Distinct effects of reward and navigation history on hippocampal forward and reverse replays

Baburam Bhattaraia,b, Jong Won Leea,b, and Min Whan Junga,b,1

Dysfunctions, Institute for Basic Science, 34141 Daejeon, Korea; and bCenter for Synaptic Brain Dysfunctions, Institute for Basic Science, 34141 Daejeon, Korea

To better understand the functional roles of hippocampal forward and reverse replays, we trained rats in a spatial sequence memory task and examined how these replays are modulated by reward and navigation history. We found that reward enhances both forward and reverse replays during the awake state, but in different ways. Reward enhances the rate of reverse replays, but it increases the fidelity of forward replays for recently traveled as well as other alternative trajectories heading toward a rewarding location. This suggests roles for forward and reverse replays in reinforcing representations for all potential rewarding trajectories. We also found more faithful reactivation of upcoming than already rewarded trajectories in forward replays. This suggests a role for forward replays in preferentially reinforcing representations for high-value trajectories. We propose that hippocampal forward and reverse replays might contribute to constructing a map of potential navigation trajectories and their associated values (a "value map") via distinct mechanisms.

Results

Behavior. Three rats were trained in a spatial sequence memory task (Fig. 1A). The maze contained a figure 8-shaped section through which the rats could navigate in 1 of 4 different spatial sequences (left–left (LL), left–right (LR), right–left (RL), and right–right (RR)). The rats performed 2 consecutive sessions each day with each session comprising 4 blocks of trials associated with 4 different correct sequences. The order of the correct block sequences was randomized. Each block consisted of 5 forced-choice trials followed by 15 free-choice trials, and a short delay of 3 s was imposed in the delay zone in each trial. A fixed amount (30 μL) of water reward was provided at a designated location in the forced-choice trials and also in free-choice trials when the animal chose the same spatial sequence as in the forced-choice trials (correct trials), but not when the animal chose a different spatial sequence (error trials). One animal’s movement trajectory in one sample recording session is shown in Fig. 1B. As shown, in well-trained rats, the movement trajectories and head directions varied across different blocks of trials in the awake (12) states induces subsequent deficits in spatial–memory performance. This suggests that SWR-associated hippocampal replays contribute to hippocampal memory functions rather than representing a passive by-product of experience-induced network activity (9, 13, 14).

Although the functions of hippocampal replays are not clearly understood, they have been proposed to play roles in memory retrieval (8, 15), memory consolidation (4, 7, 10, 16), learning from experience (5, 17–19), planning future navigation (6, 20), linking recent and remote memories (8), and building cognitive maps (21). In this study, to better understand the function of hippocampal replays, we designed a spatial sequence memory task in which rats had to take 1 of 4 different trajectories to obtain a reward at a designated location. This allowed us to examine hippocampal replays related not only to a recently experienced trajectory, but also to other trajectories that lead to a rewarding location. Using this task, we asked how reward and navigation history affect hippocampal replays during the awake state. Because hippocampal place-cell activation patterns are replayed during SWRs in the same order as they fired during the experienced spatial trajectory as well as in reverse (6, 22), we examined the effects of reward and navigation history on forward and reverse replays separately. We found that reward enhances both forward and reverse replays, but in different ways. Reward enhances the rate of reverse replays, while increasing the fidelity of forward replays for recently traveled as well as other trajectories heading toward a rewarding location. We also found that forward replays show more faithful reactivation of upcoming rewarding trajectories than already rewarded trajectories. We propose, based on these findings, that hippocampal forward and reverse replays might contribute to building a map of potential navigation trajectories with associated values via distinct mechanisms.

Significance

Sequential firing patterns of hippocampal place cells during spatial navigation are replayed in both forward and reverse directions during offline idling states, but functional roles of these replays are not clearly understood. To clarify this matter, we examined hippocampal replays in a spatial sequence memory task. We found that reward increases the fidelity of forward replays, while increasing the rate of reverse replays for all possible trajectories heading toward a rewarding location. We also found more faithful reactivation of place-cell sequences for upcoming rewarding trajectories than for already rewarded trajectories in forward replays. Our results suggest that hippocampal forward and reverse replays contribute to constructing a map of potential navigation trajectories with associated values ("value map") via distinct mechanisms.

Author contributions: B.B., J.W.L., and M.W.J. designed research; B.B. performed research; B.B., J.W.L., and M.W.J. analyzed data; and B.B. and M.W.J. wrote the paper.

The authors declare no competing interest.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1912533117/-/DCSupplemental.

First published December 23, 2019.

www.pnas.org/cgi/doi/10.1073/pnas.1912533117

PNAS | January 7, 2020 | vol. 117 | no. 1 | 689–697
Spatial Firing. Throughout the study, we linearized the animal’s spatial position by projecting each spatial position in the maze onto the mean spatial trajectory for each session (SI Appendix, Fig. S1A). We recorded 698 single units from the dorsal CA1 region, and 648 of them (92.8%) were classified as putative pyramidal cells (SI Appendix, Fig. S2). Because we were interested in how replay events relate to place-cell firing along different spatial trajectories, and because CA1 spatial firing varies across different spatial trajectories (23–26) and over time (27–29), we determined the place cells for each trial block and for each session separately using forced-choice and correct free-choice trials. We determined a total of 479 (73.9%) and 483 (74.5%) pyramidal cells as place cells in at least one block of sessions 1 and 2, respectively (the number of place cells in each session set is shown in SI Appendix, Table S1). Fig. 2A shows the spatial firing of all place cells arranged according to the position of peak firing on a linearized maze. As shown, the spatial firing pattern differed between blocks not only in the track zone, but also in the rest of the maze. Note that variations in actual trajectories and head directions can explain block (i.e., trajectory)-dependent firing in the track zone, but not in the other zones (SI Appendix, Fig. S1). The spatial firing pattern also differed between sessions 1 and 2 even for the same spatial sequences. Thus, CA1 spatial firing varied as a function of both spatial trajectory and session.

Fig. 2B and E shows 4 sample place cells that showed trajectory- and/or session-dependent firing that could not be explained by variations in spatial trajectory (additional examples are shown in SI Appendix, Fig. S3). Fig. 2B is a sample of place-cell firing that showed significant session-dependent firing (analysis of covariance [ANCOVA], main factors, session, and trajectory) in the return zone. The firing of this place cell in the return zone was stronger in session 1 than in session 2. Fig. 2C is a sample of a place cell that showed significant trajectory and trajectory × session interaction effects in the return zone. The firing of this place cell in the return zone was stronger during RR and RL than the other sequences in session 1, but not in session 2. Fig. 2D is a sample of the place-cell firing that showed significant trajectory, session, and trajectory × session interaction effects in the return zone. The firing of this place cell in the return zone was stronger in RR and LR than the other sequences in session 1, but not in session 2. Fig. 2E is a sample place-cell firing that showed significant trajectory and session effects in the reward zone. The majority of place cells showing significant trajectory-dependent firing outside the track zone showed significant trajectory × session interaction effects as well (134 of 185; 72.4%). Also, the majority of session-sensitive place cells showed abrupt changes in activity across sessions 1 and 2 (Fig. 2 and SI Appendix, Fig. S3). This indicates that the session-dependent firing that we observed in our study is more likely to reflect nonspatial context (task structure)-dependent firing rather than gradual changes in firing over time (27, 28). As shown by these examples, both global remapping and rate remapping (30) were observed across blocks and/or sessions.

We performed additional analyses to examine trajectory- and/or session-dependent firing of place cells using correct free-choice trials. We first examined the proportions of place cells that showed trajectory- and/or session-dependent firing using ANCOVA (main factors: trajectory and session; cofactors: lateral deviation of head position, head direction, and speed). The proportion of place cells with significant main effects of trajectory was ~60% in the track zone and >30% in the rest of the maze. This indicates that spatial trajectory strongly modulates CA1 spatial firing. The proportion of place cells with significant main effect of session was >30%, and those with the significant trajectory × session interaction was >40% in all parts of the maze (Fig. 2F). The majority of trajectory- and/or session-sensitive neurons were both trajectory- and session-sensitive. Only a small number of cells showed strictly trajectory- or session-dependent firing (ANCOVA, main effect of only trajectory without significant trajectory × session interaction, n = 27, 24, and 15 in the delay, reward, and return zones, respectively; main effect of session only without significant trajectory × session interaction, n = 32, 38, and 34 in the delay, reward, and return zones, respectively).

We also calculated the percentage of place cells that had place fields in 1, 2, 3, or all 4 blocks in a session using correct free-choice trials. Sixty-one and 65% of place cells maintained their firing fields inside the track zone in all 4 blocks in sessions 1 and 2, respectively, while 75 and 81% did so outside the track zone in sessions 1 and 2, respectively (Fig. 2G). In addition, we quantified
Fig. 2. Spatial firing. (A, Left) Schematic showing different spatial zones of the maze. (A, Right) Normalized spatial firing rates of all place cells (determined based on forced-choice and correct free-choice trials) are shown as heat maps on the linearized maze for each sequence block. Each row shows spatial firing of one place cell. Place cells were aligned according to the position of the maximal firing rate during the RR block of the first session. Vertical dashed lines divide different spatial zones on the maze. Rt, return; D, delay; T, track; Rw, reward. Triangles indicate sample place cells shown in the trajectory and spatial firing-rate plots (Bottom; smoothed by a Gaussian kernel with σ = 4 cm; shading, SEM) are shown along the linearized maze separately for sessions 1 (Left) and 2 (Right). Colors denote different sequence blocks. P values for the main effects of trajectory and session as well as trajectory × session interaction (ANCOVA) are indicated for each sample place cell. (F) Percentages of place cells in each zone that were significantly responsive to trajectory, session, and/or the trajectory × session interaction (ANCOVA, P < 0.05). The horizontal dashed line indicates the level of chance (binomial test, alpha = 0.05). (G) Pie charts showing percentages of place cells with peak firing inside (Top) or outside (Bottom) the track zone that had place fields in 1, 2, 3, or 4 blocks in each session (Left; session 1; Right; session 2). (F and G) Each block was divided in half (~7 trials each), and the difference in place-field peak-firing locations (H; Left; cumulative plot; Right; mean ± SEM) and spatial correlation of place-cell population vectors (I; Left; cumulative plot; Right; mean ± SEM) were computed between the halves of each block (within block), between the adjacent halves of adjacent blocks (between blocks), and between the corresponding halves of identical blocks of the 2 sessions (between sessions) for the track zone and the rest of the maze (outside track). ***P < 0.0005 (Kruskal–Wallis tests followed by Bonferroni’s post hoc tests).

### A

- **Return zone**
  - Trajectory: P = 0.425, Session: P = 0.001
  - Session 1: P = 0.235

- **Track**
  - Session 1: P = 0.015
  - Session 2: P = 2.5 x 10^-8

- **Reward zone**
  - Session 1: P = 0.6 x 10^-5
  - Session 2: P = 1.6 x 10^-5

### B

- **Return zone**
  - Trajectory: P = 0.009, Session: P = 5.9 x 10^-8

- **Track**
  - Session 1: P = 4.7 x 10^-6
  - Session 2: P = 0.001

### C

- **Return zone**
  - Trajectory: P = 0.015, Session: P = 0.147

- **Track**
  - Session 1: P = 2.1 x 10^-3
  - Session 2: P = 0.079

### D

- **Return zone**
  - Trajectory: P = 0.009
  - Session: P = 5.9 x 10^-8

- **Track**
  - Session 1: P = 4.7 x 10^-6
  - Session 2: P = 0.001

### E

- **Reward zone**
  - Trajectory: P = 0.015
  - Session: P = 0.147

- **Track**
  - Session 1: P = 2.1 x 10^-3
  - Session 2: P = 0.079

### F

- **Percentage of significant blocks**
  - Return: 16%
  - Delay: 61%
  - Track: 4%
  - Session 1: 6%
  - Session 2: 4%

### G

- **Pie charts**
  - **Track**
    - Session 1: 16%
    - Session 2: 16%

  - **Outside-track**
    - Session 1: 18%
    - Session 2: 18%

### H

- **Cumulative fraction**
  - **Track**
    - Session 1: 75%
    - Session 2: 81%

- **Outside-track**
  - Session 1: 75%
  - Session 2: 81%
the change in place field location for the place cells that did not undergo global remapping. For this, we divided each block into halves (correct trials only; ~7 trials) and compared the peak firing positions of the place fields between the halves of the same block across 2 adjacent blocks of the same session (to examine trajectory effect) and across 2 sessions for identical sequence blocks (to examine session effect; Fig. 2F). The shift in peak-firing position varied significantly across the 3 comparisons (Kruskal–Wallis test, track zone, \( \chi^2(3) = 208.686, P = 4.8 \times 10^{-36} \), \( n = 1,242, 1,147, \) and 581 place-field pairs, respectively; outside-track zone, \( \chi^2(2) = 143.294, P = 7.7 \times 10^{-32} \), \( n = 1,951, 1,886, \) and 919 place-field pairs, respectively). It was significantly greater between adjacent blocks than within the same block (Bonferroni’s post hoc test, track zone, \( P = 9.8 \times 10^{-47} \); outside-track zone, \( P = 4.4 \times 10^{-27} \); data from sessions 1 and 2 combined) as well as between sessions than within a session for the same sequence (track zone, \( P = 1.3 \times 10^{-6} \); outside-track zone, \( P = 1.7 \times 10^{-18} \); Fig. 2F). Finally, to compare the overall changes in place fields due to changes in firing rate and shifts in place fields, we calculated population vector correlations for within-block, between-block, and between-session pairs. We found that spatial correlation of place-cell population vectors (30) varied significantly across the 3 comparisons (Kruskal–Wallis test, track zone, \( \chi^2(3) = 149.788, P = 2.9 \times 10^{-33} \), \( n = 205, 179, \) and 204 population vector pairs, respectively; outside-track zone, \( \chi^2(2) = 189.753, P = 6.2 \times 10^{-42} \), \( n = 205, 179, \) and 204 population vector pairs, respectively). The spatial correlation was significantly greater within the same block than between adjacent blocks (Bonferroni’s post hoc test, track zone, \( P = 2.8 \times 10^{-33} \); outside-track zone, \( P = 1.2 \times 10^{-27} \)) and between the same blocks of each pair of sessions (track zone, \( P = 8.5 \times 10^{-16} \); outside-track zone, \( P = 9.4 \times 10^{-26} \); Fig. 2F). These results indicate strong effects of trajectory and session on the spatial firing of place cells in our task.

Replay Events. The majority of SWRs were observed in the reward zone (\( n = 76,609, 57.3\% \); delay zone, \( n = 7,603, 5.7\% \); the rest of the maze, \( n = 48,453, 36.5\% \)). Instead of waiting quietly, the rats moved their heads frequently during the delay period, and theta oscillations were stronger in the delay than in the reward zone (SI Appendix, Fig. S4). This likely explains why we observed SWRs so infrequently in the delay zone. We therefore focused our analysis on SWR-associated neural activity recorded in the reward zone.

Fig. 3A shows a neural decoding of the animal’s position on a linearized maze averaged across all sessions (SI Appendix, Materials and Methods). We found that the averaged spatial position predicted by neuronal ensemble activity in 50-ms time bins correlated well with the actual position of the animal in the track, return, and delay zones, each of which showed relatively high movement velocity (track, 15.8 ± 0.05; return, 17.5 ± 0.06; delay, 12.2 ± 0.12 cm/s). In the reward zone, however, where average movement velocity was low (7.3 ± 0.27 cm/s), the estimated position was broadly distributed throughout the maze (cf. ref. 31). This is consistent with the finding that hippocampal replays in the waking state are observed mostly during immobility (6, 20).

We identified a candidate replay event as a set of spikes emitted by at least 5 different place cells from a given block preceded by a >60-ms silent period and with at least one coincident SWR (Fig. 3B–E). Thus, we restricted our analysis to replay events related to the spatial firing of a particular block of interest. We constructed place-cell templates using forced-choice and correct free-choice trials for each block and analyzed replay events during free-choice trials unless otherwise noted. We found 2,573 candidate replay events that are related to the current trajectory (correct path in the current block), and their mean (±SD) duration was 151.9 ± 61.5 ms. We combined the analysis results from sessions 1 and 2 to increase the sample size and included in the analysis only significant replay events that were determined with a permutation test (SI Appendix, Materials and Methods). We then calculated a weighted correlation (\( r \)) (SI Appendix, Materials and Methods) (32–34) for each significant replay event, a measure of how well the decoded positions match

![Fig. 3. Candidate replay events.](https://www.pnas.org/cgi/doi/10.1073/pnas.1912533117)
to a linear trajectory corresponding to the correct path in the current block. We then divided significant replay events for the current trajectory into forward and reverse replays based on the signs of their weighted correlation coefficients (forward replays, \( r > 0 \); reverse replays, \( r < 0 \)). Forward replays were significantly more frequent than reverse replays (707 and 625, respectively, \( \chi^2 \)-test, \( \chi^2 = 6.811, P = 0.009 \)).

**Effect of Reward on Replays of the Current Trajectory.**

*Place cell reactivation during SWR.* Prior to examining effects of reward on replay, we first examined how reward affects reactivation of place cells during SWRs in the reward zone. For each block, we determined place cells using forced-choice and correct trials and compared how they are reactivated during SWRs in correct versus error trials. The animals stayed significantly longer in the reward zone in correct compared to error trials (time spent at the reward zone, mean \( 16.53 \pm 4.10 \) s and \( 6.18 \pm 9.47 \) s, respectively; signed-rank test, \( z = 4.008, P = 6.1 \times 10^{-5} \)). As a consequence, the number of SWRs was significantly larger in correct than in error trials (0.61 \( \pm 0.13 \) and 0.13 \( \pm 0.04 \) per trial, respectively; signed-rank test, \( z = 4.084, P = 4.4 \times 10^{-5} \); SI Appendix, Fig. S5). When the number of SWRs was normalized to the time spent in the reward zone, the rate of SWR was still higher in correct than in error trials (signed-rank test, \( z = 2.571, P = 0.010, n = 25 \) session sets, Fig. 4/4). We also calculated the normalized difference between correct and error trials in place-cell activation per SWR. For each place cell in the current block, we calculated its activation probability in the SWRs of correct and error trials. We then normalized the difference in activation probabilities between correct and error trials with the SE of the difference based on a binomial distribution to obtain a \( z \) score (\( \Delta pZ_{\text{reward}} \)). Place cells were activated at higher probabilities per SWR in correct than in error trials so that \( \Delta pZ_{\text{reward}} \) was significantly larger than zero (signed-rank test; \( z = 4.441, P = 8.9 \times 10^{-6}, n = 627 \) place cells; Fig. 4B).

We performed additional analyses to explore whether higher reactivation of place cells in correct than in error trials can be explained by factors other than reward. For example, higher reactivation of place cells in correct trials might be because of an experience effect. If recently activated place cells during navigation are more likely to be reactivated during SWRs, those place cells the fields of which are crossed by the animal in correct, but not error, trials would be more likely to be reactivated in correct than in error trials. This is a confounding factor in our analysis because we determined place cells using correct and forced-choice, but not error, trials. To test this possibility, we examined reactivation of place cells within the track zone and outside the track zone separately. We found similar reward effects for the 2 groups of place cells (signed-rank test; place cells inside track, \( \Delta pZ_{\text{reward}} \), \( z = 2.336, P = 0.019, n = 258 \); place cells outside track, \( \Delta pZ_{\text{reward}} \), \( z = 3.827, P = 1.3 \times 10^{-4}, n = 369 \), Fig. 4C). Because navigation trajectory outside the track zone was similar across blocks (SI Appendix, Fig. S1), enhanced reactivation of outside-track-zone place cells in correct trials cannot be attributed to an experience effect. Also, for the place cells within the track zone, we selected those fields that were crossed by the animal during both correct and error trials. These place cells were reactivated at higher probabilities per SWR in correct

**Fig. 4.** Effect of reward on replays of the current rewarding trajectory. (A–D) Effect of reward on place-cell reactivation during SWRs. (A) Mean \( \pm \) SEM rates (Hz) of SWR in correct and error trials. \( P < 0.05 \) (signed-rank test). (B) Place-cell activation probabilities (\( p_{\text{act}} \)) per SWR in correct and error trials shown as cumulative plots. (B, Middle and Right) Normalized difference in \( p_{\text{act}} \) between correct-trial and error-trial SWRs (\( \Delta pZ_{\text{reward}} \)) shown as a cumulative plot (Middle) or mean \( \pm \) SEM (Right). \( * * * P < 0.0005 \) (signed-rank test for difference from 0). (C and D) \( \Delta pZ_{\text{reward}} \) (Left, cumulative plot; Right, mean \( \pm \) SEM) is shown for place cells inside and outside the track zone separately (C) and for place cells inside the track zone the fields of which were crossed by the animal in both correct and error trials (D). \( P < 0.05 \). \( * * * P < 0.0005 \) (signed-rank test for difference from 0). (E–I) Effect of reward on forward and reverse replays. (E) Mean \( \pm \) SEM rates (Hz) of forward and reverse replays per SWR in correct and error trials. \( * * * P < 0.0005 \) (signed-rank test). (F) Mean \( \pm \) SEM numbers of forward and reverse replays per SWR in correct and error trials. (G) Absolute weighted correlations (\( r \)) of forward and reverse replays (Left, cumulative plot; Right, mean \( \pm \) SEM). \( * P < 0.05 \) (rank-sum test). (H and I, Left) Activation (\( p_{\text{act}} \); Middle and Right) Normalized difference in place cells in forward and reverse replays in correct and error trials shown as cumulative plots. (H and I, Middle and Right) Normalized difference in \( p_{\text{act}} \) (H) or \( p_{\text{coact}} \) (I) between correct-trial and error-trial replays (\( \Delta pZ_{\text{reward}} \)) shown as a cumulative plot (Middle) or mean \( \pm \) SEM (Right). \( * P < 0.005, ** P < 0.0005 \) (signed-rank test for difference from 0; rank-sum test for comparisons between forward and reverse replays).
than in error trials (signed-rank test, $\Delta p_{\text{reward}}, z = 1.994, P = 0.046, n = 102$ place cells, Fig. 4D). Again, the difference cannot be attributed to an experience effect because it is identical for the analyzed place cells. We then examined whether place-cell reactivation differs between forced-choice and error trials. We found that place-cell reactivation probability per SWR was significantly higher in forced-choice than in error trials, but similar between forced-choice and correct trials (SI Appendix, Fig. S6B). Thus, reward, rather than being correct or incorrect, seems to be a determining factor in enhancing place-cell reactivation. Collectively, these results consistently support the possibility that reward enhances not only SWR rate but also place-cell reactivation per SWR, which is consistent with previous reports (18, 19).

**Replay rate.** We then assessed the effects of reward on forward and reverse replays by comparing replays during correct and error trials. We first examined how reward affects replays of the current trajectory (i.e., sequential place-cell firing along the correct trajectory in the ongoing block). Replay rate (Hz) was significantly higher in correct than in error trials for reverse replays, but not for forward replays (signed-rank tests; forward replays, $z = 1.429, P = 0.153, n = 24$ session sets; reverse replays, $z = 3.269, P = 0.001, n = 24$ session sets; Fig. 4E). The numbers of forward and reverse replays per SWR did not differ significantly between correct and error trials (signed-rank tests; forward replay, $z = 0.254, P = 0.210, n = 17$ session sets; reverse replay, $z = 0.414, P = 0.679, n = 17$ session sets; Fig. 4F). These results indicate that reward increases the rate of reverse but not of forward replays, which is consistent with previous reports (19, 35).

**Replay fidelity.** We then compared the absolute weighted correlation ($\langle r \rangle$) (33, 34) between correct and error trials as a measure for the fidelity of replays. The absolute weighted correlation was significantly higher in correct trials than in error trials for forward replays (rank-sum test; $z = 2.666, P = 0.008, n = 500$ and 36), but not for reverse replays ($z = 0.942, P = 0.346, n = 433$ and 30; Fig. 4G).

As additional measures, we compared the probability for a place cell to be activated in replays ($p_{\text{act}}$) as well as the probability for 2 place cells to be activated together in replays ($p_{\text{coact}}$) between correct and error trials. Specifically, we calculated the normalized differences in $p_{\text{act}}$ and $p_{\text{coact}}$ between correct and error trials. For each place cell (or place-cell pair) in the current block, we calculated its probability of activation (or coactivation) in the replays of correct and error trials. Then we normalized the difference in activation (or coactivation) probabilities between correct and error trials with the SE of the difference based on a binomial distribution to obtain a $z$ score ($\Delta p_{\text{reward}}$) (36). For both $p_{\text{act}}$ and $p_{\text{coact}}$, $\Delta p_{\text{reward}}$ was significantly larger than zero in forward replays (signed-rank test, $P_{\text{act}} z = 3.367, P = 7.6 \times 10^{-4}, n = 222$ place cells; $P_{\text{coact}} z = 7.915, P = 4.6 \times 10^{-11}, n = 731$ place-cell pairs), but not in reverse replays ($P_{\text{act}} z = 0.219, P = 0.827, n = 102$ place cells; $P_{\text{coact}} z = 2.872, P = 0.004, n = 430$ place-cell pairs; Fig. 4H and I). The normalized difference in $p_{\text{coact}}$ ($\Delta p_{\text{reward}}$) was also significantly larger in forward than in reverse replays (rank-sum test; $P_{\text{coact}} z = 1.918, P = 0.055; P_{\text{coact}} z = 2.837, P = 0.005$; Fig. 4I).

Additional analyses indicated that the observed differences between correct and error trials are not because of a systematic change in the animal’s behavior over time [e.g., an ‘automation’ effect (37)] (SI Appendix, Fig. S7) or the difference in recent navigation experience between correct and error trials (SI Appendix, Fig. S8). We also found similar differences between forced-choice and correct trials (SI Appendix, Fig. S8). These results consistently support the conclusion that reward increases the rate of reverse replays, while enhancing the fidelity of forward replays.

**Effect of Reward on Replays of Alternative Trajectories.**

*Place-cell reactivation during SWR.* Hippocampal replays are not restricted to recently experienced navigation trajectories (8, 21). We therefore examined the effect of reward on replays related to the rewarding trajectories in the blocks other than the current one. Prior to examining reward effect on replay, we examined reward effect on reactivation of other-block place cells during SWRs in the reward zone. We included only track-zone place cells (SI Appendix, Materials and Methods) in this analysis and examined whether reward affects reactivation of those place cells in the fields of which were located off the current correct trajectory.

For example, if the correct trajectory in the current block is RR, the place fields located in the first or second L (determined on the basis of neural and behavioral data in LL, LR, or RL block) were analyzed. We found that these place cells were activated at higher probabilities per SWR in correct than in error trials (signed-rank test; $\Delta p_{\text{reward}}$, $z = 2.343, P = 0.019, n = 250$ place cells, Fig. 5A). This result indicates that reward enhances activation of place cells during SWRs the fields of which are not only on, but also off, the recently taken trajectory.

**Replay rate and fidelity.** We then examined the effects of reward on forward and reverse replays of the trajectories other than the current one (i.e., place-cell sequences in the other blocks in a given session). In this analysis, we used place cells in each of the blocks other than the current one to determine candidate replay events. If a candidate event matched significantly to a correct path in a given block (a template of place cells determined using forced-choice and correct free-choice trials in a given block), it was considered as a replay of that specific trajectory. We also used place cells in each of the other blocks to calculate replay rates, absolute weighted cross correlations ($\langle r \rangle$, $p_{\text{act}}$, and $p_{\text{coact}}$). All other procedures were identical to those in the analysis of current-trajectory replays. Note that place cells of the other blocks overlap substantially with those of the current block. Consequently, replays for the current block overlap substantially with those for the other blocks (SI Appendix, Fig. S9). To alleviate this problem, we excluded those other-trajectory replays that were current-trajectory replays as well (i.e., common replays) from this analysis (forward replays: 446 excluded from 857; reverse replays: 372 excluded from 760).

The analysis of other-trajectory replays (with common replays excluded) yielded results largely similar to those obtained with the analysis of current-trajectory replays. The replay rate was significantly higher in correct than in error trials for reverse, but not for forward, replays (signed-rank tests; forward replays: $z = 1.410, P = 0.159, n = 24$ session sets; reverse replays: $z = 2.832, P = 0.005, n = 24$ session sets; Fig. 5B). The absolute weighted correlation was significantly higher in correct than error trials for forward, but not reverse, replays (rank-sum test; forward replays: $z = 3.325, P = 8.9 \times 10^{-4}, n = 377$ and 34, respectively; reverse replays: $z = 0.365, P = 0.715, n = 359$ and 29, respectively; Fig. 5C). Also, for both $p_{\text{act}}$ and $p_{\text{coact}}$, the normalized difference between correct and error trials ($\Delta p_{\text{reward}}$) was significantly different from zero in forward replays (signed-rank test, $P_{\text{act}} z = 2.237, P = 0.025, n = 130$ place cells; $P_{\text{coact}} z = 3.246, P = 0.001, n = 390$ place-cell pairs; Fig. 5D and E), but not in reverse replays ($P_{\text{act}} z = -0.172, P = 0.863, n = 49$ place cells; $P_{\text{coact}} z = -0.478, P = 0.633, n = 156$ place-cell pairs), and $\Delta p_{\text{Reward}}$ for $P_{\text{coact}}$ was significantly larger in forward than in reverse replays (rank-sum test; $P_{\text{coact}} z = 1.726, P = 0.084; P_{\text{coact}} z = 4.803, P = 1.6 \times 10^{-5}$; Fig. 5D and E). These results indicate similar effects of reward on replays of the current and other alternative rewarding trajectories.

**Effect of Navigation History on Replay.** Next, we asked whether and how place-cell sequences in the previous and upcoming blocks in a given session (referred to as recent and upcoming trajectories, respectively) are reactivated differently in forward and reverse replays. In this analysis, if a candidate event matched significantly to a correct path in an upcoming (or previous) block (a template of place cells determined using forced-choice and correct free-choice trials in an upcoming or previous block), it was considered as the replay of the upcoming (or recent) trajectory. All other
Fig. 5. Effect of reward on replays of alternative rewarding trajectories. Shown are reward effects on replays for rewarding trajectories other than the current one. (A) Effect of reward on reactivation of place cells located on alternative trajectories during SWRs. Those place cells with firing fields located inside the track zone, but not in the correct trajectory of the current block, were analyzed. The same format as in Fig. 4B. *P < 0.05 (signed-rank test). (B–E) Effect of reward on forward and reverse replays for alternative rewarding trajectories. Replays were determined using place-cell activity in each of the blocks other than the current one, and those overlapping with current-block replays were excluded from the analysis. (B) Mean ± SEM rates (Hz) of forward and reverse replays for alternative rewarding trajectories. **P < 0.005 (signed-rank test). (C) Absolute weighted correlations (r) of forward and reverse replays for alternative rewarding trajectories. The same format as in Fig. 4G. **P < 0.005 (rank-sum test). (D and E) Activation (Pact) and coactivation (Pcoact) of place cells. The probabilities of place cells in forward and reverse replays for alternative rewarding trajectories. The same format as in Fig. 4H and I. *P < 0.05, **P < 0.005, ***P < 0.0005 (signed-rank test for differences from 0; rank-sum test for comparisons between forward and reverse replays).

Discussion

In this study, we identified differential effects of reward and navigation history on CA1 forward and reverse replays during the awake state in a spatial sequence memory task. Reward increased the rate of reverse replays, while increasing the fidelity of forward replays. The enhancing effect of reward was observed not only for the current, but also for other alternative trajectories heading toward a rewarding location in both forward and reverse replays. In addition, upcoming rewarding trajectories showed more faithful reactivation than did already rewarded trajectories in forward replays.

Previous studies have shown positive effects of reward on SWRs and hippocampal replays. Reward enhances the rate of waking SWRs (wSWRs) and CA3 place-cell activity during wSWRs (18). Reward increases the rate of reverse replays during wSWRs (19, 35). Place-cell coactivity during wSWRs is stronger for CA1 neurons with place fields near reward locations than for those with place fields farther from reward locations (16). Also, trajectories reconstructed from replays of CA1 place cells are preferentially directed toward previously visited as well as unvisited (but observed) reward locations during wSWRs and sleep-associated SWRs (5, 20, 21, 38). In humans, hippocampal activity patterns for high-reward contexts are preferentially reactivated during postlearning rest, and this reactivation is predictive of memory retention (39). These studies consistently indicate facilitating effects of reward on hippocampal replays. Consistent with the results of these studies, we found that reward facilitates both forward and reverse replays, but in different ways. We found that reward increases the rate of reverse but not forward replays significantly, which is consistent with previous studies (19, 35). Even though reward enhanced forward replay rate in many sessions, reward also decreased it in some sessions, suggesting that different reward conditions may have different effects on hippocampal replay patterns.
to imagine how place-cell sequences activated in reverse can contribute to strengthening forward trajectory representations, dopamine may allow reverse replays to strengthen forward associations (40, 41). In addition, we found that not all trajectories are reactivated equally in forward replays. In forward replays, upcoming rewarding trajectories are more faithfully reactivated than already rewarded trajectories. This difference is rather small, and hence its functional consequences are likely to be subtle. Also, it is currently unclear why upcoming rewarding trajectories are reactivated preferentially to already rewarded ones. Nevertheless, this may indicate preferential reactivation of the more valuable trajectories in forward replays. In our task, the animals may increase and decrease reward values associated with upcoming and recent trajectories, respectively, because they are forced to take trajectories alternative to those already experienced in a given session to obtain a reward. The reactivation of trajectories other than the most recently experienced one has been proposed as one of the mechanisms necessary for maintaining a stable representation of an environment (i.e., a cognitive map) (21). Our results raise the possibility that hippocampal replays may contribute to building not simply a spatial map, but a map of potential navigation trajectories and their associated values (i.e., a "value map"). In the future, we hope to test this possibility using behavioral tasks that are designed to more directly test value dependence of hippocampal replays.

How might forward and reverse replays contribute to building a value map? Reverse replays may contribute to reinforcing representations of all rewarding trajectories indiscriminately, while forward replays may preferentially reinforce representations for high-value trajectories. Alternatively, as proposed previously (5), hippocampal reverse replay may contribute to assigning values to hippocampal place cells according to their distances from a rewarding location. We have proposed a model (42) in which value-dependent CA1 place-cell activity during navigation (43, 44) facilitates preferential replays of high-value trajectories during SWRs. Briefly, during navigation, value-dependent CA1 neuronal activity allows preferential strengthening of synapses between those CA3 and CA1 neurons with overlapping place fields near a rewarding location via activity-dependent synaptic plasticity. During subsequent SWR episodes that occur when inhibitory tone is low and neuromodulatory signals differ from those during navigation (9), CA3 generates diverse sequences. Then, rewarding sequences are preferentially replayed in CA1 because of the enhancement of CA3-CA1 synapses for place cells near rewarding locations. In this scenario, reverse replays increase the reward values of those CA1 place cells having firing fields near a rewarding location, thereby facilitating preferential forward replays of rewarding trajectories during SWRs. In other words, reverse replays are used in constructing a map of CA1 place-cell values, while forward replays are used to construct a map of potential trajectories and their associated values. Further studies are needed to test this admittedly speculative hypothesis and to determine how forward and reverse replays affect hippocampal and extrahippocampal neural circuits to guide adaptive future navigation.

Materials and Methods
Experimental procedures for animal behavior, single-unit recording, and data analysis are described in detail in SI Appendix, Materials and Methods. The experimental protocol was approved by the Ethics Review Committee for Animal Experimentation at the Korea Advanced Institute of Science and Technology.

Data Availability. The raw data is deposited in Figshare with identifier doi.org/10.6084/m9.figshare.10032866.v2 (45).

ACKNOWLEDGMENTS. We thank A. D. Redish for his helpful comments on the initial manuscript. This work was supported by the Research Center Program of the Institute for Basic Science Grant IBS-R002-G1 (to M.W.J.).


