Mega-scale movie-fields in the mouse visuo-hippocampal network

Chinmay S. Purandare[1-3]†✉ and Mayank R. Mehta[2-4]✉

[1] Department of Bioengineering, UCLA, Los Angeles, CA, USA.

[2] W.M. Keck Center for Neurophysics, Department of Physics and Astronomy, UCLA, Los Angeles, CA, USA.

[3] Department of Neurology, UCLA, Los Angeles, CA, USA.

[4] Department of Electrical and Computer Engineering, UCLA, Los Angeles, CA, USA.

† Present Address – Department of Physiology, UCSF, San Francisco, CA, USA

✉e-mail: chinmay.purandare@gmail.com, MayankMehta@ucla.edu

Date of first submission (at *Science*): June 6 2022

Abstract: Natural behavior often involves a continuous series of related images, often while the subject is immobile. How is this information processed across the cortico-hippocampal circuit? The hippocampus is crucial for episodic memory[1–3], but most rodent single unit studies require spatial exploration[4–6] or active engagement[7]. Hence, we investigated neural responses to a silent, iso-luminant, black and white movie in head-fixed mice without any task or locomotion demands, or rewards. The activity of most neurons (97%, 6554/6785) in the thalamo-cortical visual areas was significantly modulated by the 30s long movie clip. Surprisingly, a third (33%, 3379/10263) of hippocampal –dentate gyrus, CA1 and subiculum– neurons showed movie-selectivity, with elevated firing in specific movie sub-segments, termed movie-fields. On average, a cell had more than 5 movie-fields in visual areas, but only 2 in hippocampal areas. The movie-field durations in all brain regions spanned an unprecedented 1000-fold range: from 0.02s to 20s, termed mega-scale coding. Yet, the total duration of all the movie-fields of a cell was comparable across neurons and brain regions, partly due to broader movie-fields in hippocampal areas, indicating greater sequence coding. Consistently presentation of the movie images in a scrambled sequence virtually abolished hippocampal but not visual-cortical selectivity. The enhancement of sequential movie tuning compared to the scrambled sequence was eight-fold greater in hippocampal than visual areas, further supporting visual sequence encoding. Thus, a movie was encoded in all mouse-brain areas investigated. Similar results are likely to hold in primates and humans. Hence, movies could provide a unified way to probe neural mechanisms of non-spatial information processing and memory across brain regions and species.

Introduction

In addition to the position and orientation of simple visual cues, like Gabor patches and drifting gratings[8], primary visual cortical responses are also direction selective[9], and show predictive coding[10], suggesting that the temporal sequence of visual cues influences neural firing. Accordingly, these and higher visual cortical neurons also encode a sequence of visual images, i.e., a movie[11–13]. The hippocampus is farthest downstream from the retina in the visual circuit.

The rodent hippocampal place cells encode spatial or temporal sequences[2,14–21] and episode-like responses[22–24]. However, these responses typically require active locomotion[25], and they are thought to be non-sensory responses[26]. Primate and human hippocampal responses are selective to specific sets of visual cues, e.g., the object-place association[27], their short-term[1] and long-term[28] memories, cognitive boundaries between episodic movies[29], and event integration for narrative association[30]. However, despite strong evidence for hippocampal role in episodic memory, the hippocampal encoding of a continuous sequence of images, i.e., a visual episode, is unknown.

Results

We used a publicly available dataset (Allen Brain Observatory – Neuropixels Visual Coding, © 2019 Allen Institute). Mice were monocularly shown a 30s clip of a continuous segment from the movie *Touch of Evil* (Welles, 1958)[31] (Extended Data Fig. 1). Mice were head-fixed but were free to run on a circular disk. A total of 17048 broad spiking, active, putatively excitatory neurons were analyzed, recorded using 4-6 Neuropixel probes in 24 sessions from 24 mice (See *Methods*).

The majority of neurons in the visual areas (Lateral geniculate nucleus LGN, primary visual cortex V1, higher visual areas: antero-medial and medio-lateral AM-PM) were modulated by the movie, consistent with previous reports (Extended Data Fig. 2)[11–13,32]. Surprisingly, neurons from all parts of the hippocampus (dentate gyrus DG, CA3, CA1, subiculum SUB) were also clearly modulated (Fig. 1), with reliable, elevated spiking across many trials in small movie segments. To quantify selectivity in an identical, firing rate- and threshold- independent fashion across brain regions, we computed the z-scored sparsity[33,34] of neural selectivity (See *Methods*). Cells with z-scored sparsity >2 were considered significantly (p<0.03) modulated. The areas V1 (97.3%) and AM-PM (97.1%) had the largest percentage of movie tuned cells, much higher than reported earlier[11] (~40%), perhaps because we analyzed extracellular spikes, while the previous study used calcium imaging. The majority of neurons in LGN (89.2%) too showed significant modulation by the movie. Thus, the vast majority of thalamo-cortical neurons were significantly modulated by the movie.

Movie selectivity was prevalent in the hippocampal regions too, despite head fixation, dissociation between self-movements and visual cues as well as the absence of rewards, task, or memory demands (Fig. 1a-d). Subiculum, the output region of the hippocampus, farthest removed from the retina, had the largest fraction (44.6% Fig. 1d) of movie-tuned neurons, followed by the upstream CA1 (33.6%, Fig. 1c) and dentate gyrus (33.1%, Fig. 1a). However, CA3 movie selectivity was nearly half as much (17.3%, Fig. 1b). This is unlike place cells, where CA3 and CA1 selectivity are comparable[35,36] and subiculum selectivity is weaker[37].

To confirm these findings, we did several controls. Running alters neural activity in visual areas[38,39] and hippocampus[40,41]. Hence, we used the data from only the stationary epochs (see *Methods*) and only from sessions with at least 300 seconds of stationary data (17 sessions, 24906 cells). Movie tuning was unchanged in this data (Extended Data Fig. 3). This is unlike place cells where spatial selectivity is greatly reduced during immobility[5,6]. Neurons recorded simultaneously from the same brain region also showed different selectivity patterns (Extended Data Fig. 4). Thus, nonspecific effects such as running cannot explain the brain wide movie selectivity.

Hippocampal neurons have one or two place fields in typical mazes which take a few seconds to traverse[42]. In larger arenas that take tens of seconds to traverse, the number of peaks per cell and the peak duration increases[43–46]. Peak detection for movie tuning is nontrivial because neurons have nonzero background firing rate, and the elevated rates cover a wide range (Fig. 1). We developed a novel algorithm to address this (see *Methods*). On average, V1 neurons had the largest number of movie-fields (Fig. 2a, mean±s.e.m.=10.4±0.1), followed by LGN (8.6±0.3) and AM-PM (6.3±0.07). Hippocampal areas had significantly fewer movie-fields per cell: dentate gyrus (2.1±0.1), CA3 (2.8±0.3), CA1(2.0±0.02) and subiculum (2.1±0.05). Thus, the number of movie-fields per cell was smaller than the number of place-fields per cell in comparably long spatial tracks[43–48], but a handful of hippocampal cells had more than 5 movie-fields (Extended Data Fig. 5).

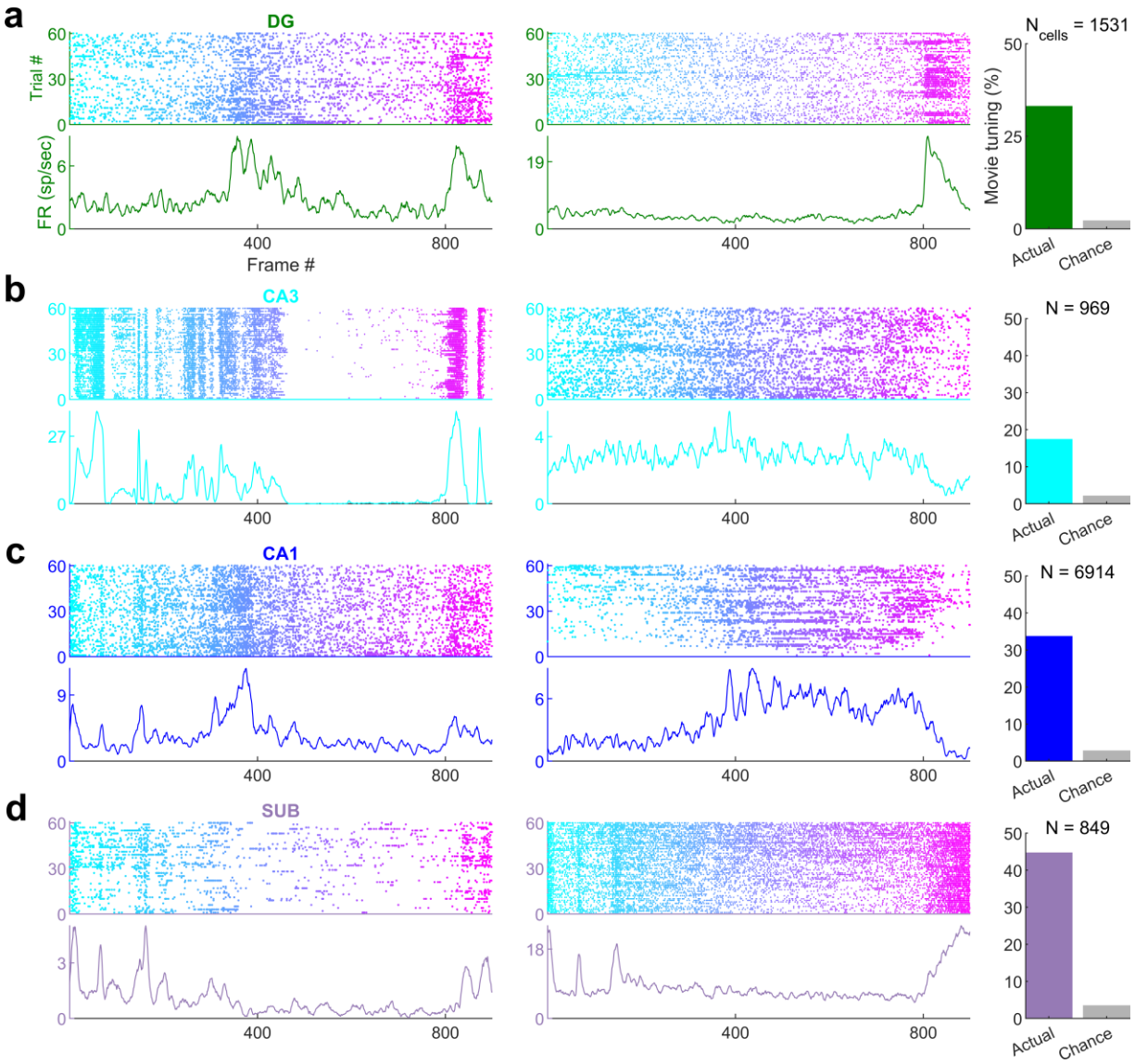


Fig. 1 | Movie frame selectivity in hippocampal neurons (a) Raster plots of two different dentate gyrus (DG) neurons as a function of the movie frame (top). The corresponding mean firing rate response over 60 trials is also shown (bottom). These two cells had significantly increased firing activity in specific parts of the movie. 33.1% of dentate neurons were significantly modulated by the movie (right, green bar),

far greater than chance (gray bar). Total active, broad spiking neurons for each brain region indicated at top (N_{tuned} /N_{cells}=506/1531). **(b)** Same as (a), for CA3 (168/969, 17.3%), **(c)** CA1 (2326/6914, 33.6%) and **(d)** subiculum (379/849, 44.6%) neurons.

Typical receptive field size increases as one moves away from the retina in the visual hierarchy[31]. A similar effect was seen for movie-field durations. On average, hippocampal movie-fields were longer than visual regions (Fig. 2b). But there were many exceptions –movie-fields of LGN (median±s.e.m., here and subsequently, unless stated otherwise, 308.5±33.9 ms) were twice as long as in V1 (156.6±9.2ms). Movie-fields of subiculum (3169.9±169.8 ms) were significantly longer than CA1 (2786.1±77.5 ms) and nearly three-fold longer than the upstream CA3 (979.1±241.1 ms). However, the dentate movie-fields (2113.2±172.4 ms) were two-fold longer than the downstream CA3. This is similar to the patterns reported for CA3, CA1 and dentate gyrus place cells[48]. But others have claimed that CA3 place fields are slightly bigger than CA1[49], whereas movie-fields showed the opposite pattern.

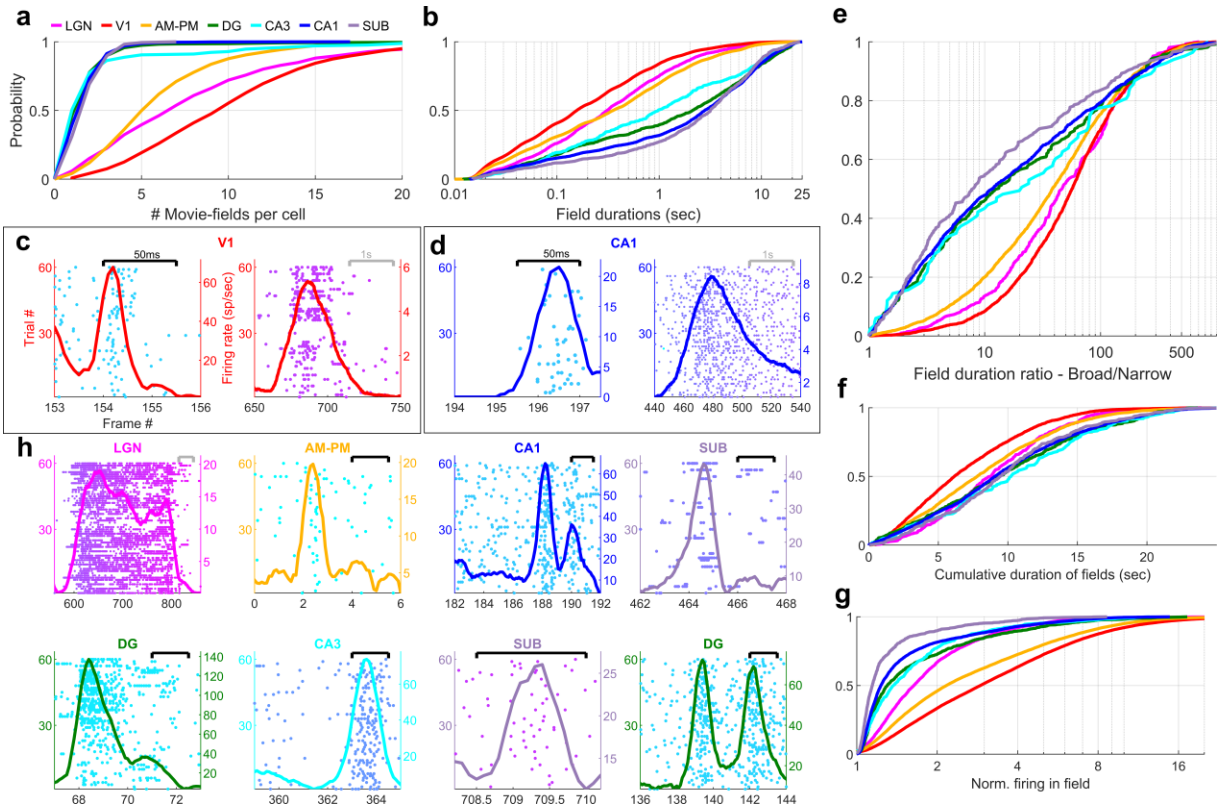


Fig. 2 | Multi-peaked, mega-scale movie-fields across all brain areas (a) Distribution of the number of movie-fields per tuned cell (See *Methods)* in different brain regions (shown by different colors, top line inset, arranged in their hierarchical order). Hippocampal regions (blue-green shades) were significantly different from each other (KS-test p<0.04), except DG-CA3. All visual regions were significantly different from each other (KS-test $p<7.0\times10^{-11}$). All visual-hippocampal region pair-wise comparisons were also significantly different (KS-test $p<1.8\times10^{-44}$). CA1 had the lowest number of movie-fields per cell (2.0±0.02, mean±s.e.m.) while V1 had the highest (10.4±0.1). **(b)** Distribution of the durations of movie-fields identified in (a), across all tuned neurons from a given brain region. These were significantly

different for all brain region pairs (KS-test $p<7.3\times10^{-3}$). The longest movie-fields were in subiculum (3169.9±169.8 ms), and the shortest in V1 (156.6±9.2ms). **(c)** Snippets of movie-fields from an example cell from V1, with 2 of the fields zoomed in, showing 60x difference in duration. Black bar at top indicates 50ms, and gray bar indicates 1s. Each frame corresponds to 33.3ms. Average response (solid trace, y-axis on the right) is superimposed on the trial wise spiking response (dots, y-axis on the left). Color of dots corresponds to frame numbers as in Fig. 1. **(d)** Same as (c), for a CA1 neuron with 54x difference in duration. **(e)** The ratio of longest to shortest field duration within a single cell, i.e., mega-scale index, was largest in V1 (56.7±2.2) and least in subiculum (8.0±9.7). All visual-visual and visual-hippocampal brain region pairs were significantly different on this metric (KS-test $p<0.02$). Among the hippocampal-hippocampal pairs, only CA3-SUB were significantly different ($p=0.03$). **(f)** For each cell, the total duration of all movie-fields, i.e., cumulative duration of significantly elevated activity, was comparable across brain regions. The largest cumulative duration (10.2±0.46s, CA3) was only 1.66x of the smallest (6.2±0.09 sec, V1). Visual-hippocampal and visual-visual brain region pairs' cumulative duration distributions were significantly different (KS-test $p<0.001$), but not hippocampal pairs ($p>0.07$). **(g)** Distribution of the firing within fields, normalized by that in the shuffle response. All fields from all tuned neurons in a brain region were used. Firing in movie-fields was significantly different across all brain region pairs (KS-test, $p<1.0\times10^{-7}$), except DG-CA3. Movie-field firing was largest in V1 (2.9±0.03) and smallest in subiculum (1.14±0.03). **(h)** Snippets of movie-fields from representative tuned cells, from LGN showing a long movie-field (233 frames, or 7.8s, panel 1), and from AM-PM and from hippocampus showing short fields (2 frames or 66.6ms wide or less).

The movie-field durations spanned a 500-1000-fold range in every brain region investigated (Fig. 2e). This mega-scale scale is unprecedentedly large, nearly 2 orders of magnitude greater than previous reports in place cells[43,45]. Even individual neurons showed 100-fold mega-scale responses (Fig. 2c & d) compared to less than 10-fold scale within single place cells[43,45]. The mega-scale tuning within a neuron was largest in V1 and smallest in subiculum (Fig. 2e). This is partly because the short duration movie-fields in hippocampal regions were typically neither as narrow and nor as prominent as in the visual areas (Extended Data Fig. 6).

Despite these differences in mega-scale tuning across different brain areas, the total duration of elevated activity, i.e., the sum of movie-field durations within a single cell, was remarkably conserved across neurons within and across brain regions (Fig. 2f). Unlike movie-field durations, which differed by more than ten-fold between hippocampal and visual regions, cumulative durations were quite comparable, ranging from 6.2s (V1) to 10.2s (CA3) (Fig. 2f, LGN=8.8±0.21sec, V1=6.2±0.09, AM-PM=7.8±0.09, DG=9.4±0.26, CA3=10.2±0.46, CA1=9.1±0.12, SUB=9.5±0.27). Thus, hippocampal movie-fields are longer and less multi-peaked than visual areas, such that the total duration of elevated activity was similar across all areas, spanning about a fourth of the movie, comparable to the fraction of large environments in which place cells are active[45,47,48]. To quantify the net activity in the movie-fields, we computed the total firing in the movie-fields (i.e., the area under the curve for the duration of the movie-fields), normalized by the expected discharge from the shuffled response. Net movie-field discharge was also comparable across brain areas, but maximal in V1 and least in subiculum (Fig. 2g).

Many movie-fields showed elevated activity spanning up to several seconds, suggesting rate-code like encoding (Fig. 2h). However, some cells showed movie-fields with elevated spiking

restricted to less than 50ms, similar to responses to briefly flashed stimuli in anesthetized cats[12,13,50]. This is suggestive of a temporal code, characterized by low spike timing jitter[51]. Such short-duration movie-fields were not only common in the thalamus, but also AM-PM, three synapses away from the retina. A small fraction of cells in the hippocampal areas, more than five synapses away from the retina, too showed such temporally coded fields (Fig. 2h).

To determine the stability and temporal-continuity of movie tuning across the neural ensembles we computed the population vector overlap between even and odd trials[52] (see *Methods*). Population response stability was significantly greater for tuned than for untuned neurons (Extended Data Fig. 7). The population vector overlap around the diagonal was broader in hippocampal regions than visual cortical and LGN, indicating longer temporal-continuity, reflective of their longer movie-fields. Further, the population vector overlap away from the diagonal was longer around frames 400-800 in all brain areas due to the longer movie-fields in that movie segment (see below).

Are all movie frames represented equally by all brain areas? The duration and density of movie-fields varied as a function of the movie frame and brain region (Extended Data Fig. 8). We hypothesized that this variation could correspond to the change in visual content from one frame to the next. Hence for comparison, we quantified the similarity between adjacent movie frames as the correlation coefficient between corresponding pixels and termed it as frame-to-frame (F2F) correlation. The majority of brain regions had substantially reduced density of movie-fields between the movie frames 400 to 800, but the movie-fields were longer in this region. This effect was greater in the visual areas than hippocampal. Using significantly tuned neurons, we computed the average neural activity in each brain region at each point in the movie (see *Methods*). Although movie-fields (Fig. 3a), or just the tallest movie-field per cell (Fig. 3b), covered the entire movie, the peak normalized, ensemble activity level of all brain regions showed significant overrepresentation, i.e., deviation from the norm, in certain parts of the movie (Fig. 3c, see *Methods*). This was most pronounced in V1 and higher visual areas AM-PM. The number of movie frames with elevated ensemble activity was higher in visual areas than hippocampal regions (Fig. 3d), and also this modulation (see *Methods*) was smaller in hippocampus and LGN, compared to visual cortical regions (Fig. 3e).

Using the significantly tuned neurons, we also computed the average neural activity in each brain region corresponding to each frame in the movie, without peak rate normalization (see *Methods*). The degree of continuity between the movie frames, quantified as above (F2F correlation), was inversely correlated with the ensemble rate modulation in all areas except DG, CA3 and CA1 (Fig. 3f and g). As expected for a continuous movie, this F2F correlation was close to unity for most frames, but highest in the latter part of the movie where the images changed more slowly. The population wide elevated firing rates, as well as the smallest movie-fields, occurred during the earlier parts (Extended Data Fig. 8). Thus, the movie-code was stronger in the segments with greatest change across movie frames. These results show differential ensemble-wide representation of the movie across brain regions.

If these responses were purely visual, a movie made of scrambled sequence of images would generate equally strong or even stronger selectivity due to the even larger change across movie frame, despite the absence of similarity between adjacent frames. To explore this possibility, we investigated neural selectivity when the same movie frames were presented in a fixed but scrambled sequence (scrambled movie). The within frame and the total visual content was

identical between the continuous and scrambled movies, and the same sequence of images was repeated many times in both experiments (see *Methods*). But there was no correlation between adjacent frames, i.e., visual continuity, in the latter (Fig.4a).

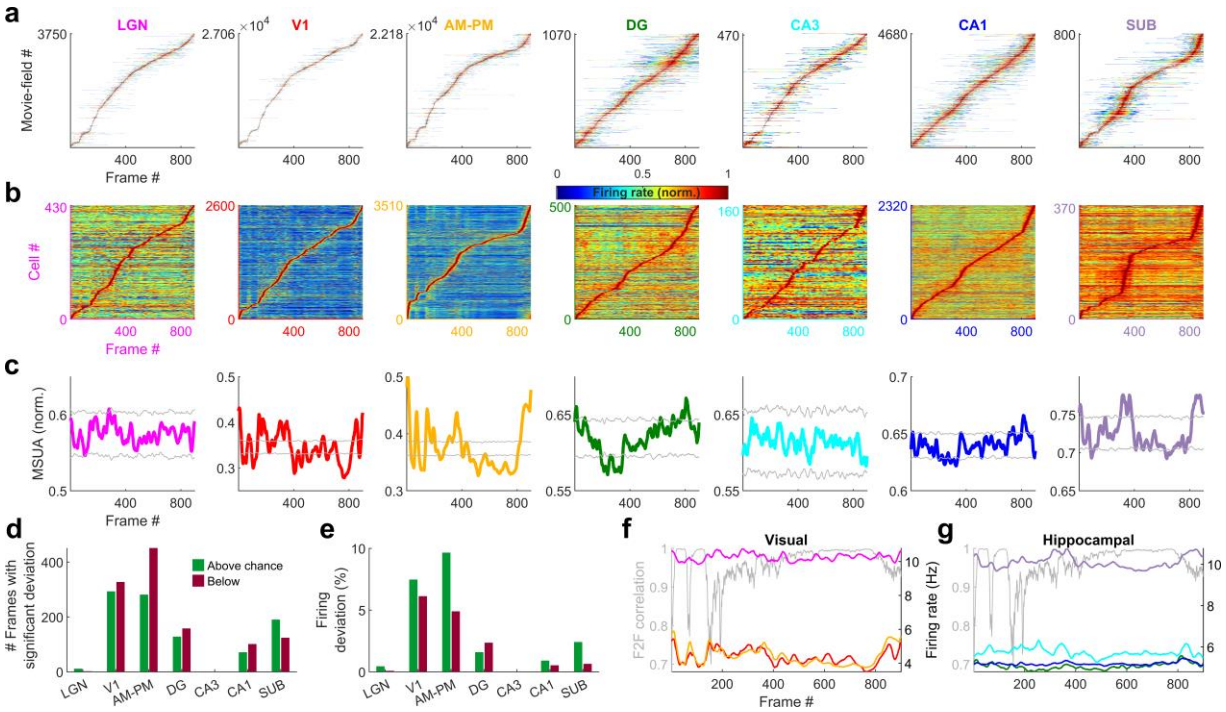


Fig. 3 | Population averaged movie-tuning varies across brain areas. (a) Stack plot of all the movie-fields detected from all tuned neurons of a brain region. Color indicates relative firing rate, normalized by the maximum firing rate in that movie-field. The movie-fields are sorted according to the frame with the maximal response. Note accumulation of fields in certain parts of the movie, especially in subiculum and AM-PM. **(b)** Similar to (a), but using only a single, tallest movie-field peak from each neuron showing a similar pattern, with pronounced overrepresentation of some portions of the movie in most brain areas. Each neuron's response is normalized by its maximum firing rate. **(c)** Multiple single unit activity (MSUA) in a given brain region, obtained as the average response across all tuned cells, by using maxima-normalized response for each cell from (b). Gray lines indicate mean±4*std response from the shuffle data corresponding to p=0.025 after Bonferroni correction for multiple comparisons (see *Methods*). AM-PM had the largest MSUA modulation (sparsity=0.01) and CA1 had the smallest (sparsity=1.8×10^{-4}). The MSUA modulation across several brain region pairs –AM&PM-DG, V1-CA3, DG-CA3, CA3-CA1 and CA1-SUB were not significantly correlated (Pearson correlation coefficient p>0.05). Some brain region pairs, DG-LGN, DG-V1, AM&PM-CA3, LGN-CA1, V1-CA1, DG-SUB and CA3-SUB, were significantly negatively correlated (r<-0.18, p<4.0×10^{-7}). All other brain region pairs were significantly positively correlated (r>0.07, p<0.03). **(d)** Number of frames for which the observed MSUA deviates from the z=±4 range from (c), termed significant deviation. V1 and AM-PM had the largest positive deviant frames (289), and CA3 had the least (zero). **(e)** Firing in deviant frames above (or below) chance level, as a percentage of the average response. Above chance level deviation was greater or equal to that below, for all brain regions, with the largest positive deviation in AM-PM (9.3%), largest negative deviation in V1 (6.0%), and least in CA3 (zero each). **(f)** Total firing rate response of visual regions across

tuned neurons. All regions had significant negative correlation (r<-0.39, p<3.4x10^{-34}) between the ensemble response and the frame-to-frame (F2F) correlation (gray line, y-axis on the left) across movie frames. **(g)** Similar to (f), for hippocampal regions. CA3 response were not significantly correlated with the frame-to-frame correlation, dentate gyrus (r=0.26, p=4.0x10^{-15}) and CA1 (r=0.21, p=1.5x10^{-10}) responses were positively correlated, and subiculum response was negatively correlated (r=-0.44, p=2.2x10^{-43}). Note the substantially higher mean firing rates of LGN in (f) and subiculum neurons in (g) (colored lines closer to the top) compared to other brain areas.

For all brain regions investigated, the continuous movie generated significantly greater modulation of neural activity than the scrambled sequence (Fig. 4b). Middle 20 trials of the continuous movie were chosen as the appropriate subset for comparison since they were chronologically closest to the scrambled movie presentation. This preference for sequential over scrambled movie was the greatest in hippocampal regions where the percentage of significantly tuned neurons (4.4%, near chance level of 2.3%) reduced more than 4-fold compared to the continuous movie (17.8%, after accounting for the lesser number of trials, see *Methods*). This was unlike visual areas where the scrambled (80.4%) and the continuous movie (92.4%) generated similar prevalence levels of selectivity (Fig. 4b). The few hippocampal cells which had significant selectivity to the scrambled sequence, did not have long-duration responses, but only very short, ~50ms long responses (Fig. 4d), reminiscent of, but even sharper than human hippocampal responses to flashed images[28]. To estimate the effect of continuous movie compared to the scrambled sequence on individual cells, we computed the normalized difference between the continuous and scrambled movie selectivity for cells which were selective in either condition (Fig. 4c, see *Methods*). This visual continuity index was more than eight-fold higher in hippocampal areas (87.8%) compared to the visual areas (10.6%).

The pattern of increasing visual continuity index as we moved up the visual hierarchy, largely paralleled the anatomic organization[53], with the greatest sensitivity to visual continuity in the hippocampal output regions, CA1 and subiculum, but there were notable exceptions. The primary visual cortical neurons showed the least reduction in selectivity due to the loss of temporally contiguous content, whereas LGN neurons, the primary source of input to the visual cortex and closer to the periphery, showed far greater sensitivity (Fig. 4c).

Many visual cortical neurons were significantly modulated by the scrambled sequence, but their number of movie-fields per cell was greater and their duration was shorter than during the continuous movie (Extended Data Fig. 9 and 10). This could occur due to the loss of frame-to-frame correlation in the scrambled sequence. The average activity of the neural population in V1 and AM-PM showed significant deviation even with the scrambled movie, comparable to the continuous movie, but this multi-unit ensemble response was uncorrelated with the frame-to-frame correlation in the scrambled sequence (Extended Data Fig. 11). A substantial fraction of visual cortical and LGN responses to the scrambled sequence could be rearranged to resemble continuous movie responses (Extended Data Fig. 12, see *Methods*). The latency needed to shift the responses was least in LGN and largest in AM-PM, as expected from the feed-forward anatomy and the model of visual information processing (Extended Data Fig. 12). This rearrangement did not recreate the continuous movie responses above chance levels in the hippocampal regions (example cells in Fig. 4e, also see Extended Data Fig. 12 for statistics and details).

Population vector decoding of the ensemble of a few hundred place cells is sufficient to decode the rat's position using place cells[54], and the position of a passively moving object[55]. Using similar methods, we decoded the movie frame number (see *Methods*). Continuous movie decoding was better than chance in all brain regions analyzed (Fig. 4f). Scrambled movie decoding was significantly weaker yet above chance level (as expected from shuffles, see *Methods*) in visual areas, but not in CA3 and dentate gyrus. But CA1 and subiculum neuronal ensembles could be used to decode scrambled movie frame number slightly above chance levels (Fig. 4g). Similarly, the population overlap between even and odd trials for scrambled sequence was strong for visual areas, and weaker in hippocampal regions, but significantly greater than untuned neurons in hippocampal regions (Extended Data Fig. 13).

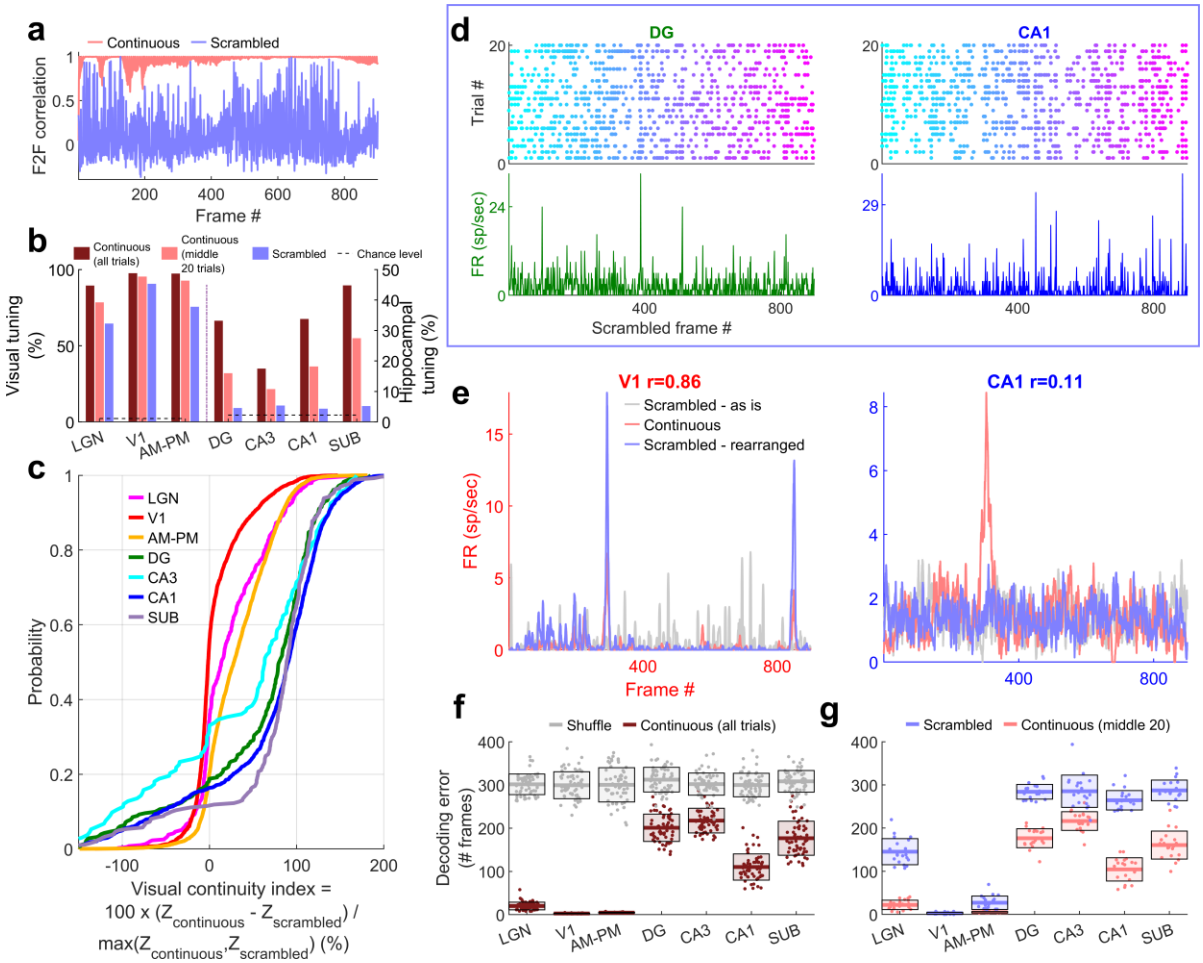


Fig. 4 | Larger reduction of selectivity in hippocampal than visual regions due to scrambled presentation. (a) Similarity between the visual content of one frame with the subsequent one, quantified as the *Pearson* correlation coefficient between pixel-pixel across adjacent frames for the continuous movie (pink) and the scrambled sequence (lavender), termed F2F correlation. Similar to Fig. 3g. For the scrambled movie, the frame number here corresponds to the chronological frame sequence, as presented. **(b)** Fraction of broad spiking neurons significantly modulated by either the continuous movie or the scrambled sequence using z-scored sparsity measures (similar to Fig. 1, see *Methods*). For all brain regions, continuous movie generated greater selectivity than scrambled sequence (KS-test $p<7.4\times10^{-4}$). **(c)** Percentage change in the magnitude of tuning between the continuous and scrambled movies for cells significantly modulated by either continuous or scrambled movie, termed visual

continuity index. Largest drop in selectivity due to scrambled movie occurred in CA1 (90.3±2.0%), and least in V1 (-1.5±0.6%). Continuous to scrambled tuning change was significantly different between all brain region pairs (KS-test $p<0.03$) and significantly greater for hippocampal areas than visual (8.2-fold, $p<10^{-100}$). **(d)** Raster plots (top) and mean rate responses (color, bottom) showing increased spiking responses to only one or two scrambled movie frames, lasting about 50ms. Tuned responses to scrambled movie were found in all brain regions, but these were the least frequent in DG and CA1. **(e)** One representative cell each from V1 (left) and CA1 (right), where the frame rearrangement of scrambled responses resulted in a response with high correlation to the continuous movie response for V1, but not CA1. *Pearson* correlation coefficient values of continuous movie and rearranged scrambled responses are indicated on top. **(f)** Average decoding error for observed data (see *Methods*), over 60 trials for continuous movie (maroon), was significantly lower than shuffled data (gray) (KS-test $p<8.2\times10^{-22}$). Solid line – mean error across 60 trials, shaded box – s.e.m. **(g)** Similar to (f), decoding of scrambled trials was significantly worse than that for the continuous movie (KS-test $p<3.6\times10^{-8}$), except V1 ($p=0.13$), where the errors were not significantly different (2.6 vs. 2.7 frames). Scrambled responses, in their "as is", chronological order were used herein. LGN decoding error for scrambled presentation was 6.5x greater than that for continuous movie, whereas the difference in errors was least for V1 (1.04x). Scrambled movie decoding error for all visual areas and for CA1 and subiculum was significantly smaller than chance level (KS-test $p<2.6\times10^{-3}$), but not DG and CA3 ($p>0.13$). Only the middle 20 trials of the continuous movie were used for comparison with the scrambled movie since the scrambled movie was only presented 20 times. Middle trials of the continuous movie were chosen as the appropriate subset since they were chronologically closest to the scrambled movie presentation.

Discussion

To understand how neurons encode a continuously unfolding visual episode, we investigated the neural responses in the rodent brain to an isoluminant human movie. As expected, neural activity in all thalamo-cortical visual areas was significantly modulated, with elevated activity in response to specific visual content. Most (96.6%, 6554/6785) of thalamo-cortical neurons showed significant movie tuning, nearly double that reported for the classic stimuli such as Gabor patches in the same dataset[31]. Remarkably, a third of hippocampal neurons (32.9%, 3379/10263) were also movie-tuned, comparable to the fraction of neurons with significant spatial selectivity in mice[56] and bats[57]. While the hippocampus is implicated in episodic memory, rodent hippocampal responses are largely studied in the context of spatial maps or place cells, and more recently in other tasks which requires active locomotion or active engagement[7,58]. However, unlike place cells[5,6], movie-tuning remained intact during immobility in all brain areas studied. This dissociation of the effect of mobility on spatial and movie selectivity agrees with the recent reports of dissociated mechanisms of episodic encoding and spatial navigation in human amnesia[59].

Across all brain regions, neurons showed a mega-scale encoding by movie-fields varying in duration by up to 1000-fold, similar to, but far greater than recent reports of 10-fold multi-scale responses in the hippocampus[43–48,60]. Importantly, mega-scale responses were found within single neurons as well, which can enable encoding of both finer and broader details of an episode in the same neuron. This could provide greatly enhanced episodic memory and recall.

The response latency was highest in AM-PM, then V1 and least in LGN, thus following the visual hierarchy. Among the visual areas, continuous movie tuning was greater in V1, more than the upstream LGN. This is surprising but can explain the recent findings that silencing V1 reduces movie tuning in the thalamus[61]. Similarly, the preference for the continuous movie over

scrambled sequence was greater in LGN and AM-PM than V1, which does not follow the expected hierarchy of visual processing. This effect was even greater in the hippocampal areas than any visual regions. Temporal integration window[62–64] as well as intrinsic timescale of firing[31] increases along the anatomical hierarchy in the cortex, with the hippocampus being farthest removed from the retina[53]. This hierarchical organization could explain the longer movie-fields and several fold greater preference for the continuous movie over scrambled sequence in the hippocampus. But, unlike reports of image-association memory in the inferior temporal cortex for unrelated images[65], only a handful hippocampal neurons showed selective responses to the scrambled sequence. These results, along with the longer duration of hippocampal movie-fields could mediate visual-chunking or binding of a sequence of events.

Could the brain-wide mega-scale tuning be an artifact of poor unit isolation, e.g., due to an erroneous mixing of two neurons, one with very short and another with very long movie-fields? This is unlikely since the LGN and visual cortical neural selectivity to classic stimuli (Gabor patches, drifting gratings etc.) in the same dataset was similar to that reported in most studies[31] whereas poor unit isolation should reduce these selective responses. However, to directly test this possibility, we calculated the correlation between the unit isolation index (or fraction of refractory violations) and the mega-scale index of the cell, while factoring out the contribution of mean firing rate (Extended Data Fig. 14). This correlation was not significant (p>0.05) for any brain areas.

Do the movie fields arise from the same mechanism as place fields? Studies have shown that when rodents are passively moved along a linear track that they had explored[6], or when the images of the environment around a linear track was played back to them[5], some hippocampal neurons generated spatially selective activity. Since the movie clip involved change of spatial view, one could hypothesize that the movie fields are just place fields generated by passive viewing. This is unlikely for several reasons. Mega-scale movie fields were found in all brain regions investigated, including the lateral geniculate nucleus and the primary visual cortex. There is no evidence of place cells in these brain regions. Further, in prior passive viewing experiments, the rodents were shown the same narrow linear track, like a tunnel, that they had previously explored actively to get food rewards at specific places. In contrast, in current experiments, these mice had never actively explored the space shown in the movie, nor obtained any rewards. Active exploration of a maze, combined with spatially localized rewards engages multisensory mechanisms resulting in increased place cell activation[17,23,66] which are entirely missing in these experiments during passive viewing of a movie, presented monocularly, without any other multisensory stimuli and without any rewards. Compared to the exploration of a real-world maze, exploration of a visually identical virtual world causes 60% reduction in CA1 place cell activation[67]. In contrast, there was no evidence of significant CA1 neural shutdown during movie viewing, compared to the blank screen presentation (Extended Data Fig. 14).

A recent study showed that CA1 neurons encode the distance, angle, and movement direction of motion of a vertical bar of light[55], consistent with the position of hippocampus in the visual circuitry[53]. Do those findings predict the movie tuning herein? There are indeed some similarities between the two experimental protocols –purely passive optical motion without any self-motion or rewards. However, there are significant differences too; similar to place cells in the real and virtual worlds[34], all the cells tuned to the moving bar of light had single receptive fields with elevated responses lasting a few seconds; there were neither punctate responses nor even 10-fold variation in neural field durations, let alone the 1000-fold change reported here. Finally, those

results were reported only in area CA1, while the results presented here cover nearly all the major stations of the visual hierarchy.

Notably, hippocampal neurons did not encode Gabor patches or drifting gratings in the same dataset, indicating the importance of temporally continuous sequences of images for hippocampal activation[31]. This is consistent with the hypothesis that the hippocampus is involved in coding spatial sequences[16,24,68]. However, unlike place cells that degrade in immobile rats, hippocampal movie tuning was unchanged in the immobile mouse. Further, the scrambled sequence too was presented in the same sequence many times, yet movie tuning dropped to chance level in the hippocampal areas. Unlike visual areas, scrambled sequence response of hippocampal neurons could not be rearranged to obtain the continuous movie response. This shows the importance of continuous, episodic content instead of mere sequential recurrence of unrelated content for rodent hippocampal activation.

These results complement recent findings of spatial modulation of visual cortical neurons and coordinated activity of visual and hippocampal neurons during navigation[69,70]. Our findings open up the possibility of studying thalamic, cortical, and hippocampal brain regions in a simple, passive, and purely visual experimental paradigm and extend comparable convolutional neural networks[11] to have the hippocampus at the apex[53]. Further, our results here bridge the long-standing gap between the hippocampal rodent and human studies[29,71–73], where natural movies can be decoded from fMRI signals in immobile humans[74]. This brain-wide mega-scale encoding of a human movie episode and enhanced preference for visual continuity in the hippocampus compared to visual areas supports the hypothesis that the rodent hippocampus could play a role in non-spatial episodic memories, consistent with classic findings in humans[1] and agrees with a more generalized, representational framework[75,76] of episodic memory where it encodes temporal patterns. Similar responses are likely across different species, including primates. Thus movie-coding can provide a unified platform to investigate the neural mechanisms of episodic coding, learning and memory.

References

1. Scoviille, W. B. & Milner, B. LOSS OF RECENT MEMORY AFTER BILATERAL HIPPOCAMPAL LESIONS. *J. Neurol. Neurosurg. Psychiat* (1957) doi:10.1136/jnnp.20.1.11.
2. MacDonald, C. J., Lepage, K. Q., Eden, U. T. & Eichenbaum, H. Hippocampal “Time Cells” Bridge the Gap in Memory for Discontiguous Events. *Neuron* **71**, 737–749 (2011).
3. Vargha-Khadem, F. *et al.* Differential effects of early hippocampal pathology on episodic and semantic memory. *Science (1979)* **277**, 376–380 (1997).
4. O’Keefe, J. & Nadel, L. *The hippocampus as a cognitive map*. (Clarendon Press, 1978).
5. Chen, G., King, J. A., Burgess, N. & O’Keefe, J. How vision and movement combine in the hippocampal place code. *Proceedings of the National Academy of Sciences* **110**, 378–383 (2013).
6. Foster, T. C., Castro, C. A. & McNaughton, B. L. Spatial Selectivity of Rat Hippocampal Neurons: Dependence on Preparedness for Movement. *Science (1979)* **244**, 1580–1582 (1989).
7. Aronov, D., Nevers, R. & Tank, D. W. Mapping of a non-spatial dimension by the hippocampal–entorhinal circuit. *Nature* **543**, 719–722 (2017).
8. Hubel, D. H. & Wiesel, T. N. Receptive fields of single neurones in the cat’s striate cortex. *J Physiol* **148**, 574–591 (1959).
9. De Valois, R. L., William Yund, E. & Hepler, N. The orientation and direction selectivity of cells in macaque visual cortex. *Vision Res* **22**, 531–544 (1982).
10. Xu, S., Jiang, W., Poo, M. M. & Dan, Y. Activity recall in a visual cortical ensemble. *Nat Neurosci* **15**, 449–455 (2012).

11. de Vries, S. E. J. *et al.* A large-scale standardized physiological survey reveals functional organization of the mouse visual cortex. *Nat Neurosci* **23**, 138–151 (2020).
12. Yen, S.-C., Baker, J. & Gray, C. M. Heterogeneity in the Responses of Adjacent Neurons to Natural Stimuli in Cat Striate Cortex. *J Neurophysiol* **97**, 1326–1341 (2007).
13. Herikstad, R., Baker, J., Lachaux, J. P., Gray, C. M. & Yen, S. C. Natural movies evoke spike trains with low spike time variability in cat primary visual cortex. *Journal of Neuroscience* **31**, 15844–15860 (2011).
14. Mehta, M. R., Quirk, M. C. & Wilson, M. A. Experience-Dependent Asymmetric Shape of Hippocampal Receptive Fields. *Neuron* **25**, 707–715 (2000).
15. Mehta and Wilson, M.A., M. R. From Hippocampus to V1: Effect of LTP on Spatio-Temporal Dynamics of Receptive Fields. *Neurocomputing* **32**, 905–911 (2000).
16. Mehta, M. R. *From synaptic plasticity to spatial maps and sequence learning*. *Hippocampus* vol. 25 756–762 (2015).
17. Mehta, M. R., Barnes, C. A. & Mcnaughton, B. L. Experience-dependent, asymmetric expansion of hippocampal place fields. *Proc Natl Acad Sci U S A* **94**, 8918 (1997).
18. Buzsáki, G. & Moser, E. I. Memory , navigation and theta rhythm in the hippocampal-entorhinal system. (2013) doi:10.1038/nn.3304.
19. Mau, W. *et al.* The Same Hippocampal CA1 Population Simultaneously Codes Temporal Information over Multiple Timescales. *Current Biology* **28**, 1499-1508.e4 (2018).
20. Kraus, B. J. *et al.* During Running in Place, Grid Cells Integrate Elapsed Time and Distance Run. *Neuron* **88**, 578–589 (2015).
21. Kraus, B. J., Robinson, R. J., White, J. A., Eichenbaum, H. & Hasselmo, M. E. Hippocampal 'time cells': time versus path integration. *Neuron* **78**, 1090–1101 (2013).
22. Pastalkova, E., Itskov, V., Amarasingham, A., Buzsaki, G. & Buzsáki, G. Internally generated cell assembly sequences in the rat hippocampus. *Science (1979)* **321**, 1322–1327 (2008).
23. Moore, J. J., Cushman, J. D., Acharya, L., Popeney, B. & Mehta, M. R. Linking hippocampal multiplexed tuning, Hebbian plasticity and navigation. *Nature 2021 599:7885* **599**, 442–448 (2021).
24. Buzsáki, G. & Tingley, D. *Space and Time: The Hippocampus as a Sequence Generator*. *Trends in Cognitive Sciences* vol. 22 853–869 (Elsevier Ltd, 2018).
25. McNaughton, B. L. *et al.* Deciphering the hippocampal polyglot: the hippocampus as a path integration system. *J Exp Biol* **199**, 173–85. (1996).
26. O'Keefe, J. & Dostrovsky, J. The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat. *Brain Res* **34**, 171–5. (1971).
27. Parkinson, J. K., Murray, E. A. & Mishkin, M. A selective mnemonic role for the hippocampus in monkeys: Memory for the location of objects. *Journal of Neuroscience* **8**, 4159–4167 (1988).
28. Quiroga, R. Q., Reddy, L., Kreiman, G., Koch, C. & Fried, I. Invariant visual representation by single neurons in the human brain. **435**, 1102–1107 (2005).
29. Zheng, J. *et al.* Neurons detect cognitive boundaries to structure episodic memories in humans. *Nature Neuroscience 2022 25:3* **25**, 358–368 (2022).
30. Cohn-Sheehy, B. I. *et al.* The hippocampus constructs narrative memories across distant events. *Current Biology* **31**, 4935-4945.e7 (2021).
31. Siegle, J. H. *et al.* Survey of spiking in the mouse visual system reveals functional hierarchy. *Nature* **592**, 86–92 (2021).
32. Hoseini, M. S. *et al.* Dynamics and sources of response variability and its coordination in visual cortex. *Vis Neurosci* (2019) doi:10.1017/S0952523819000117.
33. Acharya, L., Aghajan, Z. M., Vuong, C., Moore, J. J. & Mehta, M. R. Causal Influence of Visual Cues on Hippocampal Directional Selectivity. *Cell* **164**, 197–207 (2016).
34. Aghajan, Z. M. *et al.* Impaired spatial selectivity and intact phase precession in two-dimensional virtual reality. *Nat Neurosci* **18**, 121–128 (2015).
35. Jung, M. W. & Mcnaughton, B. L. *Spatial Selectivity of Unit Activity in the Hippocampal Granular Layer*. *HIPPOCAMPUS* vol. 3 (1993).
36. Muller, R. A Quarter of a Century of Place Cells. *Neuron* **17**, 813–822 (1996).
37. Sharp, P. E. & Green, C. Spatial correlates of firing patterns of single cells in the subiculum of the freely moving rat. *Journal of Neuroscience* **14**, 2339–2356 (1994).
38. Niell, C. M. & Stryker, M. P. Modulation of Visual Responses by Behavioral State in Mouse Visual Cortex. *Neuron* **65**, 472–479 (2010).

39. Erisken, S. *et al.* Effects of Locomotion Extend throughout the Mouse Early Visual System. *Current Biology* **24**, 2899–2907 (2014).
40. Góis, Z. H. T. D. & Tort, A. B. L. Characterizing Speed Cells in the Rat Hippocampus. *Cell Rep* **25**, 1872-1884.e4 (2018).
41. Wiener, S. I., Paul, C. A. & Eichenbaum, H. *Spatial and Behavioral Correlates of Hippocampal Neuronal Activity*. (1989).
42. O'keefe, J. & Burgess, N. *Geometric determinants of the place fields of hippocampal neurons*.
43. Eliav, T. *et al.* Multiscale representation of very large environments in the hippocampus of flying bats. *Science (1979)* **372**, (2021).
44. Kjelstrup, K. B. *et al.* Finite Scale of Spatial Representation in the Hippocampus. *Science (1979)* **321**, 140–143 (2008).
45. Harland, B., Contreras, M., Souder, M. & Fellous, J. M. Dorsal CA1 hippocampal place cells form a multi-scale representation of megaspace. *Current Biology* **31**, 2178-2190.e6 (2021).
46. Rich, P. D., Liaw, H. P. & Lee, A. K. Large environments reveal the statistical structure governing hippocampal representations. *Science (1979)* **345**, 814–817 (2014).
47. Fenton, A. A. *et al.* Behavioral/Systems/Cognitive Unmasking the CA1 Ensemble Place Code by Exposures to Small and Large Environments: More Place Cells and Multiple, Irregularly Arranged, and Expanded Place Fields in the Larger Space. (2008) doi:10.1523/JNEUROSCI.2862-08.2008.
48. Park, E., Dvorak, D. & Fenton, A. A. Ensemble Place Codes in Hippocampus: CA1, CA3, and Dentate Gyrus Place Cells Have Multiple Place Fields in Large Environments. *PLoS One* **6**, 22349 (2011).
49. Roth, E. D., Yu, X., Rao, G. & Knierim, J. J. Functional Differences in the Backward Shifts of CA1 and CA3 Place Fields in Novel and Familiar Environments. *PLoS One* **7**, e36035 (2012).
50. Xia, J., Marks, T. D., Goard, M. J. & Wessel, R. Stable representation of a naturalistic movie emerges from episodic activity with 2 gain variability. *Nature Communications 2021 12:1* **12**, 1–15 (2021).
51. Ikegaya, Y. *et al.* Synfire Chains and Cortical Songs: Temporal Modules of Cortical Activity. *Science (1979)* **304**, 559–564 (2004).
52. Resnik, E., McFarland, J. M., Sprengel, R., Sakmann, B. & Mehta, M. R. The Effects of GluA1 Deletion on the Hippocampal Population Code for Position. *J Neurosci* **32**, 8952–68 (2012).
53. Felleman, D. J. & Van, D. C. E. Distributed hierarchical processing in the primate cerebral cortex. *Cereb Cortex* **1**, 1–47 (1991).
54. Wilson, M. A. & McNaughton, B. L. Dynamics of the hippocampal ensemble code for space. *Science (1979)* **261**, 1055–8. (1993).
55. Purandare, C. S. *et al.* Moving bar of light evokes vectorial spatial selectivity in the immobile rat hippocampus. *Nature 2022* 1–7 (2022) doi:10.1038/s41586-022-04404-x.
56. Jun, H. *et al.* Disrupted Place Cell Remapping and Impaired Grid Cells in a Knockin Model of Alzheimer's Disease. *Neuron* **107**, 1095-1112.e6 (2020).
57. Yartsev, M. M., Witter, M. P. & Ulanovsky, N. Grid cells without theta oscillations in the entorhinal cortex of bats. *Nature* (2011) doi:10.1038/nature10583.
58. Danjo, T., Toyoizumi, T. & Fujisawa, S. Spatial representations of self and other in the hippocampus. *Science (1979)* **359**, 213–218 (2018).
59. McAvan, A. S., Wank, A. A., Rapcsak, S. Z., Grilli, M. D. & Ekstrom, A. D. Largely intact memory for spatial locations during navigation in an individual with dense amnesia. *Neuropsychologia* **170**, 108225 (2022).
60. Harland, B., Contreras, M. & Fellous, J.-M. A Role for the Longitudinal Axis of the Hippocampus in Multiscale Representations of Large and Complex Spatial Environments and Mnemonic Hierarchies. *The Hippocampus - Plasticity and Functions* (2018) doi:10.5772/INTECHOPEN.71165.
61. Spacek, M. A. *et al.* Robust effects of corticothalamic feedback and behavioral state on movie responses in mouse dLGN. *Elife* **11**, (2022).
62. Norman-Haignere, S. V. *et al.* Multiscale temporal integration organizes hierarchical computation in human auditory cortex. *Nat Hum Behav* **6**, (2022).
63. Gauthier, B., Eger, E., Hesselmann, G., Giraud, A. L. & Kleinschmidt, A. Temporal tuning properties along the human ventral visual stream. *Journal of Neuroscience* **32**, (2012).
64. Hasson, U., Yang, E., Vallines, I., Heeger, D. J. & Rubin, N. A Hierarchy of Temporal Receptive Windows in Human Cortex. *Journal of Neuroscience* **28**, 2539–2550 (2008).

65. Sakai, K. & Miyashita, Y. Neural organization for the long-term memory of. *Nature* **354**, 152–155 (1991).
66. Mehta, M. R. & McNaughton, B. L. Expansion and Shift of Hippocampal Place Fields: Evidence for Synaptic Potentiation during Behavior. *Computational Neuroscience* 741–745 (1997) doi:10.1007/978-1-4757-9800-5_115.
67. Ravassard, P. *et al.* Multisensory control of hippocampal spatiotemporal selectivity. *Science* **340**, 1342–6 (2013).
68. Foster, D. J. & Knierim, J. J. Sequence learning and the role of the hippocampus in rodent navigation. *Current Opinion in Neurobiology* vol. 22 294–300 Preprint at https://doi.org/10.1016/j.conb.2011.12.005 (2012).
69. Haggerty, D. C. & Ji, D. Activities of visual cortical and hippocampal neurons co-fluctuate in freely moving rats during spatial behavior. *Elife* **4**, e08902 (2015).
70. Saleem, A. B. *et al.* Coherent encoding of subjective spatial position in visual cortex and hippocampus. *Nature* **562**, 124–127 (2018).
71. Rutishauser, U., Mamelak, A. N. & Schuman, E. M. Single-trial learning of novel stimuli by individual neurons of the human hippocampus-amygdala complex. *Neuron* **49**, 805–813 (2006).
72. Silson, E. H., Zeidman, P., Knapen, T. & Baker, C. I. *Representation of contralateral visual space in the human hippocampus 1 2*. (2021). doi:10.1101/2020.07.30.228361.
73. King, J. R., Wyart, V. & King, J. R. The Human Brain Encodes a Chronicle of Visual Events at Each Instant of Time Through the Multiplexing of Traveling Waves. *Journal of Neuroscience* **41**, 7224–7233 (2021).
74. Nishimoto, S. *et al.* Reconstructing visual experiences from brain activity evoked by natural movies. **21**, 1641–1646 (2011).
75. Nadel, L. & Peterson, M. A. The hippocampus: part of an interactive posterior representational system spanning perceptual and memorial systems. *J Exp Psychol Gen* **142**, 1242–1254 (2013).
76. Nadel, L. & Hardt, O. Update on Memory Systems and Processes. *Neuropsychopharmacology 2011 36:1* **36**, 251–273 (2010).
77. Stringer, C. *et al.* Spontaneous behaviors drive multidimensional, brainwide activity. (2019) doi:10.1126/science.aav7893.
78. Wang, Q. *et al.* The Allen Mouse Brain Common Coordinate Framework: A 3D Reference Atlas ll The Allen Mouse Brain Common Coordinate Framework: A 3D Reference Atlas. *Cell* **181**, 936–953 (2020).
79. Hill, D. N., Mehta, S. B. & Kleinfeld, D. Quality Metrics to Accompany Spike Sorting of Extracellular Signals. *Journal of Neuroscience* **31**, 8699–8705 (2011).
80. Schmitzer-Torbert, N., Jackson, J., Henze, D., Harris, K. & Redish, A. D. Quantitative measures of cluster quality for use in extracellular recordings. *Neuroscience* **131**, 1–11 (2005).

Methods:

Experiments:

We used the Allen Brain Observatory – Neuropixels Visual Coding dataset (© 2019 Allen Institute, https://portal.brain-map.org/explore/circuits/visual-coding-neuropixels). This website and related publication[31] contain detailed experimental protocol, neural recording techniques, spike sorting etc. Briefly, and of relevance here, prior to implantation with Neuropixel probes, mice passively viewed the entire range of images including drifting gratings, Gabor patches and the movies of interest here. Videos of the body and eye movements were obtained at 30Hz and synced to the neural data and stimulus presentation using a photodiode. Movies were presented monocularly on an LCD monitor with a refresh rate of 60Hz, positioned 15cm away from the mouse's right eye and spanned 120°x95°. 30 trials of the continuous movie presentation were followed by 10 trials of the scrambled movie. Next was a presentation of drifting gratings, followed by a quiet period of 30 minutes where the screen was blank. Then the second block of drifting gratings, scrambled movie and continuous movie was presented.

Neural spiking data was sampled at 30 kHz with a 500Hz high pass filter. Spike sorting was automated using Kilosort2[77]. Output of Kilosort2 was post-processed to remove noise units, characterized by unphysiological waveforms. Neuropixel probes were registered to a common co-ordinate framework[78]. Each recorded unit was assigned to a recording channel corresponding to the maximum spike amplitude and then to the corresponding brain region. Broad spiking units identified as those with average spike waveform duration (peak to trough) between 0.45 to 1.5ms were analyzed herein.

Movie tuning quantification

The movie consisted of 900 frames: 30s total, 30Hz refresh rate, 33.3ms per frame. At the first level of analysis, spike data were split into 900 bins, each 33.3ms wide (the bin size was later varied systematically to detect mega-scale tuning, see below). The resulting tuning curves were smoothed with a Gaussian window of σ=66.6 ms or 2 frames. The degree of modulation and its significance was estimated by the sparsity *s* as below, and as previously described methods[55].

$$s = 1 - \frac{1}{N}\frac{(\sum_n r_n)^2}{(\sum_n {r_n}^2)}$$

where r_n is the firing rate in the nth frame or bin and N=900 is the total number of bins. Statistical significance of sparsity was computed using a bootstrapping procedure, which does not assume a normal distribution. Briefly, for each cell, the spike train as a function of the frame number from each trial were circularly shifted by different amounts and the sparsity of the randomized data computed. This procedure was repeated 100 times with different amounts of random shifts. The mean value and standard deviation of the sparsity of randomized data were used to compute the z-scored sparsity of observed data using the function *zscore* in MATLAB. The observed sparsity was considered statistically significant if the z-scored sparsity of the observed spike train was greater 2, which corresponds to $p<0.023$ in a one tailed t-test. Similar method was used to quantify significance of the scrambled movie tuning, as well as for the subset of data with only stationary epochs, or its equivalent subsample (see below). Middle 20 trials of the continuous movie were used in comparisons with the scrambled movie in Fig. 4, to ensure a fair comparison by using same number of trials, with similar time delays across measurements.

Stationary epoch identification

To eliminate the confounding effects of changes in behavioral state associated with running, we repeated our analysis in stationary epochs, defined as epochs when the running speed remained less than 2cm/sec for this period, as well as at least 5 seconds before and after this period. Analysis was further restricted to sessions with at least 5 total minutes of these epochs during the 60 trials of continuous movie presentation. To account for using lesser data of the stationary epochs, we compared the tuning using a random subsample of data, regardless of running or stopping and compared the two results for difference in selectivity.

Mega-scale movie-field detection in tuned neurons

For neurons with significant movie-sparsity, i.e., movie-tuned, the movie response was first recalculated at a higher resolution of 3.33ms (10 times the frame rate of 33.3ms). The *findpeaks* function in MATLAB was used to obtain peaks with *prominence* larger than 110% (1.1x) the

range of firing variation obtained by chance, as determined from a sample shuffled response. This calculation was repeated at different smoothing values (logarithmically spaced in 10 Gaussian smoothing schemes with σ ranging from 6.7ms to 3430ms), to ensure that long as well as short movie-fields were reliably detected and treated equally. For frames where overlapping peaks were found at different smoothing levels, we employed a comparative algorithm to only select the peak(s) with higher prominence score. This score was obtained as the ratio of the peak's prominence to the range of fluctuations in the correspondingly smoothed shuffle. This procedure was conducted iteratively, in increasing order of smoothing. If a broad peak overlapped with multiple narrow ones, the sum of scores of the narrow ones was compared with the broad one. To ensure that peaks at the beginning as well as the end of the movie frames were reliably detected, we circularly wrapped the movie response, for the observed as well as shuffle data.

Identifying frames with significant deviations in multiple single-unit activity (MSUA)

First, the average response across tuned neurons for each brain region was computed for each movie frame, after normalizing the response of each cell by the peak firing response. This average response was used as the observed "Multiple single unit activity (MSUA)" in Fig. 3. To compute chance level, individual neuron responses were circularly shifted with respect to the movie frames to break the frame to firing rate association but maintain overall firing rate modulation. 100 such shuffles were used, and for each shuffle, the shuffled MSUA response was computed by averaging across neurons. Across these 100 shuffles, mean and standard deviation was obtained for all frames, and used to compute the z-score of the observed MSUA. To obtain significance at p=0.025 level, Bonferroni correction was applied, and the appropriate z-score (4.04) level was chosen. The number of frames in the observed MSUA above (and below) this level were further quantified in Fig. 3. The firing deviation for these frames was computed as the ratio between the mean observed MSUA and the mean shuffled MSUA, reported as a percentage, for frames corresponding to z-score greater than +4 or less than -4. To obtain a total firing rate report, where each spike gets equal vote, we computed the total firing response by computing the total rate across all tuned neurons (and averaging by the number of neurons) in Fig. 3 and across all neurons in Extended Data Fig. 8.

Population Vector Overlap

To evaluate the properties of a population of cells, movie presentations were divided into alternate trials, yielding even and odd blocks[52]. Population vector overlap was computed between the movie responses calculated separately for these 2 blocks of trials. Population vector overlap between frames x of the even trials & frame y of the odd trials was defined as the Pearson correlation coefficient between the vectors $(R_{1,x}, R_{2,x}, \ldots R_{N,x})$ & $(R_{1,y}, R_{2,y}, \ldots R_{N,y})$, where $R_{n,x}$ is the mean firing rate response of the nth neuron to the xth movie frame. N is the total number of neurons used, for each brain region. This calculation was done for x and y ranging from 1 to 900, corresponding to the 900 movie frames. The same method was used for tuned and untuned neurons in continuous movie responses in Extended Data Fig. 7, and for scrambled sequence responses in Extended Data Fig. 13.

Decoding analysis

Methods similar to those previously described were used[54,55]. For tuned cells, the 60 trials of continuous movie were each decoded using all other trials. Mean firing rate responses in the 59 trials for 900 frames were used to compute a "look-up" matrix. Each neuron's response was normalized between 0 and 1. At each frame in the "observed" trial, the correlation coefficient was computed between the population vector response in this trial and the look-up matrix. The frame corresponding to the maximal correlation was denoted as the decoded frame. Decoding error was computed as the average of the absolute difference between actual and decoded frames, across the 900 frames of the movie. For comparison, shuffle data was generated by randomly shuffling the cell-cell pairing of the look-up matrix and "observed response". A similar procedure was used for the 20 trials of the scrambled sequence, and the corresponding middle 20 trials of the continuous movie were used.

Rearranged scrambled movie analysis

To differentiate the effects of visual content versus visual continuity between consecutive frames, we compared the responses of the same neuron to the continuous movie and the scrambled sequence. In the scrambled movie, the same visual frames as the continuous movie were used, but they were shuffled in a pseudo random fashion. The same scrambled sequence was repeated for 20 trials. The neural response was computed at each frame of the scrambled sequence, keeping the frames in the chronological order of presentation. Then the scrambled sequence of frames was rearranged to recreate the continuous movie and the corresponding neural responses computed. To address the latency between movie frame presentation and its evoked neural response, which can differ across brain regions and neurons, this calculation was repeated for rearranged scrambled sequences with variable delays between τ= -500 to +500 ms (i.e., -150 to +150 frames of 3.33ms resolution, in steps of 5 frames or 16.6ms). The correlation coefficient was computed between the continuous movie response and this variable delayed response at each delay as $r_{measured}(\tau) = \text{corrcoef}(R_{continuous}, R_{scramble\text{-}rearranged}(\tau))$. $R_{continuous}$ is the continuous movie response, obtained at 3.33ms resolution and similarly, $R_{scramble\text{-}rearranged}$ corresponds to the scrambled response after rearrangement, at the latency τ. The latency τ yielding the largest correlation between the continuous and rearranged scrambled movie was designated as the putative response latency for that neuron. This was used in Extended Data Fig. 12. The value of $r_{measured}(\tau_{max})$ was bootstrapped using 100 randomly generated frame reassignments, and this was used to z-score $r_{measured}(\tau_{max})$, with z-score > 2 as criterion for significance. The resultant z-score is reported in Extended Data Fig. 12.

Acknowledgements: We thank the Allen Brain Institute for provision of the dataset, Dr. Josh Siegle for help with the dataset, Dr. Krishna Choudhary for proof-reading of the text and Dr. Massimo Scanziani for input and feedback. This work was supported by grants to M.R.M. by the National Institutes of Health NIH 1U01MH115746.

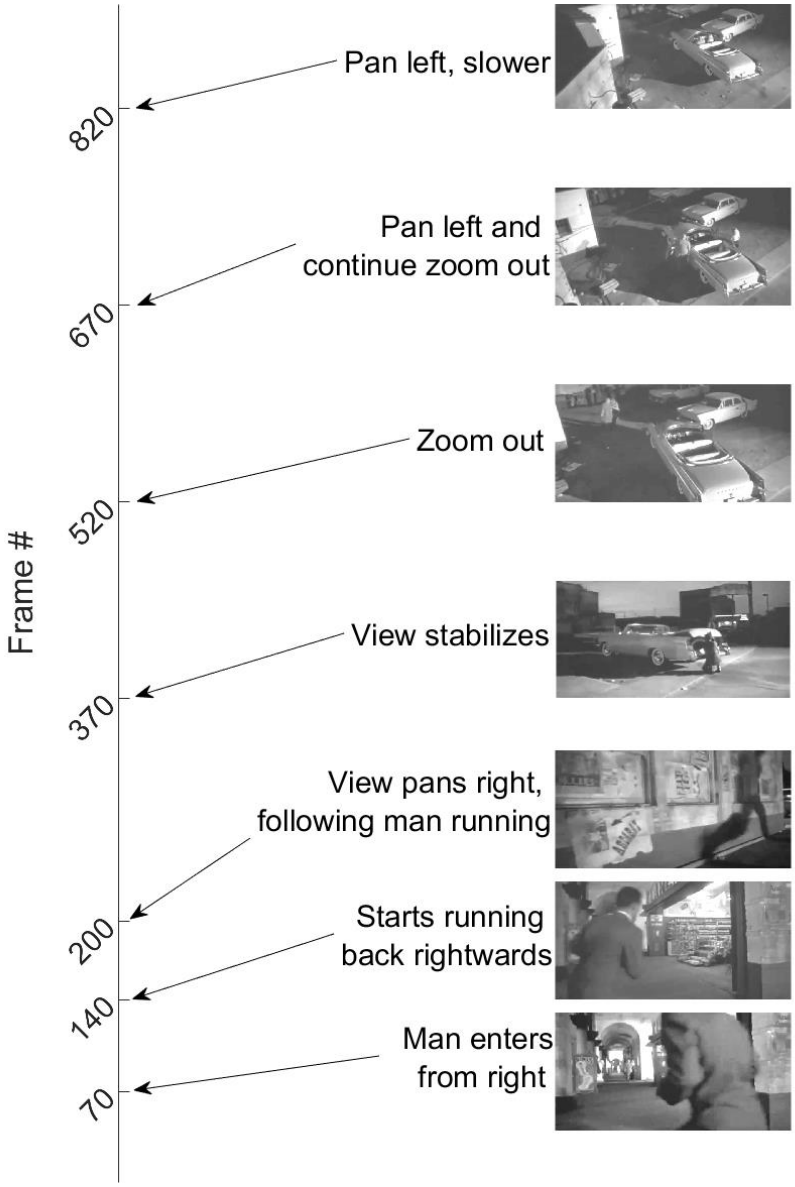
Author contributions: C.S.P. performed the analyses with input from M.R.M. C.S.P. and M.R.M. wrote the manuscript.

Competing interests: Authors declare that they have no competing interests.

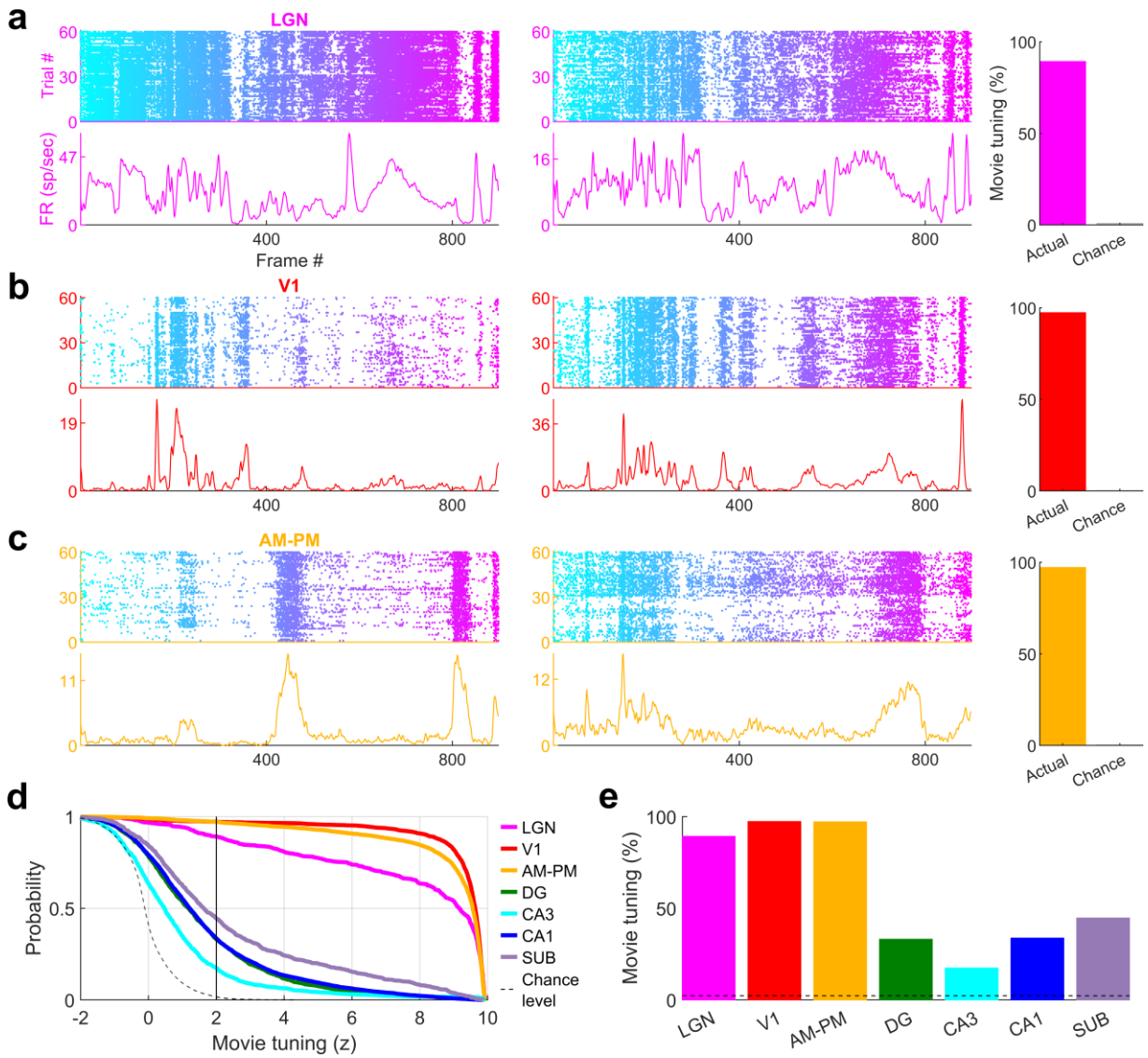
Data availability: All data is publicly available at the Allen Brain Observatory – Neuropixels Visual Coding dataset (© 2019 Allen Institute, https://portal.brain-map.org/explore/circuits/visual-coding-neuropixels).

Code availability: All analyses were performed using custom-written code in MATLAB version R2020a. Codes written for analysis and visualization are available from the corresponding authors at reasonable request.

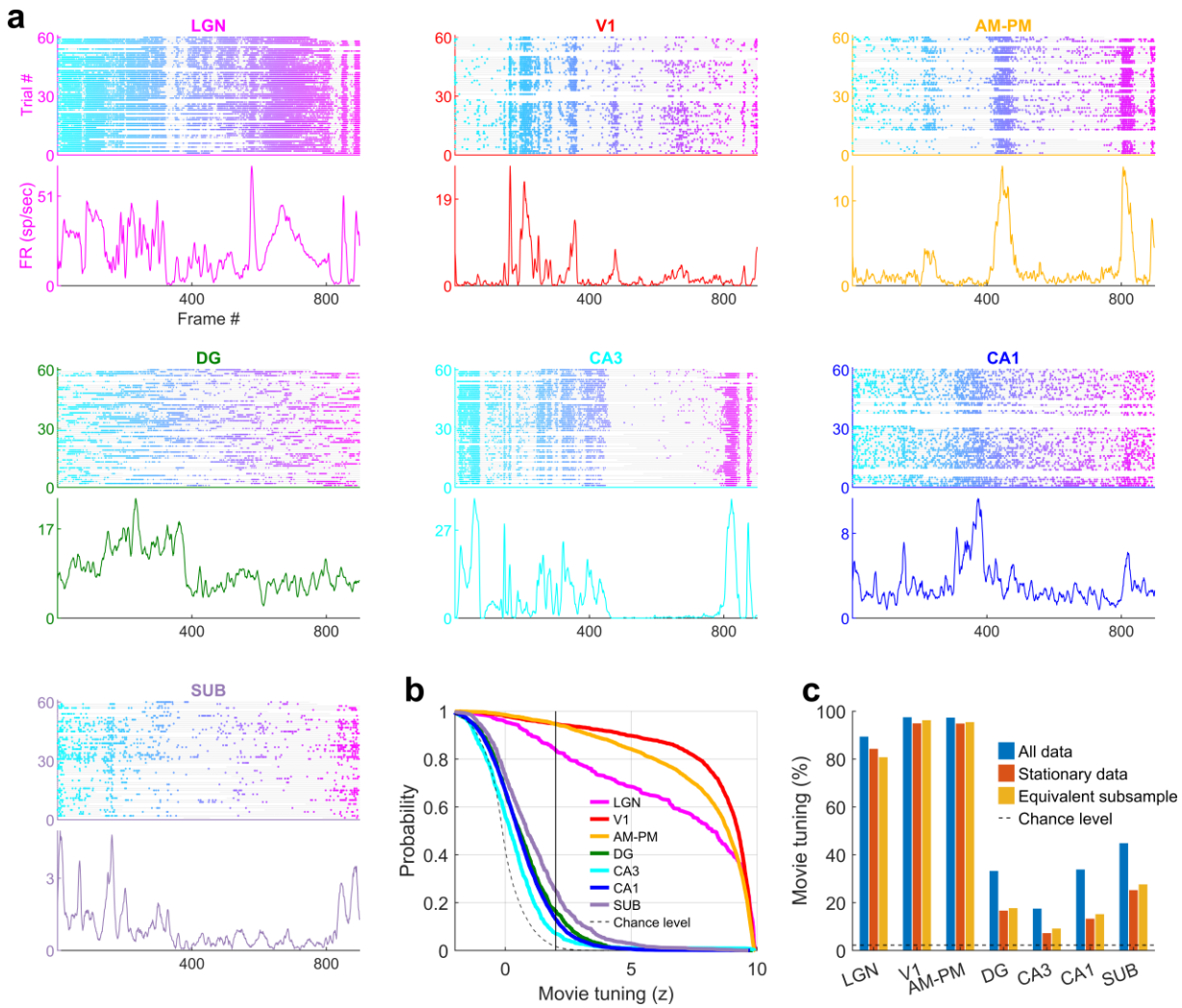
Extended Data Figures



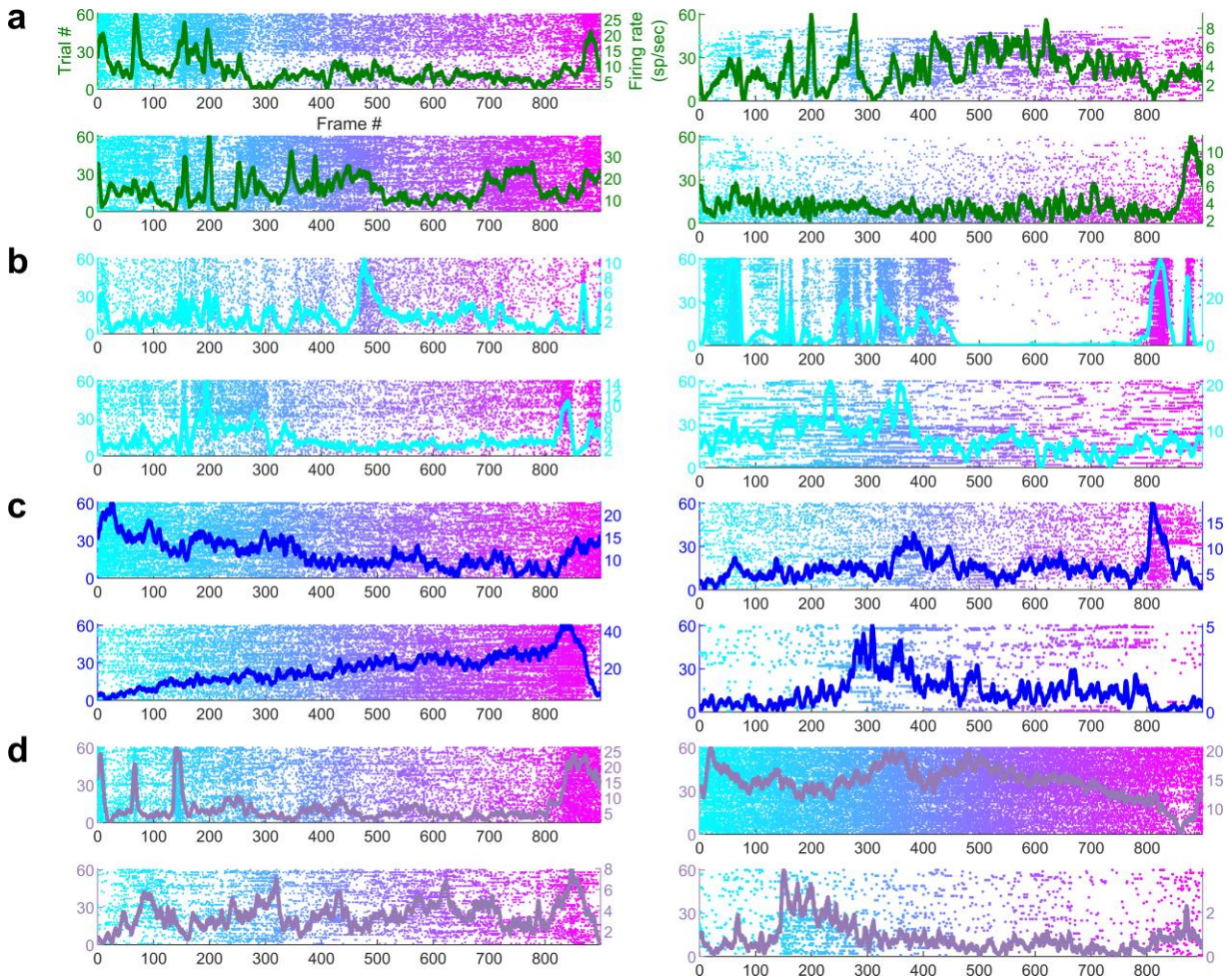
Extended Data Fig. 1 | The movie. The 30 second long, isoluminant movie with frame numbers denoting key episodes in this continuous segment.



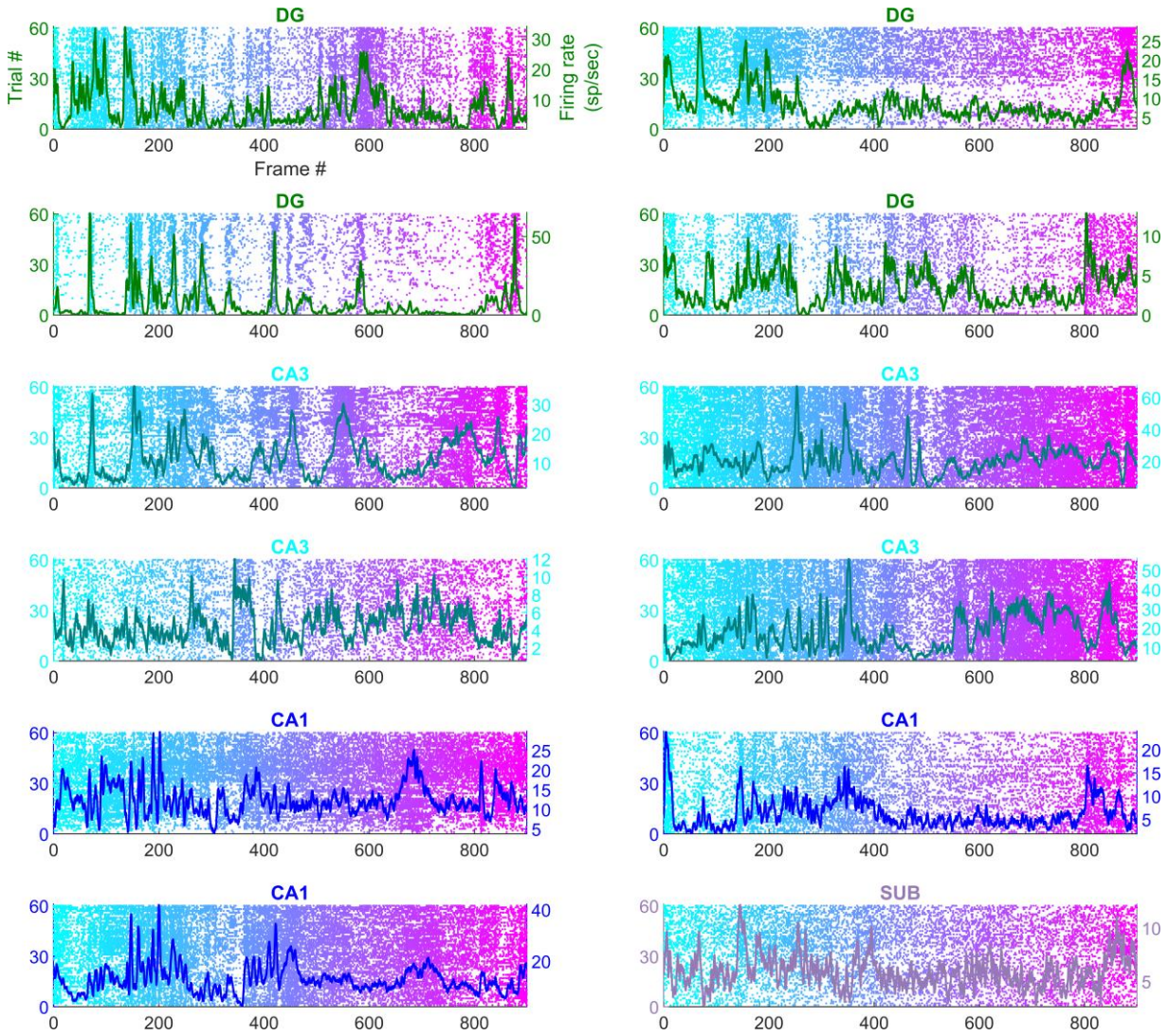
Extended Data Fig. 2 | Movie selectivity across brain areas. (a) Similar to Fig. 1, representative single cells from LGN showing selective movie responses. Fraction selective are shown by the bar chart on the right. **(b)** Same as (a), for V1. **(c)** Same as (a), for higher visual areas AM-PM. **(d)** Cumulative distribution of movie selectivity across all broad spiking cells, including tuned cells (z>2 vertical black line, see *Methods*). Largest prevalence of selectivity in broad spiking neurons was seen in primary visual cortex (V1, 97.3%, 2606 out of 2679) and least in CA3 hippocampus (17.3%, 168 out of 969). **(e)** All brain regions analyzed showed far greater selectivity than the chance level (dashed gray line).



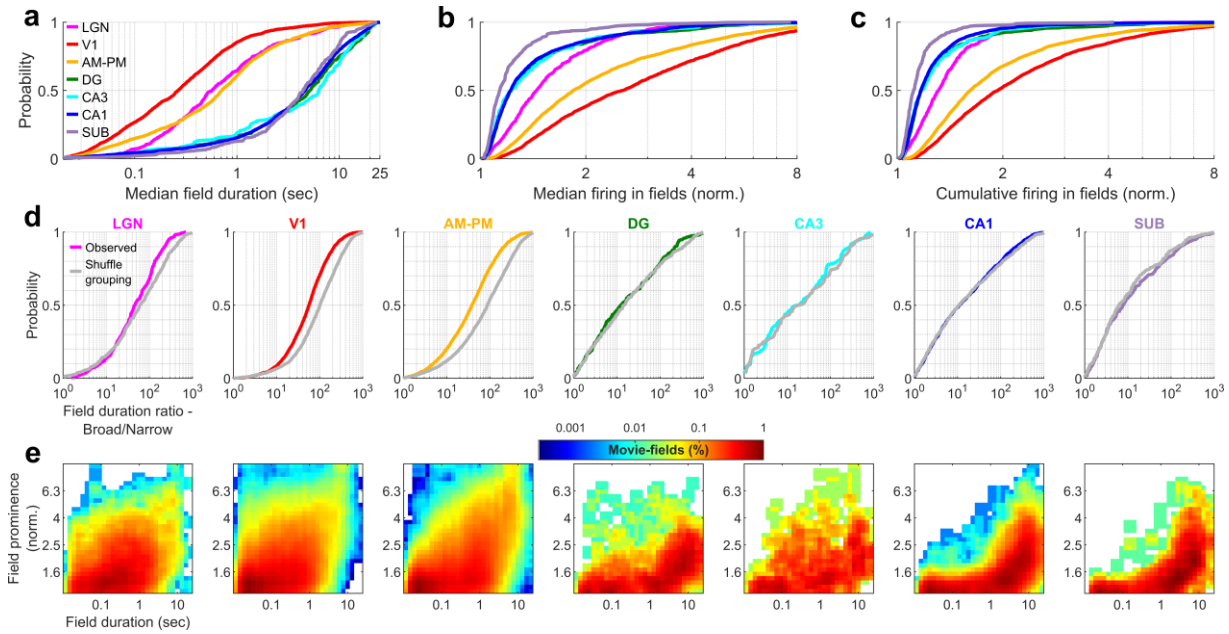
Extended Data Fig. 3 | Movie tuning is unaffected by locomotive state of the mouse (a) Similar to Fig. 1, a representative cell from each brain region showing significant modulation movie tuning using only the data when the mouse was immobile, while excluding the data when the mouse was running (stationary data, see *Methods*). All cells except for DG are from Fig. 1 or Extended Data Fig. 2. **(b)** Fraction of selective neurons was significantly above chance in all brain regions, ranging from 94.7% in V1 up to 7.1% in CA3 in the stationary data. **(c)** To explicitly test the effect of running on movie selectivity, we compared the results in (b) with a random subsample of data that included running and stationary, to control for the loss of data (see *Methods*). Prevalence of movie selectivity was not significantly different (KS-test p>0.18) in these 2 subsamples, except in CA1 (p=0.007, 13.1% in stationary data, 15.0% in the equivalent subsample). Only sessions with at least 300 seconds of stationary data were used in this analysis to ensure sufficient statistical power.



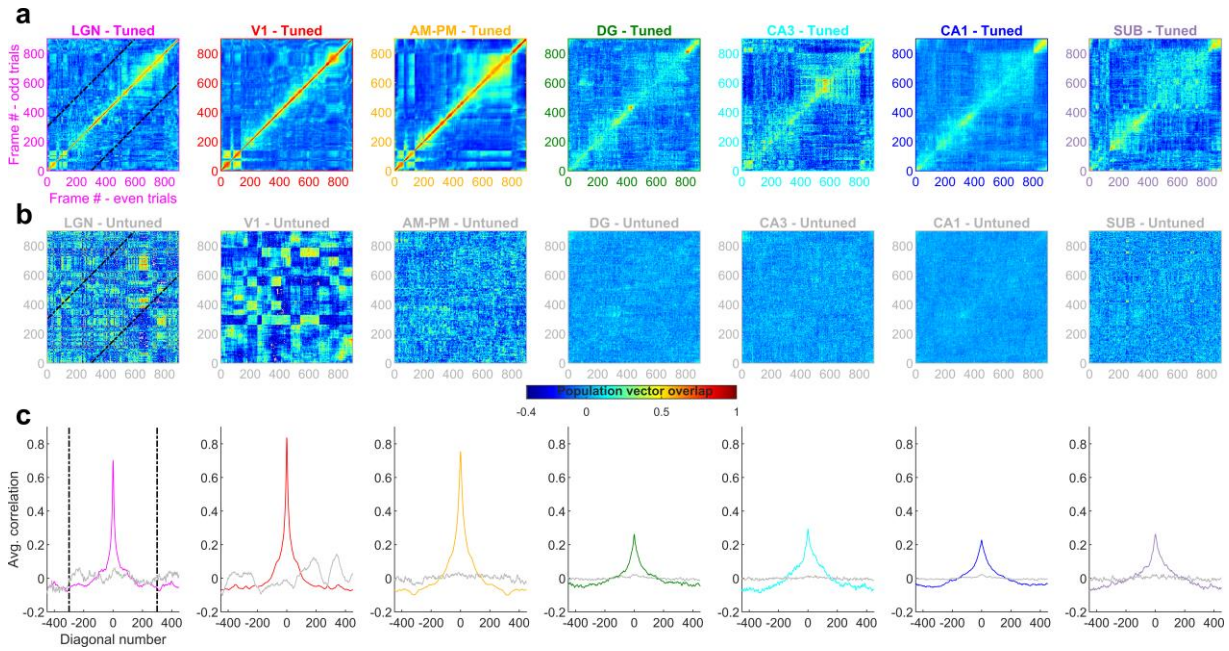
755 **Extended Data Fig. 4 | Simultaneously recorded hippocampal cells have different movie tuning.** Four simultaneously recorded and significantly movie-tuned cells from each from **(a)** Dentate gyrus, **(b)** CA3, **(c)** CA1 and **(d)** Subiculum. Each cell shows different movie selectivity. Average responses are overlaid (on raster plots), and their color corresponds to the different brain regions, described in Fig. 1 legend.



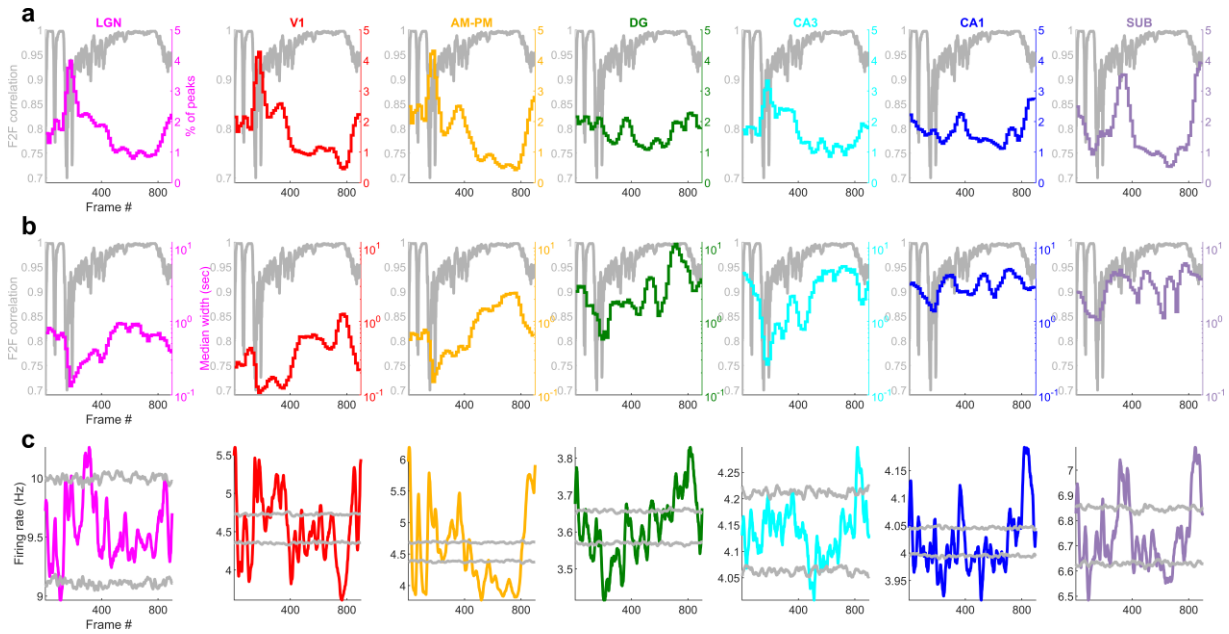
Extended Data Fig. 5 | Few hippocampal neurons had greater than 5 movie-fields. Handful of movie-tuned neurons from dentate gyrus (row 1 and 2), CA3 (row 3 and 4), CA1 and subiculum (bottom-right), had multiple movie-fields. Average responses are overlaid (on raster plots), similar to Fig. 1 and Extended Data Fig. 4.



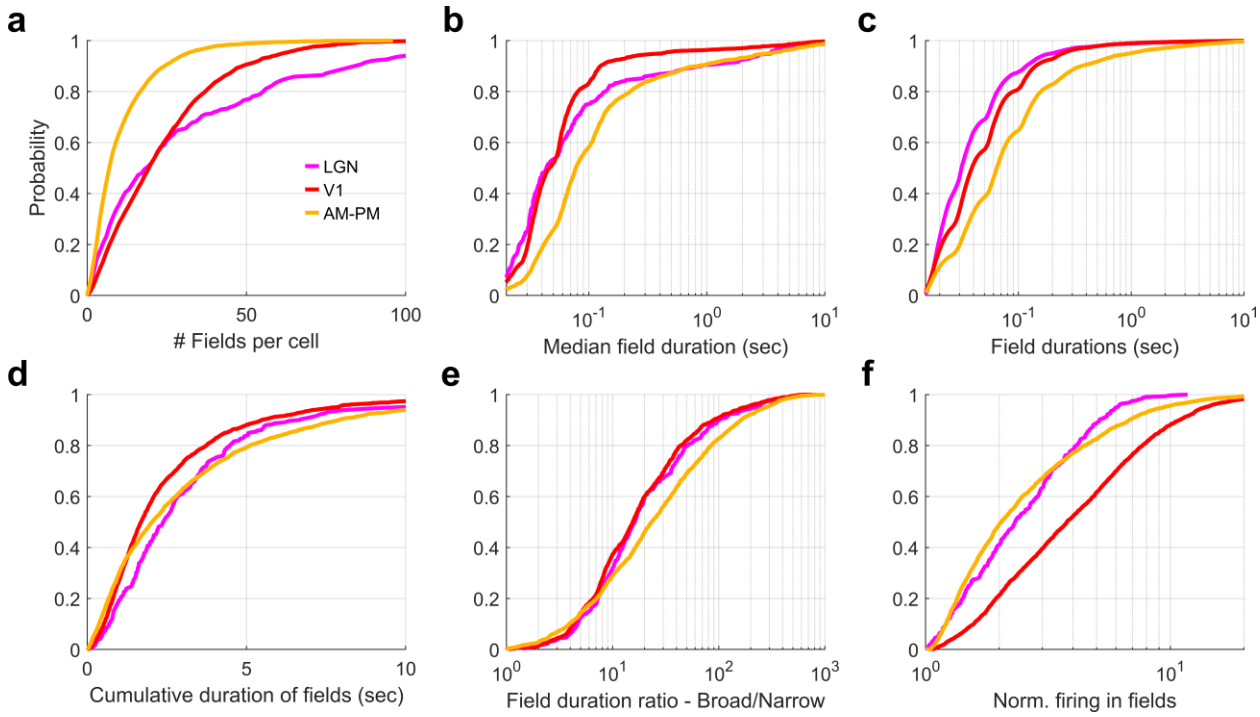
Extended Data Fig. 6 | Movie-field duration ratios are shorter than expected by chance in visual, but not hippocampal areas. (a) Distribution of median duration of movie-fields, computed across all fields of a neuron. All visual-hippocampal region pairs were significantly different (KS-test $p<7.1\times10^{-31}$). DG-CA3 and DG-CA1 were not significantly different, but other visual-visual and hippocampal-hippocampal region pairs were significantly different. (KS-test $p<0.04$). CA3 had the largest median field duration (6.3±0.48s), and V1 had the smallest (0.27±0.03sec). Surprisingly, LGN movie-field durations (0.57±0.13s) were significantly longer than V1 ($p=2.5\times10^{-21}$), and comparable to those in the higher order brain areas (0.71±0.05s). **(b)** Median firing in movie-fields, normalized by that in the shuffle response, obtained as the median from all fields of a neuron. This metric is significantly different across all brain region pairs (KS-test $p<3.4\times10^{-5}$), except DG-CA3, CA3-CA1 and DG-CA1 pairs. The largest median firing was seen for V1 (2.5±0.05), and the smallest in subiculum (1.13±0.03). **(c)** Cumulative firing in movie-fields, normalized by that in the shuffle response, obtained by adding the firing within all fields of a neuron was significantly different across all brain region pairs (KS-test $p<3.0\times10^{-7}$), except DG-CA3, CA3-CA1 and DG-CA1. V1 response was largest (1.93±0.04), and subiculum was the smallest (1.11±0.02). **(d)** For each brain region, the movie-field duration ratio was recalculated by randomly reassigning the cell ids to all the movie peaks from that brain region. Using this new assignment of movie peaks to a cell, we obtain the chance level of mega-scale index (largest/smallest peak duration) within a cell. The observed mega-scale index was lesser than chance in all the visual areas (KS-test $p<3.2\times10^{-3}$, median was 77.5%, 56.2% and 41.7% of chance for LGN, V1 and AM-PM respectively). This was not the case in hippocampal regions ($p>0.23$). **(e)** Histogram of movie-fields, binned for their durations and their prominence, on a log-log scale. The most prominent fields tended to be wider in most brain areas, and this effect was stronger in hippocampal regions, than visual. Note that the histogram color is also logarithmically scaled.



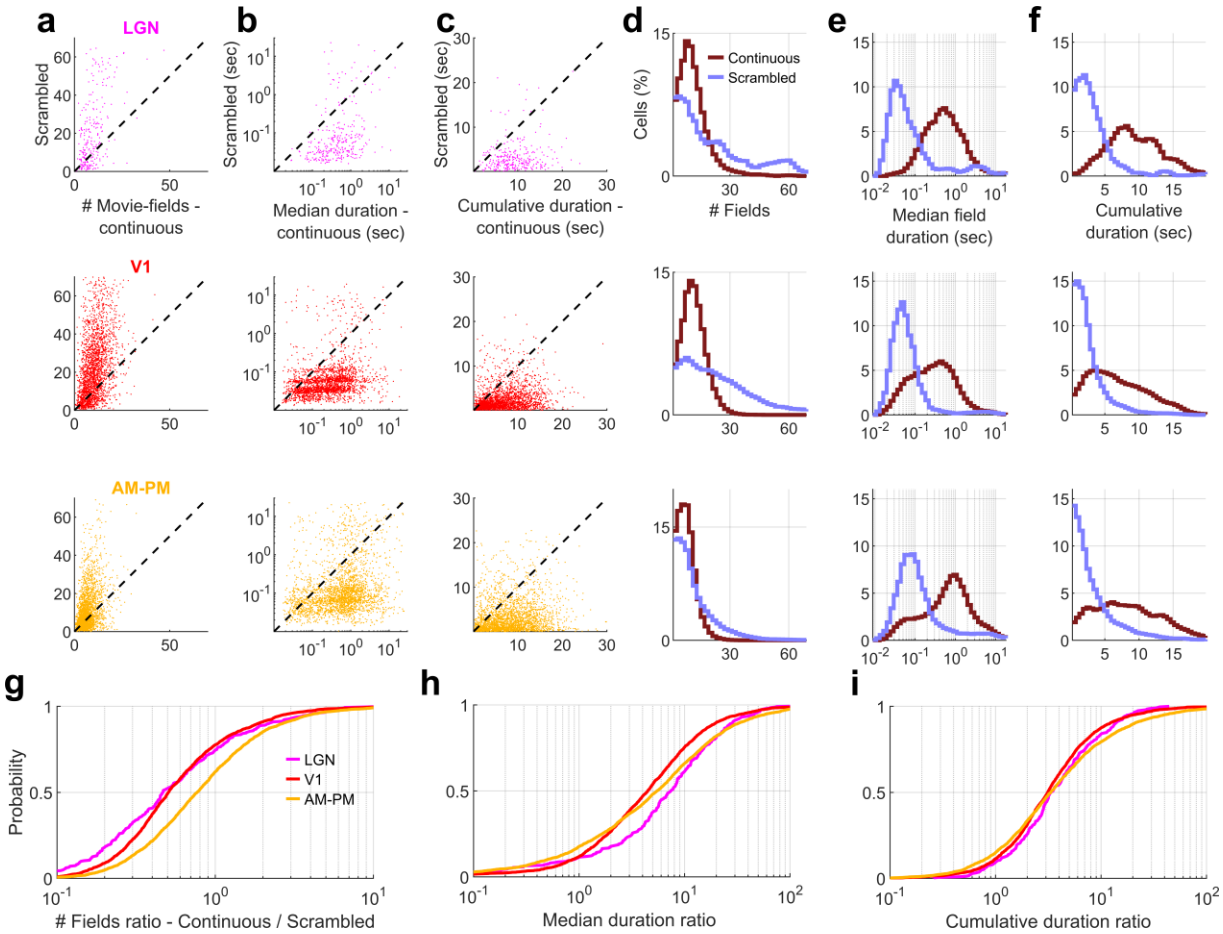
Extended Data Fig. 7 | Population vector overlap is broader in hippocampus and more stable for tuned than untuned cells. (a) Population vector overlap between even and odd trials for the population of tuned neurons show highest overlap along the diagonal, i.e. the same movie frame, for all brain regions. Each neuron's response was normalized by its mean rate and the average response in even as well as odd trials was smoothed by a Gaussian window of 2 frames (66.6ms, see *Methods*). Dashed black lines indicate the -300 and +300 frames away from the diagonal. Notice large correlations (close to unity, horizontal color bar) indicating stable responses. The correlations decay quickly to smaller values for the visual areas but slowly for hippocampal areas, due to their broader movie-fields. **(b)** Same as (a), but for untuned neurons, resulting in a salt and pepper overlap pattern and low values of correlation, indicating lesser stability than the tuned neurons. Since the majority of cells in the visual areas were tuned, the untuned population was smaller, leading to more variable population vector overlap. **(c)** The average overlap as a function of the number of movie frames away from the diagonal. It had a large value in visual regions for the 0th diagonal (colored lines) indicating stable responses, whereas the untuned neuron population (gray lines) were unstable, with values near zero, or chance level. The highest population vector overlap in hippocampal regions was smaller than visual areas but persisted for more frames, due to their broader movie-fields (Full width at half maximum of the peak – 17.3 frames for LGN, 22.7-V1, 39.0-AM&PM, 49.8-DG, 57.4-CA3, 64.7-CA1 and 59.2-subiculum).



Extended Data Fig. 8 | Distribution of movie-fields reflects frame to frame correlation structure of the movie. (a) Histogram of movie-fields across all tuned neurons in a brain region, as a function of the movie frame, showing non-uniform distribution. Frame_n to frame_{n+1} correlation coefficient (F2F correlation), indicating the similarity of 2 consecutive frames, is shown for reference in gray. All distributions were deemed significantly different than a uniform distribution based on a Chi-square goodness-of-fit test ($p<3.8\text{x}10^{-6}$). All distributions were significantly negatively correlated with F2F correlation ($r<-0.18$, $p<10^{-7}$). **(b)** Same as (a), but for the median duration of movie-fields. F2F correlation shown in gray, with large correlation between consecutive frames between frames 400-800 reflected in larger movie-field durations in visual areas. All distributions were significantly different than a uniform distribution based on a Chi-square goodness-of-fit test ($p<10^{-100}$) and all distributions were significantly positively correlated with F2F correlation ($r>0.24$, $p<2\text{x}10^{-13}$). Note- y-axes for the histogram are log-scaled and show larger median durations for hippocampal regions than visual. **(c)** Total firing rate across all broad spiking neurons in different brain regions, showing similar non-uniformity as Fig. 3. All brain regions had significantly negative correlation with the F2F correlation ($r<-0.08$, $p<0.03$), except DG, which was significantly positively correlated ($r=0.21$, $p=2.4\text{x}10^{-10}$). Largest number of above chance deviations were seen for AM-PM (340 frames), and least for CA3 (57 frames). Below chance level deviations were least common in LGN (25 frames), and most common in AM-PM (441 frames). Similar to Fig. 3c-e.

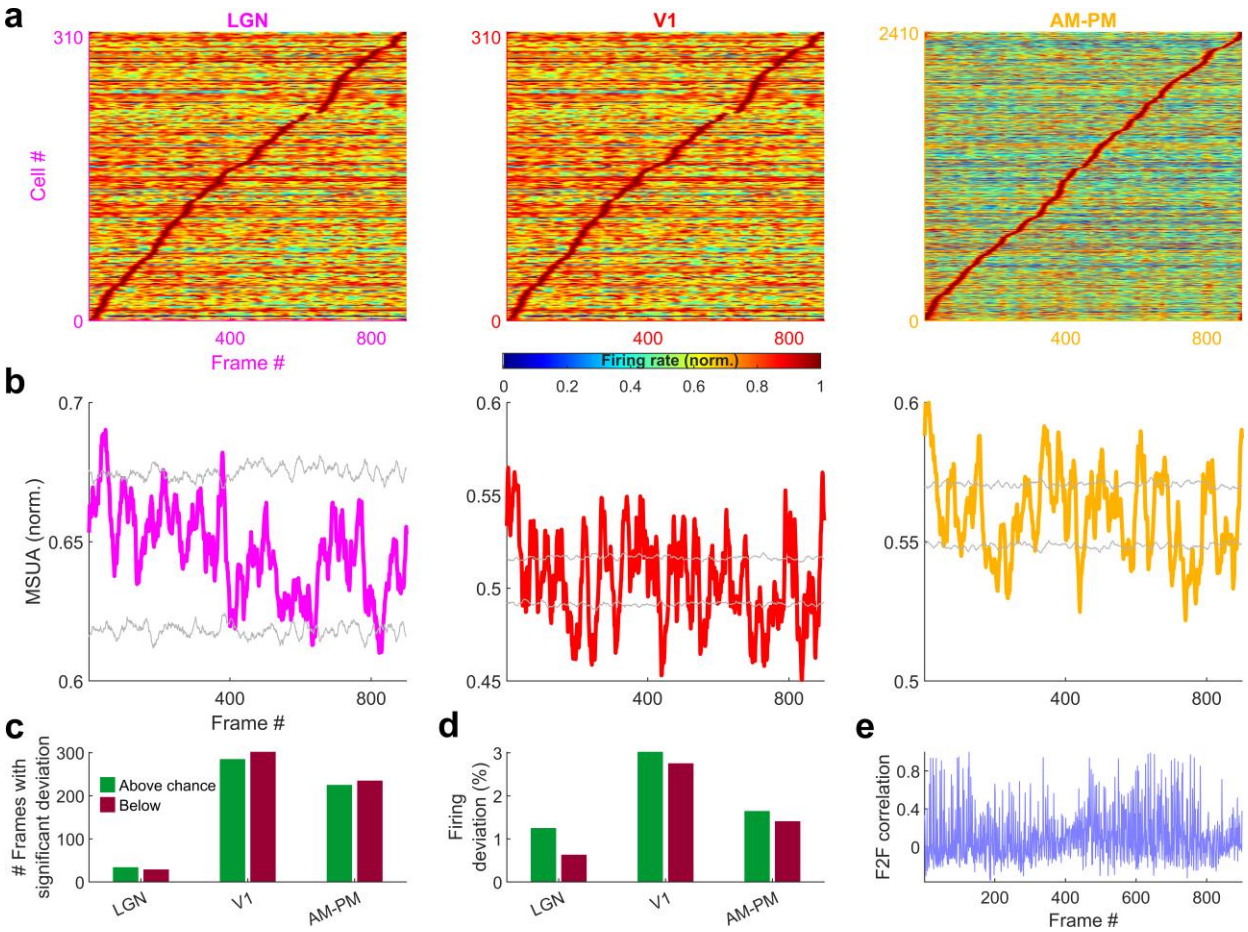


Extended Data Fig. 9 | Scrambled movie elicits narrower but more movie-fields per cell than the continuous movie in the visual regions. (a) Total number of fields per cell for the scrambled sequence were hierarchically arranged, with largest number of fields in LGN (mean±s.e.m., 31.8±2.0), followed by V1 (24.0±0.38) and last AM-PM (11.1±2.1). All three brain regions were significantly different from each other (KS-test $p<2.0\times10^{-5}$). **(b)** Median scrambled movie-field duration was shortest in LGN (43.9±131.2ms), intermediate in V1 (46.2±24.8) and widest in AM-PM (77.6±40.1ms), and differences were significant ($p<7.0\times10^{-4}$). This was much smaller than for the continuous movie (Fig. 2). **(c)** Durations of fields for scrambled sequence across all fields of all neurons from a brain region. These were narrowest in LGN (31.3±6.5ms), followed by V1 (38.6±0.2) and last AM-PM (64.3±6.7). All differences were significant (KS-test $p<7.2\times10^{-136}$). **(d)** Despite these differences, the cumulative duration of movie-fields was comparable across the three brain regions (1.69±0.05sec for V1, 2.03±0.07 for AM-PM and 2.4±0.2 for LGN), but significantly different ($p<1.7\times10^{-5}$). Note the linear scale on the x-axis in this panel compared to the log-scale in other panels. **(e)** Ratio of field durations, i.e., mega-scale index, was smallest in V1 (15.5±1.6), intermediate in LGN (16.3±5.4) and largest in AM-PM (23.4±2.1), and not significantly different between V1 and LGN (p=0.28). V1-AM&PM and LGN-AM&PM were significantly different ($p<5.7\times10^{-5}$). **(f)** Cumulative firing activity, summed across all movie-fields of a given neuron was largest in V1 (3.8±0.1), intermediate in LGN (2.3±0.1) and smallest in AM-PM (2.0±0.07), and significantly different between all brain region pairs ($p<0.02$).

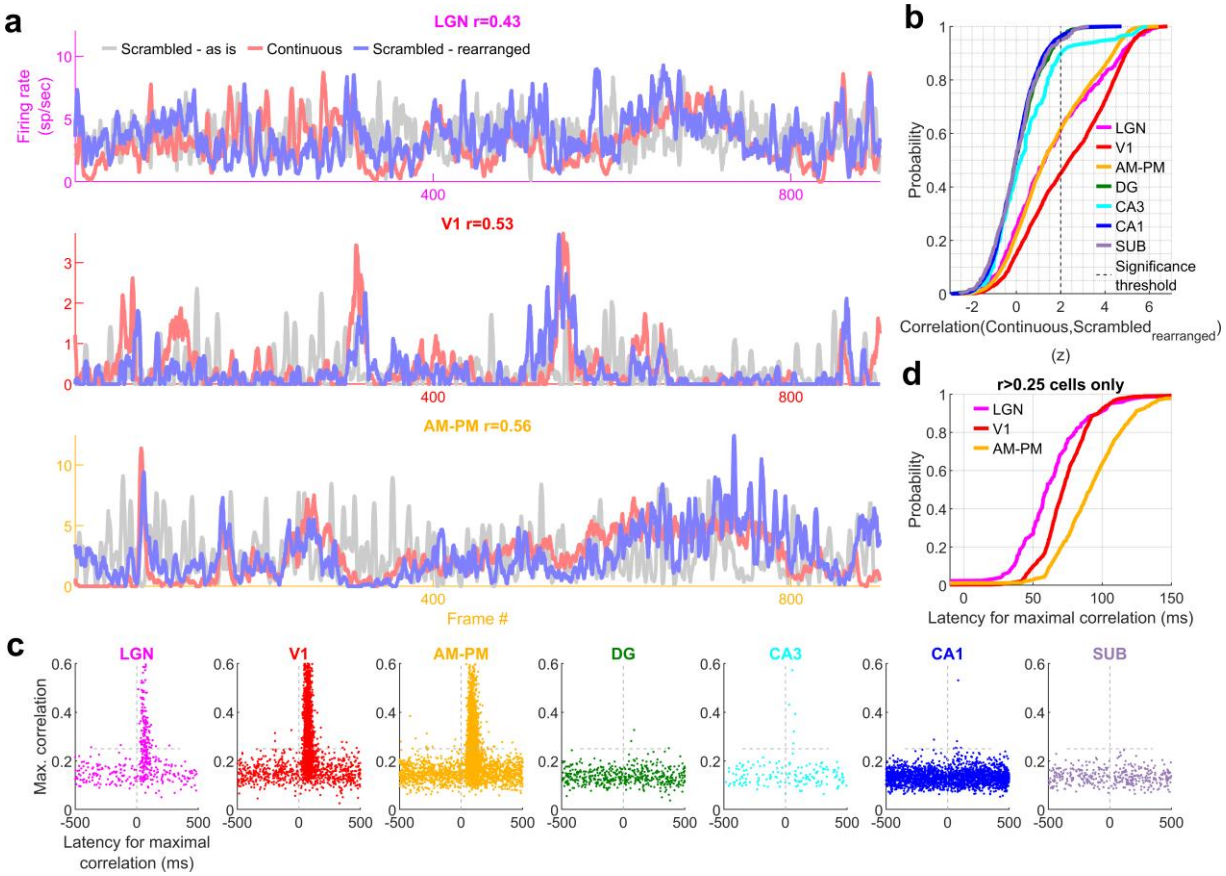


Extended Data Fig. 10 | Scrambled sequence evoked movie-fields were narrower than the continuous movie-fields in all visual regions, cell by cell comparison. Data for only those visual area neurons that were significantly modulated by both the continuous and scrambled movie were used. **(a)** The number of movie-fields per cell for the continuous movie was significantly smaller than that for scrambled sequence in all brain areas (LGN – continuous mean±s.e.m.=10.7±0.42, scrambled=31.8±2.0, KS-test p=2.0x10^{-23}, V1-10.8±0.11 vs. 24.0±0.38, KS-test p=3.7x10^{-210}, AM&PM-6.9±0.07, vs. 11.1±0.21, KS-test p=1.3x10^{-57}). Data are additionally scattered by a small random number for ease of visualization. **(b)** Median duration of movie-fields for a cell was significantly larger for continuous movie, compared to scrambled sequence in all visual regions. (LGN continuous=0.46±0.08sec, scrambled=0.04±0.13, KS-test p=7.1x10^{-65}, V1-0.25±0.03 vs. 0.04±0.02, KS-test p<10^{-150}, AM&PM-0.65±0.04, vs. 0.08±0.04, KS-test p<10^{-150}). **(c)** Cumulative duration of all movie-fields for a cell was significantly larger for continuous movie, compared to scrambled sequence in all visual regions. (continuous=8.9±0.23sec, scrambled=2.4±0.19, KS-test p=3.2x10^{-69}, V1-6.1±0.09 vs. 1.69±0.05, KS-test p=3.3x10^{-296}, AM&PM-7.8±0.1, vs. 2.0±0.07, KS-test p=9.0x10^{-318}). **(d)** Histogram of number of fields per cell, for continuous and scrambled movies. **(e)** Logarithmically spaced histogram of median field durations was significantly different between continuous and scrambled sequence. **(f)** Similar to (e), histogram of cumulative duration of movie-fields for each cell. **(g)** The ratio of number of fields per cell between continuous and scrambled movies was biased to smaller than unity values for all brain regions, with the largest bias for LGN (0.46±0.08), intermediate for V1 (0.5±0.04), and least for AM-PM (0.77±0.05). **(h)** The median field duration ratio was biased to values greater than unity, with the largest bias for LGN (7.4±1.4), least for V1 (4.5±0.68), and intermediate for AM-PM (5.5±0.82). **(i)** The cumulative field duration ratio was also

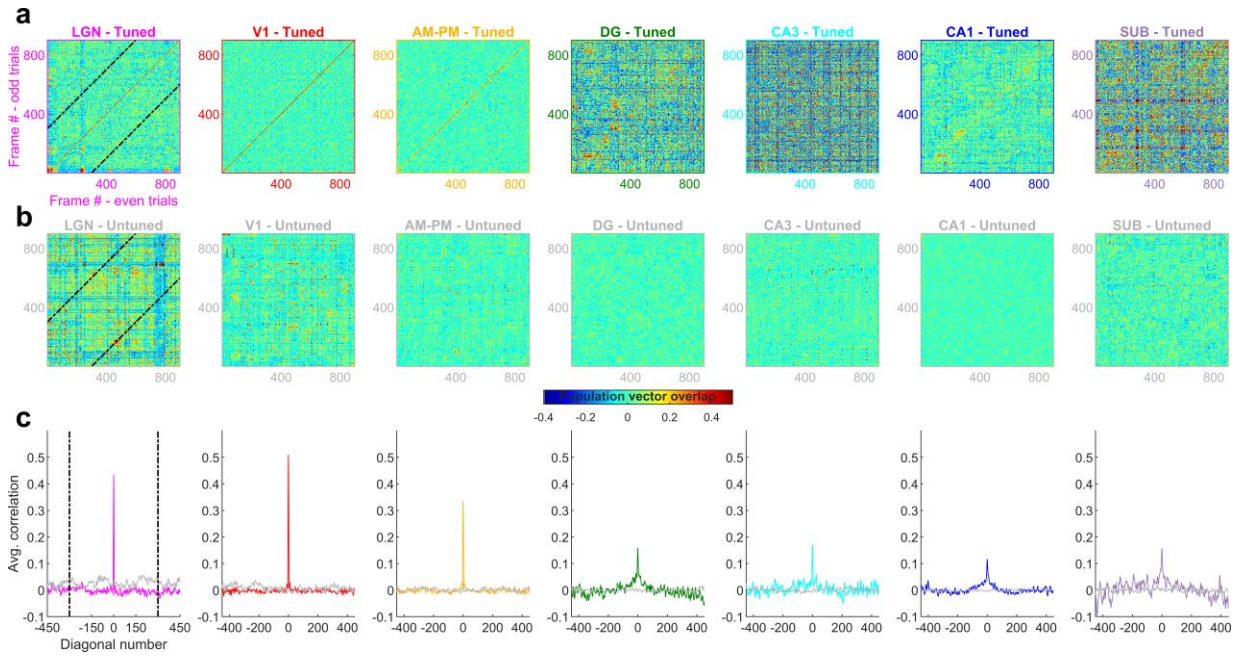
biased to values greater than unity, with similar biases for LGN (3.37±0.36), V1 (3.1±0.3), and AM-PM (3.3±0.67).



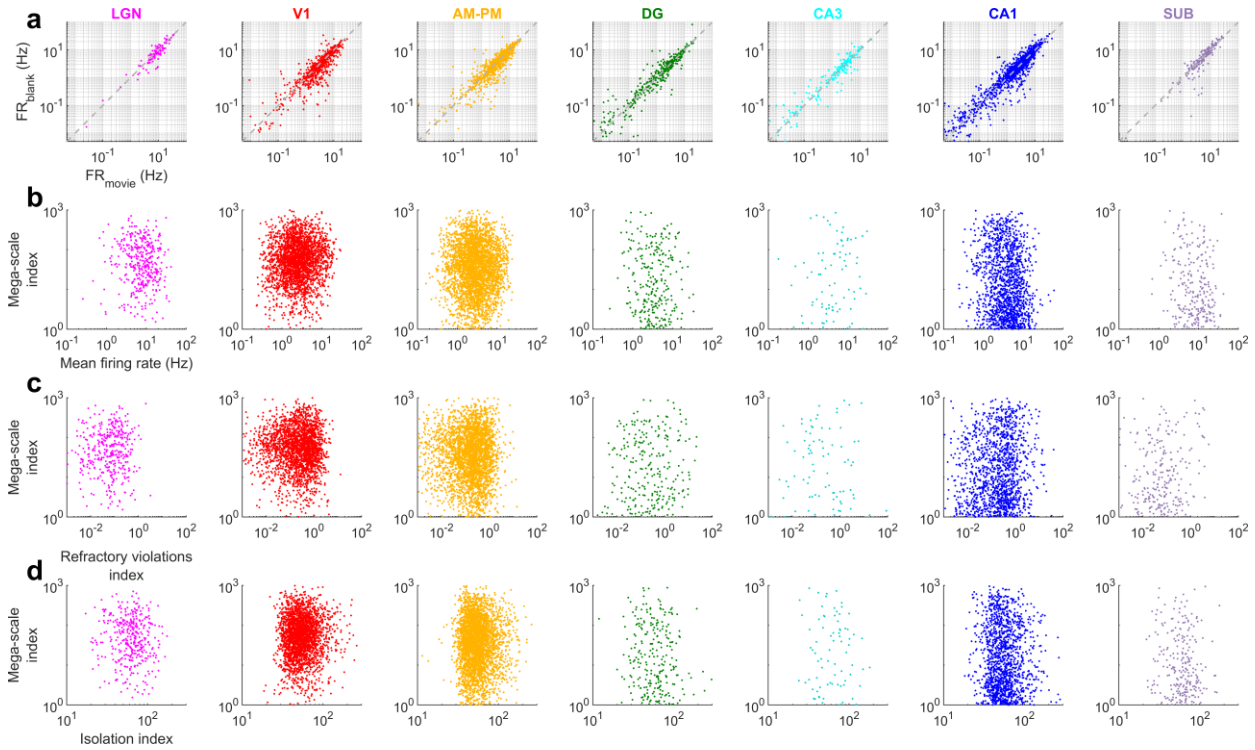
Extended Data Fig. 11 | Multiple-single unit activity (MSUA) across all movie-tuned neurons in a brain region shows greater modulation than chance levels for scrambled sequence in all visual areas. (a) Stack plot of tuned responses to the scrambled movie presentation from each brain region, arranged in their increasing order of the frame corresponding to the peak firing response. Each response is normalized by the peak response. **(b)** Colored trace- average response, across all tuned responses from (b). gray trace - chance level, z=±4, corresponding to the *p*=0.025 level after Bonferroni correction. **(c)** Number of frames for which the observed response exceeds (or falls below) z=±4 cutoff from (b), called significantly deviant frames. V1 had the largest number of positive (279 frames) and negative (297) deviant frames, AM-PM had intermediate (225 & 235), and LGN had the least (31 & 29). **(d)** Firing rate deviation above chance levels, corresponding to the significant frames, as identified in (c), normalized by the mean rate of the MSUA. Largest deviation was observed in V1 (above-3.1 and below-2.7%), and least in LGN (1.1% and 0.45%) Compare with Fig. 3. **(e)** Frame to frame correlation, from Fig. 4a for comparison. This was not significantly correlated with the MSUA responses in (b), for any of the brain regions (*Pearson* correlation coefficient LGN *p*=0.06, V1 *p*=0.07, AM-PM *p*=0.26).



Extended Data Fig. 12 | Latency of responses to scrambled-sequence correspond to the anatomical hierarchy of visual areas. (a) Average response for one representative cell from each visual region, that had high similarity between continuous movie and rearranged scrambled sequence responses (see *Methods*). Gray response in background corresponds to the chronological scrambled sequence. **(b)** Cumulative histogram of z-scored correlation between continuous and scrambled-rearranged tuning responses (see *Methods*). Dotted black line indicates significance threshold of z>2. **(c)** The latency at which continuous and scrambled-rearranged responses were maximally correlated showed high values (heuristically above 0.25) in a short range of positive latencies for LGN, V1 and AM-PM neurons. This analysis was restricted to neurons tuned in continuous as well as scrambled movies. Similar analysis for hippocampal regions resulted in almost no correlations above 0.25. **(d)** Cumulative histogram of latencies when the continuous and scrambled-rearranged responses were maximally correlated was smallest for LGN (59.5±4.6ms), and largest for higher visual areas, AM-PM (91.6±1.6ms). Hippocampal regions were excluded, owing to lack of data with correlation about 0.25.



Extended Data Fig. 13 | Population vector overlap was narrow at the diagonal with scrambled movie **(a)** Population vector overlap between even and odd trials for tuned neurons showing higher overlap along the diagonal for all brain regions. Black lines indicate the -300 and +300 diagonal, whereas the main diagonal is the 0th diagonal. **(b)** Same as (a) but for untuned neurons, resulting in a salt and pepper overlap without higher correlation around the diagonal. **(c)** The average overlap along diagonals had a large value in visual regions for the 0th diagonal, which was not true for the untuned neuron population. Average correlation in hippocampal regions was broader and lesser in magnitude compared to visual regions. Similar to Extended Data Fig. 7. Full width at half maximum of the peak – 4.4 frames for LGN, 4.8-V1, 5.2-AM&PM, 7.6-DG, 5.7-CA3, 10.8-CA1 and 15.1-subiculum.



Extended Data Fig. 14 | Movie presentation did not alter hippocampal firing rates and the mega-scale coding was unrelated to cluster quality. (a) More than 50% of hippocampal place cells shut down during maze exploration. In contrast, there was no consistent pattern of neural activation or shutdown during the movie presentation in all brain areas. To make a more conservative estimate, this comparison was restricted to units whose firing rates did not differ by more than 20% across the two movie blocks. Further only the data when the animals were immobile was used to avoid confounding effects of running. **(b)** The mega-scale index was only weakly correlated with the mean firing rate of a neuron in V1 (Pearson's correlation coefficient r=0.08, p=7.3x10^{-5}), CA1 (r=-0.14, p=3.5x10^{-8}) and subiculum (r=-0.14, p=0.02), and was uncorrelated for other brain regions (p>0.05) **(c)** The refractory violations index was uncorrelated with the mega-scale index (lower index means better cluster quality[31,79]) for all brain regions (p>0.05). To remove potential confounding effect of mean firing rates, we computed the partial correlation coefficient, by factoring out the mean firing rate). **(d)** Similar to (c), the isolation index (greater isolation index means better cluster quality[31,80]) was uncorrelated with the mega-scale index for all brain regions (partial correlation coefficient, by factoring out the mean firing rate, p>0.12). Factoring out mean firing rate was deemed necessary since the isolation index was typically positively correlated, and the refractory violations index was typically negatively correlated with mean rate. The mega-scale index comparisons were restricted to movie active, tuned neurons with at least two movie peaks. Note-log spaced axes for (a)-(d)