Title: **Prefrontal multistimulus integration within a dedicated disambiguation circuit guides interleaving contingency judgment learning**

Authors:

Justin D. Pastore [1,4], Johannes Mayer [1,4], Jordan Steinhauser [1,4], Kylene Shuler [1,4], Tyler W. Bailey [2], John H. Speigel III [2], Evangelos E. Papalexakis [3], and Edward Korzus [1,2,5,*]

Affiliations:

[1] Department of Psychology, University of California Riverside, CA 92521, USA

[2] Neuroscience Program**,** University of California Riverside, CA 92521, USA

[3] Department of Computer Science and Engineering, University of California Riverside, CA 92521, USA

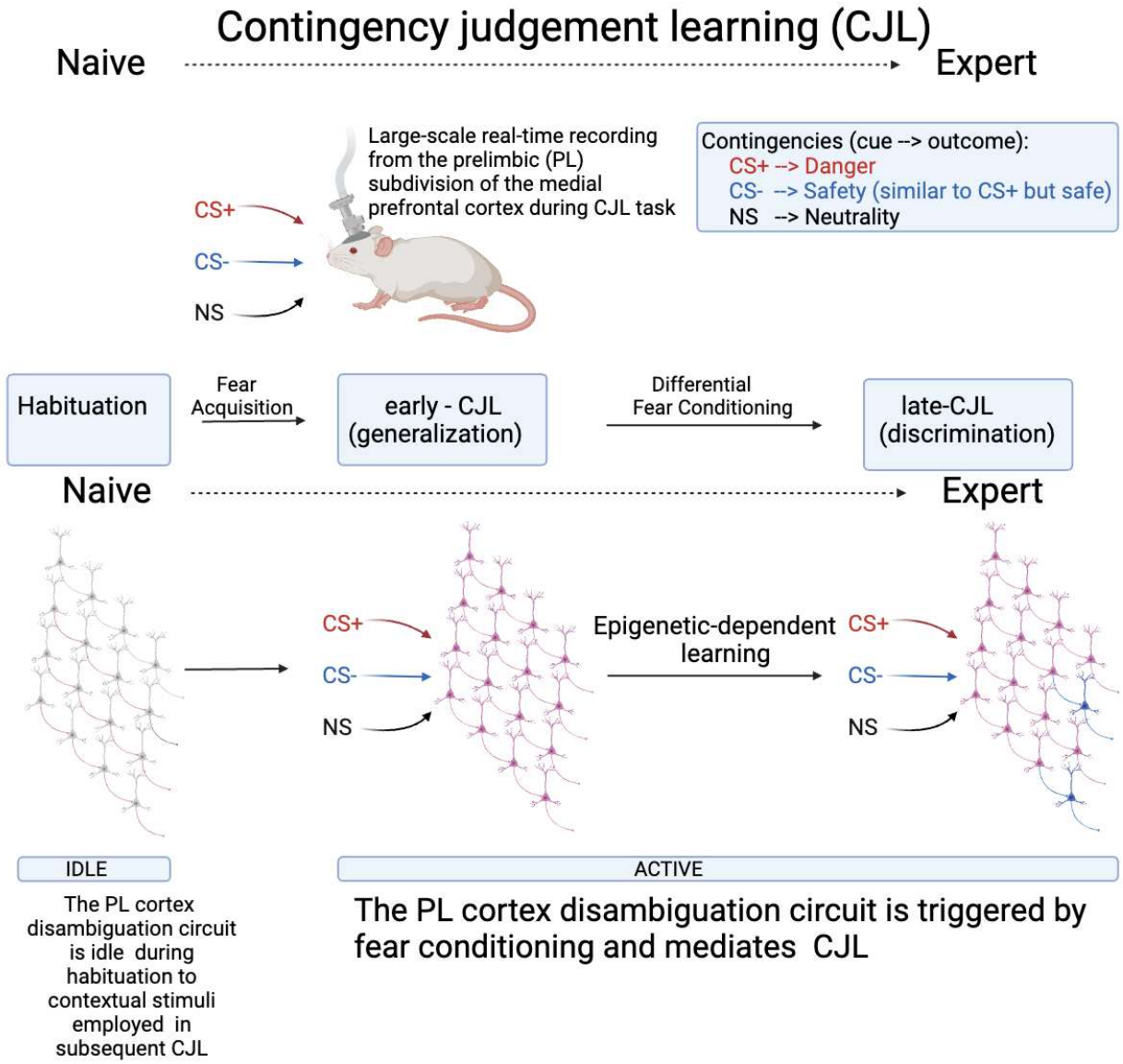
[4] These authors contributed equally

[5] Lead contact

* Correspondence: edkorzus@ucr.edu (E.K.)

Highlights:

- Tensor decomposition analytics and graph-based computational models revealed prefrontal multistimulus integration within a newly discovered prelimbic (PL) cortex disambiguation circuit, triggered by initial fear acquisition, guides interleaving contingency judgment learning (CJL).
- The prefrontal disambiguation circuit manages spurious and predictive relationships during cue–danger, cue–safety, and cue–neutrality contingencies.
- Animals with memory deficiencies in the PL cortex exhibit impaired CJL and the PL cortex disambiguation circuit function.



SUMMARY: Understanding how cortical network dynamics support learning is a challenge. This study investigates the role of local neural mechanisms in the prefrontal cortex during contingency judgment learning (CJL). To better understand brain network mechanisms underlying CJL, we introduce ambiguity into associative learning after fear acquisition, inducing a generalized fear response to an ambiguous stimulus sharing nontrivial similarities with the conditioned stimulus. Real-time recordings at single-neuron resolution from the prelimbic (PL) cortex show distinct PL network dynamics across CJL phases. Fear acquisition triggers PL network reorganization, led by a disambiguation circuit managing spurious and predictive relationships during cue–danger, cue–safety, and cue–neutrality contingencies. Subjects with PL-targeted memory deficiency show malfunctioning disambiguation circuit function, while naive subjects lacking unconditioned stimulus exposure lack the disambiguation circuit. This study shows that fear conditioning induces prefrontal cortex cognitive map reorganization, and subsequent CJL relies on the disambiguation circuit's ability to learn predictive relationships.

INTRODUCTION

Contingency judgment is widely studied in human social, legal, political, and psychological research, as it enables an effective and accurate prediction of events by inferring causal relationships. Both humans [1-3] and animals [4,5] can accurately detect contingency alongside contingency variations. However, contingency detections may be inaccurate when conflicting cues or causal illusions abound. The neural mechanism underlying contingency judgment learning (CJL), which involves the disambiguation of contextual cues followed by conversion from spurious to predictable cue-outcome relationships, is unknown. Learning to control fear is a complex behavior [6-8] that relies on CJL and has been of interest to basic and clinical researchers. Distinguishing between threats and closely related but not identical safe stimuli is vital for survival. Overcoming generalized fear is equally crucial for achieving successful behavioral performance. People with posttraumatic stress disorder (PTSD) find it difficult to subdue fear, even in safe conditions [9].

PTSD presents with impaired contextual fear processing across the hippocampus (Hip)-the basolateral amygdala (BLA)-medial prefrontal cortex (mPFC) circuit resulting in misjudgment of cue-safety *vs.* cue-danger relationships. Impaired contextual fear inhibition and overgeneralized fear are postulated to be markers for PTSD risk and treatment outcome as individuals with PTSD feel acutely threatened by generalized cues[10-13]. The hippocampus [14,15], the basolateral amygdala [16], the ventromedial prefrontal cortex, and the dorsal anterior cingulate cortex [12] have been implicated in contextual fear inhibition and PTSD risk [13]. Fear-related behavior is controlled by discrimination, generalization, and extinction[6,7,17,18]. Both rodent and human studies implicated prefrontal regions in contextual fear acquisition, expression, generalization, inhibition, and discrimination (extensively reviewed in [11,13,19-21]). While the hippocampus-amygdala-prefrontal circuit function underlying fear modulation is one of the best-understood circuit mechanisms in neuroscience, the complete understanding of mechanisms involving distinctive prefrontal regions in contextual fear modulation remains to be elucidated, and translational considerations pose concerns due to anatomical and functional differences.

Despite widespread interest in understanding the mechanisms of cortical network dynamics at a cellular resolution that underlie complex behaviors, these mechanisms remain elusive, primarily due to technological and analytical constraints. Specifically, how neurons cluster into functional groups to drive network activity and how neural network dynamic codes trigger behavioral responses remain unclear. Studies on the molecular and circuit mechanisms of how the mPFC controls fear behavior abound [6,7,13,17,18,22-33]. The prelimbic (PL) and infralimbic (IL) subdivisions of mPFC and their long-distance interaction can differentially regulate fear behavior[34-37], the result of distinct connectivity with the amygdala[38,39] and the central nucleus of the amygdala output via BLA excitatory pathways[40] or the amygdala intercalated neurons relay[41] providing an alternative mechanism for extinction[23,42,43]. Prefrontal circuitry has been strongly implicated in safety learning, the general learning process of associating a cue/stimulus with safety (e.g., PL projections to ventrolateral periaqueductal gray [44] or PL projections to BLA[30,45], PL projections to IL [46], long-range cortical interaction [47], cortical-subcortical interactions[44] and BLA projections to PL[30]). Studies have also explored prefrontal top-down control of behavior [48-51].

Considerable evidence indicates that mPFC, BLA, and Hip neurons are functionally coupled at the theta range (4-12 Hz oscillations) during fear conditioning [52,53]and discriminative fear learning[54,55]. Prior studies suggested that the PL subdivision of the mPFC monitors context–danger and context–safety relationships [24,56]. Neural correlates of fear discrimination were reported in the BLA, while abnormalities in BLA function may result in generalization [57-59]. Furthermore, PL input to the BLA is likely to drive fear specificity, while disruption of PL function during fear discrimination results in abnormal generalization [60]. Nonetheless, a fundamental question that has remained unresolved pertains to whether the prefrontal network dynamics' role in CJL and/or its dependence on the acquisition of properties within the PL network constitute essential components of a global brain mechanism that guides behavioral responses, particularly in the CJL context to choose appropriate responses to threat, safety, and neutral environmental cues.

Using head-mounted miniature microscopes (HMMs) in freely behaving mice[61,62], we studied prefrontal network dynamics at a single-cell resolution by tracking network responses to unambiguous threat-conditioned, ambiguous safety-conditioned, and unambiguous neutral stimuli across CJL learning task. Real-time recordings at a single-neuron resolution in control *vs.* genetically targeted PL hypofunction acquired from the PL cortex using HMM showed distinct PL network dynamics that manage learning on a CJL task. We used tensor decomposition analytics empowered with a graph-based computational model to demonstrate that prefrontal multistimulus integration within a dedicated disambiguation circuit, triggered by initial fear acquisition, guides interleaving CJL. Mutant animals with dysfunction limiting the PL network's ability to consolidate memories displayed an inability to learn cue-safety and cue-danger relationships associated with severe impairments of the disambiguation circuit function.

RESULTS

CJL's reliance on PL network integrity

Understanding the psychological and underlying circuit mechanisms of CJL poses theoretical and technical challenges. How brain networks manage the recognition of predictive and spurious cue-outcome relationships in response to threat and safety contingency learning remains unclear. To better understand the brain network mechanisms underlying contingency learning, we added

ambiguity to associative learning after fear acquisition in the well-studied temporal contiguity model of Pavlovian conditioning [17] by triggering a generalized fear response to an ambiguous stimulus sharing nontrivial similarity with the conditioned stimulus [26,63]. Figure 1A shows the CJL behavior paradigm used to systematically evaluate behavioral responses to the unambiguous threat-conditioned stimulus (CS+), ambiguous safety-conditioned stimulus (CS-), and unambiguous neutral stimulus (NS) across 33 trials of habituation preceding fear conditioning and differential fear conditioning. This interleaving learning strategy involving a deliberate alteration in stimulus presentations represents a learning in which cognitive schemas develop progressively through exposure, experience, and gradual integration. While CS- and NS are frequently used in contemporary research as non-reinforced control stimuli associated with "no outcome" (i.e., no US, unconditioned stimulus), a substantial theoretical and technical distinction exists between both stimuli in the current design, which demonstrates active learning of safety information. Encounters with a safety signal were treated as a significant reinforcing episode rather than an unambiguous emotionally neutral event. Notably, CS- shares nontrivial similarity with CS+, yielding an ambiguity-triggered generalization that we consider a spurious relationship as opposed to predictable cue-outcome relationships represented by CS+ (that is always reinforced with US) or predictable cue-no outcome relationships represented by NS (not associated with positive or negative valence due to a lack of meaningful similarity with CS+ or CS- and always presented with "no outcome").

To determine the function of the PL cortical network in the adult brain and specifically the role of CBP-mediated histone acetyltransferase (HAT) activity in network dynamics, we generated transgenic mice carrying a dominant negative CBP transgene (PL-CBPΔHAT) that specifically blocks HAT activity in PL cells (Fig. S1A-B). Under normal conditions, CREB phosphorylation, which depends on CBP recruitment to critical promoters, and CBP-HAT enzymatic activity are both required for memory consolidation (Fig. 1D-E) [28,64-73]. However, CBP is not considered a limiting factor due to its inherent abundance. Notably, any decrease in CBP-HAT activity selectively and dramatically lowers the probability of successful memory consolidation (i.e., conversion from short-term to long-term memory). For example, a two-amino-acid substitution mutation ($Y^{1540}/F^{1541} \rightarrow Y^{1540}/F^{1541}$) of the CBP's HAT domain (Fig. 1 D, CBPΔHAT mutant) selectively disables the binding of the primary substrate acetyl CoA and results in the inability of CBP to acetylate histone and memory consolidation [64,65]. Thus, blocking transcription on the gene promoters that require histone acetylation for synaptic activity-triggered gene expression hinders long-term memory consolidation. Cytohistological analysis of brain tissue isolated from PL-CBPΔHAT and Ctrl animals showed that conditioned PL-CBPΔHAT mice displayed decreased levels of acetylated histone H3 (AcH3, left panel, t-test PL-Ctrl vs. PL-CBPΔHAT: $p = 0.0283$, $t(20) = 2.359$, $r = 0.431$; Ctrl: 1 ± 0.1255, $n = 11$; PL-CBPΔHAT: 0.6206 ± 0.1006, $n = 14$) and decreased levels of acetylated histone H4 (Ac-H4; right panel; t-test PL-Ctrl vs. PL-CBPΔHAT: $p = 0.0338$, $t(17) = 2.303$, $r = 0.441$; Ctrl: 1 ± 0.1428, $n = 9$; PL-CBPΔHAT: 0.5746 ± 0.1172, $n = 14$) in cells expressing mCherry when compared to conditioned control animals (Fig. S1). These findings corroborated those of previous studies reporting decreased levels of acetylated histones in CBP mutant mice [28,64-73].

To assess the PL network dynamics and their role in learning CJ accuracy, we compared the CJL performance of the Ctrl mice with that of PL-CBPΔHAT mutants while assessing real-time PL network dynamics at cellular resolution using HMMs (Fig. 1A-C and Fig. 2). Behaviorally, fear responses were operationalized as the percentage of time a mouse spent freezing (i.e., motionless other than respiration and heartbeat) during the trial. Thus, higher freezing levels indicated a

greater fear response. For imaging during behavior using HMMs, all mice expressed the Ca^{2+} indicator GCaMP6f targeted to the PL neurons and had chronically installed prism lenses that allowed the collection of light signals from the PL cortex (Fig. 2A) during all 33 trials of the CJL task (Fig. 1A).

As baseline controls for contingency learning, two groups (Ctrl-w/o US and PL-CBPΔHAT-w/o US) were never exposed to US. Figure 1B–C shows the performance of mice carrying HMMs or non-functional replicas on the CJL task. Figure 1B demonstrates that HMM-carrying Ctrl and PL-CBPΔHAT mutant mice performed similarly in the contextual version of the fear conditioning (FC) task, showing strong conditioned fear acquisition (Fig. 1B; two-way RM ANOVA: US × Group: F (9, 114) = 9.104, $p < 0.0001$; US, F (2.082, 79.13) = 20.74, $p < 0.0001$; Group, F (3, 38) = 6.501, $p = 0.0012$). Sidak's comparisons between all four groups at the baseline were insignificant ($p > 0.05$). While the Ctrl and PL-CBPΔHAT groups showed strong fear acquisition, the Ctrl-w/o US and PL-CBPΔHAT-w/o US groups did not acquire fear to CS+ (Bonferroni's comparisons of Baseline vs. After 3rd US-CS+ Pairing: Ctrl, $p < 0.0001$; PL-CBPΔHAT, $p < 0.0001$; Ctrl-w/o US, $p > 0.05$; PL- CBPΔHAT -w/o US, $p > 0.05$. During the CJL portion of the paradigm, responses to NS were not significant and were similar between the test (Fig. 1C, left) [Two-way RM ANOVA: Block Trial × Group, F (9, 117) = 1.177, $p = 0.3158$; Block Trial, F (2.412, 94.08) = 1.350, $p = 0.2645$; Group, F (3, 39) = 2.618, $p = 0.0645$].

The CJL task was designed to measure the acquisition of CJ accuracy during repeated exposures to the threatening context stimulus CS+ (always paired with US) alternated with exposures to similar but not the same safe context stimulus CS- (cue-safety relationship). A three-way ANOVA showed differences in the performance of the Ctrl and PL-CBPΔHAT mice carrying HMMs during CJL (Fig. 1C, right. Ctrl and PL-CBPΔHAT, three-way ANOVA: Block x Group x Stimulus, F (3, 90) = 5.0778, P=0.0027; Block, F (3, 90) = 1.2842, P=0.2847; Group, F (1, 30) = 3.7600, P=0.0619; Stimulus, F (1, 30) = 30.337, P<0.0001; Block x Group, F (3, 90) = 0.37795, P=0.7691; Block x Stimulus, F (3, 90) = 7.4868, P=0.0002; Group x Stimulus, F (1, 30) = 0.60056, P=0.4444). Initially, the Ctrl and PL-CBPΔHAT mice generalized their conditioned responses and exhibited similar freezing levels to both the CS+ and CS- contexts (Fig. 1C, Tukey's correction for multiple comparisons: Ctrl, CS- *vs.* CS+: Block 1, $p > 0.05$; Block 2, $p > 0.05$; PL-CBPΔHAT, CS- *vs.* CS+, Block 1, $p > 0.05$; Block 2, $p > 0.05$). However, the Ctrl animals began to freeze significantly less in response to context CS- than to context CS+ during Blocks 3 and 4 trials of training (Fig. 1C; Ctrl, CS- *vs.* CS+: Block 3, $p < 0.0001$; Block 4, $p < 0.0001$), demonstrating high CJ accuracy during late (L)-CJL. Unlike the Ctrl group, the PL-CBPΔHAT mice failed to distinguish between context CS- and CS+ and continued to show bias towards a generalized CS+ response during each of the four CS- block trials on the CJL task (Fig. 1C, PL-CBPΔHAT, CS+ *vs.* CS-: Block 3, $p > 0.05$; Block 4, $p > 0.05$), demonstrating again that PL-CBPΔHAT mice have a strong deficit in CJL.

Additionally, average learning curves for learning responses to CS+ and CS− were calculated based on the performance of the Ctrl and PL-CBPΔHAT groups across the CJL training, followed by fitting the regression line to the CS+ curve (red) and the CS−curve (blue) (Fig. 2C, Block trials 1–4). The analysis of patterns of responses to CS+ and CS- in Ctrl animals showed that the improvement of the contingency judgment accuracy (the slopes CS+ *vs*. CS- are different: F=14.12, p=0.0027) was due to a decline in freezing responses to CS− (CS− slope, $p = 0.0027$), while the slope of the CS+ curve did not deviate from zero (P > 0.05, Fig. 2C, Ctrl - Block trials 1–4). The PL-CBPΔHAT group, which failed to improve contingency judgment

accuracy, showed no difference between the slopes of the CS+ and the CS− curves (F=0.510, p=0.489), and the slopes of both curves were not different from zero (P > 0.05, Fig. 2C, PL-CBPΔHAT - Block trials 1–4). In addition, the slopes CS+/Ctrl and CS+/PL-CBPΔHAT are not different (F=2.2, p=0/165). Ctrl-w/o US and PL-CBPΔHAT-w/o US showed no fear to any tested stimuli because these mice never acquired fear as a result of not being exposed to US at any time.

These data show that the CJL paradigm is effective in capturing transitions from spurious cue-outcome relationships (CS- responses during early (E)-CJL) to predictive cue-outcome relationships (CS- responses during L-CJL). In contrast, stable responses to CS+ (high) and NS (low) remain predictable during the entire CJL task (see additional details below). These findings indicate that neural hypofunction in the PL circuitry can result in imbalanced neural processes underlying CJL, suggesting that PL network might be the locus of learning contingency judgment accuracy.

Imaging of prefrontal circuit dynamics during CJL at single-neuron resolution

To assess real-time PL network dynamics associated with the performance on the CJL task, we extracted Ca^{2+} transients of neurons activated in the PL cortex in response to context CS+, CS-, and NS stimulus presentation trials across CJL from normal mice carrying HMMs (Ctrl) and compared them with Ctrl-w/o US mice and PL-CBPΔHAT-w/o US mice that were never exposed to US but otherwise underwent a full training protocol and with mutant PL-CBPΔHAT mice, which showed a strong deficit in CJL (Fig. 1C). Using this approach, we extracted the activity profile of individual neurons within CJL trials, reflecting rapid large-scale neural dynamics during responses to contextual stimuli CS+, CS-, and NS. Thus, tracking changes in large-scale neural dynamics on a trial-to-trial basis while comparing CS+ *vs.* CS- triggered rapid dynamics should show long-term circuit plasticity associated with safety learning.

Due to postsurgical tissue scaring, not all mice shown in (Fig. 1B-C) were used for calcium imaging acquisition (see Methods). Figures S2A and S2B show the performance of mice (Ctrl *vs.* PL- CBPΔHAT *vs.* Ctrl-w/o US *vs.* PL-CBPΔHAT-w/o US) used for Ca2+ imaging during CJL and network dynamics analysis (see Methods). Performance of all four groups of mice (Ctrl, PL-CBPΔHAT, Ctrl-w/o US, PL-CBPΔHAT-w/o US) used for Ca2+ imaging (Fig. S2) was similar to larger cohorts described previously (Fig. 1B-C). Figure S2A showed that Ctrl and PL-CBPΔHAT mice acquired fear, while animals that were not treated with US (Ctrl-w/o US and PL-CBPΔHAT-w/o US) showed no fear during the entire behavioral procedure (Fig. S2A-B). Analysis of CJL learning curves (CS+ vs. CS-) in Ctrl showed strong learning (Fig. S2B, Ctrl: the slope of Ctrl/CS- descended significantly, Y = -2.895*X + 65.24, F=6.663, P=0.0417; the slope of Ctrl/CS+ ascended significantly, Y = 1.666*X + 24.88, F=10, P=0,0195; and the slopes Ctrl/CS+ *vs.* Ctrl/CS- were different F=13.55, P=0.0031). However, PL- CBPΔHAT showed no learning on the CJL task (Fig. S2B, PL-CBPΔHAT: the slope of PL-CBPΔHAT /CS- did not deviate from zero, Y = -0.6589*X + 35.50, F=0.5102, P=0.5019; the slope of PL-CBPΔHAT CS+ did not deviate from zero, Y = -0.1134*X + 38.41, F=0.0309, P=0,8662; and the slopes PL-CBPΔHAT /CS+ *vs.* PL- CBPΔHAT /CS- were not different, F=0.2350, P=0.6366). In addition, a three-way ANOVA showed differences in the performance of the Ctrl and PL-CBPΔHAT mice used for Ca^{2+} imaging during CJL (Fig. S2B; Ctrl *vs.* PL-CBPΔHAT, three-way ANOVA Block x Group x Stimulus, F (3, 39) = 2.0803, P=0.1185). Tukey's correction for multiple comparisons showed that Ctrl mice learned to discriminate between CS+ and CS- during late-CJL (Fig. S2B,

Ctrl, CS- *vs.* CS+: Block 4, $P < 0.05$) after initially showing fear generalization during early-CJL (Fig. S2B, Ctrl, CS- *vs.* CS+: Block 1, $P > 0.05$). PL-CBPΔHAT showed deficits in CJL learning (Fig. S2B, PL-CBPΔHAT: CS- *vs.* CS+, Blocks 1, 2, 3 and 4, $p > 0.05$).

Ctrl-w/o US and PL-CBPΔHAT-w/o US showed no fear to any tested stimuli because these mice never acquired fear as a result of not being exposed to US at any time (Fig S2B-right, Ctrl-w/o US Two-way RM-ANOVA Block × Stimulus, no effect of Block, no effect of Stimulus and no interaction, $p > 0.05$. PL-CBPΔHAT-w/o US Two-way RM-ANOVA Block × Stimulus, (Ctrl-w/o US Two-way RM-ANOVA Block × Stimulus, no effect of Block, no effect of Stimulus and no interaction, $p > 0.05$).

Thus, expression of the inhibitor of long-term memory consolidation (i.e., CBPΔHAT) targeting PL neurons abolished learning on the CJL task. The deficit observed in PL-CBPΔHAT mice was found to be specific to learning contingency judgment accuracy. These mutant animals acquired and generalized fear to the same levels as Ctrl animals (Fig. 2SA-B) and distinguish between different contexts similarly to Ctrl (e.g., CS+ *vs.* NS during E and L); however, subtle differences between CS+ and CS- led to an inability to transition from spurious to predicted cue-no outcome relationships. This resulted in biased contingency judgment, leading to the generalization of fear.

In vivo imaging of Ca^{2+} transients using miniature fluorescence microscopes can provide insights into network activity in the brain [33,57,62,74-77]. Figure 2A shows a coronal representation of GCaMP6f expression in PL neural populations with a laterally positioned microendoscopic lens. Representative images from a PL-CBPΔHAT mouse confirmed PL targeting with GCaMP6f (Fig. 2A-B) and CBPΔHAT.mCherry (Fig. 2A-B), as well as the overlap of GCaMP6f and mCherry expression (Fig. 2A). We also found that GCaMP6f+ve cells also expressed NeuN (Fig. 2C), confirming that infected cells were neurons. Neural spatial footprints and neural temporal traces of individual Ca^{2+} transients were retrieved from motion-corrected images collected at 15 Hz after inspection for stability across the stack of frames. Figure 2D shows recordings of the PL populations' calcium activity during exposure to fear-conditioned context stimulus CS+ (before and after US) applying a constrained matrix factorization to accurately separate the background, followed by demixing and denoising the neural signals (see Methods, Fig. S3, Fig. S4). Figure 2D delineates unsorted Ca^{2+} traces of the general population, with the spatiotemporal components sorted into two subpopulations of neurons out of the cell range that are categorized based on US-responding or their unspecific response (Fig. 2D), confirming that some PL neurons respond directly to US. The accuracy of neural spatial footprints between trials was supervised using the cell registration code [78] (Fig. S5), which is described in the Methods section. The data processing protocol described in the Methods section provides an efficient and accurate extraction of single neural activities from endoscopic video data, allowing us to simultaneously track well-isolated neural signals within noisy and overlapping data regimes of neurons.

Global network activity is stable across CJL learning trials

The phenotype of the PL-CBPΔHAT mutant mice clearly demonstrated that PL integrity is indispensable for CJL. Our initial hypothesis was that there would be an obvious measurable change in global PL network dynamics, i.e. a detectable change in manifest variables of a PL state, explaining a moment of engagement of PL during CJL. With the ability of real-time large-scale neuronal activity assessment in PL using HMMs, we first tested global network activity across CJL trials. To gain insights into PL network dynamics, we tracked changes in large-scale

neural dynamics on a trial-to-trial basis while comparing CS+ *vs.* CS- *vs.* NS -triggered PL network responses.

The average Ca^{2+} spike rates did not vary significantly in Ctrl (Fig. 2E, top: two-way ANOVA of Block Trial × Stimulus, F (8, 84) = 1.077, P=0.3869; Block, F (2.329, 48.90) = 3.412, P=0.0345; Stimulus, F (2, 21) = 1.259, P=0.3044. Tukey's multiple comparisons showed no differences between consecutive trials for any of the three stimuli: CS+ responses between trial, $p > 0.05$; CS- responses between trials, $p > 0.05$; and NS responses between trials, $p > 0.05$ or between responses to any of the three stimuli (CS+ vs. CS- vs. NS) in any block trial, $p > 0.05$). Similarly, the PL-CBPΔHAT group also did not show any differences between specific trials across CJL in stimulus-triggered average Ca^{2+} spike frequency (Fig. 2E, bottom: two-way ANOVA of Block Trial × Stimulus, F (8, 72) = 0.6385, P=0.7428; Block, F (3.335, 60.04) = 9.686, P<0.0001; Stimulus, F (2, 18) = 1.693, P=0.2121). Tukey's multiple comparisons showed no differences between trials for any of the three stimuli: CS+ responses between trials, $p > 0.05$; CS- responses between trial, $p > 0.05$; and NS responses between trials, $p > 0.05$ or between responses to any of the three stimuli (CS+ *vs.* CS- *vs.* NS) in any block trial, $p > 0.05$.

Investigations showed that average Ca^{2+} spike rates did not vary significantly between all four tested groups (Fig. S6A, Ctrl *vs.* PL-CBPΔHAT *vs.* Ctrl-w/o US *vs.* PL-CBPΔHAT-w/o US, one-way ANOVA: $p > 0.05$). No significant difference was observed in the average Ca^{2+} spike rates between block trials (Fig. S6B-E, top) and within block trials (Fig. S6B-E, bottom) of CJL during exposures to different stimuli (CS-, CS+ or NS) in each of the tested four groups: Ctrl (Fig. S6B, $p > 0.05$), PL-CBPΔHAT (Fig. S6C, $p > 0.05$), Ctrl-w/o US (Fig. S6D, $p > 0.05$), and PL-CBPΔHAT-w/o US (Fig. S6E, $p > 0.05$).

While global PL network activity was not varied across trials (Fig. 2E), another important question was if the PL network's stimulus selectivity of the PL network was changing across CJL that could explain learning to distinguish between cue–danger and cue–safety predictive relationships. To assess potential network bias towards specific stimuli signaling spurious or predictive relationships during cue–danger and cue–safety contingencies across different phases CJL, we first evaluated selective tuning toward CS- or CS+ for recorded neurons by calculating the Discrimination Index (DI, see Methods). The DI was found by using Ca^{2+} spike rate responses to CS- and CS+ and plotting the cumulative probability distribution of DIs during habituation (H) or E- or L-phase trials of CJL in Ctrl (Fig. 2F, top) and compared to PL-CBPΔHAT (Fig. 2F, bottom). The Kolmogorov-Smirnov (K-S) test was used to compare cumulative frequency distributions of DIs across CJL learning. The K-S test did not show any changes in network selectivity towards stimuli (CS+ and CS-) in Ctrl during transition from habituation to E-CJL (Fig. 2F, left-top; Ctrl, K-S test; H *vs.* E; D = 0.026, p = 0.4537) nor during transition from early- to late-CJL (Fig. 2F, right-top; Ctrl, K-S test; E *vs.* L; D = 0.0385, p = 0.1098). PL-CBPΔHAT mice showed minuscule distortion of cumulative frequency distribution patterns of DIs during transition from habituation to early-CJL (Fig. 2F, left-bottom; PL-CBPΔHAT, K-S test; H *vs.* E; D = 0.074, p < 0.0001) and during transition from early- to late-CJL (Fig. 2F, right-bottom; PL-CBPΔHAT, K-S test; E *vs.* L; D = 0.076, p < 0.0001).

The decomposition of CJL-associated PL network dynamics data shows learning-specific latent neuronal patterns.

While behavioral data showed that the PL network is critical for CJL (Fig. 1B-C and Fig. S2A-B), initial investigations showed that manifest variables representing global PL network states during learning trials failed to explain CJL. Specifically, average Ca^{2+} spike rates and their cumulative frequency distributions remained stable across learning trials (Fig. 2E-F and Fig. S6). We, therefore, hypothesized that there are latent variables that represent model states that could explain CJL. Our experimental design organized neural Ca^{2+} spike rates into, at a minimum, a data array with three axes: neuron factor (neuronal populations), trial factor, and temporal factor (time within trial) (Fig. 3A). This data array is essentially a tensor (the CJL neural tensor), which is a data representation that has been successfully applied in analyzing and extracting patterns from high-dimensional high-order data. In order to show latent neuronal activity patterns that reflect interpretable trial variables (trial conditions and fear levels), we utilized tensor decomposition-based analytics [79], which is a set of analytical optimization tools that express a given tensor as a sum of its constituent components, each one of which representing a pattern that exists in the data (see Methods).

The Canonical Polyadic Decomposition (CPD) model in Ctrl showed that during the CJL task, the PL network was driven by three tensor components (TC1, TC2, TC3) that collectively explained behavioral correlates (Fig. 3, also see Methods). The first tensor component (TC1) pattern exhibited a habituation-specific response, while TC2 appeared to integrate responses to multiple stimuli but was explicitly relevant to CJL. TC1 showed neurons most responsive to all stimuli within the habituation preconditioning phase. TC2 showed neurons most responsive to CS− and CS+ stimuli postconditioning during the early, middle, and late phases of differential fear conditioning. TC3 showed neurons most responsive to the NS stimulus postconditioning during the early, middle, and late phases of differential fear conditioning. Interestingly, TC2 was triggered by fear conditioning and appeared to manage responses to multiple stimuli during all CJL trials that are critical to CJ and relevant to learning appropriate responses to safe stimuli. Thus, we will refer to this population as the PL cortex disambiguation circuit. Remarkably, these populations captured during the interleaving CJL protocol in the PL network were more likely to manage different phases of learning underlying complex behavior than typically observed neural representations associated with specific stimuli, such as those observed in sensory cortices or the hippocampus [57,74,80-84]. Thus, The CPD model detected latent populations (TC1, TC2, TC3) and their activity patterns that reflect interpretable trial variables (trial conditions and fear levels) of the CJL task in Ctrl mice. TC2, or the disambiguation circuit, integrates responses to multiple stimuli and guides CJL.

Community detection using ND-graph model shows distinctive learning-specific neuronal populations in PL during CJL

The CJL tensor decomposition model suggests that three major neural components can explain the PL network dynamics underlying CJL. However, tensor decomposition is usually used to identify latent variable models rather than address finite sample issues, and assessing the meaningful size of components remains a challenging and unresolved issue [85,86] and is rarely implemented using unconventional approaches[87,88]. To gain deeper insights into the discrete temporal structure of prefrontal activation patterns during learning using a stochastic graph-based analysis [89], which was applied to the PL network activation patterns at a single-neuron resolution across all trials of the CJL task in Ctrl mice. We hypothesized that if there are CJL-specific latent neuronal populations guiding learning within the PL network, the graph-based

computational modeling of PL network activity at a single-neuron resolution should be able to capture these latent populations both qualitatively and quantitatively and show their pattern of activity correlating with CJL. By considering neurons as vertices and their temporal interactions as edges, graph-based analytics allow visualization of large neural network recordings within and across trials so that quantitative testing of theoretical concepts relevant to network dynamics and behavior can be assessed. Supplementary Figure S8A shows the construction of an exemplar adjacency matrix for an individual trial used for graph-based computational modeling of PL network activity during CJL. The neural network dynamics graph (ND-graph also see Methods) is an application-specific weighted directed graph in which we used a 66 milliseconds time binning size and a sliding time window set to 666 milliseconds to capture the temporal features of the PL network dynamics at a cellular resolution underlying CJL (Fig. S8A). The weight reflects the time interval between neural activation events (i.e., Ca^{2+} spikes), and the probability of neural coactivation is defined as the value of summed weights. To gain insights into temporal relationships between individual neurons and/or groups of neurons sharing similar temporal activation patterns (referred to as neural communities or communities), we modeled assembly activation patterns across CJL trials. Thus, each temporally aggregated adjacency matrix representing each of the trials was all-trials temporally aggregated (summed together), collapsing all trials into a single adjacency matrix (Fig. 4A). In the next step, the Louvain Community algorithm [90] was applied to ND-graphs to determine the optimal community structure subdivision of the network with nonoverlapping neural clusters in the PL networks during CJL of the Ctrl mice. Figure 4B shows a representative image of the qualitative analysis outcome of the community detection that shares a general similarity to tensor decomposition-based computational analysis (Fig. 3B). Three latent neural groups (i.e., NC1, NC2, and NC3) were detected, showing a distinctive pattern of activity across CJL trials. Intriguingly, the NC2 community was triggered by fear conditioning and was responsive to both CS+ and CS- during CJL in a similar fashion as the TC2 component (Fig. 3B). This general qualitative similarity (Fig. 3B vs. Fig. 4B) is striking because applied tensor decomposition and graph-based community detection analytics are substantially different computational approaches. In addition, Fig. 4C showed a substantial level of overlap between neuronal population detected using the ND-graph model approach (Fig. 4B. NC1, NC2, and NC3) and CPD model approach (Fig. 3C. TC1, TC2, and TC3) in the same Ctrl mouse.

Disambiguation circuit activity explains learning of contingency judgment

The Louvain community detection algorithm (Fig. 4A) yielded a qualitative evaluation of PL network dynamics (Fig. 4B) and quantitative aspects of communities' activity patterns across CJL trials, allowing testing of the main hypothesis that PL network dynamics guides CJL. To address this hypothesis, we tracked prefrontal dynamics in Ctrl mice and compared them with PL-CBPΔHAT mutant mice displaying unsuccessful learning during CJL. First, we performed a detailed analysis of activation patterns of neuronal communities across the CJL trials in Ctrl mice. This analysis method outperforms other computational approaches owing to its temporal scope, as it enables the capturing of the dynamics of subtle temporal structures formed by populations across a large number of active cells. Figure 5A-B shows detailed patterns of neural dynamics for each of the three detected communities (NC1, NC2, NC3) plotted as a percent of total network activity during 10 trials (i.e., 5 block trials: Habituation (Block -1), E-CJL (Blocks 1-2), and L-CJL (Blocks 3-4)) for each tested stimulus (CS-, CS+, NS) in Ctrl mice compared to PL-CBPΔHAT, Ctrl-w/o US, and PL-CBPΔHAT-w/o US mice.

Figure 5A illustrates changes in responsiveness to the tested stimuli in three discovered neural communities in the Ctrl group. The patterns of neural dynamics across trials of CJL recorded in the Ctrl group were consistent with those observed in the tensor decomposition model (Fig. 3C). The NC1 activity pattern correlated specifically with habituation trials, with no bias toward any of the tested stimuli. The NC1 activity level declined after CJL started (Fig. 5A-left, NC1: Fig. 5A, Ctrl, Two-way RM ANOVA: Block Trial x Stimulus, F (8, 84) = 1.131, P=0.3513; Block Trial, F (1.420, 29.81) = 399.2, P<0.0001; Stimulus, F (2, 21) = 0.1569, P=0.8558. Bonferroni's multiple comparisons test: CS-: Block −1 *vs.* 1 or 2 or 3 or 4, $p < 0.05$. CS+: Block −1 *vs.* 1 or 2 or 3 or 4, $p < 0.05$. NS: Block −1 *vs.* 1 or 2 or 3 or 4, $p < 0.05$). In addition, NC1 shows no bias toward any tested stimuli (CS-, CS+, and NS) during any of the block trials across the entire CJL task (Fig. 5A, Bonferroni's multiple comparisons test: CS- *vs.* CS+ *vs.* NS at all Block Trials, $P > 0.05$).

Unlike NC1, the NC2 activity rapidly increased in responses to CS+, CS-, and NS after FC (Fig. 5A-middle, NC2: Ctrl, Two-way RM ANOVA: Block Trial x Stimulus, F (8, 84) = 2.782, P=0.0088, Block Trial, F (2.053, 43.11) = 80.06, P<0.0001, Stimulus, F (2, 21) = 4.625, P=0.0217
Bonferroni's multiple comparisons test: CS-: Block −1 *vs.* 1 or 2 or 3 or 4, $p < 0.05$. CS+: Block −1 *vs.* 1 or 2 or 3 or 4, $p < 0.05$. NS: Block −1 *vs.* 1 or 2 or 3 or 4, $p < 0.05$. CS+, CS-, and NS are cues that should signal a high probability of distinctive future circumstances such as danger, safety, and neutrality during L-CJL, respectively. We hypothesized that the CJL task requires temporal (across-trial) and multistimulus integration during memory updates and consolidation because animals use subtle differences between context CS+ and CS- across many days with a single exposure to each context only once per day during CS- cue - safety relationship judgment learning that requires 4–8 days of training. Thus, it was imperative to compare the responses of FC-triggered NC2 to specific cues (i.e., CS+, CS-, and NS) during each specific block trial after FC.
Figure 5A shows an unexpected pattern of NC2 responses to these three distinctive cues, demonstrating that responses to CS- and CS+ were indistinguishable during all block trials of CJL; however, relative responses to NS were variable but consistent with learning predictive (i.e., accurate) CS- cue - safety relation. Bonferroni's multiple comparisons test showed that NC2 responses to CS- were not different from those to CS+ at all block trials (CS- *vs.* CS+: $p > 0.05$ for all block trials). During E-CJL, NC2 showed significantly lower responses to NS than to CS- (CS- *vs.* NS: Block 1, $P < 0.05$; Block 2, $P < 0.05$) and CS+ (CS+ *vs.* NS: Block 1, $P < 0.05$). However, responses between all three stimuli (CS-, CS+, and NS) became indistinguishable at late-CJL during block trials 3 and 4 (Block 3 / CS+ *vs.* NS: $P > 0.05$, CS- *vs.* NS: $P > 0.05$; Block 4 / CS+ *vs.* NS: $P > 0.05$, CS- *vs.* NS, $P > 0.05$). These findings suggest that acquiring predictive CS-/safety relations may involve encoding neutral cues-no outcome relations within the NC2 circuit. PL appears to be learning contingencies, and maybe CSs are prioritized (i.e., have a significantly greater percentage of the network activity compared to NS) in the beginning (E) because they are more challenging to differentiate during the early learning phase (with the more spurious relationship for CS-) and less processing is required for the unambiguous NS but then through learning by the end of CJL (L) once animals are showing successful discrimination, all three stimuli are processed similarly by this community because each predictive relationship was learned successfully (i.e., CS- predicts safety, CS+ predicts danger, and NS predicts neutral). Thus, NC2 tracked responses to unambiguous threat-conditioned, ambiguous safety-conditioned, and unambiguous neutral stimuli across 33 trials of habituation, fear conditioning, and differential fear conditioning (Fig. 1A, C). As the NC2 dynamic pattern explains neural

computations associated with successful CJL learning, community NC2 defines the disambiguation circuit similarly to TD2 (Fig. 3C).

The community NC3 dynamic pattern appeared to correlate with coding neutrality (Fig. 4B and 5A-right) into the PL network. A two-way RM ANOVA of Block Trial × Stimulus showed effect of Block Trial (Fig. 5A – NC3: Ctrl, two-way RM ANOVA: Block Trial x Stimulus, F (8, 84) = 1.445, P=0.1899; Block Trial, F (1.688, 35.46) = 5.493, P=0.0115; Stimulus, F (2, 21) = 4.453, P=0.0244). Bonferroni's multiple comparisons test showed that NC3 show biased towards NS after fear conditioning during E-CJL, which suggests that NC3 may code initial NS recognition in the PL network (Fig. 5A-right, Bonferroni's multiple comparisons test: NS: Block −1 *vs.* 1 and Block −1 *vs.* 2, $P < 0.05$.

We performed the same community analysis in the PL-CBPΔHAT mice and compared the results with those of the Ctrl mice. Similarly, we found three distinctive communities: NC1, NC2, and NC3 in PL-CBPΔHAT (Fig. 5B), with a marked group difference in the activity patterns of NC2. Fig. 5B shows that PL-CBPΔHAT's NC1 exhibits a very similar pattern to the Ctrl's NC1. The NC1 activity pattern correlated specifically with habituation trials (Fig. 5B-left, NC1: PL-CBPΔHAT, Two-way RM ANOVA: Block Trial x Stimulus, F (8, 72) = 0.9277, P=0.4991, Block Trial, F (2.252, 40.53) = 239.7, P<0.0001; Stimulus, F (2, 18) = 0.3904. P=0.6824. Bonferroni's multiple comparisons test: CS-: Block −1 *vs.* 1 or 2 or 3 or 4, $p < 0.05$. CS+: Block −1 *vs.* 1 or 2 or 3 or 4, $p < 0.05$. NS: Block −1 *vs.* 1 or 2 or 3 or 4, $p < 0.05$). In addition, NC1 shows no bias toward any tested stimuli (CS-, CS+, and NS) during block trials across the entire CJL task (Fig. 5B-left, Bonferroni's multiple comparisons test: CS- *vs.* CS+ *vs.* NS at all Block Trials $P > 0.05$.

In PL-CBPΔHAT mice, the NC2 showed an abnormal pattern of activity, explaining the inability of PL-CBPΔHAT mice to learn during the CJL task properly. In these mutant mice expressing inhibitor of long-term coding in the PL network, the NC2 community rapidly increased in responses to CS+ and CS- after FC but only during E-CJL (Fig. 5B-middle, NC2: PL-CBPΔHAT, Two-way RM ANOVA: Block Trial x Stimulus, F (8, 72) = 0.3716, P=0.9323; Block Trial, F (1.483, 26.70) = 15.89, P=0.0001; Stimulus, F (2, 18) = 1.024, P=0.3790). Bonferroni's multiple comparisons test showed that the NC2 community showed elevated responses to CS+ and CS- only during E-CJL (NC2/CS-: Block −1 *vs.* 1 or 2, $p < 0.05$. NC2/CS+: Block −1 *vs.* 1 or 2, $p < 0.05$) but NC2 responses to CS+ and CS- during L-CJL were not distinguishable from responses before fear conditioning (NC2/CS-: Block −1 *vs.* 3 or 4, $P > 0.05$; NC2/CS+: Block −1 *vs.* 3 or 4, $P > 0.05$). The responses of NC2 to NS in PL-CBPΔHAT mice remained the same through all CJL trials (NC2/NS, Block Trial −1 *vs.* 1 or 2 or 3 or 4, $P > 0.05$). In addition, NC2 shows no bias toward any tested stimuli (CS-, CS+, and NS) during any of the block trials across the entire CJL task (Fig. 5B-middle, Bonferroni's multiple comparisons test: CS- *vs.* CS+ *vs.* NS at all Block Trials, $P > 0.05$).

The PL-CBPΔHAT mice showed a strong deficit in the function of the disambiguation circuit (i.e., NC2). The PL-CBPΔHAT's NC2 failed to engage during responses to NS, CS-, and CS+ during late CJL, coinciding with a deficiency in learning contingency judgment accuracy. Unlike the Ctrl's NC2, PL-CBPΔHAT's NC2 was not (or weakly) responsive to any stimuli during L-CJL, suggesting that long-term plasticity within PL is necessary for appropriate coding of the disambiguation circuit. Thus, the disambiguation circuit was ineffective in managing network

responses to contextual cues in PL-CBPΔHAT mice, yielding failure in learning contingency judgement accuracy.

Analysis of NC3 in PL-CBPΔHAT did not show any changes in pattern of activity across all behavioral trials (Fig. 5B-right, NC3: PL-CBPΔHAT, two-way RM ANOVA: Block Trial × Stimulus, F (8, 72) = 0.1617, P=0.9951, Block Trial, F (1.537, 27.66) = 5.134, P=0.0190; Stimulus, F (2, 18) = 0.8114, P=0.4598). Bonferroni's comparisons test found no significant differences between the Block Trials in responses to CS-, CS+, and NS (CS-: Block −1 *vs.* 1 or 2 or 3 or 4, $p > 0.05$. CS+: Block −1 *vs.* 1 or 2 or 3 or 4, $p > 0.05$. NS: Block −1 *vs.* 1 or 2 or 3 or 4, $p > 0.05$). In addition, NC3 in PL-CBPΔHAT did not show any meaningful difference in responses to CS-, CS+, and NS during any of the block trials (Fig. 5B-right, Bonferroni's multiple comparisons test: CS- *vs.* CS+ *vs.* NS at all Block Trials, $P > 0.05$).

Ctrl-w/o US (Fig. 5C) and PL-CBPΔHAT-w/o US (Fig. 5D) mice were not treated with US and did not acquire any relevant cue–outcome contingencies during behavioral testing. Instead, the detected NC1, NC2, and NC3 communities in these mice showed no changes in average activity across all block trials. This confirms that the disambiguation circuit provides a neural platform for developing learning-specific cognitive schema in subjects after experience with aversive stimulus during fear conditioning.

We have also performed group statistics on CPD models generated for each mouse separately in each tested groups: Ctrl (Fig. S10A), PL-CBPΔHAT (Fig. S10B), Ctrl-w/o US (Fig. S10C) and PL-CBPΔHAT-w/o US (Fig. S10D) and found similar results as compared to the ND-graph model. Like the ND-graph model (Fig. 5), the CDP model's TC1 component was selectively active during habituation trials for all three stimuli in the Ctrl mice (Fig. 10A, left). The Ctrl's TC2 component dynamics showed very strong alignment with the learning of contingency judgment accuracy (Fig. S10A, middle), in a similar fashion as NC2 (the PL cortex disambiguation circuit) detected in the ND-graph (Fig. 5A, middle). The Ctrl component TC3 dynamic pattern appeared to correlate with coding neutrality into the PL network in both models (Fig. 5A, right and S10A, right). Unlike the Ctrl's TC2, PL-CBPΔHAT's TC2 was not (or weakly) responsive to any stimuli during L-CJL (Fig. S10B), suggesting that long-term plasticity within PL is necessary for appropriate coding of the disambiguation circuit. Thus, the PL cortex disambiguation circuit defined using the CPD model (TC2) was ineffective in managing network responses to contextual cues in PL-CBPΔHAT mice, yielding failure in learning contingency judgment accuracy.

Ctrl-w/o US (Fig. S10C) and PL-CBPΔHAT-w/o US (Fig. S10D) mice were not treated with US and did not acquire any relevant cue–outcome contingencies during behavioral testing. Instead, the detected TC1, TC2, and TC3 components in these mice showed no changes in average activity across all block trials. This confirms that the disambiguation circuit provides a neural platform for developing learning-specific cognitive schema in subjects after experience with aversive stimulus during fear conditioning.

The pattern of activity of the PL disambiguation circuit (TC2) recovered in Ctrl explains CJL (Fig. S10A). Control mice showed a fully functional PL disambiguation circuit (TC2) managing stimulus responses and guiding successful CJL. In contrast, PL-CBPΔHAT showed a deficiency in CJL, coinciding with impairment of the performance of the PL disambiguation circuit (TC2) during late-CJL (Fig. S10B). In Ctrl, TC2 manages network responses to CS+, CS-, and NS after

FC but not before. During E-CJL, TC2 discriminates between NS stimuli and CS+/CS-, but during L-CJL, it does not; thus, TC2 dynamics across trials explain CJL learning similarly as was seen in the ND-graph model. TC1 component showed strong responses during habituation without bias towards specific stimuli in Ctrl and PL-CBPΔHAT; however, TC1 component remains highly active across all Block Trials in animals not treated with US. TC3 component acquired (transiently) bias toward NS during Block Trials 1 and 2 in Ctrl mice but was not significantly active across block trials in PL-CBPΔHAT.

The ND graph captures a static snapshot of neuron co-firing behavior while preserving temporal information on the edges and, more specifically, the weights of those edges, with neurons that co-fire more closely in time having a higher edge weight. On the other hand, the tensor decomposition treats every time moment equally and does not place any higher emphasis on neurons that fire more closely together in time. As a result, the two different approaches, even though they both try to capture temporal dynamics, do so in different ways, driven by the nature of the approach. Thus, we are expecting that different approaches may highlight somewhat different sets of neurons as part of each distinct pattern (component or community); however, the core population per pattern, as we observe experimentally, will be the same (Fig. 4C, Fig. S9). Remarkably, both CPD model and ND-graph model showed similar neuronal population dynamics (NC1 *vs.* TC1, NC2 *vs.* NC3, and NC2 *vs.* TC3) in Ctrl mice (Fig. 5A and Fig. S10A) and strong impairment of the PL cortex disambiguation circuit (TC2 or NC2, respectively) in PL-CBPΔHAT mutant mice (Fig. 5B and Fig. S10B) further supporting that the coding in PL that underlies mechanisms directing long-term memory consolidation is critical for encoding the disambiguation circuit within PL. Thus, the distinctive dynamics of the PL cortex disambiguation circuit that explained CJL was confirmed in two independent computational models.

DISCUSSION

Remarkably, the community analysis performed on Ctrl mice recovered three characteristic communities (NC1 – habituation-specific, NC2 – the disambiguation circuit, and NC3 – neutrality recognition) that showed striking resemblance in temporal patterns to the three tensor components (TC1, TC2, and TC3, respectively). Detailed analysis of detected communities and their activity patterns reflect interpretable trial variables (trial conditions and fear levels) and explain learning on the CJL task. The current data suggest that three neural communities organize PL network responses during contingency learning in discrete temporal activation patterns. Graph models demonstrated that responses of the disambiguation circuit to tested stimuli (CS+, CS-, and NS) varied significantly between E-CJL and L-CJL. At the same time, CJL-deficient PL-CBPΔHAT mutant mice showed significant and specific disruption in the function of the disambiguation circuit. In contrast, the function of the other two components driving PL network dynamics remained mostly intact. This finding suggests that the observed PL network dynamics may arise from encoding safety signals within the disambiguation circuit. Interestingly, CJL coincides with the simultaneous coding of new information about neutrality (NS), or cue-no outcome, into the disambiguation circuit. Therefore, the reconstruction of temporal relationships of functional populations embedded in the PL network across trials, followed by comparative analysis of CS-, CS+, and NS stimuli-triggered patterns across all CJL trials, showed discrete temporal patterns in the PL network associated with coding signals relevant to learning realistic contingency judgment.

These data suggest that CJL involves assessing and integrating information about real outcomes in response to ambiguous and unambiguous stimuli within the PL cortex disambiguation circuit during interleaving learning. Current data suggest that the PL cortex disambiguation circuit may be pivotal in controlling a general fear circuit and regulating fear modulation and fear responses. Generalized fear appears to be a primary phenotype of PL-CBPΔHAT mutant mice harboring hypofunction targeted to PL network. Noteworthy, generalized fear is also a well-defined symptom in disorders of trauma and anxiety and is associated with a hyperactive amygdala and a hypoactive mPFC [91]. Prior studies indicated that neural signatures of fear discrimination are also present in the BLA [58 59], and mice with abnormal BLA inhibitory circuits generalize fear [92 93]. Thus, the PL cortex disambiguation circuit dynamics described in the current study fit the PL – BLA direct interactions model as a critical component of fear control.

The complexity of fear-related CJL is underscored by evidence indicating that neurons in the mPFC, BLA, and hippocampus are functionally coupled at the theta range during fear conditioning [52,53], conditioned extinction [94], and discriminative fear learning [54]. Furthermore, prefrontal memory encoding and consolidation support CJL. In fact, well-defined key molecular mediators of synaptic plasticity and long-term memory consolidation in the mPFC, such as the N-methyl-D-aspartate receptor, endocannabinoid-dependent long-term plasticity, CREB, and CREB-binding protein's (CBP) intrinsic HAT activity, are all required for successful discriminative fear learning [26-28], consistent with the idea that new memory encoding within the prefrontal network drives the attainment of CJ accuracy.

The described neural mechanism underlying CJL is surprisingly different from well-studied redundancy reduction during hippocampal neural computations underlying orthogonalization of overlapping features in context stimulus discrimination [63,95,96], where each context stimulus triggers distinctive stimulus-specific neural representations [57,74,80-84] with variable levels of overlap. Hippocampal neurons can code context and space, but strong evidence exists that aversive stimuli present in the environment can trigger the remapping of their firing field [97]. This may contribute to storing memories of aversive experiences in the hippocampus [98]. In addition, the memory linking mechanism may explain the hippocampal integration of information about different events [62]. Thus, stimuli are likely distinguished before information is relayed to the prefrontal and amygdala networks.

Current data provide strong evidence that supports the idea that the PL network may be involved in coding functional context-danger associations during contextual fear conditioning. The CJL task requires temporal (across-trial) and multistimulus integration during memory updates and consolidation. Learning that CS- is distinguishable by subtle differences from CS+ and is predictive of safety rather than threat requires multiple training days and exposure to each context. In addition, these data demonstrate that the population, referred to as the disambiguation circuit, manages cue–danger, cue–safety, and cue-neutrality relationships. The disambiguation circuit is likely a locus of neural computations underlying the learning of CJ accuracy during CJL because this circuit manages predictive relations after learning and spurious relationships in E-CJL. Furthermore, the detected changes in responsiveness of the disambiguation circuit to different stimuli across CJL explained the learning of CJ accuracy (Fig. 5A). These data also confirm the prediction that network science can provide a robust stochastic approach to analyzing datasets capturing brain network activity [99] at a single-neuron resolution, where the functional and anatomical connectivity between neurons can be modeled to explain temporal and spatial features of cortical network dynamics, adaptation, and information coding.

Monitoring network activity at a single-neuron resolution over many days empowered with informative computational modeling allows precise tracking of brain network dynamics during interleaving learning of predictive cue-behavioral output relationships relying on cross-trial, cross-stimuli, and cross-modality prefrontal integration of input information from critical brain networks, including the amygdala and hippocampus.

Limitations of the study

Due to the scope of the study, this manuscript focuses on PL network dynamics and does not cover other critical components of the fear modulation circuit, such impact of projections from the amygdala, and hippocampus, interactions with nucleus accumbens, locus coeruleus, dopaminergic and cholinergic systems but complete understanding of PL network dynamics during CJL will require integration our results with the past and future discoveries in the field to generate comprehensive model of CJL. The major limitation of Ca2+ Imaging is temporal resolution. Nevertheless, this powerful technology is transformative in studying neural assembly behavior by allowing for unparalleled real-time imaging of circuit dynamics at cell resolution from many genetically modified and/or regulated neurons in highly specialized brain regions such as PL over weeks. Current studies were carried out using male mice; therefore, interpretations are limited because it is unclear if observed results are sex dependent. While rodent models are essential to understanding circuitry underlying defensive behavior, one cannot ignore the limitations of studying the role of the prefrontal network in non-primate animal models for abnormal fear responses and anxiety because of enormous differences in structure and cognitive function of prefrontal cortex between the primate and non-primate PFC (e.g., primate-specific high-level regulatory strategies aiming at coping with anxiety, also see [100,101]). However, even though the rodent PFC is not well differentiated (structurally and functionally), it's critical interactions across brain regions underlying fear (i.e., with the thalamus, amygdala and hippocampus) are evolutionarily conserved.

Figure Legends

Fig. 1. PL network manages learning on CJL task.

(A) Experimental design for contingency judgment learning (CJL). The entire behavioral task comprises three phases: habituation, fear conditioning, and CJL, further divided into early (E)-CJL (Block Trial 1) and late (L)-CJL (Block Trial 4) (for details see Methods)

(B) Fear conditioning. The Ctrl and PL-CBPΔHAT groups showed similar baseline and robust performance during fear conditioning, while the mice that were never exposed to US (Ctrl-w/o US and PL-CBPΔHAT-w/o US) failed to acquire fear.

(C) Left, all four groups (Ctrl, PL-CBPΔHAT, Ctrl-w/o US, PL-CBPΔHAT-w/o US) showed similar responses to NS throughout the CJL. Right, the Ctrl mice exhibited robust leaning on the CJL task and learned accurate contingency judgment, while PL-CBPΔHAT failed to learn the task. The change in freezing response to stimuli across training was calculated as the average freezing per a block trial. However, the PL-CBPΔHAT mice failed to learn appropriate contingency judgment and showed strong generalization during all four block trials (PL-CBPΔHAT, CS+ *vs.* CS-, Block 1–4, $p > 0.05$).

(D) Mechanism of histone acetylation-dependent memory consolidation (for details see [64]). This model suggests that CBP-mediated histone acetylation during learning-triggered transcriptional activation is a critical step in the molecular mechanism controlling memory stabilization. Initial steps include induction of CREB phosphorylation, CBP activation, and CBP-mediated histone acetylation at a specific transcriptional unit in response to the initial synaptic events associated with learning. Subsequently, prolonged elevated transcription required for memory consolidation could be maintained by CBP- and CREB phosphorylation-independent nuclear mechanisms even after signals to CREB and CBP are no longer present. This transient transcriptional activation would remain active until the competing phosphatase- and deacetylase-dependent repression mechanism shut off transcription.

(E) PL-CBPΔHAT mutant mice express CBPΔHAT with eliminated acetyltransferase activity (for details see [65]). Based on mutagenesis studies demonstrating that single amino acid substitutions in the acetyl coenzyme A (acetyl-CoA) binding domain of acetyltransferases result in loss of their enzymatic activity, these studies have employed a CBP dominant negative mutant (CBPΔHAT) harboring a substitution mutation of two conserved residues (Y1540/F1541 to A1540/ A1541). This mutant has no intrinsic HAT activity.

Ctrl, n = 18, PL-CBPΔHAT, n = 14, Ctrl-w/o US, n = 5, PL-CBPΔHAT-w/o US, n = 5. The data represent the mean ± SEM, and n represents the number of mice per group. Significance values were set at $p < 0.05$: *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$; ****, $p < 0.0001$ and ns indicates not significant.

Fig. 2. Head-mounted microscopes (HMMs) effectively capture PL network dynamics at a cellular resolution.

(A) Coronal representation of virus injection sites (left) and GCaMP6f expression in PL with a laterally positioned microendoscopic lens. The middle, representative coronal brain section shows targeting of PL (red, mCherry) with PL-CBPΔHAT, with HSV.LS1L.CBPΔHAT.mCherry vector expressing CBPΔHAT and mCherry (red) + AAV1.CamKII.Cre + AAV1.SynGCaMP6f (or control viruses AAV1.Syn.GCaMP6f + HSV.LS1L.mCherry + AAV1.CamKII.Cre). Right, a representative coronal brain section shows the PL cell population expressing GCaMP6f (green) in PL with a laterally positioned microendoscopic prism lens.

(B) Representative image indicating PL infection of the viruses for GCaMP6f (left), CBPΔHAT.mCherry + Cre (middle), and overlap of GCaMP6f and mCherry expression (right). Labeling indicated that most GCaMP+ infected cells were also mCherry+ infected cells. White scale bars indicate 50 μm.
(C) Representative image indicating PL infection of the viruses for GCaMP6f (left), immunohistochemistry staining of neuronal marker NeuN (middle), and overlap of GCaMP6f (green) and NeuN (red) staining (right). White scale bars indicate 50 μm. Immunohistochemistry of NeuN staining showed that most GCaMP6f+ infected cells are neurons. To determine the pattern of GCaMP6f-tagged and mCherry-tagged virus expression, the imaged tissue was compared to the Paxinos and Franklin mouse atlas (Paxinos & Franklin, 2019), and areas of maximal GCaMP6f and mCherry expression were labeled as injection sites.
(D) A representative analysis of PL neurons from a single mouse demonstrates neuronal responses to external stimuli (US). Left-top: spatial footprints of the cell population (blue) responding to CS+ during early-CJL. Left bottom: spike raster of the PL population corresponding to the general population of cells shown in blue above. The corresponding Ca^{2+} traces of red (right-top) and green (right-bottom) footprints (left-top) are shown in red and green, respectively. Right-top: calcium transient time traces of cells from the same dataset as on the left responding to US. Right-bottom: cells responding to US are marked in red, while cells not responding to US are marked in green. Black lines below Ca^{2+} traces correspond to inferred Ca^{2+} spikes. US, red dashed line. Ctrl, n = 8, PL-CBPΔHAT, n = 7.
(E) There were no differences in stimulus-triggered average Ca^{2+} spike frequency between trials across CJL in Ctrl. Similarly, PL-CBPΔHAT group also did not show any differences between specific trials across CJL in stimulus-triggered average Ca^{2+} spike frequency
(F) The analysis of global PL network activity showed network stability and lack of bias towards tested stimuli (CS+ and CS-) across CJL (i.e., trials H, E or L) in Ctrl mice (top). PL-CBPΔHAT mice (bottom) showed modest distortions in the global network stability across CJL. The index of discrimination between CS+ and CS- responses for specific neurons was calculated as follows: Index of Discrimination = (CS- – CS+) / (CS- + CS+), where CS- and CS+ represent Ca^{2+} spikes frequencies of individual neurons in response to CS- or CS+, respectively, during H, E or L Block Trials. Kolmogorov-Smirnov (K-S) test was used to compare cumulative frequency distributions of neuronal response bias towards CS- vs. CS+ across CJL learning. Number of binned values: Ctrl-H, n = 2373, Ctrl-E, n = 2011; Ctrl-L, n = 1903; PL-CBPΔHAT-H, n = 1910; PL-CBPΔHAT-E, n = 2144; PL-CBPΔHAT-L, n = 1871. The data represent the mean ± SEM, and n represents the number of mice per group. Significant values were set at $p < 0.05$: *, $p <$ 0.05, ns indicates insignificant.

Fig. 3. The CPD model detected latent neuronal populations and their activity patterns that reflect interpretable trial variables (trial conditions and fear levels) of the CJL task in a single Ctrl mouse.

The three-component tensor decomposition model of neuronal activity across all trials of the CJL task shows latent components including unveiling the PL cortex disambiguation circuit.
(A) Pipeline of the tensor decomposition analytics. Spike raster (unsorted) of all extracted neurons from CNMF-E tracked across all trials with CellReg (neurons × time) were reshaped into a rank 3 tensor (i.e., CJL neural tensor) of shape (I, neurons × J, trial × K, time within trial), i.e.,. Finally, CPD (Canonical Polyadic Decomposition) model was used to create 3 factors:

neuron factor [N], temporal factor [S] (within-trial), and trial factor [T]. The Multiplicative Updates implementation (provided in the Tensor Toolbox for MATLAB [102]) was used to fit the non-negativity-constrained CPD. For details, see Methods.
(B) CJL Tensor decomposition showed three latent populations in the PL network. Ca2+ spike recording from 1153 PL neurons of a single Ctrl mouse collected across all 33 trials of CJL was used as input data set to pipeline of the tensor decomposition analytics shown in Fig. 3A and, subsequently, to visualize N × S × T factors. Red vertical lines (shown on T factor plots) separate individual trials lasting 200 s each, while blue dots indicate 10 s timestamps within each trial (see methods). Entire CJL paradigm (Fig. 1A) involved 11 presentations of each stimulus (CS-, CS+, NS). (C) Ca2+ spike raster of detected tensor components (TC1, TC2, TC3) found in a single Ctrl mouse as a result of analytics shown in Fig. 3B (see text Methods).

Fig. 4. The ND-graph model detected the disambiguation circuit in the PL cortex (see text).
A) A pipeline of recovered neuronal communities from trial-aggregated ND-Graph. Extracted calcium spikes per trial are converted into time-aggregated adjacency matrices per trial via ND-Graph. All time-aggregated adjacency matrices are further aggregated into a single trial-aggregated adjacency matrix. Louvain community analysis was performed on a trial-aggregated ND-Graph. The presented adjacency matrices are examples with 25 nodes.
(B) A three-community ND-graph model uncovered latent populations (NC1, NC2, NC3) and their activity patterns that reflect interpretable trial variables (trial conditions and fear levels) of the CJL task in a Ctrl mouse. A characteristic pattern of NC1, NC2, and NC3 activity across CJL trials is remarkably similar to TC1, TC2, and TC3 activation patterns recovered in the three-components tensor decomposition model (Fig. 3C).
(C) Comparison of CPD model and ND-graph model based on data obtained from a single brain (shown in Fig. 3C and Fig. 4B). The overlap between components and corresponding community was calculated as follows:
% Overlap (intersect) between a tensor component (TC) and a corresponding community (NC) = 100 * [(TC ∩ NC) / total number of TC].

Fig. 5. Activity patterns of recovered latent neural communities.

Activity patterns of recovered latent neural communities (NC1, NC2 and NC3) across trials of CJL calculated using the NG-graph model in experimental groups: Ctrl (A), PL-CBPΔHAT (B), Ctrl-w/o US (C), and PL-CBPΔHAT-w/o US (D). The pattern of activity of the disambiguation circuit (NC2) recovered in Ctrl explains the learning of contingency judgment accuracy. Control mice showed a fully functional disambiguation circuit (NC2) managing stimulus responses and guiding successful CJL. In contrast, PL-CBPΔHAT showed a deficiency in CJL, coinciding with impairment of the performance of the disambiguation circuit. In Ctrl, NC2 manages network responses to CS+, CS-, and NS after FC but not before. During E-CJL, NC2 discriminates between NS stimuli and CS+/CS-, but during L-CJL, it does not; thus, CJL learning correlates with the acquisition of new responsiveness (i.e., to NS) and properties of NC2. NC1 community showed strong responses during habituation without bias towards specific stimuli in Ctrl, n = 8, PL-CBPΔHAT. However, the NC1 community remains highly active in animals not treated with US. NC3 community-acquired bias toward NS during Block Trial 2 in Ctrl mice
The NG-graph model was constructed per each mouse independently. The percentage of network activity in response to each stimulus per trial was calculated per mouse before the group statistic

was performed. Percent of network activity was defined as the ratio between the number of Ca2+ spikes originating from a defined PL neuronal population during a specific trial (stimulus presentation) and the total number of Ca2+ spikes elicited from all recorded PL neurons in individual animals during that same trial. The data represent the mean ± SEM, and n represents the number of mice per group. Ctrl, n = 8, PL-CBPΔHAT, n = 7, Ctrl-w/o US, n = 5, and PL-CBPΔHAT-w/o US, n = 4. Significant values were set at $p < 0.05$: *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$; ****, $p < 0.0001$ and ns indicates not significant.

Star*Methods

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
rabbit Anti-acetyl-Histone H3 Antibodies	Millipore	Millipore Cat# 06-599, RRID:AB_2115283
rabbit Anti-acetyl-Histone H4 Antibodies	Millipore	Millipore Cat# 06-866, RRID:AB_310270
donkey anti-Rat IgG (H+L) - highly cross-adsorbed, biotin conjugate,	Life Technologies Corporation	Thermo Fisher Scientific Cat# A18749, RRID:AB_2535526
Streptavidin, Alexa Fluor® 555 conjugate, Life Technologies Corporation,	Life Technologies Corporation	Thermo Fisher Scientific Cat# s21381, RRID:AB_2307336
Anti-NeuN, clone A60	Millipore	Millipore Cat# MAB377, RRID:AB_2298772
goat Alexa Fluor® 568 Goat Anti-Mouse IgG	Life Technologies Corporation	Thermo Fisher Scientific Cat# A-11031, RRID:AB_144696)
Bacterial and Virus Strains		
AAV1.Syn.GCaMP6f.WPRE.SV40	U Penn Core	AV-1-PV2822; RRID:Addgene_100837
HSV-hEF1a-LS1L-CBPΔHAT-mCherry	MIT Viral Core	Custom Preparation
HSV-hEF1a-LS1L-mCherry	MIT Viral Core	RN413
pENN.AAV1-CaMKIIa-Cre	U Penn Core	AV-1-PV2396; RRID:Addgene_105558
Experimental models: organisms/strains		
C57BL/6 mice	Taconic	C57BL/6NTac
Software and Algorithms		
MATLAB	Mathworks, NA	R2021a
Prism 10	GraphPad	version 10
ImageJ	Schneider and Rasband [103]	Version: 1.54a
Community detection	Brain Connectivity Toolbox	Rubinov, M., and Sporns, O. (2010). Neuroimage *52*, 1059-1069.
Cell Registration (CellReg) code	Sheintuch et al., [78]	https://github.com/zivlab/CellReg

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, *Edward Korzus* (edkorzus@ucr.edu).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- Raw files are available from the lead contact upon request.
- This paper does not report the original code
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS:

Mice:

The UC Riverside Institutional Animal Care and Use Committee approved all procedures following the NIH guidelines for the care and use of laboratory animals. C57BL/6 mice from Taconic were used for the study. The mice housed four animals per cage, with same-sex littermates. They had *ad libitum* access to food and water and were maintained on a 12 h light/dark cycle. Old bedding was exchanged for fresh autoclaved bedding every week. Male C57BL/6 mice housed with ovariectomized companion females were used for current studies. The studies used 2-5 months old mice.

Acknowledgments:

Funding:

The Army Research Office Grant Number W911NF-23-1-0145 (EK)

National Institutes of Health grant R01 MH106617 (EK)

Brain and Behavior Research Foundation grant (EK)

National Science Foundation under CAREER grant no. IIS 2046086 (EEP)

Acknowledgments:
The research was sponsored by the Army Research Office and was accomplished under Grant Number W911NF-23-1-0145. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Office or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes, notwithstanding any copyright notation herein.

Author contributions:

Justin D. Pastore, conceptualization, data curation, formal analysis, validation, investigation, visualization, methodology, writing – methodology, writing – review and editing.
Johannes Mayer, conceptualization, data curation, formal analysis, validation, investigation, visualization, methodology, writing – methodology, writing – review and editing.
Jordan Steinhauser, investigation, data curation, formal analysis, validation, writing – review and editing.
Kylene Shuler, investigation, data curation, formal analysis, validation, writing – review and editing.
Tyler W. Bailey, investigation, data Analysis, visualization, validation.
John H. Speigel III, investigation, data analysis.
Evangelos E. Papalexakis, conceptualization, resources, supervision, funding acquisition, writing – review and editing.
Edward Korzus, conceptualization, designed all experiments, resources, data curation, formal analysis, visualization, supervision, funding acquisition, validation, writing – original draft, project administration.

Declaration of Interests: The authors declare no competing interests.

Ethics:
All experimental procedures were approved by the Institutional Animal Care and Use Committee of the University of California Riverside.

METHOD DETAILS

Behavioral Testing

Figure 1A shows an outline of the behavioral testing schedule. The entire behavioral task comprises three phases: habituation, fear conditioning, and CJL, further divided into early (E)-CJL (Block Trial 1) and late (L)-CJL (Block Trial 4). During the habituation phase, mice were placed in two similar but contextually different contexts: CS+, CS-, and a third neutral context (NS). During habituation, the mice did not show any fear responses. The mice received three US (foot shocks) during fear conditioning in the context CS+. During 8 days of CJL, the mice received three trials daily: exposure to NS followed by exposures to CS+ and CS-. The behavioral response as a measure of the cue-outcome relationship was assessed as a level of fear (i.e., freezing) on a scale of 0%–100% to multiple environmental cues: 1) contextual CS+ stimulus always paired with unconditioned stimulus (US, electric foot shock); 2) contextual inhibitory CS- stimulus never reinforced but similar to CS+; and 3) contextual NS that was substantially different from threat (CS+) and safety (CS-) stimuli.

After being handled, individual mice were exposed once daily to each of the three context stimuli used in the protocol - Context A, Context B, and neutral context (NS) - during habituation trials over three days before fear conditioning. Context A was a modified fear conditioning box (Coulburn Instruments Inc.) comprising a vanilla scent, staggered floor bars, an upsweep tone (2-9kHz), and striped wall inserts and was placed inside a sound-attenuated chamber with the house light and house fan turned on. Contexts A and B were similar but distinct. Context B was an unmodified fear conditioning chamber with a lemon scent, uniform floor bars, and a constant 2.8kHz tone. It was placed inside a sound-attenuated chamber with the house light and house fan turned on. NS was similar to the home cage. Context A, Context B and NS are illustrated in Fig. S2. Fear conditioning training was done in the fear conditioning box (CS+) from Coulburn Instruments Inc. Individual mice were exposed to a fear conditioning box for 180 s and received three-foot shocks (0.75 mA, 2 s) at a 180-s interval (3x US-CS+ pairings). The animals were left for another 180 s after the last US inside the chamber. Freezing was scored and analyzed automatically. Performance was scored by measuring freezing behavior, defined as the complete absence of movement (Fanselow, 1980). Freezing was also scored and analyzed automatically using a video-based system (Freeze Frame software, ActiMetrics Inc.). The FreezeFrame software calculated the difference between consecutive frames by comparing the gray scale values for each pixel in a frame. Freezing was defined based on the experimenter's observations and set as sub-threshold activity for longer than 1 second. Freezing was expressed as % freezing, which was calculated as the percentage of freezing time per total time spent in the testing chamber. The chamber was cleaned in between trials with 70% ethanol and distilled water. Four days after fear conditioning, the mice were subjected to CJL. During the CJL, individual mice were exposed once per day to each of the tested context stimuli for 8 days. NS for 242 s, and two hours later, to Context A for 180 s, received a 0.75 mA, 2 s foot shock, and left for another 60 s inside the chamber. Two hours later, the mice were exposed to a similar Context B for 242 s and received no foot shock. The context in which animals received the US was referred to as CS+, and the context in which animals were never shocked was referred to as CS− (similar to CS+ but not the same) or NS (home cage). The CS+ and CS− were counterbalanced.

Calcium imaging

Head-mounted miniature microscopes (HMMs) (also known as miniscopes) and related parts were built and used according to the open-source project (Resource: http://miniscope.org/) [62,77,104]. Imaging calcium transients with a miniscope system comprises a four-step protocol.

1) To prepare mice for *in vivo* imaging during behavior, each mouse underwent the first surgical procedure of viral injection at 9–12 weeks. The mice were anesthetized with isoflurane/oxygen (induction: 3.5% / 2 LPM; maintenance: 1%–1.5% / 0.8 LPM), fixed in a robot stereotactic frame (Neurostar, Kopf Instruments), and stabilized at a body-core temperature of 36°C (probed temperature-controlled heat pad) and a respiratory rate of 1 Hz. Stereotaxic coordinates in mm from the bregma were PL (AP + 2.1, ML ± 0.37, DV 2.1) at a 15° angle (in the coronal plane, toward lateral) to avoid disruption of the PL integrity. The calcium-sensitive fluorescent protein GCaMP6f (Chen et al., 2013) was expressed using AAV virus (AAV1.Syn.GCaMP6f.WPRE.SV40, 6.9×10^{12} GC/ml)-mediated gene delivery to PL. In addition, CBPΔHAT expression in PL (PL-CBPΔHAT) was controlled by bilateral injection of a mixture of HSV.CBPΔHAT.mCherry (4×10^7 GC/ml) and AAV.CamKII.Cre (~5×10^{12} GC/ml). Bilateral injection of HSV.mCherry (4×10^7 GC/ml) and AAV.CamKII.Cre (5×10^{12} GC/ml) into PL served as a control for CBPΔHAT (Ctrl). While AAV1.Syn.GCaMP6f.WPRE.SV40 was injected unilaterally, all other viruses were injected bilaterally. Injection (33 G beveled steel needle; WPI) of 200 nl of the virus suspension (AAV: 6.9 × 10e12 GC/ml. HSV: ~4 × 107 GC/ml) into each PL of mPFC yielded stable expression of GCaMP6f or CBPΔHAT for the time of all experimental readouts (evaluated in prior dilution experiments). Speed of injections was set to 50 nl min^{-1}

2) One week after viral infections, the second surgical procedure involved a 4.3 mm-long, 1 mm-diameter prism microendoscopic GRIN lens (Inscopix, Inc.) being permanently implanted. The mice were anesthetized with isoflurane/oxygen (induction: 3.5% / 2 LPM; maintenance: 1%–1.5% / 0.8 LPM), fixed in a robot stereotactic frame (Neurostar, Kopf Instruments), and stabilized at a body-core temperature of 36°C (probed temperature-controlled heat pad) and a respiratory rate of 1 Hz. The skull was cleaned with hydrogen peroxide and saline three times. The skull was scored and leveled. One screw (00-96 × 1/16 (stainless steel) 1.6 mm cut length) was implanted above the hippocampus bilaterally. A 1 mm × 1 mm square was drilled into the skull above PL. Once drilled, the bone fragment was removed, and the tissue was aspirated away (~1 mm^3). Bleeding was aspirated until it stopped. The prism lens was positioned at a 10^o angle lateral to PL and lowered to (AP + 1.75, ML + 0.55, DV 3.0). The space between the skull and lens was filled with Kwik-Sil Adhesive. The lens was then cemented to the skull with dental cement (Dental Cement Powder, Coral, 1 lb. (A-M Systems 525000) & Dental Cement Solvent, 16 oz (A-M Systems 526000).

3) After two weeks of recovery, the baseplate holding the miniature microscope was cemented onto the mouse skull after the miniscope was optically aligned with the lens. The mice were anesthetized with isoflurane/oxygen (induction: 3.5% / 2 LPM; maintenance: 1%–1.5% / 0.8 LPM), fixed in a robot stereotactic frame (Neurostar, Kopf Instruments), and stabilized at a body-core temperature of 36°C (probed temperature-controlled heat pad) and a respiratory rate of 1 Hz.

4) Subsequently, the mice were habituated to the experimenters and the miniscope over several days prior to the start of Ca^{2+} imaging in freely behaving animals. The animals were provided with analgesia (Buprenorphine and Rimadyl), and their well-being was monitored throughout the experimental period.

Histology

Mice were sacrificed using Nembutal (200 mg/kg, intraperitoneal injection) and transcardially perfused with 20 ml of PBS and then with 20 ml of 4% PFA. Brains were extracted and soaked in 4% PFA overnight at 4°C. Then, they were soaked in 20% sucrose at 4°C until the brains sank. The brains were flash-frozen in embedding media (Tissue-Tek, 4583) using dry ice and ethanol, followed by storage at −80°C. Free-floating 40 µm coronal sections were sliced using a cryostat (Leica, CM1860) and stored in cryoprotectant (50% PBS, 30% ethylene glycol, and 20% glycerol) at −20°C. Free-floating immunohistochemistry (IHC) was performed by washing sections 2 times for 10 min in 1x PBS, followed by a 1-h incubation in blocking buffer (4% normal goat serum in washing buffer) and washing 3 times for 10 min in washing buffer (1x PBS w/ 0.3% Triton X-100). The sections were then incubated overnight at 4°C in an antibody diluent (2% normal goat serum in washing buffer) with primary antibodies (rat anti-mCherry Monoclonal Antibody (IgG2a isotype), Life Technologies Corporation, M11217, 1:2000; mouse Anti-NeuN, clone A60, MILLIPORE, MAB37, 1:2000; rabbit Anti-acetyl-Histone H3 Antibodies, Millipore, 06-599, 1:2000; rabbit Anti-acetyl-Histone H4 Antibodies, Millipore, 06-866, 1:2000). After three washes with washing buffer, the sections were incubated with secondary antibodies (donkey anti-Rat IgG (H+L) - highly cross-adsorbed, biotin conjugate, Life Technologies Corporation, A18749, 1:1000; goat Alexa Fluor® 568 Goat Anti-Mouse IgG, Life Technologies Corporation, A-11031, 1:1000; goat Alexa Fluor® 647 Goat Anti-Rabbit IgG (H+L), highly cross-adsorbed, Life Technologies Corporation, A-21245, 1:1000) for 3 h at room temperature in an antibody diluent. After three washes with washing buffer, the sections were incubated with tertiary antibodies (Biotin Streptavidin, Alexa Fluor® 555 conjugate, Life Technologies Corporation, S-21381, 1:1000) for 3 h at room temperature in an antibody diluent. The sections were then washed twice with 1x PBS for 10 min and mounted on glass slides (Superfrost Plus, 12-550-15) using a mounting medium (ProLong Antifade, P36965) before imaging.

Microscopy imaging

Coronal sections from 2.5 mm to 1.7 mm anterior to the bregma were imaged at 20x (20x/0.95 XLUMPlanFl objective) magnification using a semi-automatic laser-scanning confocal Olympus FV1000 microscope controlled by Olympus FV10-ASW software (v. 2.01). The gain and offset of each channel were balanced manually using Fluoview saturation tools for maximal contrast. All settings were tested on multiple slices before data collection, and brain slices were imaged using identical microscope settings once established. Each channel was acquired in "Sequential Mode, Frame." All images were acquired using the "Integration Type: Line Kalman" and "Integration Count: 2" to increase the signal-to-noise ratio. Localization of the PL and IL regions within the mPFC was performed by overlaying images from the Allen Mouse Brain Atlas. For quantification, 10 optical sections were acquired from a 30 µm Z-stack encompassing 2.7.

Calcium-imaging and data processing

The behavioral video recordings were synchronized with a PC running the data acquisition software (DAQ) for *in vivo* calcium imaging obtained from open source miniscope.org. Microendoscopic signals from HMM were recorded at ~15 Hz and saved to a hard drive as 50-s clips. Video clips recorded during the trial were concatenated into a single raw video using ImageJ. Each raw video was motion-corrected by a piecewise motion correction [105] to stabilize the frames of the video across time and to correct for physiological inner-frame transformations of the imaged tissue. The motion-corrected videos were further aligned by landmark coordinates via two-dimensional shift correction manually using ImageJ. The aligned videos, covering the

first 200 s of stimulus presentation, were stitched with ImageJ and inspected for stability and continuity across the stack of frames. The process of cell signal detection using constrained nonnegative matrix factorization for microendoscopic data (CNMF-E) [106] was semi-supervised (interventions), demanding manual pre-selection of merge-split-delete decisions about initialized components during iterative matrix updates to improve the estimates. Demixed spatial footprints and temporal traces of the identified and extracted components were scanned further for quality control. Therefore, simultaneous visual inspection of spatiotemporal characteristics was performed on each output raw dataset to identify and exclude residual false-positive components. After verification, sets of cells were used for downstream data analysis. Notably, we used only recordings of high quality to down-sample our groups. Figure S2 shows the behavioral performance of mice selected for the analysis of PL network dynamics.

Denoising Ca^{2+} traces and Ca^{2+} spike deconvolution

Both calcium traces and Ca^{2+}spike data, the peaks of calcium traces, were extracted using the open-source (CNMF-E) [106] by sampling at ~15 Hz (Fig. S3 and Fig. S4). A spike raster of all extracted neurons from CNMF-E tracked across the entire CJL (neurons × time) served as a source of data set for further processing. Each mouse then had a varying number of observable neurons across all trials, although not all these neurons were active during each trial. Active neurons were determined if at least one Ca^{2+} spike occurred during a trial. The activity of a neuron was defined as a Ca^{2+} spike count in a given trial.

Tracking cells

To track cells across the behavioral paradigm, we used the Cell Registration (CellReg) code [78]. Fig. S5 shows an output from CellReg capturing spatial footprint similarities across sessions. 33 sessions were loaded into the GUI of all stimulus presentations across all days (11 days and 3 stimulus presentations per day). An approximate micron per pixel value was calculated by dividing the sampled average diameter of neurons expressing GCaMP6f in the mPFC through histology (~15.4 microns) by the average diameter in pixels of registered neurons from CNMF-E. The diameter in pixels was calculated by finding the maximum and minimum arms of the registered neurons and averaging the two. This was done by finding the maximum distance (long axis) between any two points on the cell boundary and finding the perimeter of the boundary of the cell in pixels. Then, using the ellipse perimeter formula, a short axis was found. Lastly, the short and long axes were averaged to determine the approximate diameter of the cell. The middle session (17 out of 33) was used as a reference session. We used 12 microns as the maximal distance in microns. All other parameters were set by the best model, which was determined by the spatial correlation code. We used the cell to index map to find and use the cell IDs for tracking across the 33 sessions.

$$perimeter \sim 2\pi \sqrt{\frac{a^2 + b^2}{2}}$$

$$a = long\ axis$$

$$b = short\ axis$$

Neuronal selectivity and network stability.

The selectivity of each neuron was expressed by a Discrimination Index (DI) that was calculated based on Ca^{2+} spike rates in response to CS- and CS+ using formula DI = (CS- – CS+) / (CS- +

CS+), where CS- and CS+ represents Ca^{2+} spikes frequencies of individual neuron in response to CS- or CS+, respectively, during H, E or L Block Trials. Positive or negative values indicate larger or smaller responses to CS- than to CS+ stimulus presentations. DI values varied between −1 and 1 and were binned into 0.1 intervals before generating cumulative frequency distributions of DIs counts in each of 20 intervals during H, E- and L- phase CJL trials. Kolmogorov-Smirnov (K-S) test was used to find maximum difference between the two cumulative frequency distributions (H *vs.* E and, separately E vs. L) of neuronal response selectivity towards CS- vs. CS+.

Tensor decomposition-based analysis (CPD computational model)

Tensor decomposition analysis was performed per each mouse brain independently. Figure 3A shows a pipeline of tensor decomposition. A spike raster of all extracted neurons from CNMF-E tracked across the entire CJL with CellReg (neurons × time) served as a source of the data set. The CJL behavioral test consists of 33 stimulus presentations (CS+, CS-, NS) called trials. Each trial lasted 200 s; thus, the total time of CJL was 6,600 s. Data was reshaped from a 2D spike raster (neurons × time) into a 3D spike raster: neurons (varied in each mouse) × trial (33 behavioral trials) × time (660-time intervals of the total, each lasting 10 s). Note that instead of dealing with absolute time stamps, each timestamp is in relevance to which trial we are referring to, and even though we can consider all trials as a single time series, we posit that there is inter-trial structure and forming a tensor this way allows us to discover it. Therefore, 33 trials yielded 660 data points representing neuronal engagement throughout CJL. In detail, the trial factor consists of 660 total data points because each data point represents a value of the neuron's activity during a 10-second time bin, resulting in each 200-second trial having 20 total data points. Thus, the output of CNMF-E and CellReg was organized into a contingency judgment tensor (Fig. 3A), that is, a neuron × trial × time (i.e., tensor $\boldsymbol{\mathcal{X}}$ of dimensions IxJxK).

The Canonical Polyadic Decomposition or PARAFAC Decomposition model [107] is approximating a tensor $\boldsymbol{\mathcal{X}}$ of dimensions IxJxK as the sum of R rank-one components [108], each one being the outer product of three vectors, as shown below:

$$\boldsymbol{\mathcal{X}} \approx \sum_{r=1}^{R} \boldsymbol{\lambda}_r \boldsymbol{N}_r \circ \boldsymbol{T}_r \circ \boldsymbol{S}_r$$

where $\boldsymbol{N}$ is the *IxR* "Neuron" factor, $\boldsymbol{T}$ is the *JxR* "Trial" factor, $\boldsymbol{S}$ is the *KxR* "Time within Trial" factor, $\boldsymbol{\lambda}$ is an R-dimensional vector that absorbs the scaling of each rank-one component (assuming that the columns of matrices N, T, S are normalized to unit norm) and ∘ denotes the generalized outer product (for product $\boldsymbol{a} \circ \boldsymbol{b} \circ \boldsymbol{c}$, its (i,j,k) value is equal to $\boldsymbol{a}_i \boldsymbol{b}_j \boldsymbol{c}_k$).

The above CPD model, given that the factors have non-negative entries, essentially performs co-clustering of neurons, trials, and points of time within a trial [107]. This essentially means that for a given component, the neurons, trials, and points of time that have a high value in their corresponding tensor factor vector have a high degree of participation in that co-cluster. Co-cluster membership need not be exclusive, so neurons, trials, and time points can belong to multiple co-clusters with varying degrees of intensity.

In order to achieve non-negative factors, we need to introduce the additional constraint:

$$\boldsymbol{\mathcal{X}} \approx \sum_{r=1}^{R} \boldsymbol{\lambda}_r \boldsymbol{N}_r \circ \boldsymbol{T}_r \circ \boldsymbol{S}_r$$

subject to $\boldsymbol{N} \geq 0, \boldsymbol{T} \geq 0, \boldsymbol{S} \geq 0$, where the inequality is applied element-wise. In order to fit the non-negativity-constrained CPD, we use the Multiplicative Updates implementation [109] provided by the Tensor Toolbox for MATLAB [102] .

Finaly, we turned the component memberships from soft (i.e., multiple components membership) to hard (i.e., single component membership) assignments in a similar fashion as seen in case of community detection [87,88]. For the i-th neuron, we look at the corresponding row of the N (neuron) factor N(i,:) and find the maximum value. The maximizing column corresponds to the component to which this neuron belongs with the highest membership value, so we assign that neuron to the index of the maximum value as follows:

n_i = arg max_R N(i,:), where arg max denotes the maximizing index between 1...R for the vector N(i,:) (which denotes the i-th row of matrix N).

An open problem in tensor analysis that is crucial to the success of the current work is the determination of the number of meaningful components in the decomposition, a matter that has not been optimally addressed by previous studies. Using traditional measures of goodness of fit can be misleading and cannot show the true number of components, risking the extraction of noisy components[110]. To determine the best rank-R model for our data in the current studies, we planned to build upon specialized heuristics for tensor decomposition—starting from the so-called "Core Consistency Diagnostic," or CORCONDIA [111,112] . CORCONDIA is a heuristic model fit score that measures how well a given decomposition in a certain rank R (or more compactly we can say, given R components) adheres to the CPD model. Prior work evaluated CORCONDIA effectiveness in synthetic and real data and demonstrated that it yielded a robust, stable, and reliable estimation of the accurate number of latent components that represent meaningful patterns by eliminating noisy artifacts [113,114]. The core consistency score is always less than or equal to 100%. Empirically, decompositions that score between 80% - 100% demonstrate clear structure, while values roughly between 60% - 80% demonstrate a fair amount of noise which may dilute the purity of the so-called trilinear structure that the tensor decomposition seeks to uncover, and values below 60% may indicate a decomposition that is very noisy thus not very suitable to represent the patterns in the data. Note that those thresholds are rather empirical and a may vary across applications, therefore we use them as a guide in choosing an appropriate set of components which we subsequently evaluate for validity manually. The number of components was then based on the CORCONDIA plot, which is the best-represented model. All models were run with a tolerance of 10^{-6} and 10,000 iterations to ensure convergence. Supplemental Fig. S7 shows the CORCONDIA plot for CJL neuronal tensor and determined the three-latent factor model as an optimal model, which was used to generate data output included in the CPD model (Fig. 3B-C, Fig. S10). While CORCONDIA-based estimation of a number of latent components has been applied before in various applications, including human brain studies (fMRI and EEG) [115-119] , it is rarely considered for neural network studies at a single cell resolution.

Neuronal network dynamics graph (ND-graph)

ND-Graph was constructed independently for each mouse brain. The ND-graph was constructed to create a graph of neural activity which considers temporal dynamics within the actual network (Fig. S8A). The ND-graph is essentially a time-aggregated graph generated by summing multiple dynamic graphs into a single static graph (i.e., time-aggregated). To generate the ND-graph adjacency matrix from calcium spike data, we begin with the initial activity data in the 2-dimensional neurons × time (200 s) format for each trial. Then, to create an adjacency

matrix (i.e., the connections between neurons), we must specify the rule by which these connections will be formed. In our network graph, a connection (i.e., edge) represents the strength of the functional connectivity between two neurons (i.e., if two neurons are coactive, they should be linked). Each trial (200 s total) was segmented into smaller time intervals. We used a 66 milliseconds time binning size and a sliding time window (SW) set to 666 milliseconds to capture the temporal features of the PL network dynamics at a cellular resolution underlying CJL (Fig. S8A). With the specified time window (0.666 s), a SW is used so that each next time window slightly overlaps with the preceding window. The SW approach has the benefit of offering continuity in tracking a consistent set of coactive neurons. Assigning edge weights is an important addition to this graph that helps account for temporal properties. Neurons that fire within the same frame (i.e., bin) are given the largest weight (weight = 1). Edge weights decrease from neurons in the first frame that project to subsequent frames within the SW in steps of *f*. Thus, lower weights are assigned as two neurons' spikes become further separated in time. The edges within the same frame are considered bidirectional, and the edges from one frame to the next become directional in accordance with the forward progression of time. The edge weights are calculated as follows:

$$weight = \frac{SW - f}{SW}$$
$$f = 0, 1, 2, \dots, SW - 1$$

These weights construct the adjacency matrix for each SW. Each graph is generated using the 'digraph' function (MATLAB) and simplified (using 'simplify' function, MATLAB). The 'simplify' function was used to aggregate multiple edges between two nodes into a single edge and constrain nodes with more than one self-loop to a single self-loop in the simplified graph. This simplified graph (with summed weights) was then converted into an adjacency matrix (using the 'adjacency' function, MATLAB) and reshaped to be a 3-dimensional adjacency matrix (neurons x neurons x bin) for each SW. Then, each adjacency matrix representing each SW is summed across time to generate a cumulative adjacency matrix for a single trial (200 s).

$$\sum_{t=1}^{\#frames} adj_t = Adj$$

$$adj_1 + adj_2 + \dots + adj_t = Adj$$

The above steps are repeated for each trial (33 trials total) to generate adjacency matrices for each trial. Lastly, all adjacency matrices are summed into one overall adjacency matrix that represents the network across the entire behavioral paradigm. In summary, the ND-graph captures temporal patterns of neuronal activations with the PL network by measuring temporal proximity of neuronal co-activations within 660 ms timescale. Stronger weights correlate with neurons firing in closer temporal proximity while higher cumulative weights reflect higher probability of neurons firing together.

Community analysis (ND-graph computational mode)

Community analysis was performed per each mouse brain independently. In analogy to the tensor decomposition model, the ND-graph reflects well temporal structures across neurons and time (e.g., trial). To gain insights into the trial-based specificity of temporal neuronal coactivation patterns, we employed a community detection algorithm on data sets that include all 33 CJL trials (Fig. 4). After the cumulative ND-graph representing all 33 trials in one adjacency matrix is generated, it can be analyzed for the underlying community structure using the Louvain community detection algorithm (Brain Connectivity Toolbox [90]). This algorithm finds the optimized community structure of nonoverlapping groups of nodes that effectively maximizes the sum of weighted edges within the community while minimizing the sum of weighted edges between communities [90,120-122]. The inputs to the algorithm include the adjacency matrix and the value of a hyperparameter 'gamma'. The gamma value adjusts the resolution of community detection such that a gamma value exceeding 1 will detect smaller subsets of nodes, while gamma values between 0 and 1 will detect larger subsets of nodes within the network. The Louvain community detection algorithm then generates two outputs: community assignments (i.e., which community each node belongs to) and the modularity value of the graph. Modularity (Q) is an optimized community structure statistic representing how well a network is divisible into modules or communities.

$$Q = \frac{1}{2m} \times \sum_{i,j} \left[A_{i,j} - \frac{k_i k_j}{(2m)} \right] \times \delta(c_i, c_j)$$

$A_{i,j}$ = represents the edge weight between nodes i and j
k_i, k_j are the sum of weights of the edges attached to nodes i and j, respectively
m is the sum of all edge weights in the graph, c_i and c_j are the communities of the nodes
δ is Kronecker delta function (δ(x,y) = 1 if x =y, 0 otherwise)

The Louvain community analysis identifies nonoverlapping communities. As a routine, we (and others) start community detection at hyperparameter gamma value equal to 1 and run multiple time community detection algorithm oscillating hyperparameter gamma value in the range 0.9-1.0 with a step of 0.01. We detected three PL network communities using this approach during a CJL performance (Fig. S8B). The modularity coefficient, Q, is a measure of how well each of the communities is put together on a scale from 0 to 1, maximizing connections within the community and minimizing the connections between communities (Fig. S8B).

The probability of the intersection of a component and a community.
The overlap (%) between a component (TC) and the corresponding community (NC) was calculated as follows:

$$\%\ Overlap\ (TC\ and\ NC) = 100 \times \left[\frac{\text{TC} \cap \text{NC}}{(\text{total \# of TC})} \right]$$

Software

Data processing and analysis were performed using MATLAB (Mathworks, NA), Tensor Toolbox for MATLAB [95], Brain Connectivity Toolbox [88], GraphPad Prism 9, and ImageJ.

QUANTIFICATION AND STATISTICAL ANALYSIS

The experimenters were blind to group designations. The data represent the mean ± SEM, and N represents the number of mice per group unless stated otherwise. Percent of network activity was defined as the ratio between the number of Ca^{2+} spikes originating from a defined PL neuronal population during a specific trial (stimulus presentation) and the total number of Ca^{2+} spikes elicited from all recorded PL neurons in each animal during that same trial. Percent of network activity (based on Ca^{2+} imaging) or percent of freezing (fear behavior) was always accessed from individual mice first before calculating the mean of a specific animal group. Statistical analysis was performed using GraphPad Prism and Excel (Microsoft, Inc.). Student's *t*-test or ANOVA was used for statistical comparisons. Pearson's correlation (*r*) was used as the effect size. For one-way ANOVA, eta-squared ($\eta 2$) was used as the effect size. Welch's tests were used when the same standard deviation was not assumed. For learning assessment at the circuit or behavioral levels during behavioral testing, repeated measures (RM), two-way ANOVA or 3-way ANOVA, and post hoc analysis with appropriate multiple comparisons test were used. With a large dataset, dealing with sporadic missing data due to technological problems is not unusual. In the presence of missing values (missing completely at random), the results can be interpreted as RM ANOVA. We analyzed the data instead by fitting a mixed model (REML) as implemented in GraphPad Prism 8.0. REML uses a compound symmetry covariance matrix and is fitted using restricted maximum likelihood. In the absence of missing values, this method gives the same *p* values and multiple comparison tests as RM ANOVA. The Greenhouse-Geisser correction (GGC) in RM designs was used for correcting against violations of sphericity. The two-sample nonparametric Kolmogorov-Smirnov (K-S) test was used to compare the cumulative frequency distributions of two data sets representing discrimination indices. Significance values were set at $p < 0.05$. Asterisks indicate statistical significance: *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$, ****, $p < 0.0001$ and ns or absence of asterisk(s) indicates not significant. The statistical tests used and the number of replicates have been included in the result section and proper figure legends.

Reporting summary. Further information on research design is available in the Star*Methods Key Resource Table linked to this article.

Figures

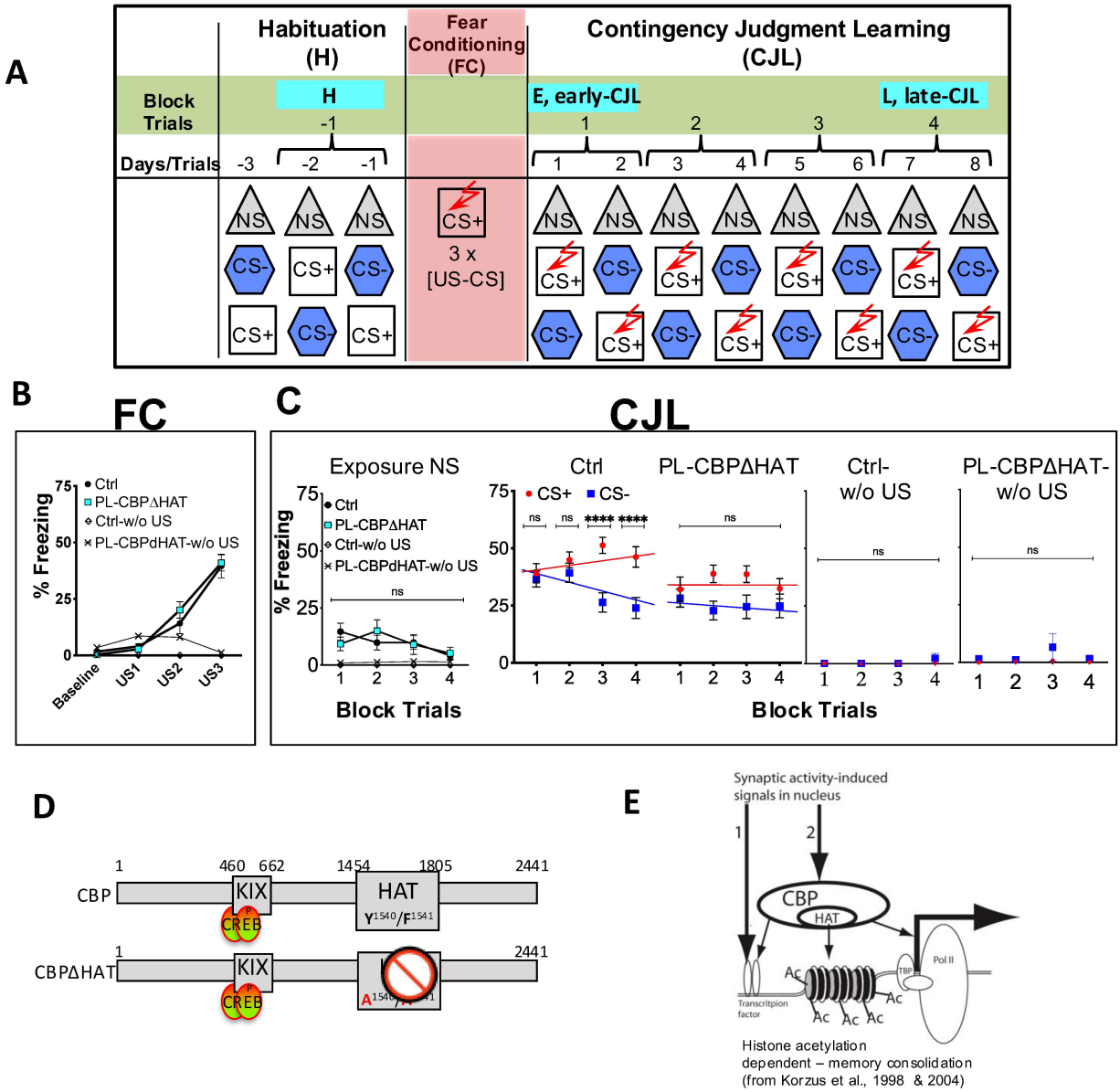


Fig. 1. PL network manages learning on CJL task.

(A) Experimental design for contingency judgment learning (CJL). The entire behavioral task comprises three phases: habituation, fear conditioning, and CJL, further divided into early (E)-CJL (Block Trial 1) and late (L)-CJL (Block Trial 4) (for details see Methods)

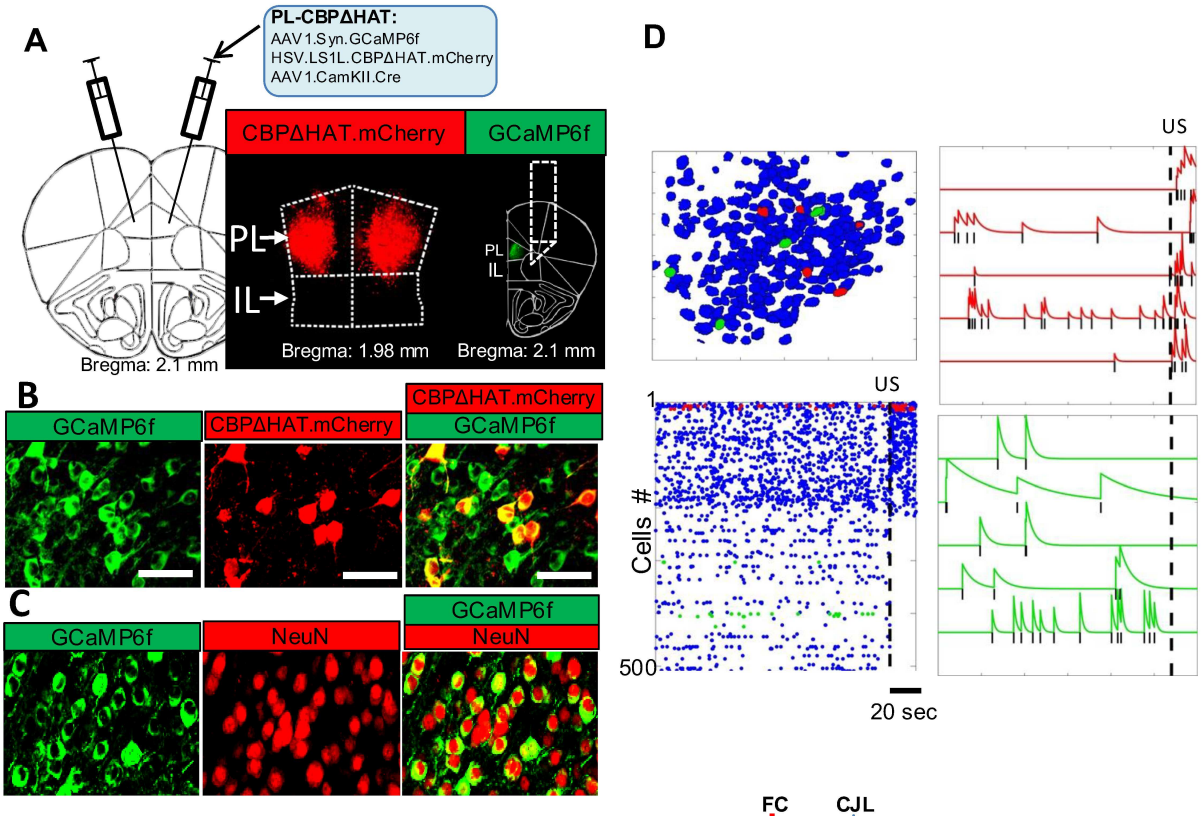
(B) Fear conditioning. The Ctrl and PL-CBPΔHAT groups showed similar baseline and robust performance during fear conditioning, while the mice that were never exposed to US (Ctrl-w/o US and PL-CBPΔHAT-w/o US) failed to acquire fear.

(C) Left, all four groups (Ctrl, PL-CBPΔHAT, Ctrl-w/o US, PL-CBPΔHAT-w/o US) showed similar responses to NS throughout the CJL. Right, the Ctrl mice exhibited robust leaning on the CJL task and learned accurate contingency judgment, while PL-CBPΔHAT failed to learn the task. The change in freezing response to stimuli across training was calculated as the average freezing per a block trial. However, the PL-CBPΔHAT mice failed to learn appropriate contingency judgment and showed strong generalization during all four block trials (PL-CBPΔHAT, CS+ *vs.* CS-, Block 1–4, $p > 0.05$).

(D) Mechanism of histone acetylation-dependent memory consolidation (for details see [64]). This model suggests that CBP-mediated histone acetylation during learning-triggered transcriptional activation is a critical step in the molecular mechanism controlling memory stabilization. Initial steps include induction of CREB phosphorylation, CBP activation, and CBP-mediated histone acetylation at a specific transcriptional unit in response to the initial synaptic events associated with learning. Subsequently, prolonged elevated transcription required for memory consolidation could be maintained by CBP- and CREB phosphorylation-independent nuclear mechanisms even after signals to CREB and CBP are no longer present. This transient transcriptional activation would remain active until the competing phosphatase- and deacetylase-dependent repression mechanism shut off transcription.

(E) PL-CBPΔHAT mutant mice express CBPΔHAT with eliminated acetyltransferase activity (for details see [65]). Based on mutagenesis studies demonstrating that single amino acid substitutions in the acetyl coenzyme A (acetyl-CoA) binding domain of acetyltransferases result in loss of their enzymatic activity, these studies have employed a CBP dominant negative mutant (CBPΔHAT) harboring a substitution mutation of two conserved residues (Y1540/F1541 to A1540/ A1541). This mutant has no intrinsic HAT activity.

Ctrl, n = 18, PL-CBPΔHAT, n = 14, Ctrl-w/o US, n = 5, PL-CBPΔHAT-w/o US, n = 5. The data represent the mean ± SEM, and n represents the number of mice per group. Significance values were set at $p < 0.05$: *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$; ****, $p < 0.0001$ and ns indicates not significant.


A
PL-CBPΔHAT:
AAV1.Syn.GCaMP6f
HSV.LS1L.CBPΔHAT.mCherry
AAV1.CamKII.Cre
CBPΔHAT.mCherry
GCaMP6f
PL
IL
Bregma: 2.1 mm
Bregma: 1.98 mm
Bregma: 2.1 mm
B
GCaMP6f
CBPΔHAT.mCherry
CBPΔHAT.mCherry
GCaMP6f
C
GCaMP6f
NeuN
GCaMP6f
NeuN
D
US
US
Cells #
1
500
20 sec

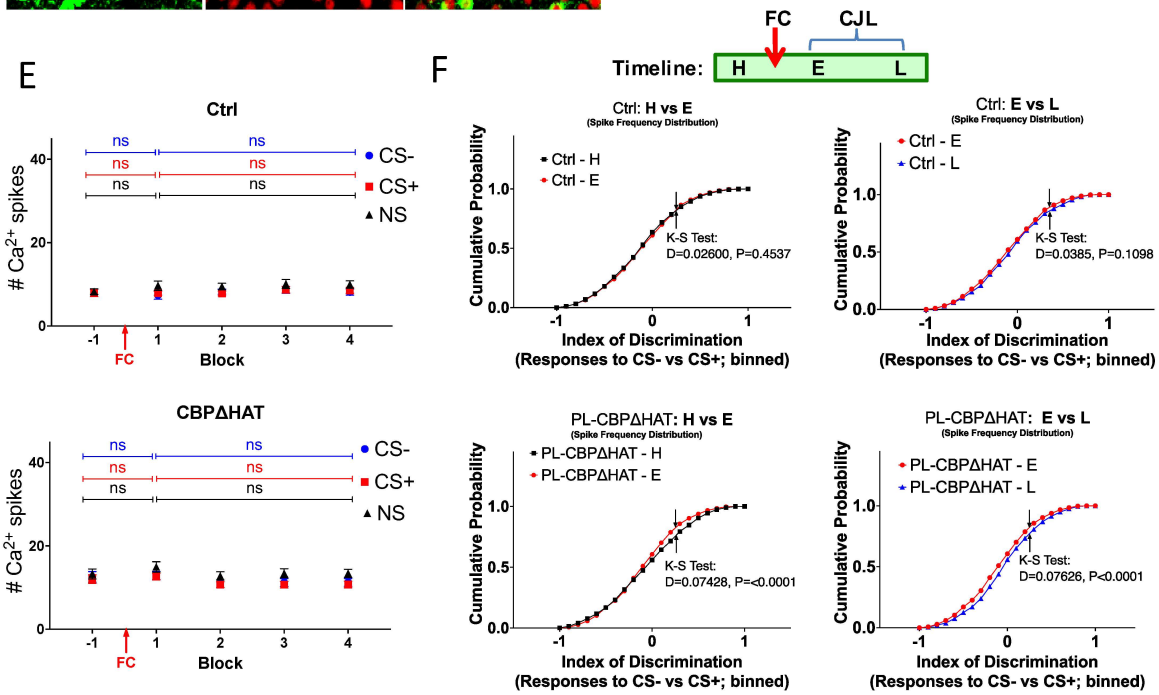

E
Ctrl
Ca2+ spikes
ns
CS-
CS+
NS
FC
Block
CBPΔHAT
F
Timeline:
FC
CJL
H
E
L
Ctrl: H vs E
(Spike Frequency Distribution)
Ctrl - H
Ctrl - E
K-S Test:
D=0.02600, P=0.4537
Cumulative Probability
Index of Discrimination
(Responses to CS- vs CS+; binned)
Ctrl: E vs L
Ctrl - E
Ctrl - L
K-S Test:
D=0.0385, P=0.1098
PL-CBPΔHAT: H vs E
PL-CBPΔHAT - H
PL-CBPΔHAT - E
K-S Test:
D=0.07428, P=<0.0001
PL-CBPΔHAT: E vs L
PL-CBPΔHAT - E
PL-CBPΔHAT - L
K-S Test:
D=0.07626, P<0.0001

Fig. 2. Head-mounted microscopes (HMMs) effectively capture PL network dynamics at a cellular resolution.

(A) Coronal representation of virus injection sites (left) and GCaMP6f expression in PL with a laterally positioned microendoscopic lens. The middle, representative coronal brain section shows targeting of PL (red, mCherry) with PL-CBPΔHAT, with HSV.LS1L.CBPΔHAT.mCherry vector expressing CBPΔHAT and mCherry (red) + AAV1.CamKII.Cre + AAV1.SynGCaMP6f (or control viruses AAV1.Syn.GCaMP6f + HSV.LS1L.mCherry + AAV1.CamKII.Cre). Right, a representative coronal brain section shows the PL cell population expressing GCaMP6f (green) in PL with a laterally positioned microendoscopic prism lens.

(B) Representative image indicating PL infection of the viruses for GCaMP6f (left), CBPΔHAT.mCherry + Cre (middle), and overlap of GCaMP6f and mCherry expression (right). Labeling indicated that most GCaMP+ infected cells were also mCherry+ infected cells. White scale bars indicate 50 μm.

(C) Representative image indicating PL infection of the viruses for GCaMP6f (left), immunohistochemistry staining of neuronal marker NeuN (middle), and overlap of GCaMP6f (green) and NeuN (red) staining (right). White scale bars indicate 50 μm. Immunohistochemistry of NeuN staining showed that most GCaMP6f+ infected cells are neurons. To determine the pattern of GCaMP6f-tagged and mCherry-tagged virus expression, the imaged tissue was compared to the Paxinos and Franklin mouse atlas (Paxinos & Franklin, 2019), and areas of maximal GCaMP6f and mCherry expression were labeled as injection sites.

(D) A representative analysis of PL neurons from a single mouse demonstrates neuronal responses to external stimuli (US). Left-top: spatial footprints of the cell population (blue) responding to CS+ during early-CJL. Left bottom: spike raster of the PL population corresponding to the general population of cells shown in blue above. The corresponding Ca^{2+} traces of red (right-top) and green (right-bottom) footprints (left-top) are shown in red and green, respectively. Right-top: calcium transient time traces of cells from the same dataset as on the left responding to US. Right-bottom: cells responding to US are marked in red, while cells not responding to US are marked in green. Black lines below Ca^{2+} traces correspond to inferred Ca^{2+} spikes. US, red dashed line. Ctrl, n = 8, PL-CBPΔHAT, n = 7.

(E) There were no differences in stimulus-triggered average Ca^{2+} spike frequency between trials across CJL in Ctrl. Similarly, PL-CBPΔHAT group also did not show any differences between specific trials across CJL in stimulus-triggered average Ca^{2+} spike frequency

(F) The analysis of global PL network activity showed network stability and lack of bias towards tested stimuli (CS+ and CS-) across CJL (i.e., trials H, E or L) in Ctrl mice (top). PL-CBPΔHAT mice (bottom) showed modest distortions in the global network stability across CJL. The index of discrimination between CS+ and CS- responses for specific neurons was calculated as follows: Index of Discrimination = (CS- – CS+) / (CS- + CS+), where CS- and CS+ represent Ca^{2+} spikes frequencies of individual neurons in response to CS- or CS+, respectively, during H, E or L Block Trials. Kolmogorov-Smirnov (K-S) test was used to compare cumulative frequency distributions of neuronal response bias towards CS- vs. CS+ across CJL learning. Number of binned values: Ctrl-H, n = 2373, Ctrl-E, n = 2011; Ctrl-L, n = 1903; PL-CBPΔHAT-H, n = 1910; PL-CBPΔHAT-E, n = 2144; PL-CBPΔHAT-L, n = 1871. The data represent the mean ± SEM, and n represents the number of mice per group. Significant values were set at $p < 0.05$: *, $p <$ 0.05, ns indicates insignificant.

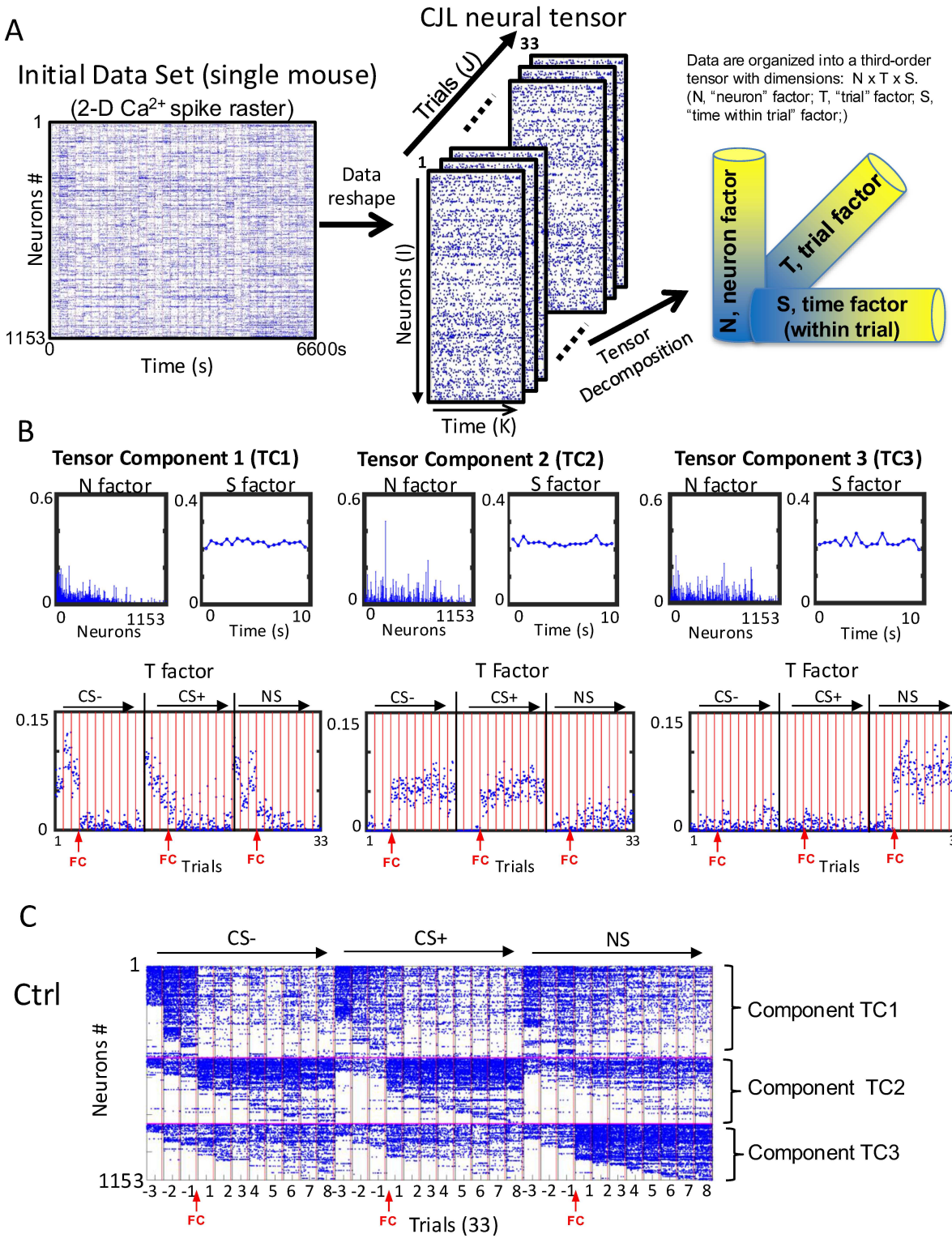

A
Initial Data Set (single mouse)
(2-D Ca2+ spike raster)
Neurons #
1
1153
0
6600s
Time (s)
Data reshape
CJL neural tensor
Trials (J)
33
1
Neurons (I)
Time (K)
Tensor Decomposition
Data are organized into a third-order tensor with dimensions: N x T x S. (N, "neuron" factor; T, "trial" factor, S, "time within trial" factor;)
N, neuron factor
T, trial factor
S, time factor (within trial)
B
Tensor Component 1 (TC1)
Tensor Component 2 (TC2)
Tensor Component 3 (TC3)
N factor
S factor
Neurons
Time (s)
T factor
T Factor
CS-
CS+
NS
FC
Trials
C
Ctrl
Neurons #
Component TC1
Component TC2
Component TC3
Trials (33)

Fig. 3. The CPD model detected latent neuronal populations and their activity patterns that reflect interpretable trial variables (trial conditions and fear levels) of the CJL task in a single Ctrl mouse.

The three-component tensor decomposition model of neuronal activity across all trials of the CJL task shows latent components including unveiling the PL cortex disambiguation circuit.

(A) Pipeline of the tensor decomposition analytics. Spike raster (unsorted) of all extracted neurons from CNMF-E tracked across all trials with CellReg (neurons × time) were reshaped into a rank 3 tensor (i.e., CJL neural tensor) of shape (I, neurons × J, trial × K, time within trial), i.e.,. Finally, CPD (Canonical Polyadic Decomposition) model was used to create 3 factors: neuron factor [N], temporal factor [S] (within-trial), and trial factor [T]. The Multiplicative Updates implementation (provided in the Tensor Toolbox for MATLAB [102]) was used to fit the non-negativity-constrained CPD. For details, see Methods.

(B) CJL Tensor decomposition showed three latent populations in the PL network. Ca2+ spike recording from 1153 PL neurons of a single Ctrl mouse collected across all 33 trials of CJL was used as input data set to pipeline of the tensor decomposition analytics shown in Fig. 3A and, subsequently, to visualize N × S × T factors. Red vertical lines (shown on T factor plots) separate individual trials lasting 200 s each, while blue dots indicate 10 s timestamps within each trial (see methods). Entire CJL paradigm (Fig. 1A) involved 11 presentations of each stimulus (CS-, CS+, NS). (C) Ca2+ spike raster of detected tensor components (TC1, TC2, TC3) found in a single Ctrl mouse as a result of analytics shown in Fig. 3B (see text Methods).

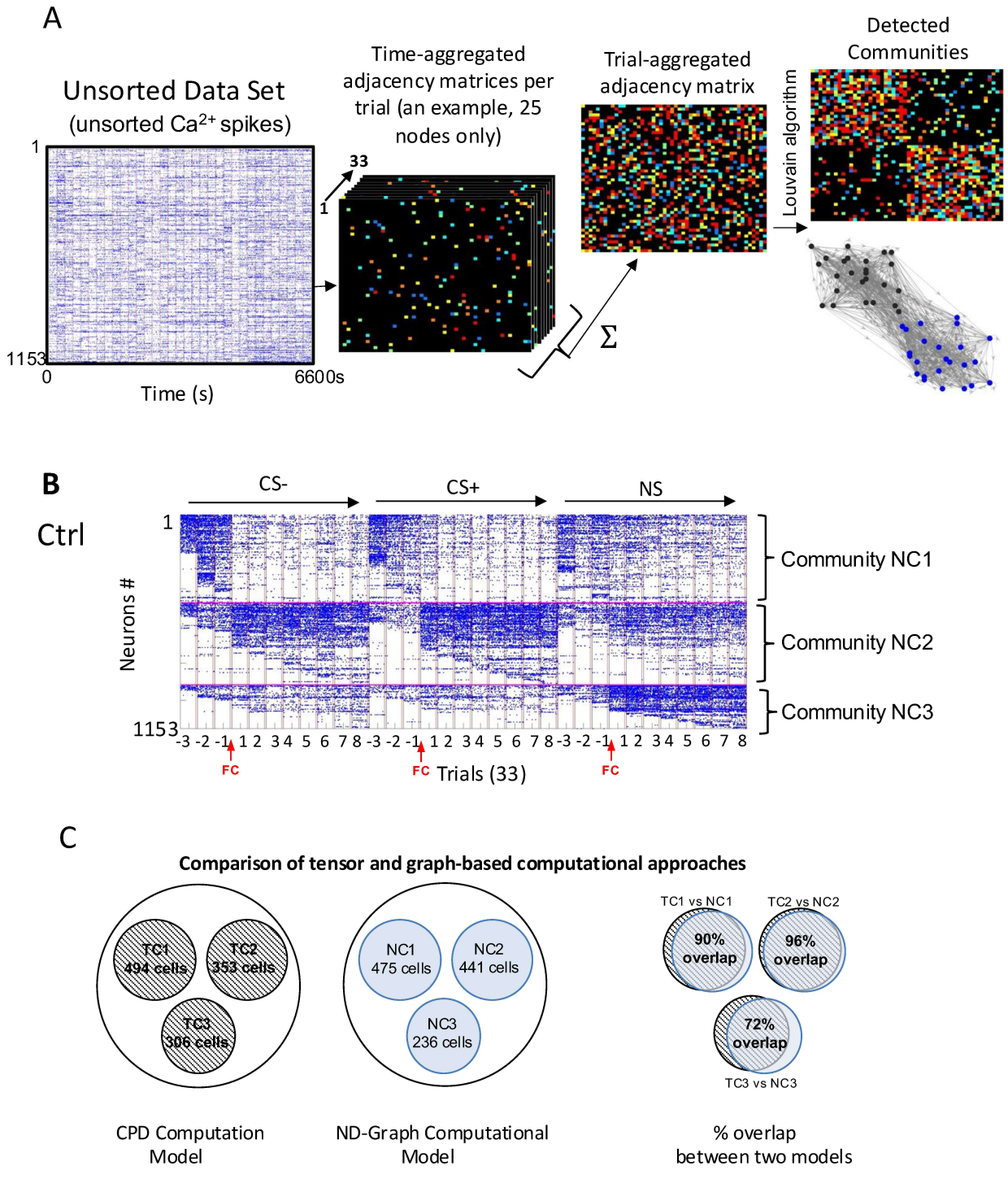
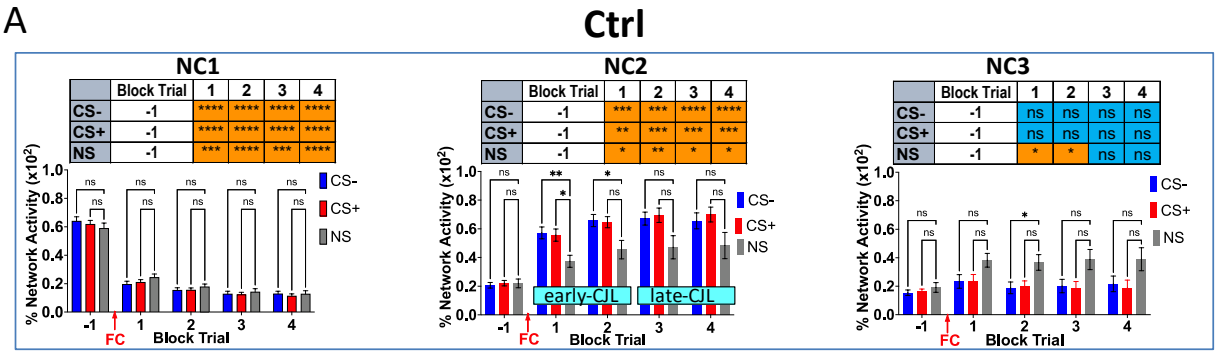


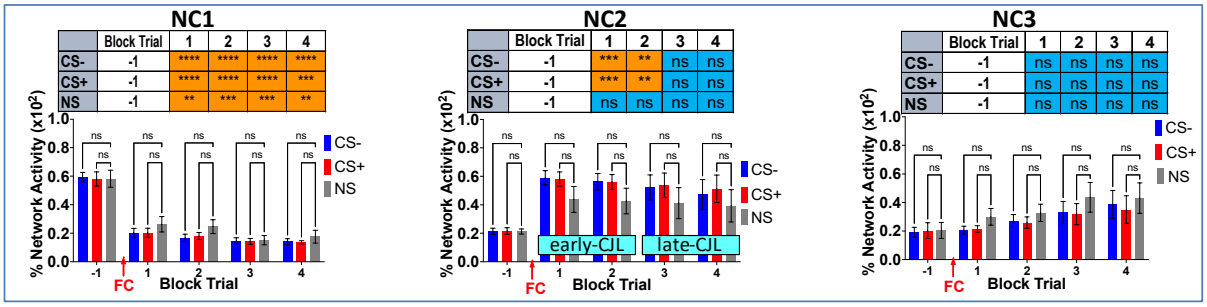
Fig. 4. The ND-graph model detected the disambiguation circuit in the PL cortex (see text).

A) A pipeline of recovered neuronal communities from trial-aggregated ND-Graph. Extracted calcium spikes per trial are converted into time-aggregated adjacency matrices per trial via ND-Graph. All time-aggregated adjacency matrices are further aggregated into a single trial-aggregated adjacency matrix. Louvain community analysis was performed on a trial-aggregated ND-Graph. The presented adjacency matrices are examples with 25 nodes.
(B) A three-community ND-graph model uncovered latent populations (NC1, NC2, NC3) and their activity patterns that reflect interpretable trial variables (trial conditions and fear levels) of the CJL task in a Ctrl mouse. A characteristic pattern of NC1, NC2, and NC3 activity across CJL trials is remarkably similar to TC1, TC2, and TC3 activation patterns recovered in the three-components tensor decomposition model (Fig. 3C).
(C) Comparison of CPD model and ND-graph model based on data obtained from a single brain (shown in Fig. 3C and Fig. 4B). The overlap between components and corresponding community was calculated as follows:
% Overlap (intersect) between a tensor component (TC) and a corresponding community (NC) = 100 * [(TC ∩ NC) / total number of TC].

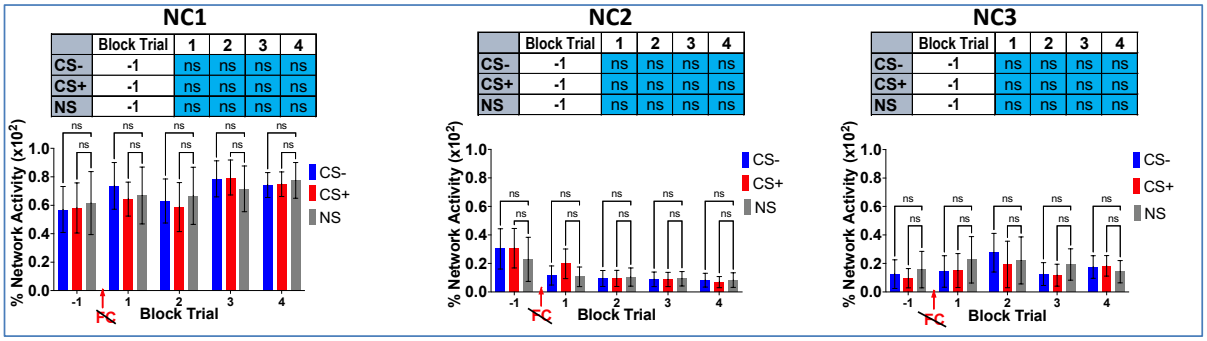
A


Ctrl
NC1
NC2
NC3
Block Trial
CS-
CS+
NS
% Network Activity (x10²)
FC
Block Trial
early-CJL
late-CJL

B


PL-CBPΔHAT
NC1
NC2
NC3
Block Trial
CS-
CS+
NS
% Network Activity (x10²)
FC
Block Trial
early-CJL
late-CJL

C


Ctrl-w/oUS
NC1
NC2
NC3
Block Trial
CS-
CS+
NS
% Network Activity (x10²)
FC
Block Trial

D

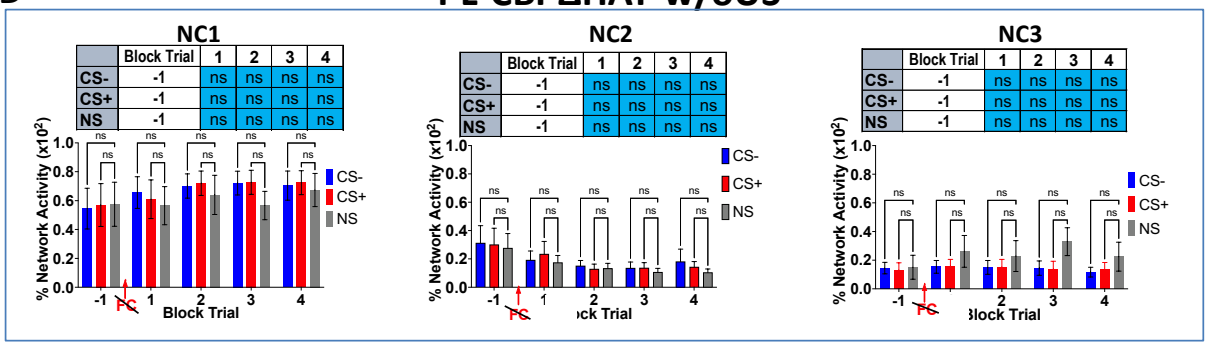

PL-CBPΔHAT-w/oUS
NC1
NC2
NC3
Block Trial
CS-
CS+
NS
% Network Activity (x10²)
FC
Block Trial

Fig. 5. Activity patterns of recovered latent neural communities.

Activity patterns of recovered latent neural communities (NC1, NC2 and NC3) across trials of CJL calculated using the NG-graph model in experimental groups: Ctrl (A), PL-CBPΔHAT (B), Ctrl-w/o US (C), and PL-CBPΔHAT-w/o US (D). The pattern of activity of the disambiguation circuit (NC2) recovered in Ctrl explains the learning of contingency judgment accuracy. Control mice showed a fully functional disambiguation circuit (NC2) managing stimulus responses and guiding successful CJL. In contrast, PL-CBPΔHAT showed a deficiency in CJL, coinciding with impairment of the performance of the disambiguation circuit. In Ctrl, NC2 manages network responses to CS+, CS-, and NS after FC but not before. During E-CJL, NC2 discriminates between NS stimuli and CS+/CS-, but during L-CJL, it does not; thus, CJL learning correlates with the acquisition of new responsiveness (i.e., to NS) and properties of NC2. NC1 community showed strong responses during habituation without bias towards specific stimuli in Ctrl, n = 8, PL-CBPΔHAT. However, the NC1 community remains highly active in animals not treated with US. NC3 community-acquired bias toward NS during Block Trial 2 in Ctrl mice

The NG-graph model was constructed per each mouse independently. The percentage of network activity in response to each stimulus per trial was calculated per mouse before the group statistic was performed. Percent of network activity was defined as the ratio between the number of Ca2+ spikes originating from a defined PL neuronal population during a specific trial (stimulus presentation) and the total number of Ca2+ spikes elicited from all recorded PL neurons in individual animals during that same trial. The data represent the mean ± SEM, and n represents the number of mice per group. Ctrl, n = 8, PL-CBPΔHAT, n = 7, Ctrl-w/o US, n = 5, and PL-CBPΔHAT-w/o US, n = 4. Significant values were set at $p < 0.05$: *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$; ****, $p < 0.0001$ and ns indicates not significant.

References

1. Baker, A.G., Mercier, P., Vallée-Tourangeau, F., Frank, R., and Pan, M. (1993). Selective associations and causality judgments: Presence of a strong causal factor may reduce judgments of a weaker one. Journal of Experimental Psychology: Learning, Memory, and Cognition, *19*, 414-432.
2. Shanks, D.R., and Dickinson, A. (1987). Associative accounts of causality judgment. In The psychology of learning and motivation: Advances in research and theory,, G.H. Bower, ed. (Academic Press), pp. 229-261.
3. Wassermen, L.A. (1990). Belief functions and statistical inference. Canadian Journal of Statistics *18*, 183-196. https://doi.org/10.2307/3315449.
4. Miller, R.R., and Matzel, L.D. (1988). The comparator hypothesis: A response rule for the expression of associations. In The psychology of learning and motivation: Advances in research and theory, , G.H. Bower, ed. (Academic Press), pp. 51-92. https://doi.org/10.1016/S0079-7421(08)60038-9.
5. Rescorla, R.A. (1968). Probability of shock in the presence and absence of CS in fear conditioning. Journal of comparative and physiological psychology *66*, 1-5. 10.1037/h0025984.
6. Fanselow, M.S., and Poulos, A.M. (2005). The neuroscience of mammalian associative learning. Annu Rev Psychol *56*, 207-234. 10.1146/annurev.psych.56.091103.070213.
7. LeDoux, J.E. (2000). Emotion circuits in the brain. Annu Rev Neurosci *23*, 155-184. 10.1146/annurev.neuro.23.1.155.
8. Ponomareva, O.Y., Fenster, R.J., and Ressler, K.J. (2023). Enhancing Fear Extinction: Pharmacological Approaches. Curr Top Behav Neurosci. 10.1007/7854_2023_443.
9. Jovanovic, T., Norrholm, S.D., Blanding, N.Q., Davis, M., Duncan, E., Bradley, B., and Ressler, K.J. (2010). Impaired fear inhibition is a biomarker of PTSD but not depression. Depression and anxiety *27*, 244-251. 10.1002/da.20663.
10. Jovanovic, T., Kazama, A., Bachevalier, J., and Davis, M. (2012). Impaired safety signal learning may be a biomarker of PTSD. Neuropharmacology *62*, 695-704. 10.1016/j.neuropharm.2011.02.023.
11. Liberzon, I., and Abelson, J.L. (2016). Context Processing and the Neurobiology of Post-Traumatic Stress Disorder. Neuron *92*, 14-30. 10.1016/j.neuron.2016.09.039.
12. van Rooij, S.J.H., and Jovanovic, T. (2019). Impaired inhibition as an intermediate phenotype for PTSD risk and treatment response. Prog Neuropsychopharmacol Biol Psychiatry *89*, 435-445. 10.1016/j.pnpbp.2018.10.014.
13. Ressler, K.J., Berretta, S., Bolshakov, V.Y., Rosso, I.M., Meloni, E.G., Rauch, S.L., and Carlezon, W.A., Jr. (2022). Post-traumatic stress disorder: clinical and translational neuroscience from cells to circuits. Nature reviews. Neurology *18*, 273-288. 10.1038/s41582-022-00635-8.
14. van Rooij, S.J., Stevens, J.S., Ely, T.D., Fani, N., Smith, A.K., Kerley, K.A., Lori, A., Ressler, K.J., and Jovanovic, T. (2016). Childhood Trauma and COMT Genotype Interact to Increase Hippocampal Activation in Resilient Individuals. Frontiers in psychiatry / Frontiers Research Foundation *7*, 156. 10.3389/fpsyt.2016.00156.
15. van Rooij, S.J.H., Ravi, M., Ely, T.D., Michopoulos, V., Winters, S.J., Shin, J., Marin, M.F., Milad, M.R., Rothbaum, B.O., Ressler, K.J., et al. (2021). Hippocampal activation during contextual fear inhibition related to resilience in the early aftermath of trauma. Behav Brain Res *408*, 113282. 10.1016/j.bbr.2021.113282.
16. Gill, J.L., Schneiders, J.A., Stangl, M., Aghajan, Z.M., Vallejo, M., Hiller, S., Topalovic, U., Inman, C.S., Villaroman, D., Bari, A., et al. (2023). A pilot study of closed-loop neuromodulation for treatment-resistant post-traumatic stress disorder. Nature communications *14*, 2997. 10.1038/s41467-023-38712-1.
17. Pavlov, I.P. (1927). Conditioned Reflexes: An Investigation of the Physiological Activity of the Cerebral Cortex. Translated and Edited by G. V. Anrep (Oxford University Press).
18. Maren, S., and Quirk, G.J. (2004). Neuronal signalling of fear memory. Nat Rev Neurosci *5*, 844-852. 10.1038/nrn1535.
19. Rozeske, R.R., Valerio, S., Chaudun, F., and Herry, C. (2015). Prefrontal neuronal circuits of contextual fear conditioning. Genes Brain Behav *14*, 22-36. 10.1111/gbb.12181.
20. Plas, S.L., Tuna, T., Bayer, H., Juliano, V.A.L., Sweck, S.O., Arellano Perez, A.D., Hassell, J.E., and Maren, S. (2024). Neural circuits for the adaptive regulation of fear and extinction memory. Front Behav Neurosci *18*, 1352797. 10.3389/fnbeh.2024.1352797.
21. Maren, S., Phan, K.L., and Liberzon, I. (2013). The contextual brain: implications for fear conditioning, extinction and psychopathology. Nat Rev Neurosci *14*, 417-428. 10.1038/nrn3492.

22. Gonzalez, S.T., and Fanselow, M.S. (2020). The role of the ventromedial prefrontal cortex and context in regulating fear learning and extinction. Psychol Neurosci *13*, 459-472. 10.1037/pne0000207.
23. Milad, M.R., and Quirk, G.J. (2002). Neurons in medial prefrontal cortex signal memory for fear extinction. Nature *420*, 70-74. 10.1038/nature01138 nature01138 [pii].
24. Knapska, E., and Maren, S. (2009). Reciprocal patterns of c-Fos expression in the medial prefrontal cortex and amygdala after extinction and renewal of conditioned fear. Learn Mem *16*, 486-493. 10.1101/lm.1463909.
25. Zelikowsky, M., Bissiere, S., Hast, T.A., Bennett, R.Z., Abdipranoto, A., Vissel, B., and Fanselow, M.S. (2013). Prefrontal microcircuit underlies contextual learning after hippocampal loss. Proc Natl Acad Sci U S A *110*, 9938-9943. 10.1073/pnas.1301691110.
26. Lovelace, J.W., Vieira, P.A., Corches, A., Mackie, K., and Korzus, E. (2014). Impaired fear memory specificity associated with deficient endocannabinoid-dependent long-term plasticity. Neuropsychopharmacology *39*, 1685-1693. 10.1038/npp.2014.15.
27. Vieira, P.A., Corches, A., Lovelace, J.W., Westbrook, K.B., Mendoza, M., and Korzus, E. (2015). Prefrontal NMDA receptors expressed in excitatory neurons control fear discrimination and fear extinction. Neurobiol Learn Mem *119*, 52-62. 10.1016/j.nlm.2014.12.012.
28. Vieira, P.A., Lovelace, J.W., Corches, A., Rashid, A.J., Josselyn, S.A., and Korzus, E. (2014). Prefrontal consolidation supports the attainment of fear memory accuracy. Learn Mem *21*, 394-405. 10.1101/lm.036087.114.
29. Marek, R., Sun, Y., and Sah, P. (2018). Neural circuits for a top-down control of fear and extinction. Psychopharmacology (Berl). 10.1007/s00213-018-5033-2.
30. Burgos-Robles, A., Kimchi, E.Y., Izadmehr, E.M., Porzenheim, M.J., Ramos-Guasp, W.A., Nieh, E.H., Felix-Ortiz, A.C., Namburi, P., Leppla, C.A., Presbrey, K.N., et al. (2017). Amygdala inputs to prefrontal cortex guide behavior amid conflicting cues of reward and punishment. Nat Neurosci *20*, 824-835. 10.1038/nn.4553.
31. Herry, C., Ciocchi, S., Senn, V., Demmou, L., Muller, C., and Luthi, A. (2008). Switching on and off fear by distinct neuronal circuits. Nature *454*, 600-606. 10.1038/nature07166.
32. Senn, V., Wolff, S.B., Herry, C., Grenier, F., Ehrlich, I., Grundemann, J., Fadok, J.P., Muller, C., Letzkus, J.J., and Luthi, A. (2014). Long-range connectivity defines behavioral specificity of amygdala neurons. Neuron *81*, 428-437. 10.1016/j.neuron.2013.11.006.
33. Kitamura, T., Ogawa, S.K., Roy, D.S., Okuyama, T., Morrissey, M.D., Smith, L.M., Redondo, R.L., and Tonegawa, S. (2017). Engrams and circuits crucial for systems consolidation of a memory. Science *356*, 73-78. 10.1126/science.aam6808.
34. Sotres-Bayon, F., Cain, C.K., and LeDoux, J.E. (2006). Brain mechanisms of fear extinction: historical perspectives on the contribution of prefrontal cortex. Biol Psychiatry *60*, 329-336. 10.1016/j.biopsych.2005.10.012.
35. Quirk, G.J., and Mueller, D. (2008). Neural mechanisms of extinction learning and retrieval. Neuropsychopharmacology *33*, 56-72. 10.1038/sj.npp.1301555.
36. Courtin, J., Bienvenu, T.C., Einarsson, E.O., and Herry, C. (2013). Medial prefrontal cortex neuronal circuits in fear behavior. Neuroscience *240*, 219-242. 10.1016/j.neuroscience.2013.03.001.
37. Sierra-Mercado, D., Padilla-Coreano, N., and Quirk, G.J. (2010). Dissociable roles of prelimbic and infralimbic cortices, ventral hippocampus, and basolateral amygdala in the expression and extinction of conditioned fear. Neuropsychopharmacology *36*, 529-538. npp2010184 [pii] 10.1038/npp.2010.184.
38. Vertes, R.P. (2004). Differential projections of the infralimbic and prelimbic cortex in the rat. Synapse *51*, 32-58. 10.1002/syn.10279.
39. Gabbott, P.L., Warner, T.A., Jays, P.R., Salway, P., and Busby, S.J. (2005). Prefrontal cortex in the rat: projections to subcortical autonomic, motor, and limbic centers. J Comp Neurol *492*, 145-177. 10.1002/cne.20738.
40. Pape, H.C., and Pare, D. (2010). Plastic synaptic networks of the amygdala for the acquisition, expression, and extinction of conditioned fear. Physiol Rev *90*, 419-463. 10.1152/physrev.00037.2009.
41. Royer, S., Martina, M., and Pare, D. (1999). An inhibitory interface gates impulse traffic between the input and output stations of the amygdala. J Neurosci *19*, 10575-10583.
42. Likhtik, E., Popa, D., Apergis-Schoute, J., Fidacaro, G.A., and Pare, D. (2008). Amygdala intercalated neurons are required for expression of fear extinction. Nature *454*, 642-645. 10.1038/nature07167.

43. Santini, E., Ge, H., Ren, K., Pena de Ortiz, S., and Quirk, G.J. (2004). Consolidation of fear extinction requires protein synthesis in the medial prefrontal cortex. J Neurosci *24*, 5704-5710. 10.1523/JNEUROSCI.0786-04.2004 24/25/5704 [pii].
44. Rozeske, R.R., Jercog, D., Karalis, N., Chaudun, F., Khoder, S., Girard, D., Winke, N., and Herry, C. (2018). Prefrontal-Periaqueductal Gray-Projecting Neurons Mediate Context Fear Discrimination. Neuron *97*, 898-910 e896. 10.1016/j.neuron.2017.12.044.
45. McGarry, L.M., and Carter, A.G. (2017). Prefrontal Cortex Drives Distinct Projection Neurons in the Basolateral Amygdala. Cell reports *21*, 1426-1433. 10.1016/j.celrep.2017.10.046.
46. Marek, R., Xu, L., Sullivan, R.K.P., and Sah, P. (2018). Excitatory connections between the prelimbic and infralimbic medial prefrontal cortex show a role for the prelimbic cortex in fear extinction. Nat Neurosci *21*, 654-658. 10.1038/s41593-018-0137-x.
47. Concina, G., Cambiaghi, M., Renna, A., and Sacchetti, B. (2018). Coherent Activity between the Prelimbic and Auditory Cortex in the Slow-Gamma Band Underlies Fear Discrimination. J Neurosci *38*, 8313-8328. 10.1523/JNEUROSCI.0540-18.2018.
48. Buschman, T.J., and Miller, E.K. (2007). Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. Science *315*, 1860-1862. 10.1126/science.1138071.
49. Tomita, H., Ohbayashi, M., Nakahara, K., Hasegawa, I., and Miyashita, Y. (1999). Top-down signal from prefrontal cortex in executive control of memory retrieval. Nature *401*, 699-703. 10.1038/44372.
50. Sharpe, M.J., and Killcross, S. (2018). Modulation of attention and action in the medial prefrontal cortex of rats. Psychological review *125*, 822-843. 10.1037/rev0000118.
51. Rajasethupathy, P., Sankaran, S., Marshel, J.H., Kim, C.K., Ferenczi, E., Lee, S.Y., Berndt, A., Ramakrishnan, C., Jaffe, A., Lo, M., et al. (2015). Projections from neocortex mediate top-down control of memory retrieval. Nature *526*, 653-659. 10.1038/nature15389.
52. Seidenbecher, T., Laxmi, T.R., Stork, O., and Pape, H.C. (2003). Amygdalar and hippocampal theta rhythm synchronization during fear memory retrieval. Science *301*, 846-850. 10.1126/science.1085818.
53. Popa, D., Duvarci, S., Popescu, A.T., Lena, C., and Pare, D. (2010). Coherent amygdalocortical theta promotes fear memory consolidation during paradoxical sleep. Proc Natl Acad Sci U S A *107*, 6516-6519. 10.1073/pnas.0913016107.
54. Likhtik, E., Stujenske, J.M., Topiwala, M.A., Harris, A.Z., and Gordon, J.A. (2014). Prefrontal entrainment of amygdala activity signals safety in learned fear and innate anxiety. Nat Neurosci *17*, 106-113. 10.1038/nn.3582.
55. Karalis, N., Dejean, C., Chaudun, F., Khoder, S., Rozeske, R.R., Wurtz, H., Bagur, S., Benchenane, K., Sirota, A., Courtin, J., and Herry, C. (2016). 4-Hz oscillations synchronize prefrontal-amygdala circuits during fear behavior. Nat Neurosci *19*, 605-612. 10.1038/nn.4251.
56. Corches, A., Hiroto, A., Bailey, T.W., Speigel, J.H., 3rd, Pastore, J., Mayford, M., and Korzus, E. (2019). Differential fear conditioning generates prefrontal neural ensembles of safety signals. Behav Brain Res *360*, 169-184. 10.1016/j.bbr.2018.11.042.
57. Grewe, B.F., Grundemann, J., Kitch, L.J., Lecoq, J.A., Parker, J.G., Marshall, J.D., Larkin, M.C., Jercog, P.E., Grenier, F., Li, J.Z., et al. (2017). Neural ensemble dynamics underlying a long-term associative memory. Nature *543*, 670-675. 10.1038/nature21682.
58. Grosso, A., Santoni, G., Manassero, E., Renna, A., and Sacchetti, B. (2018). A neuronal basis for fear discrimination in the lateral amygdala. Nature communications *9*, 1214. 10.1038/s41467-018-03682-2.
59. Sangha, S., Chadick, J.Z., and Janak, P.H. (2013). Safety encoding in the basal amygdala. J Neurosci *33*, 3744-3751. 10.1523/JNEUROSCI.3302-12.2013.
60. Stujenske, J.M., O'Neill, P.K., Fernandes-Henriques, C., Nahmoud, I., Goldburg, S.R., Singh, A., Diaz, L., Labkovich, M., Hardin, W., Bolkan, S.S., et al. (2022). Prelimbic cortex drives discrimination of non-aversion via amygdala somatostatin interneurons. Neuron *110*, 2258-2267 e2211. 10.1016/j.neuron.2022.03.020.
61. Ziv, Y., Burns, L.D., Cocker, E.D., Hamel, E.O., Ghosh, K.K., Kitch, L.J., El Gamal, A., and Schnitzer, M.J. (2013). Long-term dynamics of CA1 hippocampal place codes. Nat Neurosci *16*, 264-266. 10.1038/nn.3329.
62. Cai, D.J., Aharoni, D., Shuman, T., Shobe, J., Biane, J., Song, W., Wei, B., Veshkini, M., La-Vu, M., Lou, J., et al. (2016). A shared neural ensemble links distinct contextual memories encoded close in time. Nature *534*, 115-118. 10.1038/nature17955.

63. McHugh, T.J., Jones, M.W., Quinn, J.J., Balthasar, N., Coppari, R., Elmquist, J.K., Lowell, B.B., Fanselow, M.S., Wilson, M.A., and Tonegawa, S. (2007). Dentate gyrus NMDA receptors mediate rapid pattern separation in the hippocampal network. Science *317*, 94-99. 1140263 [pii] 10.1126/science.1140263.
64. Korzus, E., Rosenfeld, M.G., and Mayford, M. (2004). CBP histone acetyltransferase activity is a critical component of memory consolidation. Neuron *42*, 961-972. 10.1016/j.neuron.2004.06.002.
65. Korzus, E., Torchia, J., Rose, D.W., Xu, L., Kurokawa, R., McInerney, E.M., Mullen, T.M., Glass, C.K., and Rosenfeld, M.G. (1998). Transcription factor-specific requirements for coactivators and their acetyltransferase functions. Science *279*, 703-707. 10.1126/science.279.5351.703.
66. Alarcon, J.M., Malleret, G., Touzani, K., Vronskaya, S., Ishii, S., Kandel, E.R., and Barco, A. (2004). Chromatin acetylation, memory, and LTP are impaired in CBP+/- mice: a model for the cognitive deficit in Rubinstein-Taybi syndrome and its amelioration. Neuron *42*, 947-959. 10.1016/j.neuron.2004.05.021 S0896627304003022 [pii].
67. Wood, M.A., Kaplan, M.P., Park, A., Blanchard, E.J., Oliveira, A.M., Lombardi, T.L., and Abel, T. (2005). Transgenic mice expressing a truncated form of CREB-binding protein (CBP) exhibit deficits in hippocampal synaptic plasticity and memory storage. Learn Mem *12*, 111-119. 12/2/111 [pii] 10.1101/lm.86605.
68. Maddox, S.A., Watts, C.S., and Schafe, G.E. (2013). p300/CBP histone acetyltransferase activity is required for newly acquired and reactivated fear memories in the lateral amygdala. Learn Mem *20*, 109-119. 10.1101/lm.029157.112.
69. Barrett, R.M., Malvaez, M., Kramar, E., Matheos, D.P., Arrizon, A., Cabrera, S.M., Lynch, G., Greene, R.W., and Wood, M.A. (2011). Hippocampal focal knockout of CBP affects specific histone modifications, long-term potentiation, and long-term memory. Neuropsychopharmacology *36*, 1545-1556. npp201161 [pii] 10.1038/npp.2011.61.
70. Valor, L.M., Pulopulos, M.M., Jimenez-Minchan, M., Olivares, R., Lutz, B., and Barco, A. (2011). Ablation of CBP in forebrain principal neurons causes modest memory and transcriptional defects and a dramatic reduction of histone acetylation but does not affect cell viability. J Neurosci *31*, 1652-1663. 31/5/1652 [pii] 10.1523/JNEUROSCI.4737-10.2011.
71. Valor, L.M., Viosca, J., Lopez-Atalaya, J.P., and Barco, A. (2013). Lysine acetyltransferases CBP and p300 as therapeutic targets in cognitive and neurodegenerative disorders. Curr Pharm Des *19*, 5051-5064.
72. Chen, G., Zou, X., Watanabe, H., van Deursen, J.M., and Shen, J. (2010). CREB binding protein is required for both short-term and long-term memory formation. J Neurosci *30*, 13066-13077. 30/39/13066 [pii] 10.1523/JNEUROSCI.2378-10.2010.
73. Peixoto, L., and Abel, T. (2013). The role of histone acetylation in memory formation and cognitive impairments. Neuropsychopharmacology *38*, 62-76. 10.1038/npp.2012.86.
74. Roy, D.S., Kitamura, T., Okuyama, T., Ogawa, S.K., Sun, C., Obata, Y., Yoshiki, A., and Tonegawa, S. (2017). Distinct Neural Circuits for the Formation and Retrieval of Episodic Memories. Cell *170*, 1000-1012 e1019. 10.1016/j.cell.2017.07.013.
75. Resendez, S.L., Jennings, J.H., Ung, R.L., Namboodiri, V.M., Zhou, Z.C., Otis, J.M., Nomura, H., McHenry, J.A., Kosyk, O., and Stuber, G.D. (2016). Visualization of cortical, subcortical and deep brain neural circuit dynamics during naturalistic mammalian behavior with head-mounted microscopes and chronically implanted lenses. Nat Protoc *11*, 566-597. 10.1038/nprot.2016.021.
76. Pinto, L., and Dan, Y. (2015). Cell-Type-Specific Activity in Prefrontal Cortex during Goal-Directed Behavior. Neuron *87*, 437-450. 10.1016/j.neuron.2015.06.021.
77. Ghosh, K.K., Burns, L.D., Cocker, E.D., Nimmerjahn, A., Ziv, Y., Gamal, A.E., and Schnitzer, M.J. (2011). Miniaturized integration of a fluorescence microscope. Nat Methods *8*, 871-878. 10.1038/nmeth.1694.
78. Sheintuch, L., Rubin, A., Brande-Eilat, N., Geva, N., Sadeh, N., Pinchasof, O., and Ziv, Y. (2017). Tracking the Same Neurons across Multiple Days in Ca(2+) Imaging Data. Cell reports *21*, 1102-1115. 10.1016/j.celrep.2017.10.013.
79. Bader, B.W., and Kolda, T.G. (2020). MATLAB Tensor Toolbox: B. W. Bader and T. G. Kolda. Algorithm 862: MATLAB tensor classes for fast algorithm prototyping, ACM Transactions on Mathematical Software 32(4):635-653, December 2006. DOI: 10.1145/1186785.1186794. https://www.tensortoolbox.org.
80. Reijmers, L.G., Perkins, B.L., Matsuo, N., and Mayford, M. (2007). Localization of a stable neural correlate of associative memory. Science *317*, 1230-1233. 317/5842/1230 [pii]

10.1126/science.1143839.
81. Kitamura, T., Sun, C., Martin, J., Kitch, L.J., Schnitzer, M.J., and Tonegawa, S. (2015). Entorhinal Cortical Ocean Cells Encode Specific Contexts and Drive Context-Specific Fear Memory. Neuron *87*, 1317-1331. 10.1016/j.neuron.2015.08.036.
82. Josselyn, S.A., and Tonegawa, S. (2020). Memory engrams: Recalling the past and imagining the future. Science *367*. 10.1126/science.aaw4325.
83. Jimenez, J.C., Su, K., Goldberg, A.R., Luna, V.M., Biane, J.S., Ordek, G., Zhou, P., Ong, S.K., Wright, M.A., Zweifel, L., et al. (2018). Anxiety Cells in a Hippocampal-Hypothalamic Circuit. Neuron *97*, 670-683 e676. 10.1016/j.neuron.2018.01.016.
84. Ghandour, K., Ohkawa, N., Fung, C.C.A., Asai, H., Saitoh, Y., Takekawa, T., Okubo-Suzuki, R., Soya, S., Nishizono, H., Matsuo, M., et al. (2019). Orchestrated ensemble activities constitute a hippocampal memory engram. Nature communications *10*, 2637. 10.1038/s41467-019-10683-2.
85. Pasricha, R.S., Saini, U.S., Sidiropoulos, N.D., Fang, F., Chan, K., and Papalexakis, E.E. (2023). Harvester: Principled Factorization-based Temporal Tensor Granularity Estimation. In Proceedings of the 2023 SIAM International Conference on Data Mining (SDM), pp. 82-90. 10.1137/1.9781611977653.ch10.
86. Pasricha, R.S., Gujral, E., and Papalexakis, E.E. (2022). Adaptive granularity in tensors: A quest for interpretable structure. Front Big Data *5*, 929511. 10.3389/fdata.2022.929511.
87. Papalexakis, E.E., Akoglu, L., and Ienco, D. Do more views of a graph help? community detection and clustering in multi-graphs. In Proceedings of the 16th International Conference on Information Fusion, 2013.
88. Gujral, E., and Papalexakis, E.E. Smacd: Semi-supervised multi-aspect community detection. Proceedings of the 2018 SIAM International Conference on Data Mining, 2018. pp. 702-710.
89. Fornito, A., Zalesky, A., and Bullmore, E.T. (2016). Fundamentals of brain network analysis (Elsevier/Academic Press).
90. Rubinov, M., and Sporns, O. (2010). Complex network measures of brain connectivity: uses and interpretations. Neuroimage *52*, 1059-1069. 10.1016/j.neuroimage.2009.10.003.
91. Milad, M.R., Pitman, R.K., Ellis, C.B., Gold, A.L., Shin, L.M., Lasko, N.B., Zeidan, M.A., Handwerger, K., Orr, S.P., and Rauch, S.L. (2009). Neurobiological basis of failure to recall extinction memory in posttraumatic stress disorder. Biol Psychiatry *66*, 1075-1082. 10.1016/j.biopsych.2009.06.026.
92. Bergado-Acosta, J.R., Sangha, S., Narayanan, R.T., Obata, K., Pape, H.C., and Stork, O. (2008). Critical role of the 65-kDa isoform of glutamic acid decarboxylase in consolidation and generalization of Pavlovian fear memory. Learn Mem *15*, 163-171. 10.1101/lm.705408.
93. Shaban, H., Humeau, Y., Herry, C., Cassasus, G., Shigemoto, R., Ciocchi, S., Barbieri, S., van der Putten, H., Kaupmann, K., Bettler, B., and Luthi, A. (2006). Generalization of amygdala LTP and conditioned fear in the absence of presynaptic inhibition. Nat Neurosci *9*, 1028-1035. 10.1038/nn1732.
94. Lesting, J., Narayanan, R.T., Kluge, C., Sangha, S., Seidenbecher, T., and Pape, H.C. (2011). Patterns of coupled theta activity in amygdala-hippocampal-prefrontal cortical circuits during fear extinction. PLoS One *6*, e21714. 10.1371/journal.pone.0021714.
95. O'Reilly, R.C., and McClelland, J.L. (1994). Hippocampal conjunctive encoding, storage, and recall: avoiding a trade-off. Hippocampus *4*, 661-682. 10.1002/hipo.450040605.
96. Roy, D.S., Okuyama, T., and Tonegawa, S. (2017). Tagging activated neurons with light. Nat Biotechnol *35*, 827-828. 10.1038/nbt.3954.
97. Moita, M.A., Rosis, S., Zhou, Y., LeDoux, J.E., and Blair, H.T. (2003). Hippocampal place cells acquire location-specific responses to the conditioned stimulus during auditory fear conditioning. Neuron *37*, 485-497. 10.1016/s0896-6273(03)00033-3.
98. Blair, G.J., Guo, C., Wang, S., Fanselow, M.S., Golshani, P., Aharoni, D., and Blair, H.T. (2023). Hippoc ampal place cell remapping occurs with memory storage of aversive experiences. eLife *12*. 10.7554/eLife.80661.
99. Barabasi, D.L., Bianconi, G., Bullmore, E., Burgess, M., Chung, S., Eliassi-Rad, T., George, D., Kovacs, I.A., Makse, H., Nichols, T.E., et al. (2023). Neuroscience Needs Network Science. J Neurosci *43*, 5989-5995. 10.1523/JNEUROSCI.1014-23.2023.
100. Laubach, M., Amarante, L.M., Swanson, K., and White, S.R. (2018). What, If Anything, Is Rodent Prefrontal Cortex? eNeuro *5*. 10.1523/ENEURO.0315-18.2018.
101. Kenwood, M.M., Kalin, N.H., and Barbas, H. (2021). The prefrontal cortex, pathological anxiety, and anxiety disorders. Neuropsychopharmacology. 10.1038/s41386-021-01109-z.
102. Bader, B.W., Kolda, T.G., and al., e. (2023). Tensor Toolbox for MATLAB, Version 3.6, www.tensortoolbox.org. www.tensortoolbox.org.

103. Schneider, C.A., Rasband, W.S., and Eliceiri, K.W. (2012). NIH Image to ImageJ: 25 years of image analysis. Nat Methods *9*, 671-675. 10.1038/nmeth.2089.
104. Aharoni, D., and Hoogland, T.M. (2019). Circuit Investigations With Open-Source Miniaturized Microscopes: Past, Present and Future. Frontiers in cellular neuroscience *13*, 141. 10.3389/fncel.2019.00141.
105. Pnevmatikakis, E.A., and Giovannucci, A. (2017). NoRMCorre: An online algorithm for piecewise rigid motion correction of calcium imaging data. J Neurosci Methods *291*, 83-94. 10.1016/j.jneumeth.2017.07.031.
106. Zhou, P., Resendez, S.L., Rodriguez-Romaguera, J., Jimenez, J.C., Neufeld, S.Q., Giovannucci, A., Friedrich, J., Pnevmatikakis, E.A., Stuber, G.D., Hen, R., et al. (2018). Efficient and accurate extraction of in vivo calcium signals from microendoscopic video data. eLife *7*. 10.7554/eLife.28728.
107. Papalexakis, E.E., Sidiropoulos, N.D., and Bro, R. (2012). From K-means to higher-way co-clustering: multilinear decomposition with sparse latent factors. IEEE TRANSACTIONS ON SIGNAL PROCESSING, *Originally presented: Proc. IEEE ICASSP 2011, pp. 2064-2067, May 22-27, Prague - Czech Republic.*
108. Kolda, T.G., and Bader, B.W. (2009). Tensor decompositions and applications. SIAM review *51*, 455-500.
109. Lee, D.D., and Seung, H.S. (1999). Learning the parts of objects by non-negative matrix factorization. Nature *401*, 788-791. 10.1038/44565.
110. Rasmus, B., and Kiers, H.A.L. (2003). A new efficient method for determining the number of components in PARAFAC models. Journal of Chemometrics *17*, 274-286. https://doi.org/10.1002/cem.801.
111. Bro, R., and Kiers, H.A.L. (2003). A new efficient method for determining the number of components in PARAFAC models. Journal of Chemometrics *17*, 274-286. https://doi.org/10.1002/cem.801.
112. Papalexakis, E.E., and Faloutsos, C. (2015). Fast efficient and scalable Core Consistency Diagnostic for the parafac decomposition for big sparse tensors. 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP).
113. Tsitsikas, Y., and Papalexakis, E.E. (2020). NSVD: Normalized Singular Value Deviation Reveals Number of Latent Factors in Tensor Decomposition. Big Data *8*, 412-430. 10.1089/big.2020.0074.
114. Papalexakis, E.E., and Faloutsos, C. (2016). Unsupervised Tensor Mining for Big Data Practitioners. Big Data *4*, 179-191. 10.1089/big.2016.0026.
115. Liu, K., Da Costa, J.P.C., So, H.C., Huang, L., and Ye, J. (2016). Detection of number of components in CANDECOMP/PARAFAC models via minimum description length. Digital Signal Processing *51*, 110–123. https://doi.org/10.1016/j.dsp.2016.01.003.
116. Mørup, M., and Hansen, L.K. (2009). Automatic relevance determination for multi-way models. J. Chemometr. *23*, 352-363. https://doi.org/10.1002/cem.1223.
117. Acar, E., Aykut-Bingol, C., Bingol, H., Bro, R., and Yener, B. (2007). Multiway analysis of epilepsy tensors. Bioinformatics *23*, i10-18. 10.1093/bioinformatics/btm210.
118. Miwakeichi, F., Martinez-Montes, E., Valdes-Sosa, P.A., Nishiyama, N., Mizuhara, H., and Yamaguchi, Y. (2004). Decomposing EEG data into space-time-frequency components using Parallel Factor Analysis. Neuroimage *22*, 1035-1045. 10.1016/j.neuroimage.2004.03.039.
119. Papalexakis, E.E. (2016). Automatic unsupervised tensor mining with quality assessment. Proceedings of the SIAM International Conference on Data Mining, pp. 711–719.
120. Blondel, V.D., Guillaume, J.-L., Lambiotte, R., and Lefebvre, E. (2008). Fast unfolding of communities in large networks. Journal of Statistical Mechanics: Theory and Experiment *P10008*, 1-12. 10.1088/1742-5468/2008/10/p10008.
121. Rubinov, M., and Sporns, O. (2011). Weight-conserving characterization of complex functional brain networks. Neuroimage *56*, 2068-2079. 10.1016/j.neuroimage.2011.03.069.
122. Reichardt, J., and Bornholdt, S. (2006). Statistical mechanics of community detection. Physical review. E, Statistical, nonlinear, and soft matter physics *74*, 016110. 10.1103/PhysRevE.74.016110.