

**Figure S1.** Sound induced sharpening of OS in L2/3 pyramidal neurons. (Related to Figure 1)

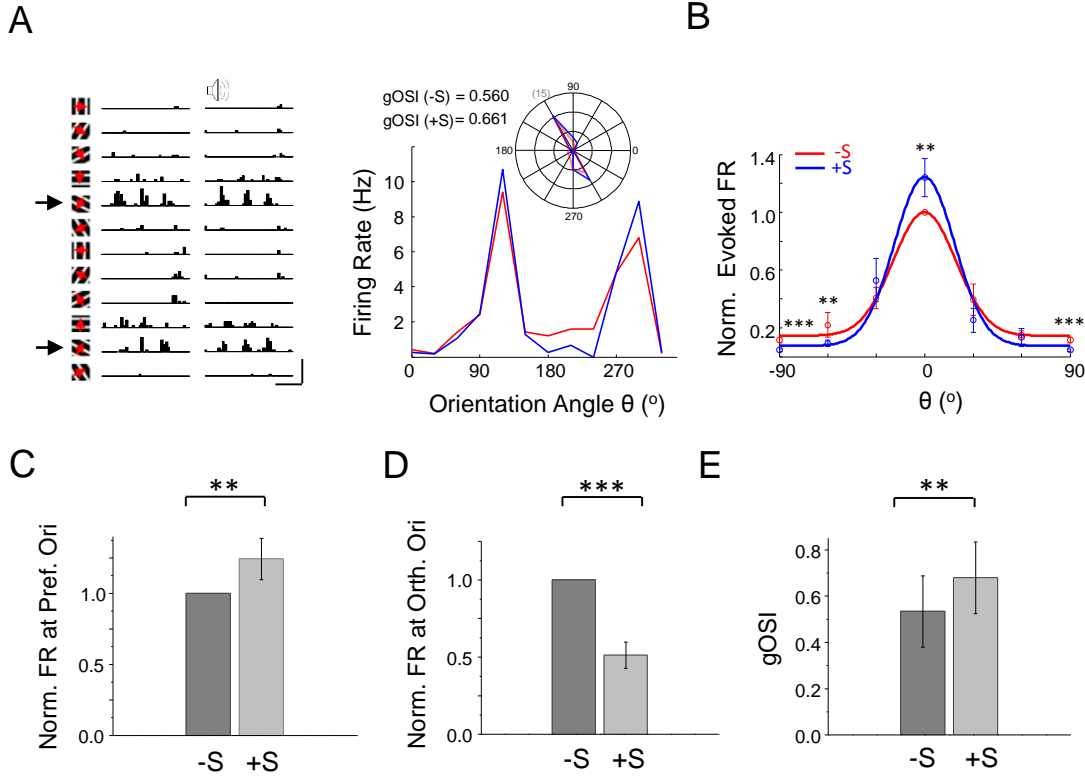
(A) Mean percentage change in gOSI value induced by sound in high contrast and low contrast visual stimulation conditions. Bar = SD, \*  $p < 0.05$ , paired t-test.

(B) Spontaneous firing rates without and with sound stimulation (left panel) or LED stimulation (right panel) plotted for individual cells. The average firing rate in no sound condition was 1.18 Hz and in the sound condition was 0.85 Hz. The average firing rate in the LED off condition was 0.83 Hz and in the LED on condition was 0.60 Hz. Bar = SEM. \*,  $p < 0.05$ , paired t-test.

(C) Plot of the evoked firing rate (FR) at the preferred orientation in the presence (+S) versus that in the absence (-S) of sound stimulation for L4 excitatory neurons. Inset, mean evoked firing rate normalized to the value in the sound off condition of all the cells recorded. 'ns',  $p > 0.5$ , paired t-test ( $n = 8$ ). Bar = SEM.

(D) Plot of the evoked firing rate at the orthogonal orientation in the presence (+S) versus that in the absence (-S) of sound stimulation for L4 cells. Inset, mean evoked firing rate normalized to the value in the sound off condition of all the cells. 'ns',  $p > 0.5$ , paired t-test. Bar = SEM.

(E) Plot of gOSI in the presence (+S) versus absence (-S) of sound stimulation for L4 cells. Inset, mean gOSI value. 'ns',  $p > 0.5$ , paired t-test. Bar = SD.

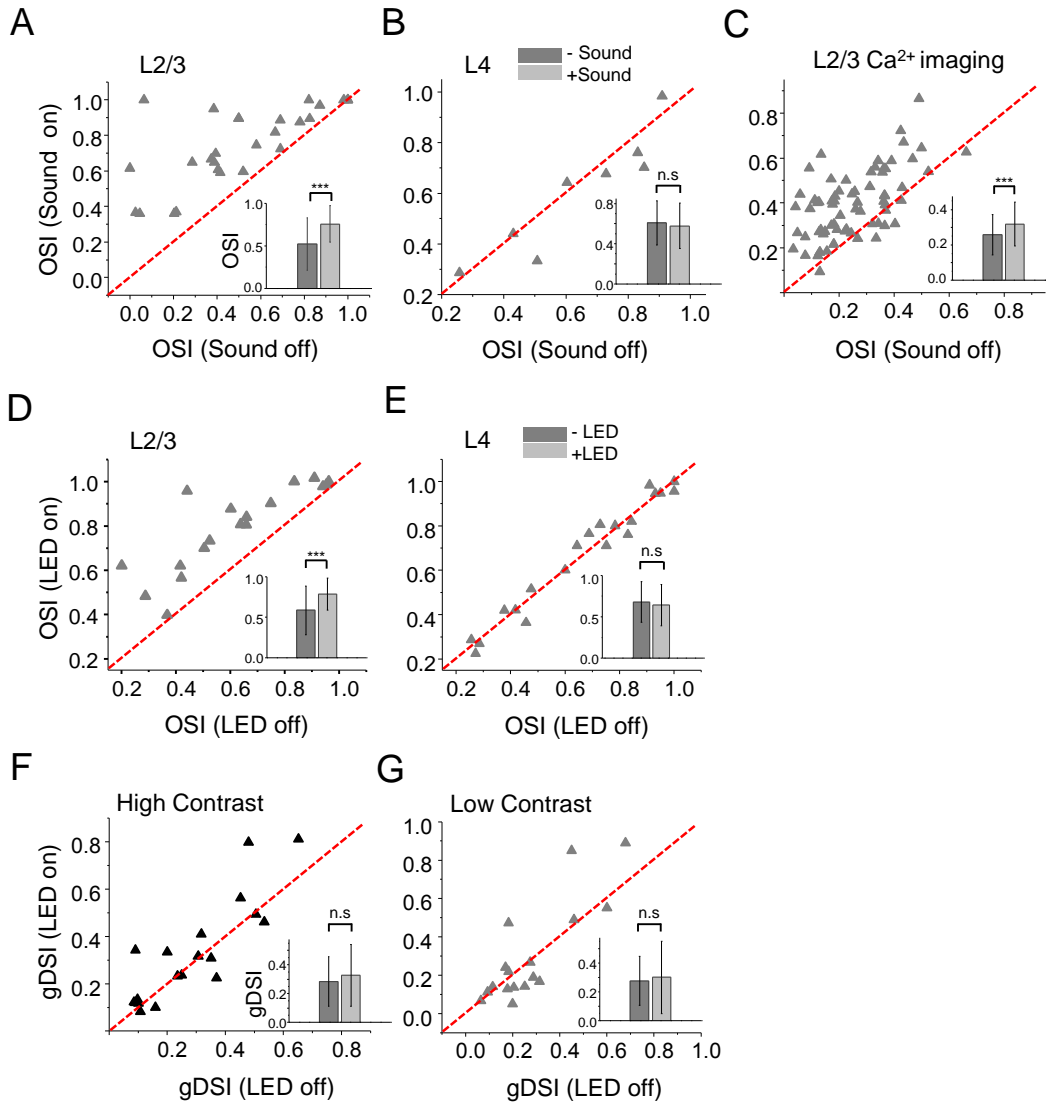


**Figure S2.** Sound induced sharpening of OS in awake mice. (Related to Figure 1)

(A) Left panel, PSTHs of spike responses of an example L2/3 pyramidal neuron to drifting sinusoidal gratings without (left) and with (right) coupling with sound stimulation. The preferred orientation of the cell was marked by arrows. Scale: 60 Hz, 500 ms. Right panel, firing rates at the different stimulus directions for the same cell. Red: visual stimulation only. Blue: visual plus sound stimulation. Top inset, polar graph plotting of orientation tuning of the cell. The axial value (number of spikes evoked) is marked within the parentheses. The gOSI values without (-S) and with (+S) are labeled.

(B) Average normalized firing rates for 10 pyramidal cells without (red) and with (blue) sound, fitted with a Gaussian function. The tuning curves were aligned according to the preferred orientation (defined as  $0^\circ$ ) in the sound off condition. Responses to the two directions of the same orientation were averaged. Error bar = SEM. \*\*,  $p < 0.01$ , \*\*\*,  $p < 0.001$ , paired t-test. Error bar = SEM.

(C-E) Mean normalized evoked firing rate of all cells at preferred orientation (C), orthogonal orientation (D), and mean gOSI (E) in the sound off and sound on conditions. \*\*,  $p < 0.01$ , \*\*\*,  $p < 0.001$ , paired t-test. Error bar = SEM (C, D) or SD (E).



**Figure S3.** Plots of OSI and gDSI measurements. (Related to Figures 1,2 and 3)

(A) Plot of the OSI value in the sound off versus sound on condition for L2/3 pyramidal neurons (from loose-patch recording data). Inset, mean OSI value of all cells. \*\*\*,  $p < 0.001$ , paired t-test.

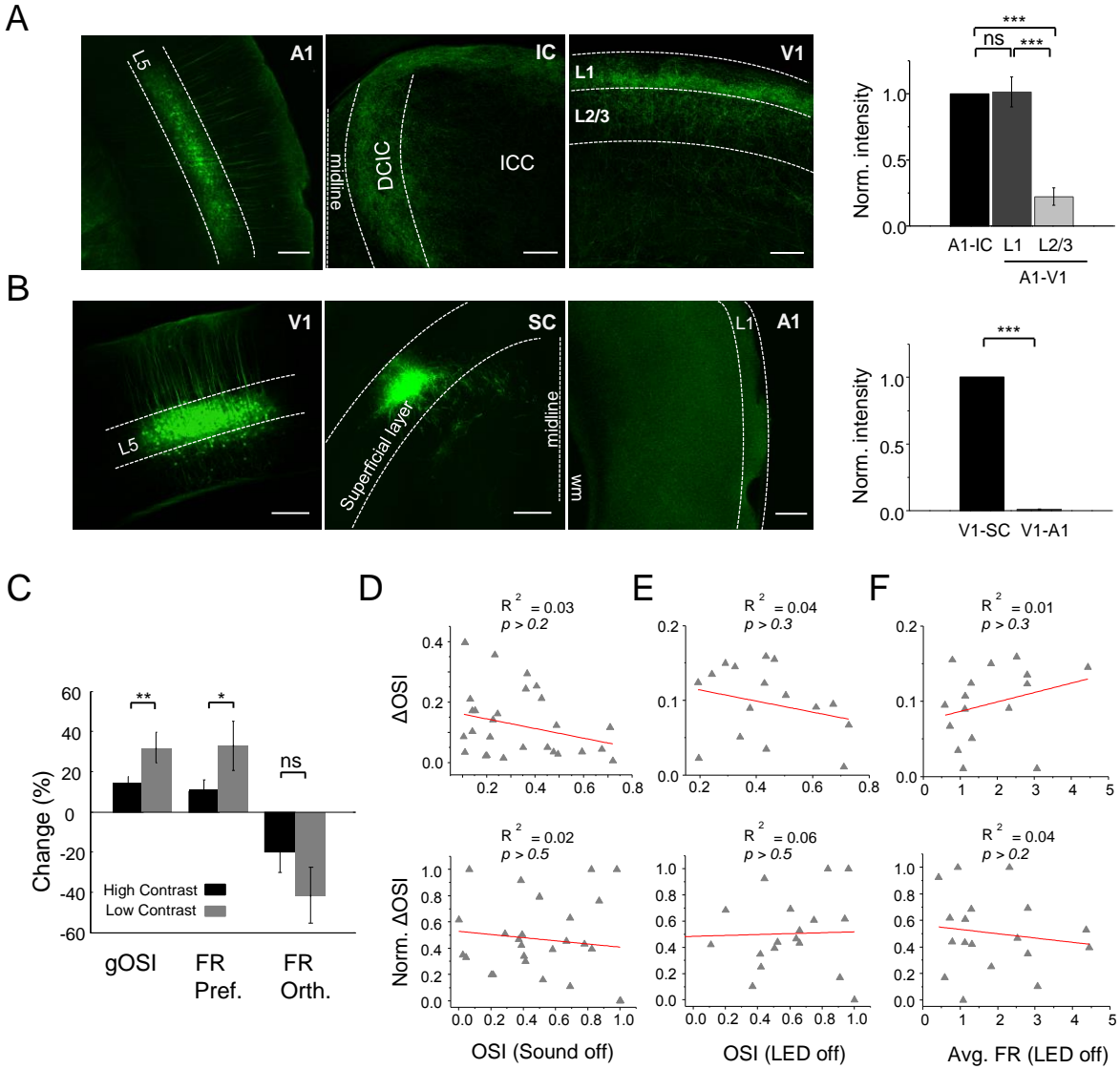
(B) Plot of the OSI values in the sound off versus sound on condition for L4 neurons (from loose-patch recording data). 'ns',  $p > 0.05$ , paired t-test.

(C) Plot of the OSI value in the sound off versus sound on condition for Ca<sup>2+</sup> responses of L2/3 neurons.

(D) Plot of the OSI value in the LED off versus LED on condition for L2/3 neurons.

(E) Plot of the OSI value in the LED off versus LED on condition for L4 neurons.

(F,G) Plots of the gDSI value in the LED off versus LED on condition for L2/3 neurons under high contrast (F) and low contrast (G) visual stimulation. Inset, mean gDSI value of all the cells.



**Figure S4.** Effects of optogenetic stimulation of A1-V1 projections. (Related to Figure 3)

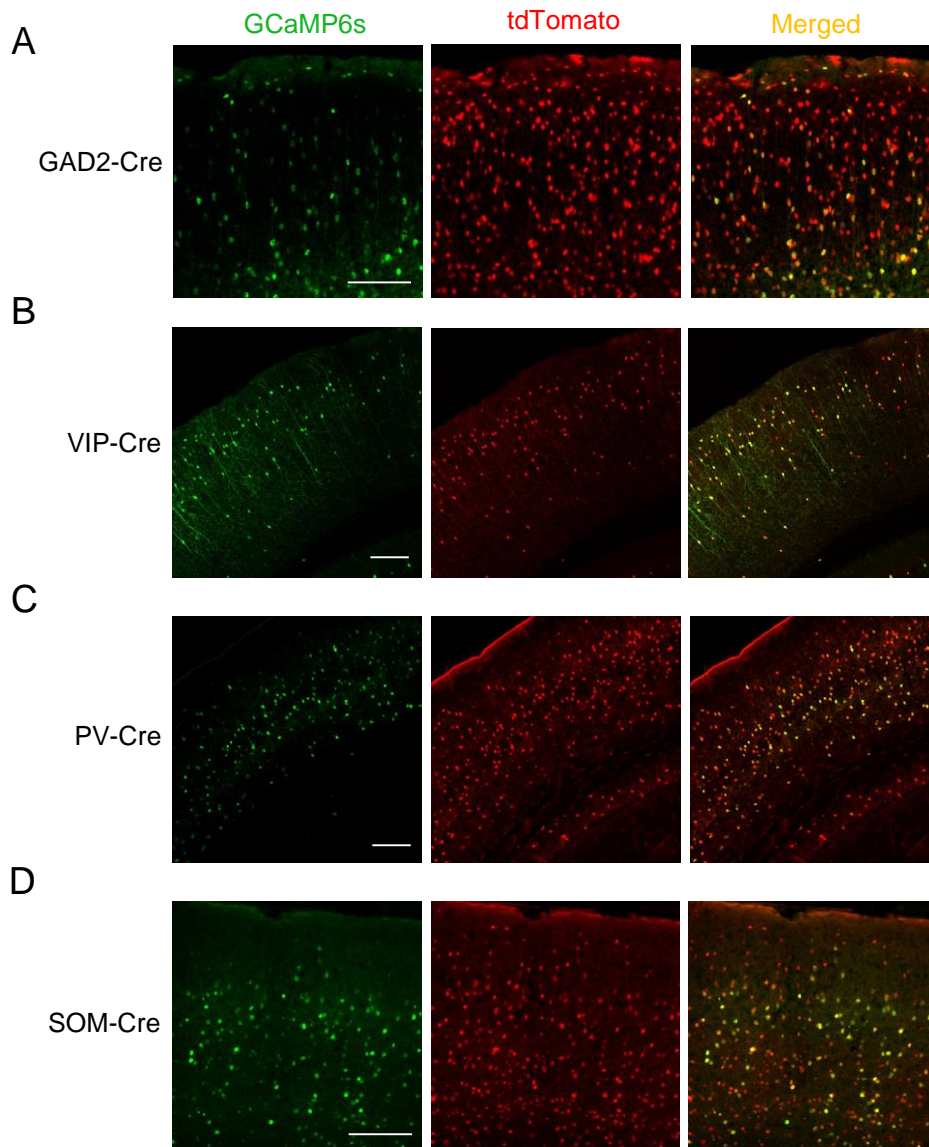
(A) Images of fluorescence labeled A1 neurons (first panel), the A1 axons projected into inferior colliculus (IC) (second panel), and that into V1 (third panel) from the same Rpb4-Cre animal injected with AAV encoding floxed hChR2-EYFP in A1. Scale Bar = 200 $\mu$ m. Right most panel, fluorescence intensity was quantified within an array of 100X100 $\mu$ m boxes along L1, L2/3 and dorsal cortex of IC (DCIC), and the values were averaged (n = 3 mice), and normalized to the average value for the A1-DCIC projection. Bar = SD. 'ns', p = 0.1; \*\*\* p < 0.001, two-sample t-test.

(B) Images of fluorescence labeled V1 neurons (first panel), of the V1 axons projected into superior colliculus (SC) (second panel), and of A1 area (third panel) from the same Rpb4-Cre animal injected with AAV-floxed-GFP in V1. Scale Bar = 200  $\mu$ m. wm = white matter. Note that only auto-fluorescence was seen in A1 under higher levels of exposure. Right most panel, quantification of projection density, normalized to the value in SC (n = 3 mice).

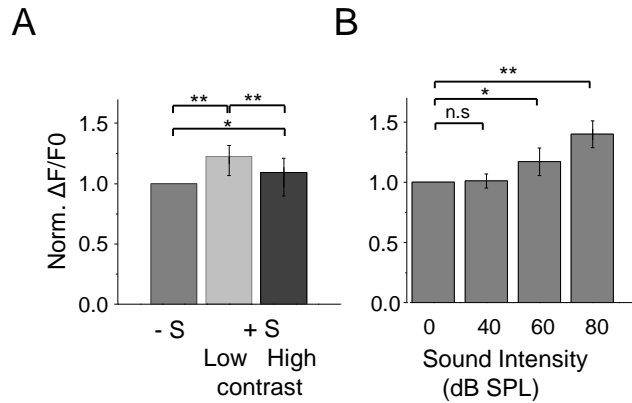
(C) Percentage changes in gOSI, firing rate at the preferred orientation, firing rate at the orthogonal orientation induced by LED stimulation, compared between high contrast and low contrast visual stimulation conditions for the same set of cells. Bar = SD. 'ns', p = 0.07; \* p < 0.05; \*\* p < 0.01, paired t-test. Note: This data is from a subset of cells shown in Fig. 3C-3J) recorded under the two visual contrast conditions for each individual cell.

(D,E) The change in OSI plotted against the original OSI value (in the sound/LED off condition) (upper panel), or the normalized change in OSI versus the original OSI value (bottom panel). Data in low visual contrast condition were used. Normalized change was calculated as  $\Delta\text{OSI}/(1-\text{OSI}_{\text{off}})$ . R is the correlation coefficient. P value is indicated.

(F) The change (upper) or normalized change (bottom) in OSI plotted against the average firing rate across orientations in the LED off condition.



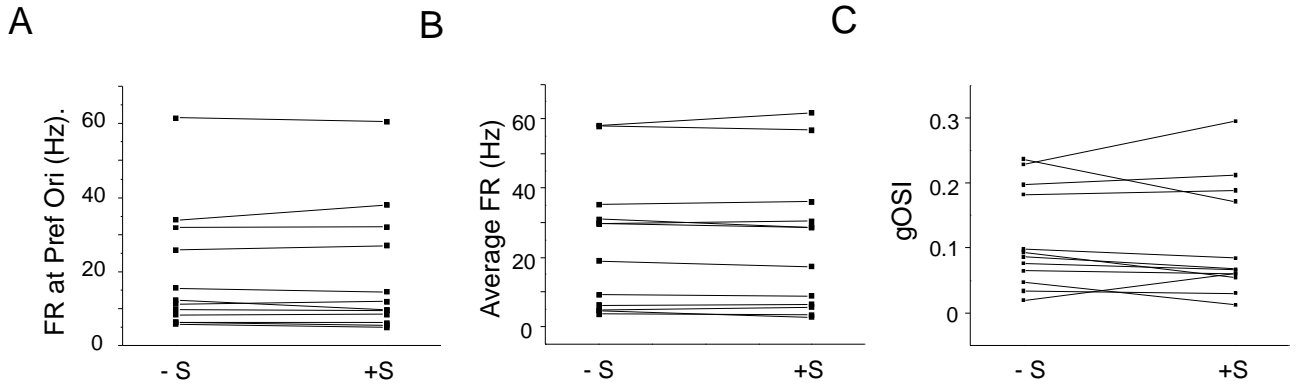
**Figure S5.** Expression of GCaMP6s in various Cre::tdTomato mice. (Related to Figures 5 and 7) Left panel, GCaMP6s expression (green) in V1 of a GAD2-Cre (A), VIP-Cre (B), PV-Cre (C) and SOM-Cre (D)::tdTomato mouse. Middle panel, tdTomato expression (red) in V1 of the corresponding animal. Right panel, superimposed image of GCaMP6s and tdTomato expression. Note that all green cells are also red. Scale = 250  $\mu$ m.



**Figure S6.** Effect of sound on visual responses of L1 neurons. (Related to Figure 5)

(A) Mean normalized peak  $\text{Ca}^{2+}$  response amplitude (to the value under no sound condition) of all the neurons ( $n = 40$ ) under low and high contrast visual stimulation. \*  $p < 0.05$ ; \*\*  $p < 0.01$ , one-way ANOVA with repeated measures.

(B) Mean normalized  $\text{Ca}^{2+}$  response amplitude (to the value at zero intensity) to flash visual stimulation at different sound intensities for all the L1 neurons tested. 'ns',  $p > 0.05$ ; \*  $p < 0.05$ ; \*\*  $p < 0.01$ , one-way ANOVA with repeated measures.



**Figure S7.** *In vivo* loose-patch recordings from fast-spiking neurons in layer 2/3. (Related to Figure 7)

(A) Evoked firing rate at the preferred orientation without (-S) and with (+S) sound, plotted for all recorded fast-spiking cells in L2/3 ( $n = 12$ ). Data points for the same cell are connected by a line ( $p > 0.05$ , paired t-test). Fast-spiking neurons were identified by their narrow spike waveforms (see Experimental Procedures).

(B) Average evoked firing rate across all orientations, plotted for all the cells ( $p > 0.05$ , paired t-test).

(C) The gOSI values under sound off and sound on conditions plotted for all the cells ( $p > 0.05$ , paired t-test).

## Supplemental Experimental Procedures

All experimental procedures used in this study were approved by the Animal Care and Use Committee of USC. In all experiments, adult (45–60 days) female mice were used.

### Viral injection

For L5-specific expression, Rbp4-Cre mice were crossed with tdTomato reporter (Ai14) mice (The Jackson Laboratory, C57BL/6J background). For injections, we anaesthetized mice with 2% isoflurane, thinned the skull over A1 (or V1) and performed  $\sim 0.2$  mm<sup>2</sup> craniotomy. We delivered the virus using a beveled glass micropipette (tip opening: 40–50  $\mu$ m in diameter) attached to a microsyringe pump (World Precision Instruments). Adeno-associated viruses (AAVs) to deliver hChR2 were acquired from the UPenn Viral Vector Core: AAV2/9.EF1 $\alpha$ .DIO.hChR2(H134R)-EYFP.WPRE.hGH (Addgene 20298). We made one injection of virus in each mouse at a volume of 50 nl and at a rate of 15 nl/min in A1 at a depth of 600  $\mu$ m (n = 15 mice). We then sutured the scalp, and administered an analgesic (0.1 mg/kg Buprenex). We made *in vivo* recordings 2–3 weeks after viral injections. We examined the expression pattern of hChR2(H134R)-EYFP in each injected mouse to make sure that the expression level was sufficiently strong before the experiment. For slice recording, non-floxed AAV9.hSyn.hChR2(H134R)-eYFP.WPRE.hGH (Addgene26973P, also from UPenn Vector Core) was injected into A1 (one injection at a depth of 600  $\mu$ m) of PV-, SOM-, VIP- or GAD2-Cre::tdTomato mice (n = 3 mice each). For calcium imaging of pyramidal neurons, non-floxed AAV2/1.Syn.GCaMP6s.WPRE.SV40 (UPenn Vector Core) was injected into V1 (at two depths of 200 and 400  $\mu$ m and at two sites) at a of wild-type C57BL/6J mice (n = 4 mice). For calcium imaging of inhibitory neurons, AAV2/1.Syn.Flex.GCaMP6s.WPRE.SV40 (UPenn Vector Core) was injected into V1 (at two depths of 200 and 400  $\mu$ m at two sites) of GAD2-, PV-, SOM- or VIP-Cre::tdTomato mice (n = 4 mice each). For inhibiting L1 neurons, AAV1-CAG-FLEX-ArchT-GFP (Addgene 28307, from UNC vector core) was injected into GAD2-Cre mice via iontophoresis using 3 $\mu$ A currents for 5 sec, at a depth of 50  $\mu$ m and at 4 different sites within V1. Cells included in the analysis were obtained only from mice that had ArchT-GFP expression largely limited to L1 only (n = 4 mice).

### Retrograde tracer injection

For retrograde tracer injections into V1, 40 nl of fluorescently conjugated Cholera Toxin subunit B (CTb Alexa 488, 0.25%; Invitrogen) was pressure injected into V1 through a pulled glass micropipette at a rate of 15 nl/min and at a depth of 400  $\mu$ m at one site (n = 3 mice in total). After 5–7 days, the animal was deeply anesthetized and transcardially perfused with 4% paraformaldehyde. Brain tissue was sliced into 150  $\mu$ m sections using a vibratome (Leica), and sections were mounted onto glass slides and imaged under a confocal microscope (Olympus).

### Animal surgery for *in vivo* recording in anaesthetized mice

We sedated the mouse with the correct expression with an intramuscular injection of chlorprothixene hydrochloride (10 mg/kg in 4 mg/ml water solution) and then anesthetized it with urethane (1.2 g/kg, i.p., at 20% w/v in saline), as previously described (Li et al., 2013; Liu et al., 2009; Niell and Stryker, 2008). We maintained the animal's body temperature at  $\sim 37.5^\circ$  by a heating pad (Havard Apparatus, MA). We performed tracheotomy, and inserted a small glass capillary tube to maintain a free airway. We performed cerebrospinal fluid draining, removed the skull and dura mater ( $\sim 1 \times 1$  mm) over the V1, and applied artificial cerebrospinal fluid solution (ACSF, containing [in mM] 140 NaCl, 2.5 KCl, 2.5 CaCl<sub>2</sub>, 1.3 MgSO<sub>4</sub>, 1.0 NaH<sub>2</sub>PO<sub>4</sub>, 20 HEPES, 11 glucose, pH 7.4) to the exposed cortical surface when necessary. We trimmed eyelashes contralateral to the recording side, and covered the eyes with ophthalmic lubricant ointment until recording, at which time we rinsed the eyes with saline and applied a thin layer of silicone oil (30,000 centistokes) to prevent drying while allowing clear optical transmission. Eye movements and the receptive field drift were negligible within the time window of our recording (Mangini and Pearlman, 1980). For optogenetic activation applied to the surface of A1, the ipsilateral eye was enucleated to eliminate potential stimulation of this eye by the blue LED light.

### Animal surgery for *in vivo* recording in awake mice

One week before electrophysiological recordings, mice were anesthetized with isoflurane (1.5% by volume) and a screw for head fixation was mounted on top of the skull with dental cement, as previously described (Xiong et al., 2013; Zhou et al., 2014). Afterward 0.1 mg/kg buprenorphine was injected subcutaneously before they were returned to home cages. During the recovery period, the mice were trained to get accustomed to the head fixation on



the recording plate, which was flat and could rotate smoothly around its center. To fix the head, the head screw was tightly fit into a metal post. One day before electrophysiological recordings the mouse was anesthetized with isoflurane and craniotomy was made over the V1 (2.6 mm, lateral to midline; 3.9 mm posterior to bregma). We then covered the opening with Vaseline and then agarose, and injected the animal with buprenorphine before returning it to the home cage. On the day of the recording, the animal was fixed to a metal post through the screw, and agarose and Vaseline were cleared. Due to potential complex effects of locomotion on visual responses (Fu et al., 2014; Niell and Stryker, 2010), we have only selected recording trials in which animals remained stationary.

### ***In vivo* electrophysiology**

We pre-penetrated the pia with a broken pipette under visual guidance before *in vivo* recordings, and then performed cell-attached recordings with an Axopatch 200B amplifier (Molecular Devices). To record from excitatory neurons, the patch pipette had a tip opening of  $\sim 2 \mu\text{m}$  (4 – 5 M $\Omega$  impedance). The intrapipette solution contained (in mM): 140 NaCl, 2.5 KCl, 2.5 CaCl<sub>2</sub>, 1.3 MgSO<sub>4</sub>, 1.0 NaH<sub>2</sub>PO<sub>4</sub>, 20 HEPES, 11 glucose, pH = 7.4. A 100–250 M $\Omega$  seal was formed on the targeted neuron. We recorded spikes in the voltage-clamp mode with a small command potential applied to achieve a zero baseline current. The spike signal was filtered at 10 kHz and sampled at 20 kHz. All the neurons recorded under this condition showed regular-spikes (the spike waveform had a trough-to-peak interval of  $0.85 \pm 0.10$  ms,  $n = 72$  cells), consistent with the sampling bias towards excitatory neurons as shown previously with cell morphology reconstructions (Liu et al., 2010; Wu et al., 2008). The recordings were done in layer 2/3 (200–350  $\mu\text{m}$  from the pia) and layer 4 (375–510  $\mu\text{m}$ ). The layer assignment of the blindly recorded neurons was made mostly according to the vertical travel distance of the electrode. The assignment was reasonably precise because our use of a high-magnification objective (40 $\times$ ) on the microscope allowed a precise identification of the cortical surface and our application of pre-penetration minimized the dimpling of the cortical surface (Li et al., 2013; Li et al., 2012a). To record from fast-spiking (FS) inhibitory neurons, smaller pipettes with a higher impedance (10 M $\Omega$ ) were used and neurons with fast-spikes (trough-to-peak interval of the spike waveform  $< 0.5$  ms) were actively searched (Wu et al., 2008).

### **Visual stimulation**

The visual stimuli were generated using Matlab with Psychophysics Toolbox and were displayed with a gamma-corrected LCD monitor (refresh rate 75 Hz, maximum luminance 280 cd/m<sup>2</sup>) placed 0.25 m away from the right eye. We placed the center of the monitor at 45° Azimuth, 25° Elevation, and it covered  $\pm 35^\circ$  horizontally and  $\pm 27^\circ$  vertically of the mouse visual field. We made recordings in the monocular zone of the V1. We recorded spontaneous activity when presenting a uniform grey background. To measure orientation tuning, we applied drifting sinusoidal gratings (spatial frequency of 0.04 cycles per degree and temporal frequency of 2Hz) of 12 directions (30° steps) in a random sequence. The visual stimulation with and without sound or LED illumination were alternated, but the stimulus sequence was randomized independently for sound/LED off and sound/LED on trials. We set the inter-stimulus interval at 10 s to allow a full recovery of ChR2 function from desensitization (Li et al., 2013). Each cell was recorded under high contrast (95%) and low contrast (25%) conditions. We applied five to ten sets of stimuli to each cell, with the sequence different between sets. The long recording time in our experimental conditions prevented us from applying different combinations of spatial and temporal frequency. To better drive as large fraction as possible of V1 neurons, we have carefully chosen our visual stimuli (0.04 cpd, 2Hz) based on a previous study of a large number of neurons (Niell and Stryker, 2008), which shows that the largest fraction of mouse V1 cells prefers the spatial frequency of 0.04 cpd, and the average preferred spatial frequency in L2/3 is  $\sim 0.04$  cpd, and that most of units prefer a temporal frequency of  $\sim 2$ Hz. Since the level of change in selectivity was not correlated with the initial selectivity level or the overall response level (Supplementary Figure 4D-4F), this suggests that both optimally driven and non-optimally driven neurons could increase their selectivity under sound/A1 stimulation.

### **Auditory stimulation**

The auditory stimulation consisted of white noise pulses at 70 dB sound pressure level (SPL) presented at 10Hz (50 ms on, 50 ms off) throughout the duration of the visual stimulus. The onset and offset of auditory stimulation were the same as visual stimulation. The sound was delivered by a single speaker located contralateral to the recording side. The visual stimuli alone and those coupled with sound, were alternated, but the stimulus sequence was randomized independently for sound off and sound on trials.

### ***In vivo* optogenetic manipulation**

To photoactivate hChR2, we used a blue (470 nm) fiber-coupled LED (0.8 mm diameter, Doric Lenses) placed on top of the exposed cortical surface. LED light spanned the entire area of V1. We applied black pigment stained agar to cover the tip of the optic fiber, as to prevent LED light leakage reaching the contralateral eye. We had verified that LED light did not directly stimulate the contralateral eye in wild-type mice (data not shown). The LED was driven by the analog output from a NIDAQ board (National Instruments). The intensity of LED was around 5 mW (measured at the tip of the fiber). To photoactivate ArchT, we used a green light (530 nm) fiber-coupled LED (0.8 mm diameter, Doric Lenses) and followed the same procedure as with hChR2.

### ***In vitro* electrophysiology**

*Slice preparation.* Viral injected mice were anesthetized with urethane. After decapitation, the brain was rapidly removed into an ice-cold oxygenated dissection buffer (60 mM NaCl, 3 mM KCl, 1.25 mM NaH<sub>2</sub>PO<sub>4</sub>, 25 mM NaHCO<sub>3</sub>, 115 mM sucrose, 10 mM glucose, 7 mM MgCl<sub>2</sub>, 0.5 mM CaCl<sub>2</sub>; bubbled with 95% O<sub>2</sub> and 5% CO<sub>2</sub>; pH = 7.4). Coronal cortical slices of 350 μm thickness were cut from the infected brain hemisphere by a vibrating microtome (Leica VT1000s). After being incubated in a warmed (at 34 °C) artificial cerebral spinal fluid (ACSF; 126 mM NaCl, 2.5 mM KCl, 1.25 mM Na<sub>2</sub>PO<sub>4</sub>, 26 mM NaHCO<sub>3</sub>, 1 mM MgCl<sub>2</sub>, 2 mM CaCl<sub>2</sub>, 0.5 mM ascorbic acid, 2 mM sodium pyruvate, and 10 mM glucose, bubbled with 95% O<sub>2</sub> and 5% CO<sub>2</sub>) for >30 min, the slice was transferred to the recording chamber at room temperature.

*Electrophysiological recording.* Recording was made under an upright fluorescence microscope (Olympus BX51WI) equipped with an infrared light source. Slices were examined under a 4X objective before recording to determine whether hChR2-EYFP was expressed in A1. In slices with good expression sites, whole-cell voltage-clamp recordings were selectively performed on fluorescence-labeled inhibitory neurons in PV-Cre, SOM-Cre, VIP-Cre or GAD2-Cre::tdTomato slices or non-fluorescent excitatory cells in GAD2-Cre::tdTomato slices under epifluorescence imaging in V1. The extracellular solution contained: 60 mM NaCl, 3 mM KCl, 1.25 mM NaH<sub>2</sub>PO<sub>4</sub>, 25 mM NaHCO<sub>3</sub>, 115 mM Glucose, 10 mM Sucrose, 7 mM MgCl<sub>2</sub>, 0.5 mM CaCl<sub>2</sub>, PH 7.2, and bubbled with 95% O<sub>2</sub> and 5% CO<sub>2</sub>. For examining monosynaptic excitatory responses only, recordings were made with TTX (a sodium channel blocker, 1 μM) and 4-aminopyridine (a potassium channel blocker, 1 mM) present in the external solution. Glass pipette (4-7 MΩ impedance) was filled with a cesium-based internal solution (125 mM Cs-gluconate, 5 mM TEA-Cl, 4 mM MgATP, 0.3 mM GTP, 10 mM phosphocreatine, 10 mM HEPES, 1 mM EGTA, 2 mM CsCl, 1% biocytin, pH = 7.2). The pipette and whole-cell capacitances were completely compensated and the initial series resistance was compensated for 50% at 100 μS lag. Recordings were made with an Axopatch 200B amplifier (Molecular Devices). Excitatory synaptic currents were recorded by clamping the cell's membrane potential at -70 mV. Signals were filtered at 2 kHz and sampled at 10 kHz. In each slice, multiple neurons were recorded. The cortical depth of each recorded cell was based on the vertical distance of the cell body from the pial surface of the cortex, which was set as 0 μm. The distance was measured with a micromanipulator coupled with a digital reader (SD Instrument DR1000). L1 was defined as 0 – 150 μm.

*Photostimulation.* The hChR2 was activated by blue light pulses from a mercury Arc lamp gated by an electronic shutter (Li et al., 2014). The excitation light was passed through a blue light filter and the objective. A calibrated aperture placed at the conjugate plane of the slice was used to control the size of the illumination area. The aperture was adjusted so that the entire V1 area was illuminated. The power of light stimulation was 3 mW measured at the focal plane. Brief light pulses (3 ms) were applied individually (0.033 Hz). For each condition, 10 – 30 trials were given and responses were averaged.

### ***In vivo* two-photon calcium imaging**

Imaging was performed after at least 2 weeks of viral expression. We used a custom built Mai Tai (Spectra-physics) based 2-photon system and recorded data using a custom-modified version of the Scanimage software (Pologruto et al. 2003). Calcium imaging from labeled cells was performed in multiple subregions each spanning ~200 x 200 up to ~400 x 400 μm. Scan lines were drawn across each of the clearly visible cell bodies (typically 5-15 cells) and then imaged continuously (rapidly alternating between all the scan lines) while presenting moving sinusoidal grating stimuli (1.5 sec) at 10 second intervals. Depending on the number of scan lines the scan rate ranged between 25-70 Hz. Within each scan line, we manually defined regions of interests (ROIs) based on the presence of fluorescence transients. The selected ROIs were compared with the 2 dimensional snapshot of the region to make sure that we were imaging the labeled cell bodies rather than neuropils. The signal within a ROI was processed using standard

methods to derive fractional change over baseline, i.e.  $\Delta F/F_0$  (Jia et al. 2011). Neurons located within 150  $\mu\text{m}$  below the pia surface were considered as L1 cells.

### Data analysis

We performed data analysis with custom-developed software (LabVIEW, National Instruments; and MATLAB, Mathworks). We counted the spikes evoked by drifting sinusoidal gratings within a time window covering the visual stimulation duration with a 70 ms delay, and subtracted the average spontaneous firing rate from the stimulus-evoked spike rate.

We quantified the strength of OS with a global orientation selectivity index (gOSI), which considers responses at all test orientations (Mariño et al., 2005):

$$\text{gOSI} = \frac{\|\sum R(\theta) \times e^{2i\theta}\|}{\sum R(\theta)}$$

$i$  is  $\sqrt{-1}$ .  $\theta$  is the angle of the moving direction.  $R(\theta)$  is the response level at angle  $\theta$ . The preferred orientation was determined from this vector sum of all the responses. We then averaged responses to the gratings of opposite directions, and obtained the orientation tuning curve between 0 – 180 degrees, and fitted it with a Gaussian function:  $R(\theta) = A \times \exp(-0.5 \times (\theta - \phi)^2 / \sigma^2) + B$ .  $\phi$  is the determined preferred orientation and  $\sigma$  is the tuning width. We also computed an orientation selectivity index (OSI) defined as  $(R_{\text{pref}} - R_{\text{orth}}) / (R_{\text{pref}} + R_{\text{orth}})$ , where  $R_{\text{pref}}$  is the response level at the angle of  $\phi$ , and  $R_{\text{orth}}$  is that at the angle of  $\phi + 90$ . The gOSI and OSI values vary between 0 and 1, with 0 being the value for an untuned neuron and 1 for a perfectly tuned neuron (Carandini and Ferster, 2000; Mariño et al., 2005; Ringach et al., 2002). To compare tuning between conditions, the preferred orientation was the peak of the Gaussian fit for the data in the sound/LED off condition (e.g. in Figure 2E, right panel), or was the test orientation closest to the peak of the Gaussian fit in the sound/LED off condition (e.g. in Figure 1B, 1C, 2E left panel, and 2G). Similarly, the orthogonal orientation (e.g. in Figure 1D, 2H) was also determined based on the data in the sound/LED off condition of each cell. Tuning curves were aligned according the preferred orientation (set as zero degree) before averaging. The global direction selectivity index (gDSI) was quantified according to:

$$\text{gDSI} = \frac{\|\sum R(\theta) \times e^{i\theta}\|}{\sum R(\theta)}$$

For slice recording data, peak response (EPSC) amplitude was measured for each cell. In order to take into account the varying levels of channelrhodopsin expression among different slices, we quantified a relative response amplitude. Specifically, in each slice, at least three L2/3 pyramidal cells were recorded, and their EPSC amplitudes were averaged to obtain a mean pyramidal cell response for that slice. The mean pyramidal cell responses of different slices were then averaged to obtain a global average pyramidal cell response. In each slice, a scaling factor was determined by the ratio between the mean pyramidal cell response for that slice and the global average pyramidal cell response. EPSC amplitudes of all of cells recorded in the same slice were then scaled according to the scaling factor for that slice.

### Statistical analysis

We first performed Lilliefors test to check whether the data were normally distributed. In the case of a normal distribution, we performed paired t-test. Otherwise, we performed a non-parametric test (Wilcoxon signed-rank test in this study). For the L1 inhibition experiment, one-way ANOVA with repeated measures with Bonferroni post-hoc test was performed. For the slice recording experiment, the relative input strengths of different groups were compared using one-way ANOVA with Bonferroni post-hoc test. No statistical method was used to pre-determine sample sizes, but our sample sizes were similar to those reported in previous publications in the field.