NEUROSCIENCE

Persistence of neuronal representations through time and damage in the hippocampus

Walter G. Gonzalez, Hanwen Zhang, Anna Harutyunyan, Carlos Lois*

How do neurons encode long-term memories? Bilateral imaging of neuronal activity in the mouse hippocampus reveals that, from one day to the next, ~40% of neurons change their responsiveness to cues, but thereafter only 1% of cells change per day. Despite these changes, neuronal responses are resilient to a lack of exposure to a previously completed task or to hippocampus lesions. Unlike individual neurons, the responses of which change after a few days, groups of neurons with inter- and intrahemispheric synchronous activity show stable responses for several weeks. The likelihood that a neuron maintains its responsiveness across days is proportional to the number of neurons with which its activity is synchronous. Information stored in individual neurons is relatively labile, but it can be reliably stored in networks of synchronously active neurons.

Memories are processed and stored by complex networks of neurons across several circuits in the brain; however, little is known about how stable information is encoded in these neurons. The hippocampus plays an essential role in the formation of memories (*1*, *2*), and neurons in this brain area show robust responses to space (via place cells) or other cues relevant to the task (*3*–*5*). Neuronal activity in the hippocampus changes during learning (*6*, *7*) and reexposure to an environment (*8*). However, what aspects of neuronal activity in the hippocampus persist during future visits to a familiar environment, how information is encoded in groups of neurons, and how lesions perturb the long-term maintenance of these neuronal patterns remain poorly understood.

Neuronal representations in the hippocampus change over time as information is transferred from other brain areas (*9*), but how these changes are regulated is unknown (*10*–*14*). Long-term imaging of neuronal activity showed that during repeated visits to a familiar environment only 31% of neurons are active in any given session and 2.8 ± 0.3% are active in all sessions (*10*). But how is it possible to recall a memory with such a small percentage of active neurons (*15*)? Technical limitations, including motion artifacts or toxicity induced by the overexpression of the genetically encoded calcium indicator GCAMP delivered via viral vectors, could have hampered the detection of active neurons over multiple days (*16*). To overcome these potential limitations, we chronically implanted custom-built microendoscopes and performed long-term simultaneous bilateral imaging of neuronal activity in the CA1 region of the hippocampus in freely moving Thy1-GCaMP6s mice (Fig. 1, A to C; figs. S1 and S2; and movies S1 and S2) (*17*–*20*). We imaged CA1 pyramidal neuron activity for several weeks spanning three phases: (i) the "learning phase," during the initial four sessions of exposure to a novel linear track with a sugary water reward at the ends; (ii) the "reexposure phase," after a 10-day period of no task-exposure; and (iii) the "damage phase," after creation of a light-induced hippocampus lesion (fig. S1F). We reliably imaged the activity of neurons in the same field of view (FOV) for up to 8 months and followed the activity of the same neurons for up

Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA, USA.
*Corresponding author. Email: clois@caltech.edu

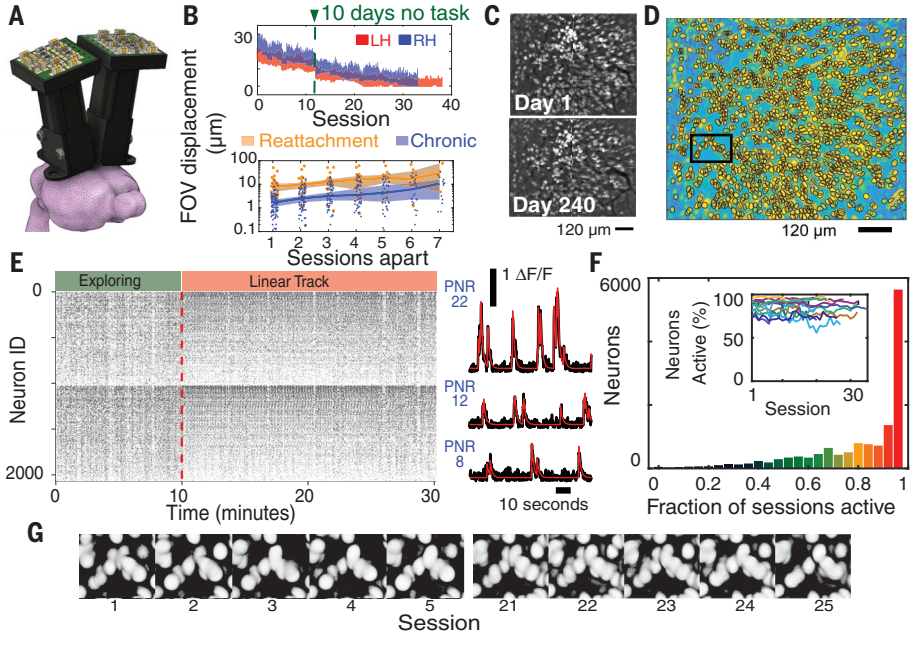

Fig. 1. Long-term simultaneous bilateral imaging of CA1 activity in freely moving transgenic mice. (**A**) Placement of custom microendoscopes for imaging activity in the left (LH) and right (RH) hemispheres. We did not notice large differences between hemispheres, and unless stated otherwise, the values reported represent combined data from both hemispheres in eight mice (table S1). (**B**) (Top) Drift in the FOV during bilateral imaging across weeks in one mouse. (Bottom) Chronic implants reduce motion artifacts (in micrometers, chronic versus reattachment, 1 day apart: 1.7 ± 2.6 versus 8.0 ± 10, $P < 0.01$; 7 days apart: 11 ± 10 versus 30 ± 31, $P < 0.01$, rank-sum test). (**C**) Maximum intensity projection in the same animal recorded for 8 months. (**D**) CNMFe (constrained non-negative matrix factorization for microendoscopic data) output from 25 sessions analyzed simultaneously in one mouse. (**E**) (Left) Deconvoluted neuronal activity in the RH (top) and LH (bottom) of a trained mouse. Neurons are ranked by peak-to-noise ratio (PNR). Between 324 and 1386 neurons were recorded per FOV (1974 ± 486 in four bilateral implants and 1081 ± 325 in five unilateral implants, a total of 13,558 neurons and 3443 place and end cells). (Right) Three representative raw traces (black) and deconvoluted transient traces (red) of neurons with the highest, middle, and lowest PNRs obtained from the data shown at left. ΔF/F, normalized fluorescence intensity change. (**F**) The majority of neurons are active in any given session (inset; 87 ± 9%, each color represents one mouse), and 47 ± 22% are active in every session (RH and LH, $n = 13$ replicates). (**G**) Intensity correlation map of a small region of interest [black rectangle in (D)] showing persistent activity of the same neurons throughout 25 sessions.

to 2 months (Fig. 1, D and E; figs. S2 to S4; and movies S3 and S4). On any given day, 87 ± 9% of all neurons in the FOV were active, and 47 ± 22% were active in every session (Fig. 1, F and G, and figs. S2 and S3). Within the same day, ~95% of neurons were active while the animal was in the cage as well as in the linear track. However, there were significant changes in neuron firing rates on the same day between environments (cage or linear track) or across days in the linear track (fig. S4).

To investigate the stability of neuronal representations in CA1, we studied both place cells and end cells (cells firing during periods of immobility at the ends of the track) (*21*, *22*) (figs. S5 to S7). Of all the place and end cells we recorded, 35 ± 13% that responded to a field on a given day did not respond to any field on the next day. Of the ~65% of place and end cells that retained responses to a field (defined here as "cell overlap"), 44 ± 13% of place cells and 29 ± 26% of end cells changed the centroid of the response field by more than 15 cm and 480 ms, respectively (fig. S5). These changes were present regardless of the animal's level of familiarity with the environment and were more prominent during the learning phase (*10*) (fig. S5). Despite the abrupt change between two consecutive days, the cell overlap between longer intervals decreased an additional ~1% per day, reaching random levels at ~50 days (Fig. 2, A and B).

The changes in neuronal representation between days could be due to spontaneous drift, or they could reflect minor changes in the environment (e.g., different personnel or odors in the room). To investigate these possibilities, trained animals were not exposed to the linear track for at least 10 days (fig. S1B). After re-exposure, place and end cell overlap before and after the 10-day gap decreased by the same amount in the aforementioned animals as in animals continuously exposed to the task in sessions separated by 10 days (Fig. 2C and fig. S8). Thus, changes in cell overlap of place and end cells are not triggered by exposure to the task.

How do neuronal representations change after learning a task? The response fields of approximately a third of place and end cells were similar across days after the learning phase (*36* ± 22% retained the same directional preference between sessions, here defined as "directional stability," with a response field correlation of 0.6 ± 0.4) (Fig. 2, E and F, and fig. S8). The responses of neurons during the first two sessions of the reexposure phase were less stable (directional stability: 30 ± 17, correlation: 0.6 ± 0.4) than the field responses that developed thereafter (directional stability: 42 ± 21, correlation: 0.8 ± 0.3). After a 10-day period in which the animals were not exposed to the task, 28 ± 15% of place and end cells recovered their fields (field correlation: 0.7 ± 0.4). Notably, in a third of the sessions, ~85% of place and end cells spontaneously reversed their direction selectivity, although they retained the centroid location of the field with respect to the reward (figs. S9 and S10). These field rotations occurred simultaneously across hemispheres and had minimal impact on behavior (figs. S9 and S10).

Are CA1 representations resilient to brain damage? We created local lesions in CA1 by increasing the microendoscope LED (light-emitting diode) power and illumination time (5- to 10-fold above the threshold needed to visualize GCaMP6s) (Fig. 3A). One day after damage, we observed a pronounced increase in synchronized firing, reminiscent of the neuronal activity associated with seizures (Fig. 3B). These abnormal patterns of activity recruited the majority of neurons in the FOV and were direction specific in the linear track (fig. S11 and movie S5) (*23*). During days with abnormal CA1 activity, the number of place and end cells increased but their fields were highly unstable (directional stability: 13 ± 13%, correlation: 0.5 ± 0.4) (Fig. 3, C to F, and fig. S11F). After 2 to 10 days, abnormal activity ceased (movie S4), response fields stabilized (directional stability: 47 ± 19%, correlation 0.8 ± 0.3), and 20 ± 10% of place and end cells retained their pre-lesion directional preference and a field correlation of 0.6 ± 0.3 (Fig. 3F and fig. S11G).

Theoretical models have suggested that groups of neurons with synchronized activity may form cell assemblies able to encode learned representations for long periods of time (*23*). Cell assemblies that encode task-dependent cues have been observed in the hippocampus (*24*, *25*), but how these groups of neurons develop during learning, persist across days, or encode stable information is not known. To investigate these questions, we first analyzed the activity of pairs of neurons (fig. S13). The number of synchronized neuronal pairs in CA1 increased within and across hemispheres as mice became familiar with the linear track and was proportional to their firing rate (Fig. 4A). Synchronized pairs of neurons also made similar errors during task performance: (i) Whenever one neuron in the pair failed to fire in its field, 50 ± 30% of the time the other neuron in the pair failed as well. (ii) When both neurons fired, their deviation from the field peak was highly correlated (0.71 ± 0.14, $P < 10^{-8}$) (fig. S13). Neuron pairs encoded higher information content (as defined by calculation of mutual information; see materials and methods) than individual place or end cells,

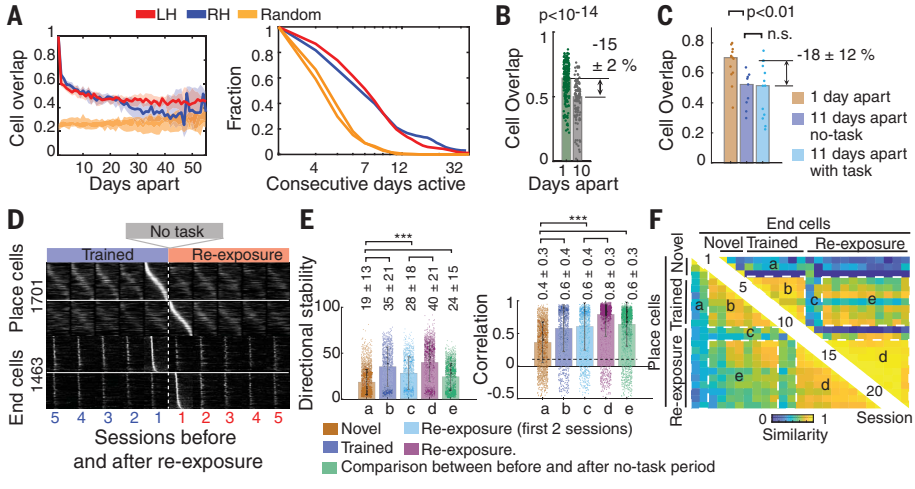

Fig. 2. Response fields in CA1 persist through periods of no exposure to the task. (**A**) (Left) Cell overlap of place and end cells on sessions *n* days apart. Lines and shading represent the median and 95% bootstrap confidence interval, respectively. (Right) Fraction of place and end cells retaining a field for more than two consecutive days (LH: $n = 5$, RH: $n = 8$). Place and end cells retained their place field response, on average, for 7 ± 4 consecutive days. (**B**) Quantification of cell overlap 1 day apart (65 ± 13%, $n = 206$ session pairs) and 10 days apart (51 ± 16%, rank-sum test, $n = 110$ session pairs). The analysis depicted in (A) and (B) is between all possible session intervals, including periods of learning and no task. (**C**) Cell overlap during the trained phase 1 day apart, 11 days apart without task, and 11 days apart with task (70 ± 12%, 52 ± 12%, and 51 ± 18%, respectively; rank-sum test, $n = 10$ session pairs). After the learning phase, there were no obvious changes in the behavior of the animals (fig. S2B). n.s., not significant. (**D**) Normalized tuning curves of all place and end cells aligned to the day before or after the no-task period (the number of place and end cells before and after the no-task period is shown on the *x* axis). Cells are sorted by centroid location on the day before (top) or after (bottom) the no-task period. (**E**) Directional stability and field correlation between sessions during different periods of the task (rank-sum test, $^{***}P < 10^{-15}$; dots are median of session pairs, dashed line is random level; see tables S1 and S2 for statistics). Error bars indicate SD. (**F**) Field similarity map during different periods of the task (data for one mouse shown). Regions a, b, c, d, and e correspond to the periods defined in (E).

and the likelihood of a neuron maintaining its responsiveness to the same or a different field was proportional to its correlation with the other neuron in the pair (fig. S13).

Do synchronized neurons form part of a larger neuronal network storing stable task information? Network graphs of correlated neuronal activity showed a clear behavior-dependent topology, evolving through learning and reorganizing upon transition from the home cage to the linear track (Fig. 4B). Graphs revealed dense clusters of neurons with preferences for specific behaviors or cues (Fig. 4C and movie S7). Extraction of these clusters by a Markov diffusion approach identified groups of neurons (defined here as cell groups) encoding direction-specific information about running, immobility, drinking, and turning (Fig. 4D and movies S8 and S9) (*26*). Synchronization within cell groups was specific to the task performed (running, staying on the end of the track, or foraging in the home cage). Cell groups that had synchronous activity in the linear track became asynchronous in the home cage, and vice versa (fig. S16). Whereas individual neurons developed response fields within minutes of task exposure, the functional connectivity of neurons in a graph and a neuron cluster developed over a much longer period of time, usually two to four sessions (Fig. 4E). The reorganization of network activity was highly correlated with task performance (clustering coefficient correlation of 0.75, $P < 10^{-3}$, all mice combined) and could not be explained by random activity of neurons with overlapping response fields (fig. S17). Information about the task encoded in cell groups was more stable than that encoded by individual place and end cells, even after 10 days of no task exposure (Fig. 4F and figs. S15 and S16).

Using graphs, we observed that the likelihood of a neuron to remain a place cell or an end cell was proportional to the number of graph connections with other neurons (Fig. 4G). The cell group participation of a neuron in the graph could also be used to classify a neuron as a place cell, an end cell, or neither, without the need to observe the animal's behavior (Fig. 4H). We could decode the field centroid location of place and end cells on the basis of the properties of near neighbors in the graph (Fig. 4I; see materials and methods). In addition, we observed that future drifts in the response field of place cells were proportional to the centroid location of their neighbors in the graph. This relationship was used to predict, without observing neuronal activity or behavior, the magnitude of future drifts in the centroid of place cells (Fig. 4J). Thus, the features of a neuron in a graph are sufficient to decode signatures specific to place and end cells, as well as the current and future location of their response fields.

In contrast to previous studies, we observed that, although the firing rate changed across sessions and tasks, the majority of neurons were active on most days (*10*). We observed high stability of place cells such that, even after 35 days, 40% of place and end cells were still participating in the encoding of the environment (*10*, *12*). Most of the changes in hippocampal representations (~35%) were observed from one day to the next but were modest (~1% per day) thereafter. These gradual changes are consistent with the reported variability of neuronal responses in the hippocampus (*27*, *28*). The representation in CA1 recovered after an extended period with no task exposure, or even after abnormal activity induced by local lesions. In both of these cases, fields underwent transient drifts and fluctuations that ultimately converged to a neuronal representation in which ~20% of place and end cells recovered a similar response field to that present before perturbations. Altogether, these results support a model with three complementary features. First, neuronal representations in the hippocampus spontaneously change over time, such that it is rare to find cells whose fields persist longer than 35 days. Second, attractor-like mechanisms ensure the persistence of representations over short periods of time (days), even if the animals are not exposed to the task or if the circuit is perturbed by lesions. Third, our results suggest that information is dynamically stored in CA1. Initially, when naïve mice are exposed to a novel environment, place and end cells become sensitive to a field, but their responses are unstable between days. As mice become familiar with the task and environment, place and end cells within and between hemispheres synchronize their activity, become more stable between days, and become resilient to perturbations. Although individual neurons with response fields on more than two sessions retain their field response for

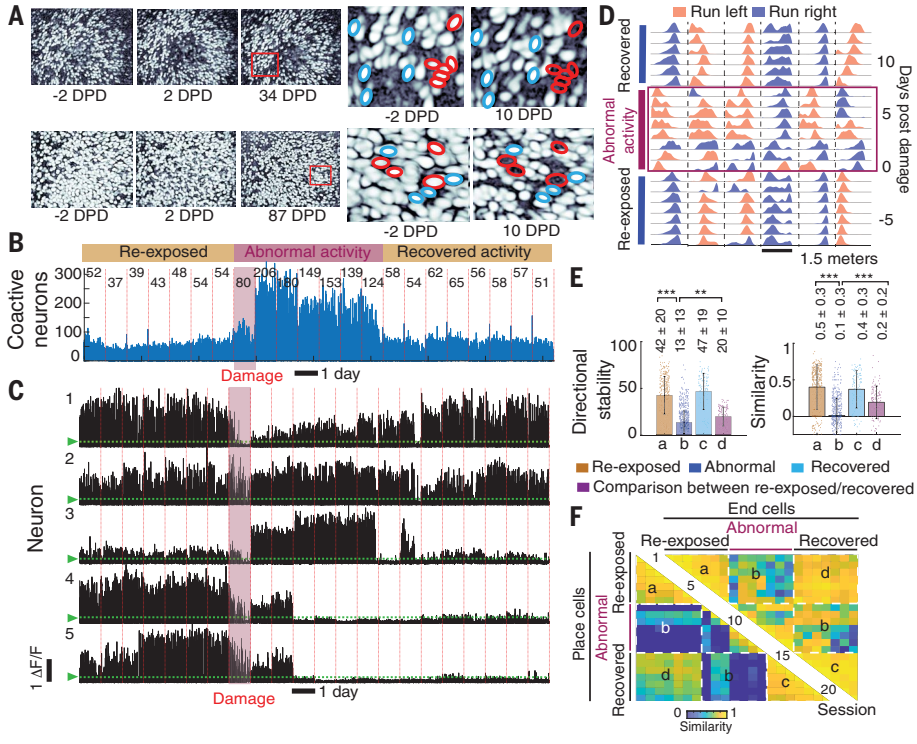

Fig. 3. Tissue damage induces direction-specific bursting activity and unstable fields followed by partial recovery of the original representation. (**A**) Intensity correlation of a session before and after formation of a lesion induced by high illumination intensity (DPD, days post damage). Top and bottom rows are from animals with representative large and small lesions, respectively. The four panels on the right show high-magnification views of the regions enclosed by the red rectangles at left, indicating neurons that remain active (blue circle) or become inactive (red circle) after lesion formation. Both animals displayed clear direction-dependent abnormal activity in the linear track (fig. S11 and movies S5 and S6). (**B**) Number of neurons active simultaneously during each session in one mouse (numerals indicate the average number of neurons in each burst). Each day is marked by a vertical red line. (**C**) Single neuron activity across days while the mouse runs in the linear track (black traces), showing changes in activity induced by tissue damage (red shaded region). Green arrowheads point to the 3σ error of the noise. The top two neurons recover from the damage, the middle neuron is highly active during the period of abnormal activity, and the bottom two neurons become inactive (dead or silent) after damage. (**D**) Tuning curves of six place cells, showing high instability during the period of abnormal activity (rectangle). (**E**) Directional stability and field similarity between sessions before and after lesion formation (dots are medians of session pair; $^{**}P < 0.05$, $^{***}P < 10^{-5}$, $n = 3$ mice, RH; see tables S3 and S4 for statistics). (**F**) Field similarity map between sessions before and after the lesion sessions (RH, one mouse). Regions a, b, c, and d correspond to the periods defined in (E).

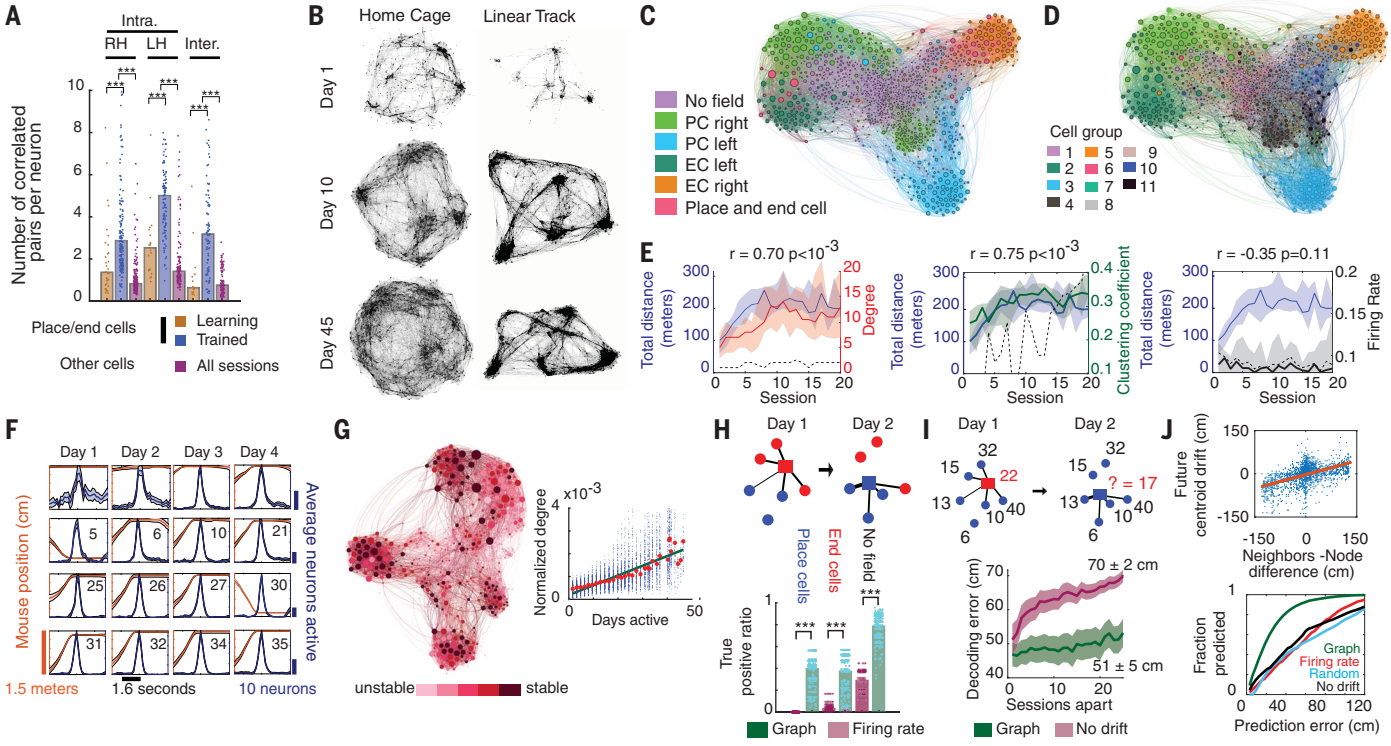

Fig. 4. Synchronous activity in groups of cells encoding stable representations. (**A**) The number of inter- and intrahemispheric synchronous neuron pairs increases during training (dots are sessions; $^{***}P < 10^{-6}$, four bilateral mice). (**B**) Graph topology during learning (day 1), trained (day 10), and reexposed (day 45) phases (lines represent correlation >0.10, neurons shown as dots, RH from one mouse). (**C** and **D**) Graphs of neuronal activity in the linear track colored by place cell (PC) or end cell (EC) classification (C) or by cell group (D). The size of each node is proportional to its degree. (**E**) The degree (left, red line) and clustering coefficient (middle, green line) increase on time scales of days and are correlated with the total distance run by the mouse in a single session (blue line). Firing rates (right, black line) remain constant throughout all sessions. Medians are shown as solid lines, dashed lines represent results from the synthetic dataset (see materials and methods), and the shaded areas denote 95% bootstrap confidence intervals. (**F**) Activity of one cell group persisting for 35 days. Medians ± SDs across laps in a session are shown as lines and shaded areas, respectively. (**G**) (Left) Graph colored by the fraction of sessions a neuron was classified as a place cell or an end cell. (Right) Quantification of the relationship between degree and stability (blue dots are individual neurons, red dots mark median; linear fit $R^2 = 0.82$, $P < 10^{-10}$). (**H**) (Top) Diagram illustrating how connectivity of a neuron (square) in a graph on day 1 can be used to decode its classification as a place cell (blue circles) or end cell (red circles) on another day. (Bottom) Classification accuracy using graph connectivity and firing rates (firing rate versus graph: 0 versus 0.36 ± 0.14 for end cells, 0.04 ± 0.04 versus 0.32 ± 0.16 for place cells, no field cells: 0.30 and 0.83; $^{***}P < 10^{-20}$, rank-sum test). (**I**) (Top) The centroid of near neighbors of a place cell (square) on one day can be used to decode the centroid on another day. (Bottom) Quantification of decoding accuracy using graph neighbors or assuming no centroid drift (median ± SEM, errors for session 25 shown, $P < 10^{-10}$). (**J**) (Top) In a session, the difference between the response field of a place cell and the average centroid of its neighbors in the graph is proportional to future drifts in its response field (linear fit $R^2 = 0.22$, $P < 10^{-10}$, $n = 2000$ centroid deviations across eight mice). (Bottom) The linear relationship can be used to make predictions on 50% of place cells with <35 ± 7 cm error. This result is more accurate than those obtained using other metrics (see materials and methods).

7 days, on average, cell groups of synchronously active neurons formed stable representation of the environment that persisted for several weeks. Our results show that stable representations of a task are stored in groups of CA1 neurons with synchronous activity. On the basis of the known connectivity of CA1 neurons, this synchrony may arise from local inhibition within CA1 or common input from CA3. We observe that the firing patterns of neurons that are neighbors in a graph can be used to decode current and future drifts in the response field of a neuron, suggesting that future changes in place cells are, to some extent, encoded during task performance. Overall, we demonstrate that activity patterns of individual neurons gradually change over time, whereas the activity of groups of synchronously active neurons ensures the persistence of representations in CA1.

REFERENCES AND NOTES

1. W. B. Scoville, B. Milner, *J. Neurol. Neurosurg. Psychiatry* **20**, 11–21 (1957).
2. N. J. Cohen, L. R. Squire, *Science* **210**, 207–210 (1980).
3. J. O'Keefe, J. Dostrovsky, *Brain Res.* **34**, 171–175 (1971).
4. A. Czurkó, H. Hirase, J. Csicsvari, G. Buzsáki, *Eur. J. Neurosci.* **11**, 344–352 (1999).
5. D. Aronov, R. Nevers, D. W. Tank, *Nature* **543**, 719–722 (2017).
6. M. R. Mehta, C. A. Barnes, B. L. McNaughton, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 8918–8921 (1997).
7. M. A. Wilson, B. L. McNaughton, *Science* **261**, 1055–1058 (1993).
8. J. Debiec, J. E. LeDoux, K. Nader, *Neuron* **36**, 527–538 (2002).
9. S. M. Zola-Morgan, L. R. Squire, *Science* **250**, 288–290 (1990).
10. Y. Ziv *et al.*, *Nat. Neurosci.* **16**, 264–266 (2013).
11. J. D. Zaremba *et al.*, *Nat. Neurosci.* **20**, 1612–1623 (2017).
12. T. Hainmueller, M. Bartos, *Nature* **558**, 292–296 (2018).
13. C. G. Kentros, N. T. Agnihotri, S. Streater, R. D. Hawkins, E. R. Kandel, *Neuron* **42**, 283–295 (2004).
14. E. A. Mankin, G. W. Diehl, F. T. Sparks, S. Leutgeb, J. K. Leutgeb, *Neuron* **85**, 190–201 (2015).
15. K. Z. Tanaka *et al.*, *Neuron* **84**, 347–354 (2014).
16. H. A. Zariwala *et al.*, *J. Neurosci.* **32**, 3131–3141 (2012).
17. D. J. Cai *et al.*, *Nature* **534**, 115–118 (2016).
18. H. Dana *et al.*, *PLOS ONE* **9**, e108697 (2014).
19. K. K. Ghosh *et al.*, *Nat. Methods* **8**, 871–878 (2011).
20. C. J. MacDonald, K. Q. Lepage, U. T. Eden, H. Eichenbaum, *Neuron* **71**, 737–749 (2011).
21. K. Kay *et al.*, *Nature* **531**, 185–190 (2016).
22. R. Boehringer *et al.*, *Neuron* **94**, 642–655.e9 (2017).
23. D. O. Hebb, *The Organization of Behavior* (Wiley, 1949).
24. K. D. Harris, J. Csicsvari, H. Hirase, G. Dragoi, G. Buzsáki, *Nature* **424**, 552–556 (2003).
25. E. Pastalkova, V. Itskov, A. Amarasingham, G. Buzsáki, *Science* **321**, 1322–1327 (2008).
26. Y. N. Billeh, M. T. Schaub, C. A. Anastassiou, M. Barahona, C. Koch, *J. Neurosci. Methods* **236**, 92–106 (2014).

27. K. Z. Tanaka *et al*., *Science* **361**, 392–397 (2018).
28. T. Kitamura *et al*., *Science* **356**, 73–78 (2017).
29. W. Gonzalez, Persistence of neuronal representations through time and damage in the hippocampus, Version 1.0, CaltechDATA (2019); https://doi.org/10.22002/d1.1229.

ACKNOWLEDGMENTS

Funding: W.G.G. was supported by the American Heart Association (grant 17POST33670151), the Della Martin Foundation, and the Burroughs Wellcome Fund PDEP (grants 1016194 and 1018228) and CASI awards. The project was funded by the National Institute of Neurological Disorders and Stroke (grant R01NS104925 to C.L.). **Author contributions:** C.L. and W.G.G. conceived and designed the experiments. W.G.G. designed, built, and implanted microendoscopes in mice. W.G.G., H.Z., and A.H. performed behavior recordings. W.G.G. and H.Z. wrote the analysis code. W.G.G. analyzed and compiled the data. C.L. and W.G.G. wrote the manuscript. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** Processes data and MATLAB code are available at the Caltech Research Data Repository (*29*).

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/365/6455/821/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S21
Tables S1 to S4
References (*30*–*36*)
Movies S1 to S9

7 March 2019; accepted 1 August 2019
10.1126/science.aav9199

Persistence of neuronal representations through time and damage in the hippocampus

Walter G. Gonzalez, Hanwen Zhang, Anna Harutyunyan and Carlos Lois

Science **365** (6455), 821-825.
DOI: 10.1126/science.aav9199

Robust memories through neuron networks

How does the brain store information over a long period of time? Gonzalez *et al.* chronically implanted custom-built high-sensitivity microendoscopes and performed long-term imaging of neuronal activity in freely moving mutant mice. The majority of neurons were active on most days, but their firing rate changed across sessions and tasks. Although the responses of individual neurons changed, the responses of groups of neurons with synchronous activity were very stable across days and weeks. In addition, the network activity in hippocampal area CA1 recovered after an extended period without performing the task or even after abnormal activity induced by local lesions. These findings indicate the presence of attractor-like ensemble dynamics as a mechanism by which the representations of an environment are encoded in the brain.

Science, this issue p. 821