HHS Public Access

Author manuscript

Nat Neurosci. Author manuscript; available in PMC 2017 September 01.

Published in final edited form as:

Nat Neurosci. 2017 March ; 20(3): 427–437. doi:10.1038/nn.4479.

REM sleep selectively prunes and maintains new synapses in development and learning

Wei Li[1,2,#], **Lei Ma**[1,2,#], **Guang Yang**[3], and **Wenbiao Gan**[1,2,*]

[1]Drug Discovery Center, Key Laboratory of Chemical Genomics, Peking University Shenzhen Graduate School, Shenzhen, 518055, China

[2]Skirball Institute, Department of Neuroscience and Physiology, New York University School of Medicine, New York, NY 10016, USA

[3]Department of Anesthesiology, New York University School of Medicine, New York, NY 10016, USA

Abstract

The functions and underlying mechanisms of rapid eye movement (REM) sleep remain unclear. Here we show that REM sleep prunes newly-formed postsynaptic dendritic spines of layer 5 pyramidal neurons in the mouse motor cortex during development and motor learning. This REM sleep-dependent elimination of new spines facilitates subsequent spine formation in development and when a new motor task is learned, indicating a role of REM sleep in pruning to balance the number of new spines formed over time. In addition, REM sleep also strengthens and maintains some newly-formed spines that are critical for neuronal circuit development and behavioral improvement after learning. We further show that dendritic calcium spikes arising during REM sleep are important for pruning and strengthening of new spines. Together, these findings indicate that REM sleep has multifaceted functions in brain development, learning, and memory consolidation by selectively eliminating and maintaining newly-formed synapses via dendritic calcium spike-dependent mechanisms.

Because of its close association with dreaming, the function of rapid eye movement (REM) sleep has been a topic of substantial interest and speculation[1–9]. REM sleep accounts for up to 50% of total sleep time during early development and its duration increases after learning later in life[1, 10]. REM sleep deprivation (REMD) during development leads to deficits in the developing visual system as well as behavioral changes in adulthood[11, 12]. Multiple lines of evidence have shown that REMD often impairs performance improvement after learning[3–6, 13]. These studies suggest that REM sleep has an important role in brain

*Corresponding author. gan@saturn.med.nyu.edu.
#These authors contributed equally to this work.

Author contributions

W.L. and L.M. contributed equally to this work. G.Y. and W.-B.G. initiated the project. W.L., L.M, G.Y. and W.-B.G. designed the experiments. W.L. and L.M. performed the experiments and analyzed the data with the help from W.-B.G. W.L., L.M, and W.-B.G. prepared the manuscript with the input from G.Y.

Competing financial interests

The authors declare no competing financial interests

maturation and memory consolidation. On the other hand, REMD has no significant effects on experience-dependent plasticity in the developing cat visual cortex[14]. The detrimental effects of REMD are not observed in some studies of declarative memories, and loss of REM sleep in humans with brainstem lesions has no obvious impact on cognitive functions[6, 15, 16]. Furthermore, REM sleep suppression by antidepressant drugs either has no significant effect or even improves memories[17, 18]. Thus, the precise functions of REM sleep in development and learning still remain unclear[3–8, 16, 18–20]. It has been suggested that REM sleep is important for removing spurious neuronal connectivity during development and learning[9, 21]. Such an "unlearning" process is thought to increase signal to noise ratio in neuronal network[9, 22], but the experimental evidence supporting the "unlearning" theory of REM sleep is lacking.

One prominent feature of brain development and plasticity is that a large number of new synapses are formed each day, but only a small fraction of them are stably maintained over time[23–26]. Because up to 5–10% of new synapses are formed daily, a selective process of pruning and maintaining new synapses is necessary for the brain to store new information continuously without disrupting previously-acquired memories[23–25]. The mechanisms underlying this selective process remain unknown. The long duration of REM sleep during early development coincides with the occurrence of extensive synapse formation and elimination. Exposure to novel experience increases the expression of synaptic plasticity-related genes during REM sleep[27–30]. These findings raise the possibility that REM sleep may affect the processes of synapse formation and maintenance that are fundamental to brain development and learning. While recent studies have shown that REM sleep has no significant effect on the formation of motor learning-induced new dendritic spines of layer V (L5) pyramidal neurons in the motor cortex[31], whether and how REM sleep is involved in the pruning or maintenance of newly-formed synapses remains unknown.

To investigate the functions and underlying mechanisms of REM sleep, we examined the impact of REM sleep on postsynaptic dendritic spines of L5 pyramidal neurons in the mouse primary motor cortex. We found that REM sleep prunes and balances the number of newly-formed spines during development and after motor learning. Concurrently, REM sleep also strengthens and maintains a subset of new spines that are critical for neuronal circuit development and performance improvement after learning. Furthermore, REM sleep-dependent spine pruning and strengthening are mediated by NMDA receptor-dependent dendritic calcium spikes. Together, our findings show that REM sleep contributes to brain development, learning and memory storage by selectively pruning and maintaining new synapses via dendritic calcium spike-mediated mechanisms.

RESULTS

REM sleep prunes new spines during development and after motor learning

Previous studies have shown that newly-formed synapses are largely pruned while only a small fraction of them are maintained[25]. It is generally believed that experience is important for pruning imprecise synaptic circuits during development and throughout adult life[23–25, 32]. The role of REM sleep in this process has not been explored. To examine whether REM sleep has a role in synapse pruning, we first used transcranial two-photon

microscopy to identify newly-formed and existing postsynaptic dendritic spines of layer V (L5) pyramidal neurons in the motor cortex of mice at postnatal day (P) 21 (Fig. 1a)[25]. At this developmental stage, the rate of new spine formation and the amount of REM sleep were significantly higher than in adolescence and adulthood ($P < 0.05$) (Fig. 1b, c)[25, 33]. We then examined the effect of REM sleep deprivation (REMD) on the survival of newly-formed and existing spines (Fig. 1a). REM sleep was deprived by gentle handling based on electroencephalography (EEG) and electromyography (EMG) recordings as described previously[31] (Supplementary Fig. 1a). The gentle handling procedure led to a significant reduction (~90%) in the amount of REM sleep in REMD mice than in non-deprived control mice over an 8 hour period ($P < 0.05$) (Fig. 1d and Supplementary Fig. 2a–c). Notably, we found that the elimination rate of new spines was significantly higher in non-deprived control mice than in REMD mice over 8 hours (65.1 ± 1.6% versus 42.3 ± 4.4%, $P < 0.01$) (Fig. 1e,f and Supplementary Table 1). In contrast, the elimination rate of existing spines was comparable between the non-deprived and REMD groups ($P = 0.701$) (Fig. 1g).

To control for potential non-specific effects of REMD related to the gentle handling procedure, a group of mice were disturbed during non-REM (NREM) sleep in a similar manner as in REMD (Supplementary Fig. 2d). We found that NREM sleep disturbance (NREM-d) did not affect the total amount of REM sleep (Fig. 1d and Supplementary Fig. 2a,b). Importantly, the elimination of new spines, but not existing spines, was significantly higher in NREM-d mice than in REMD mice ($P < 0.01$) (Fig. 1f,g). These results indicate that REM sleep has an important role in pruning newly-formed dendritic spines in the developing motor cortex.

As development proceeds, there is a substantial reduction in the rate of new spine formation (Fig. 1b). Consistent with previous studies[31, 34], new spine formation increases significantly after motor skill learning on an accelerated rotarod in adolescent mice (Fig. 1b). We found that the elimination rate of new spines formed during the first 8 hours after motor training was also significantly higher in non-deprived control mice and NREM-d mice than in REMD mice over the next 8–16 hours ($P < 0.01$) (Fig. 1h–k, 1m and Supplementary Fig. 3 and Supplementary Tables 2–3). No significant difference was observed in the elimination rate of existing spines among non-deprived control, NREM-d and REMD mice during the same period ($P > 0.05$) (Fig. 1l,n). Furthermore, the elimination of new spines formed during the first 24 hours after motor learning, but not existing spines, was significantly higher over the next 12 hours in non-deprived control mice than in REMD mice ($P < 0.01$) (Supplementary Fig. 4 and Supplementary Table 4). Taken together, these results indicate that REM sleep increases the elimination of newly-formed dendritic spines both during development and after motor learning.

REM sleep-dependent pruning of new spines facilitates subsequent spine formation during development and after learning

What might be the function of REM sleep-dependent pruning of new spines? It has been hypothesized that by removing excessive synaptic plasticity, REM sleep may free up spaces in neural circuits to facilitate memory storage[35, 36] or reduce the accessibility of spurious memories[9, 21, 22]. To test whether REM sleep may serve such functions, we first identified

new spines formed between hours 0–8 in P21 mice and examined whether the elimination or survival of these earlier-formed new spines affects spine formation in subsequent hours (Fig. 2a). We found that new spines formed between hours 16–24 were more likely to grow within 2 μm to the sites of new spines formed between hours 0–8 if these earlier-formed spines were eliminated between hours 8–16 (Fig. 2b). The population of new spines within 2 μm to earlier-formed transient new spines was significantly larger in non-deprived control mice and NREM-d mice than in REMD mice (~21–31% versus 6%, $P < 0.05$) (Fig. 2b). In contrast, the percentage of new spines formed between hours 16–24 and located >2 μm away from the sites of new spines formed between hours 0–8 was comparable among different groups, regardless of the fate of these earlier-formed new spines (Fig. 2c and Supplementary Fig. 5a). These results indicate that by eliminating earlier-formed new spines, REM sleep facilitates the subsequent formation of new spines near the vacant sites during development.

Similar to the findings in development, we found that new spines induced by a motor learning task tended to grow near the vacant sites where earlier-formed new spines induced by a different motor task were eliminated, but rarely grew close to earlier-formed, persistent new spines (Fig.2d–i and Supplementary Fig. 5b,c). The percentage of new spines induced by the second motor task and located within 2 μm to transient new spines induced by the first motor task was significantly lower in REMD mice than in non-deprived control and NREM-d mice ($P < 0.05$) (Fig. 2e,h). Furthermore, consistent with the role of new spines in motor skill learning[25, 34, 37], we found that both new spine formation and performance improvement after the second motor training task were significantly lower in REMD mice than that in non-deprived control mice ($P < 0.05$) (Fig. 2j–m and Supplementary Table 5). These findings indicate that REM sleep-dependent elimination of new spines induced by one motor task facilitates subsequent new spine formation and behavioral improvement when a new motor task is learned, thus suggesting a role of REM sleep in removing excessive synaptic plasticity to balance the number of learning-induced new synapses over time.

REM sleep strengthens a fraction of new spines during development and after learning

Despite the fact that most of newly-formed spines are eliminated, a small fraction of new spines are maintained over extended periods of time[25, 31, 34]. While our findings above show that REM sleep increases the elimination of newly-formed spines, it is unknown whether REM sleep affects the persistent new spines. To address this question, we identified new spines formed during development or after motor learning, and examined the effect of REMD on the size of persistent new spines over the next 8–16 hours (Fig. 3a–c, e-g,i and Supplementary Fig. 6). We found that a significantly larger percentage of newly-formed spines exhibited an increase in size in non-deprived control mice and NREM-d mice as compared to REMD mice ($P < 0.05$) (Fig. 3c,g,i and Supplementary Fig. 6a,c-e). The increase in the average size of persistent new spines was significantly larger in non-deprived control mice and NREM-d mice when compared with REMD mice ($P < 0.05$) (Fig. 3c,g,i). In contrast, the average size of existing spines remained unchanged in non-deprived control, NREM-d or REMD mice (Fig. 3d,h,j). Thus, REM sleep not only promotes the elimination of new spines but also increases the size of persistent new spines in both development and learning. Because the size of dendritic spines strongly correlates with the strength of

synapses[38, 39], these findings also suggest that REM sleep increases the strength of newly-formed synapses that persist over time.

When performance improvement between the control and REMD mice was compared between hours 0–24, the improvement was significantly smaller in REMD mice than that in non-deprived and NREM-d control mice (Fig. 3k and Supplementary Table 6; Forward running: 42.8 ± 9.7% versus 73.6 ± 7.2% and 73.3 ± 3.6%; $P < 0.05$). Twelve hours after recovery from REMD, performance improvement in REMD mice continued to be significantly lower than that in control mice ($P < 0.05$) (Supplementary Fig. 7). Given the importance of new spines in motor learning[25, 34, 37], these findings also suggest that strengthening new synapses during REM sleep contributes to memory consolidation.

REM sleep facilitates long-term survival of new spines during development and after learning

Our data indicate that REM sleep eliminates some of newly-formed spines while strengthening others. Because strengthening new spines may facilitate their survival over time, it is unclear whether the long-term impact of REM sleep is to enhance or impede the survival of new spines. To address this question, we examined how REM sleep deprivation affects the survival of new spines formed between hours 0–8 over 4 days in postnatal day 21 mice (Fig. 4a). Notably, although more new spines formed between hours 0–8 survived between hours 8–16 in REMD mice than in control mice (Fig. 4a), nearly all of these surviving new spines were eliminated over the next 4 days in REMD mice (Fig. 4a,b). The percentage of new spines formed during hours 0–8 and maintained over the next 4 days was significantly lower in REMD mice than in control mice ($P < 0.05$) (Fig. 4a,b). We also found that in mice with REM sleep, new spines that persisted over 4 days were strengthened in size between hours 8–16 whereas the subset of new spines that were eliminated over 4 days were not (Fig. 4c,d), suggesting that REM sleep-dependent increase in spine size contributes to the long-term survival of new spines.

Similarly, although more new spines formed during a forward running task survived after 16 h in REMD mice (Fig. 4e), these survived new spines were largely eliminated in the next 12 hours after subsequent backward running (66.7 ± 4.8 % eliminated over hours 24–36) (Fig. 4f). In contrast, new spines formed during forward running in mice with subsequent REM sleep were largely maintained in the next 12 h with or without backward running (20.0 ± 6.5% or 28.8 ± 3.8% eliminated over hours 24–36) (Fig. 4e,f). As the result, fewer new spines induced by forward running persisted after backward running in REMD mice than in non-deprived control mice (18.8 ± 2.8 % versus 30.3 ± 2.7% over hours 8–36, $P < 0.05$) (Fig. 4e). Likewise, fewer new spines induced by backward running were maintained after forward running in REMD mice than in non-deprived and NREM-d control mice (Supplementary Fig. 8). We also found that new spines persisted over hours 24–36 were strengthened in size between hours 8–24 in non-deprived control mice, but not in REMD mice (Fig. 4g,h). Taken together, these findings underscore a crucial role of REM sleep in selectively strengthening a fraction of new spines to facilitate their maintenance over long term in development and learning.

Dendritic Ca^{2+} spikes in REM sleep strengthen and prune new spines

To gain insights into the mechanisms underlying the effect of REM sleep on new spine elimination and maintenance, we used two-photon Ca^{2+} imaging to examine the activities of apical dendrites of L5 pyramidal neurons expressing the genetically encoded calcium indicator GCaMP6s in the motor cortex of head-restrained mice that were trained to run on a custom-built treadmill under a two-photon microscope[31, 40] (Fig. 5a). Notably, we observed an upsurge of dendritic Ca^{2+} transients across long stretches (> 30μm) of apical tuft dendrites during REM sleep when compared to other brain states (Fig. 5b–e and Supplementary Fig. 9; see also Supplementary Video 1). Over a period of 1 minute, the number and peak amplitude of dendritic Ca^{2+} transients during REM sleep were comparable to those during active running, but significantly higher than during NREM sleep and quiet awake state (Fig. 5c,d and Supplementary Fig. 9b). The duration of Ca^{2+} transients during REM sleep was significantly longer than during other brain states (Fig. 5e). These dendritic Ca^{2+} transients during REM sleep resemble NMDA receptor activation-dependent dendritic Ca^{2+} spikes described previously in various brain regions including the motor cortex[40–45]. Consistent with this notion, the number, duration and amplitude of dendritic Ca^{2+} transients during REM sleep were significantly reduced upon local application of NMDA receptor antagonist MK801 to the cortical surface ($P < 0.001$) (Fig. 5c–e and Supplementary Fig. 9b).

Many lines of evidence indicate that dendritic Ca^{2+} spikes are important for the induction of synaptic potentiation and depotentiation[40, 46–48]. To test whether dendritic Ca^{2+} spikes during REM sleep may affect the strength and survival of new spines on apical tufted dendrites of L5 pyramidal neurons, we first blocked dendritic Ca^{2+} spike generation by injecting MK801 locally into the layer 1 of the motor cortex at the beginning of REM sleep (Supplementary Fig. 10). We found that pulsed MK801 injection (3 pulses) over ~2s effectively blocked the generation of Ca^{2+} spikes in the motor cortex within the next 2–3 minutes (Fig. 5f). Over a period of 8 hours, > 90% of dendritic Ca^{2+} spikes during REM sleep were blocked when MK801 was injected at the beginning of each REM sleep episode, but not outside the period of REM sleep (Fig. 5g). Similar to REMD, MK801 blockade of dendritic Ca^{2+} spikes during REM sleep reduced new spine elimination and potentiation over 8 hours (Fig. 5h,i). In contrast, MK801 injection during NREM sleep had no significant effect on the elimination and size increase of new spines (Fig. 5h,i). These results suggest that dendritic Ca^{2+} spikes generated during REM sleep are involved in selective pruning and strengthening of new spines.

When dendritic Ca^{2+} spikes were compared among different brain states, we observed a substantial overlap between Ca^{2+} spikes triggered by motor training and during REM sleep (Supplementary Fig. 11). Eight hours after the initial motor training, > 60% of dendrites exhibiting Ca^{2+} spikes during 1-min REM sleep also showed Ca^{2+} spikes during 4-min retraining with comparable amplitudes (Fig. 6a and Supplementary Fig. 12). We reasoned that if dendritic Ca^{2+} spikes were critical for new spine plasticity, retraining-induced Ca^{2+} spikes may have similar effects as those occurring during REM sleep. To test this, we identified new spines formed within 8 hours after the initial training and examined the effect of 40-min retraining in REMD mice (Fig. 6b). Similar to the effect of REM sleep, 40-min retraining increased both the elimination rate and size of new spines over 8 hours in REMD

mice (Fig. 6b,c). Furthermore, when retraining-induced dendritic Ca^{2+} spikes were blocked by repeated injections of MK801 in layer 1 (Fig. 6d and Supplementary Fig. 10), the effects of retraining on new spine elimination and potentiation in REMD mice were also blocked (Fig. 6b,c). In contrast, repeated injection of MK801 immediately after retraining did not block the effect of retraining on new spine plasticity (Fig.6b,c). In addition to MK801, we also locally applied KN-62, a CaMKII inhibitor, to block the downstream targets of calcium spikes during retraining. Unlike MK801, KN-62 did not alter the number, duration and peak amplitude of Ca^{2+} spikes, but prevented the effect of retraining on new spine elimination and potentiation (Fig.6e–j). Taken together, these results suggest that dendritic Ca^{2+} spikes, generated during REM sleep or retraining, selectively prune and strengthen new spines via a process requiring NMDA receptor activation and downstream targets of calcium signaling.

DISCUSSION

By investigating how REM sleep affects postsynaptic dendritic spines of L5 pyramidal neurons, we have revealed multifaceted functions of REM sleep in development and learning via dendritic calcium spike-dependent mechanisms. In particular, our results show that REM sleep has a fundamental role in establishing stably-connected synaptic circuits during development, as nearly all newly-formed synapses would not be maintained without REM sleep. Furthermore, REM sleep also contributes to memory consolidation by selectively strengthening and maintaining a fraction of learning-induced new spines that are important for the improvement of motor skills. In addition, our studies suggest a role of REM sleep in "unlearning" as REM sleep prunes recently-formed new spines to facilitate subsequent formation of new spines and performance improvement after new learning. Together, these findings underscore important functions of REM sleep in brain development, learning and memory consolidation by selectively pruning and maintaining new synapses in neuronal circuits throughout life.

A cardinal feature of the nervous system development and plasticity is that a large number of synapses are formed daily, but only a small fraction of new synapses are strengthened and maintained[23–26, 31]. This selective maintenance of synapses is essential for progressively building up functional neuronal circuits throughout development and for continuously incorporating new information in adult life. It is well established that sensory and learning experiences are important for promoting synapse formation and elimination both during development and in adulthood[23–25, 32]. The mechanisms underlying the highly selective maintenance of new synapses, however, remain obscure. Our findings indicate that REM sleep not only prunes but also strengthens and maintains a fraction of new synapses during development and after learning. This selective pruning and strengthening of new spines are dependent on endogenously generated calcium spikes on apical dendrites of L5 pyramidal neurons during REM sleep. Because dendritic calcium spikes generated during REM sleep differ from those occurring during awake and non-REM sleep states (Fig. 5 and Supplementary Fig. 9), our findings strongly suggest that endogenous neuronal activity during REM sleep is fundamentally and perhaps uniquely important for the selective maintenance of new synapses during development and in adult life.

The abundance of REM sleep and the detrimental effects of REMD during development have led to many hypotheses on the functions of REM sleep in brain maturation[12]. It has been proposed that REM sleep substitutes for 'wakefulness' to stimulate neuronal growth and plasticity during early development when wakefulness is limited[1]. It has also been suggested that REM sleep is important for the "unlearning" process when excessive synapses formed during development need to be removed[9, 21, 22]. In addition, REM sleep is thought to be the time for genetic programming of neuronal circuits to preserve individual differences[2]. Our studies of L5 pyramidal neurons show that REM sleep is important for pruning and maintaining new synapses formed during both development and learning. Thus, REM sleep is not simply a substitute for wakefulness to stimulate plasticity, for pruning synapses during development or for persevering genetic identities of neuronal circuits. Instead, REM sleep is important for selectively incorporating new synapses into the existing circuits. It could be viewed as a "selection committee" for building and maintaining the synaptic network (or an organization). Each day, REM sleep selectively strengthens a small fraction of newly-formed synapses (new candidates) to be stably integrated into the existing network. At the same time, REM sleep also removes many of newly-formed synapses so that the synaptic network is not overloaded and newer synapses can continue to be integrated and maintained throughout the period of development.

In addition to its role in brain development, our data also show that REM sleep contributes to memory consolidation by strengthening and maintaining a subset of new spines induced by motor learning. Although REM sleep immediately after learning prunes newly-formed spines, the long-term survival rate of these new spines is higher in mice with REM sleep than without when animals are trained with multiple motor tasks. Thus, over long-term, REM sleep helps to strengthen and incorporate learning-induced new connections into existing neuronal circuits. As the degree of persistent new spines correlates with performance improvement[25, 31, 34], these findings support the function of REM sleep in the consolidation of motor memories [6, 19, 20]. Furthermore, as NREM sleep, not REM sleep, promotes the initial formation of new spines[31], the findings of REM sleep-dependent strengthening of new spines also provide insights into why REM sleep typically occurs after NREM sleep.

Given that ~5–10% of new spines are formed daily during development and after learning, pruning some of newly-formed connections seems necessary for neuronal circuits to keep integrating new connections without removing all of previously established connections. Indeed, our data show that REM sleep-dependent pruning of recently-formed new spines facilitates new spine formation during subsequent learning. Our findings of REM sleep-dependent removal of new spines resonate with the "unlearning" theory of REM sleep[9, 21, 22]. According to this theory, by removing excessive synaptic plasticity, REM sleep may reduce the accessibility of spurious memories and benefit the recall of stored information[9, 21, 22]. However, because the functional contribution of excessive new spines in REMD mice is not known, whether these new spines would cause spurious memories and interfere with memory recall remains to be examined. Nevertheless, because REM sleep prunes predominately recently-formed new spines to facilitate subsequent growth of new spines in response to new learning tasks, our studies suggest that REM sleep does not simply prevent spurious memories or free up spaces in the brain to facilitate memory

storage[9, 21, 35, 36]. Instead, REM sleep prunes synaptic circuits to balance the number of new synapses generated over time in development and learning.

Our findings also provide novel mechanistic insights into the function of REM sleep in pruning and strengthening spine plasticity. By performing *in vivo* dendritic Ca^{2+} imaging under different brain states, we found a substantial rise of dendritic Ca^{2+} spikes on apical tuft branches of L5 pyramidal neurons during the stage of REM sleep. By blocking and mimicking the generation of dendritic Ca^{2+} spikes during REM sleep, our findings indicate that these Ca^{2+} spikes are critical for selective strengthening and pruning new spines induced by learning. A variety of experimental evidence have shown that dendritic Ca^{2+} spikes regulate synapse potentiation and depotentiation, depending on relative time windows between synaptic activity and spike generation[45–49]. Recent studies in the mouse motor cortex have shown that spines are potentiated if they are active at the time of dendritic Ca^{2+} spike generation and depotentiated if they are active prior to Ca^{2+} spikes[40]. It has been shown that extracellular signal-regulated kinase (ERK) phosphorylation and the expression of immediate-early genes such as *zif-268* are elevated during REM sleep[27, 30]. Thus, a rise in intracellular calcium concentrations as the result of Ca^{2+} spike generation could influence the ERK signaling cascade and immediate-early gene expression, which in turn regulate synaptic plasticity and neuronal excitability during and after REM sleep[50] gene expression that are important for synaptic plasticity during REM sleep. Future studies are required to examine the involvement of Ca^{2+} spike-timing dependent plasticity and signaling pathways in selectively strengthening and pruning new spines in order to better understand the functions of REM sleep in development and learning.

Lastly, it is important to note that our studies focus on how REM sleep affects newly-formed spines on apical dendrites of L5 pyramidal neurons in the motor cortex. Future studies are needed to investigate whether REM sleep has similar or different roles in regulating synapse development and plasticity of other types of neurons in different cortical layers and brain regions. Furthermore, recent studies have shown that different motor tasks induce new spine formation on different dendritic branches of the same L5 pyramidal neurons[31]. It is unclear whether new spines formed in response to one task on an individual branch need to be pruned by REM sleep in order for the same branch to grow new spines in response to new tasks. Further studies of REM functions at the level of individual dendritic branches would help to better understand the fundamental roles of REM sleep in the development and maintenance of synaptic circuits.

Online Methods

Experimental animals

Transgenic mice expressing Yellow Fluorescent Protein (YFP) in L5 pyramidal cells and *Thy1*-Cre transgenic mice (FVB/N-Tg(Thy1-Cre)1Vln/J) were purchased from the Jackson Laboratory (YFP-H line) and housed either in the animal facility with maximum of 5 mice per cage at New York University School of Medicine or in the Peking University Shenzhen Graduate School (PKUSZ). C57BL6 mice were purchased from the Jackson Laboratory and from Guangdong medical laboratory animal center. Mice were maintained at 22 ± 2 °C with a 12-hour light: dark cycle (lights on 8am, lights off 8pm). Food and water were available ad

libitum. Three-week-old male and female mice were used for developmental studies and four to five-week-old male mice were used in studies of motor learning. All experimental protocols were conducted in accordance with the institutional guidelines.

Rotarod and treadmill training

The Rotarod training procedure was performed as described before[25, 31, 34]. A Rotarod system with six individual chambers (Chengdu TME Technology Co., Ltd, China) was used in this study. Animals were force to run forward on the motorized rod (30 mm in diameter) in the chamber. The rotation speed of the rod gradually increased from 0 to 100 r.p.m. over the course of 100 s. To induce new motor learning experience, animals were also forced to run backward on the accelerated rod (speed increased gradually from 0 to 50 r.p.m. over 100 s). The time latency and rotation speed were recorded when the animal was unable to keep up with the increasing speed and fell. The experimental and control groups were subjected to the identical rotarod training session. Each session consisted of either 20 or 40 trials and lasted ~30–60 minutes. Rotarod training was used in Figs. 1– 4.

A custom built free-floating treadmill (101 cm × 58 cm × 44 cm dimensions) was used for motor training under a two-photon microscope as described in Figs. 5–6. This free-floating treadmill allows head-fixed mice to move their forelimbs freely to perform motor running tasks (forward or backward)[31, 40]. To minimize motion artifact during imaging, the treadmill was constructed so that all the moving parts (motor, belt, and drive shaft) would not be in contact with either the microscope stage or the supporting air-table. Animals were positioned on a custom-made head holder device that would allow the micro-metal bars to be mounted. During motor training, the treadmill motor was driven by a DC power supply. At the onset of a trial, the motor was turned on and the belt speed gradually increased from 0 cm/s to 8 cm/s within ~3 sec, and the speed of 8 cm/s was maintained for the rest of the trial. Each mouse was trained for 6–7 trials (5min running and 1min resting for each trial). Similar to rotarod running, 40 min treadmill running also induced new spine formation over 8 hours (Supplementary Fig. 13).

REM sleep deprivation procedure

REM sleep deprivation was achieved through gentle handling over a period of ~8–16 hours after the imaging session to identify newly-formed spines. Specifically, mice were gently touched with a cotton applicator for 1–2 seconds as soon as they displayed signs of REM sleep based on EEG/EMG recordings. The EEG trace was scored by an observer in real time. The average latency from REM onset to gentle handling is ~8 seconds. On average, mice were touched ~6.9 times per hour at the age of P21 and ~4 times per hour at the age of P30 during the period of REM sleep deprivation. The same number of touches was then used during NREM sleep to control for the disturbance cause by gentle handling. The animals were not previously habituated to this gentle handling protocol. Food and water were available ad libitum throughout the entire deprivation process.

To determine the impact of the gentle handling procedure on stress response, plasma corticosterone was measured under various conditions (Supplementary Fig. 14). We found that physical restraint for in vivo imaging induced a high level of corticosterone. The gentle

handling procedure did not increase the level of plasma corticosterone in REMD and NREM-d mice.

Surgery for imaging and EEG/EMG recording

Twenty-four hours before imaging, surgery was performed to attach a head holder and to create a thinned-skull cranial window for *in vivo* imaging and to implant four electrodes for EEG/EMG recording. Specifically, mice were deeply anesthetized with an intraperitoneal injection of pentobarbital sodium (80 mg/kg). The mouse head was shaved and the skull surface was exposed with a midline scalp incision. The periosteum tissue over the skull surface was removed without damaging the temporal and occipital muscles. The forelimb region of the motor cortex to be imaged was identified based on stereotactic coordinates (1.3 mm anterior to bregma, 1.2 mm lateral from midline) and marked with a fine marker. A thin layer of cyanoacrylate-based glue was first applied to the top of entire skull surface, and a head holder composed of two parallel micro-metal bars was then mounted on top of the skull with dental acrylic cement such that the marked skull region was exposed without being covered by dental acrylic cement between the two bars. The head holder helped restrain the animal's head and reduce motion-induced artifact during imaging.

For EEG recording, the EEG electrodes were inserted above the visual cortex (3 mm lateral to midline, 3 mm posterior to bregma) in the hemisphere contralateral to the imaged forelimb region. Before the implantation, a small area of skull (~1 mm in diameter) was thinned with a high speed drill and carefully removed with forceps. Two electrodes for EEG recording were made by soldering one end of an epoxy-coated silver wire (0.003 inch in diameter, A-M Systems) to a connector pin. The epoxy of the other end of the electrodes was removed and the exposed silver wires were carefully inserted into the superficial layer of the cortex (~100 μm below the pial surface) with the tips separated by ~800 μm from each other. The electrodes were fixed first by cyanoacrylate-based glue and subsequently by dental cement. Two electrodes for EMG recording were made by polyurethane enameled copper wires (0.13 mm in diameter) and placed on the nuchal muscle and stabilized together with EEG electrodes by dental cement.

After the dental cement was completely dry, the area for imaging was covered with a drop of 1% agar (<40°C) and sealed with cyanoacrylate-based glue, and the animals were returned to their own cages to recover.

MK801 application

For local application of MK801 (200 μM in artificial cerebrospinal fluid (ACSF), M107, Sigma Aldrich), a glass microelectrode with 20-μm outer diameter was inserted through a bone flap into the superficial layer of the cortex (20–30 μm below the pial surface) with an angle of 45° towards and ~100 μm away from the imaging area. The bone flap (~50 μm in diameter) for drug delivery was made adjacent to a thinned skull window. The area for MK801 injection and imaging was covered with a 1% agar, which could be easily removed at the time of imaging. Immediately upon the detection of REM sleep based on EEG and EMG recordings, MK801 was injected via pressure injections with Picospritzer II (40 p.s.i., 50 ms per pulse, 1 Hz, either 3 or 10 pulses; General Valve Corporation) into the motor

cortex of head-fixed mice. At the same time, Ca^{2+} imaging of dendritic activities was performed in layer 1 of the forelimb region of the motor cortex. EEG/EMG recording indicated that neither two-photon imaging, nor the injection of MK801 affect normal sleep (Supplementary Fig. 15).

To map the extent of local MK801 diffusion, Congo red (0.5%) was injected to the cortex and imaged using two-photon microscope with a 0.3 N.A 10X objective. The extent of Congo red spread was estimated by the line at which the fluorescence was less than 20% of its peak level. On the basis of this definition, we determined the fluorescence signals spread to a region 426 ± 11 μm after 3 pulses and 612 ± 39 μm after 10 pulses in L1 (Supplementary Fig. 10). Because small molecules diffuse rapidly in the cortex, we estimated that the drug concentration was reduced >10 times in the imaged cortical region, such that the final effective concentration for MK801 would be < 20 μM. We found that 3-pulse injection of MK801 blocked dendritic calcium spikes for 2–3 mins (Fig. 5f) whereas 10-pulse injection of MK801 blocked dendritic calcium spikes for ~40 mins (Fig. 6d). As shown in Supplementary Fig. 10, the large difference in blocking dendritic calcium spikes between 3-pulse and 10 pulse MK801 injection was likely due to that the concentration of MK801was lower at the electrode tip (~20 μm) than at the shaft. As a result, the amount of MK801 released from the initial pulse injection was less than that from the subsequent pulse injections.

KN-62 application

CaMKII inhibitor KN-62 (I2142, Sigma Aldrich) was dissolved in DMSO and diluted with ACSF. ~0.1 ml KN-62 (100μM in 0.5% DMSO) was applied directly to the surface of the primary motor cortex through a small bone flap lateral to the imaging window for ~40 minutes during or after treadmill running trials.

Imaging dendritic spine plasticity in awake, head-restrained mice

Before imaging, mice were allowed to recover from the surgery for one day, and habituated for a few times (10 min each) in the imaging apparatus to minimize potential stress effects due to head restraining and awake imaging. To image dendritic spines in awake, head restrained *Thy1*-YFP mice, the head holder was screwed to two metal cubes attached to a solid metal base. The glue and agar covering the marked area for imaging was removed and the skull was washed by ACSF for several times. A high-speed micro-drill and microsurgical blade were used to thin a circular area over the marked region to a thickness of approximately 20 micrometers. The thinning procedure generally took less than 2 minutes. The skull was immersed in ACSF and the head-restrained animal was then placed on the stage of a two-photon microscope. Image stacks of dendritic segments projecting to superficial cortical layers were obtained using an Olympus two-photon microscope (FV1000MPE) with the laser tuned to 920 nm and with a 1.1 N.A. 60X objective or a 1.05 N.A. 25X objective immersed in artificial cerebrospinal fluid. A 3X digital zoom was used to yield high-magnification images (66.7 μm × 66.7 μm, 512 pixel × 512 pixel for 60X objective; 169 μm × 169 μm, 1024 pixel × 1024 pixel for 25X objective) suitable for quantification of dendritic spines. For multiple imaging, the above procedure was repeated and the localization of the same region was facilitated by low-magnification images stacks at

1X digital zoom and with reference to vascular landmarks under the thinned skull area. The animal was head restrained during image acquisition, which lasted ~30 minutes. The first imaging session of P21 mice and P30 mice started at midnight and 8am, respectively. Mice were immediately returned to their home cage and stayed there until the next imaging sessions.

Data analysis of spine structural plasticity

The procedure for quantifying spine dynamics has been described in earlier studies[25, 31, 34, 51]. In brief, image stacks were analyzed using NIH ImageJ software. For each dendritic segment analyzed, filopodia were identified as long, thin protrusions with ratio of head diameter to neck diameter <1.2:1 and ratio of length to neck diameter > 3:1. The remaining protrusions were classified as spines. Spines or filopodia were considered the same between views if their positions remained the same distance from relative adjacent landmarks. Spines were considered different if they were more than 0.7 μm away from their expected positions based on the prior view. Approximately 150 spines or more were analyzed from each animal to calculate spine formation and elimination. The rate of spine formation or elimination was calculated as the number of spines added or eliminated divided by the number of pre-existing spines.

Spine head size was measured according to previous studies[26, 52–55]. After background subtraction, the fluorescence intensity of the spine (the intensity of all pixels covering the spine in the best focal plane) was divided by the fluorescence intensity of the adjacent dendritic shaft. Specifically, fluorescence intensity of a spine was measured as follows, where Area is the number of pixels in an oval surrounding the head of the spine and mean optical density (MeanOD) is the mean brightness of pixels in that area:

The ratio of spine head size to adjacent dendritic shaft size = (Area (of spine) × MeanOD (of spine) - Area (of spine) × MeanOD (of background))/ (Area (of spine) × MeanOD (of dendrite) - Area (of spine) × MeanOD (of background)). The MeanOD of both the background and the dendritic shaft was calculated from measurements taken next to each spine.

Spine head size in new spines and existing spines (randomly chosen 10–20μm away from new spines) were measured on the same dendritic segment. Spine size change was calculated by comparing spine size measurement between imaging sessions.

EEG/EMG recording and analysis

Between imaging sessions, EEG/EMG was recorded using BL-420F Biological Data acquisition & Analysis system (Chengdu TME Technology Co., Ltd, China) with band pass setting of 0.1–100 Hz. EEG/EMG data were visually scored for the states of wake and sleep. Wake state was identified by lower amplitude and higher frequency (> 10 Hz) of EEG activity, and medium to high muscle activity. REM sleep was identified by lower amplitude and higher frequency (> 10 Hz) of EEG activity, and low muscle activity. NREM sleep was identified by higher amplitude and lower frequency (< 10 Hz) of EEG activity, and low muscle activity.

All REM sleep periods 5 s and longer were analyzed. REM sleep occupied ~9–11% and ~6–7% of the total time (including sleep and wake) in non-deprived mice or NREM-d mice at the age of P21 and P30, respectively (Fig. 1c,d,i and Supplementary Fig. 2–Supplementary Fig. 4).

Two-photon calcium imaging of apical dendrites of layer V pyramidal neurons in mice expressing GCaMP6s

Genetically-encoded calcium indicator GCaMP6s was used for Ca^{2+} imaging of somata and apical dendrites of L5 pyramidal neurons in the primary motor cortex. GCaMP6s were expressed with recombinant adeno-associated virus under the Synapsin promoter (AAV, serotype 2/1; $>2\times10^{13}$ (GC/ml) titer; produced by the University of Pennsylvania Gene Therapy Program Vector Core). 0.1–0.2 μl of AAV viruses was diluted 2X in ACSF and slowly injected (Picospritzer III; 30 p.s.i., 20 ms, 0.2 Hz) over 10–15 minutes into L5 (depth of 500–700 μm) of primary motor cortex using a glass microelectrode around the coordinates of 1.5 mm anterior and 1.5 mm lateral from bregma[56]. In addition to *Synapsin*-driven GCaMP6s, *Thy1*-Cre transgenic mice (FVB/N-Tg(Thy1-Cre)1Vln/J) were used in combination with Cre-dependent AAV-*CAG*-Flex-GCaMP6s. In these mice, Cre is exclusively expressed in pyramidal neurons. Mice were 3-week-old at the time of injection. Two to three weeks after the virus injection, mice were prepared for awake animal imaging.

For imaging calcium activities of dendrites, mice were head restrained in the imaging apparatus on top of a custom-built free-floating treadmill. After habituation on the treadmill for 10 min, calcium activities were recorded for 3 mins when mice were under quiet awake state (resting). Mice were then subjected to forward running for 40 mins (1-min resting after every 5-min running). Calcium imaging was performed 5 times (1~2-min recording each time) during the entire training period. Subsequently, mice were allowed to sleep on the two-photon microscope stage. Calcium signals were recording 5 times during NREM sleep and 3–5 times (1–2 min every time) during REM sleep based on EEG/EMG recording. Dendritic calcium activities were recorded through a thinned-skull window.

Motor cortex containing GCaMP-positive apical dendritic branches (from L5 neurons) at the depth of 0–50 μm below the pial surface were imaged for detecting the activities of tuft dendrites. To confirm that layer 2/3 were not labeled by GCaMP6, a glass window was placed over the region of interest following the completion of imaging experiments, and imaging of layer 2/3 was performed to ensure the absence of GCaMP6 expression in layer 2/3 neurons.

Two-photon imaging was performed with an Olympus Fluoview 1000 two-photon system (tuned to 920 nm) equipped with a Ti:Sapphire laser (MaiTai DeepSee, Spectra Physics). The average laser power on the tissue sample was ~20–30mW for imaging the L1 of the cortex. All experiments were performed using a 40X objective (N.A. 0.8) for imaging dendrites with 3X digital zoom. All images were acquired at the frame rate ~ 1 Hz (2-μs pixel dwell time). Typical imaging window for Ca^{2+} imaging of tuft dendrites was 105 μm by 105 μm. Image acquisition was performed using FV10-ASW v.3.0 software.

Calcium Imaging Data Analysis

Changes of dendritic calcium activities during quiet awake state, running, NREM and REM sleep, as indicated by GCaMP6 fluorescence changes, were analyzed post hoc using ImageJ software (NIH). Regions-of-interests (ROIs) corresponding to visually identifiable apical tuft dendrites were selected for quantification (Supplementary Fig. 9a). On average, we quantified calcium activity on dendritic segments ~30 μm in length (ROI: ~145 pixels in length and ~7 pixels in width). During running trials, images with the lateral movement less than 1 μm were used for analyses. Vertical movements were infrequent and minimized due to flexible belt design, two micro-metal bars attached to the animal's skull by dental acrylic, and a custom-built body support to minimize spinal cord movements. All imaging stacks were registered using ImageJ plugin StackReg.

As described in previous studies, dendritic Ca^{2+} spikes were defined as the events when changes of fluorescence ($\Delta F/F_0$) observed in both dendritic spines and shaft (average length > 30 μm) were > 50% for GCaMP6s during the imaging sessions. F_0 is the fluorescence intensity in dendritic segments after background subtraction. The threshold for detecting dendritic spikes was more than 3 times the standard deviation of baseline fluorescence noise for GCaMP6s (16.7%). The majority of Ca^{2+} spikes were found to have a fluorescence increase of 100 to >1,000% for GCaMP6s. The full width at half-maximum (FWHM) of an individual spike was measured as the spike duration. The number, duration and peak amplitude of calcium spikes were quantified from individual calcium spikes that did not overlap with other spikes under various brain states. The degree of overlap of dendritic Ca^{2+} spikes between running and REM sleep = number (of the same active dendrites during both running and REM) × 2/(number (of active dendrites during running) + number (of active dendrites during REM)).

Statistics—All data are presented as mean ± SEM. Tests for differences between groups were performed using Wilcoxon–Mann–Whitney nonparametric tests (two-tailed). Significant levels were set at $P \leq 0.05$. All statistical analyses were performed using the SPSS software. Sample sizes were chosen to ensure adequate power with the statistical tests while minimizing the number of animals used in compliance with ethical guidelines. No samples or animals that were successfully imaged or measured were excluded from the analysis. Experiments did not involve randomization or blinding because no dendritic spine plasticity, dendritic calcium spikes, or behavioral performance of animal groups were predefined. The variance was similar between the groups that were statistically compared. Statistical details for the experiments described in Figs 1–6 are detailed below, including the sample size (number of mice and number of spines).

The rate of spine formation over 8 hours in Fig. 1b was analyzed at the age of P21 and P30. P21: 7.9 ± 0.5% (n = 7 mice) versus P30: 3.9 ± 0.3% (n = 4 mice) ($P = 0.008$, Wilcoxon–Mann–Whitney test). The rate of spine formation after motor training in P30 mice was 6.3 ± 0.3% (n = 7 mice), significantly higher than that in P30 mice without training ($P = 0.008$, Wilcoxon–Mann–Whitney test).

The amount of REM sleep over 8 hours in Fig. 1c was 11.2 ± 1.1% (n = 4 mice) and 6.8 ± 0.6% (n = 6 mice) in P21 and P30 mice, respectively ($P = 0.03$, Wilcoxon–Mann–Whitney test).

For evaluating the effects of REM sleep deprivation and NREM sleep disturbance in P21 mice, the amount of REM sleep was calculated as the percentage of total time (including awake, NREM sleep and REM sleep) in Fig. 1d. ND: 11.2 ± 1.1% (n = 4 mice), NREM-d: 9.2 ± 0.7% (n = 4 mice) versus REMD: 1.3 ± 0.2% (n = 5 mice) ($P = 0.014$ for ND or NREM-d versus REMD, Wilcoxon–Mann–Whitney test).

After identification of new spines over 0–8 h in P21 mice, new spines were imaged over 8–16 h and their elimination rate was analyzed in Fig. 1f. ND: 65.1 ± 1.6% (n = 7 mice), NREM-d: 63.1 ± 2.1% (n = 5 mice) versus REMD: 42.3 ± 4.4% (n = 7 mice). $P = 0.002$ for ND versus REMD and $P = 0.006$ for NREM-d versus REMD, Wilcoxon–Mann–Whitney test.

In Fig. 1g, the elimination rate of existing spines was analyzed over 8–16 h in P21 mice. ND: 3.9 ± 0.8% (n = 7 mice), NREM-d: 3.0 ± 0.6% (n = 5 mice) versus REMD: 3.1 ± 0.3% (n = 7 mice). $P = 0.701$ for ND versus REMD and $P = 0.465$ for NREM-d versus REMD, Wilcoxon–Mann–Whitney test.

In Fig. 1i, the amount of REM sleep was calculated over 8–16 h and 16–24 h in P30 mice. ND: 6.4 ± 0.5% (n = 7 mice), NREM-d: 6.3 ± 0.4% (n = 5 mice) versus REMD: 1.0 ± 0.2% (n = 6 mice) over 8–16 h. $P = 0.003$ for ND versus REMD and $P = 0.006$ for NREM-d versus REMD, Wilcoxon–Mann–Whitney test. ND: 7.3 ± 0.6% (n = 7 mice), NREM-d: 6.9 ± 0.8% (n = 5 mice) versus REMD: 1.1 ± 0.2% (n = 6 mice) over 16–24 h. $P = 0.003$ for ND versus REMD and $P = 0.006$ for NREM-d versus REMD, Wilcoxon–Mann–Whitney test.

After identification of forward running-induced new spines over 0–8 h in P30 mice, new spines were imaged over 8–24 h and their elimination rate was analyzed in Fig. 1k. ND: 51.2 ± 3.5% (n = 6 mice), NREM-d: 55.3 ± 2.7% (n = 6 mice) versus REMD: 33.8 ± 2.5% (n = 6 mice) over 8–16 h. $P = 0.004$ for ND or NREM-d versus REMD, Wilcoxon–Mann–Whitney test. ND: 60.3 ± 2.2% (n = 6 mice), NREM-d: 62.7 ± 2.0% (n = 6 mice) versus REMD: 49.4 ± 2.2% (n = 6 mice) over 8–24 h. $P = 0.008$ for ND or NREM-d versus REMD, Wilcoxon–Mann–Whitney test.

In Fig. 1l, the elimination rate of existing spines was analyzed over 8–24 h in P30 mice. ND: 3.2 ± 0.4% (n = 6 mice), NREM-d: 2.0 ± 0.3% (n = 6 mice) versus REMD: 3.2 ± 0.7% (n = 6 mice) over 8–16 h. $P = 0.748$ for ND versus REMD and $P = 0.261$ for NREM-d versus REMD, Wilcoxon–Mann–Whitney test. ND: 5.0 ± 0.4% (n = 6 mice), NREM-d: 3.6 ± 0.2% (n = 6 mice) versus REMD: 5.6 ± 0.9% (n = 6 mice) over 8–24 h. $P = 0.749$ for ND versus REMD and $P = 0.078$ for NREM-d versus REMD, Wilcoxon–Mann–Whitney test.

After identification of backward running-induced new spines over 0–8 h in P30 mice, new spines were imaged over 8–16 h and their elimination rate was analyzed in Fig. 1m. ND: 61.8 ± 2.3% (n = 6 mice), NREM-d: 60.6 ± 1.9% (n = 6 mice) versus REMD: 47.1 ± 2.7% (n = 6 mice). $P = 0.006$ for ND or NREM-d versus REMD, Wilcoxon–Mann–Whitney test.

In Fig. 1n, the elimination rate of existing spines was analyzed over 8–16 h in P30 mice. ND: 1.9 ± 0.3% (n = 6 mice), NREM-d: 1.8 ± 0.2% (n = 6 mice) versus REMD: 2.0 ± 0.4% (n = 6 mice). P = 1 for ND versus REMD and P = 0.63 for NREM-d versus REMD, Wilcoxon–Mann–Whitney test.

To test if the elimination of new spines induced by REM sleep affects subsequent new spine formation, new spines formed over 16–24 h and located with 2 μm to persistent and transient new spines form over 0–8 h were examined in P21 mice in Fig. 2b. ND: 7.6 ± 3.7% (n = 7 mice), NREM-d: 5.2 ± 2.3% (n = 5 mice) versus REMD: 4.4 ± 2.9% (n = 7 mice) for new spines formed over 16–24 h and located with 2 μm to persistent new spines formed 0–8 h (P = 0.396 for ND versus REMD and P = 0.651 for NREM-d versus REMD, Wilcoxon–Mann–Whitney test). ND: 31.2 ± 7.2% (n = 7 mice), NREM-d: 21.1 ± 4.3% (n = 5 mice) versus REMD: 5.6 ± 2.9% (n = 7 mice) for new spines formed over 16–24 h and located with 2 μm to transient new spines formed 0–8 h (P = 0.01 for ND or NREM-d versus REMD, Wilcoxon–Mann–Whitney test). P = 0.02 for ND (near transient new spines) versus ND (near persistent new spines), P = 0.01 for NREM-d (near transient new spines) versus NREM-d (near persistent new spines), P = 0.71 for REMD (near transient new spines) versus REMD (near persistent new spines), Wilcoxon–Mann–Whitney test.

In Fig. 2c, new spines formed over 16–24 h and located 2–6 μm to persistent and transient new spines form over 0–8 h were examined in P21 mice. ND: 13.7 ± 5.7% (n = 7 mice), NREM-d: 13.7 ± 5.7% (n = 5 mice) versus REMD: 9.3 ± 3.6% (n = 7 mice) for new spines formed over 16–24 h and located 2–6 μm to persistent new spines formed 0–8 h (P = 0.476 for ND versus REMD and P = 0.461 for NREM-d versus REMD, Wilcoxon–Mann–Whitney test). ND: 10.5 ± 3.8% (n = 7 mice), NREM-d: 11.9 ± 4.5% (n = 5 mice) versus REMD: 13.1 ± 3.7% (n = 7 mice) for new spines formed over 16–24 h and located 2–6 μm to transient new spines formed 0–8 h (P = 0.7 for ND versus REMD and P = 1.0 for NREM-d versus REMD, Wilcoxon–Mann–Whitney test).

In Fig. 2e, new spines formed over 16–24 h after FW training and located with 2 μm to persistent and transient new spines form over 0–8 h after BW training were examined in P30 mice. ND: 5.0 ± 2.3% (n = 6 mice), NREM-d: 4.5 ± 3.1% (n = 6 mice) versus REMD: 3.4 ± 2.7% (n = 6 mice) for new spines formed over 16–24 h and located with 2 μm to persistent new spines formed 0–8 h (P = 0.592 for ND versus REMD and P = 0.849 for NREM-d versus REMD, Wilcoxon–Mann–Whitney test). ND: 30.9 ± 4.3% (n = 6 mice), NREM-d: 35.7 ± 4.5% (n = 6 mice) versus REMD: 13.9 ± 3.8% (n = 6 mice) for new spines formed over 16–24 h and located with 2 μm to transient new spines formed 0–8 h (P = 0.01 for ND or NREM-d versus REMD, Wilcoxon–Mann–Whitney test). P = 0.004 for ND (near transient new spines) versus ND (near persistent new spines), P = 0.003 for NREM-d (near transient new spines) versus NREM-d (near persistent new spines), P = 0.056 for REMD (near transient new spines) versus REMD (near persistent new spines), Wilcoxon–Mann–Whitney test.

In Fig. 2f, new spines formed over 16–24 h after FW training and located 2–6 μm to persistent and transient new spines form over 0–8 h after BW training were examined in P30 mice. ND: 8.7 ± 5.7% (n = 6 mice), NREM-d: 12.0 ± 4.4% (n = 6 mice) versus REMD: 11.2

± 4.2% (n = 6 mice) for new spines formed over 16–24 h and located 2–6 μm to persistent new spines formed 0–8 h (P = 0.507 for ND versus REMD and P = 0.871 for NREM-d versus REMD, Wilcoxon–Mann–Whitney test). ND: 9.1 ± 3.0% (n = 6 mice), NREM-d: 13.7 ± 3.4% (n = 6 mice) versus REMD: 10.2 ± 3.7% (n = 6 mice) for new spines formed over 16–24 h and located 2–6 μm to transient new spines formed 0–8 h (P = 0.87 for ND versus REMD and P = 0.746 for NREM-d versus REMD, Wilcoxon–Mann–Whitney test).

In Fig. 2h, new spines formed over 24–36 h after BW training and located with 2 μm to persistent and transient new spines form over 0–8 h after BW training were examined in P30 mice. ND: 5.4 ± 3.6% (n = 6 mice) versus REMD: 6.1 ± 3.9% (n = 6 mice) for new spines formed over 24–36 h and located with 2 μm to persistent new spines formed 0–8 h (P = 0.924, Wilcoxon–Mann–Whitney test). ND: 34.7 ± 6.1% (n = 6 mice) versus REMD: 12.8 ± 5.2% (n = 6 mice) for new spines formed over 24–36 h and located with 2 μm to transient new spines formed 0–8 h (P = 0.024, Wilcoxon–Mann–Whitney test). P = 0.009 for ND (near transient new spines) versus ND (near persistent new spines), P = 0.011 for ND (near transient new spines) versus REMD (near persistent new spines), P = 0.346 for REMD (near transient new spines) versus REMD (near persistent new spines), Wilcoxon–Mann–Whitney test.

In Fig. 2i, new spines formed over 24–36 h after BW training and located 2–6 μm to persistent and transient new spines form over 0–8 h after BW training were examined in P30 mice. ND: 10.9 ± 4.3% (n = 6 mice) versus REMD: 19.2 ± 6.9% (n = 6 mice) for new spines formed over 24–36 h and located 2–6 μm to persistent new spines formed 0–8 h (P = 0.328, Wilcoxon–Mann–Whitney test). ND: 12.8 ± 4.8% (n = 6 mice) versus REMD: 11.0 ± 3.6% (n = 6 mice) for new spines formed over 24–36 h and located 2–6 μm to transient new spines formed 0–8 h (P = 0.87, Wilcoxon–Mann–Whitney test).

FW training-induced new spines over 16–24 h were analyzed in Fig. 2j. ND: 6.6 ± 0.7% (n = 6 mice), NREM-d: 6.5 ± 0.4% (n = 6 mice) versus REMD: 4.3 ± 0.2% (n = 6 mice) (P = 0.01 for ND versus REMD and P = 0.004 for NREM-d versus REMD, Wilcoxon–Mann–Whitney test).

BW training-induced new spines over 24–36 h were analyzed in Fig. 2k. ND: 5.4 ± 0.3% (n = 6 mice) versus REMD: 3.7 ± 0.4% (n = 6 mice) (P = 0.037, Wilcoxon–Mann–Whitney test).

In Fig. 2l, the rotarod performance improvement over 16–24 h after FW running was examined. ND: 51.6 ± 7.9% (n = 7 mice), NREM-d: 54.7 ± 5.9% (n = 7 mice) versus REMD: 22.6 ± 4.4% (n = 6 mice. P = 0.004 for ND versus REMD and P = 0.007 for NREM-d versus REMD, Wilcoxon–Mann–Whitney test).

In Fig. 2m, the rotarod performance improvement over 24–36 h after BW running was examined. ND: 38.8 ± 7.4% (n = 6 mice) versus REMD: 9.5 ± 2.4% (n = 6 mice) (P = 0.016, Wilcoxon–Mann–Whitney test).

Size change of persistent new spines was analyzed over 8–16 h in P21 mice in Fig. 3c. ND: 16.4 ± 7.9% (n = 32 new spines from 7 mice), NREM-d: 19.6 ± 9.2% (n = 24 new spines

from 5 mice) versus REMD: −7.5 ± 3.6% (n = 55 new spines from 7 mice). $P = 0.006$ for ND versus REMD and $P = 0.018$ for NREM-d versus REMD, Wilcoxon–Mann–Whitney test.

Size change of existing spines was analyzed over 8–16 h in P21 mice in Fig. 3d. ND: 6.3 ± 3.8% (n = 72 spines from 7 mice), NREM-d: 2.5 ± 3.9% (n = 63 spines from 5 mice) versus REMD: 1.9 ± 2.6% (n = 91 spines from 7 mice). $P = 0.481$ for ND versus REMD and $P = 0.787$ for NREM-d versus REMD, Wilcoxon–Mann–Whitney test.

Size change of persistent new spines induced by FW running was analyzed over 8–24 h in P30 mice in Fig. 3g. ND: 30.9 ± 8.2% (n = 22 new spines from 6 mice), NREM-d: 26.8 ± 7.4% (n = 37 new spines from 7 mice) versus REMD: 0.1 ± 7.3% (n = 37 new spines from 6 mice) over 8–16 h ($P = 0.001$ for ND versus REMD and $P = 0.0001$ for NREM-d versus REMD, Wilcoxon–Mann–Whitney test). ND: 53.7 ± 8.9% (n = 54 new spines from 16 mice), NREM-d: 52.2 ± 11.7% (n = 31 new spines from 7 mice) versus REMD: −6.7 ± 4.3% (n = 75 new spines from 16 mice) over 8–24 h ($P = 1.19 \times 10^{-10}$ for ND versus REMD and $P = 4.52 \times 10^{-8}$ for NREM-d versus REMD, Wilcoxon–Mann–Whitney test).

Size change of existing spines was analyzed over 8–24 h in P30 mice in Fig. 3h. ND: 0.02 ± 4.7% (n = 45 spines), NREM-d: 1.8 ± 4.1% (n = 59 spines) versus REMD: 0.4 ± 4.9% (n = 40 spines) over 8–16 h ($P = 0.812$ for ND versus REMD and $P = 0.977$ for NREM-d versus REMD, Wilcoxon–Mann–Whitney test). ND: 4.9 ± 3.8% (n = 45 spines), NREM-d: 2.2 ± 3.5% (n = 59 spines) versus REMD: 3.8 ± 5.1% (n = 40 spines) over 8–24 h ($P = 0.544$ for ND versus REMD and $P = 0.875$ for NREM-d versus REMD, Wilcoxon–Mann–Whitney test).

Size change of persistent new spines induced by FW running was analyzed over 8–24 h in P30 mice in Fig. 3i. ND: 32.1 ± 8.5% (n = 23 new spines from 6 mice), NREM-d: 35.1 ± 10.7% (n = 23 new spines from 6 mice) versus REMD: −3.9 ± 6.1% (n = 57 new spines from 6 mice) ($P = 4.25 \times 10^{-4}$ for ND versus REMD and $P = 8.28 \times 10^{-4}$ for NREM-d versus REMD, Wilcoxon–Mann–Whitney test).

Size change of existing spines was analyzed over 8–24 h in P30 mice in Fig. 3j. ND: 0.07 ± 3.5% (n = 45 spines), NREM-d: 0.09 ± 4.0% (n = 48 spines) versus REMD: −2.1 ± 3.4% (n = 56 spines) over 8–16 h ($P = 0.428$ for ND versus REMD and $P = 0.794$ for NREM-d versus REMD, Wilcoxon–Mann–Whitney test).

In Fig. 3k, the rotarod performance improvement over 0–24 h of FW running was examined. ND: 73.6 ± 7.2% (n = 6 mice), NREM-d: 73.3 ± 3.6% (n = 6 mice) versus REMD: 42.8 ± 9.7% (n = 7 mice) ($P = 0.032$ for ND versus REMD and $P = 0.046$ for NREM-d versus REMD, Wilcoxon–Mann–Whitney test). After identification of new spines formed 0–8 h, the survival rate of new spines was analyzed over 16 h-4 d in P21 mice in Fig. 4a. ND: 34.9 ± 1.6% (n = 7 mice), NREM-d: 36.9 ± 2.1% (n = 5 mice) versus REMD: 57.7 ± 4.3% (n = 7 mice) at the time point of 16 h ($P = 0.002$ for ND versus REMD and $P = 0.006$ for NREM-d versus REMD, Wilcoxon–Mann–Whitney test). ND: 16.3 ± 2.3% (n = 4 mice), NREM-d: 20.9 ± 4.5% (n = 5 mice) versus REMD: 2.1 ± 1.3% (n = 5 mice) at the time point of 4 d (P

= 0.013 for ND versus REMD and $P = 0.008$ for NREM-d versus REMD, Wilcoxon–Mann–Whitney test).

In Fig. 4b, the elimination rate of new spines formed 0–8 h was calculated over 16 h-4 d in P21 mice. ND: 55.4 ± 6.7% (n = 4 mice), NREM-d: 45.0 ± 9.4% (n = 5 mice) versus REMD: 96.2 ± 2.4% (n = 5 mice) ($P = 0.013$ for ND versus REMD and $P = 0.008$ for NREM-d versus REMD, Wilcoxon–Mann–Whitney test).

In Fig. 4c, new spines were divided into 2 groups (survived new spines over 16 h-4 d and eliminated new spines over 16 h-4 d) and size change was analyzed over 8–16 h in P21 mice with REM sleep (ND + NREM-d). Size change of survived new spines over 16 h-4 d: 35.3 ± 9.3% (n = 21 new spines from 9 mice) versus eliminated new spines over 16 h-4 d: 9.0 ± 8.5% (n = 22 new spines from 9 mice) ($P = 0.012$, Wilcoxon–Mann–Whitney test).

In Fig. 4d, new spines were divided into 2 groups (survived new spines over 16 h-4 d and eliminated new spines over 16 h-4 d) and size change was analyzed over 8–16 h in P21 mice without REM sleep (REMD). Size change of survived new spines over 16 h-4 d: 0.6 ± 5.1% (2 new spines from 5 mice) versus eliminated new spines over 16 h-4 d: −6.1 ± 4.3% (35 new spines from 5 mice) ($P = 0.42$, Wilcoxon–Mann–Whitney test).

After identification of new spines formed 0–8 h after FW training, the survival rate of new spines was analyzed over 24–36 h in P30 mice in Fig. 4e. ND/no BW: 40.6 ± 1.7% (n = 4 mice), ND/BW: 38.4 ± 3.1% (n = 6 mice), REMD/no BW: 54.1 ± 1.5% (n = 4 mice) and REMD/BW: 56.3 ± 2.9% (n = 6 mice) at the time point of 24 h ($P = 0.02$ for ND/no BW versus REMD/no BW, $P = 0.019$ for ND/no BW versus REMD/BW, and $P = 0.013$ for ND/BW versus REMD/no BW, $P = 0.006$ for ND/BW versus REMD/BW, Wilcoxon–Mann–Whitney test). ND/no BW: 28.9 ± 1.8% (n = 4 mice), ND/BW: 30.3 ± 2.7% (n = 6 mice), REMD/no BW: 39.8 ± 2.5% (n = 4 mice) and REMD/BW: 18.8 ± 2.8% (n = 6 mice) at the time point of 36 h ($P = 0.032$ for ND/no BW versus REMD/BW, $P = 0.02$ for ND/BW versus REMD/BW, and $P = 0.011$ for REMD/no BW versus REMD/BW, Wilcoxon–Mann–Whitney test).

In Fig. 4f, the elimination rate of new spines formed 0–8 h after FW training was calculated over 24–36 h in P30 mice. ND/no BW: 28.8 ± 3.8% (n = 4 mice), ND/BW: 20.0 ± 6.5% (n = 6 mice), REMD/no BW: 26.7 ± 2.8% (n = 4 mice) and REMD/BW: 66.7 ± 4.8% (n = 6 mice) over 24–36 h ($P = 0.009$ for ND/no BW versus REMD/BW, $P = 0.004$ for ND/BW versus REMD/BW, and $P = 0.009$ for REMD/no BW versus REMD/BW, Wilcoxon–Mann–Whitney test).

In Fig. 4g new spines were divided into 2 groups (survived new spines over 24–36 h and eliminated new spines over 24–36 h) and size change was analyzed over 8–24 h in P30 mice with REM sleep (ND). Size change of survived new spines over 24–36 h: 52.9 ± 9.2% (n = 29 new spines from 10 mice) versus eliminated new spines over 24–36 h: 11.5 ± 9.6% (n = 10 new spines from 10 mice) ($P = 0.040$, Wilcoxon–Mann–Whitney test).

In Fig. 4h, new spines were divided into 2 groups (survived new spines over 24–36 h and eliminated new spines over 24–36 h) and size change was analyzed over 8–24 h in P30 mice

without REM sleep (REMD). Size change of survived new spines over 24–36 h: 2.0 ± 4.6% (n = 24 new spines from 10 mice) versus eliminated new spines over 24–36 h: −17.6 ± 6.8% (n = 23 new spines from 10 mice) (P = 0.006, Wilcoxon–Mann–Whitney test).

The number of dendritic Ca^{2+} spikes was analyzed under various brain states and normalized to quiet awake in Fig. 5c. Quiet awake: 1 ± 0.19 (n = 13 mice), running: 4.26 ± 0.68 (n = 12 mice), NREM sleep: 0.89 ± 0.18 (n = 9 mice), REM sleep: 6.70 ± 1.07 (n = 13 mice), REM sleep + MK801: 0.70 ± 0.25 (n = 8 mice). $P = 7.05 \times 10^{-5}$ for Quiet awake versus running, P = 0.867 for quiet awake versus NREM sleep, $P = 2.19 \times 10^{-4}$ for running versus NREM sleep, $P = 1.82 \times 10^{-5}$ for quiet awake versus REM sleep, P = 0.082 for running versus REM sleep, $P = 9.37 \times 10^{-5}$ for NREM sleep versus REM sleep, $P = 1.66 \times 10^{-4}$ for REM sleep versus REM sleep + MK801, Wilcoxon–Mann–Whitney test.

In Fig. 5d, the peak amplitude of dendritic Ca^{2+} spikes was analyzed. Quiet awake: 2.79 ± 0.15 (89 spikes), running: 3.46 ± 0.17 (124 spikes), NREM sleep: 2.46 ± 0.15 (88 spikes), REM sleep: 3.88 ± 0.15 (274 spikes), REM sleep + MK801: 1.67 ± 0.16 (47 spikes). P = 0.009 for quiet awake versus running, P = 0.0596 for quiet awake versus NREM sleep, $P = 1.31 \times 10^{-5}$ for running versus NREM sleep, $P = 2.82 \times 10^{-4}$ for quiet awake versus REM sleep, P = 0.373 for running versus REM sleep, $P = 9.09 \times 10^{-8}$ for NREM sleep versus REM sleep, $P = 1.77 \times 10^{-12}$ for REM sleep versus REM sleep + MK801, Wilcoxon–Mann–Whitney test.

In Fig. 5e, the duration of dendritic Ca^{2+} spikes was analyzed. Quiet awake: 1.52 ± 0.08 s (89 spikes), running: 2.37 ± 0.12 s (124 spikes), NREM sleep: 1.82 ± 0.09 s (88 spikes), REM sleep: 2.88 ± 0.09 s (274 spikes), REM sleep + MK801: 1.38 ± 0.07 s (47 spikes). $P = 9.77 \times 10^{-8}$ for quiet awake versus running, P = 0.009 for quiet awake versus NREM sleep, P = 0.002 for running versus NREM sleep, $P = 3.01 \times 10^{-17}$ for quiet awake versus REM sleep, P = 0.0007 for running versus REM sleep, $P = 1.88 \times 10^{-10}$ for NREM sleep versus REM sleep, $P = 1.47 \times 10^{-13}$ for REM sleep versus REM sleep + MK801, Wilcoxon–Mann–Whitney test.

To evaluate the effects of local MK801 application (3 pulses, 50 ms each) on dendritic Ca^{2+} spikes, the number of dendritic Ca^{2+} spikes relative to pre-injection was analyzed in Fig. 5f. 0–1 min after injection: 23.2 ± 4.1% (n = 10 mice), 1–2 min after injection: 33.9 ± 4.1% (n = 10 mice), 2–3 min after injection: 71.9 ± 1.8% (n = 10 mice), 3–4 min after injection: 104.1 ± 4.9% (n = 10 mice). P = 0.003 for 0–1 min after injection versus pre-injection, P = 0.003 for 1–2 min after injection versus pre-injection, P = 0.013 for 2–3 min after injection versus pre-injection, P = 0.44 for 3–4 min after injection versus pre-injection, Wilcoxon test (2 related samples).

To evaluate the effects of local MK801 application (3 pulses, 50 ms each) on dendritic Ca^{2+} spikes during REM sleep, the number of dendritic Ca^{2+} spikes relative to undisturbed REM sleep was analyzed in Fig. 5g. REM + MK801: 8.4 ± 0.9% (n = 46 episodes of REM sleep with MK801 injection from 4 mice), NREM + MK801: 100.5 ± 1.8% (n = 51 episodes of NREM sleep with MK801 injection from 4 mice), $P = 2.36 \times 10^{-17}$, Wilcoxon–Mann–Whitney test.

The effect of MK801 on the elimination rate of treadmill training-induced new spines was calculated in Fig. 5h. ND: 60.4 ± 3.7% (n = 6 mice), REMD: 42.0 ± 2.4% (n = 5 mice), REM + MK801: 34.7 ± 2.9% (n = 5 mice), NREM + MK801: 65.2 ± 2.2% (n = 6 mice). P = 0.008 for ND versus REMD, P = 0.006 for ND versus REM + MK801, P = 0.297 for ND versus NREM + MK801, P = 0.006 for REM + MK801 versus NREM + MK801, Wilcoxon–Mann–Whitney test.

The effect of MK801 on the size change of persistent new spines was calculated in Fig. 5i. ND: 31.1 ± 9.8% (n = 25 new spines from 6 mice), REMD: −8.8 ± 6.2% (n = 40 new spines from 5 mice), REM + MK801: −6.5 ± 4.8% (n = 26 new spines from 5 mice), NREM + MK801: 28.2 ± 7.9% (n = 29 new spines from 6 mice). P = 0.0006 for ND versus REMD, P = 0.0019 for ND versus REM + MK801, P = 0.801 for ND versus NREM + MK801, P = 0.0005 for REM + MK801 versus NREM + MK801, Wilcoxon–Mann–Whitney test.

Cumulative reactivation of REM sleep-activated dendrites during retraining was presented in Fig. 6a. 1 min-retraining: 23.6 ± 1.5%, 2 min-retraining: 38.3 ± 3.1%, 3 min-retraining: 53.0 ± 2.9%, 4 min-retraining: 62.5 ± 1.5%, n = 4 mice.

The elimination rate of new spines was calculated after retraining and REMD in Fig. 6b. REMD: 42.0 ± 2.4% (n = 5 mice), retraining + REMD: 62.5 ± 2.5% (n = 6 mice), MK801 + retraining + REMD: 40.4 ± 3.5% (n = 4 mice), retraining + MK801 + REMD: 60.4 ± 2.0% (n = 5 mice). P = 0.006 for REMD versus retraining + REMD, P = 0.01 for retraining + REMD versus MK801 + retraining + REMD, P = 0.64 for retraining + REMD versus retraining + MK801 + REMD, P = 0.014 for MK801 + retraining + REMD versus retraining + MK801 + REMD, Wilcoxon–Mann–Whitney test.

The effect of retraining on the size change of persistent new spines was calculated in Fig. 6c. REMD: −8.8 ± 6.2% (n = 40 new spines from 5 mice), retraining + REMD: 27.4 ± 9.0% (n = 26 new spines from 6 mice), MK801 + retraining + REMD: −5.8 ± 6.6% (n = 25 new spines from 4 mice), retraining + MK801 + REMD: 27.9 ± 5.9% (n = 25 new spines from 5 mice). P = 0.0009 for REMD versus retraining + REMD, P = 0.005 for retraining + REMD versus MK801 + retraining + REMD, P = 0.85 for retraining + REMD versus retraining + MK801 + REMD, $P = 9.24 \times 10^{-5}$ for MK801 + retraining + REMD versus retraining + MK801 + REMD, Wilcoxon–Mann–Whitney test.

To evaluate the effects of local MK801 application (10 pulses, 50 ms each) on dendritic Ca^{2+} spikes, the number of dendritic Ca^{2+} spikes relative to pre-injection was analyzed in Fig. 6d. 0–1 min after injection: 4.2 ± 2.0% (n = 6 mice), 10–11 min after injection: 9.1 ± 3.1% (n = 6 mice), 20–21 min after injection: 22.3 ± 1.6% (n = 6 mice), 30–31 min after injection: 31.7 ± 3.4% (n = 6 mice), 40–41 min after injection: 45.5 ± 3.1% (n = 6 mice), 50–51 min after injection: 88.0 ± 6.0% (n = 6 mice), 60–61 min after injection: 105.1 ± 8.8% (n = 6 mice). P = 0.026 for 0–1 min after injection versus pre-injection, P = 0.027 for 10–11 min after injection versus pre-injection, P = 0.028 for 20–21 min after injection versus pre-injection, P = 0.028 for 30–31 min after injection versus pre-injection, P = 0.028 for 40–41 min after injection versus pre-injection, P = 0.116 for 50–51 min after injection versus pre-injection, P = 0.6 for 60–61 min after injection versus pre-injection, Wilcoxon test (2 related samples).

The number of dendritic Ca^{2+} spikes was analyzed and normalized to retraining in Fig. 6e. Retraining: 1 ± 0.12 (n = 5 mice), KN62 + retraining: 1.09 ± 0.07 (n = 5 mice). $P = 0.345$, Wilcoxon test (2 related samples).

In Fig. 6f, the peak amplitude of dendritic Ca^{2+} spikes was analyzed. Retrain: 3.27 ± 0.20 (146 spikes), KN62 + retraining: 3.39 ± 0.16 (162 spikes). $P = 0.181$, Wilcoxon–Mann–Whitney test.

In Fig. 6g, the duration of dendritic Ca^{2+} spikes was analyzed. Retrain: 2.34 ± 0.12 s (146 spikes), KN62 + retraining: 2.63 ± 0.14 (162 spikes). $P = 0.191$, Wilcoxon–Mann–Whitney test.

In Fig. 6h, peak amplitude distribution of dendritic calcium spikes over a period of 1 min between retraining and KN62 + retraining was compared. Retraining: 10.3% (peak amplitude < 1), 46.6% (peak amplitude 1–3), 24.0% (peak amplitude 3–5), 12.3% (peak amplitude 5–7), 6.8% (peak amplitude > 7). KN62 + retraining: 7.4% (peak amplitude < 1), 43.8% (peak amplitude 1–3), 31.5% (peak amplitude 3–5), 9.3% (peak amplitude 5–7), 8.0% (peak amplitude > 7). $P = 0.408$ for retraining versus KN62 + retraining, Wilcoxon–Mann–Whitney test.

The effect of KN62 on elimination rate of new spines after retraining and REMD was evaluated in Fig. 6i. Retraining + REMD: 62.5 ± 2.5% (n = 6 mice), KN62 + retraining + REMD: 42.7 ± 3.8% (n = 6 mice), DMSO + retraining +REMD: 61.5 ± 3.2% (n = 5 mice), retraining + KN62 + REMD: 63.9 ± 4.1% (n = 6 mice). $P = 0.008$ for retraining + REMD versus KN62 + retraining + REMD, $P = 0.013$ for KN62 + retraining + REMD versus DMSO + retraining + REMD, $P = 0.006$ for KN62 + retraining + REMD versus retraining + KN62 + REMD, Wilcoxon–Mann–Whitney test.

The effect of KN62 on the size change of persistent new spines after retraining and REMD was calculated in Fig. 6j. Retraining + REMD: 27.4 ± 9.0% (n = 26 new spines from 6 mice), KN62 + retraining + REMD: −2.5 ± 5.6% (n = 31 new spines from 6 mice), DMSO + retraining + REMD: 27.8 ± 9.2% (n = 25 new spines from 5 mice), retraining + KN62 + REMD: 31.3 ± 8.1% (n = 26 new spines from 6 mice). $P = 0.007$ for retraining + REMD versus KN62 + retraining + REMD, $P = 0.008$ for KN62 + retraining + REMD versus DMSO + retraining + REMD, $P = 0.003$ for KN62 + retraining + REMD versus retraining + KN62 + REMD, Wilcoxon–Mann–Whitney test.

Data availability—The data that support the findings of this study are available from the corresponding authors upon request.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Dr. Joseph Cichon for technical help, and all the members in the Gan laboratory for comments on the manuscript. This study was supported by the funding from NIH R01 NS047325 and R01 MH111486 to W.-B.G., by

R01 GM107469, R21 AG048410 to G.Y., and by the National Natural Science Foundation of China 81100839 to W.L., and by Shenzhen Science and Technology Innovation Funds GJHS20120628101219327, JC201105170726A, JCYJ20160428154351820, JSGG20140703163838793, ZDSYS201504301539161 and KQTD2015032709315529 to W.L. and L.M.

References

1. Roffwarg HP, Muzio JN, Dement WC. Ontogenetic development of the human sleep-dream cycle. Science. 1966; 152:604–619. DOI: 10.1126/science.152.3722.604 [PubMed: 17779492]
2. Jouvet M. Paradoxical sleep and the nature-nurture controversy. Prog Brain Res. 1980; 53:331–346. DOI: 10.1016/S0079-6123(08)60073-4 [PubMed: 7005950]
3. Siegel JM. The REM sleep-memory consolidation hypothesis. Science. 2001; 294:1058–1063. DOI: 10.1126/science.1063049 [PubMed: 11691984]
4. Hobson JA. REM sleep and dreaming: towards a theory of protoconsciousness. Nat Rev Neurosci. 2009; 10:803–813. DOI: 10.1038/nrn2716 [PubMed: 19794431]
5. Karni A, Tanne D, Rubenstein BS, Askenasy JJ, Sagi D. Dependence on REM sleep of overnight improvement of a perceptual skill. Science. 1994; 265:679–682. [PubMed: 8036518]
6. Rasch B, Born J. About sleep's role in memory. Physiological reviews. 2013; 93:681–766. DOI: 10.1152/physrev.00032.2012 [PubMed: 23589831]
7. Llinas RR, Pare D. Of dreaming and wakefulness. Neuroscience. 1991; 44:521–535. [PubMed: 1754050]
8. Vertes RP. Memory consolidation in sleep; dream or reality. Neuron. 2004; 44:135–148. DOI: 10.1016/j.neuron.2004.08.034 [PubMed: 15450166]
9. Crick F, Mitchison G. The function of dream sleep. Nature. 1983; 304:111–114. [PubMed: 6866101]
10. Smith C, Lapp L. Increases in number of REMS and REM density in humans following an intensive learning period. Sleep. 1991; 14:325–330. [PubMed: 1947596]
11. Marks GA, Shaffery JP, Oksenberg A, Speciale SG, Roffwarg HP. A functional role for REM sleep in brain maturation. Behav Brain Res. 1995; 69:1–11.
12. Mirmiran M. The function of fetal/neonatal rapid eye movement sleep. Behav Brain Res. 1995; 69:13–22. [PubMed: 7546304]
13. Datta S, Mavanji V, Ulloor J, Patterson EH. Activation of phasic pontine-wave generator prevents rapid eye movement sleep deprivation-induced learning impairment in the rat: A mechanism for sleep-dependent plasticity. J Neurosci. 2004; 24:1416–1427. DOI: 10.1523/Jneurosci.4111-03.2004 [PubMed: 14960614]
14. Frank MG, Issa NP, Stryker MP. Sleep enhances plasticity in the developing visual cortex. Neuron. 2001; 30:275–287. DOI: 10.1016/S0896-6273(01)00279-3 [PubMed: 11343661]
15. Lavie P, Pratt H, Scharf B, Peled R, Brown J. Localized Pontine Lesion - Nearly Total Absence of Rem-Sleep. Neurology. 1984; 34:118–120. [PubMed: 6537835]
16. Vertes RP, Siegel JM. Time for the sleep community to take a critical look at the purported role of sleep in memory processing. Sleep. 2005; 28:1228–1229. [PubMed: 16295207]
17. Vertes RP, Eastman KE. The case against memory consolidation in REM sleep. Behav Brain Sci. 2000; 23:867–876. discussion 904–1121. [PubMed: 11515146]
18. Rasch B, Pommer J, Diekelmann S, Born J. Pharmacological REM sleep suppression paradoxically improves rather than impairs skill memory. Nat Neurosci. 2009; 12:396–397. DOI: 10.1038/nn.2206 [PubMed: 18836440]
19. Stickgold R. Sleep-dependent memory consolidation. Nature. 2005; 437:1272–1278. DOI: 10.1038/nature04286 [PubMed: 16251952]
20. Abel T, Havekes R, Saletin JM, Walker MP. Sleep, plasticity and memory from molecules to whole-brain networks. Curr Biol. 2013; 23:R774–788. DOI: 10.1016/j.cub.2013.07.025 [PubMed: 24028961]
21. Crick F, Mitchison G. REM sleep and neural nets. Behav Brain Res. 1995; 69:147–155. [PubMed: 7546306]

22. Hopfield JJ, Feinstein DI, Palmer RG. 'Unlearning' has a stabilizing effect in collective memories. Nature. 1983; 304:158–159. [PubMed: 6866109]
23. Changeux JP, Danchin A. Selective stabilisation of developing synapses as a mechanism for the specification of neuronal networks. Nature. 1976; 264:705–712. [PubMed: 189195]
24. Lichtman JW, Colman H. Synapse elimination and indelible memory. Neuron. 2000; 25:269–278. [PubMed: 10719884]
25. Yang G, Pan F, Gan WB. Stably maintained dendritic spines are associated with lifelong memories. Nature. 2009; 462:920–924. DOI: 10.1038/nature08577 [PubMed: 19946265]
26. Grutzendler J, Kasthuri N, Gan WB. Long-term dendritic spine stability in the adult cortex. Nature. 2002; 420:812–816. [PubMed: 12490949]
27. Ribeiro S, Goyal V, Mello CV, Pavlides C. Brain gene expression during REM sleep depends on prior waking experience. Learn Mem. 1999; 6:500–508. [PubMed: 10541470]
28. Ulloor J, Datta S. Spatio-temporal activation of cyclic AMP response element-binding protein, activity-regulated cytoskeletal-associated protein and brain-derived nerve growth factor: a mechanism for pontine-wave generator activation-dependent two-way active-avoidance memory processing in the rat. Journal of neurochemistry. 2005; 95:418–428. DOI: 10.1111/j.1471-4159.2005.03378.x [PubMed: 16190868]
29. Ravassard P, et al. Paradoxical (REM) sleep deprivation causes a large and rapidly reversible decrease in long-term potentiation, synaptic transmission, glutamate receptor protein levels, and ERK/MAPK activation in the dorsal hippocampus. Sleep. 2009; 32:227–240. [PubMed: 19238810]
30. Bridi MCD, et al. Rapid eye movement sleep promotes cortical plasticity in the developing brain. Science Advances. 2015; 03
31. Yang G, et al. Sleep promotes branch-specific formation of dendritic spines after learning. Science. 2014; 344:1173–1178. 344/6188/1173 [pii]. DOI: 10.1126/science.1249098 [PubMed: 24904169]
32. Katz LC, Shatz CJ. Synaptic activity and the construction of cortical circuits. Science. 1996; 274:1133–1138. [PubMed: 8895456]
33. Nelson AB, Faraguna U, Zoltan JT, Tononi G, Cirelli C. Sleep patterns and homeostatic mechanisms in adolescent mice. Brain sciences. 2013; 3:318–343. DOI: 10.3390/brainsci3010318 [PubMed: 23772316]
34. Liston C, et al. Circadian glucocorticoid oscillations promote learning-dependent synapse formation and maintenance. Nat Neurosci. 2013; 16:698–705. DOI: 10.1038/nn.3387 [PubMed: 23624512]
35. Newman EA, Evans CR. Human dream processes as analogous to computer programme clearance. Nature. 1965; 206:534. [PubMed: 5831856]
36. Gaarder K. A conceptual model of sleep. Arch Gen Psychiatry. 1966; 14:253–260. [PubMed: 5903417]
37. Hayashi-Takagi A, et al. Labelling and optical erasure of synaptic memory traces in the motor cortex. Nature. 2015; 525:333–338. DOI: 10.1038/nature15257 [PubMed: 26352471]
38. Harris KM, Stevens JK. Dendritic spines of CA 1 pyramidal cells in the rat hippocampus: serial electron microscopy with reference to their biophysical characteristics. J Neurosci. 1989; 9:2982–2997. [PubMed: 2769375]
39. Matsuzaki M, et al. Dendritic spine geometry is critical for AMPA receptor expression in hippocampal CA1 pyramidal neurons. Nat Neurosci. 2001; 4:1086–1092. [PubMed: 11687814]
40. Cichon J, Gan WB. Branch-specific dendritic Ca(2+) spikes cause persistent synaptic plasticity. Nature. 2015; 520:180–185. DOI: 10.1038/nature14251 [PubMed: 25822789]
41. Lavzin M, Rapoport S, Polsky A, Garion L, Schiller J. Nonlinear dendritic processing determines angular tuning of barrel cortex neurons in vivo. Nature. 2012; 490:397–401. DOI: 10.1038/nature11451 [PubMed: 22940864]
42. Smith SL, Smith IT, Branco T, Hausser M. Dendritic spikes enhance stimulus selectivity in cortical neurons in vivo. Nature. 2013; 503:115–120. DOI: 10.1038/nature12600 [PubMed: 24162850]
43. Xu NL, et al. Nonlinear dendritic integration of sensory and motor input during an active sensing task. Nature. 2012; 492:247–251. DOI: 10.1038/nature11601 [PubMed: 23143335]

44. Larkum ME, Nevian T, Sandler M, Polsky A, Schiller J. Synaptic integration in tuft dendrites of layer 5 pyramidal neurons: a new unifying principle. Science. 2009; 325:756–760. DOI: 10.1126/science.1171958 [PubMed: 19661433]
45. Sheffield ME, Dombeck DA. Calcium transient prevalence across the dendritic arbour predicts place field properties. Nature. 2015; 517:200–204. DOI: 10.1038/nature13871 [PubMed: 25363782]
46. Golding NL, Staff NP, Spruston N. Dendritic spikes as a mechanism for cooperative long-term potentiation. Nature. 2002; 418:326–331. DOI: 10.1038/Nature00854 [PubMed: 12124625]
47. Holthoff K, Kovalchuk Y, Yuste R, Konnerth A. Single-shock LTD by local dendritic spikes in pyramidal neurons of mouse visual cortex. J Physiol. 2004; 560:27–36. DOI: 10.1113/jphysiol.2004.072678 [PubMed: 15319420]
48. Kampa BM, Letzkus JJ, Stuart GJ. Requirement of dendritic calcium spikes for induction of spike-timing-dependent synaptic plasticity. J Physiol. 2006; 574:283–290. DOI: 10.1113/jphysiol.2006.111062 [PubMed: 16675489]
49. Nevian T, Sakmann B. Single spine Ca2+ signals evoked by coincident EPSPs and backpropagating action potentials in spiny stellate cells of layer 4 in the juvenile rat somatosensory barrel cortex. J Neurosci. 2004; 24:1689–1699. DOI: 10.1523/JNEUROSCI.3332-03.2004 [PubMed: 14973235]
50. Grosmark AD, Mizuseki K, Pastalkova E, Diba K, Buzsáki G. REM sleep reorganizes hippocampal excitability. Neuron. 2012; 75:1001–1007. [PubMed: 22998869]
51. Yang G, Pan F, Parkhurst CN, Grutzendler J, Gan WB. Thinned-skull cranial window technique for long-term imaging of the cortex in live mice. Nat Protoc. 2010; 5:201–208. DOI: 10.1038/nprot.2009.222 [PubMed: 20134419]
52. Pan F, Aldridge GM, Greenough WT, Gan WB. Dendritic spine instability and insensitivity to modulation by sensory experience in a mouse model of fragile X syndrome. Proc Natl Acad Sci U S A. 2010; 107:17768–17773. DOI: 10.1073/pnas.1012496107 [PubMed: 20861447]
53. Fu M, Yu X, Lu J, Zuo Y. Repetitive motor learning induces coordinated formation of clustered dendritic spines in vivo. Nature. 2012; 483:92–95. DOI: 10.1038/nature10844 [PubMed: 22343892]
54. Hofer SB, Mrsic-Flogel TD, Bonhoeffer T, Hubener M. Experience leaves a lasting structural trace in cortical circuits. Nature. 2009; 457:313–317. [PubMed: 19005470]
55. Zuo Y, Lin A, Chang P, Gan WB. Development of long-term dendritic spine stability in diverse regions of cerebral cortex. Neuron. 2005; 46:181–189. [PubMed: 15848798]
56. Tennant KA, et al. The organization of the forelimb representation of the C57BL/6 mouse motor cortex as defined by intracortical microstimulation and cytoarchitecture. Cerebral cortex. 2011; 21:865–876. DOI: 10.1093/cercor/bhq159 [PubMed: 20739477]

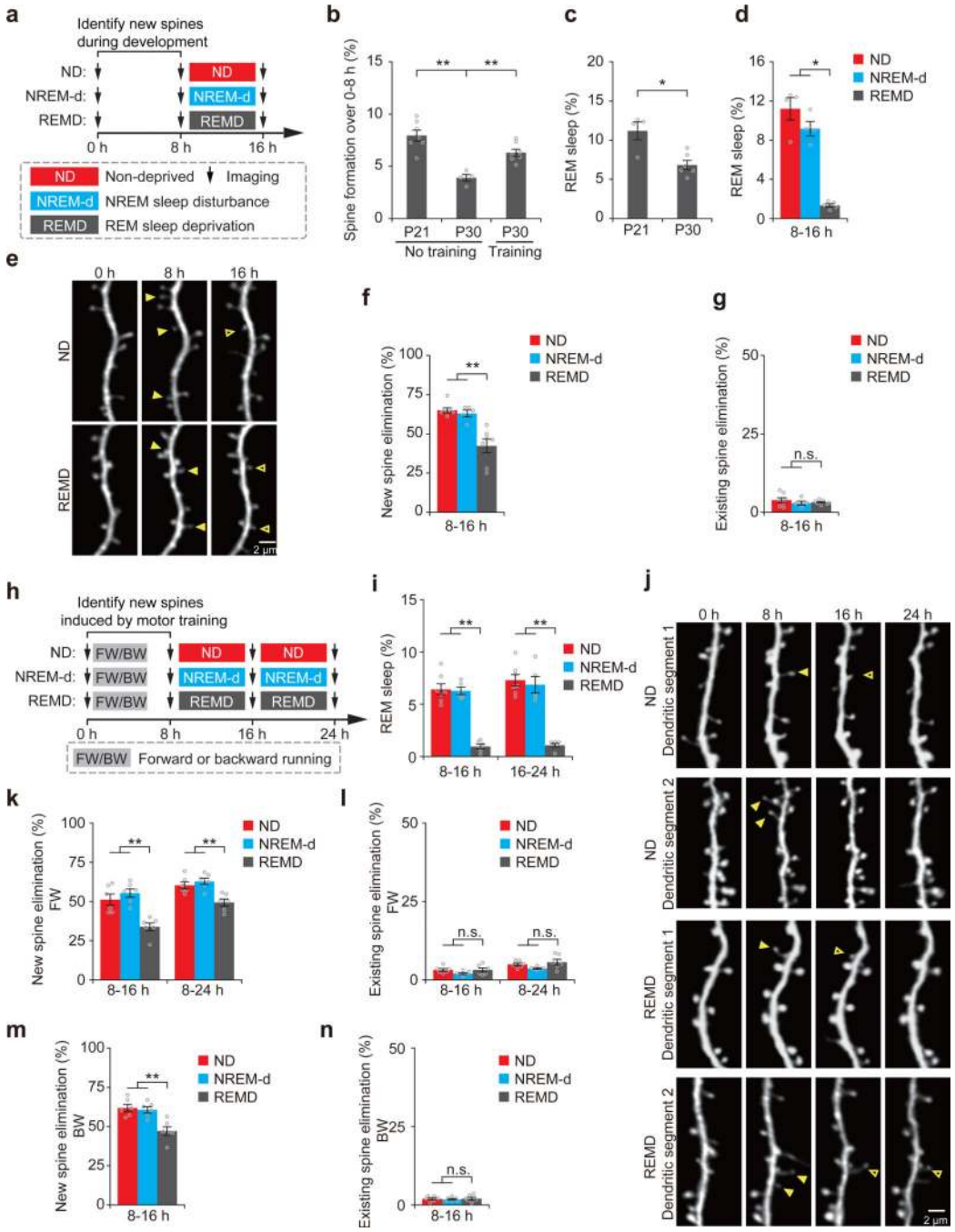

Figure 1. REM sleep prunes newly formed spines during development and after learning

(a) Schematic of experimental design. After identification of new spines formed between hours 0–8, young mice at P21 were either left undisturbed (nondeprived control, ND), subjected to NREM sleep disturbance (NREM-d) or REM sleep deprivation (REMD) and new spines were imaged over the next 8 h. **(b)** The rate of new spine formation was significantly reduced as animals matured ($P = 0.008$; n = 7 and 4 mice at P21 and P30, respectively). Motor training significantly increased the formation of new spines in P30 mice; $P = 0.008$; n = 7 and 4 mice with and without motor training, respectively). **(c)** The

amount of REM sleep was significantly lower in P30 mice than in P21 mice ($P = 0.033$, n = 6 P30 mice and 4 P21 mice). **(d)** The amount of REM sleep was significantly reduced in REMD mice as compared to ND and NREM-d mice between hours 8–16 (ND vs. REMD, $P = 0.014$; NREM-d vs. REMD, $P = 0.014$; n = 4, 4 and 5 ND, NREM-d and REMD mice, respectively). **(e)** Repeated imaging of dendritic spines on apical tuft dendrites of L5 pyramidal neurons in ND and REMD mice at P21. Filled arrowheads indicate new spines formed during hours 0–8. Some new spines (open arrowheads) persisted for the next 8 h. **(f)** The elimination rate of new spines (formed between hours 0–8) was significantly higher in ND or NREM-d mice than in REMD mice over hours 8–16 (ND vs. REMD, $P = 0.002$; NREM-d vs. REMD, $P = 0.006$; n = 7, 5 and 7 ND, NREM-d and REMD mice, respectively). **(g)** REMD did not affect the elimination of existing spines over hours 8–16 (n = 7, 5 and 7 ND, NREM-d and REMD mice, respectively). **(h)** Schematic of experimental design. After identification of motor-training-induced (forward running, FW; backward running, BW) new spines formed between hours 0–8, 1-month-old (P30) mice were either left undisturbed (ND), subjected to NREM sleep disturbance (NREM-d) or REM sleep deprivation (REMD) and new spines were imaged over the next 8–16 h. **(i)** The amount of REM sleep was significantly reduced in REMD mice as compared to ND and NREM-d mice between hours 8–16 (ND vs. REMD, $P = 0.003$; NREM-d vs. REMD, $P = 0.006$) and 16–24 (ND vs. REMD, $P = 0.003$; NREM-d vs. REMD, $P = 0.006$; n = 7, 5 and 6 ND, NREM-d and REMD mice, respectively). **(j)** Repeated imaging of dendritic spines before and 24 h after rotarod training in ND and REMD mice. Filled arrowheads indicate new spines formed during hours 0–8 after training. Some new spines (open arrowheads) persisted over the next 8–16 h. **(k)** The elimination rate of new spines (formed between 0–8 h after FW training) was significantly higher in ND or NREM-d mice than in REMD mice over the next 8–16 h ($P = 0.004$ for ND or NREM-d versus REMD over hours 8–16; $P = 0.008$ for ND or NREM-d versus REMD over hours 8–24; n = 6 mice for each group). **(l)** REMD did not affect the elimination of existing spines between hours 8–24 (over hours 8–16: ND vs. REMD, $P = 0.748$; NREM-d vs. REMD, $P = 0.261$; over hours 8–24: ND vs. REMD, $P = 0.749$; NREM-d vs. REMD, $P = 0.078$; n = 6 mice for each group). **(m)** The elimination rate of new spines formed 0–8 h after backward-running (BW) training was significantly higher in ND or NREM-d mice than in REMD mice over the next 8 h ($P = 0.006$ for ND or NREM-d versus REMD, n = 6 mice for each group). **(n)** REMD did not affect the elimination rate of existing spines over hours 8–16 (n = 6 mice for each group). Data are presented as mean ± s.e.m. Each point in b–d,f,g,i,k–n represents data from one animal. $*P < 0.05$, $**P < 0.01$, n.s. = not significant. Scale bars, 2 μm.

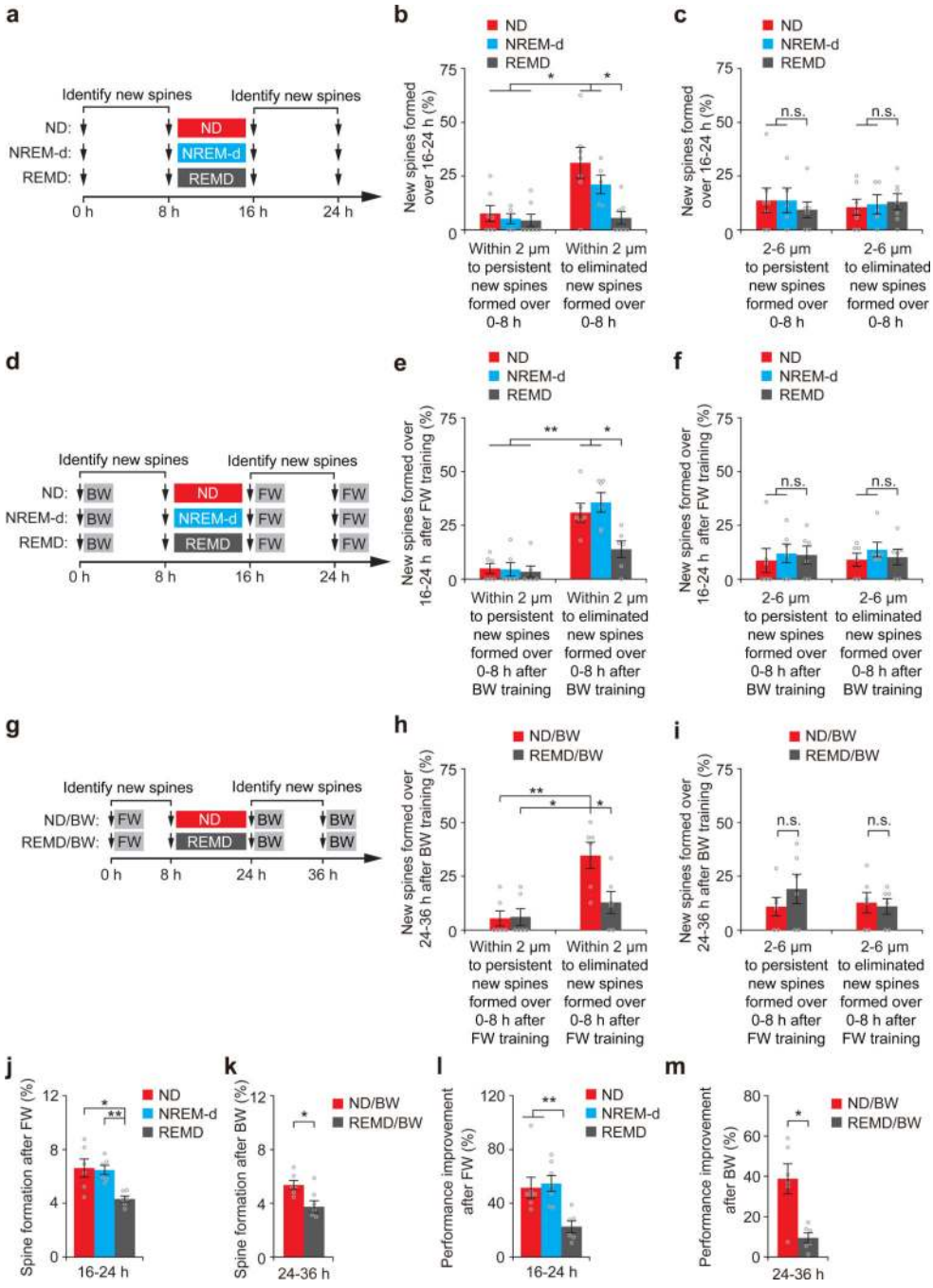

Figure 2. REM sleep-dependent spine elimination facilitates subsequent new spine formation at nearby sites

(a) Experimental design for examining the relationship between new spines formed during hours 0–8 and new spines formed over hours 16–24. New spines were identified in P21 mice over hours 0–8 and classified as persistent or eliminated based on their fate over hours 8–16. **(b)** New spines formed over hours 16–24 were rarely located within 2 µm of persistent new spines formed over hours 0–8. Significantly larger percentages of new spines formed over hours 16–24 were located within 2 µm of transient new spines formed over hours 0–8 in ND

and NREM-d mice than in REMD mice ($P = 0.01$ for ND or NREM-d versus REMD, n = 7, 5 and 7 mice for ND, NREM-d and REMD, respectively). **(c)** The percentages of new spines formed over hours 16–24 and located 2–6 μm away from persistent or eliminated new spines formed over hours 0–8 were comparable among ND, NREM-d and REMD mice (n = 7, 5 and 7 mice for ND, NREM-d and REMD, respectively). **(d)** Experimental design for examining the relationship between the elimination of new spines induced by BW and the formation of new spines induced by FW. New spines were identified in mice subjected to BW between hours 0–8. Eight hours after ND, NREM-d or REMD sleep, the animals were subjected to FW. **(e)** FW-induced new spines were rarely located within 2 μm of persistent BW-induced new spines. Significantly larger percentages of FW-induced new spines were located within 2 μm of transient BW-induced new spines in ND and NREM-d mice than in REMD mice ($P = 0.01$ for ND or NREM-d versus REMD, n = 6 mice for each group). **(f)** The percentages of new spines formed over hours 16–24 after FW and located 2–6 μm from persistent or eliminated new spines formed over hours 0–8 after BW were comparable among ND, NREM-d and REMD mice (n = 6 mice for each group). **(g–i)** Experimental design as in d–f but for examining the relationship between the elimination of new spines induced by FW and the formation of new spines induced by BW; similar results were achieved. **(j)** New spines were induced by FW between hours 16–24 at a significantly higher rate in ND and NREM-d mice than in REMD mice (ND vs. REMD, $P = 0.01$; NREM-d vs. REMD, $P = 0.004$; n = 6 mice for each group). **(k)** New spines were induced by BW between hours 24–36 at a significantly higher rate in ND mice than in REMD mice ($P = 0.037$, n = 6 mice for each group). **(l)** The rotarod performance improvement of FW was significantly smaller in REMD mice compared to in ND and NREM-d mice over hours 16–24 (ND vs. REMD, $P = 0.004$; NREM-d vs. REMD, $P = 0.007$; n = 7, 7 and 6 ND, NREM-d and REMD mice, respectively). **(m)** The rotarod performance improvement of BW was significantly lower in REMD mice compared to ND mice over hours 24–36 ($P = 0.016$, n = 6 mice for each group). Data are presented as mean ± s.e.m. $*P < 0.05$, $**P < 0.01$, n.s. = not significant.

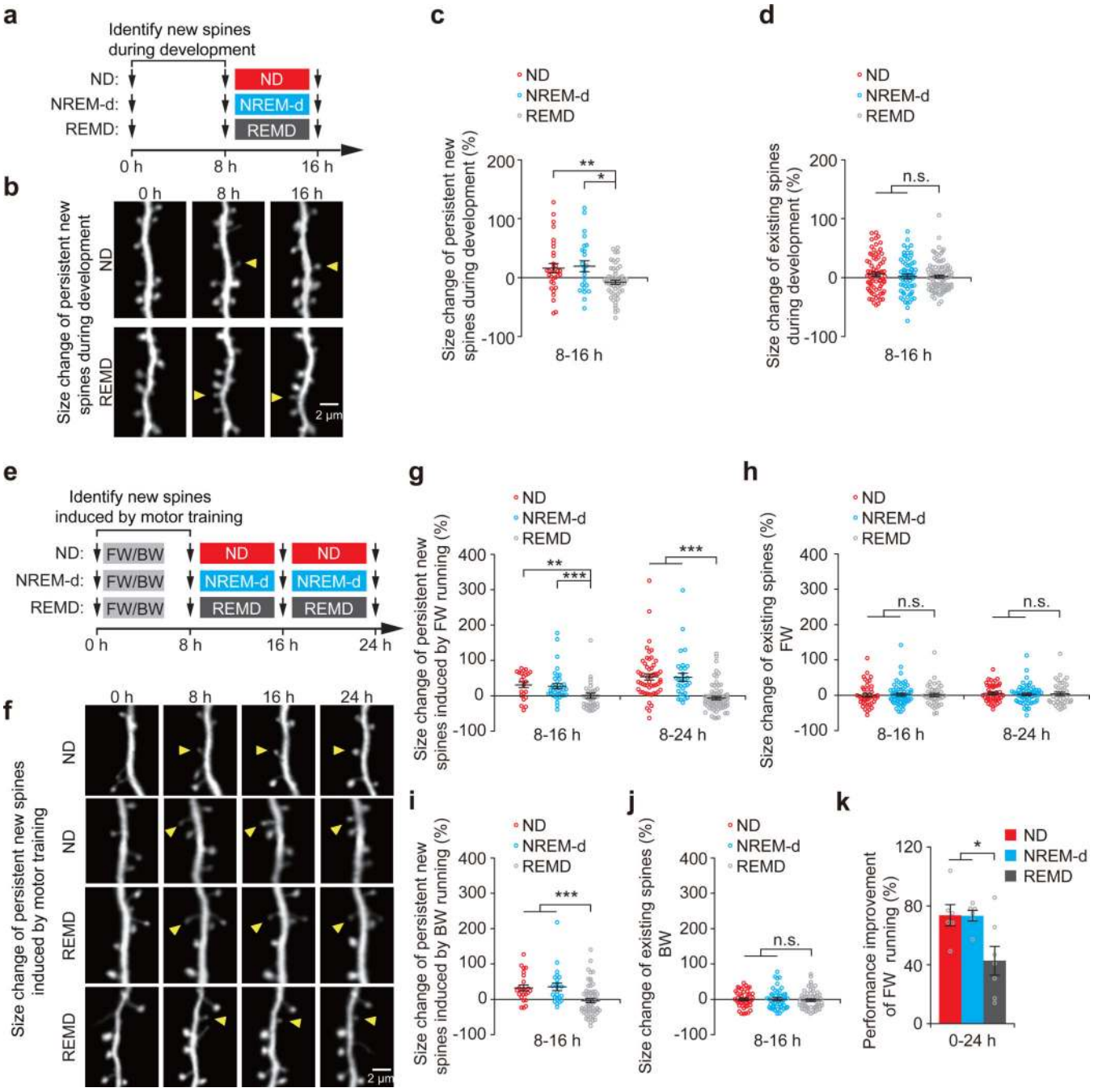

Figure 3. REM sleep strengthens a fraction of new spines formed during development and learning

(a) Schematic of experimental design. After identification of new spines formed between hours 0–8, young mice P21 were left either undisturbed (ND), subjected to NREM sleep disturbance (NREM-d) or REM sleep deprivation (REMD), and new spines were imaged over the next 8 h. **(b)** Repeated imaging of dendritic spines in ND and REMD mice at P21. Arrowheads indicate new spines that formed during hours 0–8 and persisted over time. **(c)** The average size of persistent new spines in P21 mice increased over time in ND (32 persistent new spines from 7 mice over hours 8–16) and NREM-d mice (24 persistent new

spines from 5 mice), as compared to REMD mice (55 persistent new spines from 7 mice) (ND vs. REMD, $P = 0.006$; NREM-d vs. REMD, $P = 0.018$). **(d)** REMD mice did not show substantial differences in the size of existing spines compared to ND and NREM-d mice at P21 (72, 63 and 91 randomly selected existing spines from 7, 5 and 7 ND, NREM-d and REMD mice, respectively). **(e)** Schematic of experimental design. After identification of motor-training-induced new spines formed between hours 0–8, P30 mice were left either undisturbed (ND), subjected to NREM sleep disturbance (NREM-d) or REM sleep deprivation (REMD), and new spines were imaged over the next 8–16 h. **(f)** Repeated imaging of dendritic spines before and 24 h after motor training in ND and REMD mice. Arrowheads indicate new spines that formed 0–8 h after training and persisted over time. **(g)** The average size of persistent new spines induced by FW increased over time in ND (22 new spines from 6 mice over hours 8–16, and 54 new spines from 16 mice over hours 8–24) and NREM-d mice (37 spines from 7 mice over hours 8–16, and 31 spines from 7 mice over hours 8–24) but not in REMD mice (37 spines from 6 mice over hours 8–16, and 75 spines from 16 mice over hours 8–24) (over hours 8–16: ND vs. REMD, $P = 0.001$; NREM-d vs. REMD, $P = 0.0001$; over hours 8–24: ND vs. REMD, $P = 1.19 \times 10^{-10}$; NREM-d vs. REMD, $P = 4.52 \times 10^{-8}$). **(h)** REMD mice did not show any effects on the size of existing spines as compared to ND or NREM-d mice at P30 (45, 59 and 40 randomly selected existing spines for ND, NREM-d and REMD mice, respectively). **(i)** The average size of persistent new spines induced by BW increased over time in ND (23 new spines from 6 mice over hours 8–16) and NREM-d mice (23 spines from 6 mice over hours 8–16) but not in REMD mice (57 spines from 6 mice over hours 8–16) (ND vs. REMD, $P = 4.25 \times 10^{-4}$; NREM-d vs. REMD, $P = 8.28 \times 10^{-4}$). **(j)** REMD mice did not show any effects on the size of existing spines as compared to ND or NREM-d mice (45, 48 and 56 randomly selected existing spines for ND, NREM-d and REMD mice, respectively). **(k)** The rotarod performance improvement of FW running was significantly lower in REMD mice compared to that in ND and NREM-d mice over hours 0–24 (ND vs. REMD, $P = 0.032$; NREM-d vs. REMD, $P = 0.046$; n = 7, 7 and 6 ND, NREM-d and REMD mice, respectively). Data are presented as mean ± s.e.m. $*P < 0.05$, $**P < 0.01$, $***P < 0.001$, n.s. = not significant. Scale bars, 2 μm.

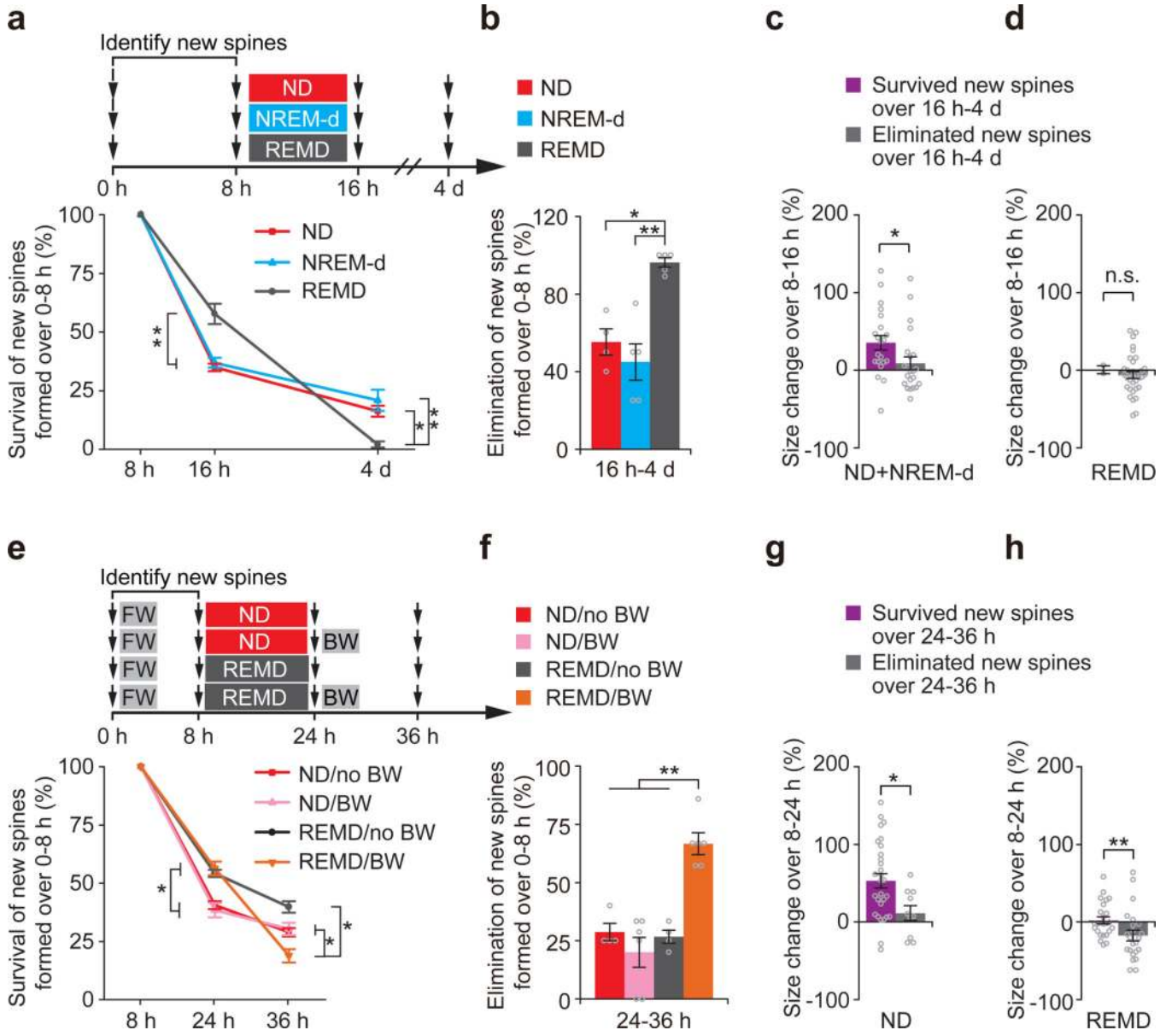

Figure 4. REM sleep facilitates long-term maintenance of new spines during development and learning

(a) Experimental design for examining the long-term survival of new spines formed over hours 0–8 in P21 mice. After undisturbed sleep (ND), NREM-d or REMD for 8 h, new spines identified over hours 0–8 were examined over 4 d. The overall survival rate of new spines formed over hours 0–8 was significantly lower in REMD mice than in ND and NREM-d mice over 4 d (16 h: ND vs. REMD, $P = 0.002$; NREM-d vs. REMD, $P = 0.006$; n = 7, 5 and 7 ND, NREM-d and REMD mice, respectively; 4 d: ND vs. REMD, $P = 0.013$;NREM-d vs. REMD, $P = 0.008$; n = 4, 5 and 5 ND, NREM-d and REMD mice, respectively). **(b)** After 8 h of REM sleep deprivation, a significantly larger percentage of new spines formed between hours 0–8 were eliminated in REMD mice compared to in ND mice or NREM-d mice over 4 d (ND vs. REMD, $P = 0.013$;NREM-d vs. REMD, $P = 0.008$;

n = 4, 5 and 5 ND, NREM-d and REMD mice, respectively). **(c)** New spines persisted over 4 d were strengthened in size between hours 8–16 in ND and NREM-d control mice, whereas new spines eliminated over 4 d were not strengthened during the same period (21 persistent and 22 eliminated new spines from 9 mice). **(d)** New spines persisted or eliminated over 4 d were not strengthened in size over hours 8–16 in REMD mice (2 persistent and 35 eliminated new spines from 5 mice). **(e)** Experimental design for examining the long-term survival of FW-induced new spines. New spines were identified in P30 mice subjected to FW on an accelerated rotarod over hours 0–8. After undisturbed sleep or REMD for 16 h, the animals were subjected to BW or no training, and the survival of the new spines identified over hours 0–8 were examined between hours 24–36. The overall survival rate of new spines formed after FW (hours 0–8) was significantly lower between hours 8–36 in REMD mice than in ND mice after BW (ND + BW vs. REMD + BW, $P = 0.02$; REMD without BW vs. REMD + BW, $P = 0.011$; n = 4, 6, 4 and 6 ND without BW, ND + BW, REMD without BW and REMD + BW mice, respectively). **(f)** After BW training, a significantly larger percentage of new spines formed over hours 0–8 were eliminated between hours 24–36 in REMD mice compared to in ND mice or nontrained mice (ND + BW vs. REMD + BW, $P = 0.004$; REMD without BW vs. REMD + BW, $P = 0.009$; n = 4, 6, 4 and 6 ND without BW, ND + BW, REMD without BW and REMD + BW mice, respectively). **(g)** Persistent new spines but not eliminated new spines over hours 24–36 were strengthened in size between hours 8–24 in ND mice (29 persistent and 10 eliminated new spines from 10 mice). **(h)** New spines persistent or eliminated over hours 24–36 were not strengthened in size between hours 8–24 in REMD mice (24 persistent and 23 eliminated new spines from 10 mice). Data are presented as mean ± s.e.m. $*P < 0.05$, $**P < 0.01$, n.s. = not significant.

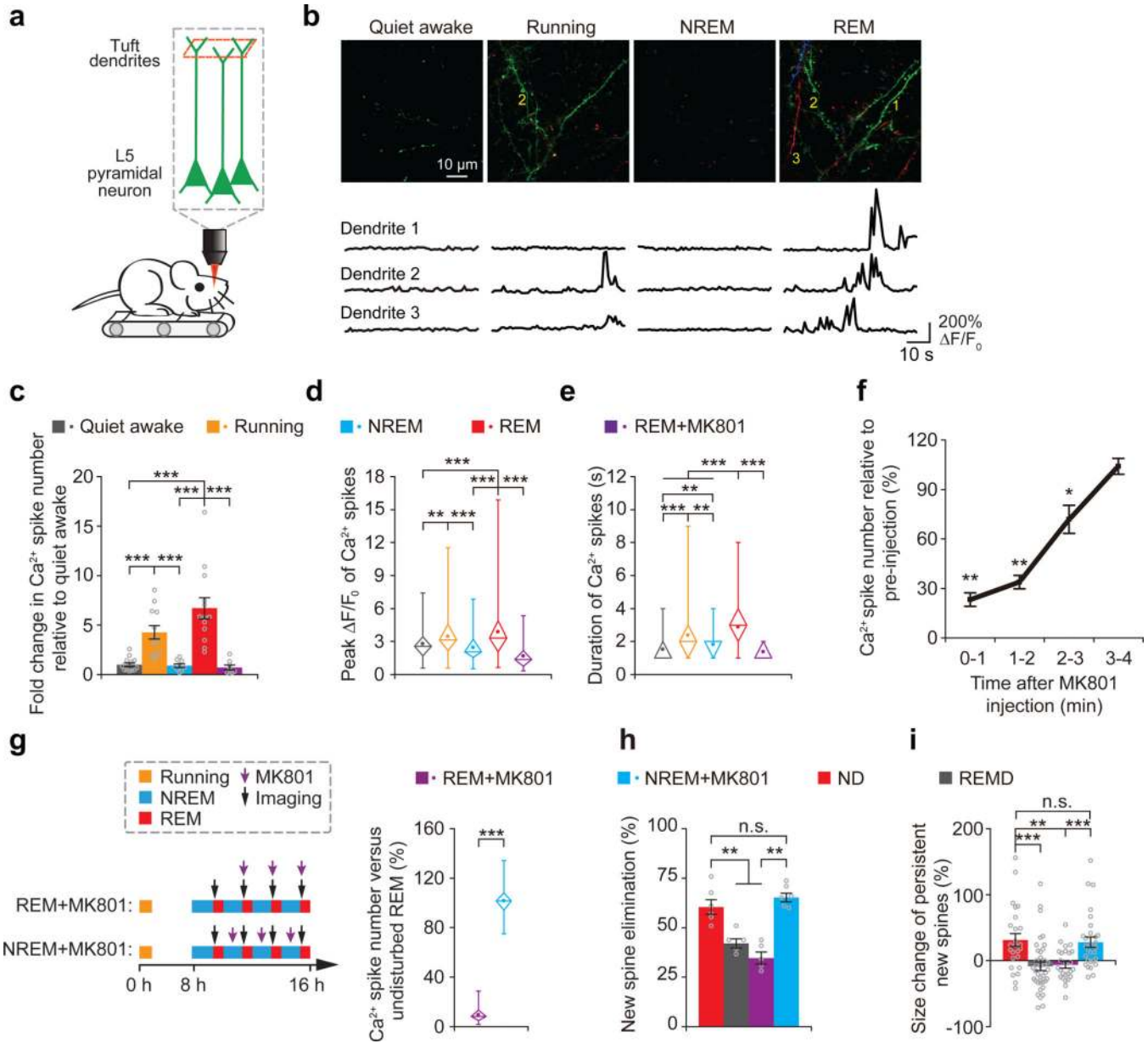

Figure 5. Dendritic Ca^{2+} spikes occurring during REM sleep are important for new spine elimination and strengthening

(a) Two-photon Ca^{2+} imaging of apical tuft dendrites of L5 pyramidal neurons in head-restrained mice on a treadmill during quiet awake, running, NREM sleep and REM sleep. **(b)** Ca^{2+} imaging of apical tuft dendrites under various states. Ca^{2+} fluorescence traces of three dendrites over 1 min are shown. Images at 3 timepoints with a 10-s interval are represented by three different colors. Scale bar, 10 μm. **(c,d)** The number and peak amplitude of dendritic Ca^{2+} spikes during REM sleep were comparable to those during running but significantly larger than those during either NREM sleep, quiet awake state or REM sleep with MK801 application (for number: $P = 1.82 \times 10^{-5}$, 0.082, 9.37×10^{-5} and 1.66×10^{-4} for REM sleep vs. quiet awake, running, NREM sleep and REM sleep + MK801, respectively; n = 13, 12, 9, 13 and 8 quiet awake, running, NREM, REM and REM

+ MK801 mice, respectively; for peak amplitude: $P = 2.82 \times 10^{-4}$, 0.373, 9.09×10^{-8} and 1.77×10^{-12} for REM sleep vs. quiet awake, running, NREM sleep and REM sleep + MK801, respectively; n = 89, 124, 88, 274 and 47 spikes for quiet awake, running, NREM, REM and REM + MK801, respectively). **(e)** The durations of dendritic Ca^{2+} spikes during REM sleep were significantly larger than those during other states ($P = 3.01 \times 10^{-17}$, 0.0007, 1.88×10^{-10} and 1.47×10^{-13} for REM sleep vs. quiet awake, running, NREM sleep and REM sleep + MK801, respectively). **(f)** Brief injection of MK801 (3 pulses, 50 ms each) into the primary motor cortex blocked dendritic Ca^{2+} spikes during quiet awake over the next 2–3 min (n = 10 mice). **(g)** More than 90% of dendritic Ca^{2+} spikes during REM sleep were blocked after pulsed injection of MK801 at the beginning of each REM sleep episode but not during NREM sleep (n = 46 episodes of REM sleep with MK801 injection from 4 mice; 51 episodes of NREM sleep with MK801 injection from 4 mice). **(h)** Injections of MK801 during REM sleep but not during NREM sleep, reduced the elimination rate of learning-induced new spines (ND vs. REM + MK801, $P = 0.006$; ND vs. NREM + MK801, $P = 0.297$; REM + MK801 vs. NREM + MK801, $P = 0.006$; n = 6, 5, 5 and 6 mice for ND, REMD, REM + MK801 and NREM + MK801, respectively). **(i)** Injection of MK801 during REM sleep, but not during NREM sleep, reduced the size increase of persistent new spines formed after treadmill training (ND vs. REM + MK801, $P = 0.0019$; ND vs. NREM + MK801, $P = 0.801$; REM + MK801 vs. NREM + MK801, $P = 0.0005$; n = 25, 40, 26 and 29 new spines from 6, 5, 5 and 6 mice for ND, REMD, REM + MK801 and NREM + MK80, respectively). Data are presented as mean ± s.e.m. In **d, e and g**, box and whisker plots show the means (central dot), medians (central line in the box), ranges between 25th and 75th percentiles (box) and minimum–maximum range (whiskers). Each point in c and h represents data from one animal. $^{**}P < 0.01$, $^{***}P < 0.001$, n.s. = not significant.

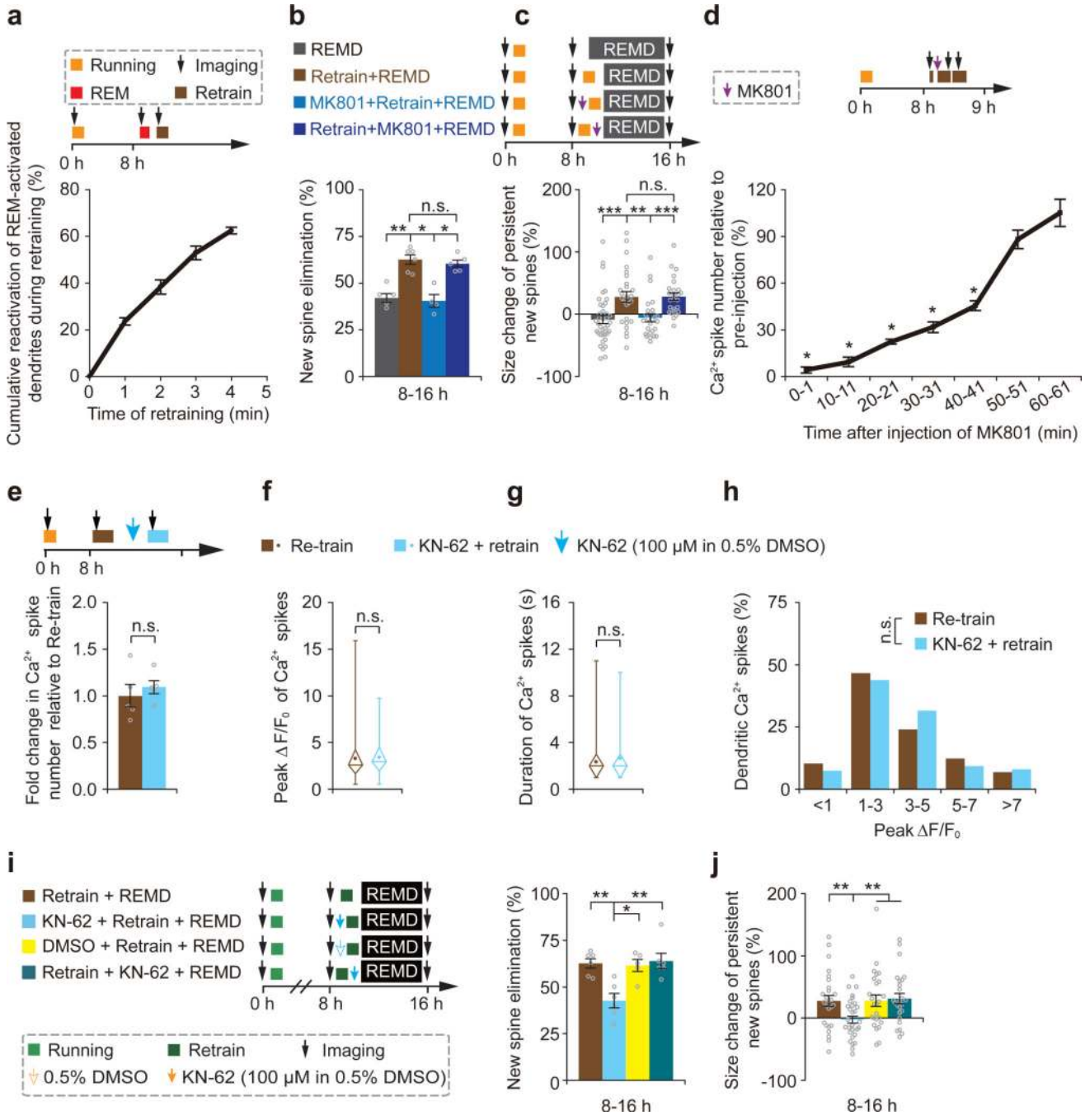

Figure 6. Motor training-induced Ca^{2+} spikes prune and strengthen new spines in REMD mice

(a) Majority of dendrites exhibiting Ca^{2+} spikes during 1-min REM sleep also showed Ca^{2+} spikes during 4-min retraining (n = 4 mice). **(b)** 40-min retraining led to increased elimination of new spines in REMD mice. Injection of MK801 before retraining, but not after retraining, blocked retraining-induced elimination of new spines (n = 5, 6, 4 and 5 mice for REMD, retraining + REMD, MK801 + retraining + REMD, and retraining + MK801 + REMD, respectively). **(c)** 40-min retraining led to the potentiation of persistent new spines in REMD mice. Injection of MK801 before retraining, but not after retraining, blocked

retraining-induced strengthening of persistent new spines (40, 26, 25 and 25 new spines from 5, 6, 4 and 5 mice for REMD, retraining + REMD, MK801 + retraining + REMD, and retraining + MK801 + REMD, respectively). **(d)** The number of retraining-induced dendritic Ca^{2+} spikes was reduced substantially for ~40 min after repeated injections of MK801 (10 pulses) at the beginning of retraining (n = 6 mice). **(e–h)** The number, average peak amplitude, duration and peak amplitude distribution of dendritic calcium spikes over a period of 1 min were not affected by local application of KN-62 (n = 5 mice). **(i)** Application of KN-62 before retraining, but not after retraining, blocked retraining-induced elimination of new spines in REMD mice (retraining vs. KN-62 + retraining, $P = 0.008$; KN-62 + retraining vs. retraining + KN-62, $P = 0.006$; n = 6, 6, 5 and 6 mice for retraining, KN62 + retraining, DMSO + retraining and retraining + KN62, respectively). **(j)** Application of KN-62 before retraining, but not after retraining, blocked retraining-induced strengthening of persistent new spines (retraining vs. KN-62 + retraining, $P = 0.007$; KN-62 + retraining vs. retraining + KN-62, $P = 0.003$; n = 26, 31, 25 and 26 new spines from 6, 6, 5 and 6 mice for retraining, KN62 + retraining, DMSO + retraining and retraining + KN62, respectively). In **f and g**, box and whisker plots show means (central dot), medians (central line in the box), ranges between 25th and 75th percentiles (box) and minimum–maximum ranges (whiskers). Data are presented as mean ± s.e.m. Each point in b, e and i represents data from one animal. $*P < 0.05$, $**P < 0.01$, $***P < 0.001$, n.s. = not significant.