Auditory corticofugal circuits drive vocalization-guided maternal behavior

Amy M. LeMessurier[1-5], Ayat A. Agha[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] Neuroscience Institute, 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 and Physiology, 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 don't 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. 1**), 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. 1a**) 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. 1d**). 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. Under saline conditions, RBP4-Cre experienced virgins retrieved pups on 88+-6% of trials (**Fig. 1e**), whereas on CNO trials, percent of pups retrieved decreased by roughly half to 47±12% ($p<0.05$, paired Student's T-test), indicating that pup retrieval depends on activity in left layer 5.

Neurons in layer 5 send corticofugal projections to areas earlier in the auditory hierarchy, as well as feedforward projections to the striatum and amygdala[29] (**Fig. 1a**). Corticostriatal projections from auditory cortex primarily innervate the posterior striatum, a sensory area that receives dense input from auditory, visual, and somatosensory cortices, as well as limbic system inputs[30]. As the striatum has been found to be involved in action selection and motor sequencing, and because the auditory corticstriatal projection in particular has been implicated in behavioral responses to acoustic cues in perceptual learning tasks[24–26], we hypothesized that this projection could serve as a direct link between auditory perception of pup vocalizations and motor planning and execution of retrieval. To test this hypothesis, we expressed HM4D(gi) selectively in auditory corticostriatal neurons using an intersectional virus approach: cre was retrogradely expressed in these neurons by injecting rgAAV-Cre into the left posterior striatum, and cre-dependent HM4D(gi)-mCherry virus or just mCherry for sham controls was injected into left auditory cortex (**Fig 1a-b**).

Surprisingly, retrieval performance under CNO administration was unchanged relative to saline on average in 4 experienced virgins and 3 dams (Retrieval probability: saline, 86+-3%; CNO, 71+-11%, $p>0.05$, paired T-test). Similarly, no effect was seen in sham-injected controls (**Fig. 1e;** Retrieval probability, all shams: saline, 76±13%; CNO, 85±10%, $p>0.05$, paired T-test; CS sham only: saline, 70±3%; CNO, 80±2%). Correct targeting of injections and spread of expression were determined by visualizing mCherry in post-hoc histology for each mouse (**Fig. 1b**). While posterior striatum was correctly targeted in 100% of mice, spread of expression varied, with several mice exhibiting expression in hippocampal CA3, which is likely due to

leakage along the injection track and spread of cre-dependent virus beyond cortex. This extended expression beyond auditory corticostriatal neurons could explain the high variance in effect of CNO across individual animals (**Fig. 1e**).

The lack of effect of silencing corticostriatal alone neurons suggested that a different subset of layer 5 neurons could be responsible for the diminished performance we observed when silencing left AC layer 5 more broadly. Corticofugal projection neurons from auditory cortex densely innervate early areas in the central auditory pathway, including the medial geniculate body of the thalamus (MGB), the inferior colliculus (IC), and auditory brainstem nuclei. We hypothesized that auditory cortex may drive pup retrieval by modulating USV representations in earlier auditory areas that project more directly to circuits that drive motor execution. IC is the first area of the auditory pathway in which neurons extensively integrate acoustic input over time, and which respond to USVs in multiple species. Thus, we assayed pup retrieval performance while suppressing activity in auditory corticocollicular neurons using the same approach as described for corticostriatal neurons, with rgAAV-Cre injected into left IC (**Fig. 1a,c**). Strikingly, CNO decreased the percentage of pups retrieved in 4/5 experienced virgins roughly in half (**Fig. 1e;** Retrieval probability: saline, 98±2%; CNO, 46±14%, $p<0.05$, paired T-test). No effect was seen in sham mCherry-injected controls (CC sham only: saline, 85±5%; CNO, 90±10%). Thus, corticofugal projections to inferior colliculus likely perform a critical function supporting retrieval.

Auditory corticocollicular neurons are particularly responsive to USVs in comparison with corticocollicular neurons. The differing effects of suppressing corticostriatal and

corticocollicular neurons during retrieval suggest that corticocollicular neurons in particular may modulate perception of USVs in ways that facilitate retrieval performance. For example, corticocollicular input could increase sensitivity to facilitate detection of calls, increase integration windows in IC neurons to support categorization of calls based on temporal and acoustic features, and/or facilitate neuromodulatory circuits that promote initiation of motor sequences. We reasoned that if corticocollicular neurons support such computations in expert retrievers, a greater proportion of them would be selective for USVs over other acoustic stimuli in comparison with corticostriatal neurons.

We performed in vivo whole-cell and cell-attached recordings in corticocollicular and corticostriatal neurons measuring responses to USVs and pure tones (**Fig. 2**). We identified recorded neurons based on projection target by expressing channelrhodopsin in either projection and testing responsiveness to blue light[31]. Channelrhodopsin-2 was expressed in CC or CS neurons in experienced retrievers (3 dams and 18 cohoused virgins) via injection of rgAAV-ChR2-GFP in IC or posterior striatum, respectively (**Fig. 2a,b**). Optic fibers were connected to a blue LED were inserted into recording pipettes for light delivery to recorded neurons. Acute craniotomies were performed over AC in mice headplated one week prior to recording. Electrodes were lowered to 400 um to target layer 5, and once a seal was obtained, sensitivity to blue light was tested using pulses of 20, 50, and 200 ms ranging intensity from 0 mW to 3 mW at the tip of the electrode. Neurons were considered to be channelrhodopsin-expressing if they spiked within 10 ms of light onset on 80% of trials, and if increasing durations of light led to increased duration of spike trains (as opposed to suppression or rebound spiking after light offset).

Current-clamp or cell-attached recordings were obtained from 8 CC neurons from 7 virgins and 5 CS neurons from one dam and 4 virgins, as well as 73 additional neurons that were not light-activated and considered unlabeled. We measured responses to a set of 6 prototypical pup isolation USVs (**Fig. 2c**), each comprised of 4-5 individual syllables and of durations ~1 second, as well as 50 ms duration pure tone pups, in half-octave intervals from 4-64 kHz. Each stimulus was repeated at least 10 times, and firing rates relative to stimulus times were averaged across trials (**Fig. 2d-f**). We then determined the best USV and best tone for each recording: Best USV was defined as the USV evoking the highest magnitude change in firing rate in the first 2 seconds following USV onset compared to baseline rate quantified in the 1.5 seconds preceding. Best tone was defined similarly, but firing rates were quantified in the first 200 ms following tone onset for evoked change in firing, and 200 ms for the preceding baseline period. USV-evoked firing rates varied across the distribution of labeled and unlabeled cells from -0.2 to 5.2 Hz with a median of 0.41(**Fig. 2e**). Tone-evoked firing rates varied from 0 to 22.7 Hz, with a median of -0.13 (**Fig. 2f**). Mean firing rates for both USVs and tones were higher in CC neurons compared to tones (**Fig. 2e-f,** 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, 0.6±0.7Hz in CS, $p<0.05$, Students T-test).

While mean rates were higher in CC neurons for both types of stimulus – reflecting the top 22nd and 15th percentiles of all recorded neurons for USVs and tones, respectively – firing rates varied greatly across these cells, ranging from the 90th to 10th percentiles of USVs responses and 55th to 5th percentiles of tone responses. Given the relatively low sample sizes that can be obtained with single-neuron recordings, we pursued in vivo 2-photon calcium imaging in labeled auditory

cortex projection neurons to obtain more highly-sampled estimates of these two populations' representations of acoustic stimuli (**Figs. 3,4**). Additionally, we wondered how representations of pup isolation USVs vary over time as virgins gain experience with alloparenting. Longitudinal 2-photon calcium imaging enabled us to track responses in individual neurons and populations over the course of co-housing (**Figs. 3a-c, 4**).

We expressed jGCaMP7f or 8f in CC neurons (8 mice) or CS neurons (6 mice) by injecting rgAAV-hSyn-jGCaMP7f or 8f in left IC (**Fig. 3b**) or left posterior striatum (**Fig 3a**), respectively. During the same surgeries, chronic cranial windows were implanted over left AC and steel headbars were attached to the skull to enable head stabilization during imaging. Imaging experiments measuring USV and tone-evoked activity were performed 3-5 weeks following surgery. Following imaging experiments, brains were sectioned and imaged to visualize injection targeting and expression in auditory cortex. In all CC and CS mice, expression in cell bodies was most dense several hundred um from the pial surface of the brain (**Fig. 3a,b**), with the bulk of labeled neurons exhibiting morphology matching that which has been described for L5 excitatory neurons; in particular, many labeled neurons in both groups had large apical dendrites protruding dorsally from somas, with the highest density of axonal branching appearing to be in superficial layers of cortex. In corticostriatal neurons, GCaMP-expressing cell bodies were also seen in more superficial layers, in accordance with studies finding corticostriatal intratelencephalic (IT) neurons across multiple layers that project to more anterior striatum and reports of long range inhibitory interneurons that also project to striatum. In CC, some expression was observed in somas deeper than the band matching L5, which were likely in L6 in accordance with other studies.

We first investigated representations of USVs and tones in expert retrieving virgins that had been cohoused for several days. At least one day following habituation to head-fixation, awake mice were stabilized under a 2-photon microscope, and auditory stimuli were played through an ultrasonic speaker as we imaged in left auditory cortex (**Fig. 3 c,d**). A set of 6 USVs were played in two blocks of ten repetitions per stimulus, with an interstimulus interval of 5 seconds, as described for single cell recordings. Between the blocks of USVs, a block of half-octave spaced pure tone pips of 250 ms duration were played at 70 dB SPL, with 10 repeats per tone at an interstimulus interval of 3 seconds (**Fig. 3d**). Fluorescence time series were extracted from ROIs representing somas following 3 days of co-housing with a dam and litter and daily retrieval testing, or for a subset of mice in which we tracked the same imaging fields throughout cohousing, ROIs were examined on the final day of imaging.

In CC neurons, we observed striking increases in the frequency and amplitude of calcium transients during blocks of USVs that were not closely time-locked to the 5 second-spaced stimuli (**Fig. 3e**). However, during blocks of pure tones, we did not observe similar levels of spontaneous activity (quantified as mean dF/F over entire blocks of each trial type, **Fig. 3e-g,** green symbols; mean dF/F in CC neurons during USV blocks: 0.03±0.001%; during tone blocks: 0.02±0.001%; $p<0.01$, Kruskal-Wallis with Tukey's HSD correction for multiple pairwise comparisons). This apparent difference in non-time-locked activity could be due to different total durations of the types of acoustic stimuli, since each USV was comprised of multiple syllables of roughly 75-200 ms each whereas each tone was a single 250 ms pip. However, we did not observe a difference in non-time-locked dF/F in CS neurons between USV and tone blocks

(0.008±0.001%; during tone blocks: 0.009±0.001%; $p>0.05$, Kruskal-Wallis with Tukey's HSD correction for multiple pairwise comparisons), suggesting that excitability in CC neurons in particular might be selective for USVs compared to tones in expert retrieving females. In accordance with single-neuron recordings described above (**Fig. 2**), overall mean dF/F was higher in CC neurons compared to CS during both blocks of USVs and tones (**Fig. 3f,g**; USV blocks, CC vs CS: $p<0.001$; tone blocks, $p<0.001$, Kruskal-Wallis with Tukey's HSD correction for multiple pairwise comparisons).

Next we examined time-locked dF/F relative to stimuli in both groups. (**Fig. 3h-n**). For each cell, dF/F was quantified for each stimulus repetition relative to the onset of the stimulus, using a pre-stimulus baseline period as baseline fluorescence (F_0). Both CC and CS neurons exhibited a diverse array of temporal response profiles and tuning for both calls and tones (**Fig. 3h,i)**, including neurons that were relatively selective for individual calls or a narrow band of tone frequencies, were responsive to calls but not tones or vice versa, were reliably activated by every USV, or were suppressed by a subset of USVs. Surprisingly many neurons were reliably activated by USVs at long delays, up to several seconds post USV onset. Thus, for each neuron we quantified 3 measures of evoked dF/F: early responses to the best USV and tone for each neuron (**Fig. 3j**), defined as the USV and tone evoking the largest magnitude mean dF/F in the first 2 seconds following stimulus onset; the best late response **(Fig. 3k**), determined by the same procedure, except that mean dF/F was computed from 2-5 seconds following stimulus onset; and most suppressed response-evoking stimuli (**Fig. 3 l**), for which mean dF/F was computed in the first 3 seconds following stimulus onset, and the greatest magnitude negative dF/F was considered the most suppressed response. The best late response was only computed for USVs,

because tones were played at an inter-stimulus interval of 3 seconds, and the delayed/late responses we observed following USVs occurred reliably up to 5 seconds post stimulus onset for some neurons.

By all three evoked response metrics, CC responses were larger magnitude on average over the populations. (Best call: CC 2.9±0.2%, CS 2.1±0.1%, $p<0.001$; Late best call: CC 4.1±3%, CS 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 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%, $p<0.001$; Kruskal Wallis test with Tukey's HSD correction for multiple pairwise comparisons).

In both CC and CS cells, responses to best tones and most suppressed tones were larger magnitude than best and most suppress USVs (Best USV vs tone: CC $p<0.001$, CS $p<0.001$; most suppressed USV vs tone: CC $p<0.01$; CS $p<0.001$). This was in contrast with our finding that in CC cells overall mean dF/F during blocks of calls was greater than during blocks of tones, and that in CS cells, overall mean dF/F during blocks of calls and tones were not different. This suggests that highly stimulus-locked responses may be independent of longer-duration modulation of activity levels. Neither facilitate nor suppressed response magnitudes were correlated between USVs and tones for either group (**Fig. 3m,n**), in agreement with previous studies finding no relationship between auditory cortex neuronal tuning for vocalizations and pure tones. This indicates that cortical representations of vocalizations are not simply predicted by frequency tuning.

Responses to USVs, but not pure tones, are plastic over alloparenting experience in both corticocollicular and corticostriatal neurons. The differences in sensitivity to USVs that we observed in CC and CS neurons in co-housed virgins could be innate or experience-dependent, reflecting plasticity in specific corticofugal circuits as female mice gain parenting experience. We tracked activity in CC and CS neurons in virgins (4 virgins with CC cells labeled, 4 CS) before and during co-housing with longitudinal 2-photon calcium imaging (**Fig. 4**). Using the same imaging approach as described for experiments in experienced virgins (**Figs. 3a-d, 4a**), we measured responses to USVs and pure tones in virgins daily, starting before cohousing and for 2-4 days following (**Fig. 4a,b**). Retrieval performance was tested daily after imaging **(Fig. 4b,c**).In each mouse the same imaging field was imaged each day, using blood vessels and ROIs as landmarks to manually match locations (**Fig. 4b**), and ROI masks were considered to belong the same neuron on multiple days if the areas across each day were at least 60% overlapping. Data were excluded if changes in window clarity (for example, due to bone re-growth or inflammation under the window) impacted the ability to image previously-imaged ROIs. In each mouse, 1-2 450 um^2 fields were tracked. Pup retrieval performance improved over days of co-housing relative to day 0 (pre co-housing) in all but one virgin, who retrieved 100% of pups on each day of imaging (**Fig. 4c**). In most mice, performance reached 80% or higher by the 3rd day following the start of co-housing.

In each mouse, USV-evoked dF/F magnitudes and tuning for individual calls were highly variable in tracked neurons (**Fig. 4d**). The majority of ROIs that were detected on multiple days were not present on every day of imaging, and most ROIs with large time-locked responses to specific calls on one day did not maintain the same response magnitude or USV preference

across all days. However, on each day of imaging, both facilitated and suppressed responses were observed for each stimulus, suggesting that despite individual variability over time, responsiveness to USVs was maintained at the population level.

We quantified best early, late, and suppressed responses to USVs (**Fig. 4e**) and tones (**Fig. 4f**) for every ROI that was imaged across at least two days of imaging. To determine whether these metrics were dependent on experience with parental care of pups in CC and CS cells, we compared responses between days with low retrieval performance (days in which ≤40% of pups were retrieved) and high performance (≥80% retrieved). In CC neurons, early best call and late best call-evoked dF/F were larger in magnitude on high-performance days than on low performance days, while suppressed responses were equivalent (Early: $p<0.001$; late: $p<0.001$; suppressed: $p>0.05$; Kruskal Wallis test with Tukey's HSD correction for multiple pairwise comparisons). In CS neurons, early best call responses and suppressed responses were equivalent on low and high performance days (Early: $p>0.05$; suppressed: $p>0.05$), but magnitudes were greater on high performance days for late best call responses (late: $p<0.001$). These results suggest that increased firing rates at a delay following USVs and/or longer durations of USV-evoked increases across in neurons projecting to multiple targets may facilitate perception of USVs in an experience-dependent manner. No differences across low and high retrieval days were observed for tone-evoked activity (**Fig. 4f**), 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 significance to calls, impaired sound localization, or decreased auditory working memory. Follow-up experiments optogenetically suppressing activity during specific phases of behavior, for example before the mouse leaves the nest or during the search for isolated pup, could distinguish between these possibilities.

In experienced virgins, corticocollicular neurons were particularly modulated by USVs compared to corticostriatal. Most strikingly, overall activity levels appeared to upregulated during blocks of USV delivery independent of time-locked, directly call-evoked responses, as well as at long delays relative to call onset and offsets. This may reflect a network-level increase in activity that could involve local microcircuits within auditory cortex, and/or an overall amplification of USV-driven activity in recurrently-linked auditory areas, including feedback loops between auditory cortex, IC, and MGB. The large number of neurons that were suppressed by USVs suggests recruitment of inhibitory circuits within cortex, while increased magnitude and duration of activity in corticocollicular neurons likely modulates IC outputs, as well as temporal integration of feedforward synaptic input to IC from auditory brainstem. Such computations could support recognition and categorization of USVs by enabling integration of acoustic features of individual syllables over the duration of multisyllabic calls, which can last seconds. Alternately, upregulation of activity in auditory cortex could modulate windows for auditory working memory, over duration needed for mice to localize pups. Future experiments recording ensemble activity across multiple stages of the auditory pathway during USV playback, as well as in freely-moving animals during retrieval, could further elucidate the relevance of auditory corticofugal modulation for perception of vocalizations and linking of perception with action.

Methods

Animals

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

2-photon in vivo calcium imaging

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

using stereotactic coordinates (IC: -5 mm 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 retrograde AAV-Syn-jGCaMP8f-WPRE virus (Addgene) was injected using a Nanoject. Auditory cortex was located using skull and blood vessel landmarks, and a 3 mm area of skull was removed. Dura was left intact, and a circular glass window (FST/Harvard apparatus, thickness 0.1 mm) was applied to the brain surface and fastened to the skull using a combination of cyanoacrylate glue and dental cement. A custom stainless steel headplate was attached to the skull using metabond and all exposed skull was covered. Mice recovered from anesthesia on a heating pad and were singly-housed for 24 hours with access to wet food. Additional doses of dexamethasone were administered daily for 2-5 days following surgery to facilitate healing.

Imaging: Imaging experiments began 3 to 4 weeks after surgery using a 2-photon microscope with resonant galvo scanners and a 20x Nikon objective (Moveable Objective Microscope, Sutter Instruments). Mice were stabilized for imaging by clamping the headplate to a custom holder (Thorlabs parts) and positioning the mouse in a length of acrylic tube. Mice were acclimated to head-fixation daily for several days prior to imaging for increasing durations starting with 5 minutes. GCaMP-expressing neurons were imaged with an excitation wavelength of 900 nm, and epifluorescence was bandpass filtered. 2D fields of corticostriatal projecting neurons were imaged at depths ranging from 250 to 500 um ventral to the dura, while corticollicular fields ranged from 350 to 800. Image stacks were collected using ScanImage (Vidrio Technologies) at a rate of 30 frames/second for 512x512 pixel images. Imaging data were collected over multiple days beginning prior to the start of co-housing, with identical stimulus sets delivered on each day

of imaging. Auditory stimuli were played using an ultrasonic speaker (TDT) positioned 10 cm from the right ear of the mouse.

In vivo whole-cell and juxtacellular recordings

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

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

Whole-cell and cell-attached electrophysiological recordings: borosilicate glass pipettes (O.D. 1.5 mm, ID 0.86 mm, Sutter Instruments) were pulled into tips of 2.5-5 MΩ resistance. Pipettes were filled with internal solution containing 127 mM K-gluconate, 8 mM KCl, 10 mM phosphocreatine, 10 mM HEPES, 4 mM Mg-ATP, 0.3 mM Na-GTP 527 (osmolality, 285 mOsm; pH 7.2 adjusted with KOH). Tips were mounted on microelectrode holders with AgCl wire. Recordings were obtained using a Multiclamp 700B amplifier and digitized with a 1440B Digidata (Molecular Devices). Data was acquired with Clampex 10.7 (Molecular Devices), low-pass filtered at 1 kHz, high-pass filtered at 100 Hz and digitized at 20 kHz. Pipette tips were lowered 0.5 – 0.8 mm from the pial surface of the brain to target layer 5. Pipette pressure was measured using a manometer (Omega) maintained at 15-20 mbar until the desired depth was reached.

Optogenetic identification of auditory cortex projection neurons: For identification of light sensitive neurons, whole-cell and cell-attached recordings were obtained using a Fiberoptic Light Stimulating Device with 465nm blue LED (A-M Systems, Item# 726500) connected to a Fiber Optic Light guide (A-M Systems, Item# 726527). The optic fiber was inserted into the patch pipette via 1.5 mm O.D. Continuous Optopatcher holder (A-M 547 Systems, Item# 663943). Pulses of blue light were delivered through the optic fiber via Digidata 1440A (Molecular Devices) while recording in cell-attached or whole-cell configuration. Light pulses of 20, 50 or 200 ms duration were delivered with increasing intensity from 20 to 100% of full LED power (3 mW at the tip of the fiber). ChR2+ neurons responded to light pulses with an increase firing rate

at low jitter <10 ms from light onset, while ChR2- neurons were not modulated by blue light, or if multi-synaptically activated, did not respond with low jitter and high reliability across trials.

Behavioral assays with chemogenetics

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

Drug administration: Clozapine-2H20 (CNO) was dissolved in sterile saline. 45-50 minutes prior to a retrieval test session, a 2 mg/kg dose of CNO or the equivalent volume of vehicle (saline) was administered via intraperitoneal injection. CNO and vehicle sessions occurred on the

same day at least 3 hours apart. The experimenter was blinded to treatment. For several mice, additional sessions were performed the following day with a 1 mg/kg dose of CNO.

2-photon imaging data analysis

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

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

computed in the first 3 seconds following stimulus onset, and the greatest magnitude negative dF/F was considered the most suppressed response.

ROI tracking: ROIs were tracked across days by registering Suite2p reference images across pairs of days and computing area overlap of ROI masks using the opensource MATLAB program ROIMatchPub (https://github.com/ransona/ROIMatchPub). ROIs were considered to be the same neuron if 60% of pixels in the masks two masks were overlapping.

Analysis of *in vivo* whole-cell and cell-attached recordings

All analysis was performed using custom MATLAB scripts except where noted. Voltage traces for each cell were acquired using clampex, then imported into MATLAB using the open source program abfload. For spike rate and timing measurements, voltage traces were highpass filtered above 100 Hz and thresholded manually to obtain binary spike trains. Spike time was recorded as the timing of the first threshold crossing. Firing rates were computed using 10 ms-wide bins and averaged over trials of repeated stimuli to obtain PSTHs. Stimulus-evoked firing rates were computed in 2 seconds following stimulus onset and a baseline period averaged over 1.5 seconds was subtracted.

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).

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

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

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

Figure Legends

Figure 1. Chemogenetic suppression of left auditory cortex projection neurons impairs pup retrieval in experts. (a) Viral injection strategy for targeting of corticostriatal (left) and corticocollicular (right) auditory cortex neurons. AAVrg-EF1a-Cre was injected into posterior striatum or inferior colliculus (IC) using stereotactic coordinates for targeting. AAV1-hSyn-DIO-hM4D(Gi)-mCherry was injected into auditory cortex with targeting relative to anatomical landmarks (cranial sutures and blood vessels). **(b)** Example histological confirmation of injection targeting to posterior striatum (P. Striatum) and expression in auditory cortex. Left, epifluorescence image of a coronal section containing posterior striatum. Right, auditory cortex. DAPI staining depicted in cyan, mCherry in red. Insets: higher magnification images of areas within white squares showing mCherry expression in fibers and cell bodies. **(c)** Example histological confirmation of expression for IC injections; panels as described in **b. (d)** Schematic of pup retrieval assay performed after administration of CNO or vehicle. **(e)** Fraction of pups retrieved under each condition. Thin lines denote performance of individual animals. Thick lines denote means for each group; vertical lines denote SEM. Black: performance in RBP4-Cre mice with AAV1-hSyn-DIO-hM4D(gi)-mCherry injected in left auditory cortex to suppress activity in the bulk of layer 5 excitatory neurons (N=6 experienced virgins). Purple: mice with hM4D(gi) targeted to corticostriatal (CS) neurons in left auditory cortex (N=7; 3 dams and 4 experienced virgins). Green: mice with hM4D(gi) targeted to corticocollicular (CC) neurons in left auditory cortex (N=5 experienced virgins). Gray: sham-injected control group. AAV1-hSyn-DIO-mCherry was targeted to corticostriatal (thin purple lines, N=2 experienced virgins) or corticocollicular (thin green lines, N=2 experienced virgins) neurons using methods as described for experimental groups. Asterisk, $p<0.05$, paired 2-tailed Student's t-test.

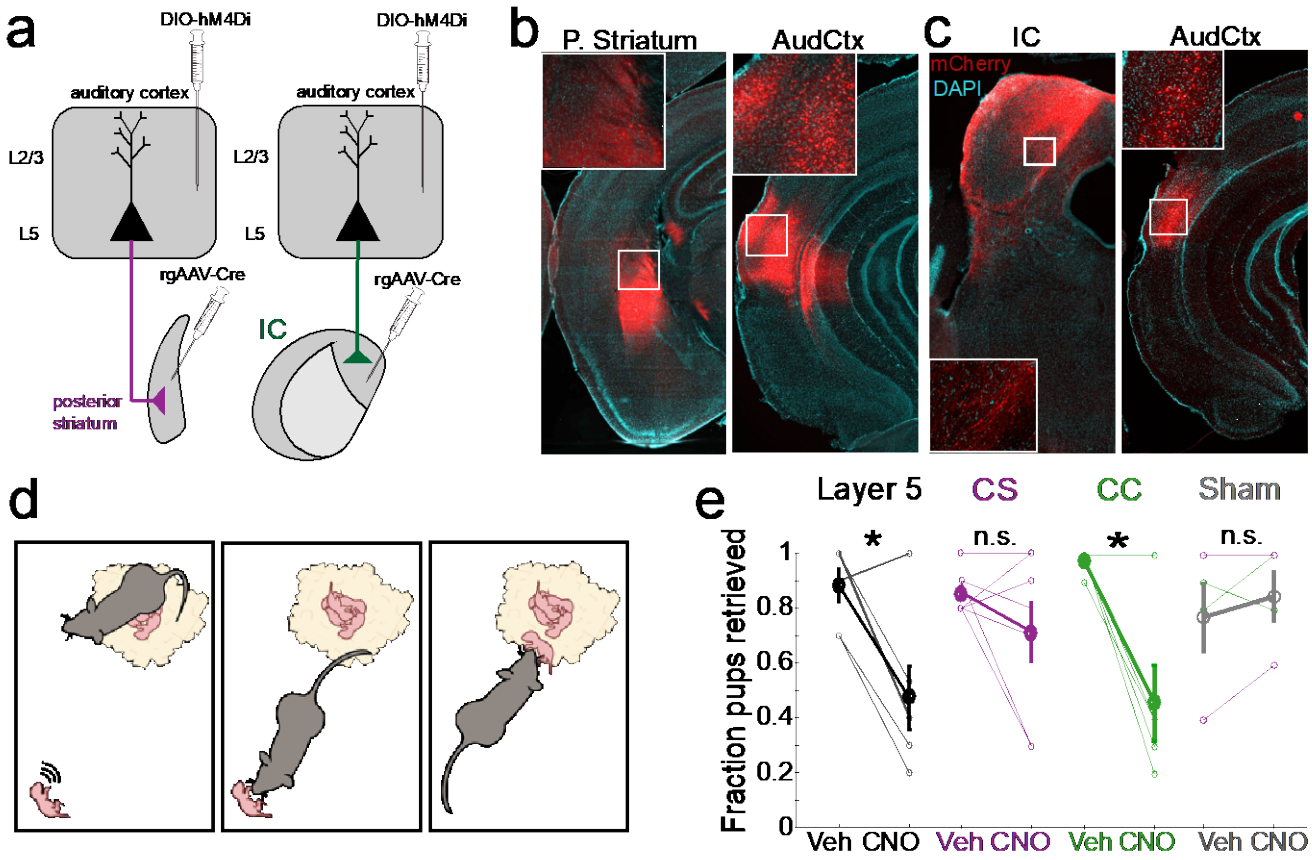
Figure 2. In vivo opto-tagged whole-cell recordings. **(a)**, Virus recording strategy in awake and anesthetized mice**. (b),** Trace and spike raster from corticostriatal opto-tagged neuron. **(c)**, Current-clamp recording from L5 auditory cortex neuron in an awake mouse. **(d),** Spike rasters and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

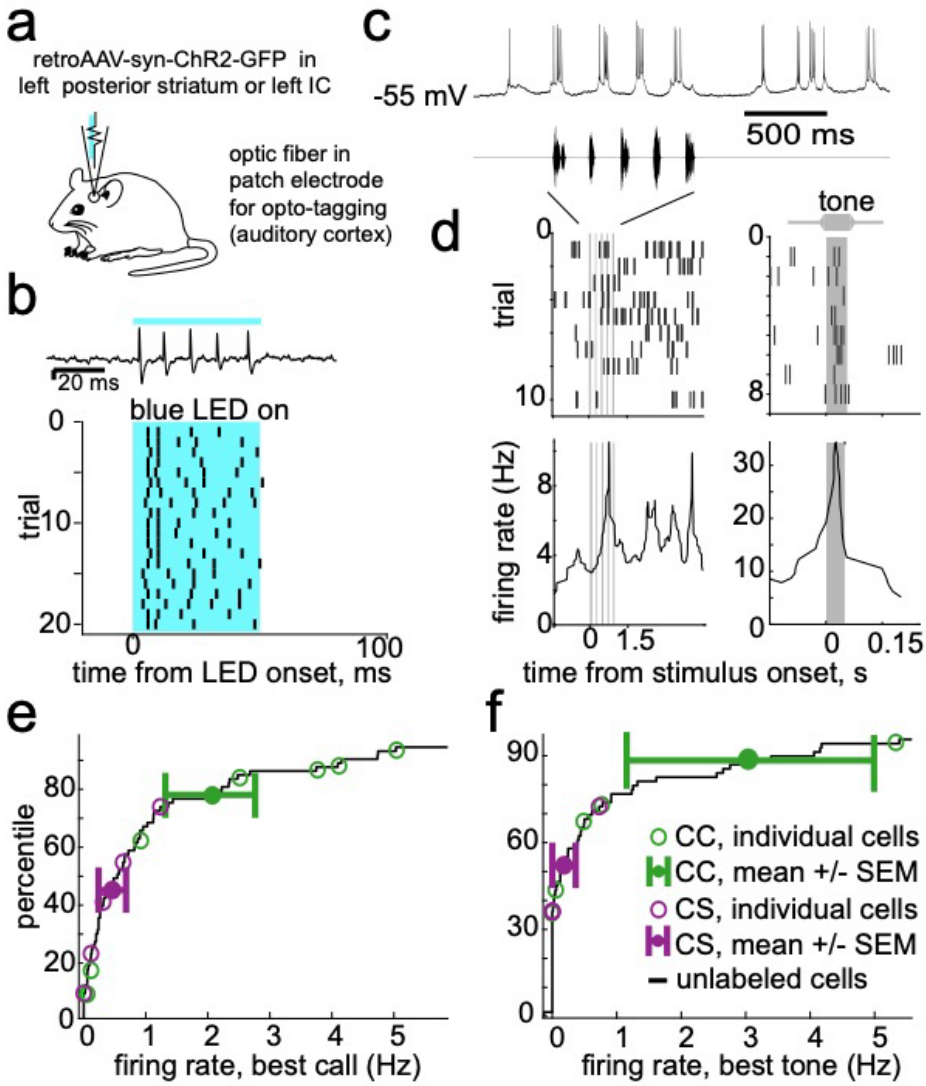
Figure 3. 2-photon in vivo calcium imaging of corticocollicular and corticostriatal-projection neurons during USV and pure tone playback. (a) Left: viral injection strategy for expression of jGCaMP7f for jGCaMP8f in auditory corticostriatal projection neurons. Right panels: histological confirmation 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)** 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. Data were collected on day 3 of cohousing or when animals displayed stable expert performance. **(g)** Mean dF/F for each cell during tone vs. call blocks. Each dot is one cell; 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

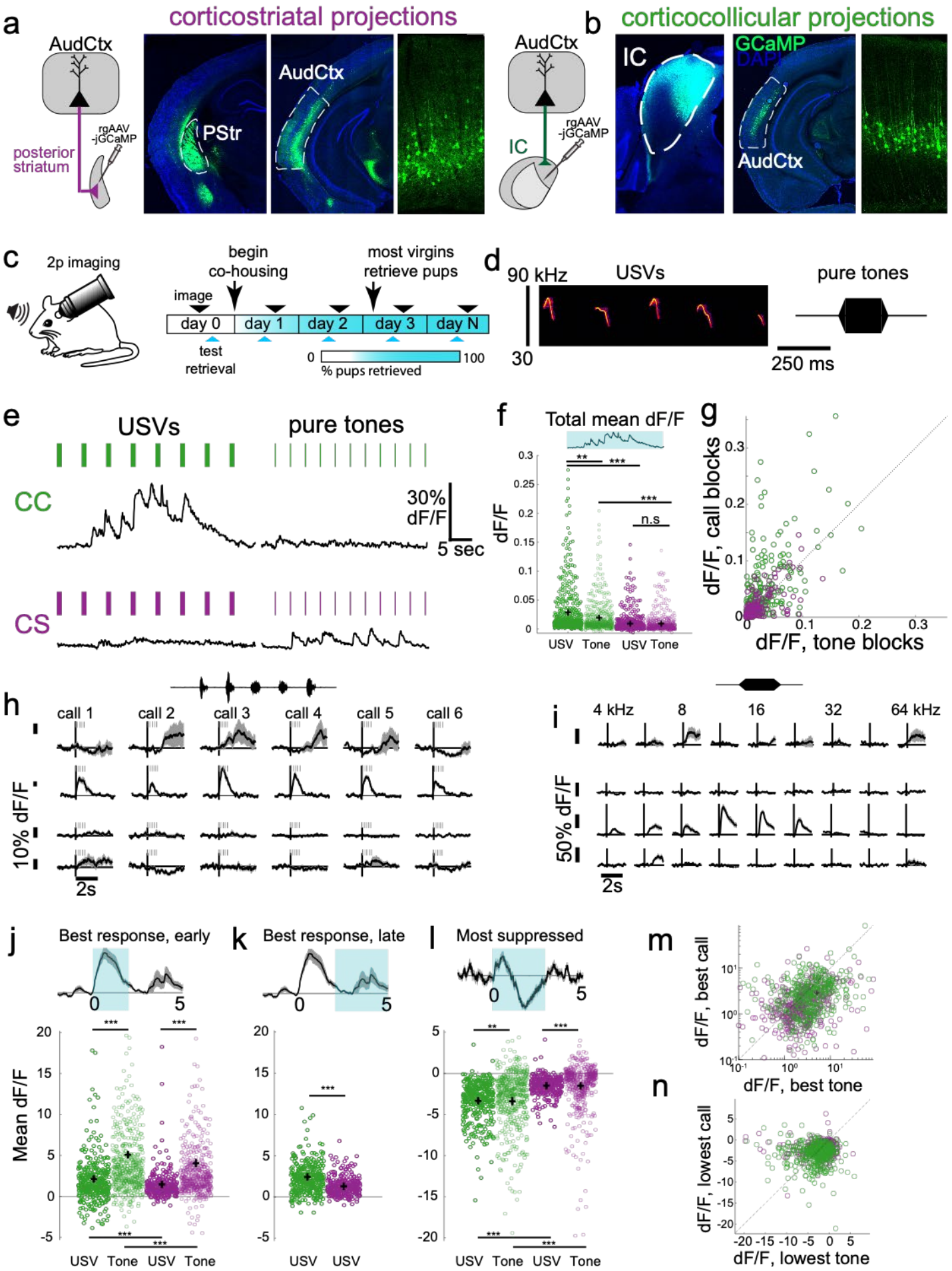
vertical hashmarks denote timing of individual USV syllables. **(i)** responses to 9 quarter-octave spaced pure tones. **(j)** Mean stimulus-evoked dF/F for each neuron's best USV and tone, defined as the USV or tone evoking the largest trial-averaged dF/F during a 2 second period beginning at stimulus 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.

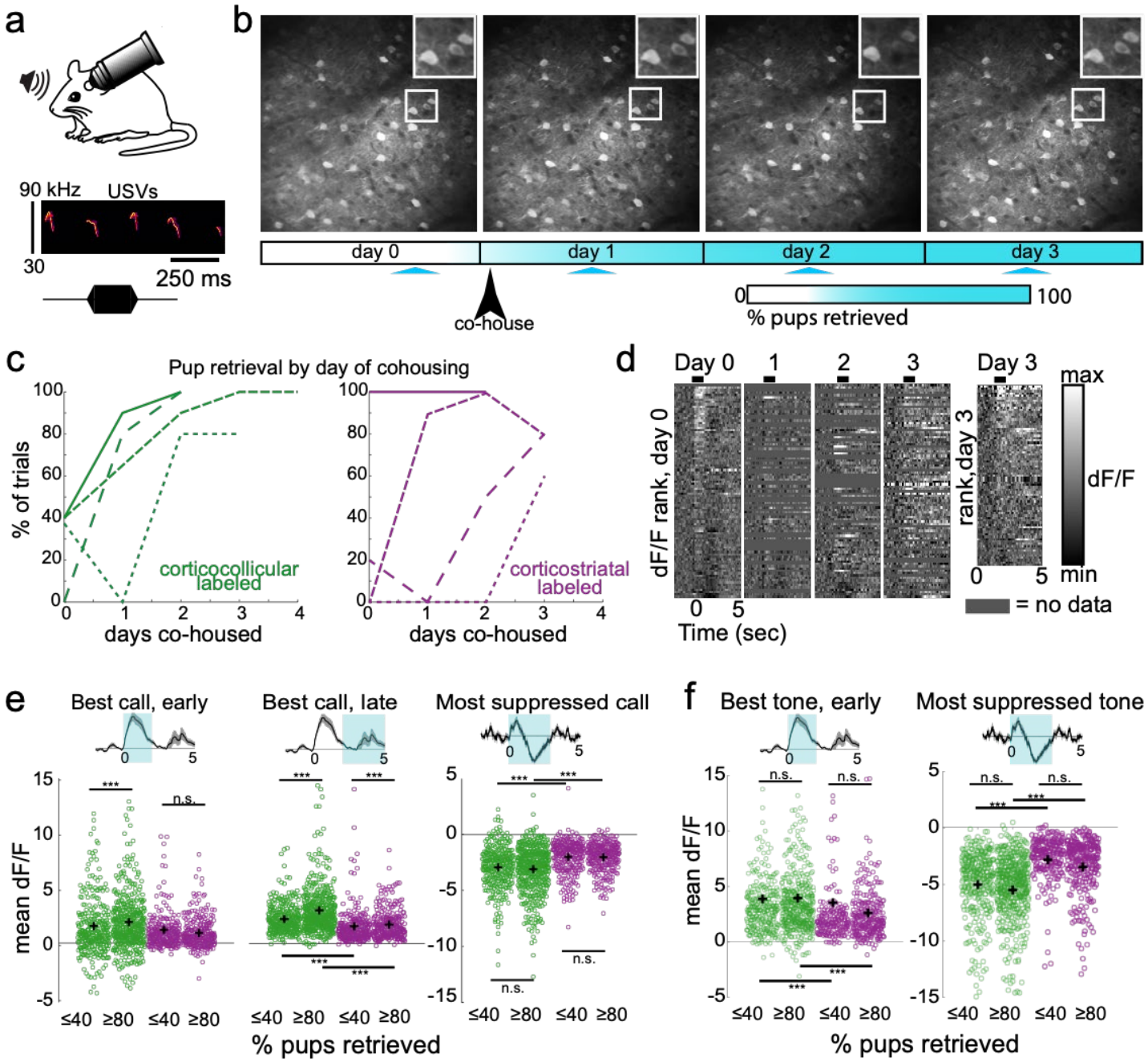
Figure 4. USV- and tone-evoked activity in corticocollicular and corticostriatal neurons tracked over co-housing reveal variation in response magnitudes relatative to retrieval performance. (a) Experiment schematic. **(b)** Average 2-photon images of an example imaging field expressing jGCaMP8f on four consecutive days of imaging. Images were averaged over entire sessions, which comprised ~10^5 individual frames. White squares denote location of zoomed-in images in insets. Lower panel: timeline of image acquitision relative to co-housing. **(c)** Pup retrieval performance in individual virgins over days co-housed with a dam and litter. Retrieval was quantified on each day as percent of pups retrieved over all trials. Left, mice with corticocollicular neurons labeled with jGCaMP8; right, corticostriatal. **(d)** Corticocollicular neurons from one example imaging field tracked over 4 days. Panels, mean responses to best calls for each neuron over 4 imaging days. Neurons are sorted by mean USV-evoked dF/F for each neurons best USV on day 0 (prior to co-housing). Rows: neurons; columns: time bins relative to call onset. Heatmap: normalized dF/F. Rightmost panel: all neurons present on day 3 of cohousing ranked by best-USV-evoked dF/F on day 3. **(e)** USV-evoked mean dF/F of all corticocollicular (green) and corticostriatal (purple) neurons grouped by pup retrieval performance on day of imaging. Low performance was defined as ≤40% of pups retrieved, high performance as ≥80%. Individual circles denote dF/F for each neuron on each day of imaging; black crosses denote population means. Left panel, best USV-evoked dF/F in first 2 seconds following call onset; middle, dF/F 2-5 seconds following onset of best late-response-evoking USV; right, dF/F in first 3 seconds following onset of USV evoking the lowest dF/F. **(f)** Tone-evoked mean dF/F of all corticocollicular and corticostriatal neurons grouped by performance as described for (e). For all panels, 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.


a
DIO-hM4Di
auditory cortex
L2/3
L5
rgAAV-Cre
posterior striatum
DIO-hM4Di
auditory cortex
L2/3
L5
rgAAV-Cre
IC
b
P. Striatum
AudCtx
c
IC
mCherry
DAPI
AudCtx
d
e
Layer 5
CS
CC
Sham
*
n.s.
*
n.s.
Fraction pups retrieved
1
0.8
0.6
0.4
0.2
0
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
IC
rgAAV -jGCaMP
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
30
pure tones
250 ms
e
USVs
pure tones
CC
30% dF/F
5 sec
CS
f
Total mean dF/F
dF/F
**
***
***
n.s
USV
Tone
USV
Tone
g
dF/F, call blocks
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
Mean dF/F
USV
Tone
USV
Tone
k
Best response, late
USV
USV
l
Most suppressed
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
corticocollicular labeled
corticostriatal labeled
days co-housed
d
Day 0
1
2
3
Day 3
dF/F rank, day 0
rank,day 3
max
dF/F
min
= no data
Time (sec)
e
Best call, early
Best call, late
Most suppressed call
mean dF/F
n.s.
≤40 ≥80 ≤40 ≥80
% pups retrieved
f
Best tone, early
Most suppressed tone
mean dF/F
% pups retrieved