

Sensory and Motor Systems

Sound-Evoked Responses of Distinct Neuron Classes from the Tail of the Striatum

 Matthew B. Nardoci,^{1,2}  Anna A. Lakunina,^{1,2} Devin C. Henderling,^{1,2} Jewlyssa C. Pedregon,^{1,3} Jennifer L. Mohn,¹ and  Santiago Jaramillo^{1,2}

<https://doi.org/10.1523/ENEURO.0201-22.2022>

¹Institute of Neuroscience, University of Oregon, Eugene, Oregon 97403, ²Department of Biology, University of Oregon, Eugene, Oregon 97403, and ³Department of Human Physiology, University of Oregon, Eugene, Oregon 97403

Abstract

Given its inputs from auditory structures and neuromodulatory systems, the posterior tail of the striatum is ideally positioned to influence behavioral responses to acoustic stimuli according to context and previous rewards. Results from previous studies indicate that neurons in this striatal region display selective responses to sounds. However, it is not clear whether different striatal cell classes code for distinct features of sounds or how different striatal output pathways may use acoustic information to guide behavior. Here we compared the sound-evoked responses of posterior striatal neurons that form the striatal direct pathway (and express the dopamine receptor D₁) to the responses of neighboring neurons in naive mice. We achieved this via optogenetic photo-identification of D₁-expressing neurons during extracellular electrophysiological recordings in awake head-fixed mice of both sexes. We found that the frequency tuning of sound-responsive direct-pathway striatal neurons is comparable with that of their sound-responsive neighbors. Moreover, we found that both populations encode amplitude-modulated sounds in a similar fashion. These results suggest that different classes of neurons in the posterior striatum of naive animals have similar access to acoustic features conveyed by the auditory system even outside the context of an auditory task.

Key words: auditory; direct pathway; dopamine receptor D₁; posterior striatum; subcortical

Significance Statement

Sound-driven decision-making is a key component of the behavioral repertoire of an animal in their quest for positive outcomes. Subsets of neurons in the striatum (a brain area associated with motor control and the integration of reward information) receive inputs from the auditory system, yet what features of sounds are accessible to specific striatal cell classes is not well understood. We found that multiple classes of posterior striatal neurons have access to detailed spectrotemporal acoustic features and could therefore potentially influence behavioral responses according to these properties of sounds.

Introduction

The striatum, as the primary input structure of the basal ganglia and a target of extensive dopaminergic inputs, is ideally positioned to influence behavioral responses to sensory stimuli according to context and previous rewards. Neurons in the posterior tail of the striatum receive numerous inputs from the auditory thalamus and the auditory cortex (Hintiryan et al., 2016; Chen et al., 2019;

Ponvert and Jaramillo, 2019; Valjent and Gangarossa, 2021), and these striatal neurons display robust responses to sounds (Bordi and LeDoux, 1992; Znamenskiy and Zador, 2013; Zhong et al., 2014; Guo et al., 2018). It is not known, however, whether different striatal cell classes code for distinct features of sounds or how different striatal output pathways may use acoustic information to guide behavior. A key step toward understanding the processing

Received May 17, 2022; accepted August 29, 2022; First published September 14, 2022.

The authors declare no competing financial interests.

Author contributions: S.J. designed research; M.B.N., A.A.L., J.C.P., and J.L.M. performed research; M.B.N., D.C.H., J.C.P., J.L.M., and S.J. analyzed data; M.B.N. and S.J. wrote the paper.

of sounds by striatal circuits and the role of different striatal cells in auditory learning is the characterization of sound-evoked responses by distinct striatal neuron classes in naive animals.

The striatum, including the posterior tail portion, is composed of a range of neuron classes with different gene expression and physiological profiles, including fast-spiking parvalbumin-expressing interneurons, spontaneously active cholinergic interneurons, and the abundant principal projection neurons (Kawaguchi, 1997; Valjent and Gangarossa, 2021). The large majority of these cells consists of medium spiny neurons that form the two main outputs of the striatum: the direct (striatonigral) pathway, composed of cells that express the dopamine receptor D₁; and the indirect (striatopallidal) pathway, with cells that express the dopamine receptor D₂ (Gerfen et al., 1990; Kreitzer and Malenka, 2008). Anatomical data suggest that, for some striatal regions, cortical sensory neurons differentially innervate each of these striatal pathways (Lei et al., 2004; Wall et al., 2013). Moreover, excitatory synapses onto direct and indirect pathway neurons exhibit different synaptic transmission and plasticity properties (Kreitzer and Malenka, 2007). Together, these observations raise the possibility that neurons from different classes in the posterior striatum are differentially influenced by sensory signals.

Using optogenetic photo-identification of specific cell populations in the posterior tail of the striatum of naive mice, we characterized the sound-evoked responses of neurons that express the dopamine receptor D₁ and compared these responses with those from neighboring neurons. We found that on average, sound-responsive D₁-expressing posterior striatal neurons have similar frequency tuning to their non-D₁ sound-responsive neighbors, and that both populations represent amplitude-modulated noise in a similar fashion. These results suggest that different classes of neurons in the posterior striatum, likely including both major output pathways, have access to acoustic features conveyed from the auditory system in naive animals even outside the context of an auditory task.

Materials and Methods

Animals

A total of 15 transgenic adult DRD1::ChR2 mice of both sexes, were used in this study. This mouse line was generated by crossing animals that express Cre recombinase in neurons positive for the dopamine receptor D₁ (MMRRC; RRID:MMRRC_036916-UCD) with mice that

This research was supported by the National Institute on Deafness and Other Communication Disorders (Grant R01DC015531) and the Office of the Vice President for Research and Innovation at the University of Oregon.

Acknowledgment: We thank Leah DeBlander and Brigid Deck for assistance with histology.

Correspondence should be addressed to Santiago Jaramillo at sjara@uoregon.edu.

<https://doi.org/10.1523/ENEURO.0201-22.2022>

Copyright © 2022 Nardoci et al.

This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International license, which permits unrestricted use, distribution and reproduction in any medium provided that the original work is properly attributed.

express the light-gated ion channel channelrhodopsin-2 (ChR2) in a Cre-dependent manner (stock #012569, The Jackson Laboratory). All procedures were conducted in accordance with the National Institutes of Health standards and were approved by the Institutional Animal Care and Use Committee of the University of Oregon.

Auditory stimuli

Experiments were performed inside a single-walled sound-isolation box (IAC Acoustics). Auditory stimuli were presented in an open-field configuration from a speaker (model MF1, Tucker-Davis Technologies) contralateral to the side of electrophysiological recordings. Speakers were calibrated using an ultrasonic microphone (model ANL-940-1, Med Associates) to obtain the desired sound intensity level for frequencies between 1 and 40 kHz. Stimuli were generated using Python software developed in-house (<https://taskontrol.readthedocs.io/>). The ensemble of auditory stimuli for evaluating frequency tuning consisted of pure-tone pips (duration, 100 ms) at 16 frequencies logarithmically spaced between 2 and 40 kHz and at 11 different intensities (15–70 dB SPL in 5 dB steps). We presented at least nine repetitions per frequency–intensity combination with interstimulus intervals randomized in the range 0.7–0.9 s. Stimuli for evaluating responses to temporal sound features were sinusoidally amplitude-modulated white noise at 11 modulation rates logarithmically spaced between 4 and 128 Hz (modulation depth, 100%; duration, 500 ms; 60 dB SPL maximum). We presented at least 20 repetitions per condition using an interstimulus interval randomized in the range 0.9–1.1 s. All stimuli had a 2 ms ramp-up and ramp-down. During sound presentation, mice were awake and head fixed on top of a freely moving wheel, leaving them free to move their limbs while their heads remained stationary.

Surgical procedure

Mice were surgically implanted with a head bar to allow for head-fixed recordings. Animals were anesthetized with isoflurane through a nose cone on a stereotaxic apparatus. Bilateral craniotomies (AP, –1 to –2 mm from bregma; ML, ±2.9 to 4 mm from midline) and durotomies were performed to allow for acute recordings from the most posterior region of the dorsal striatum. Plastic wells were attached around each craniotomy and filled with a silicone elastomer (Sylgard 170, Dow Corning) to protect the surface of the brain and retain moisture when not recording. All animals were monitored after surgery and recovered fully before electrophysiological experiments.

Electrophysiological recordings and optogenetic stimulation

Animals were habituated to the sound presentation and the head-fixed apparatus for at least 1 h, 1 d before the first day of recording. Electrical signals were collected with an acquisition system (catalog # RHD2000, Intan Technologies) and Open Ephys software (www.open-ephys.org), using 32-channel silicon probes with electrodes arranged as tetrodes (A 4 × 2-tet configuration; NeuroNexus). The shanks of the

probes were marked with a fluorescent dye (DiI, catalog #V22885; or DiD catalog #V22887, Thermo Fisher Scientific) before penetration of the brain to assist in the identification of shank location postmortem. Before neural recordings, animals were head fixed, the silicone elastomer was removed, and the electrodes were inserted through the craniotomy. The probe was held in a vertical position and lowered 2.9 mm from the brain surface. We waited for at least 15 min for the probe to settle before initiating recordings. Neural recordings were performed at multiple depths on each penetration, with recording sites typically 100–150 μm apart to avoid recording from the same cells twice. Multiple penetrations were performed for each animal. A few recording sessions included only the presentation of one stimulus type [pure tones or amplitude-modulated (AM) noise], while the large majority of recording sessions included both.

A 50- μm -core diameter Polymicro optical fiber (part #1068001596, Molex) was attached to the silicon probe between the middle shanks and $\sim 200 \mu\text{m}$ above the top tetrode. The optical fiber was connected to a 445 nm laser calibrated to deliver 2 mW at the fiber tip. To identify D_1 -expressing neurons, we ran at least 50 laser stimulation trials before the sound presentation trials.

Estimation of recording location

At the conclusion of the experiments, animals were deeply anesthetized with euthasol and perfused through the heart with 4% paraformaldehyde (PFA). Brains were extracted and left in 4% paraformaldehyde for at least 24 h before slicing. Brain slices (thickness, 50 or 100 μm) were prepared under PBS using a vibratome (model VT1000 S, Leica) and imaged using a fluorescence microscope (Axio Imager 2, Carl Zeiss) with a 1.25 \times and 2.5 \times objective (numerical aperture, 0.16). To determine the locations of our recordings, we manually registered each brain slice containing dye fluorescence from a recording track to the corresponding coronal section in the Allen Mouse Common Coordinate Framework [Common Coordinate Framework version 3, 2015, Allen Institute for Brain Science; Allen Brain Atlas API (application programming interface), <http://brain-map.org/api/index.html>]. Recordings identified to be from the cerebral cortex were excluded from further analysis.

Characterization of ChR2-enhanced yellow fluorescent protein expression

Expression of ChR2-enhanced yellow fluorescent protein (EYFP) in the striatum of experimental animals was verified during the estimation of recording locations. To further evaluate the expression of ChR2-EYFP in the posterior striatum of $DRD1::ChR2$ mice at cellular resolution, brain slices from two more mice were prepared for confocal imaging. Animals were deeply anesthetized with euthasol and transcardially perfused with 4% PFA. Brains were extracted and left in 4% PFA for 24 h, then cryoprotected in 30% sucrose for 48 h. The brains were then sectioned (20 μm thick) using a cryostat (model CM3050 S, Leica), and brain slices (1.8–2 mm posterior from bregma) were mounted and DAPI stained. Images

were acquired using a confocal laser-scanning microscope (model LSM 880, Zeiss).

Data analysis

Spike sorting and selection of D_1 -expressing neurons

Spiking activity was detected by applying a threshold (40–45 μV) to bandpass-filtered electrical signals (300–6000 Hz) measured by the electrodes. The activity from single units was isolated offline using the automated expectation maximization clustering algorithm Klustakwik (Kadir et al., 2014). Isolated clusters were only included in the analysis if $<5\%$ of interspike intervals were <2 ms. We also calculated a spike quality index, defined as the ratio between the peak amplitude of the spike waveform and the average variance, calculated using the channel with the largest amplitude. Cells were only included in the analysis if they had a spike quality index >2 . The analysis also excluded clusters identified as having noisy spike waveforms from visual inspection. Last, only neurons that had a firing rate (either spontaneous or evoked) of at least 1 spike/s were included in the analysis. Cells with lower firing rates were excluded because we considered our measurements from these potential cells to be unreliable.

To identify D_1 -expressing neurons, the spike-sorting algorithm was applied to the combination of laser trials, trials with pure tones, and trials with AM sounds, such that each neuron could be identified by its spike shape across all stimulation conditions. Neurons were classified as D_1 -expressing if their onset response to laser stimulation (first 50 ms) was statistically larger ($p < 0.01$, Wilcoxon signed-rank test) than the baseline firing estimated from 200 ms before stimulus onset. We chose the first 50 ms (of the 100 ms laser pulse) to minimize confounds resulting from network recurrence. Neurons were classified as non- D_1 if the response to the laser was a suppression in firing or if the p -value associated with the laser response was >0.1 . Neurons with positive responses to the laser and p -values between 0.01 and 0.1 were excluded from the analysis. Changing criteria slightly (e.g., requiring non- D_1 neurons to have laser-evoked p -values >0.1) did not qualitatively affect the results. Additional comparisons between neuron classes were performed while restricting the non- D_1 population to only those neurons that were recorded from sites where D_1 cells were found. That is, if no D_1 neurons were observed on a tetrode during a session, cells from that tetrode were not included.

Estimation of frequency tuning and responses to pure tones

To determine whether a neuron was responsive to pure tones, we tested whether the evoked response during sound presentation (0–100 ms) was statistically different from the baseline spontaneous firing (measured during the 200 ms before sound onset) for any of the frequencies presented, collapsed across intensities. Because this test was performed for each of the 16 frequencies presented, we performed a Bonferroni correction for multiple comparisons (resulting in $\alpha = 0.05/16 = 0.0031$). A tone response index (TRI) was calculated for the sound frequency that elicited the largest change from baseline firing

for each neuron: $TRI = (r_e - r_b)/(r_e + r_b)$, where r_e is the average evoked response for that sound frequency and r_b is the baseline spontaneous firing. To characterize the sound frequency tuning of each neuron, we first determined the frequency response area (FRA), defined as the set of frequency–intensity pairs for which the response of the cell was greater than a response threshold set to the baseline firing rate plus 20% of the difference between baseline and the maximum firing rate of the cell under any condition (Ponvert and Jaramillo, 2019). The characteristic frequency (CF) of a neuron was defined as the frequency with the lowest sound intensity inside the FRA where 85% of the intensities above were also within the FRA. Neurons were categorized as having a steep slope in the high-frequency flank of the FRA if this slope (calculated from the border of the FRA at maximum intensity and the CF) was >100 dB/octave.

The tuning bandwidth of each neuron was first estimated by fitting a Gaussian function to the average firing rate evoked by each frequency, collapsed across sound intensities, and measuring the full-width at half-maximum of this function (which for a Gaussian corresponds to 2.355σ , where σ is the SD). Only neurons with R^2 values >0.01 were included in the comparisons of tuning bandwidth. For neurons with sufficiently low thresholds, we also estimated the bandwidth at 10 dB above threshold (BW10) and BW40. In these cases, the Gaussian function was fit to the responses at a particular intensity, and the method described above for estimating the full-width at half-maximum was applied.

To estimate differences in response dynamics to pure tones, we further analyzed neural responses from the same recording sessions used for estimating frequency tuning. We first estimated the latency of response of a neuron by pooling all trials with stimuli within its FRA, calculating a peristimulus time histogram (smoothed with a Hanning window 11 ms wide), and finding the time it took this signal to reach the midpoint between baseline and peak firing. This analysis was restricted to neurons with intensity thresholds <60 dB to get a sufficient number of trials, since the FRA of high-threshold neurons is generally very small. Neurons were also excluded if the automatic method yielded a negative response latency (e.g., because of a low firing rate). We also calculated an onset-to-sustained index (OSI) as $OSI = (r_o - r_s)/(r_o + r_s)$, where r_o is the average firing rate during the early part of the stimulus (0–50 ms) and r_s is the average firing rate during the late part of the stimulus (50–100 ms). This analysis focused on stimuli that fell within the FRA of each neuron.

Estimation of responses to AM sounds

To determine whether a neuron was responsive to AM noise, we tested whether the evoked response was statistically different from the baseline spontaneous firing (measured during the 200 ms before sound onset) for any of the AM rates presented. Because neurons in the auditory system often show substantially different responses at the onset versus the sustained periods of AM stimuli, we performed separate tests for each period: onset (0–100 ms) and sustained (100–500 ms). These tests were performed for each of the 11 AM rates presented; thus, we performed a Bonferroni correction for multiple comparisons

(resulting in $\alpha = 0.05/11 = 0.0045$). A sustained response index (SRI) was calculated for the AM rate that elicited the largest change from baseline firing for each neuron: $SRI = (r_e - r_b)/(r_e + r_b)$, where r_e is the average evoked response for that AM rate during the sustained period and r_b is the baseline spontaneous firing.

We estimated the rate modulation transfer function (Eggermont, 1994) for each responsive neuron by quantifying the average firing rate during the sustained response period (see Fig. 5A–D, middle panels, examples of transfer functions). AM rate selectivity was estimated by using an index that compared the maximum and minimum evoked firing rates during the sustained period across AM rates: $(r_{\max} - r_{\min})/(r_{\max} + r_{\min})$. To evaluate how well the evoked spikes synchronized to the modulation of the stimulus, we used a vector strength metric on the sustained response for each cell (see Fig. 5A–D, right panels, examples). We then performed a Rayleigh's test to determine the AM rates for which a neuron displayed statistically significant synchronization of its firing with respect to the phase of the amplitude modulation. This test included a Bonferroni correction for multiple comparisons ($\alpha = 0.0045$). For each neuron, we estimated the highest AM rate where the firing of the neuron was significantly synchronized to the modulation period.

Statistics

Throughout the study, we used nonparametric statistical tests implemented by the Python package SciPy (Virtanen et al., 2020). When comparing evoked firing rates to spontaneous rates (e.g., for responses to laser stimulation or sound stimulation), we used a test for related paired samples (Wilcoxon signed-rank test), where each trial provides one pair. When comparing measurements across two populations of cells, we used a nonparametric test for two independent samples (Mann–Whitney U rank test). The comparisons associated with the histograms (see Figs. 3I, 5E) were performed using the absolute value of the response index in each case, while the triangles indicate the median values separately for positive and negative responses. Further details on the comparisons made and statistical tests used are presented in Table 1.

Results

Distinct classes of posterior striatal neurons respond to sound stimuli

To determine whether the representation of sounds differed across neuron classes in the posterior tail of the striatum, we recorded sound-evoked responses of photo-identified D₁-expressing neurons and neighboring (non-D₁) neurons in this brain region from naive awake mice using silicon multichannel probes that have electrodes organized as tetrodes (Fig. 1A). Photo-identification of D₁ neurons during electrophysiological recordings was made possible by using DRD1::ChR2 mice, which express the light-gated ion channel ChR2 in D₁ neurons (Fig. 1B), and evaluating the spiking responses of each recorded neuron to laser stimulation delivered via an optical fiber attached to the recording probe. This method for *in vivo* identification

Table 1: Summary of statistical analyses

Figure	Measurement	Comparison	<i>N</i> cells	Statistical test	Results
	Fraction of sound-responsive cells	D ₁ cells vs non-D ₁ cells	D ₁ = 208 of 482 nD ₁ = 135 of 465	Fisher's exact test	$p < 0.0001$
	Fraction of sound-responsive cells	D ₁ cells vs site-restricted non-D ₁ cells	D ₁ = 208 of 482 nD ₁ = 79 of 197	Fisher's exact test	$p = 0.027$
3I	Tone response index	D ₁ cells vs non-D ₁ cells	D ₁ = 400 nD ₁ = 376	Mann-Whitney <i>U</i> rank test	$p = 0.0006$ <i>U</i> = 85880
	Tone response index	D ₁ cells vs site-restricted non-D ₁ cells	D ₁ = 400 nD ₁ = 188	Mann-Whitney <i>U</i> rank test	$p = 0.156$ <i>U</i> = 34873
	Baseline firing	D ₁ cells vs non-D ₁ cells	D ₁ = 400 nD ₁ = 376	Mann-Whitney <i>U</i> rank test	$p = 0.496$ <i>U</i> = 73073
	Best frequency	D ₁ cells vs non-D ₁ cells	D ₁ = 115 nD ₁ = 79	Mann-Whitney <i>U</i> rank test	$p = 0.915$ <i>U</i> = 4583
	Characteristic frequency	D ₁ cells vs non-D ₁ cells	D ₁ = 115 nD ₁ = 79	Mann-Whitney <i>U</i> rank test	$p = 0.41$ <i>U</i> = 4852
3J	Intensity threshold	D ₁ cells vs non-D ₁ cells	D ₁ = 115 nD ₁ = 79	Mann-Whitney <i>U</i> rank test	$p = 0.0015$ <i>U</i> = 5754
3K	Frequency tuning bandwidth	D ₁ cells vs non-D ₁ cells	D ₁ = 115 nD ₁ = 79	Mann-Whitney <i>U</i> rank test	$p = 0.086$ <i>U</i> = 3883
	BW10	D ₁ cells vs non-D ₁ cells	D ₁ = 97 nD ₁ = 70	Mann-Whitney <i>U</i> rank test	$p = 0.058$ <i>U</i> = 3981
	BW40	D ₁ cells vs non-D ₁ cells	D ₁ = 18 nD ₁ = 24	Mann-Whitney <i>U</i> rank test	$p = 0.809$ <i>U</i> = 206
3L	Fraction of steep slope cells	D ₁ cells vs non-D ₁ cells	D ₁ = 17 of 75 nD ₁ = 21 of 73	Fisher's exact test	$p = 0.574$
	Response latency	D ₁ cells vs non-D ₁ cells	D ₁ = 65 nD ₁ = 64	Mann-Whitney <i>U</i> rank test	$p = 0.656$ <i>U</i> = 1985
4G	Onset-to-sustain index	D ₁ cells vs non-D ₁ cells	D ₁ = 49 nD ₁ = 48	Mann-Whitney <i>U</i> rank test	$p = 0.256$ <i>U</i> = 1334
5E	AM sustained response index	D ₁ cells vs non-D ₁ cells	D ₁ = 475 nD ₁ = 460	Mann-Whitney <i>U</i> rank test	$p = 0.0013$ <i>U</i> = 121753
	AM sustained response index	D ₁ cells vs site-restricted non-D ₁ cells	D ₁ = 475 nD ₁ = 193	Mann-Whitney <i>U</i> rank test	$p = 0.212$ <i>U</i> = 42792
	Preferred AM rate	D ₁ cells vs non-D ₁ cells	D ₁ = 81 nD ₁ = 52	Mann-Whitney <i>U</i> rank test	$p = 0.744$ <i>U</i> = 2175
5F	AM rate selectivity	D ₁ cells vs non-D ₁ cells	D ₁ = 81 nD ₁ = 52	Mann-Whitney <i>U</i> rank test	$p = 0.578$ <i>U</i> = 2227
5G	Highest synchronization rate	D ₁ cells vs non-D ₁ cells	D ₁ = 53 nD ₁ = 30	Mann-Whitney <i>U</i> rank test	$p = 0.832$ <i>U</i> = 817

of a genetically defined neuronal populations has been extensively used and validated in several brain regions, including the striatum (Lima et al., 2009; Kravitz et al., 2013; Lakunina et al., 2020). Figure 1C shows an example striatal neuron that responds reliably to laser stimulation; the quick strong response after laser onset indicates that this neuron expresses D₁. This neuron also shows reliable responses to sound, a pure tone in this case (Fig. 1E). In the same brain region, we found neurons that showed no response to laser stimulation (Fig. 1D), but reliably respond to sound (Fig. 1F).

From our sample of cells, we identified 482 neurons as having statistically significant positive laser-evoked responses (Wilcoxon signed-rank test, $p < 0.01$) and therefore classified as D₁-expressing, and 465 classified as non-D₁ neurons. From these populations, we found that 43% of D₁ neurons and 29% of non-D₁ neurons showed a reliable evoked response to at least one sound in our stimulus ensemble, a mix of pure tones of different frequencies and amplitude-modulated noise at different modulation rates. This difference in the fraction of responsive cells from each class was statistically significant (Fisher's exact test, $p < 0.0001$), although it was much

less apparent when the analysis was restricted to the 197 non-D₁ cells that were recorded from sites where D₁ cells were found (43% D₁ vs 40% non-D₁ sound-responsive neurons; Fisher's exact test, $p = 0.027$). Figure 2 shows the estimated recording locations where we found sound-responsive neurons of each type, compared with recording locations where neurons were not sound responsive. There were no major differences between the locations of responsive and unresponsive neurons.

The observations described above indicate that subsets of medium spiny neurons in the posterior tail of the striatum that express the dopamine receptor D₁, as well as neighboring non-D₁ neurons, display reliable responses to sounds. We next wanted to test whether the sound-evoked responses and acoustic features encoded by these neurons differ between the two populations.

D₁-expressing neurons and their neighbors display similar sound frequency tuning

To test whether D₁ neurons in the posterior striatum encoded the frequency of sounds with different fidelity

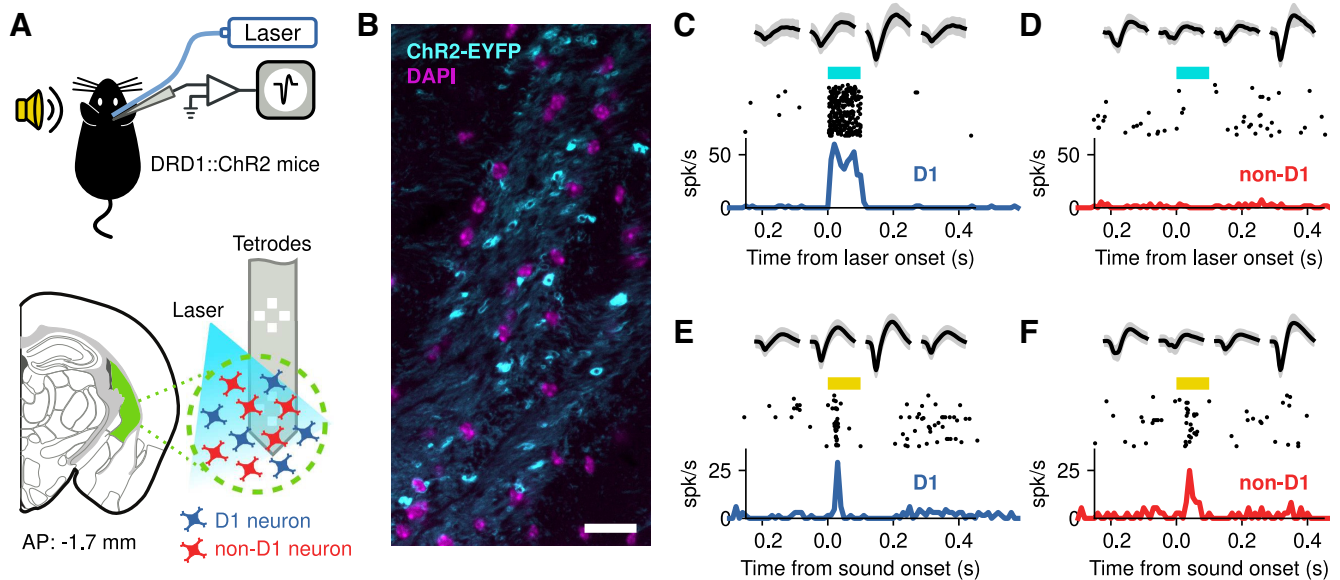


Figure 1. Photo-identification of D₁-expressing striatal neurons. **A**, Extracellular recordings of neurons from the posterior tail of the striatum (green area) of awake head-fixed DRD1::ChR2 mice during sound presentation. D₁-expressing neurons were identified during electrophysiological recordings by evaluating neural responses to blue laser stimulation. **B**, Expression of ChR2-EYFP in the posterior striatum of DRD1::ChR2 mice. Scale bar, 20 μm. **C**, Example responses to laser stimulation (cyan bar) from a striatal neuron. The top row shows the average (black) and SD (gray) of the spike shape collected from each channel of a tetrode in the silicon probe. The middle row shows the firing for each presentation of the laser, and the bottom row the peristimulus time histogram. The early and consistent response indicates that this cell expresses ChR2 and therefore is a D₁ neuron. **D**, Example of a cell that did not respond to the laser. Because of this, the cell is considered to be non-D₁. **E**, Firing of the cell in **C** evoked by a 9.9 kHz pure tone (yellow bar). Note that the spike shapes match those in **B**. **F**, Firing of the cell in **D** evoked by a 9.9 kHz pure tone. Spike shapes match those in **D**.

compared with other neurons in this region, we evaluated the evoked responses of identified D₁ and non-D₁ cells to pure tones of different frequencies and intensities. A total of 400 D₁ and 376 non-D₁ neurons were recorded during the presentation of pure tones. Figure 3A–D shows the average responses of example D₁

neurons for each frequency–intensity combination, demonstrating a clear tuning to specific frequencies and a dependence on sound intensity for these cells. As has been observed in other regions of the auditory system (Egorova et al., 2001; Palmer et al., 2013), we found neurons with different shapes of tuning, characterized

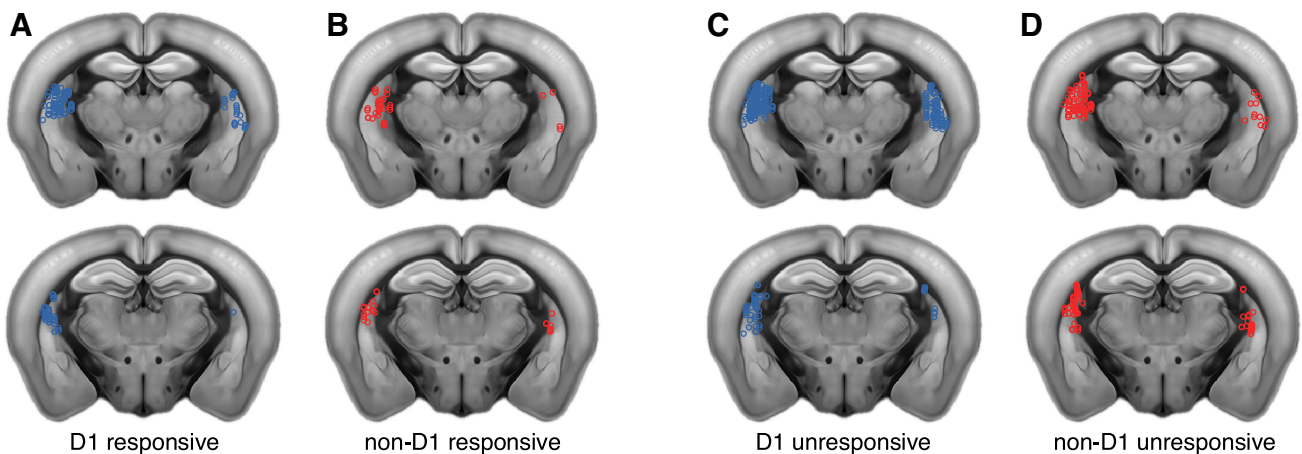


Figure 2. Location of recorded striatal neurons. **A**, Coronal slices showing sites in both hemispheres, where we found D₁ neurons that responded to at least one of the sounds in our ensemble (a combination of pure tones and amplitude modulated noise). Each circle represents one recording site. Sites are collapsed onto one of two anteroposterior locations shown: approximately –1.35 mm (top) and –1.75 mm (bottom) from bregma. The ventral region of the posterior tail of the striatum was not sampled in our experiments. **B**, Coronal slices as in **A** showing sites where we found sound-responsive non-D₁ neurons. **C**, Sites where we found D₁ neurons that did not respond to sounds. **D**, Sites where we found non-D₁ neurons that did not respond to sounds.

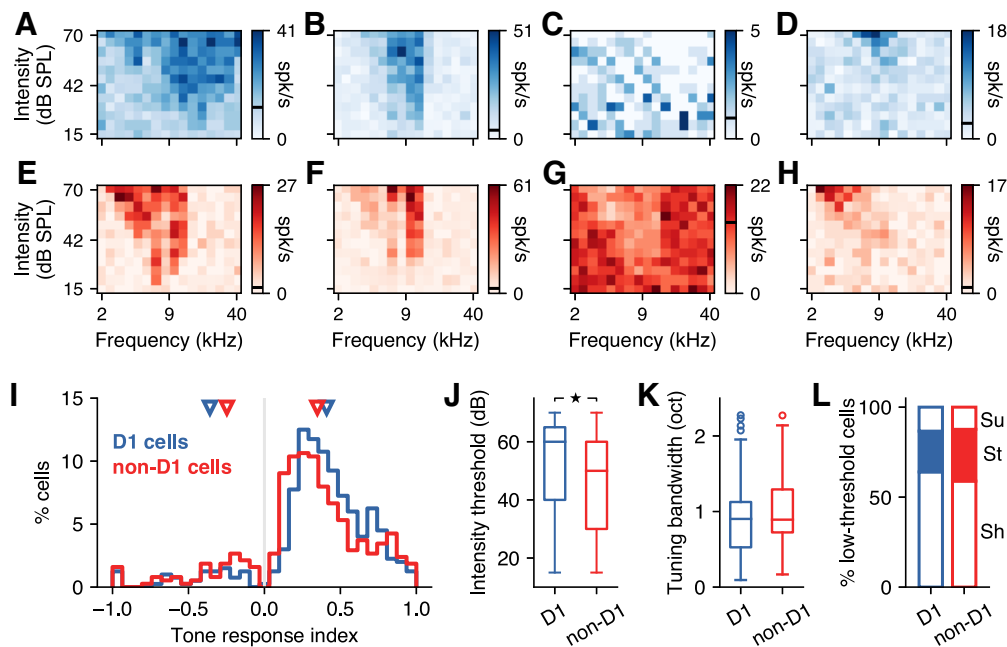


Figure 3. D₁-expressing striatal neurons have comparable sound frequency tuning to neighboring neurons. **A**, Example frequency–intensity tuning curve from a D₁ striatal neuron in response to 100 ms pure tones. The spontaneous firing rate is shown as a black line in the scale bar. **B**, D₁ neuron that has a much steeper slope at the high-frequency flank of the response area. **C**, D₁ neuron that responded to sounds by decreasing its firing. **D**, D₁ neuron with a very high-intensity threshold. **E–H**, Same as **A–D** for non-D₁ striatal neurons. **I**, Tone-evoked response index for the tone frequency that elicited the largest response in each neuron of each class. Triangles indicate the median response index calculated separately for neurons that responded by increasing or decreasing their firing. On average, D₁ neurons showed larger responses than non-D₁ neurons (Mann–Whitney *U* rank test, $p = 0.0006$). **J**, Intensity thresholds were slightly higher for D₁ neurons compared with non-D₁ neurons (Mann–Whitney *U* rank test, $p = 0.001$), but both cell types spanned the full range of possible thresholds. **K**, Frequency tuning bandwidth was similar between D₁ and non-D₁ neurons (Mann–Whitney *U* rank test, $p = 0.086$). **L**, The proportions of low-threshold cells with each shape of frequency response area were similar between D₁ and non-D₁ neurons. Sh, Shallow high-frequency slope (as in **A**); St, steep high-frequency slope (as in **B**); Su, suppressed response (as in **C**).

by the FRA. These included cells with a mostly symmetric FRA (Fig. 3A), cells with a steep slope in the high-frequency flank of the FRA (Fig. 3B), cells that responded by suppressing their firing (Fig. 3C), and cells with very high-intensity thresholds (Fig. 3D). All of these types of responses were also apparent in non-D₁ neurons (Fig. 3E–H).

To evaluate differences between D₁ and non-D₁ neurons, we first calculated a tone response index that compared the evoked response to the baseline firing for each neuron and plotted this index (Fig. 3I) for the sound frequency that elicited the most reliable evoked response in each neuron (pooled across intensities). The distribution of these best frequencies (range: D₁, 3.6–26.8 kHz; non-D₁, 3–26.8 kHz) was similar between the two populations of neurons (Mann–Whitney *U* rank test, $p = 0.9$). As expected, most neurons (even if they did not have a statistically significant response to sounds) had firing rates evoked by the best stimulus that differed from the spontaneous firing, and therefore show a response index different from zero (Fig. 3I). Positive values indicate that the evoked firing was larger than the spontaneous firing. Negative values, indicating a decrease in firing, were much less common (D₁ cells, 41 of 400 = 10.2%; non-D₁ cells, 59 of 376 = 15.7%). Cells that showed negative responses, as

those illustrated in Figure 3, C and G, had on average higher spontaneous firing rates than those with positive responses for both D₁ cells (median, 1.4 vs 0.88 spikes/s; Mann–Whitney *U* rank test, $p = 0.015$) and non-D₁ cells (median, 2.37 vs 0.77 spikes/s; Mann–Whitney *U* rank test, $p < 1e - 5$), although there was no difference in spontaneous firing rate between D₁ and non-D₁ cells (Mann–Whitney *U* rank test, $p = 0.496$).

Tone response index values largely overlapped across the two populations of cells, yet we found that the strongest sound-evoked changes in firing for D₁ neurons were larger than those for non-D₁ neurons (median absolute index: D₁, 0.41 vs non-D₁, 0.35; Mann–Whitney *U* rank test, $p = 0.0006$). Moreover, the median values (Fig. 3I, triangles) were larger for D₁ neurons that had positive evoked responses and smaller for neurons with negative evoked responses, compared with median values for non-D₁ neurons. However, this difference disappeared when we restricted our analysis to the 188 non-D₁ neurons recorded in sites (i.e., tetrodes) where D₁ cells were also found (median: D₁, 0.42 vs non-D₁, 0.44; Mann–Whitney *U* rank test, $p = 0.156$), suggesting that this result is influenced by the exact location of each recording site.

We next focused on the characteristics of the frequency response area of sound-responsive neurons. We found

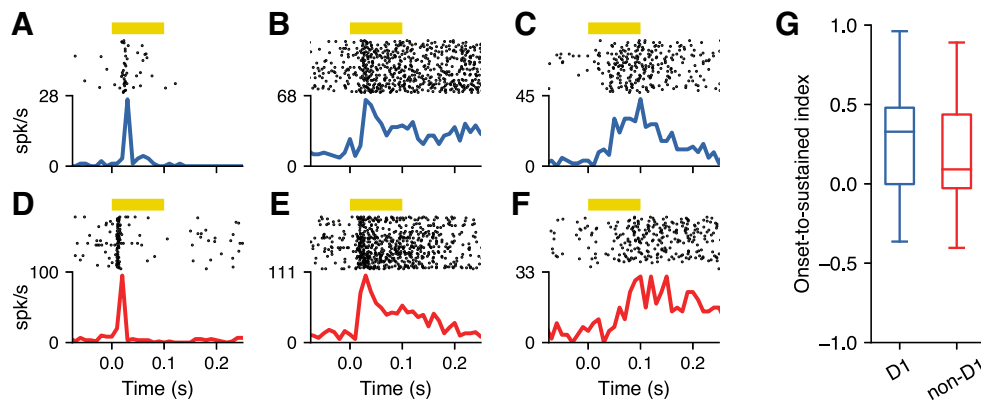


Figure 4. D₁-expressing striatal neurons have similar response dynamics to neighboring neurons. **A**, Example D₁ neuron showing an onset response to pure tones. Yellow bar indicates the duration of the sound. Trials include all frequency–intensity combinations that evoked a response. **B**, D₁ neuron showing a sustained response that continues beyond the offset of the sound stimulus. **C**, D₁ neuron showing a delayed response. **D–F**, Same as in **A–C** for non-D₁ neurons. **G**, The firing dynamics (a comparison between onset and sustained firing) for pure tones that elicited a response was similar between D₁ and non-D₁ neurons (Mann–Whitney *U* rank test, $p = 0.256$).

that the CF across neurons, which we observed in the range 3–22 kHz for D₁ and 2.4–18 kHz for non-D₁, was similar between the two populations ($p = 0.408$, Mann–Whitney *U* rank test). We then tested whether neurons were tonotopically organized and observed a small but statistically significant correlation between the CF of neurons and their location in the medial–lateral axis, similar for both populations (D₁: $r = 0.27$, $p = 0.0021$; non-D₁: $r = 0.33$, $p = 0.0059$; Spearman’s correlation), with medial neurons having higher CFs. We found no organization in either the dorsoventral or anteroposterior axes. We also found that while neurons from both classes spanned the range of tested intensity thresholds (Fig. 3J), D₁ cells had a higher threshold on average compared with non-D₁ cells (Mann–Whitney *U* rank test, $p = 0.001$). This difference was present even when restricting the set of non-D₁ cells to those from sites where we found D₁ cells (Mann–Whitney *U* rank test, $p = 0.001$). We found no topographic organization according to intensity threshold for either population. We then compared the tuning bandwidth of cells from each population. Because traditional measures of bandwidth (e.g., BW10 and BW40) are challenging to estimate for high-threshold neurons, we first evaluated tuning bandwidth by pooling responses over the full range of intensities. We identified cells for which responses across sound frequencies were well fit by a Gaussian function (115 of 166 tone-responsive D₁ neurons; and 79 of 104 tone-responsive non-D₁ neurons), and used the width at half-maximum of this curve as an estimate of tuning bandwidth (Fig. 3K). We found no significant difference between the frequency-tuning bandwidth across these populations of neurons (Mann–Whitney *U* rank test, $p = 0.086$). Similarly, we found no significant difference between cells with low enough intensity thresholds, which enable estimates of either BW10 (D₁ cells, 97; and non-D₁ cells, 70; Mann–Whitney *U* rank test, $p = 0.058$) or BW40 (D₁ cells, 18; non-D₁, 24 cells; Mann–Whitney *U* rank test, $p = 0.809$). Finally, we compared the proportion of low-threshold cells (<60 dB) with FRA shapes that matched

each of the categories presented in Figure 3, A–C and E–G. We found no major differences between these proportions (Fig. 3L). Specifically, the ratio of neurons with a steep high-frequency slope was comparable between D₁ and non-D₁ neurons (D₁, 22.7%; vs non-D₁, 28.8%; Fisher’s exact test, $p = 0.574$).

We then evaluated whether the dynamics of the responses to pure tones differed between D₁ and non-D₁ neurons. For each cell, we focused on stimuli within the FRA and measured how firing changed over time after the stimulus onset. Some D₁ neurons showed a brief response soon after the sound onset (Fig. 4A), other neurons showed sustained firing, sometimes extending beyond the end of the sound (Fig. 4B), while other neurons had a delayed response (Fig. 4C). We also found non-D₁ neurons with each of these characteristics (Fig. 4D–F). On average, we found no difference between the latency of responses from the two populations of neurons (D₁, 15.1 ms; non-D₁, 14.4 ms; Mann–Whitney *U* rank test, $p = 0.656$). To evaluate potential differences in response dynamics beyond the response latencies, we calculated an index that compares the magnitude of the onset response (0–50 ms) with the sustained part of the response (50–100 ms) for each neuron (Fig. 4G). We found no difference in this onset-to-sustained response index between the two populations (Mann–Whitney *U* rank test, $p = 0.256$).

These results suggest that the frequency tuning of D₁-expressing neurons in the posterior striatum is comparable with that of their neighboring neurons. We next wanted to evaluate the encoding of temporal acoustic features by these neuronal populations.

D₁-expressing neurons have responses to AM sounds comparable with those of their neighbors

To test whether D₁ neurons in the posterior striatum encoded temporal features of sounds with different fidelity compared with other neurons in this region, we evaluated

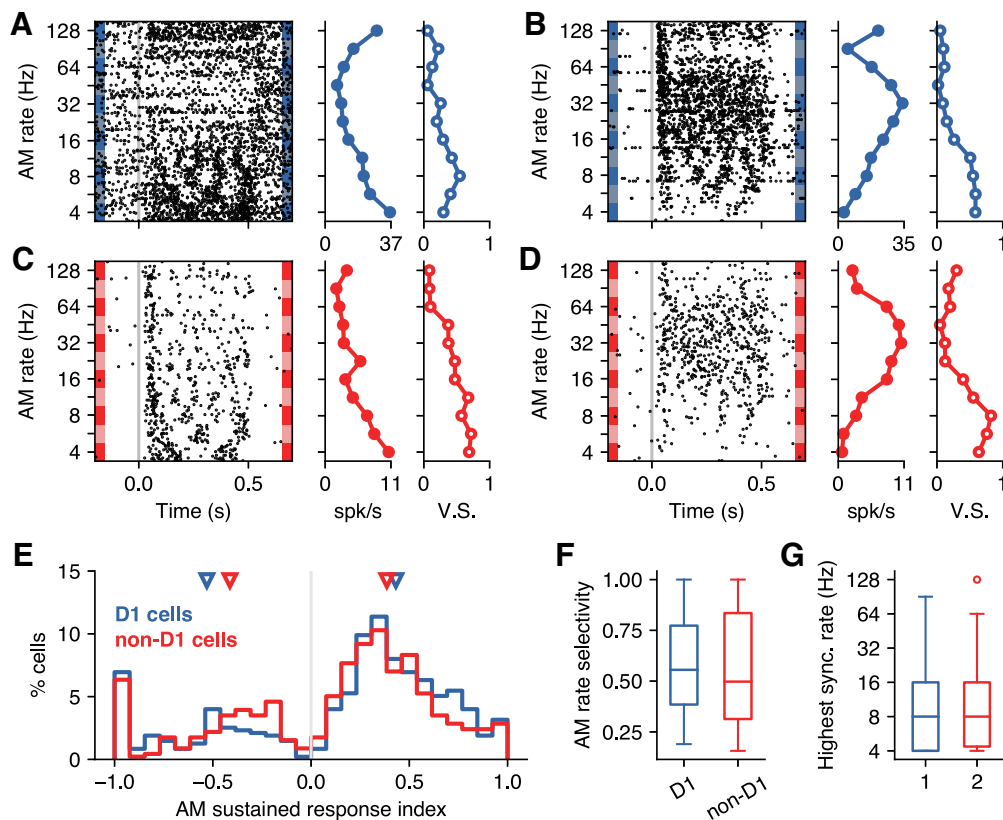


Figure 5. D₁-expressing striatal neurons display similar selectivity to AM sounds compared with their neighbors. **A**, Example responses of a D₁ neuron to 500 ms AM white noise at different modulation rates. Left, Neural firing on each trial. Middle, Average firing during the sustained period for each AM rate. Right, Vector strength for each AM rate, representing how well spikes synchronize to the stimulus. **B**, Responses to AM white noise for a different D₁ neuron. This neuron is tuned to an intermediate modulation rate, but still shows the highest synchronization to low AM rates. **C**, Responses of a non-D₁ neuron showing a high level of synchronization to the stimulus and tuning to low AM rates. **D**, Responses of a non-D₁ neuron showing tuning to an intermediate modulation rate. **E**, AM sound-evoked response index for the stimulus that elicited the largest response in each neuron of each class. Responses are calculated for the sustained portion of the response (100–500 ms). The triangles indicate the median response index calculated separately for neurons that responded by increasing or decreasing their firing. On average, D₁ neurons showed larger responses than non-D₁ neurons (Mann–Whitney *U* rank test, $p = 0.0013$). **F**, AM rate selectivity, calculated by comparing the strongest with the weakest sustained response across AM rates, was similar between D₁ and non-D₁ neurons (Mann–Whitney *U* rank test, $p = 0.578$). **G**, The highest AM rate that each neuron synchronized to was similar across cell classes (Mann–Whitney *U* rank test, $p = 0.832$).

the evoked responses of identified D₁ and non-D₁ cells to sinusoidally amplitude-modulated white noise at different modulation rates. A total of 475 D₁ and 460 non-D₁ neurons was recorded during the presentation of AM sounds. We found neurons from both classes that reliably responded to these sounds. As seen in neurons from other auditory regions, the responses of some cells were synchronized to the phase of the modulation, up to some modulation rate (Fig. 5A–D). While many neurons had the largest evoked firing for low modulation rates (Fig. 5A,C), other neurons were tuned to intermediate rates (Fig. 5B,D).

We calculated an SRI that compared the evoked response with baseline firing for each neuron and plotted this index for the modulation rate that elicited the most reliable evoked response in each neuron (Fig. 5E). For many neurons, the firing rate during the sustained portion was lower than the spontaneous firing (indicated by negative values of the index). Although the values of this index

largely overlapped across the two populations of cells, we found that the strongest sound-evoked changes in firing for D₁ neurons were larger than those for non-D₁ neurons (median absolute index: D₁, 0.47; vs non-D₁, 0.39; Mann–Whitney *U* rank test, $p = 0.0013$). The median values (Fig. 5E, triangles) were slightly larger for D₁ neurons that had positive evoked responses and smaller for neurons with negative evoked responses, compared with median values for non-D₁ neurons. This difference, however, disappeared when we restricted the analysis to the 193 non-D₁ neurons recorded from sites where D₁ cells were found (median: D₁, 0.47; vs non-D₁, 0.49; Mann–Whitney *U* rank test, $p = 0.212$), suggesting that this result is influenced by the exact location of each recording site.

We next compared the modulation rate selectivity of cells from each population (Fig. 5F), focusing on the sustained period of the response. Given the rate modulation transfer function for each cell (Fig. 5A–D, middle panels), we calculated an AM rate selectivity index by comparing

the maximum and minimum responses across modulation rates (including only neurons that were found to be responsive to AM sounds: D₁ neurons, 81; non-D₁ neurons, 52). We found no significant difference between the AM rate selectivity across these populations of neurons (Mann–Whitney *U* rank test, $p=0.578$). We also found that the most commonly preferred AM rate in both populations was the lowest rate tested (4 Hz: D₁ cells, 37%; non-D₁ cells, 42%) and found no difference in the distribution of preferred AM rate between D₁ and non-D₁ neurons (Mann–Whitney *U* rank test, $p=0.744$). Last, we tested whether neurons were organized topographically according to their AM rate preference, but found no correlation between the preferred rate and the neurons location for either D₁ or non-D₁ cells (D₁: $r < 0.2$, $p > 0.15$ in all directions; non-D₁: $r < 0.19$, $p > 0.32$ in all directions; Spearman's correlation).

We then evaluated how well the evoked spikes synchronized to the modulation of the stimulus, using a vector strength metric on the sustained response for each cell (Fig. 5A–D, right panels). We found that responses from both populations synchronized more effectively to lower AM rates, and we found no difference in the rate that elicited the highest synchronization for each neuron across the two populations (Mann–Whitney *U* rank test, $p=0.744$). We also evaluated the highest modulation rate at which responses from each neuron synchronized with the stimulus, using the Rayleigh test for periodicity (Fig. 5G) and including only cells that synchronize to at least one AM rate (D₁ neurons, 53; non-D₁ neurons, 30). We found no difference in the highest synchronized modulation rate between the two populations (Mann–Whitney *U* rank test, $p=0.832$).

These results suggest that responses to amplitude modulated sounds by D₁-expressing neurons in the posterior striatum are comparable with those of their neighboring neurons. Overall, the results above indicate that, in naive animals and outside the context of a task, different classes of neurons in the posterior tail of the striatum have access to and process acoustic features in a similar fashion.

Discussion

In this study, we quantified the activity of neurons from the posterior tail of the striatum in response to sounds of different frequency or temporal structure. Taking advantage of optogenetic methods for identifying genetically distinct cell types during extracellular recordings, we compared the representation of sounds by distinct classes of striatal neurons. Specifically, we focused on one of the major cell classes in the posterior striatum, namely one that expresses the dopamine receptor D₁ (and form the direct striatonigral pathway) and compared them to other neurons in the same striatal region. The cellular composition of the striatum is such that together D₁-expressing and D₂-expressing cells account for >90% of striatal neurons (Kreitzer and Malenka, 2008). While the proportions of these principal neurons in different subdivisions of the caudal striatum can vary greatly (Miyamoto et al., 2019), overall they are

present in similar quantities. Therefore, it is likely that most of our recorded non-D₁ neurons are in fact D₂-expressing indirect pathway neurons. The possibility of observing differences in the representation of sounds by these cell classes was motivated by previous studies that suggest differential innervation of striatal pathways by cortical neurons (Lei et al., 2004; Kreitzer and Malenka, 2007; Wall et al., 2013). We found, however, that the representation of spectral and temporal features of sounds by D₁-expressing neurons is similar to the representation of these features by neighboring neurons.

A common concern when applying methods for photo-identification during electrophysiological recordings is the possibility of observing laser-evoked responses from multi-synaptic indirect activation by other neurons. This concern is less relevant when identifying striatal D₁-expressing neurons as they are GABAergic in nature (and therefore their activation will result in the inhibition of their synaptic partners) and minimized by using only the early portion of the laser-evoked responses for identification (before recurrence can have a major impact). A more challenging limitation of these photo-identification methods results from possible changes in the spike shape if the photostimulation generates artificially large currents. Neurons in which this happens cannot be correctly identified by our algorithm and instead would be misclassified as D₁ cells that do not respond to sounds and non-D₁ cells suppressed by the laser. Visual inspection of the spike shapes and responses to laser stimulation across all cells suggests that these events were unlikely in our dataset.

Our results indicating that sound-responsive D₁ neurons encode acoustic features with similar fidelity to those of sound-responsive non-D₁ neurons was supported by multiple measurements, and it was robust to applying stricter criteria for inclusion of cells (e.g., using only cells with high firing rates to increase the reliability of response estimates; data not shown). Notably, comparisons across cell classes yielded slightly different results when we included all identified neurons or included only neurons recorded on the same electrodes. These observations cast doubt on the validity of differences observed across cell classes, as these effects could be explained by distinct levels of responsiveness by cells from different recorded locations. Variability in the recording locations across mice, together with the limited precision of our method for estimating the location of each recording make it impractical to derive further conclusions from our data regarding these potential differences. The one robust difference observed between the two cell populations was in their intensity threshold when presented with pure tones. In naive mice, we found that a larger proportion of D₁ cells had high thresholds compared with non-D₁ cells, although the responses evoked by the best tone stimuli for each neuron were comparable across the two populations. Whether this balance in thresholds changes as animals learn auditory tasks that result in striatal synaptic changes remains unknown.

Previous studies have observed strong selectivity to sound frequency in the responses of posterior striatal neurons of mice trained to perform auditory tasks (Guo et al., 2018; Chen et al., 2019). Our study complements

these observations by demonstrating that these neurons display robust responses to sounds even in naive mice. Moreover, our study illustrates that the responses of posterior striatal neurons can be synchronized to the amplitude modulation of the stimulus or be tuned to specific modulation rates, as observed in auditory thalamic and cortical neurons (Ponvert and Jaramillo, 2019). One of these earlier studies evaluated the evoked responses to pure tones from different classes of striatal neurons identified according to their spike shapes (Chen et al., 2019). They found that subsets of neurons from all identified classes (medium spiny neurons, cholinergic interneurons, and fast-spiking interneurons) displayed responses to tones with various dynamics and frequency tuning. Because the group of cells classified as medium spiny neurons contained both direct and indirect pathway neurons, their study could not derive conclusions regarding potential differences between cells from these two pathways. Our study complements these results by demonstrating that direct pathway neurons encode sounds in a similar fashion to their neighbors. Because an overwhelming majority of neurons in the striatum consists of projection medium spiny cells, split evenly between the two striatal output pathways (Kreitzer and Malenka, 2008), it is likely that this conclusion extends to the comparison between direct versus indirect pathway neurons.

Little is known about potential anatomic differences in the auditory projections to different striatal neuron classes. However, a previous study found differences in the innervation from nonauditory cortical areas to D₁- versus D₂-expressing neurons in the dorsal striatum (Wall et al., 2013). In that case, somatosensory cortex preferentially innervated D₁ neurons, while motor cortex preferentially innervated D₂ neurons. In contrast, that same study found that thalamostriatal projections were balanced between D₁ and D₂ targets. These results suggest that even if there is a preference in auditory cortical projections toward one striatal cell class, thalamic neurons could provide balanced information to D₁ and non-D₁ cells, as observed in our study. Moreover, because corticostriatal neurons differ in their representation of AM sounds compared with thalamostriatal neurons (Ponvert and Jaramillo, 2019), one would expect these differences to be reflected in the responses of D₁ and non-D₁ neurons if cortical neurons preferentially target one particular class. Our results suggest that this is not the case, and that both D₁ and non-D₁ neurons have access to similar cortical-like representations of AM sounds (e.g., a preference to intermediate AM rates in some neurons).

Our measurements also suggest a topographic organization of striatal neurons according to frequency preference. This observation, however, is at odds with results from tracing corticostriatal projections from regions of the primary auditory cortex tuned to difference sound frequencies in the mouse (Ghosh and Zador, 2021). While the tracing study found that fibers from the low-frequency region of A1 terminated more medially than those from the high-frequency region, we found that neurons located in the medial portions of the tail of the striatum preferred higher frequencies. These discrepancies could be explained by the fact

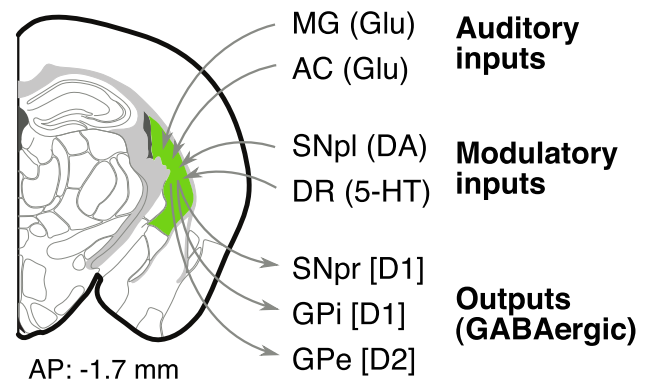


Figure 6. Inputs and outputs of the posterior tail of the striatum. The tail of the striatum (green) receives auditory glutamatergic inputs from the medial geniculate (MG) nucleus of the thalamus as well as primary and nonprimary fields of the auditory cortex (AC). It also receives dopaminergic (DA) inputs from the substantia nigra pars lateralis (SNpl) and serotonergic (5-HT) inputs from the dorsal raphe nucleus (DR). GABAergic medium spiny neurons that express either the dopamine receptor D₁ or the dopamine receptor D₂ form the main outputs of the tail of the striatum. D₁ neurons target the substantia nigra pars reticulata (SNpr) and the internal globus pallidus (GPi). D₂ neurons target the external globus pallidus (GPe). Based on the study by Valjent and Gangarossa (2021) and the Allen Mouse Brain Connectivity Atlas (Oh et al., 2014).

that the tail of the striatum receives projections not only from A1, but also from the auditory thalamus and from secondary areas of the auditory cortex (Fig. 6), which may provide frequency information organized in a different way. However, it is puzzling why projections from different regions would not be topographically aligned. Because our results come from pooling neurons from multiple antero-posterior locations and from different mice, it may be necessary to further validate these results by performing high-density recordings across the medial–lateral axis from a single mouse.

Neurons in the posterior tail of the striatum of mice are necessary for sound-driven tasks such as those that require associating a sound with a reward port (Guo et al., 2018). Moreover, the activation of D₁-expressing neurons in this region can consistently bias choices during these tasks without producing overt movements outside the task, in contrast to the movements generated by the activation of D₁ neurons in other striatal regions (Guo et al., 2018). This suggests that posterior striatal neurons play a role in sound-driven decisions beyond simply promoting or inhibiting movement. In addition, these neurons receive dopaminergic and serotonergic neuromodulatory inputs (Fig. 6), suggesting that they integrate sensory information with reward and other task-related signals. Our results show that both neuron classes that form the output pathways from the posterior striatum have access to detailed spectrotemporal acoustic features and could therefore potentially influence behavioral responses to sounds. As D₁- and D₂-expressing cells send their outputs to distinct regions of the basal ganglia (Fig. 6), which in turn have different effects on movements and choices, it seems likely

that the nervous system differentially adjusts the connections onto these pathways depending on the task at hand to implement distinct behavioral responses to different sounds. However, it remains unknown whether non-acoustic features that influence the activity of posterior striatal neurons, such as reward expectation and choice-related variables (Guo et al., 2019), are also represented similarly across cell classes, and whether differences in the representation of sounds emerge when investigating subgroups of neurons within each class (Gokce et al., 2016).

References

- Bordi F, LeDoux J (1992) Sensory tuning beyond the sensory system: an initial analysis of auditory response properties of neurons in the lateral amygdaloid nucleus and overlying areas of the striatum. *J Neurosci* 12:2493–2503.
- Chen L, Wang X, Ge S, Xiong Q (2019) Medial geniculate body and primary auditory cortex differentially contribute to striatal sound representations. *Nat Commun* 10:418.
- Eggermont JJ (1994) Temporal modulation transfer functions for AM and FM stimuli in cat auditory cortex. Effects of carrier type, modulating waveform and intensity. *Hear Res* 74:51–66.
- Egorova M, Ehret G, Vartanian I, Esser KH (2001) Frequency response areas of neurons in the mouse inferior colliculus. I. Threshold and tuning characteristics. *Exp Brain Res* 140:145–161.
- Gerfen CR, Engber TM, Mahan LC, Susel Z, Chase N, Monsma FJ, Sibley DR, Susel ZVI, Chase TN, Monsma FJ, Sibley DR (1990) D1 and D2 dopamine receptor-regulated gene expression of striatonigral and striatopallidal neurons. *Science* 250:1429–1432.
- Ghosh S, Zador AM (2021) Corticostriatal plasticity established by initial learning persists after behavioral reversal. *eNeuro* 8: ENEURO.0209-20.2021–10.
- Gokce O, Stanley GM, Treutlein B, Neff NF, Camp JG, Malenka RC, Rothwell PE, Fuccillo MV, Südhof TC, Quake SR (2016) Cellular taxonomy of the mouse striatum as revealed by single-cell RNA-Seq. *Cell Rep* 16:1126–1137.
- Guo L, Walker WI, Ponvert ND, Penix PL, Jaramillo S (2018) Stable representation of sounds in the posterior striatum during flexible auditory decisions. *Nat Commun* 9:1534.
- Guo L, Weems JT, Walker WI, Levichev A, Jaramillo S (2019) Choice-selective neurons in the auditory cortex and in its striatal target encode reward expectation. *J Neurosci* 39:3687–3697.
- Hintiryan H, Foster NN, Bowman I, Bay M, Song MY, Gou L, Yamashita S, Bienkowski MS, Zingg B, Zhu M, Yang XW, Shih JC, Toga AW, Dong HW (2016) The mouse cortico-striatal projectome. *Nat Neurosci* 19:1100–1114.
- Kadir S, Goodman D, Harris K (2014) High-dimensional cluster analysis with the masked EM algorithm. *Neural Comput* 26:2379–2394.
- Kawaguchi Y (1997) Neostriatal cell subtypes and their functional roles. *Neurosci Res* 27:1–8.
- Kravitz AV, Owen SF, Kreitzer AC (2013) Optogenetic identification of striatal projection neuron subtypes during in vivo recordings. *Brain Res* 1511:21–32.
- Kreitzer AC, Malenka RC (2007) Endocannabinoid-mediated rescue of striatal LTD and motor deficits in Parkinson's disease models. *Nature* 445:643–647.
- Kreitzer AC, Malenka RC (2008) Striatal plasticity and basal ganglia circuit function. *Neuron* 60:543–554.
- Lakunina AA, Nardoci MB, Ahmadian Y, Jaramillo S (2020) Somatostatin-expressing interneurons in the auditory cortex mediate sustained suppression by spectral surround. *J Neurosci* 40:3564–3575.
- Lei W, Jiao Y, Del Mar N, Reiner A (2004) Evidence for differential cortical input to direct pathway versus indirect pathway striatal projection neurons in rats. *J Neurosci* 24:8289–8299.
- Lima SQ, Hromádka T, Znamenskiy P, Zador AM (2009) PINP: a new method of tagging neuronal populations for identification during in vivo electrophysiological recording. *PLoS One* 4:e6099.
- Miyamoto Y, Nagayoshi I, Nishi A, Fukuda T (2019) Three divisions of the mouse caudal striatum differ in the proportions of dopamine D1 and D2 receptor-expressing cells, distribution of dopaminergic axons, and composition of cholinergic and GABAergic interneurons. *Brain Struct Funct* 224:2703–2716.
- Oh SW, et al. (2014) A mesoscale connectome of the mouse brain. *Nature* 508:207–214.
- Palmer AR, Shackleton TM, Sumner CJ, Zobay O, Rees A (2013) Classification of frequency response areas in the inferior colliculus reveals continua not discrete classes. *J Physiol* 591:4003–4025.
- Ponvert ND, Jaramillo S (2019) Auditory thalamostriatal and corticostriatal pathways convey complementary information about sound features. *J Neurosci* 39:271–280.
- Valjent E, Gangarossa G (2021) The tail of the striatum: from anatomy to connectivity and function. *Trends Neurosci* 44:203–214.
- Virtanen P, et al. (2020) SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat Methods* 17:261–272.
- Wall N, De La Parra M, Callaway E, Kreitzer A (2013) Differential innervation of direct- and indirect-pathway striatal projection neurons. *Neuron* 79:347–360.
- Zhong R, Qin L, Sato Y (2014) Auditory response properties of neurons in the putamen and globus pallidus of awake cats. *J Neurophysiol* 111:2124–2137.
- Znamenskiy P, Zador AM (2013) Corticostriatal neurons in auditory cortex drive decisions during auditory discrimination. *Nature* 497: 482–485.