iScience, Volume *24*

Supplemental Information

Optical clearing of living brains with MAGICAL to extend *in vivo* imaging

Kouichirou Iijima, Takuto Oshima, Ryosuke Kawakami, and Tomomi Nemoto

Supplemental Figures and Legends:

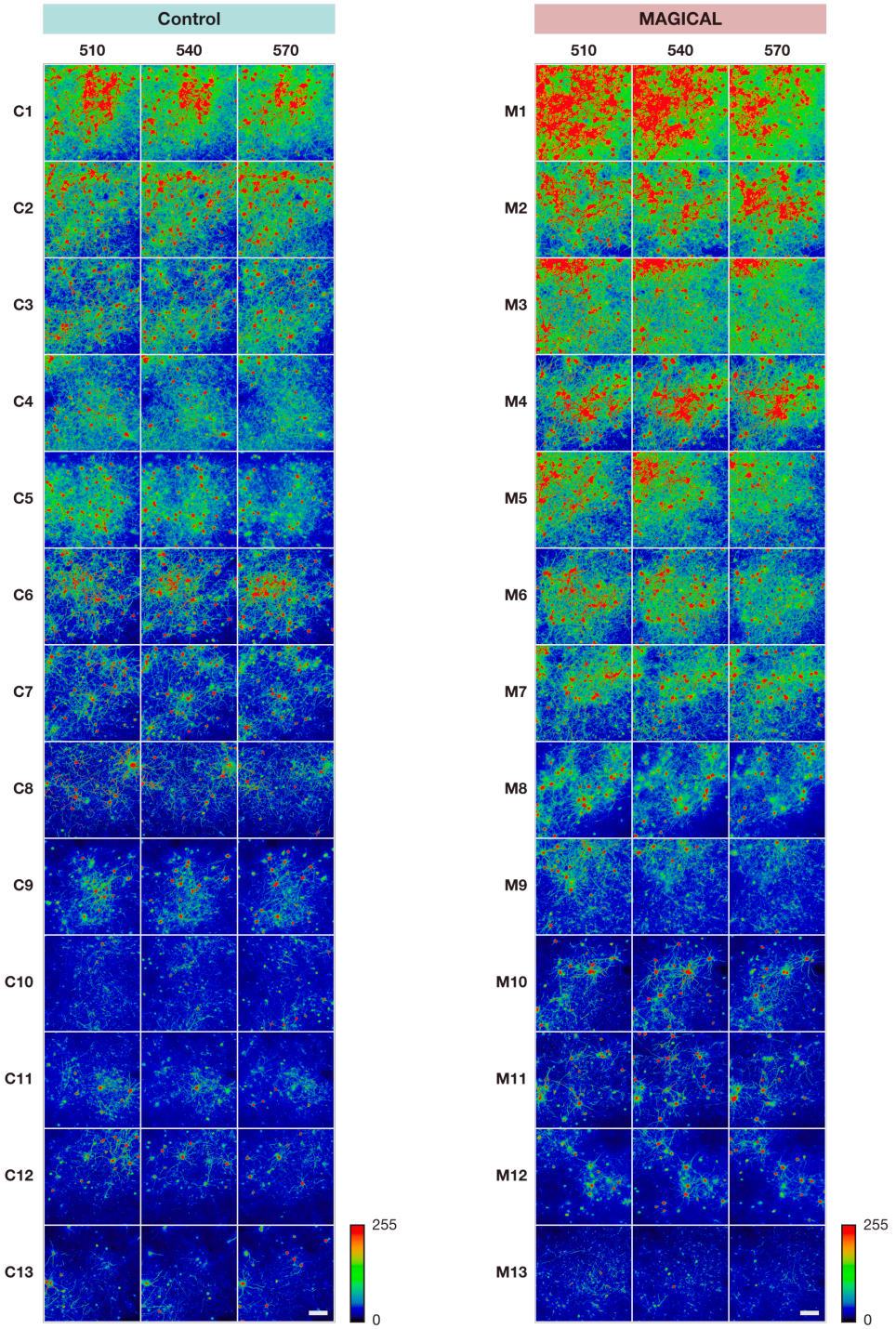

Figure S1. *In vivo* two-photon images at CxLV to evaluate fluorescence intensity. Related to Figure 1. All images were captured under the 25× Ob. with LP 66 mW and HV 30, as 3D stacks at indicated depth ± 3 μm with a 3-μm z-step size (3 sections). These 3D stacks were analyzed as FIDs (**Figure S2**) and FIPs for the GLMM (**Figure 1B–C**). All 3D stacks are displayed as MIP images, ranked by brightness scores (average fluorescence intensity per pixel) at 540 ± 12 μm depth (9 sections) captured with LP 66 mW and HV 30. scale bar 100 μm.

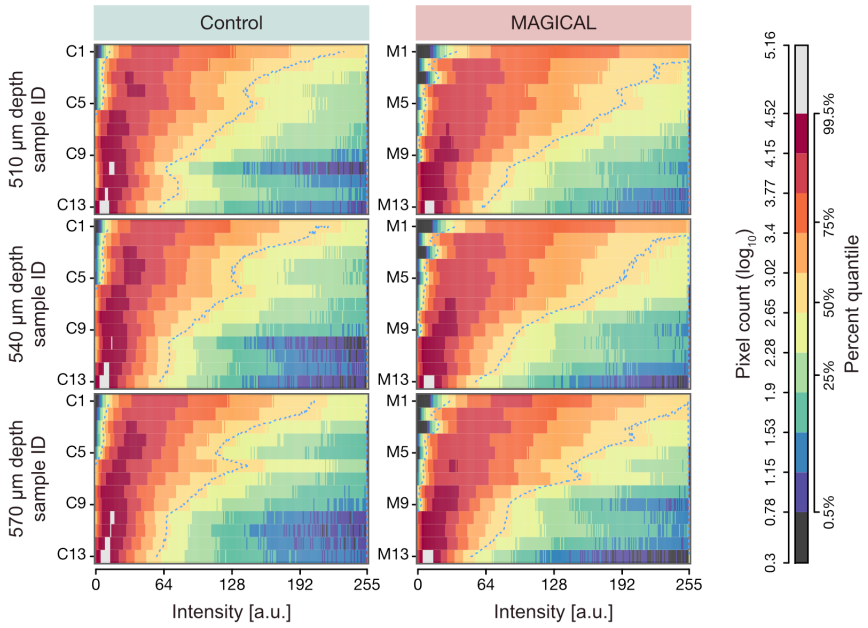

Figure S2. Heat map representation for FID of CxLV image stacks in Figure S1. Related to Figure 1. The heat map rows represent image stacks ranked by brightness scores at 540 ± 12 µm depth within each experimental group. The heat map columns represent intensity bins (8 bits). Pixel counts (z) in the FID histogram are converted with $\log_{10}(z+1)$ and assigned a color code. Pixel counts (excluding zero) were ranked in percent quantile. Cyan dotted contours are median. a.u., arbitrary unit.

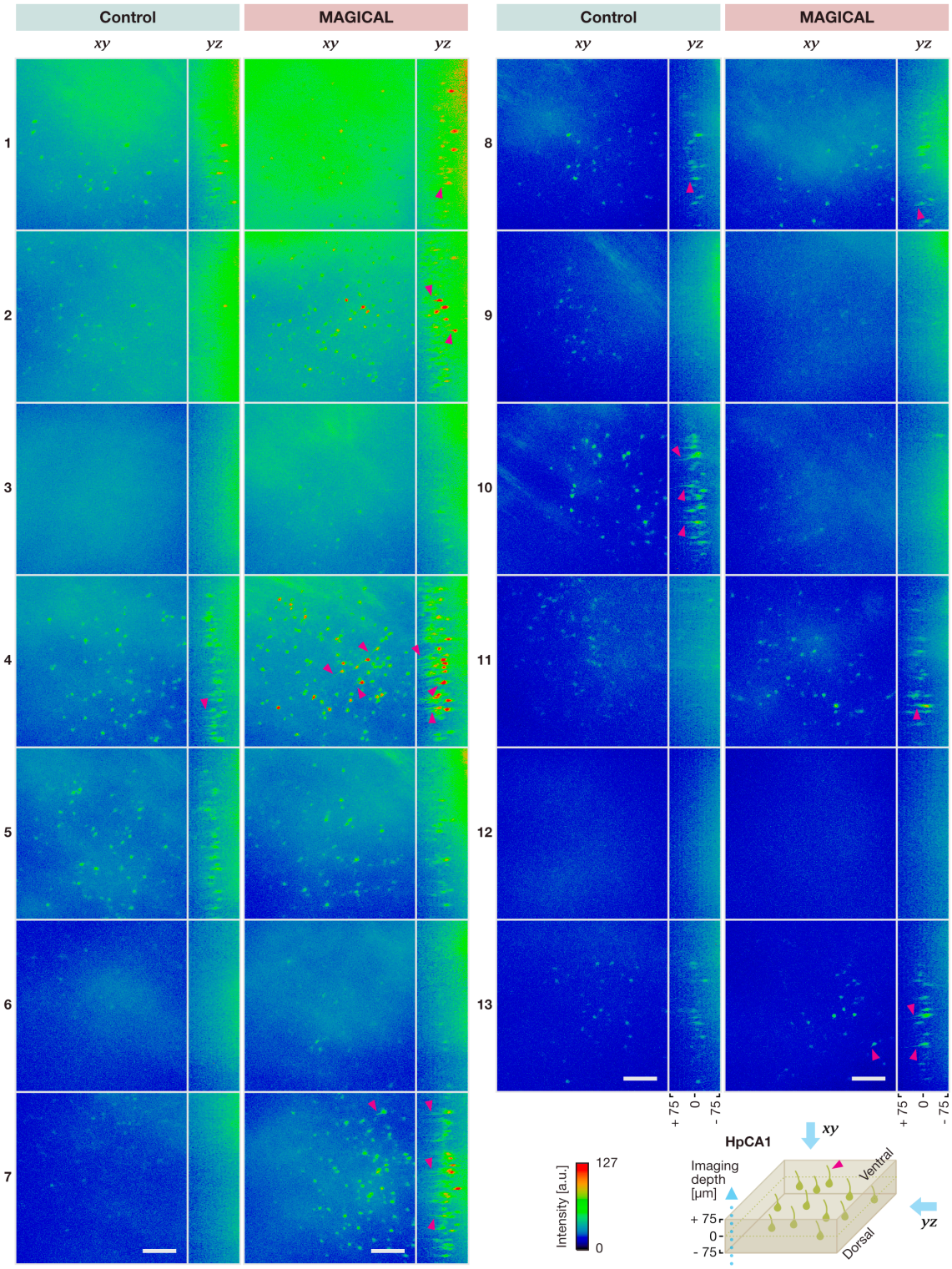

Figure S3. *In vivo* two-photon 3D rendering images at HpCA1. Related to Figure 2.
All images were captured under the 25× Ob. with LP 200 mW and HV 30, as 3D stacks at the middle of stratum pyramidale (0 μm relative depth) ± 75 μm with a 3-μm z-step size (51 sections). The 3D rendering was performed by using the Fiji's 3D project function with the brightest point mode and 50 % interior depth-cueing. Xy projection images were created from a ventral viewpoint (opposite side of observation), to avoid halation on the alveus (fibrous cloud structures) at the dorsal side. Yz projection images were created with rendering interpolation. These images are ranked by the brightness scores at 0 (center) ± 12 μm depth (9 sections). MAGICAL visualized many tiny apical dendrites (arrow-head) of pyramidal cells, although it is dependent on the signal to the background ratio. scale bar 100 μm. a.u., arbitrary unit.

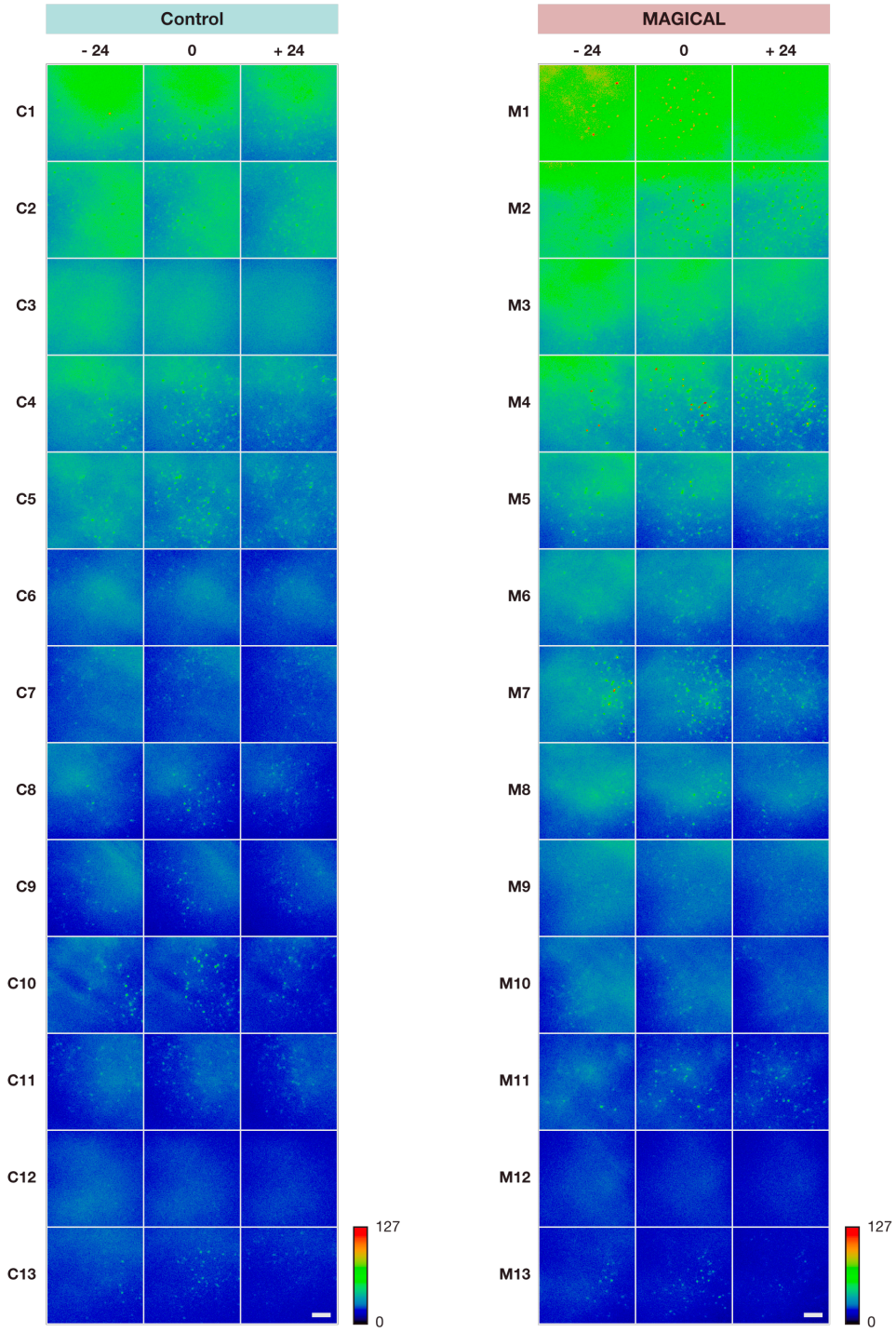

Figure S4. *In vivo* two-photon images at HpCA1 to evaluate fluorescence intensity. Related to Figure 2. All images were captured under the 25× Ob. with LP 200 mW and HV 30, as 3D stacks at indicated relative depth ± 3 μm with a 3-μm z-step size (3 sections). These 3D stacks were analyzed as FIDs (**Figure S5**) and FIPs for the GLMM (**Figure 2B–C**). All 3D stacks are displayed as MIP images, ranked by brightness scores at 0 (center) ± 12 μm depth (9 sections) captured with LP 200 mW and HV 30. scale bar 100 μm.

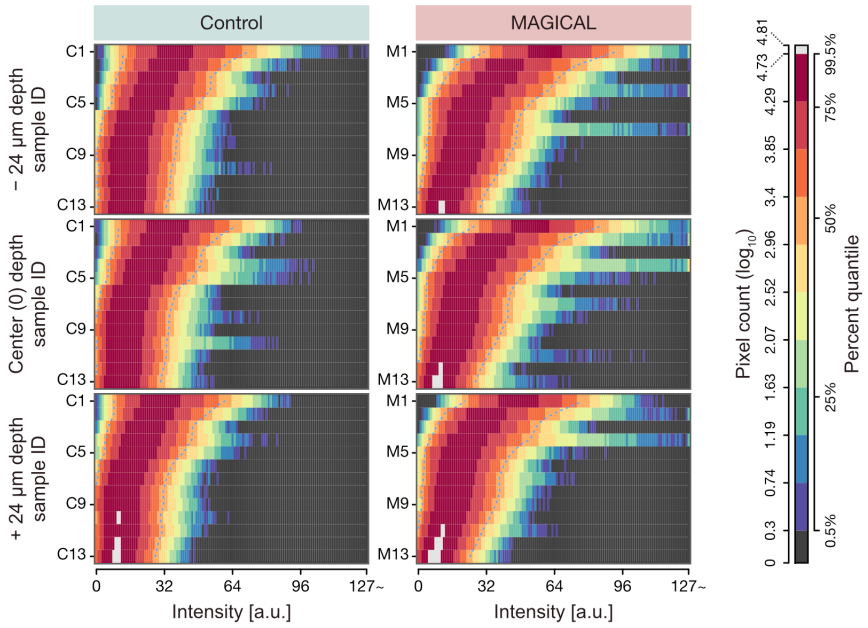

Figure S5. Heat map representation for FID of HpCA1 image stacks in Figure S4. Related to Figure 2. The heat map rows represent image stacks ranked by brightness at the center ± 12 µm depth within each experimental group. The heat map columns represent intensity bins (8 bits). Pixel counts (z) in the FID histogram are converted with $\log_{10}(z+1)$ and assigned a color code. Pixel counts (excluding zero) were ranked in percent quantile. Cyan dotted contours are median. a.u., arbitrary unit.

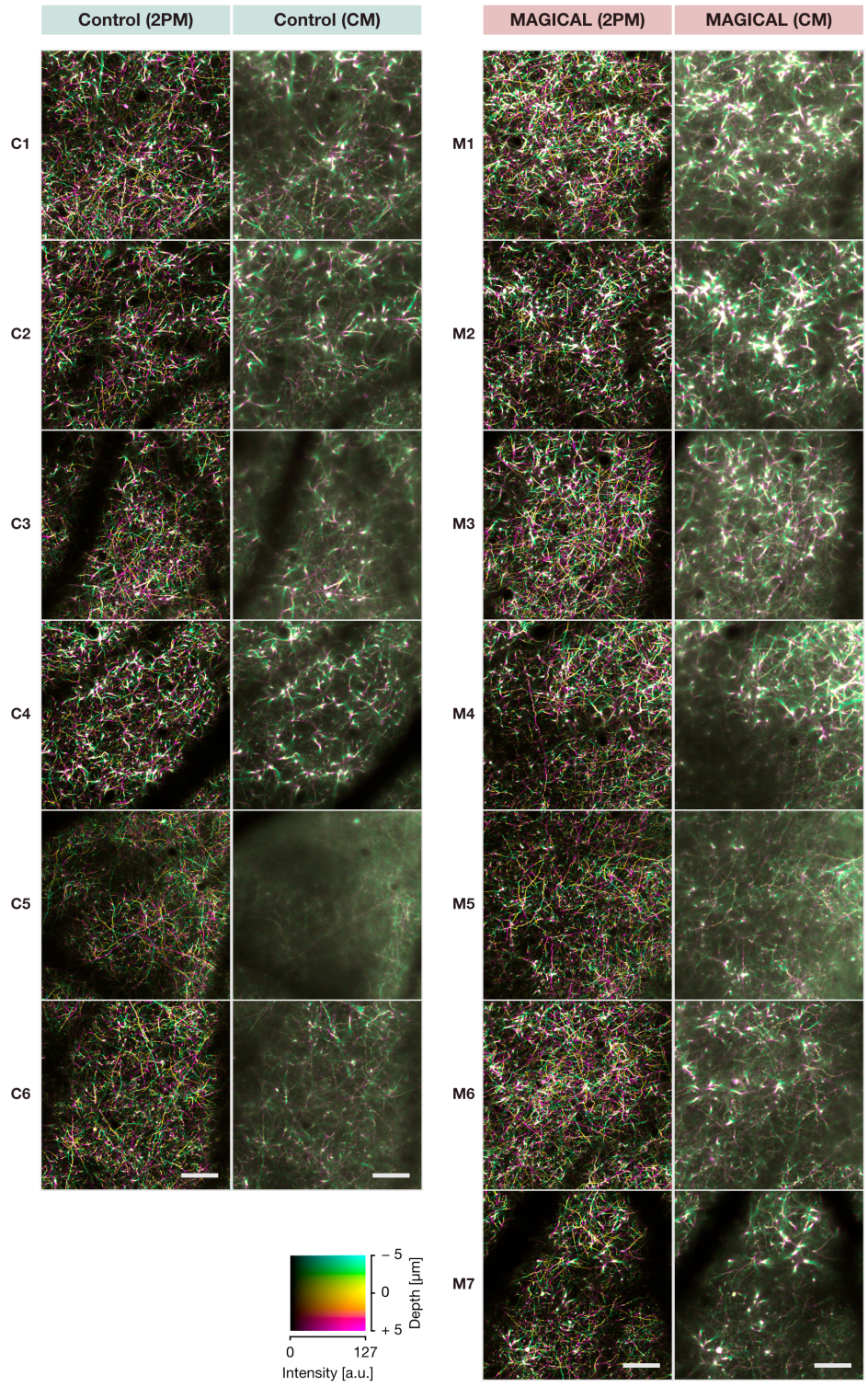

Figure S6. *In vivo* confocal images at CxLI. Related to Figure 3.
All images were captured as 3D stacks at 100 µm (center, 0 µm relative depth) ± 5 µm with a 1-µm z-step size (11 sections) under the 25× Ob., with LP 10 % and HV 90 in confocal microscopy and LP 16.1 mW and HV 30 in two-photon microscopy. All 3D stacks are displayed as DccMIP images, ranked by brightness scores in the confocal stacks. Intermediate images (C3 and M4) are shown as the representative images in **Figure 3A**. scale bar 100 µm. a.u., arbitrary unit.

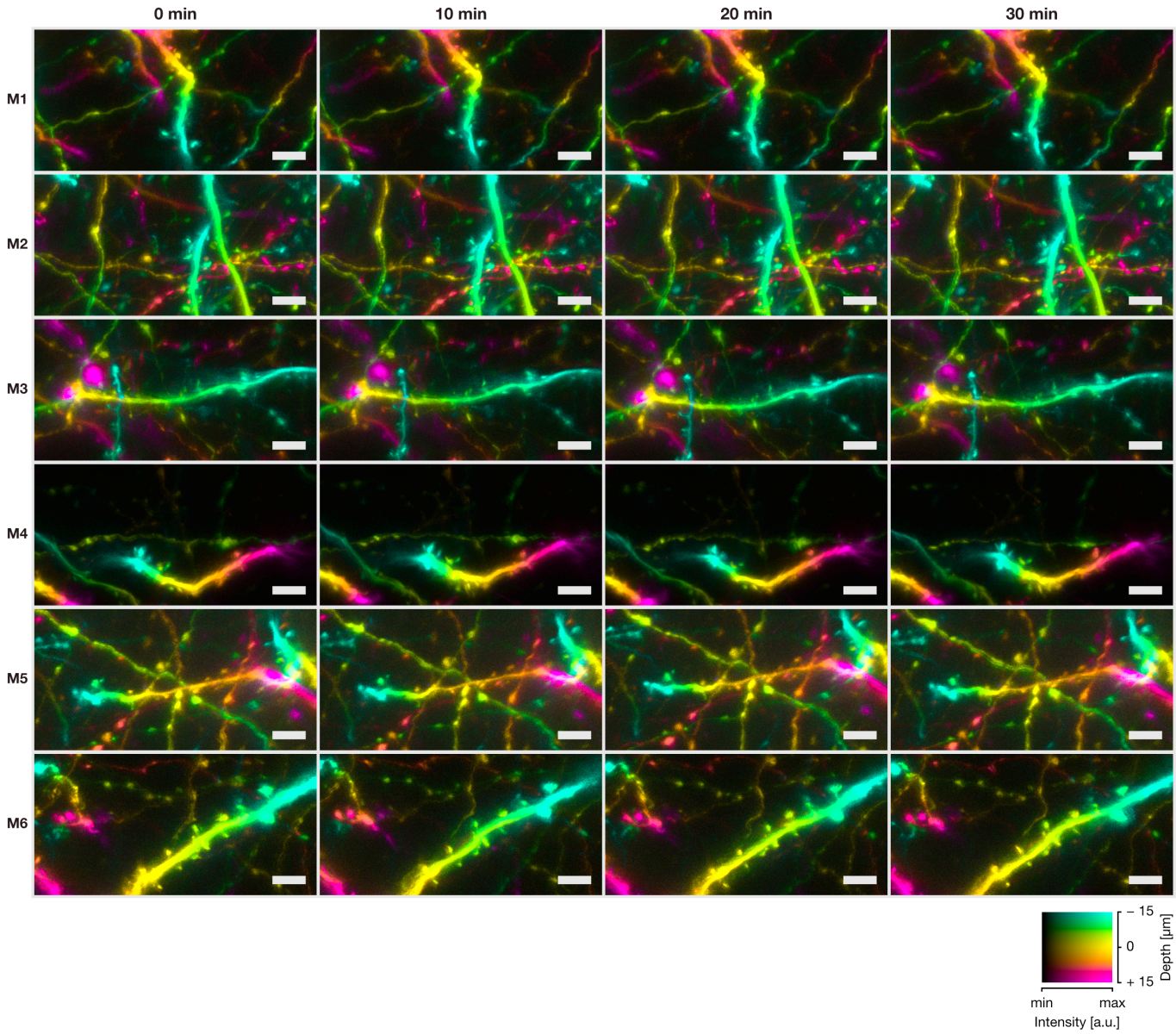

Figure S7. *In vivo* confocal high-resolution 4D images at CxLI. Related to Figure 4.
In 6 mice with MAGICAL, the 4D imaging was performed as 3D stacks with a 1-µm z-step size at approximate 100 ± 15 µm depth (31 sections, displayed as DccMIP), as a time-series at 5 minutes intervals for 30 minutes (7 time-points, but images at 5, 15, and 25 min were not shown). Laser power and detector sensitivity were set between 10 and 20 % and between HV 90 and HV 100, respectively. In the color code, each center depth is shown as 0 µm relative depth. The 3D images of M1 are shown in **Figure 4**. scale bar 5 µm. a.u., arbitrary unit.

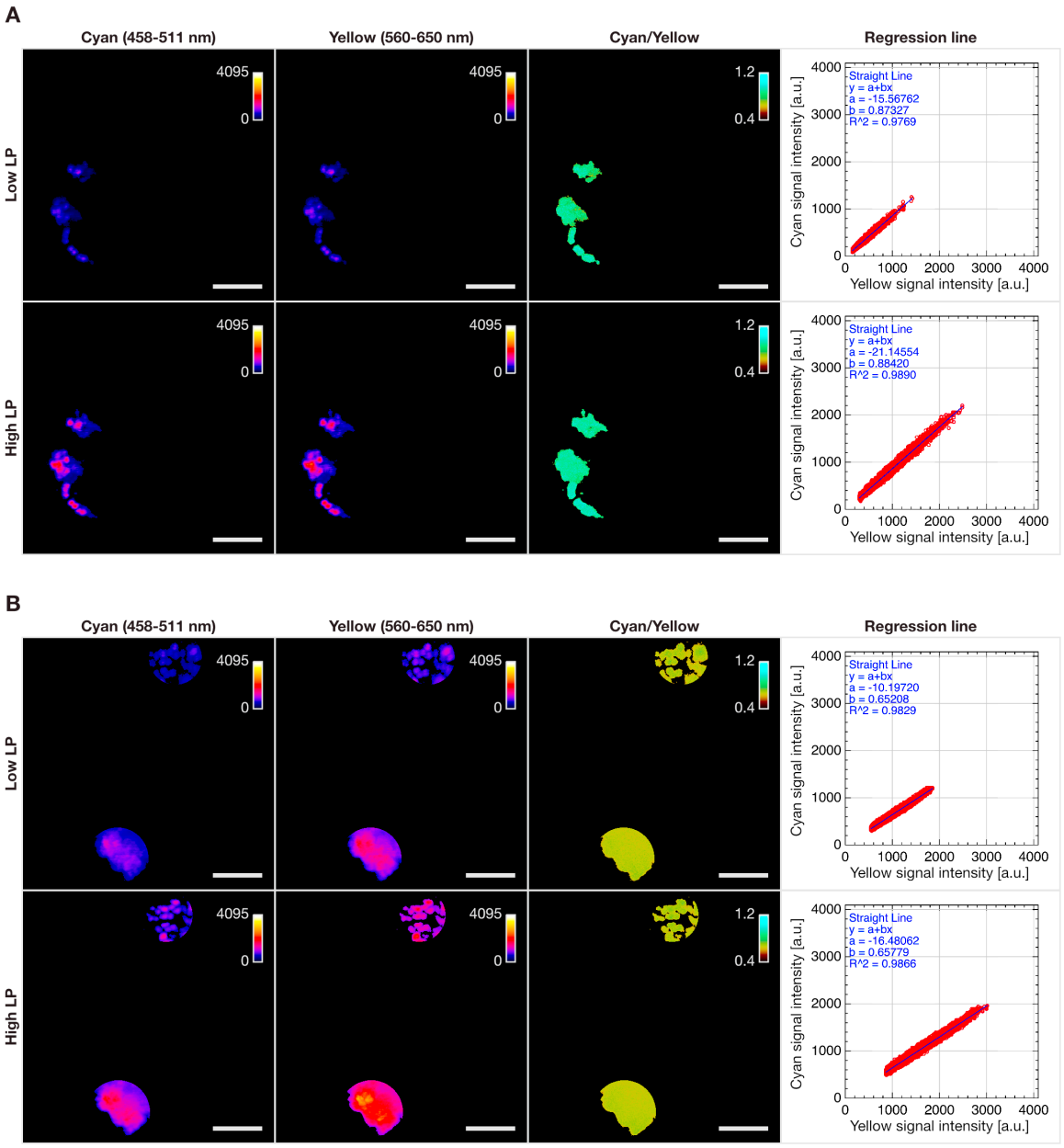

Figure S8. Cyan/Yellow ratio calculation for FESTA. Related to Figure 5.
(**A–B**) MIP images of beads clusters extracted via ROI, and respective regression lines at 100 μm (**A**) and 600 μm (**B**) depth. Although fluorescence intensity in each channel was affected by LP for two-photon excitation and the amount of fluorescence probe dependent on accumulation style of round-shaped beads, the Cyan/Yellow ratio was not affected. In contrast, the Cyan/Yellow ratio decreased with depth. scale bar 10 μm. a.u., arbitrary unit.

Table S1. The linear predictors in the GLMM. Related to Figure 1 and 2.

A

Fixed Effects :

Coefficient	Estimate	95 % CI lower	95 % CI upper	Std. Error	t value	Pr (> |z|)
(Intercept)	3.595	3.522	3.668	0.037	96.92	< 0.001
Depth_540	-0.024	-0.049	0.000	0.013	-1.92	0.0548
Depth_570	-0.044	-0.080	-0.009	0.018	-2.43	0.0153
Treatment_MAGICAL	0.259	0.164	0.355	0.049	5.33	< 0.001
Depth_540 : Treatment_MAGICAL	0.010	-0.024	0.044	0.017	0.57	0.5710
Depth_570 : Treatment_MAGICAL	-0.032	-0.080	0.016	0.024	-1.30	0.1928

Random Effects :

Group	Name	Variance	Std.Dev.
Mouse	(Intercept)	0.205	0.452
	Depth_540	0.002	0.050
	Depth_570	0.006	0.077
Residual		0.649	0.806

Number of observations: 929331, groups: Mouse, 26

B

Fixed Effects :

Coefficient	Estimate	95 % CI lower	95 % CI upper	Std. Error	t value	Pr (> |z|)
(Intercept)	3.044	2.973	3.115	0.036	84.31	< 0.001
Depth_000	-0.114	-0.122	-0.105	0.004	-27.01	< 0.001
Depth_+24	-0.230	-0.244	-0.217	0.007	-33.43	< 0.001
Treatment_MAGICAL	0.160	0.086	0.233	0.037	4.26	< 0.001
Depth_000 : Treatment_MAGICAL	-0.008	-0.020	0.004	0.006	-1.35	0.177
Depth_+24 : Treatment_MAGICAL	-0.011	-0.030	0.008	0.010	-1.15	0.250

Random Effects :

Group	Name	Variance	Std.Dev.
Mouse	(Intercept)	0.025	0.158
	Depth_000	0.000	0.008
	Depth_+24	0.000	0.013
Residual		0.159	0.398

Number of observations: 957264, groups: Mouse, 26

(**A–B**) Statistical results for the GLMMs with a Gamma error distribution and a log link function at CxLV (**A**) and HpCA1 (**B**). Wald-type 95 % CI was estimated for the coefficient of fixed effects by using the confint function.

Transparent Methods:

No statistical methods were used to predetermine sample sizes. In Generalized Linear Mixed Models (GLMMs), we used large sample sizes generally enough for modeling. For Welch's *t*-test, we confirmed that the *t*-test had adequate power, with a post-hoc test having effect size (Cohen's *d*), *p*-value (< 0.01), and the sample size.
The experiments were not randomized. *In vivo* imaging must observe individual mice convergently, not parallelly, for a long time. To reduce influence dependent on specific timing, the observation order for each experimental group should be uniformly mixed. However, the sample size is insufficient for randomization to ensure to equalize the order. Thus, we allocated mice to experimental groups so as not to localize the order.
Investigators were not blinded to allocation during experiments and data analysis. To remove observer bias during experiments, the observation area was selected by previously defined criteria, as described in "**Transparent** Methods, 'Fluorescence intensity analysis' section". Data collection and analysis were automatically performed by software with fixed parameter settings or macro programs.

Animals. All experiments were conducted in accordance with the Animal Research Committee of the Hokkaido University. The protocols were approved by the Committee on the Ethics of Animal Experiments in the Hokkaido University (No. 10-0119). Adult transgenic mice (9 to 12 weeks old, male) from Thy1-EYFP-H (H-line) were used for *in vivo* imaging of neurons. Adult C57BL/6NCr Slc mice (10 to 22 weeks old, male; Japan SLC, Shizuoka, Japan) were used for the FESTA with injected fluorescent beads. All mice were housed with food and drink *ad libitum* and were maintained in a 12 h light-dark cycle (lights on from 8:00 to 20:00), with controlled temperature (22 °C to 26 °C) and humidity (40 % to 60 %).

Glycerol administration. For MAGICAL mice, 5 % (w/v) glycerol (075-00616; Wako Pure Chemical Industries, Osaka, Japan) was administered orally in the drinking water, *ad libitum*, from 2 weeks prior to open-skull surgery to the end of the experiment, except during operations and observations. Control mice received only water.

Open-skull surgery. A cranial window was surgically created overlying the left cortex, according to a previous protocol (Holtmaat et al., 2009; Kawakami et al., 2013), with minor modifications. Briefly, 45 mg/kg-bw minocycline (Nichi-Iko Pharmaceutical, Toyama, Japan) were injected intraperitoneally (i.p.) about 4 h before surgery, to suppress bacterial infection and to protect neurons against microglial activation (neuroinflammation). Anesthesia was induced with an i.p. injection of 60 mg/kg-bw pentobarbital sodium (Somnopentyl; Kyoritsu Seiyaku Corporation, Tokyo, Japan), and was maintained with isoflurane inhalation (0.5 % to 1.5 %). To limit inflammation, 2 mg/kg-bw dexamethasone (Kyoritsu Seiyaku Corporation) was injected intramuscularly. Cranial bones were exposed, and then the left parietal bone was circularly carved about 4.2 mm in diameter with a dental drill and was removed gently. The exposed dura was washed with phosphate-buffered saline (PBS) to remove blood cells. A 4.2-mm diameter round coverslip (Micro Cover GLASS, #1S, about 0.17 mm thickness; Matsunami Glass Industry, Osaka, Japan), previously coated with a biocompatible polymer— Lipidure® (CM5206E; NOF Corporation, Tokyo, Japan), to prevent foreign-body reactions, especially to blood clotting— was placed directly over the dura without extra pressure, and was sealed on the bone edges with cyanoacrylate cement and dental adhesive resin cement (Super-Bond C&B; Sun Medical Company, Shiga, Japan). To make the foundations for chamber fixation, exposed cranial bones around the cranial window were partially coated with the adhesive resin cement. A head chamber, which has a center hole to expose the cranial window, was secured to the cranial bones with dental acrylic resin cement (Unifast III; GC Corporation, Tokyo, Japan). 5 mg/kg-bw carprofen (Rimadyl; Zoetis, Parsippany, NJ, USA) was injected i.p. to reduce inflammation and pain. After surgery, the mouse was singly housed for recovery for at least 3 hours until observation.

Beads injection. The beads injection into the cortex was performed in an open-skull surgery, with minor modifications before sealing with the coverslip, by using an oil-filled glass micropipette attached to a mechanically driven Hamilton microsyringe. The left parietal bone was elliptically resected about 3.9-mm in the anterior-posterior (A-P) axis and about 2.7-mm in the medial-lateral axis. A needle of the glass micropipette was inserted into the cortex at an angle of 54 degrees with respect to the brain surface, and proceeded 0.9 mm towards the anterior direction along the A-P axis, to reach about 0.72 mm depth. 1-μm YG beads (Fluoresbrite® YG Microspheres, Calibration Grade 1 μm, #18860; Polysciences, Incorporated, Warrington, PA, USA) were diluted at 4.6×10^8 beads/mL with PBS, and were gradually injected into the needle trace in repeated steps, in which the needle was withdrawn 0.3 mm and then was kept in place for 5 min. During and after the injection, the exposed dura was washed with PBS to remove blood cells and to prevent drying. After the surgery, the mouse was singly housed for recovery for at least one day until observation.

Setup of *in vivo* imaging. Between 3–12 hours or 1–2 days after surgery, *in vivo* fluorescence imaging was performed in H-line mice to observe neurons or in C57BL/6NCr Slc mice injected with artificial beads to do FESTA, respectively. All *in vivo* imaging sessions were conducted in mice anesthetized with isoflurane inhalation (Kawakami et al., 2013; 2015). The head chamber was mounted and glued to a customized adaptor stage, which suspended the mouse body with a harness, to reduce adverse effects due to movements. After setting under an upright microscope system, the adaptor stage was adjusted at a tilt angle, referring to 1-μm diameter fluorescent beads scattered on the coverslip as a guide, so that the cranial window was positioned horizontally, hence providing a suitable optical alignment. All *in vivo* imaging sessions were performed under a microscope system (A1R MP+ Multiphoton Confocal Microscope; Nikon, Tokyo, Japan) controlled by NIS-Elements software (version 4.13.00; Nikon) with 12-bit dynamic range. Before *in vivo* imaging, the z-position of the brain surface was set as zero for absolute depth.

Two-photon microscopy. *In vivo* two-photon imaging was performed with a water immersion objective (25× Ob., 1.10 NA) in PBS. The EYFP and the YG beads were excited by 960 nm and 900 nm, respectively, emitted from Ti:Sapphire laser (Mai Tai DeepSee; Newport Spectra-Physics, Irvine, CA, USA). By using a Galvano scanner in one-directional mode, the xy-images in the 3D stacks were acquired as 512 × 512 pixels at 1 frame per sec (fps), 2.2 μsec per pixel, which is the maximum speed in this condition. The 3D display was performed using the NIS-Elements software.

To visualize the neural circuits, all the EYFP signals under 650-nm wavelength were collected in a GaAsP type Non-Descanned Detector (GaAsP-NDD). The CxLV and HpCA1 images were captured as 3D stacks with a 3-μm z-step size. The depth of the CxLV was determined between 495 μm and 585 μm, where somas of CxLV pyramidal neurons were observed in every mouse. The depth of the HpCA1 was slightly different for each mouse, thus the middle of stratum pyramidale, approximately at 975 μm absolute depth, was set as the center depth of each HpCA1 stack. To evaluate the brightness of the images, laser power (LP) under the objective lens and detector sensitivity (HV: high voltage) were fixed. The CxLV was captured with LP 66 mW (under 25× Ob.) and HV 30. The HpCA1 was captured with LP 200 mW (under 25× Ob.) and HV 30.

For the FESTA, fluorescence signals were split into multi-colors by dichroic mirrors (DM) in front of the GaAsP-NDDs. The signals of YG beads were split by DM458, DM511, and DM560, and then separately detected at 458–511 nm (Cyan) and at 560–650 nm (Yellow). To avoid optical influence by injection trace, the FESTA applied to the beads clusters observable without crossing the trace. The beads images were captured as 3D stacks with a 0.25-μm z-step size at depth of interest ± 1.0 μm (9 sections). The detector sensitivities were fixed: HV 1 on the Cyan channel and HV 3 on the Yellow channel. The laser powers were adjusted for each bead cluster, because bead clusters show different brightness depending on cluster size and injection depth.

Confocal microscopy. *In vivo* confocal imaging was performed with a water immersion objective (25× Ob., 1.10 NA or 60× Ob., 1.20 NA) in PBS. The EYFP was excited by 488 nm, and its signal was collected at 500–550 nm in a standard detector. By using a Galvano scanner in one-directional mode, the xy-images comprising 3D stacks were acquired as 512 × 512 pixels at 1 fps or 512 × 256 pixels at 2 fps, 2.2 μsec per pixel, which is the maximum speed in this condition.

To evaluate observable depth in confocal imaging, the CxLI images were captured as 3D stacks with a 1-μm z-step size at 50 ± 5, 100 ± 5, and 150 ± 5 μm depth (11 sections) via a 25× Ob. Laser power and detector sensitivity were fixed at 10 % and HV 90, respectively. As references for confocal imaging, the same 3D stacks were also captured by using a two-photon microscope, of which detector sensitivity was fixed at HV 30, but 960-nm laser powers were adjusted to 12.5, 16.1, and 19.2 mW (under 25× Ob.) at 50, 100, and 150 μm depth, respectively.

High-resolution 4D imaging with a 60× Ob. was performed as 3D stacks with a 1-μm z-step size at approximate 100 ± 15 μm depth (31 sections), as a time-series at 5 minutes intervals for 30 minutes (7 time-points). Laser power and detector sensitivity were set between 10 and 20 % and between HV 90 and HV 100, respectively. The 3D display was performed by NIS-Elements software.

Fluorescence intensity analysis. For two-photon imaging at CxLV and HpCA1, 30 mice (Treatment: control, n = 15; MAGICAL, n = 15) were observed. For confocal imaging at CxLI, 17 mice (Treatment: control, n = 8; MAGICAL, n = 9) were observed. In the H-line mice, signal distribution is uneven, especially at the somatic layer, because the EYFP expression level varies among positive neurons, which tend to be differently clustered in different individuals. To avoid selection bias due to individual differences, the imaging area was selected according to the xy position where the hippocampus was observable at the most shallow depth.

Fluorescence intensity analysis was performed on the 3D image stacks, by using Fiji/ImageJ software (version 2.0.0-rc-49/1.51d) (Schindelin et al., 2012). For sorting of image stacks within each capture condition, mice were ranked in descending order of brightness (average fluorescence intensity per pixel) extracted from

image stacks around the center region: at 540 ± 12 μm for the CxLV, approximately at 975 ± 12 μm for the HpCA1, and at 100 ± 5 μm for the CxLI. The top or bottom brightness at each capture condition came from over- or under-exposed image stacks, respectively. Thus, mice having the top or bottom brightness within each treatment group were equally removed from the following analysis, as outliers of the capture condition. Finally, 26 mice (control, n = 13; MAGICAL, n = 13) and 13 mice (control, n = 6; MAGICAL, n = 7) were evaluated in fluorescence intensity analysis for two-photon imaging and confocal imaging, respectively. A representative image for each treatment group at each capture condition was obtained from the mouse scored as intermediate brightness.

In the two-photon imaging, we evaluated the fluorescence intensity distributions (FIDs) under HV 30, at three different depths within the observation area. For the CxLV, the three depths were set at 510 ± 3, 540 ± 3 (center), and 570 ± 3 μm depth. For the HpCA1, three depths were relatively adjusted at the middle of stratum pyramidale as the center (0 μm), approximately at 975 ± 3 μm absolute depth, and at 24 μm above and below the center. These three depths were sufficiently far away from each other to ensure separate images in the mice.

The FIDs were summarized to 8-bit intensity bins histograms by the Fiji and were visualized through the following steps by R (version 3.3.2) (R Core Team, 2016) with RStudio software (version 1.0.44). The histograms were arranged as a matrix, where the image stacks were sorted in rows according to their brightness in decreasing order, and the intensity bins were aligned in columns. Pixel counts (z) in the histogram were converted with $\log_{10}(z+1)$, where 1 is a constant to avoid $\log_{10}(0)$ contamination in the converted matrix. The matrix of FIDs was visualized as a heat map. Pixel counts (excluding zero) were ranked in percent quantile. In the heat maps, the 50 % quantile was shown as cyan dotted contours.

Image similarity analysis. Based on image features, similarities between CM images and 2PM reference images at the same CxLI areas were evaluated in 13 mice (Treatment: control, n = 6; MAGICAL, n = 7), which passed fluorescence intensity analysis to remove outliers. The 3D stack images were converted to DccMIP 2D images with a color code shown in **Figure 3A**, by using Fiji/ImageJ software. The display ranges of fluorescence intensity (min–max in 8 bits) were adjusted as follows: 2PM images at every depth, (0–127); CM images at 50, 100, and 150 μm depth, (0–255), (0–127), and (0–63), respectively.

Features extraction and comparison were performed by using OpenCV (version 3.4.2) via a python (version 3.6.6) code. In gray-scale space, a Histogram of Oriented Gradients (HOG) (Dalal and Triggs, 2005) feature was extracted from each image (512 × 512 pixels) with parameters as follows: window size, (512, 512); block size, (128, 128); block stride, (4, 4); cell size, (16, 16); nbins, 180; other parameters were used as preset in OpenCV. The raw HOG feature comprising 108,391,680 bins was normalized to 1 to convert a probability histogram. In the comparison between the paired two images, a similarity index was defined as the intersection probability of the HOG features.

Fluorescence Emission Spectrum Transmissive Analysis. For spectral ratio calculation, data processing was performed by using the Fiji/ImageJ software (version 2.0.0-rc-49/1.51d) (Schindelin et al., 2012). All xy images in raw data (multi-color 3D stack) were smoothed with the median filter (1 pixel). Pixels having faint or saturated signal values (< 32 or 4063 < in 12 bit, respectively) were excluded from the following calculation. Background area out of beads was detected by an auto-threshold function with the Otsu method on the Yellow channel and was removed from all channels. Spread bead cluster images were selected by ROI (region of interest) to exclude the injection trace from the calculation. The pixels in YG bead clusters were plotted on a scatter graph, corresponding to the Yellow intensity along the x-axis and the Cyan intensity along the y-axis. At depth of interest, a representative Cyan/Yellow ratio was determined from the slope of the best-fitted regression line (highest R^2) on the scatter plot, derived from image stacks at its depth ± 1.0 μm. To calculate the ∆Cyan/Yellow, the Cyan/Yellow ratio was divided by the average of Cyan/Yellow within the experimental group at 100 μm depth.

Statistics. All statistical analysis was performed by using R (version 3.3.2) (R Core Team, 2016) with RStudio software (version 1.0.44). Conventional two-tailed Welch's t-test and Cohen's d were calculated with the default and the MBESS package (version 4.3.0), respectively. Permuted Brunner-Munzel (Brunner and Munzel, 2000; Neubert and Brunner, 2007; Neuhäuser and Ruxton, 2009) test was executed by the lawstat package (version 3.1) with permutation code.

To assess MAGICAL effects on FIDs, we constructed GLMMs with a Gamma error distribution and a log link function. In the models, fluorescence intensities of pixels (FIPs) were used as a response variable. The FID of each image stack at the depth ± 3 μm, therefore, was re-expanded to an intensity list comprised of 786,432 pixels (512 × 512 pixels × 3 sections). To reduce modeling costs, 12,288 pixels (1.56 %) were randomly sampled without overlap from each list. Saturated pixels (255 in 8 bit) were excluded from modeling due to inaccuracies in fluorescence intensity. Zero intensity pixels were also excluded to avoid a calculation failure under the log-link

function. “Treatment” (2 levels, MAGICAL or control) and “Depth” (3 levels) were used as categorical predictor variables with fixed effects. The models included an interaction between “Treatment” and “Depth.” To overcome pseudo-replications caused by sampling multiple pixels within each mouse, “Mouse” (26 mice in total) was assigned as a random effect for intercepts and “Depth” slopes. Thus, the models were described in R as follows:
FIP ~ Depth * Treatment + (Depth | Mouse)
Consequently, the models comprised 929,331 (CxLV) or 957,264 (HpCA1) pixel observations, in 78 image stacks extracted at 3 depths from 26 mice under 2 treatments.

The model fitting was performed via the maximum likelihood according to Laplace approximation, by using the glmer function in the lme4 package (version 1.1-12). For the estimated coefficient of fixed effect, Wald-type 95 % confidence interval (CI) was calculated by using the confint function. The model prediction for each image stack was obtained through the predict.merMod function in the lme4 package. The fixed effects with asymptotic 95% CI were estimated by using the Effect function in the effects package (version 3.1-2). Pairwise comparisons for the predictor variables were examined by the lsmeans function in the lsmeans package (version 2.25-5).

As the coefficient of determination for the GLMM, two R^2 statistics are proposed (Nakagawa and Schielzeth, 2013): marginal and conditional R^2_{GLMM}. The marginal R^2_{GLMM} ($R^2_{GLMM(m)}$) describes the proportion of variance explained only by fixed effects. The conditional R^2_{GLMM} ($R^2_{GLMM(c)}$) describes the proportion of variance explained by both fixed and random effects.

$$R^2_{GLMM(m)} = \frac{\sigma^2_{fix}}{\sigma^2_{fix} + \sigma^2_{random} + \sigma^2_{residual}} \quad (1)$$

$$R^2_{GLMM(c)} = \frac{\sigma^2_{fix} + \sigma^2_{random}}{\sigma^2_{fix} + \sigma^2_{random} + \sigma^2_{residual}} \quad (2)$$

An original method for the R^2_{GLMM} was limited only to the GLMM including Poisson or binomial error distributions and not including random slope terms (Nakagawa and Schielzeth, 2013). However, two extensions had been individually provided for the R^2_{GLMM}, to calculate random effects variances in random slope models (Johnson, 2014) or to evaluate further error distribution in the residual variance (Nakagawa and Schielzeth, 2016). Thus, we calculated the R^2_{GLMM} of our models by an R code combining the two extensions as follows: in the extension for random slope models (Johnson, 2014), a provided code generates a square matrix at once but uses only diagonal elements of the matrix. When the code ran for our models, a large number of observations generated a huge square matrix, which caused the computation to fail, due to insufficient memory. Therefore, we calculated only the portion corresponding to the diagonal of the square matrix. On the other hand, the extension for residual variances was applicable to our models without modifications. We selected the trigamma function from three candidates in the paper (Nakagawa and Schielzeth, 2016), to evaluate the models including a Gamma error distribution and a log link function.

Supplemental References:

Brunner, E., and Munzel, U. (2000). The nonparametric Behrens‐Fisher problem: asymptotic theory and a small‐sample approximation. Biom. J. *42*, 17–25.

Dalal, N., and Triggs, B. (2005). Histograms of oriented gradients for human detection. 2005 IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR'05) *1*, 886–893.

Johnson, P.C.D. (2014). Extension of Nakagawa & Schielzeth's R^2_{GLMM} to random slopes models. Methods Ecol. Evol. *5*, 944–946.

Nakagawa, S., and Schielzeth, H. (2013). A general and simple method for obtaining R^2 from generalized linear mixed‐effects models. Methods Ecol. Evol. *4*, 133–142.

Nakagawa, S., and Schielzeth, H. (2016). Coefficient of determination R^2 and intra-class correlation coefficient ICC from generalized linear mixed-effects models revisited and expanded. bioRxiv 095851.

Neubert, K., and Brunner, E. (2007). A studentized permutation test for the non-parametric Behrens–Fisher problem. Comput. Stat. Data Anal. *51*, 5192–5204.

Neuhäuser, M., and Ruxton, G.D. (2009). Distribution-free two-sample comparisons in the case of heterogeneous variances. Behav Ecol Sociobiol *63*, 617–623.

R Core Team (2016). R: a language and environment for statistical computing.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. Nat Methods *9*, 676–682.