Acquisition of Auditory Discrimination Mediated by Different Processes through Two Distinct Circuits Linked to the Lateral Striatum

Susumu Setogawa[1,2,3,4,5,*], Takashi Okauchi[4], Di Hu[4], Yasuhiro Wada[6], Keigo Hikishima[7], Hirotaka Onoe[8], Kayo Nishizawa[3], Nobuyuki Sakayori[9], Hiroyuki Miyawaki[1,2], Takuma Kitanishi[2], Kenji Mizuseki[1,2], Yilong Cui[4], and Kazuto Kobayashi[3,*]

[1]Department of Physiology, Graduate School of Medicine, Osaka Metropolitan University, Osaka 545-8585, Japan

[2]Department of Physiology, Osaka City University Graduate School of Medicine, Osaka 545-8585, Japan

[3]Department of Molecular Genetics, Institute of Biomedical Sciences, Fukushima Medical University, Fukushima 960-1295, Japan

[4]Laboratory for Biofunction Dynamics Imaging, RIKEN Center for Biosystems Dynamics Research, Kobe 650-0047, Japan

[5]Japan Society for the Promotion of Science, Tokyo 102-0083, Japan

[6]Laboratory for Pathophysiological and Health Science, RIKEN Center for Biosystems Dynamics Research, Kobe 650-0047, Japan

[7]Medical Devices Research Group, Health and Medical Research Institute, National Institute of Advanced Industrial Science and Technology, Tsukuba 305-8564, Japan

[8]Human Brain Research Center, Kyoto University Graduate School of Medicine, Kyoto 606-8507, Japan

[9]Department of Physiology and Oral Physiology, Graduate School of Biomedical and Health Sciences, Hiroshima University, Hiroshima 734-8553, Japan

27 *Address for correspondence:

28 Kazuto Kobayashi, Ph.D.

29 E-mail: kazuto@fmu.ac.jp

30 Department of Molecular Genetics, Institute of Biomedical Sciences, Fukushima Medical

31 University, Fukushima 960-1295, Japan.

32 Tel. +81-24-547-1667

33

34 Susumu Setogawa, Ph.D.

35 E-mail: setogawa@omu.ac.jp

36 Department of Physiology, Graduate School of Medicine, Osaka Metropolitan University, Osaka

37 545-8585, Japan

38 Tel. +81-6-6645-3717

39

Abstract

The striatum, the central hub of cortico-basal ganglia loops, contains functionally heterogeneous subregions distinguished by the topographic patterns of structural connectivity. These subregions mediate various processes of procedural learning. However, it remains unclear when and how striatal subregions engage in the acquisition of sensory stimulus-based decision-making. We show that the anterior dorsolateral striatum (aDLS) and posterior ventrolateral striatum (pVLS) are activated in a different temporal pattern during the acquisition phase of auditory discrimination. The aDLS promotes the behavioral strategy driven by the stimulus-response association while suppressing that by the response-outcome association, and the pVLS contributes to forming and maintaining the stimulus-response strategy. These two subregions exhibit distinct event-related responses during the learning processes. Our findings demonstrate that aDLS and pVLS neurons integrate the new learning of auditory discrimination in spatiotemporally and functionally different manners.

Introduction

The neural network connecting the cerebral cortex, basal ganglia, and thalamus forms functionally heterogeneous loops distinguished by the topographic patterns of connectivity[1,2], and is implicated in various learning processes by the integration of cognitive, sensorimotor, and reward information[3,4]. The striatum is a key node in the cortico-basal ganglia loops, and its segregated subregions are considered to mediate procedural learning, such as motor skills and instrumental behavior, through parallel processing via different loops[5-7]. Previous studies in humans and non-human primates have reported that the neural activity within the striatum shifts from the associative subregions (caudate and anterior putamen) to the sensorimotor subregions (posterior putamen) as motor learning progresses[8-11]. In rodents, an instrumental behavior initiates as goal-directed actions and then transits to stimulus-response habits[12,13], requiring the dorsomedial striatum (DMS; a homologue to the caudate in primates) and the dorsolateral striatum (DLS; a homologue to the putamen in primates), respectively[14,15]. Based on these results, a model of procedural learning has been proposed, in which the functional dominance changes from the associative to sensorimotor subregions in the striatum during the learning phases[16].

In contrast, deviations from the prior model are evident in other behavioral tasks on decision-making based on external sensory cues. For example, the neural activity in the caudate in humans is associated with learning of a visuomotor association task and the execution of learned behavior[17,18]. The tail of the caudate in non-human primates contributes to the automatic, habitual behavior in decision-making task using complex visual stimuli[19]. In rodents, the DLS is involved in the acquisiton and execution of the discrimination task with visual and auditory cues, wheras the DMS madiates the performance of learned behavior[20,21]. Moreover, other striatal subregions, such as the posterior striatum and ventrolateral striatum in rodents are also engaged in the auditory discrimination[22-24]. Based on these results, a hypothesis has arisen that the acquisition of sensory cue-based decision-making requires wide-range spatiotemporal processing within the striatum. However, the neural mechanism by which the functional circuits linked to striatal subregions

change throughout learning processes is not yet fully understood.

To investigate a large-scale striatal reorganization in accordance with the acquisition of external cue-dependent decisions, we conducted a two-alternative auditory discrimination task[25]. Using a small-animal neuroimaging technique, we found that the anterior dorsolateral striatum (aDLS) and posterior ventrolateral striatum (pVLS) are activated to the highest level at the middle and late stages during the acquisition phase of auditory discrimination, respectively. Then, chronic pharmacologic manipulations confirmed that the aDLS and pVLS are necessary for discrimination learning, and the transient manipulations also ascertained the function of these subregions in the learning at the corresponding stages. The resultant effects of the manipulations on the behavioral strategy showed that the aDLS promotes the strategy based on the stimulus-response association, while suppressing that based on response-outcome associations. Additionally, the pVLS engages in the formation and maintenance of the stimulus-response association strategy. Electrophysiological recording of striatal neurons indicated that subpopulations of aDLS neurons mainly represent the outcome of specific behaviors at the initial period of discrimination learning, whereas pVLS subpopulations encode the beginning and ending of each behavior in association with the progress of the learning. In addition, other subpopulations of aDLS and pVLS neurons show sustained activation after obtaining reward with distinct patterns of the combination between the stimulus and response. These findings reveal that the aDLS and pVLS regulate the acquisition of auditory discrimination through distinct spatiotemporal and functional manners, challenging the prior model which proposed a transition from the associative to sensorimotor striatal subregions during learning.

Results

Leaning processes of auditory discrimination

We employed a two-alternative auditory discrimination task, in which animals are required to make new associations between a tone instruction cue (2 or 10 kHz) and a response (pressing the left or right lever)[25], setting the interval between the stimulus presentation and lever insertion to 3 s (Fig.

1a). The success rate gradually increased, reaching a plateau around Day 13, which persisted through Day 24 (Fig. 1b). When the success rates across Days 2, 6, 10, 13 and 24 were compared, significant differences were observed between Days 2 and 6, Days 6 and 10, as well as Days 10 and 13, showing no significant difference between Days 13 and 24 (Fig. 1c). In contrast, the response time and omission rate did not vary among the training days (Fig. 1d,e), and the variation of response bias decreased along with the progress of auditory discrimination (Fig. 1f). Based on changes observed in the success rate, we categorized the learning processes into two phases: the acquisition phase, which lasted until Day 13; and the learned phase, which began after Day 13.

Distinct brain activity patterns in striatal subregions during learning processes

To investigate dynamic changes in regional brain activity in the entire striatum during auditory discrimination in the same animals, we conducted the positron emission tomography (PET) for small animals with 2-deoxy-2-[^{18}F]fluoro-D-glucose (^{18}F-FDG), which measures cerebral glucose metabolism correlated to brain activity[26-28] (Fig. 2a). We performed ^{18}F-FDG-PET scanning on Day 4 of the single lever press task without sound-based decision as a control condition and during a series of the trainings including the acquisition (Days 2, 6, and 10) and learned (Day 24) phases (Fig. 2b). The average number of lever presses in the single lever press task exceeded 80 times (81.6 ± 1.3, mean ± s.e.m.) on Day 2 and maintained that level until Day 4 (Fig. 2c). In the following auditory discrimination task, the success rate gradually increased along with the training days (Fig. 2d), showing a significant difference among sessions in which we performed the ^{18}F-FDG-PET scanning (Fig. 2e). The response time and omission rate were consistent among the days, and the variation of response bias became smaller as the task progresses (Fig. 2e). These performances were similar to those in the intact rats (compared to Fig. 1c–f), suggesting that the procedure for ^{18}F-FDG-PET scans does not affect the acquisition of discrimination.

The regional brain activity related to the behavioral task was analyzed with a voxel-based statistical parametric analysis[26], in which the ^{18}F-FDG uptake on either day of the discrimination

task was compared with that on the single lever task. The task-related brain activity in the whole brain is summarized in Extended Data Table 1. In the present study, we focused on the task-related activity in striatal subregions. We found significant increases of the activity in the unilateral aDLS and the bilateral pVLS on Days 6 and 10 during the acquisition phase (Fig. 2f,g). However, no significant change was observed in the DMS during the acquisition phase except a unilateral decrease on Day 24 (Fig. 2f). Next, to evaluate the activity related to the progress of learning, ^{18}F-FDG uptake on Days 6, 10, and 24 was compared with that on Day 2. The learning-dependent brain activity was significantly increased in the bilateral aDLS and the unilateral pVLS on Days 6 and 10, respectively (Fig. 2h), whereas the activity decreased in the bilateral DMS on Day 24 (Fig. 2h). This decrease was also detected when ^{18}F-FDG uptakes were compared between Days 10 and 24 (Fig. 2h), suggesting the down-regulation of the DMS activity in the learned phase.

To quantitatively validate changes in the brain activity in striatal subregions, we analyzed the amount of ^{18}F-FDG uptake in the voxel of interest of these subregions (Fig. 2i). The ^{18}F-FDG uptake in the bilateral aDLS reached a peak on Day 6, and it decreased during the subsequent days (Fig. 2j), whereas the uptake in the bilateral pVLS gradually increased along with the progress of learning, showing the maximal value on Day 10 (Fig. 2k). The uptake in the DMS did not alter during the acquisition phase, and it decreased on Day 24 (Fig. 2l). These results suggest that the aDLS and pVLS contribute to learning processes in a different temporal pattern during the acquisition phase, whereas the DMS may be engaged to the execution of the behavior in the learned phase.

Excitotoxic lesion of the aDLS and pVLS, but not the DMS, disrupts the acquisition of auditory discrimination

To study whether the changed activity of striatal subregions is involved in learning processes, rats received a bilateral injection of solution containing ibotenic acid (IBO; 8 mg/mL, 0.3 μL/site) or phosphate buffered saline (PBS) into the subregions, and we conducted for the behavioral study.

158 The range of IBO lesion in the subregions was analyzed by immunostaining with a neuronal marker
159 NeuN after the behavioral tests (Fig. 3a–c). The number of single lever presses increased similarly
160 along with the training in both the PBS- and IBO-injected groups into the aDLS, pVLS, or DMS
161 along with the training (Fig. 3d–f). In the discrimination learning, increases of the success rate were
162 impaired in both IBO-injected aDLS and pVLS groups, compared to the corresponding PBS-
163 injected groups (Fig. 3g,h). The response time in the aDLS-lesioned group did not alter during the
164 discrimination training, whereas that in the pVLS-lesioned group was continuously prolonged
165 (Extended Data Fig. 1a,b), and the omission rate and response bias were similar between the
166 lesioned and the corresponding control groups (Extended Data Fig. 1d,e and 1g,h). In contrast, the
167 success rate in the discrimination task was comparable between the PBS- and IBO-injected groups
168 into the DMS (Fig. 3i). The DMS lesion lengthened the response time through the learning
169 (Extended Data Fig. 1c), whereas it did not alter the omission rate or response bias (Extended Data
170 Fig. 1f,i). These results indicate that the aDLS and pVLS are required for the acquisition of auditory
171 discrimination, whereas the DMS contributes to the execution but not learning of the discriminative
172 behavior, being consistent with previous reports[29,30].

173 Previous studies have reported that the dorsal striatum is involved in both behavioral
174 strategies based on the response-outcome association and the stimulus-response association during
175 stimulus-response learning[12,13]. To assess the impact of striatal lesions on the behavioral strategies,
176 we analyzed the proportion of response attributed to two representative strategies in all responses of
177 each session. One is the "win-shift-win (WSW)" strategy, in which, after a correct response in the
178 previous trial, rats press the opposite lever in the current trial in response to a shift of the instruction
179 cue, resulting in the correct response. The WSW is considered to reflect the behavioral strategy
180 based on the stimulus-response association[31,32]. Another strategy is the "lose-shift-lose (LSL)"
181 strategy, in which, after an error response in the previous trial rats press the opposite lever in the
182 current trial despite a shift of the instruction cue, leading to another error response. The LSL is
183 considered to appear as a consequence of the behavioral strategy based on the response-outcome

184 association[31]. In the PBS-injected group into either of three striatal subregions, the proportion of
185 WSW and LSL strategies increased and decreased, respectively, during the acquisition phase (Fig.
186 3j–l). In the IBO-lesioned group into the aDLS, the increase in the WSW proportion was markedly
187 impaired as compared to the corresponding control group, and the decrease in the LSL proportion
188 was delayed compared to the same group (Fig. 3j). In the pVLS-lesioned group, the increasing
189 WSW proportion was also impaired as compared to the corresponding controls, although the
190 change in the LSL proportion was comparable to the control response (Fig. 3k). In contrast, in the
191 DMS lesion group both WSW and LSL proportions did not show significant differences from the
192 corresponding values in their control group (Fig. 3l). These results indicate that the aDLS promotes
193 the WSW strategy through stimulus-response association and suppresses the LSL strategy rooted in
194 response-outcome association, and that the pVLS selectively facilitates the WSW strategy
195 acquisition. Thus, these two subregions not only play roles in the acquisition of auditory
196 discrimination and but also control distinct behavioral strategies during the learning processes.
197 After the completion of discrimination training, we tested locomotion and feeding behavior
198 of rats with aDLS and pVLS lesions. There was no difference in locomotor activity in the open field
199 and the amount of food intake in both lesioned groups (Extended Data Fig. 1j–m), suggesting that
200 learning deficits following excitotoxic lesions of the striatal subregions cannot be attributable to
201 abnormalities in these general motor behaviors.
202

203 **Inhibiting aDLS and pVLS activity impairs discriminative behavior at different stages of the**
204 **acquisition phase**

205 Our ^{18}F-FDG-PET imaging study indicated that the brain activity in the aDLS and pVLS was
206 increased in a different temporal pattern during the acquisition phase of auditory discrimination. To
207 examine whether neuronal activity in striatal subregions is linked with the processes at different
208 stages, we temporarily inhibited the activity in each subregion by the treatment with a $GABA_A$
209 receptor agonist muscimol (MUS). In a pilot study, we tested the effect of the bilateral injection of

210 several doses of MUS (0.1, 0.2, and 0.5 μg/μL) into the aDLS or pVLS on the single lever press
211 behavior and confirmed that the administration of the lowest dose of MUS did not affect the
212 frequency of single lever presses and the quantity of food intake (Extended Data Fig. 2a–d). In
213 addition, we divided the acquisition phase into three stages based on the success rate that included
214 the early (< 60%), middle (60–80%), and late (80–100%) stages. To achieve the stage-specific
215 inactivation of the aDLS or pVLS neurons, rats were given bilateral injections of MUS or saline
216 (SAL) at three different timings during a series of the training. The first injection was conducted on
217 Day 2 to target the early stage. The second and third injections were conducted on the day after the
218 success rate had reached 60% and 80% for the first time through the training, respectively, to target
219 the middle or late stage. After the injections, the rats were used for the behavioral test. Placement
220 locations of the injection cannula into the striatal subregions were assessed by cresyl violet staining
221 after the behavioral tests (Fig. 4a,b). The success rates on the day before the injection (day N-1) at
222 the early, middle, and late stages were similar between the two groups that received the bilateral
223 injections of MUS or SAL (Extended Data Fig. 2e,f). Rats were then administered injections into
224 the aDLS on the next day (Day N), and the success rate in the MUS-injected groups significantly
225 decreased at the middle stage compared to the corresponding SAL-injected groups, but not at the
226 early and late stages (Fig. 4c, left). In contrast, the success rate in the MUS-injected groups into the
227 pVLS significantly decreased at the late stage, but not at the early and middle stages, compared
228 with the respective SAL-injected groups (Fig. 4d, left). To normalize the variation of learning speed
229 in the individual animals, we calculated the value by subtracting the success rate on Day N-1 from
230 that on Day N in the individual animals. There were significant reductions in the subtracted rate at
231 the middle stage in the MUS-injected group into the aDLS (Fig. 4c, right) and at the late stage in the
232 MUS-injected group into the pVLS (Fig. 4d, right). These results indicate that the aDLS and pVLS
233 mainly function at the middle and late stages, respectively, in the acquisition phase of auditory
234 discrimination, supporting the results obtained from our imaging study.
235 Next, we analyzed the influence of temporal striatal inactivation on the behavioral strategy

236 at the middle and late stages of the acquisition phase. We calculated the proportion of WSW and
237 LSL strategies, and subtracted the values of WSW or LSL proportion on Day N-1 from those on
238 Day N. The proportion of both strategies on Day N-1 were comparable between the MUS- and
239 SAL-injected groups (Extended Data Fig. 2g,h). In the aDLS, the subtracted proportion of LSL
240 strategy at the middle stage significantly increased in the MUS-injected group as compared to the
241 SAL-injected group, although the subtracted proportion of WSW strategy tended to decrease in the
242 MUS-injected group as compared to the control group (Fig. 4e), and the subtracted proportion of
243 WSW or LSL strategy at the late stage was similar between both groups (Fig. 4f). In the pVLS, the
244 subtracted proportions of the WSW and LSL strategies at the middle stage were similar between the
245 MUS- and SAL-injected groups (Fig. 4g), and the subtracted WSW proportion at the late stage
246 significantly decreased in the MUS-injected group as compared to the SAL-injected group,
247 although the subtracted LSL proportion was comparable between the two groups (Fig. 4h). Our
248 results show that transient inhibition of aDLS activity at the middle stage affects the LSL strategy
249 and has a moderate effect on the WSW strategy, supporting that the aDLS mediates the acquisition
250 process of auditory discrimination at the middle stage at least partly through the suppression of the
251 behavior based on the response-outcome association. Although the impact of inhibition on the
252 WSW strategy was modest, excitotoxic lesion of the aDLS influenced the WSW strategy in addition
253 to the LSL strategy (see Fig. 3j), suggesting that the aDLS acts to enhance the auditory
254 discrimination by promoting the stimulus-response association, together with the suppression of-the
255 response-outcome association. The discrepancy in the extent of changes in the WSW strategy
256 between the two treatments may be generated from the difference in the timing or sustainability of
257 drug inhibition of the striatal subregion. In addition, our results show that pVLS inactivation at the
258 late stage selectively impairs the WSW strategy, supporting contribution of the pVLS to the process
259 via progression of the stimulus-response association.
260
261

262 **Neuronal activity in the aDLS and pVLS during the acquisition phase**

263 To examine the firing activity in striatal subregions during the acquisition phase of auditory
264 discrimination, we simultaneously recorded neuronal activity in the aDLS and pVLS at the
265 unilateral (left) side by using a multi-unit recording procedure (Fig. 5a). Using immunostaining for
266 tyrosine hydroxylase (TH) after the recordings, we verified the location of electrodes within the
267 striatal subregions (Fig. 5b). After the single lever press task, we conducted the auditory
268 discrimination task for the recordings. In correct trials, the delay period (0.5 ± 0.2 s) was inserted
269 between a lever press and a sound indicating the reward to distinguish neuronal activity related to
270 the two events (Fig. 5c). The behavioral data in individual rats were divided into three grouped
271 sessions corresponding to the early, middle, and late stages, based on a sigmoid function analysis[33]
272 of the success rate in the task. The success rates in the grouped sessions were 52.7 ± 0.8 %, 61.6 ±
273 0.7 %, and 85.3 ± 1.7 % (mean ± s.e.m.) for the early, middle, and late stages, respectively, and the
274 rate was significantly increased throughout the three stages (Extended Data Fig. 3a). However, the
275 response time, omission rate, and response bias were similar among the stages (Extended Data Fig.
276 3b–d). Thus, these stages corresponded to the timing window defined in the acquisition phase of the
277 transient MUS inhibition experiment of neuronal activity.

278 We identified 1,062 and 423 well-isolated neurons from the aDLS and pVLS, respectively,
279 across the three stages (Extended Data Fig. 3e) and focused on the major population of striatal
280 neurons, putative medium spiny neurons (MSNs; Extended Data Fig. 3f,g). The total numbers of
281 MSNs at the early, middle, and late stages were 295, 364, and 360 neurons for the aDLS, and 87,
282 110, and 190 neurons for pVLS, respectively. Based on the combinatorial pattern of the tone
283 instruction cue and lever press in our discrimination task, we categorized the electrophysiological
284 data into four trial types: (1) high-frequency tone/right lever press (HR) and (2) low-frequency
285 tone/left lever press (LL) as correct responses; and (3) low-frequency tone/right lever press (LR)
286 and (4) high-frequency tone/left lever press (HL) as error responses (see Fig. 1a). We identified HR
287 or LL type neurons showing significant changes in firing rate related to the instruction cue onset

288 (CO), choice response (CR), reward sound (RS), or first licking (FL) as compared to the baseline
289 firing rate in each of two trials (Mann-Whitney *U* test, $p < 0.01$). These neurons were further
290 divided into two groups based on increased or decreased activity (Extended Data Fig. 4a,b). In the
291 following analyses, we focused on the HR and LL type neurons with increased event-related
292 activity to explore the firing patterns of neurons underlying behavioral processes of auditory
293 discrimination.
294
295 **Subpopulations of aDLS neurons, but not of pVLS neurons, show firing activity related to the**
296 **behavioral outcome**
297 Previous studies have reported that neurons in the dorsal striatum show the outcome-related activity
298 in the instrumental conditioning tasks[24,34]. We thus investigated whether aDLS and pVLS neurons
299 exhibit a response related to the behavioral outcome during the acquisition of auditory
300 discrimination. In the aDLS, the firing activity of RS-related HR type (RS-HR) neurons showed a
301 higher extent in the increase during a period, when the RS is presented in the correct responses, in
302 the HR trial as compared with the LR, LL, or HL trial throughout the three stages (Fig. 5d, top
303 rows). The rate of these neurons also moderately increased during the same period in the LR trial
304 compared to the LL or HL trial across the three stages, although the extent of the increase was lower
305 than that in the HR trial (Fig. 5d, top rows). The area under the receiver operating characteristic
306 (auROC) values revealed a significant difference of the HR trial from the LR trial during the RS
307 period, whereas the values of HL and LL trials indicated a significant difference in the opposite
308 direction from the LR trial during the same period (Fig. 5d, bottom rows). The averaged firing rate
309 during the RS period in the HR or LR trial appeared to modestly increase at the middle stage, but no
310 significant difference existed among the stages (Fig. 5e). To directly compare the number of
311 neurons showing event-related activity among different stages, we calculated the number of RS-HR
312 neurons by matching the trial number at each stage (Fig. 5f). The proportion decreased
313 progressively from the early to middle stages, and to the late stage. In addition, the firing rate of

314 RS-related LL type (RS-LL) neurons exhibited a higher increase in the LL trial and a modest
315 increase in the HL trial while the averaged firing rate during the RS period was similar among the
316 three stages. Furthermore, the proportion of RS-LL neurons reduced progressively through the
317 stages (Extended Data Fig. 5a–c). These results indicate that the aDLS contained neurons related to
318 the outcome of choice, showing a greater increase in the firing activity for the correct responses
319 (HR and LL trials) than the error responses (LR and HL trials), and that the proportion of these
320 subpopulations was gradually reduced through the progress of learning. In contrast, during the RS
321 period at all stages, the firing rate and auROC values of RS-HR and RS-LL neurons in the pVLS
322 revealed no significant differences across the four kinds of trial (Extended Data Fig. 5d,e),
323 indicating that the firing activity of pVLS neurons did not appear to be associated with the
324 behavioral outcome.
325
326 **Subpopulations of aDLS and pVLS neurons show sustained firing activity after the reward in**
327 **different manners**
328 The dorsal striatum is known to demonstrate a response of neuronal activity after the positive
329 outcome in the instrumental behaviors[35,36]. We thus examined whether aDLS and pVLS neurons
330 show the response after obtaining the reward during the acquisition phase of auditory
331 discrimination. In the aDLS, the firing rate of FL-related HR type (FL-HR) neurons indicated a
332 sustained elevation after the FL in the HR trial as compared with any of the other three trials across
333 the early, middle, and late stages (Fig. 6a, top rows). The auROC values revealed a significant
334 difference of the LR, LL, or HL trial from the HR trial across the three stages (Fig. 6a, middle
335 rows). In addition, we calculated the averaged firing rate and increased firing period, defined as the
336 total number of time bins above 3 s.d. of the baseline firing rate for 5 s after the FL in the HR trial.
337 The averaged firing rate was comparable between the early and middle stages, but the rate was
338 significantly lower at the late stage compared to either of the other two stages (Fig. 6b). The
339 increased firing period was also comparable between the early and middle stages, showing a

340 significant reduction at the late stage (Fig. 6c). The proportions of FL-HR neurons were similar
341 between the early and late stages, although they showed a slight decrease at the middle stage (Fig.
342 6d). There were few neurons with significant correlations between the firing rate and licking
343 performance (Fig. 6e). In addition, the firing rate of the FL-related LL type (FL-LL) neurons
344 exhibited the largest increase in the LL trial at the three stages, showing a tendency to reduce the
345 average firing rate and increased firing period after the FL across the stages (Extended Data Fig.
346 6a–c). The proportion of FL-LL neurons during the learning processes displayed similar changes as
347 observed in the FL-HR neurons (Extended Data Fig. 6d). These results indicate that the aDLS
348 contained neurons showing long-lasting, increased activity after the reward in the specific correct
349 combination of the stimulus and response (HR or LL trial), and that the extent of firing activity and
350 length of firing period in the response appeared to be gradually reduced with no correlation between
351 the proportion of neurons and the progress of the learning stages.
352 In the pVLS, FL-HR neurons showed a sustained elevation in the firing rate after receiving
353 a reward in the HR trial compared to that in the LR or HL trial throughout the three stages (Fig. 6f,
354 top rows). A similar elevation in the firing rate was also observed in the LL trial throughout the
355 stages (Fig. 6f, top rows). The auROC values revealed a significant difference of the LR or HL trial,
356 but not the LL trial, from the HR trial throughout the stages (Fig. 6f, middle rows). The average
357 firing rate did not show significant changes among the three stages (Fig. 6g), and the increased
358 firing period was also similar across the stages (Fig. 6h). The proportion of FL-HR neurons
359 increased at the late stage compared to the early or middle stage (Fig. 6i). The firing rate showed
360 almost no significant correlation to the licking performance (Fig. 6j). In addition, FL-LL neurons
361 exhibited similar firing patterns and changes in the proportion of neurons across the stages
362 (Extended Data Fig. 6f–j). Therefore, the pVLS included neurons representing the long-lasting
363 increased activity after the reward in the two kinds of correct responses (HR and LL trials) with the
364 increased number of these neurons at the late stage of learning.
365

366 **Neurons in the pVLS, but not in the aDLS, exhibit firing patterns related to the beginning and**
367 **ending of the behavior**
368 The dorsal striatum plays a key role in the acquisition and execution of behaviors based on the
369 stimulus-response association, and the association is accompanied by a transient activation of
370 neuronal firing immediately after the stimulus presentation or action in trials, which is considered to
371 represent the beginning and ending of a learned behavior[37-39]. We first examined whether the firing
372 activity of CO-related HR type (CO-HR) or CO-related LL type (CO-LL) neurons corresponds to
373 the beginning of the behavior in the auditory discrimination task. In the aDLS, the firing rate of CO-
374 HR or CO-LL neurons showed a significant difference at some time points in each trial compared to
375 those in any of the other three trials, with a significant difference of auROC values between the
376 trials (Extended Data Fig. 7a,c). However, the transiently increased activity did not emerge at the
377 time points corresponding to the beginning of the behavior. In contrast, CO-HR neurons in the
378 pVLS showed a transient increase in the firing rate immediately after the CO, and the activity
379 appeared at the middle and late stages (Fig. 7a). The transient activation was observed not only in
380 the HR trial but also in the other three trials, being consistent with the results of a previous report[40].
381 The increased firing rate showed a slight but notable difference between the HR and LL trials with a
382 significant difference in the auROC values between the two trials (Fig. 7a), suggesting that different
383 combinations between the stimulus and response may affect the level of firing activity. Differences
384 in the averaged firing rate between the two kinds of trial was observed at the middle and late stages
385 (Fig. 7b). The proportion of CO-HR neurons increased at the middle and late stages compared to
386 the early stage (Fig. 7c). In addition, the firing rate of CO-LL neurons exhibited patterns similar to
387 those observed in the CO-HR neurons (Extended Data Fig. 7e). The data indicate that
388 subpopulations of pVLS neurons show firing activity related to the beginning of the behavior, with
389 a rising number of neurons throughout discrimination learning.
390 Next, we examined the firing activity of CR-related HR type (CR-HR) or CR-related LL
391 type (CR-LL) neurons, corresponding to the ending of a behavior in auditory discrimination. In the

392 aDLS, the firing rate of these neurons did not exhibit the transient increase related to the ending of
393 the behavior, instead, there was a modest increase before the lever press (Extended Data Fig. 7b,d).
394 In contrast, CR-HR neurons in the pVLS were transiently activated immediately after the lever
395 press in the four kinds of trial at the middle and late stages, and the rate appeared to gradually
396 decline toward the RS presentation (Fig. 7d). The averaged firing rate was similar between the HR
397 and LL trials though the stages (Fig. 7e). The proportion of the CR-HR neurons increased at the
398 middle and late stages compared to the early stage (Fig. 7f). In addition, the rate of CR-LL neurons
399 displayed similar firing patterns to those observed in the CR-HR neurons (Extended Data Fig. 7f).
400 Thus, subpopulations of pVLS neurons also represent the activity related to the ending of the
401 behavior in auditory discrimination, increasing the neuron number throughout the learning stages in
402 a similar pattern to neurons relevant to the beginning of the behavior in the task.
403

404 **Discussion**

405 Our findings demonstrate that the aDLS and pVLS are required for the acquisition of auditory
406 discrimination at different stages of learning, and that the DMS is not necessary for the acquisition
407 of discrimination, although it appears to contribute to the execution of actions, in particular to the
408 regulation of response time. These results not only support some recent studies reporting no
409 requirement of the DMS in the procedural learning[20,41,42], but also reveal important roles of two
410 different subregions in the lateral striatum in the external cue-dependent decision-making. We
411 propose that aDLS and pVLS neurons act to integrate the new learning of auditory discrimination in
412 spatiotemporally and functionally distinct manners, showing a remarkable difference from the prior
413 learning model that proposes the functional dominance changes from the DMS to DLS subregions
414 during procedural learning[12-16].

415 Our ^{18}F-FDG-PET imaging study demonstrated that the brain activity in the aDLS reaches
416 a peak at the middle stage of the acquisition phase of auditory discrimination (Fig. 2i). The results
417 of excitotoxic lesion experiment of striatal subregions suggest that the aDLS functions to enhance

418 the acquisition of discrimination through the promotion of behavior based on the stimulus-response
419 association and the suppression of behavior based on the response-outcome association (Fig. 3j).
420 Transient inhibition experiment of neuronal activity confirmed this aDLS function at the middle
421 stage of the learning, although the effect of drug inhibition was moderate for the promotion of the
422 stimulus-response behavior (Fig. 4e,f). In addition, we found subpopulations of aDLS neurons (RS-
423 related neurons) which represent the behavioral outcome and differentiate the correct (HR or LL)
424 and incorrect (LR or HL) choices (Fig. 5d, Extended Data Fig. 5a). The distinct responses of these
425 neurons may explain the mechanism by which the aDLS regulates the stimulus-response and
426 response-outcome associations through distinct neural circuits leading to a differential control of the
427 two associations. Alternatively, the responses mainly promote the behavior based on the stimulus-
428 response association, thereby secondarily inhibiting the response-outcome association. Indeed,
429 some previous studies have suggested that an action based on either the stimulus-response
430 association or response-outcome association is generated as a result of competition between
431 different striatal subregions[43]. We also found other subpopulations of aDLS neurons (FL-related
432 neurons) showing long-lasting increased activity after obtaining reward in response to a specific
433 combination between the stimulus and response (HR or LL trial) (Fig. 6a, Extended Data Fig. 6a).
434 This activity may function as the feedback signal to promote the stimulus and response association,
435 resulting in the enhancement of discrimination learning. The proportion of neurons related to the
436 behavioral outcome showed a gradual reduction through the progress of auditory discrimination
437 (Fig. 5e and Extended Data Fig. 5c), and the extent of the firing activity and the length of the firing
438 period of neurons showing sustained activation after reward were also decreased at the late stage of
439 learning (Fig. 6b,c). The learning-dependent changes in the properties of these subpopulations in
440 the aDLS may explain the mechanism underlying the earlier acquisition process of the auditory
441 discrimination.

442 The ^{18}F-FDG-PET imaging indicated that the pVLS activity was gradually elevated
443 through the acquisition phase of auditory discrimination, showing the maximal value at the late

stage (Fig. 2j). The lesion experiment suggested that the pVLS drives the formation of discrimination learning via the selective progression of the stimulus-response association (Fig. 3k), and the transient inhibition experiment ascertained the pVLS function at the late stage of learning (Fig. 4g,h). Previous studies have shown that the stimulus-response association is encoded in the DLS or VLS as the transient neuronal activity representing the beginning and ending of a learned behavior[37-39,44]. A recent study reported that the VLS is causally related to learned performance based on the auditory stimulus, while the DLS shows a weaker relationship to the behavior[22]. In the present study, we found subpopulations of pVLS neurons that represent the beginning and ending of a behavior (CO- and CR-related neurons, respectively) (Fig. 7a,d; Extended Data Fig. 7e,f), showing an increased number of neurons along with the progress of discrimination learning (Fig. 7c,f). These data suggest that pVLS neurons contribute to the formation of the stimulus-response association through the learning processes. In addition, we found another subpopulation of pVLS neurons (FL-related neurons) that are continuously activated after the reward outcome in response to two kinds of correct choices (HR and LL trials) (Fig. 6f; Extended Data Fig. 6f)), and the proportion of neurons was increased at the late stage (Fig. 6i; Extended Data Fig. 6i). This activity may act as a feedback signal to consolidate the stimulus and response association, leading to the maintenance of associative memory at the late stage during the acquisition phase of learning.

In the present study, we identified neurons with four kinds of event-related activity in the aDLS and pVLS, and classified them into the HR and LL type neurons in which their firing activity showed a significant increase in the HR and LL trials, respectively, as compared to the baseline (Fig. 5–7, and Extended Data Fig. 5-7). Parts of these two types of neurons with the respective event-related activity overlapped, and the numbers of overlapping neurons varied between striatal subregions (Extended Data Fig. 8a,b). In the pVLS, the overlapping neurons for the same event-related activities were 35.1%, 45.2%, and 67.0% in the CO, CR, and FL, respectively. In contrast, the aDLS had lower overlaps of 12.6% and 22.7% in the RS and FL, respectively. Thus, distinct neurons in the aDLS tend to have properties of a single trial type while the same neurons in the

470 pVLS seem to possess properties of both types. One possible explanation for the diversity in the
471 proportion of overlapping neurons between the two subregions may be the difference of input
472 patterns into striatal neurons or plastic changes in functional connectivity during learning processes.
473 The detailed mechanism by which aDLS and pVLS neurons react to the events in a different
474 manner should be investigated in the future.

475 Anatomically, striatal subregions are known to form heterogeneous cortico-basal ganglia
476 loops in mice[2,45,46]. We examined the neural circuits linked to the aDLS and pVLS in rats by using
477 bidirectional tracers. The results showed that two striatal subregions receive innervations from
478 different subregions in the cerebral cortex, intralaminar thalamus, and substantia nigra pars
479 compacta, and send projections to different subregions in the basal ganglia nuclei, including the
480 globus pallidus, entopeduncular nucleus, and substantia nigra pars reticulata (Extended Data Fig. 9),
481 confirming that the aDLS and pVLS form the heterogenous cortico-basal ganglia loops in the rat
482 brain. These data suggest that the aDLS and pVLS mediate the acquisition of auditory
483 discrimination via the parallel loops, and cooperate to achieve neural processing related to the dual
484 regulation of the stimulus-response and response-outcome associations. Further studies are needed
485 to elucidate the mechanism by which the two kinds of loops are processed separately and integrate
486 the new learning of discrimination.

487

488 ## Materials and Methods

489 **Animals**

490 Animal care and handling procedures were carried out in accordance with the guidelines established
491 by RIKEN Center for Biosystems Dynamics Research, Fukushima Medical University, and Osaka
492 City University. All procedures were approved by their Institutional Animal Care and Use
493 Committees. All efforts were made to minimize the number of animals used and their suffering
494 throughout the course of the experiments. Male Long-Evans rats (8–13 weeks old) were used for
495 the present study. The rats were maintained at 22 ± 2 °C and 60% humidity in a 12-h light/12-h dark

496 cycle, and food and water were continuously available unless stated otherwise.

497

498 **Surgery**

499 The rats underwent a stereotaxic surgery under isoflurane anesthesia (1–5%). For excitotoxic lesion
500 of striatal neurons, IBO was dissolved in PBS (0.8 mg/mL, Wako Pure Chemical), and IBO solution
501 (0.3 μL per site) was injected bilaterally with the coordinates (mm) from bregma and dura for the
502 aDLS (AP +1.6, ML ±3.3, DV -3.9), pVLS (AP -1.6, ML ±4.6, and DV -5.6), or DMS (AP -1.6, ML
503 ±1.4, and DV -4.8) according to the atlas of rat brain[47]. Injection was performed through a glass
504 pipette (diameter: 60 μm) at a constant velocity of 0.1 mL/min with the microinfusion pump
505 (EICOM). For transient inactivation of striatal neurons, two guide cannulae (length: 10 mm, outer
506 diameter: 26 gauge) were bilaterally placed on the skull and fixed using four screws and dental
507 cement. The chip of the cannula was placed at 1 mm above the target coordinates. A solution
508 containing MUS (Sigma-Aldrich) in saline (0.1 μg/μL, 0.25 μL per site) was injected into the target
509 coordinates through the internal cannula (outer diameter: 30 gauge) at a constant velocity of 0.25
510 μL/min. For anterograde and retrograde tracing, a solution containing cholera toxin subunit B (CTb;
511 1.0 mg/mL) conjugated to Alexa 488 or Alexa 555 (Invitrogen) as a bidirectional tracer was injected
512 into the aDLS and pVLS (0.3 μL/site), respectively. The coordinates for the injection into striatal
513 subregions were the same as described above. For the electrophysiological experiments, surgery
514 was performed as described previously[48]. A 64-channel silicon probe (Buzsaki64spl; NeuroNexus)
515 attaching to a 3D-printed microdrive with a movable screw was chronically implanted with the
516 coordinates (mm) from the bregma and dura into the left aDLS (AP +1.6, ML -3.3, and DV -2.5)
517 and pVLS (AP -1.6, ML -4.6, and DV -3.6). A faraday cage composed of copper mesh was secured
518 on the skull with dental cement to reduce electrical noise and avoid damage to the implants. The
519 rats were allowed to recover for 1 week after the surgeries.

520

521

522 **Behavioral analysis**

523 Behavioral tasks were conducted in operant conditioning chambers (30.5 × 24.1 × 29.2 cm; ENV-
524 007CT; MED Associates) equipped with two retractable levers on both sides of a reward port in the
525 front panel (ENV-203M-45, ENV-652AM, and ENV-201A; MED Associates). A multi-tone
526 generator (ENV-223; MED Associates) and room light were mounted on the top of the center and
527 corner panels on the rear side, respectively. For the electrophysical experiment, a 3D printed reward
528 port was used to detect the licking and to prevent contact with the head attachment and front panel.
529 The behavioral experiment in intact rats was performed using grain pellets (F0021, Bioserv) as a
530 food reward. To prevent spillover of radioactivity into the temporal cortex as a consequence of
531 excessive ^{18}F-FDG accumulation into the muscles by their activity to eat the pellets[49,50], the PET
532 imaging experiments were carried out using 0.02% saccharine solution (B0131, Tokyo Chemical
533 Industry) as a liquid reward. The electrophysiological experiment was also carried out using the
534 saccharine solution. The pharmacological experiment was performed using the pellets.

535 The rats were subjected to dietary restrictions of 12 g/day (CE-2, CLEA Japan) and water
536 access was limited to 6 to 9 h in accordance with the reward used in the experiments for 3 days
537 before the beginning of the training schedule, and the body weights of the rats were maintained at
538 ~85% of their normal weights. The training schedule consisted of four continuous steps. In the first
539 step (magazine training), the rats were habituated to the operant conditioning chambers, and the
540 reward was presented with the click every 20 s. Each daily session was started by illumination of
541 the room light, and continued for 30 min. The sessions were conducted for 3 consecutive days. In
542 the second step (shaping), only one lever on the left or right side was inserted. When the rats
543 pressed the lever, they earned a reward accompanied by a click. Each daily session was started by
544 illumination of the light, and lasted 30 min or until the criteria, in which the rats received 90
545 rewards. When the number of lever presses on either side reached the criteria, the lever was inserted
546 in the opposite side on the next day. When the rats achieved the criteria on both sides for 2 or 3
547 days, they moved to the following step. In the third step (single lever press task), each trial was

548 started by illumination of the light and insertion of one lever on either the left or right side in a
549 pseudorandom manner with equal frequencies on both sides for every set of four trials. When the
550 rats pressed the lever, the reward was presented with the sound, and the lever was immediately
551 retracted. The room light turned off 4 s after the lever press. When they did not press the lever
552 within 10 s after the start, the trial was terminated when the light turned off. The trial was repeated
553 every 20 s. Each daily session continued for 30 min, and the sessions were conducted for 4
554 consecutive days. The number of lever presses in each session was counted. In the final step (two-
555 alternative auditory discrimination task), each trial was started by the presentation of a high (10
556 kHz) or low (2 kHz) frequency tone as an instruction cue in a pseudorandom manner. After three
557 seconds, the room light was illuminated, and two retractable levers were inserted at the same time.
558 The rats were required to press the left or right levers in response to the low or high tone cues,
559 respectively. When the rats pressed the correct lever, the reward was presented with the sound, and
560 the levers were immediately retracted. Correct response allowed the rats to access the reward: the
561 retractable spout delivering water was inserted for 1 s for the PET imaging experiment; a pellet was
562 delivered by a food dispenser for the pharmacological experiment; and three drops of water were
563 delivered through a solenoid valve for the electrophysiological experiment. The room light turned
564 off 4 s after the correct response. On the other hand, error response resulted in the retraction of
565 levers and the turning off of the room light. No response within 10 s after the lever insertion was
566 recorded as an omission response. The success rate in each session was calculated by dividing the
567 number of correct response trials with the number of lever press trials minus the number of
568 omission trials. The trial was repeated every 20 s. Each daily session continued for 60 min, and the
569 sessions were conducted for 14 or 24 days according to the experiments.

570 To assess the behavioral strategies in the auditory discrimination task, a trial-by-trial
571 analysis was conducted; the WSW strategy was defined as a correct response, in which, after a
572 correct response during the previous trial, the rats press the opposite lever during the current trial in
573 response to the shift of the tone instruction cue; the LSL strategy was defined as an error response,

574 in which, after an error response during the previous trial, the rats press the opposite lever during
575 the current trial despite the shift of the instruction cue. The proportion of both strategies was
576 calculated by dividing the number of each strategy with the number of lever press trials minus the
577 number of omission trials.

578 For the transient inactivation experiment, MUS injections were bilaterally performed 20
579 min before start of the session. For habituation, all rats were injected with saline bilaterally into the
580 corresponding striatal subregions on the last day (Day 4) of the single lever press task. During the
581 auditory discrimination task, the first injection was conducted on Day 2 of the task. The second and
582 third injections were conducted the next day, after the success rate had exceeded 60% and 80%,
583 respectively. Each session continued for 15 min.

584 For the electrophysiological experiments, a delay period (0.5 ± 0.2 s) was added in the
585 auditory discrimination task between the lever press and reward sound to distinguish the motor- and
586 reward-related responses of neurons. When the rats pressed the lever, the lever was immediately
587 retracted. After the delay period, the reward was presented with a sound. The room light turned off
588 4 s after the lever press. The trial was repeated every 20 s, and each daily session continued for 60
589 min.

590

591 **PET imaging analysis**

592 An ^{18}F-FDG-PET scan was performed with a microPET Focus220 (Siemens Medical Solutions,
593 Knoxville) designed for high-resolution imaging of small animals (spatial resolution of 1.4 mm in
594 full width at half maximum at the center of the field of view) as described previously[26]. Serial ^{18}F-
595 FDG-PET scans were performed on the last day of the single lever press task and on Days 2, 6, 10,
596 and 24 of the auditory discrimination task. Under the awake condition, a solution of ^{18}F-FDG (ca.
597 74 MBq/0.4 mL) was intravenously injected just before the beginning of each behavioral session
598 through an indwelling catheter attached to the tail. After the 30-min session, the rats were returned
599 to their home cage, and then a 30-min static PET scan was performed under anesthesia with 2.0–

600 2.5% isoflurane 55-min after the ^{18}F-FDG injection. During the PET scan, the body temperature of
601 the anesthetized rats was maintained at 37°C using a small animal warmer (BWT-100A, Bio
602 Research Center). The acquired emission data were sorted into a single sinogram and reconstructed
603 by a standard 2D filtered back projection (FBP) with a ramp filter and cutoff frequency at 0.5 cycles
604 per pixel or by a statistical maximum a *posteriori* probability algorithm (MAP), 12 iterations with
605 point spread function effect.

606 Voxel-based analysis of PET images was conducted according to the method described
607 previously[26]. Briefly, individual MAP-reconstructed FDG (MAP-FDG) images were aligned to a
608 FDG template image of Long Evans rats using PMOD software package (version 3.2, PMOD
609 Technologies), and then the aligned images were transformed into the space of a T2 weighted MRI
610 templates for the rats (https://www.nitrc.org/projects/tpm_rat/). The transformation parameters
611 obtained from individual MAP-FDG images were used to match the individual FBP-reconstructed
612 FDG (FBP-FDG) images with the MRI templates. The voxel size of the template was magnified by
613 10 to a size similar to that of a human brain. All FBP-FDG images were resampled with a voxel size
614 of 1.2 × 1.2 × 1.2 mm and spatially smoothed through an isotropic Gaussian kernel (6-mm FWHM).
615 Voxel-based statistical analysis was assessed by using the Statistical Parametric Mapping software
616 (version 8; https://www.fil.ion.ucl.ac.uk/spm/). The voxel of interest (VOI) was determined as
617 voxels with more than 85% of the maximum T-value within activated areas obtained by voxel-based
618 statistical comparison of FBP-FDG images. The average regional ^{18}F-FDG uptake was calculated
619 on the basis of the values of the VOI in FBP-FDG images.

620

621 **Electrophysiological recording**

622 Extracellular multi-unit recording was conducted at a sampling rate of 20 kHz from the striatum of
623 freely moving rats by using a 512-ch acquisition board (Open Ephys) via 64-ch recording
624 headstages (C3325; Intan Technologies). The tips of the probe were lowered by turning the screw (>
625 140 μm/day) until it reached the target region (the aDLS and pVLS) during the single lever press

task. The tips were sometimes lowered by the same procedure (140 to 280 μm/day) after the daily recording session to improve cell yield.

Spike sorting was performed automatically by the Kilosort software (https://github.com/cortex-lab/KiloSort). Clear noise clusters were removed in the first manual sorting using the phy GUI (https://github.com/cortex-lab/phy). The remaining clusters were then re-clustered by using the Klusta Kwik (https://klusta.readthedocs.io/en/latest/), and the clusters were adjusted in the second manual sorting by using the phy GUI again. The sorted clusters satisfying all the following criteria were used for further analysis: L-ratio < 0.05, isolation distance > 15, and ISI index < 0.2[51,52]. Well-isolated single neurons (n = 1,485; aDLS, n = 1,062; pVLS, n = 423) were classified as putative medium spiny neurons (MSNs), fast spiking interneurons (FSIs), tonically active neurons (TANs), and unclassified interneurons (UIs). Putative MSNs and TANs were separated from FSIs and UIs by trough-to-peak duration of mean spike waveform (> 0.6 ms) according to previous studies, with slight modifications[53,54]. Putative FSIs were then separated from UIs by utilizing the proportion of time associated with long inter-spike intervals. This was achieved by summing the inter-spike intervals longer than 2 s and subsequently dividing the resulting sum by the total recording time ($Prop_{ISIs} > 2$ s)[55]. Neurons with values of $Prop_{ISIs} > 2$ s less than 0.4 were classified as putative FSIs, while neurons with $Prop_{ISIs} > 2$ s greater than 0.4 were classified as putative UIs. The putative MSNs and TANs were separated by measuring the post-spike suppression[55]. We measured the length of suppressed time during which the firing rate of a neuron was suppressed following a spike by counting the number of 1-ms bins in its autocorrelation function until the rate of the neuron was equal to or greater than its averaged rate over the 600-ms to 900-ms autocorrelation bins. Putative TANs were separated by post-spike suppression of less than 40 ms, and the remaining neurons were classified as putative MSNs. Because the numbers of FSIs, TANs, and UIs were low (79 out of 1,485, 5.3%), they were excluded from further analysis.

To divide the three stages during the acquisition phase, the sigmoidal function was fitted using the following four-parameter logistic equation:

$$f_{(x)} = \frac{1 - \gamma - (1 - \lambda)}{1 + \exp\left(-\beta(x - \alpha)\right)} + \gamma$$

Where α is the inflection point, β is a variable for slope, and λ and γ are parameters related to the lower asymptote (γ) and upper asymptote (λ) of the learning curve. The function was used to evaluate the day when the success rate reached 20% ($t_{20\%}$) or 90% ($t_{90\%}$) of the maximum rate change defined by the difference between the γ and λ. Three stages were defined according to the following criteria: early stage (up to 3 days before the middle stage); middle stage (3 days including before and after the day of $t_{20\%}$); and late stage (3 days including the day of $t_{90\%}$ and before). One rat learned the discrimination faster, and the success rate reached $t_{90\%}$ 2 days after the day of $t_{20\%}$. One day before the day of $t_{90\%}$ was assigned to the late stage. For the following analysis, spike data in each striatal subregion were separated into the early, middle, and late stages (n = 295, 364 and 360, respectively, for the aDLS; and n = 87, 110 and 190, respectively, for the pVLS).

To determine the neurons showing significant changes in firing rate related to the task events, electrophysiological data were grouped into four trial types based on the instruction cue (high or low tone) and the choice (right or left lever); HR, high tone and right lever press; HL, high tone and left lever press; LR, low tone and right lever press; and LL, low tone and left lever press. Neurons with three or more trial data were analyzed. We identified the HR or LL type neurons, detecting significant increases or decreases of the firing rate by applying the Wilcoxon rank-sum test ($p < 0.01$) to trial-by-trial spike number between the baseline period (BL, -1,500 to -500 ms or -1,100 to -500 ms before presentation of the instruction cue) and any of four peri-event periods (CO, 0 to +1,000 ms after the onset of the instruction cue; CR, -700 to +300 ms around the time of lever press; RS, -300 to 300 ms around the time of sound onset instructing the reward delivery; and FL, 0 to +1,000 ms after the time of first licking after the CR). Any spikes around the events were averaged in 50-ms bins. The event-related activity was shown as averaged firing rates aligned based on the timing of an event for each trial. The auROC values were calculated by using the spike counts between the HR or LR trials and any of the other three trials. Bins showing significant

differences in the auROC data were determined by applying the Wilcoxon rank-sum test ($p < 0.05$) to the distribution of 200 ms bin-by-bin auROC values and 0.5.

To compare the number of neurons showing event-related activity among different stages, the number of HR or LL type neurons were calculated by matching the trial number at each stage. Twenty trials were randomly selected for a neuron, and neurons with less than 20 trials were removed from the analysis. The neurons with event-related activity were identified as described above. The proportion (percentage) of neurons relative to the total number of neurons was calculated at each individual stage. The same analysis was repeated 30 times and the distribution of the proportion of neurons with event-related activity at the stage was obtained. The cumulative probability against the proportion was calculated using cumulative distribution function (cdfplot, MATLAB code from Mathworks).

Licking responses were detected with an infrared photobeam sensor placed in the 3D printed reward port, and generated timestamps for the onset of each spout contact according to the analog voltage signal. The licking responses corresponding to four kinds of trial in the FL-related neurons were shown as averaged rates (50-ms bins) aligned based on the timing of the CR. To calculate the correlation coefficient between neuronal activity and licking responses, the spikes and licking for 5 s after the FL were counted every trial for each FL-related neuron. Trials in which neurons never responded were removed from both the spike and licking data. The correlation coefficients between the two remaining data were calculated for each neuron (corr, MATLAB code from Mathworks), and neurons showing significant differences were determined by applying the Pearson correlation coefficient ($p < 0.01$).

Histology

After the behavioral experiments, the rats were anesthetized with pentobarbital (50 mg/kg body weight) and isoflurane (3%), then transcardially perfused 4% paraformaldehyde in phosphate-buffer (PB). The brains were post-fixed overnight at 4 °C and cryoprotect with 10, 20, and 30% sucrose in

703 PB. Brain tissues were cut into 30-μm sections by using a cryostat (Leica). For the lesion
704 experiment, cresyl violet staining or immunostaining were performed to define the lesioned areas in
705 the striatum. The sections were stained with a mouse monoclonal anti-NeuN IgG (1:1000 dilution;
706 Millipore Corporate), and then with a goat anti-mouse IgG conjugated with Alexa Fluor 488 (1:400
707 dilution; Invitrogen). The sections were counterstained with 0.3 μM 4',6-diamidino-2-phenylindole
708 (DAPI; Sigma-Aldrich). For transient inactivation experiment, sections were stained by cresyl
709 violet to visualize the positions of guide cannulae. For the electrophysiological experiments, small
710 lesions were made by applying an electrolytic current (3 μA, 10 s) through top and bottom channels
711 of each shank by using the stimulus isolator (A365, World Precision Instrument). Sections through
712 the striatum were stained with a mouse monoclonal anti-TH IgG (1:1,000 dilution; Sigma-Aldrich),
713 and then with a goat anti-mouse IgG conjugated with Alexa Fluor 488 (1:400 dilution; Invitrogen).
714 The recording sites were verified by the staining with DAPI and red fluorescent Nissl stain solution
715 (1:200 dilution; Thermo Fisher Scientific). Fluorescence and immunostained signals were
716 visualized under fluorescence microscopes (BZ-X800; Keyence) and confocal microscope (A1R;
717 Nikon).
718
719 **Statistical analysis**
720 Statistical analysis was performed using MATLAB (MathWorks) and SPSS Statistics (Ver. 27,
721 IBM). Comparisons of data between two groups were performed using the Student's *t*-test (paired),
722 the Wilcoxon signed rank test (paired), and the Wilcoxon rank-sum test (unpaired). Comparisons of
723 data among three groups were constructed using a one-way repeated measures ANOVA, two-way
724 repeated measures ANOVA, or the Kruskal-Wallis test (without assuming normal distribution and
725 homogeneity of variances). All statistical tests were two-sided. The Bonferroni post-hoc test and
726 Mann–Whitney U-test with Bonferroni adjustment were carried out as needed to compare means.
727 For boxplots, the data points below the lower (25th percentile) -1.5 interquartile range (IQR) and
728 above the upper (75th percentile) quartiles +1.5 IQR were regarded as outliers, and they are not

included in the boxplots.

Data availability

All source data underlying the graphs and charts are available in Supplementary Data. All other data are available upon request. The datasets generated and analyzed in this study are deposited on Mendeley data (DOI: 10.17632/ghzr838bdk.1; https://data.mendeley.com/datasets/ghzr838bdk/1).

References

1. Alexander, G.E. & Crutcher, M.D. Functional architecture of basal ganglia circuits: neural substrates of parallel processing. *Trends Neurosci* **13**, 266-271 (1990).
2. Foster, N.N., *et al.* The mouse cortico-basal ganglia-thalamic network. *Nature* **598**, 188-194 (2021).
3. Kim, H.F. & Hikosaka, O. Parallel basal ganglia circuits for voluntary and automatic behaviour to reach rewards. *Brain* **138**, 1776-1800 (2015).
4. Yin, H.H. & Knowlton, B.J. The role of the basal ganglia in habit formation. *Nat Rev Neurosci* **7**, 464-476 (2006).
5. Yin, H.H., *et al.* Dynamic reorganization of striatal circuits during the acquisition and consolidation of a skill. *Nat Neurosci* **12**, 333-341 (2009).
6. Kupferschmidt, D.A., Juczewski, K., Cui, G., Johnson, K.A. & Lovinger, D.M. Parallel, but dissociable, processing in discrete corticostriatal inputs encodes skill learning. *Neuron* **96**, 476-489 e5 (2017).
7. Thorn, C.A., Atallah, H., Howe, M. & Graybiel, A.M. Differential dynamics of activity changes in dorsolateral and dorsomedial striatal loops during learning. *Neuron* **66**, 781-795 (2010).
8. Doyon, J., *et al.* Contributions of the basal ganglia and functionally related brain structures to motor learning. *Behav Brain Res* **199**, 61-75 (2009).
9. Lehéricy, S., *et al.* Distinct basal ganglia territories are engaged in early and advanced motor

755 sequence learning. *Proc Natl Acad Sci U S A* **102**, 12566-12571 (2005).
756 10. Miyachi, S., Hikosaka, O. & Lu, X. Differential activation of monkey striatal neurons in the
757 early and late stages of procedural learning. *Exp Brain Res* **146**, 122-126 (2002).
758 11. Miyachi, S., Hikosaka, O., Miyashita, K., Kárádi, Z. & Rand, M.K. Differential roles of
759 monkey striatum in learning of sequential hand movement. *Exp Brain Res* **115**, 1-5 (1997).
760 12. Dickinson, A. Actions and habits: the development of behavioural autonomy. *Phil Trans R Soc*
761 *Lond B* **308**, 67-78 (1985).
762 13. O'Hare, J.K.*, et al.* Pathway-specific striatal substrates for habitual behavior. *Neuron* **89**, 472-
763 479 (2016).
764 14. Yin, H.H., Knowlton, B.J. & Balleine, B.W. Lesions of dorsolateral striatum preserve outcome
765 expectancy but disrupt habit formation in instrumental learning. *Eur J Neurosci* **19**, 181-189
766 (2004).
767 15. Yin, H.H., Ostlund, S.B., Knowlton, B.J. & Balleine, B.W. The role of the dorsomedial striatum
768 in instrumental conditioning. *Eur J Neurosci* **22**, 513-523 (2005).
769 16. Redgrave, P.*, et al.* Goal-directed and habitual control in the basal ganglia: implications for
770 Parkinson's disease. *Nat Rev Neurosci* **11**, 760-772 (2010).
771 17. Toni, I. & Passingham, R.E. Prefrontal-basal ganglia pathways are involved in the learning of
772 arbitrary visuomotor associations: a PET study. *Exp Brain Res* **127**, 19-32 (1999).
773 18. Seger, C.A. & Cincotta, C.M. The roles of the caudate nucleus in human classification learning.
774 *J Neurosci* **25**, 2941-2951 (2005).
775 19. Kim, H.F. & Hikosaka, O. Distinct basal ganglia circuits controlling behaviors guided by
776 flexible and stable values. *Neuron* **79**, 1001-1010 (2013).
777 20. Featherstone, R.E. & McDonald, R.J. Dorsal striatum and stimulus-response learning: lesions
778 of the dorsolateral, but not dorsomedial, striatum impair acquisition of a simple discrimination
779 task. *Behav Brain Res* **150**, 15-23 (2004).
780 21. Featherstone, R.E. & McDonald, R.J. Lesions of the dorsolateral striatum impair the

acquisition of a simplified stimulus-response dependent conditional discrimination task. *Neuroscience* **136**, 387-395 (2005).

22. Lee, J., Wang, W. & Sabatini, B.L. Anatomically segregated basal ganglia pathways allow parallel behavioral modulation. *Nat Neurosci* **23**, 1388-1398 (2020).
23. Xiong, Q., Znamenskiy, P. & Zador, A.M. Selective corticostriatal plasticity during acquisition of an auditory discrimination task. *Nature* **521**, 348-351 (2015).
24. Guo, L., Weems, J.T., Walker, W.I., Levichev, A. & Jaramillo, S. Choice-selective neurons in the auditory cortex and in its striatal target encode reward expectation. *J Neurosci* **39**, 3687-3697 (2019).
25. Nishizawa, K.*, et al.* Striatal indirect pathway contributes to selection accuracy of learned motor actions. *J Neurosci* **32**, 13421-13432 (2012).
26. Cui, Y.*, et al.* A voxel-based analysis of brain activity in high-order trigeminal pathway in the rat induced by cortical spreading depression. *Neuroimage* **108**, 17-22 (2015).
27. Kornblum, H.I.*, et al.* In vivo imaging of neuronal activation and plasticity in the rat brain by high resolution positron emission tomography (microPET). *Nat Biotechnol* **18**, 655-660 (2000).
28. Zimmer, E.R.*, et al.* [18 F] FDG PET signal is driven by astroglial glutamate transport. *Nat Neurosci* **20**, 393-395 (2017).
29. Featherstone, R.E. & McDonald, R.J. Dorsal striatum and stimulus-response learning: lesions of the dorsolateral, but not dorsomedial, striatum impair acquisition of a stimulus-response-based instrumental discrimination task, while sparing conditioned place preference learning. *Neuroscience* **124**, 23-31 (2004).
30. Featherstone, R.E. & McDonald, R.J. Lesions of the dorsolateral or dorsomedial striatum impair performance of a previously acquired simple discrimination task. *Neurobiol Learn Mem* **84**, 159-167 (2005).
31. Bergstrom, H.C.*, et al.* Dorsolateral striatum engagement interferes with early discrimination learning. *Cell Rep* **23**, 2264-2272 (2018).

807 32. Packard, M.G., Hirsh, R. & White, N.M. Differential effects of fornix and caudate nucleus
808 lesions on two radial maze tasks: evidence for multiple memory systems. *J Neurosci* **9**, 1465-
809 1472 (1989).
810 33. Bathellier, B., Tee, S.P., Hrovat, C. & Rumpel, S. A multiplicative reinforcement learning
811 model capturing learning dynamics and interindividual variability in mice. *Proc Natl Acad Sci*
812 *U S A* **110**, 19950-19955 (2013).
813 34. Nonomura, S.*, et al.* Monitoring and updating of action selection for goal-directed behavior
814 through the striatal direct and indirect pathways. *Neuron* **99**, 1302-1314 e5 (2018).
815 35. Williams, Z.M. & Eskandar, E.N. Selective enhancement of associative learning by
816 microstimulation of the anterior caudate. *Nat Neurosci* **9**, 562-568 (2006).
817 36. Smith, K.S. & Graybiel, A.M. Habit formation coincides with shifts in reinforcement
818 representations in the sensorimotor striatum. *J Neurophysiol* **115**, 1487-1498 (2016).
819 37. Jog, M.S., Kubota, Y., Connolly, C.I., Hillegaart, V. & Graybiel, A.M. Building neural
820 representations of habits. *Science* **286**, 1745-1749 (1999).
821 38. Barnes, T.D., Kubota, Y., Hu, D., Jin, D.Z. & Graybiel, A.M. Activity of striatal neurons
822 reflects dynamic encoding and recoding of procedural memories. *Nature* **437**, 1158-1161
823 (2005).
824 39. Jin, X. & Costa, R.M. Start/stop signals emerge in nigrostriatal circuits during sequence
825 learning. *Nature* **466**, 457-462 (2010).
826 40. Barnes, T.D.*, et al.* Advance cueing produces enhanced action-boundary patterns of spike
827 activity in the sensorimotor striatum. *J Neurophysiol* **105**, 1861-1878 (2011).
828 41. Smith, A.C.W.*, et al.* Opposing roles for striatonigral and striatopallidal neurons in dorsolateral
829 striatum in consolidating new instrumental actions. *Nat Commun* **12**, 5121 (2021).
830 42. Wolff, S.B.E., Ko, R. & Ölveczky, B.P. Distinct roles for motor cortical and thalamic inputs to
831 striatum during motor skill learning and execution. *Sci Adv* **8**, eabk0231 (2022).
832 43. Gremel, C.M.*, et al.* Endocannabinoid modulation of orbitostriatal circuits gates habit

formation. *Neuron* **90**, 1312-1324 (2016).

44. Chen, Z.*, et al.* Direct and indirect pathway neurons in ventrolateral striatum differentially regulate licking movement and nigral responses. *Cell Rep* **37**, 109847 (2021).
45. Hunnicutt, B.J.*, et al.* A comprehensive excitatory input map of the striatum reveals novel functional organization. *eLife* **5**, e19103 (2016).
46. Hintiryan, H.*, et al.* The mouse cortico-striatal projectome. *Nat Neurosci* **19**, 1100-1114 (2016).
47. Paxinos, G. & Watson, C. The rat brain in stereotaxic coordinates IV Ed (Academic Press, 1998).
48. Kitanishi, T., Umaba, R. & Mizuseki, K. Robust information routing by dorsal subiculum neurons. *Sci Adv* **7**, eabf1913 (2021).
49. Endepols, H.*, et al.* Effort-based decision making in the rat: an [18F]fluorodeoxyglucose micro positron emission tomography study. *J Neurosci* **30**, 9708-9714 (2010).
50. Marx, C.*, et al.* Conflict processing in the rat brain: behavioral analysis and functional μPET imaging using [F]fluorodeoxyglucose. *Front Behav Neurosci* **6**, 4 (2012).
51. Fee, M.S., Mitra, P.P. & Kleinfeld, D. Automatic sorting of multiple unit neuronal signals in the presence of anisotropic and non-Gaussian variability. *J Neurosci Methods* **69**, 175-188 (1996).
52. Schmitzer-Torbert, N., Jackson, J., Henze, D., Harris, K. & Redish, A.D. Quantitative measures of cluster quality for use in extracellular recordings. *Neuroscience* **131**, 1-11 (2005).
53. Bakhurin, K.I., Mac, V., Golshani, P. & Masmanidis, S.C. Temporal correlations among functionally specialized striatal neural ensembles in reward-conditioned mice. *J Neurophysiol* **115**, 1521-1532 (2016).
54. Kim, N.*, et al.* A striatal interneuron circuit for continuous target pursuit. *Nat Commun* **10**, 2715 (2019).
55. Peters, A.J., Fabre, J.M.J., Steinmetz, N.A., Harris, K.D. & Carandini, M. Striatal activity topographically reflects cortical activity. *Nature* **591**, 420-425 (2021).

Acknowledgements

We thank Drs. Keiji Ota, Hikaru Yokoyama, and Sho Aoki for instructing the behavioral analysis and Dr. Hideyuki Matsumoto for technical advice and discussion. We thank the Research Support Platform, Osaka Metropolitan University Graduate School of Medicine for microscopic imaging. This work was supported by grants-in-aid for JSPS Fellows (19J01997) and Scientific Research (C) (21K11556) from the Japan Society for the Promotion of Science (S.S.); a grant-in-aid for Scientific Research on Transformative Research Areas (A) Adaptive Circuit Census (21H05244) from the Ministry of Education, Science, Sports, and Culture of Japan (K.K.); the Grant for Casio Science Promotion Foundation, Takeda Science Foundation; and the Nakatomi Foundation (S.S.).

Authors' contributions

S.S., Y.C., and K.K. designed the study. S.S. performed the small-animal neuroimaging, multi-unit recordings, tracing studies, and pharmacological experiments with technical support from all authors; T.O., D.H., Y.W., H.O., K.H., and Y.C. assisted with the neuroimaging studies. K.N. performed the tracing study, and N.S. assisted with the immunostaining. H.M., T.K., and K.M. assisted with the multi-unit extracellular recordings. S.S. and K.K. wrote the manuscript with assistance from all coauthors.

Figures and legends

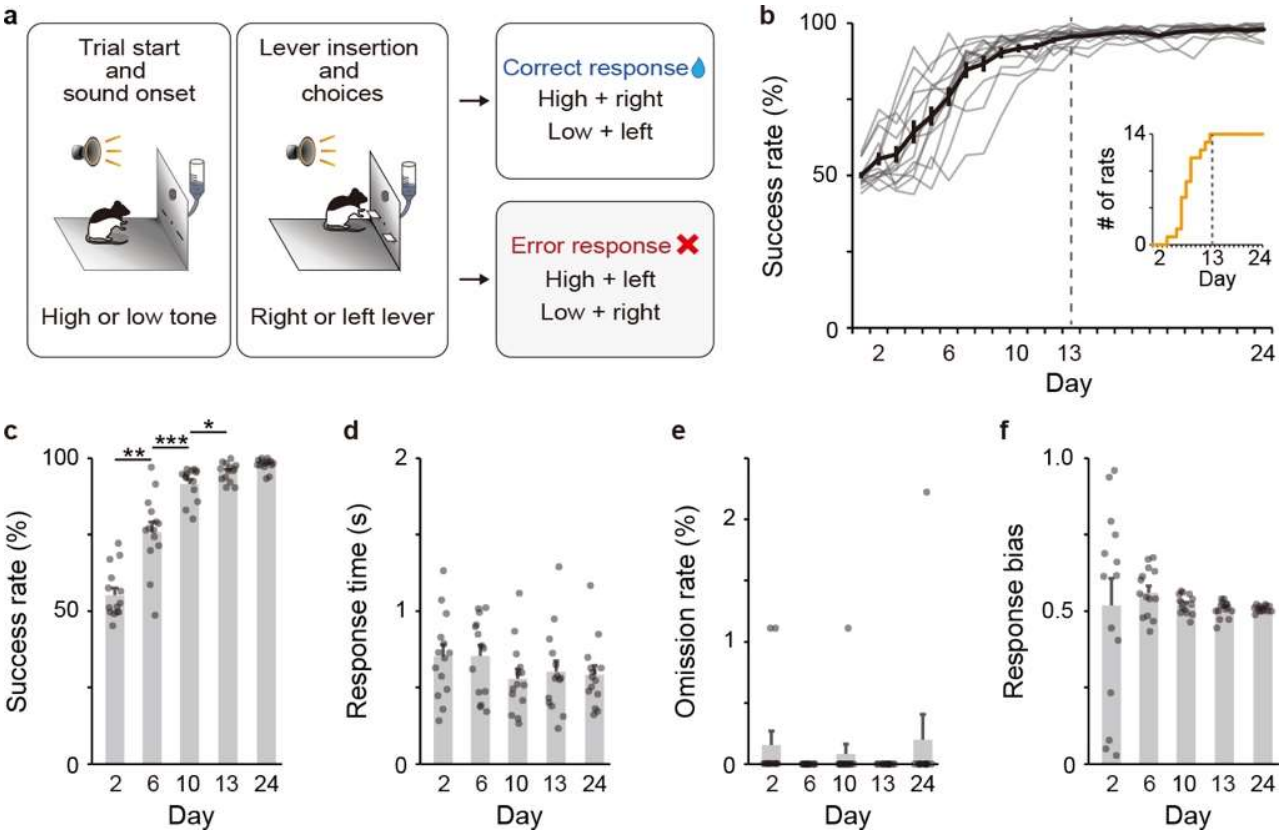

Figure 1. Behavioral performance of two-alternative auditory discrimination learning. **a**, Schematic diagram of the auditory discrimination task. Each trial started by presentation of a tone instruction cue with the high (10 kHz) or low (2 kHz) frequency. Three seconds later, the room light was illuminated, and two retractable levers were inserted at the same time. The rats were required to press the right and left levers in response to the high and low tones, respectively. **b**, Learning curve of the auditory discrimination in intact rats (n = 14). Inset indicates the cumulative curve of number achieved the success rate of more than 90%. **c**, Success rate at Days 2, 6, 10, 13, and 24 (one-way repeated ANOVA, $F[1.897,24.659] = 106.706$, $p = 1.1 \times 10^{-12}$, *post hoc* Bonferroni test, Day 2 vs. Day 6, $p = 0.001$, Day 6 vs. Day 10, $p = 3.6 \times 10^{-4}$, Day 10 vs. Day 13, $p = 0.040$, and Day 13 vs. Day 24, $p = 0.080$). **d**, Response time (one-way repeated ANOVA, $F[2.203,28.642] = 2.386$, $p = 0.105$). **e**, Omission rate ($F[4,52] = 0.699$, $p = 0.596$). **f**, Response bias (one-way repeated ANOVA, $F[1.129,14.674] = 0.250$, $p = 0.653$). Data are indicated as the mean ± s.e.m., and individual data are overlaid. *$p < 0.05$, **$p < 0.01$, and ***$p < 0.001$.

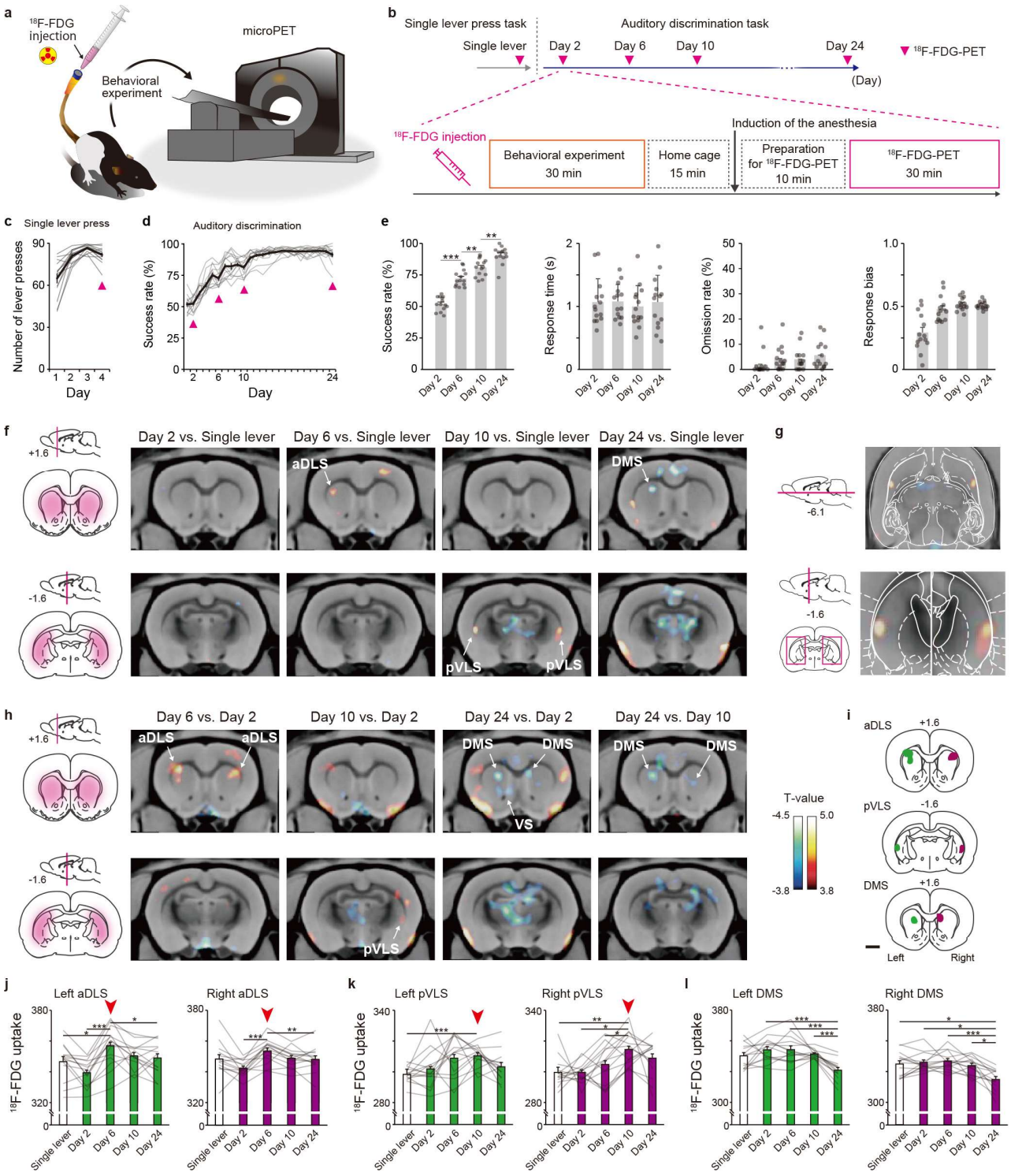

Figure 2. Dissociable brain activity patterns during auditory discrimination learning among three striatal subregions. a, Schematic illustration of an awake rat received intravenously ^{18}F-FDG injection through an indwelling catheter attached to the tail. The rat was conducted for the behavioral experiment and then used for microPET imaging. **b**, Schedule for the behavioral training and ^{18}F-FDG-PET scan. After ^{18}F-FDG injection, rats (n = 14 rats) were subjected to the behavioral

experiment (30 min) and returned to the home cage (15 min). Ten min before the PET scan started, the rats were anesthetized and then the scan was started (30 min). **c**,**d**, Learning curves of the single lever press task (**c**) and auditory discrimination task (**d**). Arrowheads indicate the PET scan days. **e**, Behavioral performance on the scan days during the auditory discrimination. Success rate (one-way repeated ANOVA, $F[3,39] = 125.012$, $p = 4.8 \times 10^{-20}$, *post hoc* Bonferroni test; Day 2 vs. Day 6, $p = 5.0 \times 10^{-7}$; Day 6 vs. Day 10, $p = 0.001$; Day 10 vs. Day 24, $p = 0.003$), response time (one-way repeated ANOVA, $F[3,39] = 0.156$, $p = 0.926$), omission rate (one-way repeated ANOVA, $F[3,39] = 1.559$, $p = 0.215$), and response bias (one-way repeated ANOVA, $F[1.779,23.130] = 17.734$, $p = 3.6 \times 10^{-5}$) are shown. **f**, Representative images of coronal sections that compared the brain activity in the single lever press task with the activity on Day 2, 6, 10, or 24 in the discrimination task. Left panels show schematic illustrations of striatal subregions. **g**, Horizontal (top) and coronal (bottom) images of striatal activation areas shown on Day 10 vs. single lever in (**f**). **h**, Representative images of coronal section that compared the brain activity on Day 2 with that on Day 6, 10, or 24 in the discrimination task. Color bars indicate the T-values, and a value of 3.8 was used as the threshold corresponding to the uncorrected threshold ($p < 0.001$). **i**, Schematic pictures showing voxel of interests for the aDLS, pVLS, and DMS (green, left hemisphere; and purple, right hemisphere). **j–l**, Regional ^{18}F-FDG uptakes in the aDLS (**j**, one-way repeated ANOVA, left aDLS, $F[4,52] = 10.322$, $p = 3.0 \times 10^{-6}$, *post hoc* Bonferroni test; single lever vs. Day 6, $p = 0.016$; Day 2 vs. Day 6, $p = 6.5 \times 10^{-5}$; Day 2 vs. Day 10, $p = 0.019$; Day 2 vs. Day 24, $p = 0.017$; Day 6 vs. Day 24, $p = 0.041$; right aDLS, $F[4,52] = 7.462$, $p = 7.9 \times 10^{-5}$, *post hoc* Bonferroni test, Day 2 vs. Day 6, $p = 2.9 \times 10^{-4}$; Day 2 vs. Day 10, $p = 0.036$; Day 6 vs. Day 24, $p = 0.005$), pVLS (**k**, one-way repeated ANOVA, left pVLS, $F[2.368,30.784] = 4.152$, $p = 0.020$, *post hoc* Bonferroni test; single lever vs. Day 10, $p = 1.8 \times 10^{-5}$; right pVLS, $F[4,52] = 5.995$, $p = 4.8 \times 10^{-4}$, *post hoc* Bonferroni test, single lever vs. Day 10, $p = 0.001$; Day 2 vs. Day 10, $p = 0.011$; Day 6 vs. Day 10, $p = 0.032$), and DMS (**l**, one-way repeated ANOVA, left DMS, $F[4,52] = 12.836$, $p = 2.4 \times 10^{-7}$, *post hoc* Bonferroni test; Day 2 vs. Day 24, $p = 1.8 \times 10^{-5}$; Day 6 vs. Day 24, $p = 4.8 \times 10^{-4}$; Day 10 vs. Day 24, $p = 5.5 \times 10^{-5}$; right DMS, $F[4,52]$

925 = 10.717, $p = 2.0 \times 10^{-6}$, *post hoc* Bonferroni test; single lever vs. Day 24, $p = 0.036$; Day 2 vs. Day
926 24, $p = 0.011$; Day 6 vs. Day 24, $p = 2.6 \times 10^{-4}$; Day 10 vs. Day 24, $p = 0.011$). Arrowheads indicate
927 the day with the most activations throughout the learning process. Data are indicated as the mean ±
928 s.e.m., and individual data are overlaid. The anteroposterior coordinates from bregma (mm) are
929 shown (**f–i**). Scale bar; 2 mm (**i**). *$p < 0.05$, **$p < 0.01$, and ***$p < 0.001$.
930

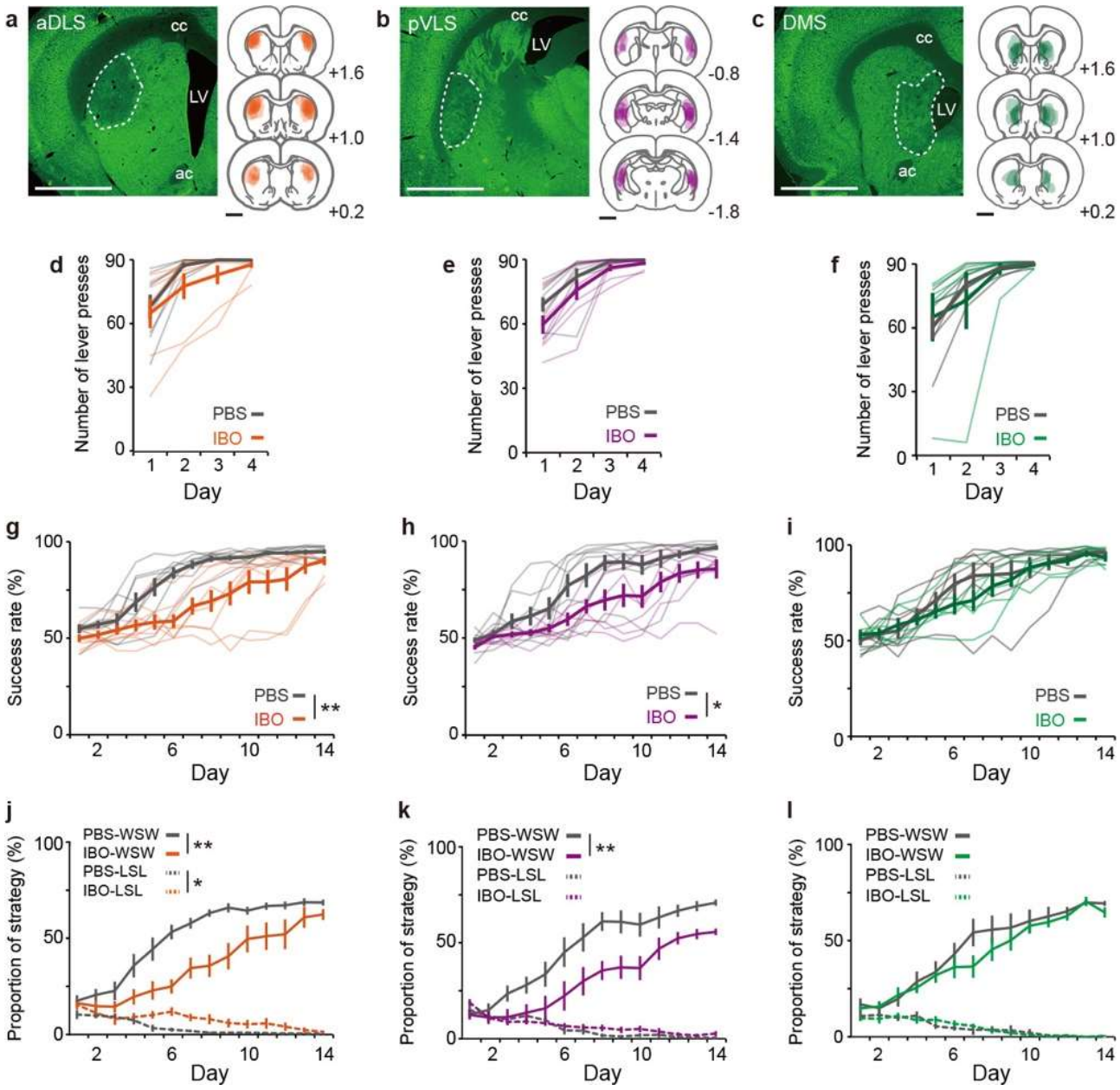

Figure 3. **Impacts of excitotoxic lesion of striatal subregions on the acquisition of auditory discrimination.** Rats were given intracranial injection of PBS or IBO solution into the aDLS (n = 8 for each injection), pVLS (n = 8 for each injection), or DMS (n = 7 for PBS injection, and n = 6 for IBO injection) before start of the single lever press task. **a–c**, Representative images of NeuN immunostaining and individual schematic illustrations showing lesioned area in the aDLS (**a**), pVLS (**b**), or DMS (**c**). Dotted lines in the images indicate the range of the lesioned area. ac, anterior commissure; cc, corpus callosum; and LV, lateral ventricle. **d–f**, Learning curves during the single lever press task in the groups injected into the aDLS **(d,** two-way repeated ANOVA, group, F[1,14]

940 = 1.275, p = 0.278, day, F[1.512,21.171] = 26.281, p = 7.0 × 10^{-6}, group × day, F[1.512,21.171] =
941 0.834, p = 0.418), pVLS (**e**, two-way repeated ANOVA, group, F[1,14] = 2.542, p = 0.133, day,
942 F[1.547,21.658] = 40.183, p = 2.2 × 10^{-7}, group × day, F[1.547,21.658] = 0.953, p = 0.380), or DMS
943 (**f**, two-way repeated ANOVA, group, F[1,11] = 0.025, p = 0.876, day, F[1.205,13.252] = 14.818, p
944 = 0.001, group × day, F[1.205,13.252] = 0.610, p = 0.478). **g–i**, Learning curves during the auditory
945 discrimination task in the groups injected into the aDLS (**g**, two-way repeated ANOVA, group,
946 F[1,14] = 11.578, p = 0.004, day, F[2.805,39.274] = 60.733, p = 1.8 × 10^{-14}, group × day,
947 F[2.805,39.274] = 3.863, p = 0.018), pVLS (**h**, two-way repeated ANOVA, group, F[1,14] = 7.221,
948 p = 0.018, day, F[3.646,51.041] = 51.794, p = 9.2 × 10^{-17}, group × day, F[3.646,51.041] = 1.612, p =
949 0.190), or DMS (**i**, two-way repeated ANOVA, group, F[1,11] = 0.469, p = 0.507, day,
950 F[2.677,29.452] = 48.860, p = 3.9 × 10^{-11}, group × day, F[2.677,29.452] = 1.094, p = 0.362). **j–l**,
951 Changes in the proportion of the WSW or LSL strategy through the discrimination learning in the
952 groups treated into the aDLS (**j**, two-way repeated ANOVA, group, F[1,14] = 13.451, p = 0.003, day,
953 F[3.048,42.673] = 54.777, p = 9.1 × 10^{-15}, group × day, F[3.048,42.673] = 3.288, p = 0.029 for the
954 WSW; group, F[1,14] = 5.039, p = 0.041, day, F[3.866,54.119] = 14.578, p = 5.2 × 10^{-8}, group × day,
955 F[3.866,54.119] = 2.471, p = 0.057 for the LSL), pVLS (**k**, two-way repeated ANOVA, group,
956 F[1,14] = 9.251, p = 0.009, day, F[3.206,44.881] = 47.168, p = 2.7 × 10^{-14}, group × day,
957 F[3.206,44.881] = 1.862, p = 0.146 for the WSW; group, F[1,14] = 0.762, p = 0.397, day,
958 F[4.773,66.815] = 24.018, p = 1.9 × 10^{-13}, group × day, F[4.773,66.815] = 1.338, p = 0.260 for the
959 LSL), or DMS (**l**, two-way repeated ANOVA, group, F[1,11] = 0.703, p = 0.420, day, F[3.052,33.568]
960 = 48.424, p = 2.3 × 10^{-12}, group × day, F[3.052,33.568] = 0.846, p = 0.480 for the WSW; group,
961 F[1,11] = 0.002, p = 0.965, day, F[3.442,37.866] = 13.787, p = 1.0 × 10^{-6}, group × day, F[3.442,
962 37.866] = 0.581, p = 0.654 for the LSL). Data are indicated as the mean ± s.e.m., and individual data
963 are overlaid (except for panels **j–l**). The anteroposterior coordinates from bregma (mm) are shown
964 (**a–c**). Scale bars; 2 mm (**a–c**). *p < 0.05 and **p < 0.01.

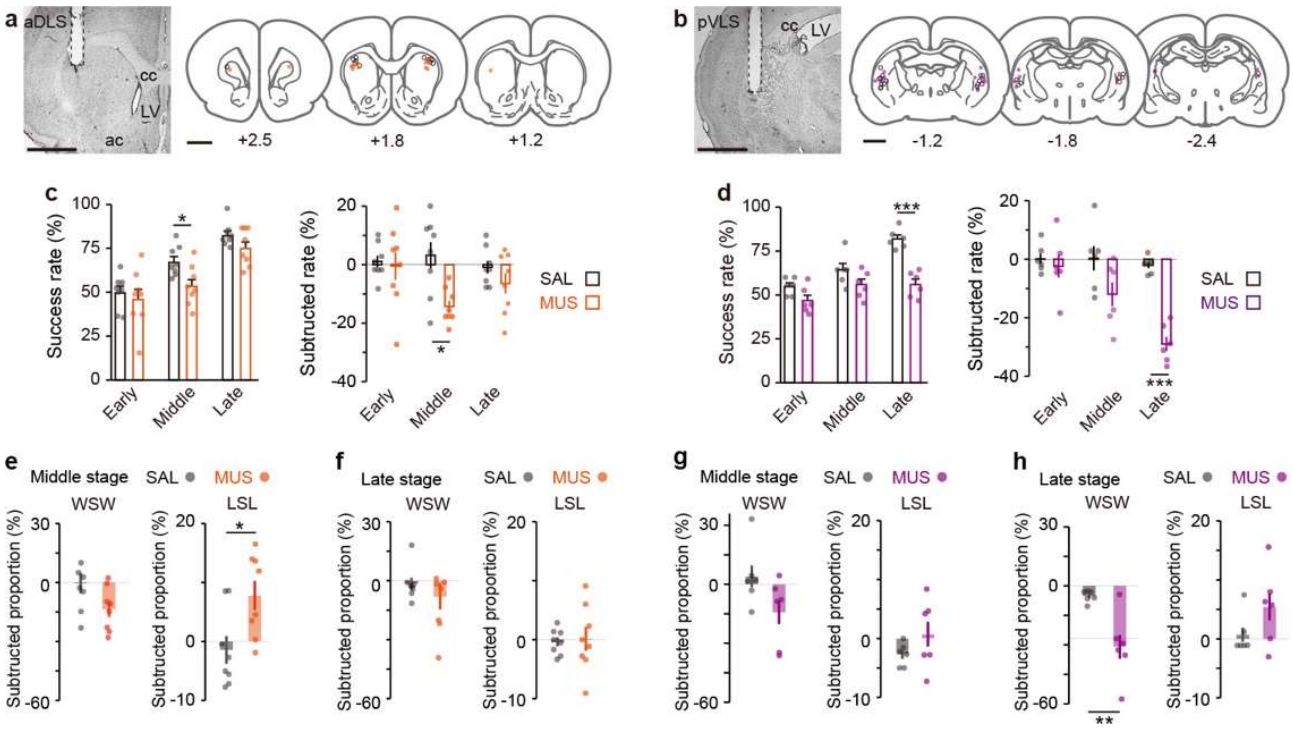

Figure 4. Influence of transient inhibition of striatal subregions at different timings on the performance of auditory discrimination. Rats received intracranial injection of SAL or MUS solution into the aDLS (n = 8 for each injection) or pVLS (n = 6 for each injection). **a**,**b**, Representative images of cresyl violet staining and schematic illustrations showing placement sites of the tip of guide cannula in the aDLS (**a**) or pVLS (**b**). Dotted lines in the images indicate the position of the cannula placement. ac, anterior commissure; cc, corpus callosum; and LV, lateral ventricle. **c**,**d**, Effects of transient striatal inhibition on the performance. For aDLS inhibition (**c**), success rate (left, two-way repeated ANOVA, stage, $F[2,28] = 28.708$, $p = 1.7 \times 10^{-7}$, group, $F[1,14] = 8.840$, $p = 0.010$, stage × group, $F[2,28] = 0.757$, $p = 0.478$; unpaired Student's t-test, early, $t[14] = 0.558$, $p = 0.586$; middle, $t[14] = 2.764$, $p = 0.015$; Welch's t-test, late, $t[12.010] = 1.617$, $p = 0.132$, $p = 0.016$ after a Bonferroni correction method) and subtracted success rate on day N-1 from that on day N (right, two-way repeated ANOVA, stage, $F[2,28] = 1.231$, $p = 0.307$, group, $F[1,14] = 10.797$, $p = 0.005$, stage × group, $F[2,28] = 2.360$, $p = 0.113$; unpaired Student's t-test, early, $t[14] = 0.285$, $p = 0.780$, late, $t[14] = 1.271$, $p = 0.225$, Welch's t-test, middle, $t[9.603] = 3.331$, $p = 0.008$, $p = 0.016$ after a Bonferroni correction method). For pVLS inhibition (**d**), success rate (left, two-way repeated ANOVA, stage, $F[2,20] = 17.642$, $p = 3.8 \times 10^{-5}$, group, $F[1,10] = 43.942$, $p = 5.9 \times 10^{-5}$, stage ×

982 group, $F[2,20] = 4.729$, $p = 0.012$; simple main effect, early, $p = 0.055$, middle, $p = 0.109$, late, $p =$
983 4.2×10^{-5}) and subtracted success rate on day N-1 from that on day N (right, two-way repeated
984 ANOVA, stage, $F[2,20] = 12.105$, $p = 3.6 \times 10^{-4}$, group, $F[1,10] = 16.310$, $p = 0.002$, stage × group,
985 $F[2,20] = 8.500$, $p = 0.002$; simple main effect, early, $p = 0.595$, middle, $p = 0.085$, late, $p = 3.0 \times 10^{-}$
986 ^{6}). **e**,**f**, Subtracted proportion of the WSW or LSL strategy in the groups injected into the aDLS at the
987 middle stage (**e**, unpaired Student's t-test, $t[14] = 2.038$, $p = 0.061$ for the WSW, and $t[14] = 2.714$,
988 $p = 0.017$ for the LSL) and the late stage (**f**, unpaired Student's t-test, $t[14] = 0.898$, $p = 0.384$ for the
989 WSW, and $t[14] = 0.226$, $p = 0.824$ for the LSL). **g**,**h**, Subtracted proportion of the WSW or LSL
990 strategy in the groups injected into the pVLS at the middle stage (**g**, unpaired Student's t-test, $t(10) =$
991 1.924, $p = 0.083$ for the WSW, and Welch's t-test, $t[6.095] = 1.364$, $p = 0.221$ for the LSL) and the
992 late stage (**h**, unpaired Student's t-test, $t[10] = 3.629$, $p = 0.005$ for the WSW, and $t[10] = 1.577$, $p =$
993 0.146 for the LSL). Data are indicated as the mean ± s.e.m., and individual data are overlaid. The
994 anteroposterior coordinates from bregma (mm) are shown (**a**,**b**). Scale bars; 2 mm (**a**,**b**). *$p < 0.05$,
995 **$p < 0.01$, and ***$p < 0.001$.

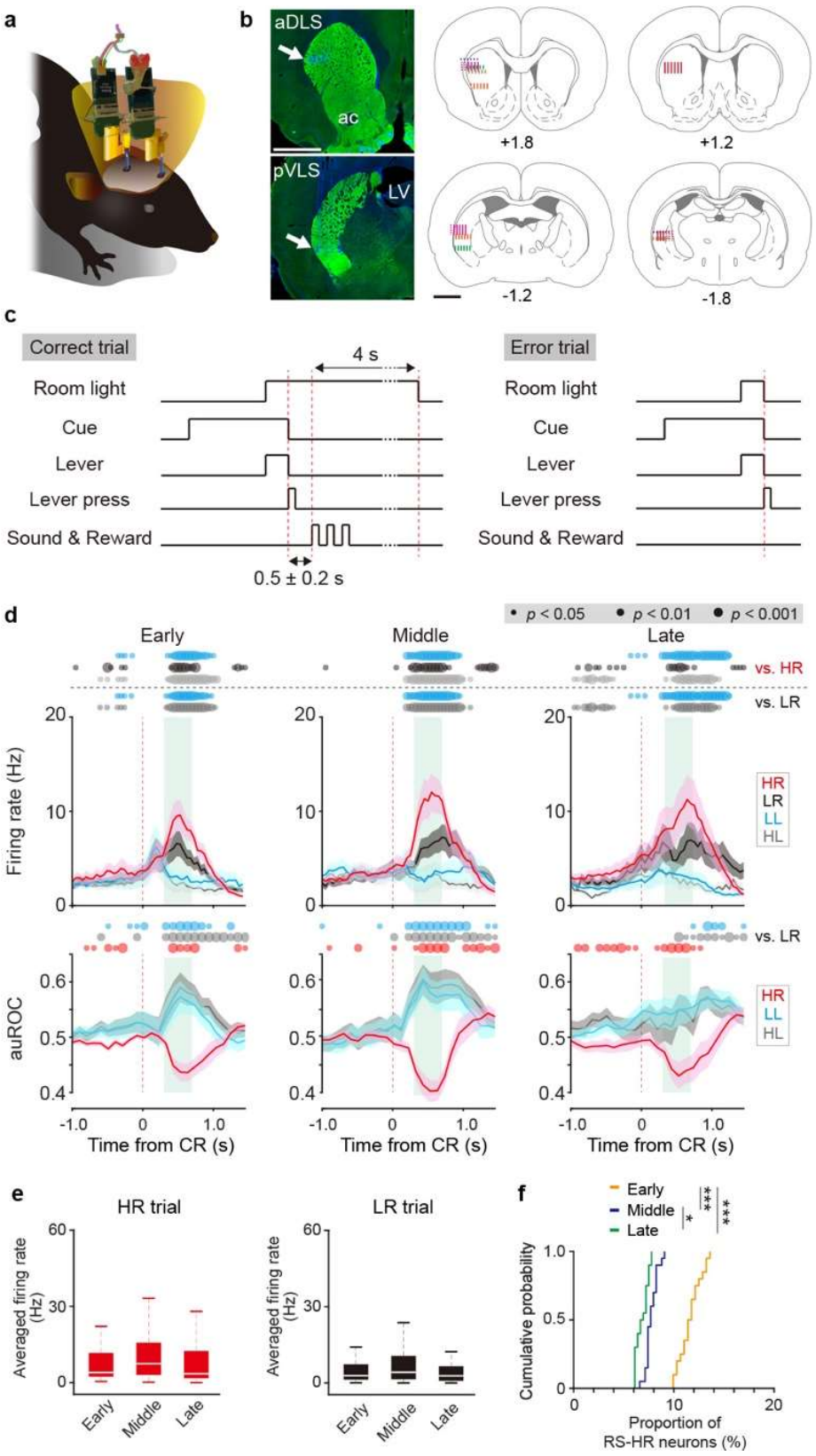

Figure 5. Multi-unit recording of neurons in striatal subregions and firing activity related to the behavioral outcome of RS-HR neurons in the aDLS. a, Schematic illustration of a freely moving rat used for simultaneous multi-unit recordings in the aDLS and pVLS. Six rats were used

for the following analysis. **b**, Representative images showing positions of electrode tips (arrows) in the aDLS and pVLS (left rows). The recording sites estimated by electrode tracks and electrical marks in individual rats are shown (center and right rows). ac, anterior commissure; LV, lateral ventricle. **c**, Sequence of some events in correct and error trials. The delay periods were pseudorandomly added between a correct lever press and the reward sound (0.5 ± 0.2 s). The room light turned off the extended time (4 s) after the correct response or immediately after the error response. **d**, Mean firing rate (top rows) and auROC values (bottom rows) of RS-HR neurons in the aDLS during the period when the RS is presented (green shadows, 500 ± 200 ms after the lever press) at the early (n = 43 neurons), middle (n = 37 neurons), and late (n = 44 neurons) stages. Time bins with significant differences between the mean firing rate in the HR or LR trial and any of the rates in other three trials (Wilcoxon signed rank test) or the distribution of auROC values and 0.5 (Wilcoxon rank-sum test) are represented by the circles at the top. **e**, Averaged firing rate during the RS period in the HR and LR trials at the early, middle, and late stages (Kruskal-Wallis test: HR trial, $\chi^2 = 3.028$, $p = 0.220$; and LR trial, $\chi^2 = 1.249$, $p = 0.536$). **f**, Cumulative probability of the proportion of RS-HR neuron number at the three stages (Kruskal-Wallis test, $\chi^2 = 45.378$, $p = 1.4 \times 10^{-10}$; *post hoc* Tukey-Kramer test, $p = 7.1 \times 10^{-5}$ for early vs. middle, $p = 1.0 \times 10^{-9}$ for early vs. late, and $p = 0.040$ for middle vs. late). Data are indicated as the mean ± s.e.m. (**d**) or the median and quartiles with the maximal and minimal values (**e**). The anteroposterior coordinates from bregma (mm) are shown (**b**). Scale bars; 2 mm (**b**). *$p < 0.05$ and ***$p < 0.001$.

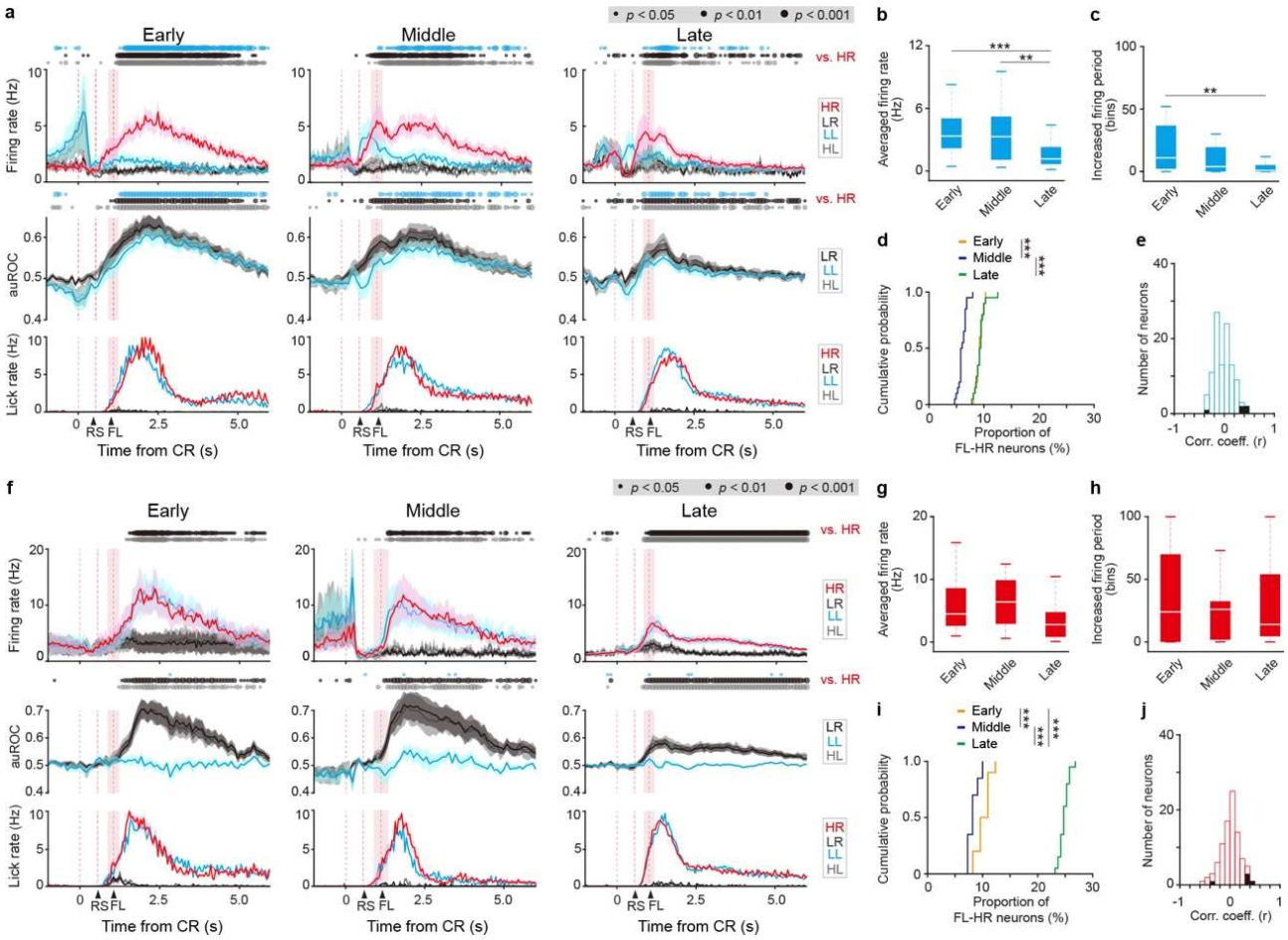

Figure 6. Sustained activity after the reward of FL-HR neurons in the aDLS and pVLS. **a**, Firing rate (top rows), auROC (middle rows), and licking rate (bottom rows) after the reward of FL-HR neurons in the aDLS at the early (n = 27 neurons), middle (n = 31 neurons), and late (n = 49 neurons) stages. Time bins with significant differences between the mean firing rate in the HR trial and any of the rate in other three trials (Wilcoxon signed rank test) or the distribution of auROC values and 0.5 (Wilcoxon rank-sum test) are represented by the circles at the top. Timings of the RS and FL are shown. **b**, Averaged firing rate for 5 s after the FL in the HR trial (Kruskal-Wallis test, χ^2 = 16.396, p = 2.8 × 10^{-4}, *post hoc* Tukey-Kramer test, early vs. middle, p = 0.726, early vs. late, p = 7.2 × 10^{-4}, middle vs. late, p = 0.009). **c**, Averaged total number of time bins above 3 S.D. of the baseline firing of aDLS neurons for 5 s after the FL in the HR trial (Kruskal-Wallis test, χ^2 = 12.803, p = 0.002, *post hoc* Tukey-Kramer test, early vs. middle, p = 0.173, early vs. late, p = 0.001, middle vs. late, p = 0.219). **d**, Cumulative probability of the proportion of FL-HR neuron number (Kruskal-Wallis test, χ^2 = 39.280, p = 3.0 × 10^{-9}, *post hoc* Tukey-Kramer test, early vs. middle, p = 1.3 × 10^{-7}, early vs. late,

$p = 0.995$, middle vs. late, $p = 2.3 \times 10^{-7}$). **e**, Distribution of the correlation coefficient between the numbers of spikes of FL-HR neurons and licking. Closed column indicates the number of neurons showing significant correlations between the two parameters (5 out of 107 neurons). **f**, Firing rate (top rows), auROC (middle rows), and licking rate (bottom rows) after the reward of FL-HR neurons in the pVLS at the early ($n = 17$ neurons), middle ($n = 11$ neurons), and late ($n = 64$ neurons) stages. Time bins with significant differences between the rate in the HR trial and any of the rate in other trials (Wilcoxon signed rank test) or the distribution of auROC values and 0.5 (Wilcoxon rank-sum test) are represented by the circles at the top. The timings of the RS and FL are shown. **g**, Averaged firing rate of pVLS neurons after the FL in the HR trial (Kruskal-Wallis test, $\chi^2 = 7.652$, $p = 0.022$; *post hoc* Tukey-Kramer test, early vs. middle, $p = 0.975$, early vs. late, $p = 0.076$, and middle vs. late, $p = 0.096$). **h**, Averaged total number of time bins above 3 S.D. of the baseline firing after the FL in the HR trial (Kruskal-Wallis test, $\chi^2 = 0.023$, $p = 0.989$). **i**, Cumulative probability of the proportion of FL-HR neuron number (Kruskal-Wallis test, $\chi^2 = 47.922$, $p = 3.9 \times 10^{-11}$; *post hoc* Tukey-Kramer test, early vs. middle, $p = 0.011$, early vs. late, $p = 1.7 \times 10^{-4}$, and middle vs. late, $p = 9.7 \times 10^{-10}$). **j**, Distribution of the correlation coefficient between the number of spikes in the pVLS and the licking. Closed column indicates the number of neurons showing significant correlations between the two parameters (5 out of 92 neurons). Data are indicated as the mean ± s.e.m. (**a**,**f**) or the median and quartiles with the maximal and minimal values (**b**,**c**,**g**,**h**). $^{**}p < 0.01$ and $^{***}p < 0.001$.

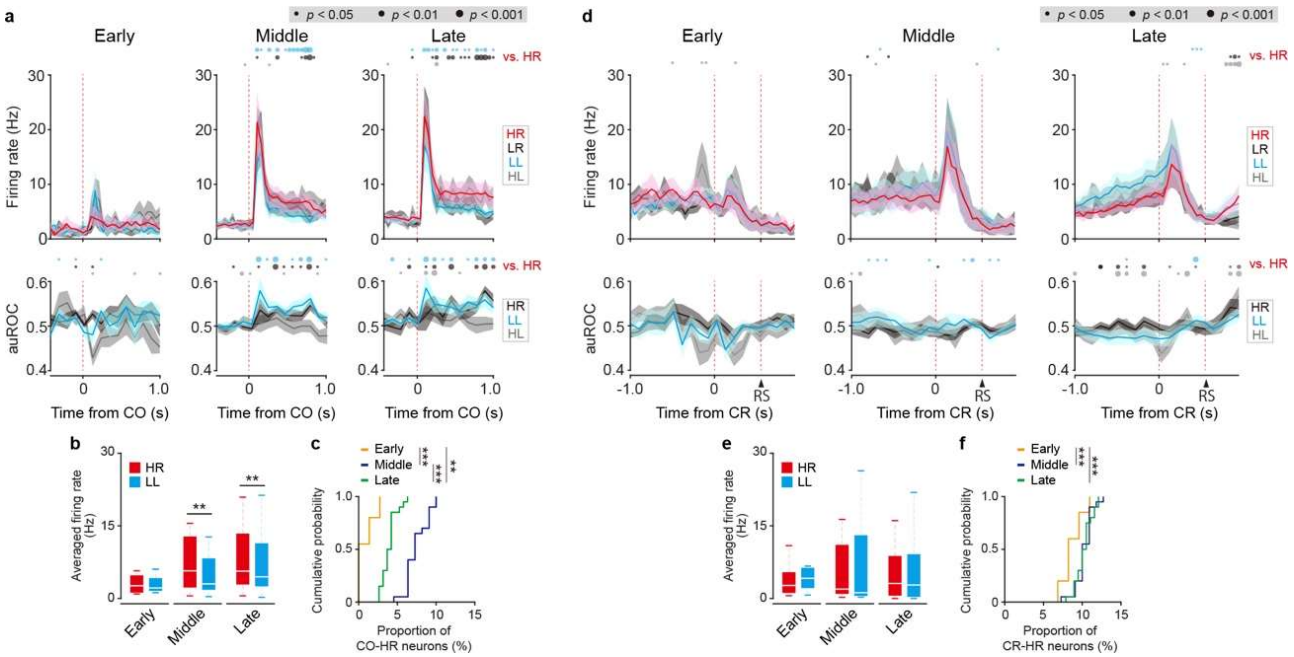

Figure 7. Transient activity related to the beginning and ending of a behavior of CO-HR and CR-HR neurons in the pVLS. a, Firing rate (top rows) and auROC values (bottom rows) of CO-HR neurons in the pVLS at the early (n = 4 neurons), middle (n = 21 neurons), and late (n = 24 neurons) stages. Time bins with significant differences between the mean firing rate in the HR trial and either of the rate in other three trials (Wilcoxon signed rank test) or the distribution of auROC values and 0.5 (Wilcoxon rank-sum test) are represented by the circles at the top. **b**, Averaged firing rate of CO-HR neurons in HR and LL trials (0-600 ms after the CO) (Wilcoxon signed rank test; early, p = 0.875, middle, p = 0.006, and late, p = 0.004). **c**, Cumulative probability of the proportion of CO-HR neuron number (Kruskal-Wallis test, χ^2 = 51.114, p = 8.0 × 10^{-12}; *post hoc* Tukey-Kramer test, early vs. middle, p = 9.6 × 10^{-10}, early vs. late, p = 0.002, and middle vs. late, p = 6.4 × 10^{-4}). **d**, Firing rate (top rows) and auROC values (bottom rows) of CR-HR neurons in the pVLS at the early (n = 11 neurons), middle (n = 22 neurons), and late (n = 41 neurons) stages. Time bins with significant differences between the mean rate in the HR trial and any of the rate in other trials (Wilcoxon signed rank test) or the distribution of auROC values and 0.5 (Wilcoxon rank-sum test) are represented by the circles at the top. The timing of the RS is presented. **e**, Averaged firing rate of CR-HR neurons in HR and LL trials (0-600 ms after the CR) (Wilcoxon signed rank test; early, p = 0.278, middle, p = 0.758, and late, p = 0.279). **f**, Cumulative probability of the proportion of CR-HR neuron number

1071 (Kruskal-Wallis test, $\chi^2 = 13.786$, $p = 0.001$; *post hoc* Tukey-Kramer test, early vs. middle, $p = 0.003$,

1072 early vs. late, $p = 0.005$, and middle vs. late, $p = 0.976$). Data are indicated as the mean ± s.e.m. (**a**,**d**)

1073 or the median and quartiles with the maximal and minimal values (**b**,**e**). $^{**}p < 0.01$ and $^{***}p < 0.001$.