

Multi-regional control of amygdalar dynamics reliably reflects fear memory age

Received: 2 May 2024

Accepted: 4 November 2024

Published online: 09 December 2024

Yuichi Makino ^{1,2,4}✉, Yi Wang ^{1,3,4}✉ & Thomas J. McHugh ¹✉

The basolateral amygdala (BLA) is crucial for the encoding and expression of fear memory, yet it remains unexplored how neural activity in this region is dynamically influenced by distributed circuits across the brain to facilitate expression of fear memory of different ages. Using longitudinal multisite electrophysiological recordings in male mice, we find that the recall of older contextual fear memory is accompanied by weaker, yet more rhythmic, BLA gamma activity which is distally entrained by theta oscillations in both the hippocampal CA1 and the anterior cingulate cortex. Computational modeling with Light Gradient Boosting Machine using extracted oscillatory features from these three regions, as well as with Transformer using raw local field potentials, accurately classified remote from recent memory recall primarily based on BLA gamma and CA1 theta. These results demonstrate in a non-biased manner that multi-regional control of BLA activity serves as reliable neural signatures for memory age-dependent recall mechanisms.

The BLA is involved in emotional information processing, particularly fear learning and memory¹. It has widely been reported that BLA is required for both newly acquired (recent) and older, long-lasting (remote) fear memory^{2–5}. These studies suggest a general participation of BLA in fear memory expression regardless of memory age, however it remains unexplored whether BLA utilizes identical or distinct mechanisms at the level of dynamic neural activity to facilitate recall at these different timepoints.

Importantly, there is accumulating evidence that the consolidation and long-term storage of memory requires multiple brain regions, such as the hippocampus and prefrontal cortex^{6–8}. Also, long-range dynamic neural interaction between these regions and the BLA has been reported, with its implicated roles in emotional memory processes^{9,10}. These observations raise the possibility that distant regions differentially entrain BLA to generate specific neural activity patterns underlying recent and remote fear memory recall, respectively. Further, if these multi-regional patterns of neural activity serve as reliable signatures for distinct forms of recall, they could be computationally analyzed to decode the age of memory. In particular, recent advancements in deep neural network modeling now may allow us to identify

neural activity patterns that are most characteristic of distinct memory ages in a non-biased manner by using raw neural signal as the input. However, such decoding has yet to be implemented.

Here, we report that recall of recent and remote contextual fear memory exhibits distinct dynamic neural activity patterns within the BLA, as well as differential cross-regional influence of those patterns. Specifically, the intensity of BLA gamma activity and its phase-amplitude coupling to theta oscillations recorded in CA1 and the anterior cingulate cortex (ACC) characterize recent and remote recall, respectively. Further, both feature-based modeling by Light Gradient Boosting Machine (LightGBM¹¹) and feature-independent modeling by Transformer could classify recent from remote memory at a high accuracy, primarily relying on BLA gamma and CA1 theta as revealed by our method of interpreting model structure.

Results

Gamma activity in the BLA is enhanced during recent memory recall

To explore neural activity in the BLA and its entrainment by distal regions associated with different memory ages, we implanted a

¹Laboratory for Circuit and Behavioral Physiology, RIKEN Center for Brain Science, Wako-shi, Saitama, Japan. ²International Research Center for Neurointelligence, UTIAS, The University of Tokyo, Bunkyo-ku, Tokyo, Japan. ³School of Mathematical and Computational Sciences, Massey University, Palmerston North, New Zealand. ⁴These authors contributed equally: Yuichi Makino, Yi Wang. ✉e-mail: makino-yuichi@g.ecc.u-tokyo.ac.jp; y.wang24@massey.ac.nz; thomas.mchugh@riken.jp

microdrive containing electrodes targeting the BLA, CA1 and ACC in freely moving mice. After recovery from surgery, mice were placed in a fear conditioning chamber and neural activity was recorded during a baseline session. Shortly after, mice were fear-conditioned in the same chamber with electrical shock, leading to the formation of a long-lasting contextual fear memory. Mice were returned to the same chamber 1 day (recent recall) and 25–27 days (remote recall) after conditioning, and memory was assessed by the level of freezing while neural activity was recorded (Fig. 1a). This longitudinal design was explicitly chosen to ensure equivalent recording positions and quality in these brain regions across time and to allow within-subject comparisons of the neurophysiology data, which could eliminate possible artefacts arising from individual variability with a cross-subject design^{12,13}. Further, we note that the raw electrophysiological recording data from CA1 and ACC, as well as the behavioral data used in this paper, are identical to those used in our previous publication¹⁴, where we reported that percentage of freezing, as well as locomotion and behavioral patterns during non-freezing periods were both comparable between the recent and remote time points (Fig. 1b and Supplementary Fig. 1).

It has been shown that oscillatory neural activity and synchrony across regions, particularly in the theta and gamma frequency ranges, play a critical role in various aspects of cognitive behavior^{9,15,16}. To probe potential activity differences in the BLA during recent and remote recall, we examined oscillatory power (power spectral density; PSD) of local field potential (LFP) in the BLA during the two recall sessions (Fig. 1c). The time course of BLA power exhibited intermittent transient peaks in the slow gamma range (30–50 Hz), particularly during non-freezing periods shortly before the onset of freezing (Fig. 1d). Consistent with this, BLA power averaged across animals showed clear peaks in this frequency range during non-freezing periods of all sessions (Fig. 1e). Additionally, similar peaks were also observed during freezing periods, albeit in a slightly higher frequency range (Fig. 1e). This slow gamma power was significantly higher during the recent compared to the baseline and remote recall sessions during non-freezing periods (Fig. 1f), suggesting that heightened slow gamma activity is associated specifically with post-learning processes and selectively at this early time point. The increase of slow gamma power in the recent session remained significant even after excluding the influence of locomotion (Supplementary Fig. 2a), indicating that it was

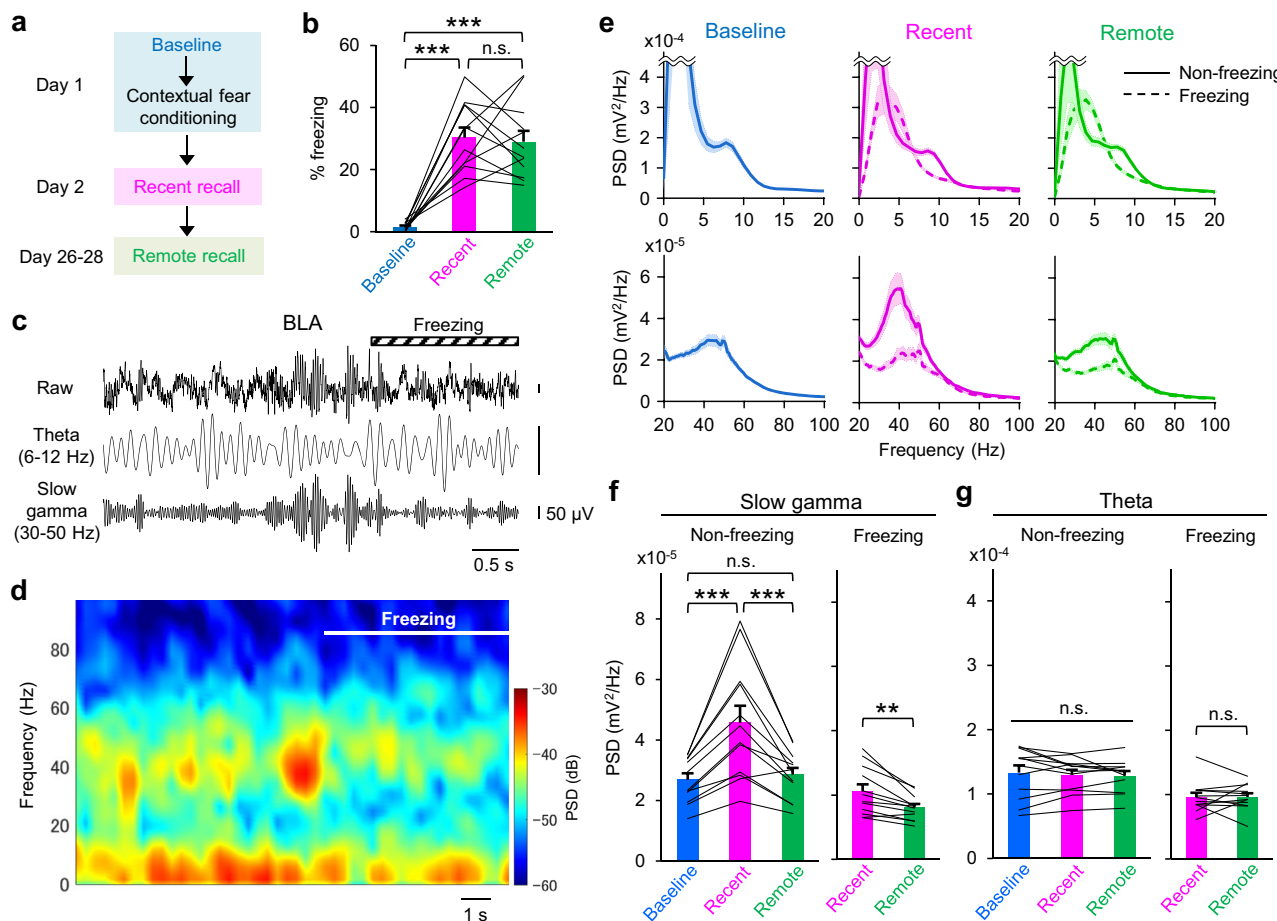


Fig. 1 | BLA slow gamma activity is augmented during recent memory recall.

a Experimental timeline. The activity of BLA, CA1 and ACC was recorded for 10 min in the fear conditioning chamber immediately before (baseline), 1 day (Day 2; recent recall) and 25–27 days (Day 26–28; remote recall) after contextual fear conditioning. **b** Percentage of freezing during the baseline, recent, and remote sessions ($F_{2,22} = 36.91$, $p = 9 \times 10^{-8}$; baseline-recent, $p = 4 \times 10^{-7}$; baseline-remote, $p = 8 \times 10^{-7}$; recent-remote, $p = 0.938$). **c** Representative raw LFP trace in the BLA as well as traces filtered at theta (6–12 Hz) and slow gamma (30–50 Hz) frequencies. **d** Transition of oscillatory power (PSD) of LFP in the BLA over time during recent recall session of a representative mouse. **e** PSD of BLA LFP during non-freezing and freezing periods at baseline, recent and remote sessions averaged over mice.

f PSD of BLA LFP averaged over slow gamma range (30–50 Hz) during non-freezing (main effect, $F_{2,22} = 21.48$, $p = 7 \times 10^{-6}$; baseline-recent, $p = 2 \times 10^{-5}$; baseline-remote, $p = 0.863$; recent-remote, $p = 6 \times 10^{-5}$) and freezing ($t_{11} = 3.97$, $p = 0.002$) periods. **g** PSD of BLA LFP averaged over theta range (6–12 Hz) during non-freezing (main effect, $F_{2,22} = 0.50$, $p = 0.614$) and freezing ($t_{11} = 0.05$, $p = 0.959$) periods. In (**b–g**), $n = 12$ mice, one-way repeated measures ANOVA with post hoc comparisons using Tukey's HSD test (non-freezing) or two-tailed paired t -test (freezing). For all bar graphs and line plots, data are represented as mean \pm SEM with values from individual mice. $**p < 0.01$ and $***p < 0.001$. Source data are available in the Source Data file.

not caused by a difference in locomotor activity. Likewise, slow gamma power was higher during recent compared to remote session during freezing periods (Fig. 1f), implying its involvement in fear expression itself. In contrast, activity at the theta frequency (6–12 Hz) during both non-freezing and freezing periods remained unchanged across sessions (Fig. 1g), demonstrating the specificity of changes in the slow gamma range. In addition, activity in the 2–6 Hz range during freezing periods, which was previously associated with fear memory¹⁷ and replicated in our data as sharp peaks in the PSD curves during those periods (Fig. 1e), remained unchanged across sessions (Supplementary Fig. 3a). These results suggest that the enhanced slow gamma activity in the BLA is associated with recall mechanisms specific to the recent timepoint, possibly reflecting activity changes of neuronal population in the BLA during recent recall¹⁸.

Entrainment of BLA gamma by theta activity of multiple regions is augmented during remote recall

The amplitude of gamma activity is known to be modulated by the phase of the theta oscillation in various regions¹⁵, and this theta-gamma phase-amplitude coupling (PAC) has also been observed in the BLA associated with fear memory¹⁹. To examine if the increase of gamma activity during recent recall is associated with altered theta-gamma coupling, we calculated the PAC of LFP in the BLA during recent and remote memory recall (Fig. 2a). PAC strength, quantified by the modulation index²⁰, was largest between theta phase and slow gamma amplitude during non-freezing periods (Fig. 2b), suggesting that theta to slow gamma coupling is predominant during these epochs. In contrast, during freezing the peak of phase-amplitude coupling was shifted to the 2–6 Hz range (Fig. 2b), suggesting distinct modes of cross-frequency coupling control BLA slow gamma during non-freezing and freezing periods.

We then examined the changes in theta to slow gamma modulation across sessions. Similar to the power of slow gamma activity, theta to slow gamma PAC increased from baseline to the recent recall session during non-freezing periods (Fig. 2c). Surprisingly, however, theta to slow gamma PAC demonstrated an even larger significant increase during remote recall (Fig. 2c), despite an overall decrease in slow gamma power (Fig. 1f). Again, this increased PAC remained significant even after excluding the influence of locomotion (Supplementary Fig. 2b). These results suggest that the timing in relation to theta, but not the intensity, of gamma activity may best reflect remote recall processes. Notably, theta to slow gamma PAC, as well as 2–6 Hz to slow gamma PAC, did not increase from recent to remote sessions during freezing periods (Fig. 2c and Supplementary Fig. 3b). Overall, these observations demonstrate that theta-entrained phase control of BLA gamma during exposure to recall cues during non-freezing periods, but not during freezing itself, may differentiate remote from recent recall, consistent with the augmented theta synchrony between CA1 and ACC during remote recall specifically observed during non-freezing periods¹⁴.

To explore how theta activity in CA1 and ACC is linked to BLA gamma, we examined cross-frequency coupling between CA1 phase and BLA amplitude, as well as between ACC phase and BLA amplitude (Fig. 2d, g). Similar to the within-BLA modulation, PAC of both regional pairs peaks at theta phase and slow gamma amplitude during non-freezing periods and at 2–6 Hz phase and slow gamma amplitude during freezing periods (Fig. 2e, h), implying the same mode of modulation within and across regions. Likewise, theta to slow gamma PAC was higher during remote compared to recent recall session for both regional pairs during non-freezing periods (Fig. 2f, i), which was again not affected by locomotion (Supplementary Fig. 2b). In contrast, PAC from either theta or 2–6 Hz to slow gamma was not increased in freezing periods of the remote recall session (Fig. 2f, i and Supplementary Fig. 3b). Taken together, these data suggest that remote recall may be supported by multi-regional phase control of BLA gamma

activity, while recent recall is better reflected in the intensity of gamma independent of distal control, implying distinct recall mechanisms mediated by BLA gamma activity.

As enhanced theta synchronization between CA1 and ACC during remote memory recall has been reported, observed as the increase in both amplitude correlation and coherence^{14,21}, it is possible that theta interaction across the three regions increases to co-entrain BLA gamma. To test this hypothesis, we first calculated the correlation of the instantaneous amplitude of the LFP at theta range between CA1 and ACC²² then looked at how this value correlates with BLA theta amplitude (Supplementary Fig. 4a). This tri-regional correlation index was significantly higher during remote recall session compared to the baseline and recent recall session during non-freezing periods (Supplementary Fig. 4b, c), but not during freezing periods (Supplementary Fig. 4d, e). These results suggest that remote recall could involve an augmented synchronization of theta oscillations across CA1, ACC and BLA, potentially controlling BLA gamma activity related to theta phase in a concerted manner.

BLA gamma intensity and its multi-regional theta-phase control differentially predict the onset of freezing for recent and remote recall

To understand how the timing of the changes in BLA gamma activity and its cross-frequency modulation relate to when memory recall occurs, we analyzed these LFP indices specifically around the onset of freezing bouts. Slow gamma power in the BLA exhibited a local peak within a 2 s window prior to freezing onset only during the recent session, which was significantly higher than the average of slow gamma power over all non-freezing epochs (Fig. 3a). On the other hand, while theta to slow gamma PAC in the BLA exhibited local peaks significantly higher than non-freezing averages within a 2 s window prior to freezing onset for both recent and remote sessions (Fig. 3b), PAC of BLA slow gamma by theta in CA1 and ACC both demonstrated the significant local peaks within the same window only for the remote session (Fig. 3c, d). These results suggest that BLA gamma and its cross-regional theta-phase control predict the onset of freezing at a short timescale during recent and remote recall, respectively, supporting the idea that these LFP phenomena are associated with recall processes themselves at each memory age.

Both feature-based and feature-free computational models accurately predict memory age based on multi-regional LFP

The distinct oscillatory activity in the BLA, CA1 and ACC associated with recent and remote recall suggests that computational models may be used to distinguish the age of the memory being recalled, even across individuals. To explore this possibility we first constructed classifiers using feature-based LightGBM to test the classification effectiveness based on oscillatory activity in different frequency bands. In this model, we used theta and slow gamma power that was associated with recent and remote recall in this study, as well as beta (20–30 Hz) and fast gamma power (60–90 Hz) implicated in various behaviors and information processing in previous studies^{19,23–25}, in LFP recorded in the BLA, CA1 and ACC to construct feature vectors (Fig. 4a). We also added PAC of BLA slow gamma amplitude by theta phase of BLA, CA1 and ACC, which was distinct between recent and remote recall, to the feature vectors. We employed between-subject cross validation to train, validate, and test this feature-based memory age classification (see Methods for details). The model achieved an average F1 score of 0.720 across the 12 mice in the test set, which was significantly higher than the label-shuffled control analysis when LFP during non-freezing periods was used (Fig. 4b), demonstrating the capability of LightGBM to accurately classify memory age. The performance of the model with locomotion included as an additional feature was comparable to that of the original model (Suppl Fig. 5a, c), indicating that locomotion does not dissociate the two memory ages.

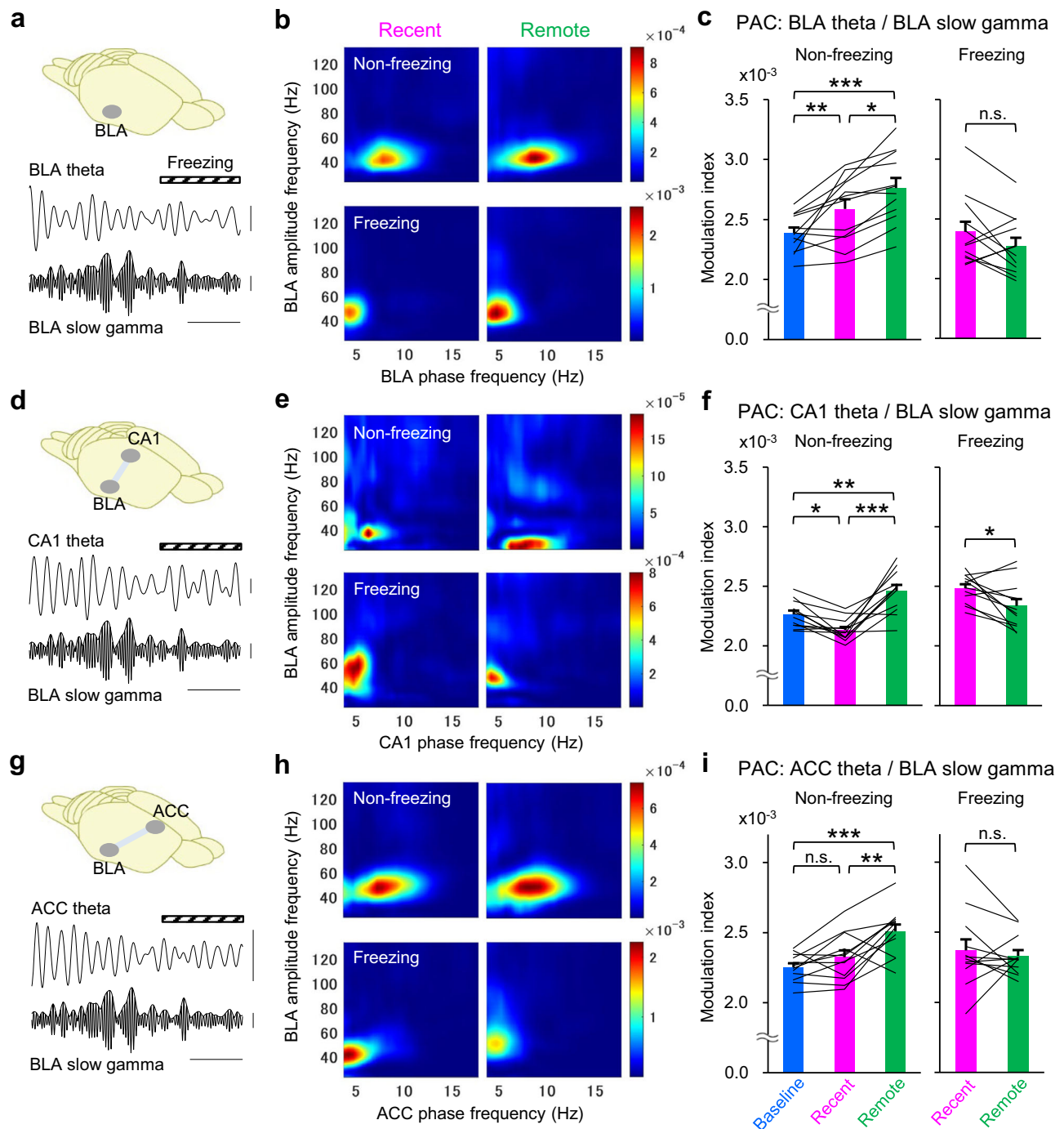


Fig. 2 | BLA slow gamma activity is theta-entrained by multiple regions during remote recall. **a** Representative BLA LFP filtered at theta and slow gamma frequencies overlaid with slow gamma envelope. **b** Representative comodulogram for cross-frequency phase-amplitude coupling (PAC) during non-freezing and freezing periods at recent and remote sessions, showing the strength of modulations of BLA LFP amplitude by BLA LFP phase using modulation index (see Methods). **c** Modulation index averaged over all mice for PAC of BLA slow gamma amplitude by BLA theta phase during non-freezing (main effect, $F_{2,22} = 19.50$, $p = 1 \times 10^{-5}$; baseline-recent, $p = 0.008$; baseline-remote, $p = 8 \times 10^{-6}$; recent-remote, $p = 0.021$) and freezing ($t_{11} = 1.82$, $p = 0.095$) periods. **d** Representative CA1 LFP filtered at theta frequency and BLA LFP filtered at slow gamma frequency overlaid with its envelope. **e** Representative comodulogram for PAC of BLA LFP amplitude by CA1 LFP phase at recent and remote sessions. **f** Modulation index for PAC of BLA slow gamma amplitude by CA1 theta phase during non-freezing (main effect,

$F_{2,22} = 22.34$, $p = 5 \times 10^{-6}$; baseline-recent, $p = 0.036$; baseline-remote, $p = 0.002$; recent-remote, $p = 3 \times 10^{-6}$) and freezing ($t_{11} = 2.63$, $p = 0.023$) periods. **g** Representative ACC LFP filtered at theta frequency and BLA LFP filtered at slow gamma frequency overlaid with its envelope. **h** Representative comodulogram for PAC of BLA LFP amplitude by ACC LFP phase at recent and remote sessions. **i** Modulation index for PAC of BLA slow gamma amplitude by ACC theta phase during non-freezing (main effect, $F_{2,22} = 16.52$, $p = 4 \times 10^{-5}$; baseline-recent, $p = 0.286$; baseline-remote, $p = 3 \times 10^{-5}$; recent-remote, $p = 0.002$) and freezing ($t_{11} = 0.59$, $p = 0.565$) periods. In (a, d, g), vertical scale bars = 50 μ V, horizontal scale bars = 0.5 s. In (c, f, i), $n = 12$ mice, one-way repeated measures ANOVA with post hoc comparisons using Tukey's HSD test (non-freezing) or two-tailed paired t -test (freezing). For all bar graphs, data are represented as mean \pm SEM with values from individual mice. * $p < 0.05$, ** $p < 0.01$ and *** $p < 0.001$. Source data are available in the Source Data file.

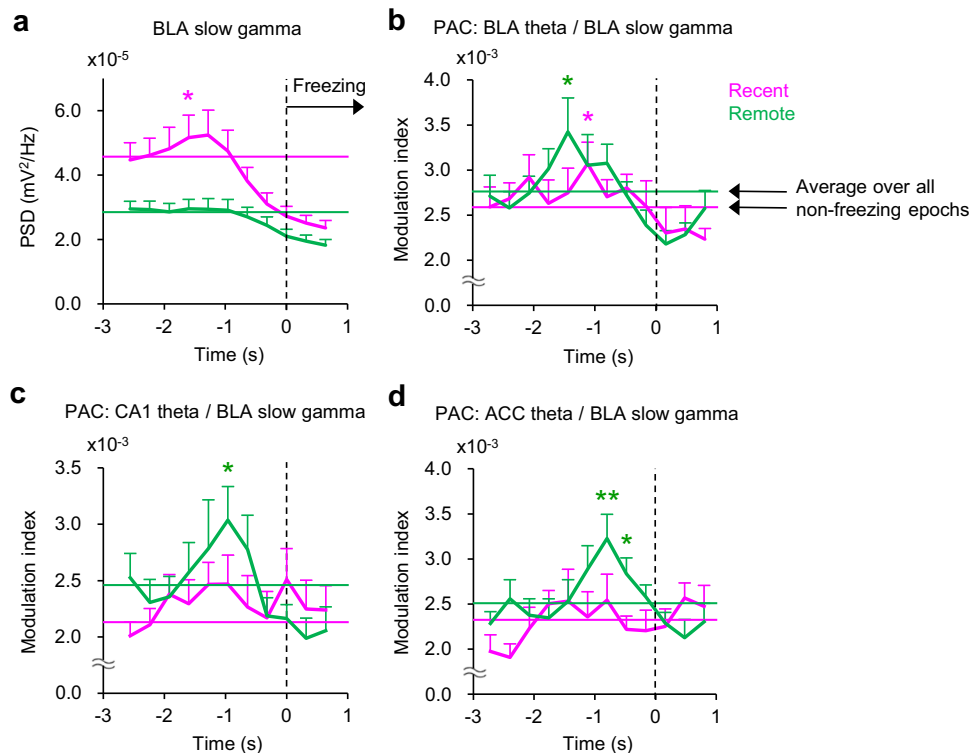


Fig. 3 | BLA slow gamma activity and its multi-regional theta entrainment differentially predict the onset of freezing for recent and remote recall. **a** Slow gamma PSD in the BLA was calculated at the segments immediately before and after the onset of freezing. Slow gamma PSD averaged over all non-freezing epochs for each of recent and remote sessions is shown by horizontal lines. **b–d** PAC of BLA slow gamma by theta of BLA, CA1 and ACC around the onset of freezing. In (**a–d**),

$n = 12$ mice. For all bar graphs, data are represented as mean \pm SEM. Time points with the PSD or modulation index significantly larger than that averaged over all non-freezing epochs at each session are indicated as * $p < 0.05$ and ** $p < 0.01$, which was tested by one-tailed paired t -test at each time point. Source data are available in the Source Data file.

Also, the model was able to perform classification when LFP during freezing periods were used at the F1 score of 0.664 (Fig. 4b), which was slightly lower than that when non-freezing periods were used, but still significantly better than the shuffled control.

The classification performance of feature-based models could be biased by the arbitrary selection of input features by researchers, which might cause technical issues such as overfitting and mask the true predictive power of other features that were not selected. To resolve this, we next employed a feature-free approach, using a large-scale Transformer to build a model that can classify memory age from raw LFP data from the three regions. In our Transformer-based approach (Fig. 4c), we concentrated on the role of the self-attention mechanism in processing multi-channel LFP data. The data were segmented and encoded with position embeddings, with a classification vector (CLS) introduced to aggregate information across the sequence. The core lies in the self-attention mechanism, where the CLS token dynamically focused on and integrated relevant segments through calculated attention weights, which were represented as $[a_{1,2} \dots a_{1,L}]$ (shown in red in Fig. 4c) and then projected back to the input raw LFP for the attention-guided feature analysis later. Self-attention enabled our model to extract essential global features for accurately predicting memory age by emphasizing the segments most pertinent to the classification task. As with the LightGBM approach, we employed a between-subject cross-validation strategy to train, validate, and test the model.

The Transformer model of the largest size ($d_t = 5$) achieved an average F1 score of 0.780 across the 12 mice in the test set when non-freezing segments were used, which was significantly higher than the label-shuffled control (Fig. 4d, e). In contrast, the average F1 score dropped to 0.645 when freezing segments were used, which was higher than the label-shuffled control, but lower than the score

obtained from non-freezing segments (Fig. 4d, e). To assess the impact of Transformer size on classification, we compared performance of Transformers with three different sizes: normal ($d_t = 3$), large ($d_t = 4$), and xlarge ($d_t = 5$). Performance using non-freezing segments was comparable across different Transformer sizes (Fig. 4d, f), implying saturated performance even at the normal size. In contrast, performance using freezing segments was improved when larger Transformer sizes were used, showing the benefit of progressively extracting higher-dimensional features, however it was still lower than the performance with non-freezing segments at all Transformer sizes (Fig. 4d, f). These results demonstrate in a non-biased manner that LFP during non-freezing periods contains more memory age-related information than that during freezing periods. Taken together, both LightGBM and Transformer models were capable of predicting memory age at high accuracy, revealing the robustness of distinct multi-regional signatures associated with memory age.

Computational models predict memory age based on multi-regional oscillatory activity

As oscillatory activity in the slow gamma and theta ranges as well as theta-gamma PAC were differentially regulated at recent and remote timepoints in our study, we asked whether LightGBM and/or Transformer have built models that extract information related to these LFP activity patterns. To first decipher the decision-making features of LightGBM, we analyzed the split feature importance, which determines the importance of each feature based on the number of times it is used to split the input feature vectors during the construction of the tree-based classifier. One common issue when analyzing the split feature importance is the noise baseline (null importance), which is a feature importance value obtained even

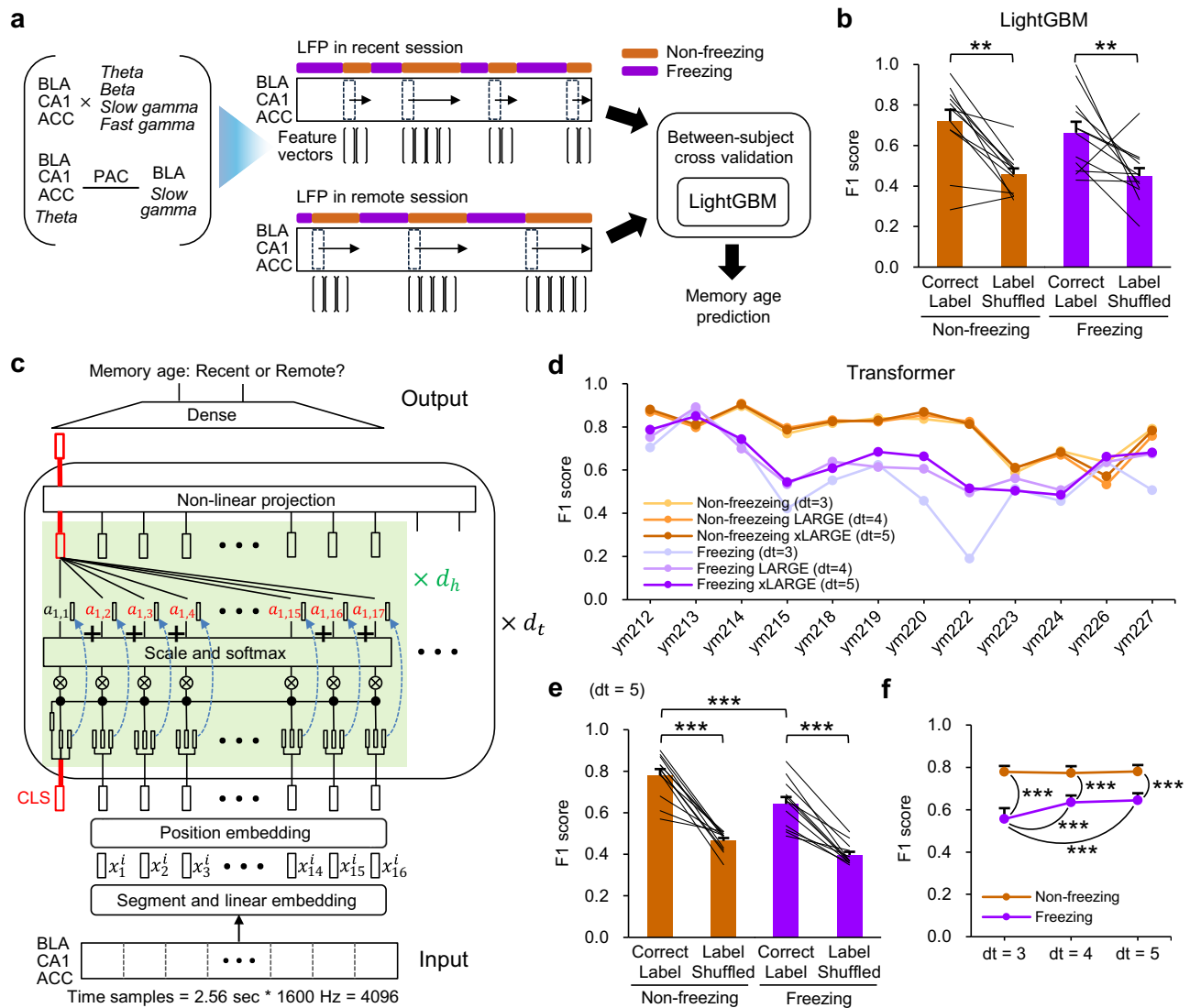


Fig. 4 | Feature-based LightGBM and feature-free Transformer both accurately classify recent from remote recall based on multi-regional LFP. **a** The pipeline diagram of the feature-based LightGBM to predict memory age (recent vs. remote), using feature vectors that consisted of mean PSD at multiple frequency ranges and regions as well as theta to slow gamma PAC at multiple regional pairs. **b** The F1 score of the test set for LightGBM to predict memory age for correctly labeled (actual performance) data and control data with randomly shuffled labels (main effect for behavioral state, $F_{1,11} = 0.61$, $p = 0.453$; main effect for label, $F_{1,11} = 33.10$, $p = 1 \times 10^{-4}$; interaction, $F_{1,11} = 0.25$, $p = 0.624$; correct vs shuffled at non-freezing, $p = 0.003$; correct vs shuffled at freezing, $p = 0.009$). **c** The schematic of the Transformer architecture with segmentation, linear embedding, positional embedding, Transformer blocks, and dense layers. It is important to note that the red path of the CLS token was utilized to aggregate global information from the LFPs. Additionally, the attention scores $[a_{1,2} \dots a_{1,17}]$ (shown in red) were employed for further analysis of feature importance. **d** The F1 score of the test set for individual mice when different sizes of Transformer were used. **e** The F1 score of the

test set for Transformer ($d_t = 5$) averaged across mice (main effect for behavioral state, $F_{1,11} = 27.86$, $p = 3 \times 10^{-4}$; main effect for label, $F_{1,11} = 115.81$, $p = 4 \times 10^{-7}$; interaction, $F_{1,11} = 2.68$, $p = 0.130$; correct vs shuffled at non-freezing, $p = 4 \times 10^{-7}$; correct vs shuffled at freezing, $p = 5 \times 10^{-6}$; non-freezing vs freezing at correct label, $p = 9 \times 10^{-4}$; non-freezing vs freezing at label shuffled, $p = 0.050$). **f** Comparison of the test F1 scores when different sizes of Transformers ($d_t = 3, 4$, or 5) were used (main effect for Transformer size, $F_{2,22} = 6.28$, $p = 0.007$; main effect for behavioral state, $F_{1,11} = 16.98$, $p = 0.002$; interaction, $F_{2,22} = 8.59$, $p = 0.002$; $d_t = 3$ vs 4 at non-freezing, $p = 0.930$; $d_t = 3$ vs 5 at non-freezing, $p = 0.997$; $d_t = 4$ vs 5 at non-freezing, $p = 0.900$; $d_t = 3$ vs 4 at freezing, $p = 3 \times 10^{-4}$; $d_t = 3$ vs 5 at freezing, $p = 9 \times 10^{-5}$; $d_t = 4$ vs 5 at freezing, $p = 0.847$; non-freezing vs freezing at $d_t = 3$, $p = 6 \times 10^{-12}$; non-freezing vs freezing at $d_t = 4$, $p = 4 \times 10^{-8}$; non-freezing vs freezing at $d_t = 5$, $p = 5 \times 10^{-8}$). In **(b, e, f)**, $n = 12$ mice, two-way repeated measures ANOVA with post hoc comparisons using Tukey's HSD test. For all bar graphs, data are represented as mean \pm SEM with values from individual mice. $**p < 0.01$ and $***p < 0.001$. Source data are available in the Source Data file.

following class label-shuffling of the training data. To circumvent this issue, we performed null importance analysis²⁶, where we considered a feature having significant contribution only if the importance value exceeds 99.7% confidence interval of null importance distribution, which was obtained by separately training the model with labels shuffled 1000 times. Among the features which made significant contributions for the model using non-freezing epochs, we found that CA1 theta power and BLA slow gamma power together accounted for 70% of the importance (Fig. 5a). In addition, PAC of

BLA slow gamma by both CA1 theta and BLA theta significantly contributed to the model (Fig. 5a), largely consistent with the differential roles of these oscillatory activities and couplings for recent and remote recall (Figs. 1f, 2c, f and Supplementary Fig. 5d). Significantly contributing features remained largely unchanged when locomotion was included as an additional feature, with locomotion itself not contributing significantly (Supplementary Fig. 5b), indicating that locomotion would not be informative to dissociate recent from remote recall.

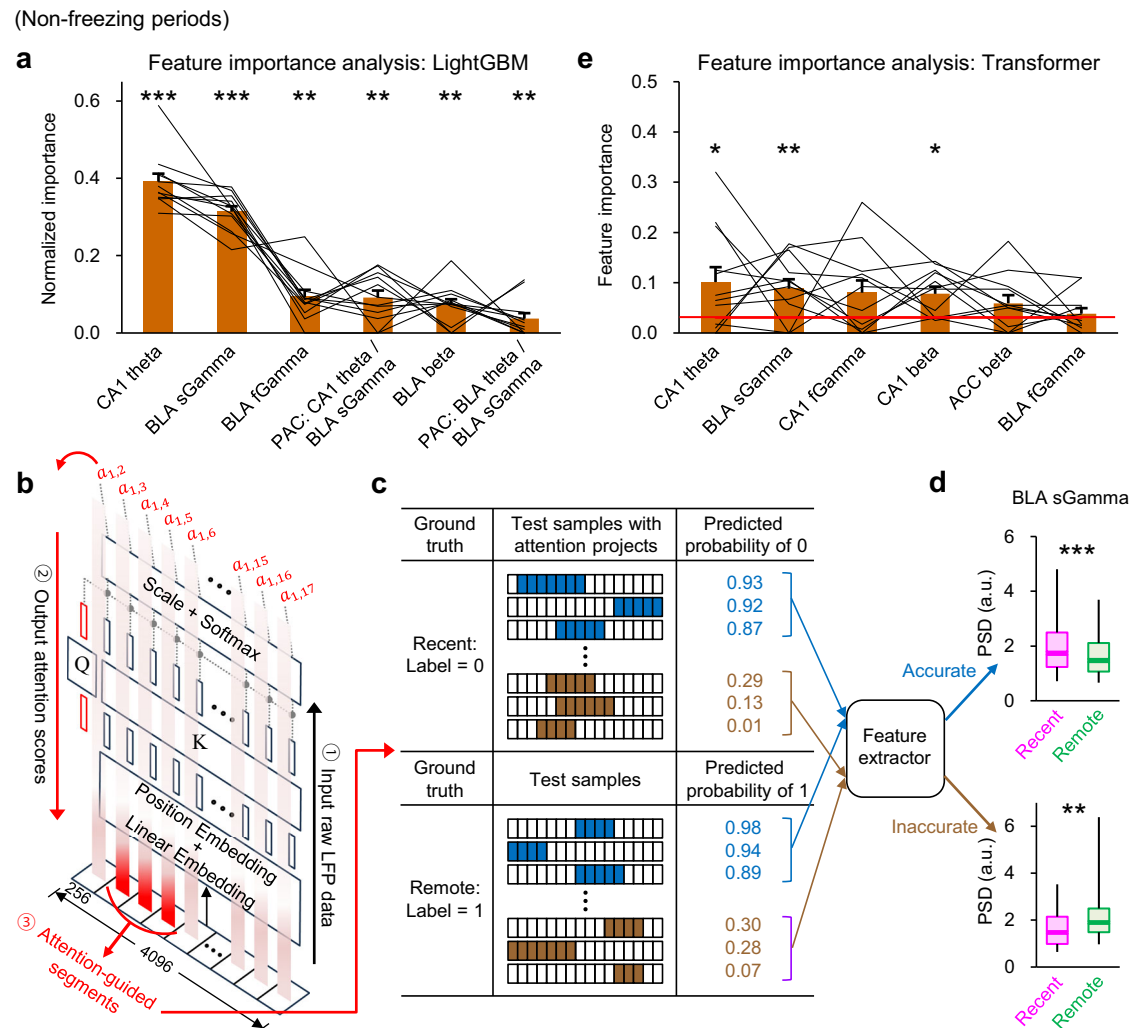


Fig. 5 | Multi-regional oscillatory activity serves as key features utilized to predict memory age by both LightGBM and Transformer. **a** The actual feature importance against the null distribution determined by LightGBM which used LFP during non-freezing periods as the input. Only the features that made significant contributions, defined as having an actual feature importance >0 (see Methods for details), are shown (CA1 theta, $W=78.0$, $p=2 \times 10^{-4}$; BLA slow gamma [sGamma], $W=78.0$, $p=2 \times 10^{-4}$; BLA fast gamma [fGamma], $W=66.0$, $p=0.002$; PAC from CA1 theta to BLA slow gamma, $W=55.0$, $p=0.003$; BLA beta, $W=55.0$, $p=0.003$; PAC from BLA theta to BLA slow gamma, $W=36.0$, $p=0.006$). **b** The schematic of the process of the attention-guided projection back to the input space. Initially, raw input LFP was fed into the Transformer block. Subsequently, attention scores were calculated using the CLS token as the query. Finally, these attention values were projected back into the input space. **c** The attention-guided feature extraction process for accurately (in blue) and inaccurately (in brown) predicted samples. For both the accurate and inaccurate cases, the theory-driven features were extracted from samples of recent and remote classes. **d** The disparity between recent-vs-remote difference of the extracted feature in accurately predicted cases and that in inaccurately predicted cases gave rise to the Attention-guided Feature Importance (AFI; see Methods for detail). In this example, within the segments guided by

attention in accurately predicted samples, PSD of BLA slow gamma was significantly higher in the recent session compared to the remote session ($p=3 \times 10^{-7}$); conversely, in inaccurately predicted samples, BLA slow gamma PSD was significantly higher in the remote session than in the recent session ($p=0.009$). These opposite recent-vs-remote differences correspond to the transition between accurate and inaccurate predictions, indicating that the Transformer utilized the BLA slow gamma feature to achieve accurate predictions. **e** Feature importance indicated by the AFI for the transformer model using LFP of non-freezing periods as the input. Only the features with the AFI higher than the threshold of null importance (indicated by the red line) are shown with statistical significance (CA1 theta, $W=66.0$, $p=0.017$; BLA slow gamma, $W=69.0$, $p=0.008$; CA1 fast gamma, $W=59.0$, $p=0.065$; CA1 beta, $W=67.0$, $p=0.013$; ACC beta, $W=56.0$, $p=0.102$; BLA fast gamma, $W=41.0$, $p=0.455$). In **(a, e)**, $n=12$ mice, one-tailed Wilcoxon signed-rank test. In **(d)**, two-tailed unpaired t -test. For all bar graphs, data are represented as mean \pm SEM with values from individual mice. For all box plots, the middle, bottom, and top lines correspond to the median, lower and upper quartiles, and the edges of lower and upper whiskers correspond to the 5th and 95th percentiles, respectively. * $p < 0.05$, ** $p < 0.01$ and *** $p < 0.001$. Source data are available in the Source Data file.

To probe what kind of information the Transformer model was extracting from raw LFP, we adopted a strategy that combines attention-guided input selection with theory-driven feature extraction, which consisted of the following three steps. First, we utilized the attention weights in the first layer of the trained model (illustrated as $[a_{1,2} \dots a_{1,17}]$ to highlight LFP patches) to identify potentially high-contributing segments of the data (Fig. 5b). Second, we performed theory-driven feature extraction on the Transformer-focused patches

in accurately or inaccurately predicted samples (Fig. 5c). Finally, by comparing recent-vs-remote differences of the feature between accurate and inaccurate samples (Fig. 5d), we calculated the Attention-guided Feature Importance (AFI) index which assessed the potential contribution of each feature to the model's decision-making (see Methods for details). This approach revealed multiple putative oscillatory features that made significant contributions for the model using non-freezing periods, including CA1 theta and BLA slow gamma

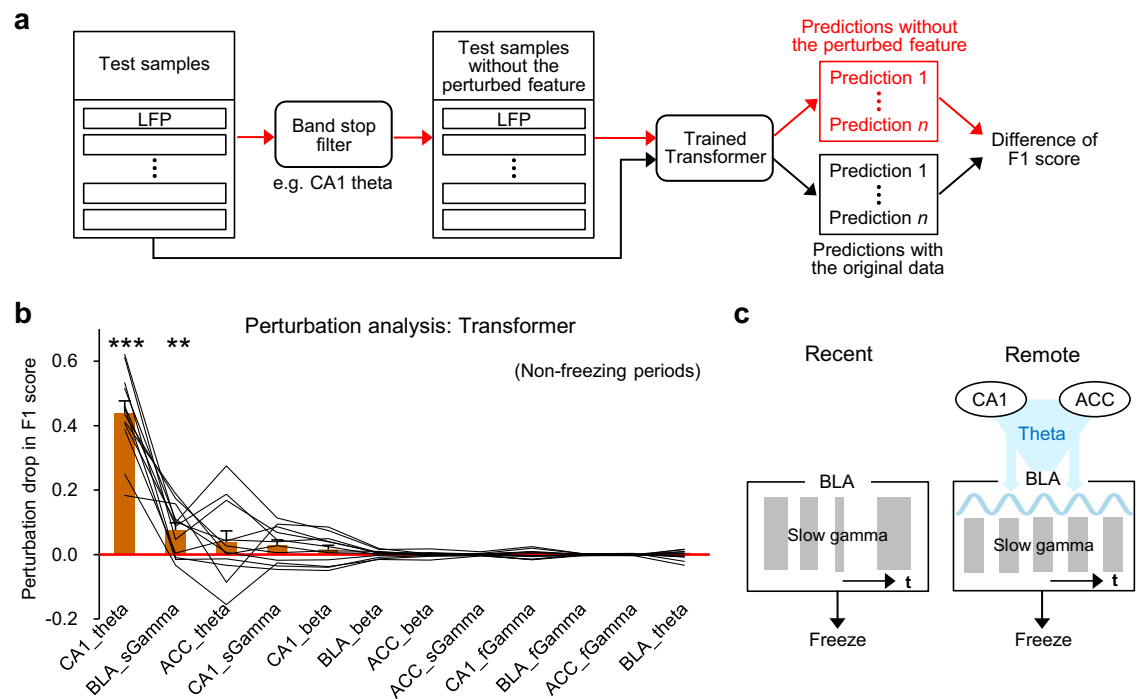


Fig. 6 | Transformer relies on CA1 theta and BLA slow gamma to predict memory age. **a** A pipeline diagram of the perturbation analysis process for the Transformer model. The red lines represent the perturbation workflow, in which a specific feature (e.g., CA1 theta) in the test samples was perturbed using a band stop filter and then input into the trained Transformer model to obtain memory age predictions without the influence of the perturbed feature. The black lines depict the workflow for testing original data, where test samples were directly fed into the trained model to obtain the actual test performance. The difference in F1 scores between these two workflows was employed to quantify the importance of the

perturbed feature. **b** Results of the perturbation analysis for the transformer model using LFP of non-freezing periods as the input. CA1 theta and BLA slow gamma exhibited significantly higher importance compared to the baseline, which is denoted by a red line ($n=12$ mice; CA1 theta, $W=78.0$, $p=2 \times 10^{-4}$; BLA slow gamma, $W=69.0$, $p=0.008$; one-tailed Wilcoxon signed-rank test). **c** Model for the modulation of BLA slow gamma associated with recent and remote recall. See Discussion for detail. For all bar graphs, data are represented as mean \pm SEM with values from individual mice. $**p < 0.01$ and $***p < 0.001$. Source data are available in the Source Data file.

(Fig. 5e), again consistent with the feature importance of LightGBM models.

To further validate significance of each putative important feature for the Transformer model, we performed perturbation analysis, in which degradation of classification performance was assessed when each oscillatory feature in the raw data was filtered out individually (Fig. 6a). These analyses demonstrated a significant drop of F1 score when CA1 theta or BLA slow gamma was perturbed (Fig. 6b), confirming that the Transformer model relied on these oscillatory activities for accurate classification. Taken together, BLA slow gamma activity and its presumable entrainment by CA1 theta serve as reliable physiological signatures dissociating remote from recent memory recall, which was consistently utilized by feature-based LightGBM and feature-free Transformer modeling.

Discussion

Despite the well-recognized role of the BLA in fear memory, its precise neural activity dynamics associated with memory age have been unexplored. Here we find that recall of recent fear memory is reflected in augmented gamma activity in the BLA, while remote recall is associated with PAC between BLA gamma amplitude and theta phase in multiple distal regions, including CA1 and ACC (Fig. 6c). Both LightGBM and Transformer modeling indicated BLA gamma and CA1 theta as most informative features to differentiate remote from recent memory, suggesting that these multi-regional oscillatory activities are key signatures for fear memory age extracted in a non-biased manner.

Previous studies have reported that the BLA is required for both recent and remote fear memory^{2,3}, suggesting its memory age-independent role. Here, we have examined the detailed activity

dynamics of BLA associated with different memory ages and found that intensity and phase control of BLA gamma are differentially recruited by recent and remote recall, suggesting distinct fine-scale activity patterns in the BLA that characterize different memory ages. Further investigation is needed to understand how higher-intensity and more phase-controlled gamma equally lead to the expression of fear memory. It has been proposed that temporally coordinated neural activity could activate target neurons more effectively by being aligned to their excitability and coincidence detection windows^{27,28}. It is possible that phase-controlled gamma in the BLA could provide efficient signaling to its downstream regions required for fear expression, such as central amygdala²⁹, potentially attaining comparable outcomes to that of high intensity gamma.

We have further shown that there is significant modulation of BLA gamma during remote recall by theta activity in both CA1 and ACC, as well as in the BLA itself. While it was traditionally thought that the hippocampus is not involved in remote memory^{30,31}, there is now accumulating evidence that it is required for both recent and remote memory particularly for retaining episodic details^{6,8,32–34}, supporting the active role of CA1 for remote recall in our study. It has also been shown that theta synchrony between higher-order auditory cortex and BLA is enhanced during recall of remote auditory fear memory³⁵, suggesting that cross-regional theta entrainment of BLA activity might underlie recall of various forms of remote memory. In addition, an enhanced correlation between CA1-ACC theta synchrony and BLA theta was observed during remote recall in our study, suggesting that the three regions may work together via theta-phase-entrainment of BLA gamma at remote time points leading to contextual fear expression and freezing.

Importantly, the increase of BLA gamma at recent recall and the enhancement of its theta phase control at remote recall was observed particularly before the onset of freezing. These dynamics suggest that intensity and phase control of BLA gamma are involved in the processes aiding recall and/or freezing at each memory age, consistent with the reported role of BLA gamma in threat assessment and response selection³⁶. Interestingly, these LFP indices as well as theta power in the BLA, CA1 and ACC were not significantly correlated with the total duration of freezing at the level of individual mice (Supplementary Fig. 6). We speculate that BLA gamma activity and its phase control support the initiation of fear memory recall or freezing, but the subsequent maintenance of sustained freezing might be regulated by separate neural mechanisms involving downstream regions, such as central amygdala and brainstem nuclei²⁹.

Here we explicitly designed a within-subject longitudinal study, recording LFP during recent and remote recall in the same cohort of mice. This approach has been repeatedly employed in the field^{14,21,37,38}, as the LFP properties we characterized, such as theta and gamma power as well as theta-gamma PAC, can exhibit high individual variability (Fig. 1f, g and Fig. 2c, f, i) due to differences such as precise electrode location and quality^{12,13}. As a result, comparing the physiology of recent and remote recall in different cohorts (between-subjects) may result in drastically reduced statistical power. On the other hand, in our within-subject approach we cannot rule out the possibility that the changes of LFP patterns observed at remote recall might be due to repeated contextual exposure, independent of memory age. However, previous similar recordings both in the hippocampus and ACC²¹, as well as in the amygdala^{39,40}, suggest this is unlikely, as cross-regional interactions as well as single-band power remain unchanged over multiple test sessions spaced closer together in time. Further, most of the LFP patterns examined in our study demonstrate differential changes from baseline to recent recall sessions and from recent recall to remote recall sessions (Fig. 1f, Fig. 2f, i and Supplementary Fig. 2), and such non-monotonic changes have not been observed for theta power, gamma power or its cross-regional synchrony in and across the hippocampus, medial prefrontal cortex and amygdala with mere repetitive exposure to memory-related context or cue, particularly while the strength of memory and the animals' behavior remains constant^{21,39–41}. Nevertheless, while we believe the changes in BLA gamma, as well as its PAC by multiple regions, from recent to remote recall sessions in our study most likely reflect the difference of memory age, we acknowledge that the effect of repeated contexture exposure on LFP of various brain regions still needs to be rigorously examined in future studies.

Finally, we asked if these oscillatory activities in multiple regions could be used to classify memory age by LightGBM and Transformer modeling. Both models demonstrated remarkably high accuracy in predicting memory age, suggesting their broad applicability in classification based on neural/electrophysiological data. Notably, the Transformer model could classify memory age at higher accuracy when data from non-freezing periods were used as the input compared to when data from freezing periods were used, demonstrating that non-freezing periods might contain more information associated with different memory ages. This is consistent with our biological data that the majority of activity differences between recent and remote recall within BLA and across regions was observed only during non-freezing periods, particularly at the time shortly before the onset of freezing. Together, these observations suggest that non-freezing periods play an active role in recall processes presumably through continuous exposure to recall cues in the fear-conditioned context, which does not occur during stationary, freezing periods.

While a universal theory or benchmarks for Transformer interpretation are still in development, we have also made an exploratory attempt to analyze the features that could be extracted by the models

using attention-based interpretation plus perturbation analysis (Figs. 5b–e, 6a, b). This approach has identified oscillatory activity in multiple regions as important features, consistent with the results of the feature importance analysis of LightGBM and highlighting the potential of attention-based methods in enhancing model interpretability. It is noteworthy that in perturbation analysis, the significance of CA1 theta far exceeded that of BLA slow gamma. This could be because the perturbation of CA1 theta, which was a major component of signal power, caused distortion of the signal, rendering any classification task challenging even when the signal contained useful features, including BLA slow gamma. Additionally, the high degree of correlation between CA1 theta and ACC theta activities¹⁴ could have rendered the utilization of just one of these signals sufficient for classification tasks, which resulted in the lower significance of ACC theta in the analysis. Although perturbation analysis involves certain challenges, importantly it enriched the decision-making process of the model by providing a double confirmation of the decision-making features discovered by the attention-based interpretation.

Overall, this study provides a more comprehensive understanding of reliable neural activity patterns in the BLA and its associated distal regions which characterize the age of fear memory. Multi-regional oscillatory dynamics have been identified as the primary features differentiating recent from remote memory, particularly by a non-feature-biased Transformer model, revealing robust and stereotypical codes of memory age dispersed across brain regions.

Methods

Experimental animals

Adult male wild-type C57BL/6J mice (Japan SLC, Inc.; C57BL/6JMSlc substrain), bred in-house, were used for all experiments. Mice were in the range of 3–7 months old over the course of the experiments. Mice were maintained on a 12 h light/dark cycle (light cycle: 8 AM–8 PM) at 22–24 °C and 30–50% humidity, provided with food and water ad libitum. All experiments were performed at the light phase. All procedures were approved by the RIKEN Institutional Animal Care and Use Committee and complied with all relevant ethical regulations.

Contextual fear conditioning apparatus

Baseline, training and recall sessions of contextual fear conditioning were all conducted in a brightly lit behavioral training room. The chamber, consisting of an acrylic plastic front and back with aluminum walls on each side, measured 30 × 25 × 21 cm (Med Associates ENV-008, Georgia, VT). The floor of the chamber consisted of 36 stainless steel rods of 3.2 mm diameter and spaced 7.9 mm apart and was connected via a cable harness to a shock generator. The chamber was cleaned between animals with 70% ethanol, and a solution of 1% acetic acid was placed beneath the chamber during the experiment to provide an olfactory cue. Freezing behavior was video recorded by the camera located behind the chamber at a frame rate of 3.75 Hz. Shock deliveries and video recordings were controlled by FreezeFrame 2 software (Actimetrics, Wilmette, IL).

Microdrive construction

Custom microdrives were manufactured with the assistance of the Advanced Manufacturing Team at RIKEN. The Microdrive was constructed on a 3D-printed plastic base which held an interface board (EIB-36; Neuralynx, Bozeman, MT) and nichrome (14 µm diameter) tetrodes that were independently adjustable along the z-axis. These tetrodes were positioned to the right BLA (center coordinate is AP −2.0, ML +3.3, DV −4.3), the right dorsal CA1 (AP −2.0, ML +1.5, DV −1.3), the right ACC (AP +0.8, ML +0.3, DV −1.6) and the right ventral hippocampal commissure (vHC; AP −0.8, ML +0.7, DV −2.4) where the reference signal was recorded. All tetrodes were gold plated to an impedance of 200–300 kΩ prior to surgery.

Surgery

Mice were anesthetized using Avertin (2, 2, 2-tribromoethanol; Sigma-Aldrich, 476 mg kg⁻¹, i.p) and placed into a stereotactic frame (Kopf). The skull was leveled using bregma and lambda landmarks. Small craniotomies were made individually over the right BLA, CA1, ACC and vHC. Tetrodes were then inserted to their respective target regions simultaneously, ensuring that all tetrodes were targeted to the stereotaxic coordinates described above. In addition, a stainless steel screw was implanted above the left cerebellum to serve as a ground. The microdrive was fixed to the skull with dental cement (Super-Bond C&B, Sun Medical). Following recovery from anesthesia, mice were returned to their home cage where they were housed individually.

After surgery, tetrodes in CA1 were slightly adjusted into the pyramidal cell layer, which was identified by numerous large amplitude pyramidal cell spikes and sharp-wave ripples during periods of immobility. Tetrodes in BLA and ACC were adjusted only minimally (± 100 μ m). Tetrodes in the vHC were left unadjusted.

Experimental procedure

At 2–3 weeks after surgery, the contextual fear conditioning experiment was conducted. On Day 1, mice were placed in the fear conditioning chamber and their LFP was recorded for 10 min (baseline session). After 30 min of rest in the housing room, they were brought back to the fear conditioning chamber and fear-conditioned with a 2 s-long, 0.5 mA foot shock delivered three times at 1 min intervals.

On Day 2, mice were brought back to the fear conditioning chamber at the same time of the day as Day 1, and their recall of contextual fear memory was assessed for 10 min while their LFP was recorded (recent session). On Day 26–28, the protocol used on Day 2 was repeated and fear recall was assessed again with LFP recordings (remote session).

Histology

At the conclusion of the experiment mice underwent terminal anesthesia and tetrode positions were marked by electrolytic lesioning of brain tissue (30 μ A current for 7 s through each electrode individually). Brains were fixed with cardiac perfusion of 4% paraformaldehyde (PFA) followed by post-fix in PFA for 24 h and coronal slices with 50 μ m thickness were prepared on a cryostat. Tetrode placement was confirmed by standard light microscopy.

Data acquisition

Recordings were obtained via a unity gain headstage preamplifier (HS-36; Neuralynx) connected to a flexible tether cable (TETH-HS-36-Flex-EXT; Neuralynx) suspended from the ceiling with a rubber band, which allowed mice to freely move their head up and down with minimal force. Data were acquired using a 32-channel Digital Lynx 4SX acquisition system with Cheetah software (v 5.7.4; Neuralynx). LFP signals were sampled at 32,000 Hz and filtered between 1–9000 Hz. After recordings, LFP was down-sampled to 1600 Hz for subsequent analyses.

Behavioral analysis

Non-freezing and freezing periods were determined by manually examining all video frames over the entire recording. The periods of each state lasting <1.28 s were excluded from analysis. More detailed behavioral categorization and calculation of animal velocity were performed by tracking each body point of the mice using the DeepLabCut algorithm⁴², as described previously¹⁴.

Power (PSD) analysis

A tetrode that was located closest to the center of each of BLA, CA1 and ACC was selected for each animal based on histology after the experiment and the same tetrodes were used across all sessions. Timestamps of non-freezing and freezing periods were aligned with

those of LFP data by referencing them to the TTL pulse sent from FreezeFrame to Cheetah Software (Neuralynx) at the beginning of the experiment. Power spectral density (PSD) from non-freezing and freezing periods was calculated with the `pmtm` function in MATLAB (Mathworks) using 1.28 s epochs. Mean PSD across frequency ranges of 2–6 Hz, 6–12 Hz (theta), 30–50 Hz (slow gamma) and 60–90 Hz (fast gamma) were then calculated. For the time course of PSD (Fig. 1d), the `pwelch` function in MATLAB was alternatively used with 0.64 s epochs that overlapped by 0.32 s.

To analyze the dynamics of PSD around the onset of freezing (Fig. 3), the episodes that include at least 3.2 s of non-freezing periods followed by at least 1.28 s of freezing periods were extracted. Within these 4.48 s segments, PSD was calculated every 0.32 s and averaged over all segments from each animal for each time point.

Cross-frequency coupling

Cross-frequency phase-amplitude coupling (PAC) of LFP oscillation within and between regions was analyzed as previously described²⁰. First, raw LFP from one region was filtered at either 2–6 Hz or theta (6–12 Hz) frequency, and the phase of the filtered LFP was computed using the Hilbert transform. Raw LFP from the same or different region was also filtered at slow gamma (30–50 Hz) frequency, and the amplitude envelope was extracted with the Hilbert transform. To examine the coupling strength of slow gamma amplitude to 2–6 Hz or theta phase, the phase values were binned into 36 bins (0–360°, 10° intervals) and the mean amplitude of slow gamma in each phase bin during the state being analyzed was calculated. These values were then used to calculate the modulation index²⁰, which indicates the strength of PAC of slow gamma amplitude by 2–6 Hz or theta phase. The dynamics of PAC around the onset of freezing (Fig. 3) were analyzed in the same way as those of PSD as described above.

To generate a comodulogram between low-frequency phase and high-frequency amplitude, the phase of the low-frequency oscillation was extracted between 4 and 18 Hz with bandpass filters of 4 Hz width (stepping by 0.5 Hz) and the amplitude of high-frequency oscillation was calculated between 25 and 135 Hz with bandpass filters of 10 Hz width (stepping by 2 Hz). Modulation index for each phase-amplitude pair was plotted as a heat map.

Correlation of LFP instantaneous amplitude

Relationship of the instantaneous amplitude of LFP oscillations at the frequency range of interest across three regions was analyzed as follows. First, raw LFP from the three regions was filtered at the target frequency range using a zero-phase-delay filter, and the amplitude envelope was extracted with the Hilbert transform. The correlation coefficients between envelopes of two of the three regions over 4 s epochs within the state being analyzed (non-freezing / freezing) was then calculated with 2 s overlaps as previously described²². Then, the correlation of these values with mean envelope of the third region within the same epochs was calculated. Mice that contained <10 epochs were excluded from analysis for each state (occurred only for freezing periods).

Computational modeling by LightGBM

Input format. For both non-freezing and freezing periods, segments >5.12 s were selected for analysis. Within each of these 5.12 s segments, a sliding window of 2.56 s with a step of 0.5 s was used to obtain a series of 2.56 s LFP segments. From each of these 2.56 s segments (segment *i*), different combinations of features were extracted as shown in each figure to form the feature vector x_i ; these features could include the PSD of theta (6–12 Hz), beta (20–30 Hz), slow gamma (30–50 Hz) and fast gamma (60–90 Hz) from BLA, CA1 and ACC, phase-amplitude coupling (PAC) of slow gamma amplitude in BLA by theta phase in BLA, CA1 and ACC, and mean velocity of the mouse. For the classes of recent and remote sessions, the feature vectors from 12 mice totaled

approximately 10,000. These feature vectors were used for training, validating, and testing the LightGBM based on between-subject cross validation.

Model architecture. As a gradient boosting framework that uses tree-based learning algorithms, LightGBM is currently the most widely used feature-based classifier in data science. It is a Gradient Boosting Decision Tree (GBDT) embedded with Gradient-based One-Side Sampling (GOSS) and Exclusive Feature Bundling (EFB), which facilitates efficient training without compromising accuracy. The loss function at iteration t can be described below.

$$L^{(t)} = \sum_{i=1}^n L(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)) + \Omega(f_t) \quad (1)$$

Where y_i is the true label of the feature vector i ; $\hat{y}_i^{(t-1)}$ is the prediction at iteration $t - 1$; $f_t(x_i)$ is the t th added tree with x_i as the input feature vector with frequency features from three regions; $\Omega(f_t)$ is the regularization term for tree f_t which prevents overfitting and improves generalization; n is the number of samples in the training data. The LightGBM loss was defined in a boosting way as the error was reduced iteratively as new trees added to the forest, which means each added new tree focused on the samples that were remaining misclassified.

Model hyperparameters. The model's hyperparameters were chosen to optimize performance. Initially, for the cross-validation phase, the model was configured with an objective of cross entropy (xentropy) and utilized the gradient boosting decision tree (GBDT) boosting. The learning rate was set to 0.01, with the number of leaves and the maximum depth established at 31 and 6, respectively. We employed 1000 estimators and enabled early stopping after 100 rounds without improvement on the validation set, to prevent overfitting and enhance model generalization. Subsequently, for the inference phase, the model adopted a similar configuration but adjusted the maximum depth to 8 and set the number of estimators based on the average best iteration determined during the cross-validation process.

Feature importance analysis. To assess the importance of each feature, split feature importance was analyzed, which measures the number of times a feature is used in a tree-based boosting model for splitting the input. One disadvantage of using raw importance value in LightGBM is the noise baseline derived from the training set. Specifically, under the influence of noise, a model might assign an importance to a feature that is not significant, termed noise feature importance, which should not be added to the actual importance. Only those importances significantly greater than the noise level should be considered as genuine contribution to the prediction. To eliminate such noise, null importance analysis was performed to assess the genuine contribution of features in decision models against noise²⁶ in the following way. First, raw feature importance was calculated. Next, labels of the training data were shuffled 1000 times and models were separately trained with those labels, which generated the distribution of null (shuffled) feature importance. Finally, the raw feature importance was compared with the null distribution. If the raw importance fell above the 99.7% confidence interval of the null distribution (greater than the mean μ of the null distribution plus three times the standard deviation σ), the feature's contribution was considered to be significantly greater than noise. *Raw importance* $-\mu - 3\sigma$ was used as the final feature importance.

Cross-validation scheme. Cross-validation (CV) utilizes mutually exclusive validation sets and an independent test set to objectively evaluate a model's performance, ensuring the model captures core, generalizable features of the data rather than biased abnormalities that lead to overfitting. In our between-subject CV scheme, one out

of 12 mice was first selected to serve as an independent test set, which did not participate in the training or validation phases. The remaining 11 mice were randomly divided into 5 folds, each containing either two or three mice. During this 5-fold CV phase, for each fold, the mice belonging to that fold were used as the validation set and the rest as the training set. Once the model was trained with the early stopping on the validation set's cross entropy, the validation set was input to obtain the final validation performance for that fold. A total of 5 stable validation performances were obtained for the 5 folds. A final model was then trained using data from all 11 mice across the 5 folds, maintaining the training parameters constant and using the mean of the early stop rounds obtained in all folds as the number of training rounds. The test set was input into the trained final model to obtain test performance.

It is important to note that a specific allocation of mice to train, validation, and test sets could potentially introduce allocation bias. To eliminate this bias, the CV scheme was executed 12 times, selecting a different mouse as the test set each time and randomly shuffling the remaining 11 mice for CV. Thus, 12 validation-test performance pairs were obtained in total. The test performance and feature importance for the 12 mice were used for further analysis.

Performance metric. The model's classification performance was assessed by the F1 score, which balances the classification effectiveness in the presence of an unbalanced class distribution. The F1 score is the harmonic mean of precision and recall:

$$F1 = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \quad (2)$$

where precision represents the proportion of true positives among all samples predicted as positive, and recall represents the proportion of true positives that were correctly predicted out of all actual positive cases:

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}} \quad (3)$$

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}} \quad (4)$$

Computational modeling by Transformer

Input format. One advantage of deep learning models is their ability to perform data mining directly from raw data. An expectation of this study is that deep learning models could spontaneously and unbiasedly identify features significantly representative of memory age. Therefore, we used the 2.56 s raw LFP segments, obtained as in LightGBM, as direct inputs to the model. Each input had 4096 time sampling points (2.56 s \times 1600 Hz) with spatial resolution across three brain regions (BLA, CA1 and ACC), thus the input data shape was 3×4096 (brain regions \times time sample points).

Model architecture. The implementation of our model was inspired by the Vision Transformer (ViT)⁴³, initially utilized in image recognition. Innovatively, ViT segments each image into small patches, reorganizing them into a sequence of vectors, aligning with the standard input format for a Transformer. In our case, the input to the Transformer was LFP at three brain regions with a time sample of 4096 (shape = 3×4096). We segmented this input LFP into 16 patches along the temporal dimension, each having a shape of 3×256 . Each patch was transposed and its three columns were then stacked to make a vector with a shape of 768×1 . All 16 vectors are fed into a projection layer to construct a sequence of vectors, described as x_j^i ($j=1 \sim 16$), where i and j were sample and patch index, respectively. The vector was embedded with position information which represented the order

of the patches. Then, a CLS token was added in front of the sequence. The CLS token is a special vector that aggregated information from all the patches over the course of multiple self-attention operations later. Then a sequence, i.e., $X^i = [\text{CLS}, x_1^i, x_2^i, \dots, x_{16}^i]$, was constructed for the Transformer block.

The Transformer block incorporated a multi-head self-attention mechanism, which was its core component. There were d_h heads in each Transformer block. In each head, the input sequence X^i was projected onto the Query, Key, and Value spaces, with the corresponding matrices Q_i , K_i , and V_i :

$$\text{Attention}(Q_i, K_i, V_i) = \text{Softmax}\left(\frac{Q_i K_i^T}{\sqrt{d_e}}\right) V_i \quad (5)$$

where d_e serves as a scaling factor. The outputs from d_h heads were then concatenated and fed into a non-linear layer to achieve the output of the Transformer block, which had the same shape of sequence as input. Then, d_t Transformer blocks were stacked to enable comprehensive feature discovery. Finally, the CLS vector output from the last block, accumulating the global representation of the entire sequence, was fed into a dense layer, followed by a softmax layer for memory age classification.

Model hyperparameters. The Transformer block was configured with 64-dimensional space to project the input vectors. It contained $d_h = 4$ attention heads in each block, with $d_t = 3, 4$, or 5 blocks stacked together. A layer normalization was performed after the blocks, followed by the final dense layers with a 64-dimensional intermediate layer and two classes as outputs.

The model was configured with a cross-entropy loss function for memory age classification. The learning rate was set to 0.001 with a weight decay of 0.0001. Early stopping was implemented with a patience value of 5, monitoring validation loss. Adam optimizer was employed to train the Transformer. The implementation is based on TorchEEG⁴⁴, an open-source deep learning library available at <https://github.com/torcheeg/torcheeg>.

Attention-guided feature importance analyses. Transformers may perform well without identifying key features in the data. Interpretability research for Transformers is categorized into attention-based⁴⁵, gradient-based⁴⁶, and relevance-based⁴⁷ methods. These approaches generate importance heatmaps by projecting calculated attentions, gradients, or relevance from the model back to the input, where higher weights signify greater decision impact. However, for the LFP data, merely projecting attention weights that signify importance onto LFP patches is insufficient because it does not reveal what theory-explainable features buried in the patches are truly highlighted. Therefore, we employed a strategy of attention-guided input selection with theory-driven extraction and perturbation, which consisted of three stages. First, we used the attention weights in the first layer of the trained model to highlight LFP patches, identifying potentially highly contributing segments (Fig. 5b). Second, we performed theory-driven feature extraction on Transformer-focused patches in both accurately and inaccurately predicted samples (Fig. 5c). Finally, we quantitatively compared the disparity between the effect size of a feature's recent-vs-remote statistical difference for accurate-predicted samples and that for inaccurate-predicted samples (Fig. 5d) by defining Attention-guided Feature Importance (AFI), thereby determining the potential feature contribution to the model. The detail of each stage is described below.

[Stage 1]. Attention weights corresponding to the CLS vector in the first block were used as guidance for interpretation (illustrated by bold in Eq. (6)). Our choice of the CLS-corresponding attention was due to the CLS being a classification backbone that aggregated information

from all patches. The selection of the first block was because the attention had a direct mapping relationship with the input data, where the weights directly indicated the degree of self-attention's selection of input patches. Mathematically,

$$\text{Attention Score}(Q_i, K_i, V_i) = \text{Softmax}\left(\frac{Q_i K_i^T}{\sqrt{d_e}}\right) = \begin{bmatrix} \mathbf{a}_{1,1} & \cdots & \mathbf{a}_{1,17} \\ \vdots & \ddots & \vdots \\ \mathbf{a}_{17,1} & \cdots & \mathbf{a}_{17,17} \end{bmatrix} \quad (6)$$

where each row represents one query, and each column represents one key. The attention weights of the keys (16 patches) that is queried by the first query (the CLS query), which is the first row of the attention matrix, were needed to be obtained. The first score $a_{1,1}$ corresponds to CLS, thus the attention weights $[a_{1,2} \cdots a_{1,17}]$ were projected to the input patches (Fig. 5b).

[Stage 2]. For each test sample, the model provided a memory age prediction with the attention projected onto the input LFP (Fig. 5b). First, samples were selected from the two categories (recent/remote) that were correctly predicted. If the probability of the ground truth category prediction was >0.55 , then it was regarded as an accurate prediction. For accurately predicted test samples (blue in Fig. 5c), the corresponding attention score was used to segment the LFP. These segmented LFPs, which have high attention, were fed into a theory-driven extractor, where the PSD of theta, beta, slow gamma, and fast gamma frequency bands in the BLA, CA1 and ACC was extracted. Each feature was grouped by its corresponding real label (recent or remote), from which statistical distributions were derived. For example, in the representative case shown in Fig. 5d, the PSD of BLA slow gamma in recent session is significantly higher than that in remote session in the high attention patches of the accurately predicted samples. The rationale is that if a feature shows a significant distribution difference between recent and remote sessions in the segments highlighted by the Transformer's attention, it was considered a putative feature potentially utilized by the attention mechanism. Furthermore, incorrectly predicted samples were selected for the same analysis with a prediction probability of <0.45 as the criterion (brown in Fig. 5c). Importantly, if the attention-guided feature in incorrectly predicted samples presents an opposite distribution to that seen in correctly predicted samples, it further indicates that the feature was potentially used for decision-making.

[Stage 3]. The Attention-guided Feature Importance (AFI) was defined to measure the significance of a feature in the decision-making process. Essentially, it assesses the disparity between the effect size of a feature's recent-vs-remote statistical difference for accurate-predicted samples and that for inaccurate-predicted samples. The greater the disparity in effect sizes for these two cases, the greater the potential importance of the feature. Effect size, specifically Cohen's d ⁴⁸, was used to measure the magnitude of statistical differences between recent and remote sessions:

$$d = \frac{M_{\text{recent}} - M_{\text{remote}}}{SD_{\text{pooled}}} \quad (7)$$

where M_{recent} and M_{remote} are the means of the two groups, SD_{pooled} is the pooled standard deviation of the two groups, as calculated below,

$$SD_{\text{pooled}} = \sqrt{\frac{(n_{\text{recent}} - 1)SD_{\text{recent}}^2 + (n_{\text{remote}} - 1)SD_{\text{remote}}^2}{n_{\text{recent}} + n_{\text{remote}} - 2}} \quad (8)$$

where n_{recent} and n_{remote} are the sample sizes, and SD_{recent} and SD_{remote} are the standard deviations of the two groups. The AFI was calculated

as follows: First, it was determined whether there was a significant difference of the feature between recent and remote in accurately predicted samples. If there was no significant difference, it indicated that the feature was not given significant attention, and thus the AFI was set to 0. If the difference was significant, then the effect size d_{accurate} was calculated. Next, it was examined whether there was a significant difference of the feature between recent and remote in inaccurately predicted cases. There were three possible scenarios: (1) If the difference was significant but the sign of the effect size was the same as in the accurate cases, it indicated that there was no disparity in the direction of the recent-vs-remote differences between the two cases, and therefore the AFI was set to 0; (2) If the difference was significant and the sign was opposite, then $\text{AFI} = |d_{\text{accurate}} - d_{\text{inaccurate}}|$; (3) If there was no significance, then $\text{AFI} = |d_{\text{accurate}}|$.

$$\text{AFI} = \begin{cases} 0 & p_{\text{accurate}} > 0.05 \text{ or } \text{sign}(d_{\text{accurate}}) = \text{sign}(d_{\text{inaccurate}}) \\ |d_{\text{accurate}} - d_{\text{inaccurate}}| & p_{\text{accurate}} < 0.05 \text{ and } p_{\text{inaccurate}} < 0.05 \text{ and } \text{sign}(d_{\text{accurate}}) \neq \text{sign}(d_{\text{inaccurate}}) \\ |d_{\text{accurate}}| & p_{\text{accurate}} < 0.05 \text{ and } p_{\text{inaccurate}} > 0.05 \end{cases} \quad (9)$$

Perturbation analysis. To further validate the importance of input features, we performed perturbation analysis which has widespread applications in the fields of deep learning-based imaging⁴⁹ and text⁵⁰. Importantly, perturbation analysis does not depend on the model's internal structure; it treats the model as a black box while manipulating input features. In this study, having identified putative features through the attention mechanism as described above, we validated the importance of these features by perturbing them in the raw LFP and observing changes in model prediction. A band-off filter was used to inhibit each putative feature identified by attention in the raw test LFP (Fig. 6a). Subsequently, the manipulated LFP without this feature was input into the trained model and the model's prediction F1 score was calculated. The importance of the feature to the model was determined by comparing the differences of model performance between the original LFP and the perturbation situation.

Cross-validation scheme. Five-fold cross-validation with exclusion of one out of 12 mice, followed by testing of the excluded mouse, were performed for the 12 validation-test pairs as described in the section for LightGBM.

Hardware. For our deep learning experiments, we leveraged an Intel Core i5-9400F CPU with 64 GB of memory, paired with an Nvidia GeForce RTX 2080 Ti GPU and an Nvidia A40 with 24 GB RAM.

Statistics

All statistical tests were performed either by MATLAB R2020a/R2022a (Mathworks), Python on Visual Studio Code 1.88.1 (Microsoft), Prism 10 (GraphPad) or Excel 2021 (Microsoft) with custom-written scripts. A two-tailed unpaired t-test was used to compare measures between two conditions. A two-tailed paired t-test or Wilcoxon signed-rank sum test was used to compare repeated measures between two conditions. A one-tailed paired t-test or Wilcoxon signed-rank sum test was used to test unidirectional changes of repeated measures between two conditions. A one-way repeated measures ANOVA was used to compare repeated measures across more than two conditions, followed by a post-hoc pairwise comparison with Tukey's HSD test. A two-way repeated measures ANOVA was used to compare repeated measures over two factors of conditions, followed by a post-hoc pairwise comparison with Tukey's HSD test. A test of no correlation (correlation t-test) was used for testing significance of correlation between different measures. All bar graphs and line plots show mean \pm SEM. All the statistically significant differences with $p < 0.05$ are described in each figure with different significance levels indicated by asterisks as $*p < 0.05$, $**p < 0.01$, $***p < 0.001$.

Reporting summary

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

Data availability

This study did not generate new unique reagents. Raw behavioral and electrophysiological recording data generated in this study are available at <https://neurodata.riken.jp/id/20241006-001> under the <https://doi.org/10.60178/cbs.20241006-001>⁵¹. Processed data shown in each figure are provided with this paper as a Source Data file. Source data are provided with this paper.

Code availability

The custom computer codes used in this study are available at https://github.com/BrainOdyssey2050/Makino-Wang-et-al-2024-Nat-Commun_Source-Code under the <https://doi.org/10.5281/zenodo.13934383>⁵².

References

1. Phelps, E. A. & LeDoux, J. E. Contributions of the amygdala to emotion processing: from animal models to human behavior. *Neuron* **48**, 175–187 (2005).
2. Maren, S., Aharonov, G. & Fanselow, M. S. Retrograde abolition of conditional fear after excitotoxic lesions in the basolateral amygdala of rats: absence of a temporal gradient. *Behav. Neurosci.* **110**, 718–726 (1996).
3. Gale, G. D. et al. Role of the basolateral amygdala in the storage of fear memories across the adult lifetime of rats. *J. Neurosci.* **24**, 3810–3815 (2004).
4. Kitamura, T. et al. Engrams and circuits crucial for systems consolidation of a memory. *Science* **356**, 73–78 (2017).
5. Liu, J., Totty, M. S., Melissari, L., Bayer, H. & Maren, S. Convergent coding of recent and remote fear memory in the basolateral amygdala. *Biol. Psychiatry* **91**, 832–840 (2022).
6. Moscovitch, M., Cabeza, R., Winocur, G. & Nadel, L. Episodic memory and beyond: the hippocampus and neocortex in transformation. *Annu. Rev. Psychol.* **67**, 105–134 (2016).
7. Goode, T. D., Tanaka, K. Z., Sahay, A. & McHugh, T. J. An integrated index: engrams, place cells, and hippocampal memory. *Neuron* **107**, 805–820 (2020).
8. Goshen, I. et al. Dynamics of retrieval strategies for remote memories. *Cell* **147**, 678–689 (2011).
9. Harris, A. Z. & Gordon, J. A. Long-range neural synchrony in behavior. *Annu. Rev. Neurosci.* **38**, 171–194 (2015).
10. Bocchio, M., Nabavi, S. & Capogna, M. Synaptic plasticity, engrams, and network oscillations in amygdala circuits for storage and retrieval of emotional memories. *Neuron* **94**, 731–743 (2017).
11. Ke, G. et al. Lightgbm: A highly efficient gradient boosting decision tree. *Adv. Neural Inf. Process. Syst.* **30**, 3149–3157 (2017).
12. Lewis, C. M. et al. Recording quality is systematically related to electrode impedance. *Adv. Healthc. Mater.* **13**, 2303401 (2024).
13. Patel, J., Fujisawa, S., Berényi, A., Royer, S. & Buzsáki, G. Traveling theta waves along the entire septotemporal axis of the hippocampus. *Neuron* **75**, 410–417 (2012).
14. Makino, Y., Polygalov, D., Bolaños, F., Benucci, A. & McHugh, T. J. Physiological signature of memory age in the prefrontal-hippocampal circuit. *Cell. Rep.* **29**, 3835–3846.e3835 (2019).
15. Lisman, J. & Buzsáki, G. A neural coding scheme formed by the combined function of gamma and theta oscillations. *Schizophr. Bull.* **34**, 974–980 (2008).
16. Benchenane, K., Tiesinga, P. H. & Battaglia, F. P. Oscillations in the prefrontal cortex: a gateway to memory and attention. *Curr. Opin. Neurobiol.* **21**, 475–485 (2011).

17. Karalis, N. et al. 4-Hz oscillations synchronize prefrontal-amygdala circuits during fear behavior. *Nat. Neurosci.* **19**, 605–612 (2016).
18. Grewe, B. F. et al. Neural ensemble dynamics underlying a long-term associative memory. *Nature* **543**, 670–675 (2017).
19. Stujenske, J. M., Likhtik, E., Topiwala, M. A. & Gordon, J. A. Fear and safety engage competing patterns of theta-gamma coupling in the basolateral amygdala. *Neuron* **83**, 919–933 (2014).
20. Tort, A. B., Komorowski, R., Eichenbaum, H. & Kopell, N. Measuring phase-amplitude coupling between neuronal oscillations of different frequencies. *J. Neurophysiol.* **104**, 1195–1210 (2010).
21. Wirt, R. A. & Hyman, J. M. ACC theta improves hippocampal contextual processing during remote recall. *Cell. Rep.* **27**, 2313–2327 e2314 (2019).
22. Adhikari, A., Sigurdsson, T., Topiwala, M. A. & Gordon, J. A. Cross-correlation of instantaneous amplitudes of field potential oscillations: a straightforward method to estimate the directionality and lag between brain areas. *J. Neurosci. Methods* **191**, 191–200 (2010).
23. Engel, A. K. & Fries, P. Beta-band oscillations—signalling the status quo? *Curr. Opin. Neurobiol.* **20**, 156–165 (2010).
24. Clarke-Williams, C. J. et al. Coordinating brain-distributed network activities in memory resistant to extinction. *Cell* **187**, 409–427.e419 (2024).
25. Colgin, L. L. et al. Frequency of gamma oscillations routes flow of information in the hippocampus. *Nature* **462**, 353–357 (2009).
26. Altmann, A., Toloşi, L., Sander, O. & Lengauer, T. Permutation importance: a corrected feature importance measure. *Bioinformatics* **26**, 1340–1347 (2010).
27. Canolty, R. T. & Knight, R. T. The functional role of cross-frequency coupling. *Trends. Cogn. Sci.* **14**, 506–515 (2010).
28. Lisman, J. E. & Jensen, O. The θ - γ neural code. *Neuron* **77**, 1002–1016 (2013).
29. LeDoux, J. E. Emotion circuits in the brain. *Annu. Rev. Neurosci.* **23**, 155–184 (2000).
30. Scoville, W. B. & Milner, B. Loss of recent memory after bilateral hippocampal lesions. *J. Neurol. Neurosurg. Psychiatry* **20**, 11–21 (1957).
31. Squire, L. R. & Alvarez, P. Retrograde amnesia and memory consolidation: a neurobiological perspective. *Curr. Opin. Neurobiol.* **5**, 169–177 (1995).
32. Riedel, G. et al. Reversible neural inactivation reveals hippocampal participation in several memory processes. *Nat. Neurosci.* **2**, 898–905 (1999).
33. Rosenbaum, R. S. et al. Remote spatial memory in an amnesic person with extensive bilateral hippocampal lesions. *Nat. Neurosci.* **3**, 1044–1048 (2000).
34. Bartsch, T., Dohring, J., Rohr, A., Jansen, O. & Deuschl, G. CA1 neurons in the human hippocampus are critical for autobiographical memory, mental time travel, and autonoetic consciousness. *Proc. Natl. Acad. Sci. USA.* **108**, 17562–17567 (2011).
35. Cambiaghi, M. et al. Higher-order sensory cortex drives basolateral amygdala activity during the recall of remote, but not recently learned fearful memories. *J. Neurosci.* **36**, 1647–1659 (2016).
36. Headley, D. B., Kyriazi, P., Feng, F., Nair, S. S. & Pare, D. Gamma oscillations in the basolateral amygdala: localization, microcircuitry, and behavioral correlates. *J. Neurosci.* **41**, 6087–6101 (2021).
37. Fitzgerald, P. J. et al. Durable fear memories require PSD-95. *Mol. Psychiatry* **20**, 901–912 (2015).
38. Lo, Y., Yi, P. L., Hsiao, Y. T., Lee, T. Y. & Chang, F. C. A prolonged stress rat model recapitulates some PTSD-like changes in sleep and neuronal connectivity. *Commun Biol* **6**, 716 (2023).
39. Narayanan, R. T. et al. Dissociated theta phase synchronization in amygdalo-hippocampal circuits during various stages of fear memory. *Eur. J. Neurosci.* **25**, 1823–1831 (2007).
40. Narayanan, V. et al. Social defeat: impact on fear extinction and amygdala-prefrontal cortical theta synchrony in 5-HTT deficient mice. *PLoS One* **6**, e22600 (2011).
41. Fitzgerald, P. J. et al. Prefrontal single-unit firing associated with deficient extinction in mice. *Neurobiol. Learn. Mem.* **113**, 69–81 (2014).
42. Mathis, A. et al. DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. *Nat. Neurosci.* **21**, 1281–1289 (2018).
43. Dosovitskiy, A. et al. An image is worth 16x16 words: transformers for image recognition at scale. In *Proc. International Conference on Learning Representations*. <https://openreview.net/forum?id=YicbFdNTTy> (2021).
44. Zhang, Z., Zhong, S.-h. & Liu, Y. TorchEEGEMO: A deep learning toolbox towards EEG-based emotion recognition. *Expert Syst. Appl.* **249**, 123550 (2024).
45. Vaswani, A. et al. Attention is all you need. *Adv. Neural Inf. Process. Syst.* <https://doi.org/10.48550/arXiv.1706.03762> (2017).
46. Selvaraju, R. R. et al. Grad-cam: Visual explanations from deep networks via gradient-based localization. In *Proc. IEEE International Conference on Computer Vision*, 618–626 (IEEE, 2017).
47. Chefer, H., Gur, S. & Wolf, L. Transformer interpretability beyond attention visualization. In *Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 782–791 (IEEE, 2021).
48. Cohen, J. A coefficient of agreement for nominal scales. *Educ. Psychol. Meas.* **20**, 37–46 (1960).
49. Fong, R., Patrick, M. & Vedaldi, A. Understanding deep networks via extremal perturbations and smooth masks. In *Proc. IEEE/CVF International Conference on Computer Vision*, 2950–2958 (IEEE, 2019).
50. Liu, S. et al. NLIZE: A perturbation-driven visual interrogation tool for analyzing and interpreting natural language inference models. *IEEE Trans Vis Comput Graph*, 651–660 (2018).
51. McHugh, T. J. *Distinct Cross-Regional Control of Amygdalar Dynamics Reliably Reflects Fear Memory Age*. <https://neurodata.riken.jp/id/20241006-001> (2024).
52. Makino, Y., Wang, Y. & McHugh, T. J. BrainOdyssey2050/makino-wang-et-al-2024-nat-commun_source-code. *Zenodo* <https://doi.org/10.5281/zenodo.13934383> (2024).

Acknowledgements

We thank all members of the Laboratory for Circuit and Behavioral Physiology (McHugh lab) at RIKEN Center for Brain Science for their support and the IT Services at Massey University for providing the A40 GPU server used for computational analyses. This work was supported by Grant-in-Aid for Scientific Research from MEXT (19H05646, 23H05478; T.J.M.), RIKEN Center for Brain Science (T.J.M.), and the RIKEN Special Postdoctoral Research Program (Y.M.).

Author contributions

Y.M. and T.J.M. conceived the study and designed the experiments. Y.M. performed the experiments and analyzed data. Y.W. performed the computational modeling and analyzed data. Y.M., Y.W. and T.J.M. wrote the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-024-54273-3>.

Correspondence and requests for materials should be addressed to Yuichi Makino, Yi Wang or Thomas J. McHugh.

Peer review information *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

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

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2024