Multiplane Mesoscope reveals distinct cortical interactions following expectation violations

Natalia Orlova[*,1], Dmitri Tsyboulski[*,1,2], Farzaneh Najafi[*,1], Sam Seid[1], Sara Kivikas[1], India Kato[1], Fiona Griffin[1], Arielle Leon[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 Groblewski[1], Jérôme Lecoq[#,1], and Peter Saggau[#,1,3,4]

* These authors contributed equally. # These senior authors contributed equally.

[1] Allen Institute for Brain Science, Seattle, 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. To monitor these interactions, we developed the Multiplane Mesoscope which combines three established microscopy technologies: time-division multiplexing, remote focusing, and random-access mesoscopy. The Multiplane Mesoscope allowed us to study cortical column interactions in excitatory and inhibitory subpopulations in behaving mice. We found that distinct cortical subnetworks represent expected and unexpected events, suggesting that expectation violations modify signal routing across cortical columns, and establishing the Multiplane Mesoscope as a unique platform to study signal routing.

Introduction

Monitoring single-neuron activity from spatially distributed neuronal populations is one of the grand challenges in neuroscience. External signals to the brain are indeed represented through distributed cortical networks[1], whose interactions are dynamic and depend on the context and predictability of those signals[2,3]. At the micro-circuit level, cortical interactions are regulated through distinct excitatory and inhibitory cell types[4]. Therefore, analysis of the activity of individual cell types across functionally connected cortical areas is necessary to achieve a mechanistic understanding of 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 *distributed networks* of neurons in brain volumes with this method has been difficult. This is because limited imaging speeds often restrict size and number of accessible regions of interest (ROIs), and typically prevent simultaneous monitoring of multiple cortical planes.

Mesoscopic imaging systems with a large field-of-view (FOV, 3 – 5 mm)[5–7] permit the investigation of information processing across multiple cortical areas in small animals. Such a large FOV two-photon random access mesoscope (2P-RAM) was recently developed[7] 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[8,9] to achieve fast and nearly aberration-free volumetric imaging. While such a system can randomly access a very large volume of the cortex[7], it remains limited in its imaging throughput, since a single laser beam is employed to sequentially scan the brain tissue.

In recent years, a variety of methods have been introduced to increase imaging throughput. Examples include imaging with Bessel beams[10,11], engineered point spread function (PSF)[12,13], targeted path galvanometer scanning[14], 3D random-access scanning with acousto-optic deflectors (AODs)[15], dual-axis two-photon imaging[16], and light sheet illumination[17]. Emerging techniques also include multiplexed TPLSM methods enabling simultaneous multi-site recordings with multiple excitation beams and either a single detector[6,18–21] or a dedicated detector associated with each laser beam[22].

In order to circumvent these limitations, we developed the Multiplane Mesoscope by combining mesoscopic random-access imaging and multi-beam TPLSM, which provides the necessary bandwidth to increase the number of simultaneously imaged areas while distributing them laterally and axially across many more interacting regions. A second imaging channel was added, using two orthogonally polarized and time-interleaved laser beams. Each laser beam was routed through a dedicated remote focusing unit (RFU) to achieve independent axial positioning of two focal planes. Temporal multiplexing of the laser pulse trains was used to encode each excitation beam and decode fluorescence signals by their arrival time at the detector. Thereby, we achieved 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.

Results

The 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 RFU, and a demultiplexing electronic circuit.

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 (PBS) 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 PBS combines both beams, creating two interleaved, temporally encoded excitation pulse trains. The input PBS 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[6] and more recently high-frequency sampling at rates near or above 1 GHz and digital unmixing with field-programmable gate arrays[13,21]. We developed a simple and efficient analog demultiplexing method which is compatible with standard data acquisition hardware commonly used for two-photon microscopy (see Online Methods and Supplementary Fig. 2). The performance of the Multiplane Mesoscope was compared to the original Mesoscope[23] by measuring inter-channel crosstalk, noise levels with and without signal demultiplexing, and point spread functions in both imaging channels across the imaging volume (see Online Methods and Supplementary Figs. 4-6).

We used this instrument to study cortical column interactions during active behavior (Fig. 2a,b). Since distinct cortical pathways are suggested to carry sensory vs. prediction signals[3,24], 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 established in previous studies[25,26]. 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. 2b). On go trials, a change in image identity occurred and mice received a water reward if they licked immediately 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. 2a; Supplementary Fig. 7).

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. 2c, top, middle). Excitatory neurons did not respond to omissions in any visual area (Fig. 2c, error-bars, top). SST neurons were slightly inhibited after omissions in all recorded locations (Fig. 2c, error-bars, middle). In sharp contrast to excitatory and SST neurons, VIP neurons, in all layers and areas, were robustly activated after omissions (Fig. 2c, bottom). They also demonstrated small anticipatory activity approximately 250 ms prior to each image presentation, and were inhibited immediately after the image (Fig. 2c, bottom), confirming previous results obtained with more conventional TPLSM imaging instruments using the same behavioral task[25].

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[27]. This increasing trend, however, was not observed in LM (Fig. 2c, top). SST responses to images were strongest at middle depths of V1, corresponding to layer 2/3 and superficial layer 4, but did not differ among cortical layers of LM (Fig. 2c, middle). VIP responses to neither images nor omissions were significantly different across cortical layers[25] or areas (Fig 2c, bottom).

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[2,3,28]. 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 owing to the multi-area, multi-plane capabilities of the Multiplane Mesoscope.

To study cortical interactions, we correlated neural responses across repeated image presentations, as well as omissions, for each cell type (pairwise "noise" correlations; Fig. 2d; Supplementary Fig. 8-10: shuffled control). We measured correlations within each area (V1-V1, LM-LM; Supplementary Fig. 8-10), as well as across areas (V1-LM; Fig. 2d). In the excitatory network, coactivation increased between V1 and LM neurons after image presentations, but only in deep layers (V1-LM: Fig. 2d, top), and also within V1 neurons in deep layers (V1-V1: Supplementary Fig. 8a,d). However, there was no significant change of coactivation within LM neurons after images, across any pair of cortical layers (LM-LM: Supplementary Fig. 8b,e). In the SST network, images increased coactivation among neurons more broadly across layers and areas, particularly in deep layer 2/3 and superficial layer 4 (Fig. 2d, middle; Supplementary Fig. 9). Omissions, on the other hand,

slightly decreased coactivation of neurons relative to the baseline level across most layers of the SST network (Fig. 2d, top, middle; Supplementary Fig. 8,9; see Fig 2d and Supplementary Fig. 8-10 captions for statistics).

In the VIP network, coactivation patterns were strikingly different compared to other cell types: neurons in *superficial* layers of V1 and LM were strongly coactivated following omissions. There was also a weak coactivation prior to images in most layers and areas (Fig. 2d, bottom; Supplementary Fig. 10). Notably, correlations were overall much weaker among excitatory neurons compared to inhibitory neurons, confirming previous results and indicating stronger local connectivity between inhibitory neurons[29,30].

Previous studies suggested that VIP neurons may carry a prediction signal[31]. Given their inputs from association areas and the neuromodulatory system, VIP neurons may also play a central role in gating sensory inputs[25,32,33]. Our results support that the VIP network is strongly recruited during unexpected events: cells from all layers are active during omissions and pairwise cell correlations are especially strong between superficial layers. This result is compatible with a shared input signal co-activating VIP cells in superficial layers across areas, thereby reorganizing the local cortical network following an expectation violation.

In summary, the Multiplane Mesoscope enables experiments requiring simultaneous imaging at different depths across multiple areas of the cortex. Importantly, we accomplished this without trading image quality for temporal resolution. Using this instrument, we demonstrated, for individual excitatory and inhibitory cell types, that interactions across cortical columns are distinct following expected and unexpected events. The exact role of each cell type and how they communicate to encode, and update prediction signals will require further studies. As a central piece of the Allen Brain Observatory two-photon data pipeline[34], the Multiplane Mesoscope will uncover new aspects of cortical computation, and will be used to generate open datasets in service of the neuroscience community. In particular, we anticipate that simultaneous recordings from distinct cell types using the Multiplane Mesoscope will give new insight into the circuit mechanisms that underlie predictive coding and inter-areal communication.

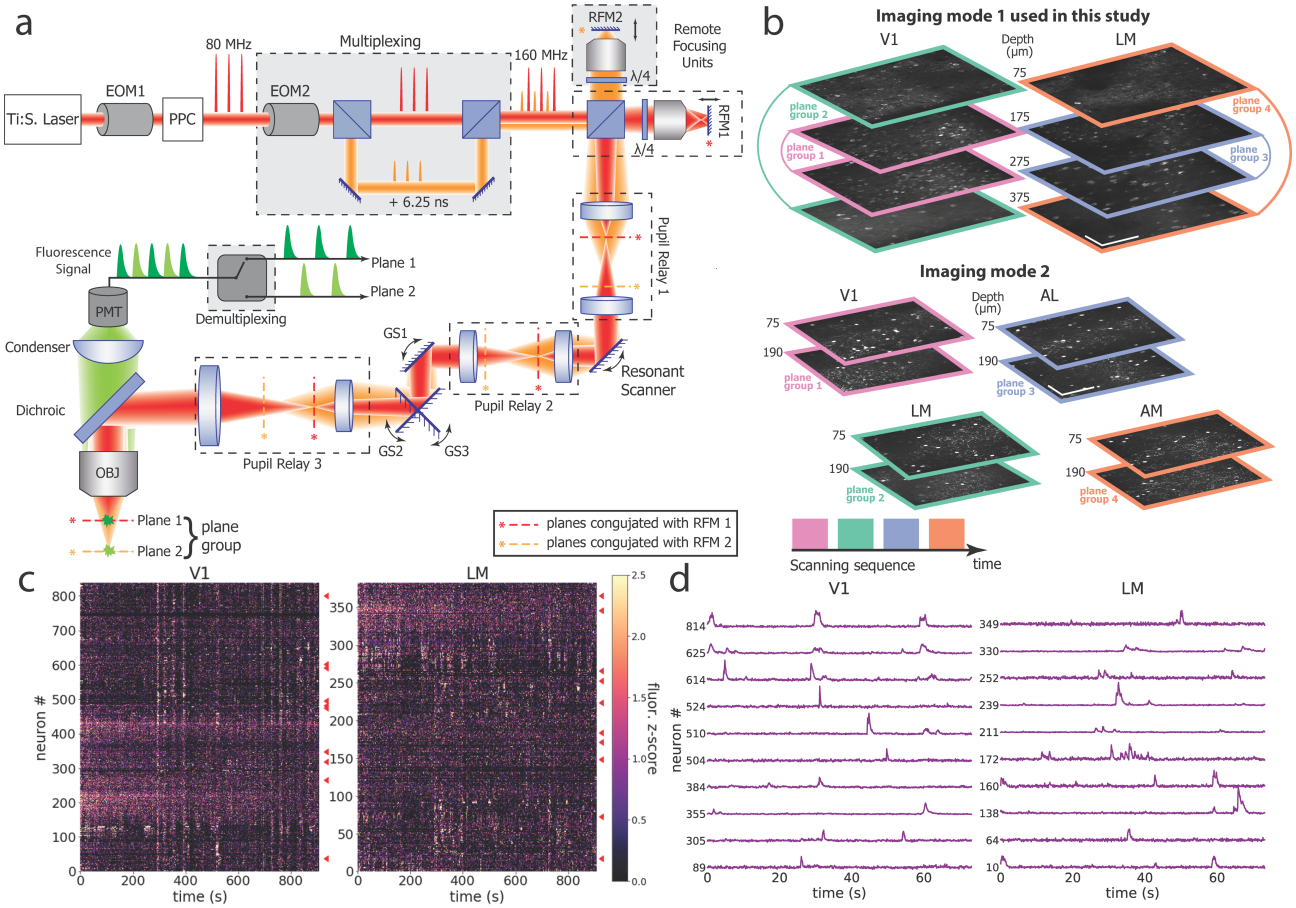

Figure 1: Dual Plane Mesoscope enables large-scale multi-plane, multi-area functional imaging. a, Instrument schematic of time-multiplexed excitation and emission. Beam power of Ti:Sapphire laser controlled by electro-optical modulator (EOM1). Multiplexing unit generates two beams using polarizing beam splitter (PBS), orange arm delayed by 6.25 ns (half a period of laser repetition rate). EOM2 acts as a dynamic waveplate, controlling power splitting ratio. Recombined beams are sent to remote focusing units (RFUs). 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. Both beams pass same pupil relays and XY-scanning blocks. **b, Top**: In vivo imaging of two cortical columns. Simultaneous imaging of 8 planes (~75 µm, ~175 µm, ~275 µm, and ~375 µm depth) in V1 and LM of a Slc17a7;Ai93 mouse. Frames 8 x time-averaged. **Bottom:** In vivo imaging of 4 cortical areas. Vip-IRES-Cre;Ai148 mouse visual cortex imaged in V1, LM, AL, and AM at two depths: 75 µm and 190 µm . In (b): Frame rate 11 Hz, FOV 400 x 400 µm^2, to match standard size of Allen Brain Observatory datasets. LUT adjusted to optimize contrast at different depths. Scale bar: 100 µm. **c,** Calcium signals across two cortical columns in V1 and LM. Slc17a7-IRES2-Cre;Camk2a-tTA;Ai93 mouse visual cortex imaged at four depths: ~100 µm, ~200 µm, ~300 µm, and ~375 µm. Analysed cells: 1150 (V1: 800, LM: 350). Many cells show correlations similar to single-plane recordings. **d,** Example Ca traces from 10 randomly-selected neurons in (c) marked with red arrow heads.

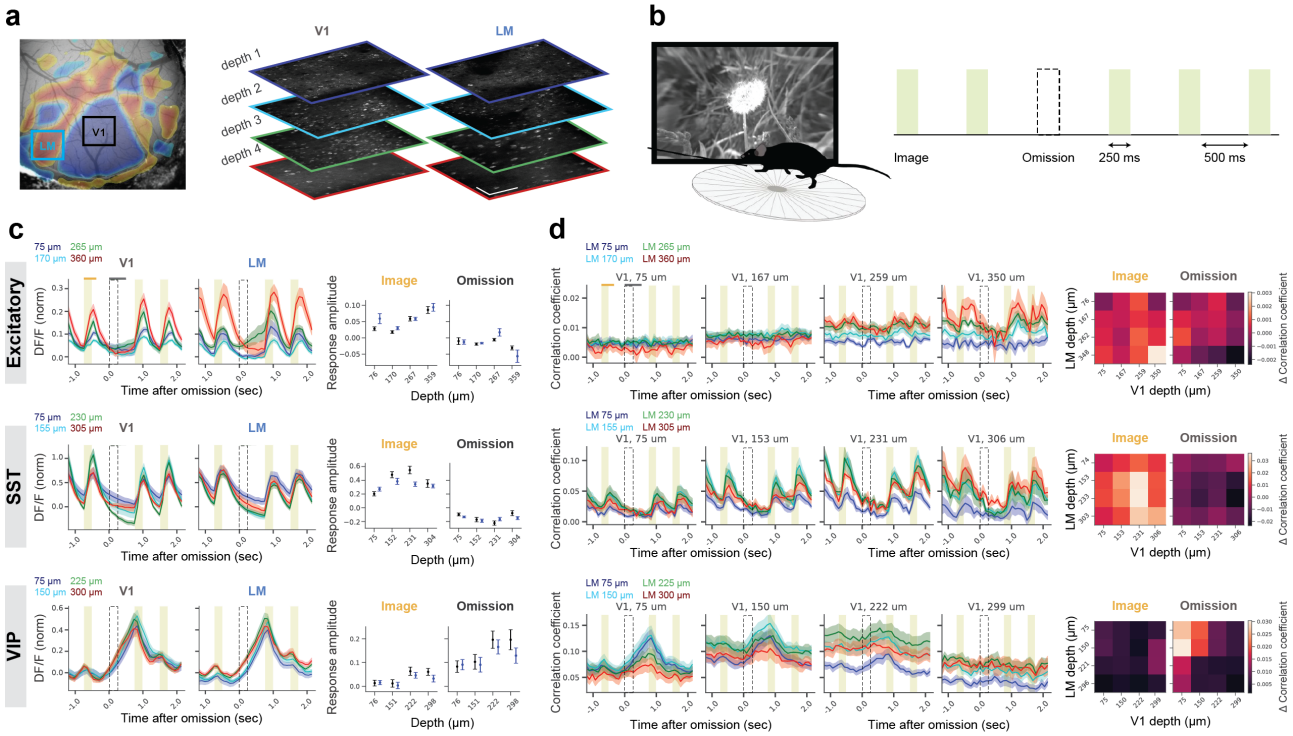

Figure 2. Distinct cortical interactions occur across primary and higher-order visual areas following expected and unexpected events. a, Left: Intrinsic imaging of visual areas V1 and LM. **Right:** Simultaneous imaging of four cortical depths of V1 and LM during image-change detection task. Scale bar: 100 µm. **b, Left:** Images presented to head-fixed mice on a running disc. **Right:** Behavioral task included repeated natural images (shaded rectangle) including 5% random omissions (dashed rectangle). **c, Left:** population-averaged calcium responses to images (shaded rectangle) 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 evoked 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). **d, Left:** Spearman correlation coefficients between V1 and LM neurons at different cortical depths (colors indicate LM depths; each subplot corresponds to a given V1 depth). **Right:** change in correlation coefficients during images (left) and omissions (right) relative to baseline correlation coefficient. 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.

METHODS

System design

The design and operation of a two-photon fluorescence Mesoscope was described previously[7]. Our modification added three custom modules to the original system, highlighted in Fig. 1a.

The first custom module controls power distribution between imaging planes and encodes excitation beams with a temporal delay. The beam from an 80 MHz femtosecond laser (Chameleon Ultra-II, Coherent), emitting at 910 nm with a pulse duration of ~120 fs and an output power of 2.5 W, passes through an EOM (350-80-LA-02, Conoptics Inc.), a prism pulse compressor (PPC), and a second EOM with the internal polarizer removed. Next, the laser beam is split in two at the polarizing beamsplitter (PBS), one beam is delayed relative to the other by 6.25 ns through free-space propagation, and the orthogonally polarized beams are recombined at the second PBS and directed towards the Mesoscope's periscopic input. EOM1 controls the total input power and EOM2 rotates the beam's polarization and thus defines the power splitting ratio of the laser beam at PBS.

The second custom module, designed in SolidWorks (Dassault Systèmes), includes an additional remote focusing (RF) unit (Fig. 1a) and enables the conversion of an original Mesoscope to a Multiplane Mesoscope (Supplementary Fig. 1). SolidWorks design files of the module are available upon request. The incoming orthogonally-polarized beams separate at the PBS and pass through $\lambda/4$ wave plates towards RF objectives and movable mirrors (RFM1, RFM2), the reflected beams, rotated by 90° polarization, are recombined at the PBS and reflected towards the optical scanners. The dual remote focusing assembly independently controls the axial positions of two focal planes and, by design, provides spherical aberration compensation at different beam defocus values associated with RF mirror positions.

The third added module is the custom demultiplexing circuit we developed to separate time-interleaved fluorescence signals from different imaging planes. A synchronization signal from the laser is passed through an 80 ± 2 MHz bandpass filter (3016, KR Electronics, Inc.). A comparator (LTC6957-HMS4, Analog Devices) converts the resulting waveform into two complimentary square wave signals with a duty cycle of ~50%. The signals are scaled to a peak-to-peak amplitude of 5 V by high-bandwidth amplifiers (GVA-83+, MiniCircuits), and shifted towards negative voltages from a common mode of 0 V to -1V – -1.5V with bias tees (ZFBT-4R2GW+, MiniCircuits). These control signals toggle the outputs of a high-frequency (HF) switch (CMD196C3, Custom MMIC), whose inputs are terminated with 100 Ohm resistors. Note that the HF switch control requires negative voltages, and excessive positive voltages may damage the circuit. Supplementary Fig. 2. shows a schematic of the demultiplexing circuit, control signals, and a demonstration of the circuit operation.

Fluorescence signal from the photomultiplier (PMT) (H11706-40, Hamamatsu) was amplified with a 400 MHz bandwidth amplifier (HCA-400M-5K-C, Femto) matched to the PMT bandwidth, and fed to the HF switch input. The length of the PMT lead cable was shortened to ~4 cm to reduce crosstalk between channels. An adjustable coaxial delay line (DB64, Stanford Research) was used

to align fluorescence signals and the HF switch control signals in time. Demultiplexed signals were passed through 50 MHz low-pass filters (BLP-50+, Mini-Circuits) and recorded with a digitizer (NI-5732 adapter module and PXIe-7962R FlexRIO module, National Instruments) at a sampling rate of 80 MHz. The system's impulse response and averaged time-resolved fluorescence signal from GCaMP6f-labeled brain tissue slices are shown in Supplementary Fig. 3. Note that fluorescence tails extend into the next 6.25 ns signal integration window, resulting in crosstalk between imaging channels.

System characterization

Time-resolved electric signals were recorded with a fast oscilloscope (204MXi-A, LeCroy) at sampling rates up to 10 GS/s. Custom MATLAB (MathWorks) routines were used to analyze the recorded data.

Point spread function (PSF) measurements were performed by acquiring and analyzing image stacks of 200 nm-beads embedded in 4% aqueous agarose gel over the 5 mm FOV at depths between 0 µm and 500 µm. Examples of measured PSFs across the imaging volume in both channels are shown in Supplementary Fig. 4. We observed nearly identical PSFs in both imaging channels, as previously reported[7].

To evaluate the influence of demultiplexing electronics on imaging quality, signal amplitudes and noise levels of GCaMP6f-labled brain tissue slices recorded with and without demultiplexing electronics were compared (Supplementary Fig. 5). We found our demultiplexing method losses ~24% of the signal intensity, part of it is attributable to the manufacturer-specified HF switch insertion loss. Noise levels measured as variances of pixel intensities in acquired images as a function of average pixel intensity were found identical.

Laser power throughput was measured using a thermal power meter in crucial locations along the optical path (Supplementary Table 1), revealing a 2% power loss in the delayed beam compared to the main beam. This difference is explained by the additional mirrors used in the delay line. Overall system throughput was 16% for non-delayed beam and 14% for the delayed one. This difference can be compensated by steering more laser power into the delayed beam, see Supplementary Table 1.

Crosstalk characterization and removal

Crosstalk is a fundamental limitation in multiplexed microscopy systems. The first 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, two beams generate a total pulse rate of 160 MHz which defines the width of the signal integration window, i.e., 6.25 ns. The detected fluorescence signal is then demultiplexed using a temporal window of ≤ 6.25 ns in the demultiplexing circuit. Second 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[35,36].

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 temporal delay while monitoring excitatory neurons in the mouse brain. 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 crosstalk of ~ 15%. Supplementary Fig 4a shows averaged images of such a test, where each row is a set of images corresponding to the plane of interest (excitation ON), coupled plane (excitation OFF), raw calcium traces of selected cells, unmixed traces of selected cells and 2D histograms of signal and crosstalk traces before and after crosstalk removal. We defined the crosstalk as the slope of the linear regression between signal and ghost trace (Supplementary Fig. 4b) and excluded from further analysis cells that were segmented purely due to crosstalk. Those neurons were called ghost cells and identified based on two factors; 1) the crosstalk value had to be >1, which corresponded to the case when the signal is stronger in the coupled plane, 2) there were no independent calcium events left in the fluorescence trace after crosstalk removal.

Crosstalk removal was performed on extracted traces using an Independent Component Analysis (ICA)-based approach, where independent components are estimated by minimizing Gaussianity of the data[37]. The code is available upon request.

Animal head-implants and cortical window implantation

Surgical headpost and cranial window implantation was performed as described elsewhere[25,34,38]. Headpost and cranial window surgery was performed on healthy mice that ranged in age from p35 to p60, weighing no less than 15 g at time of surgery. Dexamethasone (3.2 mg/kg, S.C.) injections were administered at 12 hr and 3 hr before surgery. Mice were first anesthetized with 5% isoflurane (1–3 min) and placed in a stereotaxic frame (Model #1900, Kopf, Tujunga, CA), and isoflurane levels were maintained at 1.5–2.5% for surgery. An incision was made to remove skin, and the exposed skull was levelled with respect to pitch (bregma-lambda level), roll and yaw. The stereotaxic frame was zeroed at lambda using a custom headframe holder equipped with stylus affixed to a clamp-plate. The stylus was then replaced with the headframe to center the headframe well at 2.8 mm lateral and 1.3 mm anterior to lambda. The headframe was affixed to the skull with white Metabond and once dried, the mouse was placed in a custom clamp to position the skull at a rotated angle of 23° such that visual cortex was horizontal to facilitate the craniotomy. A circular piece of skull 5 mm in diameter was removed, and a durotomy was performed. A coverslip stack (two 5 mm and one 7 mm glass coverslips adhered together) was cemented in place with Vetbond. Metabond cement was applied around the cranial window inside the well to secure the glass window. Post-surgical brain health was documented using a custom photo-documentation system and at one, two, and seven days following surgery, animals were assessed for overall health (bright, alert, and responsive), cranial window clarity, and brain health. After a 1–2 week recovery from surgery, animals underwent intrinsic signal imaging for retinotopic mapping, and then entered into behavioral training. Custom tools and implants used in the surgery were designed at the Allen Institute and are described in detail elsewhere[38].

Intrinsic imaging and mapping of the visual cortex

Retinotopic mapping was used to delineate functionally defined visual area boundaries and to target the in vivo two-photon calcium imaging to retinotopically defined locations in primary and secondary visual areas. Target maps were created from the maps of eccentricity at the centre of V1 by restricting to values of eccentricity that are within 10° from the origin and was limited to retinotopic values that are negative for both altitude and azimuth. This targeting map was overlaid on an image of the surface vasculature to provide fiducial landmarks to guide optical physiology recording sessions and to ensure that the imaged locations were retinotopically matched across areas.

Mice were lightly anesthetized with 1-1.4% isoflurane administered with a 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 Lubricant Eye Ointment; Refresh) were applied to maintain hydration and clarity of eyes during anaesthesia. 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 set of LED light sources: green (peak λ = 527 nm; FWHM = 50 nm; Cree Inc., C503B-GCN-CY0C0791) and red (peak λ = 635 nm, FWHM = 20 nm; Avago Technologies, HLMP-EG08-Y2000. All LEDs were mounted directly on the imaging lens creating a ring light. A pair of camera lenses (Nikkor 105mm f/2.8, Nikkor 35mm f/1.4, Nikon), provided 3.0x magnification (M = 105/35) onto a sCMOS camera (Andor Zyla 5.5 10tap). A bandpass filter (Semrock; FF01-630/92 nm) was used to only record reflected red light from the brain.

The visual stimulus displayed consisted of a drifting bar containing a checkerboard pattern, alternating black and white as it sweeps along a grey background. The stimulus bar sweeps across 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 square sizes measuring at 25°. The stimulus was warped spatially so that a spherical representation could be displayed on a flat monitor. To ensure maximal coverage of the field of view, 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 that the stimulus was perpendicular to the optic axis of the eye.

Visual stimulation

Visual stimuli were generated using custom scripts written in PsychoPy (Peirce, 2007, 2008) and were displayed using an ASUS PA248Q LCD monitor, with 1920 x 1200 pixels. The 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.

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.

Behavioural training

Details of training are explained in a previous study[26]. In brief, mice were trained 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, full-field, static square-wave gratings were presented. Mice received automatic water rewards once the orientation of the grating switched between 0° and 90°. In the sessions thereafter, mice were rewarded if they licked within a 750 ms time window after the identity of the visual stimulus changed. In the 1st stage, static gratings were presented as in 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 d-prime 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 d-prime >2 and mice performed at least 100 trials.

Image changes happened randomly, according to a geometric distribution (p=0.3) on go trials, 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 2a). Sessions lasted for 60 min.

Processing of calcium imaging movies

The preprocessing of all calcium imaging data was done within the imaging processing pipeline described in detail elsewhere[34] and crosstalk removal was applied after pre-processing. Briefly, all data was first corrected for brain motion by performing rigid registration in two dimensions. Then, cell-segmentation was performed to identify spatial masks of active neurons. Further, fluorescence from spatially-overlapping neuronal masks was unmixed and corrected for neuropil contamination. Finally, mask-matching and crosstalk removal was performed using FastICA[37] and ghost cells were filtered, and lastly, $\Delta F/F$ was computed on corrected masks.

Description of datasets

3 mouse lines were used in the 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, 24 sessions from 8 mice; for the SST cell line, 22 sessions from 6 mice; for the VIP cell line, 24 sessions from 9 mice were recorded.

Population averages of neuronal responses

Traces were aligned to the onset of image omission; since images were presented regularly, omission alignment also aligned the traces on images. 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. 2b, traces). 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 [-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. 2c, Error-bars).

Correlation of neural responses across cortical planes

On omission-aligned traces, Spearman correlation coefficient was computed between pairs of neurons across trials. Here, by trial we mean 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. 2d). 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 (Supplementary Fig. 8-10 d-f), 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. 2d, heatmaps; Supplementary Fig. 8-10d-f), 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". This quantity is plotted in Fig. 2d, heatmaps, and Supplementary Fig. 8-10d-f, error-bars.

The Python package scipy (scipy.stats.spearmanr) was used for computing correlations, and p-values were computed in 2 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 available from the corresponding authors upon reasonable request.

Code availability

Crosstalk removal was performed using custom routines employing FastICA which is available as part of scikit-learn Python package (https://scikit-learn.org/).

Acknowledgements

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 Paulo Chaves at 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 Christof Koch and Cliff Kerr for comments on the manuscript. We thank David Sullivan for help with designing a custom behavioral stage and Ross Hytnen for help with acquisition software.

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												■	■		
RH												■			
SC														■	
SK						■		■						■	
SL												■			
SRO											■				
SS						■		■						■	
XW														■	

Competing interests

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.

Correspondence and requests for materials should be addressed to N.O. or J.L, and for biological data analysis to F.N. or J.L.

References

1. Steinmetz, N. A., Zatka-Haas, P., Carandini, M. & Harris, K. D. Distributed coding of choice, action and engagement across the mouse brain. *Nature* (2019). doi:10.1038/s41586-019-1787-x

2. Marques, T., Nguyen, J., Fioreze, G. & Petreanu, L. The functional organization of cortical feedback inputs to primary visual cortex. *Nat. Neurosci.* (2018). doi:10.1038/s41593-018-0135-z

3. Keller, G. B. & Mrsic-Flogel, T. D. Predictive Processing: A Canonical Cortical Computation. *Neuron* (2018). doi:10.1016/j.neuron.2018.10.003

4. Gouwens, N. W. *et al.* Classification of electrophysiological and morphological neuron types in the mouse visual cortex. *Nat. Neurosci.* (2019). doi:10.1038/s41593-019-0417-0

5. Tsai, P. S. *et al.* Ultra-large field-of-view two-photon microscopy. *Opt. Express* (2015). doi:10.1364/oe.23.013833

6. Stirman, J. N., Smith, I. T., Kudenov, M. W. & Smith, S. L. Wide field-of-view, multi-region, two-photon imaging of neuronal activity in the mammalian brain. *Nat. Biotechnol.* **34**, 857–862 (2016).

7. Sofroniew, N. J., Flickinger, D., King, J. & Svoboda, K. A large field of view two-photon mesoscope with subcellular resolution for in vivo imaging. *Elife* **5**, (2016).

8. Botcherby, E. J., Juskaitis, R., Booth, M. J. & Wilson, T. Aberration-free optical refocusing in high numerical aperture microscopy. *Opt. Lett.* **32**, 2007 (2007).

9. Botcherby, E. J., Juškaitis, R., Booth, M. J. & Wilson, T. An optical technique for remote focusing in microscopy. *Opt. Commun.* (2008). doi:10.1016/j.optcom.2007.10.007

10. Lu, R. *et al.* Video-rate volumetric functional imaging of the brain at synaptic resolution. *Nat. Neurosci.* **20**, 620–628 (2017).

11. Lu, R. *et al.* Rapid mesoscale volumetric imaging of neural activity with synaptic resolution. *Nat. Methods* (2020). doi:10.1038/s41592-020-0760-9

12. Prevedel, R. *et al.* Fast volumetric calcium imaging across multiple cortical layers using sculpted light. *Nat. Methods* (2016). doi:10.1038/nmeth.4040

13. Weisenburger, S. *et al.* Volumetric Ca2+ Imaging in the Mouse Brain Using Hybrid Multiplexed Sculpted Light Microscopy. *Cell* (2019). doi:10.1016/j.cell.2019.03.011

14. Botcherby, E. J. *et al.* Aberration-free three-dimensional multiphoton imaging of neuronal activity at kHz rates. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 2919–2924 (2012).

15. Duemani Reddy, G., Kelleher, K., Fink, R. & Saggau, P. Three-dimensional random access multiphoton microscopy for functional imaging of neuronal activity. *Nat. Neurosci.* (2008). doi:10.1038/nn.2116

16. Lecoq, J. *et al.* Visualizing mammalian brain area interactions by dual-axis two-photon

calcium imaging. *Nat. Neurosci.* **17**, 1825–1829 (2014).

17. Ahrens, M. B., Orger, M. B., Robson, D. N., Li, J. M. & Keller, P. J. Whole-brain functional imaging at cellular resolution using light-sheet microscopy. *Nat. Methods* **10**, 413–420 (2013).

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

19. Chen, J. L., Voigt, F. F., Javadzadeh, M., Krueppel, R. & Helmchen, F. Long-range population dynamics of anatomically defined neocortical networks. *Elife* (2016). doi:10.7554/eLife.14679

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

21. Beaulieu, D. R., Davison, I. G., Kılıç, K., Bifano, T. G. & Mertz, J. Simultaneous multiplane imaging with reverberation two-photon microscopy. *Nat. Methods* (2020). doi:10.1038/s41592-019-0728-9

22. Rumyantsev, O. I. *et al.* Fundamental bounds on the fidelity of sensory cortical coding. *Nature* (2020). doi:10.1038/s41586-020-2130-2

23. Sofroniew, N. J., Flickinger, D., King, J. & Svoboda, K. A large field of view two-photon mesoscope with subcellular resolution for in vivo imaging. *Elife* (2016). doi:10.7554/eLife.14472

24. Bastos, A. M. *et al.* Canonical Microcircuits for Predictive Coding. *Neuron* **76**, 695–711 (2012).

25. Garrett, M. *et al.* Experience shapes activity dynamics and stimulus coding of VIP inhibitory cells. *Elife* (2020). doi:10.7554/eLife.50340

26. Groblewski, P. A. *et al.* Characterization of Learning, Motivation, and Visual Perception in Five Transgenic Mouse Lines Expressing GCaMP in Distinct Cell Populations. *Front. Behav. Neurosci.* (2020). doi:10.3389/fnbeh.2020.00104

27. Niell, C. M. & Stryker, M. P. Highly selective receptive fields in mouse visual cortex. *J. Neurosci.* (2008). doi:10.1523/JNEUROSCI.0623-08.2008

28. Yang, W. *et al.* Simultaneous Multi-plane Imaging of Neural Circuits. *Neuron* **89**, 269 (2016).

29. Najafi, F. *et al.* Excitatory and Inhibitory Subnetworks Are Equally Selective during Decision-Making and Emerge Simultaneously during Learning. *Neuron* (2020). doi:10.1016/j.neuron.2019.09.045

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

doi:10.1523/JNEUROSCI.3131-11.2011

31. Krabbe, S. *et al.* Adaptive disinhibitory gating by VIP interneurons permits associative learning. *Nat. Neurosci.* (2019). doi:10.1038/s41593-019-0508-y

32. Kuchibhotla, K. V. *et al.* Parallel processing by cortical inhibition enables context-dependent behavior. *Nat. Neurosci.* (2017). doi:10.1038/nn.4436

33. Fu, Y. *et al.* A cortical circuit for gain control by behavioral state. *Cell* (2014). doi:10.1016/j.cell.2014.01.050

34. de Vries, S. E. J. *et al.* A large-scale standardized physiological survey reveals functional organization of the mouse visual cortex. *Nat. Neurosci.* (2020). doi:10.1038/s41593-019-0550-9

35. Akerboom, J. *et al.* Optimization of a GCaMP calcium indicator for neural activity imaging. *J. Neurosci.* (2012). doi:10.1523/JNEUROSCI.2601-12.2012

36. Pliss, A., Zhao, L., Ohulchanskyy, T. Y., Qu, J. & Prasad, P. N. Fluorescence lifetime of fluorescent proteins as an intracellular environment probe sensing the cell cycle progression. *ACS Chem. Biol.* (2012). doi:10.1021/cb300065w

37. Hyvärinen, A. Fast and robust fixed-point algorithms for independent component analysis. *IEEE Trans. Neural Networks* (1999). doi:10.1109/72.761722

38. Groblewski, P. & Sullivan, D. A standardized head-fixation system for performing large-scale, in-vivo physiological recordings in mice. (2020). doi:10.1101/2020.01.22.916007