Persistently increased expression of PKMζ and unbiased gene expression profiles identify hippocampal molecular traces of a long-term active place avoidance memory and 'shadow' proteins

Jiyeon Han[1,2], Alejandro Grau-Perales[1], Rayna M. Harris[3], Hsin-Yi Kao[1], Asit Pal[2,4], Juan Marcos Alarcon[5], Todd Sacktor[6], Stefano Martiniani[1,2,4,7], Hans A. Hofmann[3], Andre A. Fenton[1,8]

[1]Center for Neural Science, New York University

[2]Center for Soft Matter Physics, Department of Physics, New York University

[3]Department of Integrative Biology and Institute for Neuroscience, The University of Texas. Austin

[4]Simons Center for Computational Physical Chemistry, Department of Chemistry, New York University

[5]Department of Pathology. State University of New York, Downstate Health Sciences University, New York

[6]Departments of Physiology and Pharmacology, Anesthesiology, and Neurology, State University of New York, Downstate Health Sciences University, New York

[7]Courant Institute of Mathematical Sciences, New York University

[8]Neuroscience Institute at the New York University Langone Medical Center, New York, NY 10016

Corresponding authors: Stefano Martiniani (stefano.martiniani@nyu.edu), Hans A. Hofmann (hofmann@utexas.edu) and André A. Fenton (afenton@nyu.edu)

Keywords: transcriptome, immunohistochemistry, tri-synaptic pathway, gene co-expression networks

ABSTRACT

Long-term memory formation transiently activates Ca^{2+}-calmodulin kinase IIα (CaMKII) and atypical protein kinase C isoform iota/lambda (PKCι/λ), whereas persistent activation of the other atypical PKC, protein kinase M zeta (PKMζ), together with its interacting partner, the scaffolding-protein KIBRA, is necessary for maintaining potentiated synapses and memory. Here, we use immediate early gene (IEG) *Arc* activation during active place avoidance memory expression to tag memory-activated neurons with EYFP, followed by PKMζ immunohistochemistry to learn which hippocampal synaptic pathways are persistently altered. For at least 1 month, PKMζ and EYFP-PKMζ colocalization persistently increase in the hippocampal tri-synaptic pathway (dentate gyrus (DG)→CA3→CA1). We thus use DG, CA3, and CA1 transcriptional profiling to identify mRNAs for PKMζ and other molecules that might contribute to maintaining the memory. We find that memory persistence correlates with the upregulation of IEGs *Arc*, *Fos*, and *NPas4* in DG. In contrast, expression of CaMKII, PKMζ, PKCι/λ, KIBRA, and most other LTP-associated proteins do not covary with memory. This result rules out strong memory-related transcriptional regulation, but not regulation of mRNA translation or stability of these "shadow proteins" that, despite being crucial for memory maintenance, evade detection by unbiased transcriptome profiling. We further examine whether covarying gene expression defines molecular ensembles that predict memory and *Prkcz* expression. Community detection, applied to both linear and non-linear pairwise gene co-expression relationships, reveals gene co-expression networks predictive of memory, memory-related IEGs, and *Prkcz* expression. These findings trace: i) a PKMζ engram in the hippocampus; and highlight: ii) sustained IEGs upregulation in >24-h memory maintenance; iii) the utility of investigating both linear and non-linear gene expression covariance patterns; iv) limits of transcriptional profiling for identifying long-term memory molecules, most of which appear to be "in shadow"; and demonstrates v) that weak pair-wise gene-gene correlations can identify strong gene co-expression networks related to behavioral and shadow molecular estimates of long-term memory.

INTRODUCTION

Learning about the significance of places empowers individuals to form and maintain (often for a lifetime) adaptive and context-appropriate memories about locations associated with specific experiences. The cognitive and neural processes underlying place preference and avoidance have been examined in great detail (Jain et al., 2017), with much of this research focused on the rodent hippocampus, which is a thoroughly studied model system for investigating how spatial memory is formed and stored (O'Keefe and Nadel, 1978; Wilson and Tonegawa, 1997; Barry et al., 2012).

The active place avoidance task is used to assess spatial navigation and long-term memory in rodents. The task crucially depends on intact hippocampus function and causes month-long changes in hippocampal synaptic function at the level of the CA3-to-CA1 population of synapses and the medial entorhinal cortex-to-suprapyramidal dentate gyrus (DG) population of synapses (Cimadevilla et al., 2000; Cimadevilla et al., 2001; Park et al., 2015; Pavlowsky et al., 2017; Levy et al., 2019; Chung et al., 2021). The task requires a rodent to explore a circular rotating arena while avoiding the location of a mild footshock. Place avoidance learning causes the transient activation of several long-term potentiation (LTP)-associated molecules in the dorsal hippocampus, including a number of kinases such as the persistent kinases Ca^{2+}-calmodulin kinase IIα (CaMKII) and the atypical protein kinase C isoform PKCι/λ for the acquisition of the conditioned avoidance (Tsokas et al., 2016; Rossetti et al., 2017; Sacktor and Fenton, 2018). Importantly, these molecules are not critical for maintaining memory. In contrast, the other atypical PKC, protein kinase M zeta (PKMζ), an autonomously active kinase whose activity is regulated by its protein level (Sacktor et al., 1993), persistently increases and is necessary for long-term memory for at least a month (Pastalkova et al., 2006; Tsokas et al., 2016; Hsieh et al., 2017), as shown across diverse brain regions, species, and memory paradigms (Drier et al., 2002; Shema et al., 2007; Serrano et al., 2008; Hardt et al., 2009; Madronal et al., 2010; Migues et al., 2010; von Kraus et al., 2010; Asiedu et al., 2011; Cai et al., 2011; Laferriere et al., 2011; Li et al., 2011; Shema et al., 2011; Barry et al., 2012; Velez-Hernandez et al., 2013; Li et al., 2014; Wang et al.,

2016). Several input-specific populations of synapses in the hippocampus undergo synaptic changes, possibly to store memory and realize experience-dependent modifications of information processing (Whitlock et al., 2006; Madronal et al., 2010; Park et al., 2015; Pavlowsky et al., 2017; Levy et al., 2019; Chung et al., 2021). However, which synaptic populations persistently change with experience remains an open question. One hypothesis is that these changes primarily occur in those synapses that are relevant for the maintenance of the memory, in which case we can predict that they will be expressed in distinct synaptic populations within the different hippocampal subfields (Whitlock et al., 2006; Takeuchi et al., 2013; Pavlowsky et al., 2017). An alternative hypothesis posits that changes occur everywhere in a non-specifically structured organization (Madar et al., 2025; Uytiepo et al., 2025). The substantial recurrent connectivity of the CA3 subfield has motivated the theory that it might act as a storage site with robust memory-related synaptic plasticity (Marr, 1971; Treves and Rolls, 1992; Li et al., 2024). Studies of hippocampal neurons that express location-specific place cell discharge identified the importance of the CA3-to-CA1 synapses in memory storage and recall. In contrast, the entorhinal cortex layer III-to CA1 temporoammonic (TA) synaptic terminals are thought to be relevant for encoding momentary information and computations (Marr, 1971; Colgin et al., 2009; Dvorak et al., 2018; Grienberger and Magee, 2022). Inputs from layer II of the lateral entorhinal cortex are thought to provide contextual information, while those from layer II of the medial entorhinal cortex provide spatial signals. These pathways differentially target dendritic compartments in the supra- and infra-pyramidal blades of dentate gyrus (Neunuebel et al., 2013; Nilssen et al., 2019; Wang et al., 2023). The specific synaptic organization of the hippocampal circuit is likely to constrain, and possibly predict, the functional organization of hippocampal information processing.

Molecules such as PKMζ that are crucial for memory maintenance, interact with other effector and regulatory molecules in multiple ways that have yet to be completely identified, but likely are best characterized as a network of interactions abstracted from the list of the molecular components (Kelly et al., 2007a; Kelly et al., 2007b; Tian et al., 2008; Yao et al., 2008; Migues et al., 2010; Sacktor, 2011;

100 Kandel, 2012; Shao et al., 2012; Eom et al., 2014; Vogt-Eisele et al., 2014; Wang et al., 2014; Tsokas et
101 al., 2016; Yu et al., 2017; Sacktor and Fenton, 2018; Tsokas et al., 2024). Although the complete
102 network of memory-mediating molecular interactions that maintain persistent place avoidance
103 memory is unknown, biochemical studies have shown that PKMζ associates with the kidney and brain
104 protein (KIBRA/WWC1) and potentiates synaptic transmission through the action of N-
105 ethylmaleimide-sensitive factor (NSF), which in turn increases the insertion of GluA2 subunit-
106 containing α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA)-type glutamate ionotropic
107 receptors into postsynaptic sites (Yao et al., 2008; Tsokas et al., 2024). This process further coincides
108 with a reduction in NUMB-mediated AMPA receptor endocytosis coordinated by protein interacting
109 with C-kinase 1 (PICK1) from the postsynaptic sites that are maintained by CaMKII-activity/structure
110 (Sacktor, 2011; Sacktor and Hell, 2017; Farah et al., 2019; Tsokas et al., 2024). It is not known,
111 however, whether these processes are largely regulated at the post-transcriptional level, including
112 regulation of mRNA translation and protein stability, or whether they are accompanied by changes in
113 the mRNA levels of these genes, possibly in opposing directions (e.g., PICK1 and NUMB functions are
114 anti-related to the memory-promoting cell biological functions of PKMζ, KIBRA, NSF, CaMKII, and
115 GluA2). Furthermore, LTP-induction is associated with transient upregulation of PKCι/λ, the other
116 atypical PKC, whereas genetic deletion of the PKMζ gene *Prkcz*, results in LTP- and memory-triggered
117 prolongation of compensatory upregulation of PKCι/λ and the conventional isoform PKCβ1 (Tsokas et
118 al., 2016). It is also unclear how this PKMζ-dependent memory-maintenance system interacts with
119 other molecular pathways that have been hypothesized to play a role in the early phases of memory
120 formation. These include mTOR, a master regulator of translation, insulin-like growth factor 2 (IGF2),
121 which promotes the strengthening of memory likely by an independent process (Chen et al., 2011;
122 Alberini and Chen, 2012), or Fragile X Messenger Ribonucleoprotein 1 (FMR1), which can lead to
123 dysregulated mRNA translation, altered synaptic function, abnormal protein synthesis-dependent
124 synaptic plasticity, and poorly coordinated neural discharge (Bassell and Warren, 2008; Thomson et
125 al., 2017; Talbot et al., 2018). These findings suggest that changes in both mRNA expression as well as

126 post-transcriptional regulation of genes involved in synaptic plasticity underlie memory maintenance.
127 Modern genomic techniques such as transcriptome profiling have made it possible to examine the
128 molecular underpinnings of plasticity in animal behavior and can be profitably used to test specific
129 hypotheses about the molecular underpinnings of memory maintenance as well as discover novel
130 pathways for subsequent experimental analysis (Harris and Hofmann, 2014).

131 In the present study, we used active place avoidance as an exemplar of long-term memory
132 maintenance in two separate experiments. Experiment 1 used immunohistochemistry to test the
133 hypothesis that PKMζ exerts its memory-maintaining role by localizing to specific cellular sub-
134 compartments of EYFP-tagged memory cells in the hippocampal subfields. Experiment 2 used RNA-
135 seq in specific hippocampal subfields to test the hypothesis that the transcriptional regulation of the
136 PKMζ system, along with several other candidate genes, are critical for memory maintenance in the
137 active place avoidance paradigm. We also tested the hypothesis that genes for the over 150
138 molecules implicated in the formation of LTP are associated with memory maintenance as well. Lastly,
139 we mined the RNA-seq data obtained from DG, CA3, and CA1 in search of covariant molecular
140 networks predictive of memory and PKMζ gene expression levels, used here as an exemplar of
141 proteins with crucial functions, such as in memory maintenance, but which are expressed at levels so
142 low that they escape detection. We refer to this class of molecules as the ’shadow proteome,' which
143 is distinct from the so-called ‘dark proteome’ (Wright et al., 2022). Using both linear and non-linear
144 methods to measure covarying expression levels, covarying gene co-expression networks were
145 identified that predict PKMζ gene expression and memory, pointing to the feasibility of detecting
146 shadow proteins and potentially identifying novel molecular interactions and pathways associated
147 with memory maintenance.

148 RESULTS

149 <u>Experiment 1 – Tracing Memory</u>

150 *PKMζ protein is persistently increased for at least 30 days in memory-tagged cells*

151 Trained and Untrained mice do not differ in entering the shock zone soon after being placed on the
152 arena during pretraining (Fig. 1c; $t_7 = 0.42$, $p = 0.7$), but only the Trained mice learn to delay entering
153 the shock zone for several minutes by the end of training ($t_7 = 4.18$, $p = 0.006$). This conditioned
154 avoidance persists and is expressed a month later in the retention test with no shock (Trained: 90 ±
155 51 s, Untrained: 5 ± 4 s). Despite the absence of shock, the time to enter the location of shock for the
156 second time is 254 ± 141 s in the Trained group and only 26 ± 17 s in the Untrained group ($t_7 = 3.60$, p
157 = 0.01), confirming place avoidance memory is maintained for a month in the Trained mice (Fig. 1c).

158 The EYFP label is detected throughout the principal cells of Ammon's horn, including the *strata*
159 *radiatum* and *lacunosum moleculare* dendritic compartments and the granule cell and the molecular
160 layers of the supra-DG and infra-DG of both Trained and Untrained mice (Fig. 1d). We first correlate
161 the somatic EYFP signal (in *stratum pyramidale* for CA1, CA2, and CA3 and in the granule cell layer for
162 DG) with the dendritic compartment signals to assess whether the EYFP label is distributed across the
163 *strata* that define synaptic compartments. Within each hippocampal subfield, we observe strong
164 positive correlations between somatic EYFP expression and each of the corresponding dendritic
165 compartments ($0.34 \geq r's \leq 0.90$; $0.89 < t_6's < 5.1$, $0.002 \geq p's \leq 0.4$). These correlations are all
166 significant ($p's < 0.05$) except between the granule cell layer and the adjacent inner molecular layer of
167 infra-DG ($r = 0.54$, $t_6 = 1.6$, $p = 0.2$) and between the pyramidal cell layer of CA3 and the distant
168 stratum lacunosum moleculare compartment ($r = 0.34$, $t_6 = 0.89$, $p = 0.4$).

169 We next determine whether memory training causes elevated EYFP expression in each hippocampal
170 subregion. Relative to control expression values (Fig. 1d), EYFP expression is significantly increased in
171 the Trained group at CA1 ($t_6 = 2.81$, $p = 0.03$), CA3 ($t_6 = 3.80$, $p = 0.008$), supra-DG ($t_6 = 8.61$, $p < 0.001$),
172 and infra-DG ($t_6 = 2.58$, $p = 0.04$), but not at CA2 ($t_6 = 1.20$, $p = 0.3$).

173 We can now investigate which EYFP-tagged synaptic populations persistently increase PKMζ
174 expression after memory training as an indicator of persistently potentiated memory-related
175 synapses. Compared to Untrained, PKMζ is significantly increased throughout the Trained supra-DG
176 (Fig. 1e). Although the M1 coefficients do not significantly differ (all p´s > 0.1), the M2 coefficient
177 estimating PKMζ-EYFP overlap exclusively in EYFP expressing regions is significantly increased in
178 Trained mice at the granule cell layer ($t_6 = 2.65$, $p = 0.02$), the middle ($t_6 = 3.11$, $p = 0.01$), and the
179 outer molecular layers ($t_6 = 4.11$, $p = 0.003$), but not at the inner molecular layer ($t_6 = 2.24$, $p = 0.06$),
180 which only reaches 1-tailed significance. This suggests that synapses of the outer two-thirds of the
181 molecular layer of supra-DG are persistently potentiated a month after memory training, coinciding
182 with where the inputs from LECII and MECII terminate in supra-DG, consistent with
183 electrophysiological observations in freely-behaving mice (Chung et al., 2021).

184 The corresponding analysis of infra-DG finds Trained versus Untrained enhancement of PKMζ only at
185 the middle and outer molecular layers (Fig. 1f) and although there are also no group differences in M1
186 coefficients, the M2 measure of PKMζ-EYFP overlap is significantly increased in Trained compared to
187 Untrained mice at the middle ($t_6 =14.4$, $p < 0.01$) and outer molecular layers too ($t_6 = 4.15$, $p < 0.01$)
188 but not elsewhere (Fig. 1f).

189 CA3 labeling for PKMζ is significantly increased in Trained compared to Untrained *stratum pyramidale*
190 ($t_6 = 3.28$, $p = 0.01$) and the *stratum radiatum* ($t_6 = 3.18$, $p = 0.01$), but not at *lacunosum moleculare* (t_6
191 $= 0.49$, $p = 0.6$). While we do not find any differences between groups in the Mander M1 coefficient
192 (all p´s > 0.4), this coincides with increase of the M2 overlap coefficient at *stratum pyramidale* ($t_6 =$
193 2.80, $p = 0.03$), and *stratum radiatum* ($t_6 = 2.59$, $p = 0.04$), but not *lacunosum moleculare* ($t_6 = 0.32$ $p =$
194 0.8), suggesting persistent potentiation of synaptic inputs from DG, but not entorhinal cortex in
195 memory-tagged CA3 (Fig. 1g).

196 PKMζ in the CA1 subfield is significantly increased in Trained compared to Untrained mice at *stratum*
197 *pyramidale* (t_6 = 3.00, p = 0.01), and *stratum radiatum* (t_6 = 3.63, p = 0.01), but not *lacunosum*
198 *moleculare* (t_6 = 1.24, p = 0.3; Fig. 1h). Similar to CA3, we do not observe group differences in the
199 Mander M1 coefficient (p´s > 0.1), but the Mander M2 coefficient is increased at *stratum pyramidale*
200 (t_6 = 2.80, p = 0.03), and *stratum radiatum* (t_6 = 2.59, p = 0.04), but not *lacunosum moleculare* (t_6 =
201 0.32, p = 0.8) of the memory-tagged cells, which replicates a prior report (Hsieh et al., 2021) and is
202 consistent with persistent memory training causing persistent electrophysiological potentiation of the
203 Schaffer collateral population synapses (Pavlowsky et al., 2017).

204 Within CA2, we do not find any differences in the expression of PKMζ in any of the synaptic
205 compartments (Fig. 1i; p´s > 0.4), the Mander M1 (p´s ≥ 0.2), and the Mander M2 coefficients (p´s ≥
206 0.3) in any of the synaptic compartments of the CA2 region, consistent with observing no evidence of
207 training-associated *Arc* activation.

208 The overall hippocampal pattern of changes in PKMζ expression in memory-tagged cells is
209 remarkable. One month after memory training, the PKMζ increase traces the synaptic compartments
210 of the trisynaptic pathway through the hippocampal circuit (Fig. 1j).

211 *Results motivating Experiment 2*

212 These findings are consistent with the hypothesis that place avoidance training causes synaptic
213 potentiation in the hippocampal tri-synaptic "model" pathway thought to be associated with memory,
214 but not in the direct "data" inputs from entorhinal cortex though to be associated with sensory
215 information processing (Colgin et al., 2009; Fries, 2009; Dvorak and Fenton, 2014; Dvorak et al.,
216 2018). This PKMζ memory trace is selectively observed in the neurons that activate *Arc* during
217 expression of the established place avoidance memory, not merely during learning (Fig. 1), in line with
218 observations that memory training increases PKMζ in CA1 (Hsieh et al., 2017; Hsieh et al., 2021) and
219 changes synaptic function selectively at CA1 *stratum radiatum* (Pavlowsky et al., 2017) as well as

220 Supra-DG, which can be measured as early as after 30 min of memory training (Chung et al., 2021),
221 see also (Whitlock et al., 2006). The observed group differences in PKMζ predict that mice with better
222 memory will express more PKMζ in the memory-tagged cells. Despite there being only four Trained
223 mice, large Pearson correlations are observed between the M2 coefficients of PKMζ in EYFP-
224 expressing cells and the time to first enter into the shock zone on Trial 3 when memory was tagged.
225 Correcting for the two calculations (M1, and M2), the M2 coefficient is significant at the middle
226 molecular layer where the medial perforant path terminates in the supra-DG ($r = 0.98$, $t_2 = 7.0$, $p =$
227 0.02), the site where Chung et al. recorded memory-induced synaptic plasticity in freely-behaving
228 mice for up to 60 days; but the correlation is not significant at infra-DG ($r = 0.93$, $t_2 = 3.7$, $p = 0.06$),
229 mirroring the electrophysiological observations (Chung et al., 2021; Dvorak et al., 2021). There is
230 strong evidence that PKMζ protein expression is persistently and detectably elevated at synaptic
231 populations, specifically in the ChR2-EYFP-tagged neurons that are an enriched sample of memory-
232 related cells. Lastly, prior work showed that the training-induced persistent increase of PKMζ in
233 memory-tagged cells and the synaptic potentiation at CA1 *str. radiatum*, are both indistinguishable
234 whether or not memory retention is tested (Pavlowsky et al., 2017; Hsieh et al., 2021).

235 <u>Experiment 2 – Transcriptional Profiling of Memory</u>

236 *Characterizing memory-related behavior*

237 We first analyze active place avoidance behavior in the memory-trained and yoked-control groups, as
238 a prerequisite for investigating transcriptional profiles related to memory (Fig. 2). During the pre-
239 training test, when the shock was off, there are no differences between the groups, as expected
240 (Table 1). Avoidance is evaluated in multiple ways, including as a reduction in the number of
241 entrances into the shock zone (Fig. 3a), an increase in the time to first enter the shock zone (Fig. 3b),
242 and a decreased likelihood of being in the shock zone (Fig. 3c). As expected, the memory-trained
243 animals show a robust conditioned avoidance whereas the yoked-controls do not, although

244 corresponding pairs of Trained and Control mice experience identical physical conditions (Fig. 2b).
245 These observations are confirmed by two-way ANOVAs, with significant effects of training treatment,
246 trial, and their interaction (Table 1).

247 The conflict-trained animals show similarly rapid avoidance learning (T1-T3) as the standard-trained
248 mice (Figure 3a-c). Initial between-day recall (Rt) of the conditioned place avoidance is also
249 indistinguishable between the two Trained groups, and as expected, the Controls do not express an
250 avoidance. Place avoidance in the two trained groups differ when the shock zone was subsequently
251 relocated for the conflict group (T4-T6), specifically on the first trial with the new shock location (T4),
252 as expected (Figure 3a-c). The conflict-trained group rapidly learns to avoid the relocated shock zone
253 in subsequent trials, so that the ability to avoid shock cannot be distinguished between the two
254 Trained groups (Table 1). Compared to the standard-trained group, the conflict-trained group learns
255 to avoid two places in the same environment and extinguishes the avoidance of one location.
256 Similarly, behavior is indistinguishable between the two yoked controls (Table 1). The day after the
257 conditioning trials, on the memory retention test (Rn) with no shock, the conditioned avoidance
258 persists and is indistinguishable between the two Trained groups, and is not expressed in the
259 Controls, when tested the next day (Rn) under identical conditions for all mice.

260 Spatial memory is a complex phenomenon to which we aim to relate gene expression. In an effort to
261 identify a memory score that does not privilege one estimate of conditioned avoidance over another
262 we perform principal component analysis (PCA) on the 26 measures of behavioral space-use obtained
263 from the automated behavioral analysis of position tracking (Table S1). Measures related to shock
264 avoidance and avoidance conditioning are captured by PC1, which together reveals a “memory
265 syndrome” that explains 37.4% of the variance when all the data from pretraining to retention are
266 analyzed together. Following the experimental design, we focus the analysis on behavior during
267 retention when the conditions were identical for all the mice, which was also just prior to euthanasia.
268 At these times, PC1 explains 52.4% of the variance and PC2 explains 14.9% of the variance (Fig. 3d).

269 The variables with strongest loading on PC1 reflect avoidance behavior (Fig. 3e), whereas PC2
270 captures measures of locomotor activity (Fig. 3f), which can be considered an "activity syndrome."
271 PC1 significantly distinguishes Trained (n = 17) and Control (n = 17) mice ($t_{32} = 6.8$, $p = 10^{-8}$), but PC2
272 does not ($t_{32} = 0.01$, $p = 0.99$). We designate PC1 as $PC1_{mem}$ and use it as a comprehensive measure of
273 the behavioral expression of active place avoidance memory, and PC2 is designated $PC2_{act}$ (Table S1).

274 *Spatial memory drives hippocampal-subfield specific patterns of transcriptome variation*

275 Transcriptomes were obtained from a subset of the mice (Fig. S2). We begin by identifying
276 differentially expressed genes (DEGs) between the four groups of trained and yoked mouse samples
277 from the DG, CA3, and CA1 hippocampal subfields; 2 DG samples (1 standard-yoked and 1 conflict-
278 yoked), 1 CA3 sample (1 conflict-trained) and 2 CA1 samples (1 standard-yoked and 1 conflict-yoked)
279 are eliminated from analysis because of insufficient RNA reads. A total of 39 samples are analyzed
280 (Fig. S2). PCA shows it is possible to distinguish the trained and control DG transcriptomes, but not the
281 CA3 and CA1 transcriptomes (Fig. 4a). Few significant DEGs are found in any subfield between
282 standard- and conflict-trained groups (Fig. 4b,c) so we combine the data into one Trained (n = 9) and
283 one Control (n = 9) group (Fig. 4d). In the ~17,000 gene-transcriptomes we cannot detect any DEGs
284 between Trained and Control CA3 samples, and only 4 in CA1 samples, whereas 263 DEGs are
285 identified in the DG samples (Fig. 4), with more increased than decreased genes in Trained compared
286 to Control samples (Fig 4d). The total DEG list in DG, CA3, and CA1 is available (see Table in Fig. 4
287 legend).

288 While *Prkcz* and related transcripts are not found to be differentially expressed, the IEGs that are also
289 DEGs are significantly correlated with $PC1_{mem}$ (Fig. 5a), and their pairwise correlations are positive
290 (Fig. 5b). These relationships are not observed in CA3, which instead only identifies negative Pearson
291 correlations between memory and *Prkci* that codes for PKC*ι*/λ and *Gria2*, the gene for the GluA2
292 AMPA receptor subunit (Fig. 5c). *Gria2* is also significantly correlated with *Gria1,* the GluR1 AMPA

293 receptor subunit, and *Grin1*, the NMDA receptor subunit GluN1 (Fig. 5d). Furthermore, within CA3,
294 the genes for all the candidate kinases forms a co-expression network (CaMKII, PKCβ, PKMζ, and
295 PKCι/λ) with *Wwc1*, the gene for KIBRA/WWC1 that targets PKMζ to activated synapses (Fig. 5d). In
296 CA1, while expression of none of the candidate genes is correlated with behavioral memory (Fig. 5e),
297 the pattern of significant pairwise correlations is intermediate to the DG and CA3 patterns, in part
298 because *Gria1* expression covaries with the mutually correlated IEGs and also covaries with *Prkcb* and
299 *Wwc1* (KIBRA), which itself correlates with *Prkcz,* the gene for PKMζ (Fig. 5e).

300 Astrocytes contribute to synaptic function and long-term memory persistence (Henneberger et al.,
301 2010; Allen and Eroglu, 2017; Adamsky et al., 2018; Curreli et al., 2022) and indeed, several astrocyte-
302 associated genes are amongst the CA3 DEGs and some of their expression values in CA3 appear to
303 covary with $PC1_{mem}$ (Fig. S3). While this is not the case for these genes in the DG or CA1 samples,
304 nonetheless, many of the astrocytic genes are correlated with each other within the samples from
305 DG, CA3, and CA1 (Fig. S3).

306 The results of the preceding analyses vary substantially across the hippocampal subfields, so we
307 examine whether the findings are robust or spurious. For one, the above analyses assume linear
308 relationships between gene expression and memory or *bona fide* memory-related molecules like
309 PKMζ, but some of these relationships are known to be non-linear due to regulatory, feedback and
310 feedforward processes (Sacktor, 2011; Tsokas et al., 2016). We therefore investigate whether
311 associating these DEGs with memory is robust across three independent information-theoretic and
312 statistical measures for ranking genes in the DG samples according to their ability to differentiate
313 between Trained and Control median-of-ratio normalized gene expression levels. Indeed, Kullback-
314 Liebler Divergence (KLD), mutual information (MI), and the Mann-Whitney tests all identify the IEGs
315 and many of the DEGs identified by DESeq2, but not PKMζ (Fig. S4). In particular, the LTP-associated
316 genes *Arc*, *Egr1*, *Npas4*, and *Sgk1* all extremize these measures, indicating that they are most strongly
317 associated with memory persistence.

318 *Might covarying gene co-expression networks predict shadow proteins better than individual genes?*

319 Our observation that both the memory/LTP-related candidate genes, as well as the astrocytic
320 candidate genes form co-expression networks indicates that instead of defining a high-dimensional
321 state space of possible gene expression patterns, gene co-expression defines a lower dimensional
322 state space. The gene-gene correlations define the intrinsic dimensions of the state space within
323 which the Training and Control conditions are defined. We next ask whether such gene co-expression
324 networks can reveal molecules in the "shadow proteome" such as PKMζ.

325 While the results of linear Pearson and non-linear Spearman correlations are largely consistent, both
326 fail to detect non-linear but biologically meaningful relationships that can be measured by
327 Chatterjee's ξ (Fig. S5; Chatterjee, 2021). Since the DEGs distinguish Trained from Control samples, it
328 is not surprising that 94% of the DEGs have significant Pearson correlations with $PC1_{mem}$. We compare
329 Pearson, Spearman, and Xi (ξ) measures of relatedness to $PC1_{mem}$ (Fig. 6). While the relationship of
330 gene expression to $PC1_{mem}$ for some genes is well characterized by the three measures, for some it is
331 not. Because of its sensitivity to outliers, subsequent analyses use Spearman instead of Pearson
332 correlations, and because of its ability to capture non-linear relationships like sigmoids and
333 thresholds, we use ξ in subsequent analyses in a combined measure, Σ, that is the norm of the two
334 correlation measures (see Methods).

335 *Shadow proteins: differential gene expression does not identify* PKMζ and related proteins in long-
336 term *memory persistence*

337 Active place avoidance memory training increases PKMζ protein in specific hippocampal subfields (Fig.
338 1). This complements findings from biochemical and genetic manipulations that *de novo* synthesis of
339 PKMζ and its interactions with KIBRA and other proteins in hippocampus are crucial for the
340 persistence of the active place avoidance memory (Pastalkova et al., 2006; Tsokas et al., 2016; Tsokas
341 et al., 2024), as well as other types of memory (Shema et al., 2007; Serrano et al., 2008; Madronal et

342 al., 2010; Migues et al., 2010; von Kraus et al., 2010; Laferriere et al., 2011; Li et al., 2011; Sacktor,
343 2011; Shema et al., 2011; Barry et al., 2012; Wang et al., 2016; Sacktor and Hell, 2017), even though
344 genetic deletion of PKMζ is compensated by increased expression of PKCι/λ (Lee et al., 2013; Volk et
345 al., 2013; Tsokas et al., 2016). Nonetheless, candidate LTP/memory-persistence genes that differ
346 significantly between the Trained and Control groups in the DG, CA3, or CA1 samples, are not related
347 to the behavioral measure of memory ($PC1_{mem}$), despite the memory-associated increases in the PKMζ
348 protein with large effect size (Cohen's d: DG = 1.3, CA3 = 1.5, CA1 = 2.5; power analysis with d = 1.5
349 predicts n's of 7 are sufficient to detect RNA differences at $\alpha = 0.05$ and $\beta = 0.8$). These data indicate
350 that detection and discovery of such 'shadow' proteins is unlikely from standard transcriptomic
351 analyses, leading us to consider innovations that can detect memory-related expression of PKMζ, and
352 related molecules.

353 Weak-pairwise correlations amongst the many (i.e., high-dimensional) components of a system can
354 indicate strong low-dimensional network states, in which single components are not crucial, but the
355 collective interactions are strong determinants of function (Schneidman et al., 2006). Such low-
356 dimensional manifold interactions are a familiar feature in studies of correlated neuronal action and
357 field potentials (Schneidman et al., 2006; Levy et al., 2023; Fenton, 2024) that have been proposed to
358 replace the reliance on individual neurons (the neuron doctrine), with a population doctrine that
359 relies on neuronal interactions to explain system functions (Ebitz and Hayden, 2021). Accordingly, the
360 positive identification of PKMζ as a shadow protein motivates exploring whether networks of
361 transcriptomic interactions rather than single genes, may better predict the importance of proteins
362 for memory persistence.

363 To test this idea, we use the Louvain Method of community detection (Blondel et al., 2008) to define
364 gene co-expression networks (communities) of strongly correlated genes that are weakly correlated
365 across communities. We also employ weighted gene co-expression network analysis (WGCNA;
366 Langfelder and Horvath, 2008) to infer gene co-expression networks that predict the behavioral

367 Trained/Control group, and PKMζ; however the results are unstably sensitive to algorithm parameters
368 (data not shown). We use the norm of both Spearman's ρ and Chattejee's ξ correlation coefficients as
369 the relationship metric to $PC1_{mem}$, and perform the Louvain clustering analysis 1000 times with
370 randomly selected seeds. Figure 7a, shows the accumulated colors for a set of 714 genes (702 genes
371 significantly correlated with $PC1_{mem}$ using norm correlation-based p-value ≤ 0.05, as well as twelve
372 additional candidate genes to yield n = 714 unique genes). We find that these genes are reliably
373 assigned to one of the three groups, each defined by the presence of the genes *Arc*, *Nsf*, and *Prkcz*, to
374 which we assign red, green, and blue colors, respectively (Hanley et al., 2002; Sacktor and Hell, 2017;
375 Weng et al., 2018). The top-20 expressed genes in each of the three communities and their associated
376 candidate genes reveal three robust gene co-expression networks (Figs. 7a, S6). Specifically, the IEGs
377 are part of the red ensemble defined by *Arc* membership. Genes associated with early-LTP and the
378 trafficking and insertion of AMPA receptors associate within the green ensemble defined by *Nsf*
379 membership. Finally, the NMDA receptor gene *Grin1* is part of the blue ensemble defined by *Prkcz*
380 membership. Note that *Wwc1* (KIBRA) is uniquely yellow, indicating that it is assigned with similar
381 probability to the red ensemble enriched with IEGs, and the green ensemble enriched with genes
382 associated with early-LTP and AMPA receptor placement. Genes associated with PKMζ-mediated
383 effects on AMPA receptors, *Pick1, Gria1, and Prkcb* are identified as turquoise because they are
384 reliably assigned to either the green ensemble or the blue ensemble. It is also remarkable that *Fmr1* is
385 gray as it is assigned approximately equally to each of the three communities. This suggests there may
386 be a detectable logic to the organization of these gene ensembles.

387 We sought to further validate this network approach by using the 714 gene–gene norm score (**Σ**) as
388 edge weights to construct an undirected graph G with genes as nodes. To assess community stability,
389 we ran Louvain clustering 1,000 times with different random seeds. For each run, we recorded the co-
390 occurrence counts of gene pairs clustering together, accumulating these in an 714 by 714 'clustering
391 probability' matrix that is a gene-wise co-occurrence probability matrix. PCA analysis of the clustering

392 probability matrix was performed to embed the genes into a 3D space (PC1, PC2, and PC3). We
393 computed the clustering probability matrices for the separate yoked and trained subgraphs to assess
394 whether clustering probabilities changed with memory training. The separate yoked and trained
395 clustering probability matrices were projected into the 3D space using the same PCA loadings of the
396 total samples to ensure a common embedding for the comparison (Fig. 7b). To enhance visual
397 contrast, we normalized the raw PCA coordinates and linearly scaled them to hex color codes. The
398 trained and yoked gene-gene relations appear distinctive because the yoked gene set is compact
399 whereas the trained set is dispersed along triangular axes (Fig. 7b).

400 We also performed PCA on the raw expression matrix (genes × samples) after z-scaling each gene's
401 expression across all samples. This raw-expression PCA failed to reveal clear separation of either the
402 candidate genes or the two training conditions, underscoring the value of investigating gene-gene
403 relationships and the clustering probability approach to analyze those relationships (Fig. S7).

404 To compare whether co-clustering patterns of each gene are lost and gained between yoked and
405 training, we generated agglomerative cluster maps using the separate yoked and trained clustering
406 probability matrices (Fig. 7c). We first subtracted the yoked matrix from the trained matrix and
407 performed hierarchical clustering with the Pearson correlation as the distance metric. The range of
408 the heatmap is -1 to 1, which corresponds to losing and gaining co-cluster likelihood, respectively. In
409 the heatmap based on all 714 genes, *Prkcz* was grouped together with *Nsf* but not with *Numb*, which
410 corresponds to the synaptic plasticity related cellular biology that PKMζ and NSF together promote
411 synaptic potentiation whereas NUMB promotes depotentiation. In the total dataset for the 714-gene
412 heatmap, *Prkcz* was grouped together with the genes associated with the endoplasmic reticulum
413 related gene ontology terms.

414 To highlight the genes that are driving the differences between the yoked and trained conditions, we
415 computed for each gene the three-dimensional shift vector $\Delta C = PC_{tr} - PC_{yo}$ using the clustering

probability matrix matrices. From ΔC, we derived: (1) the Euclidean Distance ($\| \Delta C \|$) — overall magnitude of each gene's shift; and (2) the Centroid Similarity ($\frac{\Delta C \cdot u}{\|\Delta C\|}, u = \frac{\mu_{tr} - \mu_{yo}}{\|\mu_{tr} - \mu_{yo}\|}$), which is the cosine similarity between each gene's shift and the global centroid shift between the yoked and trained conditions; and (3) the Gene Similarity ($\frac{\mathrm{PC_{yo}} \cdot \mathrm{PC_{tr}}}{\|\mathrm{PC_{yo}}\| \|\mathrm{PC_{tr}}\|}$) , which measures the cosine similarity between the yoked and trained conditions for each gene (Fig. 7d). We observed that the genes are distributed along a low-dimensional manifold defined over these three features (Fig. 7e), which we confirm by t-SNE non-linear dimensionality reduction (Fig. 7f). The distribution along the manifold reflects the Euclidean distance from the yoked condition. Notably, the KIBRA gene *Wwc1* is at the extreme of the manifold, distinct from the other candidate genes that are separately distributed along the manifold.

DISCUSSION

Summary and Limitations

To detect changes in the expression of genes for PKMζ and other proteins known to be important for long-term memory, we first assessed gene expression changes at the individual-gene level by identifying differentially expressed genes, few of which corresponded to proteins know to be important for long-term memory (Fig. 4,5). Next, we examined gene expression shifts at the population level, considering all genes significantly altered by training (Fig. 6) and a set of candidate genes hypothesized to respond to memory training by observation or their biochemical relationship with PKMζ. To capture these population-level changes, we clustered genes according to their pairwise co-expression correlations across the two conditions rather than relying on individual gene-wise differential expression. For each gene, we then estimated its cluster-membership probability and compared how those probabilities differed between conditions. Overall, our results demonstrate that active place avoidance memory training alters the covariance structure among individual gene expression changes (Fig. 7).

440 We began by evaluating long-term memory for the conditioned avoidance response in active place
441 avoidance memory. We first measured that PKMζ protein is persistently increased in the
442 somatodendritic compartments of the tri-synaptic pathway for at least a month (Fig. 1). We then
443 quantified both recall of the conditioned behavior and the associated transcriptional changes in three
444 subfields of the hippocampus a day after the end of the training and 30-min after expression of the
445 avoidance memory. The behavioral similarities and differences between the Trained and Control
446 groups reflect knowledge about the location of shock rather than the experience of shock *per se*,
447 because the 24 hours before euthanasia were physically identical across the groups, constituting time
448 in the home cage and 10 minutes on the arena with no shock, and the comparison groups
449 experienced the identical physical conditions. We cannot assert that the gene expression findings
450 associated with active place avoidance memory are due to altered gene expression specifically in
451 neurons because we sampled bulk tissue, containing approximately equal numbers of neurons and
452 other cells that are largely the glial population (Herculano-Houzel et al., 2006).

453 We started with an unbiased data-driven approach to identify genes that are sensitive to the
454 maintenance of spatial memories across at least 24 h, then conducted a hypothesis-driven analysis of
455 the data to characterize whether the memory training is associated with altered expression of dozens
456 of genes that have been previously shown to be important for learning, memory, and the cellular
457 model LTP, including PKMζ that we confirmed is increased two-fold at each of the subfields we
458 sampled for transcriptional profiling. For this hypothesis-driven approach, we examined both
459 differential gene expression and its covariance with estimates of avoidance memory behavior and
460 with other candidate genes. Our RNA-seq methods could not detect ribosomal RNA, nor did we
461 perform protein sequencing, limiting the scope of this study to findings that may related to memory-
462 related transcriptional regulation.

463 *The PKMζ engram*

464 Richard Semon introduced the term “engram” to mean the physical trace that stores memory, which
465 is a concept that Plato introduced by proposing memory is like a block of wax in which perceptions
466 and thoughts are imprinted, “just as we make impressions from seal rings” (Plato et al., 1935; Josselyn
467 et al., 2017). We observed that PKMζ is persistently increased for at least 1 month, selectively in
468 memory-tagged cells and specifically at synaptic populations of the tri-synaptic pathway (Fig. 1). We
469 did not demonstrate if this PKMζ trace is necessary or sufficient for maintaining the conditioned place
470 avoidance behavior; however, prior work discovered place avoidance training potentiates the Schaffer
471 collateral synaptic population at CA1 *stratum radiatum* but not at CA1 *lacunosum moleculare*
472 (Pavlowsky et al., 2017) and at the middle molecular layer of supraDG, complementing our current
473 findings (Fig. 1; Chung et al., 2021; Hsieh et al., 2021). Notice that the compartment-specific
474 persistent increases of PKMζ in DG, CA3, and CA1 point to an active process targeting the protein,
475 since PKMζ is observed throughout the hippocampus circuit, but is elevated selectively by memory
476 training. The compartmentalization of the PKMζ increases that we observe is hard to reconcile with a
477 somatic synthesis of PKMζ and its free diffusion in cytosol to activated synapses. However,
478 compartmentalization is compatible with local translation and recent findings that PKMζ complexes
479 with the cytoskeleton-associated postsynaptic scaffolding protein KIBRA that persistently targets the
480 kinase’s action to activated synapses to store memory (Muslimov et al., 2004; Tsokas et al., 2024).
481 These findings together suggest the PKMζ trace is a *bona fide* engram because i) PKMζ is both
482 necessary and sufficient for maintaining late-LTP in wild-type mice, and active place avoidance
483 memory in wild-type mice is eliminated by ii) inhibition of PKMζ synthesis during learning, and, for up
484 to a month after the avoidance memory is established, it is erased by iii) inhibition of PKMζ kinase
485 activity or iv) inhibition of PKMζ binding to KIBRA (Ling et al., 2002; Pastalkova et al., 2006; Serrano et
486 al., 2008; Tsokas et al., 2016; Tsokas et al., 2024). Systems consolidation, taken at face value, might
487 suggest that memory no longer depends on hippocampus after a month, but that is falsified by the
488 cited work as well as demonstration of a core and persistent hippocampal dependence of other forms
489 of memory, despite the possibility of systems consolidation (Sutherland et al., 2008; Hardt and Nadel,

490 2017; Tonegawa et al., 2018; Refaeli et al., 2023). Accordingly, the inability to detect simple
491 transcriptomic differences in LTP- and memory-associated genes cannot be attributed only to systems
492 consolidation at the early time points after training that we examined. Our observations of the PKMζ
493 trace are consistent with it being a memory trace by definition, and arguably a long-pursued engram.

494 *Memory-associated Differentially Expressed Genes*

495 Gene expression differences between trained and control groups are mostly observed in DG, with
496 substantially more increases of gene expression than decreases in the trained mice compared to the
497 yoked controls. Few genes were differentially expressed in Ammon's horn (Fig. 4), which is not
498 expected in light of theories of hippocampus memory and information processing (Marr, 1971; Treves
499 and Rolls, 1992; Lisman et al., 2017) and findings that maintenance of the active place avoidance
500 memory persistently enhances CA3-CA1 synaptic function for at least 30-days (Pavlowsky et al., 2017).
501 Nor did we find memory maintenance-related differential expression of genes for the molecular
502 translation machinery in the MEK/ERK and MTOR pathways, important for memory-related
503 transcription and translation during initial cellular memory consolidation (Adams and Sweatt, 2002;
504 Graber et al., 2013). Several genes that are differentially expressed in DG and are positively correlated
505 with avoidance behavior are IEGs, including *Arc* and the transcription factors *Fos*, *Jun*, and *Npas4*. In
506 addition to being differentially expressed, expression of these genes is highly correlated with one
507 another. In previous experiments, *Arc* was activated during contextual threat avoidance conditioning
508 and optogenetically silencing the *Arc*-activated DG neurons was sufficient to impair contextual fear
509 memory, whereas optogenetically stimulating neurons that activated *c-Fos* was sufficient for
510 expressing the memory (Liu et al., 2012). Indeed, as we also observed in DG, prior work showed that
511 *Arc* and *Fos* expression are correlated in active CA1 neurons (Mahringer et al., 2022). Because we
512 examined gene expression a day after training when training-associated IEG activation is not expected
513 (Pace et al., 2005; Pevzner and Guzowski, 2014), it is possible that the robust IEG activation we
514 observed is related to the mice learning about the absence of shock during the retention test.

515 However, this explanation predicts an inverse relationship between $PC1_{mem}$ and IEG expression,
516 because mice that expressed the weakest place avoidance by entering the shock zone the most
517 during the retention test (low $PC1_{mem}$ values) would have the greatest opportunity to learn that shock
518 was absent and the five mice that showed perfect avoidance by never entering during the retention
519 trial could not have learned that the shock was off. We reject this alternative because the correlations
520 between IEG expression in DG and $PC1_{mem}$ on the retention test are all positive. Instead, these
521 findings suggest that IEGs can respond not only to the novel experience of stimuli that are encoded
522 during the conditioning of a response, but that their activation in DG also corresponds to the
523 expression of long-term memory (O'Connell et al., 2012; Miller-Crews and Hofmann, 2024).

524 Gene ontology analysis revealed that many of the genes that are differentially expressed in the DG
525 exert their function in dendrites and are important for neuronal projections and cell-cell junctions,
526 including synapses, which is expected from the relationship between LTP and memory that forms the
527 basis of the widely-accepted synaptic plasticity and memory hypothesis (Table S2; Takeuchi et al.,
528 2013). Although, we cannot distinguish the contributions of neurons and glia to the memory-related
529 differential gene expression we observed, these findings suggest that the expression of the
530 conditioned place avoidance altered neural activity in a way that increases transcriptional activity,
531 perhaps for the purpose of promoting dendritic growth. Indeed, we note that expression of a number
532 of astrocyte-related genes, especially in CA3 were positively correlated with the $PC1_{mem}$ estimate of
533 memory expression, including *Aldh1a1*, *Aldh1l2*, *Gfap*, *Fgfr3*, and *Aqp4* (Fig. S3). Even though these
534 genes were not differentially expressed after avoidance training, the correlation with expression of
535 the conditioned avoidance behavior adds to the growing evidence that astrocytes also play an active
536 role in forming and/or stabilizing memory via a variety of potential mechanisms (Parpura et al., 1994;
537 Henneberger et al., 2010; Suzuki et al., 2011; Rusakov, 2015; Gao et al., 2016).

538 Synaptic reorganization requires that neurons polarize their dendrites to create new neural
539 connections and direct channels to specific regions within the synapse. Disruption of the processes

540 that regulate synaptic organization can impair memory or cause disease. Dysregulation of *Eif5* a DEG,
541 is associated with many neurodegenerative diseases (Kapur et al., 2017). *Cpeb4*, a DEG, encodes
542 cytoplasmic polyadenylation element binding protein 4, which regulates the translation of genes by
543 modulating their poly(A)-tail length and is a known regulator of genes associated with autism
544 spectrum disorder (Parras et al., 2018). Many of the genes that are mutated in autism spectrum
545 disorder are crucial components of the activity-dependent signaling networks that regulate synapse
546 development and plasticity (Ebert and Greenberg, 2013). Deletion of *Fmr1* can lead to dysregulated
547 mRNA translation, altered synaptic function, abnormal protein synthesis-dependent synaptic
548 plasticity, and poorly coordinated neural discharge (Bassell and Warren, 2008; Thomson et al., 2017;
549 Talbot et al., 2018).

550 Very few of the differentially expressed genes in our study overlapped with the LTP-related molecules
551 that were identified from review of the LTP literature (Table S3; Sanes and Lichtman, 1999). Some
552 classes of genes that are notably absent from the list of differentially expressed genes but present in
553 the LTP-related list are genes for receptors (calcium, glutamate, GABA, and serotonin), calcium or
554 calmodulin-binding proteins, and adhesion molecules. This is perhaps not surprising because the
555 review focused on literature examining LTP induction, the first minutes of long-term potentiation,
556 whereas our experiments were designed to investigate 24-h memory expression, which would be
557 expected to correspond to LTP maintenance. Thus, this work is consistent with the evidence (Table
558 S3; reviewed in Sacktor and Hell, 2017) that like LTP-induction, many molecules known to be crucial
559 for learning, are neither crucial nor expressed in LTP maintenance and as we observe, memory
560 persistence. This points to different molecular mechanisms for forming a memory and maintaining
561 and/or later expressing it days later.

562 *Transcriptional profiling did not identify genes known to encode proteins crucial for LTP and memory*
563 *persistence*

564 What then of the atypical PKCs and CaMKII, candidate molecules that are demonstrated to have
565 crucial roles in the maintenance of LTP and synaptic structure, as well as active place avoidance
566 memory, as demonstrated in Figure 1 (Pastalkova et al., 2006; Tsokas et al., 2016; Wang et al., 2016;
567 Hsieh et al., 2017; Rossetti et al., 2017; Sacktor and Fenton, 2018; Hsieh et al., 2021; Tsokas et al.,
568 2024)? We did not detect differential expression of these genes, nor did we detect correlations
569 between any of the corresponding mRNAs and the behavioral measures of avoidance memory
570 expression, whereas we did for IEGs as discussed above. It would, however, be naive and/or hasty to
571 conclude that despite being unbiased, this transcriptomic screen contradicts the findings of the
572 candidate molecule approaches that have identified a role for these molecules, especially in active
573 place avoidance memory. This highlights a major shortcoming of using transcriptional profiling to
574 identify molecular mechanisms of functions such as memory. First, the abundance of mRNA does not
575 relate in a straightforward way to the cellular abundance of protein because the relationship depends
576 on the translational regulation of the particular protein as well as protein translocation and stability
577 (Vogel and Marcotte, 2012; Liu et al., 2016). Indeed, as an example, while PKMζ is strongly regulated
578 by repressing translation at dendrites (Hernandez et al., 2003), there is no evidence that it is
579 transcriptionally regulated, either in late-LTP (Kelly et al., 2007a) or in active place avoidance
580 conditioning (Hsieh et al., 2017), in line with the present findings. While this highlights a limitation of
581 transcriptional profiling, it draws attention to the necessity of performing combined transcriptional
582 and proteomic profiling in tissue taken from the same subjects (Liu et al., 2016). In light of the present
583 findings, the hypothesis that PKMζ's molecular interactions maintain memory (Sacktor, 2011, 2012;
584 Sacktor and Fenton, 2018; Tsokas et al., 2024) predicts independence between *Prkcz* RNA, that
585 encodes PKMζ, and measures of memory in an unbiased screen that integrates memory-related
586 measures of behavior with circuit-specific transcriptional profiling (RNA), as we have done and
587 observed here. Somewhat unintuitively, the hypothesis also predicts that in the same tissue, memory
588 measures correlate positively with PKMζ protein levels and those of its interaction partners (e.g.,
589 KIBRA, NSF, NUMB, PICK1, GluA2), even though their RNA levels show little phenotypic variation

590 (Hanley et al., 2002; Yao et al., 2008; Makuch et al., 2011; Vogt-Eisele et al., 2014; Yu et al., 2017).
591 There is likely substantial utility in such analyses that use experimental design and concurrent multi-
592 scale measurements from single subjects to integrate information from across levels of biology (here
593 behavior, protein and gene expression), not only to discover unknown and unanticipated mechanisms
594 of complex, poorly-defined processes like memory, but to also test the validity of specific mechanistic
595 hypotheses. For example, the notion that the atypical PKCs may play complementary but competing
596 roles in the maintenance of LTP and memory, perhaps through binding to KIBRA, predicts anti-
597 correlation between *Prkcz* and *Prkci* and whichever isoform is more expressed, to be correlated with
598 KIBRA. This pattern was however not observed, suggesting a simple model of transcriptional
599 regulation does not map their functional protein-mediated functional relationships. Although we have
600 observed it is not straightforward, the findings nevertheless suggest that the transcriptional level of
601 biology can relate to biological function that is mediated by proteins, even rare ones like PKMζ.

602 *Illuminating the shadow proteome*

603 Taking a lesson from a major conceptual advance in analysis of neuronal spike train populations, we
604 also considered whether hypothesizing straightforward relationships between gene expression (or
605 even protein expression in principle) is conceptually fraught, in part because individual expression
606 changes can have manifold interactions with network-level consequences (Yuste, 2015; Ebitz and
607 Hayden, 2021). We first, demonstrated that non-linear measures of correlation like Chatterjee's ξ
608 identify *bona fide* relationships that the standard use of Pearson's correlation failed to identify (Fig.
609 6). We then used the combination of Spearman's ρ and ξ correlations to define ensembles of mutually
610 correlated genes (Figs. 5,7). Initial skepticism was assuaged because ensembles of mutually correlated
611 genes were reliably identified, i.e., their membership was meaningful according to known biochemical
612 function, and perhaps most significantly, this approach identified candidate genes independently
613 proven to be involved in memory. This last fact suggests it may be possible to identify both known (as
614 shown here) and unknown components of the shadow protein by challenging the conventional

615 hypothesis that there are unique, essential components for maintaining memory. Put simply, function
616 emerges from the subtle, cooperative tuning of many components, and not from a dramatic change
617 in any one, just as an arch bears a bridge's weight through the balanced interplay of all its stones, not
618 by any single one. Consider for example, the ongoing debate as to whether the persistent kinase
619 CaMKII or PKMζ is crucial for persistent memory storage, or whether some other molecule is crucial,
620 at the exclusion of all others (Frankland and Josselyn, 2013; Morris, 2016; Lisman, 2017). Neither
621 kinase, nor the gene for the compensating kinase *Prkci* were found to be differentially regulated
622 between the memory and control groups, nor did memory measures correlate with the expression of
623 these kinase genes or with the expression of their scaffolding partners that guide their function, like
624 KIBRA. However, by abstracting both linear and non-linear gene-gene relationships to define the
625 concept of a gene co-expression network, such distinct gene ensembles could be i) reliably defined,
626 and ii) were the dataset's best predictors of the involvement of genes that encode shadow
627 components, undetected but known proteins that are crucial for the memory being investigated (Figs.
628 1,7).

629 *Conclusion*

630 Our findings suggest a hypothesis, which contends that a low-dimensional manifold of molecular
631 interactions (genes, mRNAs, proteins) is embedded in the high-dimensional state space of biological
632 components. The manifold dimensions are defined by biological interactions/functions, not by the
633 biological components themselves. This view extends the manifold concept from our
634 electrophysiological studies to other levels of biology, where the high-dimensional space of potential
635 memory-maintaining biological components, in practice, collapses into a low-dimensional manifold of
636 biological solutions for maintaining memory (Levy et al., 2023; Fenton, 2024). In this view, the logic of
637 the functional implementation, rather than the traditional, parallel measurement of physical
638 components is best able to identify the solutions. Despite vast heterogeneity in biological elements
639 (genes, proteins, cell types with their own cell biology and electrophysiology that contribute to

640 memory), this view asserts there is a conserved computational logic to how these elements interact
641 to realize memory storage and expression. Innovative research programs will be required to validate
642 or refute this conjecture (Sinha et al., 2020), with important implications if it is supported. For
643 example, if key biological functions like memory are governed by a manifold of multiple replaceable
644 interacting genes and/or proteins, then the endeavor to find definitive "causal" biomarkers of
645 function and dysfunction is at best fraught, because key components can be replaced when they are
646 absent or defective (Harris and Hofmann, 2014; El-Brolosy and Stainier, 2017), as is the case when
647 PKMζ is genetically deleted and is certainly the case for neuronal circuit functions and the cell-specific
648 molecular phenotyping that has been compellingly demonstrated with the crustacean stomatogastric
649 ganglion (STG) model system (Tsokas et al., 2016; Northcutt et al., 2019; Tsokas et al., 2024). On the
650 other hand, as the STG as an archetype demonstrates, neuronal systems are robust and treatment
651 options can be plentiful when the system components are manifold, replaceable, and can be
652 coordinated (Goldman et al., 2001; Marder et al., 2014; Marder et al., 2015).

653 METHODS

654 Experimental Subjects

655 All animal care and use complied with the Public Health Service Policy on Humane Care and Use of
656 Laboratory Animals and were approved by the New York University Animal Welfare Committee and
657 the Marine Biological Laboratory IACUC. Transgenic ArcCreER^{T2} (originally purchased from The
658 Jackson Laboratories, Jax: 021881) and R26R-STOP^{flx}-ChR2-EYFP mice (a gift from Christine Denny)
659 were crossed to generate the ArcCreER^{T2}-ChR2-EYFP mice for Experiment 1, as previously described
660 (Denny et al., 2014; Lacagnina et al., 2019; Hsieh et al., 2021). Wildtype C57BL/6J mice (originally
661 purchased from The Jackson Laboratories, Jax: 000664) were used for Experiment 2.

662 Experiment 1 – Tracing Memory: Memory-induced increases of protein expression were studied to
663 identify changes in the strengthening of hippocampal synaptic populations that persist for a month

664 after memory training. We focused on PKMζ because it is both necessary and sufficient for the
665 maintenance of LTP in hippocampal slices over at least five hours, and PKMζ in hippocampus is
666 necessary for active place avoidance memory and its memory-initiated protein expression is
667 enhanced in CA1 for at least a month (Pastalkova et al., 2006; Tsokas et al., 2016; Hsieh et al., 2017;
668 Hsieh et al., 2021). PKMζ protein increases in synapses during memory storage, making this an ideal
669 method to investigate the subcellular distribution of the protein within the different regions of the
670 hippocampus (Hsieh et al., 2021). PKMζ visualization through immunohistochemistry (IHC) was
671 combined with Targeted Recombination in Active Populations (TRAP2) technology to identify neurons
672 that express a fluorescent marker (EYFP) once the expression is initiated by the external ligand, 4-
673 hydroxytamoxifen (TAM, Sigma, St. Louis, MO, H7904). The Cre-mediated recombination is directed
674 to neurons expressing high levels of the immediate early gene (IEG) *Arc* while TAM is circulating. The
675 combination of these two techniques allows us after a month's delay to identify the EYFP-expressing
676 neurons as likely to have been active during TAM administration, and to interpret the synaptic sites
677 within these neurons that exhibit elevated PKMζ levels as loci of persistent potentiation.

678 Thus the ArcCreER^{T2} x ChR2-EYFP TRAP2 mice (n=4 Trained and n=4 Untrained) were temporarily
679 housed in an experimental room with an isolation chamber (1 mouse/cage) on a 12:12 light/dark
680 cycle. These mice contained a ArcCreER^{T2} DNA sequence allowing initiation of transcription at the *Arc*
681 promoter to express Cre recombinase in the presence of TAM. The mice also contained a STOP codon
682 flanked by LoxP sites followed by the sequence for Channelrhodopsin 2 (ChR2) fused with EYFP (Fig.
683 1a). Thus, TAM will induce activation of Cre recombinase in *Arc* activated cells to remove the STOP
684 codon to cause perpetual membrane-targeted EYFP expression. These ArcCreER^{T2} x ChR2-EYFP mice
685 allowed us to tag neurons with EYFP if they activate *Arc* during expression of the conditioned place
686 avoidance, as in prior work (Denny et al., 2014; Hsieh et al., 2021).

687 To reduce spurious activation of *Arc* and EYFP tagging, 2-3 days before the initial exposure to the
688 behavioral arena, all mice began housing in a sound-insulated and light-controlled isolation chamber

689 that was approximately 1 m away from the behavioral apparatus. The mice remained housed in the
690 isolation chamber for two days after TAM administration and afterwards were returned to the
691 vivarium. This isolation limited the association of EYFP expression to the arena experience, as was
692 validated in prior work (Hsieh et al., 2021). Thirty minutes before the third training trial when the
693 mice expressed strong place avoidance memory (see below), all mice were injected with TAM (2 mg in
694 200 µl i.p.).

695 Experiment 2 – Transcriptional Profiling of Memory: The subjects for transcriptional profiling were a
696 subset of the 34 C57BL/6J mice that went through the behavioral training protocol. They were housed
697 at the Marine Biological Laboratory on a 12:12 (light: dark) cycle with continuous access to food and
698 water in home cages with up to five littermates. Only male mice were analyzed in this study to limit
699 gene expression variations due to estrus status.

700 *The Active Place Avoidance Task*

701 We used an active place avoidance paradigm to establish a hippocampal-, synaptic plasticity- and
702 PKMζ-dependent spatial memory that persists for months (Cimadevilla et al., 2000; Cimadevilla et al.,
703 2001; Pastalkova et al., 2006; Tsokas et al., 2016; Pavlowsky et al., 2017; Chung et al., 2021; Tsokas et
704 al., 2024). Briefly, an overhead digital video tracking system monitors the position of a mouse as it
705 moves on an elevated circular 40-cm diameter rotating (1 rpm) arena; the transparent arena wall
706 allows the mouse to observe the environment (Tracker, Bio-Signal Group Corp., Acton, MA). The
707 computer activates a mild constant current (0.2 mA, 500 ms, 60 Hz) foot shock across the parallel bars
708 of the floor whenever the mouse enters a 60° stationary sector that is designated to be the shock
709 zone. Mice rapidly learn to avoid the shock zone. Behavior in the arena is automatically quantified by
710 48 measures during each trial, as subset of which are used for formal analysis (Table S1; Tracker, Bio-
711 Signal Group Corp., Acton, MA).

712 Experiment 1 – Tracing Memory: Mice were removed from the isolation chamber and placed on the
713 behavioral arena and were subsequently returned to the chamber in a way that minimized novel or
714 distinctive extra-arena experience. Each mouse received a 30-min pretraining trial (Pretrain) with no
715 shock. After a 24-h interval back in isolation, each mouse received three 30-min training trials (Train
716 1-3), separated by 24-h intervals in isolation. Four Trained mice received memory training to avoid the
717 location of a mild foot shock in a 60° shock zone; these mice acquired a conditioned place avoidance.
718 Four other Untrained control mice received the identical experience except shock was never turned
719 on. Thirty minutes before Train 3, each animal was injected with TAM. This was intended to tag
720 conditioned avoidance memory-activated cells with ChR2-EYFP in the Trained group and control-
721 activated cells in the Untrained group. All mice were returned to their home cage for 30 days before
722 testing retention of the conditioned avoidance without shock.

723 Experiment 2 – Transcriptional Profiling of Memory: We used a three-day training protocol (Figs. 2a,
724 S2). On day 1, the mice (n = 34) were acclimated to the rotating arena during a 10-min pretraining
725 trial (P) with no shock. Each Trained mouse (n = 17) mouse then receives three 10-min training trials
726 (T1-T3), each separated by 2-h intervals, after which they are returned to their home cage. The next
727 day, the mice are retested (Rt), with the shock in the same location. This was followed by three
728 additional 10-min training trials (T4-T6), each separated by 2-h intervals. The shock zone remained in
729 the same location during T4-T6 for standard-trained animals (n = 8) but it was relocated 180° for the
730 conflict-trained mice (n = 9). The relocated position of shock coincides with the place the mice had
731 preferred to be during the T1-T3 trials (Fig. 2a). Each trained mouse (n's = 8, 9) was yoked to a Control
732 mouse (n's = 8, 9) that received the identical shock time series (Fig. 2b). Accordingly, Trained (n = 17)
733 and Control (n = 17) mice had the identical physical experiences of the environment, but the Trained
734 mice could learn to avoid the shock zone, and the Controls could not. On Day 3 all mice received a
735 memory retention trial (Rn) with no shock to evaluate the conditioned avoidance. The Trained and
736 Control mice were each divided into two groups that only differed during the T4-T6 trials on Day 2.

737 Littermates were randomly assigned to one of the 2 (Standard-Trained/Conflict-Trained) x 2
738 (Standard-Yoked/Conflict-Yoked) groups (Fig. 2a). Note that the shock is mild and compared to mice
739 on the arena but never shocked, we do not detect behavioral or plasma corticosterone expression
740 differences in these groups 24-h after the avoidance or yoked experience compared to the untrained
741 of walking on the rotating arena (Lesburgueres et al., 2016).

742 *Immunohistochemistry and confocal microscopy (Experiment 1)*

743 Detailed methods have been reported and are only briefly summarized here. Immediately after
744 testing retention, animals were anesthetized with 2% isoflurane for two minutes, and the brains were
745 fixed by cardiac perfusion then removed. Transverse 300-µm hippocampal slices were placed in ice-
746 cold 4% paraformaldehyde in 0.1 M phosphate buffer (PB, pH 7.4) and post-fixed for 48 h, as
747 previously described (Hsieh et al., 2021). Antibodies specific to EYFP (Abcam, #ab13970) and PKMζ
748 (C2 polyclonal antiserum from the Sacktor lab) were used. Slices were then washed with PBS (pH 7.4)
749 and 30-µm sections of the dorsal hippocampus were prepared using a Leica VT 1200S vibratome
750 (Leica Biosystems, Buffalo Grove, IL). Free-floating sections were permeabilized with PBS containing
751 0.1% Tween20 (PBS-T) for 1 h at room temperature and blocked with blocking buffer (BB; 10% normal
752 goat serum in PBS-T) for 2.5 h at room temperature. The sections were then incubated overnight at
753 4°C with primary antibody rabbit anti-PKMζ C-2 antisera (1:10,000) and chicken anti-EYFP (1:1,000) in
754 BB. After washing 6 times for 10 mins each in PBS-T, the sections were then incubated with the
755 secondary antibodies goat anti-chicken Alexa 488 and goat anti-rabbit-Alexa 647 (both 1:500 in BB;
756 Jackson ImmunoResearch, West Grove, PA) for 2 h at room temperature. After extensive washing
757 with PBS, the sections were mounted with 4′,6-diamidino-2-phenylindole (DAPI) Fluoromount-G
758 (Southern Biotech, Birmingham, AL) or Vectashield (Vector Laboratories, Newark, CA).

759 Three sections from each mouse were observed using an upright Leica SP8 confocal microscope (Leica
760 Biosystems, Buffalo Grove, IL) and analyzed using ImageJ (version 1.53a; Schneider et al., 2012). For

761 each section, 8.5 µm-thick Z-stacks of the dorsal CA1, CA2, CA3, supra-DG, and infra-DG were created
762 using the maximum intensity projection function in ImageJ. In each stack, two square regions of
763 interest were centered on *strata pyramidale*, *radiatum*, *lacunosum moleculare* for the CA1, CA2, and
764 CA3 subfields, and on the granule cell layer and molecular layer for both the supra- and infra-
765 pyramidal layers of DG. The molecular layer of the DG was further subdivided into three equal
766 compartments (inner, middle, and outer) along the dorsoventral axis. Measurements were made
767 from each mouse in each region of interest (ROI). The raw integrated density (defined as the sum of
768 the values for all pixels) of the Z-stack ROI expressing the fluorescent label was measured for the
769 green (EYFP), red (PKMζ), and yellow (overlap) volume of pixels, and the average of each of the 6 (2
770 ROI x 3 slices) measurements was taken as representative for the region for each mouse. These
771 measures account for the locations of the signal as well as their intensity. To rigorously compare the
772 overlap between EYFP and PKMζ signals in the Trained and Untrained brain regions, we computed the
773 Manders coefficients M1 and M2 using the JACOP plugin for ImageJ (Manders et al., 1993). M2
774 estimates how much PKMζ is in the memory-tagged cells by computing the proportion of the PKMζ
775 signal that co-localizes with the EYFP signal, while M1 estimates the proportion of EYFP signal that co-
776 localizes with the PKMζ signal. Based on prior work, the working hypothesis is that M2 will increase in
777 persistently potentiated populations of synapses in somato-dendritic compartments that are
778 maintained by a PKMζ-dependent mechanism (Pastalkova et al., 2006; Pavlowsky et al., 2017; Hsieh et
779 al., 2021).

780 *Preparation of micro-dissected hippocampal subfields (Experiment 2)*

781 A day after the last training session, and 30 min after the retention test without shock, mice were
782 anesthetized with 2% isoflurane and decapitated. Transverse 300-µm brain slices were prepared using
783 a vibratome (model VT1000 S, Leica Biosystems, Buffalo Grove, IL) and incubated at 36°C for 30 min
784 and then at room temperature for 90 min in oxygenated artificial cerebrospinal fluid (aCSF in mM:
785 125 NaCl, 2.5 KCl, 1 MgSO4, 2 CaCl2, 25 NaHCO3, 1.25 NaH2PO4, and 25 Glucose) as previously

786 prepared to measure LTP and place avoidance memory-induced synaptic plasticity (Pavlowsky and
787 Alarcon, 2012; Pavlowsky et al., 2017). Slices were cut in half so that one hemisphere could be used
788 for RNA-seq (see below) and one for *ex vivo* slice physiology (data not shown). The DG, CA3, and CA1
789 subfields were micro-dissected for later RNA extraction using a 0.25 mm punch tool and cold mat
790 (Electron Microscopy Sciences, Hatfield, PA) under a dissecting scope (Zeiss, Oberkochen, Germany;
791 Fig. 2c).

792 *RNA-sequencing and quality control (Experiment 2)*

793 RNA was isolated using the Maxwell 16 LEV RNA Isolation Kit (Promega, Madison, WI). Sequencing
794 libraries were prepared by the Genomic Sequencing and Analysis Facility at The University of Texas at
795 Austin and sequenced on the Illumina HiSeq platform following standard protocols. Reads were
796 processed on the Stampede cluster at the Texas Advanced Computing Center. The quality of raw and
797 filtered reads was checked using the program FASTQC (Wingett and Andrews, 2018) and visualized
798 using MultiQC (Ewels et al., 2016). We obtained 6.9 million ± 6.3 million reads per sample (Fig. S1a).
799 Next, we used Kallisto (Bray et al., 2016) to pseudo-align raw reads to a mouse reference
800 transcriptome (Gencode version 7; Waterston et al., 2002), which yielded 2.6 million ± 2.1 million
801 reads per sample (Fig. S1b). Mapping efficiency was about 42% (Fig. S1c).

802 *Statistical analyses*

803 To enhance statistical power and reduce the impact of low-abundance noise, we first applied gene-
804 level filtering prior to all statistical analyses. As the goals of differential expression and correlation
805 analyses differ, we employed distinct pre-processing strategies optimized for each.

806 For differential expression analysis, we followed DESeq2 guidelines to filter low-count genes by
807 retaining only those with a total count greater than 1 across all samples. Starting from 22,485 genes,
808 we retained 16,992, 16,481, and 16,840 genes in the DG, CA3, and CA1 subfields, respectively. In the

809 combined dataset, 17,911 genes remained. Variance-stabilizing transformation (VST) was applied
810 prior to downstream analyses to improve statistical inference.

811 In contrast, correlation-based analyses prioritized the inclusion of genes showing meaningful
812 variation. Gene expression values were normalized separately for each hippocampal subfield using
813 sample-wise size factor adjustment based on gene-wise geometric means. Genes were filtered to
814 exclude those with minimal expression variation, retaining only those expressed above the minimum
815 observed value in at least two samples. The final number of genes used for correlation analysis was
816 16,467 for DG, 15,778 for CA3, and 16,171 for CA1. Expression values were then log2-transformed to
817 stabilize variance and avoid undefined values during correlation calculations.

818 All statistical analyses were performed using R version 4.0.0 (2020-04-24)
819 (R_Development_Core_Team, 2013), and Python 3.12.6. For behavioral feature analysis, we use the
820 scikit-learn (v1.5.2; Pedregosa F, 2011), scipy (v 1.14.1; Virtanen et al., 2020) and Statsmodels
821 (v0.14.4; Seabold S, 2010) packages to conduct ANOVA and principal component analysis (PCA) . One-
822 and two-way ANOVAs were used to identify group differences in behavioral measures across one or
823 multiple trials, respectively (Stanley, 2018). PCA was conducted to reduce the dimensionality of the
824 data (Lê et al., 2008; Kassambara and Mundt, 2017). In particular, to define $PC1_{mem}$, we used
825 behavioral features from the retention phase, as they provided the best description of memory
826 performance. Transcript counts were imported into R and aggregated to yield gene counts using the
827 gene identifier from the Gencode transcriptome. Differential expression analysis was performed using
828 DESeq2 (Love et al., 2014) with support from tidyverse packages (Wickham, 2009; Wickham and
829 Grolemund, 2017). Genes were considered significant at a false discovery rate (FDR) adjusted p-value
830 threshold of < 0.1. Comparisons were made either among the four behavioral treatment groups
831 (standard-trained, standard-yoked, conflict-trained, and conflict-yoked) or between the combined
832 memory-trained and combined yoked-control groups.

Gene Ontology enrichment analysis was made using g:Profiler webserver/package (Kolberg et al., 2023) and MGI webserver (Baldarelli et al., 2024) to identify biological terms associated with hypothesized and differentially expressed genes.

For analysis of non-linear associations between gene expression values, as well as to $PC1_{mem}$ behavior, we defined Xicor (ξ) according to Chatterjee's method (Chatterjee, 2021). Since several methods exist for handling tied ranks, we tested multiple approaches and selected the 'dense' ranking method as default (Fig. S5). Given our data limited sample size, we applied a correction to the observed Xicor ($\xi_{observed}$) using a scaling model (Eq. 1) with sample size (N) and a scaling factor (c) derived from simulated linear data across varying sample sizes.

$$\xi_{scaled} = \frac{\xi_{observed}}{1-\frac{c}{N}} \qquad \text{(Eq. 1)}$$

The newly generated norm score (Σ) from the Spearman and Xicor correlation values follows Eq. 2.

$$\Sigma = \sqrt{{\xi_{scaled}}^2 + {\rho_s}^2} \qquad \text{(Eq. 2)}$$

To calculate the p-value of Σ, we normalized both the observed Xicor ($\xi_{observed}$)) and Spearman correlation, using their respective null variances:

$$z_\xi = \frac{\xi}{\sqrt{2/5n}} \ , \ z_{\rho_s} = \rho_s \cdot \sqrt{n-1} \qquad \text{(Eq. 3)}$$

Assuming z_ξ and z_{ρ_s} are independent standard normal variables, the combined statistic

$$z_\Sigma = \sqrt{{z_\xi}^2 + {z_{\rho_s}}^2} \qquad \text{(Eq. 4)}$$

follows a Rayleigh distribution. The p-value was computed as the upper tail probability of the Rayleigh cumulative distribution function (CDF). Genes with p-values ≤ 0.05 were used for Louvain clustering.

852 Clustering and network analyses were performed through NetworkX (v3.4.2; Hagberg et al., 2008),
853 and its visualization was generated in Cytoscape (v3.10.2; Shannon et al., 2003). For visualization in R,
854 we relied on the packages ggplot2 (Wickham, 2009), cowplot (Wilke, 2016), and png. We used
855 Matplotlib (v3.9.2; Hunter, 2007) and Seaborn (v0.13.2; Waskom, 2021) packages in python. Final
856 illustrations were created using Adobe Illustrator and Canvas X Draw.

857 *Data Availability*

858 Raw sequence data and differential gene expression data are available in NCBI's Gene Expression
859 Omnibus Database (accession: GSE100225). All data, code, and results are publicly available on
860 GitHub (https://github.com/sommerJY/APA_PKMZ).

861 ACKNOWLEDGEMENTS

862 Supported by Chan Zuckerberg Initiative grant CP2-1-0000000614 (AAF/SM), and NIH grants
863 R01MH132204 (AAF) and R01MH115304 (AAF/TCS). We are grateful to the Simons Center for
864 Computational Physical Chemistry at NYU (Simons Foundation grant 839534, MT), Helmsley
865 Foundation, a E.E. Just Endowed Research Fellowship, the Fries Trust Research Awards, Hartline
866 MacNichol Research Awards, Great Generation Fund for Research, MGF Fuortes Memorial Research
867 Awards, Frank R. Lillie Quasi-Endowment Fund, and L. & A. Colwin Summer Research Fellowship that
868 provided the support, without which these data could not be collected. This work was supported in
869 part through the NYU IT High Performance Computing resources, services, and staff expertise.

870 CONTRIBUTIONS

871 Immunohistochemistry data analysis and analysis (AG-P), bioinformatic data analysis and presentation
872 (JH, AP), bioinformatic data curation and QC (RMH), data collection and analysis (H-YK, RMH, HAH,
873 AAF), experimental design (HAH, JMA, AAF), figure preparation and data summary (JH, RMH), project

management (SM, HAH, AAF), funding (SM, AAF), data interpretation (SM, HAH, AAF), and writing (AAF, SM, TCS, HAH).

878 REFERENCES
879

880 Adams JP, Sweatt JD (2002) Molecular Psychology: Roles for the ERK MAP Kinase Cascade in
881 Memory. Annual Review of Pharmacology and Toxicology 42:135-163.
882 Adamsky A, Kol A, Kreisel T, Doron A, Ozeri-Engelhard N, Melcer T, Refaeli R, Horn H, Regev L,
883 Groysman M, London M, Goshen I (2018) Astrocytic Activation Generates De Novo Neuronal
884 Potentiation and Memory Enhancement. Cell 174:59-71.e14.
885 Alberini CM, Chen DY (2012) Memory enhancement: Consolidation, reconsolidation and insulin-like
886 growth factor 2. Trends in Neurosciences 35:274-283.
887 Allen NJ, Eroglu C (2017) Cell Biology of Astrocyte-Synapse Interactions. Neuron 96:697-708.
888 Asiedu MN, Tillu DV, Melemedjian OK, Shy A, Sanoja R, Bodell B, Ghosh S, Porreca F, Price TJ (2011)
889 Spinal Protein Kinase M ζ Underlies the Maintenance Mechanism of Persistent Nociceptive
890 Sensitization. The Journal of Neuroscience 31:6646.
891 Baldarelli RM, Smith CL, Ringwald M, Richardson JE, Bult CJ, Mouse Genome Informatics G (2024)
892 Mouse Genome Informatics: an integrated knowledgebase system for the laboratory mouse.
893 Genetics 227.
894 Barry JM, Rivard B, Fox SE, Fenton AA, Sacktor TC, Muller RU (2012) Inhibition of protein kinase
895 Mzeta disrupts the stable spatial discharge of hippocampal place cells in a familiar
896 environment. J Neurosci 32:13753-13762.
897 Bassell GJ, Warren ST (2008) Fragile X syndrome: loss of local mRNA regulation alters synaptic
898 development and function. Neuron 60:201-214.
899 Blondel VD, Guillaume J-L, Lambiotte R, Lefebvre E (2008) Fast unfolding of communities in large
900 networks. Journal of Statistical Mechanics: Theory and Experiment 2008:P10008.
901 Bray NL, Pimentel H, Melsted P, Pachter L (2016) Near-optimal probabilistic RNA-seq quantification.
902 Nat Biotechnol 34:525-527.
903 Cai D, Pearce K, Chen S, Glanzman DL (2011) Protein kinase M maintains long-term sensitization and
904 long-term facilitation in aplysia. J Neurosci 31:6421-6431.
905 Chatterjee S (2021) A New Coefficient of Correlation. Journal of the American Statistical Association
906 116:2009-2022.
907 Chen DY, Stern SA, Garcia-Osta A, Saunier-Rebori B, Pollonini G, Bambah-Mukku D, Blitzer RD,
908 Alberini CM (2011) A critical role for IGF-II in memory consolidation and enhancement.
909 Nature 469:491-497.
910 Chung A, Jou C, Grau-Perales A, Levy ERJ, Dvorak D, Hussain N, Fenton AA (2021) Cognitive control
911 persistently enhances hippocampal information processing. Nature 600:484-488.
912 Cimadevilla JM, Fenton AA, Bures J (2000) Functional inactivation of dorsal hippocampus impairs
913 active place avoidance in rats. Neurosci Lett 285:53-56.
914 Cimadevilla JM, Wesierska M, Fenton AA, Bures J (2001) Inactivating one hippocampus impairs
915 avoidance of a stable room-defined place during dissociation of arena cues from room cues
916 by rotation of the arena. Proc Natl Acad Sci U S A 98:3531-3536.
917 Colgin LL, Denninger T, Fyhn M, Hafting T, Bonnevie T, Jensen O, Moser MB, Moser EI (2009)
918 Frequency of gamma oscillations routes flow of information in the hippocampus. Nature
919 462:353-357.
920 Curreli S, Bonato J, Romanzi S, Panzeri S, Fellin T (2022) Complementary encoding of spatial
921 information in hippocampal astrocytes. PLOS Biology 20:e3001530.
922 Denny CA, Kheirbek MA, Alba EL, Tanaka KF, Brachman RA, Laughman KB, Tomm NK, Turi GF,
923 Losonczy A, Hen R (2014) Hippocampal memory traces are differentially modulated by
924 experience, time, and adult neurogenesis. Neuron 83:189-201.
925 Drier EA, Tello MK, Cowan M, Wu P, Blace N, Sacktor TC, Yin JC (2002) Memory enhancement and
926 formation by atypical PKM activity in Drosophila melanogaster. Nat Neurosci 5:316-324.
927 Dvorak D, Fenton AA (2014) On track with two gammas. Neuron 82:506-508.

928 Dvorak D, Chung A, Park EH, Fenton AA (2021) Dentate spikes and external control of hippocampal
929 function. Cell Rep 36:109497.
930 Dvorak D, Radwan B, Sparks FT, Talbot ZN, Fenton AA (2018) Control of recollection by slow gamma
931 dominating mid-frequency gamma in hippocampus CA1. PLoS Biol 16:e2003354.
932 Ebert DH, Greenberg ME (2013) Activity-dependent neuronal signalling and autism spectrum
933 disorder. Nature 493:327-337.
934 Ebitz RB, Hayden BY (2021) The population doctrine in cognitive neuroscience. Neuron.
935 El-Brolosy MA, Stainier DYR (2017) Genetic compensation: A phenomenon in search of mechanisms.
936 PLoS Genet 13:e1006780.
937 Eom T, Muslimov IA, Tsokas P, Berardi V, Zhong J, Sacktor TC, Tiedge H (2014) Neuronal BC RNAs
938 cooperate with eIF4B to mediate activity-dependent translational control. J Cell Biol
939 207:237-252.
940 Ewels P, Magnusson M, Lundin S, Kaller M (2016) MultiQC: summarize analysis results for multiple
941 tools and samples in a single report. Bioinformatics 32:3047-3048.
942 Farah CA, Dunn TW, Hastings MH, Ferguson L, Gao C, Gong K, Sossin WS (2019) A role for Numb in
943 Protein kinase M (PKM)-mediated increase in surface AMPA receptors during facilitation in
944 Aplysia. Journal of Neurochemistry 150:366-384.
945 Fenton AA (2024) Remapping revisited: how the hippocampus represents different spaces. Nature
946 Reviews Neuroscience.
947 Frankland PW, Josselyn SA (2013) Neuroscience: Memory and the single molecule. Nature 493:312-
948 313.
949 Fries P (2009) The model- and the data-gamma. Neuron 64:601-602.
950 Gao V, Suzuki A, Magistretti PJ, Lengacher S, Pollonini G, Steinman MQ, Alberini CM (2016)
951 Astrocytic beta2-adrenergic receptors mediate hippocampal long-term memory
952 consolidation. Proc Natl Acad Sci U S A 113:8526-8531.
953 Goldman MS, Golowasch J, Marder E, Abbott LF (2001) Global structure, robustness, and modulation
954 of neuronal models. J Neurosci 21:5229-5238.
955 Graber TE, McCamphill PK, Sossin WS (2013) A recollection of mTOR signaling in learning and
956 memory. Learn Mem 20:518-530.
957 Grienberger C, Magee JC (2022) Entorhinal cortex directs learning-related changes in CA1
958 representations. Nature.
959 Hagberg AA, Schult DA, Swart P, Hagberg J (2008) Exploring Network Structure, Dynamics, and
960 Function using NetworkX. Proceedings of the Python in Science Conference.
961 Hanley JG, Khatri L, Hanson PI, Ziff EB (2002) NSF ATPase and alpha-/beta-SNAPs disassemble the
962 AMPA receptor-PICK1 complex. Neuron 34:53-67.
963 Hardt O, Nadel L (2017) Systems consolidation revisited, but not revised: The promise and limits of
964 optogenetics in the study of memory. Neurosci Lett.
965 Hardt O, Migues PV, Hastings M, Wong J, Nader K (2009) PKMzeta maintains 1-day- and 6-day-old
966 long-term object location but not object identity memory in dorsal hippocampus.
967 Hippocampus 20:691-695.
968 Harris RM, Hofmann HA (2014) Neurogenomics of behavioral plasticity. Adv Exp Med Biol 781:149-
969 168.
970 Henneberger C, Papouin T, Oliet SHR, Rusakov DA (2010) Long-term potentiation depends on release
971 of d-serine from astrocytes. Nature 463:232-236.
972 Herculano-Houzel S, Mota B, Lent R (2006) Cellular scaling rules for rodent brains. Proceedings of the
973 National Academy of Sciences 103:12138.
974 Hernandez AI, Blace N, Crary JF, Serrano PA, Leitges M, Libien JM, Weinstein G, Tcherapanov A,
975 Sacktor TC (2003) Protein kinase M zeta synthesis from a brain mRNA encoding an
976 independent protein kinase C zeta catalytic domain. Implications for the molecular
977 mechanism of memory. J Biol Chem 278:40305-40316.

978 Hsieh C, Tsokas P, Serrano P, Hernandez AI, Tian D, Cottrell JE, Shouval HZ, Fenton AA, Sacktor TC
979 (2017) Persistent increased PKMzeta in long-term and remote spatial memory. Neurobiol
980 Learn Mem 138:135-144.
981 Hsieh C, Tsokas P, Grau-Perales A, Lesburgueres E, Bukai J, Khanna K, Chorny J, Chung A, Jou C,
982 Burghardt NS, Denny CA, Flores-Obando RE, Hartley BR, Rodriguez Valencia LM, Hernandez
983 AI, Bergold PJ, Cottrell JE, Alarcon JM, Fenton AA, Sacktor TC (2021) Persistent increases of
984 PKMzeta in memory-activated neurons trace LTP maintenance during spatial long-term
985 memory storage. Eur J Neurosci 54:6795-6814.
986 Hunter JD (2007) Matplotlib: A 2D Graphics Environment. Computing in Science & Engineering 9:90-
987 95.
988 Jain D, Jakhalekar IR, Deshmukh SS (2017) Navigational Strategies and Their Neural Correlates. J
989 Indian Inst Sci 97:511-525.
990 Josselyn SA, Kohler S, Frankland PW (2017) Heroes of the Engram. J Neurosci 37:4647-4657.
991 Kandel ER (2012) The molecular biology of memory: cAMP, PKA, CRE, CREB-1, CREB-2, and CPEB. Mol
992 Brain 5:14.
993 Kapur M, Monaghan CE, Ackerman SL (2017) Regulation of mRNA Translation in Neurons-A Matter
994 of Life and Death. Neuron 96:616-637.
995 Kassambara A, Mundt F (2017) factoextra: Extract and Visualize the Results of Multivariate Data
996 Analyses. In: Available at: https://cran.r-project.org/package=factoextra.
997 Kelly MT, Crary JF, Sacktor TC (2007a) Regulation of protein kinase Mzeta synthesis by multiple
998 kinases in long-term potentiation. J Neurosci 27:3439-3444.
999 Kelly MT, Yao Y, Sondhi R, Sacktor TC (2007b) Actin polymerization regulates the synthesis of
1000 PKMzeta in LTP. Neuropharmacology 52:41-45.
1001 Kolberg L, Raudvere U, Kuzmin I, Adler P, Vilo J, Peterson H (2023) g:Profiler-interoperable web
1002 service for functional enrichment analysis and gene identifier mapping (2023 update).
1003 Nucleic Acids Res 51:W207-W212.
1004 Lacagnina AF, Brockway ET, Crovetti CR, Shue F, McCarty MJ, Sattler KP, Lim SC, Santos SL, Denny CA,
1005 Drew MR (2019) Distinct hippocampal engrams control extinction and relapse of fear
1006 memory. Nat Neurosci 22:753-761.
1007 Laferriere A, Pitcher MH, Haldane A, Huang Y, Cornea V, Kumar N, Sacktor TC, Cervero F, Coderre TJ
1008 (2011) PKMzeta is essential for spinal plasticity underlying the maintenance of persistent
1009 pain. Mol Pain 7:99.
1010 Langfelder P, Horvath S (2008) WGCNA: an R package for weighted correlation network analysis.
1011 BMC Bioinformatics 9:559.
1012 Lê S, Josse J, Husson F (2008) FactoMineR: An R Package for Multivariate Analysis. Journal of
1013 Statistical Software 25:1 - 18.
1014 Lee AM, Kanter BR, Wang D, Lim JP, Zou ME, Qiu C, McMahon T, Dadgar J, Fischbach-Weiss SC,
1015 Messing RO (2013) Prkcz null mice show normal learning and memory. Nature 493:416-419.
1016 Lesburgueres E, Sparks FT, O'Reilly KC, Fenton AA (2016) Active place avoidance is no more stressful
1017 than unreinforced exploration of a familiar environment. Hippocampus 26:1481-1485.
1018 Levy ERJ, O'Reilly KC, Fenton AA (2019) Long-Lasting Input-Specific Experience-Dependent Changes
1019 of Hippocampus Synaptic Function Measured in the Anesthetized Rat. eNeuro 6.
1020 Levy ERJ, Carrillo-Segura S, Park EH, Redman WT, Hurtado JR, Chung S, Fenton AA (2023) A manifold
1021 neural population code for space in hippocampal coactivity dynamics independent of place
1022 fields. Cell Reports 42.
1023 Li W, Wang P, Li H (2014) Upregulation of glutamatergic transmission in anterior cingulate cortex in
1024 the diabetic rats with neuropathic pain. Neurosci Lett 568:29-34.
1025 Li Y, Briguglio JJ, Romani S, Magee JC (2024) Mechanisms of memory-supporting neuronal dynamics
1026 in hippocampal area CA3. Cell 187:6804-6819.e6821.

1027 Li YQ, Xue YX, He YY, Li FQ, Xue LF, Xu CM, Sacktor TC, Shaham Y, Lu L (2011) Inhibition of PKMzeta in
1028 nucleus accumbens core abolishes long-term drug reward memory. J Neurosci 31:5436-
1029 5446.
1030 Ling DS, Benardo LS, Serrano PA, Blace N, Kelly MT, Crary JF, Sacktor TC (2002) Protein kinase Mzeta
1031 is necessary and sufficient for LTP maintenance. Nat Neurosci 5:295-296.
1032 Lisman J (2017) Criteria for identifying the molecular basis of the engram (CaMKII, PKMzeta). Mol
1033 Brain 10:55.
1034 Lisman J, Buzsaki G, Eichenbaum H, Nadel L, Ranganath C, Redish AD (2017) Viewpoints: how the
1035 hippocampus contributes to memory, navigation and cognition. Nat Neurosci 20:1434-1447.
1036 Liu X, Ramirez S, Pang PT, Puryear CB, Govindarajan A, Deisseroth K, Tonegawa S (2012) Optogenetic
1037 stimulation of a hippocampal engram activates fear memory recall. Nature 484:381-385.
1038 Liu Y, Beyer A, Aebersold R (2016) On the Dependency of Cellular Protein Levels on mRNA
1039 Abundance. Cell 165:535-550.
1040 Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq
1041 data with DESeq2. Genome Biol 15:550.
1042 Madar AD, Jiang A, Dong C, Sheffield MEJ (2025) Synaptic plasticity rules driving representational
1043 shifting in the hippocampus. Nature Neuroscience 28:848-860.
1044 Madronal N, Gruart A, Sacktor TC, Delgado-Garcia JM (2010) PKMzeta inhibition reverses learning-
1045 induced increases in hippocampal synaptic strength and memory during trace eyeblink
1046 conditioning. PLoS One 5:e10400.
1047 Mahringer D, Zmarz P, Okuno H, Bito H, Keller GB (2022) Functional correlates of immediate early
1048 gene expression in mouse visual cortex. Peer Community J 2:e45.
1049 Makuch L, Volk L, Anggono V, Johnson RC, Yu Y, Duning K, Kremerskothen J, Xia J, Takamiya K,
1050 Huganir RL (2011) Regulation of AMPA receptor function by the human memory-associated
1051 gene KIBRA. Neuron 71:1022-1029.
1052 Manders EMM, Verbeek FJ, Aten JA (1993) Measurement of co-localization of objects in dual-colour
1053 confocal images. Journal of Microscopy 169:375-382.
1054 Marder E, O'Leary T, Shruti S (2014) Neuromodulation of circuits with variable parameters: single
1055 neurons and small circuits reveal principles of state-dependent and robust
1056 neuromodulation. Annu Rev Neurosci 37:329-346.
1057 Marder E, Goeritz ML, Otopalik AG (2015) Robust circuit rhythms in small circuits arise from variable
1058 circuit components and mechanisms. Curr Opin Neurobiol 31:156-163.
1059 Marr D (1971) Simple memory: a theory for archicortex. Philos Trans R Soc Lond B Biol Sci 262:23-81.
1060 Migues PV, Hardt O, Wu DC, Gamache K, Sacktor TC, Wang YT, Nader K (2010) PKMzeta maintains
1061 memories by regulating GluR2-dependent AMPA receptor trafficking. Nat Neurosci.
1062 Miller-Crews I, Hofmann HA (2024) Single-nuclei transcriptomes in the hypothalamus and POA of a
1063 highly social cichlid. bioRxiv:2024.2011.2007.622449.
1064 Morris RG (2016) Forget me not. Elife 5.
1065 Muslimov IA, Nimmrich V, Hernandez AI, Tcherepanov A, Sacktor TC, Tiedge H (2004) Dendritic
1066 transport and localization of protein kinase Mzeta mRNA: implications for molecular
1067 memory consolidation. J Biol Chem 279:52613-52622.
1068 Neunuebel JP, Yoganarasimha D, Rao G, Knierim JJ (2013) Conflicts between local and global spatial
1069 frameworks dissociate neural representations of the lateral and medial entorhinal cortex. J
1070 Neurosci 33:9246-9258.
1071 Nilssen ES, Doan TP, Nigro MJ, Ohara S, Witter MP (2019) Neurons and networks in the entorhinal
1072 cortex: A reappraisal of the lateral and medial entorhinal subdivisions mediating parallel
1073 cortical pathways. Hippocampus 29:1238-1254.
1074 Northcutt AJ, Kick DR, Otopalik AG, Goetz BM, Harris RM, Santin JM, Hofmann HA, Marder E, Schulz
1075 DJ (2019) Molecular profiling of single neurons of known identity in two ganglia from the
1076 crab Cancer borealis. Proc Natl Acad Sci U S A.

1077 O'Connell LA, Matthews BJ, Hofmann HA (2012) Isotocin regulates paternal care in a monogamous
1078 cichlid fish. Horm Behav 61:725-733.
1079 O'Keefe J, Nadel L (1978) The hippocampus as a cognitive map. Oxford: Clarendon Press.
1080 Pace TWW, Gaylord R, Topczewski F, Girotti M, Rubin B, Spencer RL (2005) Immediate–early gene
1081 induction in hippocampus and cortex as a result of novel experience is not directly related to
1082 the stressfulness of that experience. European Journal of Neuroscience 22:1679-1690.
1083 Park EH, Burghardt NS, Dvorak D, Hen R, Fenton AA (2015) Experience-Dependent Regulation of
1084 Dentate Gyrus Excitability by Adult-Born Granule Cells. J Neurosci 35:11656-11666.
1085 Parpura V, Basarsky TA, Liu F, Jeftinija K, Jeftinija S, Haydon PG (1994) Glutamate-mediated
1086 astrocyte-neuron signalling. Nature 369:744-747.
1087 Parras A, Anta H, Santos-Galindo M, Swarup V, Elorza A, Nieto-Gonzalez JL, Pico S, Hernandez IH,
1088 Diaz-Hernandez JI, Belloc E, Rodolosse A, Parikshak NN, Penagarikano O, Fernandez-Chacon
1089 R, Irimia M, Navarro P, Geschwind DH, Mendez R, Lucas JJ (2018) Autism-like phenotype and
1090 risk gene mRNA deadenylation by CPEB4 mis-splicing. Nature 560:441-446.
1091 Pastalkova E, Serrano P, Pinkhasova D, Wallace E, Fenton AA, Sacktor TC (2006) Storage of spatial
1092 information by the maintenance mechanism of LTP. Science 313:1141-1144.
1093 Pavlowsky A, Alarcon JM (2012) Interaction between long-term potentiation and depression in CA1
1094 synapses: temporal constrains, functional compartmentalization and protein synthesis. PLoS
1095 One 7:e29865.
1096 Pavlowsky A, Wallace E, Fenton AA, Alarcon JM (2017) Persistent modifications of hippocampal
1097 synaptic function during remote spatial memory. Neurobiol Learn Mem 138:182-197.
1098 Pedregosa F VG, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R,
1099 Dubourg V, Vanderplas J. (2011) Scikit-learn: Machine learning in Python. . The Journal of
1100 machine Learning research 12:2825-2830.
1101 Pevzner A, Guzowski JF (2014) Immediate-early gene transcriptional activation in hippocampus CA1
1102 and CA3 does not accurately reflect rapid, pattern completion-based retrieval of context
1103 memory. Learn Mem 22:1-5.
1104 Plato, Cornford FM, Plato (1935) Plato's theory of knowledge; the Theaetetus and the Sophist of
1105 Plato. London, New York,: K. Paul, Trench, Harcourt, Brace and company.
1106 R_Development_Core_Team (2013) R: A language and environment for statistical computing. In: R
1107 Foundation for Statistical Computing. Available at: http://www.r-project.org (Team RDC, ed).
1108 Vienna, Austria.
1109 Refaeli R, Kreisel T, Groysman M, Adamsky A, Goshen I (2023) Engram stability and maturation
1110 during systems consolidation. Curr Biol 33:3942-3950.e3943.
1111 Rossetti T, Banerjee S, Kim C, Leubner M, Lamar C, Gupta P, Lee B, Neve R, Lisman J (2017) Memory
1112 Erasure Experiments Indicate a Critical Role of CaMKII in Memory Storage. Neuron 96:207-
1113 216 e202.
1114 Rusakov DA (2015) Disentangling calcium-driven astrocyte physiology. Nature Reviews Neuroscience
1115 16:226-233.
1116 Sacktor TC (2011) How does PKMzeta maintain long-term memory? Nat Rev Neurosci 12:9-15.
1117 Sacktor TC (2012) Memory maintenance by PKMzeta--an evolutionary perspective. Mol Brain 5:31.
1118 Sacktor TC, Hell JW (2017) The genetics of PKMzeta and memory maintenance. Sci Signal 10.
1119 Sacktor TC, Fenton AA (2018) What does LTP tell us about the roles of CaMKII and PKMzeta in
1120 memory? Mol Brain 11:77.
1121 Sacktor TC, Osten P, Valsamis H, Jiang X, Naik MU, Sublette E (1993) Persistent activation of the zeta
1122 isoform of protein kinase C in the maintenance of long-term potentiation. Proc Natl Acad Sci
1123 U S A 90:8342-8346.
1124 Sanes JR, Lichtman JW (1999) Can molecules explain long-term potentiation? Nat Neurosci 2:597-
1125 604.
1126 Schneider CA, Rasband WS, Eliceiri KW (2012) NIH Image to ImageJ: 25 years of image analysis.
1127 Nature Methods 9:671-675.

1128 Schneidman E, Berry MJ, 2nd, Segev R, Bialek W (2006) Weak pairwise correlations imply strongly
1129 correlated network states in a neural population. Nature 440:1007-1012.
1130 Seabold S PJ (2010) Statsmodels: econometric and statistical modeling with python. SciPy 7:92-96.
1131 Serrano P, Friedman EL, Kenney J, Taubenfeld SM, Zimmerman JM, Hanna J, Alberini C, Kelley AE,
1132 Maren S, Rudy JW, Yin JC, Sacktor TC, Fenton AA (2008) PKMzeta maintains spatial,
1133 instrumental, and classically conditioned long-term memories. PLoS Biol 6:2698-2706.
1134 Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T
1135 (2003) Cytoscape: a software environment for integrated models of biomolecular interaction
1136 networks. Genome Res 13:2498-2504.
1137 Shao CY, Sondhi R, van de Nes PS, Sacktor TC (2012) PKMzeta is necessary and sufficient for synaptic
1138 clustering of PSD-95. Hippocampus 22:1501-1507.
1139 Shema R, Sacktor TC, Dudai Y (2007) Rapid erasure of long-term memory associations in the cortex
1140 by an inhibitor of PKMzeta. Science 317:951-953.
1141 Shema R, Haramati S, Ron S, Hazvi S, Chen A, Sacktor TC, Dudai Y (2011) Enhancement of
1142 consolidated long-term memory by overexpression of protein kinase Mzeta in the
1143 neocortex. Science 331:1207-1210.
1144 Sinha S, Jones BM, Traniello IM, Bukhari SA, Halfon MS, Hofmann HA, Huang S, Katz PS, Keagy J,
1145 Lynch VJ, Sokolowski MB, Stubbs LJ, Tabe-Bordbar S, Wolfner MF, Robinson GE (2020)
1146 Behavior-related gene regulatory networks: A new level of organization in the brain.
1147 Proceedings of the National Academy of Sciences 117:23270-23279.
1148 Stanley D (2018) apaTables: Create American Psychological Association (APA) Style Tables. In:
1149 Available at: https://cran.r-project.org/package=apaTables.
1150 Sutherland RJ, O'Brien J, Lehmann H (2008) Absence of systems consolidation of fear memories after
1151 dorsal, ventral, or complete hippocampal damage. Hippocampus 18:710-718.
1152 Suzuki A, Stern SA, Bozdagi O, Huntley GW, Walker RH, Magistretti PJ, Alberini CM (2011) Astrocyte-
1153 neuron lactate transport is required for long-term memory formation. Cell 144:810-823.
1154 Takeuchi T, Duszkiewicz AJ, Morris RG (2013) The synaptic plasticity and memory hypothesis:
1155 encoding, storage and persistence. Philos Trans R Soc Lond B Biol Sci 369:20130288.
1156 Talbot ZN, Sparks FT, Dvorak D, Curran BM, Alarcon JM, Fenton AA (2018) Normal CA1 Place Fields
1157 but Discoordinated Network Discharge in a Fmr1-Null Mouse Model of Fragile X Syndrome.
1158 Neuron 97:684-697.
1159 Thomson SR, Seo SS, Barnes SA, Louros SR, Muscas M, Dando O, Kirby C, Wyllie DJA, Hardingham GE,
1160 Kind PC, Osterweil EK (2017) Cell-Type-Specific Translation Profiling Reveals a Novel Strategy
1161 for Treating Fragile X Syndrome. Neuron 95:550-563 e555.
1162 Tian D, Dmitrieva RI, Doris PA, Crary JF, Sondhi R, Sacktor TC, Bergold PJ (2008) Protein kinase M zeta
1163 regulation of Na/K ATPase: a persistent neuroprotective mechanism of ischemic
1164 preconditioning in hippocampal slice cultures. Brain Res 1213:127-139.
1165 Tonegawa S, Morrissey MD, Kitamura T (2018) The role of engram cells in the systems consolidation
1166 of memory. Nature Reviews Neuroscience 19:485-498.
1167 Treves A, Rolls ET (1992) Computational constraints suggest the need for two distinct input systems
1168 to the hippocampal CA3 network. Hippocampus 2:189-199.
1169 Tsokas P, Hsieh C, Flores-Obando RE, Bernabo M, Tcherepanov A, Hernandez AI, Thomas C, Bergold
1170 PJ, Cottrell JE, Kremerskothen J, Shouval HZ, Nader K, Fenton AA, Sacktor TC (2024) KIBRA
1171 anchoring the action of PKMzeta maintains the persistence of memory. Sci Adv 10:eadl0030.
1172 Tsokas P, Hsieh C, Yao Y, Lesburgueres E, Wallace EJ, Tcherepanov A, Jothianandan D, Hartley BR,
1173 Pan L, Rivard B, Farese RV, Sajan MP, Bergold PJ, Hernandez AI, Cottrell JE, Shouval HZ,
1174 Fenton AA, Sacktor TC (2016) Compensation for PKMzeta in long-term potentiation and
1175 spatial long-term memory in mutant mice. Elife 5.
1176 Uytiepo M et al. (2025) Synaptic architecture of a memory engram in the mouse hippocampus.
1177 Science 387:eado8316.

1178 Velez-Hernandez ME, Vazquez-Torres R, Velasquez-Martinez MC, Jimenez L, Baez F, Sacktor TC,
1179 Jimenez-Rivera CA (2013) Inhibition of Protein kinase Mzeta (PKMzeta) in the mesolimbic
1180 system alters cocaine sensitization in rats. J Drug Alcohol Res 2:235669.
1181 Virtanen P et al. (2020) SciPy 1.0: fundamental algorithms for scientific computing in Python. Nat
1182 Methods 17:261-272.
1183 Vogel C, Marcotte EM (2012) Insights into the regulation of protein abundance from proteomic and
1184 transcriptomic analyses. Nat Rev Genet 13:227-232.
1185 Vogt-Eisele A, Krüger C, Duning K, Weber D, Spoelgen R, Pitzer C, Plaas C, Eisenhardt G, Meyer A,
1186 Vogt G, Krieger M, Handwerker E, Wennmann DO, Weide T, Skryabin BV, Klugmann M,
1187 Pavenstädt H, Huentelmann MJ, Kremerskothen J, Schneider A (2014) KIBRA (KIdney/BRAin
1188 protein) regulates learning and memory and stabilizes Protein kinase Mζ. Journal of
1189 Neurochemistry 128:686-700.
1190 Volk LJ, Bachman JL, Johnson R, Yu Y, Huganir RL (2013) PKM-zeta is not required for hippocampal
1191 synaptic plasticity, learning and memory. Nature 493:420-423.
1192 von Kraus LM, Sacktor TC, Francis JT (2010) Erasing sensorimotor memories via PKMzeta inhibition.
1193 PLoS One 5:e11125.
1194 Wang C, Lee H, Rao G, Doreswamy Y, Savelli F, Knierim JJ (2023) Superficial-layer versus deep-layer
1195 lateral entorhinal cortex: Coding of allocentric space, egocentric space, speed, boundaries,
1196 and corners. Hippocampus 33:448-464.
1197 Wang DC, Liu PC, Hung HS, Chen TJ (2014) Both PKMzeta and KIBRA are closely related to reference
1198 memory but not working memory in a T-maze task in rats. J Comp Physiol A Neuroethol Sens
1199 Neural Behav Physiol 200:77-82.
1200 Wang S, Sheng T, Ren S, Tian T, Lu W (2016) Distinct Roles of PKCiota/lambda and PKMzeta in the
1201 Initiation and Maintenance of Hippocampal Long-Term Potentiation and Memory. Cell Rep
1202 16:1954-1961.
1203 Waskom ML (2021) seaborn: statistical data visualization. Journal of Open Source Software 6:3021.
1204 Waterston RH, Lander ES, Sulston JE (2002) On the sequencing of the human genome. Proc Natl
1205 Acad Sci U S A 99:3712-3716.
1206 Weng F-J, Garcia RI, Lutzu S, Alviña K, Zhang Y, Dushko M, Ku T, Zemoura K, Rich D, Garcia-
1207 Dominguez D, Hung M, Yelhekar TD, Sørensen AT, Xu W, Chung K, Castillo PE, Lin Y (2018)
1208 Npas4 Is a Critical Regulator of Learning-Induced Plasticity at Mossy Fiber-CA3 Synapses
1209 during Contextual Memory Formation. Neuron 97:1137-1152.e1135.
1210 Whitlock JR, Heynen AJ, Shuler MG, Bear MF (2006) Learning induces long-term potentiation in the
1211 hippocampus. Science 313:1093-1097.
1212 Wickham H (2009) ggplot2: Elegant Graphics for Data Analysis. New York: Springer-Verlag.
1213 Wickham H, Grolemund G (2017) R for Data Science: Import, Tidy, Transform, Visualize, and Model
1214 Data., 1st edn. Edition: O'Reilly Media, Inc.
1215 Wilke CO (2016) cowplot: Streamlined Plot Theme and Plot Annotations for "ggplot2". In: Available
1216 at: https://cran.r-project.org/package=cowplot.
1217 Wilson MA, Tonegawa S (1997) Synaptic plasticity, place cells and spatial memory: study with second
1218 generation knockouts. Trends Neurosci 20:102-106.
1219 Wingett SW, Andrews S (2018) FastQ Screen: A tool for multi-genome mapping and quality control.
1220 F1000Res 7:1338.
1221 Wright BW, Yi Z, Weissman JS, Chen J (2022) The dark proteome: translation from noncanonical
1222 open reading frames. Trends in Cell Biology 32:243-258.
1223 Yao Y, Kelly MT, Sajikumar S, Serrano P, Tian D, Bergold PJ, Frey JU, Sacktor TC (2008) PKM zeta
1224 maintains late long-term potentiation by N-ethylmaleimide-sensitive factor/GluR2-
1225 dependent trafficking of postsynaptic AMPA receptors. J Neurosci 28:7820-7827.
1226 Yu NK, Uhm H, Shim J, Choi JH, Bae S, Sacktor TC, Hohng S, Kaang BK (2017) Increased PKMzeta
1227 activity impedes lateral movement of GluA2-containing AMPA receptors. Mol Brain 10:56.
1228 Yuste R (2015) From the neuron doctrine to neural networks. Nat Rev Neurosci 16:487-497.

1229

1230

FIGURES and LEGENDS

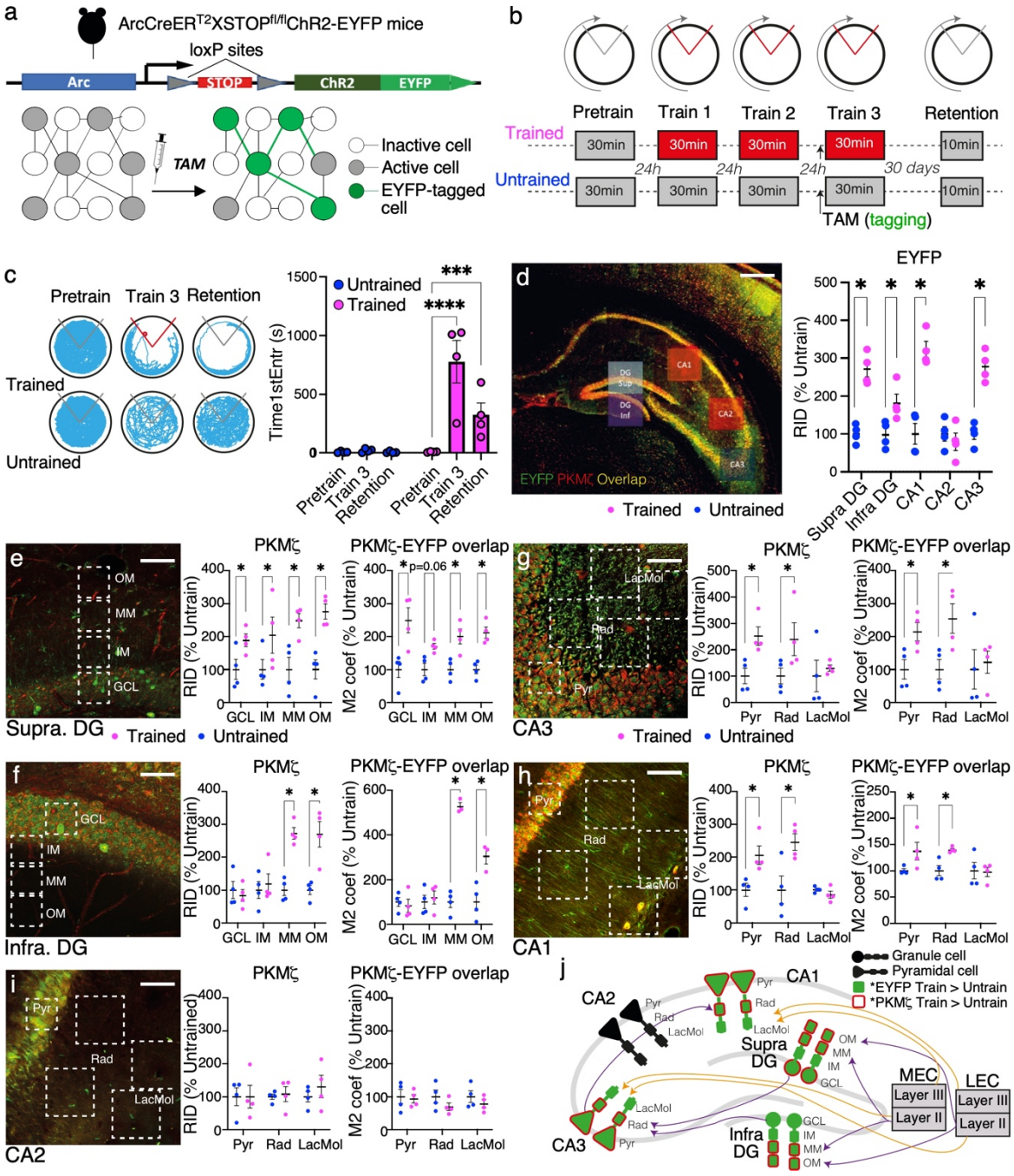


Fig. 1. Persistently elevated PKMζ protein expression in dorsal hippocampus subcellular compartments that trace the trisynaptic pathway circuit after 1-mo active place avoidance memory. a) Schematic of the TRAP2 mouse that was engineered to permanently express the EYFP-ChR2 tag in neurons that activate the immediate early gene *Arc* that is expressed primarily in excitatory neurons. b) Schematic of the behavioral experimental design. After learning an active place avoidance, EYFP-ChR2 tagging was initiated by systemic injection of 4OH-tamoxifen (TAM) 30 min prior to the third training session of the place avoidance memory and untrained control groups. Thirty days later, memory retention is tested prior to euthanasia and collection of tissue for PKMζ immunohistochemistry (n's = 4). c) Example of spatial exploration on the arena for exemplar Trained and Untrained mice to illustrates the conditioned behavior, the summary plot quantitates the active

place avoidance. d) Merged EYFP-ChR2 (green) and PKMζ (red) immunofluorescent image of dorsal hippocampus of a Trained mouse, with ROIs for quantifying Training 3 memory activation by expression of EYFP-ChR2 and memory retention by PKMζ expression computed as the raw integrated density (RID), each normalized by the control average. e) Summary of quantitation of EYFP expression. f-i) Example ROI samples targeting specific subcellular tissue compartments of the dorsal hippocamous circuit. Plots summarize control-normalized RID quantitation of PKMζ expression in each subcompartment and co-expression of EYFP and PKMζ, estimated by the control-normalized Mander's overlap coefficient (M2), which computes the ratio of PKMζ expression (RID) at image elements that express EYFP. j) Graphical summary illustrating the hippocampal circuit compartments that are persistently altered by increased PKMζ protein expression 1 month after memory training. Scale bars in panel d: 300 µm; e-I: 50 µm.

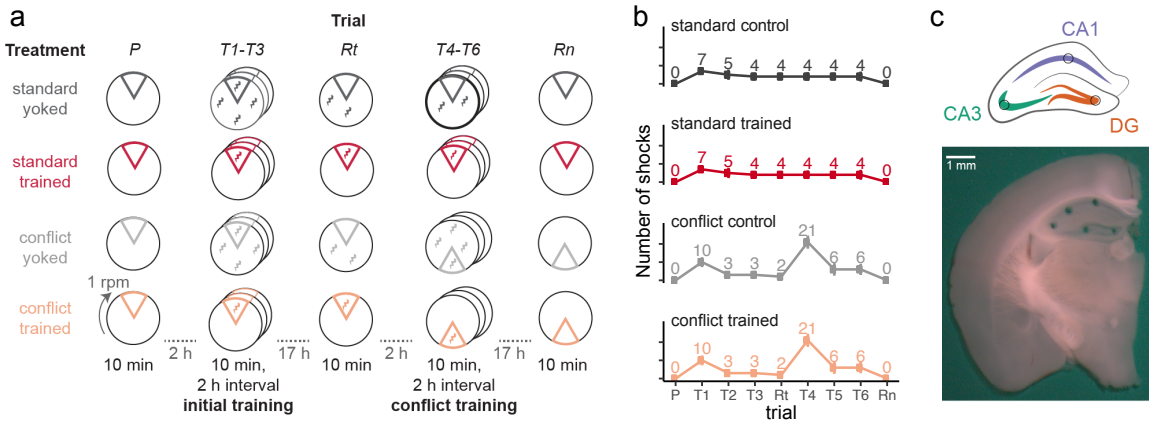


Fig. 2 Experimental Design. a) WT mice are assigned to one of four groups: standard-trained (red, n=8), standard-yoked (dark grey, n=8), conflict-trained (peach, n=9), or conflict-yoked (light grey, n=9). Mice were placed on the rotating arena for 10-min trials separated by 2-h or overnight (~17 h) intervals. Behavior was tracked during the pretraining (P), initial training (T1-T3), retest (Rt), conflict training (T4-T6), and retention (Rn) trials. In the trial schematics, the 60° sector represents the shock zone or for the yoked groups, the equivalent region that is used for evaluating place avoidance. Whereas trained mice receive shocks only in the shock zone, yoked mice experience shock throughout the environment. b) Standard-trained and standard-yoked mice received an identical number of shocks (and time series), with the most shocks occurring on trial 1. Conflict-trained and conflict-yoked mice received an identical number of shocks (and time series), with the most shocks occurring on trial 4, the first trial of conflict training. c) A schematic representation and a photo show the size and location of tissue samples collected from the dorsal hippocampus for RNA-sequencing. Tissues from five subfields (CA1, CA2, CA3, CA4, DG) were collected, but only the DG (orange), CA3 (teal), and CA1 (purple) subfields were processed for RNA-seq.

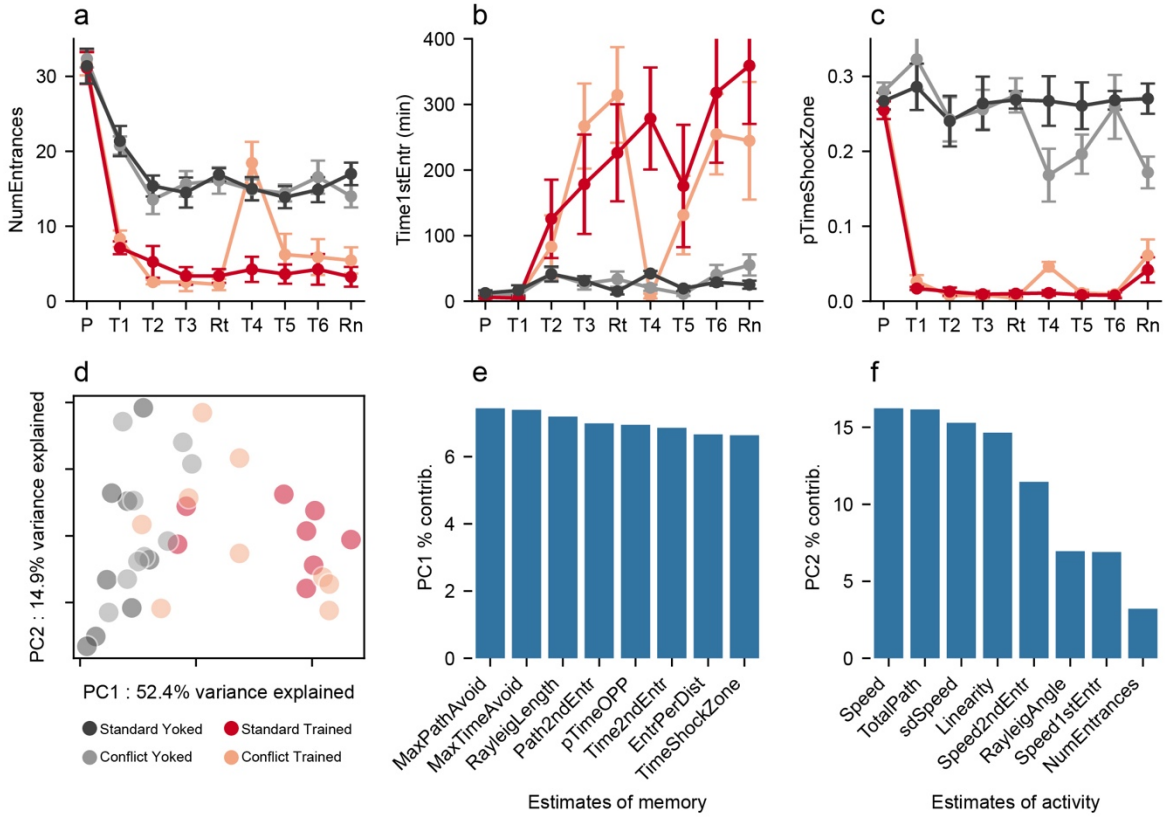


Fig. 3. Characterizing and comparing behavior amongst the groups. Standard-trained and conflict-trained groups rapidly learned to avoid the shock zone, whereas the yoked control groups did not avoid the corresponding region measured by a) reduction in the number of entrances into the shock zone, b) increased latency to enter the shock zone for the first time on each trial, and c) reduction in the proportion of time spent in the shock zone and its equivalent. d) PCA performed on the 26 measures of behavior during retention, results in PC1 that distinguishes the Trained and Control groups just prior to sacrifice when physical conditions are identical for all the mice because there is no shock. Large dots indicate group averages on the retention test while transparency is highest for pretraining and lowest for retention. Parameter loadings for e) PC1 and f) PC2 indicate that PC1 is comprised of estimates of place avoidance and PC2 estimates of locomotor behavior. Comparison statistics are provided in Table 1. MaxPathAvoid = maximum distance walked without entering the shock zone; MaxTimeAvoid = maximum time (s) spent out of the shock zone; RayleighLength = the Rayleigh vector length estimating the angular bias of the animal's dwell probability; Path2ndEntr = distance (m) walked before entering the shock zone twice; pTimeOPP = proportion of time spent in the quadrant opposite from the target quadrant; Time2ndEntr = amount of time it took the animal to enter the shock zone for the second time; EntrPerDist = number of entrances per meter walked; TimeShockZone = total time (s) spent in the shock zone; Speed = Average speed of the animal throughout the session; TotalPath = total distance (m) walked; sdSpeed = standard deviation of the speed estimates; Linearity = straightness of locomotion, estimated every 2 seconds as the ratio of the end-to-end Euclidean distance divided by the integral of the path segments; Speed2ndEntr = speed when entering the shock zone the second time; RayleighAngle = angle of the vector describing where the animal spent most of its time (the angle describes the particular preferred location. Ranges 0-

360); Speed1stEntr = speed when entering the shock zone the first time; NumEntrances = total number of entrances into the shock zone.

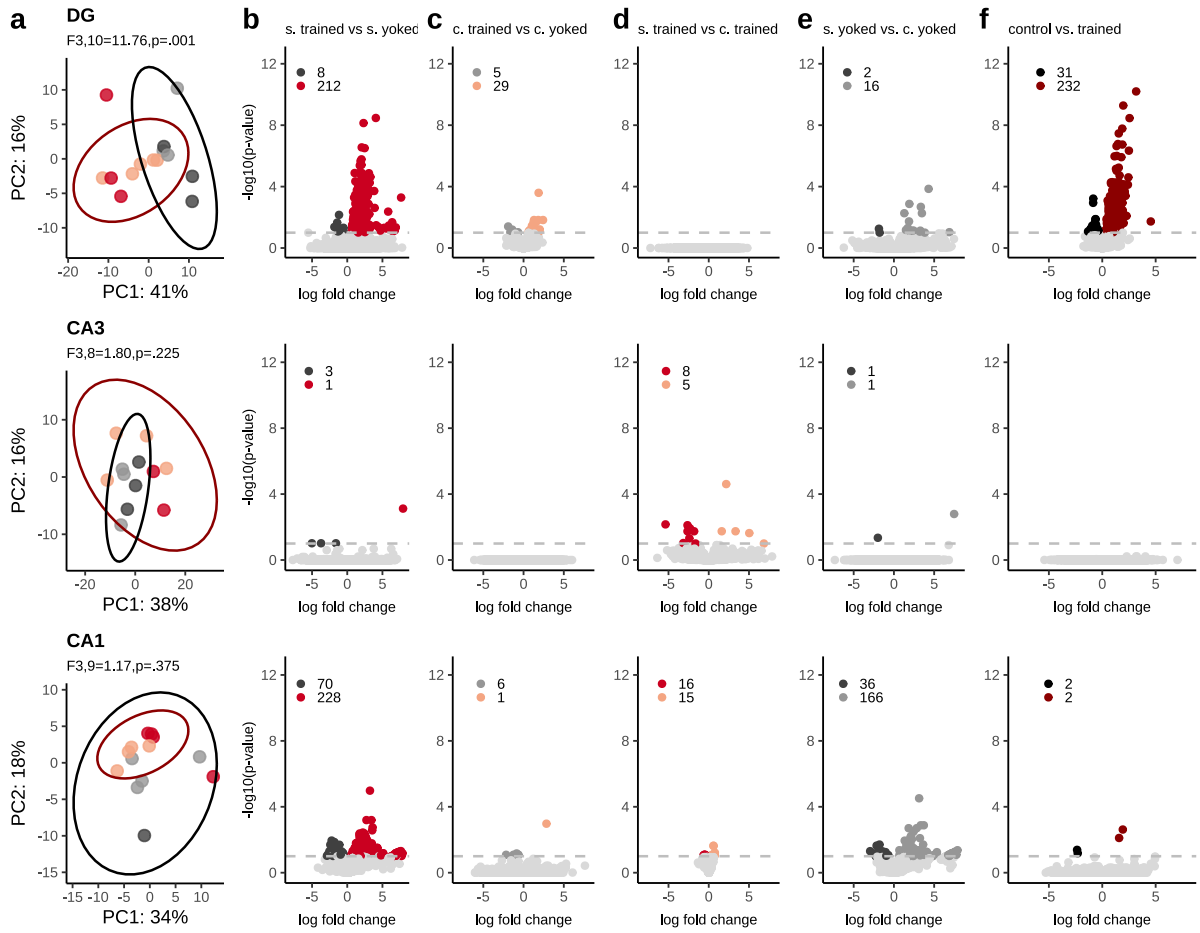
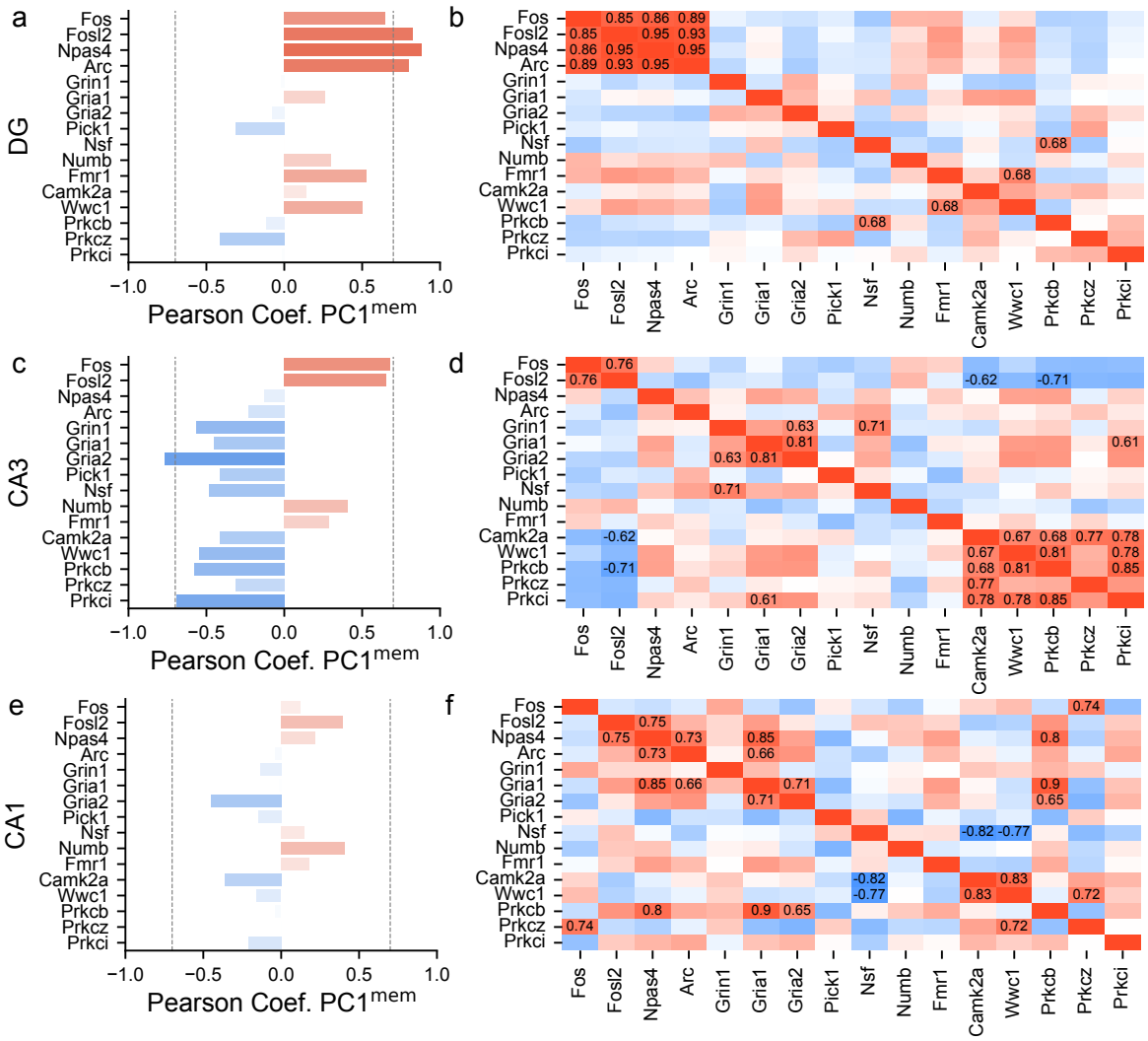


Fig. 4. Subfield and treatment-specific gene expression differences. a) PCA can only partly distinguish trained and control DG transcriptomes. Volcano plots illustrating the significant and non-significant differences in gene expression when b) the standard trained and yoked groups, c) the conflict and yoked groups, d) the two trained groups, e) the two yoked groups, and f) the combined trained and yoked groups are compared. Rows show gene expression differences by region from DG where information enters the hippocampus, to CA3 and CA1, the output of the hippocampus. Each dot represents a gene whose color corresponds to whether its expression increased, decreased, or was not significantly different between the comparison pair. The total number of differentially expressed genes (DEGs) are reported in the top left corner of each panel. The log-fold change, adjusted p-value, log-transformed adjusted p-value, and directionality of significant gene expression in the DG, CA3, and CA1 are available here: DEGlist.



1307

1308 **Fig. 5. Hippocampal subfield-specific LTP and memory-associated candidate gene analysis.**
1309 Pearson correlations of the expression levels between behavioral memory and candidate gene
1310 expression, as well as between pairs of candidate genes in a,b) DG, c,d) CA3, and e,f) CA1. Values are
1311 given for significant correlations (Absolute score ≥ 0.6). The correlation for the full set of analyzed
1312 genes is available here: PC1_DEGCorrData.

1313

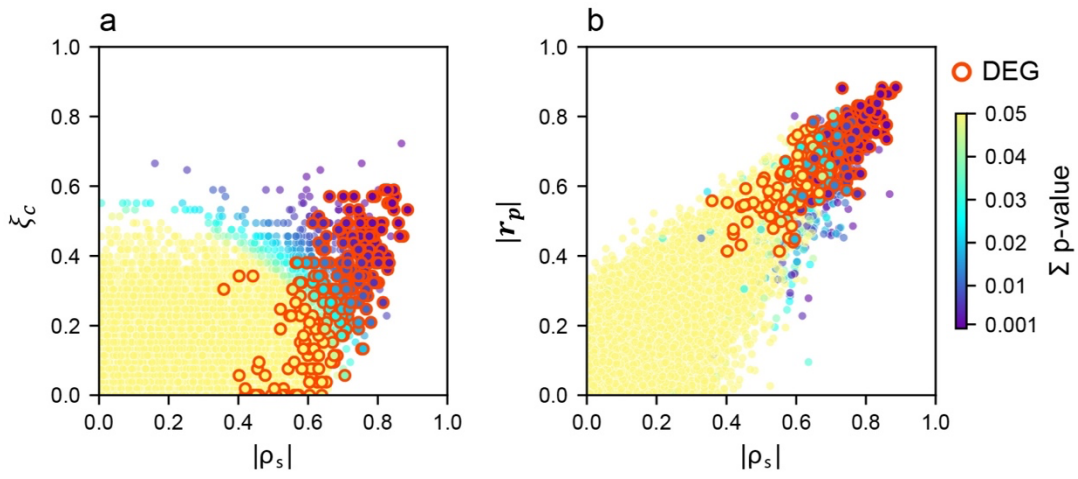


Fig. 6. Comparing Pearson (r_p, linear), and non-linear Spearman (ρ_s), and Chatterjee ξ_c correlations of gene expression and memory. The relationship between gene expression and $PC1_{mem}$ was computed by the three correlation measures. Their significance ($p \leq 0.05$) according to the norm of both ξ and ρ is color-coded. Comparison of a) Spearman and ξ correlations, and b) Spearman and Pearson correlations.

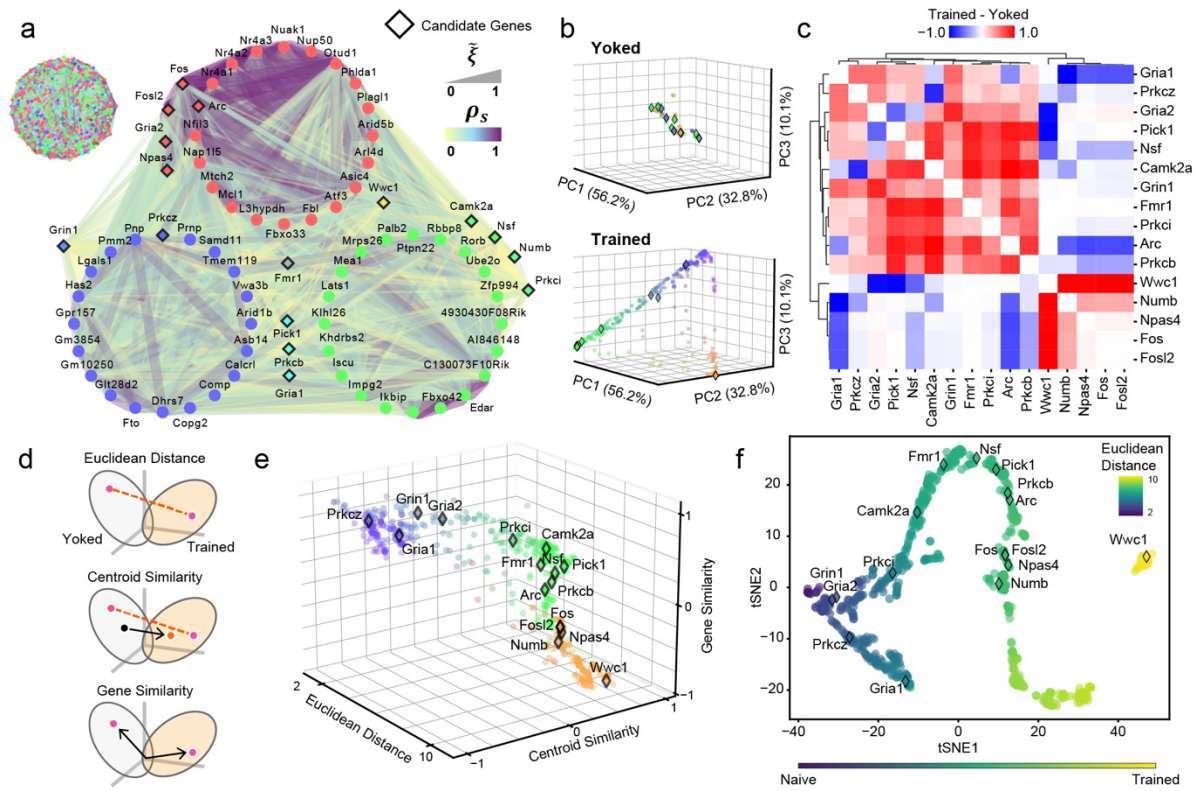


Fig. 7. Three robust communities of gene-gene correlations detected by the Louvain Method of community detection. 714 genes were organized into communities by 1000 iterations of the Louvain Method of community detection, each seeded with a different random set of genes. The community in which *Arc* is a member is always designated 'red', the community in which *Nsf* is a member is always designated 'green,' and the community in which *Prkcz* is a member is always designated 'blue.' Genes assigned to another community are designated 'gray.' After 1000 runs each gene has a red,

green, and blue score with brightness set by the gray score. a - left) gene network with community color-coded genes and edges coded according to correlation values. a - right) The top 20 genes in each community and some candidate genes are organized according to their community color-code and the set of correlation values. The candidate genes are marked by the diamond outline. b) Separate yoked and trained 714-D clustering probability matrices were computed to estimate the probability of all the gene – gene co-clustering probabilities. The two clustering probability matrices were then projected into a common 3-D PCA space. c) Hierarchical clustering of the difference between the trained and yoked clustering probability matrices, using Pearson correlations as the distance metric. d) Cartoons illustrate calculations of the Euclidean Distance, Centroid Similarity, and the Gene Similarity measures of how the yoked and trained genes changed within each condition's clustering probability matrix. e) Summary plot of how the 714 genes are distributed across the features depicted in panel d. Colors in panels b and e correspond to their positions within the value distributions. f) t-SNE plot derived from the three metrics for dimensionality reduction. The color bar represents the magnitude of Euclidean Distance changes from the yoked to the trained data set.

YOKED: Standard (n=8) vs. Conflict (n = 9)						
Trials	ANOVA	Predictor	df	F	p-value	Sig.
Pretraining (P)	NumEntrances	treat	1, 15	0.15	0.707	
	Time1stEntr	treat	1, 15	0.12	0.739	
	pTimeShockZone	treat	1, 15	0.64	0.436	
Initial training (T1-T3)	NumEntrances	treat	1, 45	0.09	0.77	
		trial	2, 45	8.89	0.001	**
		treat x trial	2, 45	0.37	0.69	
	Time1stEntr	treat	1, 45	0.28	0.598	
		trial	2, 45	5.85	0.006	**
		treat x trial	2, 45	0.09	0.915	
	pTimeShockZone	treat	1, 45	0.15	0.705	
		trial	2, 45	1.99	0.149	
		treat x trial	2, 45	0.25	0.779	
Initial recall (Rt)	NumEntrances	treat	1, 15	0.14	0.714	
	Time1stEntr	treat	1, 15	1.99	0.178	
	pTimeShockZone	treat	1, 15	0.05	0.822	
Conflict training (T4-T6)	NumEntrances	treat	1, 45	0.29	0.59	
		trial	2, 45	0.49	0.617	
		treat x trial	2, 45	0.16	0.855	
	Time1stEntr	treat	1, 45	1.01	0.32	
		trial	2, 45	3.88	0.028	*
		treat x trial	2, 45	2.45	0.097	
	pTimeShockZone	treat	1, 45	4.79	0.034	*
		trial	2, 45	1.25	0.296	
		treat x trial	2, 45	1.0	0.376	
Conflict recall (Rn)	NumEntrances	treat	1, 15	2.0	0.178	
	Time1stEntr	treat	1, 15	2.77	0.117	
	pTimeShockZone	treat	1, 15	11.29	0.004	**
All trials (Pre-Rn)	PC1 ~ treatment	treat	1, 151	4.47	0.036	*
	PC2 ~ treatment	treat	1, 151	2.97	0.087	
Retention (Rn)	PC1 ~ treatment	treat	1, 15	4.71	0.046	*
	PC2 ~ treatment	treat	1, 15	0.91	0.356	

Table 1. One-way and two-way ANOVA results of place avoidance behavior analysis of the yoked groups. One-way ANOVAs were used to determine the effects of training treatment on behavior during a single trial, including the pretraining and recall trails. Two-way ANOVAs were used to determine the effects of treatment, training, and interaction across three trials, including the initial training session and the conflict training session. Analyses were performed separately for the trained and yoked groups. Key: NumEntrances = total number of entrances into the shock zone; Time1stEntr = time (min) passes before entering the shock zone the first time; pTimeShockZone = proportion of time spent in the shock zone relative to the total time spent in the four equivalent areas centered in each quadrant. Treat = binary treatment condition (standard vs. conflict training). df: degrees of freedom for treatment and residuals. Sig.: level of statistical significance indicated as * ($p < 0.05$), ** ($p < 0.01$), and *** ($p < 0.001$).

TRAINED: Standard (n=8) vs. Conflict (n = 9)						
Trials	ANOVA	Predictor	df	F	p-value	Sig.
Pretraining (P)	NumEntrances	treat	1, 15	0.04	0.837	
	Time1stEntr	treat	1, 15	0.96	0.343	
	pTimeShockZone	treat	1, 15	0.51	0.485	
Initial training (T1-T3)	NumEntrances	treat	1, 45	0.6	0.444	
		trial	2, 45	8.9	0.001	**
		treat x trial	2, 45	1.28	0.288	
	Time1stEntr	treat	1, 45	0.15	0.697	
		trial	2, 45	9.33	0.0	***
		treat x trial	2, 45	0.85	0.433	
	pTimeShockZone	treat	1, 45	0.06	0.805	
		trial	2, 45	4.01	0.025	*
		treat x trial	2, 45	1.07	0.351	
Initial recall (Rt)	NumEntrances	treat	1, 15	0.97	0.341	
	Time1stEntr	treat	1, 15	0.72	0.411	
	pTimeShockZone	treat	1, 15	2.38	0.144	
Conflict training (T4-T6)	NumEntrances	treat	1, 45	10.7	0.002	**
		trial	2, 45	5.68	0.006	**
		treat x trial	2, 45	4.61	0.015	*
	Time1stEntr	treat	1, 45	4.56	0.038	*
		trial	2, 45	2.56	0.089	
		treat x trial	2, 45	1.51	0.233	
	pTimeShockZone	treat	1, 45	13.31	0.001	**
		trial	2, 45	14.54	0.0	***
		treat x trial	2, 45	9.62	0.0	***
Conflict recall (Rn)	NumEntrances	treat	1, 15	0.96	0.342	
	Time1stEntr	treat	1, 15	0.82	0.381	
	pTimeShockZone	treat	1, 15	0.49	0.495	
All trials (Pre-Rn)	PC1 ~ treatment	treat	1, 151	1.16	0.283	
	PC2 ~ treatment	treat	1, 151	0.0	0.980	
Retention (Rn)	PC1 ~ treatment	treat	1, 15	1.48	0.243	
	PC2 ~ treatment	treat	1, 15	0.0	0.959	

Table 2. One-way and two-way ANOVA results of place avoidance behavior analysis of the trained groups.

YOKED (n=17) vs. TRAINED (n=17)						
Trials	ANOVA	Predictor	df	F	p-value	Sig.
Pretraining (P)	NumEntrances	train	1, 32	0.07	0.789	
	Time1stEntr	train	1, 32	1.49	0.231	
	pTimeShockZone	train	1, 32	1.24	0.274	
Initial training (T1-T3)	NumEntrances	train	1, 96	199.21	0.0	***
		trial	2, 96	17.47	0.0	***
		train x trial	2, 96	0.85	0.431	
	Time1stEntr	train	1, 96	16.21	0.0	***
		trial	2, 96	10.67	0.0	***
		train x trial	2, 96	8.27	0.0	***
	pTimeShockZone	train	1, 96	366.74	0.0	***
		trial	2, 96	2.98	0.055	
		train x trial	2, 96	1.29	0.279	
Initial recall (Rt)	NumEntrances	train	1, 32	141.42	0.0	***
	Time1stEntr	train	1, 32	22.79	0.0	***
	pTimeShockZone	train	1, 32	411.4	0.0	***
Conflict training (T4-T6)	NumEntrances	train	1, 96	36.77	0.0	***
		trial	2, 96	3.25	0.043	*
		train x trial	2, 96	3.27	0.042	*
	Time1stEntr	train	1, 96	27.9	0.0	***
		trial	2, 96	2.7	0.073	
		train x trial	2, 96	2.01	0.14	
	pTimeShockZone	train	1, 96	251.64	0.0	***
		trial	2, 96	0.62	0.542	
		train x trial	2, 96	2.12	0.125	
Conflict recall (Rn)	NumEntrances	train	1, 32	49.7	0.0	***
	Time1stEntr	train	1, 32	16.36	0.0	***
	pTimeShockZone	train	1, 32	51.54	0.0	***
All trials (Pre-Rn)	PC1 ~ treatment	train	1, 304	285.77	0.0	***
	PC2 ~ treatment	train	1, 304	28.61	0.0	***
Retention (Rn)	PC1 ~ treatment	train	1, 32	46.74	0.0	***
	PC2 ~ treatment	train	1, 32	0.0	0.99	

Table 3. One-way and two-way ANOVA results of place avoidance behavior analysis comparing merged yoked and trained groups. The predictor term 'train' indicates the two binary training conditions (Yoked vs. Trained).