LETTER

doi:10.1038/nature09633

Preplay of future place cell sequences by hippocampal cellular assemblies

George Dragoi[1] & Susumu Tonegawa[1]

During spatial exploration, hippocampal neurons show a sequential firing pattern in which individual neurons fire specifically at particular locations along the animal's trajectory (place cells[1,2]). According to the dominant model of hippocampal cell assembly activity, place cell firing order is established for the first time during exploration, to encode the spatial experience, and is subsequently replayed during rest[3–6] or slow-wave sleep[7–10] for consolidation of the encoded experience[11,12]. Here we report that temporal sequences of firing of place cells expressed during a novel spatial experience occurred on a significant number of occasions during the resting or sleeping period preceding the experience. This phenomenon, which is called preplay, occurred in disjunction with sequences of replay of a familiar experience. These results suggest that internal neuronal dynamics during resting or sleep organize hippocampal cellular assemblies[13–15] into temporal sequences that contribute to the encoding of a related novel experience occurring in the future.

We recorded neuronal firing sequences from the CA1 area of the mouse hippocampus (Supplementary Fig. 1) during periods of awake rest (Fam-Rest) alternating with periods of running (Fam-Run) on a familiar track (Fam session; Supplementary Fig. 2a) that preceded the exploration of a novel linear arm in contiguity with the familiar track (Contig-Run on L-shaped track; Fig. 1, Supplementary Fig. 2a and Methods). All the place cells active on the novel arm during Contig-Run, whether previously silent[16] (19% in both directions and 31% in at least one direction; Methods and Supplementary Tables 1–3) or active during Fam-Run (subpanels a in Fig. 1), fired during Fam-Rest at the ends of the familiar track (range, 0.17–11.7 Hz; Supplementary Fig. 3) as part of a number of 'spiking events'. The spiking events were defined as epochs composed of multiple individual spikes from at least four different place cells active on the novel arm or familiar track, separated by less than 50 ms and flanked by at least 50 ms of silence[3,4]. More significantly, the temporal sequence in which the cells active on the novel arm fired during Fam-Rest (subpanels b in Fig. 1) was significantly correlated with the spatial sequence in which they fired later as place cells on the novel arm during Contig-Run (subpanels c in Fig. 1), despite being uncorrelated with their spatial sequence as place cells on the familiar track during Fam-Run. This is illustrated as place cell sequences during Contig-Run (subpanels c in Fig. 1) and Fam-Run (subpanels a in Fig. 1) compared with the firing sequences of these cells within individual spiking events observed during Fam-Rest (subpanels b in Fig. 1). We refer to this process as 'preplay' of place cell sequences because the temporal sequence of firing during Fam-Rest had occurred before the actual exploration of the novel arm in the subsequent Contig-Run and was not a replay of the place cell sequences from the previous Fam-Run.

To quantify the significance of preplay and to compare it with replay, we created place cell sequence templates according to the spatial order of the peak firing of place cells[3,4,10] on the novel arm during Contig-Run (novel arm templates; subpanels c in Fig. 1 and Methods) and on the familiar track during Fam-Run (familiar track templates) for each run direction. The spikes of all the place cells used to construct the two types of template that were emitted during Fam-Rest were sorted by time, and spiking events were determined as explained above (subpanels b in Fig. 1). For each spiking event, we calculated a rank-order correlation between the novel arm templates and the temporal sequence of firing of the corresponding cells in the spiking events during Fam-Rest. The event correlation was considered significant if it exceeded the 97.5th percentile of a distribution of correlations resulting from randomly shuffling the order of place cells in the novel arm templates 200 times ($P < 0.025$). Forward[4] and reverse[3,4] preplay refers to the cases in which the sequence of place cells during Contig-Run and the firing order of the corresponding cells in Fam-Rest were in the same and opposite directions, respectively. In 91% of the preplay cases, the spiking events were correlated with the novel arm template in one direction only. The distribution of event correlation values obtained using the original novel arm templates was significantly shifted towards higher positive or negative values in comparison with the distribution of correlation values obtained using shuffled templates (Fig. 2a and Supplementary Fig. 4). Figure 2a also shows the distribution of significant preplay events (in red). Of all the spiking events detected as above and in which at least four novel arm place cells were active, 14.2% were significant preplay events for the place cell sequence on the novel arm ($P < 10^{-32}$, binomial probability test[4]) in the forward or reverse order (Fig. 2b).

The occurrence of significant preplay events was correlated with the occurrence of high-frequency ripple oscillations in CA1 (Fig. 2c). The majority of the significant preplay events (81.1%; Fig. 2d, total, blue) took place at the junction between the familiar and novel arms, and the remaining 18.9% took place at the free end of the familiar track (Fig. 2d, total, purple). The proportion of significant preplay events among the total events at each of the two track ends was higher at the junctional end (15.2%, $P < 10^{-26}$) than at the free end (8.5%, $P < 10^{-4}$) of the familiar track ($P < 0.035$, Z-test; Fig. 2d, normalized).

We found a relatively high correlation between the place field maps (Fig. 1A, B and Supplementary Fig. 5) of the familiar track before and after the novel experience (median $r = 0.66$; Fig. 2e, familiar track, blue); it was significantly higher than the correlations obtained when the cell identities were shuffled (median $r = 0.23$, $P < 10^{-4}$; Fig. 2e, familiar track, black). A similar correlation analysis showed a relatively high stability of the newly formed place fields on the novel arm from the beginning to the end of Contig-Run (median $r = 0.62$ (newly formed) versus median $r = 0.21$ (shuffled), $P < 10^{-3}$; Fig. 2e, novel arm, blue versus grey). These results suggest that preplay of the novel arm does not occur over an entirely new (that is, remapped) representation of the whole L-shaped track but rather benefits from the relative stability of the familiar track representation across sessions and perhaps facilitates the rapid, stable encoding of the novel arm experience.

Using the familiar track templates and spiking events during Fam-Rest, constructed as above, we determined that 16.2% ($P < 10^{-91}$; data not shown) were significant replay events[3–6,17] among the spiking events in which a minimum of four familiar track place cells were active. All significant preplay events occurring during Fam-Rest ($n = 75$) were

[1]The Picower Institute for Learning and Memory, RIKEN-MIT Center for Neural Circuit Genetics, Department of Biology and Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA.

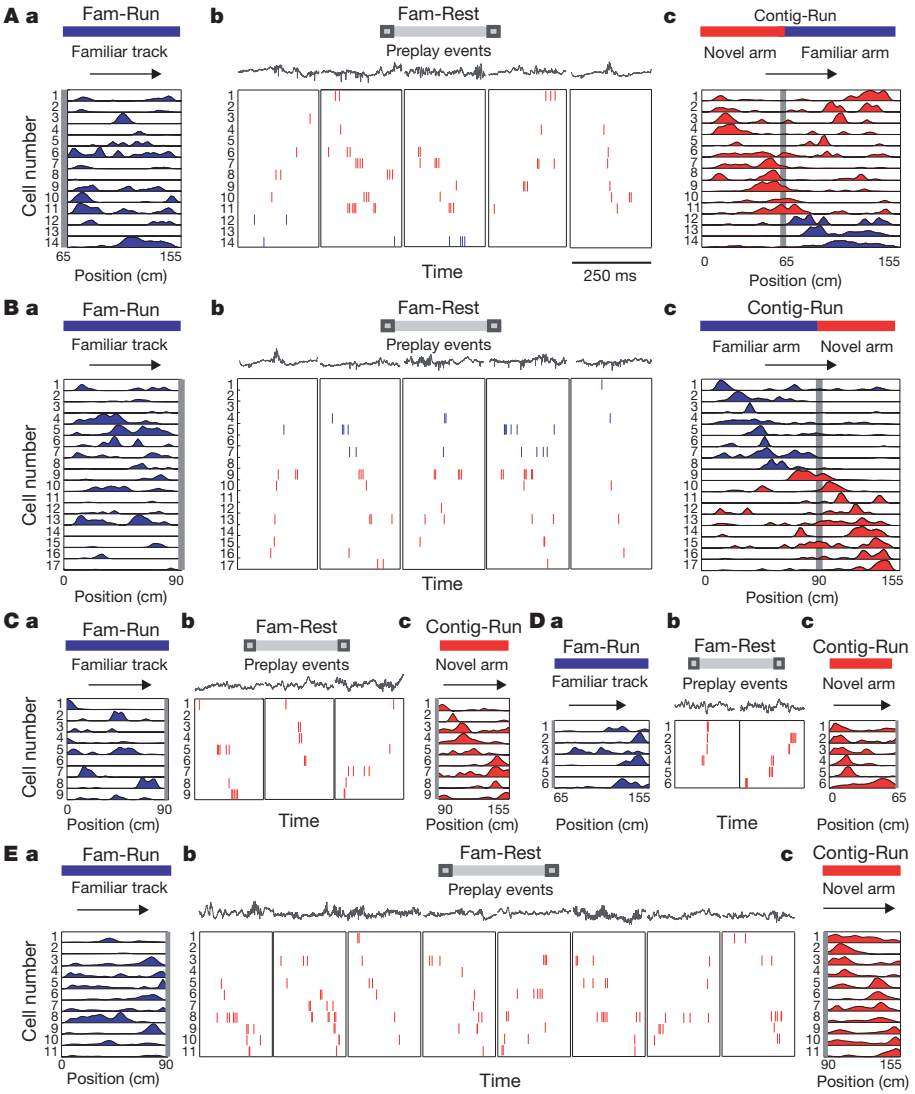

Figure 1 | Preplay of novel place cell sequences. Fam-Run and Fam-Rest respectively denote run and rest sessions on the familiar linear track before barrier removal; Contig-Run denotes run sessions on the L-shaped track after barrier removal. The L-shape track was linearized for display/analysis. **A**, **B**, mouse 1; **C**, **D**, mouse 2; **E**, mouse 3. **A**–**E**, **a**, Spatial activity on the familiar track during Fam-Run of the cells that had place fields in Contig-Run and preplayed during Fam-Rest (one cell per row); activity on the novel arm and familiar track are on the same scale. Horizontal arrows indicate run directions. Vertical grey bars indicate barrier locations during Fam-Run and Fam-Rest. **A**–**E**, **b**, Examples of representative spiking events in the forward or reverse direction during Fam-Rest in 250-ms time windows (350 ms for the second and fourth panels from left in **E**, **b**). Tick marks indicate individual spikes: red, preplay events for place cell sequences in the novel arm; blue (in **A**, **b** and **B**, **b**), additional spikes from the familiar track place cells participating in the spiking event (not shown in **C**–**E**, **b**). Numbers on the left denote cell numbers and correspond to the place cell numbers in **A**–**E**, **a**. Square boxes indicate the ends of the familiar track where preplay events occurred. Local field potentials recorded simultaneously with the spikes are shown above spiking events. **A**–**E**, **c**, Place cell sequences in the novel arm (**C**–**E**, **c**; red) or in both the novel arm (red) and the familiar arm (blue) (**A**, **c** and **B**, **c**) in Contig-Run.

tested for possible replay of the familiar track spatial sequence: these spiking events were more correlated with the novel arm template (Fig. 2f, red) than the familiar track template (Fig. 2f, blue). Seventy-two percent ($n = 54$) of the significant events previously considered to be preplay had no significant correlation with the familiar track template. An additional 16% ($n = 12$) of those events were better correlated with the novel arm templates (mean absolute $r = 0.92$) than with the familiar track template (mean absolute $r = 0.67$, $P < 10^{-3}$). Together, these findings reject the hypothesis that the preplay events simply represent a replay of the familiar track activity (see additional controls in Supplementary Information). Moreover, we found that the proportion of events exclusively composed of silent cells that perfectly matched the novel arm spatial templates was 0.67 (16 of 24 triplets), which is significantly greater ($P < 0.025$) than the proportion of by-chance perfect matches (0.33).

To illustrate the distribution and relative proportions of preplay and replay events among all significant spiking events during Fam-Rest, we calculated a 'template specificity index' (Fig. 2g and Methods) for each event. Pure preplay events (Fig. 2g, red) and pure replay events (Fig. 2g, blue) were segregated, and only a minority of events were significant for both preplay and replay (Fig. 2g, yellow). Consistent with this segregation of preplay and replay events, the novel arm and the 'corresponding familiar track' templates were not significantly correlated (Fig. 2h and Methods). The ratio between the number of pure replay events ($n = 171$) and the number of pure preplay events ($n = 54$) during Fam-Rest was about 3.1 (Fig. 2g, inset; see Supplementary Information for proportions of events). Preplay and replay events were distributed in time across Fam-Rest (Supplementary Fig. 6a–c) and their occurrences were generally uncorrelated (Supplementary Fig. 6d). The majority (79.9%) of the spiking events during Fam-Rest did not significantly correlate with either of the two templates (data not shown).

We used a Bayesian reconstruction algorithm[2,5,6,18,19] (Methods) to decode the animals' position from the spiking activity during Fam-Run (Fig. 3a) or Fam-Rest (Fig. 3b, c). For all original and shuffled[6]

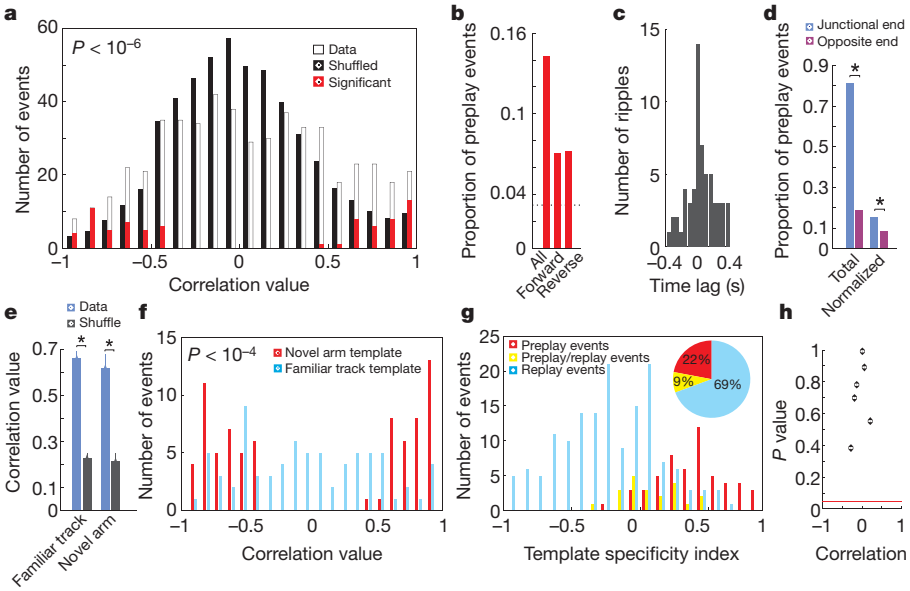

Figure 2 | Quantification of the preplay phenomenon and comparison with replay. **a**, Distribution of correlations between spiking events in Fam-Rest and spatial templates of the novel arm. Open bars indicate spiking events versus the original (unshuffled) templates; filled bars indicate spiking events versus 200 shuffled templates scaled down 200 times; red bars show the distribution of preplay (that is, significant) events. Similar distributions (not shown) of corresponding spiking events were obtained when spatial templates were constructed using all place cells active on the L-shaped track (Figs 1A, b, c and 1B, b, c; red and blue). **b**, Proportion of all, forward and reverse preplay events among the spiking events in Fam-Rest. The dotted line indicates the chance level (3.2%). **c**, Cross-correlation between preplay events and ripple epochs. **d**, Location of preplay events on the familiar track: total, proportions of preplay events at ends of the track; normalized, proportion of preplay events normalized by the number of spiking events at each end of track. Preplay events represented a trajectory running from the free end of the novel arm to the junctional end (40%) or begun near the familiar track (60%); the latter suggests that in some cases preplay events could be triggered by the activity of the familiar track place cells during Fam-Rest. **e**, Stability of place cell spatial tuning across the novel experience: familiar track, stability of the place fields active on the familiar track before (Fam-Run) versus after (Contig-Run) barrier removal; novel arm, stability of the place fields active on the novel arm at the beginning (first four laps of run) versus the end (last four laps) of the Contig-Run session. Data (blue), within-cell correlation of place cell spatial tuning for the corresponding track/arm; shuffle (black), cell identity shuffle (Supplementary Information). Error bars, s.e.m.; asterisks in **d** and **e** indicate significant differences. **f**, Distribution of preplay event correlations (red) versus distribution of these event correlations with the familiar track template (blue). Spiking events were detected using all place cells from the familiar track and novel arm templates (>1 Hz). Red bars are the same as in **a**. Correlation is strong with the novel arm template (preplay) and weak with the familiar arm template (replay). The *P* value corresponds to there being a significant difference between the two distributions. **g**, Disjunctive distribution of pure preplay (red), pure replay (blue) and preplay/replay (yellow) events during Fam-Rest over their template specificity index (Supplementary Information). Inset, proportions of pure preplay events (red), pure replay events (blue) and preplay/replay events (yellow) among all of the spiking events that were significantly correlated with at least familiar track templates or novel arm templates. **h**, Lack of correlation between the novel arm template and the corresponding familiar track template. Each of the six dots represents either a forward or a reverse run direction of one of the three mice analysed. Red horizontal line denotes a *P* value of 0.05. The correlation values were not significant in any of the cases (Supplementary Information).

probability distributions, a line was fitted to the data using a line-finding algorithm[6] to represent the decoded virtual trajectory (Methods and Supplementary Information). In 16.36% of cases representing trajectories, the reconstructed trajectory during spiking events in Fam-Rest was contained within the novel arm (Fig. 3c, top), a place the animal had not yet visited (that is, trajectory preplay). Moreover, in 79.8% of the trajectory preplay cases the shuffling procedures resulted in lines that were significantly less or not at all contained within the novel arm (that is, not preplay; Supplementary Information). The remaining trajectories decoded during Fam-Rest represented replay of the familiar track (64.15%; Fig. 3c, middle) or spanned the joint familiar track/novel arm space (19.49%; Fig. 3c, bottom). Means of absolute rank-order correlations between spiking activity and novel arm templates (Fig. 2a) restricted during epochs of trajectory preplay were significantly larger than those between spiking activity and familiar track templates calculated during the same epochs (0.75 versus 0.59, $P < 10^{-4}$). Overall, these results support the existence of the preplay phenomenon.

To investigate the possibility that preplay of novel arm place cell sequences during Fam-Rest depends on the prior run experience on the familiar track, mice with no prior experience on any linear track were placed in a high-walled sleep box and recorded while resting/sleeping. The animals were then transferred to a novel isolated linear track that was in the same room but could not be seen from inside the box, and the recording continued during *de novo* formation of place cells (Supplementary Fig. 2b, *de novo* session). We found that in a relatively large proportion (16.1%) of spiking events identified during sleep/rest in the sleep box, the neuronal firing sequences were significantly correlated with the place cell sequences observed during the first run session on the novel track (Fig. 4A, B and Methods); this was the case for all four individual mice (Supplementary Fig. 7). Preplay events were associated with the ripple occurrence (Fig. 4C). The place cells established on the novel track in the *de novo* session were more dynamic (median $r = 0.42$; Fig. 4D, blue) than in Contig-Run (median $r = 0.62$, $P < 0.016$; Fig. 2e, right, blue).

We have demonstrated that a significant number of temporal firing sequences of CA1 cells during resting periods of a familiar track exploration that preceded a novel track exploration in the same general environment were correlated with the place cell sequences of the novel track rather than the familiar track. This phenomenon, preplay, is temporally opposite to the process of replay[3–10,19,20], when activity during rest or sleep periods recapitulates place cell sequences that have already occurred during previous explorations. Preplay differs fundamentally from replay because it occurs before exploration of novel tracks.

Although our recordings were carried out in CA1, we believe that what we observed could be a reflection of the output of the recurrent cellular assemblies from upstream regions (CA3 or entorhinal cortex).

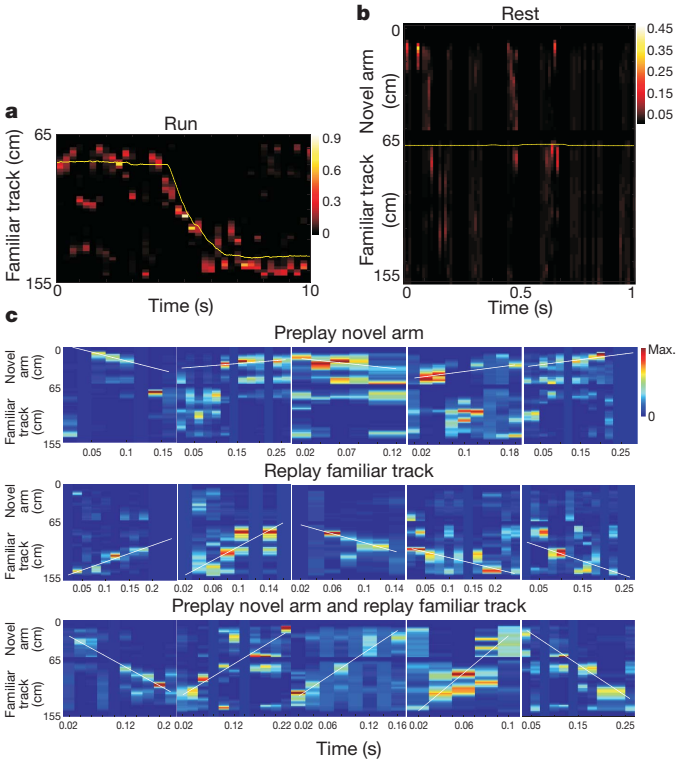

Figure 3 | Bayesian reconstruction of the animal's trajectory in the familiar track (replay) and novel arm (preplay). **a**, Position reconstruction of a one-lap run on the familiar track from the ensemble place cell activity during Fam-Run. The heat map displays the reconstructed position of the animal using ensemble place cell activity during the run (250-ms bins; animal velocity, >5 cm s^{-1}). The yellow line indicates the actual trajectory of the animal during Fam-Run. **b**, Example of virtual trajectory reconstruction (familiar track and novel arm) from the ensemble place cell activity during Fam-Rest at the ends of the familiar track (20-ms bins; animal velocity, <5 cm s^{-1}) before barrier removal and novel arm exploration. The yellow line reflects the spatial location of the animal in time: the animal was immobile at the junction end of the familiar track. The time-compressed ($\sim$5 m s^{-1}) trajectory reconstruction often 'jumps' over the barrier (top of the figure) into the novel arm area. At around 0.5 s, a preplay of the novel arm initiated from the distal (free) end of the novel arm 'propagates' towards the location of the animal. **c**, Examples of preplay of the novel arm (top), replay of the familiar track (middle) and preplay of the novel arm together with replay of the familiar track (bottom) during Fam-Rest. All conditions are the same as in **b**. The white line shows the linear fit maximizing the likelihood along the virtual trajectory. Colour bars indicate probability of trajectory reconstruction.

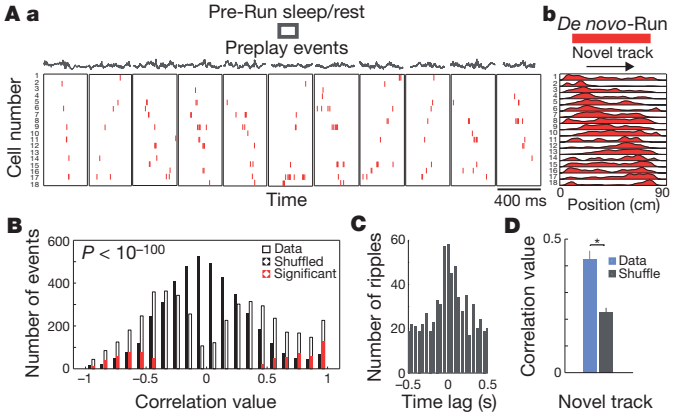

Figure 4 | Preplay of novel place cell sequences before any linear track experience. **A**, Sleep/rest session in the sleep box (Pre-Run sleep/rest) before the first run session on a linear track (*De novo*-Run). Display format is the same as in Fig. 1. **A**, **a**, Representative spiking events in the forward or reverse order during Pre-Run sleep/rest in 400-ms time windows. **A**, **b**, Place cell sequences on the novel track (red) during the *De novo*-Run session. Each row represents one cell in which the activity was normalized to the maximum firing rate. One run direction in one animal is shown. The median number of place cells active on the novel track participating in preplay events is six. **B**, Distribution of spiking events in Pre-Run sleep/rest as a function of the rank-order correlation with the place cell sequence template of the novel track. Display format is the same as in Fig. 2a. **C**, Cross-correlation between preplay events and ripple epochs during Pre-Run sleep/rest. **D**, Stability of place cell spatial tuning across the novel track experience. Display format is the same as in Fig. 2e (novel arm). Error bars, s.e.m.; asterisk indicates significant difference.

During running on a familiar track, some of the cells in the postulated upstream cellular assemblies fire sequentially at spatial locations while others, although connected anatomically to these cells, remain silent. The lack of expression of preplay sequences during Fam-Run may reflect their state-dependent suppression or subthreshold activation during these exploratory behaviours. Owing to increased net excitation during rest periods predominantly during ripples[21], some of these silent cells together with some of the familiar track cells are activated above threshold and fire in a certain sequence. Their sequence of activation may be determined in part by their functional connectivity within the hippocampal formation network. Some of these sequences may in turn be activated on a novel track as place cell sequences (Supplementary Fig. 8). The activation of the novel place cell sequences during running may strengthen their pre-existing assembly organization manifested during preplay.

It could be argued that during Contig-Run the animals simply considered the novel arm to be an extension of the familiar arm and, thus, what we considered to be preplay events were replays of the previous runs on the familiar track. If this was the case, preplay events would not be expected to be found when the experience of the familiar track run is eliminated. This idea was refuted by the demonstration of frequent preplay events in the sleep box before the mice were transferred onto a novel linear track (*de novo* condition). Under this condition, the place cell sequences were more dynamic and a higher proportion of all spiking events correlated with the place cell sequences in these runs than in the later runs on novel linear tracks. These results suggest a shift in the relative contribution of internal[22,23] and external drives in the formation of place cell sequences on encounter with a novel track. In the early phase, internal drives originating in the dynamic cellular assembly activities, which probably reflect numerous past experiences distinct from the current one and expressed as preplay, may have a greater role, whereas in the late phase, external drives that come from the specific set of stimuli of the current experience may dominate. Thus, place cell sequences on novel tracks seem to be products of a dynamic interplay between the internal and external drives.

Several previous studies did not reveal preplay[7,8,10,20]. Although it is difficult to pinpoint the apparent discrepancies between these studies and the present one, we suggest that the use of insufficiently sensitive methods (pairwise correlations) by some studies[7,8,20] and small sample sizes by others[10] might have precluded detection of preplay in previous work (see Supplementary Information for details). Data from the *de novo* condition (Fig. 4), in which we observed an even higher proportion of preplay events, have not been reported previously.

Our data showed that novel preplay events coexist in disjunction with familiar replay events during the rest periods on the familiar track. This and the finding that these preplay and replay events together make up fewer than one-quarter of all detected spiking events suggest that they are part of a dynamic repertoire of temporal sequences in the hippocampus that are past-experience dependent (replay) or future-experience expectant[24] (preplay). Post-experience replay of place cell sequences during resting[3–6] or slow-wave sleep[8–10] has been proposed to have an important role in memory consolidation[11,12]. The temporal preplay of new place cell sequences during resting or sleep is consistent with a predictive function for the hippocampal formation[25] and may contribute to accelerating learning[26]

when a new experience is introduced in multiple steps of increasing novelty.

METHODS SUMMARY

We recorded place cells from the CA1 area of the hippocampus with six independently movable tetrodes in four mice during sleep/rest sessions in the sleep box before any experience on linear tracks and during the first run session on a novel track. Following familiarization with the linear track, animals were subsequently allowed to explore a continuous (L-shaped) track in which the now familiar track and a new novel arm were made contiguous. To quantify the significance of the preplay and replay processes, spiking events in which at least four cells were active were detected during sleep/rest (speed, $<1\ \mathrm{cm\ s^{-1}}$) periods in the sleep box or awake rest (speed, $<2\ \mathrm{cm\ s^{-1}}$) periods at the ends of the familiar track and novel arm, predominantly during ripple epochs.

We calculated statistical significance at the $P<0.025$ level for each event by comparing the rank-order correlation between the event sequence and the place cell sequence (template) with the distribution of correlation values from 200 templates obtained by shuffling the original order of the place cells. Proportions of significant events were calculated as the ratio between the number of significant events and the total number of spiking events. We calculated the overall significance of preplay or replay processes by comparing the distribution of correlation values of all events with the distribution of correlation values of shuffled templates (Kolmogorov–Smirnov test). The significance of the proportion of significant events out of the total number of spiking events was determined as the binomial probability of observing the number of significant events (as successes) from the total number of spiking events (as independent trials), with a probability of success of 0.025 in any given trial. We reconstructed the position of the animal from the spiking activity emitted during resting periods using Bayesian decoding procedures[6].

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 4 December 2009; accepted 29 October 2010.
Published online 22 December 2010.

1. O'Keefe, J. & Nadel, L. *The Hippocampus as a Cognitive Map* (Oxford Univ. Press, 1978).
2. Wilson, M. A. & McNaughton, B. L. Dynamics of the hippocampal ensemble code for space. *Science* **261,** 1055–1058 (1993).
3. Foster, D. J. & Wilson, M. A. Reverse replay of behavioural sequences in hippocampal place cells during the awake state. *Nature* **440,** 680–683 (2006).
4. Diba, K. & Buzsaki, G. Forward and reverse hippocampal place-cell sequences during ripples. *Nature Neurosci.* **10,** 1241–1242 (2007).
5. Karlsson, M. P. & Frank, L. M. Awake replay of remote experiences in the hippocampus. *Nature Neurosci.* **12,** 913–918 (2009).
6. Davidson, T. J., Kloosterman, F. & Wilson, M. A. Hippocampal replay of extended experience. *Neuron* **63,** 497–507 (2009).
7. Wilson, M. A. & McNaughton, B. L. Reactivation of hippocampal ensemble memories during sleep. *Science* **265,** 676–679 (1994).
8. Skaggs, W. E. & McNaughton, B. L. Replay of neuronal firing sequences in rat hippocampus during sleep following spatial experience. *Science* **271,** 1870–1873 (1996).
9. Nádasdy, Z., Hirase, H., Czurko, A., Csicsvari, J. & Buzsaki, G. Replay and time compression of recurring spike sequences in the hippocampus. *J. Neurosci.* **19,** 9497–9507 (1999).
10. Lee, A. K. & Wilson, M. A. Memory of sequential experience in the hippocampus during slow wave sleep. *Neuron* **36,** 1183–1194 (2002).
11. Buzsáki, G. Two-stage model of memory trace formation: a role for "noisy" brain states. *Neuroscience* **31,** 551–570 (1989).
12. Nakashiba, T., Buhl, D. L., McHugh, T. J. & Tonegawa, S. Hippocampal CA3 output is crucial for ripple-associated reactivation and consolidation of memory. *Neuron* **62,** 781–787 (2009).
13. Hebb, D. O. *The Organization of Behavior: A Neuropsychological Theory* (Wiley, 1949).
14. Harris, K. D., Csicsvari, J., Hirase, H., Dragoi, G. & Buzsaki, G. Organization of cell assemblies in the hippocampus. *Nature* **424,** 552–556 (2003).
15. Dragoi, G. & Buzsaki, G. Temporal encoding of place sequences by hippocampal cell assemblies. *Neuron* **50,** 145–157 (2006).
16. Thompson, L. T. & Best, P. J. Place cells and silent cells in the hippocampus of freely-behaving rats. *J. Neurosci.* **9,** 2382–2390 (1989).
17. O'Neill, J., Senior, T. & Csicsvari, J. Place-selective firing of CA1 pyramidal cells during sharp wave/ripple network patterns in exploratory behavior. *Neuron* **49,** 143–155 (2006).
18. Zhang, K., Ginzburg, I., McNaughton, B. L. & Sejnowski, T. J. Interpreting neuronal population activity by reconstruction: unified framework with application to hippocampal place cells. *J. Neurophysiol.* **79,** 1017–1044 (1998).
19. Johnson, A. & Redish, A. D. Neural ensembles in CA3 transiently encode paths forward of the animal at a decision point. *J. Neurosci.* **27,** 12176–12189 (2007).
20. Kudrimoti, H. S., Barnes, C. A. & McNaughton, B. L. Reactivation of hippocampal cell assemblies: effects of behavioral state, experience, and EEG dynamics. *J. Neurosci.* **19,** 4090–4101 (1999).
21. Csicsvari, J., Hirase, H., Czurko, A., Mamiya, A. & Buzsaki, G. Oscillatory coupling of hippocampal pyramidal cells and interneurons in the behaving rat. *J. Neurosci.* **19,** 274–287 (1999).
22. Dragoi, G., Harris, K. D. & Buzsaki, G. Place representation within hippocampal networks is modified by long-term potentiation. *Neuron* **39,** 843–853 (2003).
23. Pastalkova, E., Itskov, V., Amarasingham, A. & Buzsaki, G. Internally generated cell assembly sequences in the rat hippocampus. *Science* **321,** 1322–1327 (2008).
24. Black, J. E. & Greenough, W. T. *Advances in Developmental Psychology* (Lawrence Erlbaum, 1986).
25. Hassabis, D., Kumaran, D., Vann, S. D. & Maguire, E. A. Patients with hippocampal amnesia cannot imagine new experiences. *Proc. Natl Acad. Sci. USA* **104,** 1726–1731 (2007).
26. Tse, D. *et al.* Schemas and memory consolidation. *Science* **316,** 76–82 (2007).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank M. A. Wilson for assistance with data acquisition, discussions and comments on an earlier version of the manuscript; J. O'Keefe, A. Siapas, F. Kloosterman, D. L. Buhl for comments on earlier versions of the manuscript; and F. Kloosterman for providing assistance with the line detection for the Bayesian decoding. This work was supported by NIH grants R01-MH078821 and P50-MH58880 to S.T., who was an HHMI Investigator in an earlier part of this study.

Author Contributions S.T. and G.D. conceived the project jointly. G.D. designed and performed the experiments and the analyses. G.D. and S.T. wrote the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to G.D. (gdragoi@mit.edu) or S.T. (tonegawa@mit.edu).

METHODS

Surgery and experimental design. Electrophysiological recordings were performed on four C57BL/6 mice (strain NR1-floxed[27]) with ages between 18 and 22 weeks. All animals were implanted under Avertin anaesthesia with six independently movable tetrodes aiming for the CA1 area of the right hippocampus (1.5–2 mm posterior to bregma and 1–2 mm lateral to the midline; Supplementary Fig. 1). The reference electrode was implanted posterior to lambda over the cerebellum. During the following week of recovery, the electrodes were advanced daily while animals rested in a small, walled sleeping box ($12 \times 20\,cm^2$, 35 cm high). The animal position was monitored by means of two infrared diodes attached to the headstage.

The experimental apparatus consisted of a $90 \times 65\,cm^2$ rectangular, walled, linear track maze. All tracks were 4 cm wide at the bottom and 8–9 cm wide at the top, and all linear track walls were 10 cm high. Experimental sessions were conducted while the animals explored for chocolate sprinkle rewards placed always at the ends of the corresponding linear tracks (one sprinkle at each end of the track on each lap). Neuronal activity was recorded in naive animals (four mice) during the sleep/rest session in the sleep box immediately preceding the first experience on linear tracks, and continued (Fig. 4) during the first run session on a novel track. After familiarization with the linear track, the animals went through a recording session of 15–60 min (Fam session), and the recordings continued for the next 34–42 min (Contig session) while the animals explored an L-shaped track for the first time. In this track, the familiar arm and the novel arm were made contiguous by removing the barrier that had separated them (Fig. 1). For the purpose of analysing the recording data, the Fam session was further divided into Fam-Run, in which the animals ran through the track (velocity of animal's movement was higher than $5\,cm\,s^{-1}$), and Fam-Rest, where the animals took awake rests at the ends of the track (velocity of animal's movement was less than $2\,cm\,s^{-1}$). During resting periods, the animals consumed the chocolate sprinkle and groomed, but mostly they were still until they self-initiated the next lap of run on the linear track. After completion of the experiments, the brains of all mice were perfused, fixed, sectioned and stained using nuclear fast red (Supplementary Fig. 1) or cresyl violet for electrode track reconstruction.

Recordings and single-unit analysis. A total of 87 neurons were recorded from the CA1 area of the hippocampus in four mice during the Fam and Contig sessions (Supplementary Tables 1–3). A total of 69 CA1 neurons were recorded from the four mice in the *de novo* condition (26, 20, 10 and 13 cells, respectively). Single cells were identified and isolated using the manual clustering method Xclust[2] and the application of cluster quality measurements[28]. Pyramidal cells were distinguished from interneurons on the basis of spike width, average rate and autocorrelations[22].

Place fields were computed as the ratio between the number of spikes and the time spent in 2-cm bins along the track, smoothed with a Gaussian kernel with a standard deviation of 2 cm. Bins where the animal spent a total of less than 0.1 s and periods during which the animal's velocity was below $5\,cm\,s^{-1}$ were excluded. Place field length and peak rate were calculated after separating the direction of movement and linearizing the trajectory of the animal. Linearized place fields were defined as areas with a localized increase in firing rate above 1 Hz for at least five contiguous bins (10 cm). The place field peak rate and location were given by the rate and location of the bin with the highest ratio between spike counts and time spent. Place field borders were defined as the points where the firing rate became less than 10% of the peak firing rate or 1 Hz (whichever was bigger) for at least 2 cm.

Local field potential analysis. Ripple oscillations were detected during sleep/rest periods in the sleep box and during rest periods at the ends of the tracks. The electroencephalography signal was filtered (120–200 Hz) and ripple-band amplitude was computed using the Hilbert transform. Ripple epochs with maximal amplitude more than 5 s.d. above the mean, beginning and ending at 1 s.d. were detected. The time of ripple occurrence (Figs 2c and 4C) was the time of its maximal amplitude. The proportion of ripples with which cells with place fields on the novel arm of the L-shaped track fired in the preceding session (Supplementary Fig. 3) was calculated for each qualifying cell as the ratio between the number of ripples during which the cell fired at least one spike and the total number of ripples during the corresponding exploratory session.

Preplay and replay analyses. To analyse the preplay and replay processes, spiking events were detected during Pre-Run sleep/rest periods in the sleep box (*de novo* condition; velocity, $<1\,cm\,s^{-1}$) or during awake rest periods at the ends of the running tracks (Contig condition; velocity, $<2\,cm\,s^{-1}$). A spiking event was defined as a transient increase in the firing activity of a population of at least four different place cells within a temporal window preceded and followed by at least 50 ms of silence. Overall, similar results were obtained using 50-, 60-, 75- and 100-ms time windows. The spikes of all the place cells active on the novel track that were emitted during the Pre-Run sleep/rest in the box for the *de novo* condition as well as the spikes of all the place cells active on the familiar track or the novel arm that were emitted during Fam-Rest session at the two ends of the familiar track for Contig condition were respectively sorted by time and further used for the detection of the spiking events.

All four animals exhibited a significant number of spiking events in the Pre-Run session of the *de novo* condition. Three of the four animals (mice 1–3) exhibited a significant number of spiking events in the Contig condition, the remaining animal (mouse 4) having a below-threshold number of simultaneously active CA1 place cells. The time of the spiking event used to compute the cross-correlation with ripple epoch occurrence (Figs 2c and 4C) was the average time of all spikes comprising the spiking event. The place cell sequences (templates) were calculated for each direction of the animal's movement and for each run session (*De novo*-Run, Fam-Run and Contig-Run) by ordering the spatial location of the place field peaks that were above 1 Hz. For place cells with multiple place fields above 1 Hz on a particular arm or track in the Contig condition (six of 52 place cells active on the novel arm in the two directions, or 12%: two for each direction in mouse 1, one in mouse 2 and one in mouse 3), only the place field corresponding to the peak firing rate of the place cell on that arm or track was considered for the construction of the template of that particular arm or track, to be consistent with all the previous studies that used spatial templates to demonstrate replay during sleep or awake rest[3,4,10]. Place cells with fields on both the novel arm in the Contig-Run session and the familiar track in the Fam-Run session participated in the construction of both the novel arm and familiar track templates.

Statistical significance was calculated for each event by comparing the rank-order correlation between the sequence of cells' firing during the event (that is, event sequence) and the place cell sequence (template), on the one hand, and the distribution of correlation values between the event sequence and 200 surrogate templates obtained by shuffling the order of place cells, on the other[4] (Fig. 2a). The significance level was set at 0.025 to control for multiple comparisons (two directions of run). The proportions of significant events (preplay novel track, preplay novel arm (Fig. 2b), replay novel arm and replay familiar track) were each calculated as the ratio between the number of significant events and the total number of spiking events in which at least four corresponding place cells were active[4]. Corresponding familiar track templates (Fig. 2h) were constructed by ordering the location of peak firing on the familiar track during Fam-Run (no minimum threshold of firing) of all place cells that subsequently fired on the novel arm. Cells comprising the corresponding familiar track templates are the same as those comprising the novel arm templates. We note that these corresponding familiar track templates are different from the ones used in Figs 1 and 2a–g, which were constructed by ordering the peak firing of all place cells active on the familiar track >1 Hz.

The overall significance of the preplay (Fig. 2a) or replay process was calculated by comparing the distribution of correlation values of all events relative to the original template with the distribution of correlation values relative to the shuffled surrogate templates, using the Kolmogorov–Smirnov test[3]. Quantification of the replay versus preplay events during the Fam-Run session (Fig. 2f, g) was performed as described above using different spatial templates for the familiar track and the novel arm. All spiking events were correlated with both the novel arm and the familiar track templates. Events significantly correlated only with familiar track or with novel arm templates were considered pure replay and pure preplay, respectively. The template specificity index was calculated for each event as the difference between the absolute value of the event's correlation with the novel arm template (preplay, high positive index) and the event's correlation with the familiar track template (replay, high negative index). For the purpose of displaying the template specificity index, events correlated with the novel arm but not with the familiar track templates were considered preplay and events correlated with the familiar track but not with the novel arm templates were considered replay (Fig. 2g). Additionally, events correlated with both the familiar track and the novel arm templates formed a third group, preplay/replay events, displayed in yellow in the inset of Fig. 2g.

Correlations between pairs of familiar track and novel arm templates (Fig. 2h) were performed using modified familiar track templates that were constructed using the location of peak firing (>0 Hz) of only those cells that had place fields on the novel arm (peak rate, >1 Hz). The lack of significant correlation in this case demonstrates that the novel arm place cell sequence is not simply a transposition of a familiar track place cell sequence on the novel arm.

We also identified neurons that did not fire during Fam-Run, that activated during Fam-Rest events and that corresponded to trajectories on the novel arm during Contig-Run (silent cells). We calculated the correlation between the order in which they fired during Fam-Rest events and their spatial sequence as new place cells on the novel arm during Contig-Run, as previously explained. Owing to the low absolute number of silent neurons, only triplets of cells were available for further analysis ($n = 24$). The proportion of events perfectly matching the spatial template was compared with the proportion of by-chance perfect matching (0.33).

Stability of place cell maps. Stabilities of place cell firing on the familiar track before and after barrier removal as well as on the novel track (*de novo* condition) and the novel arm (Contig condition) in the beginning versus the end of the run session were assessed by calculating, for each place cell and each direction, a correlation between the spatial firing in the corresponding paired situations (before versus after barrier removal for the familiar track or the first four laps versus the last four laps of the *De novo*-Run or Contig-Run session for the novel track or arm, respectively). The place cell activity was not partitioned in place fields; rather, the whole activity on the particular track or arm was considered separately for each cell and direction (average correlations are shown in Figs 2e and 4D, blue bars). In addition, we calculated the same type of correlation after shuffling the identity of the cell in one member of the correlation (once for each different cell; average correlations are in Figs 2e and 4D, black bars). Shuffle results (Figs 2e and 4D, black bars) were computed as correlation between spatial tuning of cells on the familiar track during Fam-Run and spatial tuning of all other simultaneously recorded cells on the familiar arm during Contig-Run (familiar track group; Fig. 2e, left), or correlation between spatial tuning of cells on the novel arm (or novel track) during the beginning of Contig-Run (or *De novo*-Run) and spatial tuning of all the other simultaneously recorded cells on the novel arm (or novel track) during the end of Contig-Run (novel arm group; Fig. 2e, right) or *De novo*-Run (Fig. 4D). Original and shuffled correlations were compared using the rank-sum test. The average number of laps (traversal of the novel track in both directions) per session was 20.5 in *De novo*-Run (21, 16, 27 and 18 in the four mice) and 16.3 in Contig-Run (13, 14 and 22 in the three mice).

Bayesian reconstruction of actual and virtual trajectories. For each cell, we calculated a linearized spatial tuning curve on the familiar track during the Fam-Run session and a linearized spatial tuning curve on the novel arm during the Contig-Run session. The tuning curves were constructed in 2-cm bins from spikes emitted in both run directions at velocities higher than 5 cm s^{-1}, and were smoothed with a Gaussian kernel with a standard deviation of 2 cm. We constructed a joint spatial tuning curve for each cell by juxtaposing the spatial tuning curve on the familiar track during the Fam-Run session and the spatial tuning curve on the novel arm during the Contig-Run session. We also detected for each cell all the spiking activity emitted at velocities below 5 cm s^{-1} during the Fam-Rest session, where replay and preplay events where shown to occur using the rank-order correlation method. We used a Bayesian reconstruction algorithm[6,18] to decode the virtual position of the animal from the spiking activity during Fam-Rest (Fig. 3b) in non-overlapping, 20-ms bins using the joint spatial tuning curves. We then extracted epochs of reconstructed trajectory matching the time of the spiking events as detected using multiunit activity of place cells from the familiar track and novel arm (rank-order correlation method; see 'Preplay and replay analyses', above).

We used two shuffling procedures to measure the quality of the Bayesian decoding. In the first shuffling procedure, for each event, the original time-bin columns of the probability distribution function (PDF) were replaced with an equal number of time-bin columns randomly extracted from a pool containing the time-bin columns of all PDFs of all detected events[6]. The shuffling procedure was repeated 500 times. In the second shuffling procedure, the identity of the place cells was randomly shuffled 100 times and new PDFs were calculated for all events. For all original and shuffled PDFs, a line was fitted to the data using a previously described line-finding algorithm[6]. Lines fitted to the original and shuffled data were compared using slope, spatial extent, location on the track and probability score. We defined replay and preplay as the epochs of Fam-Rest in which the reconstructed trajectory was located on the familiar track or the novel arm, respectively. The trajectory was defined across a set of position estimates during the corresponding epoch (Fig. 3c). Only epochs that lasted at least 60 ms (three bins) and which contained reconstructed trajectories spanning at least 10 cm were considered for further analysis. Trajectories for which 75% or more of their length was located on the familiar track were considered to represent replay of an animal's trajectory on the familiar track (Fig. 3c, middle), and trajectories for which 75% or more of their length was located in the novel arm were considered to represent preplay of the animal's future trajectory on the novel arm (Fig. 3c, top). The remaining events were considered preplay–replay (Fig. 3c, bottom).

An epoch was considered significant if the new line was less than 75% contained in the familiar track for replay or novel arm for preplay in at least 95% of the shuffled cases. For each epoch that was significant for replay or preplay using the reconstruction method, we retrieved the value of the rank-order correlation between the neuronal firing sequences and the familiar track and novel arm spatial templates as calculated using the rank-order correlation method. We compared the absolute correlation values between the epoch's firing sequences and familiar track templates with the absolute correlation values between the same epoch's firing sequences and novel arm templates. We also reconstructed the trajectory of the animal on the familiar track from the spiking activity during the Fam-Run session at velocities above 5 cm s^{-1} in 250-ms bins using the spatial tuning curves on the familiar track[6,18] (Fig. 3a) to validate the decoding procedure.

27. Tsien, J. Z., Huerta, P. T. & Tonegawa, S. The essential role of hippocampal CA1 NMDA receptor-dependent synaptic plasticity in spatial memory. *Cell* **87,** 1327–1338 (1996).
28. Schmitzer-Torbert, N., Jackson, J., Henze, D., Harris, K. & Redish, A. D. Quantitative measures of cluster quality for use in extracellular recordings. *Neuroscience* **131,** 1–11 (2005).