June 11, 2019

Spectral Hallmark of Auditory-Tactile Interactions in Somatosensory Cortex

by

Manning Zhang[1,†], Sung Eun Kwon[2,¥], Manu Ben-Johny[1,Ω], Daniel H. O'Connor[2], and John B. Issa[1,+,‡]

[1]Department of Biomedical Engineering
The Johns Hopkins University School of Medicine
Baltimore, MD 21205, USA

[2]Solomon H. Snyder Department of Neuroscience
The Johns Hopkins University School of Medicine, Kavli Neuroscience Discovery Institute, and Brain Science Institute
Baltimore, MD 21205, USA

[†]Current Address: Department of Biomedical Engineering, Washington University in St. Louis, St. Louis, MO 63130

[¥]Current Address: Department of Molecular, Cellular and Developmental Biology, University of Michigan, Ann Arbor, MI 48109

[Ω]Current Address: Department of Physiology and Cellular Biophysics, Columbia University, New York, NY 10032

[+]Current Address: Department of Neurobiology, Northwestern University, Evanston, IL 60201

[‡]To whom correspondence should be addressed: john.issa@gmail.com

ABSTRACT

To synthesize a coherent representation of the external world from diverse sensory information, the brain must integrate inputs across stimulus modalities. Yet the mechanistic basis of this computation at the level of neuronal populations remains obscure. Here, we investigate tactile-auditory integration in the mouse somatosensory cortex using two-photon Ca^{2+} imaging. Pairing sound with whisker stimulation modulates tactile responses in both S1 and S2, with the most prominent modulation being robust inhibition of responses in S2. The degree of inhibition depends on tactile stimulation frequency, with lower frequency responses the most severely attenuated. Alongside these neurons, we identify sound-selective neurons in S2 whose responses are inhibited by high tactile frequencies. These results are consistent with a hypothesized local mutually-inhibitory S2 circuit that spectrally selects tactile versus auditory inputs. Our findings enrich mechanistic understanding of multisensory integration and suggest a key role for S2 in combining auditory and tactile information.

INTRODUCTION

Our understanding of the external world is derived from what our senses can tell us[1]. We can enrich this representation of the environment by combining inputs across disparate sensory modalities[2]. However, maintaining a coherent percept requires active moment-to-moment integration of multisensory inputs along with resolution of inconsistencies between modalities[3]. How this active process is carried out at the level of neural populations is an open question.

Two stimulus modalities that are deeply entangled are sound and touch. Mechanical forces generated by physical contact not only activate somatosensory mechanoreceptors but also generate acoustic waves that are detected by the machinery of the auditory system[4]. Thus these stimuli are often highly correlated in the environment[5]. A significant behavioral consequence of this interaction is observed when tactile and auditory inputs are not in register: as a human subject touches a surface, if a sound is played that is different from what is expected from the tactile sensation of the surface, the reported roughness of the surface is altered in a manner dependent on the sound frequency[6]. This "parchment-skin illusion" points to the deep bond shared by these modalities and hints at the centrality of stimulus frequency as a parameter that may bind them together[7,8]. Consistent with this idea, the secondary somatosensory cortex of mice can respond to sound[9]. Furthermore, retrograde tracing in Mongolian gerbils has revealed direct synaptic projections from primary somatosensory cortex to primary auditory cortex[10] and single unit recordings in behaving monkeys have shown that neurons in anterior parietal areas respond to task-related auditory cues[11], hinting at a cohesive sound-touch integration circuit within cortex.

In this study, we leverage these observations to probe the neural correlates of multisensory integration in the neocortex. We aim to determine how stimulus frequency is encoded by neurons in the somatosensory system, whether concurrent sound can influence these responses, and how

this information is carried across a population of neurons. Thus, we performed large-scale two-photon Ca^{2+} imaging in multiple cortical regions: the primary (S1) and secondary (S2) somatosensory cortices, the insular somatosensory field (ISF), and the auditory cortex. We find that neurons in S1 and S2 encode the frequency at which a whisker is deflected. Concurrent sound modulates these responses, with responses to low tactile frequencies often completely abolished. In S2, we find a small number of sound-selective neurons whose responses are reciprocally attenuated by high-frequency whisker deflections. These frequency-dependent interactions point towards a spectrally-dependent mutually inhibitory circuit between touch-selective and sound-selective neurons and shed light on the neural circuits that may underlie the computations involved in multisensory integration.

RESULTS

Widefield Ca^{2+} Imaging to Localize Tactile and Auditory Areas of Mouse Neocortex

We first performed widefield Ca^{2+} imaging of the left temporal lobe of mouse neocortex to identify regions that respond to either tactile or auditory stimuli individually. Transgenic mice expressing the Ca^{2+}-indicator GCaMP6s under pan-neuronal promoters[12–14] were implanted with a chronic cranial window exposing up to 5 mm of the left hemisphere[15,16]. This approach allowed us to capture the majority of the auditory cortex along with the primary somatosensory (S1), the secondary somatosensory (S2), and the insular somatosensory (ISF; insular somatosensory field) cortices all in the same window. Mice were habituated to head fixation while running on a spherical treadmill. During widefield fluorescence imaging, either tactile or auditory stimuli were delivered to the contralateral target whisker or ear, respectively (Figure 1a). In response to a sinusoidal 128 Hz stimulus applied to the C2 whisker, transient increases in GCaMP6s fluorescence were observed in somatosensory areas, with smaller responses in the auditory fields (Figure 1b-c). A spatial map of these changes, which averages the change in fluorescence over 1 s of the response for stimulus frequencies ranging from 36 to 128 Hz (Figure 1d), highlights three distinct locations that were strongly responsive to tactile stimuli. Based on stereotaxic coordinates (predicted location of C2 barrel of S1 indicated by red cross), the most medial locus was identified as S1, the middle locus corresponded to S2, and the most lateral locus was consistent with the location of ISF identified previously with intrinsic optical imaging[17]. In response to auditory stimuli composed of sinusoidal amplitude modulated (SAM; Figure S1) tones, we observed significant fluorescent responses in regions corresponding to the auditory cortex (AC) (the primary auditory cortex (AI), the secondary auditory cortex (AII), and the anterior auditory field (AAF)) (Figure 1e-f). SAM tones were chosen to approximate a correspondence between the envelope of auditory stimuli and

the frequency of the sinusoidal waveform of the tactile stimuli used throughout this study (in the range of 2 to 128 Hz). Interestingly, we observed a small transient in ISF, consistent with overlap of the nearby insular auditory field (Figure 1g-h)[17,18].

Neurons Across Somatosensory Cortex Are Tuned to the Frequency of Tactile Stimulation

Having located each cortical area within our cranial window, we next examined the responses of individual layer 2/3 neurons across multiple somatosensory regions. Using two-photon Ca^{2+} imaging, we were able to monitor the activity of hundreds of neurons within a field-of-view (FOV) spanning 465×465 μm^2, a region sufficiently large to include the C2 barrel column in S1 (Figure 2a)[19]. To measure evoked tactile responses, we chose sinusoidal piezo-driven tactile stimuli with frequencies ranging from 2 to 128 Hz presented in random order. The upper limit of 128 Hz was chosen to avoid any piezo-generated sound within the hearing range of mice (see Methods and Figure S1). We found 16% of neurons in S1 were responsive to passive whisker stimulation (598 of 3723 neurons, 16 imaging fields from 8 mice). For comparison, a previous study found 16.8% of layer 2/3 excitatory neurons in the principal column responded to touch while 39%[19] or 37%[16] showed task-related tactile responses. Figure 2b displays individual baseline-normalized fluorescent responses ($\delta\hat{F} = \Delta F/F_0$) of two exemplar neurons to tactile stimuli of increasing frequency (multiple trials in gray; ensemble average in black, sorted according to stimulus frequency). A normalized tactile response for each stimulus was then estimated as the ratio of the mean fluorescence change ($\delta\hat{F}_{mean}(i)$) for the given stimulus (i) to the maximal fluorescence change ($\max_i \delta\hat{F}_{mean}(i)$) observed across all stimuli. Plotting the normalized tactile response as a function of frequency revealed a normalized tactile tuning curve (Figure 2c). The best tactile frequency (*BTF,* see Methods*)* served as a convenient measure of the tuning preference of individual neurons. Figure 2d shows the overall distribution of *BTF* for tactile-responsive neurons

in S1. We found that neurons can be classified into two categories, namely "bandpass" and "highpass" according to their response properties (Figure 2c). The bandpass neurons have biphasic tuning curves (Figure 2c, left) while the highpass neurons exhibit monotonically increasing responses (Figure 2c, right). As we limited tactile frequencies to between 2-128 Hz, we could not exclude the possibility that highpass neurons may exhibit bandpass properties when tested beyond 128 Hz. Practically, however, this classification scheme could be used to facilitate comparison of tuning preferences across regions. In S1, highpass neurons outnumbered bandpass neurons 6-fold in relative prevalence.

We further assessed tactile responses of neurons within S2 (Figure 2e), approximately ~1 mm lateral to S1 based on widefield Ca^{2+} imaging results (Figure 1d). We found that 9% of all neurons were responsive to tactile stimuli (1313 of 14051 neurons, 67 imaging fields from 19 mice). Again, examination of tactile tuning curves revealed the presence of neurons with both bandpass (left, Figure 2e-f) and highpass (right, Figure 2e-f) response profiles. Bandpass neurons were modestly but significantly more prevalent in S2 compared to S1 (19.8% for S2 and 15.9% for S1, $p < 0.05$ by Chi-square proportion test; Figure 2g). The highpass neurons nonetheless outnumbered bandpass neurons by ~4-fold. The coarse similarity in response characteristics of both S1 and S2 neurons indicates that tactile frequency tuning may be a general feature for encoding tactile sensation in multiple regions of the somatosensory cortex.

Concurrent Auditory Input Modulates Tactile Responses in S2

To identify possible multimodal integration, we explored whether concurrent auditory stimuli could modify tactile responses in the somatosensory cortex, focusing initially on S2 given its putative role in higher-order tactile processing[20–22]. Accordingly, we measured fluorescence changes of single neurons in S2 in response to three distinct stimulus conditions: (1) tactile stimuli

alone, (2) auditory stimuli (SAM tones) alone, and (3) tactile and auditory stimuli presented concurrently. Figure 3 displays responses of three tactile-selective neurons within a single FOV demonstrating various multisensory effects. Neuron 1 exhibited robust responses to high tactile frequencies (black, Figure 3b) that were enhanced upon concurrent presentation of acoustic stimuli (red, Figure 3b). The normalized responses obtained with tactile stimuli alone and with concurrent tactile and auditory stimuli are overlaid in Figure 3c to facilitate comparison. The responses here were normalized to the maximal fluorescence change observed with either tactile or concurrent tactile and auditory stimuli. By contrast, neuron 2, which was also tuned to the high-frequency tactile stimuli, demonstrated inhibition of this response when concurrent auditory stimulation was added (Figure 3b-c). Finally, neuron 3 exhibited robust bandpass responses when probed with tactile stimuli alone (black, Figure 3b) but was largely nonresponsive when both tactile and auditory stimuli were presented concurrently (red, Figure 3b). Overlay of normalized tactile responses in Figure 3c, right, across a broad range of tactile stimulus frequencies in the presence (red) and absence (black) of sound illustrates the dramatic sound-driven inhibition observed in neuron 3. All three neurons were nonresponsive to all sound stimuli tested when these sound stimuli were presented alone (blue, Figure 3b; see Methods). These results illustrate a powerful mode of multisensory processing whereby the response of individual cortical neurons to their primary sensory modality is modulated by a concurrent secondary sensory input. These effects are diverse, allowing for neuron-specific computations of multisensory interaction.

To assess the prevalence of each mode of tactile-auditory integration in S2, we combined results from 67 FOVs across 19 animals. Consequently, for all tactile-selective neurons in S2, we systematically probed for statistically significant changes in tactile responses by concurrent auditory stimulation (2-way ANOVA with a pre-established criterion of $p < 0.05$ for significance).

We found that most neurons had tactile responses that were either inhibited or unchanged by sound, with neurons tuned to lower frequencies more likely to be inhibited (Figure 3d). Furthermore, to probe the overall potency of sound modulation, we quantified the net change in aggregate tactile responses (average response elicited across all tactile frequencies) when evoked by concurrent tactile and auditory stimuli (ΔR) (Figure 3e). The relative strength of sound-driven inhibition was most potent in neurons tuned to lower frequencies (*BTF* up to 60 Hz: $\Delta R = 43\%$) in comparison to neurons tuned to higher frequencies (*BTF* > 60 Hz: $\Delta R = 31\%$; $p < 0.001$ by Wilcoxon rank-sum test) (Figure 3e). In this manner, sound-driven inhibition appeared to be both more prevalent and more prominent in neurons tuned to lower frequencies, indicating a possible tactile frequency-dependence in multisensory interactions.

Auditory stimulation could induce a behavioral response in the mouse, such as active whisking, which could in turn influence the responses of somatosensory neurons to passive whisker stimulation. To test the possibility that the observed sound-driven inhibition (Figure 3) can be attributed to an indirect pathway involving active whisking, we transected the buccal branch of the facial nerve in a subset ($n = 2$) of mice, thus severely diminishing the ability of the mouse to whisk[23]. We imaged whisker-responsive S2 neurons in these mice before and after cutting the nerve (Figure S2). Dramatic sound-driven inhibition of tactile responses was still observed (Figure S2b-c), and the proportion of neurons exhibiting sound-driven inhibition was preserved (58% before transection and 54% after transection, $p > 0.5$ by Chi-square proportion test; Figure S2d-e). Thus, active whisking is not crucial to the observed multimodal interactions.

Next, we examined the possibility that these interactions may be generic across sensory modalities. For a subset ($n = 3$) of mice, we examined the effect of a wide range of auditory stimuli on tactile responses and compared these changes to the effects of visual stimulation on tactile

responses (Figure S3). We found that the effect of sound was similar across the different sounds tested, with similar effects regardless of the specific sound (tone versus noise) or modulation frequency. Roughly 75% of neurons responded consistently; however, light produced quantitatively different effects, with roughly 50% of neurons showing a different pattern of modulation (Figure S3e). Thus, sound-driven inhibition of tactile responses appears to be a modality-specific modulation.

Finally, we examined in more detail whether sound-driven inhibition of tactile responses may depend on the specific properties of the sound stimulus used. Along with attaining the tactile frequency tuning curves of neurons in the absence and presence of a concurrent sound stimulus (Figure S4a and S4d, format similar to Figure 3-4), we measured responses to SAM tones in the absence and presence of tactile stimuli while varying either the carrier frequency (Figure S4b-c) or modulation frequency (Figure S4e-f) of the sound. Responses of sound-selective neurons in S2 were strongest for carrier frequencies between 6 and 24 kHz (Figure S4b-c, blue curves), while touch responses in tactile-selective neurons were selectively inhibited by those same carrier frequencies (orange curves). Meanwhile, the population of S2 sound-selective neurons, while individually dependent on modulation frequency (33 of 63 neurons), collectively exhibited only a small increase in response as modulation frequency increased (Figure S4e-f, blue curves). Touch responses in tactile-selective neurons were individually agnostic to modulation frequency, with only 20 of 259 neurons showing any dependence (Figure S4e-f, orange curves). Yet, as a population, a weak trend was observed, with tactile-selective neurons more strongly inhibited by higher sound-modulation frequencies, which was weakly but significantly anti-correlated to the responses of sound-selective neurons (Figure S4f, bottom row). These results indicate that the

observed sound-driven inhibition of tactile responses in S2 is proportional to the total sound-driven response and is not a peculiarity of the precise frequencies that compose the auditory stimulus.

Concurrent Auditory Input Modulates Tactile Responses in S1 and ISF

Primary sensory cortices are believed to be predominantly involved in unimodal processing, but the degree of multimodal interactions has long been debated. As a wealth of studies have illustrated potential multisensory interactions in the primary cortices[24–26], we examined whether tactile responses in mouse S1 are modulated by concurrent auditory stimuli. Figure 4B shows fluorescence recordings for three exemplar tactile-selective neurons within a single FOV in response to tactile stimuli alone or concurrent tactile and auditory stimuli. Neuron 1 exhibited minimal change to its tactile tuning relation following addition of concurrent sound (left, Figure 4b-c). By contrast, tactile responses of both neurons 2 and 3 were inhibited by sound (middle and right, Figure 4b-c), with neuron 3 exhibiting stronger inhibition. Notably, all three neurons were non-responsive to acoustic stimuli when presented alone (blue traces, Figure 4b). At the population level, unlike neurons in S2, the tactile responses for a majority of highpass (69%) and bandpass (75%) neurons in S1 were unaltered by auditory stimuli (48% and 28% for S2 highpass and bandpass neurons, respectively; $p < 1e\text{-}10$ by Chi-square proportion test for both categories when compared with S1). Nonetheless, we identified a substantial number of tactile neurons exhibiting sound-driven inhibition and a small number with sound-driven facilitation. The overall proportions of neurons exhibiting sound-modulation of tactile responses is summarized in Figure 4d. Like sound-modulated neurons in S2, the degree of inhibition appeared to be stronger for neurons with lower *BTF*s (Figure 4e) ($p < 0.01$ by Wilcoxon rank-sum test), while the degree of sound-driven facilitation appeared to be similar between the two types of tactile neurons ($p > 0.1$). These results

clearly show the presence of multisensory modulation of neurons in the primary somatosensory cortex.

We also measured response properties of neurons in the ISF, a poorly understood cortical region found lateral to S2. Here, we observed robust responses to tactile stimulation (Figure S5). In contrast to S1 and S2, we found a dearth of bandpass neurons, with 18-fold more highpass neurons (Figure S5d). The percentage of tactile neurons suppressed by auditory stimulation resembled that of S2 (Figure S5e). The presence of similar modes of tactile-auditory interactions across S1, S2, and ISF suggests that multimodal integration occurs across a distributed somatosensory cortical network.

Sound-Driven Inhibition of Tactile Neurons Preferentially Alters Low Frequency Responses

Given the prevalence of sound-driven inhibition of tactile responses observed for neurons in the somatosensory cortex, we next sought to dissect spectral features of multimodal processing. We scrutinized tactile tuning relations in the presence and absence of concurrent auditory stimuli for four neurons with varying *BTF* within a single FOV in S2. Neuron 1, tuned to tactile frequencies between 8 to 32 Hz, was inhibited across all tactile frequencies (Figure 5a-b). Interestingly, for neurons 2 and 3, sound-driven inhibition was nearly complete at low tactile frequencies (below 32 Hz) while the high-frequency tactile responses were largely spared. Similar effects were observed for neuron 4, which was responsive to high tactile frequencies, suggesting that sound-driven modulation is exquisitely dependent on the frequency of tactile stimulation: responses to low-frequency tactile stimuli are preferentially suppressed while responses to high-frequency tactile stimuli are relatively preserved.

To quantify the tactile spectral properties of sound-driven inhibition, we assigned neurons to five groups according to their best tactile frequencies (Figure 5c). Normalized tactile tuning

curves (Figure 5c) in the absence (black) and presence (red) of concurrent sound stimuli were averaged for the five categories. To quantify the strength of sound-driven inhibition, we computed the ratio of normalized tactile responses in the presence of sound stimuli to those in the absence of sound stimuli at each individual tactile frequency (r_{sound}). Across all groups, r_{sound} increased monotonically with increasing tactile stimulus frequency. Low frequency tactile responses were strongly diminished ($r_{sound} \sim 0.1$) by sound stimuli while high frequency tactile responses were more weakly perturbed ($r_{sound} \sim 0.8$) (Figure 5d). The tactile frequency for half-inhibition ($TF_{1/2}$) increased only weakly as a function of the *BTF* for the given class of neurons (Figure 5g). Overall, these results show that sound-driven inhibition depends upon tactile stimulus frequency with low frequency tactile stimuli preferentially suppressed.

This spectral dependence of sound-modulation of tactile responses has two distinct functional consequences. First, the preferential sound-driven inhibition at low tactile frequencies results in a rightward shift in the overall tactile tuning relation for individual neurons. This effect manifests as a positive change in *BTF* for a majority of neurons (Figure 5e). Second, the overall inhibition of tactile responses is higher for neurons with low *BTF* (Figure 5f).

Dissociation of Sound-Driven Inhibition from Motor-Associated Modulation of Tactile Responses

Ongoing locomotor activity can drastically alter activity in sensory cortices[27,28]. To determine whether the observed sound-driven inhibition in S2 may be associated with differences in motor activity and not directly caused by sound stimulation (Figure S6a), we recorded the animal's movement velocity on a treadmill during two-photon imaging of neural activity and used a generalized linear model (GLM) to compare the influence of sound stimulation with motor activity on tactile responses.

Tactile responses in exemplar neurons were divided into three categories based on movement velocity, two categories based on sound stimulation (off or on), and four categories of tactile stimulation. Responses stratified in this manner are shown for two exemplar neurons (Figure S6b-c, top). For neuron 1 (Figure S6b), responses are not appreciably altered by locomotion (compare rows), indicating that sound-driven inhibition of tactile responses cannot be explained by ongoing behavior. For neuron 2 (Figure S6c), running alone diminished responses; however, sound also inhibited this neuron's tactile responses. Using a GLM to quantify the contribution of sound (coefficient *c*, z-scored) and locomotion (coefficient *d*, z-scored) to each neuron's tactile responses (see Methods), we find that neuron 1 was at best weakly inhibited by running velocity ($d = -0.3$) while neuron 2 was strongly inhibited ($d = -2.9$) (Figure S6d). Importantly, in both cases sound is a strong inhibiting factor ($c = -3.4$ and -3.5, respectively).

Across the population of tactile-responsive neurons in S2, those previously identified as unperturbed by sound (Figure 3d) were found to have sound amplitudes near zero in the GLM (Figure S6e-f, left) while those identified as inhibited by sound (Figure 3d) tended to have large negative values for the sound coefficient (Figure S6e-f, right). Interestingly, for sound-inhibited tactile-responsive neurons, neurons tuned to lower tactile frequencies were inhibited by concurrent running behavior while neurons tuned to higher tactile frequencies were facilitated by running (blue curve, Figure S6g). Across all *BTF*s, however, the coefficient for sound modulation was consistently of a larger amplitude (red curve, Figure S6g).

Sound-Selective Neurons in S2 but not Auditory Cortex Are Suppressed by Touch

A fraction of neurons in S2 that were unresponsive to tactile stimuli were instead responsive to auditory stimuli (1.2% of neurons in S2, 161 neurons across 67 FOVs); for reference, 0.3% of neurons in S1 were sound-selective (12 neurons across 16 FOVs), which was significantly less

than in S2 (p < 1e-5 by Chi-square proportion test). We aimed to determine if these auditory responses were influenced by simultaneous tactile stimulation. Shown in Figure 6a-c are the responses of an exemplar sound-selective neuron located in S2. This neuron did not respond to tactile stimuli at any frequency (Figure 6b, black traces) but did respond to sound alone (Figure 6b, blue trace). When the same sound was paired with different tactile frequencies, the sound response was almost completely abolished at the highest tactile frequencies (Figure 6b, red traces). The normalized tuning curves clearly demonstrate the tactile frequency-dependent suppression of sound responses in this neuron (Figure 6c). This trend was observed for the population-averaged tuning curve of sound-selective neurons in S2, with suppression occurring for tactile frequencies above 32 Hz (Figure 6d). Note that analysis was limited not just to sound-selective neurons (132 neurons in S2) but to neurons selective to the specific auditory stimulus used in the combined stimuli (82 neurons in S2). The suppression was quantified by the response ratio, defined as $(r_{hi}-r_{lo})/(r_{hi}+r_{lo})$, where r_{hi} is the average response to sounds paired with high frequency (76-128 Hz) tactile stimulation and r_{lo} is the average response to sounds paired with low frequency (2-8 Hz) tactile stimulation. The distribution of these values for sound-selective S2 neurons (Figure 6e) shows that sound responses in most neurons were decreased by high frequency tactile stimuli. Of the 82 sound-selective neurons, 31 neurons from 19 FOVs across 7 mice met criteria for significant suppression of responses by high-frequency tactile stimulation.

To ascertain whether this effect was specific to S2, we measured the effect of tactile stimulation on the responses of sound-selective neurons in auditory cortex. Shown is an exemplar neuron that responds to sound alone but not to touch (Figure 6g-h). The addition of tactile stimulation at any frequency did not change the response of this neuron to sound (Figure 6g-h). Across the population of sound-responsive neurons in auditory cortex, neurons exhibited minimal

tactile-frequency dependent changes in their sound response (Figure 6i-j). Of 132 neurons that responded to the auditory stimulus used in the combined stimulus (out of a total of 267 neurons responsive to any auditory stimulus), only 11 met criteria for significant suppression, which was a significantly lower fraction than in S2 (p < 1e-6, Chi-square proportion test). Thus this tactile suppression of auditory responses in S2 is not unique but is much stronger than in auditory cortex.

Spatial Layout of Functional Cell Types in S2

Given the inverse relationship of multisensory integration between sound- and touch-selective neurons and the spectral tuning of this interaction, we posit a mutually inhibitory local circuit within S2 that serves to regulate the strength of tactile-auditory inhibition (Figure 7a). Under this model, sound-selective neurons inhibit tactile responses of touch-selective neurons across the spectrum of whisker deflection frequencies. At low frequencies, this inhibition is nearly total, but at high frequencies the stronger activation of touch-selective neurons as a population and their reciprocal inhibition of the sound-selective population of neurons is enough to tilt the balance towards tactile responses.

We asked whether S2 exhibited any spatial clustering of neurons based on their functional properties, which could provide indirect evidence for a mutually inhibitory circuit. This patterning should reflect the preference of cortical neurons to connect to their nearest neighbors within 100-200 μm[29]. In particular, sound-selective neurons that were not modulated by tactile stimuli should be located further away from the C2-whisker-responsive region of S2 than their touch-modulated counterparts. To facilitate comparison of FOVs across mice, we used the local (~100 μm radius) soma-excluded neuropil signal at each pixel in the image to form a map of tactile response strength (Figure 7b-d). From this map, we could identify the center of the S2 whisker-responsive region (white cross, Figure 7d), allowing us to register all S2 neurons across all FOVs (Figure 7e).

Neurons are color-coded according to whether they are touch-selective (responsive only to touch; red circles) or sound-selective (responsive only to sound; green circles). Neurons responsive to both touch alone and sound alone (n = 43) were excluded from this analysis. We found that, for this population, touch-selective neurons were more clustered (assessed by distance to the center of the S2 whisker-responsive region) than the average neuron in the field (red and black curves, Figure 7f). Sound-selective neurons that were not inhibited by touch exhibited a similar distribution to all neurons (green and black curves, top, Figure 7f), indicating that they were equally distributed across a typical FOV and perhaps even partially excluded from the whisker-responsive center of S2. However, when we examined the distribution of sound-selective neurons that showed significant suppression of responses by tactile stimuli, as in Figure 6b-c, the distribution of distance from the S2 core now closely mirrored that of touch-selective neurons (blue and red curves, bottom, Figure 7f). These results indicate that both touch-selective neurons and touch-inhibited sound-selective neurons are embedded together in a specific region of S2, hinting at a local circuit linking these two functional subtypes.

DISCUSSION

In this work, we asked how information from somatosensory and auditory inputs is integrated in the mouse neocortex. With two-photon Ca^{2+} imaging, we investigated large populations of layer 2/3 neurons across somatosensory and auditory areas with single cell resolution. We found that neurons across somatosensory cortices are tuned to the frequency of tactile stimulation. The addition of concurrent sound resulted in modulation of these tactile responses in both S1 and S2, and this modulation typically manifested as a suppression of the response. Moreover, the degree of suppression depended on tactile frequency, with responses to low frequencies more inhibited than responses to high frequencies. We also identified a population of neurons in S2 responsive to sound but not to touch. Unlike in auditory cortex, sound responses of many (31 of 82) sound-selective neurons in S2 were strongly inhibited by addition of tactile stimuli at high tactile frequencies. These neurons were spatially colocalized with S2 touch-selective neurons.

Tactile frequency tuning across somatosensory cortices. The detection of the frequency of mechanical vibrations is important for animals to discern surface texture and to handle tools[30,31]. Rodents can use rapid motion of their whiskers to distinguish smooth surfaces from rough ones[20,32–34], and neurons in S1 encode information that correlates with texture judgment[35]. As the amplitude and frequency of vibrissal oscillations depend on the roughness or smoothness of the contacted surface[36], tuning to spectral frequency in the somatosensory system can encode texture information. Indeed, previous studies have shown that neurons in rat S1[37,38] and S2[39] exhibit bandpass tuning properties. In our study, while highpass neurons outnumbered bandpass neurons in both S1 and S2 (Figures 2d and 2g), the presence of well-tuned neurons in both regions supports the notion that tactile frequency tuning may be a general organizational feature for mouse tactile

sensation, and the proportion of bandpass neurons may be an underestimate due to the limit in the upper tactile frequencies tested.

The higher proportion of neurons with bandpass tuning properties in S2 than in S1 may reflect differences in thalamocortical inputs to the two regions. S1 receives strong thalamic drive from the ventral posterior medial nucleus (VPM), while S2 receives a larger share of its thalamocortical input from the posterior medial nucleus (POm)[40,41]. Interestingly, although both POm and VPM cells show adaptation, causing decreased response amplitude under high frequency stimulation, POm cells exhibit earlier adaptation than VPM cells[42] and as a result are tuned to lower frequencies than VPM cells. Thus, the higher prevalence of neurons with bandpass tuning properties in S2 may be inherited from the response properties of thalamic neurons, although it could also reflect longer temporal integration windows in higher areas of cortex[43].

Sound-driven modulation of tactile responses in somatosensory cortices. We found that the addition of an auditory stimulus modulated tactile responses in both S1 and S2, consistent with the sound-driven hyperpolarizing currents previously observed in mouse S1[44]. This modulation has three notable features: 1) Although a similar proportion of neurons in both S1 and S2 were facilitated by sound, more neurons in S2 were inhibited than in S1 (Figure 3d, 4d). 2) Inhibition of neurons tuned to low tactile frequencies in both S1 and S2 was more severe than inhibition of neurons to high tactile frequencies in the same regions (Figure 3e, 4e). 3) Sound-driven suppression in S2 is tactile frequency dependent, with stronger inhibition occurring at lower tactile frequencies (Figure 5).

Previous studies in human and non-human primates have revealed that multimodal integration improves detection of events in the environment[3,45,46]. The optimal integration of competing sensory cues involves dominance of the most reliable sensory cue to minimize variance

in the estimation of the true stimulus[3]. This evaluation of reliability between different sensory cues is a dynamic process, with the weight or value of each stimulus modality being continuously updated[46]. Low frequency tactile stimulation is potentially less salient of a signal than high frequency tactile stimulation, since it comprises lower velocity whisker motions. Indeed, we observed more suppression of tactile responses at lower tactile stimulus frequencies than at high frequencies (Figure 5), indicating that auditory responses are more dominant when tactile stimuli are weak. This result is consistent with the prior observation that, during optimal multimodal integration, the more reliable stimulus modality dominates the response. On the other hand, this frequency-dependent integration is complementary to "inverse effectiveness," where multimodal integration is largest for weak multimodal stimuli near threshold and decreases with increasing stimulus intensity, as has been reported in the superior colliculus[47,48].

Where does this modulation arise? Previously, it was believed that multimodal influences on activities within classically defined unimodal areas are mediated by feedback from multisensory integration in higher-order cortical regions[49,50]. However, human studies using event-related potentials (ERPs) suggest that these multimodal influences may also be carried in the feedforward inputs coming from subcortical regions to unimodal regions[49,51,52]. In the present study, we identified a small (1.2%) population of sound-selective neurons within S2 itself that may be performing multimodal integration. Although prior studies have shown non-matching neurons in primary cortices that respond solely to other sensory modality inputs[53], the sound-selective neurons we found may play a special computational role in multimodal integration. The sound-driven responses in these neurons were strongly suppressed at high tactile frequencies (Figure 6a-e). In contrast, neurons in auditory cortex showed much weaker suppression by high frequency tactile stimuli (Figure 6f-j). The detection of these rare sound-selective neurons in S2 was aided

by the use of two-photon Ca^{2+} imaging, which allowed us to survey hundreds of neurons within a large FOV. We found that sound-selective neurons are not clustered in S2 (Figure 7), but touch-inhibited sound neurons are closer to the center of the whisker-responsive region of S2, similar to the spatial organization of non-matching neurons seen in other studies[52,53]. The existence of touch-inhibited sound-selective neurons in S2 indicates that they may play a role in local sound-driven suppression observed in tactile-selective neurons of S2.

Hypothetical circuit model underlying the sound-driven modulation. Unlike sound-selective neurons of the auditory cortex, auditory responses of sound-selective neurons in S2 are often robustly inhibited at high tactile frequencies (>32 Hz). This threshold tactile frequency (32 Hz) matches the inflection point of sound-driven suppression of tactile-selective neurons in S2 (Figure 5). Based on this observation, we hypothesize that local sound-selective neurons in S2 suppress responses of tactile-selective neurons during combined stimulus presentation. In a simplified circuit model (Figure 7a), tactile-selective neurons receive tactile inputs, while sound-selective neurons receive auditory inputs. Meanwhile, sound-selective neurons inhibit tactile-selective neurons, while tactile-selective neurons reciprocally inhibit sound-selective neurons with the strongest inhibition at high tactile frequencies. This winner-take-all circuit could dynamically select a stimulus modality at each moment and, under the right conditions, would be consistent with divisive normalization, a model that has been proposed as a driving force behind multisensory interactions[54–56].

MATERIALS AND METHODS

Animal Surgery and General Procedures

All animal procedures were approved by the Johns Hopkins Institutional Animal Care and Use Committee. Imaging experiments used transgenic mice expressing the genetically encoded Ca^{2+} indicator GCaMP6s[57] under a pan-neuronal promoter. Most experiments used GP4.12 Thy1-GCaMP6s mice (JAX No. 025776)[12]. A small subset of tetO-GCaMP6s mice (JAX No. 024742) were also used for imaging[14]. Finally, lox-GCaMP6s mice (JAX No. 024106)[13] were crossed with Syn1-Cre mice (JAX No. 003966)[58], resulting in lox-GCaMP6s-Syn1-Cre mice. In total, 18 males and 5 females were imaged for this study. Among them, 17 were GP4.12 Thy1-GCaMP6s mice, 3 were tetO-GCaMP6s mice, and 3 were lox-GCaMP6s-Syn1-Cre mice.

For surgery to implant headposts and chronic imaging windows, general anesthesia was provided by 1-1.5% isoflurane in 0.5 L/min O_2. Body temperature was monitored through a rectal probe and maintained at 37°C with a heating pad. Lidocaine (20 mg/ml) was injected at the incision site for local anesthesia while carprofen (5 mg/kg, i.m.) and dexamethasone (5 mg/kg, i.m.) were given for pain and inflammation. Normal saline (0.5 ml, i.p.) was given to avoid dehydration. Skin and muscle over the left hemisphere was removed to expose skull overlying auditory and somatosensory cortices and a custom headpost was attached to the skull using UV-cured primer (Kerr OptiBond) and dental cement (Heraeus Charisma A1). A 4-5 mm craniotomy spanning both auditory and somatosensory cortices was performed. After removing the skull, the craniotomy was filled with a transparent silicone gel (3-4680, Dow Corning, Midland, MI)[59] to help maintain the stability and clarity of the chronic window. Next a 4 or 5 mm (No. 0 or 1) cover glass was secured with dental cement above the window. Mice were given buprenorphine (0.1 to 0.5 mg/kg, s.c.) for

pain and allowed to recover from anesthesia on a warmed pad. After surgery, mice were periodically given buprenorphine for pain along with carprofen and dexamethasone. Mice were given 10-14 days to recover before imaging sessions were started. Location of the C2 barrel column in S1 was estimated by stereotaxic measurement of its expected location relative to lambda (1.3 mm posterior and 3.5 mm lateral). This location matched the most medial of the three regions responsive to tactile stimulation as seen with functional widefield imaging (Figure 1).

In a subset of experiments, to examine whether sound-driven inhibition was dependent on active whisking, the facial nerve was transected. The buccal branch of the facial nerve innervates the mystacial pad musculature[60], which in turn controls whisker movement[61]. General anesthesia was achieved with isoflurane (1.5% in 0.5 L/min O_2) and body temperature was maintained at 37°C. Local anesthetic (lidocaine, 20 mg/ml) was injected at the incision site, which was 1-2 mm caudal of the E1 whisker. The surrounding tissue was dissected to reveal the buccal branch of the facial nerve at the junction of the ramus superior buccolabialis and ramus inferior buccolabialis[23,62]. Micro scissors were used to transect the nerve just anterior to the junction point and proximal to entering the mystacial pad. Skin closure was achieved with nonabsorbable suture (4-0 Prolene) and Vetbond (3M). Mice were allowed 2-3 days to recover. The absence of active whisking ipsilateral to the site of the transection was used to assess the success of the surgery. Imaging was performed as described below until active whisking returned, typically 2 weeks after the initial cut.

Experimental Setup

During imaging sessions, all mice were awake and head-rotated approximately 45° to make the imaging plane perpendicular to the microscope objective. Mice were first adapted to running on a custom-built treadmill while head-fixed for 1-2 session of 20 to 40 minutes each. Once habituated,

imaging sessions began. Mice were able to tolerate head fixation for up to 2 or 3 hours. Running speed was recorded with an optical encoder (E5-2500-188-IE-S-H-D-B, US Digital). For whisker stimulation, all whiskers of the right whisker pad were trimmed except for the C2 whisker, facilitating isolated stimulation of a single whisker.

Imaging

For widefield imaging, a white light source (LED Engin LZ1-10CW00) was used for illumination. Illumination light and fluorescence signals were filtered through a GFP filter cube (460/50 excitation, 540/50 emission). A 4× 0.13 NA objective (Olympus) and a Photometrics Evolve 512 Delta camera were used to collect the emitted light. The field of view was roughly 5.5×5.5 mm^2. The frame rate was 20 Hz.

Two-photon Ca^{2+} imaging was performed with an Ultima system (Prairie Technologies) built on an Olympus BX61W1 microscope. A mode-locked laser (Coherent Chameleon XR Ti:Sapphire) tuned to 950 nm was raster scanned at 5 Hz for excitation while emitted GCaMP6s fluorescence was collected through a green filter (525/70 nm). Laser power at the sample was 20-80 mW. Dwell time was set to 2 µs. To increase imaging speed, resolution along the *y*-axis was reduced by a factor of 4. The final pixel size was 0.9x3.6 µm. A 20× 1.0 NA objective (Olympus) was used to yield a 465×465 μm^2 field of view. Imaging depths were between 150-350 µm.

Tactile Stimuli

To stimulate the target whisker, a piezoelectric bending actuator (Q220-A4-203YB, Piezo systems, Inc., MA) moving in the rostral-caudal direction was attached to a 2 cm sewing needle and the target whisker was passed through the eye of the needle at a distance of 2 mm from the face. The needle coupled the whisker to the piezo movement. The shape of the eye of the needle allowed the whisker some freedom to move in the medial-lateral direction but not in the rostral-caudal

direction. Positioning the needle close to the base of the whisker reduced flexion of the whisker during stimulation, thus minimizing non-linear mechanical transformations of the sinusoidal driving signal. To facilitate precision in the positioning of the needle, we attached the piezo to a micro-manipulator (MT-XY Compact Dovetail Linear Stage, Newport Corporation). The piezo itself was driven by a voltage driver (MDT694B, Thorlabs) with a 10× gain. Thus, the 0-10 V signal delivered by the DAQ card was expanded to a range of 0-100 V to drive the piezo.

A single tactile stimulus consisted of a 500 ms sinusoidal wave. One tactile trial consisted of a set of sixteen individual stimuli with onsets spaced by 2 s. Frequency of each sinusoidal stimulus was randomized between 2 to 128 Hz, spaced logarithmically. The amplitude of piezo deflection was 0.6 mm. With the needle 2 mm away from skin surface, a 2 Hz stimulus was equivalent to a mean angular speed of 66.8^{o}/sec, calculated as $\tan^{-1}(0.6 \text{ mm} / 2 \text{ mm}) \times 2 \text{ Hz} \times 2$ (one forward and one backward deflection per cycle). Thus, mean angular speed of our tactile stimuli is calculated from frequency as 33.4^{o}/sec × (frequency), where frequency is given in Hz. In this work, we report the stimulus parameters in frequency. The piezo was calibrated monthly by imaging piezo movement to measure maximum deflection displacement and angular velocity. Once angular velocity dropped by more than 20%, the piezo was replaced.

An important consideration in this work was whether the delivery of tactile stimulation with the piezo generated any audible noise within the hearing range of mice. We performed a noise calibration while driving the piezo at frequencies ranging from 2 to 256 Hz. Sound was recorded with a Sokolich probe placed directly next to the piezo. We found that no noise audible to mice (across the mouse hearing range of 3-100 kHz) was generated as long as the frequency of piezo deflection was kept lower than 140 Hz. Above 140 Hz, though, audible sounds were generated (Figure S1d-e). Thus, all stimuli used in this study were limited to a maximum of 128 Hz. In

support of this assertion, only 2 out of 2036 neurons recorded in auditory cortex passed criteria for a response to stimulation by the piezo alone. To test the opposite, whether acoustic stimuli could cause any vibration of the piezo, we imaged the piezo through the microscope objective at a frame rate of 363 Hz while playing sounds at the maximum amplitude possible with our speakers. We found no detectable movement of the piezo under imaging.

Auditory Stimuli

The imaging set-up was located in a sound-attenuated room (Acoustical Solutions, AudioSeal ABSC-25) with noisy equipment placed outside the room. Noise calibrations showing ambient noise to be outside the mouse hearing range have been previously published[15]. Sound stimuli were delivered through a free-field speaker (LCY K100, Ying Tai Corporation) located 15 cm away from the right ear.

The main stimuli used were sinusoidal amplitude modulated (SAM) tones, which were composed of a pure tone (the carrier frequency) with a low frequency amplitude modulation. To search for sound responses, SAM tones were delivered over a wide range of carrier frequencies (3 to 96 kHz, 16 frequencies) or amplitude modulation frequencies (2 to 256 Hz, 8 frequencies). The duration of each stimulus was 500 ms (Figure S1b). We tested a variety of additional sounds in a subset of experiments, including pure tones and broadband noise (Figure S3). Pure tones were gated by 5 ms squared cosine ramps. White noise was bandlimited between 3 to 96 kHz and, if indicated, was amplitude modulated by a sinusoidal envelope (SAM Noise). Neurons were categorized as sound-responsive if they responded to any of the auditory stimuli tested. All the stimuli were delivered in a random order and with a 2 s interval between onsets. Stimuli were played at -40 to -20 dB attenuation, corresponding to ~60-80 dB SPL using our speaker. Sound

levels were chosen by looking for reliable responses in auditory cortical areas as visualized by widefield fluorescence imaging.

Combined Stimuli

Combined stimuli consisted of simultaneous presentation of the auditory and tactile stimuli described above. The tactile stimuli were sixteen tactile deflections with different frequencies ranging from 2-128 Hz while a fixed auditory stimulus was presented concurrently with each of the sixteen tactile stimuli. For the auditory stimuli, SAM tones were used for every FOV, with the particular carrier and modulation frequencies chosen such that the stimulus had elicited a robust response in auditory regions during widefield fluorescence imaging (typically between 6-12 kHz for the carrier frequency and 8-64 Hz for the modulation frequency). As the choice of sound could have a bearing on the observed multisensory effects in somatosensory cortex, we also tested pure tones and broadband noise (with or without amplitude modulation) in some experiments (Figure S3). Each stimulus lasted 500 ms and the tactile and auditory stimuli were co-initiated and co-terminated (Figure S1c). Stimulus order was randomized with a 2 s interval between the onset of each presentation.

Stimulus Order

In general, we interleaved blocks of each stimulus condition, for example one block of tactile alone (consisting of 16 piezo frequencies) followed by one block of tactile + auditory (16 piezo frequencies in combination with one fixed sound stimulus). Within each block, the stimuli (such as the 16 piezo frequencies) were always randomized, stimuli lasted 500 ms, and intervals were 2 s. The blocks were interleaved in a semi-random order as chosen by the experimenter, with the goal being to spread each stimulus block across the duration of the imaging session to minimize the effects of habituation and bleaching. At least 4 repeats were attained for each condition,

although typically we aimed for 5 to 10 repeats. Whenever responses are shown, all individual trials are included.

Widefield Imaging Analysis

For widefield imaging, a structured sparse encoding algorithm was used to detect the baseline fluorescence F_0[15,63,64], from which $\Delta F/F_0 = (F-F_0)/F_0$ was calculated for each pixel. To facilitate comparison of changes across different regions of cortex, plotted time series data (Figure 1b-c, 1e-f) were zeroed at the start of each stimulus.

Two-Photon Imaging Analysis

We first performed a frame-by-frame image registration to correct for drift in our imaging field[65]. Assisted by a structured sparse coding algorithm[63], we then performed a semiautomatic segmentation of our movies acquired from a two-photon imaging session to acquire regions-of-interest (ROIs) corresponding to individual neurons. From each ROI, we generated brightness-over-times (BOTs). This process included neuropil subtraction, where we subtracted the signal from pixels within 75 μm of each cell (but excluding pixels containing other cell bodies). Using a low-pass filter that excluded large positive deflections (putative transients), a smooth baseline (F_0) was fit to each BOT, allowing for calculation of a normalized fluorescence signal $\Delta F/F_0$. To increase the temporal accuracy of our Ca^{2+} signal and to decrease the influence of noise, we employed a non-negative deconvolution method with a time constant of 0.7 s to generate a correlate of spike probability[66], which was then thresholded to estimate events per time bin, which loosely corresponds to the frequency of spikes or firing rates for a given neuron[67–72]. Events per time bin was used for constructing tuning curves and other measures as described in the section of "Response Analysis." Parameters for choosing neurons in the segmentation algorithm were purposefully set to be sensitive so that no responsive neurons would be missed; conversely, many

ROIs correspond to relatively inactive neurons or false detections. These ROIs would not reach significance for being responsive and thus do not alter any of the conclusions or analysis in this work other than the potential for inflating the number of nonresponsive neurons.

Registration

For most mice, the imaging window was initially mapped under two-photon imaging by tiling FOVs across the entire visible window and at multiple depths. Two-photon imaging fields were registered to widefield imaging fields by using vasculature landmarks. Information about the location of auditory (AI, AII, and AAF) and somatosensory cortices (S1, S2, and ISF) obtained from widefield imaging was then used to identify the cortical identity of each two-photon imaging field.

Response Analysis

Response analysis includes five major parts: 1) event rate normalization; 2) quality measure of responsiveness; 3) best frequency extraction; 4) construction of tuning curves; 5) calculation of sound-driven response change.

For event rate normalization, we used a 600 ms period after the start of a stimulus as our response window. The averaged baseline-corrected signal during that period yielded the event rate or response for each individual stimulus.

We used three indices to measure the responsiveness of a neuron: 1) the total area under the tuning curve; 2) the *p*-value of ANOVA to quantify whether there is a significant difference between the responses across different stimulus frequencies, which indicates a neuron is tuned to stimulus frequency; 3) the ratio of the averaged response during the response window divided by the average response outside the response window, a measure of driven rate. Each quantity was passed through a function with a saturating nonlinearity and summed to yield a unitless measure

of responsiveness. This ad hoc measure was thresholded to categorize neurons as responsive or not responsive to a particular stimulus type. We chose this approach because any single measure was inadequate in capturing whether neurons were responsive. For example, ANOVA compared responses across different frequencies but not to the baseline firing rate. Also, it does not account for magnitude differences, so a quiet neuron with incomplete neuropil subtraction may show a significant *p*-value under ANOVA even though the total area under the tuning curve is small. Individual neurons were scrutinized to test for the adequacy of this approach and the threshold chosen to best match our manual characterization of neurons. Also, adjustments in the threshold did not appreciably alter the results reported in this work. For further analysis, if the neuron responded only to tactile stimuli, we classified it as a tactile-selective neuron. Similarly, if the neuron only responded to any of the sound stimuli we tested, we classified it as a sound-selective neuron. Using a shuffle test applied to neurons from our population (100 shuffles of the times of stimulus presentation), only 1 in 4832 iterations passed criteria for touch selectivity and 1 in 14980 for sound selectivity. We did find that sound-selective S2 neurons, while not passing criteria individually for being responsive to tactile stimuli, did on average show a small level of activity in response to tactile stimuli alone (black curve, Figure 6d). This response was weak (~5× smaller than the sound response) and inconsistent, as confirmed by manual inspection of the 82 individual tuning curves, and thus may reflect weak and inconsistent tactile-related activity in these neurons.

A measure of best tactile frequency (*BTF*) was used to separate bandpass and highpass neurons. We used the average of two different approaches to measure best frequency: 1) the weighted average of all significant responses; 2) the frequency that yielded the maximal response. Neurons with best frequencies less than or equal to 60 Hz are categorized as bandpass neurons while neurons with best frequencies greater than 60 Hz are categorized as highpass neurons.

To construct tuning curves, we used two different strategies. For characterizing tuning properties of tactile-selective neurons in S1 and S2 (Figure 2), we normalized the response for different tactile frequencies to the maximum response among all the tactile frequencies. For characterizing the sound modulation effect of response to tactile stimuli (Figures 3-6, S2, and S4), we normalized the response to the maximum response of either tactile stimuli or tactile plus auditory stimuli. And, for experiments where we used a third modality (visual light stimulation in Figure S3), we normalized the response to the maximum response across the three stimuli (tactile stimuli, tactile stimuli with concurrent sound stimuli, and tactile stimuli with concurrent visual stimuli).

To quantify the sound-driven response change of tactile stimuli (Figures 3-5, S2, and S4), we calculated the ratio of response change (ΔR) as $\Delta R = \sum(R_{T+S} - R_T)/\sum(R_{T+S} + R_T)$, where R_T is the response to tactile stimuli alone, R_{T+S} the response to combined tactile plus auditory stimuli, and the summations are over all tactile frequencies tested (2 to 128 Hz).

For the registration of S2 imaging fields shown in Figure 7, we first formed smoothed neuropil maps by taking a region centered on each pixel and calculating a cell-excluded average of the signal within 100 μm of that pixel. This neuropil BOT signal was processed similarly to the neuronal BOT signal as described above, with the only difference being that neuropil subtraction was not performed on the neuropil signal. The amplitude of the whisker-induced response as quantified by summing the neuropil $\Delta F/F_0$ over a 600 ms window for each stimulus was used to generate an image of tactile response strength. The peak of this map was nominally labeled as the center of the S2 whisker region and used to register FOVs across experiments.

The generalized linear model (GLM) was used to model the response of individual neurons on a single trial basis. The output was taken as the mean response over a 600 ms window for each

stimulus, as described above, and then normalized so that the maximum response at any time is equal to one. Dependent variables consisted of tactile stimulus frequency (binned into four separate frequency channels), the presence of sound, and the running velocity. Indicator functions were used for stimulus frequency and the presence of sound, while running velocity was passed through a logarithmic function: $\log_2(6*v+1)/10$, such that the range would fall between 0 to 1 (v never went higher than 170.5 cm/s in our data). Velocity was calculated based on the ball diameter of 6.75 cm and 2500 counts per rotation on the optical encoder. A velocity of 50 cm/s resulted in a value of ~0.8234 fed into the GLM, while a velocity of 0 cm/s gave a value of 0. This transformation function was chosen as it yielded the best model fits for velocity modulated neurons in our population. The logit function was used as the link function, and the MATLAB function glmfit.m was used to perform the fitting. Resulting coefficients were then z-scored by dividing by their associated standard errors. Trials were included for tactile stimulation alone, tactile stimulation with SAM tones, SAM tones alone, and, if performed, "silent" trials with no tactile or auditory stimuli.

ACKNOWLEDGEMENTS

We thank David T. Yue for conceiving of this project and for providing the scientific foundation for this work. We also wish to thank Eric D. Young and Dwight E. Bergles for comments on the manuscript, Benjamin D. Haeffele for experimental set-up and guidance on image processing, Gordon F. Tomaselli for guidance and support, Jeffrey M. Yau and Sascha du Lac for valuable discussions, Ingie Hong for assistance with assembly of the head-fixed awake preparation, Travis Babola for assistance with computer hardware, Terry Shelley for fabrication of stereotaxic equipment, Loren Looger and the GENIE project for supplying transgenic GCaMP6s mice, and members of the O'Connor lab and the Calcium Signals Lab for helpful discussions. This work was supported by grants from the Kleberg Foundation, the Whitehall Foundation, the Klingenstein Fund, and the NIH (NIMH R01MH065531, NINDS R01NS073874, NINDS P30NS050274, NINDS R01NS089652).

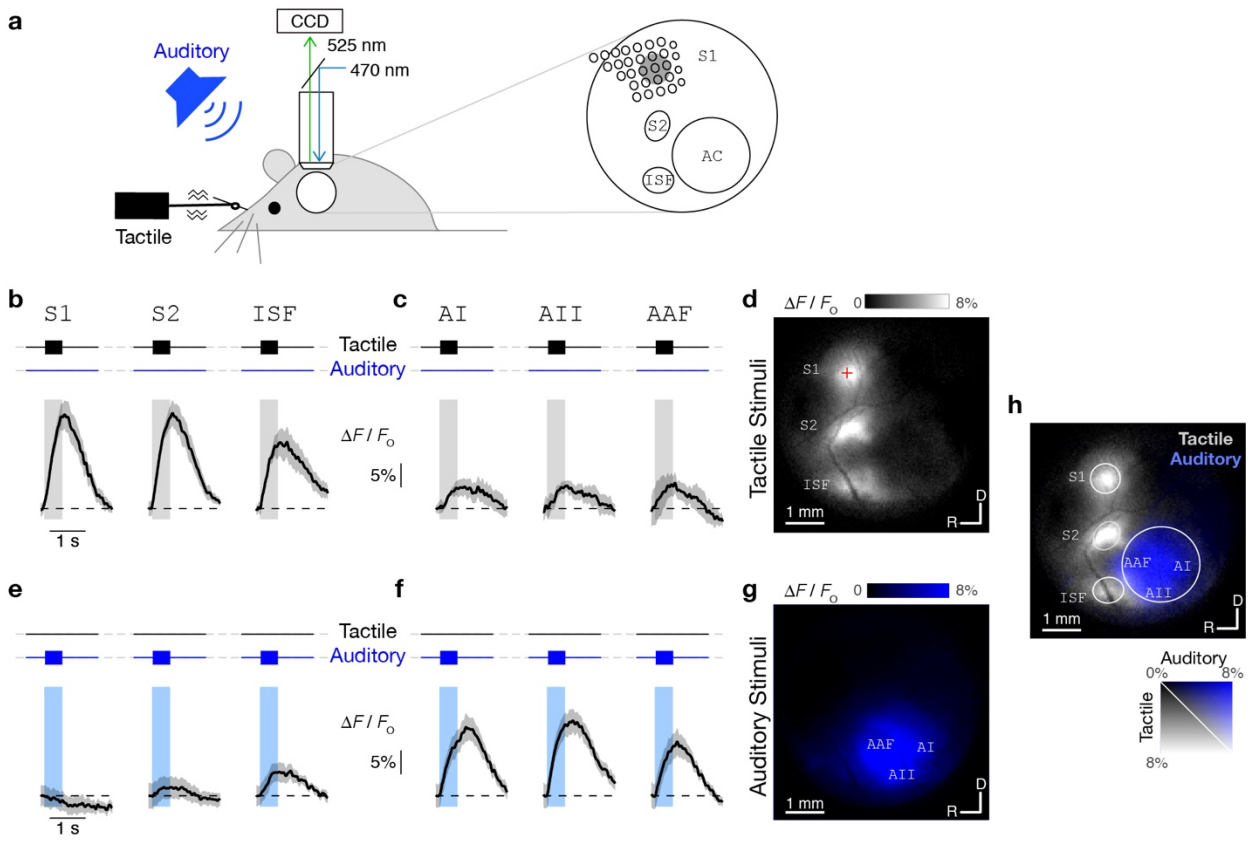


Figure 1. Responses to Tactile and Auditory Stimuli Under Widefield Ca^{2+} Imaging

(a) Widefield imaging set-up. Tactile and auditory stimuli are delivered to a head-fixed awake mouse. Piezo coupled to right C2 whisker provides vibration in rostral-caudal direction while speaker delivers sinusoidal amplitude-modulated (SAM) tones to the right ear. 5 mm chronic cranial window over the left temporal cortex (white circle) spans somatosensory and auditory areas. GCaMP6s fluorescence (470 nm illumination and 525 nm emission) collected by CCD camera through 4× objective.

(b) Responses to tactile stimuli in somatosensory areas. Average of four single-trial Ca^{2+} transients evoked in S1, S2, and ISF in response to 128 Hz whisker stimulation. Black trace shows fluorescence activity ($\Delta F/F_0$) averaged over 200×200 μm^2 regions in each of these three areas.

Standard error across four trials shown as shaded region. Vertical gray bar represents 500 ms stimulus period. Black squares represent presentation of tactile stimuli.

(c) Responses to tactile stimuli in auditory areas. Same format as **b**.

(d) Ca^{2+} activity during presentation of tactile stimuli reveals location of somatosensory areas. Spatial map depicts average response to 36-128 Hz tactile stimuli (7 individual frequencies), taken as average of $\Delta F/F_0$ signal over first 1 s after stimulus onset. Responsive regions correspond to S1, S2, and ISF, as labeled. Red cross denotes 1.3 mm posterior and 3.5 mm lateral of bregma. Dorsal-rostral orientation and 1 mm scale bar shown at bottom.

(e) Responses to 9.5 kHz SAM tone stimuli in S1, S2 and ISF averaged over six trials. Vertical blue bar represents 500 ms stimulus period. Blue squares represent presentation of SAM tones.

(f) Responses to auditory stimuli in auditory areas. Same format as **e**.

(g) Ca^{2+} activity during presentation of auditory stimuli reveals location of auditory areas. Image shows average response of 3-48 kHz SAM tone stimuli (7 individual frequencies). Responsive regions correspond to AI (primary auditory cortex), AII (secondary auditory cortex), and AAF (anterior auditory field), as labeled.

(h) Overlay of responses to tactile (gray, as shown in D) and auditory (blue, as shown in G) stimuli highlighting relative location of somatosensory and auditory fields in the same window. Rostral-caudal orientation and $\Delta F/F_0$ scale bar shown at bottom right.

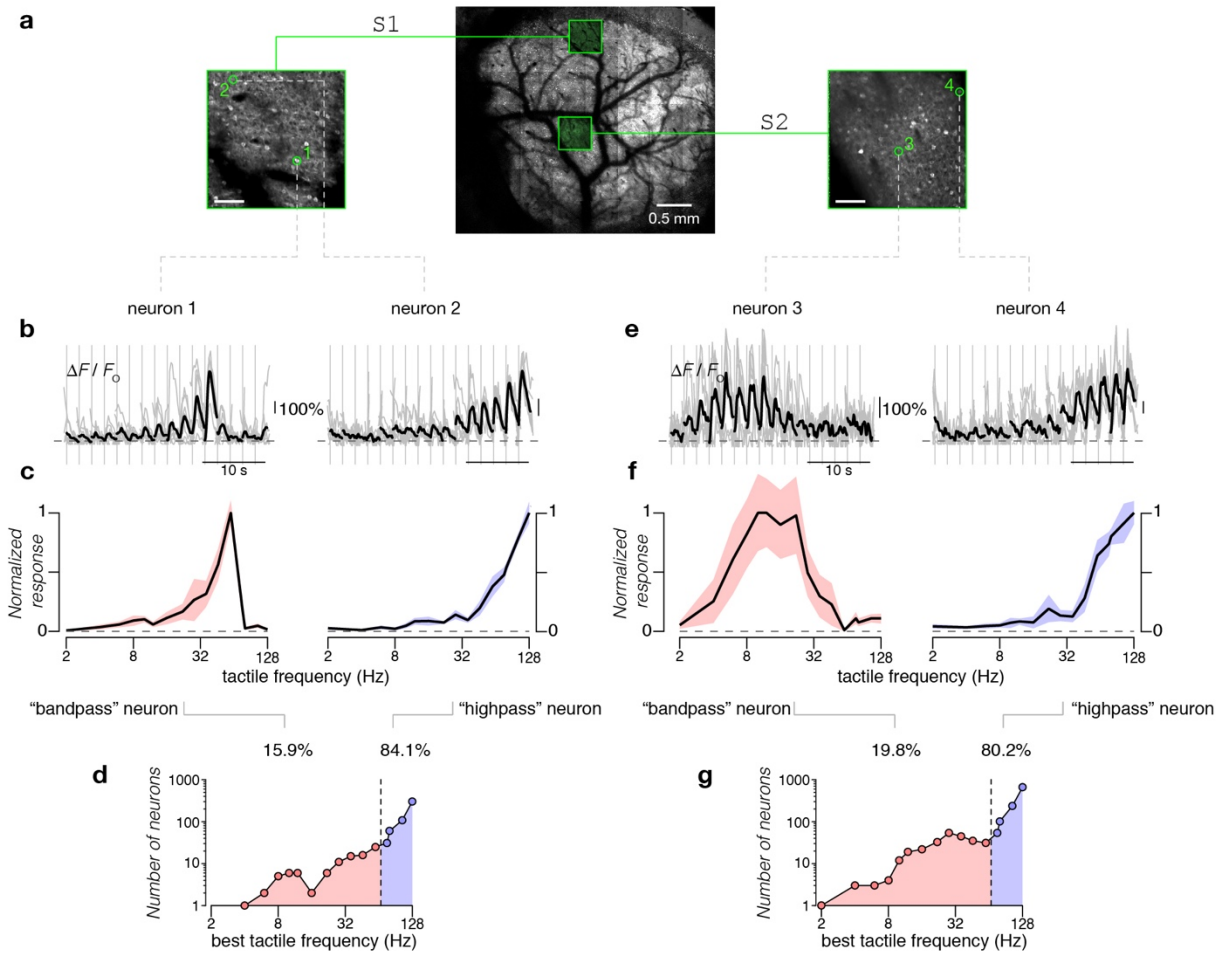


Figure 2. Responses to Tactile Stimuli in Somatosensory Cortices Under Two-Photon Ca^{2+} Imaging

(a) Map highlights location of FOVs centered on C2 barrel within S1 (left) and S2 (right) among the entire cranial window (middle, formed by tiling 7×8 individual FOVs). Scale bar on individual FOVs: 100 µm.

(b) Ca^{2+} activity ($\Delta F/F_0$) of two exemplar S1 neurons during presentation of tactile stimuli at different frequencies. Stimuli were played in random order and responses sorted by increasing frequency (2-128 Hz). Gray traces show $\Delta F/F_0$ for individual trials. Black traces show average $\Delta F/F_0$ across all trials (7 repeats). Vertical gray line denotes start of each stimulus. These exemplars depict presumptive bandpass (neuron 1) and highpass (neuron 2) neurons observed.

(c) Frequency tuning curves of individual neurons shown in **b**. Responses are averages over a 600 ms period of the deconvolved $\Delta F/F_0$ traces, which are then normalized to the maximum response across all stimuli for each neuron. Standard error shown as shaded region. Frequency of tactile stimuli shown on the *x*-axis.

(d) Population distribution of *BTF* (see Methods) for S1 neurons. Black circles show number of neurons tuned to each frequency tested. To facilitate comparison with S2, neurons with best frequencies no higher than 60 Hz were categorized as bandpass neurons (pink) while those with best frequencies above 60 Hz were categorized as highpass neurons (blue). Both *x* and *y* axes plotted on a logarithmic scale.

(e) Ca^{2+} activity induced by tactile stimuli in S2 for two exemplar neurons, again showing bandpass (left) or highpass (right) behavior. Average is across 6 trials.

(f) Frequency tuning curves for the exemplar neurons shown in **e**.

(g) Population distribution of *BTF* in S2.

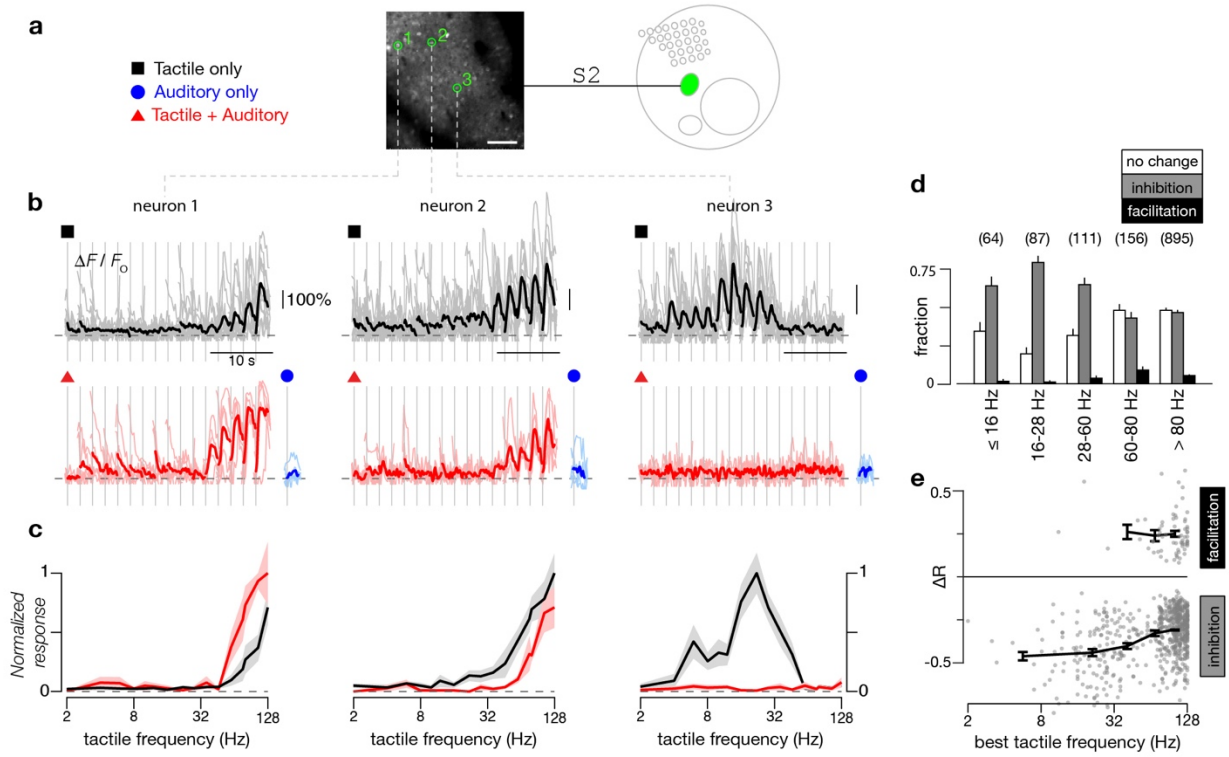


Figure 3. Sound-Modulated Response to Tactile Stimuli in S2

(a) Baseline fluorescence image of two-photon imaging field (middle) located in S2, as illustrated in schematic map (right). Symbols (left) denote different stimuli tested. Exemplar neurons to be scrutinized in panels **b-c** are highlighted by green circles (middle panel). Scale bar: 100 µm.

(b) Responses of exemplar neurons to tactile stimuli alone (black traces, top panel, 11 repeats) or to combined tactile and auditory stimuli (red traces, bottom panel, 5 repeats). In both cases, tactile frequency ranged from 2 to 128 Hz. For the combination stimulus, a concurrent auditory stimulus (10 kHz tones with 64 Hz SAM envelope at 20 dB attenuation) was added to the tactile stimuli. In addition, blue traces (bottom right) show averaged responses to the same auditory stimulus alone. For all cases, individual trials shown as light traces and response averages shown as dark traces.

(c) Frequency tuning curves of the normalized response to tactile stimuli and combined tactile plus auditory stimuli for exemplar neurons shown in **b**. Black line shows the normalized tuning curve for responses to tactile stimuli. Red line shows the normalized tuning curve for responses to tactile plus auditory stimuli. Standard error shown as shaded gray and pink regions.

(d) Fraction of neurons in S2 exhibiting either no change, a decrease (inhibition), or an increase (facilitation) in responses when SAM tones were added to the tactile stimuli, with error bars indicating standard error for a multinomial distribution. Criteria determined by ANOVA ($p < 0.05$). Neurons binned into one of five categories based on their *BTF*; number of neurons in each category indicated by numbers above the bar plots. Results pooled across 67 FOVs from 19 mice.

(e) Percentage change in response of S2 neurons to tactile stimuli (ΔR) when SAM tones were added as a function of *BTF*, shown for inhibited or facilitated neurons as described in **d**. Thick line and error bars show mean and standard error of ΔR among inhibited or facilitated neurons within each category from **d**, and are only shown if at least 3 neurons were found within that category.

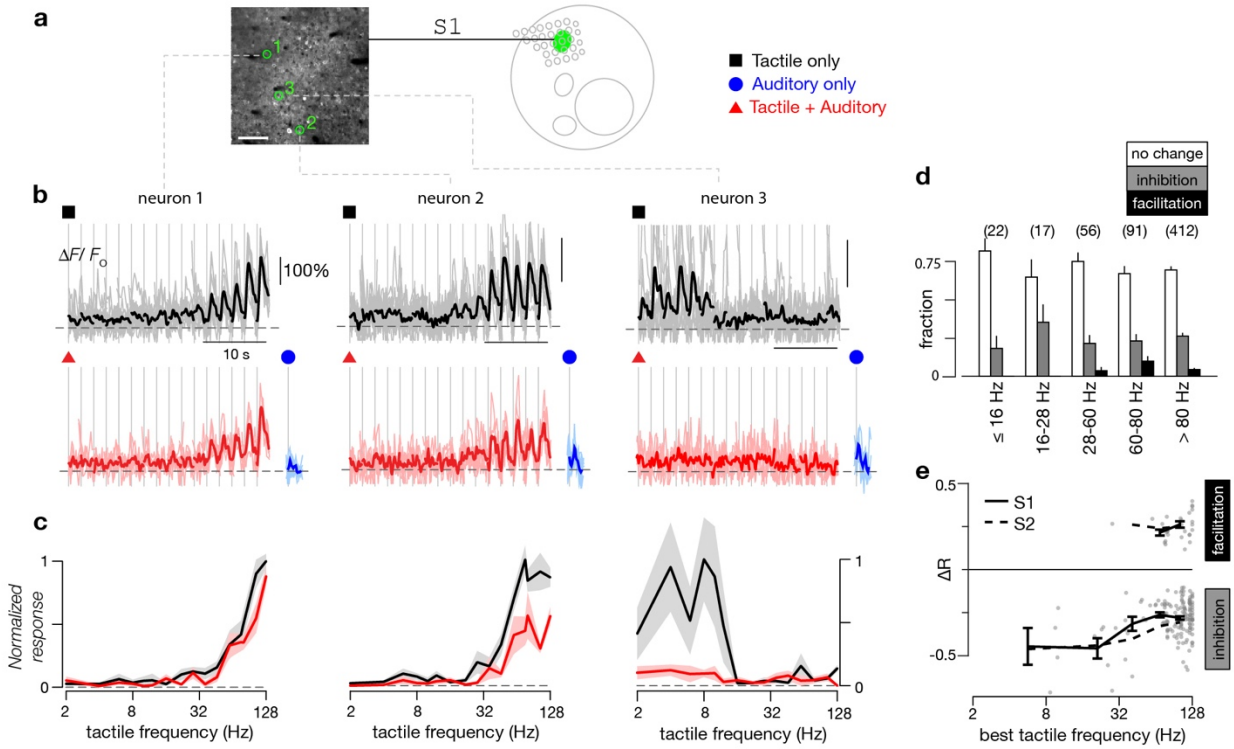


Figure 4. Sound Modulated Response to Tactile Stimuli in S1

(a) Baseline fluorescence image of two-photon imaging field (left) located in S1, as illustrated in schematic map (middle). Symbols (right) denote different stimuli tested. Exemplar neurons to be scrutinized in panels **b-c** are highlighted by green circles (left panel). Scale bar: 100 µm.

(b) Responses of exemplar neurons to tactile stimuli alone (black traces, top panel, 5 repeats) or to combined tactile and auditory stimuli (red traces, bottom panel, 5 repeats). For the combination stimulus, a concurrent auditory stimulus (10 kHz tones with 64 Hz SAM envelope at 30 dB attenuation) was added to the tactile stimuli. Blue traces (bottom right) show averaged responses to auditory stimulus alone. Some individual responses (light gray or red traces) are truncated to fit within the panels.

(c) Frequency tuning curves of the normalized response to tactile stimuli and combined tactile plus auditory stimuli for exemplar neurons shown in **b**.

(d) Fraction of neurons in S1 exhibiting either no change, a decrease (inhibition), or an increase (facilitation) when SAM tones were added to the tactile stimuli, as in Figure 3c. Results pooled across 16 FOVs from 8 mice.

(e) Percentage of change in response of S1 neurons to tactile stimuli (ΔR) when SAM tones were added as a function of *BTF*, shown for inhibited or facilitated neurons as described in **d**. Same format as Figure 3d. Inhibition curve for S2 is reproduced here from Figure 3d as a dashed line.

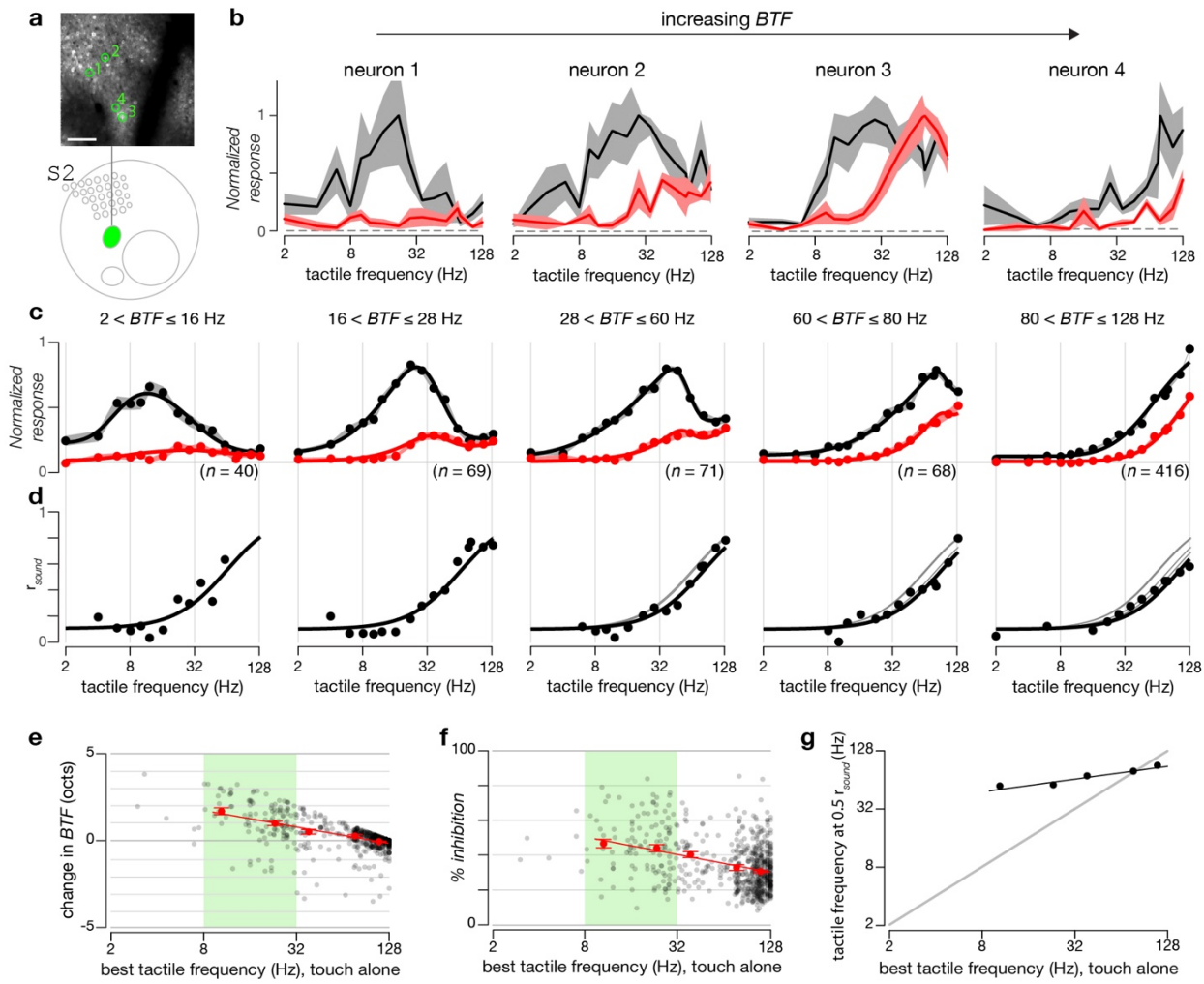


Figure 5. Sound-Driven Inhibition in S2 Depends on Tactile Frequency

(a) Baseline fluorescence image of two-photon imaging field (top) located in S2, as illustrated in schematic map (bottom). Exemplar neurons shown in panel **b** are highlighted by green circles (top panel). Scale bar: 100 µm.

(b) Frequency tuning curves of the normalized response to tactile stimuli (black traces) and combined tactile plus auditory stimuli (red traces) for exemplar neurons shown in **a**. Standard error across 6 repeats indicated by shaded gray (tactile alone) and pink (tactile plus auditory) regions. Concurrent auditory stimulation consisted of 7.8 kHz tones with 28 Hz SAM envelope at 20 dB attenuation.

(c) Population tuning curves of sound-inhibited tactile neurons in S2 grouped by their *BTF*. Responses to tactile stimuli alone and tactile plus auditory stimuli are shown in black and red, respectively. Black circles indicate average responses at each individual frequency while solid lines show smooth fits to the tuning curves. Shaded regions represent standard error across all neurons within each specific tactile frequency domain during tactile (gray) and combined (pink) stimulation. *BTF* ranges used to categorize neurons are indicated above each tuning curve. Horizontal gray line indicates normalized baseline response of 0.08.

(d) Sound inhibition ratio (r_{sound}) of neurons tuned to different *BTF*s. r_{sound} is calculated by dividing the response to combined stimuli by the response to tactile stimuli at each tactile frequency after subtracting the baseline (horizontal gray lines in panel **c**) from each. Black dots represent r_{sound} at each individual tactile frequency, with frequencies that did not evoke a significant response excluded. Black lines are the fitted curves for r_{sound} across 2-128 Hz. To facilitate comparison, gray lines reproduce the fitted curves of r_{sound} for neurons tuned to lower tactile frequencies.

(e) Change in best tactile frequencies of S2 neurons when sound was added (*y*-axis) as a function of *BTF* for touch alone (*x*-axis). Black dots represent individual neurons while red dots represent average change in *BTF* across all neurons within each *BTF* category as indicated in C. Red line is the best fit to red dots. Standard errors shown in red. Green shading highlights mid-frequency bandpass neurons.

(f) Average percentage of inhibition of sound-inhibited tactile-selective S2 neurons when sound is presented as a function of *BTF*. Black dots represent individual neurons and red dots represent average inhibition for neurons within each *BTF* bin. Red line is the best fit to red dots. Standard errors shown in red. Green shading highlights mid-frequency bandpass neurons.

(g) Tactile frequencies at half-maximum (r_{sound} = 0.5) for neurons tuned to different tactile frequencies.

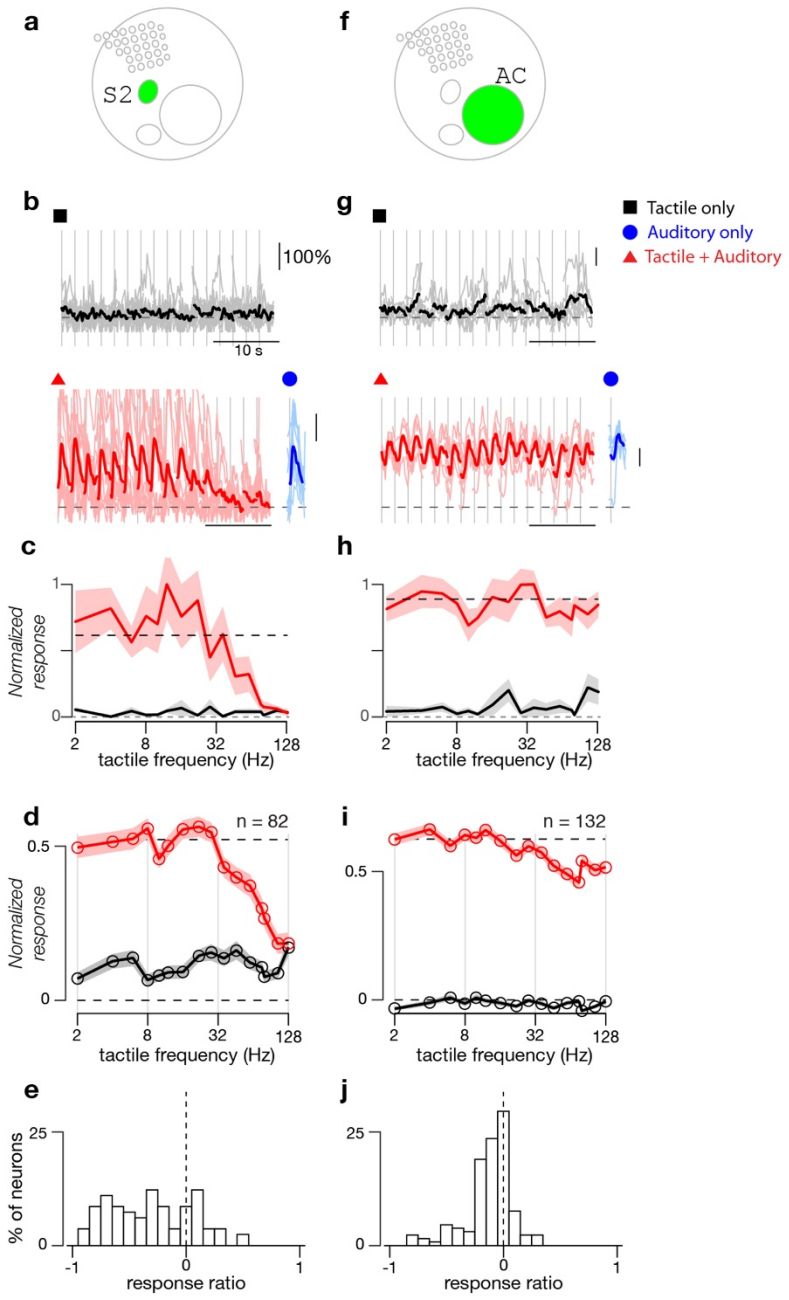


Figure 6. Sound-Selective Neurons in S2 but not in AC Are Modulated by Touch

(a) Schematic map shows location of S2.

(b) Responses of exemplar sound-selective neuron to tactile stimuli alone (black traces, top panel, 8 repeats) or to combined tactile and auditory stimuli (red traces, bottom panel, 11 repeats). For

the combined stimuli, tactile stimuli were paired with an auditory stimulus (6 kHz tone with 10 Hz SAM envelope). Blue traces (bottom right) show average response to 6 kHz SAM tones alone with 10 Hz SAM envelope. For all cases, individual trials shown as light traces and response averages shown as dark traces. Sounds played at 20 dB attenuation.

(c) Frequency tuning curves of the normalized response to tactile stimuli alone (black line) and combined tactile plus auditory stimuli (red line) for exemplar neurons shown in **b**. Standard error shown as shaded gray and red regions.

(d) Averaged frequency tuning curves of the normalized response to tactile stimuli and combined tactile plus auditory stimuli for sound selective neurons in S2 that respond to the auditory stimuli used in the combined stimuli (n = 82, identified across 29 FOVs from 10 mice). Baseline firing rate was additionally subtracted. Standard error shown as shaded gray and red regions.

(e) Distribution of response ratio at high frequency relative to low frequency for sound selective neurons in S2 shown in **c**. Response ratio is calculated as $(r_{lo}-r_{hi})/(r_{lo}+r_{hi})$, where r_{lo} is the response to sounds paired with low frequency (2-8 Hz) tactile stimuli and r_{hi} the response to sounds paired with high frequency (76-128 Hz) tactile stimuli.

(f) Schematic map shows location of auditory cortex.

(g) Exemplar neuron from auditory cortex exhibiting similar responses at all tactile frequencies when auditory stimuli (10 kHz tones with 64 Hz SAM envelope) were added to tactile stimuli. Blue traces (bottom right) show averaged responses to 10 kHz SAM tones alone with 64 Hz SAM envelope. Same format as **b**. Sounds played at 30 dB attenuation. Each stimulus repeated 5 times.

(h) Frequency tuning curves of the normalized response to tactile stimuli and combined tactile plus auditory stimuli for exemplar neuron shown in **g**. Same format as **c**.

(i) Averaged frequency tuning curves of normalized response to tactile stimuli and combined tactile plus auditory stimuli for sound responsive neurons in auditory cortex that respond to the auditory stimuli used in combined stimuli (n = 132). Same format as **d**.

(j) Distribution of response ratio (as in panel **e**) for sound responsive neurons in auditory cortex.

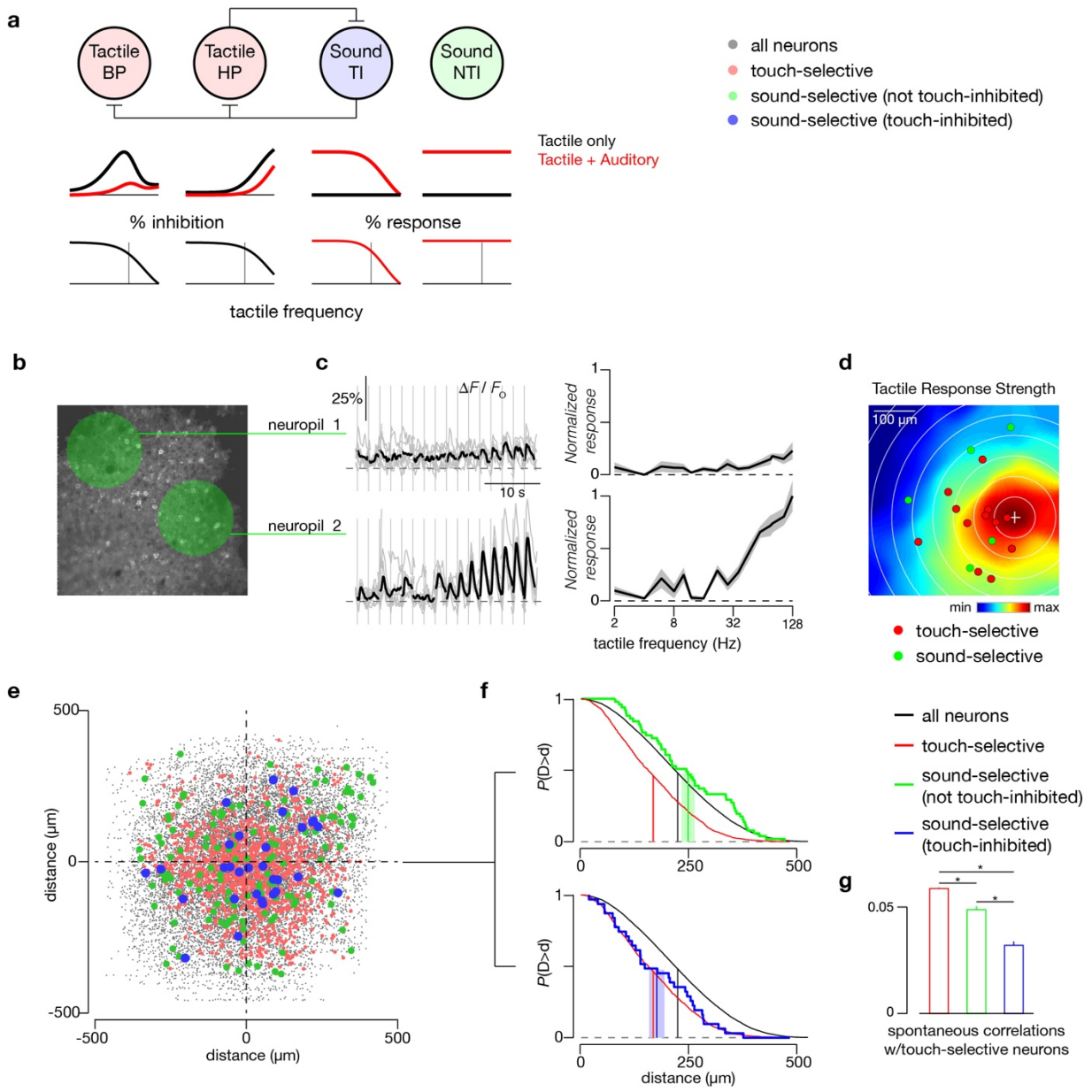


Figure 7: Spatial Layout of Multisensory Neurons in S2

(a) Model of multisensory interactions in S2. Sound-responsive and touch-responsive neurons are connected via reciprocal inhibition (top row). Touch-inhibited sound-responsive neurons ("Sound TI") are connected to touch-responsive neurons while non-touch-inhibited sound-responsive neurons ("Sound NTI") are not. Responses to tactile stimuli (black traces) and combined tactile plus auditory stimuli (red traces) are shown as a function of tactile frequency (middle row). Percent

inhibition of tactile response induced by concurrent auditory stimulation (black traces, left) or of auditory response induced by concurrent tactile stimulation (red traces, right) are shown (bottom row). Vertical gray line represents tactile frequency of 32 Hz.

(b) Baseline fluorescence image of two-photon imaging field in S2. Neuropil regions to be scrutinized in panel **d** are highlighted by green circle overlays. Cell bodies were excluded from the circular mask when calculating the neuropil signal.

(c) Fluorescence responses (left) and frequency tuning curves (right) of the two neuropil regions highlighted in panel **c**. Averages are across 6 repeats.

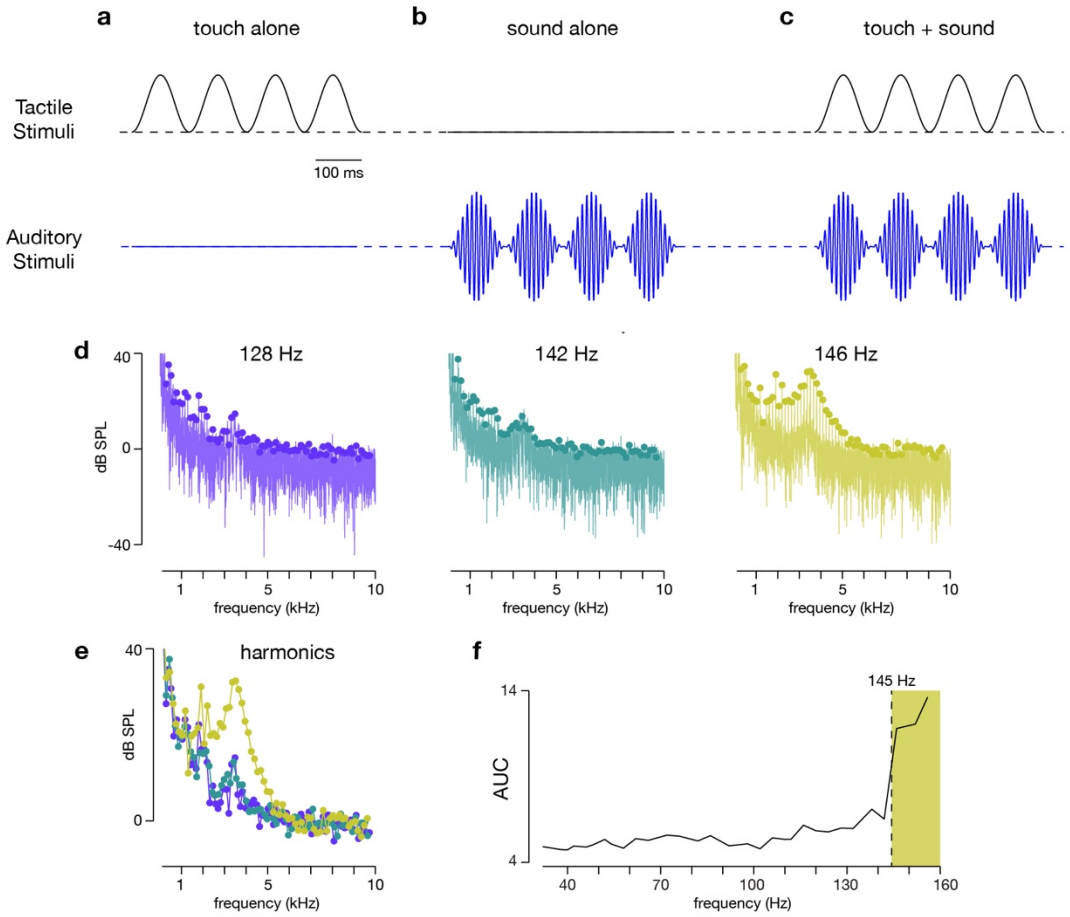
(d) Map of tactile response strength for the FOV in panel **c**. Response strength was calculated by taking the total area under the frequency tuning curve for a neuropil region of interest centered on that pixel. This strength is color-coded according to the color bar (bottom right) and the pixel with the maximum strength is marked by the white cross. Touch-selective neurons (red circles) and sound-selective neurons (green circles) are indicated as well.

(e) Registered map of all S2 neurons. Individual FOVs were centered using the peak of their tactile response strength map (as in panel **e**), allowing all neurons to be plotted on the same axis. All neurons (black dots), touch-selective neurons (red circles), NTI sound-selective neurons (green circles), and TI sound-selective neurons (blue circles) are shown.

(f) Tail distributions of distance from the tactile response strength center for different neuronal types. All neurons (black) and touch-selective neurons (red) are shown on both plots. Green trace shows NTI sound-selective neurons (top) while blue trace shows TI sound-selective neurons (bottom). Vertical lines indicate means while shaded regions represent ±SEM.

(g) Pairwise correlations during periods of spontaneous activity (absence of both tactile and auditory stimulation) between touch-selective neurons and the three groups indicated by color (red:

other touch-selective neurons; green: non-touch-inhibited sound-selective neurons; blue: touch-inhibited sound-selective neurons). These groups contained 19743, 1164, and 748 neuron pairs, respectively. Pairs within 15 μm were excluded. All groups are significantly different from each other (unpaired t-test, * indicates $p < 1e\text{-}6$).



Supplementary Figure 1. Schematic Waveforms for Three Different Stimuli

(a) Tactile stimulation alone, illustrated by a waveform of an 8 Hz sinusoidal tactile stimulus with a 500 ms stimulus period. 100 ms scale bar shown at the bottom right.

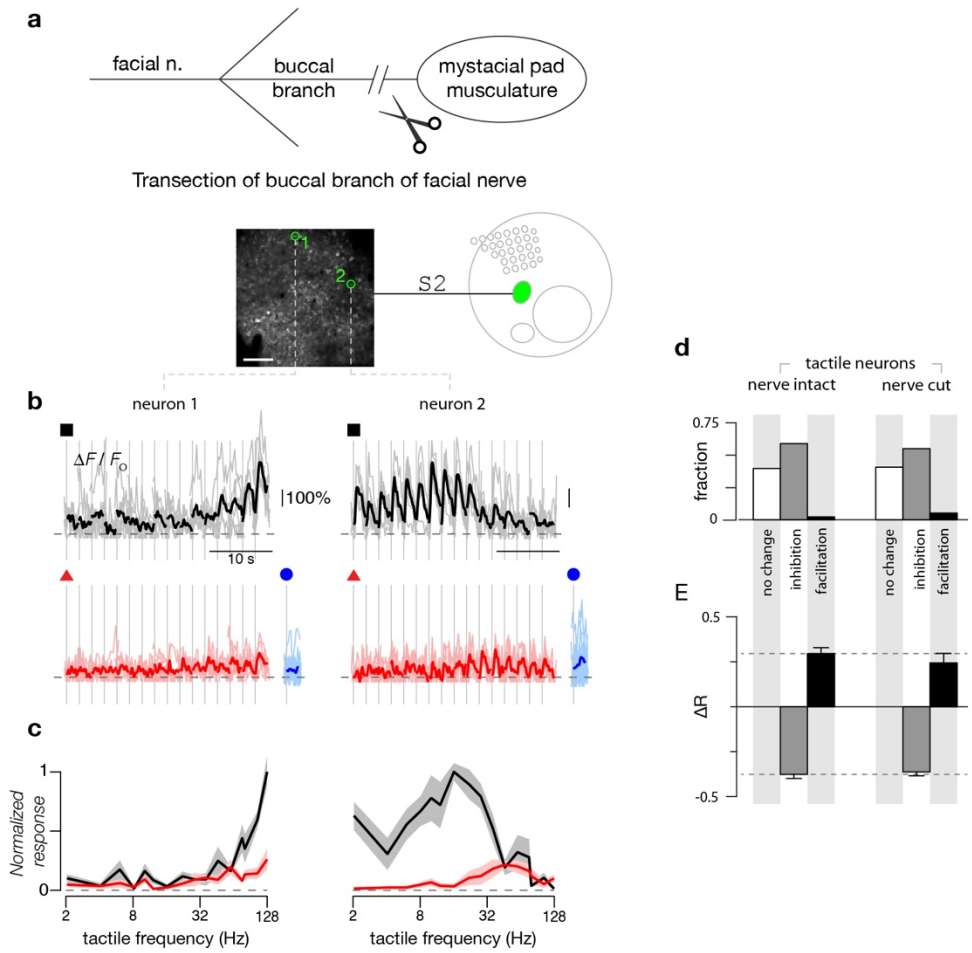
(b) Sound stimulation alone, illustrated by a waveform of a 100 Hz tone with an 8 Hz SAM envelope, again with a 500 ms stimulus period.

(c) Combined auditory-tactile stimulation, with simultaneous presentation of tactile (top) and auditory (bottom) stimuli. Here the 8 Hz frequency of the tactile stimulus is matched with the frequency of the SAM envelope of the sound stimulus.

(d) Spectrum of sound generated by piezo at 128 Hz (left), 142 Hz (middle), and 146 Hz (right) as measured by probe tube microphone. Local maxima highlighted by black and red dots along the spectrum and are observed at integer multiples of the frequency used to drive the piezo.

(e) Values of harmonic peaks are plotted on the same axis. Colors correspond to colors used in **d**.

(f) The total area under the curve (AUC) for harmonics at each piezo frequency tested (32 to 156 Hz). Shaded region illustrates the frequencies at which the piezo generated increased sound amplitudes.



Supplementary Figure 2. Preservation of Tactile-Auditory Interactions in the Absence of Active Whisking

(a) The buccal branch of the facial nerve was transected in a subset (n = 2) of mice to provide temporary (~2 weeks) unilateral abatement of active whisking. Baseline fluorescence image of two-photon imaging field (bottom) located in S2 with exemplar neurons highlighted in green. Scale bar: 100 µm.

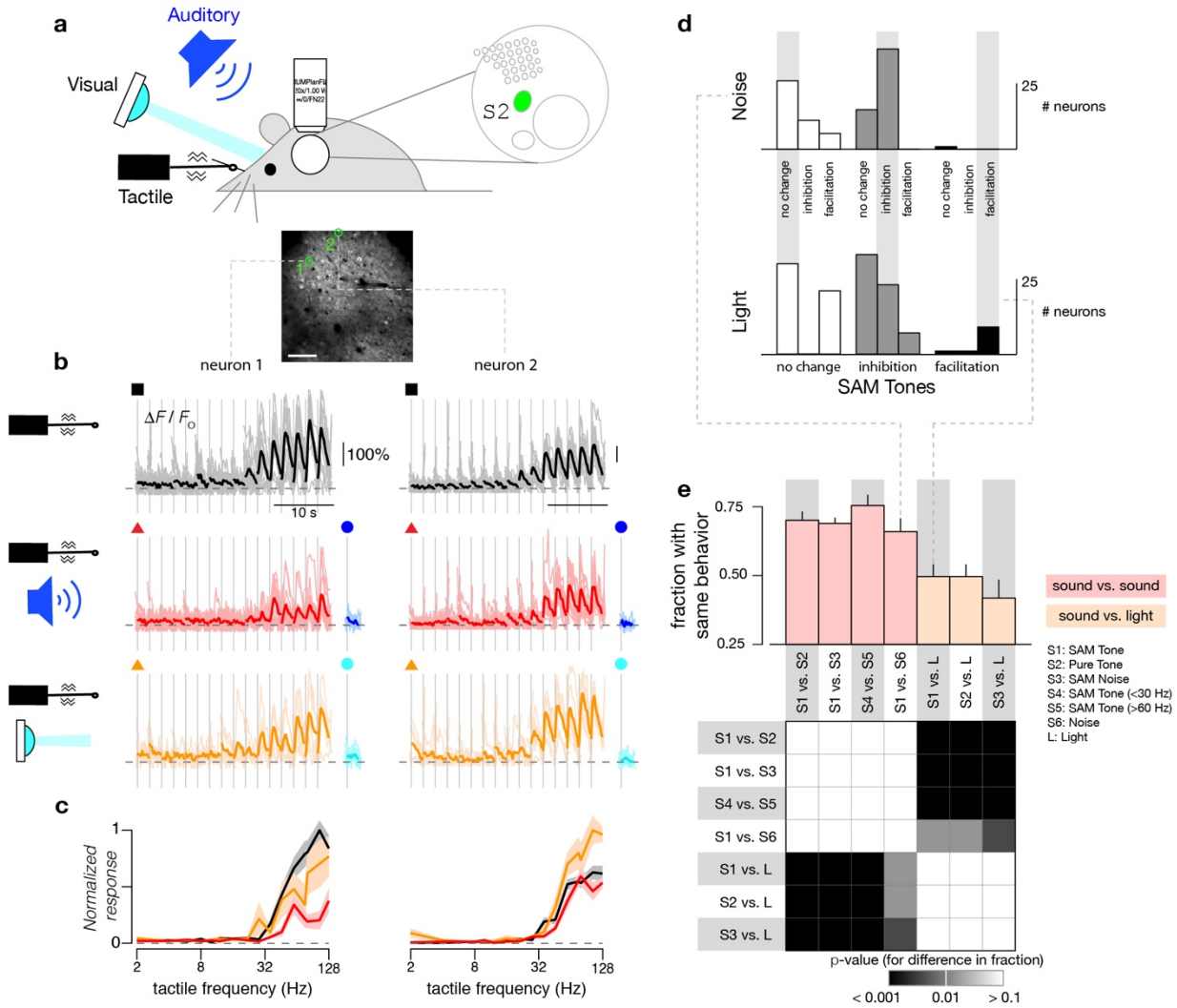
(b) Exemplar neurons observed in S2 after transection of the facial nerve. Neuronal responses to tactile stimuli alone (black traces) were significantly suppressed when auditory stimuli (8 kHz tones with 10 Hz SAM envelope at 20 dB attenuation) were presented at the same time (red traces).

Neither neuron responded to sound stimuli alone (6, 7.5, and 9.5 kHz tones with 10 Hz SAM envelope at 20 dB attenuation; blue traces). Averages are each over 6 repeats.

(c) Frequency tuning curves of the normalized response to tactile stimuli and combined tactile plus auditory stimuli for exemplar neurons in **b**.

(d) Fraction of neurons in S2 exhibiting either no change, a decrease (inhibition), or an increase (facilitation) when SAM tones were added to the tactile stimuli. Left panel shows data for FOVs imaged in S2 before the facial nerve was transected while the right panel shows data for FOVs imaged in S2 after the facial nerve was transected. Criteria determined by ANOVA ($p < 0.05$). Data taken from 7 FOVs across 2 animals (94 tactile responsive neurons) before transection and 6 FOVs across the same 2 animals (119 tactile responsive neurons) after transection.

(e) Percentage of change in response of S2 neurons to tactile stimuli (ΔR) when SAM tones were added. Neurons were divided by nerve intact (left) and nerve cut (right) conditions that showed either a decrease (gray) or increase (black) in their response. Error bars show the standard error of ΔR among inhibited or facilitated neurons.



Supplementary Figure 3. Modality-Specific Multimodal Modulation of Tactile Responses

(a) Tactile, auditory, and visual stimuli were presented either individually or concurrently while two-photon imaging was performed in S2. Baseline fluorescence image of two-photon imaging field (bottom) located in S2 with exemplar neurons highlighted in green. Scale bar: 100 µm.

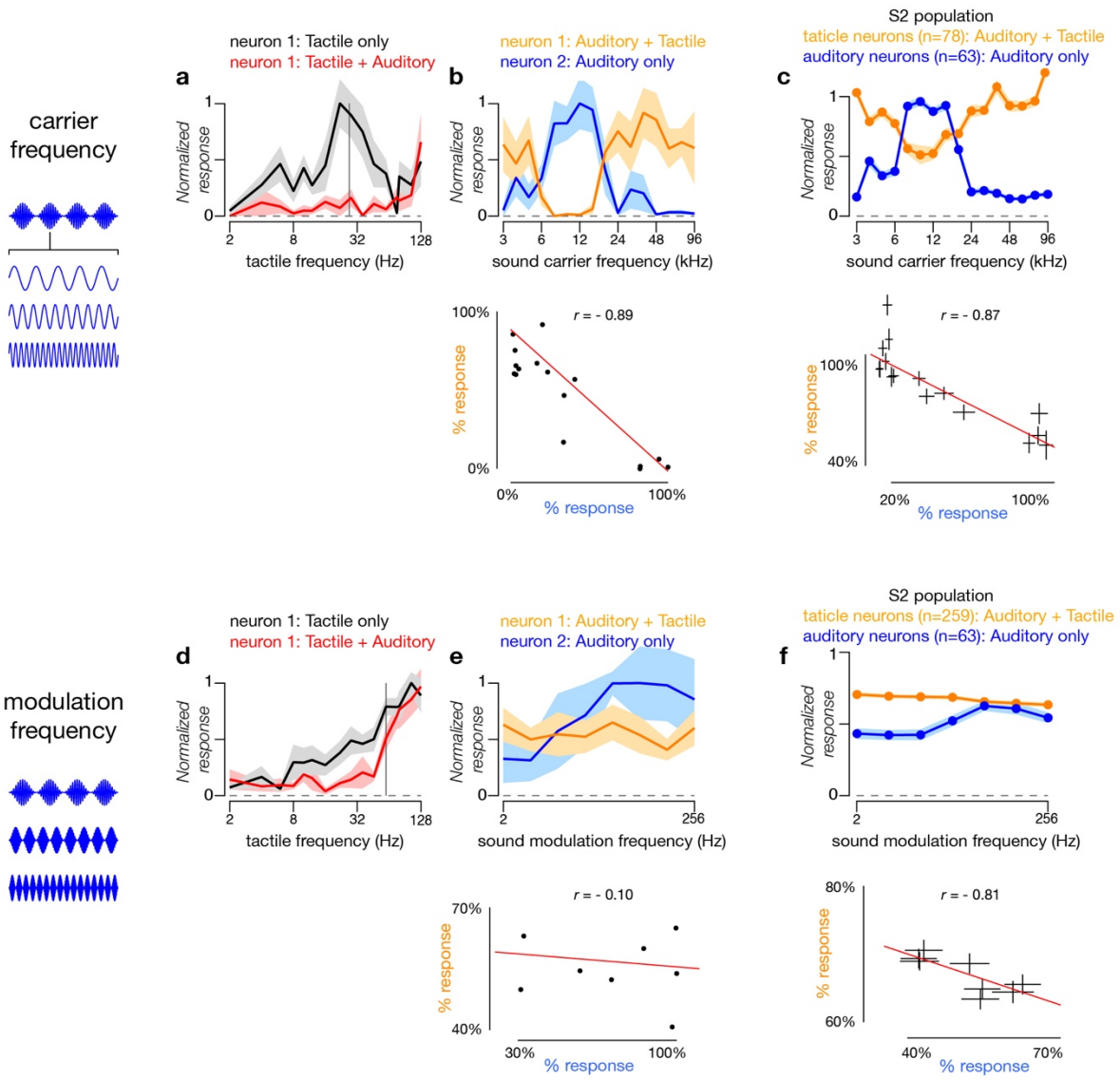
(b) Responses of exemplar neurons to tactile stimuli alone (black traces, top panel, 16 repeats), combined tactile and auditory stimuli (red traces, middle panel, 13 repeats), or to combined tactile and visual stimuli (orange traces, bottom panel, 6 repeats). Blue traces (middle right, 7 repeats) show averaged responses to 10 kHz tones with 64 Hz SAM envelope at 40 dB

attenuation and cyan traces show averaged responses to visual stimuli alone (LED pulsed on for 500 ms). Responses from individual trials are shown as light shaded traces.

(c) Frequency tuning curves of the normalized response to tactile stimuli alone (black trace), combined tactile plus auditory stimuli (red trace), and combined tactile plus visual stimuli (orange trace) for exemplar neurons shown in **b**. While tactile responses in neuron 1 were suppressed by both auditory or visual modalities, responses in neuron 2 were unchanged by sound but facilitated by visual stimuli.

(d) Comparison of the modulatory effects of various stimuli on tactile responses among individual neurons. Histograms show counts of neurons that showed the same or different behavior for the effects of noise vs. SAM tones or for light vs. SAM tones on tactile responses. Neurons are classified into three groups by whether SAM tones caused no change (left, white bars), inhibition (middle, gray bars), or facilitation (right, black bars) in their tactile responses. Neurons are then further subdivided by whether a second stimulus (noise, top; light, bottom) caused no change, inhibition, or facilitation (subdivisions within each column) in their tactile responses. Vertical gray bars highlight categories where neurons showed the same behavior for both conditions.

(e) Fraction of neurons that showed the same behavior (no change, inhibition, or facilitation in tactile response) for two different stimulus types (top). This fraction corresponds to the number of neurons that fall within the thin vertical gray bars in panel **d**. Four comparisons were performed between pairs of sounds (pale red) while three comparisons were performed between light and a sound (pale orange). Bottom plot shows chi-squared test for proportions between each fraction to check for a significant difference. All same-modal comparisons were similar (sound vs. sound compared to sound vs. sound or sound vs. light compared to sound vs. light) while all different-modal comparisons were different (sound vs. sound compared to sound vs. light).



Supplementary Figure 4. Independence of Sound-Driven Inhibition to Sound Properties

Sound-driven inhibition of tactile responses in S2 were not dependent on carrier frequency (a-c) or modulation frequency (d-f) of the SAM tone. All error bars are ±SEM.

(a) Exemplar tactile-responsive neuron inhibited by concurrent sound. Tactile responses (black curve, 10 repeats), which peaked for frequencies around 20-30 Hz, were mostly abolished by presentation of concurrent sound (red curve, 6 repeats, 12 kHz tone with 10 Hz SAM envelope at 20 dB attenuation). Vertical gray line indicates 28 Hz tactile stimulus to be used in **b**.

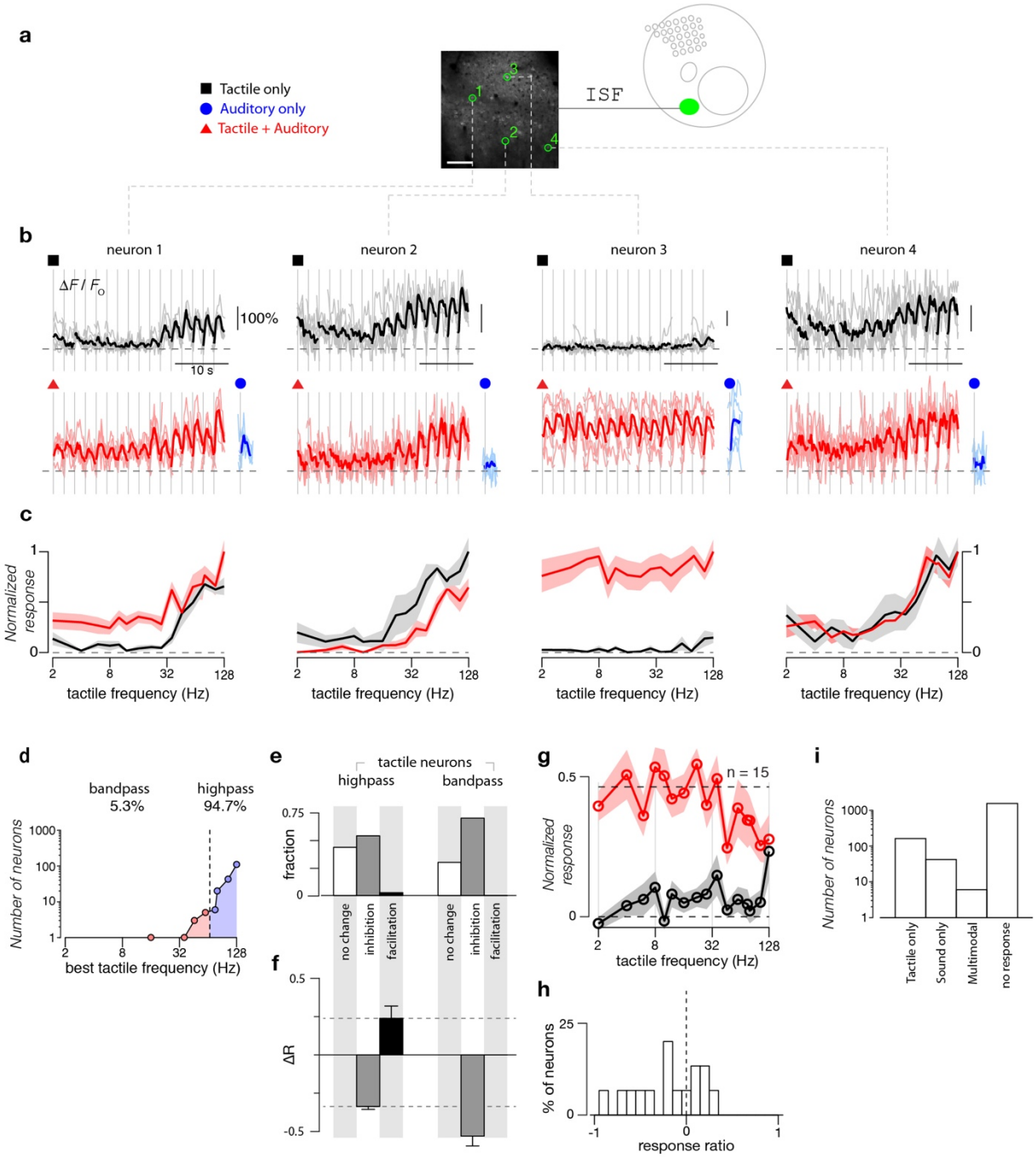
(b) Change in response as carrier frequency of SAM tone is swept from 3 kHz to 96 kHz. Response of exemplar neuron 1 (same as **a**) to SAM tone during concurrent presentation of 28 Hz tactile stimulus (upper row, orange curve, 7 repeats) shows inhibition of tactile response for sound carrier frequency around 10 kHz. Meanwhile, response of exemplar neuron 2 from the same FOV to SAM tone alone as a function of sound carrier frequency shows the strongest responses around 10 kHz (upper row, blue curve, 6 repeats). Responses of the two exemplar neurons are anti-correlated (bottom row, $p < 0.05$ for Pearson's correlation).

(c) For the population of S2 neurons (26 FOVs from 6 mice), sound-selective neurons respond most strongly to SAM tones with carrier frequencies between 6 to 24 kHz (upper row, blue curve), while touch-selective neurons are most inhibited by concurrent SAM tones with carrier frequencies in the same range (upper row, orange curve). This anti-correlation is confirmed by plotting the two responses against each other, parameterized by the sound carrier frequency (bottom row, $p < 0.05$ for Pearson's correlation).

(d) Response of exemplar neurons to tactile stimuli without (black, 9 repeats) or with (red, 7 repeats) concurrent sound stimulation (10 kHz tone with 64 Hz SAM modulation frequency at 40 dB attenuation). Vertical gray line indicates 60 Hz tactile stimulus to be used in **e**.

(e) Response of exemplar neuron from **d** to SAM tone as a function of modulation frequency during concurrent presentation of 60 Hz tactile stimulus (upper row, orange curve, 5 repeats) and for exemplar sound-selective neuron from the same FOV in response to SAM tones alone as a function of modulation frequency (upper row, blue curve, 6 repeats). The tactile-selective neuron's response is independent of sound modulation frequency ($p > 0.5$, ANOVA for modulation frequency) and not correlated with neuron 2's response (bottom row, $p > 0.5$ for Pearson's correlation).

(f) Average responses of tactile-selective neurons (upper row, orange curve) to SAM tones with concurrent tactile stimulus and of sound-selective neurons (upper row, blue curve) to SAM tones alone, both as a function of sound modulation frequency (43 FOVs from 17 mice). Of the 259 tactile neurons, 20 had a significant dependence on sound modulation frequency (ANOVA with pre-established criterion of $p < 0.05$) while 33 of 63 sound-selective neurons depended on sound modulation frequency. Across all neurons, responses of the two populations were slightly but significantly anti-correlated (bottom row, $p < 0.05$ by Pearson's correlation).



Supplementary Figure 5. Diversity of Multimodal Interactions in ISF

(a) Imaging FOV located in ISF. Baseline fluorescence image with exemplar neurons highlighted in green. Scale bar: 100 µm.

(b) Responses of exemplar neurons to tactile stimuli alone (black traces, top panel) or to combined tactile and auditory stimuli (red traces, bottom panel). For the combination stimulus, a concurrent auditory stimulus (10 kHz tones at 40 dB attenuation) was added to the tactile stimuli. Blue traces (bottom right) show averaged responses to 10 kHz SAM tones with 2 Hz SAM envelope at 40 dB attenuation. Each stimulus was repeated 5 times.

(c) Frequency tuning curves of the normalized response to tactile stimuli and combined tactile plus auditory stimuli for exemplar neurons shown in **b**.

(d) Population distribution of best tactile frequencies (*BTF,* see Methods) for ISF neurons. Black circles show number of neurons tuned to each frequency tested. Neurons with best frequencies no higher than 60 Hz were categorized as bandpass neurons (pink) while those with best frequencies above 60 Hz were categorized as highpass neurons (blue). Both x and y axes plotted on a logarithmic scale. Data taken from 9 FOVs across 4 mice.

(e) Percentage of neurons in ISF exhibiting either no change, a decrease (inhibition), or an increase (facilitation) when SAM tones were added to the tactile stimuli. Criteria determined by ANOVA ($p < 0.05$). Neurons were categorized as either highpass (left) or bandpass (right) based on their *BTF*.

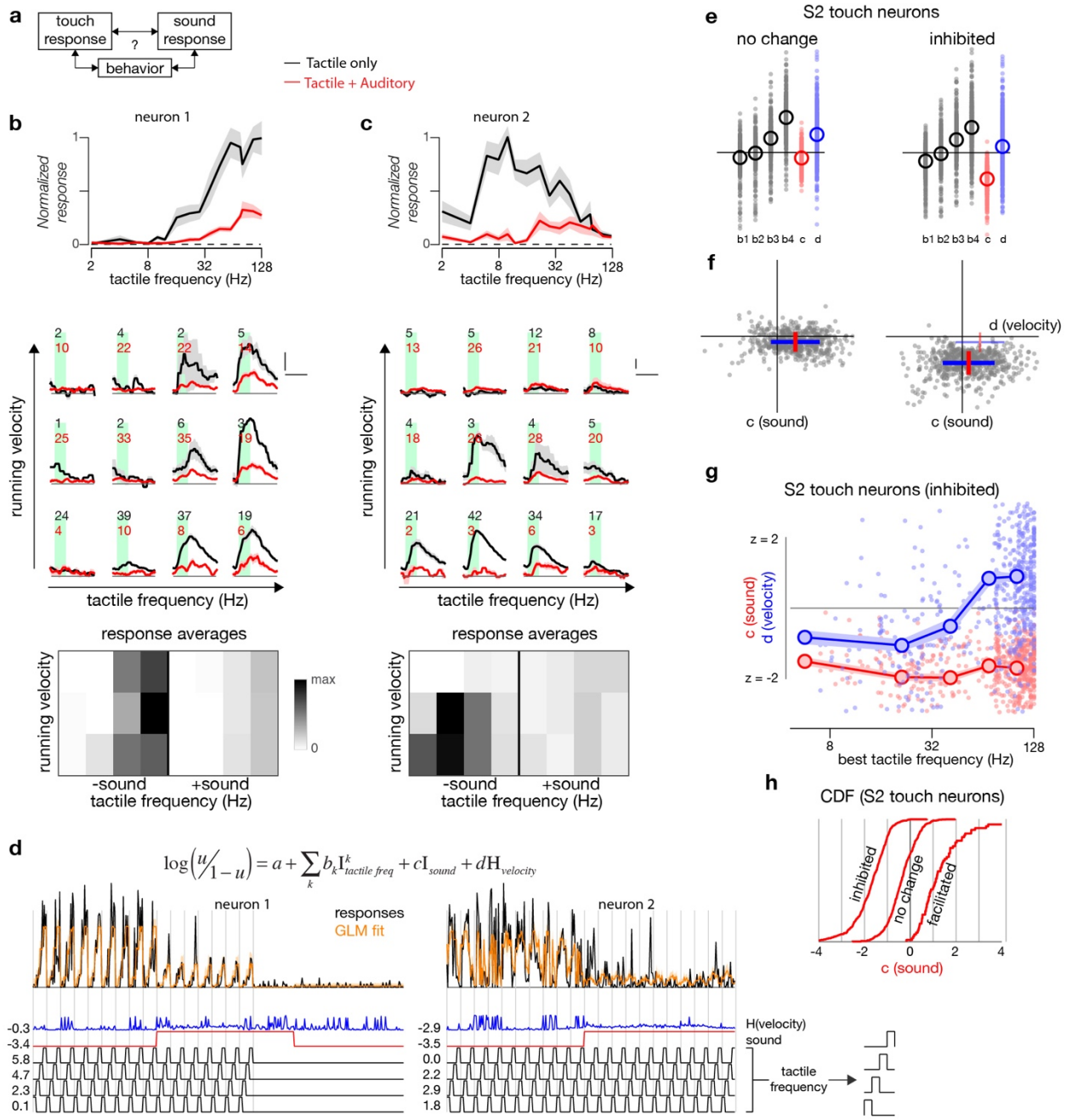
(f) Percentage of change in response of ISF neurons to tactile stimuli (ΔR) when SAM tones were added. Neurons were divided into highpass (left) and bandpass (right) neurons that showed either a decrease (gray) or increase (black) in their response, as indicated in panel **e**. Error bars show the standard error of ΔR among inhibited or facilitated neurons.

(g) Averaged frequency tuning curves of the normalized response to tactile stimuli and combined tactile plus auditory stimuli for sound selective neurons in ISF that respond to the auditory stimuli

used in the combined stimuli (n = 15). Baseline response was additionally subtracted. Standard error shown as shaded gray and red regions.

(h) Distribution of response ratio at high frequency relative to low frequency for sound selective neurons in ISF shown in **g**. Response ratio is calculated as $(r_{lo}-r_{hi})/(r_{lo}+r_{hi})$, where r_{lo} is the response to sounds paired with low frequency (2-8 Hz) tactile stimuli and r_{hi} the response to sounds paired with high frequency (76-128 Hz) tactile stimuli.

(i) Counts of ISF neurons categorized by response type.



Supplementary Figure 6: Sound-Driven Inhibition of Tactile Responses Is Not Explained by Locomotive Behavior

(a) Behavior could potentially mediate the relationship between sound and touch responses in S2. (b) Tactile responses (top, black, 9 repeats) are significantly suppressed by concurrent presentation of sound (top, red, 7 repeats) in an exemplar S2 neuron. Post-stimulus response averages (middle)

are stratified by tactile frequency (horizontal axis), running behavior (vertical axis), and the absence or presence of sound (black and red, respectively), and the average response for each category is represented as an image (bottom).

(c) Responses for a second exemplar neuron, in the same format as **b**. Tactile stimuli repeated 10 times (black curve) and tactile stimuli with concurrent sound repeated 11 times (red curve).

(d) GLM used to model the responses of the two exemplar neurons as a summation of tactile frequency (black, bottom row, divided into 4 frequency categories), running behavior (blue, passed through a nonlinear function), and the absence or presence of sound (red). Each point is the amplitude of response to a stimulus presentation (black, top) with the GLM fit in orange.

(e) Model coefficients (z-scored) for the GLM across the population of S2 neurons with no change or inhibition of tactile responses with the addition of sound (see Figure 3d). Circles represent averages for each coefficient.

(f) Sound (*c*) and running velocity (*d*) coefficients plotted against each other as a scatter plot, with each dot representing one neuron. Red and blue lines represent mean ±SEM for sound and running coefficients, respectively (thin lines on right panel reproduce lines from the left panel).

(g) Sound and running velocity coefficients for the sound-inhibited S2 population, binned by best tactile frequency.

(h) Cumulative distribution of the sound coefficient for three populations of S2 tactile-responsive neurons, as identified in Figure 3d.

REFERENCES

1. Hobbes, T. *Chapter I. Of Sense. Of Man, Being the First Part of Leviathan*. (1651).
2. Stein, B. E. & Meredith, M. A. Merging of the Senses. (1993).
3. Ernst, M. O. & Banks, M. S. Humans integrate visual and haptic information in a statistically optimal fashion. *Nature* **415**, 429–433 (2002).
4. Von Bekesy, G. Similarities between hearing and skin sensations. *Psychol. Rev.* **66**, 1–22 (1959).
5. Lederman, S. J. Auditory Texture Perception. *Perception* **8**, 93–103 (1979).
6. Jousmäki, V. & Hari, R. Parchment-skin illusion: sound-biased touch. *Curr. Biol.* **8**, R190–R191 (1998).
7. Crommett, L. E., Pérez-Bellido, A. & Yau, J. M. Auditory adaptation improves tactile frequency perception. *J. Neurophysiol.* **117**, 1352–1362 (2017).
8. Yau, J. M., Olenczak, J. B., Dammann, J. F. & Bensmaia, S. J. Temporal frequency channels are linked across audition and touch. *Curr. Biol. CB* **19**, 561–566 (2009).
9. Carvell, G. E. & Simons, D. J. Somatotopic organization of the second somatosensory area (SII) in the cerebral cortex of the mouse. *Somatosens. Res.* **3**, 213–237 (1986).
10. Budinger, E., Heil, P., Hess, A. & Scheich, H. Multisensory processing via early cortical stages: Connections of the primary auditory cortical field with other sensory systems. *Neuroscience* **143**, 1065–1083 (2006).
11. Zhou, Y.-D. & Fuster, J. M. Somatosensory cell response to an auditory cue in a haptic memory task. *Behav. Brain Res.* **153**, 573–578 (2004).
12. Dana, H. *et al.* Thy1-GCaMP6 Transgenic Mice for Neuronal Population Imaging In Vivo. *PLOS ONE* **9**, e108697 (2014).

13. Madisen, L. *et al.* Transgenic mice for intersectional targeting of neural sensors and effectors with high specificity and performance. *Neuron* **85**, 942–958 (2015).

14. Wekselblatt, J. B., Flister, E. D., Piscopo, D. M. & Niell, C. M. Large-scale imaging of cortical dynamics during sensory perception and behavior. *J. Neurophysiol.* **115**, 2852–2866 (2016).

15. Issa, J. B. *et al.* Multiscale optical Ca2+ imaging of tonal organization in mouse auditory cortex. *Neuron* **83**, 944–959 (2014).

16. Kwon, S. E., Yang, H., Minamisawa, G. & O'Connor, D. H. Sensory and decision-related activity propagate in a cortical feedback loop during touch perception. *Nat. Neurosci.* **19**, 1243–1249 (2016).

17. Gogolla, N., Takesian, A. E., Feng, G., Fagiolini, M. & Hensch, T. K. Sensory integration in mouse insular cortex reflects GABA circuit maturation. *Neuron* **83**, 894–905 (2014).

18. Sawatari, H. *et al.* Identification and characterization of an insular auditory field in mice. *Eur. J. Neurosci.* **34**, 1944–1952 (2011).

19. Peron, S. P., Freeman, J., Iyer, V., Guo, C. & Svoboda, K. A Cellular Resolution Map of Barrel Cortex Activity during Tactile Behavior. *Neuron* **86**, 783–799 (2015).

20. Carvell, G. E. & Simons, D. J. Biometric analyses of vibrissal tactile discrimination in the rat. *J. Neurosci. Off. J. Soc. Neurosci.* **10**, 2638–2648 (1990).

21. Jones, E. G. & Powell, T. P. An anatomical study of converging sensory pathways within the cerebral cortex of the monkey. *Brain J. Neurol.* **93**, 793–820 (1970).

22. Lemus, L., Hernández, A., Luna, R., Zainos, A. & Romo, R. Do sensory cortices process more than one sensory modality during perceptual judgments? *Neuron* **67**, 335–348 (2010).

23. Krupa, D. J., Matell, M. S., Brisben, A. J., Oliveira, L. M. & Nicolelis, M. A. Behavioral properties of the trigeminal somatosensory system in rats performing whisker-dependent tactile discriminations. *J. Neurosci. Off. J. Soc. Neurosci.* **21**, 5752–5763 (2001).

24. Brosch, M., Selezneva, E. & Scheich, H. Nonauditory events of a behavioral procedure activate auditory cortex of highly trained monkeys. *J. Neurosci. Off. J. Soc. Neurosci.* **25**, 6797–6806 (2005).

25. Foxe, J. J. *et al.* Auditory-somatosensory multisensory processing in auditory association cortex: an fMRI study. *J. Neurophysiol.* **88**, 540–543 (2002).

26. Kayser, C., Petkov, C. I., Augath, M. & Logothetis, N. K. Integration of Touch and Sound in Auditory Cortex. *Neuron* **48**, 373–384 (2005).

27. Niell, C. M. & Stryker, M. P. Modulation of visual responses by behavioral state in mouse visual cortex. *Neuron* **65**, 472–479 (2010).

28. Lee, S., Kruglikov, I., Huang, Z. J., Fishell, G. & Rudy, B. A disinhibitory circuit mediates motor integration in the somatosensory cortex. *Nat. Neurosci.* **16**, 1662–1670 (2013).

29. Holmgren Carl, Harkany Tibor, Svennenfors Björn & Zilberter Yuri. Pyramidal cell communication within local networks in layer 2/3 of rat neocortex. *J. Physiol.* **551**, 139–153 (2004).

30. Johnson, K. O. The roles and functions of cutaneous mechanoreceptors. *Curr. Opin. Neurobiol.* **11**, 455–461 (2001).

31. Manfredi, L. R. *et al.* Natural scenes in tactile texture. *J. Neurophysiol.* **111**, 1792–1802 (2014).

32. Diamond, M. E., von Heimendahl, M., Knutsen, P. M., Kleinfeld, D. & Ahissar, E. 'Where' and 'what' in the whisker sensorimotor system. *Nat. Rev. Neurosci.* **9**, 601–612 (2008).

33. Guić-Robles, E., Valdivieso, C. & Guajardo, G. Rats can learn a roughness discrimination using only their vibrissal system. *Behav. Brain Res.* **31**, 285–289 (1989).

34. Hartmann, M. J. Active Sensing Capabilities of the Rat Whisker System. *Auton. Robots* **11**, 249–254 (2001).

35. von Heimendahl, M., Itskov, P. M., Arabzadeh, E. & Diamond, M. E. Neuronal activity in rat barrel cortex underlying texture discrimination. *PLoS Biol.* **5**, e305 (2007).

36. Ritt, J. T., Andermann, M. L. & Moore, C. I. Embodied information processing: vibrissa mechanics and texture features shape micromotions in actively sensing rats. *Neuron* **57**, 599–613 (2008).

37. Andermann, M. L., Ritt, J., Neimark, M. A. & Moore, C. I. Neural correlates of vibrissa resonance; band-pass and somatotopic representation of high-frequency stimuli. *Neuron* **42**, 451–463 (2004).

38. Garabedian, C. E., Jones, S. R., Merzenich, M. M., Dale, A. & Moore, C. I. Band-pass response properties of rat SI neurons. *J. Neurophysiol.* **90**, 1379–1391 (2003).

39. Melzer, P., Champney, G. C., Maguire, M. J. & Ebner, F. F. Rate code and temporal code for frequency of whisker stimulation in rat primary and secondary somatic sensory cortex. *Exp. Brain Res.* **172**, 370–386 (2006).

40. Bosman, L. W. J. *et al.* Anatomical pathways involved in generating and sensing rhythmic whisker movements. *Front. Integr. Neurosci.* **5**, 53 (2011).

41. Feldmeyer, D. Excitatory neuronal connectivity in the barrel cortex. *Front. Neuroanat.* **6**, (2012).

42. Diamond, M. E., Armstrong-James, M. & Ebner, F. F. Somatic sensory responses in the rostral sector of the posterior group (POm) and in the ventral posterior medial nucleus (VPM) of the rat thalamus. *J. Comp. Neurol.* **318**, 462–476 (1992).

43. Murray, J. D. *et al.* A hierarchy of intrinsic timescales across primate cortex. *Nat. Neurosci.* **17**, 1661–1663 (2014).

44. Iurilli, G. *et al.* Sound-driven synaptic inhibition in primary visual cortex. *Neuron* **73**, 814–828 (2012).

45. Alais, D. & Burr, D. The ventriloquist effect results from near-optimal bimodal integration. *Curr. Biol. CB* **14**, 257–262 (2004).

46. Fetsch, C. R., Turner, A. H., DeAngelis, G. C. & Angelaki, D. E. Dynamic Reweighting of Visual and Vestibular Cues during Self-Motion Perception. *J. Neurosci.* **29**, 15601–15612 (2009).

47. Perrault, T. J., Vaughan, J. W., Stein, B. E. & Wallace, M. T. Neuron-specific response characteristics predict the magnitude of multisensory integration. *J. Neurophysiol.* **90**, 4022–4026 (2003).

48. Stanford, T. R., Quessy, S. & Stein, B. E. Evaluating the operations underlying multisensory integration in the cat superior colliculus. *J. Neurosci. Off. J. Soc. Neurosci.* **25**, 6499–6508 (2005).

49. Foxe, J. J. & Schroeder, C. E. The case for feedforward multisensory convergence during early cortical processing. *Neuroreport* **16**, 419–423 (2005).

50. Macaluso, E., Frith, C. D. & Driver, J. Modulation of human visual cortex by crossmodal spatial attention. *Science* **289**, 1206–1208 (2000).

51. Giard, M. H. & Peronnet, F. Auditory-visual integration during multimodal object recognition in humans: a behavioral and electrophysiological study. *J. Cogn. Neurosci.* **11**, 473–490 (1999).

52. Stein, B. E. & Stanford, T. R. Multisensory integration: current issues from the perspective of the single neuron. *Nat. Rev. Neurosci.* **9**, 255–266 (2008).

53. Wallace, M. T., Ramachandran, R. & Stein, B. E. A revised view of sensory cortical parcellation. *Proc. Natl. Acad. Sci.* **101**, 2167–2172 (2004).

54. Carandini, M. & Heeger, D. J. Normalization as a canonical neural computation. *Nat. Rev. Neurosci.* **13**, 51–62 (2011).

55. Ohshiro, T., Angelaki, D. E. & DeAngelis, G. C. A normalization model of multisensory integration. *Nat. Neurosci.* **14**, 775–782 (2011).

56. Ohshiro, T., Angelaki, D. E. & DeAngelis, G. C. A Neural Signature of Divisive Normalization at the Level of Multisensory Integration in Primate Cortex. *Neuron* **95**, 399-411.e8 (2017).

57. Chen, T.-W. *et al.* Ultrasensitive fluorescent proteins for imaging neuronal activity. *Nature* **499**, 295–300 (2013).

58. Zhu, Y. *et al.* Ablation of NF1 function in neurons induces abnormal development of cerebral cortex and reactive gliosis in the brain. *Genes Dev.* **15**, 859–876 (2001).

59. Jackson, N. & Muthuswamy, J. Artificial dural sealant that allows multiple penetrations of implantable brain probes. *J. Neurosci. Methods* **171**, 147–152 (2008).

60. Knutsen, P. M., Pietr, M. & Ahissar, E. Haptic object localization in the vibrissal system: behavior and performance. *J. Neurosci. Off. J. Soc. Neurosci.* **26**, 8451–8464 (2006).

61. Haidarliu, S., Simony, E., Golomb, D. & Ahissar, E. Muscle architecture in the mystacial pad of the rat. *Anat. Rec. Hoboken NJ 2007* **293**, 1192–1206 (2010).

62. Dörfl, J. The innervation of the mystacial region of the white mouse: A topographical study. *J. Anat.* **142**, 173–184 (1985).

63. Haeffele, B., Young, E. & Vidal, R. Structured low-rank matrix factorization: Optimality, algorithm, and applications to image processing. in *Proceedings of the 31st International Conference on Machine Learning (ICML-14)* 2007–2015 (2014).

64. Issa, J. B., Haeffele, B. D., Young, E. D. & Yue, D. T. Multiscale mapping of frequency sweep rate in mouse auditory cortex. *Hear. Res.* **344**, 207–222 (2017).

65. Guizar-Sicairos, M., Thurman, S. T. & Fienup, J. R. Efficient subpixel image registration algorithms. *Opt. Lett.* **33**, 156–158 (2008).

66. Vogelstein, J. T. *et al.* Fast nonnegative deconvolution for spike train inference from population calcium imaging. *J. Neurophysiol.* **104**, 3691–3704 (2010).

67. Bonin, V., Histed, M. H., Yurgenson, S. & Reid, R. C. Local Diversity and Fine-Scale Organization of Receptive Fields in Mouse Visual Cortex. *J. Neurosci.* **31**, 18506–18521 (2011).

68. Froudarakis, E. *et al.* Population code in mouse V1 facilitates readout of natural scenes through increased sparseness. *Nat. Neurosci.* **17**, 851–857 (2014).

69. Hofer, S. B. *et al.* Differential connectivity and response dynamics of excitatory and inhibitory neurons in visual cortex. *Nat. Neurosci.* **14**, 1045–1052 (2011).

70. Ko, H. *et al.* Functional specificity of local synaptic connections in neocortical networks. *Nature* **473**, 87–91 (2011).

71. Petreanu, L. *et al.* Activity in motor-sensory projections reveals distributed coding in somatosensation. *Nature* **489**, 299–303 (2012).

72. Rajasethupathy, P. *et al.* Projections from neocortex mediate top-down control of memory retrieval. *Nature* **526**, 653–659 (2015).