



Neural Correlates of Learning Pure Tones or Natural Sounds in the Auditory Cortex

Ido Maor^{1,2}, Ravid Shwartz-Ziv², Libi Feigin², Yishai Elyada², Haim Sompolinsky^{2,3} and Adi Mizrahi^{1,2*}

¹ Department of Neurobiology, Alexander Silberman Institute of Life Sciences, The Hebrew University of Jerusalem, Jerusalem, Israel, ² The Edmond and Lily Safra Center for Brain Sciences, The Hebrew University of Jerusalem, Jerusalem, Israel, ³ The Racah Institute of Physics, The Hebrew University of Jerusalem, Jerusalem, Israel

Associative learning of pure tones is known to cause tonotopic map expansion in the auditory cortex (ACx), but the function this plasticity sub-serves is unclear. We developed an automated training platform called the “EduCage,” which was used to train mice on a go/no-go auditory discrimination task to their perceptual limits, for difficult discriminations among pure tones or natural sounds. Spiking responses of excitatory and inhibitory parvalbumin (PV⁺) L2/3 neurons in mouse ACx revealed learning-induced overrepresentation of the learned frequencies, as expected from previous literature. The coordinated plasticity of excitatory and inhibitory neurons supports a role for PV⁺ neurons in homeostatic maintenance of excitation–inhibition balance within the circuit. Using a novel computational model to study auditory tuning curves, we show that overrepresentation of the learned tones does not necessarily improve discrimination performance of the network to these tones. In a separate set of experiments, we trained mice to discriminate among natural sounds. Perceptual learning of natural sounds induced “sparsening” and decorrelation of the neural response, consequently improving discrimination of these complex sounds. This signature of plasticity in A1 highlights its role in coding natural sounds.

Keywords: auditory, mice, automated training, interneurons, perceptual learning

OPEN ACCESS

Edited by:

David J. Margolis,
Rutgers, The State University of New
Jersey, United States

Reviewed by:

Tania Rinaldi Barkat,
University of Basel, Switzerland
Sarah L. Pallas,
University of Massachusetts Amherst,
United States

*Correspondence:

Adi Mizrahi
Mizrahi.adi@mail.huji.ac.il

Received: 23 September 2019

Accepted: 17 December 2019

Published: 28 January 2020

Citation:

Maor I, Shwartz-Ziv R, Feigin L,
Elyada Y, Sompolinsky H and
Mizrahi A (2020) Neural Correlates of
Learning Pure Tones or Natural
Sounds in the Auditory Cortex.
Front. Neural Circuits 13:82.
doi: 10.3389/fncir.2019.00082

INTRODUCTION

Learning is accompanied by plastic changes in brain circuits. This plasticity is often viewed as substrate for improving computations that sub-serve learning and behavior. A well-studied example of learning-induced plasticity is following perceptual learning where cortical representations change toward the learned stimuli (Gilbert et al., 2001; Roelfsema and Holtmaat, 2018). Whether such changes improve discrimination has not been causally tested and remains debated, and the mechanisms of change are still largely unknown.

Perceptual learning is an implicit form of lifelong learning during which perceptual performance improves with practice (Gibson, 1969). Extensive psychophysical research on perceptual learning tasks led to a general agreement on some attributes of this type of learning (Hawkey et al., 2004). For example, perceptual learning has been shown to be task specific and poorly generalized to other senses or tasks. It is also largely agreed upon that gradual training is essential for

improvement (Lawrence, 1952; Ramachandran and Braddick, 1973; Ball and Sekuler, 1987; Berardi and Fiorentini, 1987; Karni and Sagi, 1991; Irvine et al., 2000; Wright and Fitzgerald, 2001; Ahissar and Hochstein, 2004; Ericsson, 2006; Kurt and Ehret, 2010). Given the specificity observed at the behavioral level, functional correlates of perceptual learning are thought to involve neural circuits as early as primary sensory regions (Gilbert et al., 2001; Schoups et al., 2001). In auditory learning paradigms, changes are already observed at the level of primary auditory cortex (Weinberger, 2004). Learning to discriminate among tones results in tonotopic map plasticity toward the trained stimulus (Recanzone et al., 1993; Rutkowski and Weinberger, 2005; Polley et al., 2006; Bieszczad and Weinberger, 2010; reviewed in Irvine, 2017). Notably, however, not all studies could replicate the learning-induced changes in the tonotopic map (Brown et al., 2004). Furthermore, artificially induced map plasticity was shown to be unnecessary for better discrimination performances *per se* (Talwar and Gerstein, 2001; Reed et al., 2011). Our understanding of the mechanisms underlying auditory cortex plasticity remains rudimentary, let alone for more natural stimuli beyond pure tones.

To gain understanding of learning-induced plasticity at single neuron resolution, animal models have proven very useful. Mice, for example, offer the advantage of a rich genetic experimental toolkit to study neurons and circuits with high efficiency and specificity (Luo et al., 2018). Historically, the weak aspect of using mice as a model was its limited behavioral repertoire to learn difficult tasks. However, in the past decade, technical difficulties to train mice to their limits have been steadily improving with increasing number of software and hardware tools to probe mouse behavior in high resolution (de Hoz and Nelken, 2014; Egnor and Branson, 2016; Murphy et al., 2016; Aoki et al., 2017; Francis and Kanold, 2017; Krakauer et al., 2017; Cruces-Solis et al., 2018; Erskine et al., 2018). Here, we developed our own experimental system for training groups of mice on an auditory perceptual task—an automatic system called the “Educage.” The Educage is a simple affordable system that allows efficient training of several mice simultaneously. Here, we used the system to train mice to discriminate among pure tones or complex sounds.

A1 is well known for its tonotopic map plasticity following simple forms of learning in other animal models (Irvine, 2017). An additional interest in primary auditory cortex is its increasing recognition as a brain region involved in coding complex sounds (Bizley and Cohen, 2013; Kuchibhotla and Bathellier, 2018). We thus asked what are the changes single neurons undergo following training to discriminate pure tones or natural stimuli. We describe distinct changes in the long-term stimulus representations by L2/3 neurons of mice following perceptual learning and assess how these contribute to information processing by local circuits. Using two-photon targeted electrophysiology, we also describe how L2/3 parvalbumin-positive neurons change with respect to their excitatory counterparts. Our work provides a behavioral, physiological, and computational foundation to questions of auditory-driven plasticity in mice, from pure tones to natural sounds.

MATERIALS AND METHODS

Animals

A total of $n = 88$, 10- to 11-week-old female mice were used in this work as follows. Forty-four mice were C57BL/6 mice and 44 mice were a crossbreed of PV-Cre mice and tdTomato reporter mice (PV \times Ai9; Hippenmeyer et al., 2005; Madisen et al., 2010). All experiments were approved by the Hebrew University’s IACUC.

Behavioral Setup

The “Educage” is a small chamber (10 \times 10 \times 10 cm), opening on one end into a standard animal home cage where mice can enter and exit freely (Figure 1A and Supplementary Figure 1a). On the other end, the chamber contains the hardware that drives the system, hardware for identifying mice and measuring behavioral performance. Specifically, at the port entrance there is a coil radio antenna (ANTC40 which connected to LID665 stationary decoder; Dorset) followed by infrared diodes used to identify mice individually and monitor their presence in the port. This port is the only access to water for the mice. Water is delivered *via* a solenoid valve (VDW; SMC) allowing precise control of the water volume provided on each visit. Water is delivered via a water spout, which is also a lickometer (based on 1 microampere current). An additional solenoid valve is positioned next to the water spout in order to deliver a mild air puff as a negative reinforcement, if necessary. For sound stimulation, we positioned a speaker (ES1; TDT), at the center of the top wall of the chamber. Sound was delivered to the speaker at 125 kHz sampling rate *via* a driver and a programmable attenuator (ED1, PA5; TDT). For high-speed data acquisition, reliable control, on-board sound delivery, and future flexibility, the system was designed via a field programmable gate array (FPGA) module and a real-time operating system (RTOS) module (CompactRIO controller; National Instruments). A custom-made A/D converter was connected to the CompactRIO controller, mediated signal from infrared diodes and lickometer and controlled the valves. A custom code was written in Labview to allow an executable user-friendly interface with numerous options for user-input and flexibility for designing custom protocols. All software and hardware design are freely available for download at <https://github.com/MizrahiTeam/Educage>.

Training Paradigm

Prior to the training, each mouse was implanted, under light and very short period of isoflurane anesthesia, with a radio frequency identification (RFID; Trovan) chip under its scruff. RFID chips allow identification of mice individually, which is then used by the system to control the precise stimulus delivery and track behavioral performance, on a per-mouse basis. Food and water were provided *ad libitum*. While access to water was only in the Educage, mice could engage the water port without restriction. Thus, mice were never deprived of food or water. At the beginning of each experiment, RFID-tagged mice were placed in a large home cage that was connected to the Educage. Before training, we removed the standard water bottle

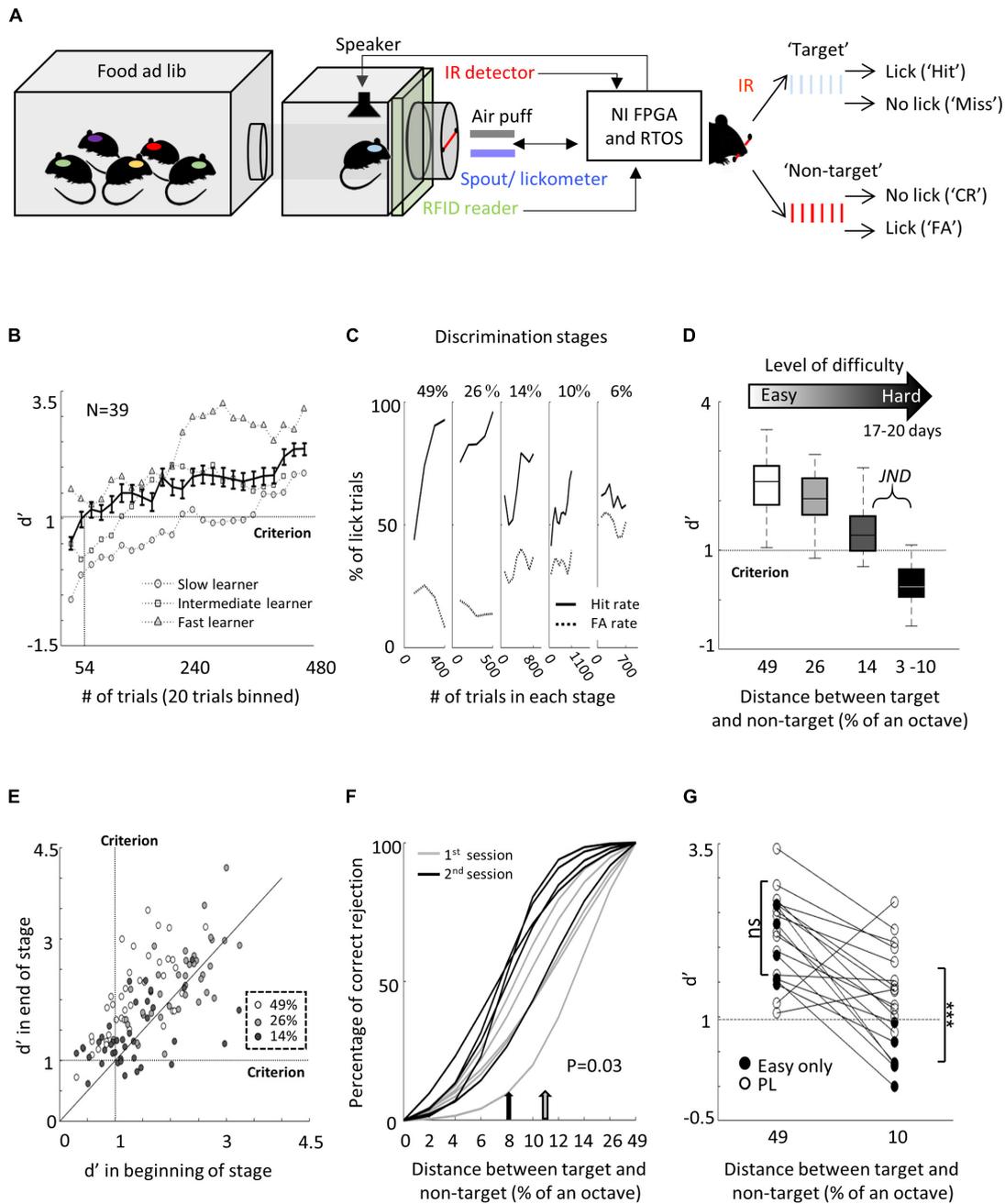


FIGURE 1 | Perceptual learning in the “Educage.” **(A)** Left: Schematic design of the “Educage” system and its components. Colored ovals on the back of each mouse represent the different radio frequency identification (RFID) chip implanted prior to the experiment, which allows identification of mice individually. For high-speed data acquisition, system control, on-board sound delivery, and future flexibility, the system was designed via a field programmable gate array (FPGA) module and a real-time operating system (RTOS) module. Right: Schematic representation of the go/no-go auditory discrimination task. CR, correct rejection; FA, false alarm. **(B)** d' values of three representative mice and the population average \pm s.e.m. for the first stage of discrimination. Learning criterion is represented as dashed line ($d' = 1$). **(C)** Lick responses to the target tone (solid line) and non-target tone (dashed line) of one representative mouse along different discrimination stages. Proportions of lick responses were calculated over 100 trials/bin. This mouse improved his discrimination (hit rate went up and FA rate went down) within a stage but his discrimination deteriorated as task became more difficult. **(D)** Population average d' values for the different discrimination stages. $N = 39$ mice (mean \pm s.e.m.). Shades denote the level of difficulty (from 49% to 4–10% octave apart). **(E)** Individual d' values at the end of each level as a function of d' in the beginning of that level. Shades denote the level of difficulty. Learning criterion represented as dashed lines ($d' = 1$). **(F)** Normalized psychometric curves of five mice calculated from the first (light curves) and the second (dark curves) 14%/octave session. Light and dark arrows indicate average decision boundaries in the first and second sessions respectively (Mann–Whitney U -test on criteria: $p = 0.03$). **(G)** d' values in easy (49%/octave) and more difficult (10% /octave) discrimination stages of individual mice from the “Easy only” group (filled circles) and from the perceptual learning group (blank circles). d' values are significantly different between groups only for the hard discrimination level (Mann–Whitney U test: $***p < 0.001$).

from the home cage. Mice were free to explore the Educage and drink the water at the behavioral port. Every time a mouse entered the behavioral port, it was identified individually by the antenna and the infrared beam and a trial was initiated. Before learning, any entry to the port immediately resulted in a drop of water, but no sound was played. Following 24 h of exploration and drinking, mice were introduced for the first time to the “target” stimulus—a series of six 10-kHz pure tones (100 ms duration, 300 ms interval; 3 ms on and off linear ramps; 62 dB SPL) played every time the mouse crossed the IR beam. To be rewarded with water, mice were now required to lick the spout during the 900-ms response window. This stage of operant conditioning lasted for 2–3 days until mice performed at >70% hit rates.

We then switched the system to the first level of discrimination when mice learned to identify a non-target stimulus (a 7.1 kHz pure tone series at 62 dB SPL) from the already known target stimulus (i.e., tones separated by 49%/octave). Thus, on each trial, one of two possible stimuli were played for 2.1 s—either a target tone or a non-target tone. Mice were free to lick the water spout anytime, but for the purpose of evaluating mouse decisions, we defined a response window and counted the licking responses only within 900 ms after the sound terminated. Target tones were played at 70% probability and non-target tones were played at 30% probability, in a pseudorandom sequence. A “lick” response to a target was rewarded with a water drop (15 μ l) and considered as a “Hit” trial. A “no lick” response to a target sound was considered as a “Miss” trial. A lick response to the non-target was considered a “False Alarm” (FA) trial, which was negatively reinforced by a mild air puff (60 PSI; 600 ms) followed by a 9-s “timeout.” A “no lick” response to the non-target was considered a correct rejection (CR) and was not rewarded.

Once mice learned the easy discrimination (reached 70% correct ratio), we switched the system to the second discrimination stage. Here, we increased task difficulty by changing the non-target tone to 8.333 kHz, thus decreasing the inter tone distance to 26%/octave. In each training stage, only one pair of stimuli was presented. Then, at the following stage, the inter tone distance was further decreased to 14%/octave and then down to 6%/octave. This last transition was often done in a gradual manner (\gg 12%/octave \gg 10%/octave \gg 8%/octave \gg 6%/octave). In some of the animals ($n = 25$), we trained mice to their just noticeable difference (JND) and then changed the task back to an easier level. In order to extract psychometric curves, for some of the mice ($n = 5$), we played “catch trials” during the first and second sessions of the 14%/octave discrimination stages. In catch trials, different tones spanning the frequency range of the whole training (7–10 kHz) were presented to the animals in low probability (6% of the total number of sounds), and were neither negatively nor positively reinforced. We have recorded from 21 mice that underwent the behavioral training. Their discriminability indexes in the 49, 26, and 14% octave discrimination levels were: 2.4 ± 0.8 , 1.9 ± 0.6 , and 1.4 ± 0.5 , respectively.

For the vocalizations task, we used playback of pups’ wriggling calls (WC) as the target stimulus. These vocalizations were

recorded with a one-quarter-inch microphone (Brüel and Kjær) from P4–P5 PV \times Ai9 pups ($n = 3$), sampled at 500 kHz, and identified offline (Digidata 1322A; Molecular Devices). As the non-target stimulus, we used manipulations of the WC. During the first stage of the operant learning, mice learned to discriminate between WC and a fully reversed version of this call. Then, the second manipulation on the non-target stimulus was a gradual change of the frequency modulation (FM) of all but the last syllable in the call while leaving the temporal structure of the full call intact. To manipulate the syllable FM, we used a dynamic linear FM ramp. This operation multiplies each sampling interval within the syllable by a dynamic speeding factor, which changed according to the relative distance from the start and end of the syllable, and generated a new waveform by interpolation from the original waveform. For example, for a 0.6 speeding factor, the beginning of each syllable was slower by a factor of 0.4 while the end of each syllable accelerated by a factor of 0.4. The range of sound modulation used here was 0.66–0.9. A value of 0.66 is away from the WC; 0.9, similar to the WC; and 1, exactly the same as the WC. The basic task design for the non-target sound was as follows: Reverse \gg 0.66 \gg 0.81 \gg 0.9.

Surgical Procedure

Mice were anesthetized with an intraperitoneal injection of ketamine and medetomidine (0.80 and 0.65 mg/kg, respectively) and a subcutaneous injection of Carprofen (0.004 mg/g). Additionally, dextrose–saline was injected to prevent dehydration. Experiments lasted up to 8 h. The depth of anesthesia was assessed by monitoring the pinch withdrawal reflex, and ketamine/medetomidine was added to maintain it. The animal’s rectal temperature was monitored continuously and maintained at $36 \pm 1^\circ\text{C}$. For imaging and recording, a custom-made metal pin was glued to the skull using dental cement and connected to a custom stage to allow precise positioning of the head relative to the speaker (facing the right ear). The muscle overlying the left auditory cortex was removed, and a craniotomy ($\sim 2 \times 2$ mm) was performed over A1 (coordinates, 2.3 mm posterior and 4.2 mm lateral to bregma) as described previously (Stiebler et al., 1997; Cohen et al., 2011; Maor et al., 2016).

Imaging and Electrophysiology

Cell-attached recordings were obtained using targeted patch-clamp recording by a previously described procedure (Margrie et al., 2003; Cohen and Mizrahi, 2015; Maor et al., 2016). For visualization, the electrode was filled with a green fluorescent dye (Alexa Flour-488; 50 μ M). Imaging of A1 was performed using an Ultima two-photon microscope from Prairie Technologies equipped with a $16 \times$ water-immersion objective lens (0.8 numerical aperture; CF175; Nikon). Two-photon excitation at wavelength of 930 nm was used in order to visualize both the electrode, filled with Alexa Flour-488, and PV⁺ somata, labeled with tdTomato (DeepSee femtosecond laser; Spectraphysics). The recording depths of cell somata were restricted to subpial depths of 180–420 μ m, documented by the multiphoton imaging. Spike waveform analysis was performed on all recorded cells, verifying that tdTomato+ cells in L2/3 had faster/narrower

spikes relative to tdTomato-negative (tdTomato-) cells (see also Cohen and Mizrahi, 2015).

Auditory Stimuli

The auditory protocol comprised 18–24 pure tones (100 ms duration, 3 ms on and off linear ramps) logarithmically spaced between 3 and 40 kHz and presented at four sound pressure levels (72–42 dB SPLs). Each stimulus/intensity combination was presented 10–12 times at a rate of 1.4 Hz. The vocalizations protocol comprised the playback of pups' wriggling calls (WC) and three additional FM calls, presented at 62 dB SPL for 16 repetitions.

Behavioral Data Analysis

To evaluate behavioral performance, we calculated, for different time bins (normally 20 trials), Hit and FA rates, which are the probability to lick in response to the target and non-target tones, respectively. In order to compensate for the individual bias, we used a measure of discriminability from signal detection theory— d' -prime (d'). d' is defined as the difference between the normal inverse cumulative distribution of the "Hit" and FA rates, $d' = z(\text{hit}) - z(\text{FA})$ (Nevin, 1969). d' for each discrimination stage was calculated based on trials from the last 33% of the indicated stage. Psychometric curves were extracted based on mouse performance in response to the catch trials. By fitting a sigmoidal function to these curves, we calculated decision boundaries as the inflection point of each curve. Detection time was calculated for each mouse individually, by determining the time in which lick patterns in the correct reject vs. the hit trials diverged (i.e., the time when significance levels crossed $p < 0.001$ in a two-sample t -test).

Data Analysis—Electrophysiology

Data analysis and statistics were performed using custom-written code in MATLAB (MathWorks). Spikes were extracted from raw voltage traces by thresholding. Spike times were then assigned to the local peaks of suprathreshold segments and rounded to the nearest millisecond. For each cell, we obtained peri-stimulus time histogram (PSTH) and determined the response window as the 100 ms following stimulus onset that evoked the maximal response integral. Only neurons that had tone-evoked response ($p < 0.05$; two sample t -test) were included in our dataset. Based on this response window, we extracted the cell's tuning curve and frequency-response area (FRA). Evoked firing rate was calculated as the average response to all frequencies that evoked a significant response. Firing rate in the training band was calculated as the response to frequencies inside the training band (7–10 kHz), averaged across all intensities. Best frequency (BF) of each cell was determined as the tone frequency that elicited the strongest responses averaged across all intensities. The selectivity of the cell is the % of all frequency-intensity combinations that evoked significant response (determined by a two-sample t -test followed by Bonferroni correction). Pairwise signal correlations (r_{sc}) were calculated as Pearson correlation between FRA's matrices of neighboring cells (<250 μm apart; Maor et al., 2016). The spontaneous firing rate of the cell was calculated based on the 100 ms preceding each stimulus presentation. Response latency

is the time point after stimulus onset at which the average spike count reached maximum.

Statistical Model Based on the Independent Basis Functions (IBF) Method

Since the measured responses before and after learning are not from the same cells, we cannot estimate the changes of individual tuning curves due to learning. Instead, we must rely on estimated learning-induced changes in the *ensemble* of single-neuron responses. Our goal, therefore, was to build a statistical model of single-neuron tuning curves before and after learning. The models were based on the statistics of each experimental group separately and used to estimate the learning-induced changes in the population of responses in each condition. Furthermore, we used this model as a generative model that allowed us to generate a large number of "model neurons" with statistically similar response properties as the measured ones.

In principle, one could use a parametric model, by fitting each observed tuning curve to a specific shape of functions (e.g., Gaussian tuning curves). However, since the tuning curves of neurons to tone frequencies do not have symmetric "Gaussian" shapes, and some are bimodal, fitting them to a parametric model has not been successful. Instead, we chose to model each single neuron response as a weighted sum of a small set of orthogonal basis functions.

$$r_i(f) = \sum_{l=1}^K a_l^i g_l(f) \quad (1)$$

Here, $r_i(f)$ is the firing rate (i.e., the trial-averaged spike count) of the i th neuron in response to the stimulus with frequency f ; K is the number of orthogonal basis functions denoted by $g_l(f)$ (dependencies in f are in log scale). In order to determine the basis functions and the coefficients, a_l^i , we performed singular value decomposition (SVD) of the matrix of the measured neuronal firing rates for the 18 values of f . Our model (1) uses a subset of the K modes with the largest singular values (the determination of K is described below). The SVD yields the coefficients $a_{i, data}^l$ for the N observed neurons and (2) smoothes the resultant SVD f -dependent vectors using a simple "moving average" technique to generate the basis functions, $g_l(f)$. (3) Importantly, to use the SVD as a generative model, for each l , we compute the histogram of the N $a_{i, data}^l$ coefficients. To generate "new neurons," we sample each coefficient independently from the corresponding histogram. In other words, we approximate the joint distribution of the coefficients by a factorized distribution. This allowed us to explore the effect of changing the number of neurons that downstream decoders use in order to perform the perceptual task.

Model (1) describes the variability of the population responses to the stimulus, in terms of tuning curves of the trial averages firing rates. Additional variability in the data is the single trial spike count. We model these as independent Poisson random variables with means given by $r_i(f)$. Since neurons are not simultaneously recorded, we do not include noise correlations in the model. We performed this procedure for the *naïve*

and *expert* measured responses separately, so that both the basis functions and the coefficient histograms are evaluated for the two conditions separately. Note that we do not make Gaussian assumptions about the coefficient histograms. In fact, the observed histograms are in general far from Gaussian.

The Choice of Number of Basis Functions

Due to a limited number of trials that we sampled for each neuron, taking a large value of K can result in overfitting the model to the noise caused by the finite number of trials. To estimate the optimal number of basis functions, we evaluated the percentage of response firing rate variability of the population (i.e., the fraction of the sum of the squared SVD eigenvalues) as a function of K . We also evaluated the parameters of model (1) based on a subset of trials and checked how well it accounts for the observed tuning curves that are calculated from the test trials. We took K that produces the smallest test error and saturated the fractional variance.

We used model (1) with the above choice of K in order to evaluate the discrimination ability of the population of A1 neurons, by creating an *ensemble* of single-neuron responses for the naïve and expert conditions. To generate the model neurons, we sampled the coefficients of the basis function independently from the corresponding histogram of the measured neurons and used these neurons for the calculations depicted below.

Fisher Information

We calculated the Fisher Information (FI) for each condition (naïve vs. expert) using our model (1). FI measure bounds the mean squared error of an (unbiased) estimator of the stimulus from the noisy single trial neuronal responses. When the neuronal population is large (and they are noise-independent) FI also determines the discriminability d' of a maximum likelihood discriminator between two nearby values of the stimulus (Seung and Sompolinsky, 1993). Under the above Poisson assumption, the FI for the i th neuron is equal to $I_i = \frac{r'_i(f)^2}{r_i(f)}$, where $r'_i(f)$ is the derivative of the firing rate with respect to the stimulus value f . The total FI is the sum of the FIs of individual neurons (Seung and Sompolinsky, 1993). This has been evaluated in both naïve and expert conditions. Note that the FI are functions of the stimulus value f , around which the discrimination task is performed.

Discrimination by Linear Readout

We applied a linear decoder to assess the ability to discriminate between nearby stimuli on the basis of the neuronal population responses. We trained a support vector machine (SVM) with a linear kernel, which finds an “optimal” linear classifier that discriminates between two nearby frequencies on the basis of single-trial vectors of spike counts generated with our generative model (1) and Poisson variability. We then evaluated the probability of classification errors to test trials, in both naïve and expert conditions. Since our training set is not linearly separable, we used SVM with slack variables (Vapnik, 1998), which incorporates a “soft” cost for classification errors. Each classification was iterated a maximum 500 times (or until

converged). In each iteration, 16 trials were used for training the classifier and four trials were used to test the decoder accuracy. Classification performance of the decoder was tested separately for discrimination between nearby frequencies which lay inside the training band or near the training band (0.4396 octave apart).

Data Analysis—Vocalization Responses

Similarity of response to different vocalizations was calculated as Pearson correlation between the PSTHs of the different stimuli. To quantify lifetime sparseness, we used the following measure: $S = (1 - [(\sum r_i/n)2/\sum (r_i^2/n)])/[1 - (1/n)]$, where r_i is the response to the i th syllable in the original vocalization (averaged across trials) and n is the number of syllables. Values of S near 0% indicate a dense code, and values near 100% indicate a sparse code (Vinje and Gallant, 2000). Population sparseness was calculated as 100—the percent of cells that evoked a significant response to each syllable in the call (Willmore and Tolhurst, 2001). Classification of vocalization identity based on population activity was determined using the SVM decoder with a linear kernel and slack variables. The decoder was tested for its accuracy to differentiate between responses to two different vocalizations. The input to the SVM consisted of the spike count of each neuron in the syllable response window. The same number of neurons (37) was used in both groups to avoid biases. We then evaluated the probability of classification errors to test trials, using leave-one-out cross-validation. Each classification was iterated 1000 times. In each iteration, 15 trials were used for training the classifier and one trial was used to test decoder accuracy. The number of syllables utilized in the decoder was increased cumulatively.

Statistical Analysis

All statistical analysis was performed with MATLAB (Mathworks). Rank-sum test was used for comparison unless otherwise noted. In cases where the same data sample was used for multiple comparisons, we used the Holm-Bonferroni correction to adjust for the increased probability of Type I error. Statistical significance was defined as $p < 0.05$.

Code Accessibility

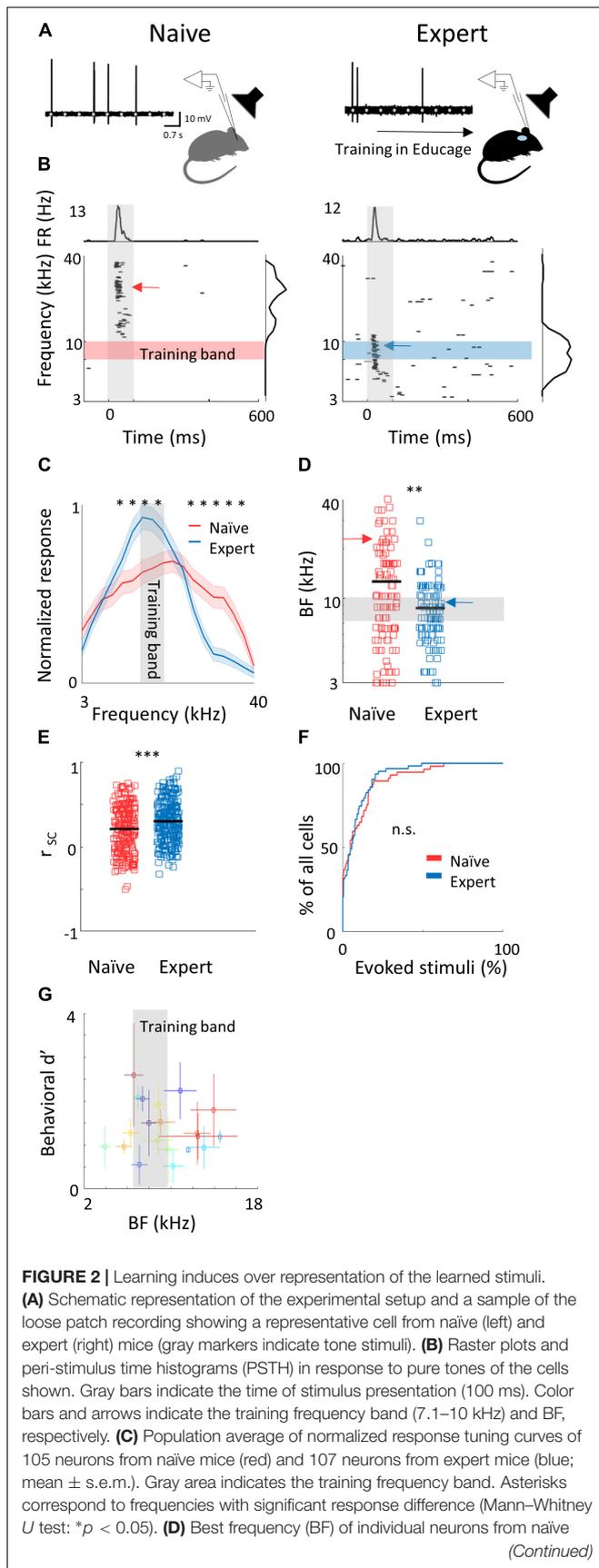
All software and hardware design for the Educage system are available for download at <https://github.com/MizrahiTeam/Educage>.

Codes used for data analysis are available from the corresponding author upon request.

RESULTS

Behavior—Discrimination of Pure Tones

To study perceptual learning in mice, we developed a behavioral platform named the “Educage” (all software and hardware design are freely available for download at <https://github.com/MizrahiTeam/Educage>). The Educage is an automated home-cage training apparatus designed to be used simultaneously with several mice (up to 8 animals). One advantage of the Educage over other procedures is that human interference is brought

**FIGURE 2 |** Continued

mice (red markers) and expert mice (blue markers; Mann–Whitney U test: $**p < 0.01$). Arrows indicate the neurons shown in **(B)**. **(E)** Pairwise signal correlations (r_{sc}) values between all neighboring neuronal pairs in naïve (red) and expert (blue) mice. Neurons in expert mice have higher r_{sc} (Mann–Whitney U test: $***p < 0.001$). **(F)** Cumulative distribution of response selectivity in naïve (red) and expert (blue) mice. Response selectivity was determined as the % of all frequency–intensity combinations that evoked a significant response. Distributions are not significantly different (Kolmogorov–Smirnov test; $p = 0.69$). **(G)** Mean BF of the neurons recorded in each mouse and its behavioral d' (during the 14% octave discrimination) shows no clear pattern of change with respect to mouse performance (mean \pm std across trials; $r^2 = -0.05$, $p = 0.7$).

to minimum and training efficiency increases. The “Educage” is a small modular chamber (10 \times 10 \times 10 cm), opening on one end into a standard animal home cage where mice can enter and exit freely (Figure 1A). On its other end, the chamber contains the hardware that drives the system, hardware for identifying mice and measuring behavioral performance (Figure 1A and Supplementary Figure 1, see section “Materials and Methods”). Mice were free to engage the behavioral port at their own will, where they consume all of their water intake. Following habituation, mice were trained on a go/no-go auditory discrimination task to lick in response to a target tone (a series of 10 kHz pure tones) and withhold licking in response to the non-target tone. A “lick” response to a target was rewarded with a water drop (15 μ l) and considered as a “Hit” trial. A “no lick” response to a target sound was considered as a “Miss” trial. A lick response to the non-target was considered a “False Alarm” (FA) trial, which was negatively reinforced by a mild air puff (60 PSI; 600 ms) followed by a 9-s “timeout.” A “no lick” response to the non-target was considered a correct rejection (CR) and was not rewarded (Figure 1A). On average, mice performed 327 ± 71 trials per day, mainly during dark hours (Supplementary Figure 1b).

The initial level of learning was to identify a non-target stimulus (a series of 7.1 kHz pure tones) from the 10 kHz target stimulus (Figure 1A right and Supplementary Movie 1). These stimuli are separated by 49% of an octave and are perceptually easily separated by mice. Discrimination performances were evaluated using d' —a measure that is invariant to individual bias (Nevein, 1969). Despite the simplicity of the task, behavioral performance varied widely between mice (Figure 1B and Supplementary Figure 1c). On average, it took mice 54 ± 38 trials to cross our criterion of learning, which was set arbitrarily at $d' = 1$ (Figure 1B; dotted line), and gradually increased to plateau at $d' = 2.35 \pm 0.64$ (Figure 1D). To extend the task to more challenging levels, we gradually increased task difficulty by changing the non-target tone closer to the target tone. The target tone remained constant at 10 kHz throughout the experiment and only the non-target stimulus changed. The lowest distance used between target and non-target was 3%/octave (9.6 kHz vs. 10 kHz). A representative example from one mouse’s performance in the Educage throughout a complete experiment is shown in Figure 1C. The JND for each mouse was determined when mice could no longer discriminate (e.g., the

JND of the mouse shown in **Figure 1C** was determined between 6 and 10%/octave). The range of JNDs was 3–14%/octave and averaged $8.6 \pm 4.7\%$ /octave. These values of JND are typical for frequency discrimination in mice (Ehret, 1975; de Hoz and Nelken, 2014). Most mice improved their performance with training (**Figures 1E,F**), showing improved perceptual abilities along the task. The duration to reach JND varied as well, ranging 3069 ± 1099 trials (14 ± 3 days). Detection times, defined as the time in which lick patterns in the correct reject trials diverged from the lick patterns of the hit trials, increased monotonically by ~ 177 ms for each step of task difficulty, demonstrating the increased perceptual load during the harder tasks (**Supplementary Figure 1d**).

To show that gradual training is necessary for perceptual learning (Ahissar and Hochstein, 2004; Ericsson, 2006; Kurt and Ehret, 2010), we trained groups of littermate mice on different protocols simultaneously. In one group of mice, we used a standard protocol and the animals were trained on the gradually increasing task difficulty described above. Simultaneously, in the second group of mice—termed “easy only”—animals were trained continuously on the easy task. Although both groups of mice trained together, only the mice that underwent gradual training were able to perform the hard task (**Figure 1G**). Taking together, these data demonstrate the efficiency of the Educage to train groups of mice to become experts in discriminating between a narrowband of frequencies in a relatively short time and with minimal human intervention.

Representation of Pure Tones in L2/3 Neurons Following Perceptual Learning

To evaluate cortical plasticity following perceptual learning, we compared how pure tones are represented in A1 of expert mice, trained to discriminate between narrowband frequencies, and age-matched naive mice who were never introduced to these sounds. We used *in vivo* loose patch recording of L2/3 neurons in anesthetized mice to record tone-evoked spiking activity in response to 3–40 kHz pure tones (**Figures 2A,B** and **Table 1**). We targeted our recording electrode to the center of A1 based on previously validated stereotactic coordinates (Maor et al., 2016). Response latencies further support that our recordings are from primary auditory cortex (range of minimal latencies: 20–44 ms; mean \pm sd: 30 ± 6 ms). Loose patch recording enables superb spatial resolution, is not biased to specific cell types, and has high signal-to-noise ratio for spike detection. However, one

caveat of this technique is a potential bias of recording sites along the tonotopic axis. In order to overcome it, we measured from neurons in a large number of animals, such that possible biases are likely averaged out. In naive mice, responses were highly heterogeneous, with best frequencies covering a significant frequency range, as expected from the heterogeneous functional microarchitecture of L2/3 neurons in A1 (**Figures 2C,D**; $n = 105$ neurons, $n = 22$ mice, red). In expert mice, best frequencies of tuning curves were biased toward the frequencies that were presented during learning (**Figures 2C,D**; $n = 107$ neurons, $n = 20$ mice, blue). These data show, as expected from previous literature, that learned frequencies in A1 become overrepresented at least as measured by the neuron’s best frequency (BF). We next showed that this overrepresentation was specific to the learned tone by training mice on 4 kHz as the target tone and recording neurons in a similar manner to the abovementioned experiment. Indeed, L2/3 neurons in A1 of mice training on 4 kHz showed BF shifts toward 4 kHz (**Supplementary Figure 2a**). These results are largely consistent with previous studies in monkeys, cats, rats, and gerbils (reviewed in Irvine, 2017), extending this phenomenon of learning-induced changes in tuning curves to the mouse, to L2/3 neurons and to local circuits.

To study neuronal changes further, we analyzed response dynamics. Temporal responses to the trained frequencies were only slightly different between naive and expert mice. Specifically, average spiking responses were slightly but significantly faster and stronger in experts (**Supplementary Figure 2b** and **Table 1**). In addition, we recorded the responses to pure tones at different intensities and constructed frequency response areas (**Supplementary Figure 2c**). The average pairwise signal correlation of neighboring neurons, calculated from these frequency response areas, was high in naive mice (0.2 ± 0.28) but even higher in experts (0.3 ± 0.24 ; **Figure 2E**). Notably, the increased signal correlation was not an artifact of differences in response properties between naive and expert group (**Supplementary Figure 2d**) but reflected true similarity in receptive fields (**Supplementary Figure 2e**). Thus, the basal level of functional heterogeneity in A1 (Bandyopadhyay et al., 2010; Rothschild et al., 2010; Maor et al., 2016) is reduced following learning. This learning-induced increase in functional homogeneity of the local circuit, emphasizes the kind of shift that local circuits undergo. Since neurons in expert mice did not have wider response areas (**Figure 2F** and **Table 1**), our data suggests that neurons shifted their response properties toward the learned tones at the expense of frequencies outside the training band.

TABLE 1 | Learning-induced physiological changes.

Group	Animals (n)	Cells (n)	BF (kHz)	Spontaneous spike rate (Hz)	Evoked firing rate in BF (Hz)	FR in T.B. (Hz)	Response latency (ms)	Selectivity (% of all stimuli)	“Lifetime sparseness” (%)
Naive	22	105	11.3 ± 8.4	0.45 ± 0.58	15.5 ± 5.3	3.4 ± 5	35.2 ± 8.7	9.7 ± 13.9	88 ± 1
Expert	20	107	8.6 ± 4	0.59 ± 0.65	13.2 ± 3.4	4.2 ± 4	32.6 ± 9.7	8.5 ± 9.9	88 ± 1
Rank-sum test			0.004	0.06	0.01	0.02	0.007	0.58	0.35

A summary table of the complete dataset of recordings from excitatory neurons in naive and expert mice after perceptual learning of pure tones. Columns show different parameters of the dataset or property tested. The third row shows the statistical *p*-value between naive and experts using a Mann–Whitney *U* test. T.B., Training Band.

Finally, the representation of sounds by the neuronal population was not clearly related to behavioral performance of individual mice (**Figure 2G**), raising the possibility that changes in BF are not necessarily related to better performance. Alternatively, the full breadth of the tuning curve is not faithfully reflected in the BF alone and a more detailed analysis of the tuning curves is necessary.

A Generative Model of A1 Population Responses to Pure Tones

To what extent does overrepresentation of a learned stimulus sub-serve better discrimination by the neural population? To answer this question, we built a statistical model of tuning curves of neurons in A1 using six basis functions (**Supplementary Figures 3a–f**) that correspond to the six largest SVD vectors of the population responses (IBF method; see section Materials and Methods). In **Figure 3A**, we show two representative examples of tuning curves and their reconstruction by our model. In contrast to Gaussian fits used previously (Vapnik, 1998; Briguglio et al., 2018), our model captures the salient features of the shape of the auditory tuning curves (asymmetry, multimodality), yet also smoothed the *raw* response vectors to reduce overfitting due to finite sampling.

In order to choose the appropriate number of basis functions, we determined the minimal number of basis functions that achieves good performance in reconstructing test single trial responses. In **Figure 3B**, we show the fraction of explained variance as a function of the number of basis functions, K . For both the naïve and expert groups, the explained variance reaches above 96% after five basis functions. **Figure 3C** shows the mean square error (MSE) on the unseen trials as a function of K in data from both the naïve and expert animals. The MSE exhibits a broad minimum for K in a range between 6 and 13. Interestingly, both naïve and experts achieve roughly the same MSE values, although for experts, the MSE values at both low and large values of K are considerably larger than that of the naïve. Taken together, we conclude that for these conditions, six basis functions are the appropriate number and we used this value for our calculations.

Based on the IBF method described above, we generated a population of tuning curves (500 “new neurons”) and estimated their total FI (see section Materials and Methods). **Figure 3D** shows the FI as a function of the stimulus f for both the naïve and expert conditions. Surprisingly, the FI of the neurons from expert animals was enhanced relative to the naïve group, but only for stimuli at both flanks of the training band. Importantly, the FI within the band of the trained frequencies remained unchanged (**Figure 3D**, within the black lines). The same result holds true for a performance of a SVM classifier. Using SVM to separate any two frequencies that are 0.2198 octave apart, discriminability (d') values derived from the classifier’s error show similar results to the FI (**Figure 3E**; Seung and Sompolinsky, 1993). The value of d' is larger in the expert groups as compared to the naïve but only outside the training band, whereas within the band, discriminability is not improved (or even slightly compromised).

The results shown in **Figure 3** do not change qualitatively if we use our SVD model for the recorded neurons, as opposed

to newly modeled neurons, nor if we compute discriminability index directly from the neuronal firing rates (**Supplementary Figure 3g**). One advantage of having a generative model for the population responses is that we can generate an unlimited number of trials and tuning curves. We took advantage of this to explore whether the results of the FI and SVM change with population size. To answer this question, we evaluated the mean discrimination performance (over test neurons) as a function of the number of sampled cells, N , which increases as expected. Consistent with the results of the SVM, the performances in the naïve and expert groups are similar with slight tendency for a higher accuracy in the naïve population at large N s (**Figure 3F**). In contrast, for frequencies near the training band, the accuracy is substantially larger in the expert than in the naïve group for virtually all N (**Figure 3G**). Thus, it seems that learning-induced changes in tuning curves do not improve discriminability of the learned stimuli.

Perceptual Learning of Natural Sounds

Natural sounds are characterized by rich spectro-temporal structures with frequency and amplitude modulations over time (Mizrahi et al., 2014). Discrimination of such complex stimuli could be different from that of pure tones. Thus, we next designed a task similar to that with the pure tones but using mouse vocalizations as the training stimuli. We used playback of pups’ wriggling calls (WC) as the target stimulus (**Figure 4A**, top). As the non-target stimuli, we used frequency modulations of the WC, a manipulation that allowed us to morph one stimulus to another by a continuous metric (**Figure 4A**). The range of sound modulation used here was indexed as a “speeding factor” (see section “Materials and Methods” for details). In short, a modulation factor of 0.66 affected the original WC more than a modulation factor of 0.9 did, and is therefore easier to discriminate (**Figure 4B**). To reach perceptual limits, we trained mice gradually, starting with an easy version of the task (WC vs. a temporally reversed version of the WC) and then gradually to modulated calls starting at 0.66 modulation. Once mice reached > 80% hit rates, we changed the non-target stimulus to more difficult stimuli until mice could no longer discriminate (**Figure 4C**). Mice ($n = 9$) learned the easy task, i.e., discriminating WC from a 0.66 modulated call, with average d' values of 2.5 ± 0.4 (**Figure 4D**). On average, mice could only barely discriminate between a WC and its 0.9 modulation (d' at 0.9 was 1 ± 0.8 ; **Figure 4D**). While these discrimination values were comparable to the performance of pure tones, detection times were substantially slower (**Supplementary Figure 4a**). For similar d' values, discriminating between the vocalizations took 300–1000 ms longer as compared to the pure tone tasks (**Figure 4E**). In addition, learning curves were slower for the vocalization task as compared to the pure tones task. The average number of trials to reach $d' = 1$ for vocalizations was 195 trials, more than three times longer as compared with pure tones (compare **Supplementary Figure 4b** and **Figure 1B**, respectively). These differences may arise from the difference in the delay, inter-syllable interval, and temporal modulation of each stimulus type, which we did not further explore. Taken together, these behavioral results demonstrate a

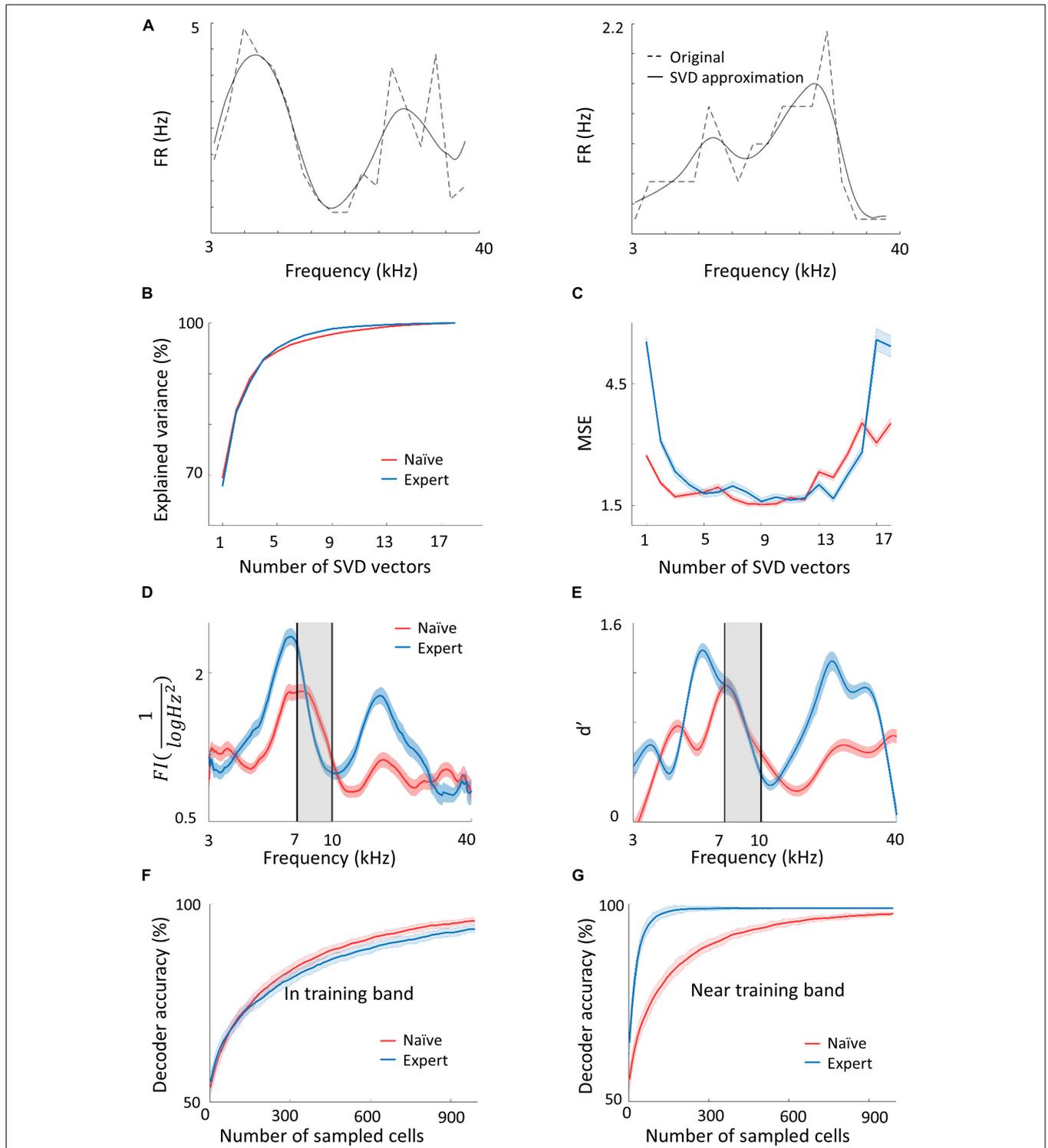
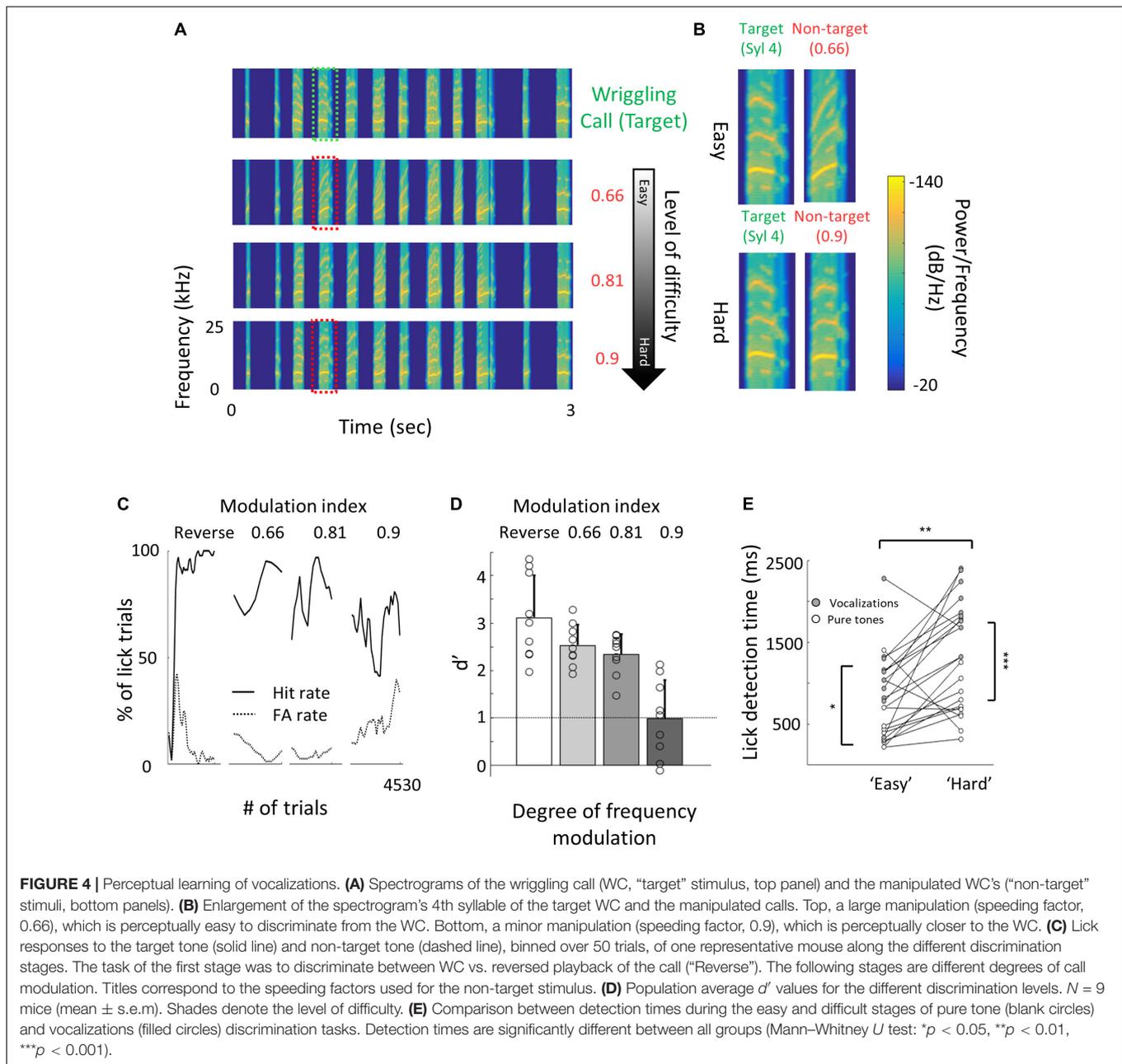


FIGURE 3 | Plasticity in A1 does not improve the discrimination of the learned tones. **(A)** Two representative examples of tuning curves of neurons recorded in A1. Average spike rates and the SVD approximation of the particular curves are shown in dashed and solid lines respectively. Note that although the SVD approximation is smooth, it captures the irregular dynamics (i.e., non-Gaussian) of the tuning curves. **(B,C)** The explained variance and error as a measure of the number of SVD vectors used in the model. **(D)** Fisher Information calculated from the tuning curves of both populations along the frequency dimension. Note the increased FI for the expert neurons in the flanks of the training band but not within it (gray band). **(E)** Discriminability (d') of SVM decoder along the range of frequencies. Pairwise comparison along the continuum is performed for frequencies 0.2198 octave apart. In accordance with d' , the decoder does not perform better in the training band (gray shade). **(F,G)** Classification performance of the decoder as a function of the number of neurons in the model. In the training band **(F)**, the performance is similar for both naïve and expert mice. Outside the training band (0.4396 octave apart; **G**) performance improved rapidly in the expert mice.

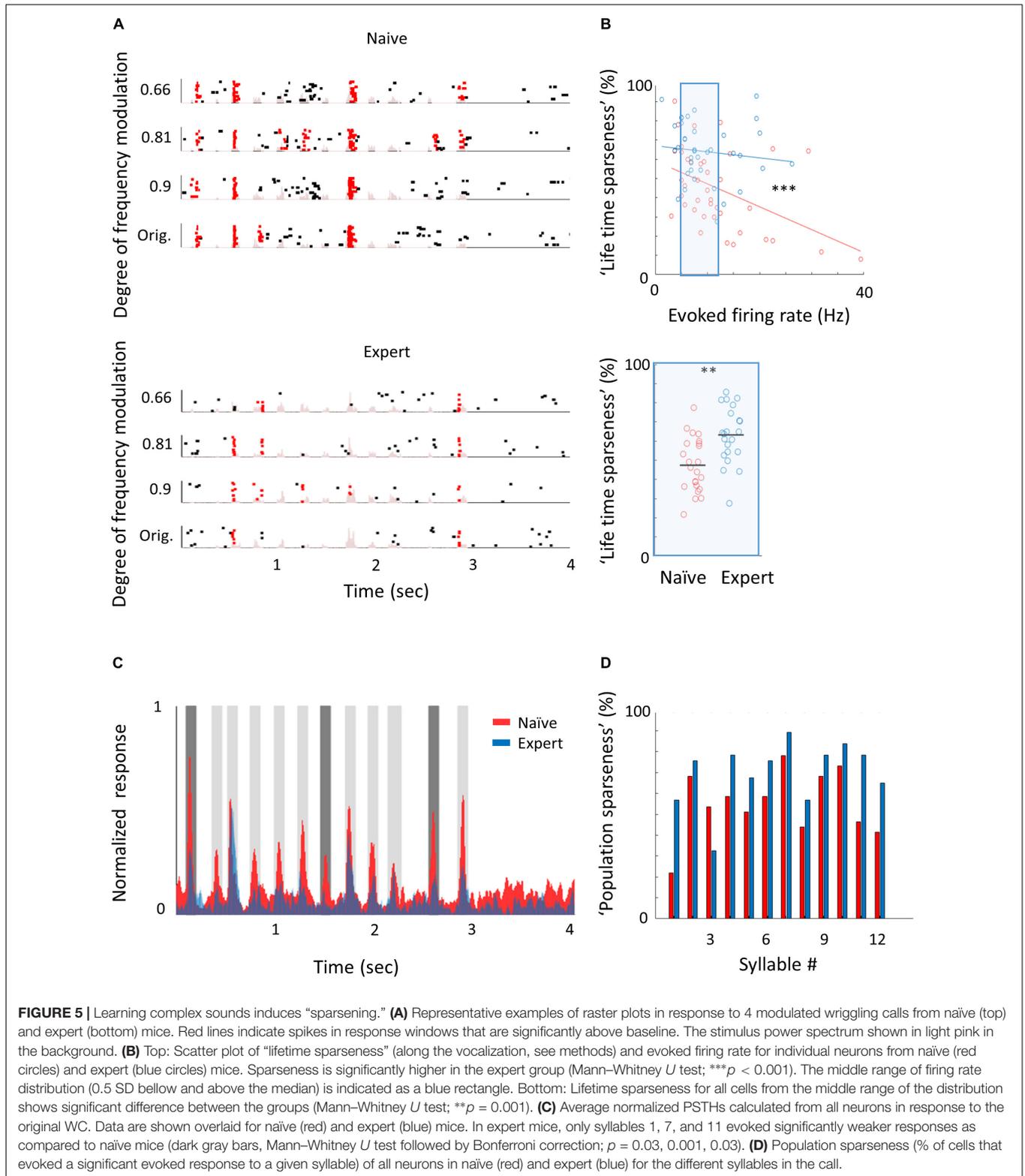


gradual increase in perceptual difficulty using a manipulation of a natural sound.

Sparser Response in L2/3 Neurons Following Perceptual Learning of Natural Sounds

To study the neural correlates in A1 that follow natural sound discrimination, we recorded L2/3 neurons in response to the learned stimuli (**Figure 5A**), expecting increased representation of these particular stimuli (e.g., that more neurons will respond to the calls or that firing rates will increase). Surprisingly, we did not find an increase in the representation of the

learned stimuli. The fraction of cells responding to the trained vocalization remained constant (**Supplementary Figure 5a**) as well as the evoked firing rate for the preferred vocalization or the preferred syllable within a vocalization (**Table 2**). Instead, representation in expert mice became sparser. Here, we measured the “lifetime sparseness” (Vinje and Gallant, 2000; Willmore and Tolhurst, 2001) of each neuron to determine how selective its responses is to a given syllable. Sparse representation can be a result of having a smaller fraction of neurons responding to a given stimulus and/or a decrease responsiveness in the call. We found that the population of neurons in expert mice were sparser (**Figure 5B**; Naïve: $45\% \pm 20\%$; Expert: $64\% \pm 15\%$; $p < 0.001$), even



regardless of their evoked firing rate (**Figure 5B**; Naïve: $47\% \pm 14\%$; Expert: $63\% \pm 15\%$; $p = 0.001$). Increased sparseness was not apparent following pure tone learning (**Table 1**). Sparseness was also evident from the average population

response to the vocalization (**Figure 5C**). Nearly all syllables had weaker responses, three of which were statistically weaker (**Figure 5C**, dark gray bars). Moreover, the population sparseness (Willmore and Tolhurst, 2001), derived from the fraction of

TABLE 2 | Vocalization learning-induced physiological changes.

Group	Animals (n)	Cells (n)	Spontaneous firing rate (Hz)	Evoked firing rate (Hz)	Response latency (ms)	Response selectivity (% of evoked stimuli)	Fano-Factor
Naïve	8	41	1.8 ± 2.9	7.7 ± 5.9	36 ± 11	25.5 ± 21.6	1 ± 0.45
Expert	6	37	0.78 ± 0.74	6.7 ± 3.6	43 ± 13	15.4 ± 11	1.1 ± 0.35
Rank-sum test			$\rho = 0.1$	$\rho = 0.6$	$\rho = 0.02$	$\rho = 0.05$	$\rho = 0.12$

A summary table of the complete dataset of recordings from excitatory neurons in naïve and expert mice after perceptual learning of vocalizations. Columns show different parameters of the dataset or property tested. The third row shows the statistical p -value between naïve and experts using a Mann–Whitney U test.

active neurons at any time, was higher in expert animals for almost all syllables in the call (**Figure 5D**; Naïve: $55 \pm 16\%$; Expert: $70 \pm 15\%$; $p = 0.026$). Notably, the smaller fraction of responses was not just an apparent sparseness due to increase in trial-to-trial variability, as reliability of responses by neurons in the expert group remained similar to reliability of responses in the naïve group (**Table 2**). Thus, “sparsening” of A1 responses is a main feature of plasticity following learning to discriminate complex sounds.

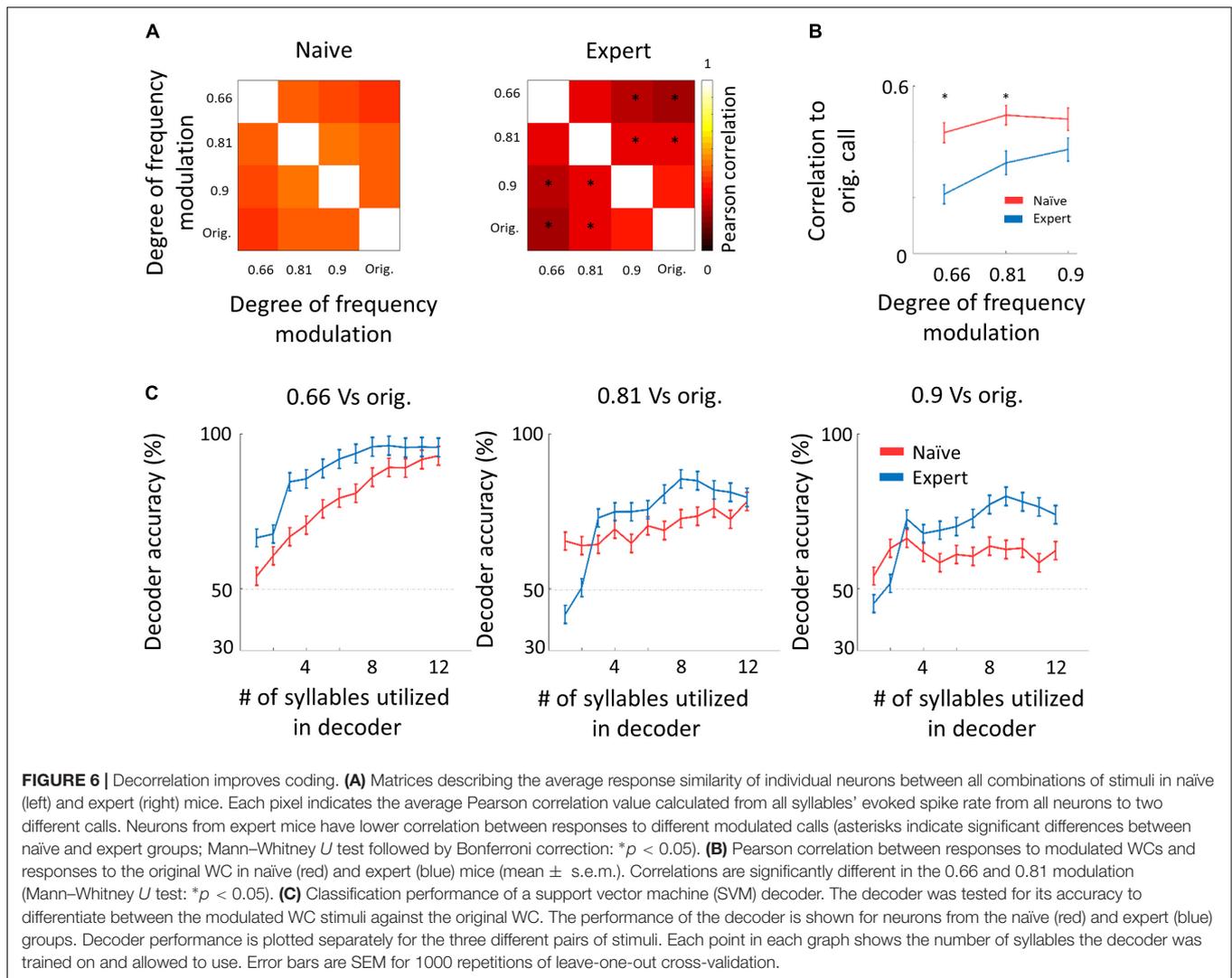
We next asked whether the “sparsening” described above bears more information to the learned stimulus. We analyzed population responses by including all neurons from all mice as if they are a single population (naïve: $n = 41$ neurons from 8 mice; expert: $n = 37$ neurons from 6 mice). We calculated the Pearson correlation of the population response to all responses in a pairwise manner (**Figure 6A** and **Supplementary Figure 5b**). As compared to naïve mice, the absolute levels of correlations in expert mice were significantly lower for nearly all pairs of comparisons (**Figure 6A**, asterisks). As expected, weaker modulations of the call and, hence, high similarity among stimuli, were expressed as higher correlations in the neuronal responses (**Figure 6B**). The pairs of stimuli that mice successfully discriminated in the behavior (0.66 vs. the original WC and 0.81 vs. the original WC) had significantly lower correlation in the expert mice (**Figure 6B**, rank sum test, $p < 0.05$). Responses to the more similar stimuli that were near perceptual thresholds (i.e., 0.9 vs. the original WC) were lower in expert mice, but not significantly (**Figure 6B**, rank sum test, $p > 0.05$). This reduced correlation suggested that plasticity in A1 supports better discrimination among the learned natural stimuli. Indeed, a SVM decoder performed consistently better in expert mice, discriminating more accurately the original WC from the manipulated ones (**Figure 6C**). As expected, the decoder performance monotonically increased when utilizing the responses to more syllables in the call. However, in the expert mice, performances reached a plateau already halfway through the call, suggesting that neuronal responses to the late part of the call carried no additional information useful for discrimination. Similarly, the correlation of the population responses along the call shows that responses were separated already following the first syllable, but that the lowest level of the correlation was in the 5th to 7th syllable range, which then rapidly recovered by the end of the call (**Supplementary Figure 5c**). These findings are also consistent with the behavioral performance of the mice as decisions are made within the first 1.5 s of the trial (corresponding to the first seven syllables of the call).

Specifically, the head of the mouse is often retracted by the time the late syllables are played (**Figure 4E**). Taken together, “sparser” responses improve neural discrimination of learned natural sounds.

Learning-Induced Plasticity of Parvalbumin Neurons

The mechanisms responsible for the learning-induced changes are currently unknown. We used mouse genetics and two-photon targeted patch to ask whether local inhibitory neurons could contribute to the observed plasticity we describe above. To this end, we focused only on parvalbumin inhibitory (PV⁺) interneurons as they are the most abundant inhibitory cell type with the strongest direct silencing effect on pyramidal cells (Avermann et al., 2012; Hu et al., 2014). In addition, recent evidence points to their role in a variety of learning-related plasticity processes (Letzkus et al., 2011; Wolff et al., 2014; Kaplan et al., 2016; Lagler et al., 2016; Goel et al., 2017; Lee et al., 2017). Here, we probe the role of inhibition in perceptual learning by measuring learning-induced plasticity in the response properties of the PV⁺ neurons. We trained PV⁺-Cre x Ai9 mice (i.e., mice with PV⁺ neurons expressing tdTomato) in the Educage and then patched single neurons under visual guidance (**Figure 7A**; Cohen and Mizrahi, 2015; Maor et al., 2016). In order to increase the sample of PV⁺ neurons, we used targeted patch and often patched both PV⁺ and PV⁻ neurons in the same mice. PV⁻ neurons were used as proxy for excitatory neurons (these neurons were also included in the analysis shown in **Figures 2–6**). All the TdTomato⁺ neurons that we patched were also verified as having a fast spike shape (**Figure 7A**), a well-established electrophysiological signature of PV⁺ cells, while PV⁻ neurons verified as having regular spike shape. PV⁺ neurons had response properties different from PV⁻ neurons in accordance with our previous work (Maor et al., 2016). For example, PV⁺ neuron responses were stronger and faster to both pure tones and natural sounds (**Figures 7B,C** and **Table 3**; see also Maor et al., 2016).

Following pure tone learning, PV⁺ neurons also changed their response profile. On average, the BF of PV⁺ neurons shifted toward the learned frequencies, similar to what we described for PV⁻ neurons (**Figures 8A,B**). This result is consistent with recent evidence from the visual cortex showing increased selectivity of PV⁺ to trained stimuli following learning (Khan et al., 2018). The shift in tuning curves of PV⁺ neurons was also accompanied by a significant widening of their receptive fields, unlike the PV⁻ population (**Supplementary Figure 6a**). When we compared the



BFs of PV^- and PV^+ neurons within the same brain (within 250 microns of each other), we found that excitatory and inhibitory neurons became more functionally homogeneous as compared to naïve mice (**Figure 8C**). As both neuronal groups show similar trends in the shift of their preferred frequencies, we rule out a simple scenario whereas parvalbumin neurons increase their responses in the sidebands of the learning frequency. In other words, plasticity does not seem to be induced by lateral inhibition *via* parvalbumin neurons, but rather maintains a strict balance between excitation and inhibition, regardless of whether they are naïves or experts (Wehr and Zador, 2003; Zhou et al., 2014). Note that the peak of the PV^+ population response and their BF distribution was on the outskirts of the training band, rather than within it (compare **Figures 8A,B** with **Figures 2C,D**), and concomitantly, the slope of the population responses at the trained frequency band increased due to learning (**Figure 8A**). To assess the computational effect of the plasticity in the inhibitory neurons' responses, we have applied on their responses the same d' and FI calculated for the PV^- neurons

(**Figures 8D,E**). Overall, the discrimination performance of the two cell populations (when equalized in size) is similar. However, the PV^+ population shows a significant learning related increase in tone discrimination performance by these cells within the training band (**Figures 8D,E**) in contrast to the results for PV^- neurons (**Figures 3D,E**). This result is consistent with the abovementioned increase in their response slopes in the trained frequency.

Following natural sound learning, we found no significant changes in basic response properties of the PV^+ neurons (**Table 3**), or in the degree of sparseness of their representation of the learned vocalizations (**Supplementary Figure 6b**). The relationship between the responses of PV^+ and their PV^- neighbors remained constant as reflected in the similar slopes of the functions describing PV^- firing versus PV^+ firing (**Supplementary Figure 6c**). This result suggests that the excitation–inhibition balance, as reflected in the responses of PV^- versus PV^+ , remains. In PV^+ neurons, the temporal correlation along the call as well as the decoding performance

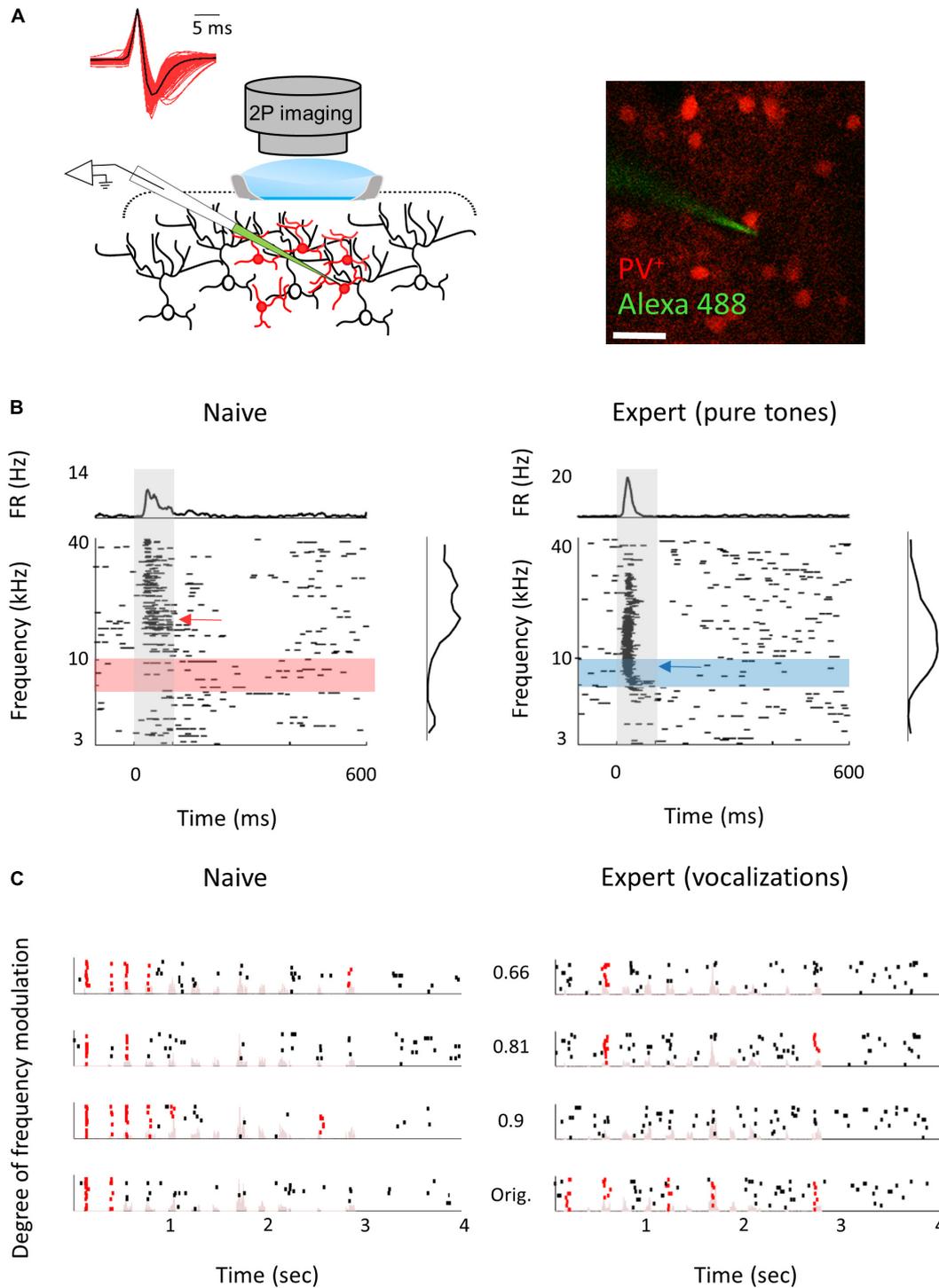


FIGURE 7 | Response properties of PV⁺ neurons. **(A)** Left: Schematic representation of the experimental setup for two-photon targeted patch and average spike waveform of 203 PV⁺ neurons. Right: Representative two-photon micrograph (projection image of 120 microns) of tdTomato+ cells (red) and the recording electrode (Alexa Fluor-488, green). **(B)** Raster plots and peri-stimulus time histograms (PSTH) in response to pure tones of a representative PV⁺ neuron from naïve (left) and expert (right) mice. Gray bars indicate the time of stimulus presentation (100 ms). Color bars and arrows indicate the training frequency band (7.1–10 kHz) and BF, respectively. **(C)** Representative examples of raster plots from PV⁺ neurons in response to four modulated wriggling calls from naïve (left) and expert (right) mice. Red lines indicate spikes in response windows that are significantly above baseline. Stimulus power spectrum shown in light pink in the background.

TABLE 3 | Physiological changes in PV⁺ neurons.

Group	Animals (n)	Cells (n)	BF (kHz)	Spontaneous firing rate (Hz)	Evoked firing rate in BF (Hz)	FR in training band (Hz)	Response latency (ms)	Response selectivity (% of evoked stimuli)
Naïve FS PT	21	120	16.2 ± 10.4	2.2 ± 2.1	21 ± 9	9.2 ± 9	31.1 ± 6.7	17 ± 17
Expert FS PT	18	83	11.8 ± 8.1	1.7 ± 1.8	22 ± 10	11.9 ± 10	33.2 ± 9.3	25 ± 18
Rank-sum test			0.006	0.05	0.7	0.02	0.2	<0.001
Naïve FS vocalizations	6	17		4 ± 3.7	14.3 ± 10.8		31.5 ± 8.6	27 ± 14.9
Expert FS vocalizations	5	18		3.4 ± 3.9	14.9 ± 12.9		32 ± 8	25.6 ± 15.6
Rank-sum test				0.4	0.97		1	0.72

A summary table of the complete dataset of recordings from PV neurons in naïve and expert mice after perceptual learning of pure tones (rows 1 and 2) and vocalizations (rows 4 and 5). Columns show different parameters of the dataset or property tested. The white color row shows the statistical *p* value between naïve and experts using a Mann–Whitney *U* test for each group separately.

from these neurons showed changes that are qualitatively similar to their PV[−] counterparts, but the data across the population were noisier (Figures 8F,G and Supplementary Figure 6d) perhaps due to the smaller sample of the PV⁺ dataset.

DISCUSSION

Plasticity in Frequency Tuning Following Perceptual Learning

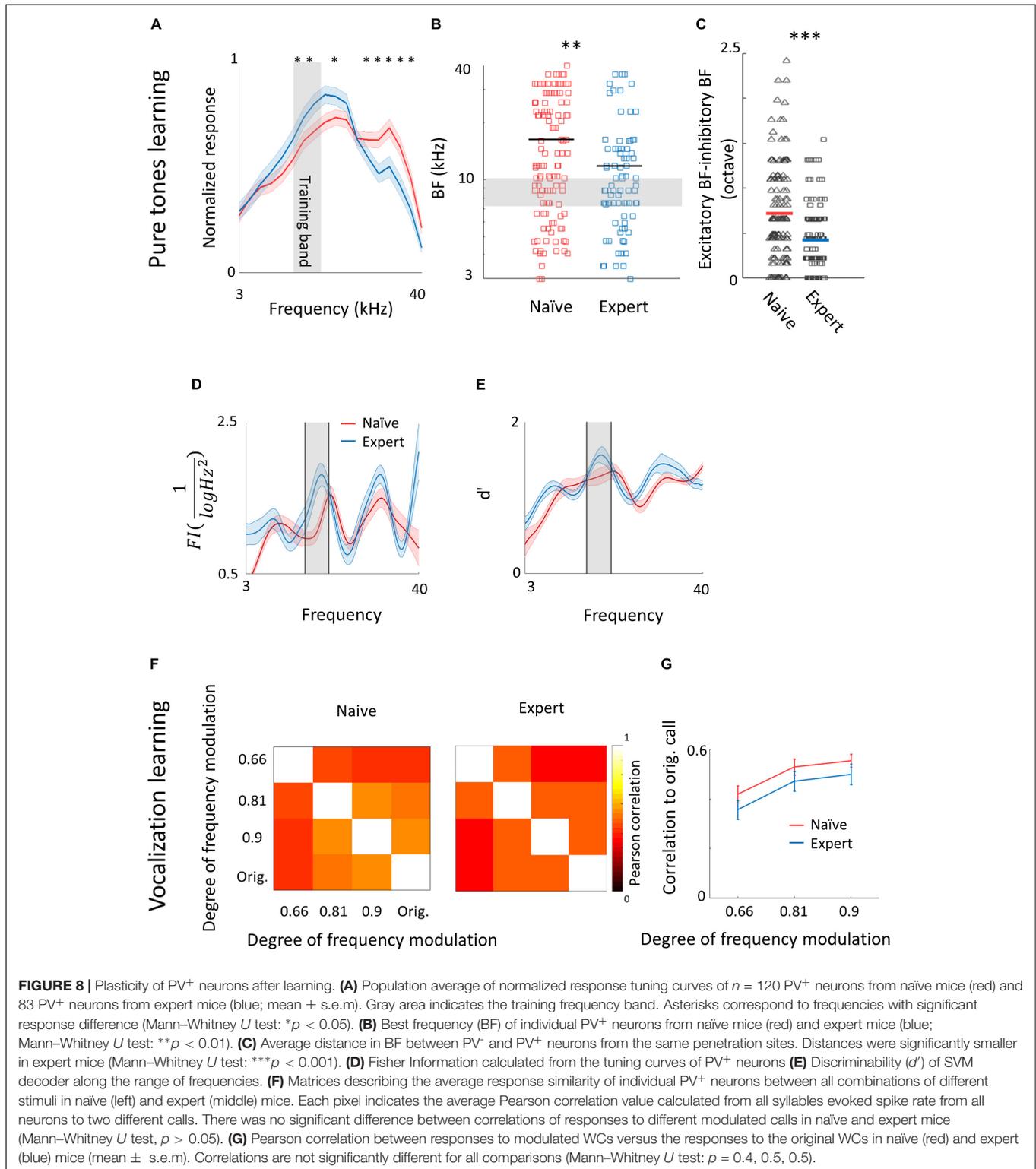
Shifts in the average stimulus representation toward the learned stimuli are not a new phenomenon. Similar findings were observed in numerous studies, multiple brain areas, animal models, and sensory systems, including in auditory cortex (Karni and Sagi, 1991; Buonomano and Merzenich, 1998; Weinberger, 2004). In fact, the model of learning-induced plasticity in A1, also known as tonotopic map expansion, is an exemplar in neuroscience (Bakin and Weinberger, 1990; Recanzone et al., 1993; Rutkowski and Weinberger, 2005; but see Crist et al., 2001; Ghose et al., 2002; Kato et al., 2015). Although we did not measure tonotopic maps, our results support the observations of others that tuning curves are plastic in primary sensory cortex. Specifically, we show here an average shift in the tuning of L2/3 neurons in A1 in mice.

Since we sampled only a small number of neurons, our observations cannot be inferred as direct evidence for tonotopic map expansion. Rather, our data emphasize that plastic shifts occur in local circuits (Figure 2E). Given that neurons in A1 are functionally heterogeneous within local circuits (Maor et al., 2016), any area in A1 that represents a range of frequencies prior to learning could become more frequency-tuned once learned. Such a mechanism allows a wide range of modifications within local circuits to enable increased representation of the learned stimuli without necessarily perturbing gross tonotopic order. One advantage of local circuit heterogeneity is that it allows circuits to maintain a dynamic balance between plasticity and stability (Mermillod et al., 2013).

It is often assumed that the learning-induced changes in tuning properties improve the accuracy of coding of the trained stimuli. In particular, perceptual learning theory predicts that sharpening the slope of the tuning curves improves the discriminability of the relevant stimuli

(Seung and Sompolinsky, 1993). However, the observed increased representation of the BFs toward the training band in expert mice may not increase over all tuning slopes and may even decrease them, especially since the slopes tend to be small at the BFs. A closer look at the tuning curves in A1 shows that they are often irregular with multiple slopes and peaks (i.e., not having simple unimodal Gaussian shapes). Furthermore, the learning-induced changes in the ensemble of tuning curves are not limited to shifting the BF; hence, a more quantitative approach was required to assess the consequences of the observed learning-induced plasticity on discrimination accuracy. Our new, SVD-based, generative model (Figure 3) allowed us to assess the combined effects of changes in BFs as well as other changes in the shapes of the tuning curves. Surprisingly, both FI analysis and estimated classification errors of an optimal linear classifier show that learning-induced changes in tuning curves do not improve tone discriminability at trained values. This conclusion is consistent with previous work on the effect of exposure to tones during development that has been argued to decrease tone discriminability for similar reasons (Han et al., 2007). However, in that work, the functional effect of tuning curve changes was consistent with an observed impaired behavioral performance, suggesting that plasticity in A1 sub-serves discrimination behavior. In contrast, the stable (or even reduced) accuracy in the coding of the trained frequency we observed occurs despite the improved behavioral performance after training. To test whether behavioral performances of individual mice are correlated with coding accuracy, we plotted neuronal *d'* with behavioral *d'* on a mouse by mouse basis. Although we recorded high *d'* values in the few mice that performed particularly well, we did not find a significant correlation between these two measures across mice (Supplementary Figure 7a). Mice with similar neuronal *d'* values often differed as much as twofold in their behavioral performance (Supplementary Figures 7a,b).

One possibility for the observed changes in the tuning of pure tones in A1 are the result of unsupervised Hebbian learning induced by overexposure to the trained tones during the training period, similar to the reported results in early overexposure (Han et al., 2007). Unsupervised learning signals are not driven by task-related reward and punishment *per se* and may increase representation rather than discriminability. Increased representation of trained stimuli may lead to improved



discriminability to untrained tones, as observed experimentally. Both reward and punishment can drive associative learning and do so *via* non-overlapping neuronal pathways (Cohen and Blum, 2002; Seymour et al., 2007). Here, we used both reward and punishment to drive mice to their perceptual

limits. Fear conditioning has been shown to induce receptive field plasticity along the auditory pathway (Diamond and Weinberger, 1986; Kim and Cho, 2017), but so did positive reinforcement (Rutkowski and Weinberger, 2005). More recently, David et al. (2012) showed that positive and negative

reinforcements drive cortical plasticity in opposite directions—negative reinforcement boosts but positive reinforcement reduces responses of the target frequency (David et al., 2012). Notably, if both mechanisms act simultaneously, as it is in our case, it would be difficult to tease out their individual contribution to discrimination. In our data, the FI of expert mice peaked near 7 kHz (**Figure 3D**), leading us to speculate that aversion effects may have predominated. Indeed, the effects of the negative reinforcement on mouse behavior were strong. Specifically, following non-target stimuli (both FA and CR trials), mice returned back to the home cage and initiated their next trial only > 100 s later (**Supplementary Figure 7c**). Thus, we speculate that both unsupervised learning and the negative reinforcement may have affected cortical responses to pure tones.

Our protocol to push mice to their limits thus involves several processes acting simultaneously. The relative contribution of each—unsupervised learning or supervised learning (positive or negative)—utilized to drive the behavior and its resultant neural correlate remains to be elucidated.

Plasticity Following Natural Sounds Learning

Unlike pure tones, learning-induced changes in A1 improved the discriminability of learned natural sounds. Although auditory cortex is organized tonotopically, it may not be critical for processing simple sounds, as these stimuli are accurately represented in earlier stages in the auditory hierarchy (Nelken, 2004; Mizrahi et al., 2014). While the causal relationships between A1 and pure tones discrimination are still debatable (Ohl et al., 1999; LeDoux, 2000; Weible et al., 2014; O’Sullivan et al., 2019), the role of A1 in processing and decoding complex sounds is becoming increasingly more evident (Letzkus et al., 2011; Ceballo et al., 2019).

Neurons in A1 respond to sounds in a non-linear fashion (Bathellier et al., 2012; Harper et al., 2016; Angeloni and Geffen, 2018). These can be selective to harmonic content that are prevalent in vocalizations (Feng and Wang, 2017). Other neurons show strong correlations to global stimulus statistics (Theunissen and Elie, 2014). Furthermore, A1 neurons are sensitive to the fine-grained spectrotemporal environments of sounds, expressed as strong gain modulation to local sound statistics (Williamson et al., 2016), as well as to sound contrast and noise (Rabinowitz et al., 2011, 2013). These features (harmonics, globally and locally rich statistics, and noise), as well as other unique attributes such as frequency range, amplitude modulations, frequency modulations, inter syllable interval and duration variability, are well represented in the WCs. Importantly, WCs have been shown to drive strong responses in mouse A1 (Maor et al., 2016; Tasaka et al., 2018). Which of these particular sensitivities changes after perceptual learning and what is the contribution of the different attributes to the plasticity observed following vocalization learning is not yet known. However, one expression of this plasticity can be the increased sparse sound representation we found here (**Figure 5**). Sparseness can take different forms (Barth and Poulet, 2012). Here, sparseness was expressed as reduced number of neurons in the network that respond to any of

the 12 syllables played (**Figures 5B–D**). Such increase could arise from disparate mechanisms, and changes in the structure of local inhibition was one suspect that we tested (Froemke, 2015).

Inhibitory Plasticity Follows Excitatory Plasticity

Cortical inhibitory neurons are central players in many forms of learning (Kullmann et al., 2012; Hennequin et al., 2017; Sprekeler, 2017). Inhibitory interneurons have been implicated as important for experience-dependent plasticity in the developing auditory system (Hensch, 2005), during fear learning in adulthood (Letzkus et al., 2011; Courtin et al., 2014), and following injury (Resnik and Polley, 2017). Surprisingly, however, and despite the numerous studies on parvalbumin neurons, we could not find any references in the literature of recordings from parvalbumin neurons after auditory perceptual learning. Two simple (non-mutually exclusive) hypotheses are naively expected. One is that plasticity in inhibitory neurons is a negative mirror of the plasticity in excitatory neurons. This would predict that inhibitory neurons would increase their responses to the stimuli for which responses of excitatory neurons are downregulated, as found in the plasticity of somatostatin expressing (SOM) neurons following passive sound exposure (Kato et al., 2015), or in multisensory plasticity in mothers (Cohen and Mizrahi, 2015). The second is that inhibitory neurons enhance their response to “lateral” stimuli, thus enhancing selectivity to the trained stimulus, as suggested by the pattern of maternal-related plasticity to pup calls (Galindo-Leon et al., 2009). Our study provides a first test of these hypotheses in the context of perceptual learning. We found no evidence for these scenarios. Instead, a common motif in the local circuit was that parvalbumin neurons changed in a similar manner to their excitatory counterparts. These results are in line with the observation that PV neurons in V1 following visual discrimination task become as selective as their pyramidal neighbors (Khan et al., 2018).

The cortex hosts several types of inhibitory cells (Hattori et al., 2017; Zeng and Sanes, 2017), presumably serving distinct roles. While PV neurons are considered a rather homogeneous pool of neurons based on molecular signature, their role in coding sounds is not (Seybold et al., 2015; Phillips and Hasenstaub, 2016). In the visual cortex, PV⁺ cells’ activity (and presumably its plasticity) was correlated with stimulus-specific response potentiation but not in ocular dominance plasticity (Kaplan et al., 2016), again suggesting that PV⁺ neurons are not necessarily involved in all forms of experience-dependent plasticity. Inhibitory neurons have been suggested to play a key role in enhancing the detection of behaviorally significant vocalization by lateral inhibition (Galindo-Leon et al., 2009). But recent imaging data argue that somatostatin interneurons rather than PV⁺ interneurons govern lateral inhibition in A1 (Kato et al., 2017). Our results are also consistent with the observation that in contrast to the SOM neurons, the changes in responses following sound exposure are similar in PV⁺ and pyramidal neurons (Kato et al., 2015; Khan et al., 2018). To what extent somatostatin or other interneurons subtypes

contribute to excitatory plasticity after auditory perceptual learning remains to be studied.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

ETHICS STATEMENT

The animal study was reviewed and approved by the Hebrew University's IACUC.

AUTHOR CONTRIBUTIONS

IM and AM designed the experiments. RS-Z and HS analyzed data. LF conducted the experiments. YE constructed the vocalization stimuli and software for their delivery. IM conducted the experiments and analyzed the data. IM, HS, and AM wrote the manuscript.

REFERENCES

- Ahissar, M., and Hochstein, S. (2004). The reverse hierarchy theory of visual perceptual learning. *Trends Cogn. Sci.* 8, 457–464. doi: 10.1016/j.tics.2004.08.011
- Angeloni, C., and Geffen, M. N. (2018). Contextual modulation of sound processing in the auditory cortex. *Curr. Opin. Neurobiol.* 49, 8–15. doi: 10.1016/j.conb.2017.10.012
- Aoki, R., Tsubota, T., Goya, Y., and Benucci, A. (2017). An automated platform for high-throughput mouse behavior and physiology with voluntary head-fixation. *Nat. Commun.* 8:1196. doi: 10.1038/s41467-017-01371-0
- Avermann, M., Tomm, C., Mateo, C., Gerstner, W., and Petersen, C. C. (2012). Microcircuits of excitatory and inhibitory neurons in layer 2/3 of mouse barrel cortex. *J. Neurophysiol.* 107, 3116–3134. doi: 10.1152/jn.00917.2011
- Bakin, J. S., and 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. doi: 10.1016/0006-8993(90)90035-a
- Ball, K., and Sekuler, R. (1987). Direction-specific improvement in motion discrimination. *Vis. Res.* 27, 953–965. doi: 10.1016/0042-6989(87)90011-3
- Bandyopadhyay, S., Shamma, S. A., and Kanold, P. O. (2010). Dichotomy of functional organization in the mouse auditory cortex. *Nat. Neurosci.* 13, 361–368. doi: 10.1038/nn.2490
- Barth, A. L., and Poulet, J. F. (2012). Experimental evidence for sparse firing in the neocortex. *Trends Neurosci.* 35, 345–355. doi: 10.1016/j.tins.2012.03.008
- Bathellier, B., Ushakova, L., and Rumpel, S. (2012). Discrete neocortical dynamics predict behavioral categorization of sounds. *Neuron* 76, 435–449. doi: 10.1016/j.neuron.2012.07.008
- Berardi, N., and Fiorentini, A. (1987). Interhemispheric transfer of visual information in humans: spatial characteristics. *J. Physiol.* 384, 633–647. doi: 10.1113/jphysiol.1987.sp016474
- Bieszczad, K. M., and Weinberger, N. M. (2010). Representational gain in cortical area underlies increase of memory strength. *Proc. Natl. Acad. Sci. U.S.A.* 107, 3793–3798. doi: 10.1073/pnas.1000159107
- Bizley, J. K., and Cohen, Y. E. (2013). The what, where and how of auditory-object perception. *Nat. Rev. Neurosci.* 14, 693–707. doi: 10.1038/nrn3565
- Briguglio, J. J., Aizenberg, M., Balasubramanian, V., and Geffen, M. N. (2018). Cortical neural activity predicts sensory acuity under optogenetic manipulation. *J. Neurosci.* 38, 2094–2105. doi: 10.1523/JNEUROSCI.2457-17.2017

FUNDING

This work was supported by an ERC consolidators grant to AM (#616063), ISF grant to AM (#224/17), and by Gatsby Charitable Foundation.

ACKNOWLEDGMENTS

We thank Israel Nelken and Livia de Hoz for help in setting up the initial behavioral experiments. We thank Yoni Cohen, David Pash, and Joseph Jubran for writing the software code of the Educage. We also thank Yonatan Loewenstein, Merav Ahissar, and the Mizrahi lab members for comments on the manuscript. This manuscript has been released as a pre-print at bioRxiv; doi: 10.1101/273342.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fncir.2019.00082/full#supplementary-material>

- Brown, M., Irvine, D. R., and Park, V. N. (2004). Perceptual learning on an auditory frequency discrimination task by cats: association with changes in primary auditory cortex. *Cereb. Cortex* 14, 952–965. doi: 10.1093/cercor/bhh056
- Buonomano, D. V., and Merzenich, M. M. (1998). Cortical plasticity: from synapses to maps. *Annu. Rev. Neurosci.* 21, 149–186. doi: 10.1146/annurev.neuro.21.1.149
- Ceballo, S., Piwkowska, Z., Bourg, J., Daret, A., and Bathellier, B. (2019). Targeted cortical manipulation of auditory perception. *Neuron* 104, 1168.e4–1179.e4. doi: 10.1016/j.neuron.2019.09.043
- Cohen, J. D., and Blum, K. I. (2002). Reward and decision. *Neuron* 36, 193–198. doi: 10.1016/s0896-6273(02)00973-x
- Cohen, L., and Mizrahi, A. (2015). Plasticity during motherhood: changes in excitatory and inhibitory layer 2/3 neurons in auditory cortex. *J. Neurosci.* 35, 1806–1815. doi: 10.1523/JNEUROSCI.1786-14.2015
- Cohen, L., Rothschild, G., and Mizrahi, A. (2011). Multisensory integration of natural odors and sounds in the auditory cortex. *Neuron* 72, 357–369. doi: 10.1016/j.neuron.2011.08.019
- Courtin, J., Chaudun, F., Rozeske, R. R., Karalis, N., Gonzalez-Campo, C., Wurtz, H., et al. (2014). Prefrontal parvalbumin interneurons shape neuronal activity to drive fear expression. *Nature* 505, 92–96. doi: 10.1038/nature12755
- Crist, R. E., Wu, L., and Gilbert, C. D. (2001). Learning to see: experience and attention in primary visual cortex. *Nat. Neurosci.* 4, 519–525. doi: 10.1038/87470
- Cruces-Solis, H., Jing, Z., Babaev, O., Rubin, J., Gur, B., Krueger-Burg, D., et al. (2018). Auditory midbrain coding of statistical learning that results from discontinuous sensory stimulation. *PLoS Biol.* 16:e2005114. doi: 10.1371/journal.pbio.2005114
- David, S. V., Fritz, J. B., and Shamma, S. A. (2012). Task reward structure shapes rapid receptive field plasticity in auditory cortex. *Proc. Natl. Acad. Sci. U.S.A.* 109, 2144–2149. doi: 10.1073/pnas.1117717109
- de Hoz, L., and Nelken, I. (2014). Frequency tuning in the behaving mouse: different bandwidths for discrimination and generalization. *PLoS One* 9:e91676. doi: 10.1371/journal.pone.0091676
- Diamond, D. M., and Weinberger, N. M. (1986). Classical conditioning rapidly induces specific changes in frequency receptive fields of single neurons in secondary and ventral ectosylvian auditory cortical fields. *Brain Res.* 372, 357–360. doi: 10.1016/0006-8993(86)91144-3
- Egnor, S. E., and Branson, K. (2016). Computational analysis of behavior. *Annu. Rev. Neurosci.* 39, 217–236. doi: 10.1146/annurev-neuro-070815-013845

- Ehret, G. (1975). Frequency and intensity difference limens and nonlinearities in the ear of the housemouse (*Mus musculus*). *J. Comp. Physiol.* 102, 321–336. doi: 10.1007/bf01464344
- Ericsson, K. A. (2006). The influence of experience and deliberate practice on the development of superior expert performance. *Camb. Handb. Expert. Expert Perform.* 38, 685–705.
- Erskine, A., Bus, T., Herb, J. T., and Schaefer, A. T. (2018). AutoNMouse: high throughput automated operant conditioning shows progressive behavioural impairment with graded olfactory bulb lesions. *bioRxiv* doi: 10.1101/291815 [Preprint].
- Feng, L., and Wang, X. (2017). Harmonic template neurons in primate auditory cortex underlying complex sound processing. *Proc. Natl. Acad. Sci. U.S.A.* 114, E840–E848. doi: 10.1073/pnas.1607519114
- Francis, N. A., and Kanold, P. O. (2017). Automated operant conditioning in the mouse home cage. *Front. Neural Circ.* 11:10. doi: 10.3389/fncir.2017.00010
- Froemke, R. C. (2015). Plasticity of cortical excitatory-inhibitory balance. *Annu. Rev. Neurosci.* 38, 195–219. doi: 10.1146/annurev-neuro-071714-034002
- Galindo-Leon, E. E., Lin, F. G., and Liu, R. C. (2009). Inhibitory plasticity in a lateral band improves cortical detection of natural vocalizations. *Neuron* 62, 705–716. doi: 10.1016/j.neuron.2009.05.001
- Ghose, G. M., Yang, T., and Maunsell, J. H. (2002). Physiological correlates of perceptual learning in monkey V1 and V2. *J. Neurophysiol.* 87, 1867–1888. doi: 10.1152/jn.00690.2001
- Gibson, E. J. (1969). *Principles of Perceptual Learning and Development*. New York, NY: Appleton-Century-Crofts.
- Gilbert, C. D., Sigman, M., and Crist, R. E. (2001). The neural basis of perceptual learning. *Neuron* 31, 681–697. doi: 10.1016/s0896-6273(01)00424-x
- Goel, A., Cantu, D., Guilfoyle, J., Chaudhari, G. R., Newadkar, A., Todisco, B., et al. (2017). Impaired perceptual learning in Fragile X syndrome is mediated by parvalbumin neuron dysfunction in V1 and is reversible. *bioRxiv* doi: 10.1101/217414 [Preprint].
- Han, Y. K., Kover, H., Insanally, M. N., Semerdjian, J. H., and Bao, S. (2007). Early experience impairs perceptual discrimination. *Nat. Neurosci.* 10, 1191–1197. doi: 10.1038/nn1941
- Harper, N. S., Schoppe, O., Willmore, B. D., Cui, Z., Schnupp, J. W., and King, A. J. (2016). Network receptive field modeling reveals extensive integration and multi-feature selectivity in auditory cortical neurons. *PLoS Comput. Biol.* 12:e1005113. doi: 10.1371/journal.pcbi.1005113
- Hattori, R., Kuchibhotla, K. V., Froemke, R. C., and Komiyama, T. (2017). Functions and dysfunctions of neocortical inhibitory neuron subtypes. *Nat. Neurosci.* 20, 1199–1208. doi: 10.1038/nn.4619
- Hawkey, D. J., Amitay, S., and Moore, D. R. (2004). Early and rapid perceptual learning. *Nat. Neurosci.* 7, 1055–1056. doi: 10.1038/nn1315
- Hennequin, G., Agnes, E. J., and Vogels, T. P. (2017). Inhibitory plasticity: balance, control, and codependence. *Annu. Rev. Neurosci.* 40, 557–579. doi: 10.1146/annurev-neuro-072116-031005
- Hensch, T. K. (2005). Critical period plasticity in local cortical circuits. *Nat. Rev. Neurosci.* 6, 877–888. doi: 10.1038/nrn1787
- Hippmeyer, S., Vrieseling, E., Sigrist, M., Portmann, T., Laengle, C., Ladle, D. R., et al. (2005). A developmental switch in the response of DRG neurons to ETS transcription factor signaling. *PLoS Biol.* 3:e159. doi: 10.1371/journal.pbio.0030159
- Hu, H., Gan, J., and Jonas, P. (2014). Fast-spiking, parvalbumin+ GABAergic interneurons: from cellular design to microcircuit function. *Science* 345:1255263. doi: 10.1126/science.1255263
- Irvine, D. R. (2017). Plasticity in the auditory system. *Hear. Res.* 362, 61–73. doi: 10.1016/j.heares.2017.10.011
- Irvine, D. R., Martin, R. L., Klimkeit, E., and Smith, R. (2000). Specificity of perceptual learning in a frequency discrimination task. *J. Acoust. Soc. Am.* 108, 2964–2968. doi: 10.1121/1.1323465
- Kaplan, E. S., Cooke, S. F., Komorowski, R. W., Chubykin, A. A., Thomazeau, A., Khibnik, L. A., et al. (2016). Contrasting roles for parvalbumin-expressing inhibitory neurons in two forms of adult visual cortical plasticity. *eLife* 5:e11450. doi: 10.7554/eLife.11450
- Karni, A., and Sagi, D. (1991). Where practice makes perfect in texture discrimination: evidence for primary visual cortex plasticity. *Proc. Natl. Acad. Sci. U.S.A.* 88, 4966–4970. doi: 10.1073/pnas.88.11.4966
- Kato, H. K., Asinof, S. K., and Isaacson, J. S. (2017). Network-level control of frequency tuning in auditory cortex. *Neuron* 95, 412.e4–423.e4. doi: 10.1016/j.neuron.2017.06.019
- Kato, H. K., Gillet, S. N., and Isaacson, J. S. (2015). Flexible sensory representations in auditory cortex driven by behavioral relevance. *Neuron* 88, 1027–1039. doi: 10.1016/j.neuron.2015.10.024
- Khan, A. G., Poort, J., Chadwick, A., Blot, A., Sahani, M., Mrsic-Flogel, T. D., et al. (2018). Distinct learning-induced changes in stimulus selectivity and interactions of GABAergic interneuron classes in visual cortex. *Nat. Neurosci.* 21, 851–859. doi: 10.1038/s41593-018-0143-z
- Kim, W. B., and Cho, J.-H. (2017). Encoding of discriminative fear memory by input-specific LTP in the amygdala. *Neuron* 95, 1129.e5–1146.e5. doi: 10.1016/j.neuron.2017.08.004
- Krakauer, J. W., Ghazanfar, A. A., Gomez-Marín, A., MacIver, M. A., and Poeppel, D. (2017). Neuroscience needs behavior: correcting a reductionist bias. *Neuron* 93, 480–490. doi: 10.1016/j.neuron.2016.12.041
- Kuchibhotla, K., and Bathellier, B. (2018). Neural encoding of sensory and behavioral complexity in the auditory cortex. *Curr. Opin. Neurobiol.* 52, 65–71. doi: 10.1016/j.conb.2018.04.002
- Kullmann, D. M., Moreau, A. W., Bakiri, Y., and Nicholson, E. (2012). Plasticity of inhibition. *Neuron* 75, 951–962. doi: 10.1016/j.neuron.2012.07.030
- Kurt, S., and Ehret, G. (2010). Auditory discrimination learning and knowledge transfer in mice depends on task difficulty. *Proc. Natl. Acad. Sci. U.S.A.* 107, 8481–8485. doi: 10.1073/pnas.0912357107
- Lagler, M., Ozdemir, A. T., Lagoun, S., Malagon-Vina, H., Borhegyi, Z., Hauer, R., et al. (2016). Divisions of identified parvalbumin-expressing basket cells during working memory-guided decision making. *Neuron* 91, 1390–1401. doi: 10.1016/j.neuron.2016.08.010
- Lawrence, D. H. (1952). The transfer of a discrimination along a continuum. *J. Comp. Physiol. Psychol.* 45, 511–516. doi: 10.1037/h0057135
- LeDoux, J. E. (2000). Emotion circuits in the brain. *Annu. Rev. Neurosci.* 23, 155–184. doi: 10.1146/annurev-neuro.23.1.155
- Lee, K., Holley, S. M., Shobe, J. L., Chong, N. C., Cepeda, C., Levine, M. S., et al. (2017). Parvalbumin interneurons modulate striatal output and enhance performance during associative learning. *Neuron* 93, 1451.e4–1463.e4. doi: 10.1016/j.neuron.2017.02.033
- Letzkus, J. J., Wolff, S. B., Meyer, E. M., Tovote, P., Courtin, J., Herry, C., et al. (2011). A disinhibitory microcircuit for associative fear learning in the auditory cortex. *Nature* 480, 331–335. doi: 10.1038/nature10674
- Luo, L., Callaway, E. M., and Svoboda, K. (2018). Genetic dissection of neural circuits: a decade of progress. *Neuron* 98, 256–281. doi: 10.1016/j.neuron.2018.03.040
- Madisen, L., Zwingman, T. A., Sunkin, S. M., Oh, S. W., Zariwala, H. A., Gu, H., et al. (2010). A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. *Nat. Neurosci.* 13, 133–140. doi: 10.1038/nn.2467
- Maor, I., Shalev, A., and Mizrahi, A. (2016). Distinct spatiotemporal response properties of excitatory versus inhibitory neurons in the mouse auditory cortex. *Cereb. Cortex* 26, 4242–4252. doi: 10.1093/cercor/bhw266
- Margrie, T. W., Meyer, A. H., Caputi, A., Monyer, H., Hasan, M. T., Schaefer, A. T., et al. (2003). Targeted whole-cell recordings in the mammalian brain *in vivo*. *Neuron* 39, 911–918. doi: 10.1016/j.neuron.2003.08.012
- Mermillod, M., Bugajska, A., and Bonin, P. (2013). The stability-plasticity dilemma: investigating the continuum from catastrophic forgetting to age-limited learning effects. *Front. Psychol.* 4:504. doi: 10.3389/fpsyg.2013.00504
- Mizrahi, A., Shalev, A., and Nelken, I. (2014). Single neuron and population coding of natural sounds in auditory cortex. *Curr. Opin. Neurobiol.* 24, 103–110. doi: 10.1016/j.conb.2013.09.007
- Murphy, T. H., Boyd, J. D., Bolaños, F., Vanni, M. P., Silasi, G., Haupt, D., et al. (2016). High-throughput automated home-cage mesoscopic functional imaging of mouse cortex. *Nat. Commun.* 7:11611. doi: 10.1038/ncomms11611
- Nelken, I. (2004). Processing of complex stimuli and natural scenes in the auditory cortex. *Curr. Opin. Neurobiol.* 14, 474–480. doi: 10.1016/j.conb.2004.06.005
- Nevin, J. A. (1969). Signal detection theory and operant behavior: a review of David M. Green and John A. Swets' Signal detection theory and psychophysics I. *J. Exp. Anal. Behav.* 12, 475–480. doi: 10.1901/jeab.1969.12-475
- Ohl, F. W., Wetzel, W., Wagner, T., Rech, A., and Scheich, H. (1999). Bilateral ablation of auditory cortex in Mongolian gerbil affects discrimination of frequency modulated tones but not of pure tones. *Learn. Mem.* 6, 347–362.
- O'Sullivan, C., Weible, A. P., and Wehr, M. (2019). Auditory cortex contributes to discrimination of pure tones. *eNeuro* 6:ENEURO.340-19.2019. doi: 10.1523/ENEURO.0340-19.2019

- Phillips, E. A., and Hasenstaub, A. R. (2016). Asymmetric effects of activating and inactivating cortical interneurons. *Elife* 5:e18383. doi: 10.7554/eLife.18383
- Polley, D. B., Steinberg, E. E., and Merzenich, M. M. (2006). Perceptual learning directs auditory cortical map reorganization through top-down influences. *J. Neurosci.* 26, 4970–4982. doi: 10.1523/jneurosci.3771-05.2006
- Rabinowitz, N. C., Willmore, B. D., King, A. J., and Schnupp, J. W. (2013). Constructing noise-invariant representations of sound in the auditory pathway. *PLoS Biol.* 11:e1001710. doi: 10.1371/journal.pbio.1001710
- Rabinowitz, N. C., Willmore, B. D., Schnupp, J. W., and King, A. J. (2011). Contrast gain control in auditory cortex. *Neuron* 70, 1178–1191. doi: 10.1016/j.neuron.2011.04.030
- Ramachandran, V., and Braddick, O. (1973). Orientation-specific learning in stereopsis. *Perception* 2, 371–376. doi: 10.1068/p020371
- Recanzone, G. H., Schreiner, C. E., and Merzenich, M. M. (1993). Plasticity in the frequency representation of primary auditory cortex following discrimination training in adult owl monkeys. *J. Neurosci.* 13, 87–103. doi: 10.1523/jneurosci.13-01-00087.1993
- Reed, A., Riley, J., Carraway, R., Carrasco, A., Perez, C., Jakkamsetti, V., et al. (2011). Cortical map plasticity improves learning but is not necessary for improved performance. *Neuron* 70, 121–131. doi: 10.1016/j.neuron.2011.02.038
- Resnik, J., and Polley, D. B. (2017). Fast-spiking GABA circuit dynamics in the auditory cortex predict recovery of sensory processing following peripheral nerve damage. *eLife* 6:e21452. doi: 10.7554/eLife.21452
- Roelfsema, P. R., and Holtmaat, A. (2018). Control of synaptic plasticity in deep cortical networks. *Nat. Rev. Neurosci.* 19, 166–180. doi: 10.1038/nrn.2018.6
- Rothschild, G., Nelken, I., and Mizrahi, A. (2010). Functional organization and population dynamics in the mouse primary auditory cortex. *Nat. Neurosci.* 13, 353–360. doi: 10.1038/nn.2484
- Rutkowski, R. G., and Weinberger, N. M. (2005). Encoding of learned importance of sound by magnitude of representational area in primary auditory cortex. *Proc. Natl. Acad. Sci. U.S.A.* 102, 13664–13669. doi: 10.1073/pnas.0506838102
- Schoups, A., Vogels, R., Qian, N., and Orban, G. (2001). Practising orientation identification improves orientation coding in V1 neurons. *Nature* 412, 549–553. doi: 10.1038/35087601
- Seung, H. S., and Sompolinsky, H. (1993). Simple models for reading neuronal population codes. *Proc. Natl. Acad. Sci. U.S.A.* 90, 10749–10753. doi: 10.1073/pnas.90.22.10749
- Seybold, B. A., Phillips, E. A., Schreiner, C. E., and Hasenstaub, A. R. (2015). Inhibitory actions unified by network integration. *Neuron* 87, 1181–1192. doi: 10.1016/j.neuron.2015.09.013
- Seymour, B., Daw, N., Dayan, P., Singer, T., and Dolan, R. (2007). Differential encoding of losses and gains in the human striatum. *J. Neurosci.* 27, 4826–4831. doi: 10.1523/jneurosci.0400-07.2007
- Sprekeler, H. (2017). Functional consequences of inhibitory plasticity: homeostasis, the excitation-inhibition balance and beyond. *Curr. Opin. Neurobiol.* 43, 198–203. doi: 10.1016/j.conb.2017.03.014
- Stiebler, I., Neulist, R., Fichtel, I., and Ehret, G. (1997). The auditory cortex of the house mouse: left-right differences, tonotopic organization and quantitative analysis of frequency representation. *J. Comp. Physiol. A* 181, 559–571. doi: 10.1007/s003590050140
- Talwar, S. K., and Gerstein, G. L. (2001). Reorganization in awake rat auditory cortex by local microstimulation and its effect on frequency-discrimination behavior. *J. Neurophysiol.* 86, 1555–1572. doi: 10.1152/jn.2001.86.4.1555
- Tasaka, G.-I., Guenther, C. J., Shalev, A., Gilday, O., Luo, L., and Mizrahi, A. (2018). Genetic tagging of active neurons in auditory cortex reveals maternal plasticity of coding ultrasonic vocalizations. *Nat. Commun.* 9:871. doi: 10.1038/s41467-018-03183-2
- Theunissen, F. E., and Elie, J. E. (2014). Neural processing of natural sounds. *Nat. Rev. Neurosci.* 15, 355–366. doi: 10.1038/nrn3731
- Vapnik, V. (1998). *Statistical Learning Theory*. New York, NY: Wiley.
- Vinje, W. E., and Gallant, J. L. (2000). Sparse coding and decorrelation in primary visual cortex during natural vision. *Science* 287, 1273–1276. doi: 10.1126/science.287.5456.1273
- Wehr, M., and Zador, A. M. (2003). Balanced inhibition underlies tuning and sharpens spike timing in auditory cortex. *Nature* 426, 442–446. doi: 10.1038/nature02116
- Weible, A. P., Liu, C., Niell, C. M., and Wehr, M. (2014). Auditory cortex is required for fear potentiation of gap detection. *J. Neurosci.* 34, 15437–15445. doi: 10.1523/JNEUROSCI.3408-14.2014
- Weinberger, N. M. (2004). Specific long-term memory traces in primary auditory cortex. *Nat. Rev. Neurosci.* 5, 279–290. doi: 10.1038/nrn1366
- Williamson, R. S., Ahrens, M. B., Linden, J. F., and Sahani, M. (2016). Input-specific gain modulation by local sensory context shapes cortical and thalamic responses to complex sounds. *Neuron* 91, 467–481. doi: 10.1016/j.neuron.2016.05.041
- Willmore, B., and Tolhurst, D. J. (2001). Characterizing the sparseness of neural codes. *Netw. Comput. Neural Syst.* 12, 255–270. doi: 10.1088/0954-898x/12/3/302
- Wolff, S. B., Gründemann, J., Tovote, P., Krabbe, S., Jacobson, G. A., Müller, C., et al. (2014). Amygdala interneuron subtypes control fear learning through disinhibition. *Nature* 509, 453–458. doi: 10.1038/nature13258
- Wright, B. A., and Fitzgerald, M. B. (2001). Different patterns of human discrimination learning for two interaural cues to sound-source location. *Proc. Natl. Acad. Sci. U.S.A.* 98, 12307–12312. doi: 10.1073/pnas.211220498
- Zeng, H., and Sanes, J. R. (2017). Neuronal cell-type classification: challenges, opportunities and the path forward. *Nat. Rev. Neurosci.* 18, 530–546. doi: 10.1038/nrn.2017.85
- Zhou, M., Liang, F., Xiong, X. R., Li, L., Li, H., Xiao, Z., et al. (2014). Scaling down of balanced excitation and inhibition by active behavioral states in auditory cortex. *Nat. Neurosci.* 17, 841–850. doi: 10.1038/nn.3701

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Maor, Shwartz-Ziv, Feigin, Elyada, Sompolinsky and Mizrahi. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.