Multiplane Mesoscope reveals distinct cortical interactions following expectation violations

Natalia Orlova[*,1], Farzaneh Najafi[1,*], Dmitri Tsyboulski[1,2], Sam Seid[1], Sara Kivikas[1], India Kato[1], Fiona Griffin[1], Arielle Leon[1], Ross D. Hytnen[1], Quinn L'Heureux[1], Kat North[1], Jackie Swapp[1], Chelsea Nayan[1], Nicole Hancock[1], Ruweida Ahmed[1], Emily Gelfand[1], Andrew Cho[1], Kyla Mace[1], Robert Howard[1], Linzy Casal[1], Sophie Lambert[1], Eric Kenji Lee[1], Shiella Caldejon[1], Xana Waughman[1], Allison Williford[1], Marina Garrett[1], Doug Ollerenshaw[1], Shawn R. Olsen[1], Peter A. Groblewski[1], Peter Saggau[1,3,4], Jérôme Lecoq[1]

[*] These authors contributed equally.

[1]Allen Institute for Brain Science, Seattle, 98109, WA, USA.

[2]Janelia Research Campus, Ashburn, VA, USA

[3]Baylor College of Medicine, Houston, TX, USA

[4]Italian Institute of Technology, Genoa, Italy

Cortical columns interact through dynamic routing of neuronal activity. Monitoring these interactions in animals performing a behavioral task as close as possible to real time will advance our understanding of cortical computation. We developed the Multiplane Mesoscope which combines three established concepts in microscopy: spatio-temporal multiplexing, remote focusing, and random-access mesoscopy. With the Multiplane Mesoscope, we recorded excitatory and inhibitory neuronal subpopulations simultaneously across two cortical areas and multiple cortical layers in behaving mice. In the context of a visual detection of change task, we used this novel platform to study cortical areas interactions and quantified the cell-type specific distribution of neuronal correlations across a set of visual areas and layers. We found that distinct cortical subnetworks represent expected and unexpected visual events. Our findings demonstrate that expectation violations modify signal routing across cortical columns and establish the Allen Brain Observatory Multiplane Mesoscope as a unique platform to study signal routing across connected pairs of cortical areas.

Introduction

Monitoring spatially distributed neuronal populations with single-neuron resolution is one of the grand challenges in experimental neuroscience. External signals to the brain are indeed represented through distributed, often spread-out cortical networks (Steinmetz et al., 2019), whose interactions are dynamic and depend on the context and predictability of those signals (Keller and Mrsic-Flogel, 2018; Marques et al., 2018). At the micro-circuit level, cortical interactions are regulated through distinct excitatory and inhibitory cell types (Gouwens et al., 2019). Therefore, analysis of the activity of individual cell types across functionally connected cortical areas is necessary to understand cortical interactions in different behavioral contexts. While in vivo calcium imaging with two-photon laser scanning microscopes (TPLSM) is an effective tool for monitoring the activity of specific cell types, sampling *spread-out networks* of neurons in brain volumes with this method has been difficult. This is because TPLSM, if not limited to a single field-of-view (FOV), often restrict size and number of accessible regions-of-interest (ROIs) and doesn't allow adequate sampling rates for simultaneous recordings at multiple cortical planes.

Mesoscopic imaging systems with a large FOV of 3 – 5 mm (Tsai et al., 2015; Sofroniew et al., 2016; Stirman et al., 2016), permit to investigate information processing across multiple cortical areas in small animals. Such a large FOV two-photon random access mesoscope (2P-RAM) was recently developed (Sofroniew et al., 2016) and since commercialized by Thorlabs Inc. It combines several advanced imaging features and technical solutions. In addition to the FOV being increased to 5 mm, it utilizes remote focusing (Botcherby et al., 2007, 2008) to achieve fast and nearly aberration-free axial scanning. While such a system can access a very large volume of the cortex, it remains limited in its imaging throughput, since a single laser beam is employed to *sequentially* scan the brain tissue.

In the past two decades, a variety of methods have been introduced to increase imaging throughput. Examples include imaging with extended focus beams or Bessel beams (Lu et al., 2017, 2020), imaging with engineered point spread function (PSF) (Prevedel et al., 2016; Weisenburger et al., 2019), targeted path galvanometer scanning (Botcherby et al., 2012), 3D random-access scanning with acousto-optic deflectors (AODs) (Duemani Reddy et al., 2008), dual-axis two-photon imaging (Lecoq et al., 2014), and techniques based on light sheet illumination (Ahrens et al., 2013). Throughput-pushing techniques also include multiplexed TPLSM methods enabling simultaneous multi-site recordings with multiple excitation beams and either a single detector (Cheng et al., 2011; Ducros et al., 2013; Chen et al., 2016; Stirman et al., 2016; Beaulieu et al., 2020) or a dedicated detector associated with each laser beam (Rumyantsev et al., 2020).

To overcome these limitations, we developed Multiplane Mesoscope by combining mesoscopic random-access imaging and spatiotemporal multiplexing. Our system increases imaging bandwidth by a factor of two, which allows doubling the number of simultaneously imaged planes, while distributing them laterally and axially. Using this Multiplane Mesoscope, we simultaneously imaged 2 visual areas, and 4 cortical depths in mice performing an image-change detection task,

which involved expected and unexpected events. We found that distinct cell-class activity patterns and cortical interactions occur when visual expectations are violated.

MATERIALS AND METHODS

System design

The design and operation of the original 2P-RAM was described previously (Sofroniew et al., 2016). Our modification added three custom modules to the original system, highlighted in **Fig. 1a**. We detailed this system in a previous publication along with *in vitro* characterization (Tsyboulski et al., 2018). The first custom module controls power distribution between imaging planes and encodes excitation beams with laser pulse delay and polarization. The beam from an 80 MHz femtosecond laser (Chameleon Ultra-II, Coherent), emitting at 910 nm with a pulse duration of ~120 fs and output power of 2.5 W, passes through an electro-optical modulator (EOM1), a prism pulse compressor (PPC), and a second modulator (EOM2) with the internal polarizer removed. Next, two orthogonal laser beams are generated by a polarizing beam splitter, one beam is delayed by 6.25 ns by free-space propagation before the beams are recombined by a polarizing beam splitter and directed towards the Mesoscope's periscopic input. EOM1 controls the total input power, EOM2 rotates the beam polarization and defines the power ratio between both beams.

The second highlighted custom module, includes an additional remote focusing (RFM2) unit (**Fig. 1a**), converting an original Mesoscope to a Multiplane Mesoscope. SolidWorks design files of the module are available upon request. The incoming orthogonally polarized beams are separated by another polarizing beam splitter and pass through wave plates, remote focusing objectives, and are reflected by movable mirrors (RFM1, RFM2). The reflected beams, rotated by 90° polarization, are then recombined by the same polarized beam splitter and directed to the optical scanners. This dual remote focusing assembly independently controls the axial positions of two focal planes and, provides spherical aberration compensation at different beam defocus values associated with remote focusing mirror positions (Botcherby et al., 2007, 2008).

The third added module is the custom demultiplexing circuit we developed to electronically separate time-interleaved fluorescence signals from different imaging planes detected by a single photomultiplier (PMT). This module was described in details previously (Tsyboulski et al., 2018). Circuit diagrams are available upon request.

System characterization

To obtain point spread functions (PSF) we acquired and analyzed image stacks of 200 nm -sized fluorescent beads embedded in 4% aqueous agarose gel over the 5 mm FOV at depths between 0 µm and 500 µm (**Fig. 1b**). Nearly identical PSFs were observed in both imaging channels and were similar to reported data for the 2P-RAM system (Sofroniew et al., 2016).

Laser power throughput was measured with a thermal power meter in selected locations along the optical path of each beam, revealing a 2% power loss in the delayed compared to the main beam due to the additional mirrors in the delay line. Overall system throughput was 16% for non-delayed beam and 14% for the delayed one. This difference can be easily compensated for by steering more laser power into the delayed beam.

Animal head-implants and cortical window implantation

Surgical headpost and cranial window implantation was performed as described elsewhere (de Vries et al., 2020; Garrett et al., 2020; Groblewski and Sullivan, 2020). Headpost and cranial window surgery was performed on healthy mice ranging in age from p37 to p63, weighing no less than 15 g. Dexamethasone (3.2 mg/kg, S.C.) was injected at 12 hrs and 3 hrs before surgery. Mice were initially anesthetized with 5% isoflurane (1–3 min) and placed in a stereotaxic frame (Model #1900, Kopf, Tujunga, CA). During surgery isoflurane levels were maintained at 1.5–2.5%. After a 1- to 2-week-recovery, animals underwent intrinsic signal imaging for retinotopic mapping before behavioral training. Custom surgery tools and implants were designed at the Allen Institute and are described in detail elsewhere (Groblewski and Sullivan, 2020).

Intrinsic imaging and mapping of the visual cortex

Intrinsic signals were used to delineate functionally defined boundaries of visual areas and to target the in vivo two-photon calcium imaging to retinotopically defined locations in primary and secondary areas. Target maps were created from eccentricities at the center of V1 within 10º from the origin and were limited to negative retinotopic values for both altitude and azimuth. These maps were overlaid on an image of the surface vasculature providing fiducial landmarks to guide optical recording sessions and to ensure imaged locations were retinotopically matched across areas.

Mice were lightly anesthetized with 1-1.4% isoflurane (SomnoSuite model #715; Kent Scientific, CT) at a flow rate of 100 ml/min supplemented with ~95% O_2-containing air (Pureline OC4000; Scivena Scientific, OR). Eye drops (Lacri-Lube) maintained hydration and clarity of eyes during anesthesia. Mice were placed on a lab jack platform and head-fixed for imaging normal to the cranial window.

The brain surface was illuminated with two independent sets of LEDs: green (peak/FWHM 527/50 nm) and red (635/ 20 nm) mounted directly on the imaging lens creating a ring light. A pair of camera lenses (Nikkor 105mm f/2.8, 35mm f/1.4, Nikon), provided 3.0x magnification onto a sCMOS camera (Andor Zyla 5.5 10tap). A bandpass filter (Semrock; FF01-630/92 nm) selected red light reflected from the brain.

The visual stimulus consisted of a drifting bar containing a checkerboard pattern, alternating black and white as it sweeps on a grey background along the four cardinal axes 10 times in each direction at a rate of 0.1 Hz. The drifting bar measures 20° x 155°, with individual squares of 25º. To account

for the close viewing angle of the mouse, a spherical warping was applied to all stimuli to ensure apparent size, speed, and spatial frequency were constant across the monitor as seen from the mouse's perspective. To ensure maximal FOV coverage, a 24" monitor was positioned 10 cm from the right eye. The monitor was rotated 30° relative to the animal's dorso-ventral axis and tilted 70° off the horizon to ensure the stimulus was perpendicular to the optic axis of the eye.

Visual stimulation for behavioral experiments

Visual stimuli were generated using custom scripts written in PsychoPy (Peirce, 2007, 2008) and displayed using an ASUS PA248Q LCD monitor, with a 1920 x 1200 display resolution. Visual stimuli were spherically warped as for all intrinsic imaging experiments. The monitor's screen spanned 120° x 95° of visual space without accounting for stimulus warping. Each screen was gamma calibrated using a USB-650 Red Tide Spectrometer (Ocean Optics). Luminance was measured using a SpectroCAL MKII Spectro-radiometer (Cambridge Research Systems). Monitors brightness (30%) and contrast (50%) corresponded to a mean luminance of 50 cd/m^2.

Behavioral training

Details of training are explained in a previous study (Groblewski et al., 2020). In brief, animals learned to detect image changes using the following procedure. Water-restricted mice were habituated to progressively longer duration of head fixation in the behavior enclosure over a five-day period. On day 1 of training, full-field, static square-wave gratings were presented. Mice received automatic water rewards whenever the grating orientation switched between 0o and 90o. In the sessions thereafter, mice were rewarded if they licked within a 750 ms time window after the visual stimulus changed. In stage 1, static gratings were presented as on day 1. In stage 2, static gratings were presented for 250 ms interleaved with 500 ms gray screens. Stage 3 was similar to stage 2, except static gratings were replaced with natural images. Progression through each stage required mice to achieve a peak dprime of 2 during two of the last three sessions. Once in stage 3, mice were considered 'ready for imaging' when two out of three sequential sessions had a dprime >2 and mice had performed at least 100 trials.

Image changes happened randomly during GO trials, according to a geometric distribution (p=0.3) after 5-11 image repetitions following a period where mice consistently withheld licking. 5% of the images were randomly omitted, excluding "change images" and the preceding image to avoid interfering with behavior performance ("omissions", see Fig. 3b). Training sessions lasted for 60 min.

EXPERIMENTAL DESIGN AND STATISTICAL ANALYSES

Description of datasets

Three mouse lines were used in this study: excitatory mouse line: Slc17a7-IRES2-Cre;Camk2a-tTA;Ai93(TITL-GCaMP6f), inhibitory subpopulation VIP: Vip-IRES-Cre;Ai148(TIT2L-GC6f-ICL-tTA2), and inhibitory subpopulation SST: Sst-IRES-Cre;Ai148(TIT2L-GC6f-ICL-tTA2).

For the excitatory cell line, 22 sessions from 6 mice; for the SST cell line, 17 sessions from 4 mice; for the VIP cell line, 25 sessions from 7 mice were recorded.

Processing of calcium imaging movies

The preprocessing of all calcium imaging data was done within the Allen Institute's image processing pipeline described in detail elsewhere (de Vries et al., 2020), and crosstalk removal was applied after pre-processing. Briefly, all data were corrected for brain motion by performing rigid registration in two dimensions. Then, cell-segmentation identified spatial masks of active neurons. Further, mask-matching and crosstalk removal was done by independent component analysis (ICA) and ghost cells were excluded. Finally, fluorescence from spatially-overlapping neuronal masks was unmixed and corrected for neuropil contamination, and lastly, $\Delta F/F$ was computed on corrected masks.

Crosstalk characterization and removal

Crosstalk is a fundamental limitation in multiplexed microscopy systems. One factor defining the amount of inter-plane crosstalk is the combined pulse rate of the multiplexed beams since it limits the maximum temporal separation between fluorescent signals. In our case, multiplexing generates a total pulse rate of 160 MHz which defines the width of the signal integration window, i.e., 6.25 ns. Thus, the detected fluorescence signals are decoded in the demultiplexing circuit using a temporal window of $\leq$ 6.25 ns. Another fundamental factor determining the amount of crosstalk is the duration of fluorescence signals. The fluorescence lifetime τ of calcium indicators derived from GFP is in the range of 2.7 – 3.2 ns (Akerboom et al., 2012; Pliss et al., 2012).

To minimize crosstalk between channels, fluorescence signals in two planes were compared while adjusting the temporal alignment of the multiplexing gates. We fine-tuned the amount of crosstalk for in vivo imaging by adjusting the pulse delay while monitoring excitatory neurons. We recorded signals in both imaging channels, exciting with only one beam. Optimal alignment with respect to the fluorescent signals was observed at 4.5 ns delay, with a residual crosstalk of ~ 15%.

Crosstalk removal was performed on fluorescence traces by ICA, using the FastICA method from the scikit-learn package (Pedregosa et al., 2011), where independent components are estimated by minimizing Gaussianity of the data (Hyvärinen, 1999). The assumption is that observations at the two planes are a linear mix of two clean sources. We assume a mixing matrix of the form [[1-a, a], [b, 1-b]], where a and b are in [0,1]. After FastICA, we transform the resulting mixing matrix to be of this form to recover the proper scaling and polarity of the mixed signals. Prior to FastICA, data

undergo a whitening transformation; performed independently of the scikit module to scale the output signals. Details about this crosstalk-removal algorithm can be found in the platform whitepaper on https://portal.brain-map.org/explore/circuits/visual-behavior-2p.

Population averages of neuronal responses

Traces were aligned to the onset of image omission. The median of neural responses was computed across trials. The average response was then computed across neurons, for each session. A grand average was then computed across sessions (see **Fig. 3c**, left). To quantify the image-evoked responses, the calcium trace of each neuron was averaged over 350 ms after image onset for excitatory and SST neurons, and over a window of [-250, 100] ms relative to image onset for VIP neurons to account for their anticipatory response. Response quantification was done on the mean (across trials) trace of each neuron. The same quantification was performed for omission-evoked responses, except a 500 ms window was used for quantification. The responses were averaged across neurons, for each session, and then a grand average was computed across sessions (see **Fig. 3c**, right).

Correlation of neural responses across cortical planes

On omission-aligned traces, the Spearman correlation coefficient was computed between pairs of neurons across trials, i.e. omissions. Correlations were computed for every individual frame, over [-1, 2] sec relative to the omission. This procedure was done for all pairs of neurons; then an average value was computed across all pairs (**Fig. 4a**). Neuron pairs were present within the same plane, or in 2 different planes. This analysis allowed studying how the response of neurons (within the same plane or in different areas/layers) covaried across trials, and how this coactivation changed at different moments (e.g. after images vs. omissions).

Spearman correlation coefficients were also computed on shuffled traces which were driven by independently shuffling trial orders for each neuron. For each neuron pair, shuffling was repeated 50 times, resulting in a distribution of correlation coefficients for shuffled data.

To compute "noise" correlations, we measured pairwise correlations of "signal"-removed traces. "Signal" was computed by taking the average neural response to each image type (there were 8 images in each session), and subtracting the average from the response of individual trials of that image type.

To quantify coactivation of neurons (**Fig. 4b**), correlations coefficients ("cc") were first averaged over 500 ms after images for excitatory and SST neurons, and over [-250, 250] ms relative to image onset for VIP neurons, accounting for their anticipatory response. Then, correlation coefficients were averaged across baseline frames. We call this quantity "image cc". To quantify omission-evoked coactivation, we averaged correlation coefficients over 750 ms after omissions. We call this

quantity "omission cc". Next we quantified baseline coactivation by averaging correlation coefficients across baselines frames. Baseline was defined as the frame immediately preceding each image presentation, for excitatory and SST neurons, and 250 ms earlier than each image presentation for VIP neurons. We call baseline quantification of correlation "baseline cc". Finally, we measured the change in coactivation during images or omissions by subtracting out "baseline cc" from "image cc", or "omission cc".

The Python package scipy (Virtanen et al., 2020) (scipy.stats.spearmanr) was used for computing correlations. P-values were computed in two ways: 1) using the p-value output of the spearmanr package; 2) manually computing the p-value by comparing the correlation coefficient of real (non-shuffled) data with the shuffled distribution using 2-sided, 1-sample t-test (using the Python package scipy.stats.ttest_1samp).

Statistical tests

We used two-way ANOVA, followed by Tukey HSD to compare population averages across cortical layers. Two-sided t-test was used to compare correlations between real and shuffled data, for each cortical plane. A p-value of 0.05 was used as the significance threshold. For comparison of the correlation coefficients, we used two-tailed, two-sample Kolmogorov–Smirnov tests.

Instrument availability

Instrument CAD and optical design files are available upon request for non-commercial use. Additionally, the dual-beam add-on module for Mesoscope was licensed to Thorlabs. Inc. and is available there.

Data availability

The data that support the findings of this study are publicly available on Allen Institute website (https://portal.brain-map.org/explore/circuits/visual-behavior-2p).

CODE AVAILABILITY

The code used for processing calcium imaging data is publicly available as part of the Allen Software Development Kit (SDK) at https://github.com/AllenInstitute/AllenSDK/. Crosstalk removal was performed using custom routines employing FastICA which is available as part of scikit-learn python package (https://scikit-learn.org/) and as part of the Allen SDK at https://github.com/AllenInstitute/AllenSDK/. Code for the analysis of calcium responses (**Fig. 3** and **Fig 4**) can be found at https://github.com/AllenInstitute/mesoscope_manuscript

RESULTS

Optical design of the Allen Brain Observatory Multiplane Mesoscope.

The detailed layout of the Multiplane Mesoscope is shown in **Fig. 1a**, with the components added to the original Mesoscope highlighted in gray: a Multiplexing Unit containing both an electro-optical modulator and a pathway to delay the orthogonally polarized second excitation beam, a second remote focusing unit (RFU), and a demultiplexing electronic circuit.

To achieve doubling of the imaging throughput, we added a second imaging channel by splitting the laser beam in two and encoding each beam with its own polarization and pulse time-of-arrival (delay). Using polarization encoding, each beam was routed through a dedicated remote focusing unit to achieve independent axial positioning of two focal planes. Fluorescence generated by each laser beam was decoded in a custom demultiplexing circuit using it's known pulse time-of-arrival. Thereby, we achieved nearly aberration-free *simultaneous* imaging from two focal planes, independently positioned in the axial direction. As a result, the Multiplane Mesoscope achieves a unique balance of optical resolution, optical field of view, and imaging throughput.

Two sequential electro-optical modulators (EOM) control the excitation power of both imaging planes, distributing it efficiently between imaging depths. EOM1 controls the total amount of laser power, while EOM2 and a polarizing beam splitter distribute the power between two beams and convert the incoming laser pulses into two orthogonally polarized trains. The delay line separates both trains by half a pulse period. Another polarizing beam splitter combines both beams, creating two interleaved, temporally encoded excitation pulse trains. The input polarizing beam splitter of the dual-plane RFU steers those trains to two separate RFUs, controlling the beams' collimation independently. Beams are then recombined and directed to the shared XY scanner. Thus, the two imaging planes are positioned independently in axial dimension while remaining laterally coupled during scanning.

Time-interleaved fluorescence signals from the two imaging planes are detected by a single PMT. The signal separation in the 100 MHz range is challenging and required two critical issues to be resolved. First, since the 80 MHz laser pulse frequency slightly fluctuates in time, detection electronics must be synchronized with excitation. Second, signals at the PMT output corresponding to neural activity vary significantly, from high-bandwidth single-photon detection events to significantly higher and longer bursts from high-count photon fluxes. Previous reports utilized different approaches for signal demultiplexing, including gated photon counting (Stirman et al., 2016) and more recently high-frequency sampling at rates near or above 1 GHz and digital unmixing with field-programmable gate arrays (Weisenburger et al., 2019; Beaulieu et al., 2020). We developed a simple and efficient analog demultiplexing method which is compatible with standard data acquisition hardware commonly used for two-photon microscopy (see **Methods**).

We validated that the Multiplane Mesoscope maintained a high optical resolution throughout the full FOV for both beams (lateral PSF < 1 µm and axial PSF < 6 µm for all measured imaging locations, see **Fig. 1b-c**).

Imaging modalities and flexibility of the Allen Brain Observatory Multiplane Mesoscope imaging platform.

The Multiplane Mesoscope was designed to provide a large flexibility of experimental set-up as our intent was to create the future basis of the Allen Brain Observatory imaging platform. We can match imaging modes with the density and spatial distribution of neural labeling in various Cre-lines. Multiple ROIs can be positioned in arbitrary X-Y locations within the total 5 mm optical access FOV, while axially, simultaneously scanned planes can be within the 2 mm range of the remote focusing units. For denser excitatory cell lines, one could benefit from imaging a single column with as many z-planes as possible. When imaging a layer-specific cell line, given the desired scan rate, the operator can decide to image 2 or more lateral areas with axial planes positioned in the labeled cortical layer.

Two configurations currently used on the Allen Brain Observatory platform are shown in Fig. 2. In imaging mode 1, (**Fig. 2A**), we aimed to image two cortical columns (for example V1 and one adjacent higher-visual areas) with 4 axial planes in each. To optimize laser power distribution, we coupled the most superficial plane with the deepest one and imaged the two middle planes as a pair. In this configuration, the resulting scan rate for each plane was ~10.7 Hz. **Fig 2.c** shows single planes acquired in this mode with single-cell resolution along with example extracted calcium traces.

Imaging mode 2 used in the Brain Observatory consisted of four lateral ROIs distributed across four areas of the visual cortex, and two axial planes. This mode was used primarily for Cre lines where neural labeling was confined to a smaller cortical depth and was sampled with the same scan-rate (10.7 Hz).

Simultaneous characterization of cell-type specific neuronal responses across brain areas during a visual behavior task.

To establish this imaging platform as a tool to study cortical computation, we integrated this microscope with a standardized behavior training cluster (Garrett et al., 2019; Groblewski et al., 2020). In this study, we used this integration to study cortical column interactions during active behavior (**Fig. 3**). Since distinct cortical pathways are suggested to carry sensory vs. prediction signals (Bastos et al., 2012; Keller and Mrsic-Flogel, 2018), we hypothesized that cortical interactions between pairs of areas and layers should be distinct in response to expected vs. unexpected events. To address this question, we trained mice on a go/no-go, image-change detection task that was previously established (Garrett et al., 2020; Groblewski et al., 2020). In brief,

water-restricted animals were presented with a constant stream of natural images (250 ms) interleaved with gray screens of matched luminance (500 ms) (**Fig. 3b**). On go trials, a change in image identity occurred and mice received a water reward if they licked within 500 ms after the image change. During imaging sessions in well-trained mice, a pseudo-random 5% of non-change images were replaced with a gray screen ("omissions"). As the mice had extensive experience with the highly regular timing of image presentations in the task (every ~750 ms), the rare stimulus omissions were unexpected events, while the frequent, repeated image presentations were expected events. We recorded the activity of excitatory cells as well as inhibitory somatostatin-expressing neurons (SST) and vasoactive intestinal polypeptide-expressing (VIP) neurons, using 3 distinct mouse lines (see **Methods**). We imaged at 4 cortical depths in 2 visual areas (V1 and LM; **Fig. 3a**).

To test if expected and unexpected events are encoded by different neuronal circuitries, we first studied how the activity of excitatory and inhibitory subtypes (VIP, SST) is modified in response to image presentations and omissions, and how it depends on cortical area and depth in the same mice. The population average of neural activities demonstrated a clear difference across cell classes: images activated excitatory neurons and, more robustly, SST neurons (**Fig. 3c**, top, middle). Excitatory neurons did not respond to omissions in any visual area (**Fig. 3c**, bar plots, top; black: V1; blue: LM). SST neurons were slightly inhibited after omissions in all recorded locations (**Fig. 3c**, bar plots, middle). In sharp contrast to excitatory and SST neurons, VIP neurons, in all layers and areas, were robustly activated after omissions (**Fig. 3c**, bottom). They also demonstrated small anticipatory activity approximately 250 ms prior to each image presentation and were inhibited immediately after the image (**Fig. 3c**, bottom), confirming previous results obtained with more conventional TPLSM imaging instruments using the same behavioral task (Garrett et al., 2020).

The Multiplane Mesoscope allowed unbiased comparison of neuronal responses across cortical depths of V1 and LM, all recorded simultaneously in each experiment. We found excitatory neuronal responses to images became progressively stronger in deeper layers of V1(Niell and Stryker, 2008) and LM (**Fig. 3c**, top). SST responses to images were strongest in V1 layer 2/3 compared to other layers but did not differ among cortical layers of LM (**Fig. 3c**, middle; V1: one-way ANOVA: $p=0.005$, Tukey HSD: $p<0.05$ for 1^{st} vs. 2^{nd} depth and 1^{st} vs. 3^{rd} depth comparison, $p>0.05$ for all other pairwise depth comparisons; LM: one-way ANOVA: $p=0.32$). VIP responses to omissions were significantly stronger in deeper layers of V1 but did not differ among LM layers (**Fig 3c**, bottom; V1: one-way ANOVA: $p=0.001$, Tukey HSD: $p<0.05$ for 3^{rd} vs. 1^{st} and 3^{rd} vs. 2^{nd} depth comparison, $p>0.05$ for all other pairwise depth comparisons; LM: one-way ANOVA: $p=0.12$).

Our analyses indicated that image presentations were more robustly represented by excitatory and SST neurons, particularly in deeper layers; in contrast, unexpected omissions were represented in VIP neuron activity. Previous studies have suggested that feedback pathways convey prediction signals from higher order cortical areas to superficial layers of V1(Yang et al., 2016; Keller and Mrsic-Flogel, 2018; Marques et al., 2018). In light of these studies, we used simultaneous multi-plane recording to investigate if cortical interactions might be different when expectations are

violated, i.e. during omissions. Importantly, this analysis could only be carried out with the multi-area, multi-plane capabilities of the Multiplane Mesoscope.

Neural correlations across cortical areas and layers following expected and unexcepted events.

To study cortical interactions, we correlated neural responses across repeated image presentations, as well as omissions, for each cell type (pairwise "noise" correlations; **Fig. 4**). We measured correlations within each area (V1-V1, LM-LM; **Fig. 4**, middle, right), as well as across areas (V1-LM; **Fig. 4**, left). In the excitatory network, after image presentations, correlations in neural activity were significantly stronger between neurons in deep layers of V1 and LM compared to those in the superficial layers (V1-LM: **Fig. 4a,b**, top left; V1-V1: **Fig. 4b**, top middle; LM-LM: **Fig. 4b**, top right; see Table 1 for statistical details). In the SST network, correlations increased broadly across all layers and areas after image presentations (**Fig. 4b**, middle. One-way ANOVA: $P>0.05$). Omissions did not modify neural correlations for the excitatory or the SST network (**Fig. 4b**, top and middle for excitatory and SST, respectively).

In the VIP network, coactivation patterns were strikingly different compared to other cell types: neural correlations were significantly increased after omissions. This occurred for neuron pairs across all layers and areas. On the other hand, the coactivation pattern of the VIP network did not change following images, among any of the cortical planes (**Fig. 4b**, bottom). Notably, correlations were overall much weaker among excitatory neurons compared to inhibitory neurons, confirming previous results and indicating stronger local connectivity between inhibitory neurons (Packer and Yuste, 2011; Najafi et al., 2020).

DISCUSSION

While they are common in many fields, integrated high-throughput *in vivo* imaging platforms are scarce in systems neuroscience. This is due to the complex integration of experimental designs and instrumentation. It is currently more practical for individual laboratories to purchase dedicated equipment and specialize its use to a few ongoing experiments, sometimes in collaboration with neighboring laboratories. This practice incentivizes the development of new technologies at the cost of their integration into robust and scalable data collection workflows. Here our goal was to combine novel technology (dual-remote focusing) and existing imaging technologies (2p-RAM and temporal multiplexing) with a scalable behavioral task. We shared both the data and the experimental platform. Indeed, this imaging platform was used to generate the dataset recently released to the public (https://portal.brain-map.org/explore/circuits/visual-behavior-2p) and will be available in the coming years for experimental proposals as part of the OpenScope project funded by the NIH (NS113646-01A1). We hope that future developments in imaging techniques will be fueled by this approach, allowing the integration of complex, nascent technologies (Lecoq et al.,

2019) for high throughput neuronal recordings, and involving modern machine learning tools (Lecoq et al., 2020).

Here, we combined three existing technical solutions to double the throughput of a large field of view two-photon mesoscope and study interactions of neuronal networks distributed across visual areas and layers. The system combines benefits of random accessing regions of interest across an ultra-large FOV and dual-beam scanning which increases imaging throughput. Similar systems have been developed to enable imaging of distributed cortical networks. For example, a dual-beam system with fully independent scan engines and an ultra-large FOV (Stirman et al., 2016) was reported and is disseminating it under open-source license. Our approach is distinct, as we combined our system with a behavioral platform, and eventually made it available as an integrated experimental platform (NIH OpenScope project).

While temporal multiplexing was introduced to two-photon imaging a decade ago (Cheng et al., 2011), it is not yet widely used, given the challenge of unmixing high-frequency PMT signals and the associated technical complexity. Here we integrated this technology with a commercially available instrument as an add-on module. This modularity was central to our design. This approach proves to be effective as external laboratories have already integrated our module in their research on cortical computation (Stringer et al., 2021). Similarly a Bessel beam module was recently introduced to the mesoscope (Lu et al., 2020). A recent modification of our approach focused additional beams along the axial direction, further scaling up population recording (Demas et al., 2021). Once challenges associated with residual plane crosstalk in these instruments are fully addressed, we anticipate that disseminating and integrating multiplexing approaches will unlock our access to cortex-wide recordings with single-cell resolution.

To validate our imaging platform, we integrated our instrument with a previously established versatile behavior task. The image-change detection task we used in our study allowed studying the representation of expectation violation signals in the brain: First, it included repeated presentations of the same image, hence forming a prediction signal; second, it involved unexpected image omissions, hence generating a prediction violation signal.

Predictive coding requires communication across brain areas for generating and updating predictions (Rao and Ballard, 1999; Hamm et al., 2018; Keller and Mrsic-Flogel, 2018; Keller et al., 2020). Therefore, to investigate the neural circuit mechanisms that underlie predictive coding, we need to study signal flow across cortical areas while animals make and update predictions. The Multiplane Mesoscope allowed us to study cortical columns during our visual task which involved predictive signals.

We found that the correlation among deeper layers of excitatory neurons goes up after image presentations, aligned with the bottom-up sensory inputs from the thalamus to deeper layers of the cortex. Following omissions on the other hand, excitatory neurons in most layers were barely active, and the coactivation patterns among excitatory neurons did not change either. This is at

odds with the deviant detection studies that show excitatory neurons encode deviant stimuli (Hamm et al., 2018). This difference could be due to the fact that omissions and deviant stimuli, although both representing expectation violation events, are encoded via distinct circuits.

We previously found that VIP neurons in the visual cortex show robust responses after unexpected omissions (Garrett et al., 2019). In the current study, we further characterized the response of VIP neurons across cortical layers of V1 and LM. Our results demonstrate that VIP neurons in V1 and LM are broadly coactivated in both superficial and deep cortical layers when unexpected omissions occur. Previous studies suggested that VIP neurons may carry a prediction signal (Krabbe et al., 2019). Given their inputs from association areas and the neuromodulatory system, VIP neurons may also play a central role in gating sensory inputs and driving context-dependent behavior (Fu et al., 2014; Kuchibhotla et al., 2017; Garrett et al., 2020). But, what is the source of the VIP omission signal in our study? The omission response occurs in the VIP network in all visual areas and layers, suggesting that it may arise from broad shared inputs onto the VIP network. The neuromodulatory system represents a good candidate for this shared input. Future studies examining the role of the neuromodulatory system in generating the VIP omission response will shed light on the mechanisms underlying the VIP omission response.

Conclusion

In summary, the Allen Brain Observatory Multiplane Mesoscope enables experiments requiring in-vivo simultaneous imaging at different depths across multiple areas of the cortex. Importantly, we accomplished this without trading image quality for temporal resolution. Imaging with this instrument during a behavioral task, we demonstrated, for individual excitatory and inhibitory cell types, that interactions across cortical columns are distinct following expected and unexpected events. As a central piece of the Allen Brain Observatory two-photon data pipeline (de Vries et al., 2020) **and of the OpenScope project (NIH U24 resource,** https://alleninstitute.org/what-we-do/brain-science/research/products-tools/openscope/) the Multiplane Mesoscope will be used to generate open datasets in service of the neuroscience community.

Corresponding authors:

Correspondence and requests for materials should be addressed to N.O. (nataliao@alleninstitute.org) or J.L. (jeromel@alleninstitute.org), and for biological data analysis to F.N. (farzaneh.najafi@alleninstitute.org) or J.L.

Conflict of interest statement

The dual-beam add-on module (D.T., N.O., J.L and P.S.) intellectual property has been licensed to Thorlabs. Inc., by the Allen Institute.

Acknowledgments

We wish to thank the founder of the Allen Institute for Brain Science, Paul G. Allen, for his vision, encouragement, and support. We thank Daniel Kapner and Wayne Wakeman for help with processing datasets. We thank Marius Pachitariu for feedbacks and comments on data and the instrument. We thank Vidrio Technologies for the development of add-on modules for ScanImage. We thank Thorlabs, Inc. for microscope support. We thank Karel Svoboda, Daniel Flickinger and Nicholas Sofroniew for support through the dissemination of the original Mesoscope design and their technical support. We thank David Sullivan for help with designing a custom behavioral stage. We thank Christof Koch and Cliff Kerr for comments on the manuscript.

Author Contributions

	Project conception & supervision	Microscope add-on design and build	Biological analysis conceptualization	Experimental software design and build	Data analysis	2 photon data acquisition	2 photon data pre-processing	Pipeline integration	Project coordination	Surgical, behavioral pipeline management	Behavioral task design & initial characterization	Mouse surgeries	Behavioral training	Intrinsic imaging	Writing paper
AC													■		
AL				■				■							
AW										■					
CN												■	■		
DO											■				
DT		■		■											■
EG												■			
EKL														■	
FG						■		■							
FN			■		■		■								■
LC									■						
IK						■								■	
JL	■		■				■	■	■						■
JS												■	■		
KM												■			
KN												■	■		
MG											■				
NH												■	■		
NO		■		■			■	■							■
PG										■	■				
PS	■	■													■
QL				■				■							
RA												■	■		
RDH				■											
RH												■			
SC														■	
SK						■		■						■	
SL												■			
SRO											■				
SS						■		■						■	
XW														■	

References

Ahrens MB, Orger MB, Robson DN, Li JM, Keller PJ (2013) Whole-brain functional imaging at cellular resolution using light-sheet microscopy. Nat Methods 10:413–420.

Akerboom J et al. (2012) Optimization of a GCaMP calcium indicator for neural activity imaging. J Neurosci.

Bastos AM, Usrey WM, Adams RA, Mangun GR, Fries P, Friston KJ (2012) Canonical Microcircuits for Predictive Coding. Neuron 76:695–711 Available at: https://www.sciencedirect.com/science/article/pii/S0896627312009592 [Accessed November 27, 2019].

Beaulieu DR, Davison IG, Kılıç K, Bifano TG, Mertz J (2020) Simultaneous multiplane imaging with reverberation two-photon microscopy. Nat Methods.

Botcherby EJ, Juskaitis R, Booth MJ, Wilson T (2007) Aberration-free optical refocusing in high numerical aperture microscopy. Opt Lett 32:2007.

Botcherby EJ, Juškaitis R, Booth MJ, Wilson T (2008) An optical technique for remote focusing in microscopy. Opt Commun.

Botcherby EJ, Smith CW, Kohl MM, Deďarre D, Booth MJ, Juškaitis R, Paulsen O, Wilson T (2012) Aberration-free three-dimensional multiphoton imaging of neuronal activity at kHz rates. Proc Natl Acad Sci U S A 109:2919–2924.

Chen JL, Voigt FF, Javadzadeh M, Krueppel R, Helmchen F (2016) Long-range population dynamics of anatomically defined neocortical networks. Elife.

Cheng A, Gonçalves JT, Golshani P, Arisaka K, Portera-Cailliau C (2011) Simultaneous two-photon calcium imaging at different depths with spatiotemporal multiplexing. Nat Methods 8:139–142.

de Vries SEJ et al. (2020) A large-scale standardized physiological survey reveals functional organization of the mouse visual cortex. Nat Neurosci.

Demas J, Manley J, Tejera F, Kim H, Traub FM, Chen B, Vaziri A (2021) High-Speed, Cortex-Wide Volumetric Recording of Neuroactivity at Cellular Resolution using Light Beads Microscopy. bioRxiv:2021.02.21.432164 Available at: https://doi.org/10.1101/2021.02.21.432164 [Accessed June 24, 2021].

Ducros M, Houssen YG, Bradley J, De Sars V, Charpak S (2013) Encoded multisite two-photon microscopy. Proc Natl Acad Sci U S A 110:13138–13143.

Duemani Reddy G, Kelleher K, Fink R, Saggau P (2008) Three-dimensional random access multiphoton microscopy for functional imaging of neuronal activity. Nat Neurosci.

Fu Y, Tucciarone JM, Espinosa JS, Sheng N, Darcy DP, Nicoll RA, Huang ZJ, Stryker MP (2014) A cortical circuit for gain control by behavioral state. Cell.

Garrett M, Manavi S, Roll K, Ollerenshaw D, Groblewski P, Kiggins J, Jia X, Casal L, Mace K, Williford A, Leon A, Mihalas S, Olsen S (2019) Experience shapes activity dynamics and stimulus coding of VIP inhibitory and excitatory cells. bioRxiv.

Garrett M, Manavi S, Roll K, Ollerenshaw DR, Groblewski PA, Ponvert ND, Kiggins JT, Casal L, Mace K, Williford A, Leon A, Jia X, Ledochowitsch P, Buice MA, Wakeman W, Mihalas S, Olsen SR (2020) Experience shapes activity dynamics and stimulus coding of VIP inhibitory cells. Elife.

Gouwens NW et al. (2019) Classification of electrophysiological and morphological neuron types in the mouse visual cortex. Nat Neurosci.

Groblewski P, Sullivan D (2020) A standardized head-fixation system for performing large-scale, in-vivo physiological recordings in mice.

Groblewski PA, Ollerenshaw DR, (2020) Characterization of Learning, Motivation, and Visual Perception in Five Transgenic Mouse Lines Expressing GCaMP in Distinct Cell Populations. Front Behav Neurosci.

Hamm J, Shymkiv Y, Han S, Yang W, Yuste R (2018) Cortical subnetworks encode context of visual stimulus. bioRxiv:452219 Available at: https://doi.org/10.1101/452219 [Accessed June 29, 2021].

Hyvärinen A (1999) Fast and robust fixed-point algorithms for independent component analysis. IEEE Trans Neural Networks.

Keller AJ, Roth MM, Scanziani M (2020) Feedback generates a second receptive field in neurons of the visual cortex. Nature 582:545–549 Available at: https://www.nature.com/articles/s41586-020-2319-4 [Accessed June 29, 2021].

Keller GB, Mrsic-Flogel TD (2018) Predictive Processing: A Canonical Cortical Computation. Neuron.

Krabbe S, Paradiso E, d'Aquin S, Bitterman Y, Courtin J, Xu C, Yonehara K, Markovic M, Müller C, Eichlisberger T, Gründemann J, Ferraguti F, Lüthi A (2019) Adaptive disinhibitory gating by VIP interneurons permits associative learning. Nat Neurosci.

Kuchibhotla K V., Gill J V., Lindsay GW, Papadoyannis ES, Field RE, Sten TAH, Miller KD, Froemke RC (2017) Parallel processing by cortical inhibition enables context-dependent behavior. Nat Neurosci.

Lecoq J, Oliver M, Siegle JH, Orlova N, Koch C (2020) Removing independent noise in systems neuroscience data using DeepInterpolation. bioRxiv.

Lecoq J, Orlova N, Grewe BF (2019) Wide. Fast. Deep: Recent Advances in Multiphoton Microscopy of In Vivo Neuronal Activity. J Neurosci 39.

Lecoq J, Savall J, Vučinić D, Grewe BF, Kim H, Li JZ, Kitch LJ, Schnitzer MJ (2014) Visualizing mammalian brain area interactions by dual-axis two-photon calcium imaging. Nat Neurosci 17:1825–1829.

Lu R, Liang Y, Meng G, Zhou P, Svoboda K, Paninski L, Ji N (2020) Rapid mesoscale volumetric imaging of neural activity with synaptic resolution. Nat Methods.

Lu R, Sun W, Liang Y, Kerlin A, Bierfeld J, Seelig JD, Wilson DE, Scholl B, Mohar B, Tanimoto M, Koyama M, Fitzpatrick D, Orger MB, Ji N (2017) Video-rate volumetric functional imaging of the brain at synaptic resolution. Nat Neurosci 20:620–628.

Marques T, Nguyen J, Fioreze G, Petreanu L (2018) The functional organization of cortical feedback inputs to primary visual cortex. Nat Neurosci.

Najafi F, Elsayed GF, Cao R, Pnevmatikakis E, Latham PE, Cunningham JP, Churchland AK (2020) Excitatory and Inhibitory Subnetworks Are Equally Selective during Decision-Making and Emerge Simultaneously during Learning. Neuron.

Niell CM, Stryker MP (2008) Highly selective receptive fields in mouse visual cortex. J Neurosci.

Packer AM, Yuste R (2011) Dense, unspecific connectivity of neocortical parvalbumin-positive interneurons: A canonical microcircuit for inhibition? J Neurosci.

Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J, Passos A, Cournapeau D, Brucher M, Perrot M, Duchesnay É (2011) Scikit-learn: Machine Learning in Python. J Mach Learn Res 12:2825–2830 Available at: http://scikit-learn.sourceforge.net. [Accessed June 23, 2021].

Pliss A, Zhao L, Ohulchanskyy TY, Qu J, Prasad PN (2012) Fluorescence lifetime of fluorescent proteins as an intracellular environment probe sensing the cell cycle progression. ACS Chem Biol.

Prevedel R, Verhoef AJ, Pernía-Andrade AJ, Weisenburger S, Huang BS, Nöbauer T, Fernández A, Delcour JE, Golshani P, Baltuska A, Vaziri A (2016) Fast volumetric calcium imaging across multiple cortical layers using sculpted light. Nat Methods.

Rao RPN, Ballard DH (1999) Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. Nat Neurosci 2:79–87 Available at: http://www.nature.com/articles/nn0199_79 [Accessed November 27, 2019].

Rumyantsev OI, Lecoq JA, Hernandez O, Zhang Y, Savall J, Chrapkiewicz R, Li J, Zeng H, Ganguli S, Schnitzer MJ (2020) Fundamental bounds on the fidelity of sensory cortical coding. Nature.

Sofroniew NJ, Flickinger D, King J, Svoboda K (2016) A large field of view two-photon mesoscope with subcellular resolution for in vivo imaging. Elife 5 Available at: https://elifesciences.org/articles/14472 [Accessed December 8, 2019].

Steinmetz NA, Zatka-Haas P, Carandini M, Harris KD (2019) Distributed coding of choice, action and engagement across the mouse brain. Nature.

Stirman JN, Smith IT, Kudenov MW, Smith SL (2016) Wide field-of-view, multi-region, two-photon imaging of neuronal activity in the mammalian brain. Nat Biotechnol 34:857–862.

Stringer C, Michaelos M, Tsyboulski D, Lindo SE, Pachitariu M (2021) High-precision coding in visual cortex. Cell 184:2767-2778.e15 Available at: http://www.cell.com/article/S0092867421003731/fulltext [Accessed June 24, 2021].

Tsai PS, Mateo C, Field JJ, Schaffer CB, Anderson ME, Kleinfeld D (2015) Ultra-large field-of-view two-photon microscopy. Opt Express.

Tsyboulski D, Orlova N, Griffin F, Seid S, Lecoq J, Saggau P (2018) Remote focusing system for simultaneous dual-plane mesoscopic multiphoton imaging. bioRxiv:503052 Available at: https://www.biorxiv.org/content/10.1101/503052v1 [Accessed December 8, 2019].

Virtanen P et al. (2020) SciPy 1.0: fundamental algorithms for scientific computing in Python. Nat Methods 17:261–272 Available at: https://doi.org/10.1038/s41592-019-0686-2 [Accessed June 23, 2021].

Weisenburger S, Tejera F, Demas J, Chen B, Manley J, Sparks FT, Martínez Traub F, Daigle T, Zeng H, Losonczy A, Vaziri A (2019) Volumetric Ca2+ Imaging in the Mouse Brain Using Hybrid Multiplexed Sculpted Light Microscopy. Cell.

Yang W, Miller J eun K, Carrillo-Reid L, Pnevmatikakis E, Paninski L, Yuste R, Peterka DS (2016) Simultaneous Multi-plane Imaging of Neural Circuits. Neuron 89:269.

Excitatory, V1-LM correlations	**Omission responses**: one-way ANOVA: F=1.96, p=0.30; TUKEY HSD: p>0.05 for all pairwise comparisons. **Image responses:** one-way ANOVA: F=5.01, p<0.001; TUKEY HSD: p<0.05 for the following pairwise comparisons: [LM depth, V1 depth] vs [LM depth, V1 depth]: [0 0] vs [2 2] [0 0] vs [2 3] [1 0] vs [2 2] [1 0] vs [2 3] [3 0] vs [2 2] [0 1] vs [2 2] [0 1] vs [2 3] [1 1] vs [2 2] [1 1] vs [2 3] [2 1] vs [2 2] [2 1] vs [2 3] [3 1] vs [2 2] [3 1] vs [2 3] [2 2] vs [0 3] [2 2] vs [1 3] [1 3] vs [2 3]
Excitatory, V1-V1 correlations	**Omission responses**: one-way ANOVA: F=0.59434, p=0.79; TUKEY HSD: p>0.05 for all pairwise comparisons. **Image responses:** one-way ANOVA: F=9.849076, p<0.001; TUKEY HSD: p<0.05 for the following pairwise comparisons: [LM depth, V1 depth] vs [LM depth, V1 depth]: [0 0] vs [2 2] [0 0] vs [2 3] [0 0] vs [3 3] [0 1] vs [2 2] [0 1] vs [2 3] [0 1] vs [3 3] [1 1] vs [2 2] [1 1] vs [2 3] [1 1] vs [3 3] [0 2] vs [2 2] [0 2] vs [3 3] [1 2] vs [2 2] [1 2] vs [3 3] [2 2] vs [0 3]

	[2 2] vs [1 3] [0 3] vs [3 3] [1 3] vs [2 3] [1 3] vs [3 3]
Excitatory, LM-LM correlations	**Omission responses**: one-way ANOVA: F=5.76883, p<0.001; TUKEY HSD: p<0.05 for the following pairwise comparisons: [LM depth, V1 depth] vs [LM depth, V1 depth]: [0 0] vs [3 3] [0 1] vs [3 3] [1 1] vs [3 3] [0 2] vs [3 3] [1 2] vs [3 3] [2 2] vs [3 3] [0 3] vs [3 3] [1 3] vs [3 3] [2 3] vs [3 3] **Image responses:** one-way ANOVA: F=2.822217, p=0.01; TUKEY HSD: p<0.05 for the following pairwise comparisons: [LM depth, V1 depth] vs [LM depth, V1 depth]: [0 0] vs [2 2] [0 1] vs [2 2] [2 2] vs [3 3]
SST, V1-LM correlations	**Omission responses**: one-way ANOVA: F=0.470705, p=0.93; TUKEY HSD: p>0.05 for all pairwise comparisons. **Image responses:** one-way ANOVA: F=0.439109, p=0.95; TUKEY HSD: p>0.05 for all pairwise comparisons.
SST, V1-V1 correlations	**Omission responses**: one-way ANOVA: F=0.270538, p=0.97; TUKEY HSD: p>0.05 for all pairwise comparisons. **Image responses:** one-way ANOVA: F=1.114853, p=0.39; TUKEY HSD: p>0.05 for all pairwise comparisons.

SST, LM-LM correlations	**Omission responses**: one-way ANOVA: F=2.368675, p=0.06; TUKEY HSD: p<0.05 for the following pairwise comparisons: [LM depth, V1 depth] vs [LM depth, V1 depth]: [0 3] vs [3 3] **Image responses:** one-way ANOVA: F=1.901994, p=0.12; TUKEY HSD: p>0.05 for all pairwise comparisons.
VIP, V1-LM correlations	**Omission responses**: one-way ANOVA: F=0.47, p=0.95; TUKEY HSD: p>0.05 for all pairwise comparisons. **Image responses:** one-way ANOVA: F=0.58, p=0.88; TUKEY HSD: p>0.05 for all pairwise comparisons.
VIP, V1-V1 correlations	**Omission responses**: one-way ANOVA: F=0.22, p=0.99; TUKEY HSD: p>0.05 for all pairwise comparisons. **Image responses:** one-way ANOVA: F=1.67, p=0.13; TUKEY HSD: p<0.05 for the following pairwise comparisons: [LM depth, V1 depth] vs [LM depth, V1 depth]: [1 1] vs [3 3]
VIP, LM-LM correlations	**Omission responses**: one-way ANOVA: F=0.64, p=0.76; TUKEY HSD: p>0.05 for all pairwise comparisons. **Image responses:** one-way ANOVA: F=0.50, p=0.87; TUKEY HSD: p>0.05 for all pairwise comparisons.

Table 1. Statistical tests comparing correlation coefficients of neuronal activity among pairs of cortical planes. Depth 0 to 3 denote the 4 depths recorded in V1 and LM (Fig. 4b).

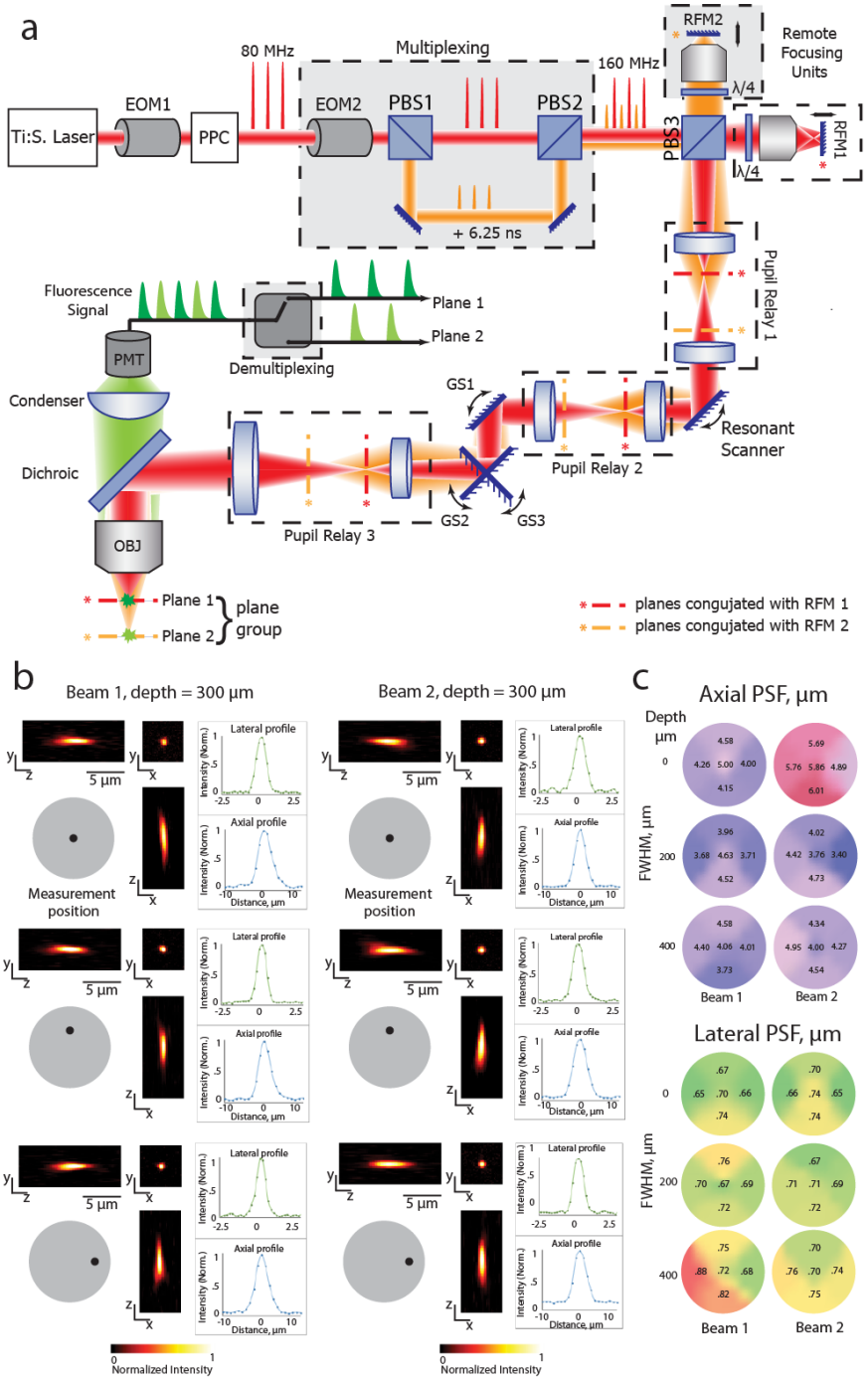

Figure 1. Dual Plane Mesoscope enables large-scale multi-plane, multi-area functional imaging. **a,** Instrument schematic for time-multiplexed excitation and detection. Beam power of Ti:Sapphire laser controlled by electro-optical modulator (EOM1). Multiplexing unit generates doubles laser pulse rate using polarizing beam splitters, orange arm delayed by 6.25 ns. EOM2 acts as dynamic waveplate, controlling power splitting ratio. Remote focusing mirrors RFM1 and RFM2 control axial positioning of focal planes by changing beams' collimation. Quarter-wave plate rotates beam polarizations by 90º at RFU output, steering beams towards scanners. **b.** Optical resolution of Multiplane Mesoscope. Profiles correspond to Point Spread Function (PSF) of 200 nm beads in agarose gel, imaged 300 µm deep at different FOV locations. **c.** PSF measurement in Z (top) and X-Y (bottom) for different FOV positions and depths. Colored circles represent full (5 mm) FOV, depth represents position from the surface, color represent optical resolution at different positions of the FoV.

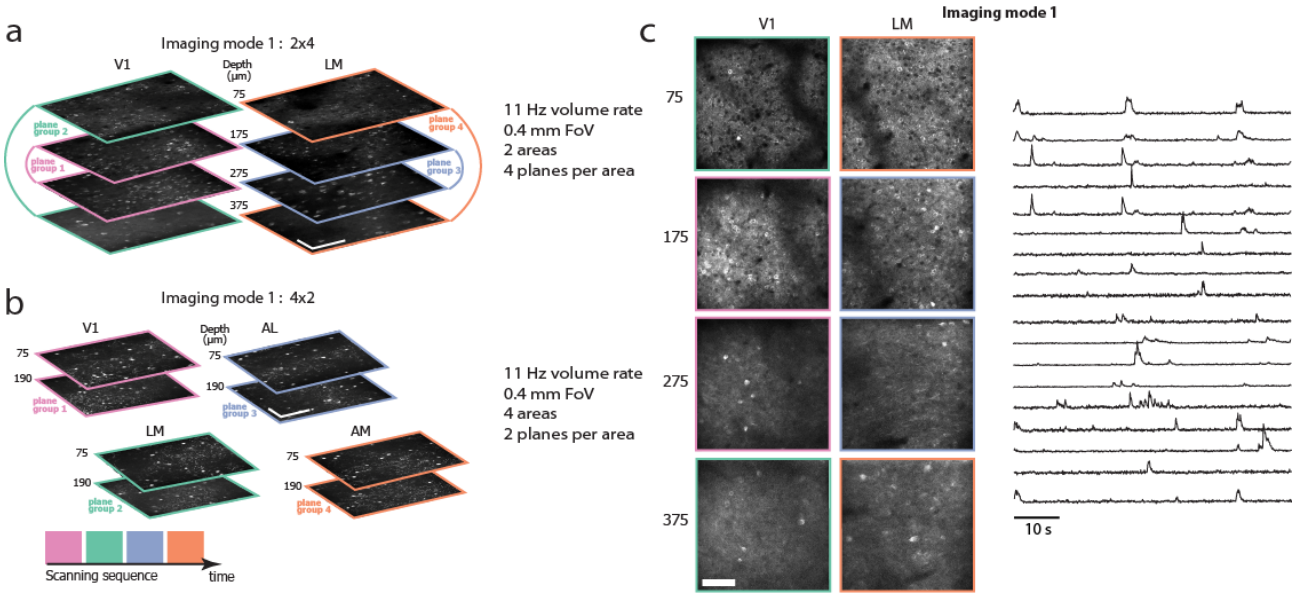

Figure 2. Dual Plane Mesoscope pipeline enables two complementary imaging modes. a. Imaging mode 1 samples two cortical areas with 4 planes in each area. Superficial planes are coupled with deeper planes to balance the laser power budget. **b.** Imaging mode 2 samples four cortical areas with a pair of axial planes in each. In both modes, pairs of planes are imaged simultaneously. **c.** Example two-photon images and set of calcium traces recorded in imaging mode 1 in V1 and LM in an awake Slc17a7-IRES2-Cre;Camk2a-tTA;Ai93 mouse visual cortex. Scale bars 100 µm.

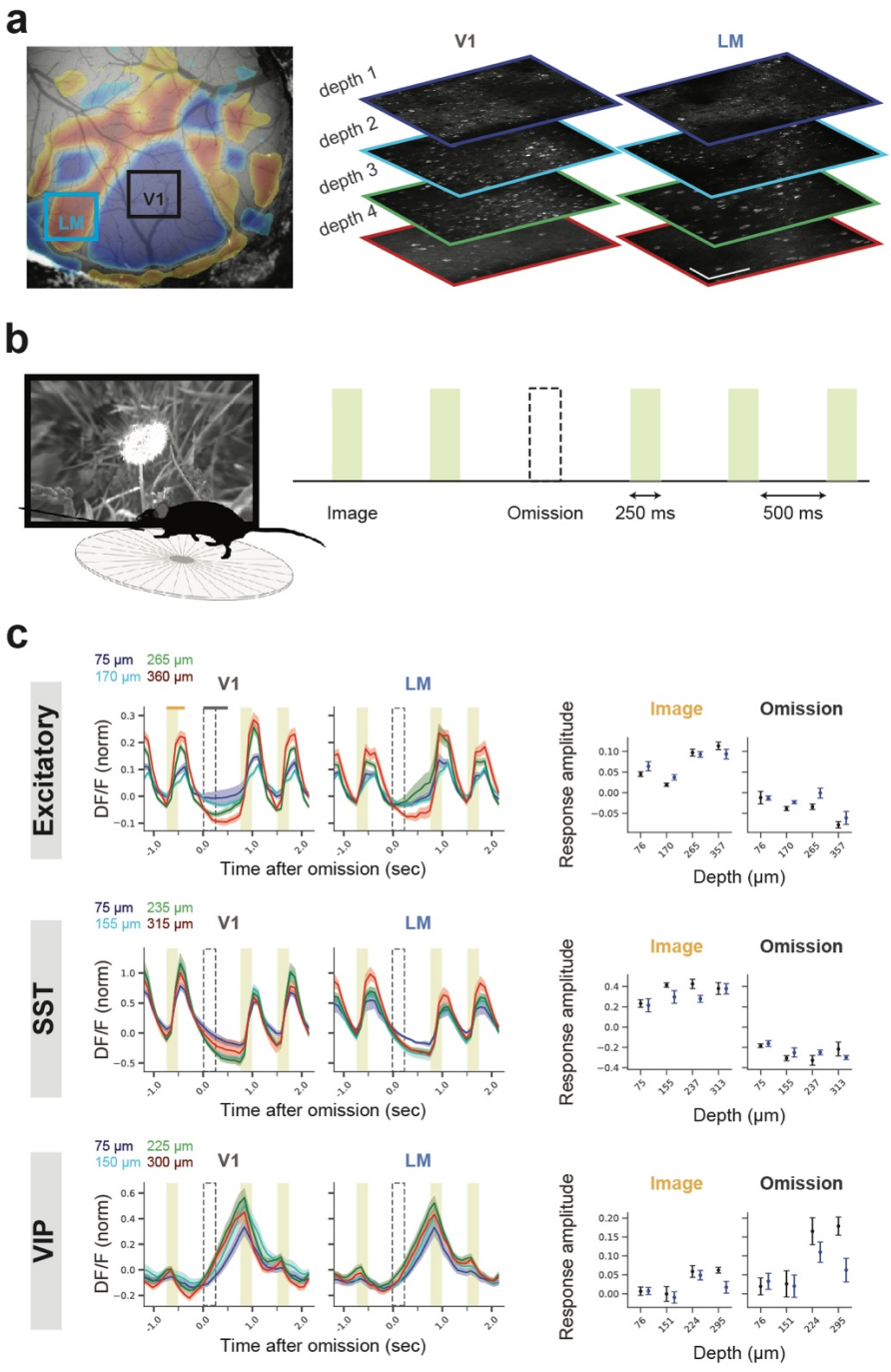

Figure 3. Excitatory and SST neurons in cortical columns of V1 and LM represent images, while VIP neurons represent omissions. a, Left: Intrinsic imaging of visual areas V1 and LM. **Right:** Simultaneous imaging of four cortical depths of V1 and LM during image-change detection task. Scale bar: 100 µm. **b, Left:** Images presented to head-fixed mice on a running disc. **Right:** Behavioral task included repeated natural image presentations ("images"; shaded rectangles), 5% of which were randomly omitted ("omissions"; dashed rectangle). **c, Left:** population-averaged calcium responses to images (shaded rectangles) and omissions (dashed rectangle) in 3 different mouse lines tagged for excitatory and two inhibitory subpopulations (mean +/- SEM; n = 24, 22, 24 sessions for excitatory, SST, and VIP, respectively). ΔF/F traces normalized to baseline standard deviation. **Right:** quantification of neural responses averaged over 350 ms after images, and 500 ms after omissions (time-windows used for quantification of image and omission-evoked responses are indicated by orange and gray horizontal lines above top left panel).

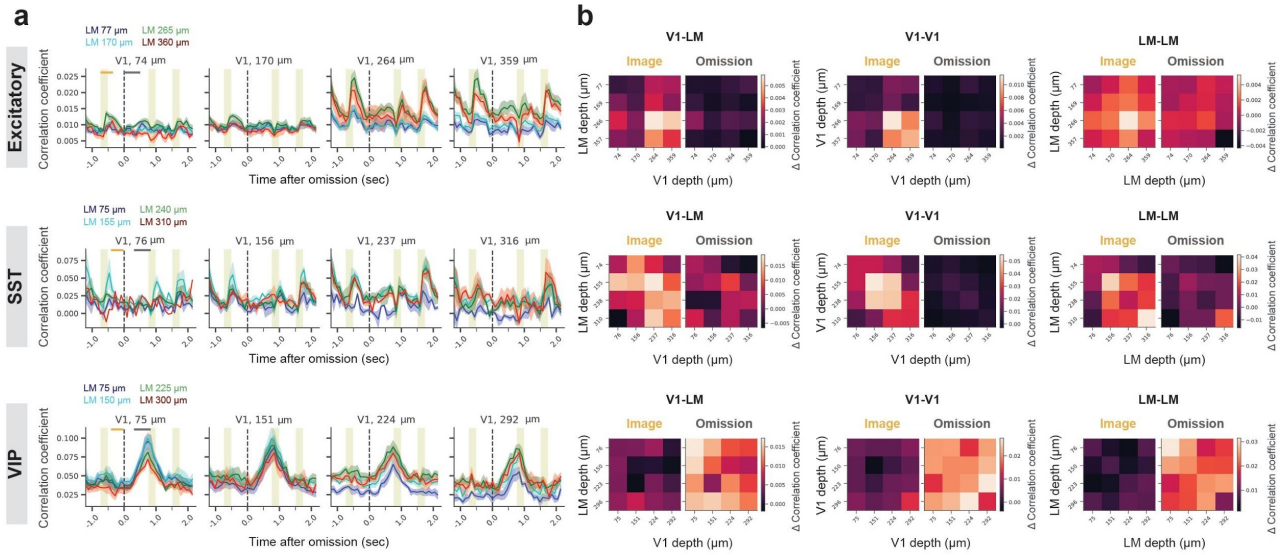

Figure 4. Distinct cortical interactions across visual areas following expected and unexpected events. a, Spearman correlation coefficients between V1 and LM neurons at different cortical depths (colors indicate LM depths; each subplot corresponds to a given V1 depth). **b,** change in correlation coefficients during images (left) and omissions (right) relative to baseline, shown for pairs of neurons across V1 and LM (left), within V1 (middle), or within LM (right). Correlation coefficients quantified over 500 ms after images, and 750 ms after omissions. Traces and heatmaps: mean +/- SEM; n = 8, 6, 9 mice for excitatory, SST, and VIP cell types, respectively.