PNAS
www.pnas.org

Supplementary Information for

Emergence of stable striatal D1R and D2R neuronal ensembles with distinct firing sequence during motor learning

Meng-jun Sheng, Di Lu, Zhi-ming Shen, Mu-ming Poo

Correspondence should be addressed to Mu-ming Poo

Email: mpoo@ion.ac.cn

This PDF file includes:

Materials and Methods
Figs. S1 to S13
Tables S1
Captions for movies S1 to S5

Other supplementary materials for this manuscript include the following:

Movies S1 to S5

www.pnas.org/cgi/doi/10.1073/pnas.1901712116

Materials and Methods

Animals. Animal use procedures were approved by the Animal Use Committee of the Institute of Neuroscience, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences. Calcium imaging and chemogenetic suppression experiments were performed on D1-Cre mice (B6.FVB(Cg)-Tg(Drd1a-cre) EY262Gsat/Mmcd, Stock Number: 030989-UCD) for targeting the direct pathway SPNs and D2-Cre mice (B6.FVB(Cg)-Tg(Drd2-cre) ER44Gsat/Mmcd, Stock Number:032108-UCD) for targeting the indirect pathway SPNs (>7 weeks, male and female). To verify the expression pattern of D1- and D2-Cre mice, we crossed the D1- and D2-Cre mice with Ai9 (B6.Cg-Gt(ROSA)26Sortm9(CAG-tdTomato)Hze/J). The inactivation experiments using muscimol were performed on adult C57BL/6J mice (Slac Laboratory Animals, Chinese Academy of Sciences) and D1-Cre mice. Sample size is described in the results or figure legends.

Virus injection. Adult mice (>7 weeks, male or female) were anesthetized with isoflurane (1.5%) and placed in a stereotaxic frame (RWD). Ophthalmic ointment was applied to prevent eye drying and body temperature was maintained at 37℃ using a heating pad (Harward). The mouse head was cleaned with 75% ethanol and then shaved. A mid-line scalp incision was made to expose the skull. A small hole was drilled at the coordinate (AP: 0.30 mm, ML: 3.75-4.00 mm). Cre-dependent virus (AAV-FLEX-GCaMP6s) was pressure-injected at an angle of 25 degree using an electrically gated pressure injector (Picospritzer III, 3-6 psi) at two sites at a depth between 1.60 and 1.85 mm under the pia with a total volume of ~ 500 nl. The skin was sutured and mice returned to home cage after recovery from anesthesia under a heating lamp.

Surgery. 3-10 days after virus injection, mice were re-anesthetized with isoflurane (1.5%) then mounted in a stereotaxic frame. After cutting the scalp, a titanium head plate was glued to the skull. A ~2.4mm circular craniotomy was performed above the dorsal lateral striatum, centered at the virus injection coordinate (AP: 0.3 mm, ML: 3.75-4.0 mm). The cortical tissue was aspirated by a 27-gauge needle along a 30 degree slowly to the surface of corpus callosum (1). A thin layer of Kwik-Sil (WPI) was applied and a cannula was inserted

along a 30 degree. The cannula consists of a stainless-steel tube and a round coverslip attached to one tube end with adhesive (Norland optical adhesive). The gap between the tube and skull was filled with Kwik-Sil, then the cannula and exposed skull were cemented with dental acrylic.

Rightward cued lever-pushing task. Two weeks after surgery, mice (20 D1-Cre, 20 D2-Cre) were water-restricted at 1 ml per day. After three days of habituation, during which mice received water from the water port, the task training under the two photon microscope was initiated. The lever-pushing setup consists of a custom-made 3D-printed plastic frame, potentiometer, and repulsive magnet pairs. The lever position was monitored by the potentiometer which transformed the lever displacement to the voltage signal, which was recorded using a PCI-6023E card (National Instrument). The repulsive magnet pairs were used to keep the lever at a relatively stable position in the resting state and provide the resistance for the active motor action by the mouse. The training paradigm was running on a custom-made program written with LabVIEW, controlling the timing to present the cue, reward, punishment and determine the threshold displacement for lever pushing. Each trial began with a 500 ms, 6-kHz pure tone. After cue onset, mice that pushed the lever across the set threshold (1 mm at the beginning of training on day 1 and increased to 3 mm after the mice learned to push the lever) received a drop of water (approximately 8 μl) as the reward within the task period. No pushing or failure to push the lever exceeding the threshold during the task period triggered a loud white noise. The inter-trial interval (ITI) was 4 s or random timing between 4 and 6 s, each lever pushing in the ITI caused additional time-out equal to the ITI for that trial. The task period was 30 s in the first session and 10 s in the following sessions.

Leftward cued lever-pushing/Downward cued lever-pressing task. Mice (8 D1-Cre, 10 D2-Cre) that had learned the rightward lever-pushing task were subjected to the lever device which was designed for unidirectional leftward-pushing or downward-pressing in the subsequent days. Both new tasks involved the same forelimb and shared the same task structure (cue, reward and ITI) with the rightward cued lever-pushing task as mentioned above, except for the direction of lever movement.

Dual-task experiment. Two weeks after surgery, mice (5 D1-Cre mice, 9 D2-Cre mice) were water-restricted at 1 ml per day. After three days of habituation, during which mice received water from the water port, the dual-task training under the two photon microscope was initiated. Mice were subjected on the same day to two tasks in tandem throughout learning, the first task was cue-reward without lever-pushing in which the lever device was replaced with a platform for the mice forelimbs to rest. Water reward (approximately 8 µl) was delivered after a random delay of 1-2.5 s following the onset of cue (500 ms, 6-kHz pure tone). The inter-trial interval (ITI) was 4 s or random timing between 4 and 6 s. After 30-40 trials of the cue-reward task, mice were subjected to the rightward cued lever-pushing task.

Passive Rolling Task training. Mice (9 D1-Cre, 10 D2-Cre) that had learned the rightward lever-pushing task were subjected to a cylinder on which mice put their forelimbs on. The cylinder was passively rotated clockwise with a speed of about 30.5 cm/s by an electric motor. Mice had to move their forelimbs with the rolling cylinder in order to keep balance. The time to initiate the cylinder was random during imaging, and the duration was from 2.15 ± 1.28 s (mean ± std). Mice were returned to the right cued lever-pushing task on the subsequent day.

Chemogenetic inhibition. Cre-dependent virus (AAV-EF1a-DIO-hM4D-mCherry) was pressure-injected at an angle of 25 degree using an electrically gated pressure injector (Picospritzer III, 3-6 psi) at a depth of 2 mm under the pia (AP: 0.30 mm, ML: 3.75-4.00 mm) with a total volume of ~1 µl in adult D1- and D2-Cre mice (>7 weeks, male or female). Three weeks after virus injection, mice were trained to perform the cued lever-pushing task. After the mice fully learned the task (consecutively three days reached the criteria), CNO (1mg/kg) was injected intraperitoneally 0.5-1.5 hour before performing the lever pushing task. Virus expression in DLS in each mouse was verified post-mortem. Vehicle (saline + 4% DMSO) was used in the control experiment.

Muscimol inactivation. Pharmacological silencing of the dorsolateral striatum (DLS) was performed by injecting muscimol into DLS using a syringe pump (Kent Scientific Corporation). The infusion guide

cannula (Plastics-one, C315T) was implanted into DLS (AP, 0 mm; ML, 3.75-4 mm; Depth, 1.7mm). Mice were trained to become an expert with a stable and high success rate. On the injection day, muscimol was diluted with sterilized saline into a 5 μg/μl solution and a total volume of 70 nl was infused into DLS 30 min before performing the lever pushing task. Saline was used in the control experiment. Each cannula position was verified post-mortem.

Two-photon calcium imaging. Imaging was performed in awake mice on a commercial two-photon microscope (Bscope, Thorlabs) operated with Thorimage software. We used a 16× objective (NIKON, NA = 0.8), the excitation wavelength was 910 nm or 925 nm (Ti-Sa laser, Spectra-physics). A CED power1401 was used to record the image frame-out timing and behavior events (cue, reward, punishment, licking and lever displacement). Time-lapsed movies were acquired at approximately 15 Hz for approximately 2 min. Size of field of view (FOV) ranged from 500 × 500 to 1108 × 1108 μm (1024 ×1024 pixels).

Image analysis. Lateral motion artifacts associated with calcium imaging movies were corrected using imageJ plugin Turboreg (2) frame-by-frame. Neurons in FOV were manually drawn using custom MATLAB code session-by-session. Only neurons that could be identified in all sessions by visual inspection were selected and further analyzed. Neurons without clear ring-shaped GCaMP6s expression were excluded from analysis.

Fluorescence analysis. For each neuron, we averaged fluorescence intensity of all labeled pixels and obtained the raw fluorescence trace. Neurons were classified into three groups according to the following criteria as reported previously (3). First, we calculated the probability distribution function (PDF) of each neuron's raw fluorescence trace and set the peak value as 0 (minus the peak value in PDF). A cutoff value was defined as the point above which 90% of the points below PDF's peak was located. A threshold was set as two times the absolute cutoff value. Cells were classified as 'moderately active' if more than 1% of their fluorescence was above the threshold, and 'highly active' were classified if more than 10% of their fluorescence was above threshold, if less than 1% of their fluorescence was above the threshold, cells were

classified as 'sparsely active'. F_0 was determined similar to a previous study (4). For "highly active" cells, F_0 was the 5th percentile of raw fluorescence. For "moderately active" cells, the 30th percentile of raw fluorescence within the window was calculated as F_0. For sparsely active cells, the median value was used as F_0. $\Delta F/F$ was calculated as $(F-F_0) / F_0$.

Identification of responsive neurons in rightward lever-pushing task. Neurons were defined as "responsive" cells by the following criteria modified from a previous study (5). The lever-pushing duration is 0.37 ± 0.31 s (mean ± std) in D1-Cre mice and 0.34 ± 0.25 s (mean ± std) in D2-Cre mice. The lever-pushing duration was defined as 0 - 0.67 s (approximately the duration of lever pushing for all D1- and D2-Cre mice) after lever-pushing movement onset, and 0.67 - 3 s after movement onset was defined as post-movement period. Thus all cells were preliminarily classified into two groups according to time of the averaged peak activity (averaging all trials on each day). For cells whose peak position was in -0.13 to 0.67 s relative to movement onset, the responsive zone was defined as -0.13 to 0.67 s relative to movement onset and baseline zone was -1 to -0.13 s relative to movement onset. In this group of cell, we tested each frames of all trials in one day in responsive region with baseline, the baseline was all the frames in baseline zone of all trials. If at least 50% frames were significantly ($P < 0.05$, Wilcoson rank sum test) modulated and the mean $\Delta F/F$ difference between responsive and baseline regions was larger than 0.5, then those cells were classified as movement responsive cells. For cells whose peak position was within 0.67 to 3 s after movement onset, the responsive zone was defined as 0.67 to 3 s relative to movement onset and the baseline zone was -1 to 0.33 s relative to movement onset. In this group of cell, we tested each frames of all trials in one day in responsive region with baseline, the baseline was all the frames in baseline zone of all trials. If at least 20% frames were significantly ($P < 0.05$, Wilcoson rank sum test) modulated and the mean $\Delta F/F$ difference between responsive and baseline regions was larger than 0.3, then those cells were classified as post-movement responsive cells. Responsive cells were the summation of movement responsive and post-movement responsive cells.

Identification of cue, reward, movement, post-movement related cells in the dual-task. For those mice that performed the dual-task (cue-reward task without lever-pushing together with cued lever-pushing task), we classified cells response to cue and reward as following criteria: first, for both cue and reward cells, the baseline zone was defined as -1 to -0.13 s relative to cue onset or reward delivery, and responsive zone was 0 to 1 s relative to cue onset or reward delivery. Then each frame of all trials in one day in responsive region was statistically tested with baseline, the baseline was all the frames in baseline zone of all trials. If at least 50% frames were significantly ($P < 0.05$, Wilcoson rank sum test) modulated and the mean $\Delta F/F$ difference between responsive and baseline regions was larger than 0.5, then those cells were defined as "Cue cells" or "Reward cells". After subtracting the "cue cells" and "reward cells" from all responsive neurons identified in the lever-pushing task, those responsive cells related to lever pushing were then classified into two groups: fired during the lever-pushing period (with peak time within -0.13 to 0.67 s relative to movement onset; i.e., movement-related "M cells") and after the pushing action (with the peak time within 0.67 to 3 s relative to movement onset; i.e., post-movement "PM cells").

Activity correlation analysis. For calculating the correlation coefficient between trials or the averages patterns between days, the temporal fluorescence traces of different cells (1 s before and 3 s after rewarded movement onset) were concatenated to a one dimension array (6), and this vector was then used to calculate the Pearson correlation coefficient using MATLAB function 'corrcoef' . Since our task there was a 4-6 s of inter-trial interval (ITI), in which lever-pushing was punished with additional timeout, there was no lever-pushing for 1 s before task onset, allowing the assessment of baseline fluorescence. We thus take a time window of 1 s before and 3 s after the rewarded movement onset for the analysis of the neural activity during the movement and post-movement periods.

Movement analysis. Movement bouts were identified in the lever displacement traces (voltage recordings from the potentiometer) at 5 kHz sampling rate. Bouts separated by less than 500 ms were considered continuous, and then the start and end times were identified as follows. The start time was defined by

finding when the lever position crossed a threshold exceeding the resting period before the movement, and the end time was defined by finding the time when the lever position went below a threshold value defined by the resting period following the movement. For trial-based analysis, the trials in which animals were moving the lever within 3 seconds before the onset of cue were excluded.

Statistics. Wilcoson rank sum test and one-way ANOVA were used to test the significance. The sample size was not predetermined, similar to commonly accepted practice in this area.

References

1. Dombeck DA, Harvey CD, Tian L, Looger LL, & Tank DW (2010) Functional imaging of hippocampal place cells at cellular resolution during virtual navigation. *Nat Neurosci* 13(11):1433-1440.
2. Thevenaz P, Ruttimann UE, & Unser M (1998) A pyramid approach to subpixel registration based on intensity. *IEEE transactions on image processing* 7(1):27-41.
3. Huber D*, et al.* (2012) Multiple dynamic representations in the motor cortex during sensorimotor learning. *Nature* 484(7395):473-478.
4. Peron SP, Freeman J, Iyer V, Guo C, & Svoboda K (2015) A Cellular Resolution Map of Barrel Cortex Activity during Tactile Behavior. *Neuron* 86(3):783-799.
5. Chu MW, Li WL, & Komiyama T (2016) Balancing the Robustness and Efficiency of Odor Representations during Learning. *Neuron* 92(1):174-186.
6. Peters AJ, Chen SX, & Komiyama T (2014) Emergence of reproducible spatiotemporal activity during motor learning. *Nature* 510(7504):263-267.

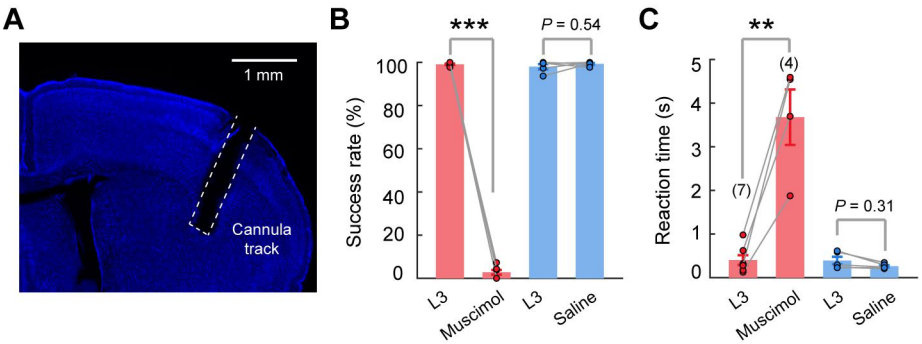

Fig. S1 (related to Fig. 1)

Dorsolateral striatum is required for performing the learned motor task. (*A*) Fluorescence image of a brain section from a mouse showing the location of cannula implantation for muscimol injection into the DLS (white dashed lines: cannula track). (*B*, *C*)

Effects of muscimol injection on the success rates (% of rewarded trials) and reaction times (from cue onset to rewarded movement) in performing the trained cued lever-pushing task (n = 7 mice, except those indicated in the parenthesis, red bars, ***P < 0.001, **P < 0.001, Wilcoxon rank sum test). Error bars, s.e.m. Gray lines connect data from the same mouse. Saline injection instead of muscimol injection was used as a control (n = 5, blue bars, Wilcoxon rank sum test).

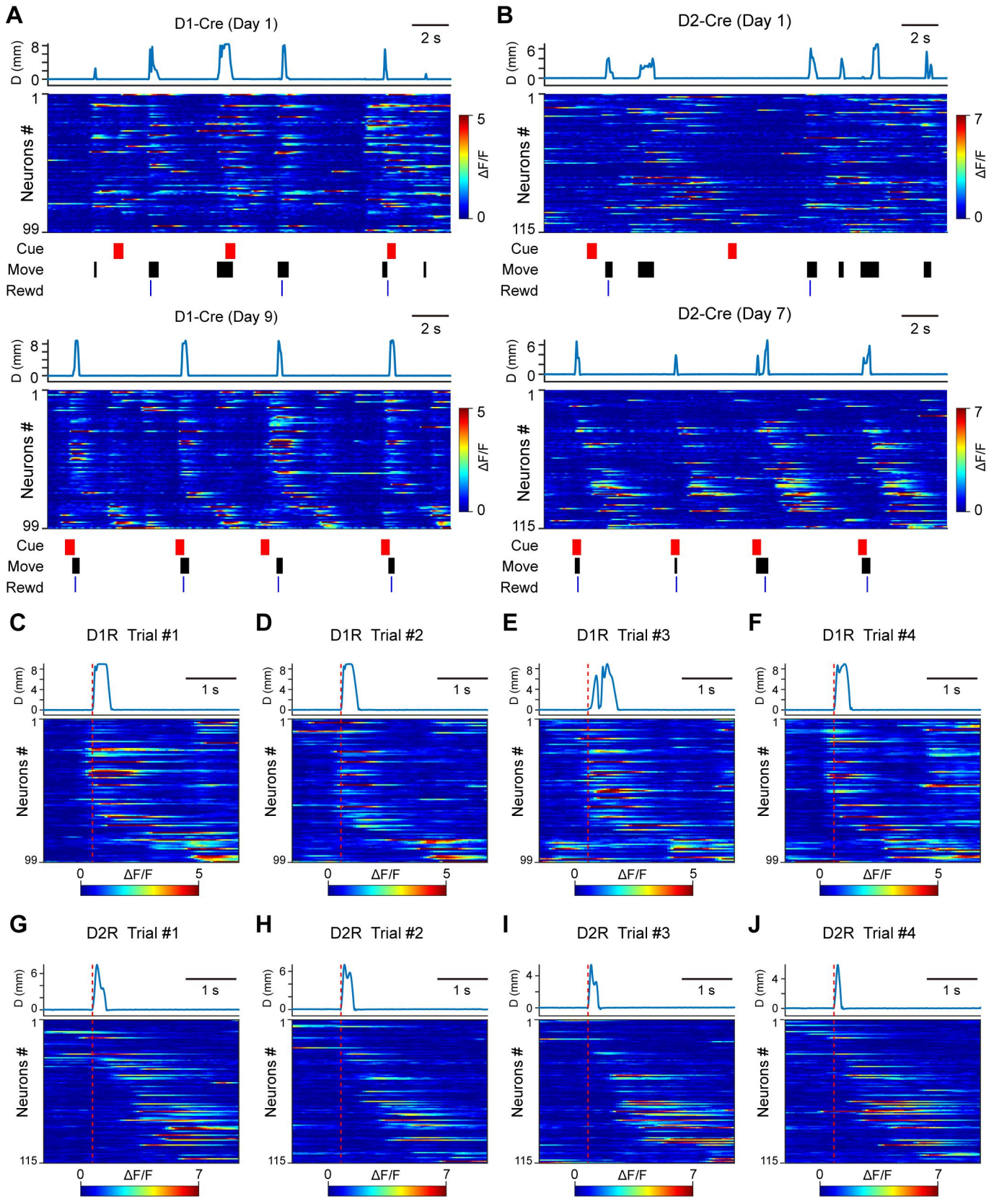

Fig. S2

Examples of single-trial activity of neuronal ensembles after learning cued lever-pushing task. (*A*, *B*) Samples of consecutive trials, with Ca^{2+} transients of the same populations of D1R neurons (*A*, 99 cells) and D2R neurons (*B*, 115 cells) recorded on the first (upper) and last (lower) day of training (ordered by the time of average peak activity on the last day), with simultaneous recordings of lever movement trajectory (D: distance of lever displacement) and behavioral events. Red, "Cue"; black, lever movement ("Move"); blue, reward ("Rewd"). (*C*-*J*)Four examples of single-trial activity depicted by heatmap of ΔF/F, for the same D1R (*C*-*F*, 99 cells) and D2R (*G*-*J*, 115 cells) neuron ensembles, together with associated lever movement trajectories (D: Distance of lever displacement). Red dashed line: lever-pushing movement onset.

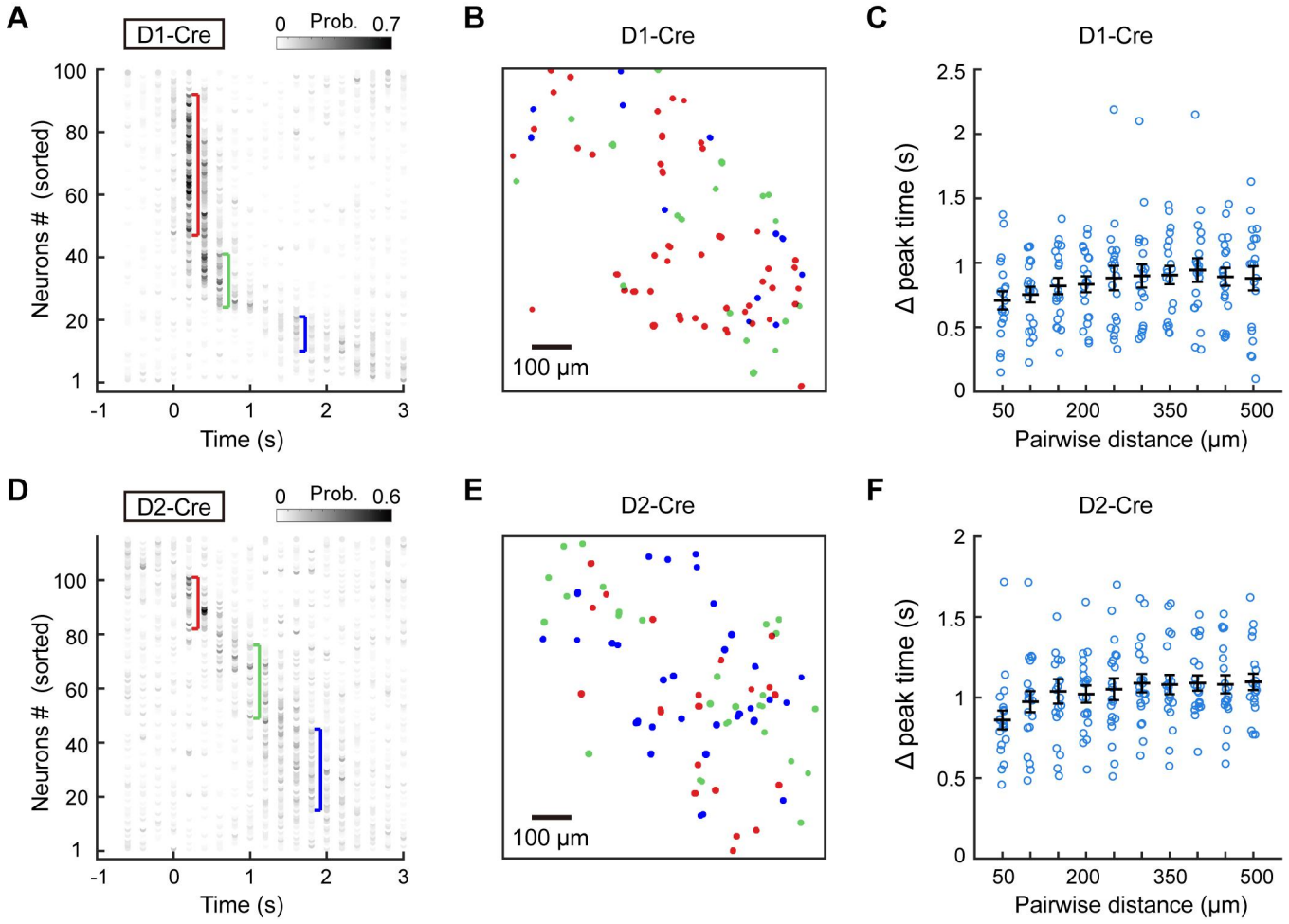

Fig. S3 (related to Fig. 3*A* and *B*)

Emergence of temporally specific and spatially dispersed neuronal ensembles. (*A*, *D*) Cell firing probability map for each time zone on L3 in an example D1-Cre mouse (*A*, 99 neurons) and an D2-Cre mouse (*D*, 115 neurons). Cells were ordered by the time of average peak activity. The 4-sec time window was binned into 20 time zones (200 ms per zone). The gray scale indicates the firing probability of an individual neuron within each time zone in all trials on that day. Considering the slow decay of calcium signal due to the previous trial, activity in the first time zone was excluded from the analysis. Three ensembles (marked by the red, green, and blue brackets) were chosen as examples for their higher firing probability within a specific time zone. (*B*, *E*) Spatial distribution of the three ensembles chosen in *A* and *D*, marked by the red, green and blue colors (corresponding to the color of the bracket) for the example D1-Cre and D2-Cre mouse. (*C*, *F*) Pair-wise difference in the time of peak activity plotted against the distance between each neuron pair (*C*, D1-Cre, n = 20, $P = 0.54$; *F*, D2-Cre, n = 20, $P = 0.15$; One-way ANOVA). Horizontal lines indicate the averages of all mice. Error bars, s.e.m. Note that the appearance of neuronal ensembles fired with higher firing probabilities at a specific time zone is in line with the sequential firing of neurons when they were sorted by the time of their

average peak activity. Quantitative analysis of differences in the activity peak times of all neuron pairs did not show significant difference among neuronal pairs with different distances, in both D1 and D2-Cre mice. This supports the idea that neurons with similar activity patterns are not distributed in clusters.

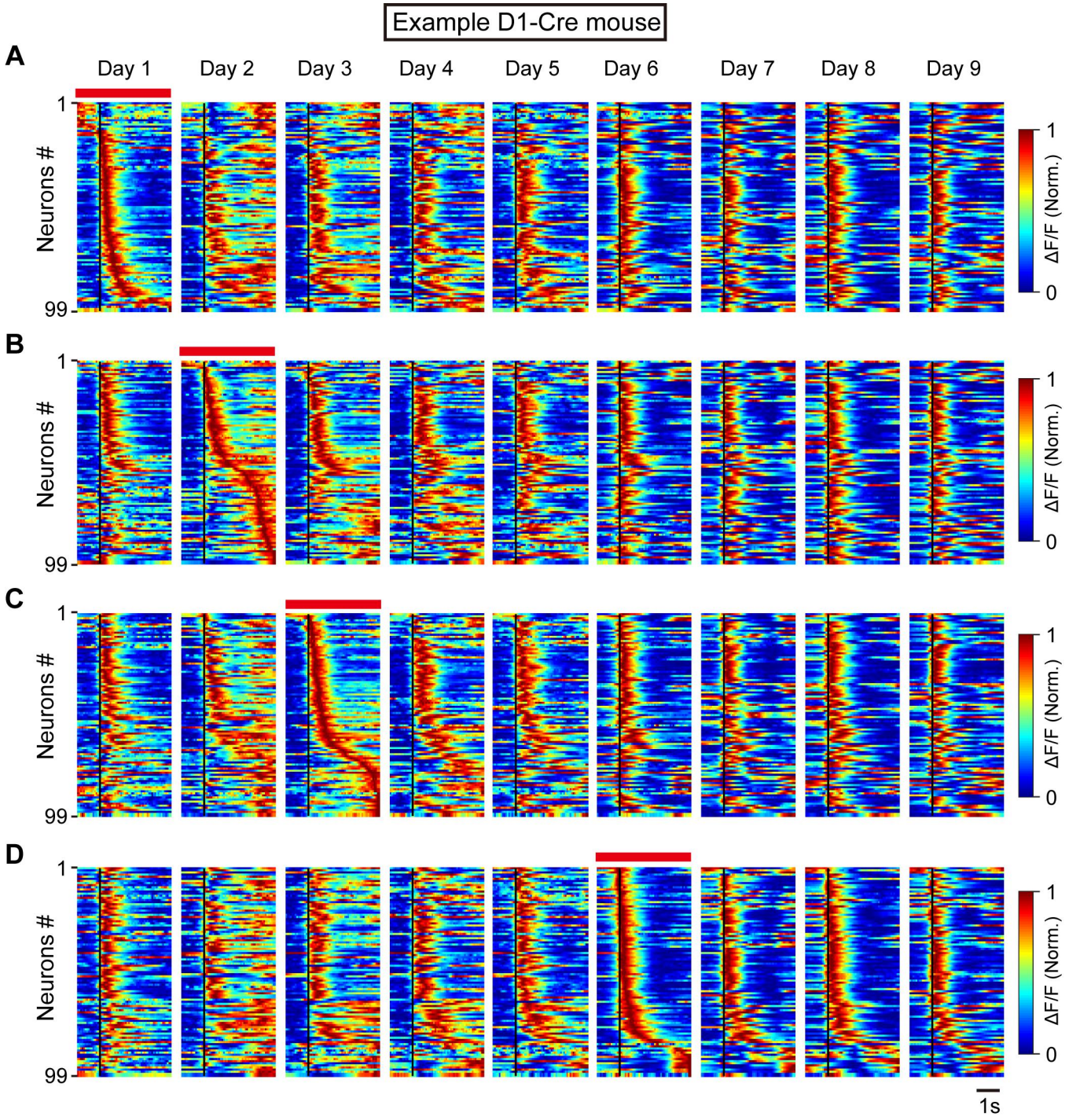

Fig. S4 (related to Fig. 3*A*)

D1R neuronal ensemble activity of an example mouse, sorted by the order of average peak activity on different days (marked by red bar on the top). (*A*-*D*) Averaged activity of the same ensemble of all responsive neurons (99 cells) from day 1 to 9 of training, ordered by the firing sequence on day 1 (*A*) , day 2 (*B*), day 3 (*C*) and day 6 (*D*). Black line: lever-pushing movement onset. Note that sequence sorted by the order on the early stage (day 1-3) was not stable, but that on the late stage (day 6) was stable.

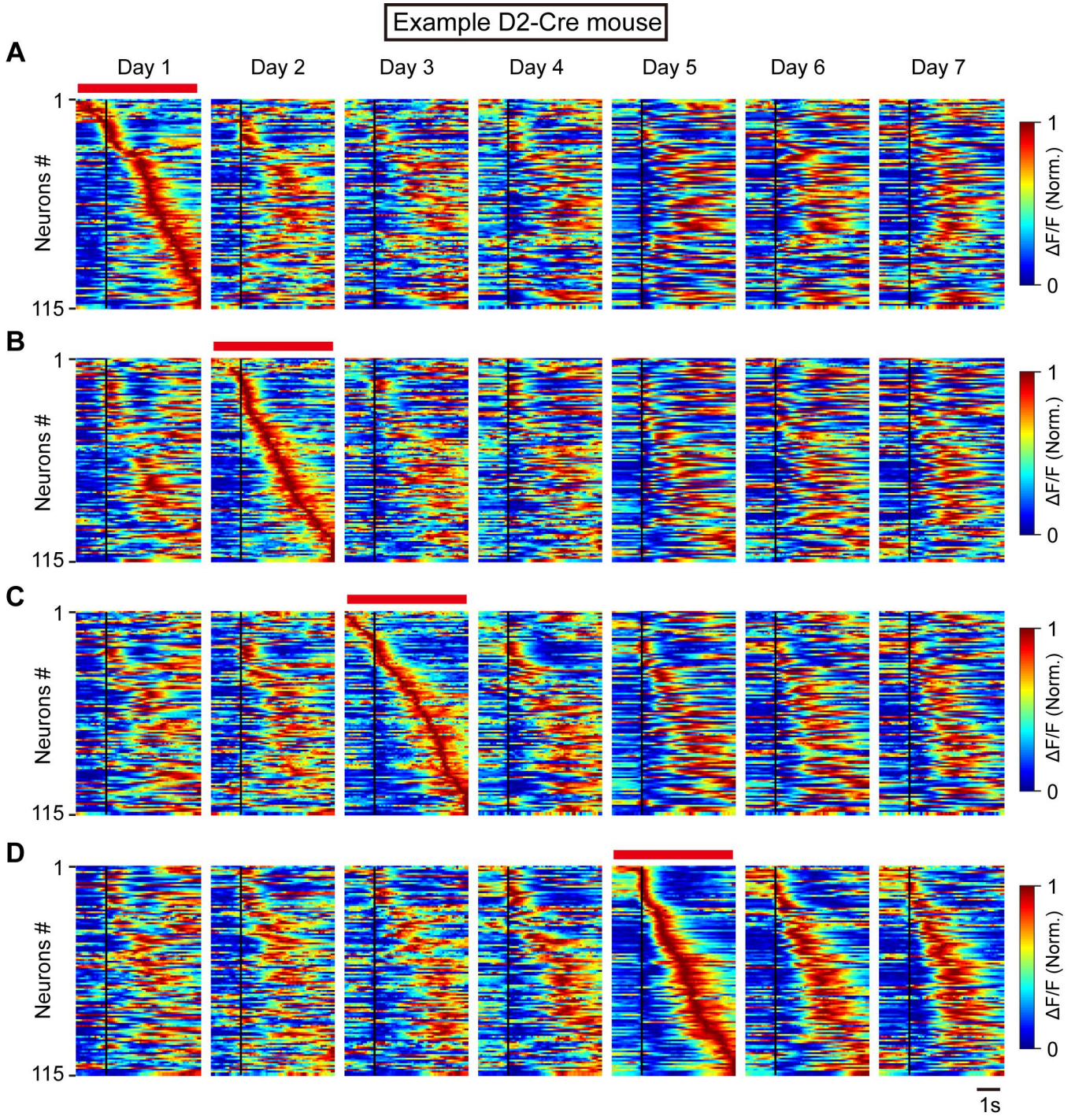

Fig. S5 (related to Fig. 3*C*)

D2R neuronal ensemble activity of an example mouse, sorted by the order of average peak activity on different days (marked by red bar on the top). (*A*-*D*) Averaged activity of the same ensemble of all responsive neurons (115 cells) from day 1 to 7 of training, ordered by the firing sequence on day 1 (*A*) , day 2 (*B*), day 3 (*C*) and day 5 (*D*). Black line: lever-pushing movement onset. Note that sequence sorted by the order on the early stage (day 1-3) was not stable, but that on the late stage (day 5) was stable.

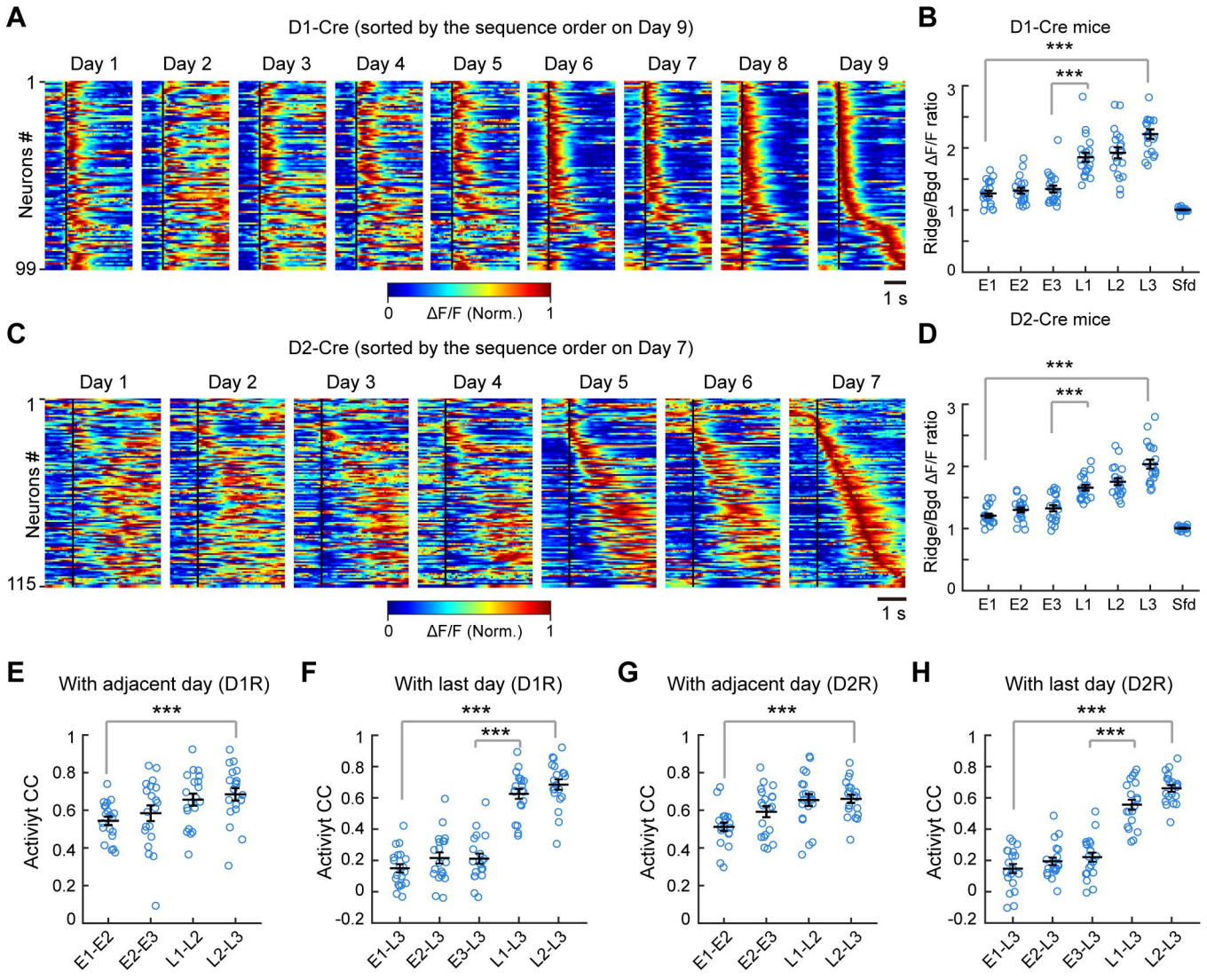

Fig. S6 (related to Fig. 3)

Changes of activity patterns and ridge-to-background ΔF/F ratio (for sequence firing) during motor learning. Data presented in the same manner as in Fig. 3, except that the same total number of trials (rather than all trials collected, the minimum number of trials on all days were used for analysis, if number of trials was larger, then the front part were used) throughout learning was used in the analysis. (*A*, *C*) Averaged activity of all responsive neurons (aligned by the movement onset time, black line) from day 1 to last day of training (D1R, 99 cells, 45 trials for each day; D2R, 115 cells, 19 trials for each day). (*B*, *D*) The ridge-to-background ("Ridge/Bgd") ΔF/F ratio during the early and late stage. (D1-Cre, n = 20; D2-Cre, n=20; ***$P < 0.001$, Wilcoxon rank sum test). Sfd: shuffled data. (*E*, *G*) Similarity of activity patterns quantified by the correlation coefficient (CC) of activity of the same population of neurons between adjacent days. (D1R, n = 20; D2-Cre, n=20; ***$P < 0.001$, Wilcoxon rank sum test). (*F*, *H*) Similar analysis as in *E* and *G*, except the CCs were calculated by comparing the activity on the last L3 (D1R, n = 20; D2-Cre, n=20; ***$P < 0.001$, Wilcoxon rank sum test). Error bars indicate s.e.m. Note that our results were not caused by differences in the total number of trials used in different days, since the same conclusion could be reached when the same total trial number was used throughout the days.

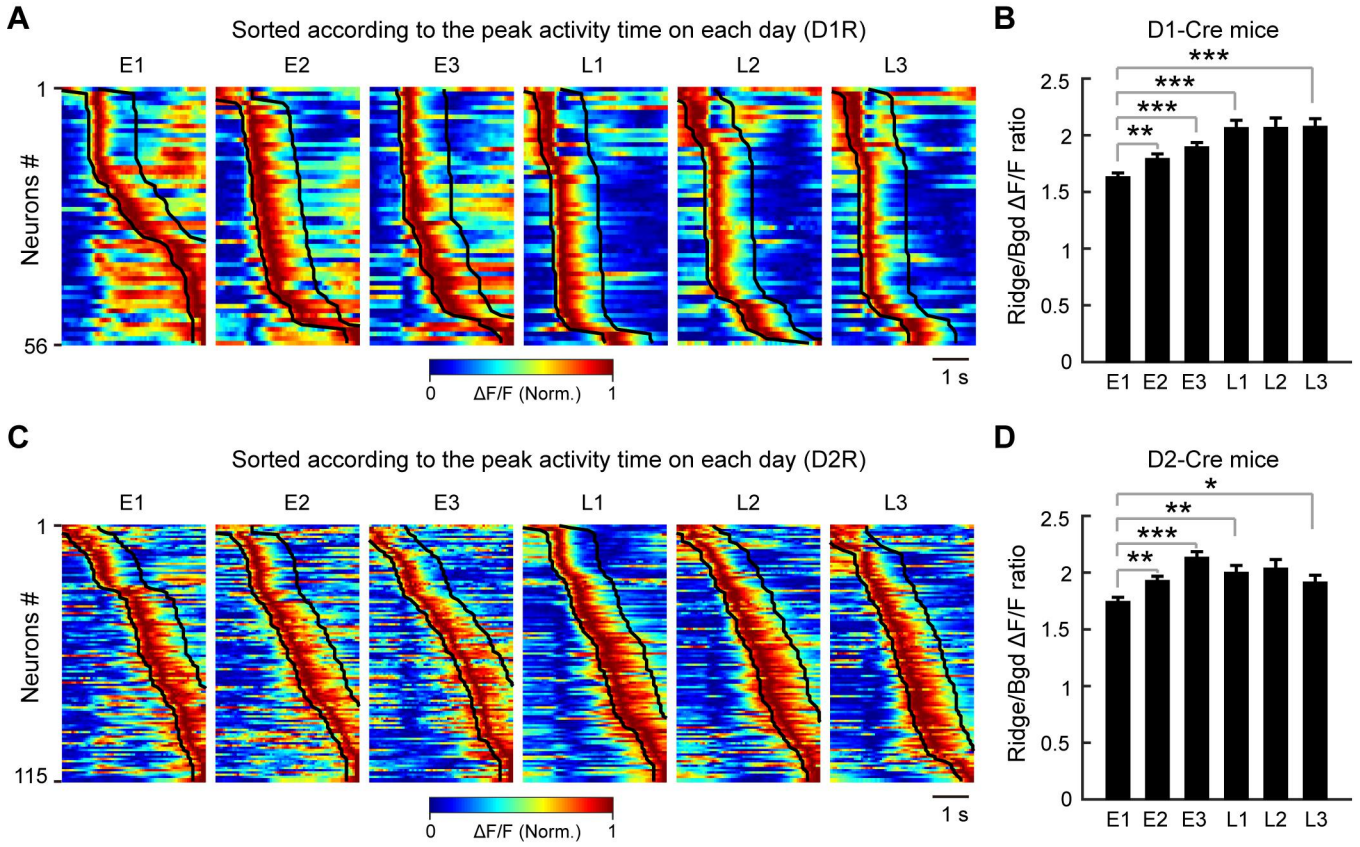

Fig. S7 (related to Fig. 3)

Ridge to background ΔF/F ratio calculated using the ridge defined by firing sequence order on each day. (*A*, *C*) Population activity on E1-E3 and L1-L3 of one example D1-Cre (*A*) and D2-Cre (*C*) mouse. Cells were sorted by the order of the time of average peak activity on each day, with the ridge area marked by the black lines. The ridge was defined as 5 frames (0.33 s, sampling rate: 15 frames per second) before and 15 frames (1 s) after the movement onset, and the rest of the frames were taken as the background. (*B*, *D*) The ridge-to-background ΔF/F ratio ("Ridge/Bgd") of each cell during the early and late stages calculated as the mean ΔF/F value of the ridge frames divided by that of the background frames. The ratio on L3 was significantly higher than that on E1 , and showed a trend of gradual increase during E1-E3 (*B*, D1R, n=20 mice; *D*, D2R, n=20 mice. $*P < 0.05$, $**P < 0.01$, $***P < 0.001$, Wilcoxon rank sum test). Error bars indicate s.e.m. Note that the ratio showed a trend of gradual increase during E1-E3, consistent with a reorganization of neuronal firing pattern with increasing reduction of background activity.

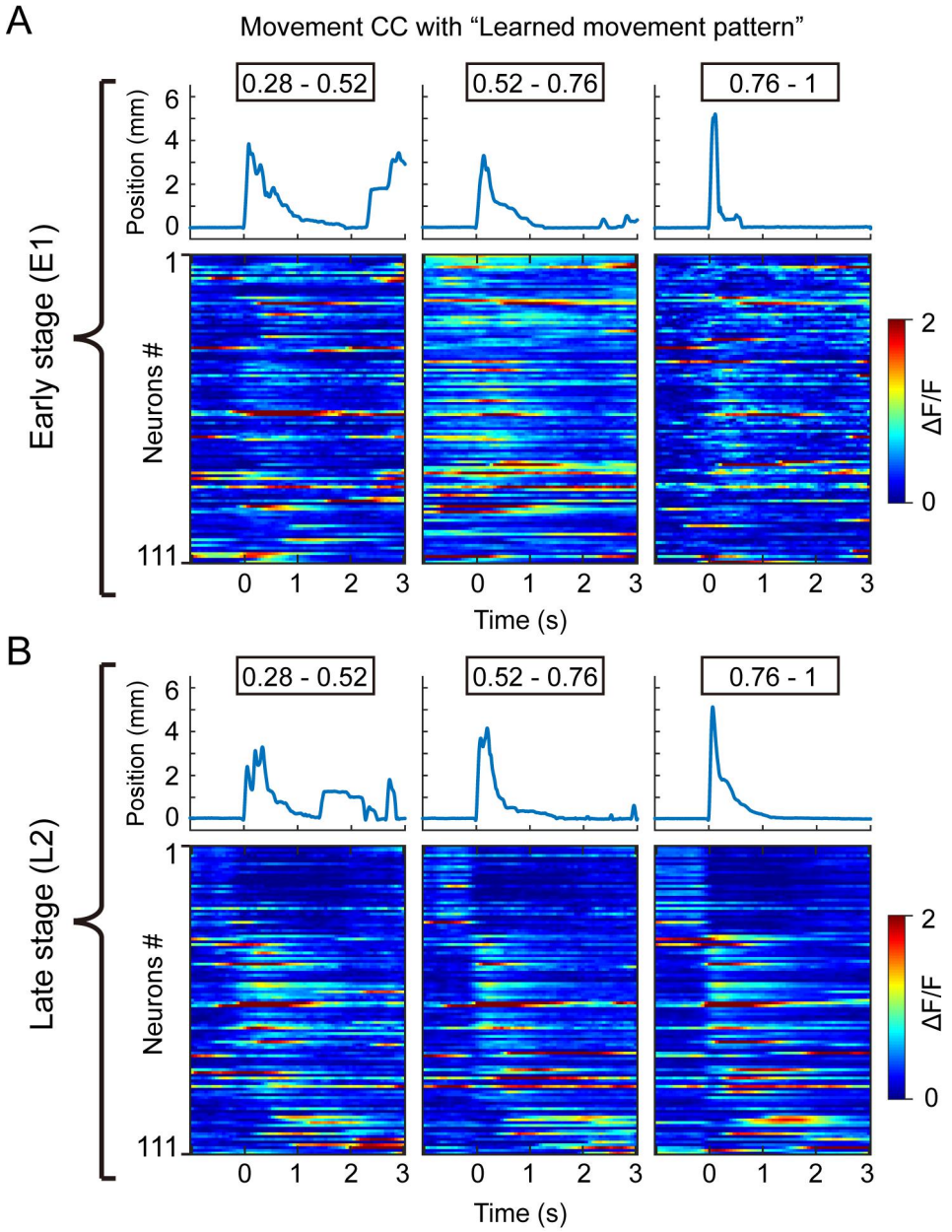

Fig. S8 (related to Fig. 4)

Examples of activity patterns corresponding to movements with low, medium, and high movement CCs. (*A*, *B*) Trials were binned into three groups according to the level of movement CC: high (0.76 - 1), medium (0.52 - 0.76) and low (0.28 - 0.52). Activity and movement trajectory were averaged for all trials chosen in these three groups. Top: mean lever movement trajectories. Bottom: heatmap of averaged activity of the same neuronal ensemble. Note that the activity patterns of the same population of neurons corresponding to the average movements trials in the three CC groups were irregular during the early stage (E1, *A*) and became more stereotyped during the late stage (L2, *B*). Furthermore, even at the late stage, the stereotyped sequential ensemble activity were more clear for stereotyped movements than less stereotyped movements (lower movement CCs). These results support the notion that the emergence of stable sequential activity patterns among the same ensemble of neurons after learning is related to the selection of the stereotyped movement, leading to the high percentage of movements with high movement CCs (see Fig. 4).

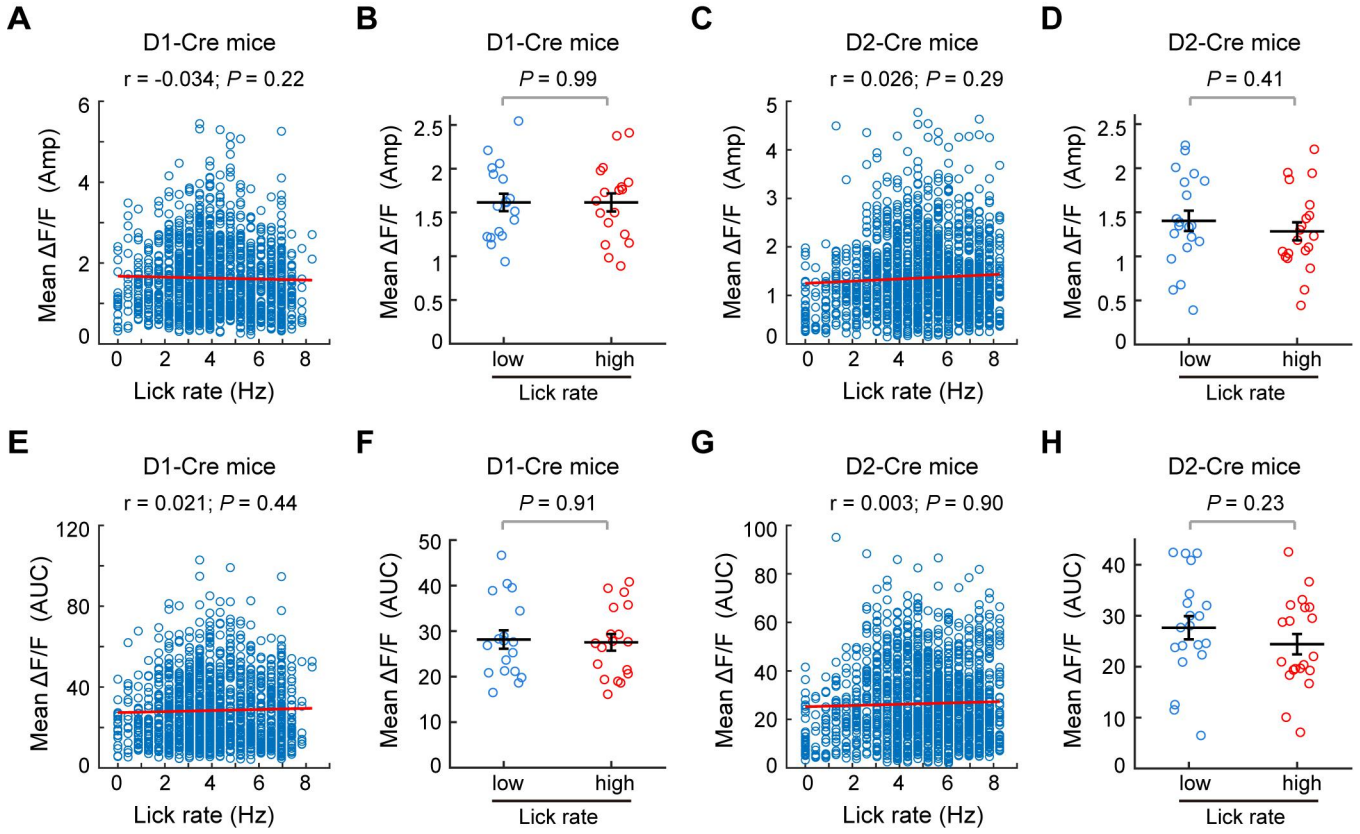

Fig. S9 (related to Fig. 5)

Relationship between DLS neuronal activity and lick rates. (*A*, *C*) The amplitude (Amp) of the mean ΔF/F of single-trial Ca^{2+} transients was plotted against the lick rate during that trial. Red line, linear regression of all data points. (*B*, *D*) Amplitudes of mean ΔF/F of high lick-rate trials and low-lick trials. High and low lick-rate trials were trials that had rates among the top and bottom 30% of all trials, respectively. Each dot represents the mean ΔF/F amplitude of high or low lick-rate trials for one mouse (D1-Cre, n = 18 mice; D2-Cre, n=20 mice, Wilcoxon rank sum test, error bar, s.e.m.). (*E-H*) Same as *A-D*, except that area under curve (AUC, integrated ΔF/F) of mean ΔF/F was used to quantify the activity.

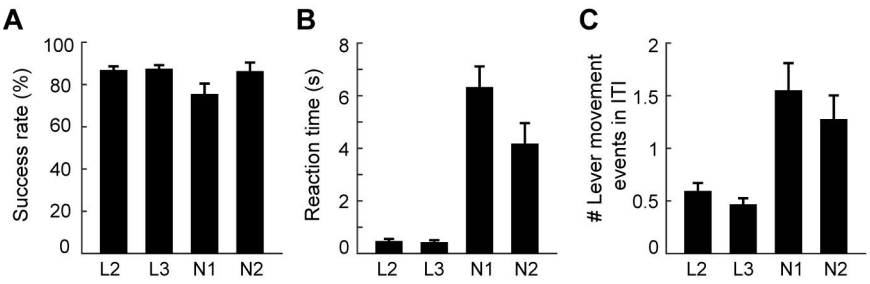

Fig. S10 (related to Fig. 6)

Behavioral performance of mice subjected to new motor tasks. (*A-C*) Success rate (*A*), reaction time (*B*) and number of lever-movement events during ITI (*C*) on L2, L3, N1 and N2. Note that mice could perform the new motor tasks with high success rate, but the reaction time was longer and the frequency of lever-movement events during ITI was higher than those associated with the learned rightward lever-pushing task (n = 18, with 8 D1- and 10 D2-Cre mice).

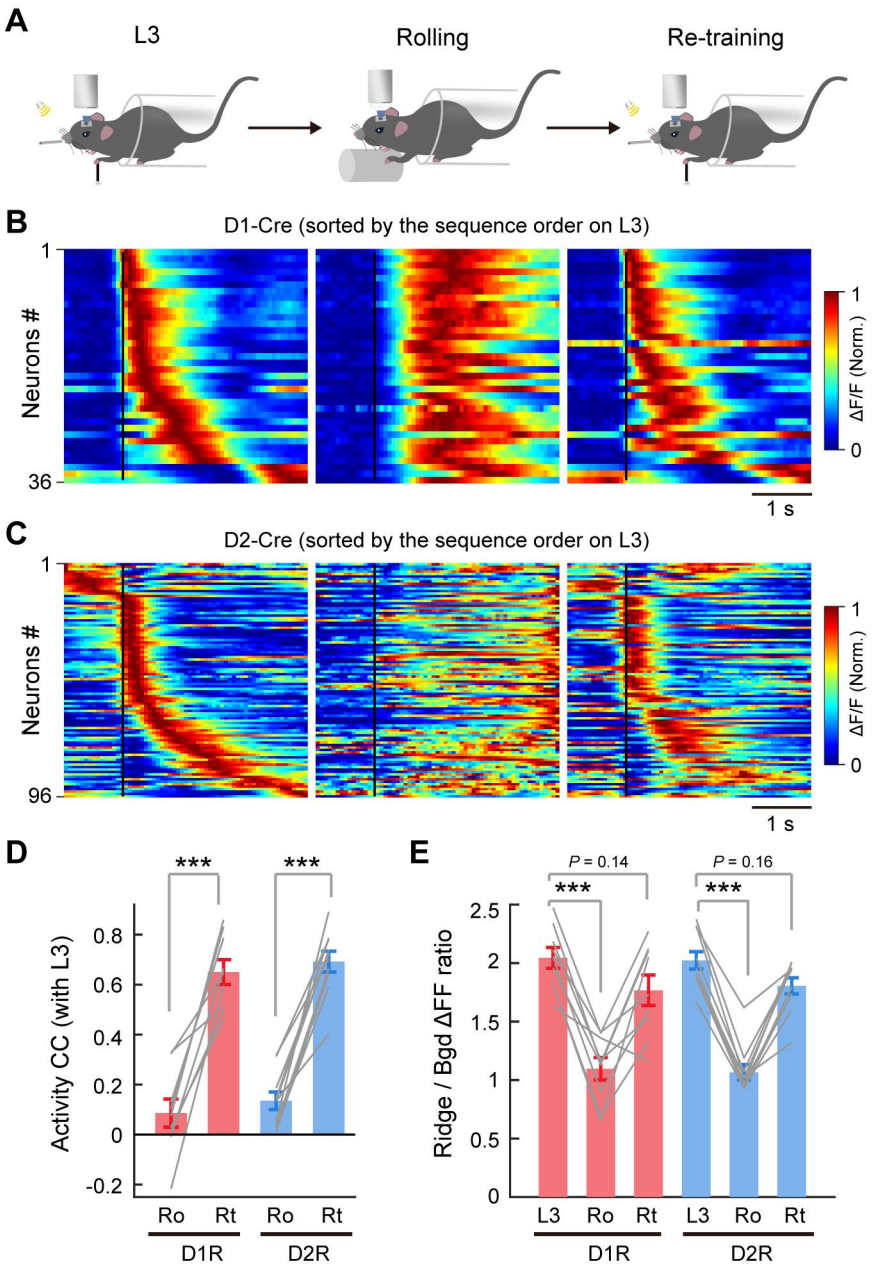

Fig. S11 (related to Fig. 6)

Changes in the sequential firing pattern when learned mice were subjected to passive rolling task. (*A*) Schematic diagram for the change in motor tasks. Left (L3): A mouse that had learned the lever-pushing task on day L3; Middle (Rolling): The learned mouse was transferred to a cylinder-rolling device and passively rolled the cylinder using forelimbs; Right (Re-training): 1 day after performing the rolling task, the mouse was subjected to the lever-pushing again. (*B*, *C*) Averaged activity of the same population of responsive neurons on day L3, "Rolling" day and "Re-training" day of one example D1-Cre mouse (*B*, 36 cells) and D2-Cre mice (*C*, 96 cells). Black line: onset time of lever-pushing or rolling. (*D*) Correlation coefficients (CCs) of average activity for each neuron between L3 and Rolling day ("L3-Ro") or between L3 and Re-training day ("L3-Rt"), with CCs calculated by using the temporal population activity vector (D1-Cre, n = 9; D2-Cre, n = 10; ***$P < 0.001$, Wilcoxon rank sum test). Gray lines connect data from the same mouse. Error bar, s.e.m. (*E*) Ridge to background ("Ridge/Bgd") ΔF/F ratio on L3, Rolling day and Re-training day (D1, n = 9, NS, $P = 0.14$; D2, n = 10, NS, $P = 0.16$; ***$P < 0.001$, Wilcoxon rank sum test). Error bar, s.e.m.

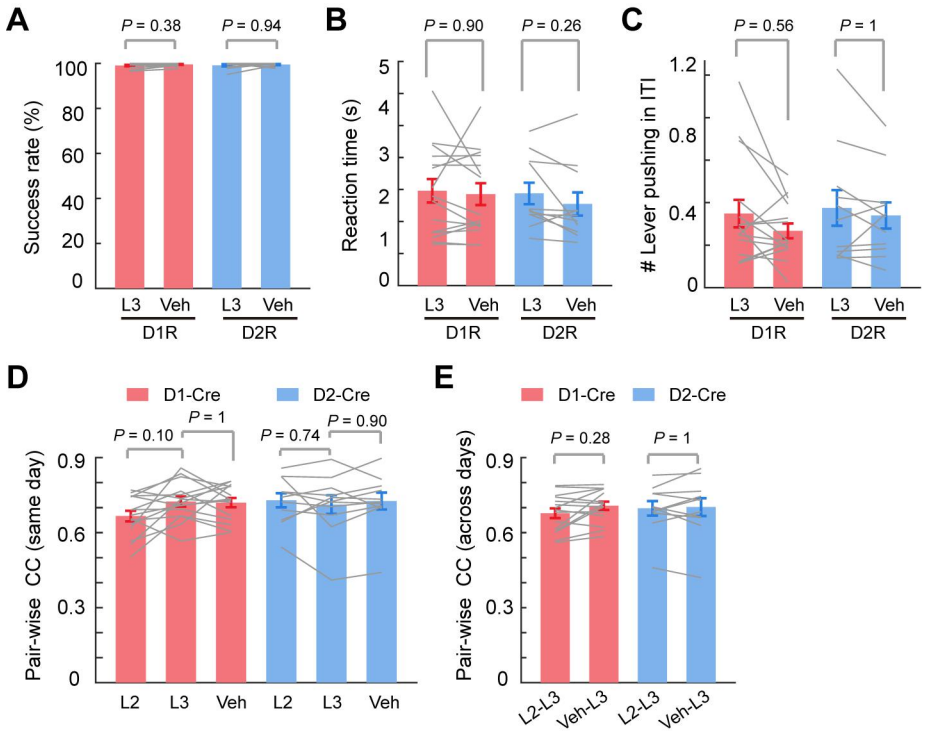

Fig. S12 (related to Fig. 7)

Control for chemogenetic suppression experiment, showing the effect of injecting vehicle instead of CNO. (*A*-*C*) Success rate (*A*), reaction time (*B*), and number of lever-pushing events during ITI (*C*), in individual D1- and D2-Cre mice expressing Cre-dependent hM4D on the last day (L3) of learning and on the next day with vehicle (saline + 4% DMSO) injection. Gray lines connect data from the same mouse (D1-Cre, n = 15 mice; D2-Cre, n = 11 mice; Wilcoxon rank sum test). (*D*, *E*) Pair-wise correlation coefficients (CCs) of movement trajectories between single-trial trajectories within the same day (*D*) and between two different days (*E*). No effect on behavioral performance was found ($P > 0.05$, Wilcoxon rank sum test). Error bar, s.e.m

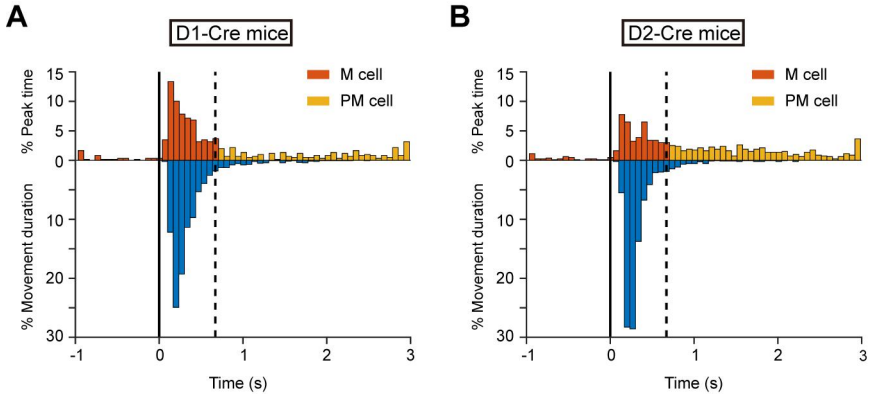

Fig. S13

Distribution of lever-pushing duration and peak activity time of M cells and PM cells. (*A*, *B*) Histogram showing the distribution of lever-pushing duration (blue bars) and peak activity time of M cells (red bars) and PM cells (yellow bars) for D1-Cre (*A*, n = 20) and D2-Cre (*B*, n = 20) mice. For D1R M cells, the time of peak ΔF/F is 0.24 ± 0.32 s (mean ± std), and the lever-pushing duration is 0.37 ± 0.31 s (mean ± std). For D2R M cells, the time of peak ΔF/F is 0.25 ± 0.34 s (mean ± std), and the lever-pushing duration is 0.34 ± 0.25 s (mean ± std). Thus, the average time for ΔF/F value to reach its peak for movement-related D1R and D2R neurons was shorter than the average duration of lever pushing. Black line: lever pushing movement onset; Black dashed line: boundary of M cells and PM cells. Cells that fired during the period of 0 - 0.67 s and after 0.67 s following movement onset were identified as M cells and PM cells, respectively.

Movie S1

This movie shows a mouse performing the rightward lever-pushing task during the late stage of learning. Yellow patch: time of presentation of the auditory cue. Top: lever-pushing movement trajectory recorded using a potentiometer. Middle: left forelimb movement obtained from the video analysis.

Movie S2

This movie shows a mouse performing the rightward lever-pushing task during the early stage of learning. Note that the reaction time was longer, since the mouse did not push the lever immediately after the cue onset.

Movie S3

This movie shows a mouse performing the leftward lever-pushing task.

Movie S4

This movie shows a mouse performing the downward lever-pressing task.

Movie S5

This movie shows a mouse performing the passive rolling task.

Table S1

The total number of trials used for averaging on each day for each D1- and D2-Cre mouse was listed in the following table. These trials were those with on extra lever pushing movement during 2.8 s before the rewarded movement onset. Column 1: animal; Column 2-7: number of trials on E1, E2, E3, L1, L2 and L3 for each mouse. Column 8: number of trials used when the data were analyzed using the same number of trials (**Fig. S6**), if the number of trials in those days larger than this value, the front part were used.

Animal	#Trial (E1)	#Trial (E2)	#Trial (E3)	#Trial (L1)	#Trial (L2)	#Trial (L3)	# Min Trial
D1_#1	45	60	58	85	102	75	45
D1_#2	46	42	44	72	95	94	42
D1_#3	40	50	49	69	55	60	40
D1_#4	26	50	48	107	93	96	26
D1_#5	43	45	49	92	92	107	43
D1_#6	22	16	47	79	69	71	16
D1_#7	20	39	56	79	60	73	20
D1_#8	46	31	40	87	65	62	31
D1_#9	33	38	27	54	54	56	27
D1_#10	39	65	54	55	47	65	39
D1_#11	40	54	31	60	70	61	31
D1_#12	34	42	27	82	104	68	27
D1_#13	40	58	51	54	80	44	40
D1_#14	18	22	53	32	46	64	18
D1_#15	37	22	27	75	53	86	22
D1_#16	27	49	53	61	54	66	27
D1_#17	72	58	47	84	72	80	47
D1_#18	61	33	43	57	50	65	33
D1_#19	67	51	51	64	94	98	51
D1_#20	32	42	43	48	61	61	32
D2_#1	44	29	42	85	83	77	29
D2_#2	38	30	45	52	31	60	30
D2_#3	52	59	56	50	71	93	50
D2_#4	54	15	57	42	74	81	15
D2_#5	37	31	63	62	72	79	31
D2_#6	44	37	30	47	74	74	30
D2_#7	40	36	22	67	68	75	22
D2_#8	52	24	19	52	73	90	19
D2_#9	19	23	26	68	50	46	19
D2_#10	23	31	28	34	59	76	23
D2_#11	74	72	85	62	90	103	62
D2_#12	55	63	61	74	60	92	55
D2_#13	82	69	57	75	87	96	57
D2_#14	47	67	82	58	79	92	47
D2_#15	77	54	63	70	69	89	54
D2_#16	85	62	62	81	70	46	46
D2_#17	54	46	60	84	101	84	46
D2_#18	45	61	51	71	94	92	45
D2_#19	63	55	73	71	76	94	55
D2_#20	77	78	62	135	112	119	62