

Subcortical coding of predictable and unsupervised sound-context associations

Chi Chen^{a,b,c,d}, Hugo Cruces-Solís^{a,b,c,1}, Alexandra Ertman^{d,e}, Livia de Hoz^{a,d,f,*}

^a Department of Neurogenetics, Max Planck Institute for Multidisciplinary Sciences, Göttingen, Germany

^b International Max Planck Research School for Neurosciences, Göttingen, Germany

^c Göttingen Graduate School of Neurosciences and Molecular Biosciences, Germany

^d Charité Medical University, Neuroscience Research Center, Berlin, Germany

^e International Graduate Program Medical Neurosciences, Charité Medical University, Berlin, Germany

^f Bernstein Center for Computational Neuroscience, Berlin, Germany

ARTICLE INFO

Keywords:

Auditory
Inferior colliculus
Statistical learning
Mouse
Prediction

ABSTRACT

Our environment is made of a myriad of stimuli present in combinations often patterned in predictable ways. For example, there is a strong association between where we are and the sounds we hear. Like many environmental patterns, sound-context associations are learned implicitly, in an unsupervised manner, and are highly informative and predictive of normality. Yet, we know little about where and how unsupervised sound-context associations are coded in the brain. Here we measured plasticity in the auditory midbrain of mice living over days in an enriched task-less environment in which entering a context triggered sound with different degrees of predictability. Plasticity in the auditory midbrain, a hub of auditory input and multimodal feedback, developed over days and reflected learning of contextual information in a manner that depended on the predictability of the sound-context association and not on reinforcement. Plasticity manifested as an increase in response gain and tuning shift that correlated with a general increase in neuronal frequency discrimination. Thus, the auditory midbrain is sensitive to unsupervised predictable sound-context associations, revealing a subcortical engagement in the detection of contextual sounds. By increasing frequency resolution, this detection might facilitate the processing of behaviorally relevant foreground information described to occur in cortical auditory structures.

1. Introduction

Our soundscape is laden with patterned information, much of which can be learned and then predicted. This statistical learning (Reber, 1967) is evident in the strong association that our brain makes between places and sounds. Indeed, as animals move through their environment, they implicitly, without reinforcement, learn and can often predict sound-context pairings. As we enter an environment, being able to quickly assess whether these sound-context pairings match our prediction is important for our course of action, yet we know little about the circuits and mechanisms underlying their representation. In this study, we investigated sound coding in the auditory midbrain of mice living in an enriched environment with different degrees of predictability in the sound-context pairings. The auditory midbrain (inferior colliculus) was chosen because it is a hub of sensory input and multimodal feedback

(Nakamoto et al., 2013; Yudintsev et al., 2021), known to be involved in the processing of soundscapes (Cruces-Solís et al., 2018).

Reinforced learning of sound-reinforcement pairings leads to plasticity in auditory structures. It takes the form of frequency-specific tuning shifts towards a reinforced sound in the auditory cortex (David et al., 2012; Bieszczad and Weinberger, 2010; Blake et al., 2002; Letzkus et al., 2011) and auditory midbrain (Gao and Suga, 2000; Yan et al., 2005; Zhang and Suga, 2005). Evidence for the detection and coding of unsupervised (non-reinforced) contextual sounds also exists. For example, prior exposure to the sounds in a room can improve speech understanding in listeners (Brandewie and Zahorik, 2018) and influence our perception of the world in a pre-attentive manner (McPherson and McDermott, 2020). Similarly, cortical auditory structures show noise-invariant responses, which implies knowledge of the structure of the background noise (Khalighinejad et al., 2019; Mesgarani et al., 2014;

* Corresponding author. Department of Neurogenetics, Max Planck Institute of Experimental Medicine, Göttingen, Germany.
E-mail address: livia.dehoz@charite.de (L. de Hoz).

¹ Current address: CNS Diseases Research, Boehringer Ingelheim Pharma GmbH & Co. KG, Biberach an der Riss, Germany.

Rabinowitz et al., 2012). Our own work has determined that consistent exposure to a sound played predictably in a context as the animal enters (establishing a predictable sound-context association), but without any associated rewards, leads to changes in frequency representation in the inferior colliculus (IC) (Cruces-Solís et al., 2018). Here we aimed to further understand the role of the auditory midbrain in coding the predictability of sound-context associations. To achieve this, we studied the role of different contextual factors in the development of associative plasticity in the auditory midbrain. Specifically, across different replications involving different groups of mice, we investigated the influence of the presence of a reinforcer, the predictability of the sound onset, and the sound exposure time.

2. Results

2.1. Contextual learning elicits IC plasticity independently of reinforcer

To better guide the reader through the results, we will first present in detail the effect that predictable sound-context exposure, with and without reward, has on frequency coding in the inferior colliculus. We separately exposed three groups of mice in the Audiobox (Fig. 1A; TSE, Switzerland), which is an enriched and automatic apparatus made of 2 main compartments connected by corridors. Here, mice live in groups for days with little interaction with the experimenter. Access to both food and water was *ad libitum* and movement across compartments was volitional. Each group had a different sound-context pairing experience

in the sound corner, as detailed in the Methods section and illustrated in Fig. 1A. Importantly, none of the animals were conditioned. We compared two groups of mice exposed to a contextually predictable sound, with a control group that was not exposed to a specific contextual sound. Predictably exposed mice, in the “pred-W” (W: water) and “pred-noW” groups (noW: no water; Fig. 1A middle and right), heard 16 kHz tone pips for the duration of each visit to the corner. Thus, the act of entering the corner (referred to hereafter as “the corner”) was predictive of sound onset. The control group (Fig. 1A left) lived in the Audiobox for an equivalent duration and had access to water in the corner, but was not exposed to the 16 kHz sound. Typically, if the water is in the corner, mice drink for about 60% of the visit time and in about 70% of the visits (Cruces-Solís et al., 2018; de Hoz and Nelken, 2014). Unlike the control and pred-W groups, the pred-noW group had no access to water in the corner and, instead, water bottles were available in the food area, where no sound was ever played (Fig. 1A right). This ensured that, for this group, no association could be made between the location of the sound and that of the water. Then, to ascertain auditory midbrain plasticity, we acutely recorded from the left inferior colliculus (IC) of anesthetized animals that had been sound-exposed in the Audiobox for 6–12 days and compared the activity to that of controls that had been in the Audiobox for a comparable amount of time. For this experiment, we recorded multiunit activity using NeuroNexus multi-electrode arrays (NeuroNexus, MI, USA) with 16 sites linearly arranged and separated by 50 μm (Fig. 1B). The electrode was dorso-ventrally inserted in IC along the tonotopic axis and was placed in two

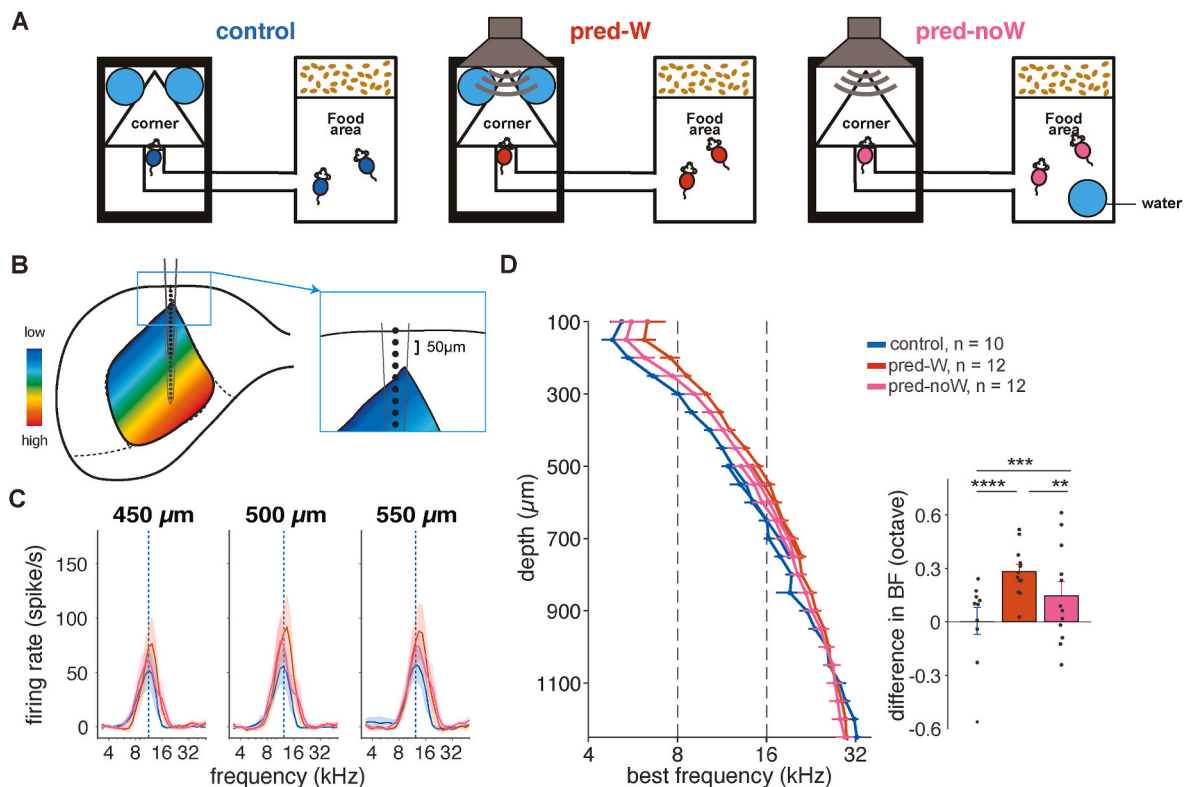


Fig. 1. Exposure to predictable sound, regardless of reinforcer, elicits plasticity in the IC

A, Schematic representation of the Audiobox design for the control (left), pred-W (middle), and pred-noW exposure groups (right). Food was available *ad libitum* in the food area. Water was available in the corner (control and pred-W groups) or in the food area (pred-noW group). **B**, Schematic representation of recording using NeuroNexus 16-electrode arrays from two positions. Insert: Schematic representation of the positioning of the most superficial recording site, aligned to the dura. **C**, Average tuning curves of simultaneously recorded sound-evoked activity at 60 dB for the depth of 450, 500, and 550 μm in the IC of the control (blue), pred-W (red), and pred-noW (pink) groups. **D**, Left: Average best frequency (BF) as a function of recording depths (two-way ANOVA, group $F(2, 761) = 6.88, p = 0.001$, depth $F(25, 761) = 211.4, p = 0$, interaction $F(59, 761) = 0.85, p = 0.77$). Right: Difference in BF relative to the average BF of the control group across the depths between 100 and 900 μm (two-way ANOVA, group $F(2, 525) = 35.64, p = 0$, depth $F(16, 525) = 0.55, p = 0.92$, interaction $F(32, 525) = 0.52, p = 0.99$). Corrected pair comparisons: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$). Error bars represent standard error. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

consecutive steps to cover a range of depths between 0 and 1250 μm (see Methods; Fig. 1B). To characterize frequency tuning, we presented sweeps of 30 ms long pure tones of different frequency-intensity combinations. We computed the count of spikes (threshold: 6 mad below the highpass filtered signal) over an 80 ms window from stimulus onset

(baseline subtracted; see Methods and individual examples of spiking activity in Supplementary Fig. 1A) and built tuning curves for each recording depth and sound intensity (see mean spiking activity across frequency, intensity, and selected depths in Fig. 2C and Supplementary Fig. 1B for these 3 groups and others described later).

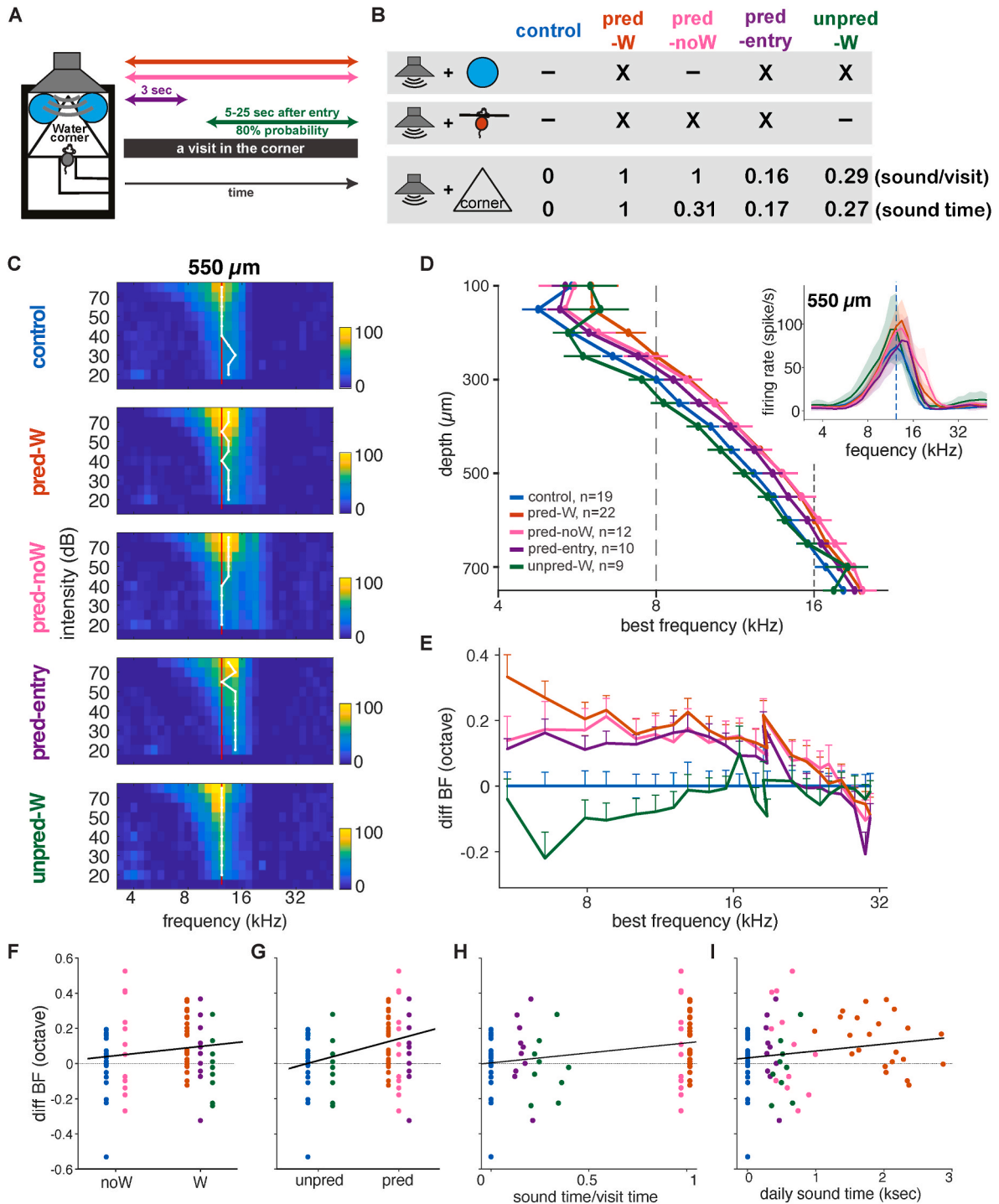


Fig. 2. Predictability of the contextual sound triggers plasticity

A, Schematic representation of the sound exposure protocol. **B**, Table illustrating the sound-context association pattern in different groups indicating whether water available in the corner (top), whether sound was triggered by visit onset (middle), and the duration of sound exposure (bottom). **C**, Group mean tuning curves across intensity (dB-SPL) for all 5 groups in **A** at depth 550 μm . Red line aligns to the control BF at the highest intensity (80 dB). White line indicates intensity-specific BF for that group. **D**, Mean BF per depth for all 5 groups in **A**. Inset: mean BF across mice at depth 550 μm . **E**, Mean difference in BF relative to the mean control BF across BFs. Error bars represent standard error. **F-I** Difference in BF per mouse relative to control mean, as a function of **F**, the presence or absence of water in corner (W vs noW respectively); **G**, the absence or presence of sound at visit onset (unpred vs pred respectively); **H**, the length of sound exposure relative to visit time; or **I**, the total length of sound exposure per day. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

The plastic changes we observed in the pred-W group replicated and extended those we reported in Cruces-Solís et al. (2018). We again found an increase in response gain and a shift in frequency tuning in this group compared to the control group. The increase in sound-evoked responses was evident when comparing the tuning curves of pred-W and control animals, shown in Fig. 1C for 60 dB across example depths of 450, 500, and 550 μm (Fig. 1C; blue vs. red). The increase in gain was observed across all frequency regions (Supplementary Fig. 2G; blue vs. red). The shift in best frequency (BF), calculated here as the mean frequency across intensities that elicited the strongest response at each depth, is also evident in the tuning curves and quantified in Fig. 1D. In addition, we could ascertain that the upward shift, while visible across many depths beyond the exposed frequency band, did not extend throughout the inferior colliculus as we had previously assumed (Cruces-Solís et al., 2018). In fact, below 950 μm (Fig. 1D), the shift was no longer significant (two-way ANOVA, group $F(2, 217) = 1.14$, $p = 0.3$, depth $F(6, 217) = 1.99$, $p = 0.07$, interaction $F(12, 217) = 0.51$, $p = 0.9$).

The key result, however, is that the patterns were similar for animals that were predictably exposed but had no water in the corner (pred-noW group). Indeed, the increase in response gain and shift in BF were also present in the pred-noW group (Fig. 1C and D, Supplementary Fig. 2G). While the presence of water in the corner did not imply a tight temporal association between the sound and the water, for mice only drank sporadically during a given visit and only in about two thirds of the visits (de Hoz and Nelken, 2014), the results confirm that the change in IC responses was not caused by an association between the 16 kHz contextual sound and the presence of water in the corner. Quantification of average evoked responses across tone frequencies confirmed that the gain increase in both pred-W and pred-noW animals relative to control appeared along the IC tonotopic axis (Supplementary Fig. 2G; linear mixed model [two factors: group \times depth]: ctrl vs. pred.-noW $p = 0.016$, depth $p < 0.0001$, ctrl/pred.-W \times depth $p = 0.02$). The increase in evoked responses at greater depths is likely related to the decrease in shank width of the probe with depth (probe: A1x16-3mm-50-177; Claverol-Tinture and Nadasdy, 2004). Nevertheless, probe positioning was consistent across animals and cannot explain the differences between the groups. Similarly, quantification of the changes in BF relative to the control across recording depths between 100 and 750 μm (the first recording, to be comparable with (Cruces-Solís et al., 2018)) also confirmed that the shift in BF was not dependent on the presence of water (Fig. 1D; linear mixed effect model [two factors: group \times depth]: overall group $p = 0.03$, ctrl vs. pred.-W $p = 0$, pred.-W vs. pred.-noW $p = 0.0002$, ctrl vs. pred.-noW $p = 0.0005$). The BF shift in the pred-noW group, while significantly different from the control, was not as strong as in the pred-W group, indicating that while plasticity is not dependent on the water association, its presence might strengthen it. Indeed, while the mice were living in a task-free environment, the mice made less visits to the corner when the water was not there (Supplementary Fig. 2J).

These results replicated and substantially extended our previous finding (Cruces-Solís et al., 2018). We confirmed that exposure to predictable sound-context associations led to an enhancement in overall response gain and a global shift in BF. Importantly, since the effect was still present when the water was removed from the corner, this plasticity was unsupervised as corroborated by the finding that it did not depend on reinforcement.

2.2. Predictability of contextual sounds triggers plasticity

The obvious question is whether the observed changes in gain and tuning are triggered by sound exposure in general or specifically by the predictability of the context-sound association driven by the mouse movement into the corner. To answer this question, we focused on the predictability of the sound onset, which had previously been controlled by the mouse's entry into the corner. We trained two new groups of mice separately in the Audiobox, each run as a separate replication with its own control group. The replication-specific comparisons are shown in

Supplementary Fig. 2A–I. In Fig. 2 we combine different replications. One group was exposed again to pure tones of 16 kHz, but this time in an unpredictable manner ("unpred-W" group; Fig. 2A, green). To make the onset of the sound unpredictable, the tone pips were activated with a variable latency between 5 and 25 s from visit onset and only in 80% of the visits (Fig. 2A green). This means that sound, crucially still present in the corner, was not controlled by the mouse movement and was also not present for the visit's duration. This combination of criteria was crucial to dissociate the effect of mere sound exposure in the corner from the predictability driven by entry into the corner. The other group ("pred-entry" group) only heard the sound at the onset of the visit for a duration of 3 s (Fig. 2A purple), meaning that while the sound was fully predictable by entry into the corner, the mice were exposed to it for shorter amounts of time than mice in the pred-W and pred-noW groups. These two groups showed different patterns of neuronal activity. The response magnitude and BF across depths in the unpred-W group did not differ from those in a control group that was trained in parallel (Supplementary Figs. 2A, B, C, and H) and showed no BF shift across intensities and depths compared to the mean control (Fig. 2C–E). Quantification of the average difference in BF with respect to the control group across the depths between 100 and 750 μm confirmed that there was no shift in BFs in the unpred-W group (Supplementary Fig. 1C insert; linear mixed model, $p = 0.69$). The pred-entry group, on the other hand, did show an increase in sound-evoked responses and BF shift in the IC comparable to that of the pred-W group trained in parallel (Supplementary Figs. 2D, E, F, and I) and comparable in intensity and depth to the mean pred-W group (Fig. 2C–E). The effect of sound intensity on BF (Tao et al., 2017) is observed in Fig. 2C for the control group. Indeed sound intensity had an effect on BF across groups. Using a linear mixed effect model to predict the change in BF with two factors [(water + onset + sound/visit time) \times dB], where "onset" means the sound was present upon entry, reveals a significant effect of dB ($p < 0.0001$), water ($p < 0.0001$), and onset ($p < 0.0001$) and interaction between dB and water ($p < 0.001$), onset ($p < 0.0135$) and sound/visit time ($p = 0.0216$). The shift upward in BF for the predictable groups was generally present across depths and intensities. The average BFs across depths were also comparable (Fig. 2G–F; Supplementary Fig. 2F; linear mixed model, $p = 0.08$).

The BF depth profile of all the groups is shown together in Fig. 2D. We investigated how BF shifted along the tonotopic axis by plotting the change in BF as a function of the BF of the control group (Fig. 2E). Despite varying durations spent in the corner or hearing the sound (Supplementary Fig. 2J and K), the BF shift in the predictable groups (pred-W, pred-noW, pred-entry) is similar in magnitude for frequencies below 16 kHz (>1 octave) and extends for frequencies above 16 kHz until 25 kHz ($>1/2$ octave). Using a linear mixed effect model to predict the change in BF with two factors [(water + onset + context) \times BF], where "onset" means the sound was present upon entry, and "context" means the sound could be present until the end of the visit, reveals a significant effect of onset ($p < 0.0001$) and interaction between onset and BF ($p < 0.0001$). A similar pattern of change was observed when comparing the pred-noW with the unpred-noW group (Fig. 2F; linear mixed model [two factors: group \times BF]: group $p < 0.0001$, BF $p = 1$, interaction $p < 0.0001$). When we analyzed the data by comparing groups trained with different sound-context association patterns (Fig. 2D–E), we found that predictability (sound triggered by visit onset; Fig. 2G), but neither reinforcement (water in the corner, Fig. 2F) nor the duration of sound exposure (Fig. 2H and I), determined the broad shift in BF in the IC (Fig. 2E; the linear mixed model on three factors [water \times onset \times context], a significant effect of onset $p < 0.0001$). This shift was also visible in the tuning curves (Fig. 2D inset).

To sum up, we exposed five groups of mice in the Audiobox with different sound exposure paradigms and examined the effect of three possible factors: reinforcement (water in the corner), mouse entry (sound-visit onset association), and contextual value (sound present for the duration of the visit), on IC plasticity induced by sound exposure in

the Audiobox (Fig. 2A). Again, it is important to note that the groups were neither positively nor negatively conditioned, only exposed. Behaviorally, consistent with previous data (Cruces-Solís et al., 2018), sound exposure, independent of its duration and predictability, did not change the animals' behavior when compared to the control group (Fig. 2G–H). To conclude, sound exposure elicited a broad and upward frequency shift in BFs along the IC tonotopic axis only when the sound onset was predicted by visit onset, and this shift was independent of the presence of a reinforcer in the corner and the duration of the sound exposure (Fig. 2F–I).

2.3. The BF shift is broad but not ubiquitous

To better understand the frequency-specificity of the BF shift, we compared the effect of predictable and unpredictable sound-context associations in the absence of water, using Neuropixels probes (Imec, Belgium) to sample the IC in depth (Fig. 3A). We trained two new groups of mice without water in the corner, namely a pred-noW and an "unpred-noW" exposure group (Fig. 3B). In addition to allowing the sampling of deep regions of the IC, the Neuropixels recordings also revealed other interesting features (Fig. 3C). For example, at greater depths, neurons often responded to low frequencies in addition to the depth-expected frequencies, similar to the long tail of the auditory nerve (Kiang and Moxon, 1974). Consistent with the previous results, mice in the pred-noW group showed an upwards shift in BF across a wide range of depths compared to the unpred-noW group (Fig. 3C). It was, however, difficult to quantify the shift from this heterogeneous patterns of BFs. We therefore used a Gaussian mixture model (see Methods) to isolate the cluster of units that fell on the tonotopic map and construct Fig. 3D–E. A linear mixed model [two factors: group \times depth]: group $p = 0.016$, depth $p = 0.002$, interaction $p = 0.013$, again confirming predictability of the sound-context association as the key variable determining the presence of plasticity. As is evident in the plots (Fig. 3C–E), the shift is constrained in depth, and therefore, given the depth alignment of the tonotopic map, in frequency range.

2.4. Increased frequency discriminability coupled to BF shifts

We were surprised that the shift in BF was always upward independently of whether the BFs were below or above the exposed 16 kHz tone, and wondered whether this reflected some other, possibly more critical, change in frequency representation. First, we looked at structural population vectors: the response to a given frequency across the IC (Cruces-Solís et al., 2018), which reflect the ongoing responses to sounds in the IC. For example, when the 16 kHz sound is presented, cells at several depths in the IC respond simultaneously with different strengths (Fig. 4A). A different pattern will appear when a frequency farther away from 16 kHz is presented. Population vectors obtained with NeuroNexus probes, plotted side by side for each group (Fig. 4A), illustrate the increasing BF with depth characteristic of the tonotopic map of IC, but they also reveal further interesting features. For example, at greater depths, neurons often responded to low frequencies in addition to the depth-appropriate frequencies, an effect that is, as mentioned, even more evident in the more extensive Neuropixels recordings (Fig. 3C). To quantify the dissimilarity of activity patterns between responses to different frequencies, we measured in each mouse the Euclidean distance (Edelman, 1998) between frequency-specific population vectors at different ΔF (frequency difference in octave; Fig. 4B) and found increased separation in the predictable groups compared to the control and unpredictable groups (Fig. 4B and Supplementary Fig. 3A). This effect was frequency- and ΔF -dependent. Modeling a linear fixed effect model to predict the boundaries where Euclidean distance is above 2 (about 80% of the max value; Fig. 4B) with two factors [(water + onset + context) \times frequency], revealed significant effects of frequency ($p < 0.0001$ for both), onset ($p < 0.0016$), and context ($p = 0.0495$) and interactions between the to-be-compared frequencies ($p = 0$), and onset and test frequency ($p = 0.0007$). This increase was not limited to comparisons with the exposed frequency and, like the BF shift, extended above and below. Independently of ΔF , the discriminability in all the groups, including the control, was larger for frequencies near and above 16 kHz but weaker for frequencies below (Supplementary Fig. 3B). The

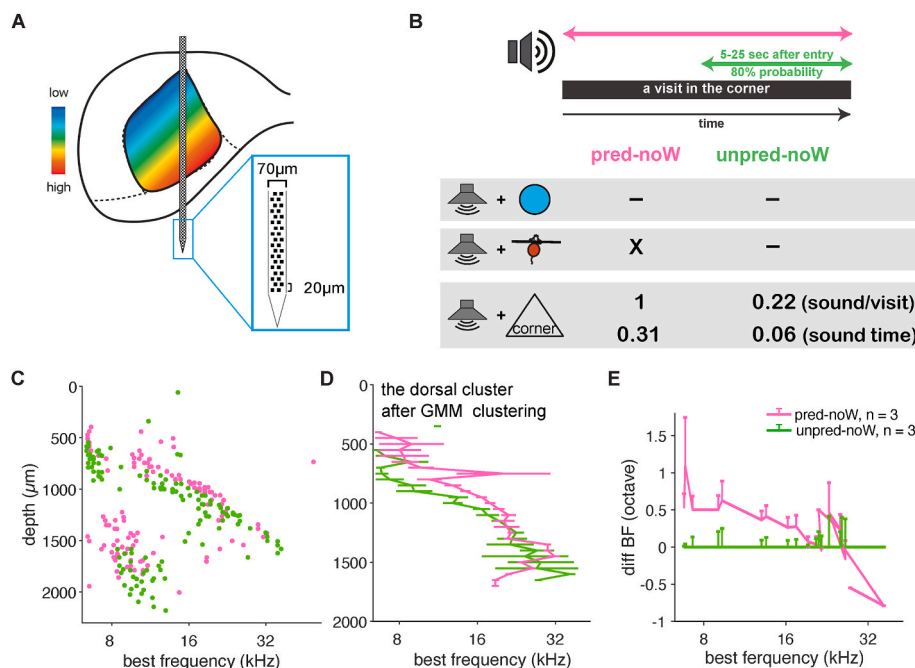


Fig. 3. BF shift is not ubiquitous.

A, Schematic representation of recording using Neuropixels. **B**, Table illustrating the sound-context association pattern indicating whether water was present in the corner (top), whether sound was triggered by visit onset (middle), and the relative duration of sound exposure (bottom). **C**, BF of individual units recorded from different depths of the IC. **D**, Average BF as a function of depths for GMM clustered units along the tonotopic axis. Linear mixed effect model with two factors [group \times depth]: $p_{\text{group}} = 0.016$, $p_{\text{depth}} = 0.002$, $p_{\text{group} \times \text{depth}} = 0.013$. **E**, Average change in BF relative to the average BF of the control group was frequency specific.

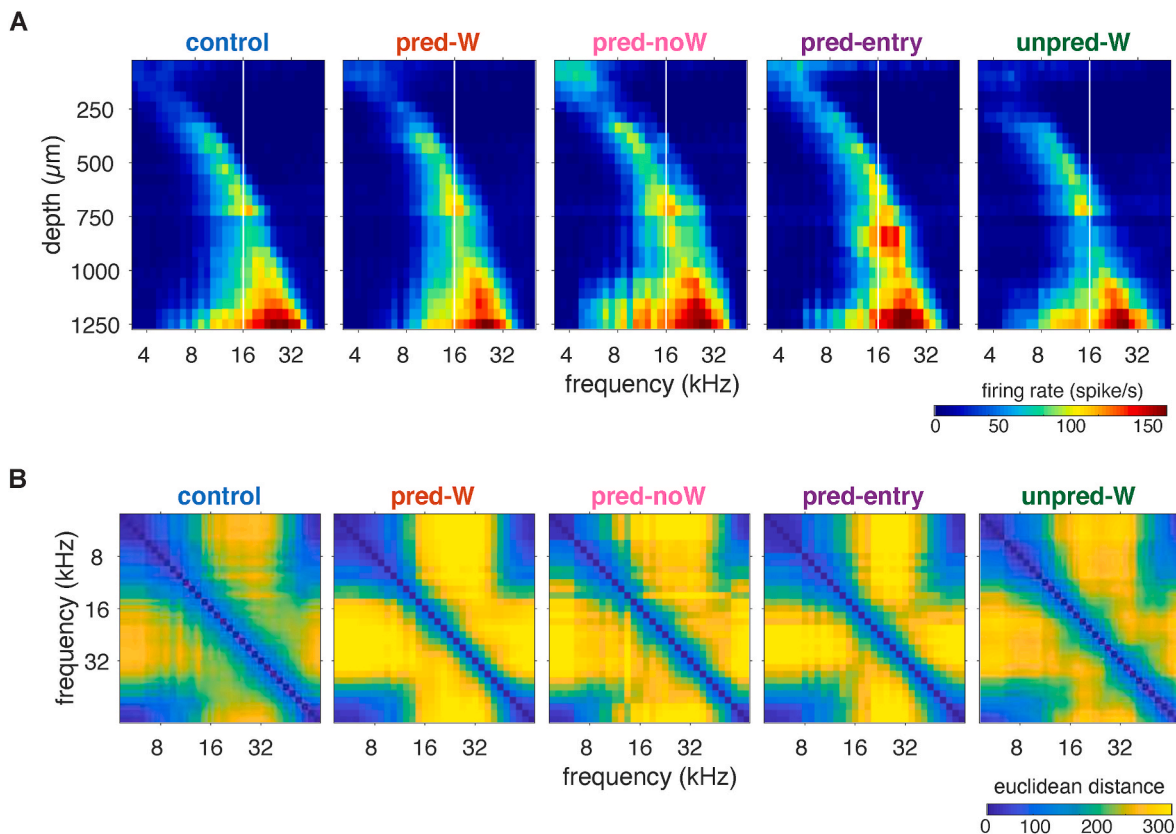


Fig. 4. BF Shifts were accompanied by an increase in neuronal frequency discrimination

A, Mean population vectors, as response magnitude in spikes/s to a given frequency across the IC for the 5 groups as in Fig. 2A. **B,** The dissimilarity of activity patterns measured by Euclidean distance between the population vectors of frequency pairs for the same groups.

dissimilarity of activity patterns not surprisingly increased with ΔF , but it did so more in the predictable groups (Supplementary Fig. 3F). We conclude that predictable exposure increases frequency discriminability.

2.5. The BFs remained unchanged during the initial days of sound exposure

All the data shown thus far were obtained from mice that had experienced sound-context associations for 6–12 days in the Audiobox. We were interested in determining whether plasticity, in the form of response gain changes and shifts in frequency representation, was observable already in the early days of exposure. For that purpose, we recorded from mice that had been in the Audiobox for only 1–5 days. We compared control and pred-W groups and found that both the differences in gain and BF described earlier showed a different pattern in the early days. During the initial exposure, the control group showed generally stronger evoked activity than the pred-W group across a wide range of depths (Fig. 5A and Supplementary 4D; linear mixed model for the ventral area [two factors: group \times depth]: group $p = 0.005$, depth $p = 0.007$, interaction $p = 0.036$). The BF, on the other hand, was not significantly different across depths between groups (Fig. 5B). A linear mixed model [two factors: group \times depth] reveals no significant effect of group ($p < 0.05$). This demonstrates that changes in response gain occur independently of changes in BF shifts. For comparison, in Fig. 5C we include data from the late exposure mice published in Cruces-Solís et al. (2018), which were run in a parallel experiment. The combined analysis of recordings made during early (1–5 days) and late (6–12 days) exposure revealed that, indeed, the BF shift in the pred-W exposed group developed over days and had different dynamics from that of the control group (Fig. 5C; linear mixed model with three factors [exposure \times training days \times depth]: exposure $p = 0.03$, depth $p < 0.0001$, training

day $p = 0.31$, exposure \times days $p = 0.13$). Importantly, the absence of a BF shift in the early days was also observed in mice belonging to the pred-noW group (Supplementary Fig. 4B). Additionally, the progression of the BF shift over time (Supplementary Fig. 4C), indicated that the time course of plasticity was independent of reward. Thus, the plastic changes observed in the late days (6–12) of exposure developed slowly and showed a different pattern compared to the early days.

2.6. Corticofugal input selectively suppressed collicular activity during early days of predictable sound exposure

Direct cortical feedback through descending projections from deep layers of the auditory cortex (Nakamoto et al., 2013; Yudintsev et al., 2021) has been shown to induce collicular plasticity (Gao and Suga, 2000; Suga et al., 2002; Yan and Ehret, 2001), modulate collicular sensory filters (Syka and Popelár, 1984), and play a role in learning (Bajo et al., 2010; Bajo and King, 2012). We previously found that while auditory cortical inactivation (Fig. 5D) led to disinhibition in the inferior colliculus, it had no differential effect on predictable and control groups exposed for 6–12 days (Cruces-Solís et al., 2018). In both, it led to a comparable increase in response gain (current Fig. 5F and H) in the absence of an effect on the upward BF shift (Fig. 5J). Thus, we concluded that corticofugal input modulated response gain without affecting BF shifts, and that the maintenance of the plastic changes triggered by predictable sound exposure was not dependent on cortical feedback. To test whether corticofugal inputs affect predictable learning during the initial sound exposure, we performed simultaneous inactivation of the auditory cortex with muscimol and recordings in the inferior colliculus as before, on a subset of control and predictable animals that had been exposed for only 1–5 days (see Methods, Fig. 5D). Unlike in the late exposure days, the effect of cortical inactivation on IC activity was

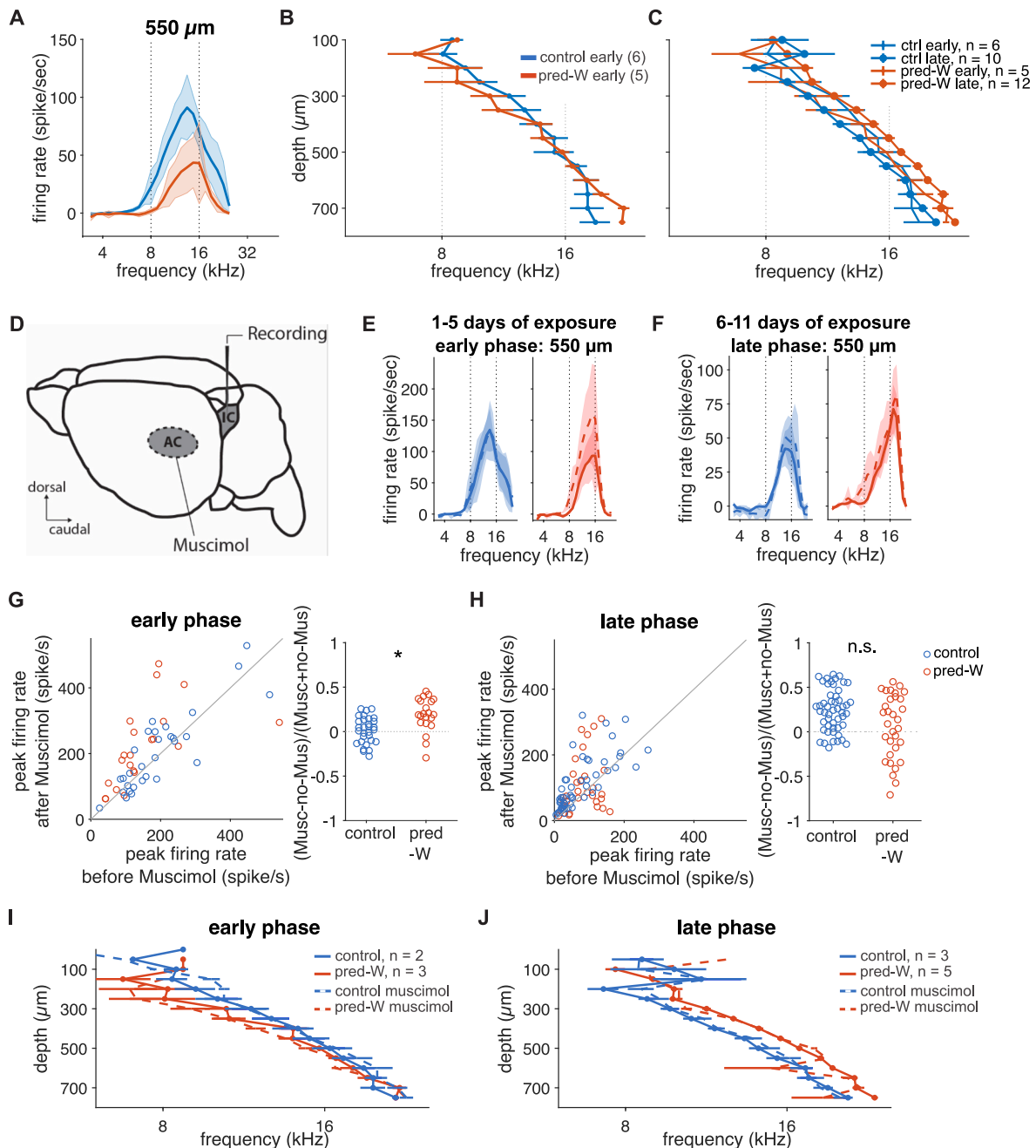


Fig. 5. Corticofugal input suppressed collicular activity during the early days of predictable sound exposure

A, Average tuning curves at 70 dB, at 550 μm for control (blue) and predictable (right) groups during the early days of sound exposure. **B**, Average BF as a function of recording depths for mice during the early days of sound exposure. **C**, Average BF as a function of recording depths for mice during the early (same as **B**) and late days of sound exposure. **D**, Schematic representation of simultaneous inactivation of the AC and recording in the IC. **E**, Average tuning curves of recorded sound-evoked activity, before (solid lines) and after (dashed lines) muscimol application in the AC, at 70 dB for 550 μm in the IC of the control (blue) and predictable (red) group. The recording was performed in mice with no more than 5 days of sound exposure. **F**, Similar to **E**, for mice with 6–11 days of sound exposure. **G**, **Left**, pairwise comparison of peak activity before and after cortical inactivation, for mice in the early phase of sound exposure. Wilcoxon signed rank test comparing before and after cortical inactivation: control $p = 0.30$, predictable $p = 0.003$. **Right**, cortical inactivation led to stronger disinhibition in the predictable group (linear mixed effect model with the group factor: $p_{\text{group}} = 0.016$). Animals and recording sites: control $n = 2$ and 21, predictable $n = 3$ and 30. **H**, Similar to **G**, for the late phase. Wilcoxon signed rank test comparing before and after cortical inactivation: control $p < 0.0001$, predictable $p = 0.19$. Cortical inactivation led to a comparable increase in response gain (linear mixed effect model with the group factor: $p_{\text{group}} = 0.24$). Animals and recording sites: control $n = 3$ and 32, predictable $n = 5$ and 51. **I**, **J**, Average BF as a function of recording depths for mice during the early and late phase of sound exposure respectively, before (solid lines) and after (dashed lines) muscimol application in the AC. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

different in control and predictable groups during the early days of exposure (Fig. 5E and G). Response gain was differently affected by cortical inactivation (Fig. 5E and G-left; Wilcoxon signed rank test comparing before and after cortical inactivation: control $p = 0.30$, predictable $p = 0.003$), leading to stronger disinhibition in the

predictable group (Fig. 5G right; linear mixed effect model with the group factor: $p = 0.02$). Regarding the BF shift, similar to during the late phase of exposure (Fig. 5J), cortical inactivation had no effect on BF during the early phase of exposure (Fig. 5I), supporting again the idea that BF shifts are independent of response gain changes as well as

independent from cortical feedback. Thus, the auditory cortex had a stronger effect on IC responses of mice exposed to predictable context-sound associations than in control mice, but only during the early days of exposure. This suggests that the weaker responses, relative to control, that we observed in these early phases (Fig. 5A and Supplementary 4D) might be due in part to cortical inhibition that is stronger in the predictable group.

3. Discussion

What we hear is often predicted by where we are. Sound-context associations are informative and predictive of normality. These associations are learned implicitly, in an unsupervised manner. Aiming to gain a deeper understanding of how unsupervised learning of sound-context associations is coded, we measured changes in sound evoked responses in the inferior colliculus, a hub of auditory input as well as multimodal feedback, in mice that lived in an enriched environment, the Audiobox. These mice were exposed to sound-context associations with group-specific degree of predictability over days. We found that living in an enriched environment with predictable sound-context associations led to an upward and broad shift in best frequency and an increase in response gain within the inferior colliculus. These plastic changes were accompanied by improved neuronal frequency discrimination and were not qualitatively dependent on reinforcement. Reward did however increase the magnitude of the observed changes. Importantly, plasticity was observed in all exposure conditions where the sound was activated as the mouse entered the corner, regardless of sound exposure time or the presence of reinforcement. Plasticity was not observed when the sound was activated at random times after visit onset. Therefore, we conclude that the auditory midbrain is sensitive to the predictability of contextual sounds, and we discuss the consequences this might have for processing background and foreground acoustic information.

The results expand upon previous findings from the laboratory where we found an increase in gain and, what we interpreted as a frequency-unspecific BF shift, in animals exposed to predictable sounds in the presence of water (Cruces-Solís et al., 2018). The sampling of a wider depth range with both Neuronexus and Neuropixels probes in the current study revealed that the BF shift is broad but not ubiquitous. Similarly, the comparison between the effects of different forms of sound-context association in this series of experiments, in the presence and absence of reinforcement, allows us to conclude that plasticity is associated specifically with the unsupervised learning of predictable sounds. While the presence of narrow tubes in the Audiobox contributed to the natural exploration and behaviour of the mice, it precluded us from performing the recordings in awake mice. Recordings were performed in anesthetized animals removed from the enriched environment at different time points from the beginning of sound exposure (early groups: 1–5 days; late groups: 6 and 12 days for Neuronexus recordings and 19–21 days for Neuropixels). Anesthesia can affect neuronal responses (Jing et al., 2021), but is unlikely to have different effects in different exposure groups. Our conclusions thus are drawn from the comparison across groups of mice that experienced different degrees of sound-context predictability and reflect comparative changes in sensory processing.

We have shown before that, when mice are exposed to non-reinforced sound-context associations, as in the current study, they do not modify their behaviour in a sound-dependent manner (Cruces-Solís et al., 2018; de Hoz and Nelken, 2014). Also in this study, animals exposed to a predictable or an unpredictable sound spent comparable daily time lengths in the corner as mice exposed to no sound at all. The absence of water in the corner did influence the motivation to go to the corner, leading to fewer visits and shorter exposure time. While the frequency coding changes in this group did not differ qualitatively from the group exposed predictably with water, the observed effects were smaller. This is probably not likely due to reduced sound exposure resulting from fewer visits to the corner in this group. The group that

only heard the sound upon entry in the corner (pred-entry) was also exposed to the sound for a shorter duration, yet showed effects comparable to those of the predictable group with water. Thus, reinforcement might modulate the size of the effect. The lack of noticeable change in behavior after unsupervised predictable sound exposure should not be confused with a lack of learning about the sound. We have observed across many studies that non-reinforced exposure in the Audiobox leads to latent learning that is only expressed during subsequent reinforced learning (Cruces-Solís et al., 2018; de Hoz and Nelken, 2014), with different time courses than reinforced learning in the same apparatus (Chen et al., 2019, Chen and de Hoz, 2023; de Hoz and Nelken, 2014), and effects on reinforced associative plasticity in the auditory cortex (de Hoz et al., 2022). Unsupervised exposure does lead to unsupervised learning and we believe that the changes in midbrain auditory responses are a correlate of this learning. Indeed, we found that neuronal generalization during unsupervised exposure paralleled behavioral generalization measured through pre-pulse inhibition of the startle reflex (Cruces-Solís et al., 2018).

In the experiments described here, the sound is controlled by the mouse as it enters the corner. In the predictable groups, this entry activates the sound, whether for the length of the visit or only for 3 s. In the unpredictable groups, this entry is not coupled with sound onset but activates a clock that determines the random interval after which the sound will begin. A BF shift was only observed in animals in which the sound was activated upon entry, even when only for the first 3 s from visit onset. While we did not test a group that activated the sound at a fixed delay upon entry, the data overall suggest that predictability could be determined by sound changes triggered through entry into a defined environment, since this is how sound-context associations are typically defined in natural situations. Time-averaged statistics of contextual sounds, such as textures, converge within 2.5–3 s (McDermott et al., 2013), supporting the notion that very short contextual sounds, more complex than a tone, might lead to the same effects. We did not test shorter sounds or longer sounds with slowly converging statistics.

Frequency-specific shifts in tuning are a hallmark of auditory plasticity (Bakin and Weinberger, 1990; Blake et al., 2002; Yan et al., 2005; Zhang and Suga, 2005). Typically, when an animal is trained in a frequency conditioning paradigm, a shift in BF is observed that is localized to the population of neurons that respond to the conditioned frequency. This shift enables and parallels an increase in the representation of the relevant frequency (Blake et al., 2002). What is surprising in our results is that the shift also occurs in populations of neurons that do not respond to the exposed frequency (16 kHz) and is always an upward shift, independently of whether the BF is below or above 16 kHz. We have shown before that the upward shift could also be triggered through exposure to frequencies between 13 and 16 kHz but not by exposure to 8 kHz (see Cruces-Solís et al., 2018; cf Supplementary Fig. 4A), but the relationship between the identity of the exposure frequency and the effect remains to be understood. In the current work, by sampling across a wider range of depths, we could conclude that the shift is frequency-specific but broad, extending to about $2/3$ of an octave above and more than 1 octave below the exposed frequency (Fig. 3D–E). Unlike BF shifts triggered by reinforced learning in the inferior colliculus (Zhang and Suga, 2005), the shift we observed was independent of cortical feedback at all time points. Overall, the data suggest that the mechanisms underlying this positive and broad shift are different from those observed under reinforced learning. The shift is likely tightly related to the observed changes in frequency representation in the form of increased frequency discriminability for the predictable groups. Thus, while reinforced learning leads to a local shift towards the relevant frequency, with a concomitant decrease in discriminability for nearby frequencies (Bao et al., 2013; de Villers-Sidani et al., 2007), unsupervised learning resulted in a broad upward shift that leads to overall increased frequency discriminability (this study) but wider frequency response overlap (Cruces-Solís et al., 2018). These studies, including the current one, used pure tones as a useful way to measure the

frequency-specificity of the observed changes. It is essential to study how these changes translate to a spectrally more complex and contextually more dynamic natural world.

Broad frequency shifts and increase in gain developed slowly, showing more variability during the initial days of exposure, when the control group tended towards larger responses than the predictable group. Exposure was sparse, in the sense that unlike in manually trained animals, exposure events in the form of visits to the corner could be separated by many minutes (Cruces-Solís et al., 2018). However, this exposure was embedded in a rich continuous environment in which many other stimuli were present, meaning that the mice living in the Audiobox were learning more patterns than just those specifically associated with the sound-context predictability. This initial disentangling of the different patterns present in the environment is likely to engage either different circuits or the same circuits differently. For example, cortico-collicular feedback projections are known to play an important role in the development of inferior colliculus plasticity in reinforced learning (Yan and Suga, 1998) and to facilitate learning when input-response mappings change (Bajo et al., 2010). Here, while we confirm the influence of cortico-collicular projections on inferior colliculus levels of activity, what was interesting is that this cortical feedback had a differential effect on the inferior colliculus of mice exposed to predictable context-sound association compared to control mice during the initial days of exposure. This is in contrast with what happens in animals that lived in the Audiobox for at least 6 days (Cruces-Solís et al., 2018 and this study), when cortical inactivation no longer has a differential effect on the two groups. Inactivating cortex did not have an effect on the BF at any time point in neither the predictable nor the control group, demonstrating that changes in gain are independent of shifts in BF. Thus, it seems that cortico-collicular feedback is more sensitive to the specifics of the environment during the initial days of exposure and might play a more relevant role in shaping inferior colliculus activity then.

Finally, we observe these changes in the inferior colliculus, an auditory midbrain hub previously associated with short lasting (up to 60 min) plasticity that is dependent on cortical feedback (Gao and Suga, 2000; Yan et al., 2005). The plasticity we observe here, which developed slowly over the first 6 days and lasted up to 21 days of predictable contextual sound exposure, might have different roles. Evidence for early detection of context can be deduced from noise-invariance in cortical structures. Indeed, using naturalistic background sounds, several labs have found evidence for noise-invariant responses in cortical structures (Khalighinejad et al., 2019; Mesgarani et al., 2014; Norman-Haignere and McDermott, 2018; Rabinowitz et al., 2012), supporting the idea that context is coded separately. One possibility is that predictable acoustic context is learned through cortico-subcortical loops and encoded in subcortical structures for the rapid processing that would facilitate noise-invariant responses higher up.

We conclude that the auditory midbrain codes the predictability of contextual auditory information learned over sparse and discontinuous exposure, independently of reward or sound exposure time and in the absence of a behavioural task. The auditory midbrain is thus sensitive to contingencies that must be integrated over wider temporal and spatial scales than previously thought. This is in striking contrast with the traditional view of the auditory midbrain as a structure that represents, for short time windows (tens of minutes), information about stimulus-reward associations that are inherited from cortex (Gao and Suga, 2000; Yan et al., 2005). The current findings place the auditory midbrain in a critical position to represent the contextual auditory information that often determines the task-contingencies and contributes to foreground vs background segregation.

4. Methods

4.1. Animals

Female C57BL/6JOLA^{Hsd} (Janvier, France) mice were used for experiments. All mice were 5–6 weeks old (NeuroNexus experiments) or 8 weeks old (Neuropixels experiments) at the beginning of the experiment. Animals were housed in groups and in a temperature-controlled environment (21 ± 1 °C) on a 12 h light/dark schedule (7am/7pm) with access to food and water *ad libitum*. All animal experiments were approved and performed in accordance with the Niedersächsisches Landesamt für Verbraucherschutz und Lebensmittelsicherheit, project license number 33.14-42502-04-10/0288 and 33.19-42502-04-11/0658; and in accordance with the Landesamt für Gesundheit und Soziales (LaGeSo), project license number G0140/19.

4.2. Apparatus: The Audiobox

The Audiobox is a device developed for auditory research and based on the Intellicage (TSE, Switzerland). Mice were kept in groups of 6–10 animals. At least one day before experimentation, each mouse was lightly anesthetized with Avertin *i. p.* (0.1 ml/10 g), Ketamin/Xylazine (13 mg/mL/1 mg/mL; 0.01 mL/g), or isoflurane and a sterile transponder (PeddyMark, 12 mm × 2 mm or 8 mm × 1.4 mm ISO microchips, 0.1 gr in weight, UK) was implanted subcutaneously in the upper back. Histoacryl (B. Braun, Germany) was used to close the small hole left on the skin by the transponder injection. Thus, each animal was individually identifiable through the use of the implanted transponder. The Audiobox served both as living quarters for the mice and as their testing arena.

The Audiobox was placed in a soundproof room which was temperature regulated and kept in a 12 h dark/light schedule. The apparatus consists of three parts, a food area, a "corner", and a long corridor connecting the other two parts (Fig. 1A). The food area serves as the living quarter, where the mice have access to food *ad libitum*. Water is available either in two bottles situated in the corner or in the food area. The presence of the mouse in the corner, a "visit", is detected by an antenna located at the entrance of the corner. The antenna reads the unique transponder carried by each mouse as it enters the corner. A heat sensor within the corner senses the continued presence of the mouse. An antenna located at the entrance of the corner detects the transponder in the back of the mouse as it enters. The mouse identification is then used to select the correct acoustic stimulus. Once in the corner, specific behaviors (nose-poking and licking) can be detected through other sensors. All behavioral data is logged for each mouse individually. A loudspeaker (22TAF/G, Seas Prestige) is positioned right above the corner, for the presentation of the sound stimuli. During experimentation, cages and apparatus were cleaned once a week by the experimenter.

4.3. Sound exposure in the audiobox

Sounds were generated using Matlab (Mathworks) at a sampling rate of 48 kHz and written into computer files. Intensities were calibrated for frequencies between 1 and 18 kHz with a Brüel & Kjær (4939 ¼" free field) microphone.

All mice had *ad libitum* food in the home cage, as well as free access to water through nose-poking in the corner, except the "no-water" groups (noW; see below) which had water in the food area instead. There was no conditioning in the Audiobox at any stage. All groups were first habituated to the Audiobox for 3 days without sound presentation. After the habituation phase, with the exception of the control group, mice were exposed to 16 kHz pure tone pips, roving between 63 and 73 dB, in the corner with different exposure protocols.

The "pred-W" group had the same sound exposure paradigm as in a previous study (Cruces-Solís et al., 2018). The mice were exposed to 16 kHz tone pips for the duration of each visit, regardless of nose-poke

activity and water intake.

The "pred-noW" group was treated the same as the pred-W group, i.e., it was exposed to 16 kHz tone pips for the duration of each visit. However, there was no water in the corner, and water was accessible in the home cage.

For the "pred-entry" group, 16 kHz tone pips were triggered by the entry into the corner, but only lasted for 3 s. For visits with a duration of less than 3 s, the sound stopped once the mouse left the corner.

For the "unpred-W" and "unpred-noW" groups, 16 kHz tone pips were presented in 80% of the visits with random delay of between 5 and 25 s from the visit onset. Once the sound started, it continued until the mouse left the corner. Thus, the onset of the sound could not be predicted by the animal.

The control group lived in the Audiobox for the same amount of time as the other groups, but without sound presentation. All animals can hear environmental sounds, i.e., their own vocalization and the sound of the ventilation fan. All animals were in the Audiobox in the exposure phase for at least 5 days: 6–12 days in the Neuronexus data set and 19–21 days in the Neuropixels data set.

4.4. Electrophysiology

After at least 5 days in the exposure phase, animals were removed from the apparatus one at a time and immediately anesthetized for electrophysiology. Animals were initially anesthetized with Avertin (0.15 ml/10 g; Neuronexus recordings) or Ketamine/Xylazine (13 mg/mL/1 mg/mL; 0.01 mL/g; Neuropixels recordings). Additional smaller doses of drug were delivered as needed to maintain anesthesia during surgery. The surgical level of anesthesia was verified by the pedal-withdrawal reflex. Body temperature was maintained at 36 °C with a feedback-regulated heating pad (ATC 1000, WPI, Germany).

Anesthetized mice were fixed with blunt ear bars on a stereotaxic apparatus (Kopf, Germany). Vidisic eye gel (Bausch + Lomb GmbH, Germany) was used to protect the eyes from drying out. Lidocaine (0.1 mL) was subcutaneously injected at the scalp to additionally numb the region before incision. The scalp was removed to expose the skull. Periosteum connective tissue that adheres to the skull was removed with a scalpel. The bone surface was then disinfected and cleaned with hydrogen peroxide. Bone suture junctions Bregma and Lambda were aligned. A metal head-holder was glued to the skull 1.0 mm rostral to Lambda with methyl methacrylate resin (Unifast TRAD, GC), or a combination of Histoacryl (B. Braun) and dental cement (Lang Dental and Kulzer). A craniotomy of 0.8 mm × 1.0 mm with the center 0.85 mm to the midline and 0.75 caudal to Lambda was made to expose the left inferior colliculus. After the craniotomy, the brain was protected with Saline (B. Braun, Germany). The area which is posterior to the transverse sinus and anterior to the sigmoid sinus was identified as the inferior colliculus.

The borders of the left inferior colliculus became visible after the craniotomy. The electrode was placed in the center of the inferior colliculus after measuring its visible rostrocaudal and mediolateral borders. The aim was to target the same position of the inferior colliculus across animals. Before lowering the electrode, a ground wire was connected to the neck muscle as a reference for the recordings. Neuronexus electrodes (single shank 16-channel silicon probe; 177 μm² area/site and 50 μm spacing between sites; Neuronexus, USA) were vertically lowered with a micromanipulator (Kopf, Germany) and slowly advanced (2–4 μm/s, to minimize damage to the tissue) to a depth of 750 μm from the brain surface. The final depth was confirmed by ensuring that the most dorsal channel was aligned with the dura (Fig. 1B). After one recording session, the electrode was further advanced 500 μm to a depth of 1250 μm. A final set of experiments was performed in Berlin using Neuropixels probes (Imec, Belgium, 1.0 with cap). These probes were inserted 3000 μm deep into the central nucleus of IC measured from the surface of the dura, at a speed of 62 μm per second using an electronic manipulator (Luigs & Neumann). A silver ground was attached to the probe and

positioned at the caudal end of the skin incision.

During data acquisition with Neuronexus probes, the electrophysiological signal was amplified (HS-18-MM, Neuralynx, USA), sent to an acquisition board (Digital Lynx 4SX, Neuralynx, USA), and recorded with a Cheetah 32 Channel System (Neuralynx, USA). The voltage traces were acquired at a 32 kHz sampling rate with a wide band-pass filter (0.1 ± 9,000 Hz). Neuropixels recordings were performed using a PXIe-6341 board (National instrument), sampled at 30 kHz, and recorded with open-source software SpikeGLX (<http://billkarsh.github.io/SpikeGLX/>).

4.5. Acoustic stimulation during recording

The sound was synthesized using MATLAB, produced by a USB interphase (Octa capture, Roland, USA), amplified (Portable Ultrasonic Power Amplifier, Avisoft, Germany), and played with a free-field ultrasonic speaker (Ultrasonic Dynamic Speaker Vifa, Avisoft, Germany). The speaker was 15 cm away from the right ear. The sound intensity was calibrated with a Brüel & Kjaer microphone. For measuring the frequency response area, we used sound stimuli which consisted of 30 ms pure tone pips with 5 ms rise/fall slopes repeated at a rate of 2 Hz. Thirty-two frequencies were played (3.4 kHz–49.4 kHz, 0.125-octave spacing) at different intensities (no sound and 10dB–80 dB SPL with steps of 5 or 10 dB) played in a pseudorandom order. Each frequency-intensity combination was played five times. For Neuropixels recordings, 51 frequencies were played (6.5–65 kHz), at 70 dB SPL, each repeated five times in a given sweep. Two or three sweeps were played per recording.

4.6. Analysis of electrophysiological recordings

The recorded voltage signals were high-pass filtered at 500 Hz to remove the slow local field potential signal. To improve the signal-to-noise ratio of the recording, the common average reference was calculated from all the functional channels and subtracted from each channel (Ludwig et al., 2009).

For Neuronexus data multi-unit analysis, spikes were detected as local minima below a threshold of 6 times the median absolute deviation of each channel and later summed at 1 ms bins. If the calculated value was higher than −40 μV, the threshold was set to −40 μV. Multi-unit sites were included in the analysis only if they were classified as sound-driven, defined by significant excitatory evoked responses (comparison of overall firing rate in 80 ms windows before and after stimulus onset for all frequencies and intensities; paired *t*-test, *p* < 0.05).

To analyze the sound-driven responses, signals were aligned at 200 ms before each stimulus onset. The temporal pattern of responses was compared using peri-stimulus time histograms (PSTHs). Multiunit spikes were counted over an 80 ms window before and after each stimulus onset and averaged across frequency-specific trials to calculate the spontaneous activity and evoked firing rates respectively. Firing rates used for the analysis were trial by trial baseline subtracted (evoked minus spontaneous).

Sound-driven responses at a given intensity as a function of frequency were used to generate the iso-intensity tuning curves. The combined frequency-intensity responses were used to generate the tonal receptive field (TRF).

The best frequency (BF) was selected as the frequency which elicited the highest responses when averaging across all intensities. In the rare cases where more than one frequency elicited the highest response, BF was calculated as the mean of those frequencies.

The population vector was characterized as the response to a given frequency across all depths. Discriminability between frequency pairs was calculated as Euclidean distance between their population vectors.

Neural activities at the depth from 500 μm to 750 μm were measured in two recording sessions, first when the probe tip was at 750 μm depth and then when it was at 1250 μm. When analyzing responses at a given

depth, recordings for the same depth were averaged.

Spike trains from the Neuropixels data set were obtained from 2 to 3 frequency sweeps recorded up to 2 h apart from the same animal, sorted via the open-source platforms Kilosort3 and CatGT. Frequency tuning was derived using custom scripts built in MATLAB 2020a. Evoked activity was calculated as all tuned post-trigger spikes summed over an 80 ms time window. Only activity which was significantly different from an 80 ms pre-trigger window was considered evoked. A set of selection criteria were applied to these evoked units: 1) we imposed a minimum spike amplitude $\geq 18 \mu\text{V}$, which corresponds to 3 times the baseline noise (Steinmetz et al., 2018); 2) a minimum spike count of 25, across responses to all frequencies within a frequency sweep; 3) tuning, meaning a smooth curve with no more than 4 crossings of the median spike count line.

To remove the tail units from Neuropixels datasets in order to construct the mean responses plotted in Fig. 3C, a two-cluster Gaussian mixture model (GMM; *fitgmdist*, MATLAB) initialized with the k-means algorithm was applied.

4.7. Simultaneous cortical inactivation and recordings in the inferior colliculus

After the craniotomy to expose the inferior colliculus, another 4×3 mm rectangular craniotomy medial to the squamosal suture and rostral of the lambdoid suture was performed over the left hemisphere to expose the ipsilateral auditory cortex (AC) (Fig. 4A). Vaseline was applied to the boundaries of the craniotomy to form a well. Initially, recording in the inferior colliculus was performed while 3–5 μL phosphate-buffered saline solution (B. Braun, Germany) was applied to the AC every 15 min. Then, saline was removed from the well over the AC with a sterile sponge. 3–5 μL muscimol (1 mg/mL, dissolved in phosphate-buffered saline solution) was applied over the AC every 15 min. Previously, we found that the muscimol application inactivated the AC usually in 15–20 min (Cruces-Solís et al., 2018). Therefore, we recorded the neuronal activity of the IC again 20 min after muscimol application.

4.8. Statistical analysis

Group comparisons were made using multiple-way ANOVAs after testing for normality distribution using the Shapiro-Wilk test. Samples that failed the normality test were compared using a Kruskal-Wallis test or Wilcoxon signed rank test for paired datasets. Multiple comparisons were adjusted by Bonferroni correction. For the analysis of data consisting of two groups, we used either paired t-tests for within-subject repeated measurements or unpaired t-tests otherwise. For data consisting of more than two groups or multiple parameters we used, repeated-measures ANOVA. All multiple comparisons used critical values from a t distribution, adjusted by Bonferroni correction with an alpha level set to 0.05.

Certain analyses were performed using linear mixed-effects models (*fitlme*, MATLAB), as these models are appropriate for data with repeated measurements from the same subject. The mouse was used as the random effect. We checked whether the residual of the linear mixed-effects model was normally distributed. A significance test was performed by comparing the full model with a similar model with only random effects using the standard likelihood ratio test.

Means are expressed \pm SEM. Statistical significance was considered if $p < 0.05$.

CRedit authorship contribution statement

Chi Chen: Conceptualization, and designed the study, were involved in datacollection and processing, analyzed the data and prepared the figures, wrote the manuscript. **Hugo Cruces-Solís:** were involved in datacollection and processing. **Alexandra Ertman:** were involved in

datacollection and processing, and all coauthors revised and approved it. **Livia de Hoz:** Conceptualization, and designed the study, wrote the manuscript.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

We are grateful to Markus Krohn, head of the technical workshop at the MPI-MS, and his team, to Harry Scherer, and to the Charite technical workshop for technical support.

CC is funded by Max Plank Institute Ph.D. scholarship and partially by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation), grant number 327654276 (SFB 1315, A09) to LdH; AE was funded by an Einstein Foundation Berlin grant EVF-2021-618 to LdH.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.crneur.2023.100110>.

References

- Bajo, V.M., King, A.J., 2012. Cortical modulation of auditory processing in the midbrain. *Front. Neural Circ.* 6, 114.
- Bajo, V.M., Nodal, F.R., Moore, D.R., King, A.J., 2010. The descending corticocollicular pathway mediates learning-induced auditory plasticity. *Nat. Neurosci.* 13, 253–260.
- Bakin, J.S., Weinberger, N.M., 1990. Classical conditioning induces CS-specific receptive field plasticity in the auditory cortex of the Guinea pig. *Brain Res.* 536, 271–286.
- Bao, Y., Szymaszek, A., Wang, X., Oron, A., Pöppel, E., Szelag, E., 2013. Temporal order perception of auditory stimuli is selectively modified by tonal and non-tonal language environments. *Cognition* 129, 579–585.
- Bieszczad, K.M., Weinberger, N.M., 2010. Representational gain in cortical area underlies increase of memory strength. *Proc. Natl. Acad. Sci. U.S.A.* 107, 3793–3798.
- Blake, D.T., Strata, F., Churchland, A.K., Merzenich, M.M., 2002. Neural correlates of instrumental learning in primary auditory cortex. *Proc. Natl. Acad. Sci. U.S.A.* 99, 10114–10119.
- Brandewie, E.J., Zoharik, P., 2018. Speech intelligibility in rooms: disrupting the effect of prior listening exposure. *J. Acoust. Soc. Am.* 143, 3068.
- Chen, C., de Hoz, L., 2023. The perceptual categorization of multidimensional stimuli is hierarchically organized. *iScience* 26 (6), 106941.
- Chen, C., Krueger-Burg, D., de Hoz, L., 2019. Wide sensory filters underlie performance in memory-based discrimination and generalization. *PLoS One* 14, e0214817.
- Claverol-Tinture, E., Nadasdy, Z., 2004. Intersection of microwire electrodes with proximal CA1 stratum-pyramidale neurons at insertion for multiunit recordings predicted by a 3-D computer model. *IEEE (Inst. Electr. Electron. Eng.) Trans. Biomed. Eng.* <https://doi.org/10.1109/tbme.2004.834274>.
- Cruces-Solís, H., Jing, Z., Babaev, O., Rubin, J., Gür, B., Krueger-Burg, D., Strenze, N., de Hoz, L., 2018. Auditory midbrain coding of statistical learning that results from discontinuous sensory stimulation. *PLoS Biol.* 16, e2005114.
- David, S.V., Fritz, J.B., Shamma, S.A., 2012. Task reward structure shapes rapid receptive field plasticity in auditory cortex. In: *Proceedings of the National Academy of Sciences of the United States of America*, vol. 109, pp. 2144–2149, 6.
- de Hoz, L., Nelken, I., 2014. Frequency tuning in the behaving mouse: different bandwidths for discrimination and generalization. *PLoS One* 9, e91676.
- de Hoz, L., Barniv, D., Nelken, I., 2022. Prior unsupervised experience leads to long lasting effects in sensory gating during discrimination learning. *bioRxiv*. <https://doi.org/10.1101/2022.09.30.510296>.
- de Villers-Sidani, E., Chang, E.F., Bao, S., Merzenich, M.M., 2007. Critical period window for spectral tuning defined in the primary auditory cortex (A1) in the rat. *J. Neurosci.* 27, 180–189.
- Edelman, S., 1998. Representation is representation of similarities. *Behav. Brain Sci.* 21, 449–467.
- Gao, E., Suga, N., 2000. Experience-dependent plasticity in the auditory cortex and the inferior colliculus of bats: role of the corticofugal system. *Proc. Natl. Acad. Sci. U.S.A.* 97, 8081–8086.
- Jing, Z., Pecka, M., Grothe, B., 2021. Ketamine-xylazine anesthesia depth affects auditory neuronal responses in the lateral superior olive complex of the gerbil. *J. Neurophysiol.* <https://doi.org/10.1152/jn.00217.2021>.

- Khalighinejad, B., Herrero, J.L., Mehta, A.D., Mesgarani, N., 2019. Adaptation of the human auditory cortex to changing background noise. *Nat. Commun.* 10, 2509.
- Kiang, N.Y., Moxon, E.C., 1974. Tails of tuning curves of auditory-nerve fibers. *J. Acoust. Soc. Am.* 55, 620–630.
- Letzkus, J.J., Wolff, S.B.E., Meyer, E.M.M., Tovote, P., Courtin, J., Herry, C., Lüthi, A., 2011. A disinhibitory microcircuit for associative fear learning in the auditory cortex. *Nature* 480, 331–335.
- Ludwig, K.A., Miriani, R.M., Langhals, N.B., Joseph, M.D., Anderson, D.J., Kipke, D.R., 2009. Using a common average reference to improve cortical neuron recordings from microelectrode arrays. *J. Neurophysiol.* 101, 1679–1689.
- McDermott, J.H., Schemitsch, M., Simoncelli, E.P., 2013. Summary statistics in auditory perception. *Nat. Neurosci.* 16, 493–498.
- McPherson, M.J., McDermott, J.H., 2020. Time-dependent discrimination advantages for harmonic sounds suggest efficient coding for memory. *Proc. Natl. Acad. Sci. U.S.A.* 117, 32169–32180.
- Mesgarani, N., David, S.V., Fritz, J.B., Shamma, S.A., 2014. Mechanisms of noise robust representation of speech in primary auditory cortex. *Proc. Natl. Acad. Sci. U.S.A.* 111, 6792–6797.
- Nakamoto, K.T., Sowick, C.S., Schofield, B.R., 2013. Auditory cortical axons contact commissural cells throughout the Guinea pig inferior colliculus. *Hear. Res.* 306, 131–144.
- Norman-Haignere, S.V., McDermott, J.H., 2018. Neural responses to natural and model-matched stimuli reveal distinct computations in primary and nonprimary auditory cortex. *PLoS Biol.* 16, e2005127.
- Rabinowitz, N.C., Willmore, B.D.B., Schnupp, J.W.H., King, A.J., 2012. Spectrotemporal contrast kernels for neurons in primary auditory cortex. *J. Neurosci.* 32, 11271–11284.
- Reber, A.S., 1967. Implicit learning of artificial grammars. *J. Verb. Learn. Verb. Behav.* [https://doi.org/10.1016/s0022-5371\(67\)80149-x](https://doi.org/10.1016/s0022-5371(67)80149-x).
- Steinmetz, N.A., Koch, C., Harris, K.D., Carandini, M., 2018. Challenges and opportunities for large-scale electrophysiology with Neuropixels probes. *Curr. Opin. Neurobiol.* 50, 92–100.
- Suga, N., Xiao, Z., Ma, X., Ji, W., 2002. Plasticity and corticofugal modulation for hearing in adult animals. *Neuron* 36, 9–18.
- Syka, J., Popelár, J., 1984. Inferior colliculus in the rat: neuronal responses to stimulation of the auditory cortex. *Neurosci. Lett.* 51, 235–240.
- Tao, C., Zhang, G., Zhou, C., Wang, L., Yan, S., Zhou, Y., Xiong, Y., 2017. Bidirectional shifting effects of the sound intensity on the best frequency in the rat auditory cortex. *Sci. Rep.* 7 (1) <https://doi.org/10.1038/srep44493>.
- Yan, J., Ehret, G., 2001. Corticofugal reorganization of the midbrain tonotopic map in mice. *Neuroreport* 12, 3313–3316.
- Yan, W., Suga, N., 1998. Corticofugal modulation of the midbrain frequency map in the bat auditory system. *Nat. Neurosci.* 1, 54–58.
- Yan, J., Zhang, Y., Ehret, G., 2005. Corticofugal shaping of frequency tuning curves in the central nucleus of the inferior colliculus of mice. *J. Neurophysiol.* 93, 71–83.
- Yudintsev, G., Asilador, A.R., Sons, S., Vaithiyalingam Chandra Sekaran, N., Coppinger, M., Nair, K., Prasad, M., Xiao, G., Ibrahim, B.A., Shinagawa, Y., Llano, D. A., 2021. Evidence for layer-specific connective heterogeneity in the mouse auditory corticocollicular system. *J. Neurosci.* 41, 9906–9918.
- Zhang, Y., Suga, N., 2005. Corticofugal feedback for collicular plasticity evoked by electric stimulation of the inferior colliculus. *J. Neurophysiol.* 94, 2676–2682.