizes were chosen on the basis of previous studies (17, 19, 24, 26, 27, 29, 35).

Supplemental Figures

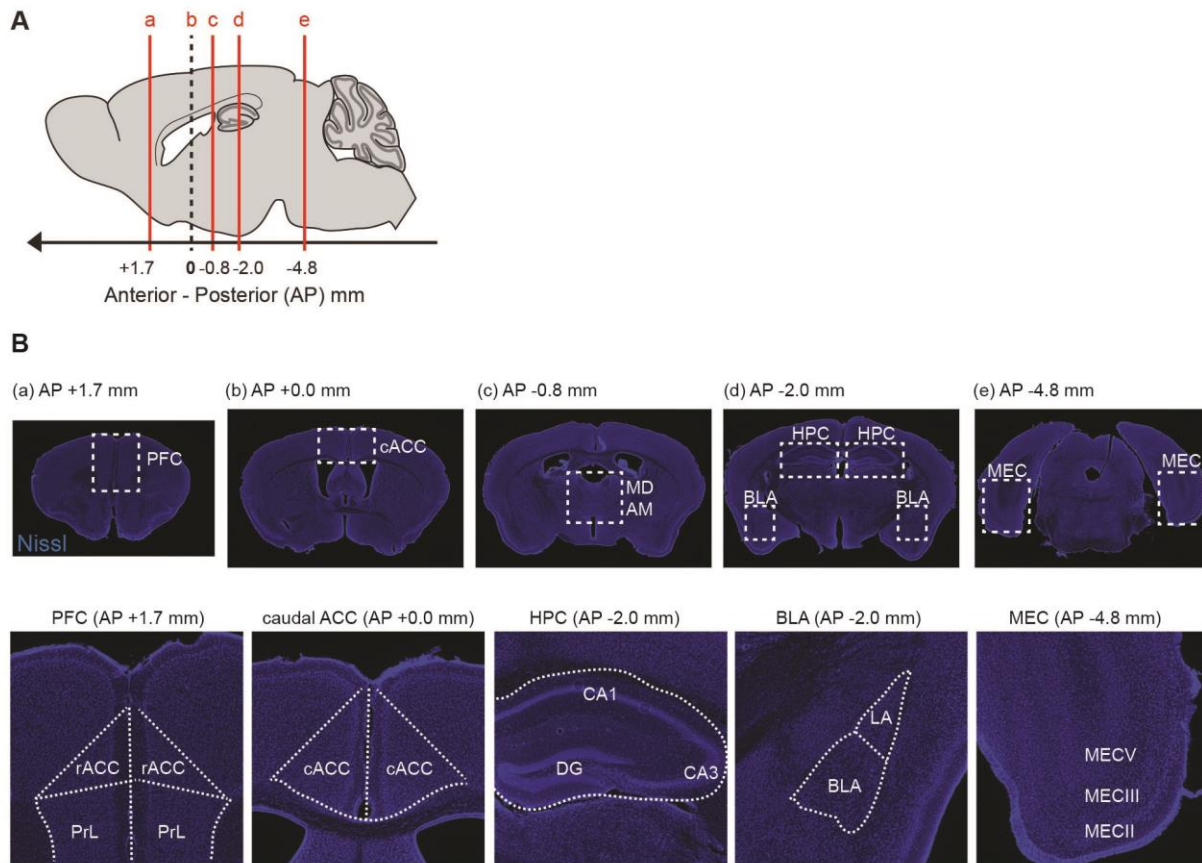


Fig. S1 Brain region references for systems consolidation of memory.

(A) Sagittal section of mouse brain. (B) Coronal sections at each AP positions (AP = +1.7mm; PFC, AP = 0.0 mm; caudal part of ACC (cACC), AP = -0.8 mm; MD, AM, AP = -2.0 mm; HPC, BLA, AP = -4.8 mm; MEC). PFC includes rostral part of ACC (rACC) and prelimbic cortex (PrL). Hippocampus (HPC) includes DG, CA3 and CA1. LA; lateral amygdala, BLA; basolateral amygdala.

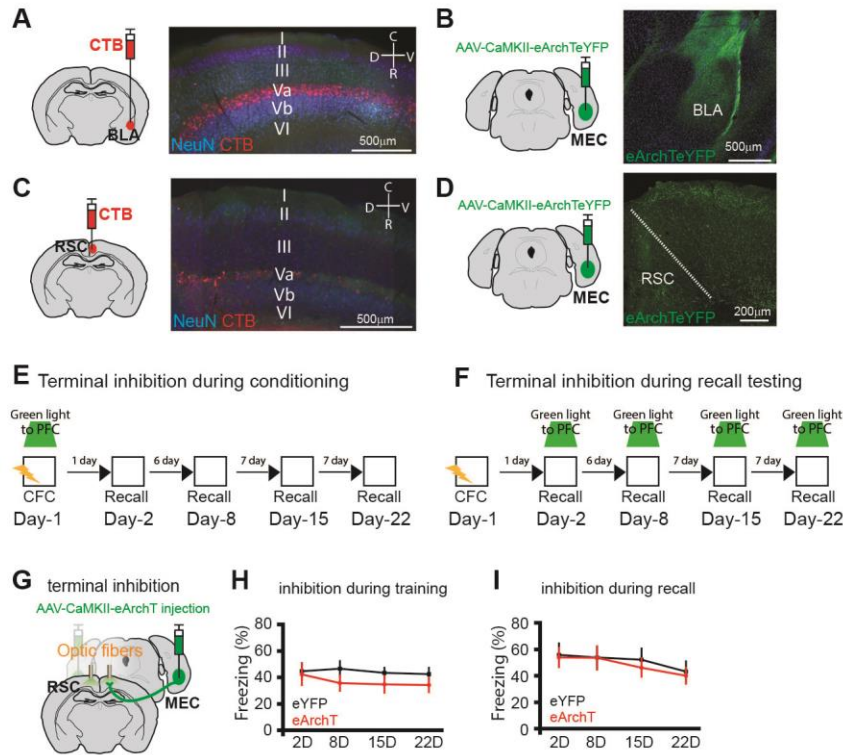


Fig. S2

MEC-Va input to RSC during the conditioning or testing periods is dispensable for contextual fear memory.

(A-B) CTB was injected into BLA. Sagittal section of MEC visualized with CTB-labeled cell bodies (red) and immunostained with anti-NeuN (blue) (N = 10 mice) (A). Coronal sections of BLA visualized with MECVa axons by eYFP (green) (B). (C-D) CTB was injected into RSC. Sagittal section of MEC visualized with CTB-labeled cell bodies (red) and immunostained with anti-NeuN (blue) (N = 5 mice) (C). Coronal sections of RSC visualized with MECVa axons by eYFP (green) (D). (E-F) Experimental schedules. (G-I) Mice were injected with the AAV₈-CaMKII:eArchT-eYFP bilaterally into MEC and implanted with an optic fiber bilaterally into the RSC (G). Time courses of freezing during recall tests on 1 day (Day-2), 7 days (Day-8), 14 days (Day-15) and 21 days (Day-22) after CFC. Green light was shone into RSC during conditioning (H) or testing (I). (H) eYFP; N = 8 mice, eArchT; N = 8 mice. There were no difference between eYFP and eArchT groups in freezing response on Day-2 ($t_{14} = 0.23$, $P = 0.82$), Day-8 ($t_{14} = 1.21$, $P = 0.25$), Day-15 ($t_{14} = 1.03$, $P = 0.32$) and Day-22 ($t_{14} = 0.98$, $P = 0.34$), analyzed by two-tailed unpaired t-test. (I) eYFP; N = 6 mice, eArchT; N = 6 mice. There were no difference between eYFP and eArchT groups in freezing response on Day-2 ($t_{10} = 0.16$, $P = 0.88$), Day-8 ($t_{10} = 0.02$, $P = 0.99$), Day-15 ($t_{10} = 0.52$, $P = 0.62$) and Day-22 ($t_{10} = 0.31$, $P = 0.77$), analyzed by two-tailed unpaired t-test. Error bars, mean \pm s.e.m.

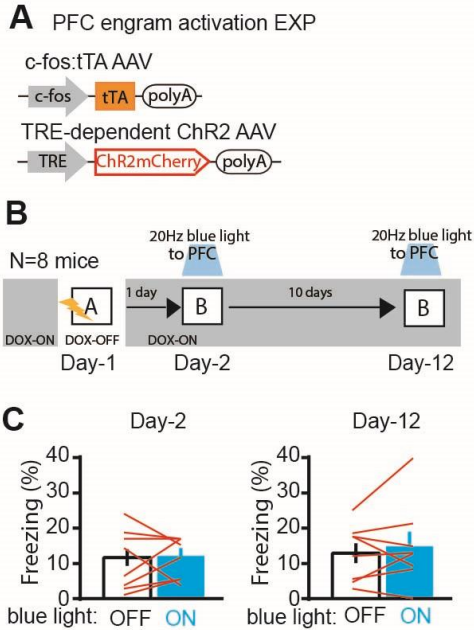


Fig. S3

Conventional 20 Hz blue light stimulation of ChR2-mCherry-expressing cells in the PFC does not induce the freezing behavior. (A) Virus-mediated engram labelling strategy using a cocktail of AAV₉-c-fos:tTA and AAV₉-TRE:ChR2-mCherry into PFC. (B) Experimental schedule for optogenetic activation of c-Fos expressed cells in PFC. (C) Freezing by blue light stimulation 1 day and 11 days after CFC. Average freezing for light-off and light-on epoch (N = 8 mice). There were no difference in freezing score between light-off and light-on epochs on Day-2 ($t7 = -0.7$, $P = 0.51$) and Day-12 ($t7 = -0.77$, $P = 0.47$), analyzed by two-tailed paired t-test. Error bars, mean \pm s.e.m.

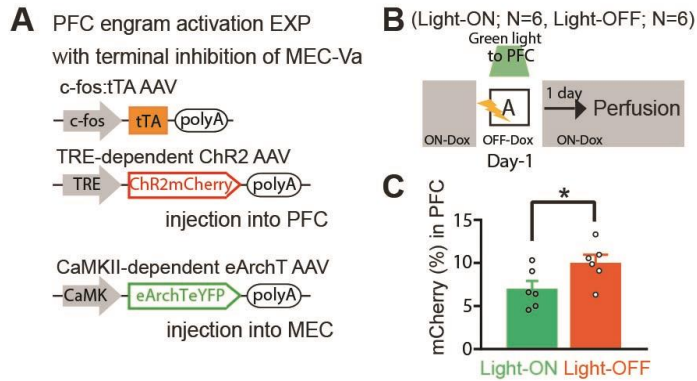


Fig. S4 Terminal inhibition of MEC-Va fibers at the PFC reduced the number of ChR2-mCherry⁺ cells in the PFC

(A) Virus-mediated engram labelling strategy using a cocktail of AAV₉-c-fos:tTA and AAV₉-TRE:ChR2-mCherry into PFC with injection of AAV₈-CaMKII-eArchTeYFP into the MEC. (B) Experimental schedule. (C) Percentages of mCherry⁺ cells in PFC of green light-ON group and green light-OFF group. There was a significant difference in the percentages of mCherry⁺ cells in the PFC between green light-ON group (N=6 mice) and green light-OFF group (N=6 mice) ($t_{10} = -2.29$, $P = 0.045$).

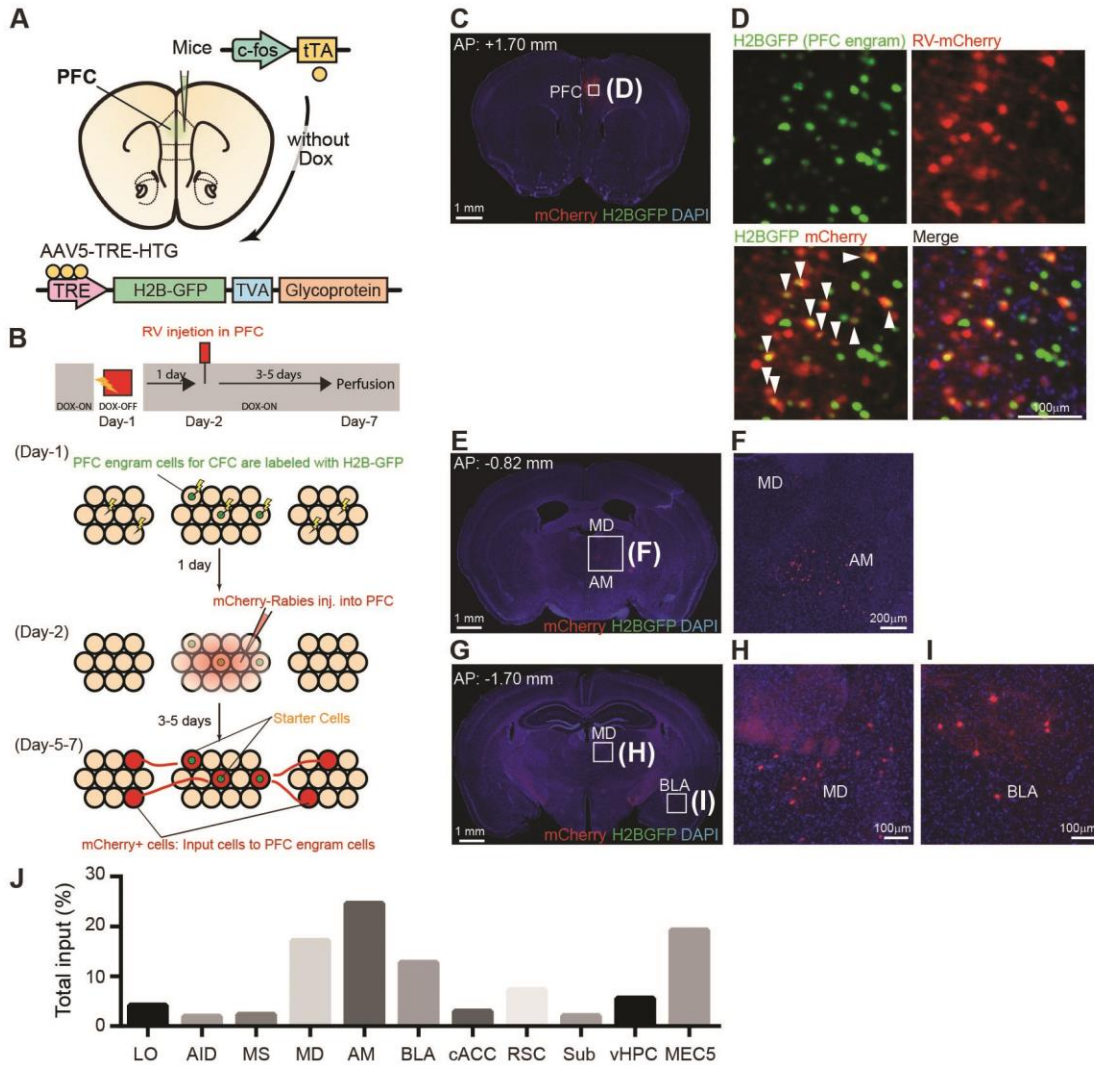


Fig. S5

Identification of monosynaptic inputs onto PFC engram cells. (A-B) Experimental schedules. To visualize the brain areas that provide monosynaptic inputs to PFC engram cells, we used TRE-dependent AAV (AAV₅-TRE-HTG) (Miyamichi et al., 2011) encoding H2B-GFP, the avian cell-surface receptor TVA, and the rabies virus glycoprotein (G), and injected this stereotaxically into the PFC area of c-fos-tTA transgenic mice. One day after CFC while mice are DOX-Off, a deletion-mutant rabies virus encoding mCherry and packaged with the TVA-specific envelope protein EnvA (RV Δ G-mCherry(EnvA)), was injected into the same area and the brain was analyzed 3-5 days later (B). (C-I) With this method, starter cells of PFC engram cells will be labeled by both mCherry (red) and H2B-GFP (green) and will therefore appear yellow (arrow heads), while the cells that are directly presynaptic to PFC engram cells will be labeled by mCherry only and hence appear red (B-D). mCherry⁺ cells were observed not only in MEC-Va (Fig. 1W), but also in medial dorsal thalamus (MD) (E, F, G, H), anteromedial thalamus (AM) (G, H), and BLA (G, I). (J) Distribution of mCherry⁺ cells in the brain (n=770 mCherry⁺ cells). LO; lateral orbital cortex, AID; agranular insular cortex, MS; medial septum, Sub; subiculum, vHPC; ventral hippocampus.

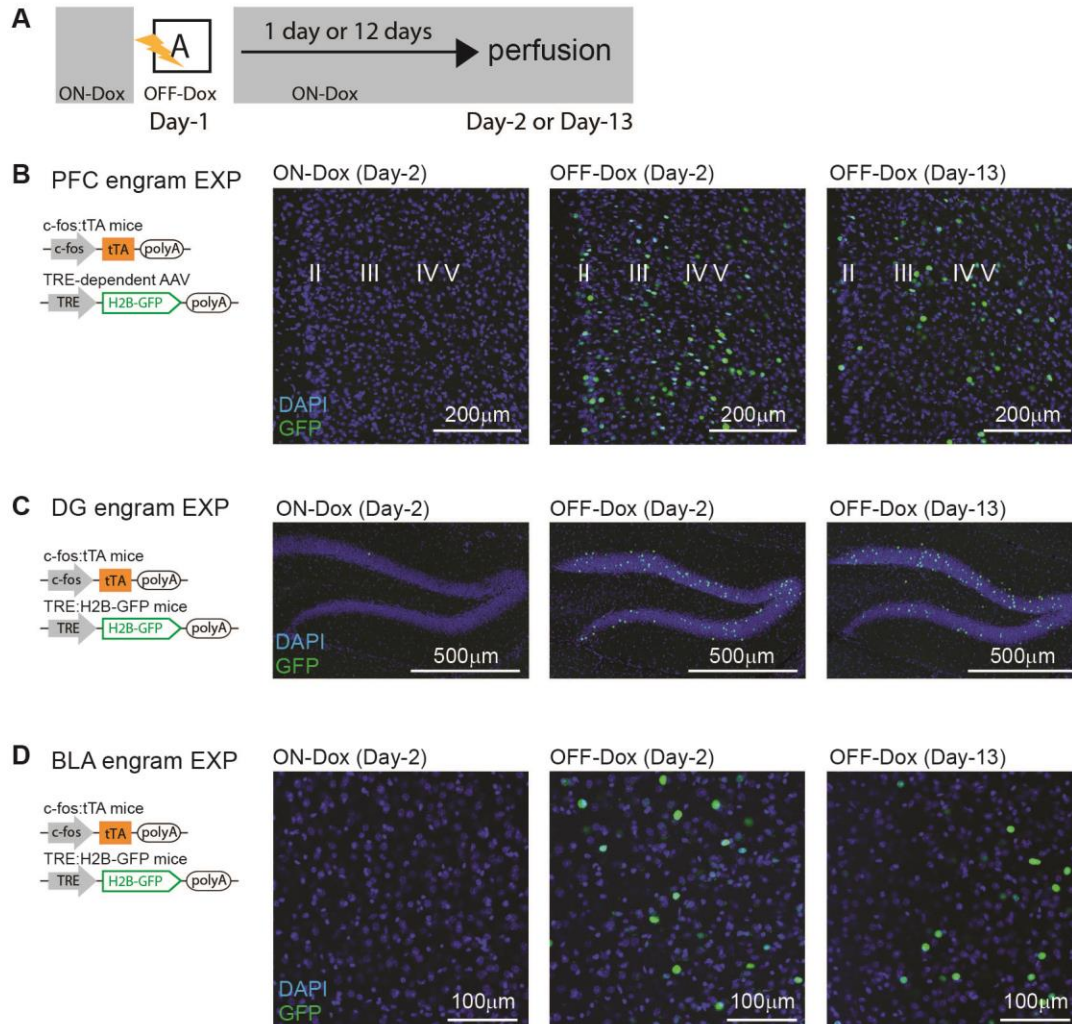


Fig. S6

Engram cell labeling in PFC, DG and BLA with H2B-GFP is DOX-dependent. (A) Experimental schedule. (B) PFC engram cells labelling injecting AAV₉-TRE:H2B-GFP into PFC of c-fos:tTA transgenic mice. Coronal sections of PFC visualized with H2B-GFP (green) and DAPI (blue) under ON-Dox and OFF-Dox conditions (1 day and 12 days after CFC). (C-D) Transgenic strategy of engram labelling using c-fos:tTA transgenic crossed with TRE:H2B-GFP transgenic mice. Coronal sections of DG (C) and BLA (D) visualized with H2B-GFP (green) and DAPI (blue) under ON-Dox and OFF-Dox conditions (1 day and 12 days after CFC) (N = 3 mice, each group).

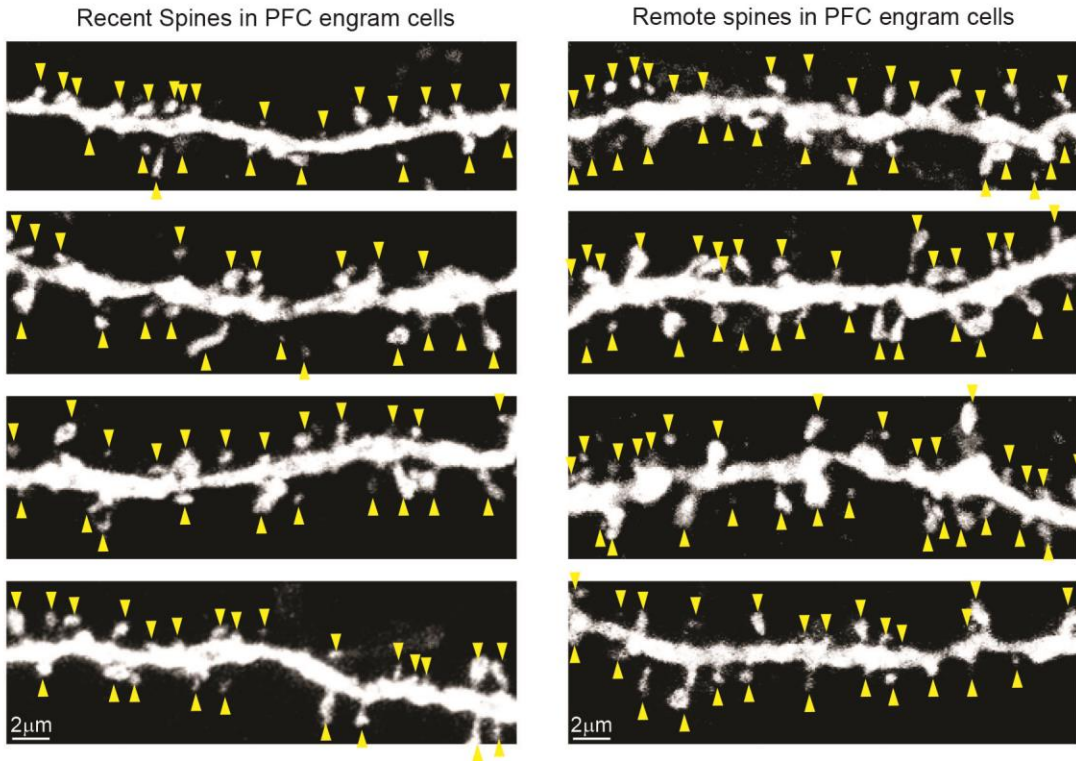


Fig. S7
Confocal Images of Dendritic Spines of PFC memory engram cells in recent and remote time points. Four series of confocal images of dendritic spines of recent PFC engrams (left) and remote PFC engrams (right) group. Arrow heads (yellow) indicate dendritic spines.

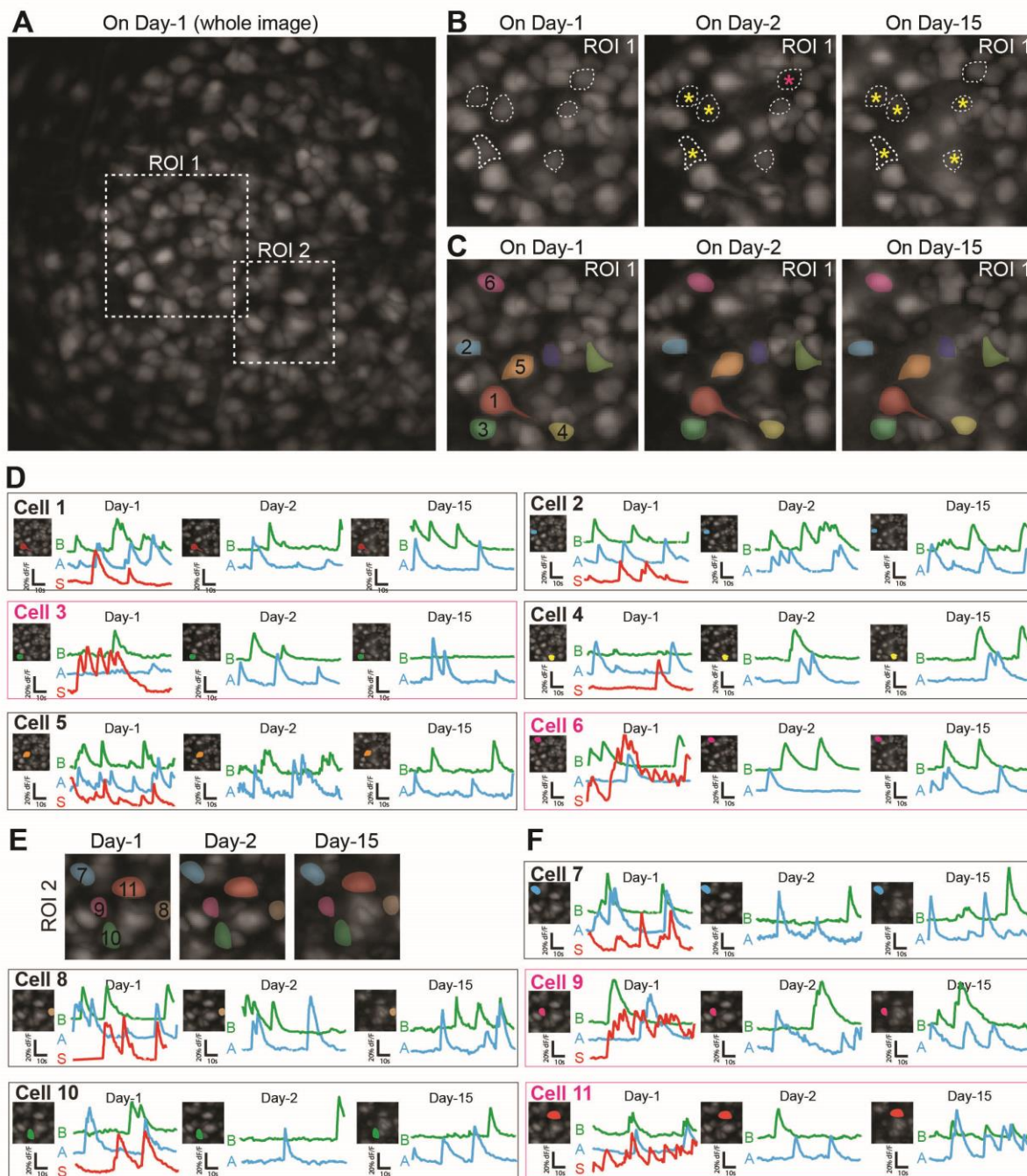


Fig. S8 Longitudinal imaging of PFC cells during CFC, recent and remote recall (A) Stacked image acquired from PFC through the microendoscope on Day-1. Region of interest (ROI). **(B)** Some cells disappear during 2 weeks imaging. Six dot lined cells were active on Day-1, but with time they lost their GCaMP6 activity. In this case, it is difficult to case individual cells through 2 weeks. Asterisks mean disappear of cells. **(C)** Stacked image acquired from PFC in ROI2 through the microendoscope on Day-1, Day-2 and Day-15. Five colored cells indicate that was successfully cased through 2 weeks. Labeled 8 cells constantly showed their GCaMP activity for 2 weeks, were successfully cased individual cells longitudinally. **(D)** GCaMP activity of PFC cells during Day-1, Day-2, and Day-15 in ROI1. B; Context-B, A; Context-A, S; shock

period. Cell 3 and Cell 6 were Shock Responding cells. **(E)** Stacked image acquired from PFC in ROI2 through the microendoscope on Day-1, Day-2 and Day-15. Five colored cells indicate that was successfully cased through 2 weeks. **(F)** GCaMP activity of 5 PFC cells during Day-1, Day-2, and Day-15 in ROI 2. B; Context-B, A; Context-A, S; shock period. Cell 9 and Cell 11 were Shock Responding cells.

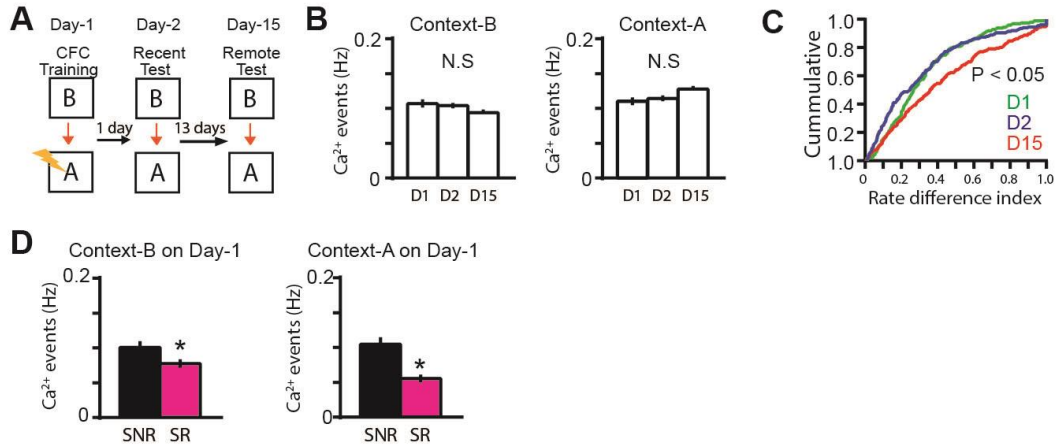


Fig. S9

GCaMP activity of PFC cells during CFC and recall periods. (A) Experimental schedule. (B) Average Ca^{2+} event frequency of PFC cells during Context-B and Context-A on Day1, 2 and15. (C) Cumulative probability of rate difference index of Ca^{2+} activity of PFC cells during Context-B and Context-A on Day1 (median; 0.2683), Day-2 (median; 0.2389) and Day-15 (median; 0.3560) ($n = 220$ cells, $N = 3$ mice). There were significant differences in the cumulative distribution curves of a rate difference index (see Methods) between Day-1 conditioning and Day-15 recall, and between Day-2 recall and Day-15 recall analyzed by Kolmogorov-Smirnov Comparison, $*P = 0.001$. (D) Average Ca^{2+} event frequency of SNR cells ($n = 28$ cells, $N = 3$ mice) and SR cells ($n = 26$ cells, $N = 3$ mice) during Context-B and Context-A on Day-1 (not including the shock delivery period). There was significant difference between SR and SNR groups in Context-B ($t_{52} = 2.11$, $P = 0.039$) and Context-A ($t_{52} = 3.99$, $P = 0.0002$). $*P < 0.05$, by two-tailed t-test. N.S means not significant. Error bars, mean \pm s.e.m.

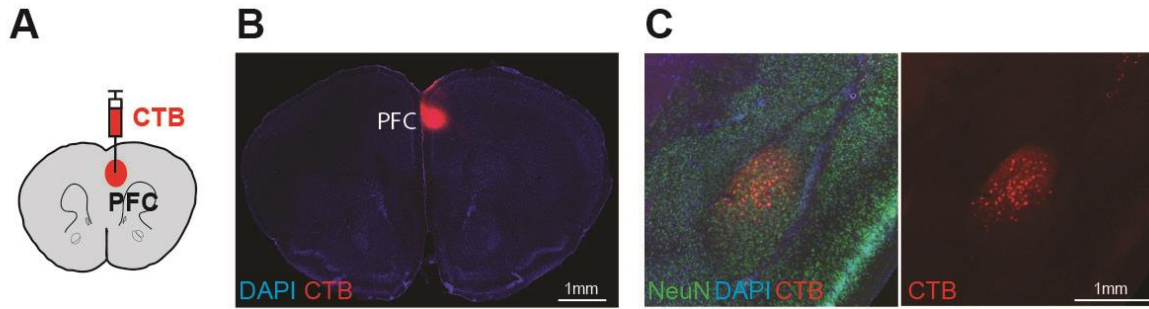


Fig. S10

BLA cells project to PFC. (A) Experimental schedule. (B-C) CTB was injected into PFC (B). Coronal section of BLA visualized with CTB-labeled cell bodies (red) and immunostained with anti-NeuN (green) and DAPI (blue) (N = 4 mice) (C).

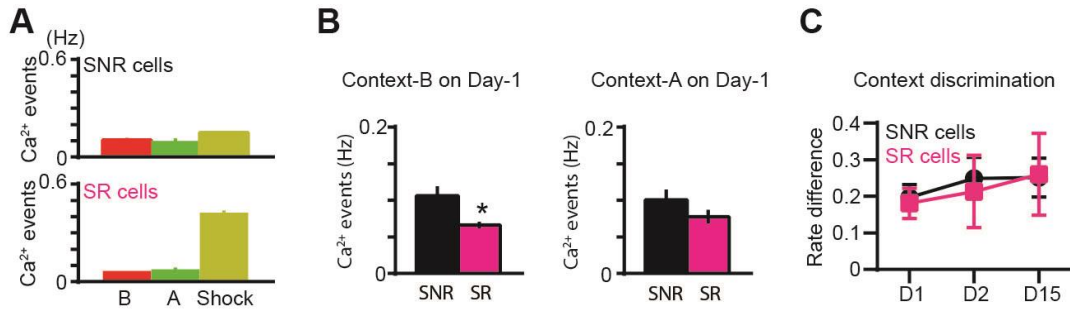


Fig. S11

GCaMP activity of PFC cells during CFC and recall periods with TeTX expression in HPC engram cells. (A) Average Ca²⁺ event frequency of SNR cells and SR cells during Context-B, Context-A and following shock periods in Context-A on Day1 under the condition which HPC engram cells express TeTX. (B) Average Ca²⁺ event frequency of SNR cells (n = 10 cells, N = 2 mice) and SR cells (n = 8 cells, N = 2 mice) during Context-B and Context-A on Day-1 (not including the shock delivery period). There was significant difference between SR and SNR groups in Context-B ($t_{16} = 2.62$, $P = 0.018$) but not in Context-A ($t_{16} = 1.29$, $P = 0.21$), analyzed by two-tailed unpaired t-test. * $P < 0.05$, by two-tailed unpaired t-test. (C) Average rate difference index of Ca²⁺ activity during Context-B and Context-A of SNR cells (black, n = 10 cells, N = 2 mice) and SR cells (pink, n = 8 cells, N = 2 mice) on Day-1, Day-2 and Day-15. There was no significant rate difference between SNR and SR cells on Day-1 ($t_{16} = 0.29$, $P = 0.78$), Day-2 ($t_{16} = 0.33$, $P = 0.74$) and Day-15 ($t_{16} = -0.08$, $P = 0.94$), analyzed by two-tailed unpaired t-test. Error bars, mean \pm s.e.m.

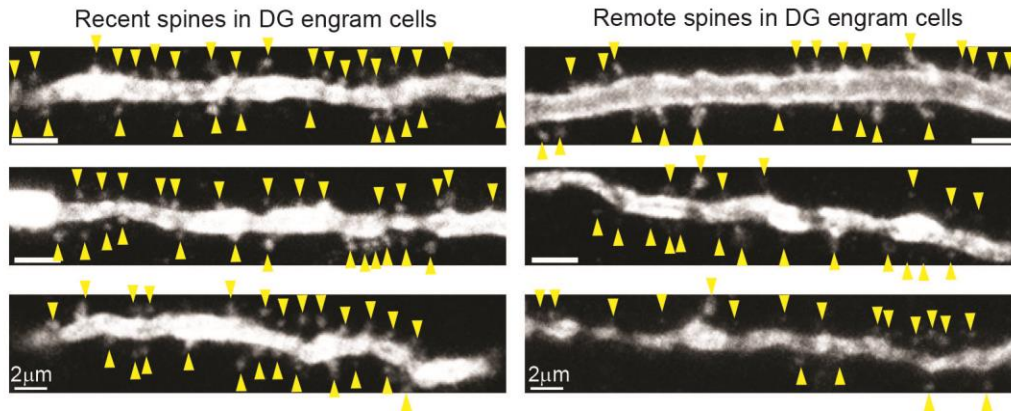


Fig S12

Confocal Images of Dendritic Spines of hippocampal DG memory engram cells in recent and remote time points. Three series of confocal images of dendritic spines of recent DG engrams (left) and remote DG engrams (right) group. Arrow heads (yellow) indicate dendritic spines.

Statistics and Sample Sizes

Figure 1:

(A) CTB injection into PFC (N = 10 mice). (B) Coronal sections of PFC with MECVa axons visualized by eYFP (N = 10 mice). (C, D) CTB injection into BLA (N = 10 mice). (E) eYFP; N = 10 mice, eArchT; N = 10 mice. There were no difference between eYFP and eArchT groups in freezing response on Day-2 ($t_{18} = -0.28$, $P = 0.78$) and Day-8 ($t_{18} = 0.39$, $P = 0.70$) analyzed by two-tailed unpaired t-test. However, there were significant differences between eYFP and eArchT groups in freezing response on Day-15 ($t_{18} = 4.85$, $P = 0.0001$) and Day-22 ($t_{18} = 2.96$, $P = 0.008$), analyzed by two-tailed unpaired t-test. (G) eYFP; N = 8 mice, eArchT; N = 10 mice. There were no difference between eYFP and eArchT groups in freezing response on Day-2 ($t_{16} = 0.026$, $P = 0.98$), Day-8 ($t_{16} = 0.74$, $P = 0.47$), Day-15 ($t_{16} = -0.82$, $P = 0.43$) and Day-22 ($t_{16} = 0.71$, $P = 0.49$), analyzed by two-tailed unpaired t-test. (H), CTB injection into ACC (N = 8 mice). (I) Coronal sections of ACC with MECVa axons visualized by eYFP (N = 10 mice). (J) eYFP; N = 10 mice, eArchT; N = 12 mice. There were no difference between eYFP and eArchT groups in freezing response on Day-2 ($t_{20} = 1.06$, $P = 0.30$), Day-8 ($t_{20} = 0.96$, $P = 0.35$), Day-15 ($t_{20} = 1.56$, $P = 0.14$) and Day-22 ($t_{20} = 1.49$, $P = 0.15$), analyzed by two-tailed unpaired t-test. (L) eYFP; N = 8 mice, eArchT; N = 8 mice. There were no difference between eYFP and eArchT groups in freezing response on Day-2 ($t_{14} = -0.84$, $P = 0.41$), Day-8 ($t_{14} = 0.05$, $P = 0.96$), Day-15 ($t_{14} = 0.15$, $P = 0.88$) and Day-22 ($t_{14} = 0.43$, $P = 0.67$), analyzed by two-tailed unpaired t-test. (O) HC; N = 4 mice, CTX; N = 4 mice, CFC with eYFP; N = 4 mice, and CFC with eArchT; N = 4 mice. There was significant differences in the percentage of c-Fos+ cells in the PFC among the 4 groups ($F(3, 12) = 5.425$, $P = 0.013$, one-way ANOVA). Tukey-Kramer test revealed there were significant difference between HC and CFC-eYFP ($P = 0.016$) and between CFC-eYFP and CFC-eArchT ($P = 0.03$) (Q) Coronal section of PFC visualized with ChR2-mCherry of ON-DOX (N = 3 mice), OFF-DOX (N = 3 mice), and OFF-DOX (N = 3 mice) conditions. (S) N = 6 mice. There were significant differences between light-OFF and light-ON conditions in freezing response on Day-2 ($t_5 = -2.89$, $P = 0.034$) and Day-12 ($t_5 = -4.69$, $P = 0.005$), analyzed by two-tailed paired t-test. (U) N = 5 mice. There was no difference between light-OFF and light-ON conditions in freezing response on Day-2 ($t_4 = -0.45$, $P = 0.67$)

and Day-12 ($t_4 = -1.55$, $P = 0.19$), analyzed by two-tailed paired t-test. **(W)**. Distribution of mCherry⁺ cells in the MEC ($n = 212$ mCherry⁺ cells, $N = 3$ mice).

Figure 2:

(C) $N = 5$ mice, each group, There was no difference in percentages of c-Fos⁺ cells between GFP⁺ and GFP⁻ cells in the condition of Recent A-A ($N = 5$ mice, total number of DAPI⁺ cells was 1,155 cells, $t_8 = -0.82$, $P = 0.43$), Recent A-B ($N = 5$ mice, total number of DAPI⁺ cells was 1,100 cells, $t_8 = -0.54$, $P = 0.60$) and Remote A-B ($N = 5$ mice, total number of DAPI⁺ cells was 830 cells, $t_8 = -0.52$, $P = 0.61$), but a significant difference in the condition of Remote A-A ($N = 5$ mice, total number of DAPI⁺ cells was 889 cells, $t_8 = 3.45$, $P = 0.0087$), analyzed by two-tailed unpaired t-test. **(E)** Day-2; $n = 507$ spines, 62 fragments of dendrites, $N = 3$ mice, Day-12; $n = 587$ spines, 62 fragments of dendrites, $N = 3$ mice, There was significant difference in spine density between recent (Day-2) and remote (Day-12) conditions, analyzed by Kolmogorov-Smirnov Comparison ($D = 0.2419$, $P = 0.044$). **(F)** $N = 10$ mice. There were significant differences between green light-OFF and green light-ON conditions in freezing response on Day-12 ($t_9 = 3.05$, $P = 0.013$), but not on Day-2 ($t_9 = 1.04$, $P = 0.32$), analyzed by two-tailed paired t-test. **(G-O)** Total 220 cells in the PFC are monitored from 3 mice. **(L-O)** SNR cells; $n = 28$ cells, $N = 3$ mice, SR cells; $n = 26$ cells, $N = 3$ mice. **(M)** There were no difference in Ca²⁺ event frequency between Context-B and Context-A on Day-1 ($t_{27} = -0.30$, $P = 0.77$), Day-2 ($t_{27} = -0.49$, $P = 0.63$), and Day-15 ($t_{27} = 1.42$, $P = 0.17$), analyzed by two-tailed paired t-test. **(N)** There was no difference in Ca²⁺ event frequency between Context-B and Context-A on Day-2 ($t_{25} = 1.13$, $P = 0.27$), but significant difference on Day-15 ($t_{25} = 2.59$, $P = 0.016$), analyzed by two-tailed paired t-test. **(O)** There were no difference in rate difference index between SNR cells and SR cells on Day-1 ($t_{52} = -0.68$, $P = 0.50$), Day-2 ($t_{52} = 1.06$, $P = 0.29$), but was significant difference on Day-15 ($t_{52} = -2.19$, $P = 0.03$), analyzed by two-tailed unpaired t-test. **(P)** Coronal sections of PFC visualized with BLA axons by eYFP ($N = 4$ mice). **(Q)** HC; $N = 4$ mice, Shock; $N = 4$ mice, CFC with eYFP; $N = 4$ mice, and CFC with eArchT; $N = 4$ mice, There was significant difference in the percentage of c-Fos⁺ cells in the PFC among the 4 groups ($F(3, 12) = 5.435$, $P = 0.014$, one-way ANOVA). Tukey-Kramer test revealed there were significant difference between HC and CFC-eYFP ($P = 0.014$) and between CFC-eYFP and CFC-eArchT ($P = 0.033$). **(R)** eYFP; $N = 6$ mice, eArchT; $N = 6$ mice. There were no difference between eYFP

and eArchT groups in freezing response on Day-2 ($t_{10} = 0.66$, $P = 0.74$), Day-8 ($t_{10} = 0.30$, $P = 0.77$) and Day-15 ($t_{10} = 1.96$, $P = 0.08$), analyzed by two-tailed unpaired t-test. However, there was a significant differences between eYFP and eArchT groups in freezing response on Day-22 ($t_{10} = 2.35$, $P = 0.04$), analyzed by two-tailed unpaired t-test.

Figure 3:

(B-C) Sagittal sections of HPC visualized with anti-VAMP2 in ON-Dox (B) and OFF-Dox (C), $N = 3$ mice, each group. **(D)** HC; $N = 4$ mice, blue-light-ON mice with eYFP; $N = 4$ mice, and blue-light-ON mice with TeTX; $N = 5$ mice. There was significant difference in the freezing responses among the 4 groups ($F(3, 10) = 7.119$, $P = 0.012$, one-way ANOVA). Tukey-Kramer test revealed there were significant difference between HC and DG-Stim ($P = 0.013$) and between DG-Stim with eYFP and TeTX ($P = 0.039$). **(F)** $N = 5$ mice, each group. There was a significant difference in percentages of c-Fos⁺ cells between GFP⁺ and GFP⁻ cells in eYFP control group ($N = 5$ mice, total number of DAPI⁺ cells was 1,051 cells, $t_8 = 2.37$, $P = 0.045$), but there was no difference in TeTX group ($N = 5$ mice, total number of DAPI⁺ cells was 1,040 cells, $t_8 = 0.87$, $P = 0.40$), analyzed by two-tailed unpaired t-test. **(G)** eYFP; $n = 1,020$ spines, 104 fragments of dendrites, $N = 3$ mice, TeTX; $n = 809$ spines, 104 fragments of dendrites, $N = 3$ mice. There was a significant difference in spine density between eYFP and TeTX groups, analyzed by Kolmogorov-Smirnov Comparison ($D = 0.2981$, $P = 0.0001$). **(H)** In SNR cells ($n = 10$ cells, $N = 2$ mice), there were no difference in Ca²⁺ event frequency between Context-B and Context-A on Day-1 ($t_9 = 0.42$, $P = 0.69$), Day-2 ($t_9 = -0.31$, $P = 0.76$) and Day-15 ($t_9 = 0.24$, $P = 0.81$), analyzed by two-tailed paired t-test. In SR cells ($n = 8$ cells, $N = 2$ mice), there were no difference in Ca²⁺ event frequency between Context-B and Context-A on Day-1 ($t_7 = -1.02$, $P = 0.34$), Day-2 ($t_7 = -0.35$, $P = 0.74$) and Day-15 ($t_7 = -0.53$, $P = 0.61$), analyzed by two-tailed paired t-test. **(L)** There were no difference in percentages of c-Fos⁺ cells between GFP⁺ and GFP⁻ cells in the condition of Recent A-B ($N = 5$ mice, total number of DAPI⁺ cells was 1,568 cells, $t_8 = -0.57$, $P = 0.59$) and Remote A-B ($N = 5$ mice, total number of DAPI⁺ cells was 1,749 cells, $t_8 = -0.42$, $P = 0.68$), but there were significant difference in the condition of Recent A-A ($N = 5$ mice, total number of DAPI⁺ cells was 2,220 cells, $t_8 = 3.02$, $P = 0.017$) and Recent A-A ($N = 5$ mice, total number of DAPI⁺ cells was 2,357 cells, $t_{10} = -2.24$, $P = 0.049$), analyzed by two-tailed unpaired t-test. **(N)** Coronal sections of DG visualized with Chr2-mCherry (red) under

ON-Dox and OFF-Dox conditions (N = 3 mice, each group). **(O)** Day-5 (recent); n = 623 spines, 60 fragments of dendrites, N = 3 mice, Day-15 (remote); n = 423 spines, 50 fragments of dendrites, N = 3 mice. There was a significant difference in spine density between recent and remote groups, analyzed by Kolmogorov-Smirnov Comparison (D = 0.26, P = 0.041). **(Q)** N = 4 mice, There were significant difference between blue light-OFF and light-ON conditions in freezing response on Day-5 ($t_3 = -4.75$, P = 0.018) and Day-15 ($t_3 = -4.57$, P = 0.02), analyzed by two-tailed paired t-test.

Figure 4:

(A) Coronal sections of BLA visualized with MECVa axons by eYFP (N = 4 mice). **(B)** eYFP; N = 12 mice, eArchT; N = 12 mice. There were significant differences between eYFP and eArchT groups in freezing response on Day-2 ($t_{22} = 3.28$, P = 0.003), Day-8 ($t_{22} = 3.72$, P = 0.001), Day-15 ($t_{22} = 2.45$, P = 0.03) and Day-22 ($t_{22} = 2.39$, P = 0.03), analyzed by two-tailed unpaired t-test. **(C)** eYFP; N = 10 mice, eArchT; N = 10 mice. There were significant differences between eYFP and eArchT groups in freezing response on Day-2 ($t_{18} = 2.80$, P = 0.01) and Day-8 ($t_{18} = 3.80$, P = 0.001), but no difference on Day-15 ($t_{18} = 0.96$, P = 0.35) and Day-22 ($t_{18} = 0.95$, P = 0.35), analyzed by two-tailed unpaired t-test. **(D)** Coronal sections of BLA visualizing axons of PFC engram cells with eYFP (N = 3 mice). **(E)** N = 10 mice. There was a significant difference between green light-OFF and green light-ON conditions in freezing response on Day-12 ($t_9 = 2.80$, P = 0.02), but not on Day-2 ($t_9 = 0.72$, P = 0.42), analyzed by two-tailed paired t-test compared to chance level. **(H)** N = 5 mice, each group. Chance level for overlap in BLA was calculated as $(c\text{-Fos}^+/\text{DAPI}^+) \times (\text{H2B-GFP}^+/\text{DAPI}^+)$. There were significant differences in the percentages of $c\text{-Fos}^+ \text{GFP}^+$ cells in BLA in recent group (Day-2; N = 5 mice, total number of DAPI^+ cells was 216 cells, $t_4 = -3.37$, P = 0.028) and remote group (Day-13; N = 5 mice, total number of DAPI^+ cells was 219 cells, $t_4 = -2.83$, P = 0.047) by two-tailed paired t-test. **(K)** N = 5 mice, each group. There were significant differences in the percentages of $c\text{-Fos}^+ \text{GFP}^+$ cells in BLA in recent group (Day-4; N = 5 mice, total number of DAPI^+ cells was 218 cells, $t_4 = -2.83$, P = 0.047) and remote group (Day-14; N = 5 mice, total number of DAPI^+ cells was 213 cells, $t_4 = -4.97$, P = 0.008) by two-tailed paired t-test.