Generative Predictive Codes by Multiplexed Hippocampal Neuronal Tuplets

Highlights

● Hippocampus generates an internal predictive model for rapid encoding of new spaces

● Prediction-error-like signals update the internal model during spatial encoding

● Hippocampal organization into tuplet motifs results in generative predictive codes

● Tuplet-based predictive codes increase the network capacity for distinct rapid encoding

Authors

Kefei Liu, Jeremie Sibille, George Dragoi

Correspondence

george.dragoi@yale.edu

In Brief

Liu et al. show that hippocampal neurons are functionally organized into short tuplet (mostly triplet) motifs whose combination and recombination into extended sequences and subsequent editing result in predictive and replay codes with large capacity for rapid encoding and recall of distinct novel experiences.
Generative Predictive Codes by Multiplexed Hippocampal Neuronal Tuplets

Kefei Liu, Jeremie Sibille, and George Dragoi

INTRODUCTION

Cortical neuronal ensembles form rapid internal representations about the external world through the interplay between internally generated spontaneous neuronal dynamics and externally driven activity about the specific sensory dynamics of the world (Dragoi and Tonegawa, 2013c; Kenet et al., 2003; Luczak et al., 2009). A stronger network reliance on the recruitment of internally generated neuronal activity would greatly enhance the speed and efficiency of encoding at the cost of limiting its capacity to the combinatorial size of the fixed repertoire of internal neuronal motifs (Dragoi and Tonegawa, 2013b). On the other hand, a reduced dependency on an internal network model of the world would increase the flexibility and capacity for encoding independent representations that would instead require multiple exposures and additional time to create stable novel representations (McClelland et al., 1995). The neuronal ensemble mechanisms that could enable rapid encoding without constraining the capacity for multiple distinct representations are not clearly understood.

A possible framework for optimizing speed and capacity of encoding is provided by the predictive coding theory, which proposes that the brain is a Bayesian interpreter that largely computes the difference between external stimuli and a prior, unconscious, internal model of the world (Bialek et al., 2001; Friston, 2010; Heeger, 2017; Pezzulo et al., 2017; Wolpert et al., 1995). An important component of predictive coding is the hierarchical organization of the cortex that enables prediction and prediction-error signals (Adams et al., 2013; Friston and Buzsáki, 2016; Heeger, 2017; Stachenfeld et al., 2017) to be processed in multiple feed-forward and feedback loops (Rao and Ballard, 1999). At the top of the cortical hierarchy lays the hippocampal formation (Fellman and Van Essen, 1991), an associative cortical area that integrates multimodal sensorimotor information and is crucial for expressing internally generated contextual representations (Dragoi and Tonegawa, 2013c; Pezzulo et al., 2017), such as cognitive mapping (O’Keefe and Nadel, 1978; Tolman, 1948), episodic memory (Scoville and Milner, 1957; Tulving, 1969), planning (Diba and Buzsáki, 2007; Miller et al., 2017; Pfeiffer and Foster, 2013), and imagining (Hassabis et al., 2007; Schacter et al., 2008).

Mammalian navigation uses internal models to rapidly form a cognitive map of a novel external space encoded within ensembles of hippocampal “place cell” sequences (Dragoi and Buzsáki, 2006; O’Keefe and Dostrovsky, 1971). The hippocampal network was shown to spontaneously generate a limited repertoire of extended temporal sequence preplay patterns that correlate with multiple future novel place cell sequences (Dragoi and Tonegawa, 2011, 2013b). The preplay phenomenon demonstrates that novel place cell sequences are not entirely created de novo during the experience and provides evidence that pre-existing extended temporal sequences of firing can be rapidly selected to form novel representations, consistent with the existence of an internal model. Due to their reliance on extended temporal sequences that correlate with extended (i.e., full) navigational experiences and trajectories, preplay and especially the post-experience replay phenomenon (Lee and Wilson, 2002) have painted the image of a rather rigid hippocampal network, with a reduced capacity for simultaneously representing multiple future (Dragoi and Tonegawa, 2013b) and, in particular, past
experiences (Ji and Wilson, 2007; Karlsson and Frank, 2009; Lee and Wilson, 2002) in a given sleep/rest session. This reduced capacity is in stark disagreement with the role of the hippocampus in navigation, planning, imagining, and memory formation and recall, where the capacity for multiple sequential or parallel representations appears much larger (Aime et al., 2014; Rich et al., 2014). This suggests that, aside from the preplay and replay of extended sequences and trajectories, the hippocampal network must embrace fundamentally different organizational and representational principles to enable a flexible, large capacity for coding.

Here, we reveal the existence of a generative predictive statistical model for the statistical regularities of the contextual space in the hippocampus and define the principles underlying its use in navigation and its update by inferred intrinsic–unlikely neuronal functional connectivity and by selective plasticity. We show that these generative predictive codes rely on functional connectivity within and between high-repeat, short motifs of sequential neurons called “tuplets.” We propose that network organization into neuronal tuplets segments the extended temporal sequences into shorter modules, analogous to segmentation of words into syllables and of genes into codons. This novel neuronal organizational principle could vastly expand the generative capacity of the network by rapidly multiplexing neuronal tuplets into multiple independent extended temporal sequences. This would enable the hippocampal network to achieve high-speed and high-capacity encoding at the same time, as required by the cognitive functions it supports.

RESULTS

Generative Predictive Codes for Future Place Cells Sequences

We recorded sequential neuronal ensemble activity from the CA1 area of the hippocampus while six experimentally naive rats performed de novo exploration of a linear track (Run1), repeated over several sessions and followed by exploration of 2–3 additional novel tracks (Run2...n), all within 1 day. Each Run session was preceded and followed by a session of sleep in a small box located in the same room (Figure 1A). The spontaneous activity of CA1 pyramidal neurons displayed a repertoire of extended firing sequences detectable within “frames” of increased multunit synchronous activity flanked by neuronal silence (STAR Methods) during the first slow-wave sleep session in the naive state (Figure 1A, Pre-Run sleep) and in the slow-wave sleep sessions following a Run session (Post-Run1...n sleep). In order to investigate the nature of temporal sequence motifs, we applied a probabilistic generative Markov chain model to the sequential activity during sleep. The order of the Markov model represents the number of preceding cells a given cell depended on. We found that probability of observing a certain cell within frames from all sleep sessions (and separately from Pre-Run and Post-Run1...n sleep) depended on the activity of 1.55 ± 0.19 (range of 1–3) cells before it, a rather small subgroup of the total cell population (Figure 1B; order was between 1 and 3, estimated by multiple methods; Figure S1A). This suggests that extended neuronal temporal sequences during sleep, while typically correlated on individual bases with extended place cell sequences during Runs as reflected by the preplay and replay phenomena, are too variable and different from each other to generate more extended dependency lengths. Therefore, we conservatively applied a first-order Markov chain to model the network activity during sleep. The parameters of the model were the unconditional probability (P1) and the conditional probability (P2) for each sleep session (e.g., P2AB = probability cell B will fire, given cell A fires; Figure 1C for Pre-Run sleep). This model fitted the data well in that it had robust estimation of the probability of entire slow-wave sleep sequences and of short 3-cell motifs (Figures 1D and S1B; explained variance [EV] = 0.62 ± 0.05, range 0.45–0.78 across rats 1–6 depending on sleep duration; EV = 0.41 ± 0.1 when sleep epochs were reduced to 1 hr; EV = 0.40 ± 0.05 on cross-validated ~2-hr-long sleep sessions).

The hippocampal network is thought to rapidly express place cell sequences during run through an interplay between (1) a selection of neuronal temporal sequences from an internally generated preexisting repertoire and (2) an externally driven activity representing specific local and distal multisensory cues. The extent to which the internal dynamics alone contribute and can predict the expression of place cell sequences during future runs as opposed to simply being correlated with them above chance is currently unknown. To investigate whether the spontaneous firing sequence motifs during sleep can internally generate predictive codes for the future run sequences, we applied this generative Markov model exclusively to the network activity during the naive Pre-Run sleep (entire slow-wave sleep or sharp-wave ripple epochs only) and estimated the probability of future place cell sequences during the first de novo Run session (Run1). For each Run1 session and forward direction of movement on the linear track, we constructed Run1 sequences by ordering the location of the primary field of all place cells active along animal’s trajectory (Figure 1E). This resulted in two independent sequences of place cells for each animal during Run1, one for each direction of movement, and a total of 12 Run1 sequences (2 directions of running by each of the 6 rats). For each Run1 sequence, we generated one million potential sequences with the same cell length by randomly sampling from all cells active in the associated Pre-Run sleep (Figures S1C and S2A). Using the Markov model of Pre-Run sleep (Figure 1C), we estimated the probability of occurrence for all 10^6 potential sequences, which formed a log-normal distribution, and determined the percentile of the corresponding Run1 sequence probability (Figure 1F). Remarkably, we found that, in all animals, the predictability of future Run1 sequences from the intrinsic functional connectivity in the preceding sleep was among the highest of all potential sequences (median percentile = 98.8; 10/12 sequences >90th percentile, p < 10^-8, binomial test; 8/12 > 95th, p < 10^-7; Figure 1G, left). Furthermore, using the same Markov model and activity in the naive-state Pre-Run sleep, we found that the distribution of percentiles of prediction probability for all Run sequences (i.e., Run1...n) was highly skewed toward high values (p < 10^-15, Kolmogorov-Smirnov test; Figure 1G, right), with 52.7% (39/74) Run sequence probabilities above the 90th percentile and 41.9% (31/74) above the 95th percentile (p < 10^-10, binomial test). Predictability of the first run sequence, Run1, was stronger compared with those of later, Run2...n...
sequences (p = 0.027, Wilcoxon rank sum test; Figure 1G, inset). This difference indicates that increased temporal proximity to Pre-Run sleep and/or degree of uncertainty and novelty can increase the contribution of the internal factors over that of external factors to future spatial encoding.

The strong predictability of future place cell sequences from the preceding sleep should come in large part from predicting the identity of future place cells and in part from predicting their exact future sequence during run. To specifically evaluate the predictability of the exact place cell order from the preceding sleep, we eliminated all the contribution coming from the variability associated with cell identity by applying the Markov model to the Pre-Run sleep activity of only future place cells. We estimated the probability of Run sequences as their percentile within the distribution of 10^6 sequences obtained by shuffling the order of only future place cells in the corresponding Run sequence. We found that the probability of predicting the exact cell order from the internal dynamics of only the future place cells during the Pre-Run sleep was higher than chance (median percentile = 61.6, chance is at 50th percentile;)
with 200-Hz sharp-wave ripple oscillations (Figure S2C). To was restricted to only Pre-Run sleep frames that co-occurred and order during Runs was maintained when prediction analysis Furthermore, the high predictability of future place cell identity due to the inclusion of only future place cells in this analysis. Select group of order-shuffled real Run-place cell sequences, sleep scored significantly high even when compared with the dictability of the exact order of place cell activation from Pre-Run (Adams et al., 2013; Friston and Buzsáki, 2016; Heeger, 2017; Stachenfeld et al., 2017). To determine the most probable novel functional connections created in response to spatial exploration (i.e., unlikely to be predicted from the preceding sleep), we “edited” the Run1 sequences to make them more consistent with the preceding sleep (Figures 2A and S4A). We compared the edited Run1 sequence having the highest percentile of predictability from the Pre-Run sleep (called sRun1) with the original sequence (Run1) to infer the potentially novel (i.e., Intrinsic-unlikely), potentially lost (i.e., Intrinsic-likely) and the remaining (i.e., Unedited) functional connections between adjacent place cells (Figure 2A). We found that editing locations in Run1 sequences occurred progressively and preferentially toward the ends of the linear tracks, consistent with an increased density of external cues and rewards compared with the middle of the tracks (p = 0.0003, chi square test; Figure 2B, dark line, and Figure S4B). This preference could not be simply explained by the distribution of place cell density (editing locations normalized by place cell density, end > middle, p < 0.02, chi-square test; Figure S4C) and of trial-by-trial variability in firing rate across the linear track (p = 0.56, Wilcoxon rank sum test; Figure S4D). The excitability of the edited place cells, presumably the most plastic cells in the network, was not different from the average single-cell excitability (i.e., conditional probability, P1) during Pre-Run sleep (Figure 2C), although their place field peak firing rates during Run1 exceeded the median of peak rates of all cells (Figure S4E). In addition, the edited cells contributed to the whole repertoire of functional connection types (i.e., Intrinsic-unlikely, Intrinsic-likely, and Unedited). All Run sequences reached the 99.9th percentile of predictability from Pre-Run sleep within 1–3 rounds of successive editing (Figure 2D).

For all place cell pairs, we computed conditional probability and cell preference and calculated Run1 prediction percentiles using the neuronal activity recorded post-experience, during post-Run1 sleep. We found that, similar to Pre-Run sleep, the Post-Run1 sleep predictability for the combined cellular identity and order in Run1 was very strong for all animals and that predictability for the exact place cell order in the recent Run1 was significantly higher than chance (Figure 2E). Surprisingly, both combined and order-only predictability percentiles for Run1 appeared similar between Post-Run1 sleep and Pre-Run sleep...
Figure 2. Preferential Strengthening of Intrinsic-Unlikely Connections during Post-experience Sleep

(A) Diagram of place cell sequence editing during Run1 and of determination of different groups of functional connections.
(B) Distribution of editing locations during Run1. Black lines represent distributions normalized by place cell density.
(C) Normalized unconditional probability (P1) during Pre-Run sleep of cells edited during Run1 and all other cells (p = 0.55, Wilcoxon rank sum test; mean and SEM: 1.12 ± 0.19, 1.00 ± 0.02).
(D) Number of editing steps applied to Run1 to reach a 99.9 prediction percentile for sRun1 as a function of the prediction percentile of the original Run1 sequences.

(legend continued on next page)
Intrinsic-unlikely connections (Figure S4H). This indicates that a stronger reactivation observed in the Intrinsic-unlikely connections was not simply due to their systematic deviations from the whole population.

Overall, these findings indicate a synaptic principle by which the newly formed functional connections (i.e., Intrinsic-unlikely), which specifically compute the difference between the place cell sequences during Run1 experience and the internal statistical model, are preferentially strengthened by the following sleep session, while the potentially lost connections (i.e., Intrinsic-likely) become weaker (Figure 2L). The selective potentiation of connections encoding this difference efficiently updates the internal model, and the network “learns” the statistics of the novel environment. The update of the internal statistical model should be expressed in the spontaneous activity during the sleep following the experience and persist into the next experience. Consistent with this idea, the Intrinsic-unlikely connections of Run1, which were selectively strengthened after the experience, remained stable across the sleep sessions preceding and following a second run experience on the same track, Run2 (Figure S4K). Moreover, the number of editing steps during sleep required to reach the 99.9th percentile of prediction probability for Run2 sequences decreased significantly from Pre-run (i.e., before the first experience) sleep to Post-Run1 (i.e., after the first experience) sleep (Figure S4L). We hypothesize this type of network learning is achieved via Hebbian plasticity in the

(Figure 2E, insets; p > 0.3 for identity and order, p > 0.68 for order, Wilcoxon sign rank test; Pre-run to Post-Run1 percentiles for place cell order across Run sequences were highly correlated: r = 0.6, p = 0.02, Pearson correlation). Combining all Run sessions together, we revealed a stronger predictability in Post-run1...n sleep compared with Pre-run sleep (Figure 2F; p < 10^{-3}, p < 0.02, Wilcoxon sign rank test; Pearson correlation of Pre-run and Post-run1 percentile for place cell order: r = 0.3, p = 0.005).

Intrigued by the similarity in prediction percentile from Pre-run to Post-Run1 sleep, we compared predictability for place cell order in Run1 from Post-run1 with the one from Pre-run sleep separately for each of the three groups of functional pairs (i.e., Intrinsic-unlikely, Intrinsic-likely, and Unedited). We found that conditional probability and cell preference were stronger in Post-run1 sleep than Pre-run sleep, specifically in the Intrinsic-unlikely group (p = 6 × 10^{-5}, p = 0.007, signed rank tests), but not in the Intrinsic-likely (p = 0.9, p = 0.2) and Unedited groups (p = 0.5, p = 0.4; Figures 2G and S4F). The magnitude of increase from Pre-run to Post-run1 sleep was higher in the Intrinsic-unlikely group than in the two other groups for both conditional probability and cell preference (Figures 2H–2K). Compared to the Unedited connections, the strength of Intrinsic-unlikely connections was initially lower during Pre-run sleep but reached similar levels in Post-run1 sleep (Figure S4G). Interestingly, when cumulative sequential rounds of editing were applied to reach the 99.9th percentile for sRun1 sequences, the cell preference for the potentially lost Intrinsic-likely functional connections became significantly weaker in Post-run1 sleep compared with Pre-run sleep (Figure 2K).

To control for our results on editing and updating of the internal model, we first detected the subgroup of Unedited and All connections that had the same amplitude as the Intrinsic-unlikely connections in Pre-run sleep (i.e., same deviation from the rest of the sample). We found these connections had significantly smaller Post-run1 to Pre-run difference compared with Intrinsic-unlikely connections (Figure S4H). This indicates that larger deviation from sample in Pre-run is not sufficient to explain the increased reactivation observed in the Intrinsic-unlikely connections. Second, we selected Unedited connections that showed Post-run1 to Pre-run differences as high as the Intrinsic-unlikely connections (50 independent samplings; Figure S4I) and used them to edit the Run1 sequences to increase their prediction percentile from Pre-run. We found that editing of these connections decreased the prediction percentile of Run1 sequences from Pre-run sleep, in stark contrast to the original increase in percentile by editing the Intrinsic-unlikely connections (Figure S4J). This indicates that a stronger reactivation in Post-run1 versus Pre-run sleep is not sufficient to indicate that the respective connections have been edited during Run1. Altogether, these results further confirm that the preferentially increased reactivation of Intrinsic-unlikely connections was not simply due to their systematic deviations from the whole population.

Overall, these findings indicate a synaptic principle by which the newly formed functional connections (i.e., Intrinsic-unlikely), which specifically compute the difference between the place cell sequences during Run1 experience and the internal statistical model, are preferentially strengthened by the following sleep session, while the potentially lost connections (i.e., Intrinsic-likely) become weaker (Figure 2L). The selective potentiation of connections encoding this difference efficiently updates the internal model, and the network “learns” the statistics of the novel environment. The update of the internal statistical model should be expressed in the spontaneous activity during the sleep following the experience and persist into the next experience. Consistent with this idea, the Intrinsic-unlikely connections of Run1, which were selectively strengthened after the experience, remained stable across the sleep sessions preceding and following a second run experience on the same track, Run2 (Figure S4K). Moreover, the number of editing steps during sleep required to reach the 99.9th percentile of prediction probability for Run2 sequences decreased significantly from Pre-run (i.e., before the first experience) sleep to Post-run1 (i.e., after the first experience) sleep (Figure S4L). We hypothesize this type of network learning is achieved via Hebbian plasticity in the
High-Repeat Short Neuronal Tuplets during Sleep

The existence of generative predictive codes for contiguous sequential place cells and the neuronal dependence on 1–3 preceding cells within the hippocampal network during sleep reducing the size of the modules to single cells completely abolished it (Figures S5A and S5B). These findings suggest that the predictive codes and the internal model of contiguous sequential space likely depend on the organization of hippocampal neurons into short sequences of 2–4 neurons, consistent with the dependency length of 1–3 neurons we observed during sleep (Figure 1B).

In line with this hypothesis, we identified populations of neurons displaying repeating identical sequential patterns of 2–4 contiguous neurons during sleep in proportions significantly different from chance (p < 0.05; Figure 3A) and that further reducing the size of the modules to 4, 3, and 2 sequential neurons dramatically affected the generative predictive codes for future place cell sequence from the preceding sleep (p = 3 x 10⁻⁶; all runs: 35/74, p = 3.3 x 10⁻¹⁵; cell order: Run1, 4/12, p = 0.02; all runs, 17/74, p = 0.0013; binomial tests). See also Figures S5 and S6.
higher than the chance level (Figure 3B), and we called them neuronal tuplets. The tuplets had higher counts \( (p = 1.9 \times 10^{-8}, \text{Wilcoxon signed rank test}) \), repeats \( (p = 2.3 \times 10^{-6}) \), and sparseness \( (p = 2.3 \times 10^{-8}) \) during Pre-Run sleep compared with shuffled sleep activity (Figure 3C). Remarkably, the tuplet length (Figure 3, inset) was independent of the total number of neurons active during sleep (range, 30–120 neurons per recording; Table S1) and of the size of frame-based neuronal population (range, 6–26 neurons per frame) and remained at an average of \( \sim 3 \) neurons despite a 4-fold increase in the size of recorded neuronal population \( (p > 0.05; \text{Figure 3D}) \). These results suggest a functional organization of neurons into modules of short temporal sequences expressed at milliseconds precision in naive animals spontaneously during sleep. Given the average length of the high-repeat tuplet motifs is \( \sim 3 \) neurons, we asked whether a higher-order (order = 2, i.e., neuronal dependency on activity of 2 other neurons) or a variable-order Markov model (Begleiter et al., 2004) for the Pre-Run sleep could also reveal a generative predictive model for the future Runs. In agreement with this hypothesis, we found that both second-order and the variable-order Markov models resulted in very high and similar prediction percentiles (Figures 3E and S5C). This suggests that statistically inferred tuplet organization could functionally determine the organization of place cell sequences during the future Runs.

Run (Figure 4), but were preferentially recruited in higher numbers \( (p = 7.4 \times 10^{-8}, \text{Wilcoxon signed rank test}) \) and proportions \( (p = 2.3 \times 10^{-8}) \) to contribute to the representation of future Run experiences (Figure 5A). Tuplets composed of cells with higher firing rate in Pre-Run sleep had higher probability of being recruited during Run (Figure 5B), while the number of repeats during Pre-Run sleep was higher for tuplets recruited during future Run (Figure 5C) compared with tuplets that “turned off” during Run. Furthermore, the prediction percentile of future Run sequences increased significantly when an increased number of Pre-Run sleep tuplets were recruited into the Run sequences (Figure 5D; \( r = 0.4, p < 0.0005 \)). We found that neuronal tuplet activity was important for the high probability prediction of Run1 sequences (Figure 1H), which remained unchanged when the Markov model was applied only to the subgroup of Pre-Run frames containing tuplets but dropped to chance levels when only the frames containing no tuplets were used (Figures 5D, S5E, and S3F). Consistent with this result, the predictive codes for the truncated place cell sequences from sleep (Figure 3A) correlated with the proportion of affected tuplets \( (r = -0.97, p = 1.6 \times 10^{-9}) \) and dropped dramatically when truncation in small modules of 2–4 contiguous place cells affected more than 75% of the tuplets of the preceding sleep (data not shown).

Finally, the spatial extent of tuplets recruited as place cell
sequences on the novel tracks during Run correlated with their temporal extent during the preceding Pre-Run sleep (2 tuplet: $r = 0.1$, $p = 3 \times 10^{-6}$; 3 tuplet: $r = 0.1$, $p = 3 \times 10^{-7}$; 4 tuplet: $r = 0.2$, $p = 0.015$, Pearson's correlation; Figure 5E) as tuplets decompressed from Pre-Run sleep to Run (Figures 4B; notice the tuplet organization in Run sequence that was not always captured by the correlation value).

We tested whether the update of the internal model by the Intrinsic-unlikely signal during Run (Figure 2) was associated with plastic changes in tuplet dynamics. The tuplets in Pre-Run sleep containing cells whose rank in place cell sequences during Run1 was edited (Figure 2A, red cells) were found to represent this update by displaying increased repeat rates (i.e., amplification) in Post-Run1 sleep (Figure 5F). In contrast, tuplets containing only place cells of Unedited functional connections during Run1 (Figure 2A, blue cells) occurred at similar rates in Pre-Run and Post-Run1 sleep (Figure 5F). Despite this selective increase in repeat rate, the overall tuplet length did not change in response to Run1 experience (Figure 5G), while the edited and unedited cells displayed similar average firing rates (Figure 2C). These findings indicate that specific tuplets are selectively involved in the predictive codes and the Intrinsic-unlikely signal within the hippocampus by generating and selectively updating the internal model of the external space.

We replicated our findings of generative predictive codes and sequence editing and tuplet organization and recruitment into future Run in an independent dataset (Figure S6) where Pre-Run sleep and de novo Run1 sessions were recorded across two different experimental rooms (Grosmark and Buzsáki, 2016) (Collaborative Research in Computational Neuroscience).
result in over $10^{43}$ different spatial sequences composed of 10–30 non-repeating place cells using the parameters provided in STAR Methods. Bottom: internally extended seq.

DISCUSSION

We have demonstrated that the hippocampal network of adult naive animals spontaneously generates internal predictive models used in future spatial navigation. During the active encoding of novel representations, internally unpredictable novel connections, likely encoding the statistically unpredictable, novel information from the external world as a difference from the model, are preferentially consolidated during post-experience sleep, and the internal model is updated. The internal model is probabilistically predictive of the future network dynamics encoding a large repertoire of future possible experiences in the linear Euclidian space, regardless of the unpredictability of exactly which of these experiences the animal will encounter next. The predictive power of the internal model stems from two factors. The first one is the similarity in some of the organizational principles for time and space, by which both appear quantal, continuous, and sequential in nature and have been proposed to conceptually emerge a priori, before explicit experience (Kant, 1781; Langston et al., 2010; Wills et al., 2010). The second factor is the similarity in the sequential activation of individual, functionally contiguous hippocampal neurons across different modalities (i.e., from time to space) and across brain states and behaviors (i.e., from sleep [temporal sequences] into navigation [spatial sequences]). We propose the internal neuronal dynamics during sleep could be seen as a generative (internal) model for the internal dynamics during future runs, largely due to their correlative and predictive relationship, from sleep and time to run and space. We propose it is the similarity in the spatial (during run) and temporal (during sleep) features that enables this model of sequence motifs to work. At the same time, this internal model could serve the role of an index code binding together multimodal future episodic experiences that could later be reactivated and recalled via hippocampal-neocortical interactions (Teyler and DiScenna, 1986). These predictive codes and their updating in the hippocampal CA1 area, at the top of the cortical hierarchy (Fellman and Van Essen, 1991), could signal initiation of brain-wide (Logothetis et al., 2012) top-down neuronal loops implicated in navigation (O’Keefe and Nadel, 1978) and memory encoding and recall (Scoville and Milner, 1957; Tulving, 2002).

The building blocks of the hippocampal internal model for predictive codes emerge from the spontaneous functional organization of neurons into millisecond-timescale-sequence triplets, which are selected, decompressed, and flexibly multiplexed into longer sequences during future exploratory Run states to encode novel spatial information (Figure 6). The consolidated novel connections encoding the statistical difference from the external world are represented in the updated internal model.
We expected that the length of tuplets would increase with the total number of recorded neurons. However, despite a 4-fold increase in the total number of recorded neurons (from 30 to 120) and 4-hr-long sleep sessions, the average length of the significant tuplets remained unchanged. While a more drastic increase in the total number of recorded neurons and sleep duration might result in an augmentation in tuple length, we pose that the concept of segmentation of extended sequences into multiplexing tuplets will continue to apply. We propose the short, recurring sequential tuplet motif could be conceptually thought of as a basic unit of neuronal sequential coding that operates in modular, generative manner as a “neuro-codon.”

Our results suggest two novel high-efficiency sequence-coding schemes for the hippocampal operations. First, the existence of a generative internal network model for space enables the rapid encoding of the difference between externally driven activity and the internal model during exploratory states. Remarkably, only the edited functional sequential connections that underlie the Intrinsic-unlikely signal are being strengthened and consoli-dated post-experience, while the unedited sequential network connections are largely unaltered. The selective consolidation of the Intrinsic-unlikely signal represents a synaptic plasticity principle that allows a more rapid and efficient way to update the internal model and ensure network-level learning. This selective consolidation principle, demonstrated here for sequential functional connectivity, generalizes to other forms of neuronal coding, like neuronal cofiring (Wilson and McNaughton, 1994), cell-assembly coactivation (van de Ven et al., 2016), and ensemble rate-code (Roux et al., 2017); however, it appears at odds with an earlier view stating that the hippocampal sequential network activity indiscriminately undergoes an almost complete reorganization during and after a run on a linear track (Lee and Wilson, 2002; Silva et al., 2015). The diversity in functional connections (i.e., Intrinsic-likely versus Intrinsic-unlikely and Unedited) in our data does not emerge from a difference at the individual-cell level between the firing rates of edited and unedited cells (Figure 2C) or from an intrinsic difference in their plasticity level (i.e., rigid versus plastic cells) (Grosmark and Buzsáki, 2016). Rather, it seems to emerge from the flexible but specific interplay between internally generated and externally driven connectivity, as largely the same cells can participate in all of these types of functional connections in different experiences. Second, the segmentation of neuronal ensemble activity into high-repeat tuplet modules enables a rapid and efficient encoding and “re-mapping” of a spatial environment by a flexible multiplex combination and recombination, respectively, of neuronal tuplets into extended place cell sequences. This principle is relevant for defining a generative grammar for the neural code, which appears analogous in nature to the recombination of existing syllables into new words and different words into new sentences, as well as to the somatic recombination of gene segments (Tonge-
wa, 1983) to increase antibody diversity.

Our study applied first- and second-order Markov chain models to fit the activity of short sequential motifs during slow-wave sleep and used these generative models to predict future spatial sequences during navigation on linear tracks. A subset of past probabilistic approaches focused on prediction of place cell activity from the network dynamics during run using peer prediction (Harris et al., 2003), expected reward (Stachenfeld et al., 2017), and hidden Markov model methods (Chen et al., 2016; Maboudi et al., 2018). The latter approach required a proper estimation of a hidden latent state to obtain a clear interpretation of the prediction, a step not required by the Markov chain model. Other studies focused on characterization of correlated extended sequences of firing depicting future spatial trajectories to novel spaces expressed during the preceding sleep (Chen et al., 2016; Dragoi and Tonegawa, 2011; Grosmark and Buzsáki, 2016) and rest (Dragoi and Tonegawa, 2011) and to familiar spaces during the preceding awake rest epochs (Diba and Buzsáki, 2007; Olafsdóttir et al., 2015; Pfeiffer and Foster, 2013). However, none of the previous studies applied generative models of hippocampal neuronal sequence activity during sleep to estimate future spatial sequences and specific Intrinsic-unlikely signals during spatial navigation. Our findings in the hippocampus are consistent with and could be interpreted from the framework of predictive coding, a concept originally used to explain encoding of external stimuli in the sensory cortical systems. Under the predictive coding framework, the Intrinsic-unlikely signal during run serves to encode the difference between the external world and the generative predictive model while the selective replay of this difference serves to update the internal model as parts of a prediction-error mechanism.

Previous research on preplay and replay during sleep identified a reduced repertoire of extended temporal sequences correlated with future and, respectively, past extended (i.e., full) spatial experiences. This approach estimated a rather limited and rigid predictive and recall capacity of the hippocampal network for spatial navigation and memory (Azizi et al., 2013; Dragoi and Tonegawa, 2013b; Karlsson and Frank, 2009; Lee and Wilson, 2002; Silva et al., 2015), in apparent contrast with the increased demands for a neuronal network serving cognitive functions like lifelong episodic memory and imagining. In fundamental contrast (Figure 6), the principle of hippocampal organization into neuronal tuplets suggests that the flexible multiplexing of tuplets and their subsequent editing during Run could result into an essentially unlimited repertoire of novel, distinct extended sequences (estimated at over 10⁶⁰ different sequences of 10–30 neurons out of a 60-neurons population, larger when considering the whole hippocampus, Figure 6). This neuronal organization could vastly increase the hippocampal network capacity to generate variability and predictive codes via internally generated (i.e., tuplet combination and recombination) and externally driven (i.e., sequence editing) mechanisms. These mechanisms could vastly increase at the same time the speed and the efficiency of hippocampal network for encoding and storage of novel information (Figure 6). The network ability to internally generate novel temporal sequences by tuplet multiplexing and recombination could underlie the hippocampal general role in a variety of internally generated novel representations that include imagining (Hassabis et al., 2007), planning (Miller et al., 2017), navigational shortcut (Gupta et al., 2010; Tolman, 1948), false memories (Ramirez et al., 2013), problem solving (Wagner et al., 2004) and, when abnormal, hallucinations (Adams et al., 2013; Powers et al., 2017; Scoville and Milner, 1957).
STAR METHODS

Detailed methods are provided in the online version of this paper and include the following:

- **KEY RESOURCES TABLE**
- **CONTACT FOR REAGENT AND RESOURCE SHARING**
- **EXPERIMENTAL MODEL AND SUBJECT DETAILS**
- **METHOD DETAILS**
  - Surgery and experimental design
  - Electrophysiology data acquisition
  - Sleep sequences and Run sequence generation
  - Parameters estimation for Markov chain model
  - Probability estimation of sequences from Markov model
  - Preference of cell interaction
  - Cell pattern occurrence estimation
  - Shuffle sleep generation
  - Prediction percentile estimation and cell identity, cell order separation
  - Order of Markov model
  - Primary, secondary and tertiary place fields
  - Run sequence editing
  - Run sequence segmentation
  - Tuple extraction and analysis
  - Capacity and efficiency estimation
  - Cross-validation in an independent dataset
- **QUANTIFICATION AND STATISTICAL ANALYSIS**
- **DATA AND SOFTWARE AVAILABILITY**

SUPPLEMENTAL INFORMATION

Supplemental Information includes six figures and one table and can be found with this article online at https://doi.org/10.1016/j.neuron.2018.07.047.

ACKNOWLEDGMENTS

We thank D. Lee, C. Sun, and G. Pezzulo for discussions and U. Farooq for pre-processing the CRNS data and discussions. This work was supported by a Whitall Foundation grant, a Brain and Behavior Foundation NARSAD Young Investigator Fellowship, and a Charles H. Hood Foundation Award to G.D. The reported data are archived on file servers at the Yale Medical School.

AUTHOR CONTRIBUTIONS

G.D. conceived and designed the study. J.S. and G.D. collected the data. K.L. and G.D. analyzed the data. G.D. and K.L. wrote the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES


McClelland, J.L., McNaughton, B.L., and O'Reilly, R.C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. Psychol. Rev. 102, 419–457.


STAR METHODS

KEY RESOURCES TABLE

<table>
<thead>
<tr>
<th>REAGENT or RESOURCE</th>
<th>SOURCE</th>
<th>IDENTIFIER</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chemicals, Peptides, and Recombinant Proteins</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Isoflurane</td>
<td>Zoetis</td>
<td>ANADA # 200-070</td>
</tr>
<tr>
<td>Grip cement</td>
<td>Teets denture material</td>
<td>N/A</td>
</tr>
<tr>
<td>Carprofen</td>
<td>Norbrook</td>
<td>55529-131-11</td>
</tr>
<tr>
<td>Experimental Models: Organisms/Strains</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rat: Long-Evans</td>
<td>Charles River</td>
<td></td>
</tr>
<tr>
<td>Software and Algorithms</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Python</td>
<td>Python software foundation</td>
<td><a href="https://www.python.org/">https://www.python.org/</a></td>
</tr>
<tr>
<td>PathTools</td>
<td>Singer et al., 2014</td>
<td><a href="https://github.com/psinger/PathTools">https://github.com/psinger/PathTools</a></td>
</tr>
<tr>
<td>VMM Algorithms</td>
<td>Begleiter et al., 2004</td>
<td><a href="http://www.cs.technion.ac.il/~ronbeg/vmm/code_index.html">http://www.cs.technion.ac.il/~ronbeg/vmm/code_index.html</a> and custom code.</td>
</tr>
<tr>
<td>Xclust3</td>
<td>Wilsonlab</td>
<td><a href="https://github.com/wilsonlab/mwsoft64">https://github.com/wilsonlab/mwsoft64</a></td>
</tr>
<tr>
<td>Data acquisition software</td>
<td>Neuralynx</td>
<td>Cheetah</td>
</tr>
<tr>
<td>Other</td>
<td></td>
<td></td>
</tr>
<tr>
<td>128 Channel Digital Amplifier</td>
<td>Neuralynx</td>
<td>Digilynx</td>
</tr>
<tr>
<td>12.7 μm tungsten wires</td>
<td>Kanthal</td>
<td>PX000004</td>
</tr>
<tr>
<td>Silicon Probes</td>
<td>Neuronexus</td>
<td>Buz-64</td>
</tr>
</tbody>
</table>

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to the Lead Contact, George Dragoi (george.dragoi@yale.edu).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Six Long-Evans adult male rats weighing ~350 g were used for data collection. Animal handling and experimental procedures were approved by the IACUC at Yale University and were performed in agreement with the NIH guidelines for ethical treatment of animals.

METHOD DETAILS

Surgery and experimental design

Animals were implanted under isoflurane anesthesia bilaterally with 32 independently movable tetrodes (2 rats) or two independently movable 64-channel 8-shank silicon octrodes (Buz-64 Neuronexus probes, 2 rats), or unilaterally with 24 independently movable tetrodes (2 animals) for extracellular recordings using 128-channel Neuralynx amplifiers (Digilynx). Craniotomy was performed above area CA1 of the hippocampus (centered at 4 mm post-bregma, 1.5-3 mm lateral to midline). The reference electrode was implanted posterior to lambda over the cerebellum. During the following several weeks of recovery, the tetrodes and silicon probes were advanced daily while animals rested and slept in a high-wall opaque sleeping box (30 x 45 x 40 h cm).

The experimental apparatus consisted of 150 x 150 cm rectangular elevated linear track maze. All tracks were 6.25 cm wide and 75 cm above the floor. 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 during the Pre-Run sleep session in the sleep box for ~2-4h, after which the linear maze was brought into the room and installed. Subsequently, the animals were transferred onto the linear maze for the first time and allowed to explore a 150 cm-long linear track whose ends were blocked by 20 cm-high, 10 cm-wide barriers (Run1 session) followed by ~2h sleep session in sleep box (Post-Run1 sleep). Animals were familiarized with the linear track via 1-3 additional Run – Post-Run sleep sessions (Run2–n and Post-Run2–n sleep sessions). Finally, while the animals were on the linear track, the track-end barriers were lifted in succession allowing the animals to explore for the first time 2-3 additional 150 cm-long linear tracks completing the square maze to which Track 1
was part of (Run2...n) followed by ~2-4h sleep sessions in sleep box (Post-Run2...n). The details of all the Run and sleep sessions are provided in Table S1 and Table S1 caption. After completion of all experiments, the brains of all rats were perfused, fixed, sectioned, and stained using Cresyl violet for electrode tracks reconstruction.

**Electrophysiology data acquisition**

Electrophysiological data acquisition was performed using digital Neurolynx data acquisition system with Cheetah software. Raw signals were recorded at 30,000 Hz and digitally filtered between 1 and 6000 Hz. Spikes were obtained by high pass filtering raw signal above 600 Hz and triggered by passing a 50 μV threshold. The animal’s position was monitored by an overhead camera whose signal was recorded by Cheetah software. Animal’s position was recognized via two diodes attached to the headstage. A total of 309 neurons were recorded from the hippocampal area CA1 in the six rats (Table S1, Rats1-6).

Single cells were identified and isolated offline using the manual clustering method Xclust3 (Dragoi and Tonegawa, 2013b; Wilson and McNaughton, 1993). Pyramidal cells were distinguished from interneurons based on spike width, average rate, and autocorrelations as published before (Dragoi and Tonegawa, 2013b).

**Sleep sequences and Run sequence generation**

Spiking frames were detected during slow-wave sleep periods in the sleep box determined based on continuous long animal immobility (velocity below 1 cm/s for at least 5 minutes) and low theta/delta ratio (below 2, after Hilbert transform for respective frequencies, 6-12 Hz for theta and 1-4 Hz for delta, and smooth with a 5 s Gaussian) to exclude epochs of rapid-eye movement sleep. In addition, sharp-wave ripples were detected using a Hilbert transform as detailed earlier (Dragoi and Tonegawa, 2013b) and spiking frames containing a ripple event (160-220 Hz) were separated and used to test the robustness of the generative predictive model (Figure S2C). A spiking frame was defined as a transient increase in the multiunit firing activity of a population of at least four different pyramidal cells within a temporal window preceded and followed by at least 100 ms of silence that delimitated the beginning and end of the event (Dragoi and Tonegawa, 2013b). The duration of spiking frames was limited to 80-1200 ms. The spikes of all cells that were emitted during the sleep and resting were sorted by time and further used for the detection of the spiking frames. The spike time center of mass (COM) was calculated for each cell in each frame and was used to determine the cells’ order in each spiking frame.

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 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 2 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 (Dragoi and Buzsáki, 2006). Place cell Run sequences (two sequences/session/animal) were calculated by ranking cells based on their primary place field location along animals’ trajectory, except in Figure S3 when secondary and tertiary place fields were also used when they could be determined using same criteria as for primary fields (second and third highest place field peak firing).

To rule out the possibility that spike sorting errors contributed to our results, we performed the prediction analysis on reduced Run sequences formed by the selection of one place cell from each tetrode (Figure S2D). For each Run sequence, we randomly sampled one cell from each tetrode. The sampling was performed until the generated reduced Run sequence contained more than 4 cells or more than the number of tetrodes in the Run sequences minus two. This approach was repeated 50 times to generate 50 reduced Run sequences for each original Run sequence.

The lap by lap variability in place cell firing rate was calculated using the bootstrap method. 90% of the laps were taken each time to calculate the place field, which was repeated 20 times and generated place fields populations for each place cell; their standard deviation was calculated at each spatial location in 15 cm bins as firing rate variability.

**Parameters estimation for Markov chain model**

To estimate the parameters for the probability model, each sleep frame was represented by a sequence of cell IDs. The maximum likelihood estimation was used to estimate the parameters. Conditional probability in the transition matrix was given by:

\[
pr(x_i|x_{i-k}...x_{i-2}x_{i-1}) = \frac{n(x_i,...,x_{i-k}...x_{i-2}x_{i-1})}{n(x_{i-k}...x_{i-2}x_{i-1})}
\]

[1]

where \(x_i\) is the \(i^{th}\) cell in the sleep sequence, \(n\) is the count of sequence, and \(k\) is the order of the model, which was set to 1 in most cases in this study (except for second and variable order Markov models in Figures 3E and S5C) and conditional probability \(pr\) is represented as \(p^2\) (P2 in the main text and figures):

\[
p^2(x_i|x_{i-1}) = \frac{n(x_i,x_{i-1})}{n(x_{i-1})}
\]

[2]
For the direct comparison of P2 across all animals, the influence of cell number on probability was eliminated by normalizing P2 by the number of cells in each animal.

To alleviate the noise induced by low firing cells, 0 in the transition matrix was reset to the minimum non-zero value in the transition matrix and 1 was reset to maximum non-one value in the matrix. This operation was performed to avoid 0 in transition matrix caused by finite sample size of real dataset since during probability estimation of long sequences, one 0 could make the probability of whole sequence to be 0. Another option to deal with this problem would be to use a Bayesian method (calculating probability by finite sample size of real dataset since during probability estimation of long sequences, one 0 could make the probability of whole matrix and 1 was reset to maximum non-one value in the matrix. This operation was performed to avoid 0 in transition matrix caused much as possible.

The alphabet of the model is cell ID.

The unconditional probability (P1) of each cell was given by:

$$p_1(x_i) = \frac{n(x_i)}{N}$$  \[3\]

where N is the length of sequential activity during sleep. For the direct comparison of P1 across all animals, to eliminate the influence of cell number on probability, P1 was normalized by the number of cells in each animal.

We used 4 methods to estimate the optimal order (dependence length) of Markov chain model: likelihood ratio, Akaike information criteria (AIC), Bayesian model selection, and comparison with shuffle. The first three methods were implemented as previously (Singer et al., 2014). Briefly, for the likelihood ratio method, likelihood was computed as the probability of observing the data given specific parameters, after which log likelihood ratio was calculated by comparing log likelihood of one model with another alternative model. AIC method was implemented based on minimization of AIC among models of different order. Bayesian model selection method determined the optimal order by maximizing the probability of data in condition of a certain order model. To estimate the dependent length by comparison with shuffle, we computed the two-dimensional transition matrix of cell pairs interleaved by certain number of cells. Variable interval conditional probability of two cell was calculated as:

$$p_r(x_i|x_{i-m}) = \frac{n(x_{i-m}x_i)}{n(x_{i-m})}$$  \[4\]

where m is the interval between two cells, which was set from 1 to 10.

The dependence of a cell on the mth cell before it was quantified by the variance of variable interval conditional probability with interval m:

$$d = \frac{1}{N} \sum_{x_i} (p_r(x_i|x_{i-m}) - \mu)^2$$  \[5\]

where N is number of unique cells, and $\mu$ is the average $p_r$ of all $x_{i-m}$.

If there was no dependence on a previous cell, $p_r(x_i|x_{i-m})$ would be equal to $p_r(x_i)$, then $d = 0$. The larger the $d$, the stronger was the dependence. However, there existed noise in the real data. To determine the significance of a meaningful dependence, we constructed random sleep sequences by shuffling cells’ order within each spiking frames. We computed dependence $d_{shuffle}$ of the random sleep sequences, and pairwise compared it with dependence $d$ of real sleeps by Wilcoxon signed rank test. Dependence existence was considered at 5% significance level.

### Probability estimation of sequences from Markov model

The probability of any sequence whose alphabet was included in the alphabet of the model could be estimated. For a given sequence $x$, the probability was estimated by

$$p_r(x) = p_r(x_1)p_r(x_2|x_1)\ldots p_r(x_n|x_1\ldots x_{n-1}) \prod_{i=k+1}^{n} p_r(x_i|x_{i-k}\ldots x_{i-2k+1})$$  \[6\]

where $x$ is the sequence of interest, $x_i$ is the ith cell, $n$ is the length of sequence, and $k$ is the order of model. For a first-order model, $k = 1$,

$$p_r(x) = p_r(x_1) \prod_{i=2}^{n} p_r(x_i|x_{i-1})$$  \[7\]

### Preference of cell interaction

To quantify cell interaction excluding single cell property influence, we calculated the preference of connections by:

$$\text{preference}(x_i|x_{i-1}) = \log_{10} \frac{p_r(x_i|x_{i-1})}{p_r(x_i)}$$  \[8\]
Cell pattern occurrence estimation
To evaluate the Markov model for sleep sequences, the occurrence of a cell pattern with three cells combination was estimated. The three-cell pattern was chosen because its actual occurrence in sleep sequences could be reasonably observed from the data. The probability of each pattern was estimated by Equation 7, then scaled by the length of sleep sequences.

The explained variance was calculated as:

\[ ev = 1 - \frac{\text{var}(n_{\text{est}}) - \text{var}(n_{\text{obs}})}{\text{var}(n_{\text{obs}})} \]  

where \( \text{var}() \) represents variance calculation, \( n_{\text{est}} \) is the estimated occurrence of 3-cell patterns and \( n_{\text{obs}} \) is their observed occurrence in the sleep sequences data.

Shuffle sleep generation
To generate shuffled sleep sequences, the participation probability of each cell (P1) was calculated for the sleep. For each sleep frame, frame length was set to be the same as in the original sleep; cells were selected by weighted random sample without replacement based on P1 of original sleep. Only P1 was maintained in shuffled sleep sequences.

Prediction percentile estimation and cell identity, cell order separation
To evaluate Run sequences, for each Run sequence, one million random sequences with the same length as the Run sequence were generated by uniform random sampling from all cells in the model. The probability of each of the random sequences and the Run sequence were estimated by the Markov model by Equation 7 after which the percentile of Run sequence probability among the distribution of probabilities of corresponding random sequences was calculated.

To investigate Run cell order prediction from sleep, the random sequences for a Run sequence were generated by shuffling the order of place cells in the Run sequence. All random sequences consisted of the same (place) cell population as the Run sequence, with different order (no information from specific cell ID was used for this estimation as same cells participated in all computations).

Order of Markov model
To test the influence of model’s order on prediction, an order = 2 model was tested, in which we estimated \( p(r_k \mid x_{i-3}, x_{i-2}, x_{i-1}) \). A variable order Markov model was also used (Figure S3C) in which the order was variable rather than constant. We used the algorithms, implementation and default parameters of Begleiter et al. (Begleiter et al., 2004).

Primary, secondary and tertiary place fields
To test the effect of place sub-fields in prediction of Run from sleep, for each Run sequence we constructed a population of sequences of same cell count using combinations (Combo, Figure S3) of primary, secondary and tertiary place fields (one field/place cell). For the cells with multiple place fields, each place field was used to construct different sequences and obtain all possible combinations of place fields. We used first order Markov model of preceding sleep to estimate the prediction percentile of these newly generated Run sequences and computed a percentile distribution for each Run sequence. The peak of the probability percentile distribution was compared with the probability percentile of the original Run sequence (Figure S3).

Run sequence editing
To investigate the intrinsic-unlikely functional connections in Run sequences, we edited the Run sequences to maximize their estimated probability from sleep. We used ‘move’ for cell editing. For each Run sequence, one cell was moved from its original Run location in the sequence and inserted into another location in the sequence to generate a new sequence. For example, ABcDEFG → ABDEFcG, where c was relocated from rank 3 to 6. This step was performed for all cells and all possible locations in the Run sequence to generate \( n(n-1) \) new sequences, where \( n \) is the count of cells in the Run sequence. These were all the possible sequences that could be generated by this operation. The probability of the new sequences was estimated based on the preceding sleep model. The edited sequence with the maximal probability from sleep (sRun) was considered to contain a mix of Intrinsic-likely functional connections (i.e., present in sRun, but not in the original Run) and Unedited connections (present in both sRun and Run). The functional connections present in Run, but not sRun were termed Intrinsic-unlikely.

To extract subgroups of functional connections with strength similar to that of the group of Intrinsic-unlikely connections from the Unedited and from all connections, a conditional probability distribution of Intrinsic-unlikely connections in Pre-Run sleep was built and was used as weight function to perform weighted random sampling without replacement from the Unedited and from all connections. The average change in conditional probability across two sleep sessions was calculated for the selected connections. This approach was repeated 1000 times and generated the distribution of changes in conditional probability to be compared with that of Intrinsic-unlikely connections.

To extract the Unedited-high connections, which were subsequently used for the editing procedure, we searched for subgroups of Unedited connections that exhibited changes in strength between two sleep sessions similar to those exhibited by the Intrinsic-unlikely connections. We calculated the distribution of changes in strength for the Intrinsic-unlikely connections and used it as a weight function to perform weighted random sampling from the Unedited connections, which resulted in the Unedited-high group.
our sequence editing method, we searched for three functional connections from each Run sequence within which at least two connections should contain one common cell, at the end of one connection and at the beginning of the other. The selected connections were used as targets for editing the Run sequences (as shown in Figure 2A for the Intrinsic-unlikely group). These sampling and editing steps were repeated 50 times. The prediction percentile of the edited sequence from Pre-Run sleep was calculated and compared with that of the original Run sequences.

### Run sequence segmentation

For each Run sequence whose prediction percentile was higher than 80, the sequence was truncated into several equal chunks containing 'n' adjacent cells. The order of chunks was shuffled to generate chunk-shuffled Run sequences. The smaller the n, the more different Run sequences were generated. A population of random sequences was also generated by shuffling the individual cells' order in the original Run sequences. The probabilities of all chunk-shuffled sequences and those of random sequences were estimated and the probability percentiles of the chunk-shuffled sequences of the distribution of randomly shuffled sequences were determined. The percentiles of chunk-shuffled sequences were compared with those of the original Run sequences (Figure 3A).

### Tuplet extraction and analysis

Recurring cell patterns during sleep activity were extracted from sleep spiking frames. Each pattern was defined as the cell sequence which appears in more than 2 spiking frames. We extracted all patterns by a ‘length growing’ method. First, all possible 2-cell patterns in sleep activity were extracted and their number of repeats (number of sleep spiking frames in which they appeared) were recorded. The 2-cell sequences which repeated more than 2 times were saved as 2-cell patterns. One additional cell was added to the 2-cell patterns to construct 3-cell sequences. The number of repeats of these 3-cell sequence patterns during sleep were calculated and the 3-cell sequences that repeated more than two times were saved. One additional cell was added to the 3-cell patterns and calibrated to search for 4-cell patterns. This procedure for pattern detection and extraction was repeated for increasingly longer sequences until no sequences repeating more than 2 times could be found.

In parallel, five hundred shuffled sleep data were generated for each sleep session, from which cell patterns and their repeats were extracted by the same length growing method. Tuplets were determined by comparison with shuffled sleep activity and were defined as the cell pattern whose repeat was higher than 95% of its repeat in shuffled sleeps. The duration of tuplets in sleep was determined to be from the spiking COM time of the first cell to that of last cell of the tuplet. The tuplet count was quantified as the number of tuplets of each length (2-tuplet: tuplets consisted of 2 cells, 3-tuplet: tuplets with 3 cells, etc.). The repeat was determined for each tuplet and the normalized repeat was calculated by dividing the repeat number by the number of spiking frames to normalize the impact of the length of sleep session. The sparseness of tuplets was calculated by the method of (Rolls et al., 1997) and adjusted by subtraction from 1 to make the parameter increase when sparseness increases:

$$\text{sparseness} = 1 - \frac{\langle V \rangle^2}{\langle V^2 \rangle}$$  \[10\]

where $\langle V \rangle$ represents the average of the repeats of all tuplets in a sleep session.

The occurrence of tuplets extracted from sleep was quantified as the number of tuplets per Run sequence and proportion in Run sequence. The number of tuplets per Run sequence was calculated by counting the number of tuplets (same cells and same order as in sleep tuplets) in each Run sequence and averaging over all future Run sequences. Proportion of tuplets in Run sequence was determined by dividing the number of tuplets participating in Run sequences by the total number of tuplets of a certain length. Tuplet’s span on track in Run was calculated as the distance from the on-track location of the first cell in that tuplet to the on-track location of the last cell of that tuplet. Tuplet duration in sleep was determined as the time interval between the activation of the first and the last cell of that tuplet.

To investigate the firing rate property of neurons in tuplets participating in Run sequences, neurons in at least one tuplet participating in at least one Run sequence were extracted and compared with neurons whose tuplets did not participate in any Run sequences.

During the comparison of sleeps before or after run session, to eliminate the impact of the length difference of sleeps, Pre-Run sleep and Post-Run1 sleep were limited to same number of spiking frames, the minimum of the two sessions. Tuplets were extracted from the adjusted sleep sessions and used for further comparisons.

To investigate the impact of tuplets on prediction performance, the sleep spiking frames were separated into two groups depending on whether they contained tuplets. To obtain a larger and comparable number of frames for each group, 2-cell tuplets were used to apply this separation. The two groups of frames were individually used to fit to Markov model and estimate future Run1 sequences.

### Capacity and efficiency estimation

To estimate the capacity and efficiency of a neuronal population (Figure 6), we calculated the number of Run sequences that the neuron population is capable to generate as:

$$N_{seq} = \frac{(N-n)!}{(N-L)!} \frac{m!}{(m-n)!}$$  \[11\]
where $N$ is the total number of neurons, $m$ is the number of tuplets in sleep, $l$ is the tuplet length, $L$ is length of Run sequence and $n$ is the number of tuplets in run sequences. When $l = 1$, Run sequences were generated by single cells, when $l = L$, Run sequences were generated by full, extended sequences. The capacity was defined as:

$$C = \frac{\log(N_{seq})}{\log(N_{seq,l=1})}$$  \[12\]

where $N_{seq,l=1}$ represents the $N_{seq}$ when $l = 1$.

The efficiency was defined as the normalized number of new connections required to generate Run sequences:

$$E = \frac{L - 1 - nl + n}{L - 1}$$  \[13\]

We used $N = 60$, $L = 10-30$, and $l = 1-9$ or $l = L$.

The performance contribution of each additional neuron was calculated by:

$$Z = \frac{\partial EC}{\partial l}$$  \[14\]

**Cross-validation in an independent dataset**

The main results in this manuscript were successfully cross-validated using the same data analysis tools in a comparable independent dataset collected from 4 adult Long-Evans male rat (Table S1, GBRat1-4) by Grosmark and Buzsáki (Grosmark and Buzsáki, 2016), available on the CRCNS - Collaborative Research in Computational Neuroscience data sharing website at http://crcns.org/NWB. The results of our analysis of these data are presented in Figures 3D and S6. The main difference in the experimental design between the two datasets is that Grosmark and Buzsaki recorded all sleep data in one experimental room and all their run data (Run1 only) in a different, novel room.

**QUANTIFICATION AND STATISTICAL ANALYSIS**

Analyses were performed using customized code written in MATLAB (R2015b; R2017b, MathWorks) and Python (2.7). The comparison between multiple populations was tested by Kruskal-Wallis tests. The comparison between two populations of different sample size was tested by Wilcoxon ranksum test, while two populations of same sample sizes were tested pairwise by Wilcoxon signed rank test. The distribution shapes were tested by Kolmogorov-Smirnov test and the significance of counts in data was tested by the Binomial test (Diba and Buzsáki, 2007). In figures, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Data are represented as mean ± standard error of the mean (SEM). P values lower than 10$^{-10}$ were cut off at 10$^{-10}$.

**DATA AND SOFTWARE AVAILABILITY**

Available data and codes are at https://medicine.yale.edu/lab/dragoi/.