Auditory corticofugal circuits drive vocalization-guided maternal behavior

Amy M. LeMessurier[1-5], Gurket Kaur[1-5], Julia I. Paraiso[1-5], Ayat A. Agha[1-5], Janaye Stephens[1-5], and Robert C. Froemke[1-5*]

[1] Skirball Institute for Biomolecular Medicine, New York University School of Medicine, New York, NY 10016, USA.

[2] Institute for Translational Neuroscience, New York University School of Medicine, New York, NY 10016 USA.

[3] Department of Otolaryngology, New York University School of Medicine, New York, NY 10016, USA.

[4] Department of Neuroscience, New York University School of Medicine, New York, NY 10016, USA.

[5] Center for Neural Science, New York University, New York, NY 10003, USA.

* Correspondence: robert.froemke@med.nyu.edu

Abstract

Perception of vocalizations is crucial for social behavior. A conserved example of this is mothers responding to distress calls from infants. In mice, experienced mothers (dams) find and retrieve isolated pups into the nest when pups emit ultrasonic vocalizations (USVs). Virgin females generally do not retrieve pups until they gain experience, for example by co-housing with a dam and litter. The onset of retrieval behavior is correlated with heightened sensitivity to USVs in left auditory cortex (AC). This plasticity may support learning via projections from cortex to early structures in the auditory pathway. To test whether projections from left AC are required for retrieval, we chemogenetically silenced activity in layer 5 during retrieval. In expert retrievers, silencing only neurons projecting to inferior colliculus (corticocollicular) led to impairment in retrieval. However, silencing neurons projecting to striatum had no effect. We used optically-tagged in vivo whole-cell recordings and 2-photon Ca^{2+} imaging in awake mice to compare encoding of USVs in corticostriatal and corticollicular neurons. Corticocollicular neurons in expert retrievers exhibited sustained increases in activity during USV playback compared to presentation of pure tones, while activity was equivalent during USV and pure tone presentation in corticostriatal neurons. The sustained activity we observed in corticollicular neurons may reflect increased excitability in a dedicated network of recurrently-linked cortical and subcortical areas. We tracked activity in corticocollicular and corticostriatal neurons over several days before and during co-housing as retrieval performance improved. This revealed robust population responses to USVs on each day in both groups. In both groups, delayed responses to USVs were larger in magnitude on days in which mice had reached expert performance, which may reflect network-level plasticity upregulating activity in recurrent auditory circuits to support perception of USVs.

Introduction

Perception of vocalizations is one of the most important functions of the auditory system in social animals. Accurately perceiving and responding to vocalizations requires segmentation and categorization of distinct vocal cues, as well as communications between auditory and motor planning and execution circuits during vocalization-guided behaviors[1]. In mammals, the central auditory pathway consists of both feedforward connections that progress from the ears to higher cortical areas, as well as descending cortical projections that innervate much of the feed-forward pathway, including brainstem[2,3], midbrain[4,5], and thalamus. Synapses from these top-down projections are numerous[6], and are key for predictive coding computations[1,7,8] that may support attending to and understanding speech[9]. In humans, the role of top-down modulation in predictive coding has primarily been studied within the cortex[1], but evidence from animal studies suggests that corticofugal projections to early areas in the auditory system can selectively process behaviorally-important sounds while suppressing noise, to drive representational plasticity enhancing discrimination/detection of the most important cues[10,11].

Vocalizations made by infants to indicate distress are common in altricial species, and survival depends on care-takers accurately identifying and responding to the category of call[12–14]. Young mouse pups emit a specific category of ultrasonic vocalizations (USVs) when they are isolated outside of the warmth of the nest. Experienced moms (dams) will respond to these calls by searching for and retrieving the vocalizing pup[15–18]. Retrieval can also be learned by virgins that have previously never interacted with pups by co-housing with a dam and litter[16–19], and previous work has identified plasticity in left auditory cortex representations of USVs as a key correlate of this learning[17,19–23]. How information in auditory cortex is transformed and communicated with motor circuits that directly control retrieval behavior is unknown.

We investigated the neural circuits through which auditory cortex facilitates behavioral responses to pup USVs in parental animals with a focus on projection neurons with downstream synaptic targets representing both feedforward and feedback auditory circuits. Previous studies have found that auditory cortex excitatory projections to the posterior tail of the striatum are involved in recognition of trained acoustic cues in auditory perceptual learning tasks[24], in which animals are trained to associate arbitrary sounds with reward delivery, and facilitation of auditory corticostriatal synapses has been correlated with learning in this context[25,26]. However, to our knowledge the involvement of this projection in auditory perception and learning has yet to be examined in social behaviors that depend on perception of vocalizations. Corticofugal projections, through which auditory cortex modulates activity in earlier steps of the central auditory pathway, have also been found to be involved in both trained and innate auditory-guided behaviors, for example, re-learning to localize sounds following partial hearing loss[10], appetitive perceptual training[27], and flight responses to looming sounds[28].

Thus, we focused our study on pup USV encoding in corticocollicular neurons projecting to inferior colliculus (IC) and corticostriatal neurons projecting to the posterior striatum, and tested the involvement of these projections in pup retrieval. We found that in experienced alloparents, corticocollicular, but not corticostriatal projections are needed for performance of pup retrieval, and that corticocollicular neurons are particularly modulated by USVs compared with corticostriatal over multiple durations relative to USVs. Measuring responses to USVs over alloparenting experience revealed facilitation and increased durations of call-evoked activity with improved behavior in both corticostriatal and corticocollicular projections.

Results

Chemogenetic suppression of left auditory cortex layer 5 and corticocollicular neurons alone, but not corticostriatal neurons alone, impairs pup retrieval in experienced females. To determine whether subcortically-projecting neurons in auditory cortex are required for pup retrieval, we assayed retrieval performance during chemogenetic suppression of layer 5 neurons (**Fig. 1a-e**), since the bulk of subcortically-projecting neurons arise in layer 5 [29]. A virus carrying cre-dependent inhibitory DREADDs receptor (AAV1-DIO-hSyn-HM4D(gi)) (**Fig. 1b**) was injected into left auditory cortex of 6 virgin female RBP4-Cre mice (GENSAT, UC-Davis), which express Cre selectively in layer 5 excitatory neurons. Retrieval performance was tested at least 2 weeks following surgery to allow for sufficient expression of the DREADDs construct. Mice were co-housed with wildtype (C57BL6j) dams with new litters beginning at postnatal day 0 or 1, and were cohoused for at least 3 days prior to pup retrieval testing. After confirming that co-housed virgins had become expert retrievers (retrieving pups on ≥80% of trials), we tested pup retrieval performance after i.p. administration of CNO or saline.

To test retrieval, mice were habituated in the testing arena with 2-3 pups situated in a nest in the corner of the arena for at least 30 minutes, following which single pups were introduced into a corner of the cage one at a time (**Fig. 1a**). On each trial mice were given up to 2 minutes to approach the pup and retrieve it to the nest, and this was repeated 10 times. Retrieval took place under dim red light to reduce reliance on visual cues. Successful retrieval trials often occurred after the onset of USV emission by the displaced pup, and the latency to retrieve the pup was negatively correlated with the frequency of USV emission during the trial (n=5 expert retrieving virgins) (**Fig. 1c,d**). Under saline conditions, RBP4-Cre experienced virgins retrieved pups on 88+-6% of trials (**Fig.**

114 **1e**), whereas on CNO trials, percent of pups retrieved decreased by roughly half to 47±12%
115 (p=0.03, Wilcoxon Rank Sum Test), indicating that pup retrieval depends on activity in left layer
116 5.
117
118 Neurons in layer 5 send corticofugal projections to areas earlier in the auditory hierarchy, as well
119 as feedforward projections to the striatum and amygdala[29] (**Fig. 1b**). Corticostriatal projections
120 from auditory cortex primarily innervate the posterior striatum, a sensory area that receives dense
121 input from auditory, visual, and somatosensory cortices, as well as limbic system inputs[30]. As the
122 striatum has been found to be involved in action selection and motor sequencing, and because the
123 auditory corticostriatal projection in particular has been implicated in behavioral responses to
124 acoustic cues in perceptual learning tasks[24–26], we hypothesized that this projection could serve as
125 a direct link between auditory perception of pup vocalizations and motor planning and execution
126 of retrieval. To test this hypothesis, we expressed HM4D(gi) selectively in auditory corticostriatal
127 neurons using an intersectional virus approach: cre was retrogradely expressed in these neurons
128 by injecting rgAAV-Cre into the left posterior striatum, and cre-dependent HM4D(gi)-mCherry
129 virus or just mCherry for sham controls was injected into left auditory cortex (**Fig 1b**).
130
131 Surprisingly, retrieval performance under CNO administration was unchanged relative to saline
132 on average in 7 experienced virgins and 3 dams (Retrieval probability: saline, 88+-3%; CNO, 75+-
133 9%, p=0.08, Wilcoxon's Rank Sum Test). Similarly, no effect was seen in sham-injected controls
134 (**Fig. 1e;** Retrieval probability, all shams: saline, 76±13%; CNO, 85±10%, p>0.05, paired T-test;
135 CS sham only: saline, 70±3%; CNO, 80±2%). Correct targeting of injections and spread of
136 expression were determined by visualizing mCherry in post-hoc histology for each mouse (**Fig.**

137 **1b**). While posterior striatum was correctly targeted in 100% of mice, spread of expression varied,
138 with several mice exhibiting expression in hippocampal CA3, which is likely due to leakage along
139 the injection track and spread of cre-dependent virus beyond cortex. This extended expression
140 beyond auditory corticostriatal neurons could explain the high variance in effect of CNO across
141 individual animals (**Fig. 1e**).
142
143 The lack of effect of silencing corticostriatal alone neurons suggested that a different subset of
144 layer 5 neurons could be responsible for the diminished performance we observed when silencing
145 left AC layer 5 more broadly. Corticofugal projection neurons from auditory cortex densely
146 innervate early areas in the central auditory pathway, including the medial geniculate body of the
147 thalamus (MGB), the inferior colliculus (IC), and auditory brainstem nuclei. We hypothesized that
148 auditory cortex may drive pup retrieval by modulating USV representations in earlier auditory
149 areas that project more directly to circuits that drive motor execution. IC is the first area of the
150 auditory pathway in which neurons extensively integrate acoustic input over time, and which
151 respond to USVs in multiple species. Thus, we assayed pup retrieval performance while
152 suppressing activity in auditory corticocollicular neurons using the same approach as described for
153 corticostriatal neurons, with rgAAV-Cre injected into left IC (**Fig. 1a,b**). Strikingly, CNO
154 decreased the percentage of pups retrieved in 5/6 experienced virgins and 1/3 dams roughly in half
155 (**Fig. 1e;** Retrieval probability: saline, 94±3%; CNO, 56±13%, p=0.03, Wilcoxon's Rank Sum
156 Test). No effect was seen in sham mCherry-injected controls (CC sham only: saline, 85±5%; CNO,
157 90±10%). Thus, corticofugal projections to inferior colliculus likely perform a critical function
158 supporting retrieval.
159

160 **Auditory corticocollicular neurons are particularly responsive to USVs in comparison with**
161 **corticocollicular neurons.** The differing effects of suppressing corticostriatal and
162 corticocollicular neurons during retrieval suggest that corticocollicular neurons in particular may
163 modulate perception of USVs in ways that facilitate retrieval performance. For example,
164 corticocollicular input could increase sensitivity to facilitate detection of calls, increase integration
165 windows in IC neurons to support categorization of calls based on temporal and acoustic features,
166 and/or facilitate neuromodulatory circuits that promote initiation of motor sequences. We reasoned
167 that if corticocollicular neurons support such computations in expert retrievers, a greater
168 proportion of them would be selective for USVs over other acoustic stimuli in comparison with
169 corticostriatal neurons.
170
171 We performed in vivo whole-cell and cell-attached recordings in corticocollicular and
172 corticostriatal neurons measuring responses to USVs and pure tones (**Fig. 2**). We identified
173 recorded neurons based on projection target by expressing channelrhodopsin in either projection
174 and testing responsiveness to blue light[31]. Channelrhodopsin-2 was expressed in CC or CS neurons
175 in experienced retrievers (3 dams and 18 cohoused virgins) via injection of rgAAV-ChR2-GFP in
176 IC or posterior striatum, respectively (**Fig. 2a,b**). Optic fibers were connected to a blue LED were
177 inserted into recording pipettes for light delivery to recorded neurons. Acute craniotomies were
178 performed over AC in mice headplated one week prior to recording. Electrodes were lowered to
179 400 µm to target layer 5, and once a seal was obtained, sensitivity to blue light was tested using
180 pulses of 20, 50, and 200 ms ranging intensity from 0 mW to 3 mW at the tip of the electrode.
181 Neurons were considered to be channelrhodopsin-expressing if they spiked within 10 ms of light

182 onset on 80% of trials, and if increasing durations of light led to increased duration of spike trains
183 (as opposed to suppression or rebound spiking after light offset).
184
185 Current-clamp or cell-attached recordings were obtained from 8 CC neurons from 7 virgins and 5
186 CS neurons from one dam and 4 virgins, as well as 73 additional neurons that were not light-
187 activated and considered unlabeled. We measured responses to a set of 6 prototypical pup isolation
188 USVs (**Fig. 2c**), each comprised of 4-5 individual syllables and of durations ~1 second, as well as
189 50 ms duration pure tone pups, in half-octave intervals from 4-64 kHz. Each stimulus was repeated
190 at least 10 times, and firing rates relative to stimulus times were averaged across trials (**Fig. 2d-f**).
191 We then determined the best USV and best tone for each recording: Best USV was defined as the
192 USV evoking the highest magnitude change in firing rate in the first 2 seconds following USV
193 onset compared to baseline rate quantified in the 1.5 seconds preceding. Best tone was defined
194 similarly, but firing rates were quantified in the first 200 ms following tone onset for evoked
195 change in firing, and 200 ms for the preceding baseline period. USV-evoked firing rates varied
196 across the distribution of labeled and unlabeled cells from -0.2 to 5.2 Hz with a median of 0.41(**Fig.**
197 **2e**). Tone-evoked firing rates varied from 0 to 22.7 Hz, with a median of -0.13 (**Fig. 2f**). Mean
198 firing rates for both USVs and tones were higher in CC neurons compared to tones (**Fig. 2e-f,**
199 USVs: 1.0±0.4Hz in CC, 0.4±0.2 Hz in CS, $p<0.05$, Students T-test; tones: 2.9±2.1 Hz in CC,
200 0.6±0.7Hz in CS, $p<0.05$, Students T-test).
201
202 While mean rates were higher in CC neurons for both types of stimulus – reflecting the top 22nd
203 and 15th percentiles of all recorded neurons for USVs and tones, respectively – firing rates varied
204 greatly across these cells, ranging from the 90th to 10th percentiles of USVs responses and 55th to

205 5th percentiles of tone responses. Given the relatively low sample sizes that can be obtained with
206 single-neuron recordings, we pursued in vivo 2-photon calcium imaging in labeled auditory cortex
207 projection neurons to obtain more highly-sampled estimates of these two populations'
208 representations of acoustic stimuli (**Figs. 3,4**). Additionally, we wondered how representations of
209 pup isolation USVs vary over time as virgins gain experience with alloparenting. Longitudinal 2-
210 photon calcium imaging enabled us to track responses in individual neurons and populations over
211 the course of co-housing (**Figs. 3a-c, 4**).
212
213 We expressed jGCaMP7f or 8f in CC neurons (8 mice) or CS neurons (6 mice) by injecting
214 rgAAV-hSyn-jGCaMP7f or 8f in left IC (**Fig. 3b**) or left posterior striatum (**Fig 3a**), respectively.
215 During the same surgeries, chronic cranial windows were implanted over left AC and steel
216 headbars were attached to the skull to enable head stabilization during imaging. Imaging
217 experiments measuring USV and tone-evoked activity were performed 3-5 weeks following
218 surgery. Following imaging experiments, brains were sectioned and imaged to visualize injection
219 targeting and expression in auditory cortex. In all CC and CS mice, expression in cell bodies was
220 most dense several hundred μm from the pial surface of the brain (**Fig. 3a,b**), with the bulk of
221 labeled neurons exhibiting morphology matching that which has been described for L5 excitatory
222 neurons; in particular, many labeled neurons in both groups had large apical dendrites protruding
223 dorsally from somas, with the highest density of axonal branching appearing to be in superficial
224 layers of cortex. In corticostriatal neurons, GCaMP-expressing cell bodies were also seen in more
225 superficial layers, in accordance with studies finding corticostriatal intratelencephalic (IT) neurons
226 across multiple layers that project to more anterior striatum and reports of long range inhibitory

227 interneurons that also project to striatum. In CC, some expression was observed in somas deeper
228 than the band matching L5, which were likely in L6 in accordance with other studies.
229
230 We first investigated representations of USVs and tones in expert retrieving virgins that had been
231 cohoused for several days. At least one day following habituation to head-fixation, awake mice
232 were stabilized under a 2-photon microscope, and auditory stimuli were played through an
233 ultrasonic speaker as we imaged in left auditory cortex (**Fig. 3 c,d**). A set of 6 USVs were played
234 in two blocks of ten repetitions per stimulus, with an interstimulus interval of 5 seconds, as
235 described for single cell recordings. Between the blocks of USVs, a block of half-octave spaced
236 pure tone pips of 250 ms duration were played at 70 dB SPL, with 10 repeats per tone at an
237 interstimulus interval of 3 seconds (**Fig. 3d**). Fluorescence time series were extracted from ROIs
238 representing somas following 3 days of co-housing with a dam and litter and daily retrieval testing,
239 or for a subset of mice in which we tracked the same imaging fields throughout cohousing, ROIs
240 were examined on the final day of imaging.
241
242 In CC neurons, we observed striking increases in the frequency and amplitude of calcium transients
243 during blocks of USVs that were not closely time-locked to the 5 second-spaced stimuli (**Fig. 3e**).
244 However, during blocks of pure tones, we did not observe similar levels of spontaneous activity
245 (quantified as mean dF/F over entire blocks of each trial type, **Fig. 3e-g,** green symbols; mean dF/F
246 in CC neurons during USV blocks: 0.03±0.001%; during tone blocks: 0.02±0.001%; $p<0.01$,
247 Kruskal-Wallis with Tukey's HSD correction for multiple pairwise comparisons). This apparent
248 difference in non-time-locked activity could be due to different total durations of the types of
249 acoustic stimuli, since each USV was comprised of multiple syllables of roughly 75-200 ms each

250 whereas each tone was a single 250 ms pip. However, we did not observe a difference in non-time-
251 locked dF/F in CS neurons between USV and tone blocks (0.008±0.001%; during tone blocks:
252 0.009±0.001%; $p>0.05$, Kruskal-Wallis with Tukey's HSD correction for multiple pairwise
253 comparisons), suggesting that excitability in CC neurons in particular might be selective for USVs
254 compared to tones in expert retrieving females. In accordance with single-neuron recordings
255 described above (**Fig. 2**), overall mean dF/F was higher in CC neurons compared to CS during
256 both blocks of USVs and tones (**Fig. 3f,g**; USV blocks, CC vs CS: $p<0.001$; tone blocks, $p<0.001$,
257 Kruskal-Wallis with Tukey's HSD correction for multiple pairwise comparisons).
258
259 Next we examined time-locked dF/F relative to stimuli in both groups. (**Fig. 3h-n**). For each cell,
260 dF/F was quantified for each stimulus repetition relative to the onset of the stimulus, using a pre-
261 stimulus baseline period as baseline fluorescence (F_0). Both CC and CS neurons exhibited a diverse
262 array of temporal response profiles and tuning for both calls and tones (**Fig. 3h,i)**, including
263 neurons that were relatively selective for individual calls or a narrow band of tone frequencies,
264 were responsive to calls but not tones or vice versa, were reliably activated by every USV, or were
265 suppressed by a subset of USVs. Surprisingly many neurons were reliably activated by USVs at
266 long delays, up to several seconds post USV onset. Thus, for each neuron we quantified 3 measures
267 of evoked dF/F: early responses to the best USV and tone for each neuron (**Fig. 3j**), defined as the
268 USV and tone evoking the largest magnitude mean dF/F in the first 2 seconds following stimulus
269 onset; the best late response **(Fig. 3k**), determined by the same procedure, except that mean dF/F
270 was computed from 2-5 seconds following stimulus onset; and most suppressed response-evoking
271 stimuli (**Fig. 3 l**), for which mean dF/F was computed in the first 3 seconds following stimulus
272 onset, and the greatest magnitude negative dF/F was considered the most suppressed response. The

273 best late response was only computed for USVs, because tones were played at an inter-stimulus
274 interval of 3 seconds, and the delayed/late responses we observed following USVs occurred
275 reliably up to 5 seconds post stimulus onset for some neurons.
276
277 By all three evoked response metrics, CC responses were larger magnitude on average over the
278 populations. (Best call: CC 2.9±0.2%, CS 2.1±0.1%, $p<0.001$; Late best call: CC 4.1±3%, CS
279 2.2±0.1%, $p<0.001$; Most suppressed call CC -3.3±0.1%, CS -1.8±0.1%, $p<0.001$; Best tone: CC
280 5.1±0.3%, CS 4.3±0.4%, $p<0.001$; Most suppressed tone CC -3.5±0.3%, CS -1.6±0.1%, $p<0.001$;
281 Kruskal Wallis test with Tukey's HSD correction for multiple pairwise comparisons).
282
283 In both CC and CS cells, responses to best tones and most suppressed tones were larger magnitude
284 than best and most suppressed USVs (Best USV vs tone: CC $p<0.001$, CS $p<0.001$; most
285 suppressed USV vs tone: CC $p<0.01$; CS $p<0.001$). This was in contrast with our finding that in
286 CC cells overall mean dF/F during blocks of calls was greater than during blocks of tones, and that
287 in CS cells, overall mean dF/F during blocks of calls and tones were not different. This suggests
288 that highly stimulus-locked responses may be independent of longer-duration modulation of
289 activity levels. Neither facilitate nor suppressed response magnitudes were correlated between
290 USVs and tones for either group (**Fig. 3m,n**), in agreement with previous studies finding no
291 relationship between auditory cortex neuronal tuning for vocalizations and pure tones. This
292 indicates that cortical representations of vocalizations are not simply predicted by frequency
293 tuning.
294

295 **Responses to USVs, but not pure tones, are plastic over alloparenting experience in both**
296 **corticocollicular and corticostriatal neurons.** The differences in sensitivity to USVs that we
297 observed in CC and CS neurons in co-housed virgins could be innate or experience-dependent,
298 reflecting plasticity in specific corticofugal circuits as female mice gain parenting experience. We
299 tracked activity in CC and CS neurons in virgins (4 virgins with CC cells labeled, 4 CS) before
300 and during co-housing with longitudinal 2-photon calcium imaging (**Fig. 4**). Using the same
301 imaging approach as described for experiments in experienced virgins (**Figs. 3a-d, 4a**), we
302 measured responses to USVs and pure tones in virgins daily, starting before cohousing and for 2-
303 4 days following (**Fig. 4a,b**). Retrieval performance was tested daily after imaging **(Fig. 4b,c**). In
304 each mouse the same imaging field was imaged each day, using blood vessels and ROIs as
305 landmarks to manually match locations (**Fig. 4b**), and ROI masks were considered to belong the
306 same neuron on multiple days if the areas across each day were at least 60% overlapping. Data
307 were excluded if changes in window clarity (for example, due to bone re-growth or inflammation
308 under the window) impacted the ability to image previously-imaged ROIs. In each mouse, 1-2 450
309 μm^2 fields were tracked. Pup retrieval performance improved over days of co-housing relative to
310 day 0 (pre co-housing) in all but one virgin, who retrieved 100% of pups on each day of imaging
311 (**Fig. 4c**). In most mice, performance reached 80% or higher by the 3rd day following the start of
312 co-housing.
313
314 In each mouse, USV-evoked dF/F magnitudes and tuning for individual calls were highly variable
315 in tracked neurons (**Fig. 4d**). The majority of ROIs that were detected on multiple days were not
316 present on every day of imaging, and most ROIs with large time-locked responses to specific calls
317 on one day did not maintain the same response magnitude or USV preference across all days.

318 However, on each day of imaging, both facilitated and suppressed responses were observed for
319 each stimulus, suggesting that despite individual variability over time, responsiveness to USVs
320 was maintained at the population level.
321
322 We quantified best early, late, and suppressed responses to USVs (**Fig. 4e-g**) and tones (**Fig. 4h-**
323 **j**) for every ROI that was imaged across at least two days of imaging. To determine whether these
324 metrics were dependent on experience with parental care of pups in CC and CS cells, we compared
325 responses between days with low retrieval performance (days in which ≤40% of pups were
326 retrieved) and high performance (≥80% retrieved). In CC neurons, early best call and late best call-
327 evoked dF/F were larger in magnitude on high-performance days than on low performance days,
328 while suppressed responses were equivalent (Early: $p<0.001$; late: $p<0.001$; suppressed: $p>0.05$;
329 Kruskal Wallis test with Tukey's HSD correction for multiple pairwise comparisons). In CS
330 neurons, early best call responses and suppressed responses were equivalent on low and high
331 performance days (Early: $p>0.05$; suppressed: $p>0.05$), but magnitudes were greater on high
332 performance days for late best call responses (late: $p<0.001$). These results suggest that increased
333 firing rates at a delay following USVs and/or longer durations of USV-evoked increases across in
334 neurons projecting to multiple targets may facilitate perception of USVs in an experience-
335 dependent manner. No differences across low and high retrieval days were observed for tone-
336 evoked activity (**Fig. 4h-j**), indicating that these effects are specific to USVs.

Discussion

Overall, our findings indicate that activity in auditory corticofugal projection neurons is particularly important for successful pup retrieval in experienced alloparental females, but that call encoding appears plastic with respect to learning in both corticocollicular and corticostriatal neurons, suggesting that exposure to pups results in plasticity in multiple cortical output circuits.

Our findings from chemogenetic inactivation of layer 5, corticocollicular, and corticostriatal neurons during retrieval suggest that the corticocollicular circuit in particular is crucial for successfully responding to pup calls. However, it is possible that due to our labeling strategy, more corticocollicular neurons were labeled than corticostriatal, representing a larger percentage of the overall layer 5 excitatory population. If this is the case, the difference in effects could be due to more complete suppression of layer 5 in the corticocollicular mice. Additionally, individual layer 5 projection neurons can send collateral axons to multiple downstream targets, so there is likely some degree of overlap in the two populations we labeled. Future experiments with dual injections will quantify the degree of overlap. Additionally, projecting neurons in these populations can target MGB, amygdala, and auditory brainstem nuclei, so it possible that some of the effects of suppression are mediated by these synapses.

While suppressing corticofugal projection neurons during the entirety of a retrieval testing session clearly suppresses performance, it is unclear what aspect of behavior is affected. Because the circuits we manipulated are primarily sensory, we hypothesize that the reduction in performance was due to impaired perception rather than motor output. If perception is indeed impaired, this could involve overall reduced detection of pup USVs, inability to properly categorize and assign

360 significance to calls, impaired sound localization, or decreased auditory working memory. Follow-
361 up experiments optogenetically suppressing activity during specific phases of behavior, for
362 example before the mouse leaves the nest or during the search for isolated pup, could distinguish
363 between these possibilities.
364
365 In experienced virgins, corticocollicular neurons were particularly modulated by USVs compared
366 to corticostriatal. Most strikingly, overall activity levels appeared to upregulated during blocks of
367 USV delivery independent of time-locked, directly call-evoked responses, as well as at long delays
368 relative to call onset and offsets. This may reflect a network-level increase in activity that could
369 involve local microcircuits within auditory cortex, and/or an overall amplification of USV-driven
370 activity in recurrently-linked auditory areas, including feedback loops between auditory cortex,
371 IC, and MGB. The large number of neurons that were suppressed by USVs suggests recruitment
372 of inhibitory circuits within cortex, while increased magnitude and duration of activity in
373 corticocollicular neurons likely modulates IC outputs, as well as temporal integration of
374 feedforward synaptic input to IC from auditory brainstem. Such computations could support
375 recognition and categorization of USVs by enabling integration of acoustic features of individual
376 syllables over the duration of multisyllabic calls, which can last seconds. Alternately, upregulation
377 of activity in auditory cortex could modulate windows for auditory working memory, over duration
378 needed for mice to localize pups. Future experiments recording ensemble activity across multiple
379 stages of the auditory pathway during USV playback, as well as in freely-moving animals during
380 retrieval, could further elucidate the relevance of auditory corticofugal modulation for perception
381 of vocalizations and linking of perception with action.

382 **Methods**

383 **Animals**

384 Experiments were performed on wild-type virgin female C57BL/6J mice, or in the case of some
385 behavior and chemogenetics assays, RBP4-Cre heterozygous virgin females. For all experiments,
386 mice were between the ages of 5 and 12 weeks old. Virgins were cohoused with wild-type
387 C57BL/6J dams with litters beginning at P0. Timed-pregnant dams were purchased (Taconic,
388 Charles River) and arrived between E10-E17. All experiments and animal use were performed
389 following NYU Medical IACUC and NIH guidelines for the proper use of animals in research.

390

391 **2-photon in vivo calcium imaging**

392 *Cranial window implant and viral transfection surgeries:* Survival surgeries were performed using
393 antiseptic surgery technique to place cranial windows over auditory cortex and to inject viral
394 vectors for expression of jGCaMP8f. Briefly, mice were anesthetized with 1.5-2% isoflurane
395 vaporized in oxygen delivered at a rate of 2 liters/minute. Anesthesia was monitored by regularly
396 checking breathing rate and foot-pinch reflex throughout surgery, and mice were kept at an internal
397 temperature of 37° C with a heating pad. Mice were stabilized with a bite bar and nose cone
398 attached to a stereotax. At the onset of surgery buprenorphine (0.1 mg/kg) and dexamethasone (2
399 mg/kg) were administered intraperitoneally to manage pain and inflammation. Prior to making an
400 incision, the scalp was trimmed of hair and topical lidocaine (2%) was applied. The scalp was
401 cleaned with iodine and 70% ethanol according to proper sterile protocols, and a patch of the scalp
402 was removed to expose the skull. The skull was cleaned of connective tissue and the left temporalis
403 muscle was separated from the skull to expose auditory cortex. Small craniotomies were performed
404 targeting inferior colliculus or posterior striatum using stereotactic coordinates (IC: -5 mm

405 posterior to bregma, 0.8 mm lateral, 0.7 m ventral; PS: -1.7 mm P, 3.4 mm L, 2.5 mm V) and
406 retrograde AAV-Syn-jGCaMP8f-WPRE virus (Addgene) was injected using a Nanoject. Auditory
407 cortex was located using skull and blood vessel landmarks, and a 3 mm area of skull was removed.
408 Dura was left intact, and a circular glass window (FST/Harvard apparatus, thickness 0.1 mm) was
409 applied to the brain surface and fastened to the skull using a combination of cyanoacrylate glue
410 and dental cement. A custom stainless steel headplate was attached to the skull using metabond
411 and all exposed skull was covered. Mice recovered from anesthesia on a heating pad and were
412 singly-housed for 24 hours with access to wet food. Additional doses of dexamethasone were
413 administered daily for 2-5 days following surgery to facilitate healing.
414
415 *Imaging:* Imaging experiments began 3 to 4 weeks after surgery using a 2-photon microscope with
416 resonant galvo scanners and a 20x Nikon objective (Moveable Objective Microscope, Sutter
417 Instruments). Mice were stabilized for imaging by clamping the headplate to a custom holder
418 (Thorlabs parts) and positioning the mouse in a length of acrylic tube. Mice were acclimated to
419 head-fixation daily for several days prior to imaging for increasing durations starting with 5
420 minutes. GCaMP-expressing neurons were imaged with an excitation wavelength of 900 nm, and
421 epifluorescence was bandpass filtered. 2D fields of corticostriatal projecting neurons were imaged
422 at depths ranging from 250 to 500 µm ventral to the dura, while corticollicular fields ranged from
423 350 to 800. Image stacks were collected using ScanImage (Vidrio Technologies) at a rate of 30
424 frames/second for 512x512 pixel images. Imaging data were collected over multiple days
425 beginning prior to the start of co-housing, with identical stimulus sets delivered on each day of
426 imaging. Auditory stimuli were played using an ultrasonic speaker (TDT) positioned 10 cm from
427 the right ear of the mouse.

428 **In vivo whole-cell and juxtacellular recordings**

429 In vivo recordings were performed in 21 female WT mice (3 dams and 18 experienced virgins
430 confirmed to be expert retrievers). One week prior to recording, stainless steel headplates were
431 implanted as described above for cranial window surgeries. A patch of skull covering left auditory
432 cortex was left exposed. For awake recordings, mice were habituated to head fixation daily for
433 three days prior to recording for progressively longer durations, up to one hour the day before
434 recording. On the day of recording, mice were anesthetized with isoflurane (1.5-2% in 2 liter/min
435 oxygen) and a small craniotomy (<2 mm) was performed over auditory cortex. The craniotomy
436 was covered in a layer of sterile saline, then a glass coverslip was temporarily fixed over the
437 craniotomy using kwik-sil. Mice were allowed to recover fully from anesthesia prior to head
438 fixation and recording. For anesthetized recordings, mice were anesthetized prior to craniotomy
439 surgery with an initial dose of ketamine (70 mg/kg) and xylazine (10 mg/kg) delivered via
440 intraperitoneal injection. Additional doses of ketamine were given as needed to maintain
441 anesthesia throughout the craniotomy and recording session (approximately every 30 minutes).
442 Anesthesia level was monitored via tail-pinch reflex and breathing rate measurement.

443

444 *Viral injections:* For optotagging experiments, 2-3 weeks prior to recording, AAV1-FLEX-ChR2-
445 GFP was injected in auditory cortex (targeted as described for GCaMP injections) and retro-AAV-
446 Cre was injected in either left posterior striatum or left inferior colliculus using stereotactic
447 coordinates (PS: 3.4 mm ML, -1.7 mm AP relative to bregma and -2.5 mm from brain surface; IC:
448 0.8 mm Ml, -5.0 mm AP relative to bregma and 0.7 mm from brain surface).

449

450 *Whole-cell and cell-attached electrophysiological recordings:* borosilicate glass pipettes (O.D. 1.5
451 mm, ID 0.86 mm, Sutter Instruments) were pulled into tips of 2.5-5 MΩ resistance. Pipettes were
452 filled with internal solution containing 127 mM K-gluconate, 8 mM KCl, 10 mM phosphocreatine,
453 10 mM HEPES, 4 mM Mg-ATP, 0.3 mM Na-GTP 527 (osmolality, 285 mOsm; pH 7.2 adjusted
454 with KOH). Tips were mounted on microelectrode holders with AgCl wire. Recordings were
455 obtained using a Multiclamp 700B amplifier and digitized with a 1440B Digidata (Molecular
456 Devices). Data was acquired with Clampex 10.7 (Molecular Devices), low-pass filtered at 1 kHz,
457 high-pass filtered at 100 Hz and digitized at 20 kHz. Pipette tips were lowered 0.5 – 0.8 mm from
458 the pial surface of the brain to target layer 5. Pipette pressure was measured using a manometer
459 (Omega) maintained at 15-20 mbar until the desired depth was reached.
460
461 *Optogenetic identification of auditory cortex projection neurons:* For identification of light
462 sensitive neurons, whole-cell and cell-attached recordings were obtained using a Fiberoptic Light
463 Stimulating Device with 465nm blue LED (A-M Systems, Item# 726500) connected to a Fiber
464 Optic Light guide (A-M Systems, Item# 726527). The optic fiber was inserted into the patch
465 pipette via 1.5 mm O.D. Continuous Optopatcher holder (A-M 547 Systems, Item# 663943).
466 Pulses of blue light were delivered through the optic fiber via Digidata 1440A (Molecular Devices)
467 while recording in cell-attached or whole-cell configuration. Light pulses of 20, 50 or 200 ms
468 duration were delivered with increasing intensity from 20 to 100% of full LED power (3 mW at
469 the tip of the fiber). ChR2+ neurons responded to light pulses with an increase firing rate at low
470 jitter <10 ms from light onset, while ChR2- neurons were not modulated by blue light, or if multi-
471 synaptically activated, did not respond with low jitter and high reliability across trials.
472

473 **Behavioral assays with chemogenetics**

474 *Pup retrieval assay:* 2-3 weeks following viral injection surgery, virgins were cohoused with a
475 dam and P0-P1 litter of at least four pups in the dam's home cage. Virgins were co-housed for at
476 least 3 days before retrieval was tested. Prior to test sessions drug administration (CNO or vehicle),
477 virgins were confirmed to be expert retrievers (retrieving pups with at least an 80% success rate).
478 Pup retrieval performance was tested using a trial-based assay as previously described[19,21,22].
479 Briefly, in each session of retrieval, the test subject was habituated for 45-60 minutes in an arena
480 with corncob bedding, food and hydrogel and 2-3 pups age P0-P3 were placed in a corner of the
481 arena with nesting material from the test subject's homecage. Habituation and testing took place
482 in a dark room under infrared light and minimal red light to enable the experimenter to perform
483 the assay. During testing, on each trial a single pup was placed in a corner of the cage away from
484 the nest. Trials were initiated only when the test subject was in the nest. Retrieval was scored as
485 successful if the subject brought the pup back to the nest within 2 minutes. If after 2 minutes the
486 pup was not retrieved, the pup was removed from the cage. Video of each session was collected
487 with a USB webcam and ultrasonic audio recorded with an ultrasonic speaker (Avisoft). In a subset
488 of mice, ethograms of behavior were hand-scored using Datavyu and aligned with the timing of
489 USVs, detected using Deepsqueak.

490

491 *Drug administration:* Clozapine-2H20 (CNO) was dissolved in sterile saline. 45-50 minutes prior
492 to a retrieval test session, a 2 mg/kg dose of CNO or the equivalent volume of vehicle (saline) was
493 administered via intraperitoneal injection. CNO and vehicle sessions occurred on the same day at
494 least 3 hours apart. The experimenter was blinded to treatment. For several mice, additional
495 sessions were performed the following day with a 1 mg/kg dose of CNO.

496

497 **2-photon imaging data analysis**

498 *Pre-processing:* Tif stacks from 2-photon imaging experiments were pre-processed using Suite2P.
499 Imaging frames were registered over entire sessions from individual days, then regions of interest
500 (ROIs) corresponding to somas were identified semi-automatically using an activity-based
501 algorithm and manual curation. Fluorescence time series were extracted from each ROI over the
502 duration of imaging. Using custom MATLAB scripts, ROIs with surrounding neuropil that had
503 much higher fluorescence at baseline were excluded from further analysis. Time series were
504 filtered using a bi-directional median filter with kernel width of 5 frames (each frame +/- 2 frames).

505

506 *Data analysis:* All further analysis was done using custom MATLAB scripts except where noted.
507 dF/F was calculated as $(F_t-F_0)/F_0$ for each stimulus trial (tone or USV) using a 2 second period
508 prior to stimulus onset as F0. All neurons tracked longitudinally in **Figure 4** were included in
509 analysis. Best early responses to stimuli (**Figs. 3, 4**) were defined as follows: for each ROI,
510 responses to each stimulus were averaged over trials, then the mean dF/F was computed over the
511 first 2 seconds following stimulus onset. The "best" USV or tone were defined as the USV and
512 tone evoking the largest magnitude mean dF/F. The same procedure was used to determine the
513 best late response-evoking USV, except that mean dF/F was computed from 2-5 seconds following
514 stimulus onset. For the most suppressed response-evoking stimuli, mean dF/F was computed in
515 the first 3 seconds following stimulus onset, and the greatest magnitude negative dF/F was
516 considered the most suppressed response.

517

518 *ROI tracking:* ROIs were tracked across days by registering Suite2p reference images across pairs
519 of days and computing area overlap of ROI masks using the open source MATLAB program
520 ROIMatchPub (https://github.com/ransona/ROIMatchPub). ROIs were considered to be the same
521 neuron if 60% of pixels in the masks two masks were overlapping.
522
523 **Analysis of *in vivo* whole-cell and cell-attached recordings**
524 All analysis was performed using custom MATLAB scripts except where noted. Voltage traces
525 for each cell were acquired using Clampex, then imported into MATLAB using the open source
526 program abfload. For spike rate and timing measurements, voltage traces were highpass filtered
527 above 100 Hz and thresholded manually to obtain binary spike trains. Spike time was recorded as
528 the timing of the first threshold crossing. Firing rates were computed using 10 ms-wide bins and
529 averaged over trials of repeated stimuli to obtain PSTHs. Stimulus-evoked firing rates were
530 computed in 2 seconds following stimulus onset and a baseline period averaged over 1.5 seconds
531 was subtracted.
532

533 **Acknowledgements**

534 We thank M.A. Kirchgessner and K.A. Martin for comments, discussions, and technical assistance.

535 This work was funded by the BRAIN Initiative (NS107616 to R.C.F. and K99DC021581 to

536 A.M.M.), NICHD (HD088411 to R.C.F.), and NIDCD (DC12557 to R.C.F.).

References

1. Asilador, A. & Llano, D. A. Top-Down Inference in the Auditory System: Potential Roles for Corticofugal Projections. *Frontiers in Neural Circuits* **14**, (2021).
2. Doucet, J. R., Rose, L. & Ryugo, D. K. The cellular origin of corticofugal projections to the superior olivary complex in the rat. *Brain Res* **925**, 28–41 (2002).
3. Schofield, B. R. & Coomes, D. L. Auditory cortical projections to the cochlear nucleus in guinea pigs. *Hearing Research* **199**, 89–102 (2005).
4. Fitzpatrick, K. A. & Imig, T. J. Projections of auditory cortex upon the thalamus and Midbrain in the owl monkey. *Journal of Comparative Neurology* **177**, 537–555 (1978).
5. Winer, J. A., Chernock, M. L., Larue, D. T. & Cheung, S. W. Descending projections to the inferior colliculus from the posterior thalamus and the auditory cortex in rat, cat, and monkey. *Hearing Research* **168**, 181–195 (2002).
6. Malmierca, M. S. & Ryugo, D. K. Descending Connections of Auditory Cortex to the Midbrain and Brain Stem. in *The Auditory Cortex* (eds. Winer, J. A. & Schreiner, C. E.) 189–208 (Springer US, 2011). doi:10.1007/978-1-4419-0074-6_9.
7. Blackwell, J. M., Lesicko, A., Rao, W., De Biasi, M. & Geffen, M. N. Auditory cortex shapes sound responses in the inferior colliculus. *eLife* **9**, 1–26 (2020).
8. Lesicko, A. M. H. & Geffen, M. N. Diverse functions of the auditory cortico-collicular pathway. *Hear Res* **425**, 108488 (2022).
9. Marsh, J. E. & Campbell, T. A. Processing Complex Sounds Passing through the Rostral Brainstem: The New Early Filter Model. *Frontiers in Neuroscience* **10**, (2016).
10. Bajo, V. M., Nodal, F. R., Moore, D. R. & King, A. J. The descending corticocollicular pathway mediates learning-induced auditory plasticity. *Nature Neuroscience* **13**, 253–260 (2010).

561 11. Bajo, V. M. *et al.* Silencing cortical activity during sound-localization training impairs
562 auditory perceptual learning. *Nature Communications* **10**, (2019).
563 12. Dulac, C., O'Connell, L. A. & Wu, Z. Neural control of maternal and paternal behaviors.
564 *Science* **345**, 765 LP – 770 (2014).
565 13. Lingle, S., Wyman, M. T., Kotrba, R., Teichroeb, L. J. & Romanow, C. A. What makes a cry
566 a cry? A review of infant distress vocalizations. *Current Zoology* **58**, 698–726 (2012).
567 14. Swain, J. E., Kim, P. & Ho, S. S. Neuroendocrinology of Parental Response to Baby-Cry.
568 *Journal of Neuroendocrinology* **23**, 1036–1041 (2011).
569 15. Ehret, G. Infant Rodent Ultrasounds – A Gate to the Understanding of Sound
570 Communication. *Behavior Genetics* **35**, 19–29 (2005).
571 16. Ehret, G., Koch, M., Haack, B. & Markl, H. Sex and parental experience determine the onset
572 of an instinctive behavior in mice. *Naturwissenschaften* **74**, 47 (1987).
573 17. Elyada, Y. M. & Mizrahi, A. Becoming a mother-circuit plasticity underlying maternal
574 behavior. *Current Opinion in Neurobiology* **35**, 49–56 (2015).
575 18. Koch, M. & Ehret, G. Estradiol and parental experience, but not prolactin are necessary for
576 ultrasound recognition and pup-retrieving in the mouse. *Physiology & Behavior* **45**, 771–776
577 (1989).
578 19. Marlin, B. J., Mitre, M., D'Amour, J. A., Chao, M. V. & Froemke, R. C. Oxytocin enables
579 maternal behaviour by balancing cortical inhibition. *Nature* **520**, 499–504 (2015).
580 20. Liu, R. C., Linden, J. F. & Schreiner, C. E. Improved cortical entrainment to infant
581 communication calls in mothers compared with virgin mice. *European Journal of*
582 *Neuroscience* **23**, 3087–3097 (2006).

583 21. Schiavo, J. K. *et al.* Innate and plastic mechanisms for maternal behaviour in auditory cortex.
584 *Nature* **587**, 426–431 (2020).
585 22. Carcea, A. I. *et al.* Oxytocin Neurons Enable Social Transmission of Maternal Behavior.
586 *Nature* **596**, 553–557 (2021).
587 23. Cohen, L. & Mizrahi, A. Plasticity during Motherhood: Changes in Excitatory and Inhibitory
588 Layer 2/3 Neurons in Auditory Cortex. *Journal of Neuroscience* **35**, 1806–1815 (2015).
589 24. Znamenskiy, P. & Zador, A. M. Corticostriatal neurons in auditory cortex drive decisions
590 during auditory discrimination. *Nature* **497**, 482–485 (2013).
591 25. Xiong, Q., Znamenskiy, P. & Zador, A. M. Selective corticostriatal plasticity during
592 acquisition of an auditory discrimination task. *Nature* **521**, 348–351 (2015).
593 26. Guo, L., Weems, J. T., Walker, W. I., Levichev, A. & Jaramillo, S. Choice-selective neurons
594 in the auditory cortex and in its striatal target encode reward expectation. *Journal of*
595 *Neuroscience* **39**, 3687–3697 (2019).
596 27. Ford, A. N., Czarny, J. E., Rogalla, M. M., Quass, G. L. & Apostolides, P. F. Auditory
597 corticofugal neurons transmit non-auditory signals to support discriminative learning.
598 2022.08.08.503214 Preprint at https://doi.org/10.1101/2022.08.08.503214 (2022).
599 28. Xiong, X. R. *et al.* Auditory cortex controls sound-driven innate defense behaviour through
600 corticofugal projections to inferior colliculus. *Nature Communications* **6**, 1–12 (2015).
601 29. Chen, X. *et al.* High-Throughput Mapping of Long-Range Neuronal Projection Using In Situ
602 Sequencing. *Cell* **179**, 772-786.e19 (2019).
603 30. Hunnicutt, B. J. *et al.* A comprehensive excitatory input map of the striatum reveals novel
604 functional organization. *eLife* **5**, 1–32 (2016).

605 31. Muñoz, W., Tremblay, R. & Rudy, B. Channelrhodopsin-Assisted Patching: InVivo
606 Recording of Genetically and Morphologically Identified Neurons throughout the Brain. *Cell*
607 *Reports* **9**, 2304–2316 (2014).
608

609 **Figure Legends**

610 **Figure 1. Chemogenetic suppression of left auditory cortex projection neurons impairs pup**
611 **retrieval in experts. (a)** Schematic of pup retrieval assay performed after administration of CNO
612 or vehicle. **(b)** Viral injection strategy for targeting of corticostriatal (left) and corticocollicular
613 (right) auditory cortex neurons and example histological confirmation. AAVrg-EF1a-Cre was
614 injected into posterior striatum or inferior colliculus (IC) using stereotactic coordinates for
615 targeting. AAV1-hSyn-DIO-hM4D(Gi)-mCherry was injected into auditory cortex with targeting
616 relative to anatomical landmarks (cranial sutures and blood vessels). Left, example histological
617 confirmation of injection targeting to posterior striatum (P. Striatum) and expression in auditory
618 cortex. Middle left, epifluorescence image of a coronal section containing posterior striatum.
619 Right, auditory cortex. DAPI staining depicted in cyan, mCherry in red. Insets: higher
620 magnification images of areas within white squares showing mCherry expression in fibers and cell
621 bodies. Right 3 panels, example histological confirmation of expression for IC injections**. (c)**
622 Example ethogram of 10 pup retrieval trials from a single mouse. Lower panel, entire duration;
623 upper panel, zoom in of the first trial. Purple denotes trial start, defined as the time of experimenter
624 placing pup in the arena. Pink lines indicate a bout of calls isolated by at least 1 second of silence.
625 Green indicates time of retrieval, from the time the pup is picked up until dropped in or near the
626 nest. Purple arrows in lower panel indicate trial starts. **(d)** Negative correlation of frequency of
627 USV emission (individual USVs per second) during a trial with the latency to retrieve on that trial.
628 N=2 CS virgins and 4 CC virgins, all successful trials from saline block included in calculation of
629 correlation. R=-0.27, p=0.035, Pearson's R. **(e)** Fraction of pups retrieved under each condition.
630 Thin lines denote performance of individual animals. Dashed lines indicate dams and continuous
631 indicate virgins. Thick black lines denote means for each group; vertical lines denote SEM. Black:

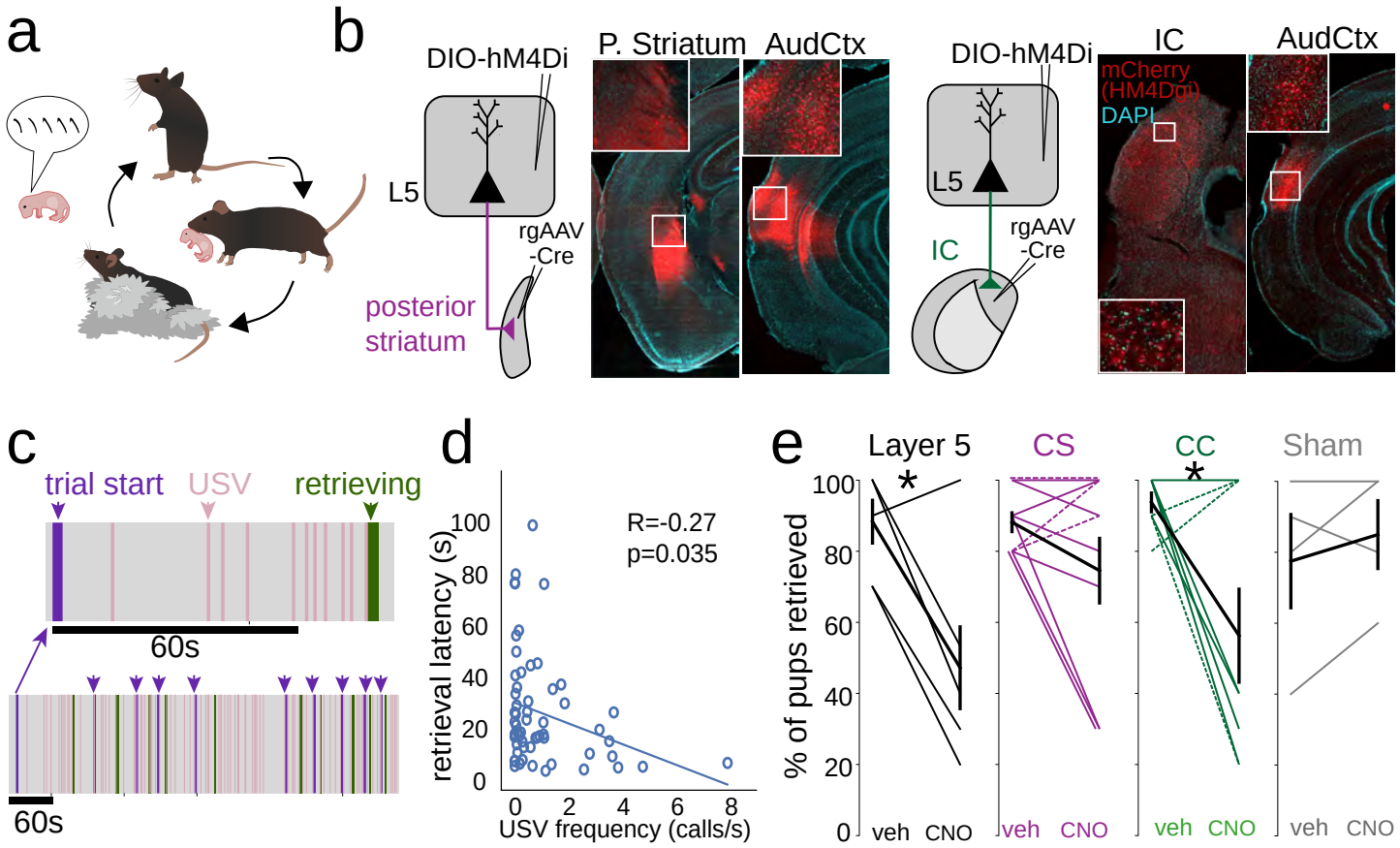
632 performance in RBP4-Cre mice with AAV1-hSyn-DIO-hM4D(gi)-mCherry injected in left
633 auditory cortex to suppress activity in the bulk of layer 5 excitatory neurons (N=6 experienced
634 virgins). Purple: mice with hM4D(gi) targeted to corticostriatal (CS) neurons in left auditory cortex
635 (N=10; 3 dams and 6 experienced virgins). Green: mice with hM4D(gi) targeted to
636 corticocollicular (CC) neurons in left auditory cortex (N=9; 3 dams and 6 experienced virgins).
637 Gray: sham-injected control group. AAV1-hSyn-DIO-mCherry was targeted to corticostriatal
638 (N=2 experienced virgins) or corticocollicular (N=2 experienced virgins) neurons using methods
639 as described for experimental groups. Asterisk, $p<0.05$, Wilcoxon Rank Sum Test.
640

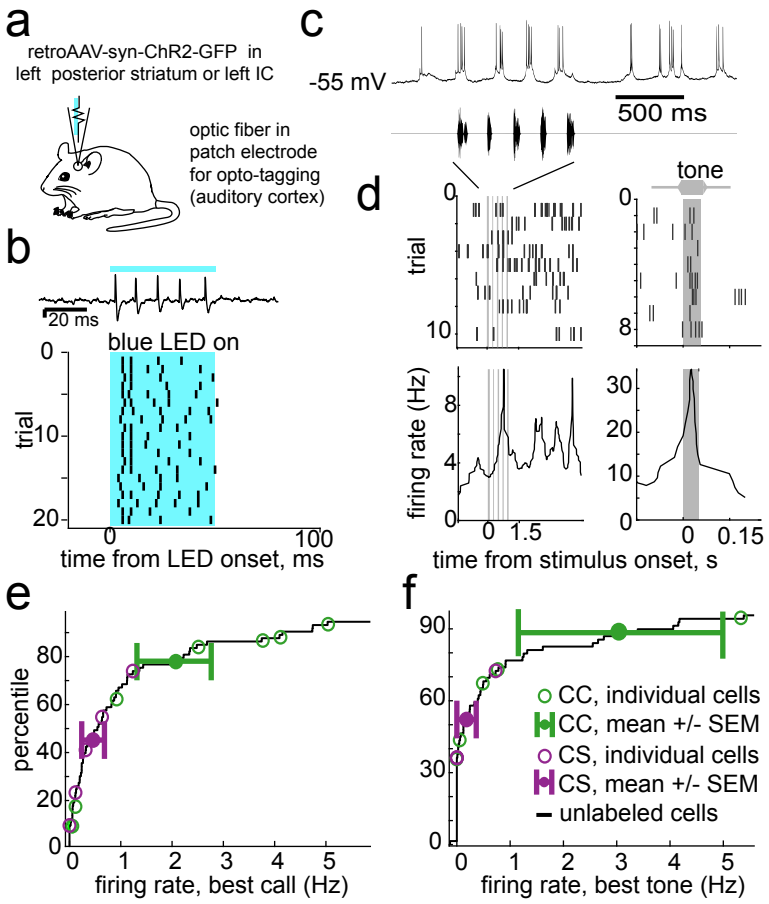
641 **Figure 2.** In vivo opto-tagged whole-cell recordings. **(a)**, Virus recording strategy in awake and
642 anesthetized mice**. (b),** Trace and spike raster from corticostriatal opto-tagged neuron. **(c)**, Current-
643 clamp recording from L5 auditory cortex neuron in an awake mouse. **(d),** Spike rasters and PSTHs
644 from a corticocollicular neuron. Left: responses to best call (USV that evoked highest firing rate
645 for that cell); right, best tone. **(e),** Cumulative distribution of best call-evoked firing rate for 77
646 putative L5 neurons recorded in current-clamp or cell-attached mode. Black line: all neurons
647 including non-tagged (N=77 neurons from 21 retrieving females; 8 corticocollicular neurons
648 indicated in green; 5 corticostriatal in purple). **(f),** As in **(e)** but for best tones.

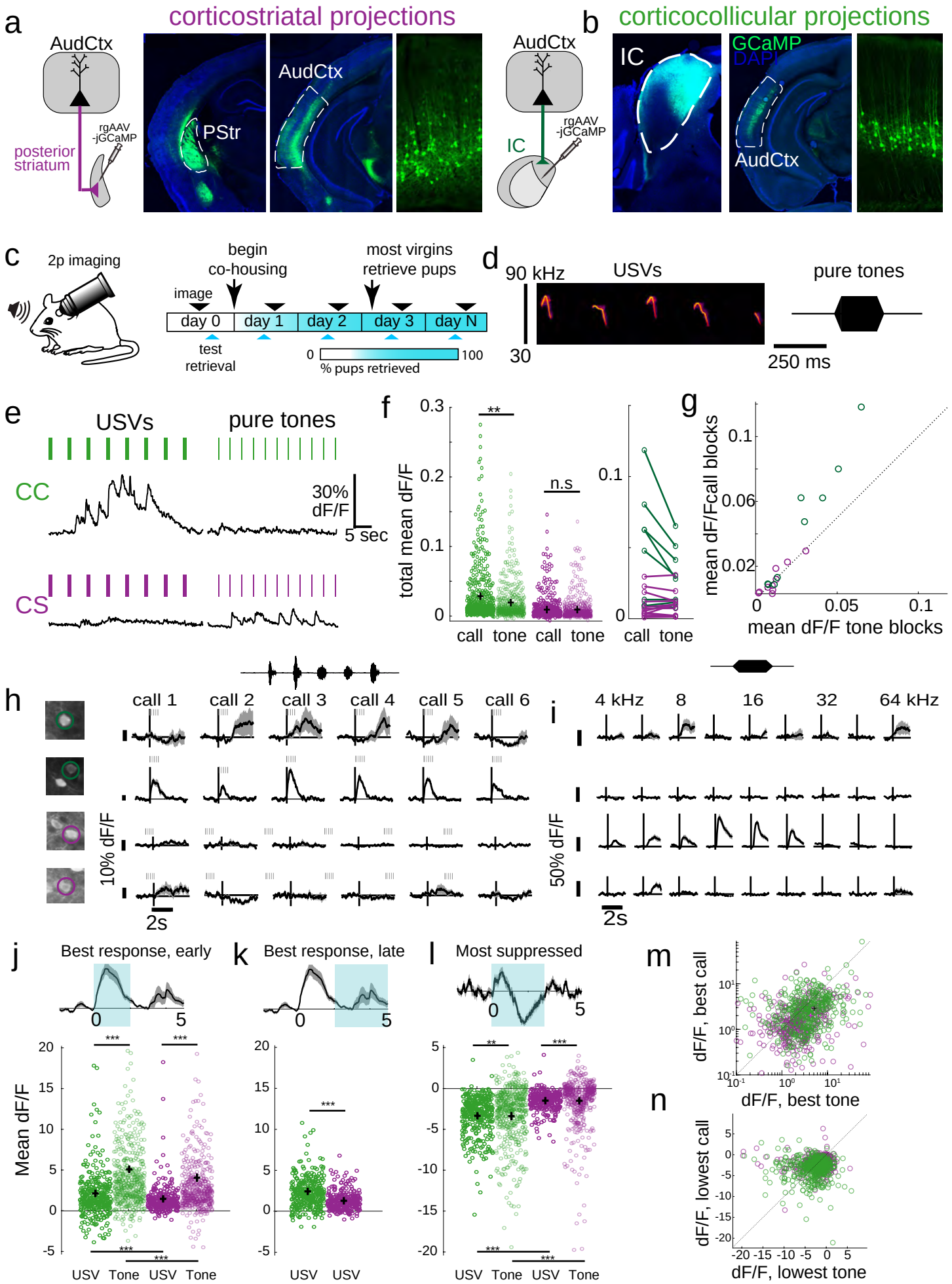
Figure 3. 2-photon in vivo calcium imaging of corticocollicular and corticostriatal-projection neurons during USV and pure tone playback. (a) Left: viral injection strategy for expression of jGCaMP7f for jGCaMP8f in auditory corticostriatal projection neurons. Right panels: histological confirmation of injection targeting in fixed coronal sections from one example brain following 2p imaging experiments. Fluorescence images of injection site, expression in auditory cortex, and high-magnification image of jGCaMP-expressing neurons in auditory cortex, respectively. PStr: posterior striatum; AudCtx: auditory cortex; rgAAV-jGCaMP: rgAAV-hSyn-jGCaMP7f or 8f. **(b)** As described in (a) for auditory corticocollicular projection neurons. IC: inferior colliculus. **(c)** Schematic describing timeline of imaging experiments relative to cohousing and pup retrieval performance testing. **(d)** Auditory stimuli. Left, spectrogram of one of 6 pup ultrasonic vocalizations (USVs) used in experiments. Right: duration of pure tone pip. Half-octave spaced pure tones ranging from 4 kHz to 64 kHz were used. **(e)** dF/F traces from example neurons imaged during epochs of call presentations (left) and tone presentations (right). Upper row: corticocollicular neurons; lower row: corticostriatal neurons. Vertical bars denote timing of stimulus delivery. **(f)** Left panel: Mean dF/F in all ROIs from N=8 corticocollicular labeled mice (purple, left) and N=7 corticostriatal mice during blocks of USV or tone presentations (green, right). Each circle is one cell's mean; black crosses indicate population means. Right panel: dF/F averaged within mouse. Each line depicts on mouse. Data were collected on day 3 of cohousing or when animals displayed stable expert performance. **(g)** Mean dF/F averaged over all ROIs during tone vs. call blocks. Each circle is one mouse; dotted line is x=y. **(h-i)** Mean USV- and tone-evoked dF/F relative to timing of USV and tone playback for 4 example neurons. Top two rows: corticocollicular neurons; bottom rows: corticostriatal. 10-20 repetitions were played per stimulus. Black lines denote means over all trials; gray shading indicates SEM. Vertical black lines indicate

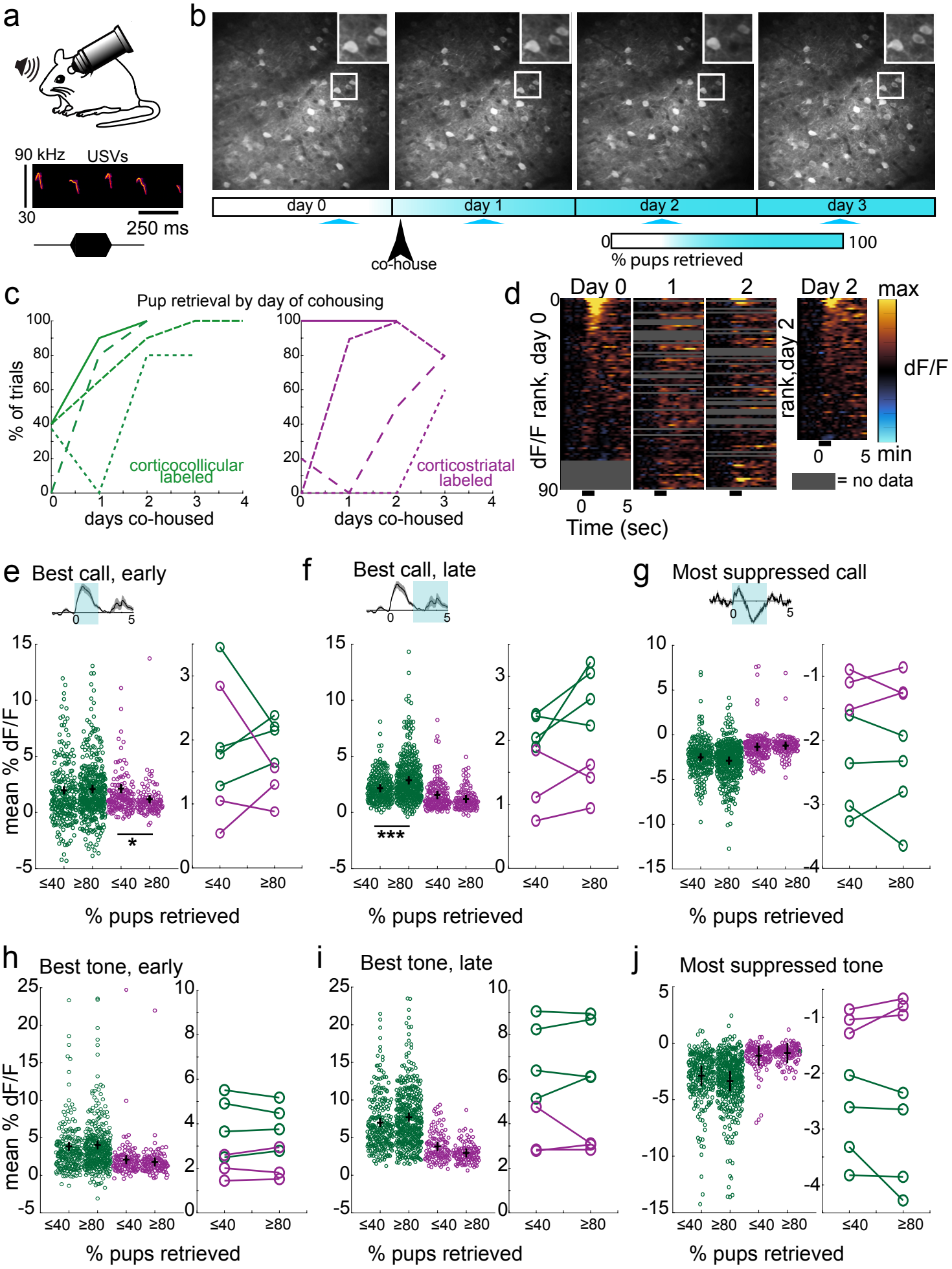
stimulus onset. **(h)** Responses to 6 USVs; gray vertical hashmarks denote timing of individual USV syllables. **(i)** responses to 9 quarter-octave spaced pure tones. **(j)** Mean stimulus-evoked dF/F for each neuron's best USV and tone, defined as the USV or tone evoking the largest trial-averaged dF/F during a 2 second period beginning at stimulus onset. Purple: corticostriatal; green: corticocollicular; symbols as described for (g). **(k)** Mean stimulus-evoked dF/F for each neuron's best late-response-evoking stimulus, measured by trial-averaged dF/F 2-5 seconds following stimulus onset. **(l)** Mean stimulus-evoked dF/F for stimuli evoking each neuron's most-suppressed response, defined by trial-averaged dF/F in 0-3 seconds following stimulus onset. All data in panels j-l were collected on day 3 of cohousing or when animals displayed stable expert performance. Asterisks denote P-values as follows: ** = $p<0.01$; *** = $p<0.001$; n.s. = not significant. P-values were obtained using Kruskal-Wallis method, with pairwise comparisons corrected for multiple comparisons post-hoc using Tukey's HSD test.

684 **Figure 4. Auditory responses in corticocollicular and corticostriatal neurons tracked over**
685 **co-housing. (a)** Experiment schematic. **(b)** Average 2-photon images of an example imaging field
686 expressing jGCaMP8f on four consecutive days of imaging. Images were averaged over entire
687 sessions, which comprised ~10^5 individual frames. White squares denote location of zoomed-in
688 images in insets. Lower panel: timeline of image acquisition relative to co-housing. **(c)** Pup
689 retrieval performance in individual virgins over days co-housed with a dam and litter. Retrieval
690 was quantified on each day as percent of pups retrieved over all trials. Left, mice with
691 corticocollicular neurons labeled with jGCaMP8; right, corticostriatal. **(d)** Corticocollicular
692 neurons from one example imaging field tracked over 3 days. Panels, mean responses to best calls
693 for each neuron over 3 imaging days. Neurons are sorted by mean USV-evoked dF/F for each
694 neurons best USV on day 0 (prior to co-housing). Rows: neurons; columns: time bins relative to
695 call onset. Heatmap: normalized dF/F. Rightmost panel: all neurons present on third day (day 2)
696 of cohousing ranked by best-USV-evoked dF/F on day 3. **(e)** USV-evoked mean dF/F of all
697 corticocollicular (green) and corticostriatal (purple) neurons grouped by pup retrieval performance
698 on day of imaging. Low performance was defined as ≤40% of pups retrieved, high performance
699 as ≥80%. Individual circles denote dF/F for each neuron on each day of imaging; black crosses
700 denote population means. Left panel, best USV-evoked dF/F in first 2 seconds following call onset;
701 right, data averaged within each mouse. (**f**) As described in e for dF/F 2-5 seconds following onset
702 of best late-response-evoking USV. **(g)** dF/F in first 3 seconds following onset of USV evoking
703 the lowest dF/F. **(h-j)** Tone-evoked mean dF/F of all corticocollicular and corticostriatal neurons
704 grouped by performance as described for (**e**). **, $p<0.01$; ***, $p<0.001$; n.s., not significant. p-
705 values were obtained using Kruskal-Wallis method, with pairwise comparisons corrected for
706 multiple comparisons post-hoc using Tukey's HSD test.


a
b
DIO-hM4Di
L5
rgAAV
-Cre
posterior
striatum
P. Striatum
AudCtx
DIO-hM4Di
L5
IC
rgAAV
-Cre
IC
AudCtx
mCherry
(HM4Dqi)
DAPI
c
trial start
USV
retrieving
60s
60s
d
R=-0.27
p=0.035
retrieval latency (s)
USV frequency (calls/s)
e
Layer 5
CS
CC
Sham
% of pups retrieved
veh
CNO
veh
CNO
veh
CNO
veh
CNO


a
retroAAV-syn-ChR2-GFP in
left posterior striatum or left IC
optic fiber in
patch electrode
for opto-tagging
(auditory cortex)
b
20 ms
blue LED on
trial
time from LED onset, ms
c
-55 mV
500 ms
d
tone
trial
firing rate (Hz)
time from stimulus onset, s
e
percentile
firing rate, best call (Hz)
f
CC, individual cells
CC, mean +/- SEM
CS, individual cells
CS, mean +/- SEM
unlabeled cells
firing rate, best tone (Hz)


a
corticostriatal projections
AudCtx
rgAAV
-jGCaMP
posterior
striatum
PStr
AudCtx
b
corticocollicular projections
AudCtx
rgAAV
-jGCaMP
IC
IC
GCaMP
DAPI
AudCtx
c
2p imaging
begin
co-housing
most virgins
retrieve pups
image
day 0
day 1
day 2
day 3
day N
test
retrieval
0
100
% pups retrieved
d
90 kHz
USVs
pure tones
30
250 ms
e
USVs
pure tones
CC
30%
dF/F
5 sec
CS
f
total mean dF/F
0.3
0.2
0.1
0
**
n.s
call
tone
call
tone
0.1
0
call
tone
g
mean dF/Fcall blocks
0.1
0.06
0.02
0
0.05
0.1
mean dF/F tone blocks
h
call 1
call 2
call 3
call 4
call 5
call 6
10% dF/F
2s
i
4 kHz
8
16
32
64 kHz
50% dF/F
2s
j
Best response, early
0
5
Mean dF/F
20
15
10
5
0
-5
***
***
***
***
USV
Tone
USV
Tone
k
Best response, late
0
5
20
15
10
5
0
-5
***
USV
USV
l
Most suppressed
0
5
5
0
-5
-10
-15
-20
**
***
***
***
USV
Tone
USV
Tone
m
dF/F, best call
dF/F, best tone
n
dF/F, lowest call
dF/F, lowest tone


a
90 kHz
USVs
30
250 ms
b
day 0
day 1
day 2
day 3
co-house
0
100
% pups retrieved
c
Pup retrieval by day of cohousing
% of trials
days co-housed
corticocollicular labeled
corticostriatal labeled
d
Day 0
1
2
Day 2
max
dF/F
dF/F rank, day 0
rank,day 2
min
= no data
Time (sec)
e Best call, early
f Best call, late
g Most suppressed call
mean % dF/F
% pups retrieved
h Best tone, early
i Best tone, late
j Most suppressed tone