Sst+ GPi output neurons provide direct feedback to key nodes of the basal ganglia and drive behavioral flexibility

Weglage M.*[1,2], Ährlund-Richter S.*[1,3], Contestabile A.*[1], Skara V.*[1], Fuzik J.[1], Santos I.[1], Lazaridis I.[#1,4], Meletis K.[#1].

*equal contribution

[#]corresponding author

[1] Department of Neuroscience, Karolinska Institutet, Stockholm, Sweden.

[2] Current address; Department of Organismic and Evolutionary Biology and Center for Brain Science, Harvard University, Cambridge, MA, USA

[3] Current address; Picower Institute of Learning and Memory, Massachusetts Institute of Technology, Cambridge, MA, USA.

[4] Current address; McGovern Institute, Massachusetts Institute of Technology, Cambridge, MA, USA.

Abstract

The internal globus pallidus (GPi) is a basal ganglia output nucleus with separate projections to the thalamus and the lateral habenula (LHb). Here, we show a GPi subtype with projections to LHb (GPi-LHb), genetically defined based on glutamate/GABA co-transmission and somatostatin (Sst) expression, also projects back to key nodes in the basal ganglia: the external globus pallidus (GPe), the striatal striosomes, and dopamine neurons in the substantia nigra. We found that the Sst+ GPi population showed strong movement and direction-specific selectivity in a goal-directed choice task, but not during self-paced exploration, and did not signal prediction errors or outcome modulation. During goal-directed behavior, the Sst+ GPi activity slowly evolved with learning of correct choice actions and genetic silencing disrupted the ability to update choice behavior following a task rule reversal. In summary, we have found that Sst+ GPi neurons establish a wide feedback network in the basal ganglia and drive behavioral flexibility.

Introduction

The basal ganglia are important for the generation of voluntary movements, decision-making and motivated behaviors[1,2]. The selection of specific actions, and the process of updating the value of chosen actions and strategies, ultimately produces motivated behaviors that optimize gain from the environment. Changes in the environmental conditions, such as resource availability (e.g. food), require animals to dynamically update their actions to survive. From this perspective, the choice to continue performing the same rewarded action (exploitation), versus the exploration of alternative options, is a critical computation. The basal ganglia and the habenula form an evolutionary conserved circuit optimized to support dynamic decision-making and choice updating[3]. The inability to adapt and explore (i.e., behavioral inflexibility) is manifested as habitual or compulsive actions[4], a condition associated with psychiatric disorders in humans and considerable cost in terms of quality of life[5].

The internal segment of the globus pallidus (GPi) and the substantia nigra pars reticulata (SNr) are the output nuclei of the basal ganglia, which in the canonical basal ganglia model regulate behavior through projections to thalamus and other downstream motor centers[6,7]. The mouse GPi (also called entopeduncular nucleus) can be further subdivided into two main output pathways: i) the parvalbumin-expressing (Pv+) GPi neurons that project to thalamus and brainstem[8] and ii) the somatostatin-expressing (Sst+) GPi neurons that project to the lateral habenula (LHb)[9–12]. The Sst+ GPi-LHb neurons co-release GABA and glutamate in LHb[13–15] and have been suggested to carry information related to the outcomes of actions[16], although optogenetic activation of the Sst+ GPi-LHb neurons does not modulate action value[13]. Thus, the GPi-LHb pathway holds a unique neuroanatomical position to integrate motor and non-motor information from the basal ganglia to guide and update motivated behavior.

Here, we used genetic targeting of the GPi-LHb neurons, based on their unique molecular profile: the co-expression of the transporters Vglut2 and Vgat (i.e., based on their glutamate/GABA co-transmission profile), or expression of the defining cellular marker Sst[13,15]. We found that visualization of the brain-wide projections of GPi-LHb neurons revealed an unexpected organization: in addition to the well-known axonal projection to LHb, GPi neurons also send projections to the external segment of the globus pallidus (GPe), the striatal striosomes, and the substantia nigra pars compacta (SNc) - three key nodes of the basal ganglia. This organization suggested that GPi-LHb neurons could shape motivated behaviors by calibrating the activity of functional nodes in the basal ganglia circuitry. We therefore recorded the activity of Sst+ GPi-LHb neurons in a serial reversal task using fiber photometry to define their role during motivated behaviors. These recordings showed no correlation of GPi-LHb pathway activity and

reward prediction errors - instead, the Sst+ GPi-LHb activity was strongly modulated during motor-related events, showing activation specifically during vigorous movements and ipsiversive turning. Functional perturbation experiments, using cell-type specific tetanus toxin-mediated silencing demonstrated that Sst+ GPi-LHb neurons are critical for behavioral adaptation following reversal after overtraining in a single reversal maze-based choice task. In summary, we propose that a neuromodulatory feedback signal from Sst+ GPi neurons onto key nodes of the basal ganglia is central to flexible behavior by shifting the network away from well-established and reinforced choices to drive exploration of alternative choices.

Results

GPi-LHb neurons have feedback connections to key nodes of the basal ganglia

GPi output neurons that project to the LHb (GPi-LHb neurons) can be defined by their GABA and glutamate co-transmission profile and their specific expression of Sst[13,15]. We mapped the brain-wide projection patterns of GPi-LHb neurons, employing three complementary cell-type specific adeno-associated virus (AAV) genetic labeling strategies. First, we used Vglut2-Cre/Vgat-Flp transgenic mice to selectively target GPi-LHb neurons that co-transmit GABA and glutamate with injection into the LHb of a retrogradely transported AAV (AAVrg) that induced Cre- and Flp-dependent expression of ChR2-EYFP (AAVrg-Con/Fon-ChR2-EYFP; experimental strategy referred to as Vglut2/Vgat$^{GPi\text{-}LHb}$, N = 7 mice, **Figure 1A**). As expected, this strategy resulted in genetic labeling of GPi neurons with a dense projection to the LHb (**Figure 1B**). In addition to labeling the known GPi-LHb pathway, we found extensive and dense collateral projections from the labeled Vglut2/Vgat$^{GPi\text{-}LHb}$ neurons to other subcortical and cortical brain regions, most prominently targeting the GPe, the caudate putamen (CPu), the SNc, and the prefrontal cortex (PFC; **Figure 1C**, **Supplementary Figure 1A**, **Supplementary Figure 2A-J**). We confirmed that the Vglut2/Vgat$^{GPi\text{-}LHb}$ axon terminals targeted the GPe using the parvalbumin (Pv) marker expression (**Figure 1D**). The Vglut2/Vgat$^{GPi\text{-}LHb}$ collaterals extended further into the CPu, displaying distinct patterns of concentrated axonal arborizations primarily in the dorsal CPu. Due to the discretely organized innervation pattern, we investigated whether the labeled axons preferentially targeted the mu-opioid receptor-expressing (i.e., Oprm1, MOR) striosomes in CPu[17]. Indeed, we found that the axonal arborization of GPi-LHb neurons was limited to MOR+ dense striosomes (**Figure 1E**, **Supplementary Figure 2I-J**). The Vglut2/Vgat$^{GPi\text{-}LHb}$ projections were also found in posterior subcortical regions, with the densest projections in the SNc, confirmed by tyrosine hydroxylase (TH) immunolabelling (**Figure 1F**). In addition, we observed axonal labeling in the substantia innominata (SI), but further analysis suggested that this axonal labeling could result from a different population of Vglut2/Vgat-expressing SI neurons (**Supplementary Figure 1B**). We found that the Vglut2/Vgat$^{GPi\text{-}LHb}$ strategy also labeled some cell bodies in the ventral tegmental area (VTA) and dorsal raphe nucleus (DRN, **Supplementary Figure 1C**) consistent with prior findings[14,18,19].

To validate these findings, we next used a second genetic targeting strategy based on the specific expression of Sst in GPi-LHb neurons[13,15]. In this strategy, we used Sst-Cre mice and injected retrograde AAVs into the LHb to achieve Cre-dependent labeling of GPi-LHb neurons based on their Sst expression (experimental strategy referred to as Sst$^{GPi\text{-}LHb}$, N = 3 mice; **Figure**

105 **1G-I and Supplementary Figure 2K-T**). We found that this strategy confirmed our findings on
106 the extensive collateralized network of GPi-LHb neurons (i.e. their projections to GPe, CPu
107 striosomes, SNc), and also eliminated the unspecific labeling of sparse neurons in SI, VTA, or DR
108 compared to the Vglut2/Vgat$^{GPi\text{-}LHb}$ neuron targeting strategy. We further quantified the GPi-LHb
109 axonal intensity in the targeted regions using the Sst$^{GPi\text{-}LHb}$ labelling strategy (N = 3), which
110 showed the highest projection density in the LHb, followed by projections density in the GPe, and
111 with lower projection densities in the SNc and in striosomes (**Figure 1J-M** and **Supplementary**
112 **Figure 3A-D**).
113 To further confirm our findings on the feedback organization of GPi-LHb projections, we
114 established a third genetic labeling strategy: we injected a Cre- and Flp-dependent viral construct
115 (AAV8-Con/Fon-ChR2-EYFP) directly into the GPi of Vglut2-Cre/Vgat-Flp mice (experimental
116 strategy referred to as Vglut2/Vgat^{GPi}; N = 6 mice; **Supplementary Figure 2U**). This direct
117 injection approach again labeled projections to the GPe and striosomes (**Supplementary Figure**
118 **2V-AD**). However, we observed some additional axonal labeling in the subthalamic nucleus (STN)
119 and substantia nigra pars reticulata (SNr), which was inconsistent with the labeling profile in the
120 other two strategies (**Supplementary Figure 4A-C**). To determine this projections profile in SNr
121 was the result labeling of Vglut2/Vgat-expressing neurons in the STN, we injected a retrograde
122 Cre- and Flp-dependent viral construct (AAVrg-Con/Fon-ChR2-EYFP) in the SNc of Vglut2-
123 Cre/Vgat-Flp mice. We found that this strategy resulted in the retrograde labeling of cells in the
124 STN, confirming that Vglut2/Vgat+ STN neurons form a direct STN-SNr projection, which is
125 distinct from the GPi-LHb pathway (**Supplementary Figure 4D-F**).
126 To investigate whether individual GPi-LHb neurons collateralized to multiple basal ganglia
127 target regions, we labeled GPi-LHb neurons to determine their collateral projections to two
128 targets: the SNc and the GPe. In this strategy, we injected mice with three AAVs: a retrograde
129 Cre-expressing virus (AAVrg-Cre) in the LHb, a retrograde Cre-dependent EGFP-expressing
130 virus (AAVrg-DIO-EGFP) in the SNc, and a retrograde Cre-dependent mCherry-expressing virus
131 (AAVrg-DIO-mCherry) in the GPe (N=3; **Supplementary Figure 5A**). Using this triple-virus
132 retrograde labeling strategy, we found single Sst+ GPi-LHb neurons labeled with both EGFP (i.e.,
133 projecting to SNc) and mCherry (i.e., projecting to GPe), as well as their EGFP+/mCherry+ axons
134 in LHb (**Supplementary Figure 5A-B**). To further confirm the collateralization of single GPi-LHb
135 neurons, we established another intersectional viral strategy in Sst-Cre mice (N = 3) using
136 injections of three different AAVs: a retrograde AAV with Cre-dependent expression of tetanus
137 toxin (AAVrg-DIO-TetLCh-dTomato) in the LHb, a retrograde AAV with Cre-dependent expression
138 of the Flp recombinase (AAVrg-DIO-FLPO) in the GPe, and a retrograde AAV with Cre- and Flp-

139 dependent EGFP expression (AAVrg-CreOn-FlpOn-EGFP) in the SNc (**Supplementary Figure**
140 **5C**). This strategy resulted in expression of EGFP and dTomato in single GPi neurons and labeled
141 axons in both the LHb and GPe, supporting the extensive collaterals of GPi-LHb neurons to the
142 GPe and the SNc (**Supplementary Figure 5D-E**). In conclusion, our results show that GPi-LHb
143 neurons form an extensive and complex feedback network in the basal ganglia, with collateral
144 projections that target key circuit nodes: the GPe, the CPu striosomes, and the SNc.
145
146 **GPi-LHb neurons synapse and co-transmit glutamate and GABA in basal ganglia targets**
147 We further examined the postsynaptic targets of GPi-LHb neurons performing a series of
148 monosynaptic retrograde tracing experiments using modified rabies virus and patch-clamp
149 recordings. To assess the postsynaptic specificity of GPi-LHb neurons across distinct striatal
150 populations, investigate whether GPi-LHb neurons preferentially form synapses on striatal
151 projection neurons (SPNs) in striosomes (i.e. Oprm1+) versus D2-expressing SPNs (iSPNs), we
152 injected a helper virus (AAV-DIO-TVA-V5-RbG), followed by a modified rabies virus expressing
153 EGFP into the CPu of Oprm1-Cre (N = 6 mice) or A2A-Cre mice (N = 4 mice; **Figure 2A**). The
154 Oprm1-Cre line labels striosomal SPNs, whereas the A2A-Cre line labels iSPNs distributed
155 across both striosome and matrix compartments, allowing us to determine whether GPi-LHb
156 neurons preferentially target striosomal versus D2-expressing SPNs independent of
157 compartmental localization. We found that Sst+ GPi neurons synapse directly and predominantly
158 onto striosomal Oprm1+ SPNs, but not onto A2A-expressing SPNs, highlighting a clear cell-type
159 and compartment-specific preference (**Figure 2A-C**, **Supplementary Figure 6A-B**, and
160 **Supplementary Table 1**). These results are consistent with the projection pattern observed in
161 striosomes and confirm that GPi-LHb neurons preferentially target D2-negative (putative D1)
162 SPNs. Similarly, we investigated whether GPi-LHb neurons directly target dopamine neurons in
163 the SNc, using the modified rabies virus monosynaptic tracing strategy in DAT-Cre mice (N = 3
164 mice; **Figure 2D**). We found that dopamine neurons (i.e. DAT+) in SNc received monosynaptic
165 inputs from Sst+ GPi neurons thereby confirming a direct synaptic connection (**Figure 2D-F**, and
166 **Supplementary Table 1**).
167 GPi-LHb neurons have been shown to co-transmit GABA and glutamate in the LHb[14]. To
168 establish the GABA/glutamate synaptic co-transmission profile of GPi-LHb neurons in other target
169 regions, we investigated the synaptic transmission of Vglut2/Vgat$^{GPi\text{-}LHb}$ axons terminals in the
170 GPe. We used optogenetic activation of Vglut2/Vgat^{GPi} axon terminals with electrophysiology
171 recordings from GPe neurons. A Cre- and Flp-dependent viral construct expressing ChR2 (AAV8-
172 Con/Fon-ChR2-EYFP) was injected directly into the GPi of Vglut2-Cre/Vgat-Flp mice, and GPe

173 neurons were recorded using cell-attached and whole-cell patch clamp recordings (**Figure 2G**).
174 The electrophysiological response to optogenetic Vglut2/Vgat^{GPi} axon activation, location, and
175 molecular profile of each patched GPe neuron was characterized (**Figure 2H-J)**. In cell-attached
176 recordings, the optogenetic activation of Vglut2/Vgat^{GPi} axons (3 ms, blue light) resulted in an
177 action potential doublet followed by firing pause, which interrupted fast tonic firing of GPe neurons
178 (**Figure 2H**). This indicates the presence of both excitation and inhibition in an intact GPe neuron.
179 The whole-cell recorded postsynaptic response of GPe neurons to optogenetic activation was
180 monosynaptic (i.e., in TTX, 4-AP) with a delay of 2-3 ms. The response exhibited both excitatory
181 and inhibitory components, and repeated optogenetic stimulation at 10 Hz induced short-term
182 synaptic depression in both (**Figure 2K-L**). Notably, the strong correlation in the jitter of GABA
183 and AMPA currents suggests a high likelihood of co-release, indicating that excitatory and
184 inhibitory neurotransmission originate from the same presynaptic terminals (**Supplementary**
185 **Figure 6C**). Sequential pharmacological dissection (Gabazine 10uM, NBQX 20uM, AP-V 50uM)
186 confirmed the synaptic co-transmission of glutamate and GABA from Vglut2/Vgat^{GPi} terminals in
187 all GPe neurons with optogenetically evoked postsynaptic response (N = 9 neurons; **Figure 2M-**
188 **N**). We further probed the subtype identity of the responsive GPe neurons, and we found that all
189 neurons were Pv+, lacked Foxp2 expression (**Figure 2J** and **Supplementary Figure 6D**), and
190 displayed intrinsic electrophysiological properties characteristic of prototypic PV+ GPe neurons[20]
191 (**Supplementary Figures 6E-F**). These recordings establish the synaptic connection and co-
192 transmission of GABA and glutamate between Vglut2+/Vgat+ GPi neurons and GPe neurons.
193 To investigate the glutamate/GABA profile of Vglut2+/Vgat+ GPi neurons projections to
194 other basal ganglia targets, we examined the expression of the vesicular transporters for
195 glutamate (i.e., Vglut2) and GABA (i.e., Vgat) in the synaptic boutons of genetically labeled
196 Vglut2/Vgat$^{GPi\text{-}LHb}$ axons. To visualize the boutons of Vglut2/Vgat^{GPi} axons, we injected a Cre- and
197 Flp-dependent AAV virus (AAV8-Con/Fon-GCaMP6m) into the GPi of Vglut2-Cre/Vgat-Flp mice
198 and mapped the localization of Vglut2 and Vgat using immunostaining (**Supplementary Figure**
199 **6G-H**). In all target regions, most GFP-labeled boutons from Vglut2/Vgat^{GPi} axons contained both
200 Vglut2 and Vgat (co-labeled boutons/total boutons: LHb 433/465; GPe 274/292; CPu 67/75; PFC
201 31/34; **Supplementary Figure 6I**), supporting that Vglut2/Vgat^{GPi} axons can co-release glutamate
202 and GABA in all target regions. In addition, to estimate the physiological influence of GPi-LHb
203 transmission we quantified the density of Vglut2 and Vgat-expressing boutons in each target
204 region. The most boutons per μm^2 were found in the LHb, followed by the GPe, CP and lastly the
205 PFC (**Supplementary Figure 6I**). These results further demonstrate that GPi-LHb neurons have

206 the potential to modulate the activity of prominent basal ganglia regions by local co-transmission
207 of glutamate and GABA.
208
209 **Sst+ GPi-LHb activity is not essential for movement execution**
210 Based on the projection profile of GPi-LHb neurons targeting key basal ganglia nodes, we
211 first aimed to define their role in spontaneous motor behaviors. We therefore recorded the activity
212 of Sst+ GPi-LHb neurons during exploration in an open field arena using fiber photometry and the
213 genetically encoded calcium indicator GCaMP7s[21] (**Figures 3A-B**). Sst+ GPi-LHb neurons were
214 labeled by injecting a retrograde AAV driving Cre-dependent expression of jGCaMP7s (AAVrg-
215 DIO-jGCaMP7s) into the LHb of Sst-Cre mice (N = 5, **Figure 3A** and **Supplementary Figure 7A**),
216 as this cell-type specific strategy resulted in the most precise targeting of the GPi-LHb population
217 (**Supplementary Figure 1-4**). The classically motor-associated GPi is composed of Pv+ neurons
218 that project to the thalamus and brainstem[15], and we therefore used the activity of Pv+ GPi
219 neurons for comparison in our assessment of movement-related activity of Sst+ GPi-LHb
220 neurons. Pv+ GPi neurons were targeted using cell type-specific labeling following injections of
221 an AAV with Cre-dependent jGCaMP7s expression (AAV-DIO-jGCaMP7s) into the GPi of Pv-Cre
222 mice (N = 3, **Figure 3A** and **Supplementary Figure 7B**). Mouse behavior was recorded with a
223 bottom-up camera and pose estimation of three body parts (snout, gravity center, and tail base)
224 was performed using DeepLabCut[22] (DLC) to identify basic locomotor-related activity in the two
225 discrete neuronal populations (**Figure 3B**).
226 We first observed that Sst+ GPi population activity increased monotonically with
227 movement velocity, similar to the Pv+ GPi population activity (**Figure 3C** and **F**). However, while
228 the Pv+ GPi population activity showed a strong and consistent preference for contraversive over
229 ipsiversive turns (relative to the recorded hemisphere), no clear preference could be established
230 for the Sst+ GPi population (**Figure 3C-H**). The motor-related signals of Sst+ GPi were therefore
231 discretely different from those of the canonical motor GPi output, suggesting that the two GPi
232 outputs may shape actions and motor programs through distinct mechanisms.
233 We therefore decided to investigate the activity of Sst+ GPi-LHb neurons in more detail
234 during specific actions and motor motifs. We trained a second DLC model with six markers (four
235 paws, snout, and tail base) and the behavioral repertoire was categorized into discrete motifs
236 through unsupervised learning using a variational autoencoding motif extractor[23] (VAME, **Figure**
237 **3I**). This classification identified 45 unique motor motifs during open field exploration. Motifs
238 typically lasted less than a second and were ranked by their average speed (**Figure 3J-K**). When
239 we aligned the calcium signals from Sst+ GPi-LHb neurons with the onset of each motif, we found

240 that low-speed motifs showed a strong decrease in neuronal activity, whereas the high-speed
241 motifs showed an increase in activity, independently of the type of motor action performed
242 (**Figures 3L-M**). In summary, recording of Sst+ GPi neurons during self-paced exploration
243 revealed a preferential activation during vigorous and high-speed actions.
244 To directly test the causal role of Sst+ GPi-LHb neurons in the execution of high-speed
245 motor motifs, we used a neurotransmitter release inactivation strategy based on cell type-specific
246 viral expression of the tetanus toxin light chain[24] (TeTxLC). This method was chosen for its ability
247 to achieve robust and sustained suppression of synaptic transmission in targeted neuronal
248 populations, making it particularly suitable for long-term behavioral studies (see **Supplementary**
249 **information** for more details). Specifically, we injected a retrograde AAV with Cre-dependent
250 expression of TeTxLC (AAVrg-DIO-TeTxLC-dTomato) bilaterally into the LHb of Sst-Cre mice (N
251 = 16 mice, **Figure 3N** and **Supplementary Figure 7C**). This experimental group was compared
252 with a control group injected with a retrograde AAV with dTomato expression (N = 9, AAVrg-DIO-
253 dTomato). Unexpectedly, we found that genetic inactivation of the Sst+ GPi-LHb neurons resulted
254 only in a small increase in the linear and angular speed as well as a small reduction in immobility
255 (**Figure 4O**). We then analyzed the behavior based on the unsupervised classification and found
256 a significant reduction in the usage of low-speed motifs following the inactivation of Sst+ GPi-LHb
257 neurons (**Figure 4P**). Overall, this inactivation experiment suggests that although Sst+ GPi-LHb
258 neurons are specifically activated during high-speed motor motifs, their activity is not necessary
259 for initiating movements and does not directly control the execution of vigorous motor actions.
260
261 **Sst+ GPi-LHb activity does not code prediction errors in a serial reversal task**
262 Since we found that the activity of Sst+ GPi-LHb neurons was positively correlated with
263 vigorous movements but not necessary for their execution, we aimed to further investigate their
264 involvement in action selection and reinforcement learning. Specifically, we sought to investigate
265 whether the Sst+ GPi-LHb population conveys an “anti-reward” outcome evaluation signal—such
266 as negative valence or inverse reward prediction error—as previously proposed[12,16]. To directly
267 test this hypothesis, we recorded the activity of Sst+ GPi-LHb and Pv+ GPi neurons during a two-
268 choice probabilistic serial reversal task using fiber photometry (**Figure 4A**).
269 In the probabilistic reversal task, water restricted mice faced a choice between two nose-
270 poke ports which flanked a central initiation port (**Figure 4B**). On any given trial, only one of the
271 choice ports yielded a water reward with 75% probability. The location of the reward reversed
272 unpredictably every 7-23 rewarded trials, forcing mice to evaluate trial outcomes in the light of
273 recent history in an ongoing fashion to infer the correct choice[13,16,25,26]. The mice performed this

task successfully, allowing us to fit a Q-learning model that accurately captured mouse behavior and generated trial-by-trial value (Q) as well as value update (ΔQ, i.e., prediction error) estimates (**Figure 4C-D**).

To test whether Sst+ GPi-LHb neurons encode negative valence or inverse ΔQ, we split both positive valence (rewarded) and negative valence (non-rewarded) outcome trials into respective low and high ΔQ trials. In low ΔQ trials, the current outcome matched the outcome obtained on the previous trial, in high ΔQ trials it did not. Given the blocked task structure, the current outcome is the best predictor of future outcomes, and matching outcomes are therefore expected to result in minor value updates, whereas mismatched outcomes require reevaluation of the choice–value mapping. This pattern of reevaluation was reflected in the behavior-fit ΔQ estimates, showing that both GPi-recorded mouse groups updated their choice preferences accordingly (**Figure 4E** and **L**).

To define how neuronal activity relates to the inferred value updates, we aligned Sst+ GPi-LHb calcium signals to entry into the choice port (i.e., where the outcome was presented). We found that ipsiversive port entries ("ipsi port", i.e., ipsilateral relative to the recorded hemisphere) that were rewarded showed a decrease in population activity, while trials with non-rewarded ipsiversive choices showed no marked signal modulation (**Figure 4F-G**). Notably, responses were independent of ΔQ, as both low and high ΔQ trial outcomes showed signals of similar magnitude during the choice port entry (**Figure 4E-G**). Rewarded contraversive port entries ("contra port") were likewise followed by a signal decrease (albeit weaker than at the ipsiversive port), while non-rewarded contraversive entries were followed by a moderate increase in population activity (**Figure 4F-G**). Like ipsiversive port responses, contraversive port responses were also unaffected by ΔQ (**Figure 4E-G**). To thoroughly evaluate these observations statistically, we conducted a repeated-measures ANOVA examining the effects of the previous and current trial outcomes, and of port choice, on the activity changes (ΔF) recorded in the second following choice port entries (data in **Figure 4G**). We found a significant main effect of the current trial outcome ($F_{(1,4)} = 228.13$, $p = .0001$), indicating that Sst+ activity was indeed relatively down-modulated following reward versus non-reward outcomes. Additionally, there was a significant main effect of the port choice ($F_{(1,4)} = 35.19$, $p = 0.004$), indicating that Sst+ activity levels were relatively up-modulated in the second after mice entered the contra- versus the ipsiversive choice port. Finally, there was no significant main effect of the previous trial outcome ($F_{(1,4)} = 4.65$, $p = 0.1$), nor were there any significant interaction effects among the factors (all $p > 0.3$). The fact that the previous trial outcome did not significantly modulate Sst+ activity supports that Sst+ GPi activity is not shaped by expectation and therefore does not signal ΔQ.

We hypothesized that the Sst+ distinctive activity peak detected after non-rewarded contraversive choice trials – and indeed the main effects of current outcome and port choice generally – could be motor- rather than valence-related, given our prior finding of a strong correlation between neuronal activity and vigorous movements in the open field. A strong negative correlation between the activity changes captured after choice port entry and the duration of the nose poke (**Figure 4I**) further suggested a link between movement (or immobility) and activity in the reversal task**.** To directly test this, we analyzed single-trial, outcome-locked fluorescence traces independent of outcome valence. We found that Sst+ GPi-LHb population activity consistently increased whenever mice exited the contra port, irrespective of whether the trial outcome was positive or negative (**Figure 4H**). This supports that the Sst+ GPi activity is associated with the action sequence following the entry in the contra port instead of signaling the negative (i.e., unrewarded) condition. We therefore hypothesized that Sst+ GPi-LHb population activity increased specifically during ipsiversive turns. To test this, we examined how population activity correlated with task-engaged ipsiversive versus contraversive turning. Unexpectedly, we found that the Sst+ GPi-LHb population activity showed marked differences depending on turn direction. Specifically, the activity of Sst+ GPi neurons increased appreciably preceding ipsiversive turns and moderately decreased preceding contraversive turns (**Figure 4J**). This robust hemispheric lateralization of the turn signals was qualitatively different from our observations in the open field, where the Sst+ GPi population was not significantly activated during ipsiversive turns (**Figures 3C-E**).

A comparison of the population activity of Pv+ GPi and Sst+ GPi-LHb neurons in response to choice port entries and exits, revealed qualitatively similar activity following rewarded versus non-rewarded trials. Similar to the Sst+ GPi population, the Pv+ GPi population showed decreased activity following rewarded port entries, with little evidence supporting inverse ΔQ or prediction error coding (**Figure 4L-N**). Accordingly, ANOVA analysis of the Pv+ population activity revealed a significant main effect of the current ($F_{(1,2)} = 33.24$, $p = 0.03$), but not of the previous trial outcome ($F_{(1,2)} = 3.80$, $p = 0.19$), and no significant interactions between the factors (all $p > 0.3$; **Figure 4N**). However, unlike Sst+ GPi-LHb neurons, Pv+ GPi neurons showed moderate increase after non-rewarded entries (**Figure 4M**) or exits from the contra or ipsi port (**Figure 4O**), nor did the ANOVA analysis evidence a significant main effect of port choice ($F_{(1,2)} = 1.00$, $p = 0.4$). Correspondingly, Pv+ GPi activity decreased during nose poke-related immobility (**Figure 4P**) but lacked significant direction-selective modulations during turns in the serial reversal task (**Figure 4Q**)**.**

341 Our findings suggest that Sst+ GPi-LHb neurons are specifically activated during task-
342 relevant directional movements, without significant modulation by value and value-updating
343 signals. To directly test whether prediction error (ΔQ) or outcome valence signals can explain any
344 activity modulation in the Sst+ or Pv+ GPi population following choice port-entry, and which is not
345 instead predicted by immobility during nose poking (i.e. the fraction of the 1s averaging window
346 spent nose-poking), we applied a series of ridge regression models (**Figures 4K** and **R**). Across
347 both GPi pathways and port choices, "ΔQ", "outcome", and "% in port" single-variable models
348 performed roughly equivalently, with the "ΔQ" model tending to perform slightly worse. Critically,
349 the proportion of the variation explained (R^2) by the better-performing "outcome" and "% in port"
350 models did not differ significantly (all $p > 0.6$; multiple paired t-tests). Moreover, the "ΔQ",
351 "outcome" and "% in port" variables proved interchangeable: scrambling any single variable in
352 models which combined all three had minimal impact on the explanatory power of the combined
353 models (all $p > 0.07$; multiple 1-sample t-tests with Holm-Sidak correction). Supporting the
354 relevance of immobility-related signals in the GPi pathways, we found that the activity of both GPi
355 populations was inversely correlated with immobility during fear conditioning, and that the
356 aversive shock-predicting auditory cues did not modulate either pathway, also supporting the
357 absence of clear valence signals (**Supplementary Figure 8A**).
358 Overall, we conclude that the activity of the Sst+ GPi-LHb pathways in the serial reversal
359 task was modulated during specific motor actions rather than by the outcome value or prediction
360 errors, and that this activity was strongly lateralized specifically in goal-directed and task-engaged
361 directional movements.
362

363 **Adaptation of Sst+ GPi-LHb activity after reward location changes in a single-reversal**
364 **maze task**

365 Based on our findings, we hypothesized that Sst+ GPi-LHb neurons selectively engage to
366 promote movement or movement-related decisions in a context or task-dependent manner. We
367 therefore aimed to investigate how Sst+ GPi activity developed during the development of goal-
368 directed strategies. In the serial reversal task, mice could have established a switching strategy
369 prior to the start of neural recording, and the fast and constrained movements in the operant
370 chamber could result in overlap of the slower signals from the genetically encoded calcium
371 sensors. To facilitate the identification of behaviorally relevant calcium signals during relevant
372 actions and over learning, we decided to record Sst+ GPi-LHb and Pv+ GPi population activity in
373 a single-reversal task in a larger maze to extend the duration of movement and decision-making
374 signals (**Figure 5A-B**). In this task, water-restricted mice run through a center corridor and make

375 sharp turns into one of the two choice corridors. Mice received rewards when they chose the
376 correct corridor and when they initiated a new trial at the start location. In the first part of this task,
377 water reward was available in the choice corridor requiring contraversive turning relative to the
378 recorded hemisphere, followed by a switch of the reward location requiring ipsiversive turning on
379 the 15th session (**Figure 5C**).

380 We again investigated the GPi activity during rewarded and non-rewarded outcomes to
381 measure negative valence or inverse reward prediction error signaling (**Supplementary Figures**
382 **9A**). We found post-reversal Sst+ GPi-LHb activity incompatible with either evaluation signal. If
383 the Sst+ population transmitted inverse reward prediction error, activity should markedly decrease
384 in response to rewards and increase in response to non-reward outcomes following the reversal.
385 Instead, session-average outcome-aligned Sst+ activity remained stable throughout the
386 experiment following rewards and slightly decreased (rather than increased) following non-reward
387 outcomes post-reversal (**Supplementary Figure 9B-C**). Notably, this post-reversal decrease was
388 once again explained by the turning preference: in the maze, mice tended to make contraversive
389 turns after failing to obtain rewards in the contra choice corridor end zone, and contraversive
390 turning was associated with decreased Sst+ GPi-LHb activity as was observed in the serial
391 reversal task (**Supplementary Figures 9B-C**; also see **Figure 4F**, ipsi port). Critically, any
392 putative reward and non-reward modulation of the Sst+ GPi activity remained similar in magnitude
393 during the entire task, arguing against a reward prediction error signaling (**Supplementary**
394 **Figures 9B-C**). Analysis of the Pv+ GPi activity in reward and non-reward outcomes showed the
395 same pattern as in the serial reversal task (**Supplementary Figures 9D-E**). These findings further
396 support our conclusion that Sst+ GPi-LHb activity does not encode negative valence or prediction
397 error signals.

398 To test the hypothesis that Sst+ GPi activity encoded the motivational or goal-directed
399 dimension of specific motor actions, we compared the relationship of GPi subpopulation activities
400 with the kinematic profiles across motivational contexts. To do so, we categorized motor motifs in
401 the serial reversal task, and we again found that Sst+ GPi-LHb population activity showed a
402 preference for ipsiversive turns (**Figures 5D-E**). In comparison, Pv+ GPi neurons displayed
403 increased activity during contraversive turns, although the activity amplitude during turning was
404 less pronounced than in the open field (**Figures 5F-G**). Using unsupervised classification, we
405 segmented the behavior into discrete motifs using VAME. When we ranked motifs by their
406 velocity, we surprisingly observed that the activity of Sst+ GPi-LHb neurons showed no relation
407 to the velocity or vigor of action motifs, which was in stark contrast to our observation in the open

408 field (**Figures 5H-I**). These findings further support that the action-specific modulation of Sst+
409 GPi-LHb activity is defined by the motivational or task context.
410 We then investigated whether the observed action and sequence-specific modulation of
411 Sst+ GPi-LHb neurons changed following task reversal, focusing on the key choice period
412 captured by the motif sequences in the central arm that described straight running and
413 contraversive turning (i.e., motif 41 "straight run" followed by motif 5 "contraversive turn", **Figure**
414 **5H-J** and **Supplementary Figure 9F** and **10A**). When we aligned calcium signals to the onset of
415 motif 41 followed by motif 5, we found that Sst+ GPi-LHb neurons showed a significantly
416 increased activity after task reversal even if mice performed motor actions at comparable linear
417 and angular speeds during pre- and post-reversal (**Figure 5K**). Notably, the contraversive turn
418 was the rewarded correct choice before the task reversal, whereas it became the incorrect,
419 unrewarded action after the reversal. The observed increase in activity specifically during this
420 transition suggests that Sst+ GPi-LHb neurons contribute to action reversal or behavioral flexibility
421 by signaling when an action that was previously rewarded must be updated to fit the new task
422 contingencies.
423 To investigate whether Sst+ GPi-LHb neurons play a role in shifting a persistent choice to
424 an alternative option and to establish a potential causal link between neuronal activity and
425 behavior, we conducted cell-type specific silencing experiments in the reversal task. Specifically,
426 we inactivated Sst+ GPi neurons by bilaterally injecting a retrograde AAV with Cre-dependent
427 expression of the tetanus light chain (TeTxLC) into the LHb of Sst-Cre mice (N = 15, **Figure 5L**
428 and **Supplementary Figure 7C**). We compared this experimental group with a control cohort (N
429 = 16) injected with a AAV expressing dTomato. We used the single-reversal maze task to study
430 the effects of Sst+ GPi pathway silencing, and similarly to the motor-related changes in the open
431 field (**Figure 4O**) we found an increased motif speed and reduced immobility time (**Figure 5M**
432 and **Supplementary Figure 11B**). The cell-type specific silencing also increased the number of
433 trials performed by the Sst+ GPi TeTxLC mice, in particular during the first training days and
434 immediately after the reversal. During the pre-reversal phase, mice with TeTxLC silenced Sst+
435 GPi-LHb neurons reached accurate and asymptotic performance (**Figure 5N** and
436 **Supplementary Figures 11C-E**). However, in the post-reversal phase, we found that silencing
437 Sst+ GPi-LHb neurons resulted in an increased number of error trials (**Supplementary Figures**
438 **11A, E, D**). As a result, mice with silenced Sst+ GPi-LHb neurons showed a decline in task
439 efficiency during the first days after reversal indicating an impairment in the ability to adjust their
440 choice strategy and suppress the well-trained pre-reversal action response (**Figure 5N**).
441 Importantly, after reversal we observed that TeTxLC mice showed increased variability in

performance, and their impaired behavioral efficiency was strongly correlated with the number of Sst+ GPi-LHb neurons expressing TeTxLC (**Figure 5O-P**). These findings demonstrate that Sst+ GPi-LHb neurons play a central role in adjusting goal-directed behavior and driving choice flexibility, through suppressing well-trained actions and promoting the alternative choice after task reversal.

Flexibility-driven modulation of Sst+ GPi-LHb activity in a single-reversal maze task

As the most prominent activation of Sst+ GPi-LHb neurons occurred in the central arm and entering the choice junction, we further analyzed the neuronal activity during this behavioral trajectory (**Figure 6A**). Mice naturally showed a decline in motivation and task-engagement in the later phases of the session (**Figure 6A** and **Supplementary Figures 9G-I**), and we therefore divided trials into four quartiles (Q1-4) to assess how neuronal activity changed with the motivational state. We found that Sst+ GPi-LHb activity became suppressed as mice performed choice runs with contraversive turning early in the session (i.e. Q1, Q2), while activity increased during the same trajectory in the last trials as mice decreased their velocity (i.e. Q4; **Figure 6B**). In contrast, in sessions after reversal of the reward location (i.e. post-reversal sessions), the Sst+ GPi-LHb activity during the same non-rewarded trajectory increased even in early trials (**Figure 6B**). Similar to our findings in the serial reversal task, Sst+ GPi-LHb neurons were more active during choice runs with ipsiversive turning, and this ipsiversive turning related activity consistently increased as mice approached the turn across all session quartiles (**Figure 6B**).

We then explored how the suppressed Sst+ GPi activity during the choice phase was related to task performance over sessions. To examine this, we plotted neuronal activity during the turn against the fraction of completed trials in each session (i.e., trial number/session total). Notably, as mice learned the task and developed a preference for correct contraversive turning, the suppression of Sst+ GPi activity in early trials became more pronounced (**Figures 6C-D**). After task reversal, when mice began to redirect their turning toward the ipsiversive choice corridor (i.e. the new rewarded side), the contraversive turning instead showed positive transients, resulting in similar Sst+ GPi activity during both ipsi- and contraversive turning (**Figures 6C-D**). Remarkably, Sst+ GPi activity during turning to the start zone (i.e., during trial initiation runs) showed no such modulation, as activity during contraversive “return” turns remained moderately suppressed and distinguishable from ipsiversive turn-related signals (**Figures 6E**). This indicates that the loss of the contraversive choice-related suppression is choice-specific and likely attributable to reversal learning.

To further investigate the relationship between Sst+ GPi-LHb neuron activity and choice behavior, we binned contraversive turning trials by session and quartile and averaged the neuronal activity. When plotting these averages against the percentage of ipsiversive choices within the same bin, we observed that higher activity during contraversive trials correlated with greater ipsiversive choice probability (**Figure 6F**). This suggests that elevated activity in Sst+ GPi-LHb neurons during contraversive trials, and by extension a reduction in directional selectivity of the Sst+ neurons, reflect a state more permissive of ipsiversive turning. In other words, whereas low Sst+ activity preceding a turn reliably signals a contraversive turn, high activity may be followed by turns in either direction, likely depending on the level of suppression in the opposite hemisphere. Unlike Sst+ GPi-LHb neurons, the Pv+ GPi population maintained consistently high activity levels during both contraversive and ipsiversive runs, showing little sensitivity to turn direction, task-engagement, or reversal condition (**Figures 6G** and **Supplementary Figure 12A-C**). In summary, we found that the activity of Sst+ GPi neurons during the choice phase reflects an updating of decision signals related to the reward location, which supports their role in adjusting goal-directed behavior and choice flexibility.

Discussion

The basal ganglia play a central role in selecting and evaluating actions, functions that are mediated via parallel but distinct GPi output pathways[12,15,16]. In this framework, actions are selected by GPi neurons that project to motor thalamus and brainstem, and action outcomes are evaluated by GPi neurons that project to the LHb[12,16]. Here, we have expanded the anatomical and functional characterizations of the GPi subtype that is characterized by projections to the LHb (GPi-LHb subtype) to define their role in goal-directed behavior. We report that the main GPi-LHb pathway – genetically defined by Vgat/Vglut2/Sst co-expression (Sst+ GPi)[13,15] – possesses extensive collateral projections that target the GPe, the striatal striosomes, as well as dopamine neurons in the SNc. Our data further support that Sst+ GPi neurons co-release GABA and glutamate in all these target regions, similar to their co-release profile in the LHb[13,14,27]. In light of this unique neuroanatomical organization, which positions Sst+ GPi neurons to broadcast a wide modulatory feedback signal in the basal ganglia, we decided to investigate their role in shaping learning and execution of choice actions.

The GPi projections to LHb have been proposed to contribute to inverse reward prediction errors and reinforcement learning through a LHb activity-mediated suppression of dopamine release in the basal ganglia action selection circuit [12,16,28–30]. However, recent studies indicate that the negative valence and reinforcement learning signals found in the LHb originate – to a functionally significant degree – from excitatory inputs from the LHA[13,31] and other regions[18,32,33], rather than from the GPi. Importantly, our findings showing that Sst+ GPi neurons do not encode inverse reward prediction errors nor negative outcome value agrees with results of precise optogenetic activation of Sst+ GPi neurons and their LHb projections at the time of the trial outcome that fails to change decision-making[13].

As highlighted in prior studies, the LHb plays a pivotal role in processing aversive signals[34,35], primarily by increasing activity in response to negative outcomes, which suppresses midbrain dopamine release[28,36]. This function is critical for guiding adaptive behaviors, as LHb inactivation impairs outcome-based decision-making[37], including effort-based choices[38] and adaptive avoidance learning[39]. While our data suggests that Sst+ GPi neurons do not directly encode negative valence, their projections to the LHb may instead modulate the activity of LHb circuits and their specialized outputs involved in adaptive behaviors. This hypothesis aligns with the growing evidence of functional heterogeneity within the habenula and its complex connectivity[40]. We propose that distinct subpopulations of LHb neurons and their discretely organized inputs may mediate these dual or more complex roles of the LHb, with Sst+ GPi inputs

523 shaping the flexibility of behavioral outputs without directly engaging in aversive signal
524 processing.

525 A key finding of our study is the context-dependent activity of the Sst+ GPi pathway. We
526 show that Sst+ GPi activity reflected the turn direction in mice that performed a spatial reversal
527 task for water rewards, more so than when they spontaneously explored a reward-free
528 environment. Specifically, the Sst+ GPi population activity was strongly lateralized and correlated
529 with ipsiversive turning choices in the reversal tasks, and this activity profile was absent during
530 exploration of an open field without reinforcement. Similar examples of reward or task context-
531 dependent modulations of movement-related signals have previously been found in GPi
532 neurons[41–43]. In comparison, the Pv+ GPi population displayed the opposite pattern, as Pv+ GPi
533 neurons were strongly activated during contraversive turning in the open field but not in the choice
534 tasks. Our findings on the task-specific activity of Sst+ GPi neurons, showing their activation
535 during choice turning direction but not outcome-related prediction errors, supports their role in
536 progressively shaping action selection instead of immediate action evaluation. This distinction is
537 supported by our experiments showing that inactivation of Sst+ GPi neurons (i) increases overall
538 locomotor activity without impairing motor execution per se, yet (ii) degrades optimal performance
539 in the single reversal task. Notably, studies have shown that LHb lesions or inhibition produce
540 similar effects on locomotion and reversal performance[38,39,44].

541 The activity profile of Sst+ GPI neurons and the behavioral effects following their genetic
542 silencing together support an important role in the development of a choice-turning activity that
543 evolves with training during task learning. Specifically, we found that the Sst+ GPi activity during
544 contraversive choice turning gradually increased after reversal of the reward location, whereas
545 the activity recorded during ipsiversive choice turning remained stationary. As a result, the
546 lateralized turning-selective activity gradually decreased as mice learned to shift their turning
547 preference from contra- to ipsiversive actions. Interestingly, the striatum also shows population-
548 level activation during contraversive choices[45]. We propose that this gradual rebalancing of
549 contraversive and ipsiversive choice-related activity during adaptation to a new task rule reflects
550 the key role of the Sst+ GPi population in suppressing previously learned and preferred actions.
551 Here, we have shown how the organization of the Sst+ GPi pathway expands the complexity of
552 connectivity diagrams that depict the basal ganglia circuit as a set of one-directional parallel
553 pathways. Our results show that the genetically defined Sst+ GPi population forms an unexpected
554 feedback projection in the basal ganglia, and we propose that this organization can underline a
555 neuromodulatory signal that suppresses the commitment to repeat the same action, thereby
556 promoting the exploration of alternative actions in the environment.

Acknowledgements

We thank the Animal Behavior Core Facility (ABCF) at Karolinska Institutet for access to infrastructure and Alexander Wolthon for mouse colony management. Funding for this study was provided to K.M. by the Swedish Research Council (Vetenskapsrådet MH), the Swedish Brain Foundation (Hjärnfonden), Karolinska Institutet (KID doctoral funding for V.S. and M.W.), and the Knut och Alice Wallenberg Foundation. Additional funding was provided by the Swedish Society for Medical Research (Svenska Sällskapet för Medicinsk Forskning) and the Wenner Gren Foundation to S.Ä-R.; the Swiss National Science Foundation (project number: P500PB_214355) and the Amicitia Foundation to A.C.

Author Contributions

SÄ-R, VS and IL designed, interpreted and performed the anatomical characterization. SÄ-R and VS analyzed and visualized the anatomical data with the help of IS and AC. MW and IL designed the behavioral experiments. MW, IL, AC and VS performed the behavioral experiments with the help of IS. MW and AC analyzed and visualized all behavioral data. JF performed all electrophysiology experiments and analysis. MW, SÄ-R, VS, AC and K.M. wrote the manuscript. IL and KM supervised and conceptualized the work. KM was responsible for project funding and management. All of the authors discussed and commented on the manuscript.

Competing Interests statement

The authors declare no competing interests.

Data availability

All data presented in this paper is available upon request to the corresponding authors.

Code availability

Custom code is available upon request.

584 **References**

585 1. Cox, J., and Witten, I.B. (2019). Striatal circuits for reward learning and decision-making.
586 Nat Rev Neurosci *20*, 482–494. https://doi.org/10.1038/s41583-019-0189-2.
587 2. Graybiel, A.M., Aosaki, T., Flaherty, A.W., and Kimura, M. (1994). The basal ganglia and
588 adaptive motor control. Science (1979) *265*. https://doi.org/10.1126/science.8091209.
589 3. Grillner, S., and Robertson, B. (2016). The Basal Ganglia Over 500 Million Years. Preprint,
590 https://doi.org/10.1016/j.cub.2016.06.041 https://doi.org/10.1016/j.cub.2016.06.041.
591 4. Jog, M.S., Kubota, Y., Connolly, C.I., and Graybiel, A.M. (1999). Building neural
592 representations of habits. Science (1979) *286*.
593 https://doi.org/10.1126/science.286.5445.1745.
594 5. Robbins, T.W., Vaghi, M.M., and Banca, P. (2019). Obsessive-Compulsive Disorder:
595 Puzzles and Prospects. Neuron *102*, 27–47. https://doi.org/10.1016/j.neuron.2019.01.046.
596 6. Alexander, G.E., and Crutcher, M.D. (1990). Functional architecture of basal ganglia
597 circuits: neural substrates of parallel processing. Trends Neurosci *13*, 266–271.
598 https://doi.org/10.1016/0166-2236(90)90107-L.
599 7. Mink, J.W. (1996). The basal ganglia: Focused selection and inhibition of competing motor
600 programs. Prog Neurobiol *50*. https://doi.org/10.1016/S0301-0082(96)00042-1.
601 8. Saunders, A., Oldenburg, I.A., Berezovskii, V.K., Johnson, C.A., Kingery, N.D., Elliott, H.L.,
602 Xie, T., Gerfen, C.R., and Sabatini, B.L. (2015). A direct GABAergic output from the basal
603 ganglia to frontal cortex. Nature *521*, 85–89. https://doi.org/10.1038/nature14179.
604 9. Rajakumar, N., Elisevich, K., and Flumerfelt, B.A. (1994). Parvalbumin-containing
605 GABAergic neurons in the basal ganglia output system of the rat. Journal of Comparative
606 Neurology *350*. https://doi.org/10.1002/cne.903500214.
607 10. Van Der Kooy, D., and Carter, D.A. (1981). The organization of the efferent projections and
608 striatal afferents of the entopeduncular nucleus and adjacent areas in the rat. Brain Res
609 *211*. https://doi.org/10.1016/0006-8993(81)90064-0.
610 11. Vincent, S.R., and Brown, J.C. (1986). Somatostatin immunoreactivity in the
611 entopeduncular projection to the lateral habenula in the rat. Neurosci Lett *68*.
612 https://doi.org/10.1016/0304-3940(86)90134-5.
613 12. Hong, S., and Hikosaka, O. (2008). The Globus Pallidus Sends Reward-Related Signals
614 to the Lateral Habenula. Neuron *60*. https://doi.org/10.1016/j.neuron.2008.09.035.
615 13. Lazaridis, I., Tzortzi, O., Weglage, M., Märtin, A., Xuan, Y., Parent, M., Johansson, Y.,
616 Fuzik, J., Fürth, D., Fenno, L.E., et al. (2019). A hypothalamus-habenula circuit controls

617 aversion. Molecular Psychiatry 2019 24:9 *24*, 1351–1368. https://doi.org/10.1038/s41380-
618 019-0369-5.
619 14. Shabel, S.J., Proulx, C.D., Piriz, J., and Malinow, R. (2014). GABA/glutamate co-release
620 controls habenula output and is modified by antidepressant treatment. Science (1979) *345*.
621 https://doi.org/10.1126/science.1250469.
622 15. Wallace, M.L., Saunders, A., Huang, K.W., Philson, A.C., Goldman, M., Macosko, E.Z.,
623 McCarroll, S.A., and Sabatini, B.L. (2017). Genetically Distinct Parallel Pathways in the
624 Entopeduncular Nucleus for Limbic and Sensorimotor Output of the Basal Ganglia. Neuron
625 *94*. https://doi.org/10.1016/j.neuron.2017.03.017.
626 16. Stephenson-Jones, M., Yu, K., Ahrens, S., Tucciarone, J.M., Van Huijstee, A.N., Mejia,
627 L.A., Penzo, M.A., Tai, L.H., Wilbrecht, L., and Li, B. (2016). A basal ganglia circuit for
628 evaluating action outcomes. Nature *539*. https://doi.org/10.1038/nature19845.
629 17. Märtin, A., Calvigioni, D., Tzortzi, O., Fuzik, J., Wärnberg, E., and Meletis, K. (2019). A
630 Spatiomolecular Map of the Striatum. Cell Rep *29*, 4320-4333.e5.
631 https://doi.org/10.1016/J.CELREP.2019.11.096.
632 18. Root, D.H., Barker, D.J., Estrin, D.J., Miranda-Barrientos, J.A., Liu, B., Zhang, S., Wang,
633 H.L., Vautier, F., Ramakrishnan, C., Kim, Y.S., et al. (2020). Distinct Signaling by Ventral
634 Tegmental Area Glutamate, GABA, and Combinatorial Glutamate-GABA Neurons in
635 Motivated Behavior. Cell Rep *32*. https://doi.org/10.1016/j.celrep.2020.108094.
636 19. Root, D.H., Mejias-Aponte, C.A., Zhang, S., Wang, H.L., Hoffman, A.F., Lupica, C.R., and
637 Morales, M. (2014). Single rodent mesohabenular axons release glutamate and GABA.
638 Nat Neurosci *17*. https://doi.org/10.1038/nn.3823.
639 20. Magill, P.J., Bolam, J.P., and Bevan, M.D. (2001). Dopamine regulates the impact of the
640 cerebral cortex on the subthalamic nucleus-globus pallidus network. Neuroscience *106*.
641 https://doi.org/10.1016/S0306-4522(01)00281-0.
642 21. Dana, H., Sun, Y., Mohar, B., Hulse, B.K., Kerlin, A.M., Hasseman, J.P., Tsegaye, G.,
643 Tsang, A., Wong, A., Patel, R., et al. (2019). High-performance calcium sensors for imaging
644 activity in neuronal populations and microcompartments. Nat Methods *16*.
645 https://doi.org/10.1038/s41592-019-0435-6.
646 22. Mathis, A., Mamidanna, P., Cury, K.M., Abe, T., Murthy, V.N., Mathis, M.W., and Bethge,
647 M. (2018). DeepLabCut: markerless pose estimation of user-defined body parts with deep
648 learning. Nat Neurosci *21*. https://doi.org/10.1038/s41593-018-0209-y.

649 23. Luxem, K., Mocellin, P., Fuhrmann, F., Kürsch, J., Miller, S.R., Palop, J.J., Remy, S., and
650 Bauer, P. (2022). Identifying behavioral structure from deep variational embeddings of
651 animal motion. Commun Biol *5*. https://doi.org/10.1038/s42003-022-04080-7.
652 24. Pashkovski, S.L., Iurilli, G., Brann, D., Chicharro, D., Drummey, K., Franks, K., Panzeri, S.,
653 and Datta, S.R. (2020). Structure and flexibility in cortical representations of odour space.
654 Nature *583*. https://doi.org/10.1038/s41586-020-2451-1.
655 25. Weglage, M., Wärnberg, E., Lazaridis, I., Calvigioni, D., Tzortzi, O., and Meletis, K. (2021).
656 Complete representation of action space and value in all dorsal striatal pathways. Cell Rep
657 *36*. https://doi.org/10.1016/J.CELREP.2021.109437.
658 26. Tai, L.H., Lee, A.M., Benavidez, N., Bonci, A., and Wilbrecht, L. (2012). Transient
659 stimulation of distinct subpopulations of striatal neurons mimics changes in action value.
660 Nat Neurosci *15*. https://doi.org/10.1038/nn.3188.
661 27. Kim, S.A., Wallace, M.L., El-Rifai, M., Knudsen, A.R., and Sabatini, B.L. (2022). Co-
662 packaging of opposing neurotransmitters in individual synaptic vesicles in the central
663 nervous system. Neuron *110*, 1371-1384.e7.
664 https://doi.org/10.1016/J.NEURON.2022.01.007.
665 28. Matsumoto, M., and Hikosaka, O. (2007). Lateral habenula as a source of negative reward
666 signals in dopamine neurons. Nature *447*. https://doi.org/10.1038/nature05860.
667 29. Christoph, G.R., Leonzio, R.J., and Wilcox, K.S. (1986). Stimulation of the lateral habenula
668 inhibits dopamine-containing neurons in the substantia nigra and ventral tegmental area of
669 the rabbit. Journal of Neuroscience *6*. https://doi.org/10.1523/jneurosci.06-03-00613.1986.
670 30. Lammel, S., Lim, B.K., Ran, C., Huang, K.W., Betley, M.J., Tye, K.M., Deisseroth, K., and
671 Malenka, R.C. (2012). Input-specific control of reward and aversion in the ventral tegmental
672 area. Nature *491*, 212–217. https://doi.org/10.1038/nature11527.
673 31. Trusel, M., Nuno-Perez, A., Lecca, S., Harada, H., Lalive, A.L., Congiu, M., Takemoto, K.,
674 Takahashi, T., Ferraguti, F., and Mameli, M. (2019). Punishment-Predictive Cues Guide
675 Avoidance through Potentiation of Hypothalamus-to-Habenula Synapses. Neuron *102*.
676 https://doi.org/10.1016/j.neuron.2019.01.025.
677 32. Ottenheimer, D.J., Bari, B.A., Sutlief, E., Fraser, K.M., Kim, T.H., Richard, J.M., Cohen,
678 J.Y., and Janak, P.H. (2020). A quantitative reward prediction error signal in the ventral
679 pallidum. Nat Neurosci *23*. https://doi.org/10.1038/s41593-020-0688-5.
680 33. Tooley, J., Marconi, L., Alipio, J.B., Matikainen-Ankney, B., Georgiou, P., Kravitz, A. V.,
681 and Creed, M.C. (2018). Glutamatergic Ventral Pallidal Neurons Modulate Activity of the

682 Habenula–Tegmental Circuitry and Constrain Reward Seeking. Biol Psychiatry *83*.
683 https://doi.org/10.1016/j.biopsych.2018.01.003.
684 34. Matsumoto, M., and Hikosaka, O. (2009). Representation of negative motivational value in
685 the primate lateral habenula. Nat Neurosci *12*. https://doi.org/10.1038/nn.2233.
686 35. Wang, D., Li, Y., Feng, Q., Guo, Q., Zhou, J., and Luo, M. (2017). Learning shapes the
687 aversion and reward responses of lateral habenula neurons. Elife *6*.
688 https://doi.org/10.7554/elife.23045.
689 36. Tian, J., and Uchida, N. (2015). Habenula Lesions Reveal that Multiple Mechanisms
690 Underlie Dopamine Prediction Errors. Neuron *87*, 1304–1316.
691 https://doi.org/10.1016/J.NEURON.2015.08.028.
692 37. Nuno-Perez, A., Trusel, M., Lalive, A.L., Congiu, M., Gastaldo, D., Tchenio, A., Lecca, S.,
693 Soiza-Reilly, M., Bagni, C., and Mameli, M. (2021). Stress undermines reward-guided
694 cognitive performance through synaptic depression in the lateral habenula. Neuron *109*.
695 https://doi.org/10.1016/j.neuron.2021.01.008.
696 38. Stopper, C.M., and Floresco, S.B. (2013). What's better for me? Fundamental role for
697 lateral habenula in promoting subjective decision biases. Nature Neuroscience 2013 17:1
698 *17*, 33–35. https://doi.org/10.1038/nn.3587.
699 39. Baker, P.M., Oh, S.E., Kidder, K.S., and Mizumori, S.J.Y. (2015). Ongoing behavioral state
700 information signaled in the lateral habenula guides choice flexibility in freely moving rats.
701 Front Behav Neurosci *9*, 167383. https://doi.org/10.3389/FNBEH.2015.00295/PDF.
702 40. Michel, L., Molina, P., and Mameli, M. (2024). The behavioral relevance of a modular
703 organization in the lateral habenula. Neuron *112*, 2669–2685.
704 https://doi.org/10.1016/J.NEURON.2024.04.026.
705 41. Verma Rodriguez, A.K., Ramírez-Jarquin, J.O., Rossi-Pool, R., and Tecuapetla, F. (2025).
706 Basal ganglia output (entopeduncular nucleus) coding of contextual kinematics and reward
707 in the freely moving mouse. Elife *13*. https://doi.org/10.7554/ELIFE.98159.
708 42. Gdowski, M.J., Miller, L.E., Parrish, T., Nenonene, E.K., and Houk, J.C. (2001). Context
709 dependency in the globus pallidus internal segment during targeted arm movements. J
710 Neurophysiol *85*, 998–1004. https://doi.org/10.1152/JN.2001.85.2.998.
711 43. Turner, R.S., and Anderson, M.E. (2005). Context-dependent modulation of movement-
712 related discharge in the primate globus pallidus. J Neurosci *25*, 2965–2976.
713 https://doi.org/10.1523/JNEUROSCI.4036-04.2005.

44. Lee, E.H.Y., and Huang, S.L. (1988). Role of lateral habenula in the regulation of exploratory behavior and its relationship to stress in rats. Behavioural Brain Research *30*, 265–271. https://doi.org/10.1016/0166-4328(88)90169-6.

45. Lauwereyns, J., Watanabe, K., Coe, B., and Hikosaka, O. (2002). A neural correlate of response bias in monkey caudate nucleus. Nature *418*, 413–417. https://doi.org/10.1038/NATURE00892.

46. Ährlund-Richter, S., Xuan, Y., van Lunteren, J.A., Kim, H., Ortiz, C., Pollak Dorocic, I., Meletis, K., and Carlén, M. (2019). A whole-brain atlas of monosynaptic input targeting four different cell types in the medial prefrontal cortex of the mouse. Nat Neurosci *22*. https://doi.org/10.1038/s41593-019-0354-y.

47. Kim, C.K., Yang, S.J., Pichamoorthy, N., Young, N.P., Kauvar, I., Jennings, J.H., Lerner, T.N., Berndt, A., Lee, S.Y., Ramakrishnan, C., et al. (2016). Simultaneous fast measurement of circuit dynamics at multiple sites across the mammalian brain. Nat Methods *13*. https://doi.org/10.1038/nmeth.3770.

48. Nath, T., Mathis, A., Chen, A.C., Patel, A., Bethge, M., and Mathis, M.W. (2019). Using DeepLabCut for 3D markerless pose estimation across species and behaviors. Nat Protoc *14*. https://doi.org/10.1038/s41596-019-0176-0.

49. Lopes, G., Bonacchi, N., Frazão, J., Neto, J.P., Atallah, B. V., Soares, S., Moreira, L., Matias, S., Itskov, P.M., Correia, P.A., et al. (2015). Bonsai: An event-based framework for processing and controlling data streams. Front Neuroinform *9*. https://doi.org/10.3389/fninf.2015.00007.

50. Steine, H.C. (2009). Firmata: Towards making microcontrollers act like extensions of the computer. In Proceedings of the International Conference on New Interfaces for Musical Expression.

51. Musall, S., Kaufman, M.T., Juavinett, A.L., Gluf, S., and Churchland, A.K. (2019). Single-trial neural dynamics are dominated by richly varied movements. Nat Neurosci *22*. https://doi.org/10.1038/s41593-019-0502-4.

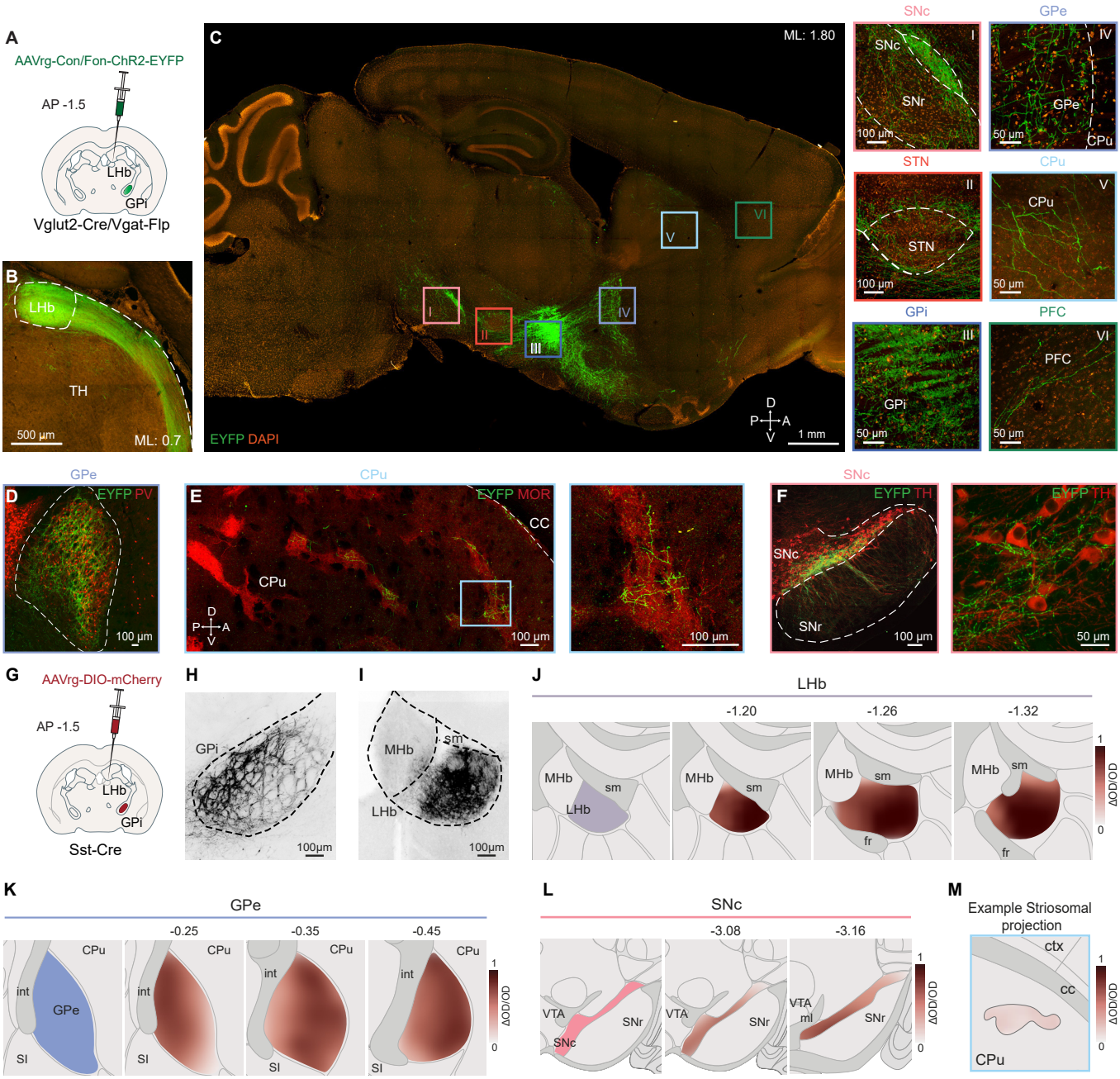
A
AAVrg-Con/Fon-ChR2-EYFP
AP -1.5
LHb
GPi
Vglut2-Cre/Vgat-Flp
B
LHb
TH
500 µm
ML: 0.7
C
ML: 1.80
EYFP DAPI
1 mm
SNc
SNr
GPe
CPu
STN
GPi
PFC
100 µm
50 µm
D
EYFP PV
E
EYFP MOR
CC
F
EYFP TH
G
AAVrg-DIO-mCherry
Sst-Cre
H
I
MHb
sm
J
LHb
-1.20
-1.26
-1.32
fr
ΔOD/OD
K
GPe
-0.25
-0.35
-0.45
int
SI
L
SNc
-3.08
-3.16
VTA
ml
M
Example Striosomal projection
ctx
cc

Figure 1

743 **Figure 1 | Mapping of Collateral projections of LHb-projecting Vglut2+/Vgat+/Sst+ GPi**
744 **neurons.**
745 **A**, Schematic representation of the experimental strategy (Vglut2/Vgat$^{GPi\text{-}LHb}$): AAVrg-Con/Fon-
746 ChR2-EYFP virus was injected in the Lhb of Vglut2-Cre/Vgat-Flp animals.
747 **B**, Sagittal section showing the projections of Vglut2+/Vgat+$^{GPi\text{-}LHb}$ neurons in LHb. N=7 animals.
748 **C**, Left: sagittal section showing the collateral projections of Vglut2+/Vgat+$^{GPi\text{-}LHb}$ neurons in
749 different brain regions. Roman numerals indicate regions of interest. Right: close-up images of
750 Vglut2+/Vgat+$^{GPi\text{-}LHb}$ cell bodies in the GPi (III) and axonal projections in the SNc (I), STN (II), GPe
751 (IV), CPu (V) and PFC (VI). N=7 animals.
752 **D**, ISH showing the Vglut2+/Vgat+$^{GPi\text{-}LHb}$ axons in the GPe confirmed by parvalbumin (PV)
753 immunostaining.
754 **E**, Left: ISH demonstrating Vglut2+/Vgat+$^{GPi\text{-}LHb}$ axonal collaterals targeting striosomes in the CPu.
755 Striosomes identified by μ-Opioid Receptor 1 (MOR) staining. Right: Close up of Vglut2+/Vgat+$^{GPi\text{-}}$
756 ^{LHb} axons projecting in one striosome area.
757 **F**, Left: ISH depicting Vglut2+/Vgat+$^{GPi\text{-}LHb}$ axons in the SNc confirmed by tryptophan hydroxylase
758 (TH) staining. Right: Close up of Vglut2+/Vgat+$^{GPi\text{-}LHb}$ axons surrounding TH-labelled neurons.
759 **G**, Schematic representation of an alternative experimental strategy (SST$^{GPi\text{-}LHb}$) to achieve
760 selective expression in LHb-projecting GPi neurons: injection of AAVrg-DIO-mCherry into the LHb
761 of Sst-Cre mice.
762 **H**, Coronal section showing the labelling of SST$^{GPi\text{-}LHb}$ neurons in GPi. N=3 animals.
763 **I**, Coronal section showing the projections of SST$^{GPi\text{-}LHb}$ axons in LHb. N=3 animals.
764 **J**, Heatmap showing the average fluorescence intensity of the projections of SST$^{GPi\text{-}LHb}$ axons in
765 three anteroposterior levels of LHb.
766 **K**, Heatmap showing the average fluorescence intensity of the projections of SST$^{GPi\text{-}LHb}$ axons in
767 three anteroposterior levels of GPe.
768 **L**, Heatmap showing the average fluorescence intensity of the projections of SST$^{GPi\text{-}LHb}$ axons in
769 two anteroposterior levels of SNc.
770 **M**, Example heatmap showing the fluorescence intensity of the projections of SST$^{GPi\text{-}LHb}$ axons in
771 one striosome of one Sst animal.
772 Abbreviations: AP: anteroposterior, GPi: globus pallidus interna, LHb: Lateral habenula, TH:
773 thalamus, ML: mediolateral, EYFP: enhanced yellow fluorescent protein, SNc: substantia nigra
774 pars compacta, SNr: substantia nigra pars reticulata, GPe: globus pallidus externa, STN:
775 subthalamic nucleus, CPu: Caudate putamen, PFC: prefrontal cortex, MHb: Medial habenula, sm:

776 stria medullaris, fr: fasciculus retroflexus, OD: optical density, int: internal capsule, SI: substantia
777 innominata, VTA: ventral tegmental area, ml: medial lemniscus, ctx: cortex, cc: corpus callosum.

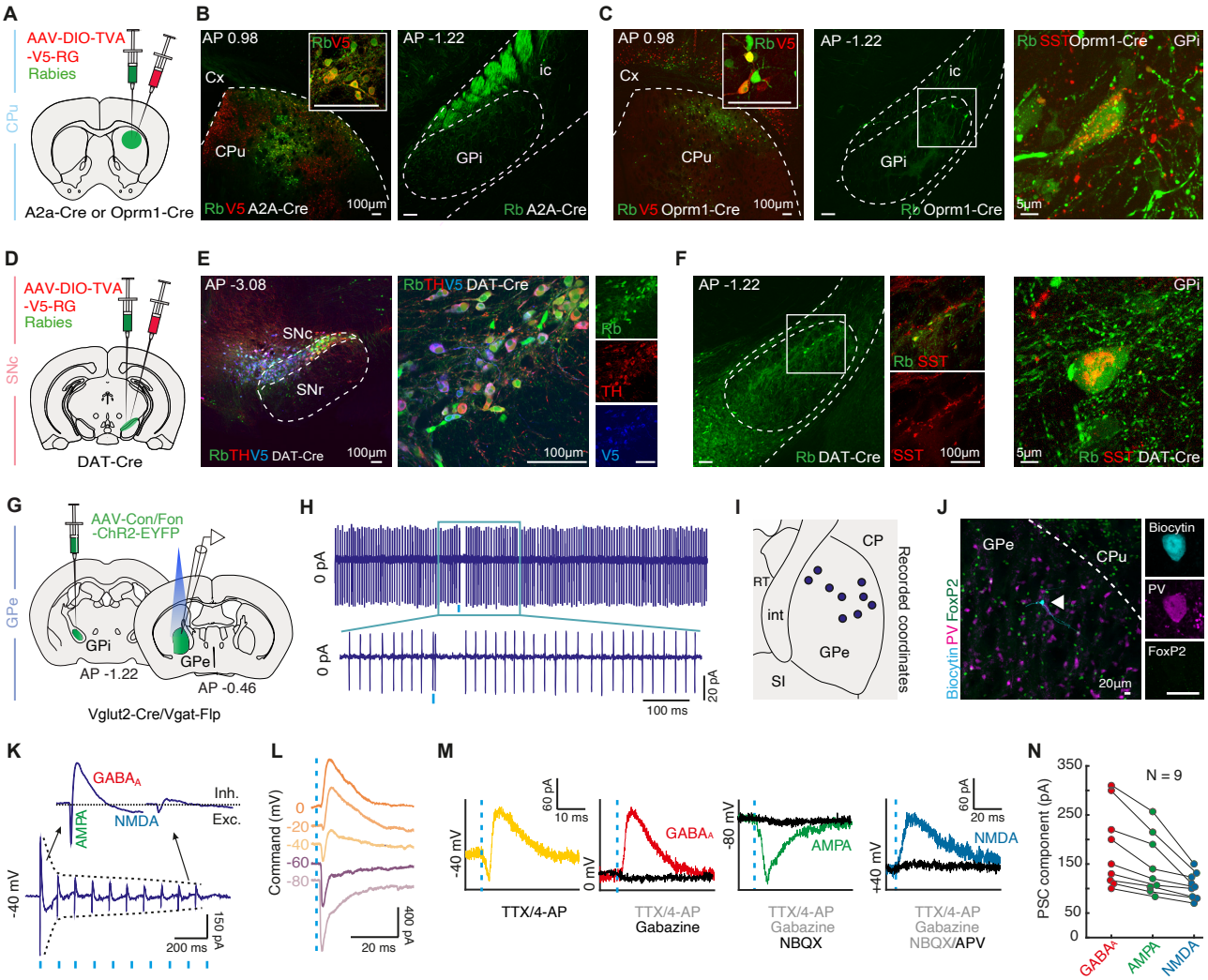
A
AAV-DIO-TVA
-V5-RG
Rabies
CPu
A2a-Cre or Oprm1-Cre
B
AP 0.98
Rb V5
Cx
CPu
Rb V5 A2A-Cre
100µm
AP -1.22
ic
GPi
Rb A2A-Cre
C
AP 0.98
Rb V5
Cx
CPu
Rb V5 Oprm1-Cre
100µm
AP -1.22
ic
GPi
Rb Oprm1-Cre
Rb SST Oprm1-Cre
GPi
5µm
D
AAV-DIO-TVA
-V5-RG
Rabies
SNc
DAT-Cre
E
AP -3.08
SNc
SNr
Rb TH V5 DAT-Cre
100µm
Rb TH V5 DAT-Cre
100µm
Rb
TH
V5
F
AP -1.22
Rb DAT-Cre
Rb SST
SST
100µm
GPi
5µm
Rb SST DAT-Cre
G
AAV-Con/Fon
-ChR2-EYFP
GPe
GPi
GPe
AP -1.22
AP -0.46
Vglut2-Cre/Vgat-Flp
H
0 pA
0 pA
20 pA
100 ms
I
CP
RT
int
GPe
SI
Recorded coordinates
J
GPe
CPu
Biocytin PV FoxP2
20µm
Biocytin
PV
FoxP2
K
GABA$_A$
Inh.
Exc.
AMPA
NMDA
-40 mV
200 ms
150 pA
L
Command (mV)
0
-20
-40
-60
-80
20 ms
400 pA
M
-40 mV
60 pA
10 ms
TTX/4-AP
0 mV
GABA$_A$
TTX/4-AP
Gabazine
-80 mV
AMPA
TTX/4-AP
Gabazine
NBQX
+40 mV
60 pA
20 ms
NMDA
TTX/4-AP
Gabazine
NBQX/APV
N
N = 9
PSC component (pA)
350
250
150
50
0
GABA$_A$
AMPA
NMDA

Figure 2

Figure 2 | GPi-LHb neurons make monosynaptic connections in the basal ganglia and co-transmit glutamate and GABA.

A, Schematic illustration of the experimental strategy: AAV-DIO-TVA-V5-RbG (V5) and Rabies-EGFP (Rb) were injected into the dorsolateral CPu of A2a-Cre and Oprm1-Cre mice, 21 days apart.

B, Representative ISH image of starter neurons co-expressing V5 (red) and Rb (green) in the CPU of an A2a-Cre mouse. Insert: close-up of starter neurons. Right: No Rb+ neurons in the GPi. Scale bars: 100μm. Replicated in N = 4 A2a-Cre mice; for cell counts see **Supplementary Table 1**. See also **Supplementary Figure 6A**.

C, Representative image of starter neurons co-expressing V5 (red) and Rb (green) in the CPU of an Oprm1-Cre mouse. Insert: closeup of starter neurons. Middle: Rb+ presynaptic neurons in the GPi. Right: Sst+ (red) Rb+ (green) neuron in the GPi. Scale bars: 100μm (left, middle); 5μm (right). Replicated in N = 6 Oprm1-Cre mice; for cell counts see **Supplementary Table 1**. See also **Supplementary Figure 6B**.

D, Schematic illustration of the experimental strategy: AAV-DIO-TVA-V5-RbG (V5) and Rabies-EGFP (Rb) were injected into the SNc of DAT-Cre mice, 21 days apart.

E, Representative ISH image of starter neurons co-expressing the V5 (blue), Rb (green), and TH (red) in the SNc of a DAT-Cre mouse. Right: closeup of starter neurons. Scale bars: 100μm.

F, Representative Rb+ GPi neurons (green) targeting DAT+ SNc neurons. Close-up images to the right show Sst+ (red) Rb+ (green) GPi neurons. Right: Close-up of one Sst+ Rb+ GPi neuron. Scale bars: 100μm (left), 5μm (right). Replicated in N = 3 DAT-Cre mice; for cell counts see **Supplementary Table 1**.

G, Schematic illustration of the experimental strategy: AAV-Con/Fon-ChR2-EYFP was unilaterally injected directly into the GPi of Vglut2-Cre/Vgat-Flp mice. ChR2-expressing GPi axon terminals were activated by blue light and postsynaptic currents in GPe neurons were recorded in whole-cell voltage-clamp mode.

H, Optogenetic activation (3ms, blue tick mark) of GPi axons results in an action potential doublet followed by a silent phase in a postsynaptic GPe neuron. Replicated in n = 3 neurons.

I, Schematic illustration showing the location of each recorded GPe neuron responsive to GPi axonal activation. n = 9 neurons.

J, ISH image showing a responsive GPe neuron (white arrowhead) filled with biocytin (cyan), co-expressing PV (pink) but not FoxP2 (green). Right: Close-up ISH images of the responsive PV+ FoxP2- GPe neuron. Scale bars: 20μm. Replicated in n = 9 recorded responsive GPe neurons. See also **Supplementary Figure 6D**.

812 **K**, Graph showing the response of a GPe neuron to optogenetic stimulation (10 Hz, 3 ms, blue
813 tick marks) of GPi axon terminals. The envelope (dashed line) indicates a symmetrical short-term
814 depression of the GPi-GPe synapse. Upper traces: Excitatory (Exc, AMPA, NMDA) and inhibitory
815 (Inh, GABAA) postsynaptic current components of the first and tenth pulse. Replicated in n = 3
816 neurons. Holding potential was -40 mV. See also **Supplementary Figure 6C.**
817 **L**, Graph depicting the optogenetically evoked postsynaptic currents at different holding
818 potentials. Averages of 5 traces for n = 9 neurons.
819 **M**, Graphs depicting the pharmacological dissection of postsynaptic currents evoked by opto-
820 stimulation (blue dashed line). Gabazine (10 μM), NBQX (20 μM) and APV (50 μM) were applied
821 sequentially. TTX (1 μM) and 4-AP (5 mM) were applied throughout the entire recording. Averages
822 of 5 traces for n = 9 neurons.
823 **N**, Graph showing the maximal current amplitude of GABAA, AMPA- and NMDA-currents
824 measured at the holding potentials shown in **M**.
825 Abbreviations: CPu: Caudate putamen, Cx: Cortex, GPi: Globus pallidus interna, ic: internal
826 capsule, SNc: substantia nigra pars compacta, SNr: substantia nigra pars reticulata, GPe: Globus
827 pallidus externa, RT: Reticular nucleus of the thalamus, int: internal capsule, SI: Substantia
828 innominata, CP: Caudate putamen, FoxP2: Forkhead box P2.

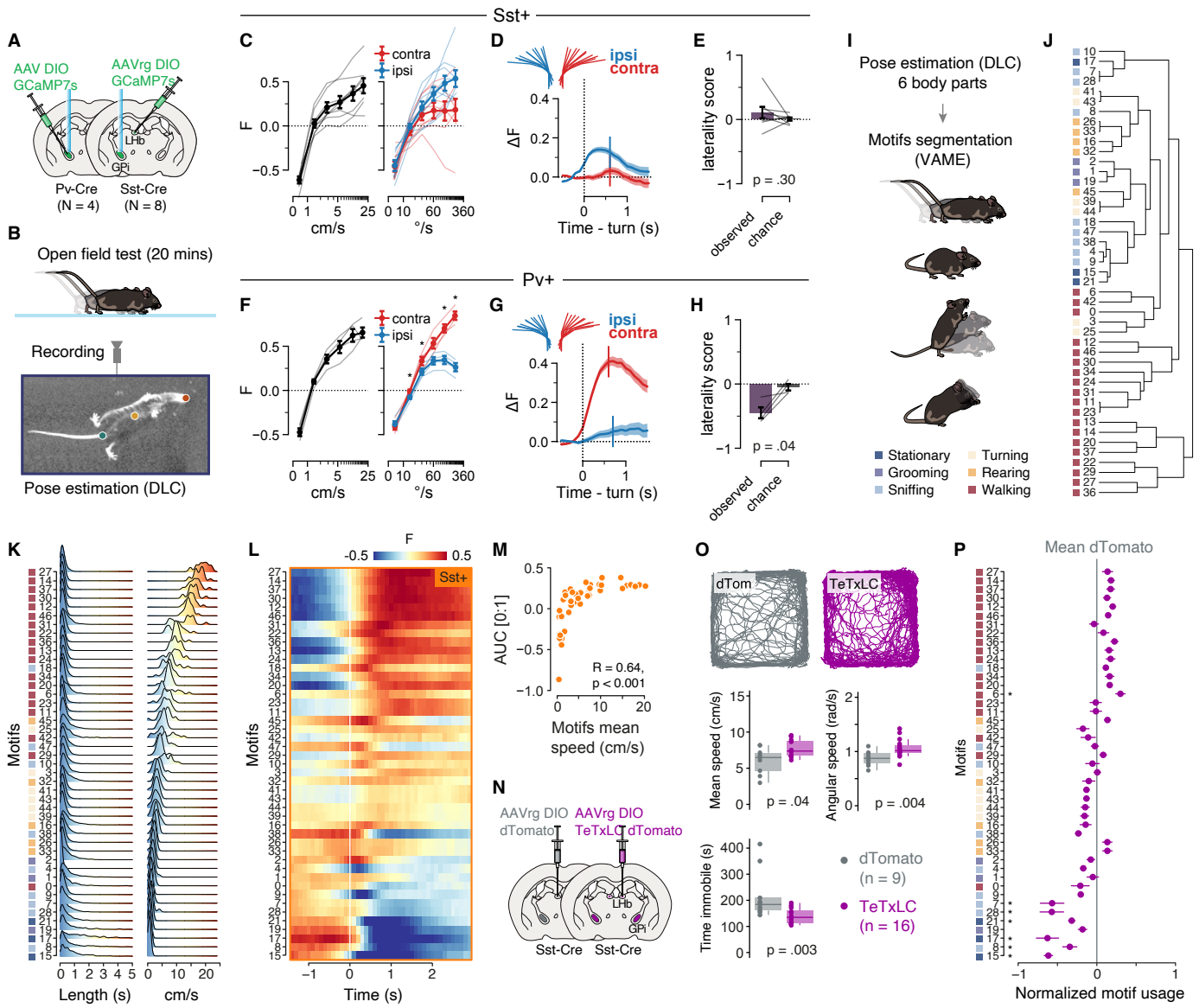
Sst+
Pv+
A
AAV DIO GCaMP7s
AAVrg DIO GCaMP7s
LHb
GPi
Pv-Cre (N = 4)
Sst-Cre (N = 8)
B
Open field test (20 mins)
Recording
Pose estimation (DLC)
C
contra
ipsi
F
cm/s
°/s
D
ipsi
contra
ΔF
Time - turn (s)
E
laterality score
p = .30
observed
chance
F
G
p = .04
H
I
Pose estimation (DLC)
6 body parts
Motifs segmentation
(VAME)
Stationary
Grooming
Sniffing
Turning
Rearing
Walking
J
K
Motifs
Length (s)
cm/s
L
Sst+
Time (s)
M
AUC [0:1]
R = 0.64,
p < 0.001
Motifs mean
speed (cm/s)
N
AAVrg DIO dTomato
AAVrg DIO TeTxLC dTomato
Sst-Cre
O
dTom
TeTxLC
Mean speed (cm/s)
Angular speed (rad/s)
Time immobile (s)
p = .04
p = .004
p = .003
dTomato (n = 9)
TeTxLC (n = 16)
P
Mean dTomato
Normalized motif usage

Figure 3

Figure 3 | Sst+ GPi neuronal activity correlates with action speed.

A, Schematic illustration showing the experimental strategy: AAVrg-DIO-jGCaMP7s was injected unilaterally in the LHb of Sst-Cre mice (N = 8). AAV-DIO-jGCaMP7s was injected unilaterally in the GPi of Pv-Cre mice (N = 4). 200µm fibers were implanted above the GPi. See also **Supplementary Figure 7A-B** and **Supplementary Table 2**.

B, Schematic representation of the open field test setup with a bottom-up camera (top), example tracking of mouse behavior for video analysis and pose estimation of three body parts through DeepLabCut (bottom).

C, Population activity (F) at different velocities (left) and angular velocities split by turn direction (right) in the open field. Bins spaced evenly along the log-scaled x-axes. Thin lines: individual Sst-Cre mice. Circles: mean ± SEM. Contra- vs ipsiversive turn F: all bins $p > 0.3$; paired t-tests, Holm-Sidak correction. N = 8 mice.

D, ΔF during spontaneous ipsi- and contraversive turns in the open field, aligned to turn onset. Baseline for ΔF: mean F of the 0.5s preceding the onset of the turn. Traces: mean ± SEM. Vertical lines: median turn offset times. Ipsi/contra n = 1235/1360 turns (pooled, N = 8 mice).

E, Measure indicating how well single-turn mean ΔF discriminates turn direction. The laterality score was defined as the area under the ROC curve (AUC), scaled from 1 to 1. Positive vs negative scores indicate relatively higher activation during ipsi- vs contraversive turns. Chance level established by shuffling the turn direction labels prior to score computation. Gray lines: individual Sst-Cre mice. Bars: mean ± SEM. $p = 0.30$; paired t-test. N = 8 mice.

F-H, same as C-E for Pv-Cre mice.

F, *: $p < 0.05$; paired t-tests, Holm-Sidak correction. N = 4 mice.

G, Ipsi/contra n = 355/395 turns (pooled, N = 4 mice).

H, $p = 0.04$; paired t-test. N = 4 mice.

I, Schematic illustration of the video analysis: pose estimation of six body parts and segmentation of behavior motifs.

J, Dendrogram reporting the clustering and behavioral annotations of the 45 identified motifs.

K, Densities plots reporting the motifs length (left) and the mouse speed (right) during the execution of the motifs.

L, Mean average of calcium transients signals of Sst+ GPi-LHb population recorded through fiber photometry and aligned to motifs onset. The motifs were ordered according to the motif's speed (cf. panel **J**).

M, AUC of the signal (period between 0 and +1 sec) correlated with the motif’s speed for Sst+ GPi-LHb recording. Spearman’s correlation coefficients are reported.

863 **N**, Illustration showing the experimental strategy: AAVrg-DIO-dTomato (N = 9) or AAVrg-DIO-
864 TeTxLC-dTomato (N = 16) were injected bilaterally in the LHb of Sst-Cre mice (see also
865 **Supplementary Figure 7C**).
866 **O**, Upper panel: example traces of a control and TeTxLC mouse. Bottom panels: boxplots
867 reporting linear and angular speed, and the time passed immobile for both control and TeTxLC
868 mice. The p-values of the Mann-Whitney U test are reported in each panel.
869 **P**, Normalized (on the performance of control mice) motifs usage for TeTxLC mice. Multiple Mann-
870 Whitney U test followed by Holm-Sidak correction. Mean ± SEM. *: $p<0.05$.
871 See **Supplementary Table 3** for more information about statistical tests**.**
872 Abbreviations: ΔF: delta fluorescence, ROC: Receiver operating characteristic, AUC: area under
873 curve.

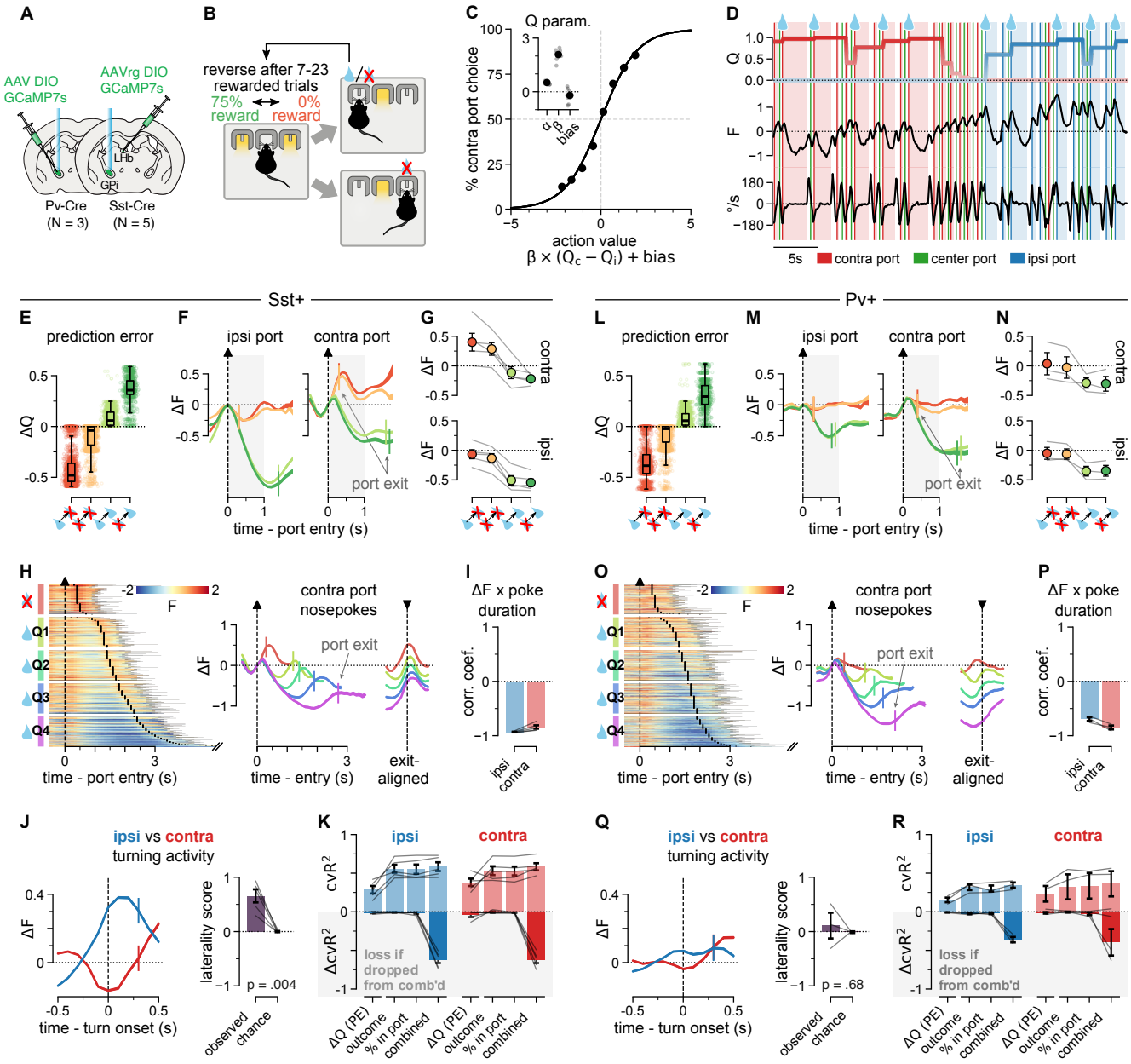
A
AAV DIO GCaMP7s
AAVrg DIO GCaMP7s
LHb
GPi
Pv-Cre (N = 3)
Sst-Cre (N = 5)
B
reverse after 7-23 rewarded trials
75% reward
0% reward
C
Q param.
% contra port choice
action value
β × (Qc − Qi) + bias
D
Q
F
°/s
5s
contra port
center port
ipsi port
Sst+
Pv+
E
prediction error
ΔQ
F
ipsi port
contra port
ΔF
port exit
time - port entry (s)
G
contra
ipsi
L
M
N
H
Q1
Q2
Q3
Q4
contra port nosepokes
time - port entry (s)
time - entry (s)
exit-aligned
I
ΔF x poke duration
corr. coef.
ipsi
contra
O
P
J
ipsi vs contra turning activity
time - turn onset (s)
laterality score
p = .004
observed
chance
K
ipsi
contra
cvR²
ΔcvR²
loss if dropped from comb'd
ΔQ (PE)
outcome
% in port
combined
Q
p = .68
R

Figure 4

874 **Figure 4 | Sst+ GPi-LHb population activity is movement-related and ipsiversive turn-**
875 **preferring.**
876 **A**, Top: Schematic representation of the experimental strategy: AAVrg-DIO-jGCaMP7s was
877 injected unilaterally in the LHb of Sst-Cre mice (N = 5). AAV-DIO-jGCaMP7s was injected
878 unilaterally in the GPi of Pv-Cre mice (N = 3). 200µm optic fibers (OF) were implanted above the
879 GPi. Bottom: jGCaMP7s (green) expressed in the GPi of a Sst-Cre mouse. See also
880 **Supplementary Figure 7A-B** and **Supplementary Table 2**.
881 **B**, Probabilistic reversal task. After trial initiation at the center nose-poke port, a reward becomes
882 available (with 75% probability) at one of the two peripheral ports. The rewarded choice reverses
883 unpredictably every 7-23 rewarded trials.
884 **C**, Graph showing the probability of contraversive port choice (relative to implanted hemisphere)
885 as a function of the outcome history-based action value. Action values were fitter per mouse using
886 Q-learning. Line: logistic regression fit, trials pooled over mice. n = 27053 trials (pooled, N = 8
887 mice; 3 Pv-Cre + 5 Sst-Cre). Circles: mean ± SEM of mice for 8 evenly sized action value bins. N
888 = 8 mice. Inset: Q-learning parameters. Gray circles: individual mice. Black circles: mean ± SEM.
889 N = 8 mice.
890 **D**, Exemplary session snippet showing Q value estimates for both port choices, Sst+ GPi-LHb
891 population activity, and the mouse's angular velocity over the course of several trials. Lines: nose-
892 poke port entries. Light shading: port occupation. All color coding as indicated in the panel legend.
893 The water drop icon marks rewarded choice port entries; unmarked entries were not rewarded.
894 Data from one Sst-Cre mouse.
895 **E**, Graph depicting trial-by-trial changes in the fitted port choice Q values (ΔQ, i.e. prediction
896 error). Trials split by current and previous trial outcomes. Open circles: single trials. Boxplots:
897 center lines represent the distribution median; edges, the upper and lower quartiles; whiskers, the
898 minimum and maximum values within the quartiles ± 1.5 x IQR. n = 1220/1867/1157/606 trials
899 (pooled, N = 5 mice).
900 **F**, Graph showing the fluorescence change (ΔF) following contraversive (left) and ipsiversive
901 (right) choice port entries (time 0, dashed arrows). Trials split by current and previous trial
902 outcomes. Baseline for ΔF: F at port entry. Color coding as in panel E. Traces: mean ± SEM.
903 Vertical lines: median port exit times. Gray shading: 1s time window used to compute means in
904 panel G. Ipsi n = 613/983/498/296, contra n = 607/884/659/310 trials (pooled, N = 5 mice).
905 **G**, Graph demonstrating the mean ΔF in the second after choice port entry (cf. panel F). Trials
906 split by the current and previous trial outcome. Gray lines: individual Sst-Cre mice. Circles: mean
907 ± SEM. A repeated-measures ANOVA showed significant main effects for the current trial

outcome ($p < 0.001$) and the port side ($p = 0.004$). The main effect of the previous trial outcome ($p = 0.097$) and all interaction effects were insignificant (all $p > 0.3$). N = 5 mice.

H, Raster: single-trial fluorescence (F) time-locked to contraversive choice port entry. Rewarded trials were split into nose-poke duration-based quartile bins (Q1-4) and a fifth, equivalently sized bin of unrewarded trials was randomly sampled. All bins sorted by poke duration. Black ticks mark port exit times. Traces: ΔF means ± SEM for each trial bin of the raster plot, plotted aligned to port entry (left, dashed up-arrow) and port exit times (right, dashed down-arrow). Baseline for ΔF: F at port entry. Vertical lines: median port exit times. n = 242 trials per bin (pooled, N = 5 animals).

I, Pearson correlation between mean ΔF and time spent nose-poking for all ipsiversive and contraversive choice port nose-pokes, irrespective of trial outcome. Gray lines: individual Sst-Cre mice. Bars: mean ± SEM. N = 5 mice.

J, Left: ΔF during task-engaged ipsi- and contraversive turns, aligned to turn onset. Includes choice and initiation turns. Baseline for ΔF: mean F of the 0.5s preceding the onset of the turn. Traces: mean ± SEM. Vertical lines: median turn offset times. Ipsi/contra n = 4658/4612 turns (pooled, N = 5 mice). Right: ROC-AUC score-based laterality score indicating how well single-turn mean ΔF discriminates turn direction. Positive vs negative scores indicate higher activation during ipsi- vs contraversive turns. Chance level established by shuffling the turn direction labels prior to score computation. Gray lines: individual Sst-Cre mice. Bars: mean ± SEM. $p = 0.004$; paired t-test. N = 5.

K, Predictive power of ridge regressions predicting trial-by-trial mean ΔF in the 1s window following ipsi or contra choice port entry (see panel F), fit per mouse and turn direction. cvR^2: cross-validated variance explained by each variable individually (ΔQ, outcome, % in port) and combined. ΔcvR^2: loss of explanatory power of the combined 3-variable model when variables are individually and jointly scrambled. Gray lines: individual Sst-Cre mice. Bars: mean ± SEM. Pairwise comparisons of the single-variable models' cvR^2 scores were insignificant (all $p > 0.3$), except for the ipsi port comparisons involving the "ΔQ" model (both $p < 0.03$; paired t-tests, Holm-Sidak correction). Exclusion of individual variables from the combined model failed to affect its performance (all $p > 0.07$; t-tests, Holm-Sidak correction). N = 5 mice.

L- R, same as E-K for Pv-Cre mice.

L, n = 986/1509/1174/534 trials (pooled, N = 3 mice).

M, ipsi n = 543/1054/626/249, contra n = 443/455/548/285 trials (pooled, N = 3 mice).

N, Repeated-measures ANOVA: significant main effect for the current trial outcome ($p = 0.029$). The main effect of the previous outcome ($p = 0.191$) and all other main and interaction effects were insignificant (all $p > 0.3$). N = 3 mice.

O, n = 208 trials per bin (pooled, N = 3 mice).

P, N = 3 mice.

Q, Left: Ipsi/contra n = 4099/4250 turns (pooled, N = 3 mice). Right: p = 0.68; paired t-test. N = 3.

R, Pairwise comparisons of the single-variable models' cvR^2 scores were insignificant (all p > 0.05; paired t-tests, Holm-Sidak correction). Exclusion of individual variables from the combined model failed to affect its performance (all p > 0.1; t-tests, Holm-Sidak correction). N = 5 mice.

See **Supplementary Table 3** for more information about statistical tests**.**

Abbreviations: F: fluorescence, Q: trial-by-trial value estimate of the Q-learning model, PE: prediction error.

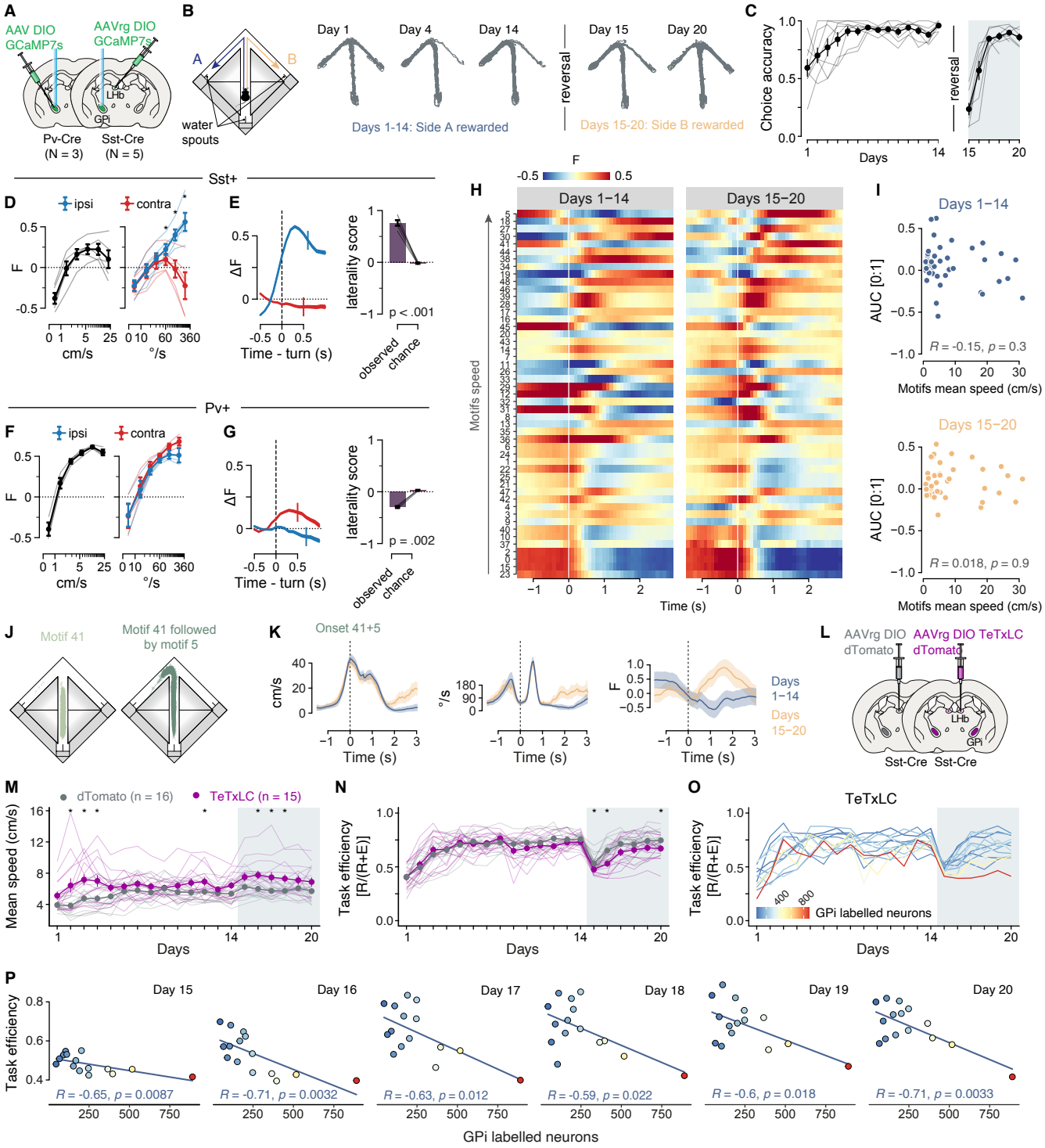
A
AAV DIO GCaMP7s
AAVrg DIO GCaMP7s
LHb
GPi
Pv-Cre (N = 3)
Sst-Cre (N = 5)
B
A
B
water spouts
Day 1
Day 4
Day 14
Days 1-14: Side A rewarded
reversal
Day 15
Day 20
Days 15-20: Side B rewarded
C
Choice accuracy
Days
reversal
Sst+
D
ipsi
contra
F
cm/s
°/s
E
ΔF
Time - turn (s)
laterality score
p < .001
observed
chance
Pv+
F
ipsi
contra
G
p = .002
H
F
Days 1−14
Days 15−20
Motifs speed
Time (s)
I
Days 1−14
AUC [0:1]
R = -0.15, p = 0.3
Motifs mean speed (cm/s)
Days 15−20
R = 0.018, p = 0.9
J
Motif 41
Motif 41 followed by motif 5
K
Onset 41+5
cm/s
°/s
Time (s)
Days 1−14
Days 15−20
L
AAVrg DIO dTomato
AAVrg DIO TeTxLC dTomato
Sst-Cre
Sst-Cre
M
dTomato (n = 16)
TeTxLC (n = 15)
Mean speed (cm/s)
Days
N
Task efficiency [R/(R+E)]
Days
O
TeTxLC
Task efficiency [R/(R+E)]
GPi labelled neurons
Days
P
Day 15
Day 16
Day 17
Day 18
Day 19
Day 20
Task efficiency
R = -0.65, p = 0.0087
R = -0.71, p = 0.0032
R = -0.63, p = 0.012
R = -0.59, p = 0.022
R = -0.6, p = 0.018
R = -0.71, p = 0.0033
GPi labelled neurons

Figure 5

Figure 5 | Changes in Sst+ GPi-LHb activity during a maze choice task.

A, Schematic representation of the experimental strategy: AAVrg-DIO-jGCaMP7s was injected unilaterally in the LHb of Sst-Cre mice (N = 5). AAV-DIO-jGCaMP7s was injected unilaterally in the GPi of Pv-Cre mice (N = 3). 200µm fibers were implanted above the GPi. See also **Supplementary Figure 7A-B** and **Supplementary Table 2**.

B, Illustration showing the single-reversal maze task. For 14 days the choice to turn contraversive relative to the imaged hemisphere was reinforced with a water drop (13 µl) at the end of the corridor (blue arrow). On day 15 and thereafter, the rewarded choice was reversed (orange arrow). Example traces of one mouse's paths through the maze are reported for days 1, 4, 14, 15 and 20. See **Methods** for details.

C, Choice accuracy of the fiber photometry-recorded mice in the single-reversal maze task across days. Gray lines: individual Sst-Cre and Pv-Cre mice. Circles: mean ± SEM. N = 8 mice.

D, Population activity (F) at different velocities (left) and angular velocities split by turn direction (right) in the open field. Bins spaced evenly along the log-scaled x-axes. Thin lines: individual Sst-Cre mice. Circles: mean ± SEM. *: $p < 0.05$; paired t-tests, Holm-Sidak correction. N = 5 mice.

E, Left: ΔF during task-engaged ipsi- and contraversive turns, aligned to turn onset. Baseline for ΔF: mean F of the 0.5s preceding the onset of the turn. Traces: mean ± SEM. Vertical lines: median turn offset times. Ipsi/contra n = 3031/2901 turns (pooled, N = 5 mice). Right: ROC-AUC score-based measure indicating how well single-turn mean ΔF discriminates turn direction. Positive vs negative scores indicate higher activation during ipsi- vs contraversive turns. Chance level established by shuffling the turn direction labels prior to score computation. Gray lines: individual Sst-Cre mice. Bars: mean ± SEM. $p < 0.001$; paired t-test. N = 5 mice.

F-G, same as D-E for Pv-Cre mice.

F, all bins $p > 0.2$; paired t-tests, Holm-Sidak correction. N = 3 mice.

G, Left: Ipsi/contra n = 1601/1628 turns (pooled, N = 3 mice). Right: $p = 0.002$; paired t-test. N = 3 mice.

H, Mean average of calcium transients of Sst+ GPi-LHb neurons aligned to motifs onset during the pre- and post-reversal phase. The motifs were ordered according to the motif's speed.

I, Area under the curve (AUC) of the signal (period between 0 and +1 sec) correlated with the motif's speed during the pre- and post-reversal phase. Spearman's correlation coefficients are reported.

J, Traces of the movements of the mice during motif 41 (in light green, "straight run") and 41 followed by 5 (in dark green, "contraversive turn") in the maze arena.

984 **K**, Calcium transients of Sst+ GPi-LHb neurons, angular and linear speed of recorded mice
985 aligned on the end of motif 41 and the beginning of motif 5 during the pre- and post-reversal
986 phase. Mean ± SEM.
987 **L**, AAVrg-DIO-dTomato (N = 16) or AAVrg-DIO-TeTxLC-dTomato (N = 15) were injected
988 bilaterally in the LHb of Sst-Cre mice (see also **Supplementary Figure 7C**).
989 **M**, Mean speed and **N,** task efficiency of AAVrg-DIO-dTomato and AAVrg-DIO-TeTxLC-dTomato
990 injected mice in the single-reversal maze task across 20 days of test. Mean ± SEM. *: $p < 0.05$,
991 multiple comparisons using the Mann-Whitney U test followed by Holm-Sidak correction.
992 **O**, AAVrg-DIO-TeTxLC-dTomato injected mice's task efficiency across days color-coded with the
993 number of TeTxLC-expressing cells into the GPi.
994 **P**, AAVrg-DIO-TeTxLC-dTomato injected mice's task efficiency during post-reversal days
995 regressed on the number of TeTxLC-expressing cells. Pearson's correlation coefficients are
996 reported.
997 See **Supplementary Table 3** for more information about statistical tests**.**
998 Abbreviations: F: fluorescence, AUC: area under curve, R: rewards, E: errors, GPi: globus pallidus
999 interna.

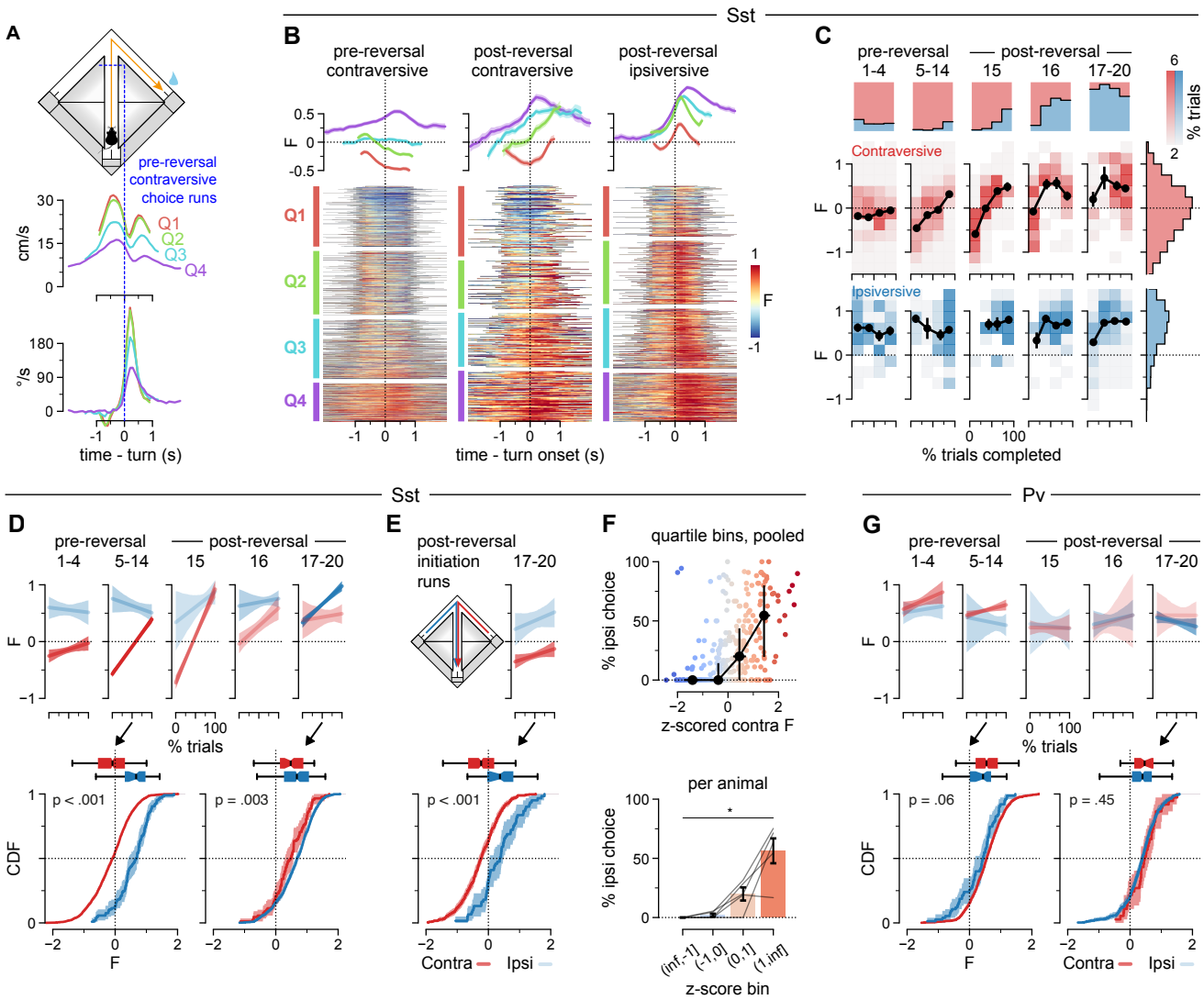
A
pre-reversal
contraversive
choice runs
cm/s
Q1
Q2
Q3
Q4
°/s
time - turn (s)
Sst
B
pre-reversal
contraversive
post-reversal
contraversive
post-reversal
ipsiversive
F
time - turn onset (s)
C
pre-reversal
post-reversal
1-4
5-14
15
16
17-20
% trials
Contraversive
Ipsiversive
% trials completed
Sst
D
pre-reversal
post-reversal
% trials
CDF
p < .001
p = .003
E
post-reversal
initiation
runs
17-20
p < .001
Contra
Ipsi
F
quartile bins, pooled
% ipsi choice
z-scored contra F
per animal
z-score bin
Pv
G
p = .06
p = .45

Figure 6

Figure 6 | Dynamic modulation of Sst+ GPi-LHb neuronal activity during strategy shifts.

A, Average velocities (middle) and contraversive angular velocities (bottom) of Sst-Cre mice during pre-reversal contraversive choice runs in each session-trial quartile bin (Q1-4, see also panel **B**), aligned to turn entry (top). Traces are clipped at the median initiation zone exit (run start) and reward zone entry (run end) time points. Data from sessions 1-4 excluded. Traces: mean ± SEM. Q1-4 n = 424/446/418/271 turns (pooled, N = 5 mice).

B, Raster: Sst+ GPi-LHb population activity (F) for individual pre- and post-reversal choice runs. Trials were grouped by task stage and turn direction, sorted by their relative position in their session (i.e., the percentage of session trials completed; trial number/session total), and finally binned into quartiles (Q1-4). The bins are unevenly sized due to the exclusion of trials in which mice returned to the center corridor before reaching a reward zone. Data from sessions 1-4 and pre-reversal ipsiversive trials omitted, the latter due to the low number of trials available. Traces: mean ± SEM for each raster bin; clipped as in panel A. n = 1559/397/826 trials (pooled, N = 5 mice).

C, Top row: stacked histograms indicating the fraction of contra- vs ipsiversive choices, by session quarter, over the course of the experiment. Quarters are based on trial quartiles, not time. Trials pooled over Sst-Cre mice. Middle and bottom rows: density plots visualizing Sst+ activity (F) during contra- (middle, in red) and ipsiversive (bottom, in blue) choice turns, by session quarter, over the course of the experiment (same as above). One data point, mean F, included per turn. Circles: mean ± SEM of turns by session quarter; mean omitted for bins with fewer than 5 data points total. Right-hand histograms: distribution of mean turn F pooled over all sessions. Contra/ipsi n = 2416/1013 trials (pooled, N = 5 mice; cf. panel D).

D, Top row: Linear regression analysis predicting single-trial Sst+ GPi-LHb choice turn activity based on session progress (i.e., the percentage of session trials completed; trial number/session total) and turn direction. Lines: regression fits and 95% confidence intervals (bootstrapped); increased transparency indicates smaller data sets. Contra n = 460/1559/216/93/88 and ipsi n = 82/105/45/107/674 (pooled, N = 5 mice).

Bottom row: distributions of Sst+ GPi-LHb activity (F) during choice turns, plotted by turn direction, for session bins “5-14” (left) and “17-20” (right). Boxplots: center lines represent the distribution median; edges, the upper and lower quartiles; whiskers, the 2.5% and 97.5% points of the distribution. Lines: cumulative distribution functions (CDFs) and 95% confidence intervals (bootstrapped). The CDF indicates the fraction of turns during which F was at or below a given value. p-values: Kolmogorov-Smirnov test, Holm-Sidak correction. “5-14" bin contra/ipsi n = 1559/105 and "7-20" bin contra/ipsi n = 88/674 (pooled, N = 5 mice).

E, Plots as described in D, but for post-reversal “return” runs from the choice zones towards the initiation zone (days 17-20). Bottom panel p-value computed and corrected with those in D. Contra/ipsi n = 369/61 (pooled, N = 5 mice).

F, Top: probability of ipsiversive choices plotted against the average F during contraversive turns for all session-quartile bins. Small circles: individual session-quartile averages, contra turn F z-scored per mouse. Black circles: median and IQR of binned session-quartiles; bin edges as indicated in bottom panel. n = 332 session-quartiles (pooled, N = 5 mice). Bottom: same as above, plotted per mouse. Thin lines: median of binned session-quartiles for individual mice; bin edges are indicated on the x axis. Bars: mean ± SEM. *: $p < 0.001$, repeated-measures ANOVA. N = 5 mice.

G, Same as D for Pv-Cre mice. Top row: contra n = 211/841/79/34/37 and ipsi n = 63/75/40/97/345 (pooled, N = 3 mice). Bottom row: “5-14” bin conta/ipsi n = 841/75 and "17-20" bin contra/ipsi n = 37/345 (pooled, N = 3 mice).

See **Supplementary Table 3** for more information about statistical tests**.**