



HHS Public Access

Author manuscript

Cell. Author manuscript; available in PMC 2019 June 28.

Published in final edited form as:

Cell. 2018 June 28; 174(1): 32–43.e15. doi:10.1016/j.cell.2018.06.012.

Optogenetic editing reveals the hierarchical organization of learned action sequences

Claire E. Geddes^{1,2}, Hao Li¹, and Xin Jin^{1,3,*}

¹Molecular Neurobiology Laboratory, The Salk Institute for Biological Studies, La Jolla, CA 92037, USA

²Neurosciences Graduate Program, University of California, San Diego, La Jolla, CA 92093, USA

SUMMARY

The organization of action into sequences underlies complex behaviors that are essential for organismal survival and reproduction. Despite extensive studies of innate sequences in relation to central pattern generators, how learned action sequences are controlled, and whether they are organized as a chain or hierarchy remain largely unknown. By training mice to perform heterogeneous action sequences, here we demonstrate that striatal direct and indirect pathways preferentially encode different behavioral levels of sequence structure. State-dependent closed-loop optogenetic stimulation of the striatal direct pathway can selectively insert a single action element into the sequence without disrupting the overall sequence length. Optogenetic manipulation of the striatal indirect pathway completely removes the ongoing subsequence while leaving the following subsequence to be executed with the appropriate timing and length. These results suggest that learned action sequences are not organized in a serial but rather a hierarchical structure that is distinctly controlled by basal ganglia pathways.

ETOC BLURB

Interrogation of basal ganglia circuits during complex behavior unveil the hierarchical structure of learned action sequences supported distinctly by striatal direct and indirect pathways.

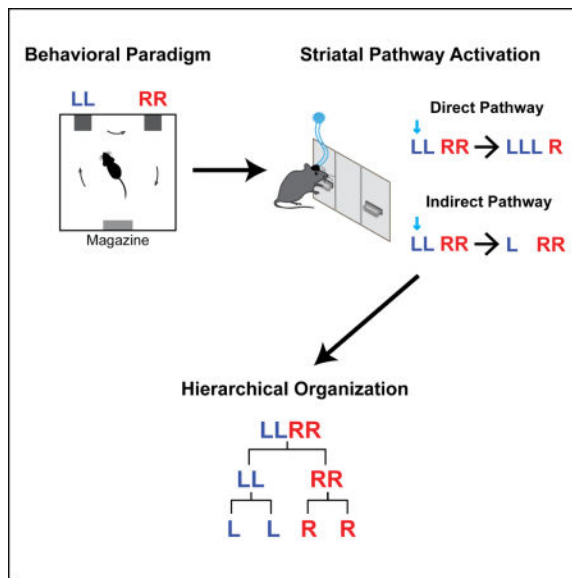
*Correspondence to: xjin@salk.edu.

³Lead Contact

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Author Contributions: X.J. conceived the project. X.J. and C.G. designed the experiments. C.G. performed the experiments and analyzed the data. H.L. conducted and analyzed the electrophysiological recordings. C.G. and X.J. wrote the paper.

Declaration of interests: The authors declare no competing interests.



INTRODUCTION

Action sequences form the basic functional units of behavior and contribute to the numerous acquired motor repertoires observed in animals and humans (Brainard and Doupe, 2002; Gallistel, 1980; Graybiel, 1998; Hikosaka et al., 1999; Jin and Costa, 2015). Many motor disorders, including Parkinson's and Huntington's diseases, are compromised in both learning new action sequences and executing previously learned sequences (Agostino et al., 1992; Hikosaka et al., 1999; Jin and Costa, 2015; Vinter and Gras, 1998). Early theories suggested that action sequences are organized as response chains, activated in series by reflex-like processes based on sensory feedback or efference copies (Sherrington, 1906). In contrast, other theories propose that action sequences might be organized hierarchically with multiple layers of control at the individual element, intermediate subsequence, and overall sequence levels (Gallistel, 1980; Jin and Costa, 2015; Lashley, 1951). Compared to a serial chain, a hierarchical organization is more error-tolerant at the cost of requiring multiple controllers at different hierarchies (Gallistel, 1980; Jin and Costa, 2015; Lashley, 1951). Still, exactly how a learned action sequence is organized remains inconclusive, and the neural substrates supporting this sequence structure are largely unknown (Graybiel, 1998; Hikosaka et al., 1999; Jin and Costa, 2015).

Here we developed a novel behavioral paradigm by training mice to perform spatiotemporally heterogeneous action sequences. It was found that sequence learning takes place in a non-back-propagation manner that critically depends on striatal NMDA receptors. By employing *in vivo* neuronal recording with cell-type-specific optogenetic-tagging, we found that although both striatal pathways are involved in element-level action control, the direct pathway preferentially signals sequence-level initiation/termination while the indirect pathway encodes the switch between subsequences. Consistently, selective diphtheria toxin-mediated ablation of neurons in the striatal direct or indirect pathway impairs sequence initiation and subsequence transitions, respectively. Using state-dependent closed-loop

optogenetic stimulation of striatal direct or indirect pathways, we can selectively insert or remove actions within a learned sequence, without necessarily disrupting the overall sequence structure or the execution of the remaining sequence. These results show that learned action sequences are organized in a hierarchical structure that is dually supported by basal ganglia direct and indirect pathways in distinct ways and have important implications for a wide range of neurological diseases from Parkinson's disease to speech disorders (Brainard and Doupe, 2002; Graybiel, 1998; Hikosaka et al., 1999; Jin and Costa, 2015; Lai et al., 2001; Vinter and Gras, 1998).

RESULTS

Learning heterogeneous action sequences requires striatal NMDA receptors

We developed a new self-paced operant task to investigate the learning and organization of heterogeneous action sequences in mice. In a customized operant chamber with two levers placed opposite a food magazine, mice were trained to press the left and right levers in the specific spatiotemporal combination 'left-left-right-right (LLRR)' (denoted as 'Penguin Dance' sequence (Dance)) to earn a food pellet as reward (Figure 1A, Movie S1, see STAR Methods for details). Extra presses besides this combination did not exclude the reward as long as the sequence contained the consecutive 'LLRR' pattern. The task follows a completely self-paced design, with no experimentally provided cues signaling sequence correctness or reward availability (see STAR Methods for details). At the behavioral level, this 'Penguin Dance' sequence could be organized as either a serial chain 'L → L → R → R', or in a hierarchical structure where L or R action elements are organized into two subsequences – 'LL' and 'RR', which are then concatenated into the target sequence 'LLRR' (Figure 1B) (Gallistel, 1980; Jin and Costa, 2015; Lashley, 1951). With training, mice gradually chunked their lever pressing into robust spatiotemporal sequences with a significant increase in performance speed and decreases in sequence variability (Figures 1C–1E and S1) (Jin and Costa, 2010, 2015). The animals' performance efficiency, measured as the percentage of rewarded lever presses ('LLRR') out of total presses, significantly increased with training (Figure 1F), indicating a progressive learning of the specific 'Penguin Dance' sequence with time.

Analysis of the sequence substructure indicated that the 'RR' subsequence was acquired first, followed by the slow acquisition of the 'LL' subsequence (Figure 1G). Further analysis of the sequence microstructure across learning revealed that at the element level, animals first identified the final and then the penultimate sequence elements as 'R' presses (Figure 1H). The identification of the first element of the sequence as 'L' took place afterwards, followed lastly by correctly identifying the second sequence element as 'L' (Figure 1H). These data indicate that animals tended to chunk actions into subsequences ('RR' and 'LL') before crystalizing these subsequences into the complete target sequence ('LLRR'). Noticeably, the order of element-level action learning is inconsistent with the classic back-propagation rule in reinforcement learning theory (Sutton and Barto, 1998), which predicts action sequence learning takes place in the reverse order of execution. Furthermore we observed that the behavioral pattern 'L – RR' is acquired before '– LRR' ('–' denotes either 'L' or 'R' press in the sequence, Figure S1I). The same order of action learning is also

observed when mice acquire a lever-retraction version of the LLRR task (Figures S1J-S1N; see STAR Methods for details). Together, these learning data are incompatible with the simple back-propagation rule in reinforcement learning theory and underscore the significance of the start and stop elements of a sequence (Jin and Costa, 2010; Murdock, 1962).

NMDA receptors in striatum have been shown to be critical for action learning and corticostriatal plasticity (Calabresi et al., 1992; Dang et al., 2006; Jin and Costa, 2010; Shen et al., 2008). Notably, mice with a striatal-specific deletion of NMDA receptors (referred to as striatal NR1-KO mice) (Dang et al., 2006; Jin and Costa, 2010) showed no improvement in performance efficiency across the four weeks of ‘Penguin Dance’ sequence training (Figure 1F) and did not demonstrate the crystalized spatiotemporal action pattern observed in either wildtype or littermate control animals (Figure S2). Instead, striatal NR1-KO mice developed a consistent right lever bias. While the frequency of executing the right press as the final and penultimate elements of the sequence increased across training, the frequency of executing the left press in the first and second positions decreased rather than increased (Figure 1H). Thus, unlike their littermate controls, striatal NR1-KO mice were not able to chunk action elements into the correct subsequences and crystalize them into the target sequence (Figures 1F and 1G). This selective impairment of sequence learning in striatal NR1-KO mice was not due to differences in reinforcement history or lack of practicing the target sequence, because the same chunking deficits were evident compared to a separate control cohort trained with the amount of reinforcers matched to striatal NR1-KO mice (Figures S2E and S2F). Together, these data suggest that mice learn to chunk actions into heterogeneous sequences in a non-back-propagation manner and NMDA receptors in the striatum are critical for this modular process of sequence learning.

Striatal pathways encode various levels of sequence structure

Impairments in action sequences could result from deficits in sequence initiation and termination (i.e. sequence level), failure to switch from one subsequence to another (i.e. subsequence level), or incorrect execution of a specific action within the sequence (i.e. element level). Since the striatal direct and indirect pathways have been shown to play distinct yet complementary roles in controlling actions (Jin et al., 2014; Kravitz et al., 2010; Tecuapetla et al., 2016), we sought to determine how striatal D1- vs. D2-expressing spiny projection neurons (referred to as dSPNs and iSPNs, respectively) encode a heterogeneous action sequence across different behavioral levels. A ChR2-aided photo-tagging method was employed to record and identify dSPNs vs. iSPNs during the execution of the ‘Penguin Dance’ sequence in D1- and A2a-ChR2 mice (Figures 2A–2H, S3A, and S3B, see STAR Methods for details) (Jin et al., 2014; Lima et al., 2009). Among the task-related neurons (84% of all positively identified neurons, $n = 50$ dSPNs and $n = 44$ iSPNs), over half of dSPNs showed sequence-level start/stop-related activity, which was less frequently observed in iSPNs (Figures 2I–2K). At the element level, over one-quarter of dSPNs showed phasic activation related to each individual press within the sequence, while more iSPNs were inhibited instead (Figures 2L–2N). Notably, some SPNs were selectively active during the transition from the left to the right subsequence (Figure 2O). This “switch-related” activity appeared after the last press in the left subsequence, terminated before the initiation of the

first press in the right subsequence and spanned most of this transition period (Figure 2P). In particular, 31% of iSPNs, compared to only 6% of dSPNs, demonstrated “switch-related” activity (Figure 2Q).

Together, these data suggest that dSPNs and iSPNs encode information related to distinct levels of the sequence structure. Specifically, while dSPNs and iSPNs are both involved in element-level action execution, dSPNs more likely signal sequence initiation and termination, whereas iSPNs preferentially encode the switch between subsequences.

Striatal pathway ablations differently impair learned sequence

We next determined how the different activity patterns observed in dSPNs and iSPNs contribute to action sequence execution. We first verified that ongoing neuronal activity in the dorsal striatum was required for correct execution of the learned ‘Penguin Dance’ sequence by bilateral intra-striatal infusion of a small volume of muscimol (Figures 3A–3C, see STAR Methods for details). Striatal inactivation impaired sequence performance at the sequence (Figure 3A), subsequence (Figure 3B) and element levels (Figure 3C), suggesting striatal activity is necessary for appropriate organization of learned action sequences. To further elucidate the role of specific striatal pathways during sequence performance, we selectively ablated dorsal striatal dSPNs or iSPNs in trained D1- and A2a-cre mice by virally expressing diphtheria toxin receptors (AAV-FLEX-DTR-eGFP) in a cre-dependent manner, followed by diphtheria toxin (DT) injections (Figures 3D, 3E, S3F, and S3G, see STAR Methods for details) (Saito et al., 2001). Bilateral dSPN or iSPN ablation markedly altered sequence behavior, such that dSPN-ablation mice had difficulty initiating the left subsequence while iSPN-ablation mice were impaired in switching from the left to the right subsequence (Figure 3F). Thus, dSPN or iSPN ablation significantly reduced the efficiency of performing the learned LLRR sequence (Figure 3G). Noticeably, the behavioral efficiency of dSPN-ablation mice, but not iSPN-ablation mice, was similar to the day 1 performance of control animals (Figure 3G, statistics; also see Figures S3H–S3K). These data suggest that ablating dSPNs, but not iSPNs, completely abolishes the learned LLRR sequence and underscore the role of dSPNs in controlling the overall sequence.

Further analyses of the sequence microstructure revealed that dSPN-ablation mice showed a significant impairment in the initiation of the sequence (Figures 3H and 3I), which also resulted in a reduction in the overall frequency of L-R subsequence switches during a sequence (Figure 3J; see Figures S3H and S3I for more detailed analyses). In contrast, iSPN-ablation mice showed much less impairment on average in initiating or terminating the sequence with the correct element (Figures 3H and 3I; see Figures S3J and S3K for more detailed analyses). Rather, iSPN-ablation mice suffered from a significant reduction in the number of switches per action sequence (Figure 3J). Together, these results suggest that dSPNs and iSPNs play distinct roles in controlling action sequences and preferentially mediate sequence- vs. subsequence-level sequence execution, respectively.

Striatal pathways distinctly control sequence execution

The encoding of the action sequence at different behavioral levels by dSPNs and iSPNs does not necessarily imply whether the sequence is organized serially or hierarchically. To gain

further insights into the organization of learned action sequences, we next employed optogenetics to perturb the animals' ongoing actions within the sequence in a state-dependent closed-loop manner and investigate its effects on the subsequent sequence structure. D1- and A2a-cre mice expressing ChR2 were bilaterally implanted with optic fibers into the dorsal striatum (Figures S3C and S3D, see STAR Methods for details). After mice learned the 'Penguin Dance' sequence, a brief 500 ms pulse of constant blue light was delivered upon the first left, second left, first right or second right lever press during sequence performance (Figure 4A, see STAR Methods for details). The sustained firing shown by a large proportion of dSPNs during sequence execution suggested that the direct pathway might play an important role in maintaining sequence elements. Indeed, we observed that brief optogenetic stimulation of dSPNs after the 1st or 2nd left press facilitated ongoing actions and frequently inserted an additional left press into the left subsequence (Figures 4B and 4C). This effect of dSPN stimulation could not simply be explained as a "re-initiation" of the sequence, since stimulation on the 2nd press of the right subsequence also resulted in one additional right press (Figures 4D and 4E). Stimulation on the 1st right press did not have any obvious behavioral effect, suggesting strong state-dependent effects of optogenetic modulation of the sequence (due to an almost 100% natural likelihood of pressing right again, see Figures 1G and 1H; also see Figures S4A-S4D). Notably, the insertion of an additional left press into the left subsequence following dSPN stimulation was counterbalanced by the shortening of the right subsequence, so that the overall sequence length did not change between control and stimulated sequences (Figures 4J-4Q). These data suggest that dSPN stimulation facilitates ongoing action and inserts an additional element into the current subsequence. Yet, sequence-level properties like total sequence length can be maintained by additional levels of control that adjust the length of the following subsequence.

In contrast, iSPN stimulation after the 1st left or right lever press, through the elimination of the following action, consistently shortened the left and right subsequences, respectively (Figures 4F, 4H, and 4J-4Q). However, when a natural switch was expected after the 2nd left or right press, iSPN stimulation exerted no behavioral effects and the total sequence length remained intact (Figures 4G, 4I, and 4J-4Q), excluding the possibility that iSPNs act through general inhibition. This is consistent with what one would predict from the neuronal recording data in which iSPNs are largely inhibited during action execution but highly active during between-subsequence switching. Noticeably, when iSPN stimulation following the 1st left press removed the following action in the left subsequence, mice continued to execute the right subsequence normally (Figures 4F and 4J). Unlike the case of dSPN stimulation, the right subsequence did not compensate to maintain the same total sequence length after iSPN stimulation, resulting in a reduction in total sequence length (Figures 4J and 4N). Stimulation of iSPNs following the 1st right press instead caused animals to immediately run to the magazine to check for reward. Additional optogenetic experiments with 14 Hz frequency stimulation further confirmed these optogenetic effects (Figures S4I-S4Q). These results thus suggest that optogenetic stimulation of dSPNs or iSPNs can add or remove single actions in the sequence respectively, with distinct effects on the global sequence structure.

Optogenetic editing unveils the hierarchical structure of learned sequences

While iSPN stimulation following the 1st left press largely eliminated the next press of the left subsequence, the following right subsequence remained largely unchanged in terms of both its subsequence length and temporal structure (Figures 4F, 4J; see Figures S4E-S4H for more analyses). This observation is inconsistent with the serial chain model, which predicts that disrupting an early action would result in the termination of the whole sequence. Furthermore, these data raise the possibility that not only are element- and sequence-level structures maintained independently (Figures 4B and 4C), but that the left and right subsequences are also controlled separately. If so, one would predict that after sequence initiation, the execution of the right subsequence might remain largely normal even in the absence of the entire left subsequence. To test this hypothesis, we optogenetically stimulated dSPNs or iSPNs right before the initiation of the whole sequence. An infrared beam was placed in front of the left lever and used to trigger optogenetic stimulation when the animals transitioned from the magazine to the left lever for sequence initiation (Figure 5A) (Tecuapetla et al., 2016). While optogenetic stimulation of dSPNs delayed sequence initiation without disrupting the overall sequence structure (Figures S5A-S5E) (Tecuapetla et al., 2016), optogenetic stimulation of iSPNs during sequence initiation completely abolished the entire left subsequence (Figure 5B). These results suggest that iSPN stimulation can trigger a behavioral transition to the next subsequence in the motor program, whether by removing a single action element (Figure 4F) or the complete subsequence (Figure 5B). Notably, despite zero to few presses in the left subsequence following iSPN stimulation (Figure 5C), animals still executed the right subsequence with the usual length, timing, and duration as in control sequences (Figures 5C-5E). These data demonstrate that the left and right subsequences can be controlled independently. In addition, in the experiments of dSPN stimulation during the left subsequence, the right subsequence adjusted to maintain the appropriate total sequence length (Figures 4B and 4C). Together, these data suggest that the learned action sequence is likely organized in a hierarchical manner, with both local subsequence-level and global sequence-level controls. Importantly, the basal ganglia direct and indirect pathways interact distinctly with these different controllers.

To further confirm the hierarchical organization of learned action sequences, we trained a separate group of mice to perform an even more complicated heterogeneous sequence composed of three left followed by three right presses ('LLLRRR') (Figure 5F). Similar to the observations in the LLRR sequence, brief stimulation (100 ms) of iSPNs after the 1st press of the LLLRRR sequence ablated the entire left subsequence, removing multiple upcoming left presses well beyond the stimulation period (Figures 5G and 5H). Still, the right subsequence was executed at the expected time with its normal structure, including both the subsequence length and duration (Figures 5I-5J). These behavioral effects were consistently observed with various optogenetic stimulations spanning a wide range of durations (Figures S5F-S5J and S6). Together, these data support the notion that learned action sequences are organized hierarchically with separate modes of control at the element, subsequence, and sequence levels.

DISCUSSION

Here we developed a novel heterogeneous action sequence task in mice and investigated the organizational structure of learned action sequences. Differing from the popular back-propagation algorithms in reinforcement learning theory (Rumelhart et al., 1986; Schraudolph et al., 1994; Sutton and Barto, 1998), we found that heterogeneous action sequences are learned in a non-back-propagation manner where the start and stop actions represent highly significant elements (Jin and Costa, 2010; Murdock, 1962; Roediger and DeSoto, 2014). NMDA receptors in the striatum are critical for sequence learning and chunking distinct elements into the target sequence. Recent studies have suggested that the striatal direct and indirect pathways, instead of working antagonistically as the canonical model describes (Albin et al., 1989; DeLong, 1990; Kravitz et al., 2010), might work in a complementary manner for controlling actions (Cui et al., 2013; Isomura et al., 2013; Jin et al., 2014; Tecuapetla et al., 2016). By using a novel heterogeneous sequence task, the current study suggests that rather than simply competing or cooperating for individual motor output, the direct and indirect pathways might coordinate to dynamically control action sequences at different behavioral levels. Specifically, while the direct pathway is involved in initiating or facilitating actions, whether at the sequence or element level, the indirect pathway functions to terminate the ongoing subsequence and control the transition between subsequences in the motor program.

Several lines of evidence suggest that the optogenetic effects we observed cannot be attributed to short-term reinforcement of behavior by dSPN or iSPN stimulation (Kravitz et al., 2012; Yttri and Dudman, 2016). First, our within-subject design allows us to compare the sequence performance of the same subject with or without optogenetic stimulation in a given session. We do not observe any gradual changes in the structure of inter-leaved control sequences within each optogenetic session (Figures 4B–4I). In addition, optogenetic stimulation of dSPNs following the 1st right press, in contrast with stimulation on other presses, has no effect on the sequence structure (Figure 4D). Similarly, optogenetic stimulation of iSPNs following the 2nd left or the 2nd right press, unlike the 1st left or right presses, has no obvious effect on the sequence structure (Figures 4G and 4I). These results suggest that the optogenetic effects we observed following dSPN and iSPN stimulation are highly sequence state-dependent and are unlikely to result from simply positive or negative reinforcement.

The use of a heterogeneous action sequence further revealed a population of SPNs preferentially encoding the transition between subsequences (Figures 2O and 2P). These dynamics were preferentially expressed in iSPNs (Figure 2Q) and ablation of iSPNs specifically impaired animals' ability to link distinct subsequences (Figures 3F and 3J). The switch-related neuronal dynamics we observed in iSPNs during the transition between the left and right subsequences do not appear to reflect locomotion. In fact, optogenetic stimulation of iSPNs results in freezing behavior as mice locomote and elicits bradykinesia (Kravitz et al., 2010), likely through the inhibition of glutamatergic neurons in the mesencephalic locomotor region (Roseberry et al., 2016). In addition, optogenetic stimulation of iSPNs after the 2nd left press does not trigger any behavioral changes in our experiments (Figure 4G), contrasting with the complete removal of the current subsequence

after stimulating iSPNs on the 1st left or 1st right press or during sequence initiation and further excluding the possibility that iSPNs are simply involved in locomotion.

When the left subsequence was removed during iSPN stimulation, the right subsequence was executed at a similar time and with a similar duration as in the control sequences (Figure 5). Since right lever pressing does not occur immediately following iSPN stimulation, one might thus argue that iSPNs are only involved in action inhibition but not necessarily in directly mediating subsequence switching. Indeed, we found that about a quarter of iSPNs were inhibited during sequence execution (Figure 2N), suggesting that activation of these iSPNs might be involved in the inhibition of ongoing actions (Jin et al., 2014; Kravitz et al., 2010). However, inhibition of actions alone cannot reconcile how very brief (100) stimulation of iSPNs can remove multiple upcoming actions well beyond the stimulation period (Figures 5F–5J). A pure inhibition effect also fails to explain the observation that long durations (5 s) of iSPN stimulation, which cover the duration of the whole sequence, do not inhibit all action elements in the sequence (Figure S6). Instead, it produces an ablation of the left subsequence while leaving the entire right subsequence to be executed normally after stimulation offset (Figure S6). The data presented here suggest that iSPNs, in addition to action inhibition, might be directly involved in action switching. In fact, optogenetic stimulation of iSPNs during sequence initiation removes the left subsequence but not the whole sequence, again leaving the entire right subsequence to be executed normally (Figures 5A–5E). The switch among action repertoires is one of the most fundamental features of behavior (Brainard and Doupe, 2002; Gallistel, 1980; Graybiel, 1998; Hikosaka et al., 1999; Jin and Costa, 2015). We posit that the neural implementation of a switch requires the coordination of the basal ganglia with the current state of the network, specifically the timing information carried by other behavioral hierarchies, which together determine the actual execution of the next component in the motor program (Gallistel, 1980).

Our findings suggest that the basal ganglia direct and indirect pathways distinctly support different levels of the sequence structure (Figure 6). More specifically, we observed that one subpopulation of both dSPNs and iSPNs can encode the sequence-level start/stop (Cui et al., 2013; Isomura et al., 2013; Jin et al., 2014) while another subpopulation of dSPNs and iSPNs show sustained or inhibited activity during sequence execution, respectively (Jin et al., 2014). In addition, a selective group of iSPNs are active specifically during the transition between subsequences. These results emphasize the functional heterogeneity within each striatal dSPN or iSPN cell type. In fact, selective dSPN ablation not only impairs the correct initiation of the sequence but also appears to abolish the learned action sequence altogether (Figure 3). Ablation of iSPNs, though not noticeably affecting sequence initiation, strongly impairs the transition between left and right subsequences and sequence performance efficiency (Figure 3). Furthermore, the optogenetic experiments reveal that the two basal ganglia pathways also interact with different levels of the sequence hierarchy. Activation of the direct pathway by dSPN stimulation can insert an additional action into the sequence while maintaining the total sequence length through compensation of the right subsequence length. Activation of the indirect pathway, on the other hand, was sufficient to terminate the entire ongoing subsequence while leaving the next components in the motor program to be executed normally. These findings thus emphasize the importance of studying neural circuits under more complicated behavioral contexts, which better permits some of the complexity

and diversity of circuit functions to fully unfold. These results also underscore the much more complicated functions of the basal ganglia pathways in controlling actions than previously appreciated, and it is likely an oversimplification to assign one singular function to one striatal cell type or pathway (Albin et al., 1989; Calabresi et al., 2014; DeLong, 1990).

The classical model of the basal ganglia suggests that the direct and indirect pathways play antagonistic roles in controlling action (Albin et al., 1989; DeLong, 1990; Kravitz et al., 2010). More recent models, however, propose that the indirect pathway co-activates with the direct pathway to inhibit competing actions (Calabresi et al., 2014; Cui et al., 2013; Hikosaka et al., 2000; Isomura et al., 2013; Jin et al., 2014; Mink, 1996; Tecuapetla et al., 2016). The results presented here reveal the distinct yet complementary roles of the direct and indirect pathways (Jin et al., 2014; Tecuapetla et al., 2016) and, importantly, a more dynamic picture of their temporally precise interactions during sequence execution. In the current study, different subpopulations of neurons in each pathway encode different levels of the behavioral hierarchy, and they fire in an antagonistic or co-activated manner, depending on and evolving with the exact moment of ongoing execution of the sequence (Figure 6). This might explain why either inhibiting or activating dSPNs during lever approach delayed the start of the whole sequence, presumably due to the interference of the temporally precise physiological activity in dSPNs required for appropriate sequence initiation (Jin et al., 2014; Tecuapetla et al., 2016). It also provides mechanistic insights into the significant action sequence execution deficits observed in Parkinson's and Huntington's diseases (Agostino et al., 1992; Vinter and Gras, 1998). For instance, the Parkinsonian brain is dominated by abnormally synchronized population activity across the basal ganglia networks (Costa et al., 2006; Goldberg et al., 2004; Hammond et al., 2007), which are deprived of generating the dynamically ordered neuronal activity in both striatal pathways required for sequence execution. Proper organization of action sequences thus requires precisely coordinated activity between the direct and indirect pathways, likely through interactions with cortical/thalamic inputs (Hikosaka et al., 1999; Kupferschmidt et al., 2017; Smith et al., 2011; Tanji, 2001) as well as the dynamic release of dopamine in the striatum (Howard et al., 2017).

Taking advantage of the closed-loop optogenetic editing of a single action element or individual subsequence, the current study reveals that learned heterogeneous action sequences are likely organized hierarchically. Accordingly, we observed that the total sequence length (sequence level), the timing and length of subsequences (subsequence level), and the individual actions within the sequence (element level) can all be maintained separately. One major advantage of a hierarchical organization is error tolerance (Gallistel, 1980; Jin and Costa, 2015; Lashley, 1951). Indeed, we found that changes in one subsequence do not necessarily affect the execution of the following subsequence with regard to its proper timing, length, and duration. In addition, a hierarchical organization will also support more behavioral flexibility by facilitating module-based new learning (Gallistel, 1980; Hikosaka et al., 1995; Jin and Costa, 2015; Lashley, 1951). A hierarchical structure requires multiple levels of control, which are likely implemented by a distributed yet interconnected brain network (Dehaene et al., 2003; Gallistel, 1980; Graybiel, 1998; Hamaguchi et al., 2016; Hikosaka et al., 1999; Jin and Costa, 2015; Long et al., 2010; Tanji, 2001). We have shown that the basal ganglia are not only required for sequence learning but

also for appropriately organizing action sequences at different hierarchies. Previous studies have suggested that various cortical regions are involved in encoding sequence order (Tanji, 2001), number (Dehaene et al., 2003) or controlling sequence timing (Hamaguchi et al., 2016; Long et al., 2010). Future work will aim to elucidate how cortico-basal ganglia circuits work in coordination to control different aspects of sequence organization. Nevertheless, the current study underscores the importance of basal ganglia circuitry in relation to the functional organization of learned action sequences and has important implications from Parkinson's disease to speech disorders, in which the proper organization of action sequences is largely compromised (Brainard and Doupe, 2002; Graybiel, 1998; Hikosaka et al., 1999; Jin and Costa, 2015; Lai et al., 2001; Lashley, 1951; Vinter and Gras, 1998).

STAR METHODS

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Xin Jin (xjin@salk.edu).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Mice—All experiments were approved by the Salk Institute Animal Care and Use Committee and conformed to NIH Guidelines for the Care and Use of Laboratory Animals. Experiments were performed on both male and female mice, at least two months old, housed on a 12-hour light/dark cycle. C57BL/6 (Envigo/Harlan) mice were used in the wild-type experiments. Striatal-specific NMDAR1 knockout and control littermates were generated by crossing RGS9-cre mice with NR1 floxed (also denoted as *Grin1^{lox/flox}* in the Jackson Laboratory database) mice as previously described (Dang et al., 2006; Howard et al., 2017; Jin and Costa, 2010). RGS9-NR1 KO (referred to as striatal NR1-KO) mice and their littermate controls including RGS9-NR1 heterozygous, NR1 floxed, and RGS9-cre mice were used for behavioral experiments. BAC transgenic mice expressing cre recombinase under the control of the dopamine D1 receptor (GENSAT: EY217) or the A2a receptor (GENSAT: KG139) promoter were obtained from MMRRC and either crossed to C57BL/6 or Ai32 (012569) mice obtained from Jackson Laboratory (Cui et al., 2013; Jin et al., 2014; Madisen et al., 2012; Tecuapetla et al., 2016). To determine the extent of cell loss using the DTR ablation strategy, D1- and A2a-cre mice were crossed to the BAC reporter lines D1-eGFP (MMRRC: MMRRC_000297-MU; GENSAT: X60) and D2-eGFP (MMRRC: MMRRC_000230-UNC; GENSAT: S118) (Gong et al., 2007).

METHOD DETAILS

Behavioral Training—Behavioral training took place in standard mouse operant chambers as described previously (Howard et al., 2017; Jin and Costa, 2010; Jin et al., 2014). Briefly, operant chambers (21.6 cm × 17.8 cm × 12.7 cm; Med Associates, VT) were housed in sound-attenuating boxes and each chamber was equipped with a food magazine, a house light (3 W, 24 V) placed opposite the food magazine, and two retractable levers flanking the house light. Food pellets (20 mg; Bio-Serv, NJ) were delivered through a dispenser into the magazine as reinforcers and magazine entries were recorded using an

infrared beam. Behavioral chambers were controlled by behavioral software (MED-PC IV, Med Associates, VT) that recorded all timestamps of lever presses and magazine entries for each animal with a 10 ms resolution. All behavioral programs were custom written. Mice were food-restricted prior to behavioral training and were maintained at ~85% of normal body weight by receiving ~2.5 g of food pellets and normal chow per animal daily.

Behavioral training began with continuous reinforcement (CRF) as previously described (Howard et al., 2017; Jin et al., 2014). Briefly, CRF sessions started with the illumination of the house light and extension of either the left or right lever. Mice underwent two consecutive sessions of CRF each day (one session per lever) and the order of left and right sessions alternated daily. Mice received up to 5, 10, and 15 reinforcers per session on days one, two, and three of CRF, respectively. Following CRF, mice began training in the left-left-right-right (LLRR) sequence task ('Penguin Dance' – Self-Paced Version). Sessions started with the illumination of the house light and the extension of both the left and right levers. Reinforcers were delivered any time the behavioral program identified the consecutive 'left-left-right-right' lever press pattern. Therefore, extra presses in addition to the 'LLRR' pattern did not exclude the reward. No cues were presented to signal sequence correctness or reward availability and daily sessions lasted for up to three hours or until the mouse received 40 reinforcers. Training sessions ended with retraction of both levers and offset of the house light. For learning experiments, all mice were trained in the LLRR sequence task for 28 consecutive days. Since there were no significant differences in the learning of the LLRR sequence between WT mice and the littermate controls of RGS9-NR1 KO mice, the data were thus combined. In addition, since RGS9-NR1 KO mice received significantly less reinforcers than littermate controls, a separate cohort of littermate control mice underwent the same training protocol described above except the number of reinforcers was limited to 20 pellets per day to match the RGS9-NR1 KO mice. For the LLLRRR task, training followed the same design as the LLRR sequence task described above except reward contingency was based on the identification of the 'left-left-left-right-right-right' lever press pattern.

For the lever-retraction version of the LLRR sequence task ('Penguin Dance' – Lever-Retraction Version), training took place in the same boxes as described above but the left and right levers were placed on the same side of the magazine. Training began with CRF as described above. Following CRF, mice began training in a fixed-ratio four schedule. Sessions started with the illumination of the house light and extension of both the left and right levers. After every four presses, levers retracted for a 5 s inter-trial interval. Reward delivery only occurred when the four-press sequence was composed of left-left-right-right. Daily sessions lasted for up to three hours or until the mouse received 60 reinforcers. Training sessions ended with the retraction of both levers and offset of the house light. C57BL/6 wildtype mice ($n = 10$) were trained in the lever-retraction version of the LLRR sequence task for 28 consecutive days.

Behavioral quantification—The beginning of a sequence was defined as the first press following a magazine entry. For all learning, recording, and ablation data, the termination of a sequence was defined by magazine entry. All optogenetic data went through an additional post hoc analysis process to better identify discrete sequences for the quantification of the

optogenetic effects on element-, subsequence-, and sequence-level changes. The termination of the left or right subsequence was further determined based on the distribution of inter-press intervals for each animal. The inter-press intervals often follow a multimodal distribution corresponding to chunked bouts of pressing at the shortest intervals (within subsequence), shorter intervals (switch between subsequences) and longer intervals (sequences separated by magazine checking). Left and right subsequences were identified as the first peak in the distribution of the inter-press intervals (Jin and Costa, 2010; Jin et al., 2014). Behavioral efficiency (%) was defined as the percentage of rewarded lever presses ('LLRR'; 4 presses/reward) out of the total number of presses within a behavioral session. The learning of the left (LL --) ('-' denotes either 'L' or 'R' press in the sequence) and right (-- RR) subsequences was defined as the percentage of sequences with four or more presses beginning with LL or ending with RR, respectively. The learning of each element of the sequence was defined as the percentage of sequences with four or more presses containing a left press in the first (L --) or second (-- L --) sequence positions or containing a right press in the penultimate (-- R --), or final right (-- -- R) positions. For the ablation data analyses, the percentage of "Start" and "Stop" elements as well as the average number of left-right switches per sequence were determined from all sequences composed of two or more presses. Sequence quantification in the LLLRRR sequence task was similarly defined as in the LLRR sequence task. Sequences were first defined by the occurrence of magazine entries and further refined by the distribution of inter-press intervals as described above. Behavioral efficiency (%) was defined as the percentage of rewarded lever presses ('LLLRRR'; 6 presses/reward) out of the total number of presses within a behavioral session.

Sequences in the lever-retraction version of the LLRR sequence task were defined as the four presses between lever extension and lever retraction. The percentage of correct sequences was defined as the percentage of left-left-right-right (LLRR) sequences out of the total number of sequences within a behavioral session. The learning of the left (LL --) and right (-- RR) subsequences was defined as the percentage of sequences beginning with LL or ending with RR, respectively. The learning of each element of the sequence was defined as the percentage of sequences containing a left press in the first (L --) or second (-- L --) sequence positions or containing a right press in the penultimate (-- R --), or final right (-- -- R) positions.

Surgery and implantation—All intracranial injections/implantations were conducted in mice at least two months of age under general ketamine (100 mg/kg) and xylazine (10 mg/kg) or isoflurane (~4% induction; 1-2% sustained) anesthesia. The head was shaved, cleaned with 70% ethanol and povidone-iodine, and then placed in a Kopf stereotaxic frame. For cannula, fiber, or array implantation, two skull screws were placed posterior to bregma to better affix the dental cement to the skull. For muscimol experiments, 22 gauge guide cannulas (Plastics One, VA) were implanted into dorsal striatum at a 4° angle to ensure enough separation between the two cannulas using the following coordinates: +0.5 mm AP, ±2.55 mm ML, -2.16 mm DV. Cannulas were cemented in place with dental acrylic (Contemporary Ortho-Jet powder and liquid, Lang Dental, IL). Dummy cannulas fitted to the length of the guide cannulas were inserted following surgery. For DTR ablation

Author Manuscript

Author Manuscript

Author Manuscript

experiments, D1-cre and A2a-cre mice were stereotaxically injected with a cre-inducible adeno-associated virus carrying the diphtheria toxin receptor (Azim et al., 2014) (AAV9-FLEX-DTR-GFP; Salk GT3 Core, CA). Virus was injected in eight different sites. We used two different AP/ML sites for each hemisphere followed by two DV coordinates at each AP/ML site. The coordinates were +0.9 mm AP, ± 1.6 mm ML, -2.2 and -3.0 mm DV and 0.0 mm AP, ± 2.1 mm ML, -2.2 and -3.0 mm DV. A Hamilton syringe was used to inject 1 μ L at the four -3.0 mm DV sites and another 0.5 μ L at the four -2.2 mm DV sites for a total of 3 μ L injected per hemisphere. Following each injection, the needle was left in place for ~5 minutes and then raised over ~5 minutes. This same protocol was used for each injection site. All optogenetic viral injections or fiber implants were performed as previously described (Howard et al., 2017; Tecuapetla et al., 2016). Briefly, mice expressing only D1- or A2a-cre were stereotaxically injected with a cre-inducible adeno-associated virus carrying channelrhodopsin (AAV9-EF1a-DIO-hChR2(H134R)-eYFP, University of Pennsylvania vector core, PA or AAV5-EF1a-DIO-hChR2(H134R)-mCherry, University of North Carolina vector core, NC) into dorsal striatum (+0.5 mm AP, ± 2.0 – 2.4 mm ML, -2.2 mm DV) using a Hamilton syringe (1 μ L per side) (Howard et al., 2017; Tecuapetla et al., 2016). Following viral injections or for mice genetically expressing ChR2 under cre control (D1-Ai32, A2a-Ai32), optic fibers constructed as previously described (Howard et al., 2017; Tecuapetla et al., 2016) (200 μ m optic fiber) were lowered into dorsal striatum using the same coordinates as for viral injections. Fibers were cemented in place with dental acrylic (Contemporary Ortho-Jet powder and liquid, Lang Dental, IL).

Author Manuscript

Array implants for optogenetic-assisted identification recordings were performed as previously described (Howard et al., 2017; Jin et al., 2014). Briefly, we utilized electrode arrays (Innovative Neurophysiology Inc., NC) of 16 tungsten contacts (2×8) with each contact 35 μ m in diameter and spaced 150 μ m apart. Each array was also equipped with a cannula located 300 μ m from the electrode tips, allowing for insertion of an optic fiber to deliver 473-nm light. Arrays targeting dorsal striatum (+0.5 mm AP, ± 1.5 mm ML, -2.2 mm DV) were unilaterally implanted into D1-Ai32 or A2a-Ai32 mice. The hemisphere for implantation was pseudorandomized across animals. Silver grounding wire was attached to skull screws. Once the array was lowered into dorsal striatum, the grounding wire and array were affixed using dental acrylic. Following viral injections and/or implantation, mice received buprenorphine (0.5–1 mg/kg) as an analgesic, and mice were allowed to recover for at least 1 week in their home cage before food-restriction and behavioral training (Howard et al., 2017; Jin et al., 2014).

Muscimol experiments—Mice implanted with cannulas were re-trained until they achieved at least 40% behavioral efficiency to ensure stable behavior. The following day, we started our three-day infusion protocol in which mice received consecutive days of saline, muscimol, and saline infusions. Muscimol was dissolved in saline before infusion (Sigma-Aldrich; 0.05 μ g/ μ L). For the infusions, mice were briefly anesthetized with isoflurane and injection cannulas (Plastics One, VA) were bilaterally inserted into the cannulas, with the injection cannulas projecting 0.1 mm beyond the implanted guide cannulas. Each injection cannula was attached to an infusion pump (BASi, IN) via polyethylene tubing. Animals were bilaterally infused with 200 nL of liquid (saline or muscimol) followed by a five-minute

waiting period before removal of the infusion cannulas. Mice were returned to their home cage and started in the behavioral task 30 minutes after infusion. Behavioral sessions lasted until the animal received 80 reinforcers or 3 hours had passed.

DTR-mediated cell ablation—To determine the ablation efficiency of the AAV9-FLEX-DTR-GFP virus and diphtheria toxin strategy in striatum, adult D1-cre;D1-eGFP ($n = 2$) and A2a-cre;D2-eGFP mice ($n = 2$) were injected with AAV9-FLEX-DTR-GFP in one hemisphere and sham-injected in the other using the same coordinates described above. Two weeks later, mice were administered 1 μ g of diphtheria toxin (DT) dissolved in 300 μ L of phosphate buffered saline (PBS) via intraperitoneal (I.P.) injection on two consecutive days (Azim et al., 2014). Mice were perfused two weeks later and tissue was processed for immunohistochemistry. For ablation behavioral experiments, mice were food-restricted and, following completion of CRF, underwent training in the LLRR behavioral paradigm for three weeks. Immediately after day 21 of LLRR training, mice were pseudorandomly divided into control and treatment groups. Treatment mice were administered DT via I.P. injection whereas control mice received I.P. injections of PBS. The same injections were given on the following day. To allow for neuronal ablation, animals were stopped in behavioral training and placed back on normal chow. Animals resumed LLRR sequence training 14 days after the first DT or PBS injection.

Histology and cell counting—For tissue collection, mice were deeply anesthetized with ketamine/xylazine and transcardially perfused with 0.01 M PBS followed by 4% paraformaldehyde (PFA) using a peristaltic pump. Brains were removed and post-fixed in 4% PFA overnight at 4° C. Tissue was then transferred to 30% sucrose in 0.1 M phosphate buffer for cryoprotection and kept at 4° C until the brains sunk. Tissue was sectioned with a microtome into 40-50 μ m sections and either mounted onto glass slides and cover-slipped with mounting media (Aqua-Poly/Mount, Polysciences, PA) and DAPI (1:1000, Sigma-Aldrich) or used for antibody labeling. Amplification of the eGFP signal in D1-cre;D1-eGFP and A2a-cre;D2-eGFP mice was carried out via immunohistochemistry as previously described (Smith et al., 2016). Briefly, sections were washed 3 \times 15 min in tris-buffered saline (TBS) and then incubated for 1 hour in blocking solution (3% normal horse serum and 0.25% Triton-X-100 in TBS). Sections were transferred to primary antibody diluted in blocking solution (Green fluorescent protein, Rabbit polyclonal, 1:400, Invitrogen Molecular Probes, IL) overnight at 4° C and, the following day, washed 2 \times 15 min in TBS. Sections were transferred to blocking solution for 30 minutes then placed in secondary antibody diluted in blocking solution (AlexaFluor 647 Donkey anti-rabbit, 1:250, Jackson ImmunoResearch, PA) for 2-3 hours. Sections were then washed 3 \times 15 min in TBS before being mounted onto glass slides and cover-slipped with mounting media and DAPI. For each D1-cre;D1-eGFP and A2a-cre;D2-eGFP animal, three sections were imaged on a Zeiss LSM 710 laser scanning microscope with a 10 \times objective. For cell counting, confocal images of GFP expression in sham-injected and DTR-injected hemispheres were imported into Fiji, overlaid with a grid, and counted using the plugin Cell Counter. Cells were determined to be positive for GFP based on clear somal expression. Following counting, each hemisphere was divided into dorsomedial and dorsolateral regions. Cell counts for each region of the ablated

hemisphere were then expressed as a percentage by dividing by the cell counts of the corresponding region in the sham-injected hemisphere.

Optogenetic experiments—*In vivo* optogenetic stimulation was delivered with a 473 nm laser (LaserGlow Technologies, Canada). The laser was controlled by a TTL output programmed in the behavioral software (MED-PC IV, Med Associates, VT). Following implantation, mice were re-trained in operant chambers while tethered to two fiber-optic patch cords extending from a commutator (Doric, Canada) to allow for free rotation within the behavioral chamber. Optogenetic stimulation began once mice reached 40% behavioral efficiency to ensure sufficient trials for analysis. During every optogenetic session, stimulation was only delivered once per stimulation sequence. The likelihood of stimulation was ~50% and randomized so that non-stimulated (control sequences) and stimulated sequences were randomly interleaved. Some stimulation conditions were repeated across multiple stimulation days to ensure enough trials for robust analysis (Howard et al., 2017).

For the element editing optogenetic experiments, we defined four stimulation conditions based on the four presses of the LLRR sequence. On any given stimulation day, mice only underwent stimulation triggered by one press of the sequence—1st (left), 2nd (left), 3rd (right), or 4th (right)—and the order of stimulation conditions was pseudo-randomized across mice. During a stimulated sequence, lever pressing triggered one constant 500 ms pulse of 473-nm light. In the case of 2nd press stimulation, mice displayed a range of probabilities in pressing the left lever again. Stimulation of iSPNs following the 2nd left press was focused on stimulating when the natural switch of the animal was expected to occur. Therefore, to maintain a consistent state across animals, only mice with control left subsequences close to 2 presses were used for data analysis.

To evaluate the consistency of stimulation effects across varying stimulation parameters, some mice also went through additional stimulation sessions in which lever pressing triggered 10 ms light pulses delivered at 14 Hz for 500 ms. For the beam break experiments, mice were tethered to the commutator and also trained with the beam break apparatus (custom built) located within the behavioral chamber. The infrared beam device consisted of an emitter placed above the left lever and facing downward. An infrared sensor was placed in the behavioral tray below the emitter to establish a beam of infrared light. When the infrared beam was broken by the animals' approach to the left lever, a TTL input was sent to the MED-PC software to trigger 500 ms of constant blue light (Tecuapetla et al., 2016). For the LLLRRR optogenetic experiments, A2a-cre mice injected with cre-dependent Chr2 and A2a-Ai32 mice were first trained in the LLLRRR sequence as described above. Once animals reached 40% efficiency, mice underwent stimulation in which the first press of the LLLRRR sequence triggered 50, 100, 200, or 500 ms of constant 473-nm light. Since there were no significant differences for the optogenetic effects in mice with viral expression of Chr2 in A2a-cre and A2a-Ai32 mice, the data were thus combined (same for mice with viral expression of Chr2 in D1-cre and D1-Ai32). To construct the peri-event time histograms for control and stimulated sequences, lever pressing in both the control and stimulated conditions were aligned to the stimulated press, averaged in 100 ms bins, and filtered with a Gaussian low-pass filter (window size = 5, standard deviation = 5). Due to the narrow smoothing window, all the PETHs in the optogenetic experiments were plotted by

excluding the referenced press in both the control and stimulated conditions for illustration clarity.

***In vivo* neuronal recording with ChR2-aided cell type identification**—All D1-Ai32 and A2a-Ai32 mice were pre-trained in the LLRR sequence task as described above. Following implantation, mice were allowed to recover approximately one week before food-restriction and behavioral training. Recording mice followed the same tethering procedure as optogenetic mice but were instead tethered via a recording cable. In order to ensure stable behavior for data analysis, recordings were only performed for mice that reached 40% efficiency. *In vivo* recording during freely moving behavior and neuronal identification was performed as previously described (Howard et al., 2017; Jin and Costa, 2010; Jin et al., 2014). Briefly, an optic fiber was inserted into a cannula affixed to the recording array and neural activity was recorded using the MAP system (Plexon Inc., TX). Spike activity was first sorted online with a built-in algorithm (Plexon Inc., TX) and only spikes with stereotypical waveforms distinguishable from noise and a high signal-to-noise ratio were saved. Following completion of the behavioral task, varying durations of constant or 14 Hz blue light from a 473-nm laser were delivered to verify the identity of recorded units. Following the recording, all spikes were further sorted into individual units using an offline sorting software (Offline Sorter, Plexon Inc., TX). Identified units displayed a clear refractory period with no spikes during the refractory period (larger than 1.3 ms). To determine light-evoked responses, neuronal firing was aligned to laser onset and averaged across all stimulation trials in 1 ms bins. Baseline firing was defined by averaging neuronal firing -1000 to 0 ms before laser onset in 1 ms bins. The latency to respond to light stimulation was defined as the start of a significant firing rate increase and the threshold for significance was defined as > 99% of baseline activity (3 standard deviations). Only units showing very short response latencies (< 10 ms) to light stimulation and a strong correlation between spike waveforms occurring during behavior and those generated by optogenetic stimulation ($R = 0.95$) were considered cre-positive units (Howard et al., 2017; Jin and Costa, 2010; Jin et al., 2014).

Analysis of neuronal activity—Given the self-paced nature of the task, neural activity occurring prior to the initiation of the LLRR sequence was confounded by animals' consumption of the reward at the magazine, lever pressing, or transitions from the magazine to initiate left lever pressing. Therefore, neural activity following the start of the task but prior to the initiation of lever pressing was randomly sampled with a 10-s time window 50 times to estimate the baseline firing rate for each unit. Neuronal activity in the 10-s window was binned with 10 ms time bins, averaged across all 50 samples, and filtered with a Gaussian low-pass filter (window size = 5, standard deviation = 5) to define baseline activity. Neuronal activity referenced to lever pressing was aligned to lever press onset, averaged across all trials in 10 ms bins, and smoothed using the same Gaussian filter described above to construct peri-event time histograms (PETH). We then determined which smoothed 10 ms bins occurring 1,000 ms before and after each lever press met the criteria for sequence-related activity (Jin et al., 2014). A significant increase in firing rate was defined as at least 5 consecutive bins with activity exceeding 95% (2 standard deviations) of the baseline activity and an inhibitory response was defined as at least 5 consecutive bins with activity 68% (1

standard deviation) below baseline activity (Barnes et al., 2005; Jin and Costa, 2010; Jin et al., 2014).

To evaluate element-related or sequence-related activity, we generated four PETHs, one for each action of the LLRR sequence—first left, final left, first right, and final right presses. Sequence-related start/stop neurons were defined as those with a significant firing rate modulation before the first press (start) and/or after the final press (stop) of the sequence that was significantly different than the firing rate modulations associated with the remaining presses within the sequence. Inhibited or sustained activity was defined as a significant negative or positive firing rate modulation constantly associated with multiple lever presses in the sequence (Jin and Costa, 2010; Jin et al., 2014). To identify between-subsequence switch-related neuronal activity, PETHs were constructed by aligning to the termination of the left subsequence or initiation of the right subsequence. Switch neurons were defined as showing a significant firing rate modulation during this transition period compared to the baseline. All analyses were performed with custom-written scripts in Matlab (MathWorks, MA).

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistics—Statistics for the wildtype and RGS9-NR1 KO learning data as well as the DTR ablation experiments were performed on the basis of values for each mouse per session. Statistics for the optogenetic data were performed on the basis of control and stimulated values for each mouse per stimulation condition. Normality was tested using the Shapiro-Wilk normality test. Control and wildtype learning data were analyzed using repeated-measures one-way ANOVA. RGS9-NR1 KO data were analyzed using repeated-measures two-way ANOVA. Muscimol data were analyzed using repeated-measures one-way ANOVA. To determine the efficiency of the DTR ablation strategy by cell type and striatal region, a two-way ANOVA was used. For evaluation of dSPN- and iSPN-ablation behavioral experiments, one-way ANOVA or two-tailed unpaired t-tests were used as indicated. The neuronal recording data was analyzed using a z-test for the comparison of proportions (Sheskin, 2004). For analysis of the optogenetic data, two-tailed paired t-tests were used. Sidak and Tukey post-hoc multiple comparisons were performed as indicated. All data were first analyzed in Matlab (Mathworks, MA) and all statistics were performed in GraphPad Prism 7 (GraphPad Software, CA). Results are presented as mean \pm S.E.M. except for the neuronal recording data, which are presented as the percentage within the task-related positively identified units. $P < 0.05$ was considered significant. All statistical details are located within the figure legends. The number of animals (n) used in each experiment is reported in the figure legends and the number of identified dSPNs or iSPNs is specified in the text.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors would like to thank Drs. Ed Callaway, Rusty Gage, Tom Jessell, Chris Kintner, Terry Sejnowski and members of the Jin lab for discussion and comments on the manuscript. This research is supported by grants from the US National Institutes of Health (R01NS083815 and R01AG047669), the Dana Foundation, Ellison Medical Foundation and Whitehall Foundation to X.J.

References

- Agostino R, Berardelli A, Formica A, Accornero N, Manfredi M. Sequential arm movements in patients with Parkinson's disease, Huntington's disease and dystonia. *Brain*. 1992; 115:1481–1495. [PubMed: 1422799]
- Albin RL, Young AB, Penney JB. The functional anatomy of basal ganglia disorders. *Trends Neurosci*. 1989; 12:366–375. [PubMed: 2479133]
- Azim E, Jiang J, Alstermark B, Jessell TM. Skilled reaching relies on a V2a propriospinal internal copy circuit. *Nature*. 2014; 508:357–363. [PubMed: 24487617]
- Barnes TD, Kubota Y, Hu D, Jin DZ, Graybiel AM. Activity of striatal neurons reflects dynamic encoding and recoding of procedural memories. *Nature*. 2005; 437:1158. [PubMed: 16237445]
- Brainard MS, Doupe AJ. What songbirds teach us about learning. *Nature*. 2002; 417:351–358. [PubMed: 12015616]
- Calabresi P, Picconi B, Tozzi A, Ghiglieri V, Di Filippo M. Direct and indirect pathways of basal ganglia: a critical reappraisal. *Nature Neurosci*. 2014; 17:1022. [PubMed: 25065439]
- Calabresi P, Pisani A, Mercuri NB, Bernardi G. Long-term potentiation in the striatum is unmasked by removing the voltage-dependent magnesium block of NMDA receptor channels. *Eur J Neurosci*. 1992; 4:929–935. [PubMed: 12106428]
- Costa RM, Lin SC, Sotnikova TD, Cyr M, Gainetdinov RR, Caron MG, Nicolelis MA. Rapid alterations in corticostriatal ensemble coordination during acute dopamine-dependent motor dysfunction. *Neuron*. 2006; 52:359–369. [PubMed: 17046697]
- Cui G, Jun SB, Jin X, Pham MD, Vogel SS, Lovinger DM, Costa RM. Concurrent activation of striatal direct and indirect pathways during action initiation. *Nature*. 2013; 494:238–242. [PubMed: 23354054]
- Dang MT, Yokoi F, Yin HH, Lovinger DM, Wang Y, Li Y. Disrupted motor learning and long-term synaptic plasticity in mice lacking *NMDAR1* in the striatum. *Proc Nat'l Acad Sci USA*. 2006; 103:15254–15259. [PubMed: 17015831]
- Dehaene S, Piazza M, Pinel P, Cohen L. Three parietal circuits for number processing. *Cogn Neuropsychol*. 2003; 20:487–506. [PubMed: 20957581]
- DeLong MR. Primate models of movement disorders of basal ganglia origin. *Trends Neurosci*. 1990; 13:281–285. [PubMed: 1695404]
- Gallistel CR. *The Organization of Action: A New Synthesis* Lawrence Erlbaum Associates; 1980
- Goldberg JA, Rokni U, Boraud T, Vaadia E, Bergman H. Spike synchronization in the cortex-basal ganglia networks of parkinsonian primates reflects global dynamics of the local field potentials. *J Neurosci*. 2004; 24:6003–6010. [PubMed: 15229247]
- Gelato. Penguin's Game 1999 <https://www.youtube.com/watch?v=S7WNgpXPeTc>. [https://en.wikipedia.org/wiki/Bunny_hop_\(dance\)](https://en.wikipedia.org/wiki/Bunny_hop_(dance)); Lyrics: "Left left right right, go turn around, go go go."
- Gong S, Doughty M, Harbaugh CR, Cummins A, Hatten ME, Heintz N, Gerfen CR. Targeting cre recombinase to specific neuron populations with bacterial artificial chromosome constructs. *J Neurosci*. 2007; 27:9817–9823. [PubMed: 17855595]
- Graybiel AM. The basal ganglia and chunking of action repertoires. *Neurobiol Learn Mem*. 1998; 70:119–136. [PubMed: 9753592]
- Hamaguchi K, Tanaka M, Mooney R. A distributed recurrent network contributes to temporally precise vocalizations. *Neuron*. 2016; 91:680–693. [PubMed: 27397518]
- Hammond C, Bergman H, Brown P. Pathological synchronization in Parkinson's disease: networks, models and treatments. *Trends Neurosci*. 2007; 30:357–364. [PubMed: 17532060]

- Hikosaka O, Nakahara H, Rand MK, Sakai K, Lu X, Nakamura K, Miyachi S, Doya K. Parallel neural networks for learning sequential procedures. *Trends Neurosci.* 1999; 22:464–471. [PubMed: 10481194]
- Hikosaka O, Rand MK, Miyachi S, Miyashita K. Learning of sequential movements in the monkey: process of learning and retention of memory. *J Neurophysiol.* 1995; 74:1652–1661. [PubMed: 8989401]
- Hikosaka O, Takikawa Y, Kawagoe R. Role of the basal ganglia in the control of purposive saccadic eye movements. *Physiol Rev.* 2000; 80:953–978. [PubMed: 10893428]
- Howard CD, Li H, Geddes CE, Jin X. Dynamic nigrostriatal dopamine biases action selection. *Neuron.* 2017; 93:1436–1450. [PubMed: 28285820]
- Isomura Y, Takekawa T, Harukuni R, Handa T, Aizawa H, Takada M, Fukai T. Reward-modulated motor information in identified striatum neurons. *J Neurosci.* 2013; 33:10209–10220. [PubMed: 23785137]
- Jin X, Costa RM. Start/stop signals emerge in nigrostriatal circuits during sequence learning. *Nature.* 2010; 466:457–462. [PubMed: 20651684]
- Jin X, Costa RM. Shaping action sequences in basal ganglia circuits. *Curr Opin Neurobiol.* 2015; 33:188–196. [PubMed: 26189204]
- Jin X, Tecuapetla F, Costa RM. Basal ganglia subcircuits distinctively encode the parsing and concatenation of action sequences. *Nature Neurosci.* 2014; 17:423–430. [PubMed: 24464039]
- Kravitz AV, Freeze BS, Parker PR, Kay K, Thwin MT, Deisseroth K, Kreitzer AC. Regulation of parkinsonian motor behaviours by optogenetic control of basal ganglia circuitry. *Nature.* 2010; 466:622–626. [PubMed: 20613723]
- Kravitz AV, Tye LD, Kreitzer AC. Distinct roles for direct and indirect pathway striatal neurons in reinforcement. *Nature Neurosci.* 2012; 15:816–818. [PubMed: 22544310]
- Kupferschmidt DA, Juczewski K, Cui G, Johnson KA, Lovinger DM. Parallel, but dissociable, processing in discrete corticostriatal inputs encodes skill learning. *Neuron.* 2017; 96:476–489. [PubMed: 29024667]
- Lai CSL, Fisher SE, Hurst JA, Vargha-Khadem F, Monaco AP. A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature.* 2001; 413:519–523. [PubMed: 11586359]
- Lashley KS. The problem of serial order in behavior In *Cerebral Mechanisms in behavior* Jeffress LA, editor John Wiley Press; 1951
- Lima SQ, Hromadka T, Znamenskiy P, Zador AM. PINP: a new method of tagging neuronal populations for identification during in vivo electrophysiological recording. *PLoS One.* 2009; 4:e6099. [PubMed: 19584920]
- Long MA, Jin DZ, Fee MS. Support for a synaptic chain model of neuronal sequence generation. *Nature.* 2010; 468:394–399. [PubMed: 20972420]
- Madisen L, Mao T, Koch H, Zhuo JM, Berenyi A, Fujisawa S, Hsu YWA, Garcia AJ, Gu X, Zanella S, et al. A toolbox of Cre-dependent optogenetic transgenic mice for light-induced activation and silencing. *Nature Neurosci.* 2012; 15:793–802. [PubMed: 22446880]
- Mink JW. The basal ganglia: focused selection and inhibition of competing motor programs. *Prog Neurobiol.* 1996; 50:381–425. [PubMed: 9004351]
- Murdock BB Jr. The serial position effect of free recall. *J Exp Psychol.* 1962; 64:482–488.
- Roediger H, DeSoto K. Forgetting the presidents. *Science.* 2014; 346:1106–1109. [PubMed: 25430768]
- Roseberry TK, Lee AM, Lalive AL, Wilbrecht L, Bonci A, Kreitzer AC. Cell-type-specific control of brainstem locomotor circuits by basal ganglia. *Cell.* 2016; 164:526–537. [PubMed: 26824660]
- Rumelhart DE, Hinton GE, Williams RJ. Learning representations by back-propagating errors. *Nature.* 1986; 323:533.
- Saito M, Iwawaki T, Taya C, Yonekawa H, Noda M, Inui Y, Mekada E, Kimata Y, Tsuru A, Kohno K. Diphtheria toxin receptor-mediated conditional and targeted cell ablation in transgenic mice. *Nature Biotechnol.* 2001; 19:746–750. [PubMed: 11479567]
- Schraudolph NN, Dayan P, Sejnowski TJ. Temporal difference learning of position evaluation in the game of Go. *Adv Neural Inf Process Syst.* 1994; 6:817–824.

- Shen W, Flajolet M, Greengard P, Surmeier DJ. Dichotomous dopaminergic control of striatal synaptic plasticity. *Science*. 2008; 321:848–851. [PubMed: 18687967]
- Sherrington CS. *The Integrative Action of the Nervous System* Yale University Press; 1906
- Sheskin DJ. *Handbook of Parametric and Nonparametric Statistical Procedures* CRC Press; 2004
- Smith JB, Klug JR, Ross DL, Howard CD, Hollon NG, Ko VI, Hoffman H, Callaway EM, Gerfen CR, Jin X. Genetic-based dissection unveils the inputs and outputs of striatal patch and matrix compartments. *Neuron*. 2016; 91:1069–1084. [PubMed: 27568516]
- Smith Y, Surmeier DJ, Redgrave P, Kimura M. Thalamic contributions to basal ganglia-related behavioral switching and reinforcement. *J Neurosci*. 2011; 31:16102–16106. [PubMed: 22072662]
- Sutton R, , Barto AG. *Reinforcement Learning: An Introduction* MIT Press; 1998
- Tanji J. Sequential organization of multiple movements: involvement of cortical motor areas. *Annu Rev Neurosci*. 2001; 24:631–651. [PubMed: 11520914]
- Tecuapetla F, Jin X, Lima SQ, Costa RM. Complementary contributions of striatal projection pathways to action initiation and execution. *Cell*. 2016; 166:703–715. [PubMed: 27453468]
- Vinter A, Gras P. Spatial features of angular drawing movements in Parkinson’s disease patients. *Acta Psychol*. 1998; 100:177–193.
- Yttri EA, Dudman JT. Opponent and bidirectional control of movement velocity in the basal ganglia. *Nature*. 2016; 533:402. [PubMed: 27135927]

HIGHLIGHTS

- Non-back-propagation learning of sequences depends on striatal NMDA receptors
- Striatal direct pathway facilitates actions and controls sequence start/stop
- Striatal indirect pathway inhibits actions and mediates subsequence switch
- Optogenetic manipulations unveil the hierarchical structure of learned sequences

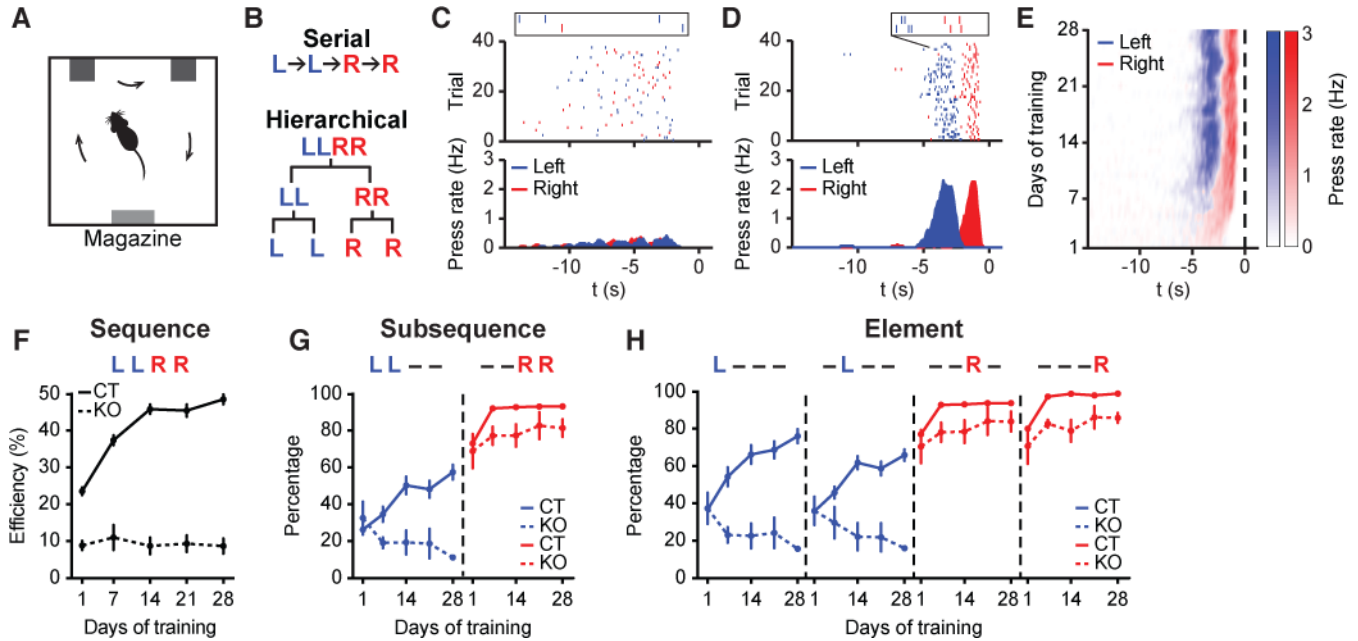


Figure 1. NMDA receptors in striatum are critical for learning heterogeneous action sequences (A) Operant chamber schematic. (B) Serial (Top) vs. hierarchical (Bottom) organization of LLRR sequence. (C and D) Example of typical wildtype mouse behavior on day 1 (C) and day 28 (D) of training. Top Panels: Left and right lever presses indicated by blue and red dashes, respectively, and aligned to magazine entry at time zero. Bottom Panels: Averaged left and right lever press rate indicated by blue and red lines, respectively. Insets show two representative sequences. (E) Development of a stereotypical action sequence across 28 days of training in a wildtype mouse. (F) Behavioral efficiency for control ($n = 22$ mice; main effect of training $F_{4,84} = 81.71$, $P < 0.0001$) and striatal NR1-KO mice ($n = 5$ mice; no effect of training $F_{4,16} = 0.1531$, $P = 0.9588$) across training (main effect of genotype $F_{1,25} = 257$, $P < 0.0001$). (G) Percentage of sequences beginning with the ‘LL’ subsequence (LL --, ‘-’ denotes either ‘L’ or ‘R’ press in the sequence) or ending with the ‘RR’ subsequence (-- RR) for control (LL -- : main effect of training $F_{4,84} = 18.27$, $P < 0.0001$; -- RR: main effect of training $F_{4,84} = 30.26$, $P < 0.0001$) and striatal NR1-KO mice (LL -- : no effect of training $F_{4,16} = 1.704$, $P = 0.1982$; -- RR: no effect of training $F_{4,16} = 0.6898$, $P = 0.6096$) across training (LL -- : main effect of genotype $F_{1,25} = 11.32$, $P = 0.0025$; -- RR: main effect of genotype $F_{1,25} = 26.93$, $P < 0.0001$). (H) Percentage of sequences containing each appropriate element position for control (L --- : main effect of training $F_{4,84} = 24.17$, $P < 0.0001$; - L --- : main effect of training $F_{4,84} = 23.88$, $P < 0.0001$; -- R - : main effect of training $F_{4,84} = 29.97$, $P < 0.0001$; --- R: main effect of training $F_{4,84} = 62.05$, $P < 0.0001$; L --- vs. - L --- : main effect of element position $F_{1,21} = 9.616$, $P = 0.0054$; first day of significant difference, Day 7, $P = 0.0027$; -- R - vs. --- R: main effect of element position $F_{1,21} = 41.56$, $P < 0.0001$; first day of significant difference, Day 1, $P = 0.013$) and striatal NR1-KO mice (L --- : no effect of training $F_{4,16} = 1.747$, $P = 0.189$; - L --- : no effect of training $F_{4,16} = 1.077$, $P = 0.4006$; -- R - : no effect of training $F_{4,16} = 0.7249$, $P = 0.5877$; --- R: no effect of training $F_{4,16} = 1.412$, $P = 0.275$) across training (L --- : main effect of genotype $F_{1,25} = 22.77$, $P < 0.0001$; - L --- : main effect of genotype $F_{1,25} =$

32.21, $P < 0.0001$; -- R -- : main effect of genotype $F_{1,25} = 27.01$, $P < 0.0001$; --- R: main effect of genotype $F_{1,25} = 60.26$, $P < 0.0001$). Data were analyzed using repeated-measures one-way or two-way ANOVA followed by Tukey or Sidak post-hoc comparisons, respectively. Error bars denote S.E.M., same for below unless stated otherwise. See also Figures S1 and S2.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

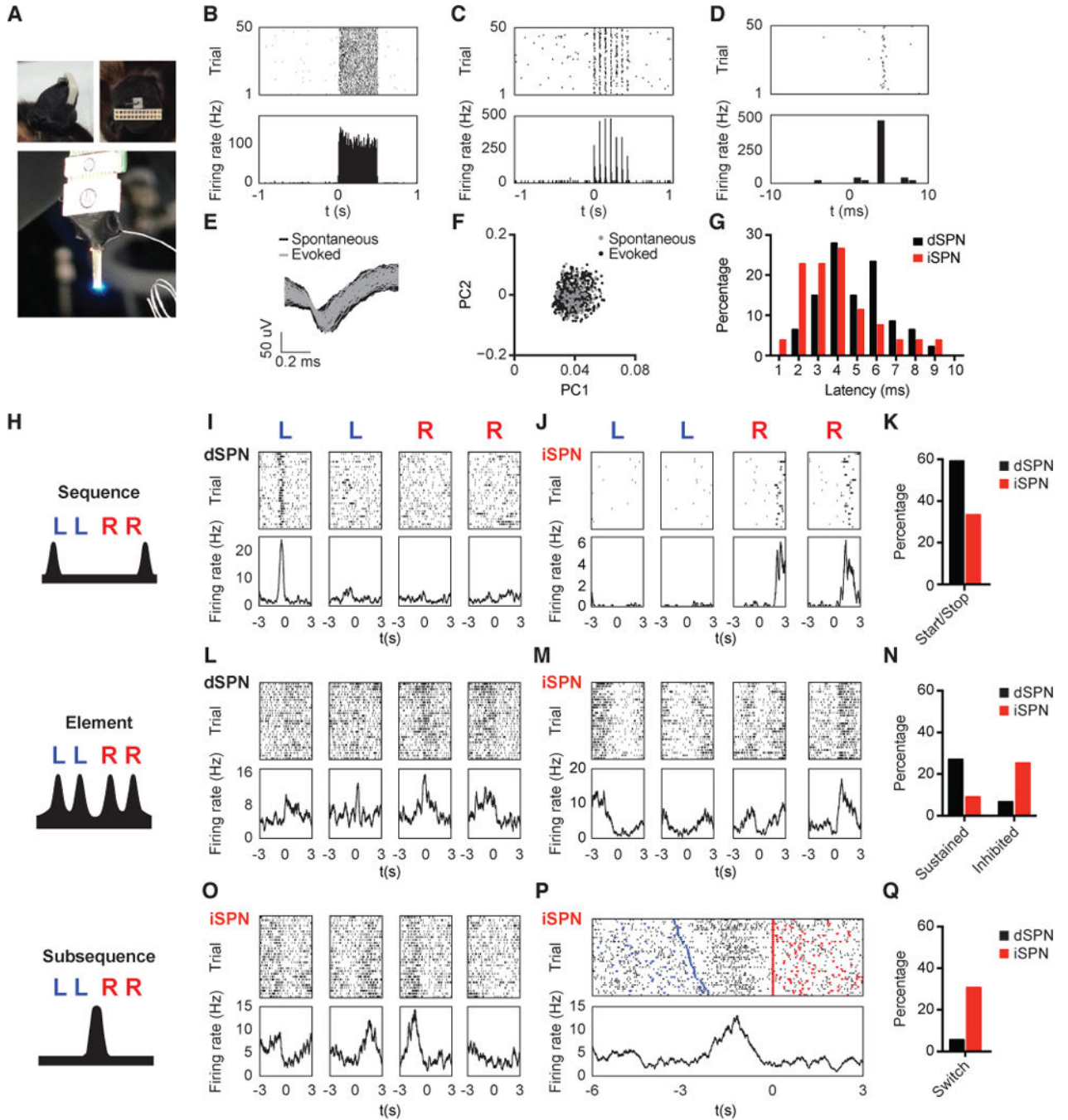


Figure 2. Striatal pathways differentially encode sequence structure

(A) Left: Side (top image) and top-down (bottom image) view of a recording array affixed with a cannula implanted in a D1-Ai32 animal. Right: Light emitted by optic fiber placed through the attached cannula is in close proximity to the tips of the recording array. (B and C) Representative dSPN response to 500 ms of constant (B) or 14 Hz (C) laser stimulation. Top Panels: Each row represents one trial and the black dashes indicate spikes. Bottom Panels: Average firing rate (Hz) aligned to laser stimulation at time zero. (D) Same unit as in (C) but with a finer timescale. (E) Waveforms from the same dSPN in (A-C) for spontaneous

and laser-evoked spikes. (F) Principal component analysis of spontaneous and laser-evoked waveforms demonstrates the overlapped clustering of spontaneous and evoked spikes. (G) Distribution of light response latencies for dSPNs and iSPNs. (H) Schematic of sequence-, element-, and subsequence-related neural activity. (I) Representative dSPN showing sequence start activity. Top Panels: Each dash indicates a spike. Bottom Panels: Neuronal activity is aligned (time zero) to the 1st left, final left, 1st right and final right lever presses within the sequence, respectively. (J) Representative iSPN showing sequence stop activity. (K) Proportion of dSPNs and iSPNs showing sequence start/stop activity (Two-sample z-test, $Z = 2.28$, $P = 0.0226$). (L) Representative dSPN showing sustained activity to each action element. (M) Representative iSPN showing inhibited activity during each action element. (N) Proportion of dSPNs and iSPNs showing element sustained (Two-sample z-test, $Z = 2.03$, $P = 0.0424$) or inhibited (Two-sample z-test, $Z = -2.25$, $P = 0.0244$) activity. (O) Representative iSPN showing subsequence switch-related activity. (P) Peri-event time histogram (PETH) of the same iSPN as shown in (O) with trials sorted by left-right subsequence switch intervals. Top Panel: Each dash indicates a spike. The left and right presses are marked by inverted blue and red triangles, respectively. Bottom Panel: Neuronal activity is aligned to the first right press at time zero. (Q) Proportion of dSPNs and iSPNs showing subsequence switch-related activity (Two-sample z-test, $Z = -2.92$, $P = 0.0035$). See also Figure S3.

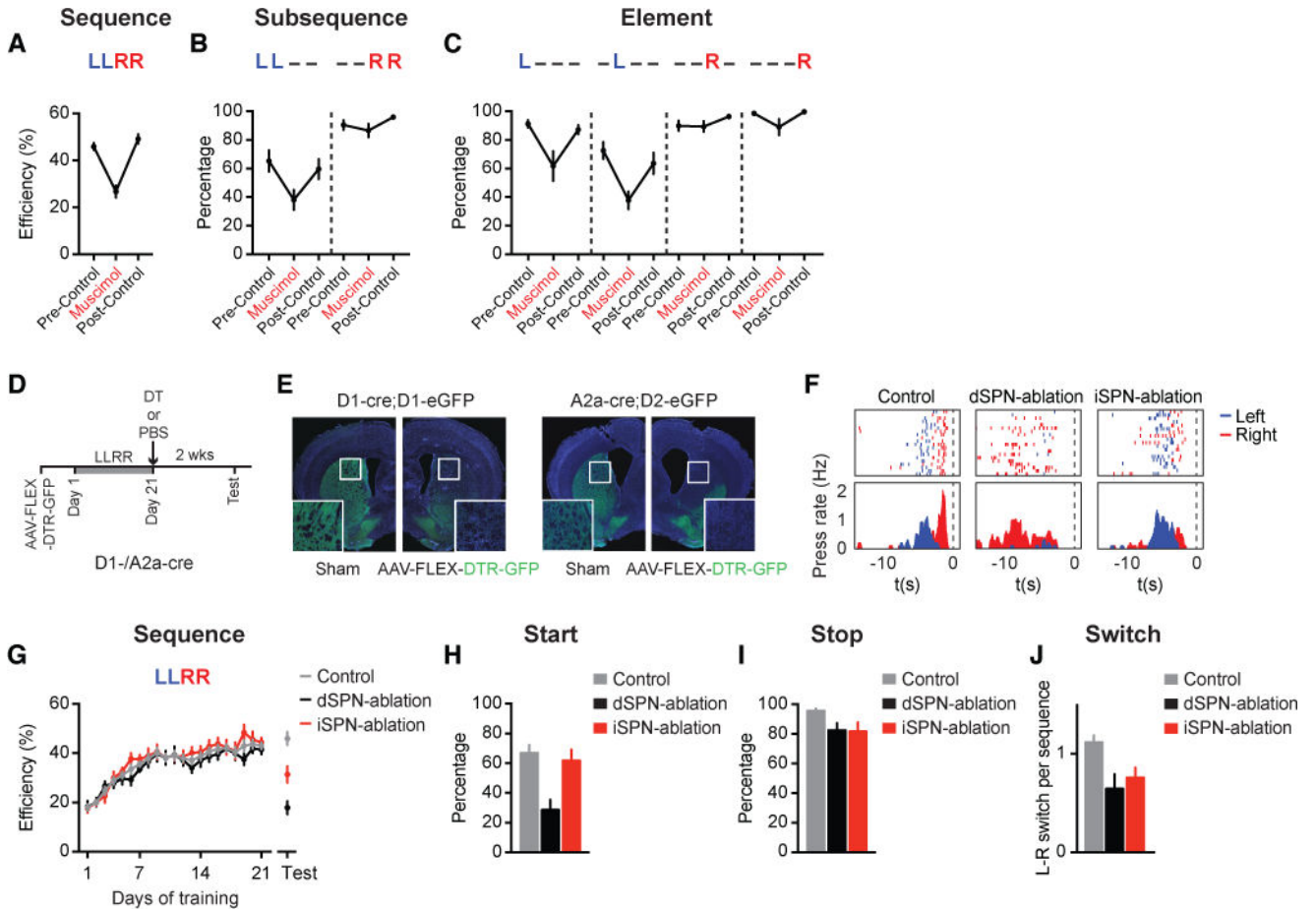


Figure 3. Dorsal striatum is necessary for sequence execution, which is distinctly controlled by dSPNs and iSPNs

(A) Behavioral efficiency of sequence performance in trained mice during the muscimol infusion day and the pre-/post-control days ($n = 7$ mice; main effect of treatment $F_{2,12} = 32.44$, $P < 0.0001$; muscimol vs. pre-/post-control, $P < 0.0001$ and $P < 0.0001$, respectively). (B) Percentage of sequences beginning with the ‘LL’ subsequence (LL-- : main effect of treatment $F_{2,12} = 7.859$, $P = 0.0066$; muscimol vs. pre-/post-control, $P = 0.007$ and $P = 0.0309$, respectively) or ending with the ‘RR’ subsequence (--RR: no main effect of treatment $F_{2,12} = 2.958$, $P = 0.0903$). (C) Percentage of sequences containing each appropriate element position (L-- : main effect of treatment $F_{2,12} = 6.604$, $P = 0.0116$; muscimol vs. pre-/post-control, $P = 0.0147$ and $P = 0.0338$, respectively; -L-- : main effect of treatment $F_{2,12} = 7.59$, $P = 0.0074$; muscimol vs. pre-/post-control, $P = 0.0072$ and $P = 0.0407$, respectively; --R- : no main effect of treatment $F_{2,12} = 2.138$, $P = 0.1606$; -- --R: no main effect of treatment $F_{2,12} = 3.754$, $P = 0.0542$). (D) Timeline for animal training and DT-mediated dSPN or iSPN ablation. (E) Cell ablation in sham- or AAV-FLEX-DTR-GFP-injected hemispheres following I.P. DT injection in D1-cre;D1-eGFP or A2a-cre;D2-eGFP mice. (F) Example of control, dSPN-ablation, and iSPN-ablation mouse behavior on the day of testing. Data are aligned to magazine entry at time zero. (G) Behavioral efficiency for control ($n = 8$), dSPN-ablation ($n = 7$), and iSPN-ablation ($n = 8$) mice (Test Day: main effect of treatment $F_{2,20} = 22.28$, $P = 0.0041$; Tukey’s multiple comparison test, control vs.

dSPN-ablation, $P < 0.0001$; control vs. iSPN-ablation, $P = 0.0051$; dSPN-ablation vs. iSPN-ablation, $P = 0.0118$) (Day 1 control vs. dSPN-ablation, unpaired t-test, $t_{21} = 0.0302$, $P = 0.9762$). (H) Percentage of sequences starting with a left press for control, dSPN-ablation, and iSPN-ablation mice on the day of testing (Main effect of treatment $F_{2,20} = 9.452$, $P = 0.0013$; control vs. dSPN-ablation, $P = 0.0018$; control vs. iSPN-ablation, $P = 0.8483$; dSPN-ablation vs. iSPN-ablation, $P = 0.0059$). (I) Percentage of sequences ending with a right press for control, dSPN-ablation, and iSPN-ablation mice on the day of testing (No main effect of treatment $F_{2,20} = 2.929$, $P = 0.0766$). (J) Averaged number of L-R switches per sequence for control, dSPN-ablation, and iSPN-ablation mice on the day of testing (Main effect of treatment $F_{2,20} = 5.379$, $P = 0.0135$; control vs. dSPN-ablation, $P = 0.0057$; control vs. iSPN-ablation, $P = 0.0241$; dSPN-ablation vs. iSPN-ablation, $P = 0.4671$). Muscimol and ablation data were analyzed using repeated-measures one-way ANOVA and one-way ANOVA, respectively, followed by Tukey's multiple comparison test. See also Figure S3.

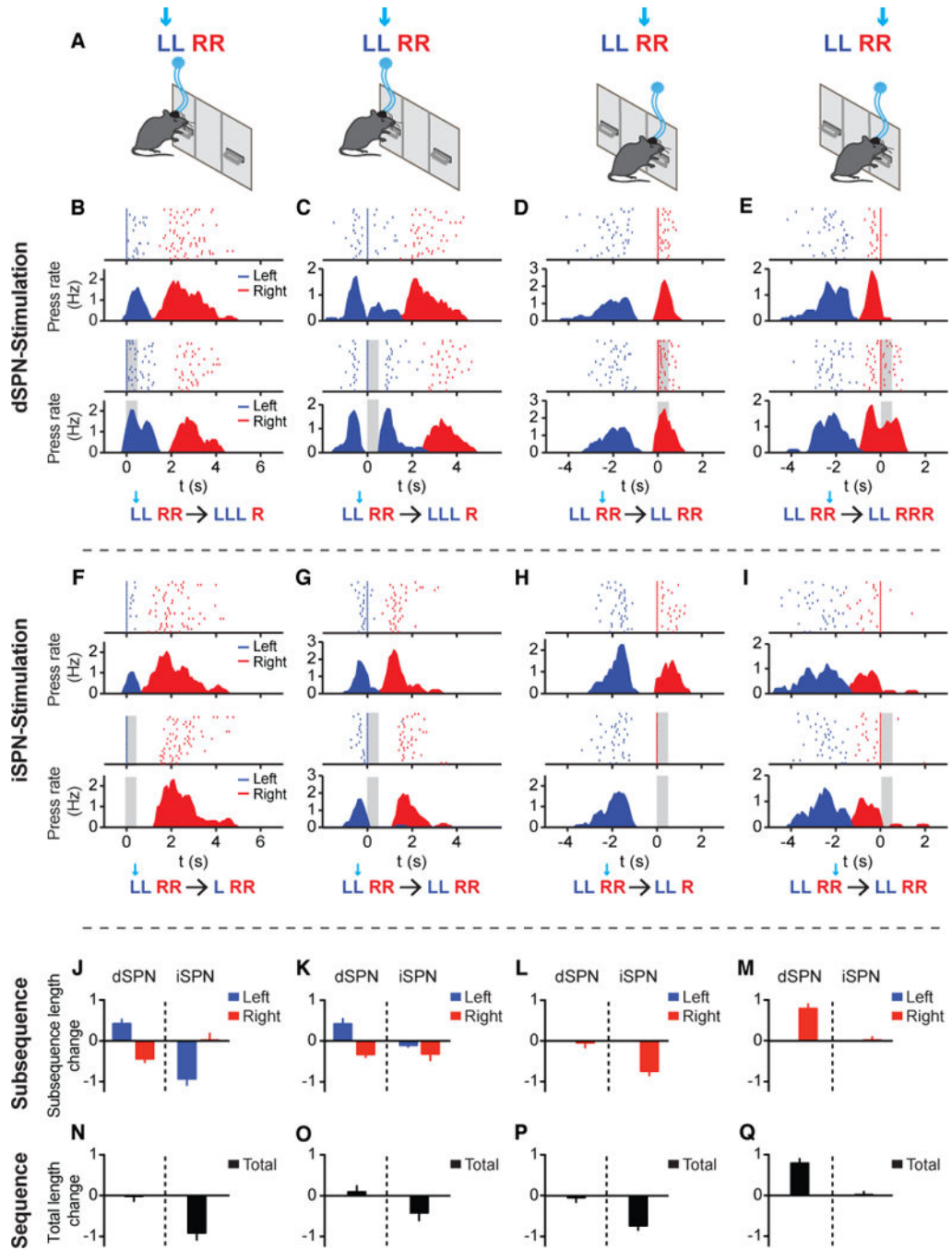


Figure 4. Different modulation of sequence structure by optogenetic stimulation of dSPNs or iSPNs

(A) Optogenetic experiment protocol for delivering 500 ms light stimulation triggered by the 1st left, 2nd left, 1st right and 2nd right lever press within the sequence, respectively, of randomly chosen 50% of trials. (B-E) Behavioral examples of dSPN stimulation following the 1st left (B), 2nd left (C), 1st right (D), and 2nd right (E) press of the sequence in the control (Top Panels) and stimulated (Bottom Panels) conditions. A representation of the change in the LLRR sequence following optogenetic stimulation is shown below each behavioral example. Left and right presses are shown as blue and red lines, respectively. The

period of stimulation (500 ms) is covered with gray shadow. In each case, lever pressing is aligned to the stimulated press at time zero. Note that the PETHs in all optogenetic experiments were plotted by excluding the referenced lever press in both control and stimulated conditions for illustration clarity, same for below. (F-I) Same as (B-E) except for iSPN stimulation following the 1st left (F), 2nd left (G), 1st right (H), and 2nd right (I) press of the sequence. (J-M) Change in the left and right subsequence lengths following dSPN or iSPN stimulation on the 1st left (J, paired t-tests, dSPN Left: $t_{12} = 4.03$, $P = 0.0017$, Right: $t_{12} = 4.951$, $P = 0.0003$; iSPN Left: $t_9 = 6.116$, $P = 0.0002$, Right: $t_9 = 0.113$, $P = 0.9125$; dSPN, $n = 13$; iSPN, $n = 10$ mice), 2nd left (K, paired t-tests, dSPN Left: $t_{12} = 3.309$, $P = 0.0062$, Right: $t_{12} = 4.477$, $P = 0.0008$; iSPN Left: $t_7 = 1.88$, $P = 0.1021$, Right: $t_7 = 2.005$, $P = 0.085$; dSPN, $n = 13$; iSPN, $n = 8$ mice), 1st right (L, paired t-tests, dSPN Right: $t_{12} = 0.3975$, $P = 0.698$; iSPN Right: $t_8 = 7.177$, $P < 0.0001$; dSPN, $n = 13$; iSPN, $n = 9$ mice) and 2nd right (M, paired t-tests, dSPN Right: $t_{12} = 7.445$, $P < 0.0001$; iSPN Right: $t_9 = 0.226$, $P = 0.8263$; dSPN, $n = 13$; iSPN, $n = 10$ mice) lever presses within the sequence. (N-Q) Change in the total sequence length following dSPN or iSPN stimulation on the 1st left (N, paired t-tests, dSPN: $t_{12} = 0.123$, $P = 0.9042$; iSPN: $t_9 = 5.404$, $P = 0.0004$), 2nd left (O, paired t-tests, dSPN: $t_{12} = 0.6434$, $P = 0.5321$; iSPN: $t_7 = 2.284$, $P = 0.0563$), 1st right (P, paired t-tests, dSPN: $t_{12} = 0.3975$, $P = 0.698$; iSPN: $t_8 = 7.177$, $P < 0.0001$) and 2nd right (Q, paired t-tests, dSPN: $t_{12} = 7.445$, $P < 0.0001$; iSPN: $t_9 = 0.226$, $P = 0.8263$) lever presses within the sequence. See also Figures S3 and S4.

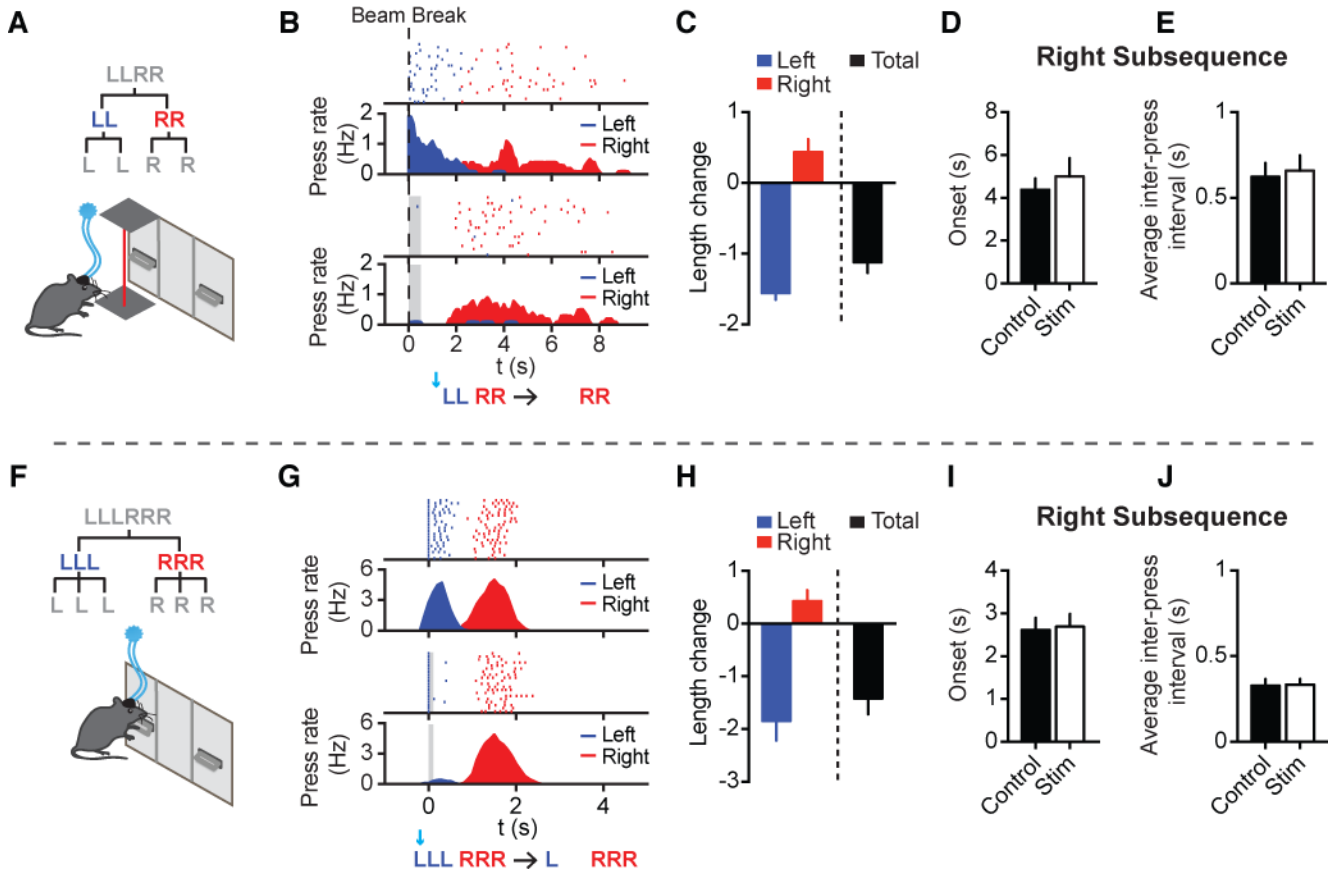


Figure 5. Optogenetic editing unveils a hierarchical structure of learned action sequences (A) Optogenetic stimulation right before LLRR sequence initiation triggered by infrared beam break ($n = 6$ mice). (B) Behavioral effect of optogenetic iSPN activation prior to sequence initiation. Lever pressing is aligned to beam break at time zero in both the control (Top Panels) and stimulated (Bottom Panels) conditions. The period of stimulation (500 ms) is covered with gray shadow. (C) Change in the length of the left subsequence (paired t-test, $t_5 = 16.84$, $P < 0.0001$), right subsequence (paired t-test, $t_5 = 2.431$, $P = 0.0593$), and the whole sequence (paired t-test, $t_5 = 7.474$, $P = 0.0007$) between control and stimulated sequences. (D) Averaged time of onset for the right subsequence in the control and stimulated sequences (paired t-test, $t_5 = 1.365$, $P = 0.2306$). (E) Averaged inter-press interval of the right subsequence in the control and stimulated sequences (paired t-test, $t_5 = 0.9858$, $P = 0.3695$). (F) Schematic of optogenetic stimulation triggered by the 1st left press of the LLLRRR sequence ($n = 12$ mice). (G) Behavioral effect of optogenetic iSPN activation following the 1st left press of the LLLRRR sequence in the control (Top Panels) and stimulated (Bottom Panels) conditions. Left and right presses shown as blue and red lines, respectively. The period of stimulation (100 ms) is covered with gray shadow. (H) Change in the length of the left subsequence (paired t-test, $t_{11} = 5.011$, $P = 0.0004$), right subsequence (paired t-test, $t_{11} = 2.069$, $P = 0.0628$), and the whole sequence (paired t-test, $t_{11} = 4.773$, $P = 0.0006$) in the control and stimulated sequences. (I) Averaged time of onset for the right subsequence in control and stimulated sequences (paired t-test, $t_{11} = 0.4149$, $P = 0.6862$). (J)

Averaged inter-press interval of the right subsequence in the control and stimulated sequences (paired t-test, $t_{11} = 0.4275$, $P = 0.6773$). See also Figures S5 and S6.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

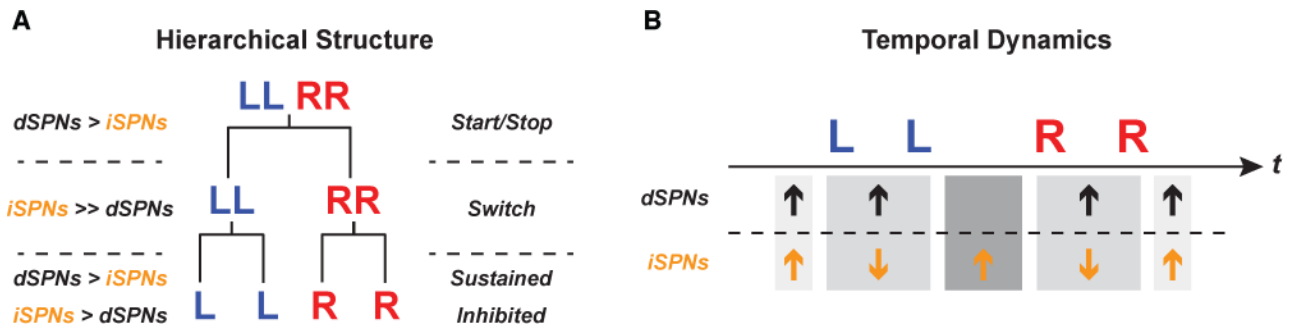


Figure 6. Dynamically coordinated activity between dSPNs and iSPNs supports the hierarchical organization of learned action sequences

(A) Summary diagram of the different roles of dSPNs and iSPNs at each level of the behavioral hierarchy. At the sequence level, both dSPNs and iSPNs signal sequence start/stop. At the subsequence level, the indirect pathway preferentially encodes between-subsequence switch. At the element level, both direct and indirect pathways are involved in action execution with different neuronal dynamics. The magnitude difference between the proportions of dSPNs or iSPNs at each hierarchical level is indicated with greater-than ('>'), much greater-than ('>>') and less-than ('<') signs. (B) Striatal direct and indirect pathways dynamically coordinate their activity during sequence execution. The different subpopulations of dSPNs and iSPNs coordinate their activity to support the start/stop of the sequence, the execution of the elemental actions, and the switch between subsequences. The 'up' or 'down' arrows indicate the positive or negative modulation of firing rate in each neuronal subpopulation, respectively.