Supporting Information

Wess et al. 10.1073/pnas.1710793114

SI Materials and Methods

Animals. All procedures were approved by the University of Maryland, College Park IACUC. Because of their immaturity at birth, we used ferrets to evaluate the emergence of sensory-driven responses in vivo. All ferrets were obtained from Marshall Farms.

In Vivo Physiology. Ferret kits (n = 48) of both sexes were anesthetized with 2–4% isoflurane–oxygen mixture. The patency of the ear canal was verified by visual inspection. Body temperature was maintained at 37 °C using a heating pad coupled with a rectal thermometer. A midline incision was made in the scalp, and the right temporal muscle was retracted to expose the skull. A steel post was cemented to the frontal portion of the skull. After which, a fine drill bit and surgical scalpel were used to make a small craniotomy (>2 mm) above the ACX. The dura mater was carefully removed and the craniotomy was thoroughly rinsed with sterile saline before recordings. The animal was head-fixed inside a sound-attenuating chamber. During the recordings, the isoflurane level was kept at 1–2%, and the depth of anesthesia was repeatedly monitored by paw withdrawal reflex. Extracellular electrode arrays (MEAs) (6 MΩ impedance; Microprobe) were inserted into A1, orthogonally to the cortical surface, using a motorized manipulator (MP285; Sutter Instruments). A ground electrode was placed within the temporalis muscle. Recording depth was identified by the z travel of the Sutter micromanipulator. Care was taken to minimize tissue dimpling at electrode insertion. The MEAs were advanced slowly and allowed to settle ∼10 min after each advancement before recording commenced. Recording locations were categorized as being in the CP (L4) if they were recorded in mid-cortical locations (depending on age 300–700 μm from pia) and putative SPNs if they were recorded >700–1,000 μm from pia based on histology (71–73) (Fig. S1).

Electrode signals were amplified and digitized using a 32-channel recording system (Neuralynx Cheetah). We recorded both spiking activity (filtered at 300–6,000 Hz, sampled at 33 kHz) and wideband LFP (filtered at 0.1–6,000 Hz, sampled at 5,988 Hz) simultaneously. All stimuli were generated using a computer-controlled digital signal processor (RX6; TDT), attenuated (PA5; TDT), and presented using a power amplifier (Crown) and high-output speakers (Fostex). Pure-tone sound stimuli, ranging between 1 and 35 kHz in 21 log-spaced steps, were presented pseudorandomly at intensities ranging from 64 to 94 dB SPL in 10-dB steps, for a total of 10 repetitions. Calibration was performed in situ. We observed that sound-evoked responses at young ages disappeared when repetition rates were increased, possibly due to immature synapses; thus, we presented stimuli at lower repetition rates. Spike sorting was carried out using a standard model of unsupervised clustering, and significant neuronal responses were identified using a binless algorithm, following which standard metrics were calculated. The detailed description of analysis methods is published elsewhere (74). These extracted standard metrics included the latency of first spikes and FRA characteristics.

After the recording session, ferrets were deeply anesthetized with >4% isoflurane and perfused with 4% paraformaldehyde in 0.1 M sodium phosphate buffer (PBS). Brains were subsequently removed and placed in the same solution for postfixation. Horizontal sections were cut (50 μm) on a microtome and Nissl stained to recover the electrode tracks. Cells were classified as driven if their firing rate was at least 2 SDs above the mean of the distribution of spontaneous rates for the population of cells. All data, categorized by age, and further divided by depth (CPN or SPN) were compared using n-way analysis of variance (ANOVA).

An MDA was used to assign neural responses to individual stimuli according to intensity. MDA was performed on the spike count vectors, forming the feature space obtained for each unit at each of the frequency-level combinations for n stimulus repetitions. The MDA estimates the probability with which each element of the stimulus matrix was assigned to the correct value based on the similarity between the observed responses to multiple repeats of the same stimulus. The centroid of each repeat was calculated, and every subsequent repeat was compared with the previous (random) repeat of the stimulus. The relationship between the predicted and predictor variables was graphically examined using confusion matrices and the mutual information calculated. The detailed description of analysis methods is published elsewhere (74).

For simultaneous laminar recordings in CP and SP, custom electrodes (2–6 MΩ impedance; Microprobe) consisting of two rows of either eight or two electrodes per row with the rows staggered in depth by 600 μm and rows spaced by 250 μm were used to record LFP signals (analog filtered at either 0.1–6,000 Hz or 10–6,000 Hz and sampled at 5,988 Hz). The shank of the array was advanced so that deep electrodes were at a depth of 800–1,300 μm (corresponding to shallow electrode depths of 200–700 μm). Spontaneous LFP activity was measured during 5 s preceding the stimulus onset and stimulus LFP activity during the stimulus presentation (0.5 s pure tones). Raw LFPs were digitally low pass-filtered (10th order Butterworth filter, 240 Hz cutoff), notch-filtered at 60 Hz and harmonics (third order Butterworth filter, cutoffs ± 2 Hz around 60 Hz and harmonics), and down-sampled by a factor of 6. Coherence was calculated using a multitaper method (75–77) essentially as the normalized zero-lag cross-correlation in frequency domain:

$$S_{ij} = S_i\, S_j^*$$

$$C_{XY} = \left| \frac{S_{XY}}{\sqrt{S_{XX}\, S_{YY}}} \right|,$$

where S_i and S_j are power spectral densities (PSDs) estimated using the multitaper method and C_{xy} is the coherence. PSDs were estimated using 0.5-s time windows, trial averaging (5–10 trials per stimulus), and 10 Hz smoothing parameter for the tapers. The down-sampled rate of 998 Hz and 0.5-s windows resulted in a frequency resolution of 2 Hz and was only estimated above 14 Hz to avoid potential breathing and heartbeat artifacts. Coherence variance was calculated by jackknifing across tapers. Mean coherence was calculated across all stimuli, all electrodes (grouped by deep or shallow), and across penetrations.

LFP-driven activity was measured by z scoring the stimulus interval activity relative to LFP variance measured during the 5-s prestimulus interval. Sliding windows spaced by 2 ms of either 50-, 100-, or 200-ms duration were used to find the maximum variance during the stimulus interval and then normalized by the prestimulus variance to arrive at the response amplitude z score. This allowed for both long latency slow oscillation responses and more adult-like auditory LFP responses (78) to be included as driven responses. Any z scores above 2 were taken as significant, and latencies were calculated for these responses as the center position of the first sliding window with a z score of 2 or greater. Amplitude was measured as

the root-mean-square (rms), and each FRA was calculated for each sound presentation at three sound levels.

The identification of local tuning similarity and global organization typically relies on the best frequency (BF) at the recording site. However, since FRAs can be broad or multipeaked, BF comparisons are somewhat noisy, and comparing the similarity of the whole FRA is more robust. The similarity of simultaneously recorded LFP FRAs was measured as the Cd between *z*-scored LFP responses at each FRA stimulus (frequency and sound level):

$$Cd_{st} = 1 - \frac{(x_s - \overline{x_s})(x_t - \overline{x_t})}{\sqrt{(x_s - \overline{x_s})(x_s - \overline{x_s})'}\sqrt{(x_t - \overline{x_t})(x_t - \overline{x_t})'}}$$

where x_s and x_t are the *z*-scored LFP FRAs obtained by the same sound frequency and level combinations for the pairs of electrodes s and t. Cds were then tabulated as a function of the distance between electrodes (in number of electrodes) and averaged for each age group and recording depth. Cds can vary between 0 and 2, with 1 indicating no correlation, 0 indicating perfect correlation, and 2 indicating anticorrelation.

Statistics. Results are plotted as means ± SD unless otherwise indicated. Populations are compared with a rank sum or Student's *t* test (based on Lilliefors test for normality) unless indicated otherwise. Post-stimulus time histogram (PSTH) variance comparison is done with *F* test and deemed significant if $P < 0.05$.

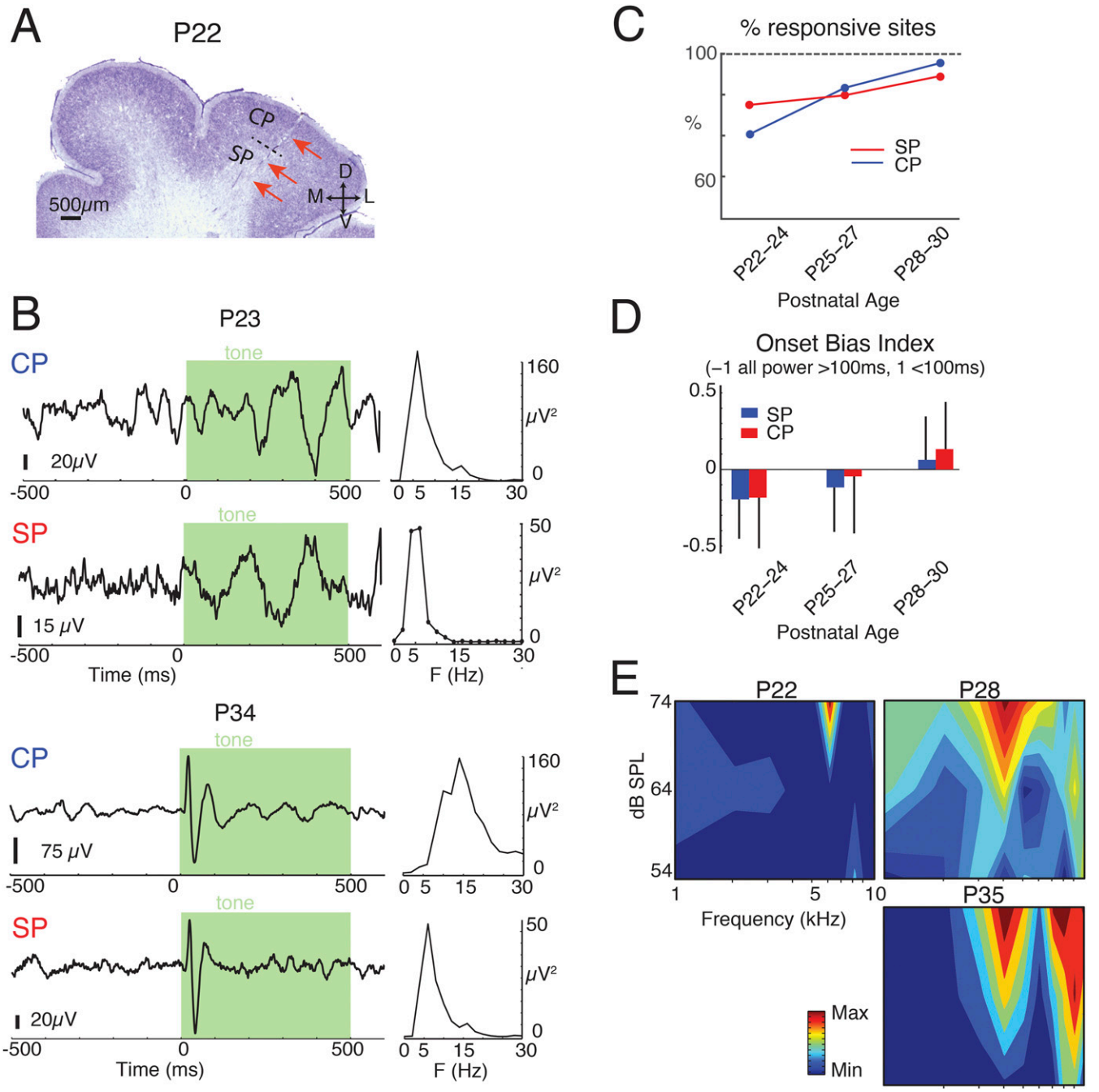

Fig. S1. Sound-evoked LFPs are present in ferret ACX after P21. (*A*) Cresyl violet section showing a recording track (red arrows) through ACX at P22. (Scale bar, 500 μm.) Dashed line indicates border of CP and SP. (*B*) Traces show LFP recordings during sound stimulation (green shading) in CP (*Top*) and SP (*Bottom*) at P23 and P34. Right graphs show PSD for each recording. Typical LFP responses in SP at P21–P24 demonstrated long latency coherent oscillatory activity during the duration of the stimulus with most oscillatory power around 5 Hz. This activity is similar to spindle burst activity recorded in neonatal cortex. (*C*) Fraction of recording sites in each layer that showed significant (z score > 2) sound-evoked responses (see *Materials and Methods*). (*D*) Ratio of LFP power in the first versus second 100 ms of sound stimulus. Sound responses in older animals are briefer. (*E*) Exemplar SP LFP FRAs at three different ages.

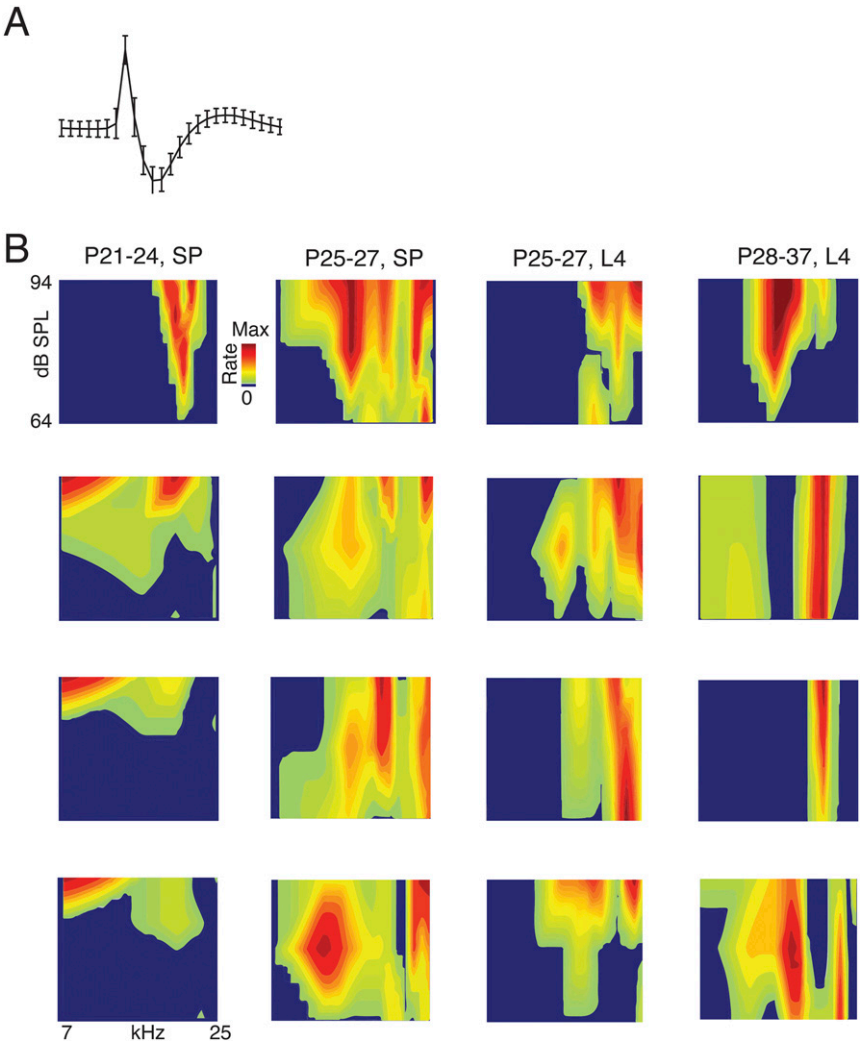

Fig. S2. Sound-evoked single unit activity in developing ferret ACX. (*A*) Trace shows action potential after spike sorting from a single unit. (*B*) Exemplar single unit FRAs from CP (putative L4) and SP. Max rates (*Top* to *Bottom*): SP P21–P24 4.0, 0.6, 0.6, 0.6 Hz; SP P25–27 3.2, 1.2, 1.2, 2.2 Hz; L4 P25–P27 4.4, 2.0, 2.6, 2.8 Hz; L4 P28–P37 5.2, 4.1, 3.4, 2.4 Hz. Note that while single-peaked tuning curves can be observed, many tuning curves have multiple response areas.

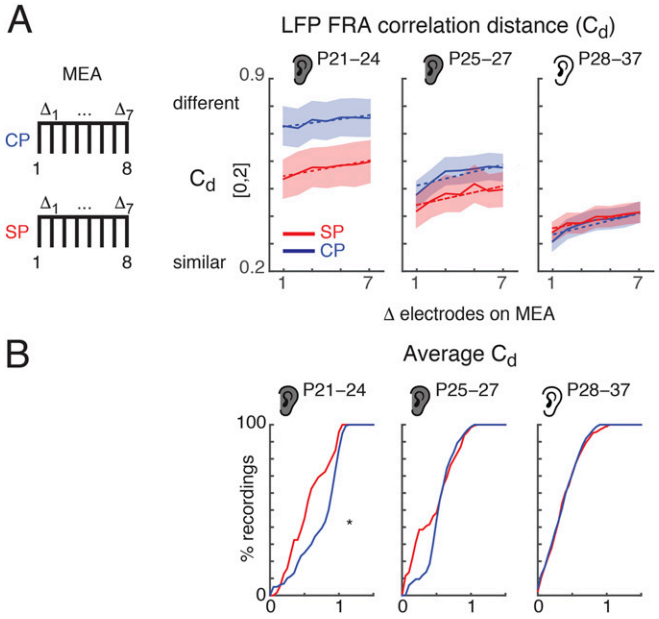

Fig. S3. LFP correlations decrease with distance. (*A*) Average pairwise Cds of the LFP FRA; lower values indicate more similarity between FRAs. Averages are grouped for pairs compared at different physical electrode distances (Δ electrodes) and at the different ages. Schematic on *Left* shows MEAs in SP and CP. Graphs on *Right* show Cd as a function of electrode separation. Plotted are means ± SEM. Dashed line shows linear regression to the data (*y* intercepts: P22–P24 CP, SP: 0.72, 0.53; P25–P27 CP, SP: 0.49, 0.43; P28–P30 CP, SP: 0.43, 0.32). Slopes are positive, indicating that Cd increases with increasing electrode distance (for slopes: P22–P24 CP, SP: $P = 0.0023$, $P = 0.00018$; P25–P27 CP, SP: $P = 0.006$, $P = 0.007$; P28–P30 CP, SP: $P = 0.002$, $P = 0.001$). (*B*) Cumulative distribution function (CDF) of the Cd averaged over all electrode distances (Δ electrodes in *A*). Cd is smaller at P21–24 in SP than CP ($P < 0.001$, rank sum).

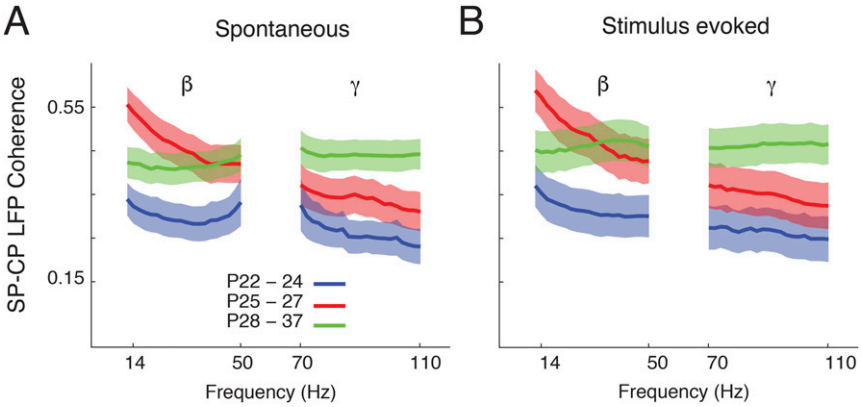

Fig. S4. Coherence of LFP activity in SP and CP. Graphs show coherence (normalized spectral cross-correlation) of LFP activity during simultaneous recordings in SP and CP under spontaneous or stimulus-evoked conditions. Coherence essentially measures phase alignment. Coherence was measured in the beta (14–30 Hz) and gamma band (30–110 Hz). At a particular frequency, if two signals have an identical phase, then coherence would be one. Conversely, if two signals have completely different phases, coherence would be zero. A strong interaction between the layers could be caused by a common input or interactions between the two layers. Coherence was present at all ages for both spontaneous (*A*) and sensory-evoked (*B*) LFPs. Coherence was strongest in the beta band under both conditions. Plotted are means and shaded ± STD over the population of animals.