PNAS

Supporting Information for
Learning critically drives parkinsonian motor deficits through imbalanced striatal pathway recruitment

Timothy H. C. Cheung, Yunmin Ding, Xiaoxi Zhuang, Un Jung Kang[*]

*Corresponding author: Un Jung Kang
Email: Un.Kang@nyulangone.org

This PDF file includes:

> Supporting text
> Figures S1 to S10
> Legends for Movies S1 to S4
> Legend for Dataset S1
> SI References

Other supporting materials for this manuscript include the following:

> Movies S1 to S4
> Dataset S1

Supporting Information Text

Material and Methods

Mice

Adult mice of both sexes (≥3 months old at the beginning of the experiments) were grouped housed, up to five to a cage and kept on a 12-h light/dark cycle with *ad libitum* food and water. Except otherwise stated, C57BL/6J mice (Jackson Laboratory, Bar Harbor, ME, stock# 000664) were used. To visualize iSPNs, heterozygous mice expressing green fluorescent protein (GFP) under the regulation of the *Drd2* promotor were used (Drd2-EGFP; STOCK Tg(*Drd2-EGFP*)S118Gsat/Mmnc from the GENSAT Project obtained from the MMRRC, stock# 000230-UNC) (1). To examine the role of D1Rs and D2Rs in parkinsonian motoric decline, mice with homozygous D1R deletion (D1R knockout; D1R-KO) (2) or with homozygous D2R knockdown (D2R-KD) (3) were used. D2R-KD mice had ~97% reduction in D2R mRNA expression and lower locomotor activity (3). Wild type (WT) littermates of the respective mouse line (D1R-WT and D2R-WT) were used as controls. To restrict Cre-dependent viral expression to cholinergic interneurons, heterozygous mice expressing Cre-recombinase under the regulation of the *ChAT* promotor were used (ChAT-IRES-Cre; B6;129S6-*Chat* tm2(cre)Lowl/J, Jackson Laboratory, Bar Harbor, ME, stock# 006410) (4). To restrict Cre-dependent viral expression to iSPNs, heterozygous mice expressing Cre-recombinase under the regulation of the *Adora2a* promotor were used (Adora2a-Cre; STOCK Tg(*Adora2a-cre*)KG139Gsat/Mmucd from the GENSAT Project obtained from the MMRRC, stock# 036158-UCD) (5). All transgenic and mutant mouse lines were backcrossed with C57BL/6J mice for ≥5 generations prior to use.

Behavioural procedures

Step task (forepaw adjusting treadmill stepping task). In this task (modified from 6), the mouse was gently held by its lower body by the experimenter in a stationary position above a moving treadmill. Its weight-bearing forepaws were in contact with the moving treadmill, and the number of adjusting steps was videoed and subsequently counted for each forepaw separately. The belt of the treadmill was set to move at a speed of 6.1 cm/s, in the posterior-to-anterior direction relative to the mouse. Each trial consisted of approximately 47.3 cm of stepping (i.e., lasting approximately 7.8 s).

      Pole task. In this task, the mouse was placed at the top of a pole (diameter: 1.3 cm; length: 89 cm) facing upwards and descended the pole by itself (7). The surface of the pole was covered with surgical tape to enhance friction for the mouse. The time taken to descend the pole (descend latency) and the net contralateral turn during the descent (number of contralateral turns − number of ipsilateral turns) were recorded. Each trial consisted of one descent.

      While the above tasks are typically used to study motor deficits over short timescales, we found that they also revealed experience-dependent worsening of deficits. Unless otherwise

stated, each daily session for a task consisted of 5 trials, with consecutive trials separated by ≥30 s. For each mouse, the behavioural outcome was averaged across the 5 trials for each daily session before downstream analysis (see Statistical analysis below). For experiments that used both Step task and Pole task (Fig. 1 and S1), at least 4 hours separated the two tasks for each daily session. Within these experiments, for experiments presented in Fig. 1D and S1A-B, Step task was always run before Pole task for each daily session. For experiments presented in Fig. 1E-H and S1C-D, the order of Step task vs. Pole task was randomized across daily sessions.

Unless otherwise stated, all "DA-depleted baseline" and "Baseline" performance denotes task performance in the final treatment-free session before repeated treatments began. For all long-term rescue experiments (except for experiments presented in Fig. 1 and S1), Step task long-term rescue was monitored during the repeated treatment phase by probing the mouse on the Step task for 2 trials before each day's drug treatment, i.e., when the mouse was drug-free and after ~23 h washout from the previous drug treatment. Stepping was averaged across the 2 probe trials before downstream analysis.

For one-day parkinsonian decline experiments (Fig. 3B, 5B, S6B, S6D), on test day, mice were treated with their designated drug/vehicle treatment 20-30 minutes before the Step task. Mice in parkinsonian decline groups were then tested on the Step task for 5 blocks, each block consisting of 5 trials. The start of each block was separated by approximately 10 minutes. Mice were euthanized and perfused 90-110 minutes after their drug/vehicle treatment. This allowed neuronal activation associated with parkinsonian decline to be examined with Fos and p-rpS6$^{S235/236}$ immunohistochemistry. Mice who took part in the experiment in Fig. 3B were additionally probed on the Step task for 2 trials just before euthanasia.

Behavioural analysis
Quantitative model of parkinsonian decline
To obtain the rate of parkinsonian decline of motor performance (both post-6-OHDA and after induction of long-term rescue), we used quantitative modelling to estimate the rate of performance decline. For the Step task, a two-parameter exponential decay model was used to fit the observed stepping data (averaged step/trial) for each daily session from an individual mouse:

$$Observed\ stepping\ on\ day\ d = Step_1 \times e^{\ slope\ (d-1)}$$

where $Step_1$ is the fitted stepping on day 1, *slope* is the fitted slope, and *d* is the day number. Contralateral vs. ipsilateral stepping was fitted separately. For the Pole task, a two-parameter linear model was used to fit the observed net contralateral turn (averaged net turn/trial) for each daily session from an individual mouse:

$$Observed\ net\ contra\ turn\ on\ day\ d = Turn_1 + slope \times (d-1)$$

where $Turn_1$ is the fitted net contralateral turn on day 1, *slope* is the fitted slope, and *d* is the day number. We did not fit an exponential model to the Pole task because unlike stepping, net

3

contralateral turn could be negative numbers, and to fit it with an exponential decay would require a third parameter. We therefore used a linear model for the Pole task for parsimony. Note that in both models, a negative slope implies performance is declining, a slope of zero implies steady performance (neither increasing nor decreasing), and a positive slope implies performance is increasing (improving). The models were fitted using R (function "lm" to fit the linear decline model using linear least square; function "nlsLM" from the package "minpack.lm" to fit the exponential decay model using non-linear least square). The fitted slope for each mouse was used for downstream statistical analysis. Furthermore, a predicted performance was generated for each mouse for each day using the fitted parameters. The resulting "decline fit" group means ± s.e.m. were plotted as transparent bands in the corresponding figures, to allow visual confirmation that they largely overlapped with observed data's group means ± s.e.m., consistent with reasonable fit.

Comparison of contralateral vs. ipsilateral stepping
Unilateral 6-OHDA can also cause mild ipsilateral motor function changes (6), possibly due to impaired bilateral co-ordination between the impaired contralateral limb, controlled by the dopamine-depleted striatum, and the ipsilateral limb whose stepping could not fully compensate. Alternatively, mild ipsilateral motoric deficits may reflect contribution from cross-hemispheric dopamine projections (8). To examine the effects of unilateral 6-OHDA on contralateral vs. ipsilateral stepping, for every experiment that involved 6-OHDA or sham lesion, we compared the rate of post-6-OHDA performance decline for contralateral vs. ipsilateral stepping (Fig. S2B) by calculating a Step decline index as: (Contra step on Day 4 / Contra step on Day 1) − (Ipsi step on Day 4 / Ipsi step on Day 1). Thus, a negative Step decline index means that contralateral stepping declined faster than ipsilateral stepping post-6-OHDA, which would be consistent with unilateral dopamine depletion primarily affecting contralateral movements. Furthermore, for every experiment that involved a long-term rescue probe session with no drug treatment (after 72 h drug washout), we compared the long-term rescue amplitude for contralateral vs. ipsilateral stepping. Long-term rescue amplitude for each side was calculated as: long-term rescue probe session stepping – post-6-OHDA baseline stepping (last session before treatment began). The difference in long-term rescue amplitude between contralateral vs. ipsilateral stepping was analysed and shown in Fig. S3A, with positive value representing greater long-term rescue amplitude for contralateral vs. ipsilateral stepping. Groups that did not receive Step task training paired with DA agonist or iSPN inhibition were excluded, as they did not acquire long-term rescue. Finally, we compared the decline slope of long-term rescue for contralateral vs. ipsilateral stepping, by calculating the difference in fitted decline slopes (Fig. S3B-C). A negative value means that contralateral stepping declined faster than ipsilateral stepping.

Surgical procedures

Unilateral lesion of dopamine neurons with 6-OHDA

Anaesthesia was induced by intraperitoneal (IP) injection of ketamine and xylazine, followed by a subcutaneous (SC) injection of bupivacaine for local anaesthesia at the incision site. Animal were head-fixed in a stereotaxic apparatus (Kopf Instruments), an incision was made to expose the skull, and 6-hydroxydopamine (6-OHDA, Sigma, St. Louis, MO; Cat# H-116; 4.5 µg dissolved in 1.5 µL of 0.05% ascorbic acid in 0.9% saline) was injected unilaterally into either left or right medial forebrain bundle (MFB) (coordinates: AP −1.3 mm and ML +1.3 or −1.3 mm from Bregma, and DV −5.0 mm from dura) through a small borehole in the skull. Injections into either left vs. right MFB were randomized between mice. 6-OHDA was infused at a rate of 0.2 µL/min through a stainless-steel cannula (Braintree Scientific, Braintree, MA; Cat# RM-SBL STD), which was left in place for an additional 5 min before withdrawal and incision closure. Desipramine (Sigma Cat# D3900; 25 mg/kg dissolved in water, IP) was given 30 min prior to 6-OHDA infusion to block uptake of 6-OHDA by noradrenergic neurons. Intensive post-operative care included providing supplemental nutrition (Bacon Softies; Bio-Serv, Flemmington, NJ; Cat# F-3580,) and extra fluids (saline, SC and dextrose saline, IP). The health status of the animals was monitored daily until stabilization of body weight. Sham-lesioned control mice received the same volume of vehicle (0.05% ascorbic acid in 0.9% saline) unilaterally into either left or right MFB (randomized between mice).

Infusion of adeno-associated virus (AAV)

AAV was injected unilaterally into the striatum through stereotaxic surgery as described for 6-OHDA. Briefly, anaesthesia was induced by ketamine and xylazine (IP), with local bupivacaine (SC) at incision site. Animal were head-fixed in the stereotaxic apparatus, an incision was made to expose the skull, and a small borehole was drilled in the skull above the microinjection site (see below for coordinates). A 33-gauge stainless-steel injector (PlasticsOne, Roanoke, VA; Cat# C315IA/SPC) was slowly lowered to the target region. The injector remained in place for 5 min before AAV was injected at a rate of 0.08 µL/min. The injector was left in place for 10 min after the injection to allow the virus to diffuse away. The injector was then slowly retracted, and the incision site closed. Injection of specific viruses are described below.

      Selective expression of hM4Di in iSPNs. To selectively and reversibly inhibit iSPNs, we infused AAV encoding Cre-dependent hM4Di-DREADD (9) into the striatum of Adora2a-Cre mice two weeks after unilateral 6-OHDA lesion. Specifically, AAV5-hSyn-hM4D(Gi)-mCherry (AAV-FLEX-hM4Di; pAAV-hSyn-hM4D(Gi)-mCherry was a gift from Bryan Roth, Addgene viral prep # 50475-AAV5; http://n2t.net/addgene: 50475; RRID: Addgene_50475) was unilaterally infused into the striatum (1.5 µL, left or right matching the same side as previous 6-OHDA injection; coordinates AP+0.3 mm and ML +2.4 or −2.4 mm from Bregma, and DV −2.8 mm from dura).

Control mice received infusion of AAV5-hSyn-DIO-mCherry instead (AAV-FLEX-mCherry; pAAV-hSyn-DIO-mCherry was a gift from Bryan Roth, Addgene viral prep # 50459-AAV5; http://n2t.net/addgene: 50459; RRID:Addgene_50459). Mice in the experiment shown in Fig. 4A-C were implanted with intra-striatal cannula during the same surgery (see below). Clozapine-*N*-oxide (CNO) treatments started 3-4 weeks after AAV infusion.

Selective ablation of cholinergic interneurons (ChIs). To selectively ablate ChIs, we infused AAV encoding Cre-dependent diphtheria toxin A subunit (DTA) into the striatum of ChAT-IRE-Cre mice, which led to ChI ablation (10). Specifically, AAV5-EF1a-FLEX-DTA-mCherry-WPRE (AAV-FLEX-DTA; University of North Carolina Gene Therapy Center) was unilaterally infused into the striatum (2 µL, left or right randomized between mice; coordinates AP+0.3 mm and ML +2.2 or −2.2 mm from Bregma, and DV −2.8 mm from dura). Control mice received infusion of AAV5-EF1a-DIO-EYFP instead (AAV-FLEX-EYFP; pAAV-Ef1a-DIO EYFP was a gift from Karl Deisseroth, Addgene viral prep # 27056-AAV5; http://n2t.net/addgene:27056; RRID: Addgene_27056). For comparison, in control ChAT-IRES-Cre mice that virally expressed a Cre-dependent EYFP reporter in the striatum, 88.4±1.5% of EYFP-labelled cells also expressed the ChI marker ChAT (mean ± s.e.m., N=13 mice), consistent with ChI selectivity. Step task training started after a 3-week recovery period. For the experiment examining the effect of ChI ablation on 6-OHDA-induced motor deficits (Fig. 4E-I, S7F-G), 6-OHDA was injected into the MFB on the same side as the previous AAV injection about 4 weeks after AAV injection.

Intra-striatal cannula placement (for experiment shown in Fig. 4A-C)

To allow for clozapine *N*-oxide (CNO) infusion into the striatum, an intra-striatal cannula was implanted unilaterally into the dopamine-depleted striatum two weeks after 6-OHDA surgery. Briefly, anaesthesia was induced by ketamine and xylazine (IP), with local bupivacaine (SC) at incision site. Animal were head-fixed in the stereotaxic apparatus, the skull was exposed, and two anchor screws were attached to the skull. A small borehole was drilled above the targeted site (left or right matching the same side as previous 6-OHDA injection; coordinates AP+0.3 mm and ML +2.4 or −2.4 mm from Bregma), and a 26-gauge stainless-steel cannula (PlasticsOne, Roanoke, VA; Cat# C315GA/SPC) was slowly lowered to DV−2.3 mm from skull surface. Because the internal injector projected 0.9 mm past the tip of the cannula, the drug would be infused to DV−3.2 mm from skull surface. Dental cement (Ortho-Jet, Lang Dental, Wheeling, IL; Cat# 1323CLR) was used to secure the cannula to the anchor screws and the head. A dummy cap (PlasticsOne, Roanoke, VA; Cat# C315DC/SPC) was screwed onto the cannula for protection. At least one week of recovery time was given before behavioural testing began (i.e., three weeks after 6-OHDA lesion).

Drug treatments

All drugs were dissolved in 0.9% w/v saline (vehicle) unless otherwise stated. Systemic treatments were given SC instead of IP, because IP but not SC injections resulted in highly variable drug absorption that could result in "dose failure" (11). The dopamine precursor L-DOPA (L-3,4-dihydroxyphenylalanine methyl ester, 1 mg/kg; Sigma-Aldrich Cat# D-1507) was given as a cocktail with the peripheral aromatic L-amino acid decarboxylase inhibitor benserazide (12.5 mg/kg; Sigma-Aldrich Cat# D-7283), injected SC 20 min before Step or Pole task. The D1R antagonist SCH23390 (0.1 mg/kg; Tocris Cat# 0925) was injected SC 30 min before Step task. The D2R antagonist eticlopride (0.16 mg/kg; Sigma-Aldrich Cat# E-101) was injected SC 30 min before Step task. The D1R agonist SKF81297 (0.2 mg/kg; Tocris Cat# 1447) was injected SC 15-20 min before Step task. The D2R agonist quinpirole (0.8 mg/kg; Tocris Cat# 1061) was injected SC 20-30 min before Step task. The doses of SKF81297 and quinpirole were chosen based on pilot studies that found that these doses reversed parkinsonian ipsilateral turning bias. The hM4Di agonist CNO (clozapine N-oxide, 3 mg/kg; from NIMH Chemical Synthesis and Drug Supply Program; dissolved in 5% DMSO in saline [vehicle]) was injected SC 45-50 min before Step task for the experiment shown in Fig. 2E. The $A_{2A}R$ antagonist istradefylline (3 mg/kg; Tocris Cat# 5147; dissolved in 8% Tween80 in 0.9% saline [vehicle]) was injected SC 25 min before Step task.

Intra-striatal infusions

Intra-striatal CNO infusion (for experiment shown in Fig. 4A-C). To avoid potential off-target effects of systemic CNO treatment (12), CNO was infused into the striatum unilaterally via the implanted cannula for this experiment, using an automated syringe pump (Chemyx, model: Fusion 100), Hamilton syringe (5 µL, Cat# 75 RN SYR), and a 33-gauge stainless-steel injector with 0.9 mm projection past the tip of the cannula (PlasticsOne, Roanoke, VA; Cat# C315LI/SPC) targeting a final coordinate of AP+0.3 mm and ML +2.4 or −2.4 mm from Bregma, and DV−3.2 mm from skull surface. The dummy cap was removed from the cannula, and the injector was inserted. The injector remained in place for 5 min before drug infusion began. Drug was infused at a rate of 62.5 nL/min, and the injector remained in place for 5 min after drug infusion had ended before retraction. The dummy cap was then screwed back onto the cannula. The mouse was allowed to move freely in a cage during the infusion. The hM4Di agonist CNO (clozapine N-oxide dihydrochloride, 500 nL, 50 µM, Tocris Cat# 6239; dissolved in 1× phosphate-buffered saline [vehicle], Corning Cat# 21040CV) or its vehicle was infused for the experiment shown in Fig. 4A-C, and Step task was tested 30-35 min after the start of infusion.

Group assignments

Mice were randomly assigned into treatment groups. For experiments that manipulated parkinsonian decline of long-term rescue, mice were counterbalanced according to both their

dopamine-depleted baseline performance and their long-term rescue probe performance (last session before washout) prior to being assigned to treatment groups (statistics included in corresponding Dataset S1). Drug treatment groups were masked (blinded) for data collection and analysis, except for the following (not blinded): data shown in Fig. 2C (long-term rescue by Veh/SKF/Quin), Fig. 4C (parkinsonian decline of long-term rescue with Veh/CNO), Fig. S8B (LDR decay with Veh/SKF/Quin). It was not possible to blind Sham vs. 6-OHDA lesion due to the obvious ipsilateral turning bias shown by 6-OHDA-lesioned mice. Genotypes were not blinded – it was difficult to mask D1R-KO mice as they were noticeably physically smaller. Virus infusions were not blinded.

Immunohistochemistry and image analysis

Mice were deeply anaesthetized with ketamine and xylazine, then fixed by transcardial perfusion with saline followed by 4% paraformaldehyde (PFA). Brains were dissected from the skull and post-fixed in 4% PFA at 4°C overnight. Brains were then sectioned at 30 μm with a vibratome (VT1000S, Leica Biosystems, Buffalo Grove, IL) and stored at -20°C in cryoprotectant (glycerol 25% v/v; ethylene glycol 30% v/v; phosphate buffer 0.05M). For immunofluorescence staining, 3-6 free-floating brain slices spanning the striatum (approximately between AP+1.18 to −0.58 mm from bregma) were washed 3×10 min at room temperature (RT) in Tris buffer saline (TBS, 0.05M), blocked in 5% normal donkey serum (Jackson ImmunoResearch, Cat# 017-000-121) in 0.3% Triton X-100 (Fisher Scientific) at RT for 1 h, then incubated in one or several primary antibodies for 72 h at 4°C on an orbital shaker. Brain slices were then washed 3×10 min at RT in 0.1% Tween 20 (Fisher Scientific) in TBS, blocked in 5% normal donkey serum at RT for 30-60 min, then incubated in one or several secondary antibodies (1:500, except otherwise noted) for 1 h at RT on an orbital shaker. Brain slices were then washed 3×10 min at RT in TBS, then mounted on Superfrost Plus coated slides (Fisher Scientific) and coverslipped with ProLong Gold Mountant with DAPI (Invitrogen Cat# P36931) or with Fluoromount-G (SouthernBiotech Cat# 0100-01). Slides were stored at -20°C in the dark until imaging.

Fluorescent images (8 bit) were acquired using a Leica DM6 confocal microscope with a 20x/0.7 numerical aperture oil immersion objective (Leica HCX PL APO CS, for experiments in Fig. 1-4 and S4-S7), or Leica TCS SP2 confocal microscope scanner with a 5x air objective (for tyrosine hydroxylase [TH] optical density), or ZEISS LSM 800 with a 20x/0.8 numerical aperture air objective (for the experiment shown in Fig. 3H-J, 5D-F, S9C-E and S10B-D). Images from each individual immunohistochemical experiment were acquired using the same microscope.

Primary and secondary antibodies

Primary antibodies used were rabbit anti-tyrosine hydroxylase (1:500, Pel-Freez Cat# P40101), rabbit anti-GFP (1:1000, Invitrogen Cat# A-11122; this antibody also recognizes EYFP), chicken

anti-GFP (1:1000, Abcam Cat# ab13970), goat anti-c-Fos (1:500, Santa Cruz Biotechnology Cat# sc-52-G), rabbit anti-c-Fos (1:500, Santa Cruz Biotechnology Cat# sc-52), goat anti-ChAT (1:100, R&D Systems Cat# AF3447), rat anti-CTIP2 (1:500, Sigma-Aldrich Cat# MABE1045), guinea pig anti-c-Fos (1:500, Synaptic Systems Cat# 226004. This was used together with goat anti-ChAT antibody [R&D Systems Cat# AF3447] because we found that rabbit anti-c-Fos [Santa Cruz Cat# sc-52] cross reacted with the goat anti-ChAT antibody; this was also used together with rat anti-CTIP2), rat anti-mCherry (1:1000, Invitrogen Cat# M11217), rabbit anti-DARPP32 (1:500, Cell Signaling Cat# 2306), and rabbit anti-p-rpS6$^{S235/236}$ (1:500, Cell Signaling Cat# 4858).

Secondary antibodies used were donkey anti-rat Alexa Fluor Plus 405 (1:250, Invitrogen Cat# A48268), donkey anti-rat DyLight 405 (1:100, Jackson ImmunoResearch, Cat# 712-475-153), donkey anti-rabbit Alexa Fluor 488 (Invitrogen Cat# A-21206), donkey anti-chicken Alexa Fluor 488 (Jackson ImmunoResearch, Cat# 703-545-155), donkey anti-goat Alexa Fluor 555 (Invitrogen Cat# A-21432), donkey anti-rabbit Alexa Fluor 555 (Invitrogen Cat# A-31572), donkey anti-rat Alexa Fluor 594 (Invitrogen Cat# A-21209), donkey anti-rabbit Alexa Fluor 647 (Invitrogen Cat# A-31573), donkey anti-goat Alexa Fluor 680 (Invitrogen Cat# A-21084), donkey anti-rabbit Alexa Fluor 680 (Invitrogen Cat# A10043), and donkey anti-guinea pig Alexa Fluor 680 (Jackson ImmunoResearch, Cat# 706-625-148).

Image analysis

Image analysis for all figures except Fig. 3H-J, 5, S9 and S10. Immunohistochemistry (IHC) images were analysed using Fiji (ImageJ). Tyrosine hydroxylase (TH) optical density for each striatal side (Fig. 1A) was calculated by first subtracting the background optical density (measured from the same side's corpus callosum). For each subject, the TH optical density (% Injected side/Non-injected side) was averaged across striatal slices (anterior to posterior), then $\log_{10}$-transformed before analysis. For cell counting experiments (iSPNs, ChIs, Fos-positive cells) except those in Figs. 3H-J, S9, and S10, labelled cells were automatically segmented and detected using an ImageJ macro written in-house that performed de-speckling via median filtering, local background subtraction, watershed segmentation, and thresholding based on particle circularity, edge sharpness, size, and fluorescence intensity. Each label type analysed had different set of thresholding parameters (e.g., size threshold was bigger for ChAT$^+$ than for Fos$^+$) that were fixed across images for a given experiment. For all cell-counting analyses, density of labelled cells was calculated for each subject by pooling across dorsolateral striatum (DLS) and dorsomedial striatum (DMS), for up to 6 anterior to posterior striatal slices, calculated as: (total number of labelled cells summed across DLS and DMS regions of interest in all slices) / (total area of regions of interest summed across DLS and DMS in all slices).

Image analysis for Fig. 3H-J, 5, S9 and S10. For these experiments, in order to identify dSPNs vs. iSPNs, we used IHC to label for CTIP2 in addition to GFP in Drd2-EGFP mice. In the

striatum, CTIP2 selectively labels dSPNs and iSPNs (13), allowing us to identify dSPNs (CTIP2$^+$GFP$^-$) and iSPNs (CTIP2$^+$GFP$^+$). Furthermore, to better characterize SPN activation, we co-labelled for Fos and ribosomal protein S6 Ser235/236 phosphorylation (p-rpS6$^{S235/236}$). This is because p-rpS6$^{S235/236}$ is a proxy marker for cAMP/protein kinase A (PKA) activation in SPNs (14, 15), and the time course of Fos protein induction and p-rpS6$^{S235/236}$ were similar, facilitating IHC co-labelling. Images were segmented using Cellpose – a convolutional neural net that had been pretrained for cell segmentation (16) – to generate mask proposals separately for each channel (GFP, CTIP2, Fos, p-rpS6$^{S235/236}$). The "cytoplasm" model of Cellpose (model_type='cyto') was used for all channels, even for channels that labelled the nuclei (Fos, p-rpS6$^{S235/236}$), because we found that the "cytoplasm" model segmented nuclei protein immunofluorescence better. Cellpose performance was assessed with the standard machine learning metric of precision and recall, using test sets that were manually segmented by human experts and previously unseen by Cellpose. Assessment results were: GFP (precision = 0.95, recall = 0.96), CTIP2 (precision = 0.96, recall = 0.91), Fos (precision = 0.92, recall = 0.93), p-rpS6$^{S235/236}$ (precision = 0.93, recall = 0.94). To select SPNs that were activated above background levels, a mask proposal for activation marker channels (Fos and p-rpS6$^{S235/236}$) was accepted if it satisfied the following criterion: $Intensity_{in} - Intensity_{out} \geq Intensity_{threshold}$. For both Fos and p-rpS6$^{S235/236}$, $Intensity_{out}$ was defined as the average pixel intensity of a 3-pixel-wide ring immediately surrounding the mask (excluding pixels that were part of another mask proposal). For Fos, $Intensity_{in}$ was the average pixel intensity of all pixels inside the mask, and $Intensity_{threshold}$ was 20. For p-rpS6$^{S235/236}$, $Intensity_{in}$ was the average pixel intensity of a 2-pixel-wide ring immediately inside the mask, including the mask circumference. A ring was used for p-rpS6$^{S235/236}$ because p-rpS6$^{S235/236}$ labelling was cytoplasmic but excluded the nucleus, resulting in ring-like staining (Fig. S5B). $Intensity_{threshold}$ for p-rp$^{S6235/236}$ was 40.

      We used the following thresholding heuristic to determine whether a given mask from channel A overlapped with masks from channel B. The pixels, if any, which overlapped with the given mask from channel A and a mask from channel B were identified as proposed mask overlap. The proposed mask overlap area (number of pixels) was totalled as $area_{overlap}$. The areas of the original masks (number of pixels) from channels A and B were also counted, and the smaller area was designated $area_{original}$. If $area_{overlap}/area_{original} \geq 0.6$, then the proposed mask overlap is accepted as real overlap, otherwise it is rejected as spurious overlap. We used the area of the smaller of the original mask as $area_{original}$, because channels that have nuclei masks (CTIP2, Fos) are expected to have smaller masks than channels that have cytoplasmic masks (GFP, p-rpS6$^{S235/236}$). We used 0.6 as the overlap area proportion threshold to allow for Cellpose's variability in demarcating the precise outlines of cells and nuclei. For thresholding mask overlaps across >2 channels, the above algorithm was used, with the change that $area_{original}$ is the area of the smallest original mask across all channels. Examples of Cellpose's

segmentation for each channel and outcomes of the overlap mask thresholding algorithm are shown in Fig. S5B. Activation measure for each subject was calculated by pooling across dorsolateral striatum (DLS) and dorsomedial striatum (DMS), for up to 6 anterior to posterior striatal slices. For example, dSPN %Fos$^+$ was calculated as (total number of CTIP2$^+$GFP$^-$Fos$^+$ across DLS and DMS in all slices)/ (total number of CTIP2$^+$GFP$^-$ across DLS and DMS in all slices) × 100%. Other combinations of SPN activation measures were similarly calculated. SPN pathway bias index for Fos$^+$ was calculated as $(P_{dSPN:Fos+}{}^{\alpha} - P_{iSPN:Fos+}{}^{\alpha}) / (P_{dSPN:Fos+} + P_{iSPN:Fos+})$, where $P_{dSPN:Fos+}$ vs. $P_{iSPN:Fos+}$ denotes the proportion of dSPN vs. iSPN activation respectively for Fos (i.e., % activated / 100). SPN pathway bias index for other activation measures were similarly calculated. Two key components of the SPN bias index – exponentiation thresholding (α) and divisive normalization – are often used in neural computational modelling (17).

Detailed statistical analysis

Unless otherwise stated, statistical tests were conducted using R (R core team, version 4.0.5). Linear regression was fitted using linear least square in R (function "lm"). Nonlinear least square (exponential decay function) was conducted with the R package "minpack.lm" (version 1.2-1). ANOVA was conducted with the R package "ez" (version 4.4-0) using Type III Sum-of-Squares method. Likelihood ratio test of nested models was conducted with the R package "lmtest" (version 0.9-40). All tests of significance were conducted at $\alpha = 0.05$, two-tailed, unless otherwise stated. For independent samples (between-subject data) with >2 groups and/or multiple factors, independent sample ANOVA was conducted, with homogeneity of variance tested by Levene's test. If homogeneity of variance was violated for one-way ANOVA, Kruskal-Wallis test (non-parametric ANOVA) was conducted instead. For repeated-measure samples (within-subject data) with >2 groups and/or multiple factors, repeated-measure ANOVA was conducted. Mauchly's test of sphericity was used to verify homogeneity of variance for repeated-measure samples. If sphericity was violated, Huynh–Feldt ε correction (HFε) was used. For samples with both independent and repeated measures, mixed ANOVA was conducted. Significant interactions from ANOVAs were followed up with tests of simple effects using unifactorial ANOVAs. Significant main effects from ANOVAs were followed up with the appropriate post-hoc tests of simple effects: for independent sample factors, Tukey HSD tests were conducted; for Kruskal-Wallis test (non-parametric ANOVA), independent sample t-tests for unequal variance (Welch's t-tests) were used; for repeated measures factors, paired t-tests were used. The Holm-Bonferroni correction was used to control for family-wise error rate from multiple comparisons with post-hoc t-tests.

For comparisons involving two independent samples, Shapiro-Wilk test of normality was first conducted for each group. If normality assumption was violated, Mann-Whitney-Wilcoxon test (non-parametric t-test) was conducted. If normality assumption was satisfied, independent sample t-test was conducted, and variances were assumed to be equal or not according to the F-

test of variance equality. For comparisons involving two related samples, paired-sample t-tests were conducted. When needed, the Holm-Bonferroni correction was used to control for family-wise error rate from multiple comparisons.

For figures showing long-term rescue induced by repeated treatments (Fig. 2B, 2D, S4B, S5H, S7G), between-group comparisons of DA-depleted parkinsonian baseline (BL) are shown. Furthermore, filled arrows show the first day of detectable treatment-induced acute rescue vs. BL for each group, defined as the first day of at least two consecutive days during which on-treatment stepping was significantly higher than BL (Holm-Bonferroni corrected paired t-tests). Open arrows show the first day of detectable long-term rescue vs. BL for each group, defined as the first day of at least two consecutive days during which off-treatment stepping (pre-treatment probe) was significantly higher than BL (Holm-Bonferroni corrected paired t-tests). Finally, between-group comparisons of acute rescue in the final on-treatment session during the repeated treatment phase are shown in the figures. The acute rescue averaged across days was also statistically analysed in the corresponding Dataset S1.

Analysing acute and long-term rescue induced by targeting D1R/dSPNs vs. D2R/iSPNs vs. both pathways (Fig. 2F-H). Subjects from long-term rescue induction experiments in Fig. 2B and 2D were pooled into groups according to the pathway targeted during repeated treatments, either D1R/dSPNs (L-DOPA+eticlopride), D2R/iSPNs (L-DOPA+SCH23390, or iSPN hM4Di), or both (L-DOPA+vehicle), as shown in Fig. 2E. We examined whether the following measures differed between pathway-targeted groups: acute rescue (final on-treatment session), long-term rescue (No-treatment session post-washout), and long-term/acute rescue ratio. Non-parametric statistics were used to test differences between pathway-targeted groups (Dataset S1).

Linear regression models of parkinsonian decline slope against different SPN activation measures (Fig. S9C-E and S10B-D), and model comparison (Fig. S9G and S10F). We used linear regression models to examine the relationship between parkinsonian decline (and its prevention by dopamine receptor agonists) vs. SPN activation. For SPN activation, we examined 3 different activation measures (Fos$^+$ vs. p-rpS6$^{S235/236+}$ vs. double-labelled Fos$^+$p-rpS6$^{S235/236+}$). For dSPN-alone and iSPN-alone activation measures, their directly observed measures (% of SPN activated; Fig. 3H-J and 5D-F) were used as linear predictors (Fig. S9C-D and S10B-C, x-axes) to predict parkinsonian decline slope for the corresponding subject (Fig. S9C-D and S10B-C, y-axes). dSPN-alone and iSPN-alone linear models were fitted using linear least square (R function "lm"). Additionally, for each activation measure, dSPN activation and iSPN activation were combined using a SPN bias index: $(P_{dSPN+}{}^\alpha - P_{iSPN+}{}^\alpha) / (P_{dSPN+} + P_{iSPN+})$, where P_{dSPN+} vs. P_{iSPN+} denotes the proportion of dSPN vs. iSPN activation respectively for a particular activation measure (i.e., [% activated]/100) (Fig. S9A). SPN bias index ranges from negative (more iSPNs activated than dSPNs), to 0 (dSPNs and iSPNs equally activated), to positive (more dSPNs activated than iSPNs). The exponent α acts as a sensitivity threshold that modulates SPN

activation in the numerator. SPN bias indices were used as linear predictors (Fig. S9E and S10D, x-axes) to predict parkinsonian decline slope (Fig. S9E and S10D, y-axes). While P_{dSPN+} and P_{iSPN+} are direct observations (IHC data), α is a free parameter – similar to regression intercept and regression slope – that was shared between mice. Due to the presence of α, SPN bias indices were fitted to parkinsonian decline slope using non-linear least square (R package "minpack.lm", function "nlsLM"), yielding maximum likelihood (ML) estimates of α, regression intercept and regression slope, and their standard error (s.e.). The ML estimates of α for each activation measure (shown in Fig. S9F and S10E, with ±s.e.) were substituted into the equation of SPN bias index to produce the resultant SPN bias indices shown as the x-axes of Fig. S9E and S10D. For the experiment in Fig. S10, α for the Sham-lesioned (dopamine-intact) group were fixed to values of α estimated from the dopamine-intact mice in Fig. S9 (i.e., α = 1.45, 1.34, and 1.33 for Fos$^+$, p-rpS6$^{S235/236+}$, and Fos$^+$p-rpS6$^{S235/236+}$ respectively), while α for the other 3 groups (dopamine-depleted) were free parameters that were fitted to the data. This allowed α from dopamine-depleted mice to be estimated without being "contaminated" by a potentially different α from the dopamine-intact group. Forcing α to be the same across all 4 groups in Fig. S10 led to the same conclusion: α in Fig. S10E were not significantly different from 1.

Least-square fitting yielded Holm-Bonferroni-corrected p-values for the slope of the linear fits, and log-likelihood of each model. R^2 for linear models with SPN bias indices were calculated by first substituting the ML estimate of α into the SPN bias index equation. Log-likelihood of each model was used to calculate the model's Akaike Information Criterion corrected for small sample size (AICc) [18]. AICc allows comparison of relative model likelihood, penalizing model complexity (the number of free parameters) to discourage overfitting. Models that used dSPN-alone or iSPN-alone activation measures had 3 free parameters (intercept, slope, residual variance). Models that use SPN bias index had one additional free parameter (α). The AICc difference, $\Delta AICc_i$, for a given model i is calculated as $AICc_i - AICc_{min}$, where $AICc_i$ is the AICc for model i and $AICc_{min}$ is the minimum AICc across all models considered. $\Delta AICc_i$ was then used to calculate relative model likelihood shown in Fig. S9G and S10F for model comparison: *relative likelihood of model i* = exp($-\Delta AICc_i/2$) [18]. Finally, for Fig. S9F and S10E, we tested whether α were significantly different from 1 (default value) by comparing the log-likelihood of models that allowed α to vary as a free parameter vs. the log-likelihood of their nested models with α fixed to 1, using the likelihood ratio test (R package "lmtest", function "lrtest") to compute a χ^2 statistic.

Figures were plotted using R (R core team, version 4.0.5) and the R package "ggplot2" (version 3.3.3). Reported values are means ± s.e.m. unless otherwise stated.
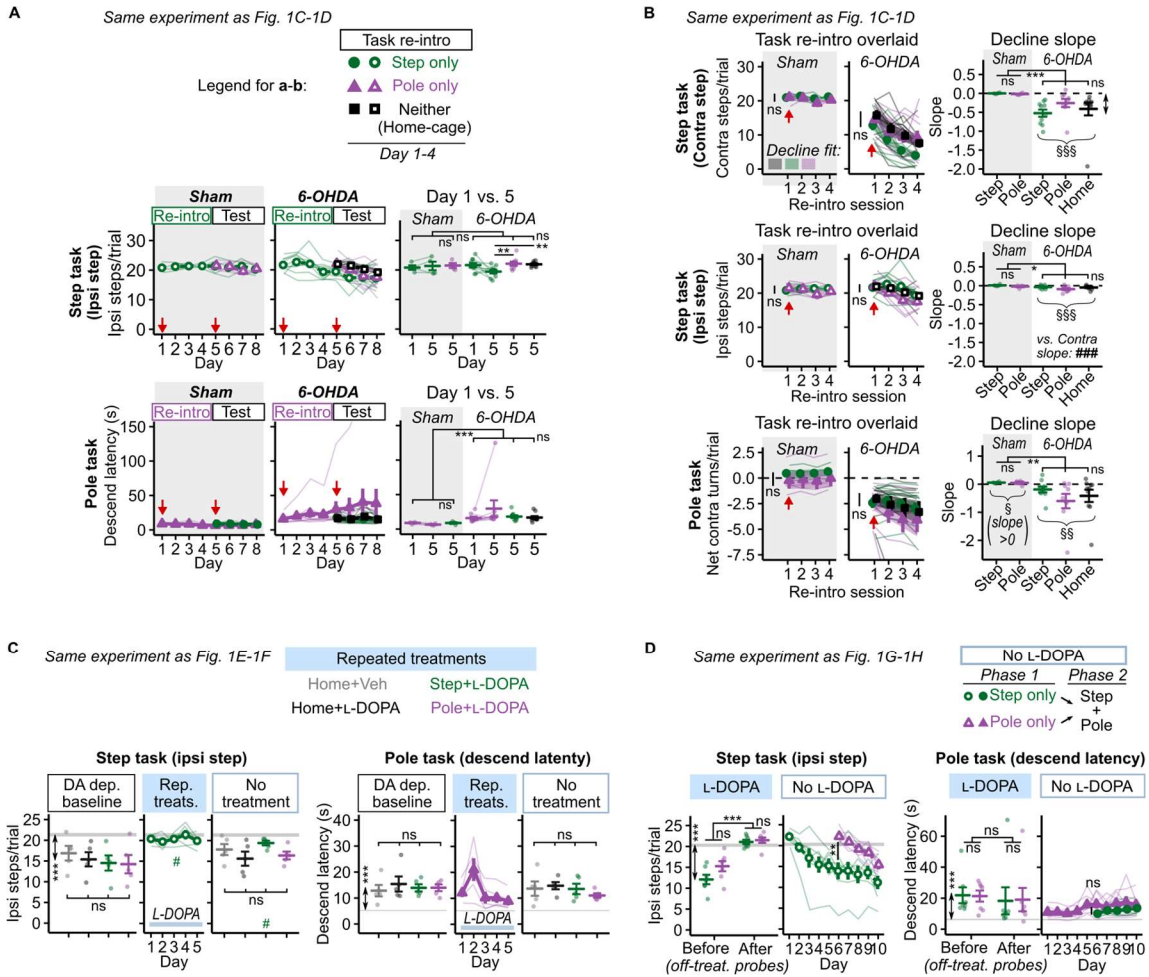
Fig. S1. Additional parkinsonian performance measures on Step task and Pole task (related to Fig. 1)

(**A**) Step task ipsilateral stepping (top) and Pole task descend latency (bottom) from same experiment as in Fig. 1C-1D. Right panels, Day 1 (first session for the group re-introduced to Step/Pole task during Day 1-4) vs. Day 5 (first session for remaining groups). (**B**) Post-6-OHDA motor performance from same experiment as in Fig. 1C-1D. "Task re-intro overlaid" panels show parkinsonian decline during the first 4 sessions upon re-introduction to the task post-6-OHDA. "Decline slope" panels show fitted decline rates of "Task re-intro overlaid" panels (Methods); negative values indicate worsening performance. Note that although 6-OHDA also caused a mild decline in ipsilateral stepping, contralateral stepping declined much more steeply: [###]$P<0.001$, paired t-test. (**C**) Step task ipsilateral stepping (left) and Pole task descend latency (right) from same experiment as in Fig. 1E-1F. [#]$P<0.05$ vs. same group's DA-depleted baseline, paired t-test. (**D**) Step task ipsilateral stepping (left) and Pole task descend latency (right) from same experiment as in Fig. 1G-1H. Gray horizontal bands in (C) and (D): pre-6-OHDA performance. Data are mean ± s.e.m.. [§]$P<0.05$, [§§]$P<0.01$, [§§§]$P<0.001$, vs. zero slope (no decline), Holm-Bonferroni-corrected one-sample t-tests. *$P<0.05$, **$P<0.01$, ***$P<0.001$. ANOVAs or Holm-Bonferroni-corrected t-tests were used for between-group comparisons, followed by Tukey HSD post-hoc tests when needed. See Dataset S1 for full statistics. See also Fig. S2.
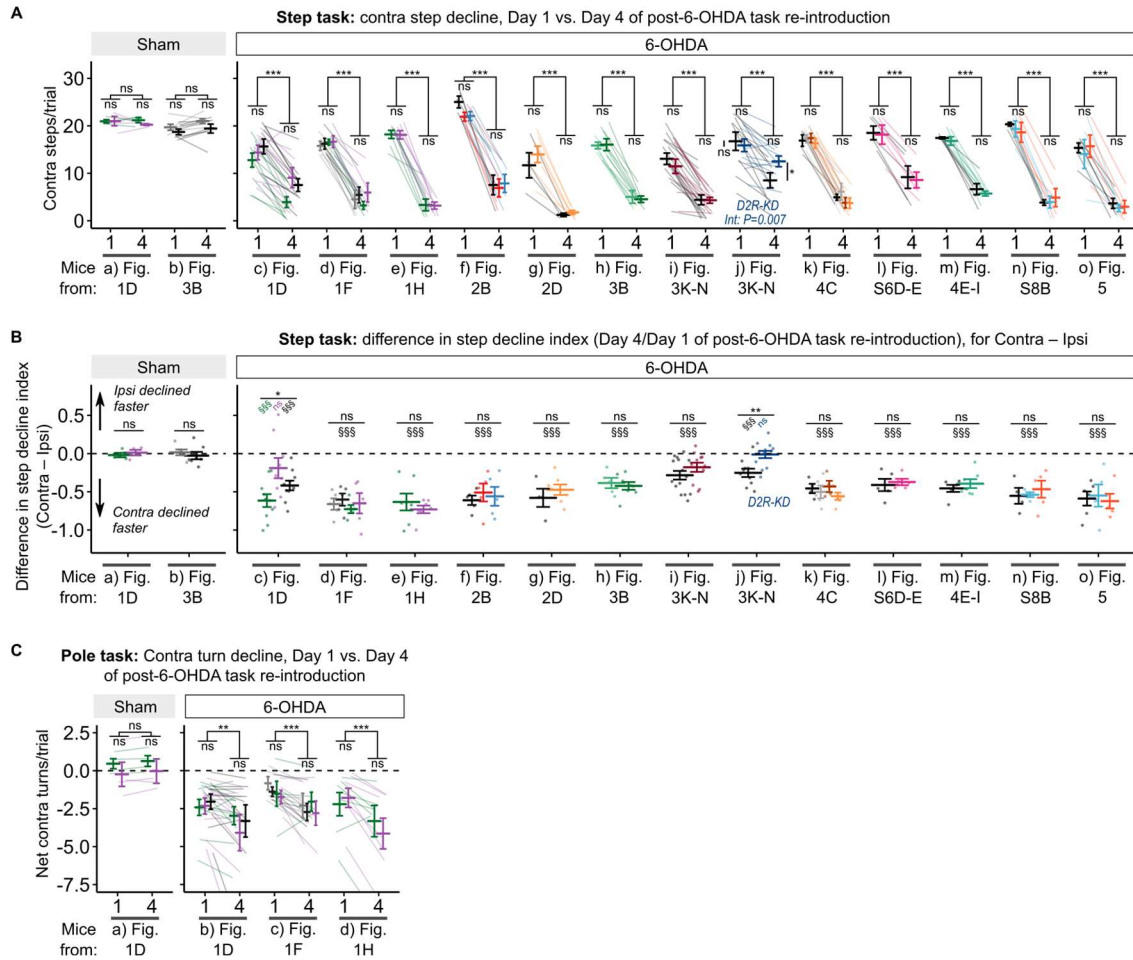
Fig. S2. Post-6-OHDA parkinsonian decline is highly replicable, and contralateral stepping declined more than ipsilateral stepping.
(**A**) Step task contralateral performance post-6-OHDA (≥3 weeks post-6-OHDA), Day 1 vs. Day 4 of task re-introduction (prior to group-assigned drug treatments). x-axis also shows the experiment to which the mice belonged in the present paper (Sham: total 21 mice across 2 independent experiments. 6-OHDA: total 221 mice across 13 independent experiments). (**B**) Post-6-OHDA Step task contralateral performance declined more than ipsilateral performance. The same mice as in (A) were analysed. Step decline index was calculated as: (Contra steps on day 4 / Contra steps on day 1) − (Ipsi steps on day 4 / Ipsi steps on day 1) (Methods). Only two groups failed to show more severe contralateral vs. ipsilateral decline: Pole re-intro group (from Fig. 1D), likely due to sampling variability, and D2R-knockdown (D2R-KD) group (from Fig. 3K-N), caused by D2R-KD occluding 6-OHDA-induced motoric worsening (see main text). (**C**) Same as in (A), but for post-6-OHDA Pole task contralateral turning decline (Sham: total 8 mice across 1 experiment. 6-OHDA: total 63 mice across 3 independent experiments). Data are mean ± s.e.m.. §§§$P<0.001$ vs. zero (no contralateral vs. ipsilateral difference), one-sample t-test. *$P<0.05$, **$P<0.01$, ***$P<0.001$. One-way ANOVAs or independent t-tests were used for between-group comparisons. Mixed ANOVAs were used to compare between-groups and repeated factors. Group colours: same as referred-to figures. See Dataset S1 for full statistics.
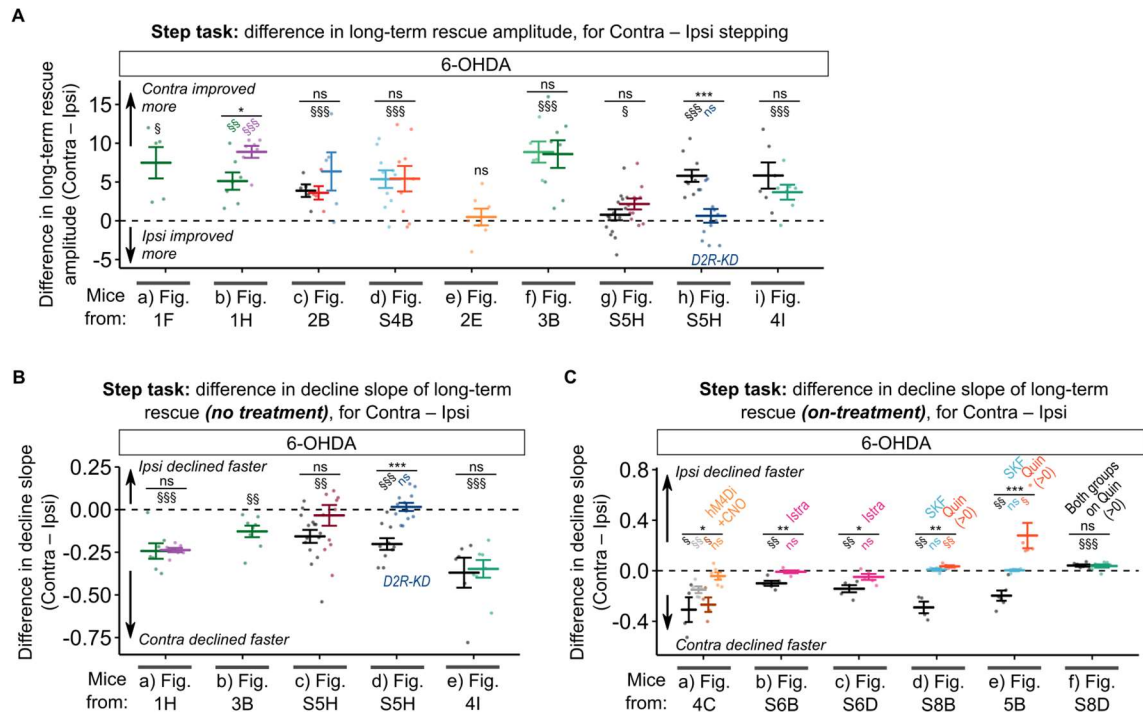
Fig. S3. Long-term rescue and its parkinsonian decline were more prominent for contralateral vs. ipsilateral stepping.

(**A**) Higher long-term rescue amplitude (Methods) for contralateral vs. ipsilateral stepping (total 134 mice across 9 independent experiments). Only groups that received Step task training paired with dopamine receptor agonist or iSPN-selective inhibition were included. Only two groups failed to show bigger contralateral vs. ipsilateral long-term rescue amplitude: D2R-KD group (from Fig. S5H), caused by D2R-KD occluding baseline parkinsonian motor deficits, and iSPN-hM4Di group (from Fig. 2D), possibly due to the modest induced long-term rescue amplitude, or sampling variability. (**B**) Long-term rescue's decline slope (no treatment, post-washout) was more severe for contralateral vs. ipsilateral stepping (total 90 mice across 5 independent experiments). Only experiments that examined parkinsonian decline while off-drug-treatment were included. The only exception was the D2R-KD group (from Fig. S5H), caused by D2R-KD occluding parkinsonian decline. (**C**) The more severe contralateral vs. ipsilateral decline was acutely blocked by treatments that blocked contralateral stepping decline (highlighted groups) (total 81 mice across 6 independent experiments). Quinpirole (Quin) unexpectedly caused stronger ipsilateral vs. contralateral decline, possibly due to quinpirole stimulating dopamine neurons' D2R autoreceptors in the intact hemisphere, causing mild acute ipsilateral motor deficits. Data are mean ± s.e.m.. §$P<0.05$, §§$P<0.01$, §§§$P<0.001$, difference score vs. zero, one-sample t-test. *$P<0.05$, **$P<0.01$, ***$P<0.001$. ANOVAs or independent sample t-tests were used for between-group comparisons, except in (B) "e) Fig. 4I" where Wilcoxon rank-sum test was used instead due to normality test violation, and in (C) "a) Fig. 4C" where Kruskal-Wallis test was used instead due to Levene's variance equality test violation. Group colours: same as referred-to figures. See Dataset S1 for full statistics.
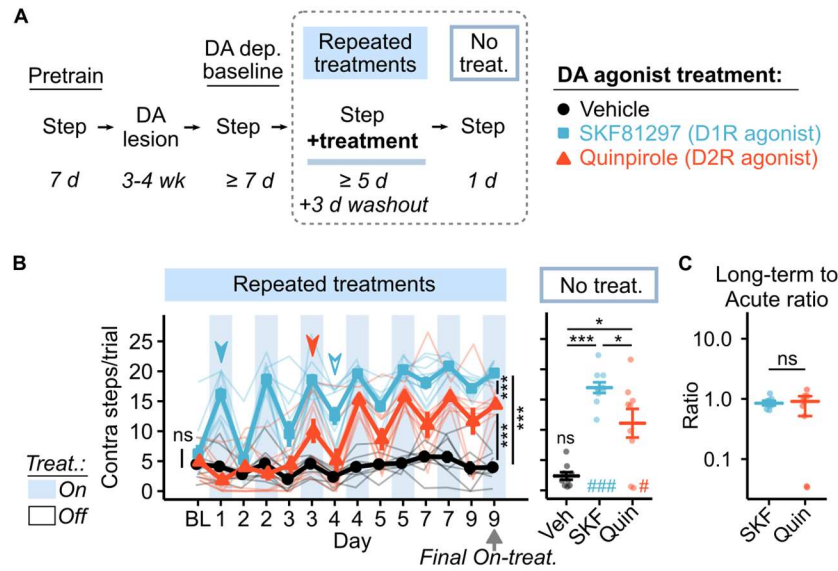
16

Fig. S4. D1R agonist potently induces acute and long-term rescue of parkinsonian deficits (related to Fig. 2).

(**A**) Strategy to examine acute vs. long-term rescue induced by stimulating D1R vs. D2R. (**B**) Effect of repeated vehicle vs. D1R agonist vs. D2R agonist. BL: parkinsonian baseline final session. Filled arrow: first day of detectable acute rescue vs. BL. Open arrow: first day of detectable long-term rescue vs. BL (Holm-Bonferroni-corrected paired t-tests, Methods). N=9, 10, 9. (**C**) Median ± IQR of long-term rescue to acute rescue ratio (note $\log_{10}$ y-axis). Data are mean ± s.e.m. unless specified. #P<0.05, ###P<0.001, vs. same group's DA-depleted baseline, paired t-tests. *P<0.05, ***P<0.001. ANOVAs or Holm-Bonferroni-corrected t-tests were used for between-group comparisons, followed by Tukey HSD post-hoc tests when needed. For (C), Kruskal-Wallis test followed by Wilcoxon rank sum tests was used instead. See Dataset S1 for full statistics.
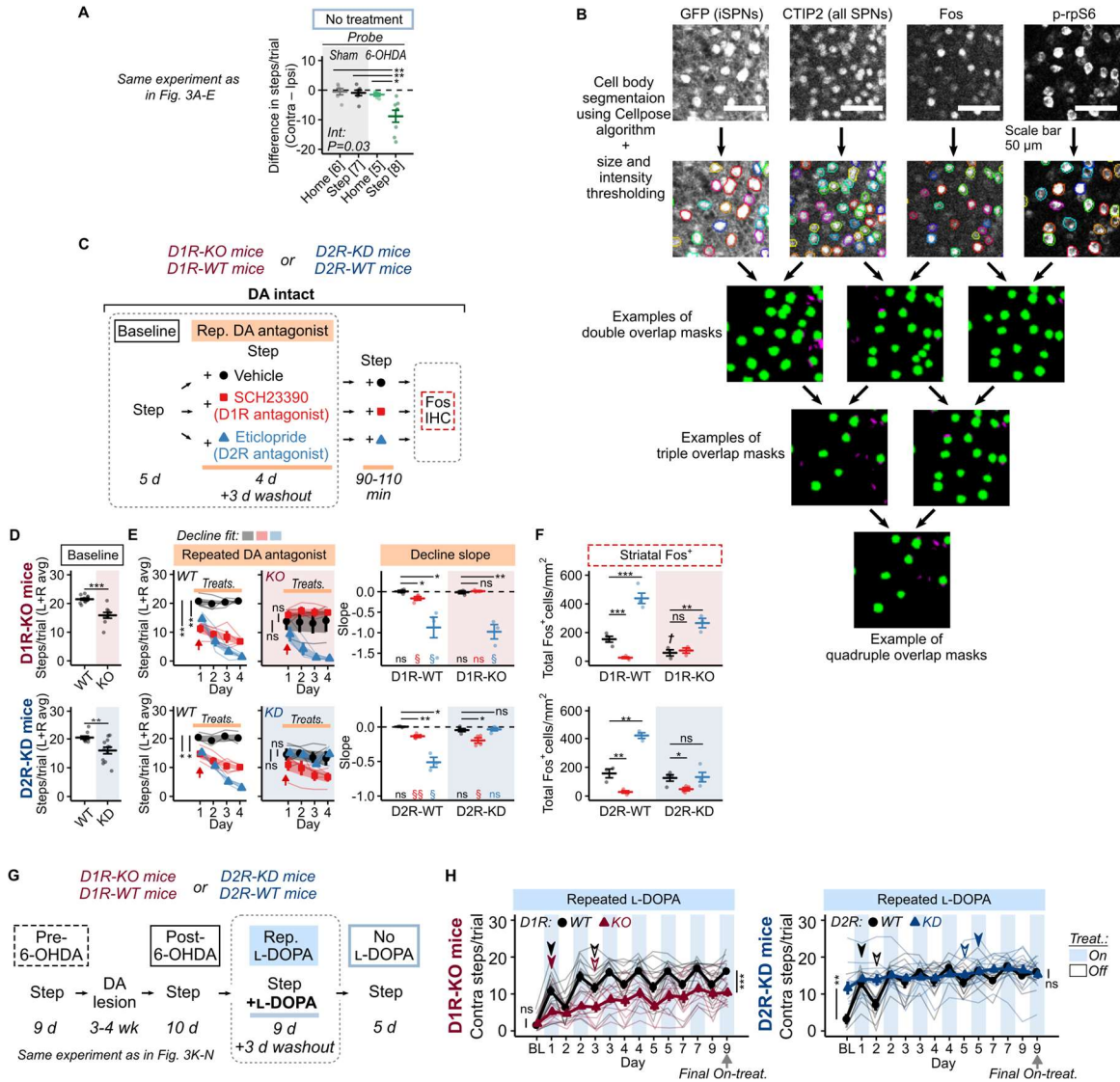
Fig. S5. Detecting SPN activation using immunohistochemistry and automatic cell segmentation, and occlusion of parkinsonian motoric deficits by D1R-KO vs. D2R-KD (related to Fig. 3).

(**A**) From the same experiment as Fig. 3A-E, showing differences in contralateral vs. ipsilateral stepping during off-treatment probe trials. (**B**) A representative example of automatic cell segmentation using Cellpose algorithm followed by size and intensity thresholding (second row, each segmented cell is outlined with a random colour), followed by mask overlap thresholding (green: accepted mask overlap; magenta: rejected mask overlap). (**C-F**) Strategy to examine occlusion of DA antagonist-induced parkinsonian deficits in DA-intact, D1R-knockout (D1R-KO) vs. D2R-knockdown (D2R-KD) mice (C), baseline stepping without treatment (D), stepping performance during repeated DA antagonist treatments (E), and striatal Fos+ quantification (F). N=4, 4, 4 (D1R-WT); 3, 3, 3 (D1R-KO); 3, 4, 3 (D2R-WT); 4, 4, 4 (D2R-KO). (**G**) Strategy to test whether D1R-knockout (D1R-KO) or D2R-knockdown (D2R-KD) occlude L-DOPA-induced acute vs. long-term rescue. Form the same experiment as Fig. 3K-N. (**H**) D1R-KO blocked L DOPA's acute rescue, resulting in reduced long-term rescue during repeated L DOPA phase (off-treatment probes averaged across Days, main effect of D1R genotype, P=0.024). D2R-KD occluded parkinsonian baseline deficit (BL), thereby also occluding subsequent rescues by L DOPA. BL: parkinsonian baseline final session. Filled arrow: first day of detectable acute rescue

18

vs. BL. Open arrow: first day of detectable long-term rescue vs. BL (Methods). Stepping declines are fitted with exponential decay, negative slope indicates decline (Methods). Data are mean ± s.e.m.. $^{§}P<0.05$, $^{§§}P<0.01$, $^{§§§}P<0.001$, vs. zero slope (no decline), Holm-Bonferroni-corrected one-sample t-tests. $^{*}P<0.05$, $^{**}P<0.01$, $^{***}P<0.001$. ANOVAs or Holm-Bonferroni-corrected t-tests were used for between-group comparisons, followed by Tukey HSD post-hoc tests when needed. Wilcoxon rank-sum test was used for (D) "D2R-WT/KD mice" and (H) BL for both "D1R-WT/KD mice" and "D2R-WT/KD mice" instead, due to normality test violation. See Dataset S1 for full statistics.
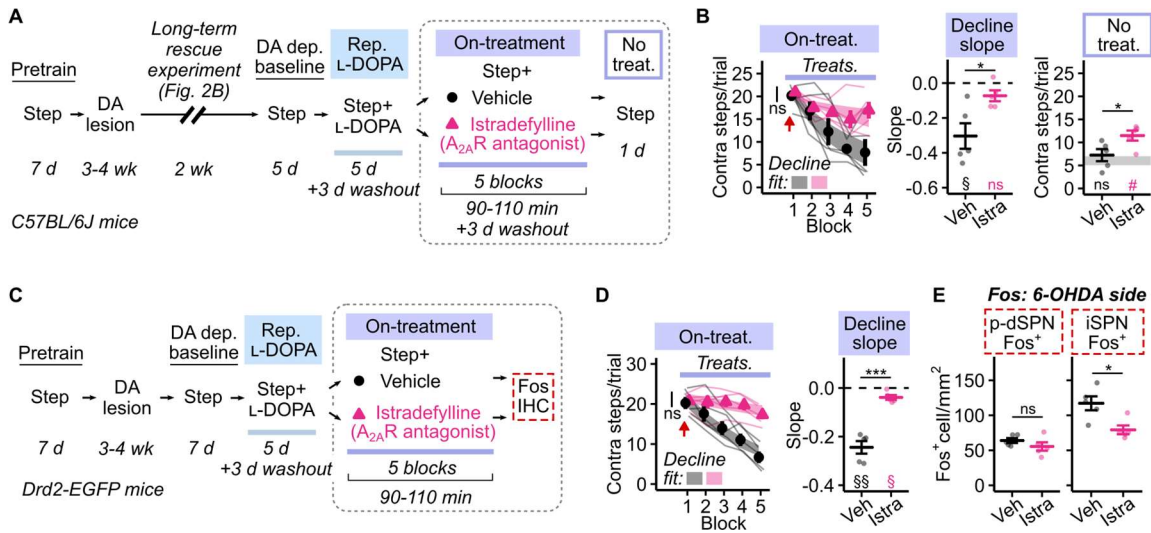
Fig. S6. A₂AR antagonist block parkinsonian decline of long-term rescue, and rescue pathological iSPN over-recruitment but not dSPN under-recruitment (related to Fig. 4).
(**A**-**B**) Strategy to test if istradefylline (A₂AR antagonist) attenuate parkinsonian decline acutely and long-term (A), and the associated contralateral stepping performance (B). Gray horizontal band in "No treat." panel: dopamine-depleted baseline stepping (groups pooled). [#]$P<0.05$ vs. same group's DA-depleted baseline, paired t-test. N=5, 5. (**C**-**D**) Strategy to test if istradefylline rescue pathological SPN Fos⁺ recruitment (C), the associated contralateral stepping (D), and SPN Fos⁺ recruitment (E). N=5, 5. Stepping declines are fitted with exponential decay, negative slope indicates decline (Methods). Data are mean ± s.e.m.. [§]$P<0.05$, [§§]$P<0.01$ vs. zero slope (no decline), Holm-Bonferroni-corrected one-sample t-tests. *$P<0.05$, ***$P<0.001$. ANOVAs or independent-samples t-tests were used for between-group comparisons, followed by Holm-Bonferroni-corrected post-hoc t-tests when needed. Wilcoxon rank-sum test was used for iSPN Fos⁺ data in (E) due to normality test violation. See Dataset S1 for full statistics.
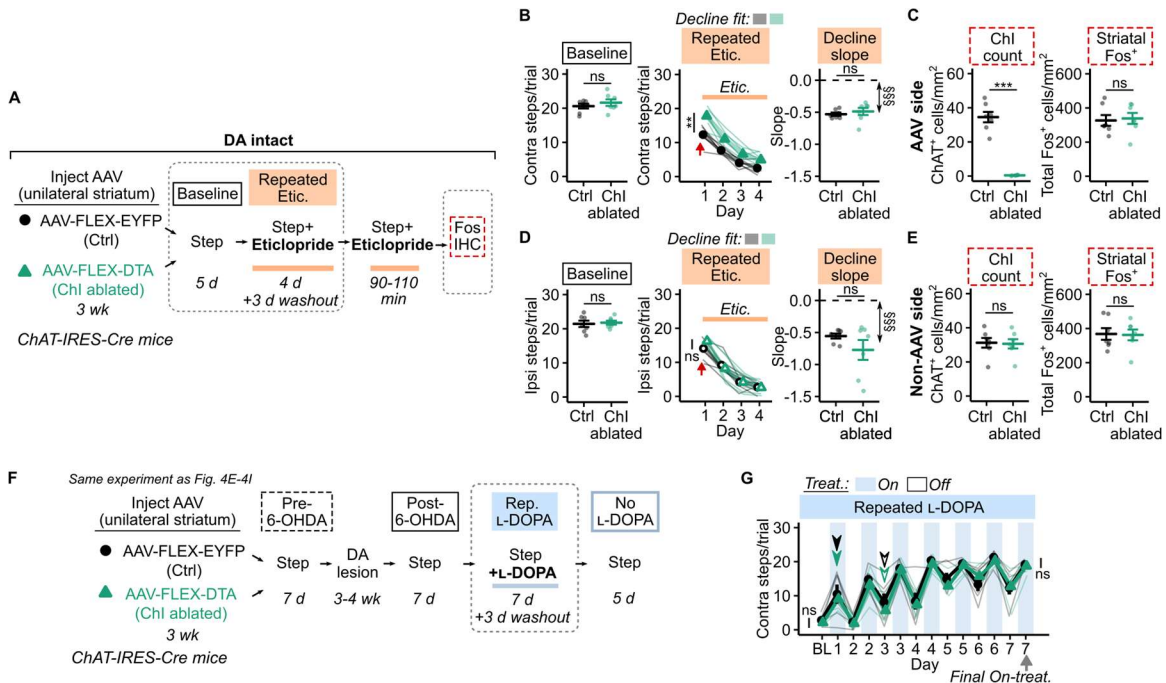
Fig. S7. Cholinergic interneuron ablation attenuates acute motor impairment from D2R blockade in DA-intact mice, but not L-DOPA-induced motoric rescue in DA-depleted mice (related to Fig. 4).
(**A-E**) Strategy to examine whether ChI ablation occlude parkinsonian decline induced by D2R antagonist (eticlopride) (A), the associated contralateral stepping performance (B) and IHC quantification for the AAV-injected side (C), and the associated ipsilateral stepping performance (D) and IHC quantification for the non-AAV-injected side (E), N=7, 7. Stepping declines are fitted with exponential decay, negative slope indicates decline (Methods). §§§P<0.001 vs. zero slope (no decline), Holm-Bonferroni-corrected one-sample t-tests. ***P<0.001, independent-samples t-tests. (**F-G**) Strategy to examine whether ChI ablation occlude L-DOPA-induced motoric rescue (F), and acute vs. long-term rescue during repeated L-DOPA phase (G), from the same experiment as Fig. 4E-4I. BL: parkinsonian baseline final session. Filled arrow: first day of detectable acute rescue vs. BL. Open arrow: first day of detectable long-term rescue vs. BL (Holm-Bonferroni-corrected paired t-tests, Methods). N=6, 7. Data are mean ± s.e.m.. See Dataset S1 for full statistics.
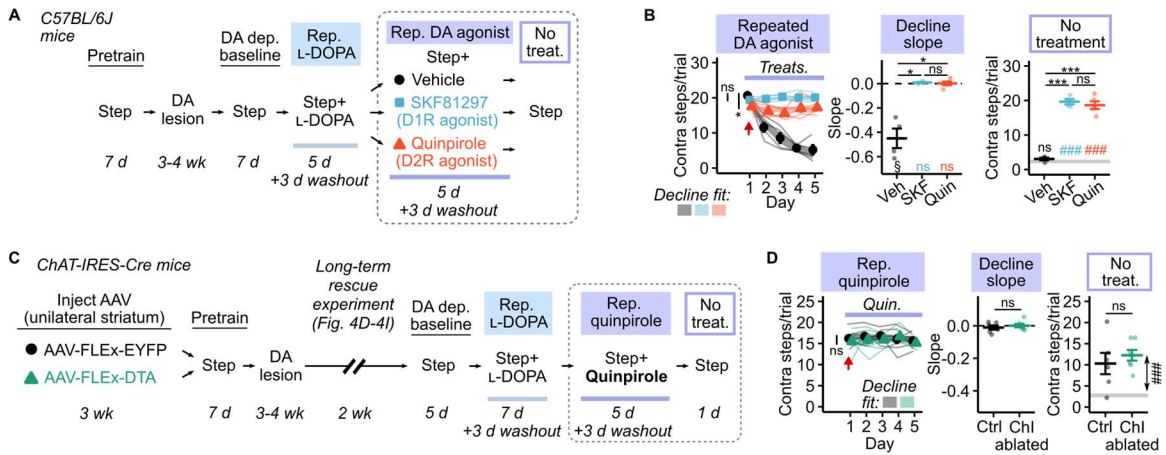
Fig. S8. D1R and D2R agonists are each sufficient to prevent parkinsonian decline of long-term rescue, and D2R agonist's effect did not require cholinergic interneurons (related to Fig. 5).

(**A-B**) Strategy to examine whether DA receptor agonist prevented parkinsonian decline of L-DOPA-induced long-term rescue (A), and the associated contralateral stepping performance (B). ###$P<0.001$ vs. same group's DA-depleted baseline, Holm-Bonferroni-corrected paired t-tests. $N=4, 4, 5$. (**C-D**) Strategy to examine whether ChI ablation attenuated quinpirole's (D2R agonist) ability to prevent parkinsonian decline (C), and the associated contralateral stepping performance (D). $N=6, 7$. Gray horizontal band in "No treatment" panels in (B) and (D): dopamine-depleted baseline stepping (groups pooled). §$P<0.05$ vs. zero slope (no decline), one-sample t-tests. ###$P<0.001$, vs. DA-depleted baseline (groups pooled), paired t-test. Data are mean ± s.e.m.. *$P<0.05$, ***$P<0.001$. ANOVAs were used for between-group comparisons, followed by appropriate post-hoc tests when needed. See Dataset S1 for full statistics.
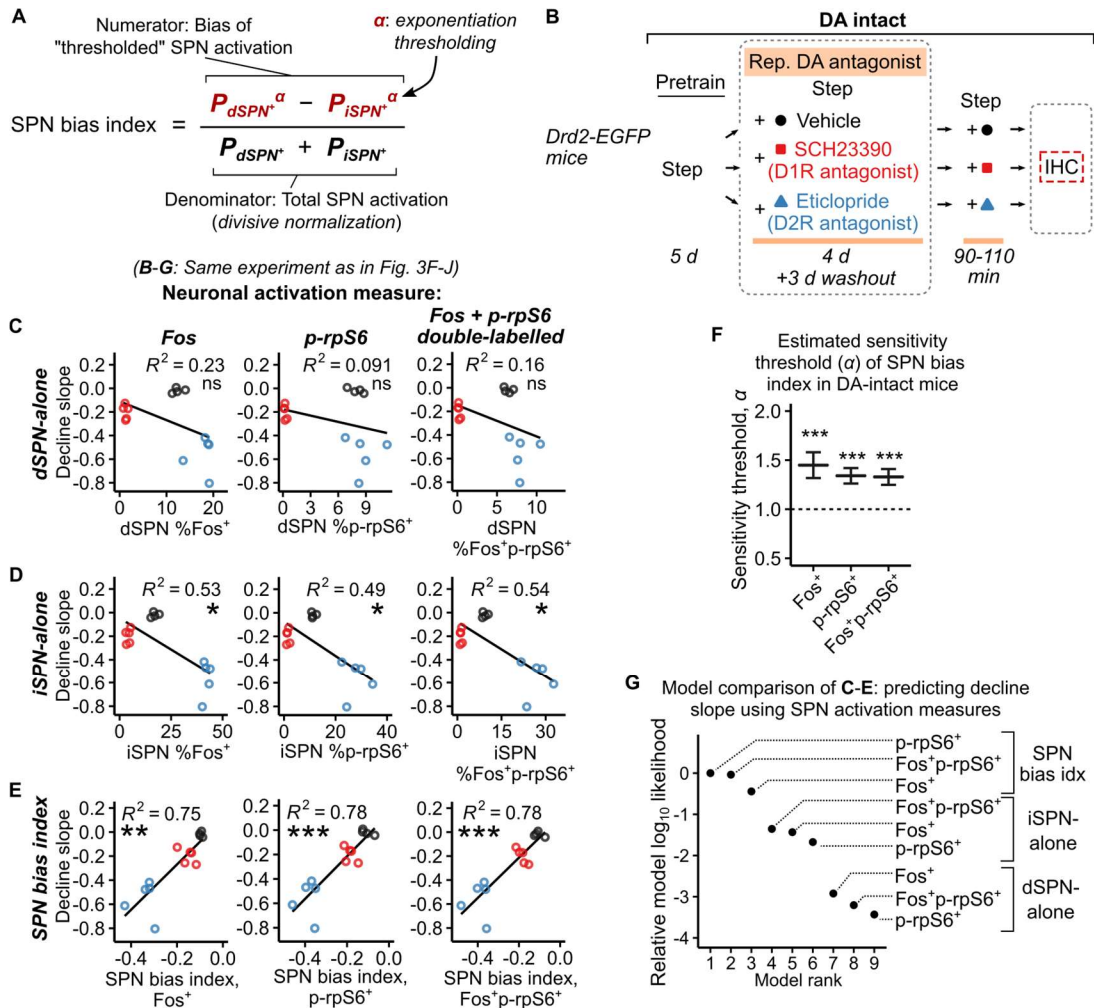
Fig. S9. Reduced dSPN:iSPN activation balance best predict parkinsonian decline in dopamine-intact mice treated with D1R vs. D2R antagonists (related to Fig. 3 and 5).
(**A**) A SPN bias index to summarize dSPN:iSPN activation balance. P_{dSPN+} vs. P_{iSPN+} denotes the proportion (i.e., between 0 and 1) of dSPNs vs. iSPNs that are positively labelled for a particular activation measure (Fos, p-rpS6$^{S235/236}$, or double-labelled). Two key components of the SPN bias index – exponentiation thresholding (α) and divisive normalization – are often used in neural computational modelling (17). SPN bias index ranges from negative (more iSPNs activated than dSPNs), to 0 (dSPNs and iSPNs equally activated), to positive (more dSPNs activated than iSPNs). The exponent α ($\alpha>0$) acts as a sensitivity threshold. When α is large (>1), sparse SPN activation are suppressed and filtered out. (**B**) Experimental design to use SPN activation to predict parkinsonian decline slope with linear regression, from the same experiment as Fig. 3F-J. N=4, 5, 5. (**C-E**) Linear models to predict parkinsonian decline slopes, using Fos (left panels), phosphorylated-rpS6$^{S235/236}$ (p-rpS6, middle panels), or Fos + p-rpS6 (right panels) as the neuronal activation measure, and using dSPN-alone (C), iSPN-alone (D), or SPN bias index (E) as the pathway predictor (Methods). Note that the decline slope (y-axis) of all panels in (C-E) is the same as the decline slope in Fig. 3G, and the dSPN-alone and iSPN-alone activation measures (x-axes) of (C-D) are from Fig. 3H-J. Solid lines: linear fits. Stars: regression slope significantly different from 0. (**F**) Sensitivity threshold α of SPN bias index, estimated (± s.e.) from each linear model in (E). Stars: difference from 1 (likelihood ratio tests). (**G**) Relative model likelihood of (C-E), calculated based on Akaike Information Criterion (Methods). Note log$_{10}$ scale of y-axis. Models that use SPN bias indices are 10-47 more likely than the best models that used dSPN- or iSPN-alone activation measures. *$P<0.05$, ***$P<0.001$. See Dataset S1 for full statistics.
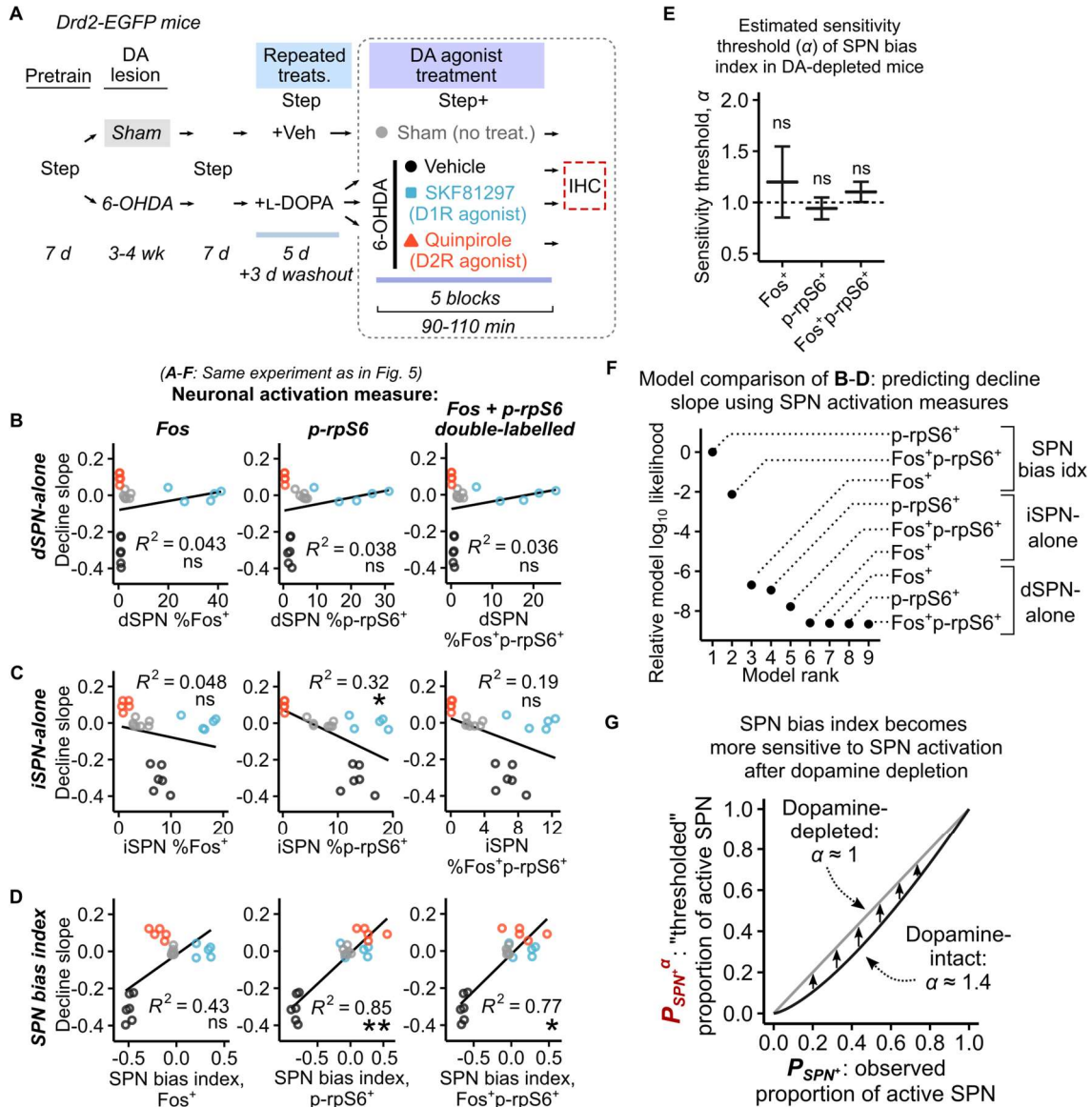
Fig. S10. Restoring balanced dSPN:iSPN activation best predict prevention of parkinsonian decline in dopamine-depleted mice by D1R vs. D2R agonists (related to Fig. 5).

(**A**) Experimental design to use SPN activation to predict parkinsonian decline slope with linear regression, from the same experiment as Fig. 5. The Sham group is taken from Figure 3A "Sham+Step" group. *N*=7 (Sham); 6, 5, 5 (6-OHDA). (**B-D**) Linear models to predict parkinsonian decline slopes, using Fos (left panels), phosphorylated-rpS6$^{S235/236}$ (p-rpS6, middle panels), or Fos + p-rpS6 (right panels) as the neuronal activation measure, and using dSPN-alone (B), iSPN-alone (C), or SPN bias index (D) as the pathway predictor (Methods). Note that the decline slope (y-axis) of all panels in (B-D) is the same as the decline slope in Fig. 5B, and the dSPN-alone and iSPN-alone activation measures (x-axes) of (B-C) are from Fig. 5D-F. Solid lines: linear fits. Stars: regression slope significantly different from 0. (**E**) Sensitivity threshold α of SPN bias index, estimated (± s.e.) from each linear model in (D). Stars: difference from 1 (likelihood ratio tests). (**F**) Relative model likelihood of (B-D), calculated based on Akaike Information Criterion (Methods). Note log$_{10}$ scale of y-axis. The top two models use SPN bias indices of p-rpS6$^{235/236+}$ and Fos$^+$p-rpS6$^{S235/236+}$ as the predictor, and they are >10^4 times more likely than the best models

25

that used dSPN- or iSPN-alone activation measures. (**G**) The effect of reducing the sensitivity threshold, α, from ≈1.4 in dopamine-intact mice (Fig. S9F) to ≈1 in dopamine-depleted mice (E). The larger α in dopamine-intact mice causes SPN activation (P_{SPN+}) to be filtered out by the thresholded activation ($P_{SPN+}{}^{\alpha}$) before it could modulate the SPN bias index (see Fig. S9A). By contrast, the smaller α in dopamine-depleted mice allows the thresholded activation, and thus the SPN bias index, to be more sensitive to SPN activation (P_{SPN+}).*$P<0.05$, **$P<0.01$, ***$P<0.001$. See Dataset S1 for full statistics.

Movie S1 (separate file). Step task: unilateral 6-OHDA led to contralateral deficit in stepping. Video shows Step task performance from the same mouse before 6-OHDA (dopamine intact) vs. post-6-OHDA after severe parkinsonian performance decline.

Movie S2 (separate file). Pole task: unilateral 6-OHDA led to contralateral deficit in turning during pole descent. Video shows Pole task performance from the same mouse before 6-OHDA (dopamine intact) vs. post-6-OHDA after severe parkinsonian performance decline.

Movie S3 (separate file). Step task: acute vs. long-term rescue from L-DOPA. Video shows post-6-OHDA Step task performance from the same mouse as in Supplemental Video 1, (i) after severe parkinsonian decline, (ii) while acutely on L-DOPA, and (iii) off treatment after a 3-day (72 hr) washout from repeated L-DOPA + Step Task training.

Movie S4 (separate file). Pole task: acute vs. long-term rescue from L-DOPA. Video shows post-6-OHDA Pole task performance from the same mouse as in Supplemental Video 2, (i) after severe parkinsonian decline, (ii) while acutely on L-DOPA, and (iii) off treatment after a 3-day (72 hr) washout from repeated L-DOPA + Pole Task training.

Dataset S1 (separate file). Supplementary tables of statistics.

SI References

1. Gong S, *et al.* (2003) A gene expression atlas of the central nervous system based on bacterial artificial chromosomes. *Nature* 425(6961):917-925.

2. Caine SB, *et al.* (2007) Lack of Self-Administration of Cocaine in Dopamine D1 Receptor Knock-Out Mice. *The Journal of Neuroscience* 27(48):13140-13150.

3. Beeler JA, Faust RP, Turkson S, Ye H, & Zhuang X (2016) Low Dopamine D2 Receptor Increases Vulnerability to Obesity Via Reduced Physical Activity, Not Increased Appetitive Motivation. *Biol. Psychiatry* 79(11):887-897.

4. Rossi J, *et al.* (2011) Melanocortin-4 Receptors Expressed by Cholinergic Neurons Regulate Energy Balance and Glucose Homeostasis. *Cell Metabolism* 13(2):195-204.

5. Gong S, *et al.* (2007) Targeting Cre Recombinase to Specific Neuron Populations with Bacterial Artificial Chromosome Constructs. *The Journal of Neuroscience* 27(37):9817-9823.

6. Olsson M, Nikkhah G, Bentlage C, & Bjorklund A (1995) Forelimb akinesia in the rat Parkinson model: differential effects of dopamine agonists and nigral transplants as assessed by a new stepping test. *J. Neurosci.* 15(5):3863-3875.

7. Ogawa N, Hirose Y, Ohara S, Ono T, & Watanabe Y (1985) A simple quantitative bradykinesia test in MPTP-treated mice. *Research communications in chemical pathology and pharmacology* 50(3):435-441.

8. Fox ME, *et al.* (2016) Cross-hemispheric dopamine projections have functional significance. *Proceedings of the National Academy of Sciences* 113(25):6985-6990.

9. Krashes MJ, *et al.* (2011) Rapid, reversible activation of AgRP neurons drives feeding behavior in mice. *The Journal of Clinical Investigation* 121(4):1424-1428.

10. Won L, Ding Y, Singh P, & Kang UJ (2014) Striatal Cholinergic Cell Ablation Attenuates L-DOPA Induced Dyskinesia in Parkinsonian Mice. *J. Neurosci.* 34(8):3090-3094.

11. Lindgren HS, Rylander D, Ohlin KE, Lundblad M, & Cenci MA (2007) The "motor complication syndrome" in rats with 6-OHDA lesions treated chronically with l-DOPA: Relation to dose and route of administration. *Behav. Brain Res.* 177(1):150-159.

12. Gomez JL*, et al.* (2017) Chemogenetics revealed: DREADD occupancy and activation via converted clozapine. *Science* 357(6350):503-507.

13. Arlotta P, Molyneaux BJ, Jabaudon D, Yoshida Y, & Macklis JD (2008) Ctip2 Controls the Differentiation of Medium Spiny Neurons and the Establishment of the Cellular Architecture of the Striatum. *The Journal of Neuroscience* 28(3):622-632.

14. Valjent E*, et al.* (2011) Haloperidol Regulates the State of Phosphorylation of Ribosomal Protein S6 via Activation of PKA and Phosphorylation of DARPP-32. *Neuropsychopharmacology* 36(12):2561-2570.

15. Biever A, Valjent E, & Puighermanal E (2015) Ribosomal Protein S6 Phosphorylation in the Nervous System: From Regulation to Function. *Frontiers in Molecular Neuroscience* 8(75).

16. Stringer C, Wang T, Michaelos M, & Pachitariu M (2021) Cellpose: a generalist algorithm for cellular segmentation. *Nature Methods* 18(1):100-106.

17. Carandini M & Heeger DJ (2012) Normalization as a canonical neural computation. *Nat. Rev. Neurosci.* 13(1):51-62.

18. Burnham KP & Anderson DR (2002) *Model selection and multimodel inference: A practical information-theoretic approach* (Springer, Secaucus, NJ, USA) 2nd Ed.