ORIGINAL ARTICLE

Changes in the Proportion of Inhibitory Interneuron Types from Sensory to Executive Areas of the Primate Neocortex: Implications for the Origins of Working Memory Representations

Santiago Torres-Gomez[1], Jackson D. Blonde[1], Diego Mendoza-Halliday[2], Eric Kuebler[1], Michelle Everest[1], Xiao Jing Wang[3], Wataru Inoue[1] and Michael O. Poulter[1] and   Julio Martinez-Trujillo ⓘ[1,4]

[1]Department of Physiology and Pharmacology, Schulich School of Medicine and Dentistry, Robarts Research Institute and the Brain and Mind Institute, Western University, London, Ontario, N6A 5B7, Canada, [2]McGovern Institute for Brain Research, Massachusetts Institute of Technology, Cambridge, MA 02139, USA, [3]Center for Neural Science, New York University, New York, NY 10003, USA and [4]Department of Psychiatry, Schulich School of Medicine and Dentistry, Western University, London, Ontario, N6A5B7, Canada

Address correspondence to Julio Martinez-Trujillo. Email: julio.martinez@robarts.ca

Abstract

Neuronal spiking activity encoding working memory (WM) is robust in primate association cortices but weak or absent in early sensory cortices. This may be linked to changes in the proportion of neuronal types across areas that influence circuits' ability to generate recurrent excitation. We recorded neuronal activity from areas middle temporal (MT), medial superior temporal (MST), and the lateral prefrontal cortex (LPFC) of monkeys performing a WM task and classified neurons as narrow (NS) and broad spiking (BS). The ratio NS/BS decreased from MT > MST > LPFC. We analyzed the Allen Institute database of ex vivo mice/human intracellular recordings to interpret our data. Our analysis suggests that NS neurons correspond to parvalbumin (PV) or somatostatin (SST) interneurons while BS neurons are pyramidal (P) cells or vasoactive intestinal peptide (VIP) interneurons. We labeled neurons in monkey tissue sections of MT/MST and LPFC and found that the proportion of PV in cortical layers 2/3 decreased, while the proportion of CR cells increased from MT/MST to LPFC. Assuming that primate CR/CB/PV cells perform similar computations as mice VIP/SST/PV cells, our results suggest that changes in the proportion of CR and PV neurons in layers 2/3 cells may favor the emergence of activity encoding WM in association areas.

Key words: CR interneurons, inhibition–excitation balance, medial superior temporal, middle temporal, narrow and broad spiking, prefrontal cortex, PV interneurons, working memory (WM)

Introduction

Early electrophysiological studies in macaque monkeys showed that persistent neuronal firing in macaque monkey lateral prefrontal cortex (LPFC) neurons encodes the contents of working memory (WM) (Goldman-Rakic 1990; Fuster 2008). Persistent neuronal firing during the delay period of WM tasks has been observed in association areas of the macaque monkey neocortex, including prefrontal, posterior parietal, and inferior temporal cortices, but it is weak or absent in early sensory areas (Leavitt et al. 2017 for a review). A previous study has shown that persistent activity encoding WM representations of motion direction during a delayed match-to-sample task is present in the LPFC and area medial superior temporal (MST), but absent in area middle temporal (MT) (Mendoza-Halliday et al. 2014; Mendoza-Halliday and Martinez-Trujillo 2017). Differences in task-related activity between MT and LPFC have also been observed in other studies of WM (Zaksas and Pasternak 2006), as well as during categorization tasks (Freedman and Assad 2016). The source of these differences remains unclear.

Several mechanisms have been proposed to explain how persistent activity encodes WM representations (see Zylberberg and Strowbridge 2017 for a review). Some studies have proposed that persistent activity integrated over time windows of hundreds of milliseconds are the main coding substrate of WM. Others have proposed that persistent activity has a temporal structure, i.e., oscillations in certain frequencies carry the trains of action potentials encoding the contents of WM (see Constantinidis et al. 2018; Lundqvist et al. 2018 for reviews). Models of how persistent activity is generated have proposed that interconnected pyramidal cells (P cells) tuned for similar locations or features form network clusters that maintain recurrent excitation. These P cells activate inhibitory interneurons, mainly parvalbumin (PV) positive basket cells, which result in the inhibition of different clusters of P cells tuned for other locations or features (Goldman-Rakic 1995; Gonzalez-Burgos et al. 2000, 2005).

Modeling studies have attempted to explain the difference in the ability of microcircuits to produce persistent activity between sensory and association areas. Some have proposed the existence of "temporal receptive windows" that are progressively enlarged along the cortical hierarchy (Chaudhuri et al. 2015). Other studies have further proposed that increases in the ratio of excitatory to inhibitory elements (excitatory and/or inhibitory neuron types) within a microcircuit may favor recurrent excitatory dynamics between cells encoding similar representations and contribute to the emergence of persistent activity in association areas during WM tasks (Wang 2006; Wang and Yang 2018; inset in Fig. 1a). In this study we will examine the latter issue by obtaining functional measurements of putative cell types during electrophysiological experiments in behaving monkeys as well as histological measurements of cell types in brain slices.

The majority of the neurons in primate neocortical microcircuits can be classified into four categories defined by their neurotransmitter content and their expression of calcium-binding proteins (Conde et al. 1994; Tremblay et al. 2016). These are excitatory P cells and inhibitory calretinin (CR), calbindin (CB), and PV-expressing interneurons. CR, CB, and PV are calcium-binding proteins, and their concentrations can vary across interneuron types (Ascoli et al. 2008). These cell types may have similar roles as the vasoactive intestinal peptide (VIP), somatostatin (SST), and PV-expressing interneurons in the mouse (Xu et al. 2010; Kennedy et al. 2016). In general, it has been estimated that P cells

make up 70–80% of primate neocortex neurons while interneurons make up the remaining 20–30% (DeFelipe and Fariñas 1992; Markram et al. 2004; Elston et al. 2011). P cells are glutamatergic and can exert their effects more globally, while interneurons are generally GABAergic and contribute to local processing (DeFelipe and Fariñas 1992; Markram et al. 2004; Roux and Buzsáki 2015). In primate neocortex, CR, CB, and PV interneurons account for more than 90% of the inhibitory interneuron population, overlapping very little with one another (Conde et al. 1994; DeFelipe 1997; Dombrowski et al. 2001; Raghanti et al. 2010); they also vary in morphology and the neurons that they target (Wang et al. 2004; Raghanti et al. 2010; Wang and Yang 2018).

Assuming that CB, CR, and PV interneurons in primates perform similar computations as SST, VIP, and PV neurons in the mouse neocortex, one could elaborate on the following interactions. CB interneurons provide inhibitory input to the apical dendrites of P cells, allowing them to control the flow of excitatory inputs (Zaitsev et al. 2009; Wang 2013). Since their activation may suppress P cell activity, CB interneurons can be considered inhibitory to P cells (Wang et al. 2004). CB interneurons are targeted and inhibited by CR interneurons (i.e., inhibition of inhibition); therefore, activation of CR neurons facilitates activation of the P cell by sensory or recurrent inputs into their dendrites (Conde et al. 1994; Meskenaite 1997; Zaitsev et al. 2004; Runyan et al. 2010; Zhu et al. 2015). Lastly, PV interneurons provide inhibitory input to CR interneurons and to P cell bodies (Williams et al. 1992). In summary, CR interneurons may potentiate the activation of P cells via inhibition of CB interneurons, while PV interneurons may attenuate the activation of P cells via direct inhibition and/or via inhibition of CR interneurons. This hypothetical microcircuit arrangement is depicted in Figure 1a (adapted from Wang and Yang 2018).

It is possible that differences in the relative proportions of PV, CR, CB, and P cells across areas of the visual processing pathway cause a shift in the excitation–inhibition balance that controls P cell firing, influencing the ability of P cell clusters to produce persistent firing across areas. For instance, assuming that the influence of a cell type "A" into other cell type "B" projects to increases with increases in the proportion of "A," one could hypothesize the following. A decrease in the proportion of PV and/or CB interneurons, as well as an increase in CR interneurons, would increase the activation of P cells either by allowing inputs to activate their dendrites (e.g., decrease in CB interneuron activity by a large population of CR interneurons) or by reducing direct inhibition (e.g., decrease in PV interneuron population). These changes could favor recurrent excitatory dynamics between interconnected P cells and result in an increased ability to produce persistent activity encoding WM (Wang 2001; Wang and Yang 2018) (Fig. 1b). The opposite effect may occur if the proportion of PV and/or CB interneurons increases, while the proportion of CR interneurons decreases (Fig. 1c). Such changes along the cortical pathways may cause persistent firing encoding WM to emerge in areas such as LPFC. One must consider that the models illustrated in Figure 1a–c oversimplify cell type diversity in the neocortex; however, they allow to illustrate our hypothesis.

To investigate whether the proportions of CR, CB, and PV neurons change along the cortical pathways, we first recorded single-cell responses from areas MT, MST, and LPFC of two macaque monkeys during a WM task. We classified neurons according to their action potential waveform width into narrow spiking (NS) and broad spiking (BS). We found that the ratio of NS to BS neurons decreased from MT to MST to LPFC. We
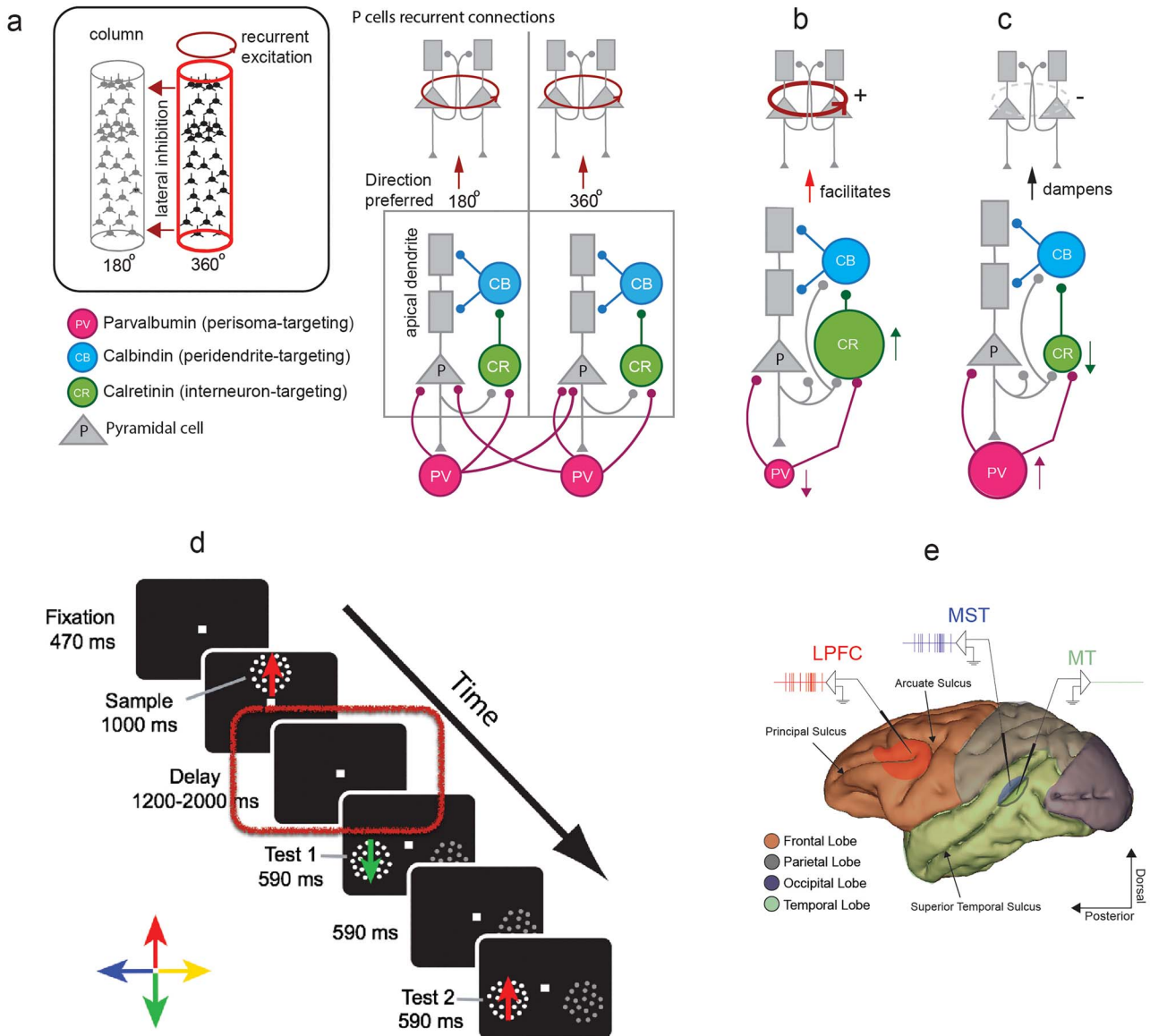
Figure 1. (*a*) Hypothetical circuit illustrating the different cell types and their connectivity in the primate neocortex. The basic circuit involving different cell types is at the bottom. The inset at the top illustrates the recurrent connections between P cells encoding the same feature (motion direction in this case) and inhibitory connection between neurons/columns encoding different features. (*b*) Consequences on recurrent excitation of groups of P cells when increasing the proportion of CR interneurons (increase in the size of the green circle) and decreasing the proportion of PV interneurons (decrease in the size of the red circle). (*c*) Consequences on recurrent excitation of groups of P cells when increasing the proportion of PV interneurons (increase in the size of the red circle) and decreasing the proportion of CR interneurons (decrease in the size of the green circle). (*d*) Visual sequence of events for delayed match-to-sample task (DMTS). (*e*) Image of the macaque brain reconstructed from MRI showing the recorded cortical areas, with MT in green, MST in blue, and LPFC in red.

used the human and mouse cell types database from the Allen Institute (https://celltypes.brain-map.org) to analyze the relationship between neuronal types and waveform (NS and BS). We showed that while NS neurons are in general either PV or SST interneurons, BS neurons can be P cells or VIP interneurons. Finally, we used immunohistochemistry to label PV, CR, CB, and P cells in layers 2/3 of MT, MST, and LPFC of two other macaque monkeys. We focused on these layers because it has been shown they contain the majority of neurons with persistent activity encoding WM (Bastos et al. 2018). We found a decrease in the relative proportion of PV interneurons and an increase in the

relative proportion of CR interneurons from MT/MST to LPFC. Finally, the relative proportion of P cells and CB interneurons remained similar across areas.

Materials and Methods

All animal procedures and protocols were performed in accordance with the Canadian Council of Animal Care guidelines and were approved by the McGill University and the University of Western Ontario Animal Care Committees.

Behavioral Task

We trained two adult male rhesus monkeys (*Macaca mulatta*; monkey M and monkey S) to perform a motion direction delayed match-to-sample task, in which they had to memorize the motion direction of a visual stimulus, while single-neuron spiking activity was recorded in areas MT, MST, and LPFC. The monkeys were given fruit juice as a reward for correctly performing the task trials (Fig. 1*d*). During a trial, the monkey maintained gaze on a white fixation square (size $0.25° \times 0.25°$) at the center of the screen and pressed a button to initiate the trial. After 470 ms, a sample stimulus with motion in one of four orthogonal directions was presented for 1000 ms. The sample was then removed, and after a delay period of variable duration (from 1200 to 2000 ms), two test stimuli were sequentially presented for 590 ms each, with 590 ms in between. To receive a juice reward, the monkey was required to release the button during the presentation of a test with the same motion direction as the sample. This occurred in half of the trials for each test. If the monkey failed to do so, the trial was terminated without a reward. The trial was also terminated without a reward if gaze position moved outside a fixation point centered window (area of 1°) before the end of the trial.

Sample and test stimuli were composed of random dots moving linearly with 100% coherence within a virtual circular aperture (13 cd/m^2 dot luminance contrast, 0.17° dot size, density of 4 dots/degree^2). High motion coherence ensured that encoding of the sample direction in working memory was not ambiguous. The motion directions of the sample and tests were chosen from a set of four orthogonal directions aligned so that one of them matched the neuron's preferred direction (often differing from neuron to neuron within the same day). This way, the monkeys could neither use long-term memory representations of these directions nor simply learn four fixed categories across sessions to solve the task. In trials in which the first test did not match the sample, the direction of this stimulus was randomly chosen from the other three directions in the set.

A behaviorally irrelevant stimulus with 0% coherent motion and lower luminance contrast was presented simultaneously with the two tests on the opposite hemifield. The test and irrelevant stimulus locations were randomly swapped from trial to trial, preventing the monkeys from predicting the test location before its presentation. The irrelevant stimulus also served other experimental purposes unrelated to the results presented here. The variable delay prevented monkeys from anticipating the timing of the test onset.

Anatomical Localization of Recording Sites

An MRI scan was conducted on each monkey involved in the behavior-electrophysiological recording experiments before the surgery in order to guide the positioning of the chamber. After chamber implantation, a plastic grid (Crist Instruments) was positioned, and an additional MRI scan was conducted to precisely locate the areas of interest in the brain with respect to the electrode trajectories (Supplementary Fig. S1*a*) (see also Mendoza-Halliday et al. 2014). The majority of LPFC neurons were recorded by placing the electrode tip in positions in the prearcuate region, around the principal sulcus.

Electrode trajectories reached areas MT and MST in the right hemisphere according to MRI reconstructions. To record from MT, we positioned the electrode tip ventral and posterior to the superior temporal sulcus. To record from MST, the electrode tip was placed dorsal and anterior to the superior temporal sulcus. The recorded MST neurons were located within the dorsal region (MSTd).

Electrophysiological Recordings

During each recording session, we made transdural penetrations with standard epoxy-insulated extracellular tungsten electrodes (FHC Inc; shank diameter = 500 μm in PFC and shank diameter = 125 μm in MT and MST; impedance = 2–4 MΩ at 1 kHz). For MT and MST recordings, a guide tube was lowered through the posterior craniotomy until it penetrated the dura. The electrode was then lowered through the guide tube until it reached the desired depth. During each session, we recorded with one electrode placed in MT or MST and/or with one to four electrodes simultaneously in PFC (separated by at least 2 mm). A Plexon data acquisition system (MAP) was used to record, store, and sort spike data (Lennert and Martinez-Trujillo 2011).

Characterization of Neuronal Response Properties

For MT and MST, once a neuron was isolated, a brief manual characterization of the motion direction responsiveness and receptive field location ensued, and exploration continued until a stable, identifiably responsive neuron signal was found. The spatial and motion tuning properties of MT and MST neurons were then further characterized during trials in which the monkey responded to a contrast change in the fixation point, and random dot patterns (RDPs) varying in location, size, motion type (linear and spiral motion directions), and speed were presented (Treue and Martinez-Trujillo 1999). The preferred direction was then determined based on firing rate increase when exposed to RDPs of various motion directions as compared with baseline (no stimulus).

MT neurons were identified based on their linear motion direction selectivity, receptive field (RF) size, and laterality (i.e., contralateral to the recorded hemisphere). MST neurons were identified based on their linear and spiral motion direction selectivity and receptive field size and position (considerably larger than for MT neurons and often spanning both hemifields). The distributions of receptive field sizes of MT and MST neurons were clearly segregated, and the percentage of neurons with RFs including ipsilateral regions was far higher in MST (56%) than in MT (12%) (Supplementary Fig. S1*b*, see also Mendoza-Halliday et al. 2014).

Because the activity of LPFC neurons is highly task-dependent, we did not use the mapping task to characterize their visual motion response properties. Instead, we recorded while the monkey performed the delayed match-to-sample task immediately after isolating a neuron. Finally, to avoid a bias in the selection of neurons due to the differences in the isolation procedure between brain areas, neurons were further selected for this study only if they showed significant direction selectivity in the sample or delay period of the trial determined by a two-factor ANOVA ($P < 0.05$), as described below in "direction selectivity."

Electrophysiological Data Analysis: Waveform Classification

Waveform classification is a limited, yet still useful, analysis tool that takes advantage of the well-established action potential dynamics of a main type of inhibitory interneuron,

PV-expressing inhibitory interneurons (most likely large basket cells). PV interneurons substantially correspond with NS profiles, target perisomatic regions of P cells, and provide widespread inhibition (McCormick et al. 1985; Kawaguchi et al. 1987; Cauli et al. 1997; Galarreta and Hestrin 2002).

Preprocessing
Neuronal responses were isolated and recorded using a Plexon data acquisition system (MAP). During recording, the electrical signal was initially high-pass filtered at 300 Hz, and action potentials were then isolated by setting a manual threshold based on the amplitude of the filtered signal, such that threshold-crossing spikes were recorded. Time stamps for action potentials were sampled every 25 µs (40 kHz) with each waveform containing a trace including 200 ms before and 600 ms after crossing the threshold. Spike sorting was manually performed with Plexon sorting software (Plexon Inc), where single units were determined if they formed a well-isolated and clear cluster separable from noise or other unit clusters. For all well-isolated neurons in a given recording, the mean action potential (AP) waveform template for each sorted unit was obtained, and templates were then aligned at the point of threshold crossing and normalized to the largest peak amplitude.

Waveform Width
The AP waveform was further interpolated with a spline fit to an increased precision of *100 for a final resolution of 0.25 µs per tick (Mitchell et al. 2007; Ardid et al. 2015). Irregular waveforms were detected by visual inspection and manually discarded if they had shapes uncharacteristic of APs. Units were also discarded if their waveform template deviated drastically from the overall waveform distribution. From the resulting average AP waveform template, we obtained the time stamps at the trough and peak (minimal and maximal voltage values), and the AP waveform width was computed as the difference between the two time stamps. AP waveform widths for all neurons in a given area were plotted in a histogram.

Waveform Width Bimodality Testing and Thresholding
To discriminate between NS and BS neurons, the AP waveform widths were fitted using a Gaussian Mixed Model (GMM), and a cutoff threshold was defined to divide neurons into two groups by setting a boundary at the inflection point of the two distributions. The GMM was evaluated using Akaike's information criterion (AIC) (Akaike 1974), a measure that compares between statistical models to determine the best model from a set of candidate models (unimodal vs. bimodal distribution in this case). Bimodality was inferred if the AIC value was lower for the bimodal fit than the unimodal fit. Bimodality of the waveform distributions was further corroborated, in addition to the AIC, by using Hartigan's dip test for unimodality (Hartigan and Hartigan 1985; Freeman and Dale 2013). For Hartigan's dip test, a P value below 0.05 indicates significant multimodality, while P values between 0.05 and 0.1 suggest multimodality with a marginal significance. The trough of the bimodal Gaussian fit was used as the cutoff point for the distribution. We classified neurons in the upper range of trough-to-peak duration, BS neurons, as P cells, CB, and CR interneurons. Likewise, neurons in the lower range, NS neurons, were classified as PV interneurons (McCormick et al. 1985; Kawaguchi et al. 1987; Cauli et al. 1997; Galarreta and Hestrin 2002).

Naka–Rushton Fit
To further validate the neuron type classification, a sigmoid function was employed to fit the initial segment of the AP waveform, a Naka–Rushton fit:

$$R = [R_{max}^* X_t^n] / [X_t^n + X50^n] + M$$

where X_t stands for the different values of the waveform as a function of time, R_{max} stands for the maximum voltage value of the AP waveform after subtraction of M, the lowest voltage value, and n represents the slope coefficient of the fit. This function is normally used to fit contrast-response functions (Sclar et al. 1990; Khayat et al. 2010). The slope coefficient n was then used to establish a distribution of values that underwent a similar bimodality testing and thresholding as that performed on the waveform width distributions.

Spike Density Function
All analyses of spike data and statistical tests were performed with custom software written in MATLAB. We computed spike density functions by convolving the spike train with a Gaussian probability function, using a normal Gaussian kernel ($\sigma = 40$ ms). Results obtained from both monkeys were qualitatively similar. All neurons from both monkeys were pooled together for analysis. We analyzed neuronal activity during the sample and delay periods.

Direction Selectivity
To assess whether a given neuron encoded the sample direction over time, we tested for significant differences in firing rates between trials with different sample directions across time bins of 120 ms. We used a mixed between-within two-factor analysis of variance (ANOVA) with sample direction as a "between" factor and time bin as a "within" factor. A neuron was considered direction-selective if it had a main effect of direction ($P < 0.05$) in at least one of two ANOVAs: one using time bins from the sample presentation period (sensory selectivity) and the other using time bins from the delay period (delay selectivity).

To determine the percentage of sensory-selective and delay-selective neurons that would be expected by chance, we randomly shuffled, for each neuron, the sample direction labels of all trials. We then performed the same two-factor ANOVA to obtain a surrogate percentage of sensory-selective and delay-selective neurons. The analysis was repeated 500 times to obtain 500 surrogate values of percentage of selective neurons. Percentages were considered significantly higher than chance if they were ranked within the top 95th percentile among all 500 surrogate values.

Receiver Operating Characteristic Analysis
To quantify neurons' ability to discriminate between sample motion directions, we performed a Receiver operating characteristic (ROC) analysis. For each neuron, we computed the area under the ROC curve (auROC) to measure the separability of the distributions of firing rates between all possible pairs of sample directions. The auROC was computed across a sliding time window of 200 ms shifted by increments of 40 ms. auROC values between 0 and 0.5 were rectified to their corresponding values in the range between 0.5 and 1. Among all pairs of sample directions, the two directions for which the mean auROC across the sample and delay periods was highest were chosen as preferred and least-preferred directions. The auROC computed between these two sample directions was used to measure each

neuron's direction discriminability and to compute population averages.

To test whether auROC values were significantly higher than expected by chance, we performed a permutation test in which we shuffled preferred-sample and least-preferred-sample trial labels and computed the auROC between the shuffled trials. This procedure was repeated 500 times. An auROC was considered significant if it reached or exceeded the 99th percentile of the distribution of the 500 shuffled surrogates. This was performed through all steps of the sliding time window to detect periods of significant direction discriminability.

Intracellular Recording Analysis

We downloaded data from the Allen Institute cell types database (https://celltypes.brain-map.org). For mice recordings were obtained in V1 slices from male or female animals between ages of P45 and P70. Interneuron-specific or layer-specific Cre driver lines were used to target subpopulations (i.e., PV, VIP, and SST). Human brain tissue was from neurosurgical origin (https://celltypes.brain-map.org for details). The membrane potential of cells in the database was sampled at either 50 or 200 kHz (explained in white paper). For our analysis, we specifically used the rheobase action potential triggered by 1 s square pulse current injection protocol. The spike width of intracellular action potentials was computed by finding the maximum (i.e., peak) and minimum voltage (i.e., trough) and then computing the difference between these points in time (measured in ms). Cells with spike widths >3 ms were removed from the analysis (mostly VIP and spiny cells).

Immunohistochemistry

Tissue Extraction and Preparation

Two adult male rhesus monkeys, monkey T and monkey R (different from those used for electrophysiology), were anesthetized with ketamine (100 mg/mL at a dose of 7–10 mg/kg) and medetomidine (1 mg/mL at a dose of 0.02–0.03 mg/kg) and intracardially perfused with heparinized saline. Their brains were then removed and placed in paraformaldehyde.

After removing each brain, we located areas MT and MST along the superior temporal sulcus (STS) and area LPFC in the prefrontal cortex using the Paxinos et al. (2000) brain atlas. The LPFC tissue was taken from areas 8A/46, which are anterior to the arcuate sulcus and around the third posterior end of the principal sulcus (Petrides 2005). To extract each area, the brains were first cut midsagittally, and the hemispheres were placed on their medial surface to provide stability and accuracy during subsequent cutting. Areas MT and MST were extracted by making coronal cuts at approximately interaural 2.05 mm, bregma −19.75 mm (see Figure 86, Paxinos et al. 2000 brain atlas) and interaural −7.80 mm, bregma −29.70 mm (see Figure 108 of atlas). Since these areas are adjacent, they were both present within the same block of tissue (Fig. 1e). Areas 8A/46 were extracted by making coronal cuts at approximately interaural 32.70 mm, bregma 10.80 mm (see Figure 18 of atlas) and interaural 23.25 mm, bregma 1.35 mm (see Figure 38 of atlas). Visual inspection and comparison of the coronal faces of the MT/MST and LPFC tissue blocks to the atlas corroborated that the correct regions were extracted. All brain areas were taken from the left hemisphere, except for monkey R's LPFC, which was taken from the right hemisphere due to a contralateral electrode implantation and potential distortion of the anatomy within that area.

The MT/MST and LPFC tissue blocks were then placed in LANA's fixative (4% paraformaldehyde, 20% picric acid) for 24 h and subsequently transferred to a 30% sucrose phosphate buffer until they sank (∼48 h). The tissue was then flash frozen in −80 °C isopentane for 30 s and stored at −20 °C until sectioning. Using a Leica CM1950 cryostat (Leica Biosystems), each block of tissue was sectioned at −20 °C in the coronal plane at a slice thickness of 50 μm. The slices were placed in 24-well trays, free-floating in cryoprotectant.

Staining

The immunohistochemistry procedures used in the present study are similar to those described in Pollock et al. (2014). Six MT/MST and LPFC tissue slices from each primate were first washed with 1× PBS for 5 min followed by two 5-min washes with 1× PBS + 0.5% Triton X-100. The six slices progressed from anterior to posterior and were inspected to ensure that they contained the regions of interest. The slices were then incubated at room temperature for 1 h in blocking solution (0.025% Triton X-100 + 1% BSA in 1× PBS) + 10% goat or donkey serum. Next, primary antibodies diluted in blocking solution were added to the tissue. Four primary antibodies were used: mouse anti-PV, rabbit anti-CR, rabbit anti-CB, and sheep anti-Neurogranin. Neurogranin was used to selectively stain for P cells (Singec et al. 2004; Guadaño-Ferraz et al. 2005). See Table 1 for a summary of the primary antibodies. The tissue was then incubated overnight at 4 °C.

The following day, two 5-min 1× PBS washes were performed. The secondary antibodies diluted in blocking solution were then added to the tissue. These included Alexa Fluor 647 conjugated donkey anti-rabbit IgG, 1:500 (Invitrogen, catalog number: A31573), Alexa Fluor 488 conjugated goat anti-mouse IgG, 1:500 (Invitrogen, catalog number: R37120), and Alexa Fluor 488 conjugated donkey anti-sheep IgG, 1:500 (Invitrogen, catalog number: A11015). The tray was then immediately covered with tinfoil to prevent photobleaching and incubated at room temperature for 1 h. Next, a 5-min 1× PBS + 1:5000 DAPI (4′, 6-diamidino-2-phenylindole) wash was performed, followed by a second 5-min 1× PBS wash. The tissue was then incubated in 1% Sudan Black B dissolved in 70% ethanol for 2 min to attenuate lipofuscin autofluorescence (Pollock et al. 2014). Lastly, the tissue was washed twice for 1 min with 70% ethanol and then washed twice for 5 min with 1× PBS. Note that the tray containing the tissue was placed on a rocker for all washes and incubations to ensure maximal antibody binding and thorough washing.

Each tissue section was then transferred onto a Fisher Super-Frost Plus slide using a pipette, and excess fluid was removed using a suction system. The coverslip was then mounted using Prolong gold anti-fade with DAPI mounting medium (Molecular Probes). Slides were stored at −4 °C in the dark until they were imaged.

Image Acquisition

Wide field images of MT, MST, and LPFC were acquired using an EVOS FL Auto 2 Imaging System (Thermo Fisher Scientific) outfitted with a 10× objective lens. The individual 10× images were aligned and merged using EVOS's automated stitching function. Stitching the images allowed for the visualization of the entire tissue slice and was used to ensure that the correct brain areas were isolated, to visualize the distribution of each cell type across cortical layers, and to direct subsequent imaging to layers 2 and 3. Cortical layers were identified using DAPI

Table 1 Primary antibodies

1° Antibody	Animal	Dilution	Mono/polyclonal	Source	Catalog number
Calbindin D28K	Rabbit	1:500	Polyclonal	Chemicon	AB1778
Calretinin	Rabbit	1:500	Polyclonal	Abcam	AB702
Neurogranin	Sheep	1:500	Polyclonal	R&D Systems	AF7947
Parvalbumin	Mouse	1:500	Monoclonal	Swantz	235

staining, which revealed differences in neuronal soma size and density across the layers (Elston and Rosa 1997). For instance, layers 2 and 4 contain relatively small, tightly packed somata and appear as light blue bands running parallel to the edge of the tissue. In contrast, layer 3 contains larger and more dispersed somata and appears as a dark blue band between layers 2 and 4.

The images used for cell counting were taken from layers 2 and 3 of MT, MST, and LPFC. These images were acquired using an Olympus IX 60 inverted microscope outfitted with a Perkin Elmer spinning disk confocal attachment, 20× objective lens (numerical aperture = 0.5), and Hamamatsu Orca ER CCD camera (1300 × 1010 pixels; pixel size 6.5 μm). Volocity 6.3 software was used to operate the microscope (PerkinElmer Inc). Six layer 2 and six layer 3 images were taken for each of the four cell types (P cell, PV, CR, CB) within each slice of MT/MST and LPFC. Recall that there were six MT/MST tissue slices and six LPFC tissue slices; therefore, a total of 288 images (12 images × 6 slices × 4 cell types) were taken for each brain area in each monkey. Considering that three brain areas were analyzed across two monkeys, a total of 1728 images were acquired.

For each layer 2 image taken, a "matched" layer 3 image was taken within the same cortical column. The cell counts across each pair of layer 2/3 images were later summed to get an accurate representation of "layer 2/3." To ensure consistency in imaging, the six pairs of layer 2/3 images were acquired in roughly the same positions across slices, starting at one end of the brain area and ending at the other. Layers 2 and 3 were chosen for analysis because they contain the highest density of interneurons (Gabbott and Bacon 1996; Dombrowski et al. 2001) as well as intra-layer horizontal connections that may provide the anatomical substrate for recurrent circuits (Wang 2013).

Cell Counting
Following layer 2 and 3 image acquisition, all 1728 images (Fig. 7a,b) were imported into Fiji (Wayne Rasband, National Institute of Mental Health). Cells were counted manually using the "Cell Counter" tool found within Fiji's "Analyze" plugin. Data were recorded in an Excel spreadsheet.

Results

Electrophysiological In Vivo Experiments

A total of 631 neurons were recorded from monkey M and monkey S while performing a motion direction delayed match-to-sample task (DMS) (Fig. 1d,e), with both monkeys' performance well above chance (79% and 76%, respectively). One hundred twelve neurons (57 from monkey M and 55 from monkey S) were from MT, 247 neurons (145 from monkey M and 102 from monkey S) were from MST, and 272 neurons (118 from monkey M and 154 from monkey S) were from the LPFC (see Supplementary Fig. S1). After removing outliers that did not fit the criteria for the waveform classification analysis, a total of 96 neurons in MT

(86%), 222 neurons in MST (90%), and 229 neurons in LPFC (84%) remained and were included in the analysis.

For each neuron, individual action potential waveforms were recorded, and an average waveform was obtained. Neurons were categorized as either NS or BS by measuring the width of their average action potential (AP) waveform or peak-to-trough time interval (Fig. 2a–c; Mitchell et al. 2007; Hussar and Pasternak 2009). The distributions of these values are plotted for all areas (Fig. 2d–f). Bimodality of the distributions was assessed by fitting the data with a single Gaussian function (1-Gaussian) and a sum of two Gaussian functions (2-Gaussian) and determining whether the 2-Gaussian model provided a better fit. The goodness of fit was determined based on Akaike information criterion (AIC). We also evaluated bimodality using the Hartigan's dip test (see Methods). A threshold was set by finding the inflection point where the two Gaussian distributions intersected. For all three areas, AIC was lowest for the 2-Gaussian model (1-Gaussian vs. 2-Gaussian mean AIC values MT: 1189 vs. 1170; MST: 2774 vs. 2768; LPFC: 2857 vs. 2779). The Hartigan's dip test P value was below 0.1 for all three areas (MT, $P = 0.086$; MST, $P = 0.039$; LPFC, $P = 0.094$), "suggesting bimodality with marginal significance" (Freeman and Dale 2013).

To further corroborate these results, we quantified the slope of a Naka–Rushton (sigmoidal) fit to the average AP waveform in each neuron and plotted the distribution of slopes (see Supplementary Fig. S2a). Again, we fitted Gaussian mixed models (GMM) to the data and computed the goodness of fit using the AIC. Both AIC and Hartigan's dip test values indicated bimodality. AIC values were lower for the 2-Gaussian model (1-Gaussian vs. 2-Gaussian AIC values MT: 5.343 vs. 5304; MST: 2598 vs. 2566; LPFC: 861 vs. 842) and the Hartigan's dip test P value below 0.05 for MT and MST (MT: $P = 0.018$; MST: $P = 0.009$) and below 0.1 for LPFC ($P = 0.091$), suggesting that distributions were bimodal.

We categorized neurons into NS or BS based on the deflection point between the 2-Gaussians. Both methods (using spike width or using the slope of the Naka–Rushton function) yielded similar proportions of neuron types. Of all recorded neurons, the percentage of NS was 52% in MT, 39% in MST, and 30% in LPFC, whereas the percentage of BS was 48% in MT, 61% in MST, and 70% in LPFC. The width values and the estimated slope of the sigmoidal fit were strongly negatively correlated (R values: MT = −0.762; MST = −0.752; LPFC = −0.736) (Supplementary Fig. S2b).

In order to compute the proportion of NS and BS neurons, we first computed the integral of the predicted values of each Gaussian function provided by the fit of the 2-Gaussians to the width measurement. The Gaussian with the lowest mean value represents the NS distribution, while the Gaussian with largest mean value represents the BS distribution (Fig. 2g). The ratio of NS to BS neurons shows a clear decreasing trend from MT to MST to LPFC (Fig. 2h). Indeed, we found significant differences in the proportion of NS neurons between areas (Chi-square test P values: MT-MST = 0.0108; MT-LPFC = 0.0001;
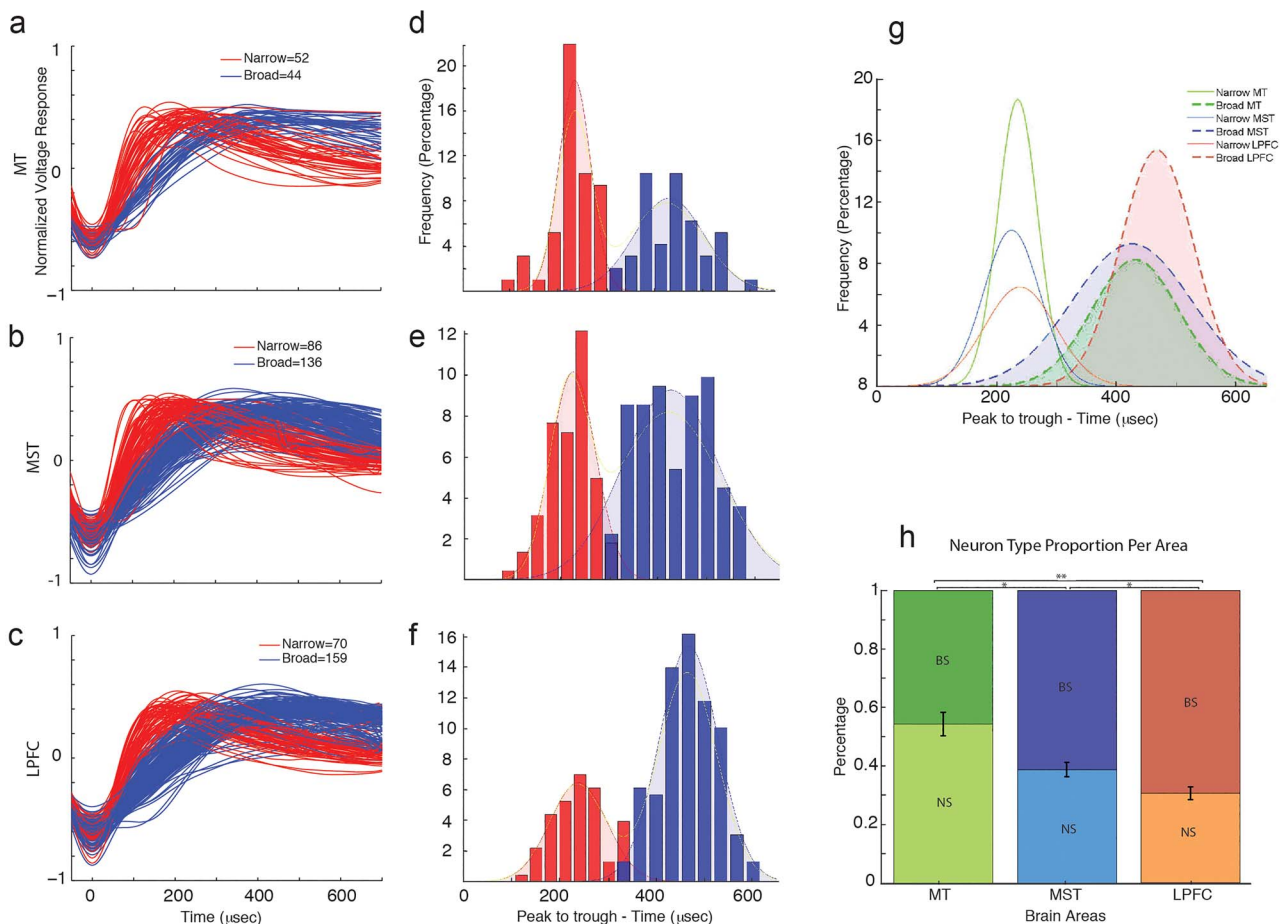
Figure 2. Classification of neurons into narrow spiking (NS) and broad spiking (BS). (*a–c*) Normalized average action potential waveforms for MT, MST, and LPFC, displaying NS (red) and BS (blue) waveforms. (*d–f*) Histograms showing the distribution of the trough-to-peak distance of the AP waveform. A Gaussian fit for each distinct population, with NS neurons in red and BS neurons in blue, is shown, with a bimodal distribution being observed for all areas. Hartigan's dip and *P* value are shown for each area. (*g*) Gaussian functions fit to the different populations of neurons per area. Filled dashed lines represent BS distributions, solid lines represent NS distributions, and colors represent different brain areas. (*h*) Proportion of NS and BS neurons per area.

MST-LPFC = 0.0482). Moreover, the cumulative distribution functions of the waveform peak-to-trough widths are shifted to the right with the rightmost corresponding to LPFC, followed by MST, and then MT (Kolmogorov–Smirnov test *P* values: MT-MST, 0.0181; MT-LPFC, 1.1233e − 5; MST-LPFC, 7.7637e − 7; Supplementary Fig. S2c). Thus, MT contains a higher relative proportion of NS neurons, followed by MST, and finally by LPFC.

Motion Direction Discrimination During the WM Task in NS and BS Neurons

We next explored whether NS and BS neurons differ in their ability to discriminate motion direction during the sample and delay periods of the WM task. We used a receiver operating characteristic (ROC) analysis to determine the ability of NS and BS neurons to discriminate between their preferred and anti-preferred motion directions. We computed the area under the ROC curve (auROC) during the sample and delay period for each neuron in each brain area in time bins of 240 ms by steps of 40 ms. The plots in Figure 3a–d show these results at the level of single units. A large number of neurons discriminated motion direction during the sample period in MT (d, light and dark green for NS and BS, respectively), followed by MST (e, light

and dark blue for NS and BS, respectively), and finally by the LPFC (f, orange and red for NS and BS, respectively). During the delay period, this trend reverses. Notably, for all areas, the results are similar for the subpopulations of NS and BS neurons. This indicates that NS and BS neurons show a similar ability to discriminate motion direction in the different periods.

In order to further quantify these results, we computed the mean auROC as a function of neuron type and area. Mean auROC values during the sample period were higher for MT, followed by MST, and finally LPFC. During the delay period, the results show an inverse trend across areas. (Fig. 4a,b). There was no distinction between NS and BS neurons within any area, with both neuron types showing similar direction discrimination during both the sample and delay period (*t*-test *P* values: sample period—MT, 0.2560; MST, 0.1692; LPFC, 0.0708; delay period— MT, 0.7539; MST, 0.6765; LPFC, 0.5308). This result was further confirmed when analyzing the distributions of auROC values for all areas and for the sensory and delay periods (Supplementary Fig. S3).

To compare the ability of neurons to discriminate motion direction during the sample and delay periods, we developed a period coding index [PCI = (auROCsample—auROCdelay)/ (auROCsample + auROCdelay)]. PCI > 0 means that the neurons
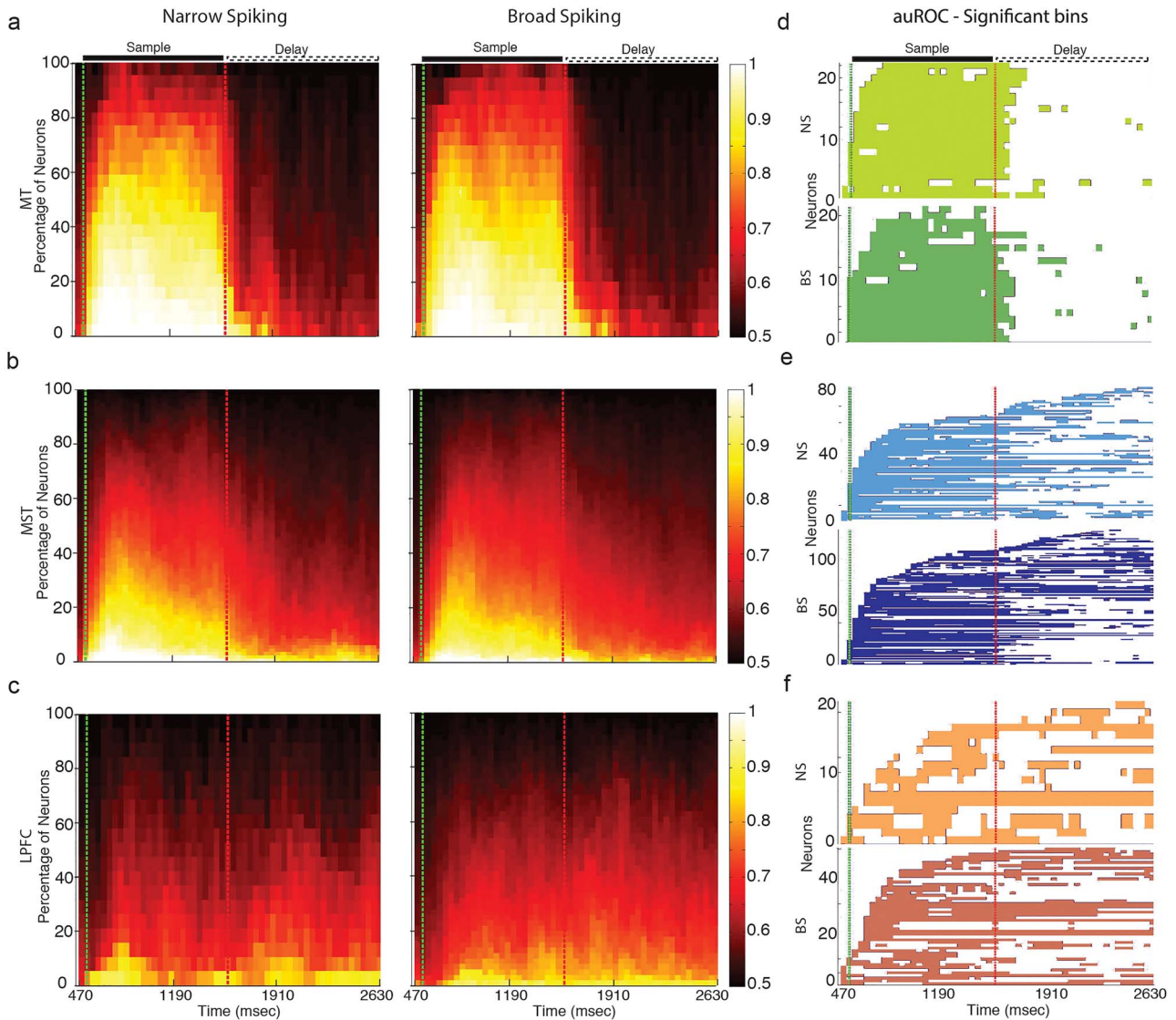
Figure 3. Proportions of neuron type (NS and BS) per area ((*a*) MT, (*b*) MST, and (*c*) LPFC) as a function of task period (abscissa). (*d–f*) Single neuron significance of auROC for NS and BS neurons for all three areas during the sample and delay periods.

are better at encoding the direction during the sample period than during the delay period; PCI < 0 means the opposite. PCI = 0 means that the neurons equally encode information during both periods. The boxplots in Figure 4c show the distributions of the different indices for NS and BS neurons across areas. MT shows the highest PCI values, followed by MST, and finally by LPFC. We found an effect of area on the PCI index (two-way ANOVA: area F, 75.75; $P = 1.33e - 27$) but no effect of cell type (cell type F: 2.44; $P = 0.1192$), indicating that cell type did not influence the neurons' ability to encode sensory and mnemonic signals (Tukey–Kramer post hoc test P values between NS and BS: MT, 0.3469; MST, 0.796; LPFC, 0.9937). Interestingly LPFC neurons showed indices centered at "0" suggesting that the ability of these neurons to encode sensory and mnemonic signals is somewhat similar. On the other hand, neurons in MST and MT showed progressively larger PCI, indicating that coding of sensory signals relative to mnemonic signals increases in strength from LPFC to MST to MT (Tukey–Kramer post hoc test

P values: MT-MST, $2.7518e - 9$; MT-LPFC, $9.56e - 10$; MST-LPFC, $5.2965e - 9$).

Spike Width as a Measurement to Discriminate Between Cell Types

Spike width (NS and BS) has been used to classify macaque monkey neurons in several brain areas into putative interneurons and P cells (Mitchell et al. 2007; Hussar and Pasternak 2009) as well as mouse neurons into P cells, PV, and SST (Ghaderi et al. 2018). However, some studies have shown that some inhibitory interneurons are BS (Casale et al. 2015) and some P cells can be NS (Vigneswaran et al. 2011). Thus, it is possible that our classification into NS and BS does not reliably reflect the intended categories of P cells and inhibitory interneurons.

In order to investigate this issue, we used data from the Allen Institute cell types database (https://celltypes.brain-map.org). We used rheobase spikes evoked by 3 ms square pulse
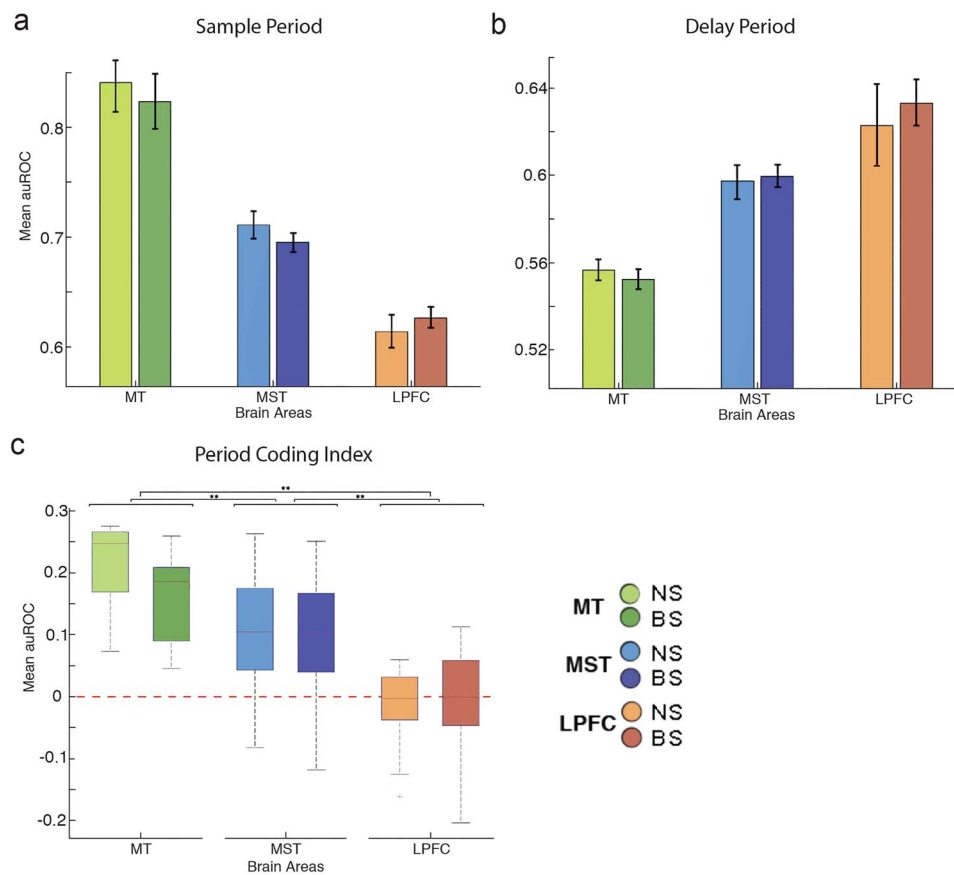
Figure 4. Period coding index (PCI). (*a*) Boxplot figures for PCI for all areas and neuron types. A value above 0 indicates a predominance of sensory coding over mnemonic coding. (*b*) Bar plot showing the proportion of neurons with reverse tuning per area. (*c*) Percentage of neurons displaying reverse tuning per neuron type for all areas.

current injection steps recorded from a sample of 256 human neurons and 941 mouse neurons using whole-cell patch clamp. The human neurons were classified as spiny and aspiny based on post-recording morphological reconstruction. We used the classification conducted by the Allen Institute (https://celltype s.brain-map.org anatomical white paper). In the cortex, spiny neurons are generally P cells, while aspiny neurons are generally interneurons, particularly when reaching adulthood (Ascoli et al. 2008). The mouse neurons were also classified as spiny and aspiny, and aspiny neurons were further classified based on their genetic labeling as PV, VIP, and SST.

Figure 5*a* shows the waveforms of all human neurons. The waveforms of aspiny neurons are in red and the waveforms of spiny neurons are in blue. Note that the intracellular waveforms peak-to-trough does not necessarily correspond to the peak-to-trough of the extracellular waveforms (see Henze et al. 2000). However, the profile of the waveform in intracellular and extracellular recordings is similar; it has a maximum deflection and an afterhyperpolarization. One can use these deflections to measure waveform peak-to-trough in a similar manner as in the extracellular recordings (see Fig. 2 and Methods), shown in Fig. 5*b*.

A sum of two Gaussians was fitted to the data showing bimodality. Using the inflection point between the two Gaussians, we classified the neurons in NS and BS similar as in Figure 2. Because we had the labels of spiny and aspiny for these neurons, we could quantify the number of BS and NS neurons

that were spiny or aspiny. The table on Figure 5*c* shows that almost all spiny neurons fell within the BS category (189/192). However, only half of the NS neurons were classified as spiny (34/64). We conducted the same analysis for the mouse neurons (Fig. 5*d–f*). We found that the large majority of spiny neurons were classified as BS (235/247). On the other hand, only 3/4 of the aspiny neurons (526/694) were classified as NS. Although these numbers are not as dramatic as in the human data, they suggest that a considerable proportion of aspiny neurons (putative inhibitory interneurons) are BS. At this point, it is important to consider that the classification of spiny and aspiny may not fully segregate excitatory and inhibitory interneurons.

To further investigate this issue, a subset of mouse neurons were labeled as PV, SST, and VIP. These data were obtained from studies using transgenic mouse lines (see Allen Institute white paper attached to the Supplementary Material). The waveforms for these neurons are plotted in Figure 5*g–i* with different colors. When plotted, we see that the majority of PV and SST (red and green bars in Fig. 5*h*) neurons are NS, while VIP neurons (blue) fall within the BS category. The table in Figure 5*i* clearly illustrates this result. Most PV neurons (206/207) and the large majority of SST (97/112) were NS, but only the minority of VIP neurons (17/80) were NS. This result strongly suggests that while PV and SST neurons are generally NS, the third category of interneurons, VIP, is usually BS. In order to extrapolate these results to extracellular recordings, one must make two assumptions. First, that the waveform distributions are
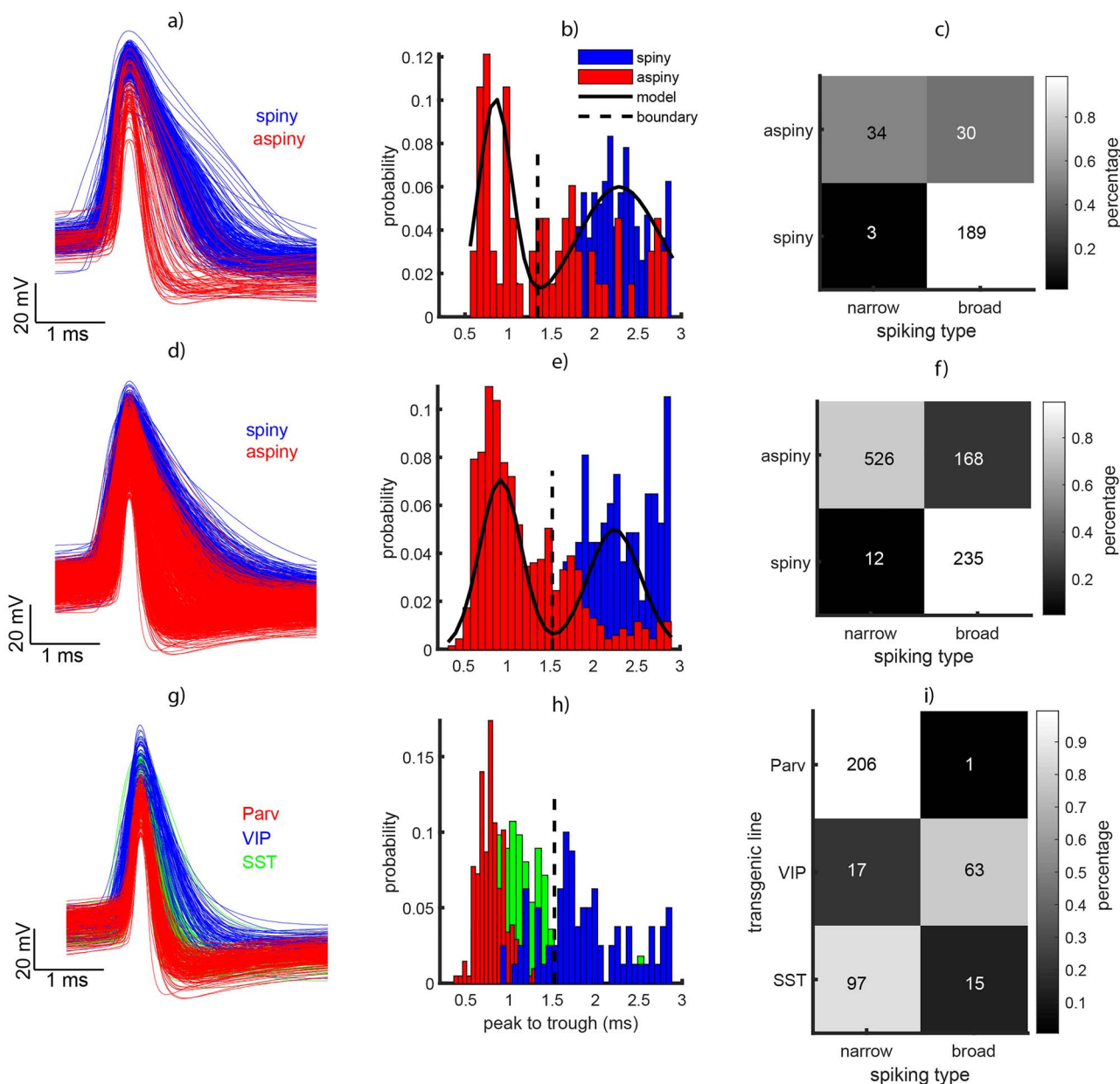
Figure 5. Analysis of waveform type in a sample of human and mouse neurons obtained during intracellular recordings (patch clamp) experiments. (*a, d, g*) Action potential waveforms at rheobase aligned to the deepest deflection of the waveform for human, mouse, and transgenic, respectively. In *a* and *d* narrow spiking (NS) is shown in red and broad spiking (BS) neurons are shown in blue. In *g* the red indicates PV/PV cells, the green SST cells, and the blue VIP cells. (*b, e, h*) Histograms of waveform peak-to-trough for the data in the first column. (*c, f*) Dendritic type (spiny and aspiny) as a function of waveform type (NS and BS). The colors indicate the percentage of neurons. (i) Interneuron type from the transgenic lines (Parv/PV, VIP, and SST) as a function of spiking phenotype (NS and BS).

similar in intracellular and extracellular recordings, and only the duration of the spikes undergoes a transformation that does not affect whether a neuron falls in the NS or BS categories relative to the intersection point between the Gaussians that describe each distribution. Second, that SST and VIP neurons in mice can be considered homologues of CB and CR neurons in primates, respectively (see Discussion section). Considering these are true, the current results suggest that spike peak-to-trough measured during extracellular recordings, as we did in Figure 2, only allows discrimination between two groups of neurons, NS cells that are mainly interneurons and BS cells that are a mix of interneurons and P cells.

Ex Vivo Analysis of Cell Types: Immunohistochemistry

The previous analysis has the limitation that BS cells are likely a mixed population of P cells and CR interneurons. Moreover, our extracellular recordings may introduce, at least theoretically, sample biases in favor of sampling neurons with larger cell bodies (i.e., P cells) and higher spike rates. Finally, extracellular single electrode recordings do not provide precise information about the cortical layer neurons are located. The latter is important considering that models of persistent activity propose that the microcircuitry underlying this phenomenon is located in layers 2/3 of the neocortex (Wang 2013).
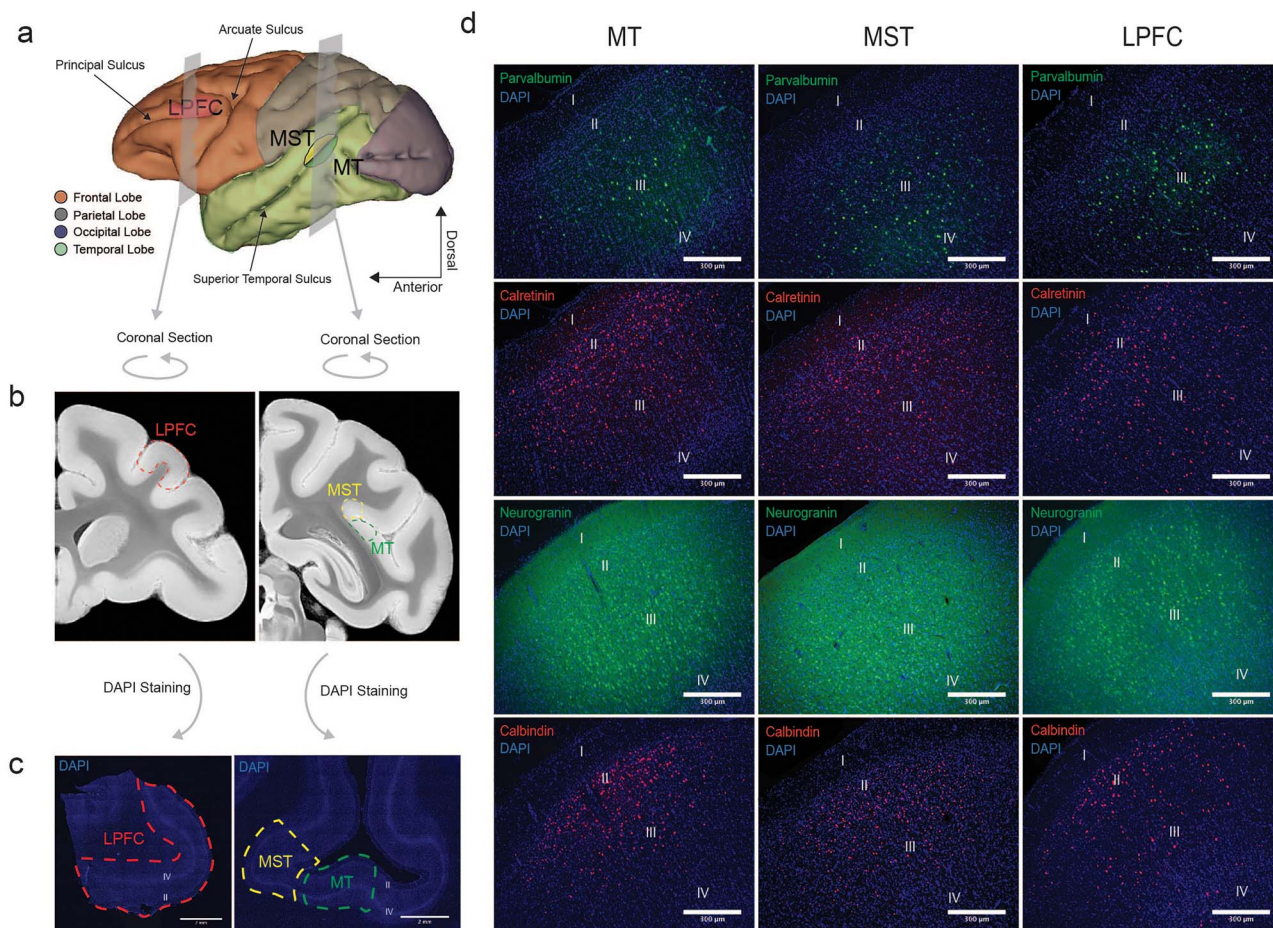
Figure 6. Landmarks for tissue extraction and immunohistochemistry images identifying the areas, layers, and neuronal types. (*a*) MRI reconstruction of the macaque monkey brain with coronal planes representing the regions of MT, MST, and areas 8A/46 that were extracted. (*b*) coronal images taken from the Calabrese et al. (2015) brain atlas that represent the cut surfaces of the extracted tissue blocks. (*c*) DAPI images of the sectioned tissue. Layers 2 and 4 can be identified as lighter bands. (*d*) Example immunohistochemistry images of PV, CR, P cells (Neurogranin), and CB (rows, from top to bottom) across areas MT, MST, and LPFC (columns, from left to right). The different cortical layers are labeled in each image.

In order to address these issues, we performed immunohistochemistry on postmortem tissue slices of MT, MST, and LPFC obtained from two other macaque monkeys. The close proximity of MT and MST areas in the cortical surface allowed us to fit them into one slice. We used antibodies against PV, CR, CB, and Neurogranin (P cells) in layers 2 and 3 of each area (see methods). Layers 2/3 were analyzed because they have been identified by circuit models as the potential location of microcircuits generating persistent activity (Gabbott and Bacon 1996; Dombrowski et al. 2001; Wang 2013; Fig. 6a,b). The cell counts across each pair of layer 2/3 images (see methods) were added to obtain a representation of both layers.

We calculated the proportion of each neuron type within each area. This was taken as the mean count of a given neuron type divided by the added means of all neuron types. The proportions were then compared across brain areas using a one-factor ANOVA (n = 36). Tukey–Kramer post hoc tests were subsequently performed. We found a significant main effect of brain area for PV, $F(2, 105) = 19.61$, $P = 5.81886e - 8$ (Tukey–Kramer post hoc test P values: MT vs. MST, 0.4447; MT vs. LPFC, 0; MST vs. LPFC, 0), and CR, $F(2, 105) = 11.44$, $P = 3.20334e - 5$ (Tukey–Kramer post hoc test P values: MT vs. MST, 0.9988; MT vs. LPFC, 2.0e − 4; MST vs.

LPFC. 2.0e − 4). MT and MST contained similar proportions of PV and CR cells. We observed a small trend for a larger proportion of PV neurons in MT relative to MST, but it was not statistically significant (Fig. 7c). On the other hand, LPFC contains a smaller proportion of PV and a larger proportion of CR neurons than MT and MST.

There was no significant main effect of brain area for CB, $F(2, 105) = 0.63$, $P = 0.535$, or P cells (Neurogranin), $F(2, 105) = 0.81$, $P = 0.449$. Thus, the average percent of CB and P cells remained similar across brain areas (Fig. 7c). In summary, our immunohistochemistry findings suggest a decrease in the relative proportion of PV neurons from MT/MST to LPFC as well as an increase in the relative proportion of CR interneurons from MT/MST to LPFC.

Discussion

In the current study, we report differences in the proportion of different neuronal types classified according to their waveform phenotype (BS and NS) and their immunohistochemical profile (CB, CR, PV, and NG/P cell) in different areas of the primate dorsal visual processing stream. In terms of our electrophysiology
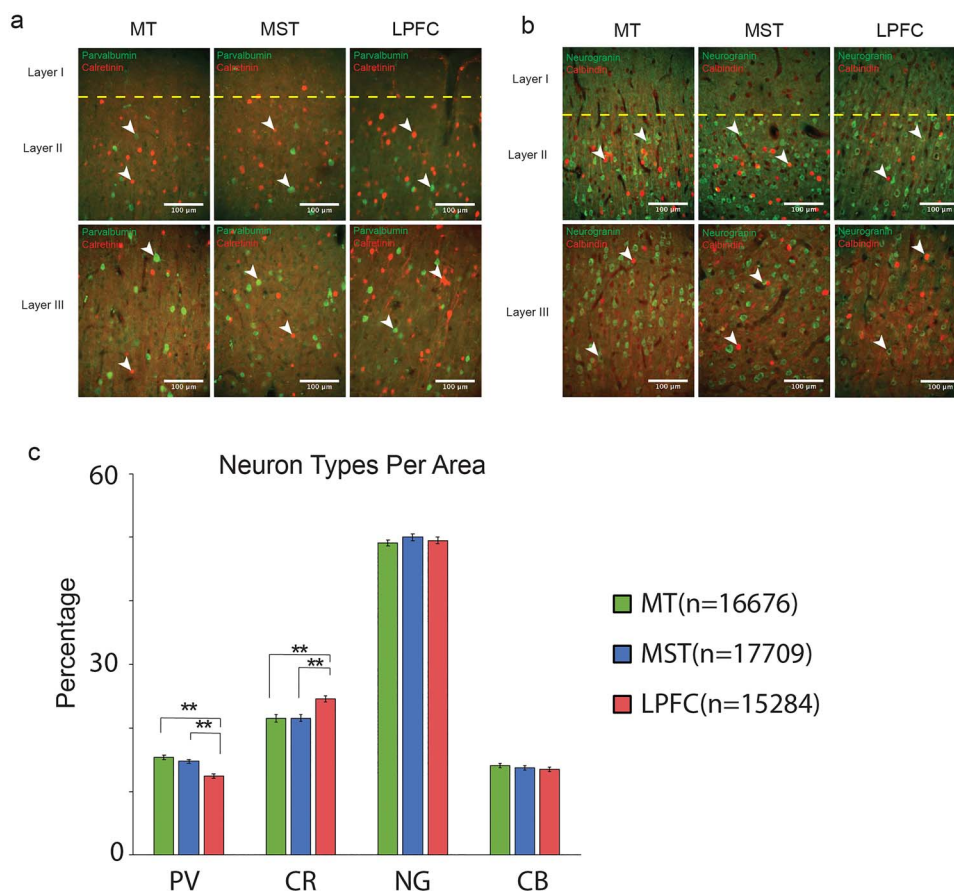
Figure 7. Immunohistochemistry, cell counting, and proportions. (*a*) Example magnified images of PV and CR used for cell counting. The different cortical layers and brain regions are labeled accordingly. (*b*) Example magnified images of P cells (Neurogranin) and CR interneurons used for counting. (*c*) Average percentages of PV, CR, P cells, and CB across MT, MST, and LPFC.

experiments, we found a progressive decrease in the ratio NS/BS neurons from MT to MST to LPFC. A close exploration of the correspondence between the width of the action potential waveform and the neuronal type indicated that while NS neurons were usually PV and SST (putative CB), BS neurons were usually P cells or VIP (putative CR) interneurons. We further conducted immunohistochemistry analysis of cell types in layers 2/3, where persistent activity has been previously reported and found that the proportion of PV interneurons decreased while the proportion of CR interneurons increased from MT and MST to LPFC. The proportions of P cells and CB interneurons remained similar across the different areas.

Previous Studies of Cell Types in Macaques

At least two studies have described electrophysiological properties of interneurons in the monkey LPFC. The first (Zaitsev et al. 2005) studied 81 interneurons in supragranular layers 2/3 of the macaque monkey. Thirty-eight of the neurons showed immunoreactivity for either PV (fast-spiking basket or chandelier cells), CR (not fast-spiking cells with vertically oriented axons), or CB (also not fast spiking with dense axon collaterals in cortical layer 1). They used multiple electrophysiological markers to produce categories of cells via statistical analyses. Among these measurements, the duration of the action potential was one that is more related to the measurements used in our study.

The mean duration of the PV cell action potential was almost half of that of CB and CR cells. The latter two cell types had similar action potential durations. Therefore, according to this parameter, the CB and CR cells could be discriminated from PV cells, but could not be distinguished from one another.

In our study we used the Allen database of mouse data to examine how the width of action potentials relates to cell types. We found that PV and SST cells were NS, but VIP cells were BS, the same as most P cells. Based on the fact that CR neurons have a resemblance to VIP neurons in that they preferentially target other inhibitory neurons in layers 2/3 and are both non-fast spiking, we made the assumption that these cell types may play similar roles in cortical microcircuits in these layers (Meskenaite et al. 1997; Xu et al. 2010; Kennedy et al. 2016). Considering this, the results of Zaitsev et al. (2005) would partially agree with those derived from the transgenic mouse cells shown in our Figure 5.

However, if we were to establish a correspondence between SST and VIP cells in mice to CB and CR cells in monkeys, respectively, the results of our analyses in Figure 5 deviate from those of Zaitsev et al. (2005). In the transgenic mice data from the Allen database, we found that PV cells have shorter peak-to-trough than SST. This agrees with Zaitsev et al.' findings. However, we also found that SST have shorter peak-to-trough than VIP. One possible explanation for this latter discordance is that the duration of the action potential and peak-to-trough duration are

not the same measurement. The former measures the width of the action potential depolarization phase at half height, while the latter includes part of the afterhyperpolarization time of the action potential (trough). Here we use peak-to-trough rather than action potential duration, as Zaitsev et al. used. Another possible explanation is that the sample size used in the study by Zaitsev et al. (81 neurons) was much smaller than in the data extracted from the Allen Institute database ($n = 941$). A larger sample size may have revealed differences between the SST and VIP neurons (putative CB and CR, respectively). To clarify the latter explanation, one must repeat the analyses using a larger sample of monkey neurons. A third possible explanation is that mouse and monkey neurons simply differ in basic features (e.g., the duration of their action potentials) such that the correspondence between SST/CB and VIP/CR neurons cannot be established.

In another study by Zaitsev et al. (2009), the researchers used a larger sample size and grouped 194 interneurons in the LPFC into three different electrophysiological classes. The first class consisted of fast-spiking PV chandelier or harbor cells, which likely correspond to the PV category of the 2005 study, and the PV category of the Allen database mouse data that we present in Figure 5. The second class includes a wide variety of morphological interneuron types, basically the CR and CB cells of the 2005 study; however, here they also observed that some of these interneurons had short spike duration. We speculate that these short-spiking neurons may correspond to the SST neurons with short spike duration identified in the Allen database of waveforms used in our study. The third class is the neurogliaform cell that co-expressed CB and neuropeptide Y. These cells also show a duration of action potential longer than that of PV cells. This study offered more detail about electrophysiological cell types than their previous 2005 study. It also shows the importance of increasing sample size when studying cell types in the neocortex.

It is difficult to compare the results of these studies with our extracellular data and with the data from the Allen Institute database. Here we make the assumption that at least some of the features of the action potential across similar cell types have been preserved through evolution (Hodge et al. 2019). Interestingly, a recent study using extracellular waveform has described four subclasses of neurons in monkey neocortex but did not show a correspondence between these subclasses and specific neuronal types defined by neurotransmitter type, morphology, or protein expression patterns (Trainito et al. 2019). Future work in primates is needed to clarify these outstanding issues.

When interpreting our results, one must take into consideration that intracellular and extracellular recordings are different techniques that use different methods to record neural activity (see Henze et al. 2000). In intracellular recordings, the data contains much less noise than extracellular recordings. This is because intracellular recordings directly measure changes in membrane potential, while extracellular recordings measure changes in the voltage of the extracellular environment "close" to the neuron. These fundamental differences make it difficult to compare the two measurements. Nevertheless, both methods measure action potential features to some extent, and we assume that the data they generate are unique enough to different cell types that the waveform distributions and peak-to-trough measurements of extracellular and intracellular recordings can be mapped into one another. One method to aid in this comparison is the proposal that intracellular waveforms can approximate extracellular ones by computing its first derivative. These assumptions are likely approximations (see Henze et al. 2000 for an extensive discussion of this topic); however, they may allow us to compare the extracellular recordings with intracellular recordings.

In our study we did not directly compare the waveforms in extracellular monkey (our data) and intracellular (mouse and human from the Allen Institute database). We generated peak-to-trough distributions for the different data sets and observed that the distributions are in all cases bimodal (Figs 2 and 5). We assume that the bimodality, and therefore the grouping of NS and BS cells, and the cell types within each one of these two categories are preserved across extracellular and intracellular recordings. One may argue that this is not necessarily the case. For example, PV cells may be NS during intracellular recordings but BS during extracellular recordings. We cannot fully reject this argument; however, some studies have found that during extracellular recordings NS cells produce high baseline firing rates and spike counts, similarly as PV cells during intracellular recordings (Rao et al. 1999; Constantinidis and Goldman-Rakic 2002; Trainito et al. 2019). At least one study in monkeys has examined whether for the particular case of fast-spiking (FS) and non-FS cells the relationship between the width of the action potential during extracellular and intracellular recordings is preserved (Gonzalez-Burgos et al. 2005). They have found this is the case, suggesting that changes in action potential duration from intracellular to extracellular recordings are similar across cell types, and thus the bimodality of the distributions shown in Figs 2 and 5 should be preserved. Future studies are needed to investigate the exact relationship between action potential morphology in different cell types during in vivo and ex vivo recordings.

Changes in Electrophysiological Phenotypes along the Dorsal Visual Pathway

Several studies have exploited differences in action potential dynamics to classify neuron subtypes based exclusively on extracellular recordings (Wilson et al. 1994; Constantinidis and Goldman-Rakic 2002; Nowak et al. 2003; Bartho et al. 2004). Some studies have used the terms putative pyramidal neurons and putative interneurons to refer to the BS and NS neurons (Mitchell et al. 2007; Hussar and Pasternak 2009). This has been put into question by studies showing that some inhibitory interneurons are BS (Casale et al. 2015) and some P cells are NS (Vigneswaran et al. 2011). However, these results have been difficult to reconcile since the different studies were conducted using different techniques (e.g., in vivo and in vitro electrophysiology) or in different areas (e.g., V4 and LPFC vs motor cortex), and the data have been analyzed using different quantification strategies. We used the NS/BS classification in our extracellular recordings based on different parameters of the waveform. Our electrophysiological data revealed a decrease in NS relative to BS neurons from MT to MST to LPFC, leading us to the interpretation that putative interneurons were less abundant than putative pyramidal neurons as one progresses downstream from sensory areas (i.e., MT-MST-LPFC).

We further conducted an examination of waveform width (peak-to-trough) and its correspondence to neuronal types using in vitro data. We used the Allen cell types database that contains a large sample of human neurons and mouse neurons that have been tested using a standard patch clamp protocol. According to their dendrite type, neurons can be classified in spiny or aspiny, being the former putative pyramidal and the

latter putative GABAergic interneurons (Ascoli et al. 2008). In the Allen database, neurons have been classified into spiny and aspiny using fluorescent microscopy (http://celltypes.brain-ma p.org). Moreover, the database contains transgenic mouse lines data which allows the identification of PV, SST, and VIP positive GABAergic interneurons. We found that while PV and SST cells are generally NS, VIP neurons are BS, similar to P cells (Fig. 5). Interestingly, at least one study in the rat has documented that NS neurons could have at least two different waveform phenotypes (Quirk et al. 2009). These may correspond to the PV and SST neurons in Figure 5h. A recent study has also proposed that neurons can be classified into four different phenotypes (Trainito et al. 2019). These may correspond to the phenotypes PV, SST, VIP, and P cells of our study. However, this issue needs further investigation.

One somewhat surprising finding of our study is that in area MT, about half of the cells are NS and the other half are BS. Considering that NS cells are putative inhibitory interneurons and BS cells can be either inhibitory or excitatory, these findings contradict the widely spread idea that the proportion of excitatory to inhibitory cell types in the neocortex is 80% excitatory and 20% inhibitory. One may speculate this is because recordings with extracellular electrodes favor sampling a larger number of interneurons in early sensory areas (particularly PV cells) compared with associative areas such as the LPFC. This is because P cells in the LPFC are substantially larger than P cells in V1 (Luebke 2017; Elston 2003 for a review) (and likely MT), and the extracellular recording technique favors larger cells that produce larger electrical fields. Thus, LPFC electrodes will oversample large P cells that are BS and under sample smaller, inhibitory interneurons that are NS. The opposite would occur in early sensory areas such as MT. In order for this hypothesis to be veridical, we would need to assume that NS interneuron size does not change across different areas, or at least not in the same manner as the P cell. We found no data indicating that this is the case; therefore this hypothesis remains to be tested.

It is also possible that the sampling with single electrodes, as in the case of MT and MST recordings, is biased towards certain cortical layers, where PV cells are more frequently found (see section below), while the microelectrode arrays used in LPFC may sample neurons from more superficial layers where CR cells seem to be more abundant (see Fig. 7). Our results highlight the limitations and challenges of identifying neuronal types based solely on extracellular recordings. Thus, anatomical reconstruction or molecular characterization of neurons may be necessary to accurately classify neuron types (González-Burgos et al. 2005; Sharpee 2014). The latter reveals a gap between electrophysiological studies in vivo and in vitro that needs to be addressed in the future.

Immunohistochemical Profile of Single Neurons

We used immunohistochemistry to physically stain and visualize the cell bodies of PV, CR, CB, and P cells (Neurogranin). About 90% of interneurons in the primate cortex express one type of these calcium-binding proteins (DeFelipe 1997; Raghanti et al. 2010). We concentrated in layers 2/3 of each area, where neurons showing persistent activity encoding working memory have been isolated (Bastos et al. 2018). We reported a P cell composition of ~50% and CB of ~14%. These findings in layers 2/3 do not necessarily contradict previous reports of CB interneurons comprising ~5% of the neurons across the cortex (Conde et al. 1994) and P comprising ~70–80% (DeFelipe and

Fariñas 1992; Markram et al. 2004). They can be explained by a higher concentration of interneurons in layers 2/3 relative to other layers (Gabbott and Bacon 1996; Dombrowski et al. 2001).

We found that CR was the most abundant interneuron in layer 2/3 of MT, MST, and LPFC (Fig. 6d). This was anticipated in the LPFC (Conde et al. 1994; Gabbott and Bacon 1996), but it is a novel finding in MT and MST. MT is close to the primary visual cortex, which has previously been reported to have an interneuron composition of ~75% PV, ~10% CR, and ~10% CB (Brederode et al. 1990; Meskenaite 1997; Whissell et al. 2015). However, our results can be explained by the fact that the aforementioned studies did not conduct a layer-specific analysis. The latter highlights the importance of conducting layer-specific analyses of neuronal type densities in the primate neocortex.

Interestingly, Kim et al. (2017) described a principle of cortical organization in mice in which sensory-motor areas are dominated by output-modulating PV interneurons, while association areas are dominated by input-modulating somatostatin (STT) positive interneurons. The SST interneuron in mice corresponds to the CB neuron in primates (Ascoli et al. 2008). Our results partially support this principle, in that we found a decrease in the relative proportion of PV cells from MT/MST to LPFC. However, we found no change in the relative proportion of CB interneurons across areas. This may be due to differences between rodents and primates in terms of neuronal type densities, or because of the fact we did not explore as many areas as the other authors did, or because the correspondence between SST, VIP in mice, and CB and CR in primates is not clear.

Nevertheless, based on the microcircuit proposed by Wang (2018) and the findings in these studies, one could reasonably argue that a candidate for the changes in microcircuits that contribute to the emergence of persistent activity supporting WM is changes in the proportions of PV and CR interneurons. Whether this is a trend that generalizes across species and across cortices containing neurons encoding different sensory modalities (Hwang and Romanski 2015; Vergara et al. 2016) requires further investigation. Interestingly, a recent study has reported that pyramidal neurons in the parietal and prefrontal cortex of monkeys show different intrinsic properties (Gonzalez-Burgos et al. 2019), further suggesting that changes in circuit composition are more the norm than the exception across areas of the primate neocortex.

One issue that remains unclear is the lack of differences in the proportion of different interneuron types between areas MT and MST. There was a tendency for the PV interneurons to be in a smaller proportion in MST than in MT, but it was not significant (Fig. 7c). MST neurons have been shown to produce persistent activity encoding WM representations as well as sensory responses encoding sensory features of stimuli that are visually available. However, when comparing MST and LPFC, it is clear that MST coding of sensory features is stronger than the coding of mnemonic features, while in LPFC neurons similarly encode sensory and mnemonic features (Fig. 4c). Thus, it may be that the proportion of CR to PV interneurons cannot solely account for the ability to encode WM representations, but other factors may play a role. MST area may represent a transition in architecture and functional properties of cortical microcircuits between early sensory areas such as MT and executive areas such as LPFC.

Our results do not demonstrate a causal relationship between changes in the proportion of interneuron types in layers 2/3 across cortical areas and the origin of persistent activity encoding WM; however, the fact that sensory and

association areas differ in this respect may be linked to the origins of persistent activity encoding WM (Wang 2018). One issue that needs to be clarified is the specific role of particular interneuron types (e.g., CR) in the origin of persistent activity encoding WM. Future studies should address this issue.

Other Variables Potentially Contributing to the Emergence of Persistent Activity Encoding WM

Variables such as changes in P cell morphology, interneuron connectivity, and network biophysics may change across different brain areas and contribute to the emergence of persistent activity downstream from early sensory cortices. For instance, previous studies have reported an increase in size and structural complexity of the dendritic arbor, spike discharges, and spine densities of P cells from early visual cortices to prefrontal cortices (Elston 2000; Elston and Rosa 2000; Elston et al. 2005; Amatrudo et al. 2012). These changes can increase the strength of input signals to P cells within the microcircuits, which may increase positive feedback activation and result in persistent activity. In favor of this hypothesis, it has been reported that LPFC P cells have an increase in the number of dendritic spines relative to neurons in the lateral intraparietal area LIP (Gonzalez-Burgos et al. 2019).

Changes in interneuron connectivity may also contribute to the emergence of persistent activity. For example, increased connectivity between CR and CB would result in greater inhibition of CB interneurons, thereby increasing the excitability of P cells by incoming inputs and contributing to the development of persistent activity. Future research could investigate differences in interneuron and P cell connectivity across brain areas in primates by co-staining neuron cell markers (e.g., CB, PV, CR, Neurogranin) with excitatory and inhibitory vesicular transporters, such as VGAT and VGLUT2. This would provide an idea of excitatory and inhibitory connectivity within area microcircuits.

Quantitative changes in network biophysics (the molecular and cellular processes of a cortical microcircuit) may also cause one brain area to exhibit persistent activity. For example, a recent study has reported differences in the timescales in which neurons across the macaque visual processing stream integrate information (Murray et al. 2014). The researchers analyzed data from several serially connected brain areas and applied autocorrelation analysis to show that the time constant of temporal integration is shorter in sensory visual areas relative to association areas. They proposed that longer timescales in neurons with strong excitatory connections favor recurrent dynamics and lower current leakage, which allows neural activity to locally reverberate within a circuit.

It has also been suggested that slow excitatory dynamics relative to inhibitory feedback dynamics, likely mediated by the ratio of NMDA to AMPA receptors, may facilitate persistent activity and WM coding (Goldman et al. 2009). Such slow dynamics are primarily modulated by NMDA receptors, whose relatively slow decay time constant (50–100 ms) and unbinding help to stabilize reverberatory activity without leading to runoff excitatory drive (Wang 1999, 2013; Lim and Goldman 2013). Consistent with this model, antagonists for the NMDA receptor have been linked to impairments in WM performance (Krystal 1994; Wang et al. 2013). Interestingly, prefrontal P cells express a larger amount of mRNA for NMDA receptor subunits than P cells in primary visual cortices (Elston and Rosa 2000). These factors may contribute to the emergence of persistent activity in areas such as the LPFC, downstream from sensory cortices.

Conclusion

Our study shows changes in the proportion of different interneurons types, defined according to electrophysiological and immunohistochemical profiles, along different areas of the macaque dorsal visual pathway. Our results help to bridge the gap between functional observations obtained from neurophysiological extracellular recordings and the structural organization of microcircuits across brain areas obtained using histological and anatomical techniques. Finally, our results allow the refinement of computational models of cortical microcircuits underlying WM.

Supplementary Material

Supplementary material is available at *Cerebral Cortex* online.

Notes

Conflict of Interest: The authors declare no competing interests.

References

Ascoli GA, Alonso-Nanclares L, Anderson SA, Barrionuevo G, Benavides-Piccione R, Burkhalter A, Buzsaki G, Cauli B, DeFelipe J, Fairen A, et al. 2008. Petilla terminology: nomenclature of features of GABAergic interneurons of the cerebral cortex. *Nat Rev Neurosci*. 9(7):557.

Akaike H. 1974. A new look at the statistical model identification. *IEEE T Automat Contr*. 19:716–723.

Amatrudo JM, Weaver CM, Crimins JL, Hof PR, Rosene DL, Luebke JI. 2012. Influence of highly distinctive structural properties on the excitability of pyramidal neurons in monkey visual and prefrontal cortices. *J Neurosci*. 32:13644–13660.

Ardid S, Vinck M, Kaping D, Marquez S, Everling S, Womelsdorf T. 2015. Mapping of functionally characterized cell classes onto canonical circuit operations in primate prefrontal cortex. *J Neurosci*. 35:2975–2991.

Bartho P, Hirase M, Monoconduit L, Zugaro M, Harris KD, Buzsaki G. 2004. Characterization of neocortical principle cells and interneurons by network interactions and extracellular features. *J Neurophysiol*. 92:600–608.

Bastos AM, Loonis R, Kornblith S, Lundqvist M, Miller EK. 2018. Laminar recordings in frontal cortex suggest distinct layers for maintenance and control of working memory. *Proc Natl Acad Sci USA*. 115(5):1117–1122.

Brederode JF, Mulligan VKA, Hendrickson AE. 1990. Calcium-binding proteins as markers for subpopulations of GABAergic neurons in monkey striate cortex. *J Comp Neurol*. 298:1–22.

Calabrese E, Badea A, Coe CL, Lubach GR, Shi Y, Styner MA, Johnson A. 2015. A diffusion tensor MRI atlas of the post-mortem rhesus macaque brain. *Neuroimage*. 117:408–416. doi: 10.1016/j.neuroimage.2015.05.072.

Casale AE, Foust AJ, Bal T, McCormick DA. 2015. Cortical interneuron subtypes vary in their axonal action potential properties. *J Neurosci*. 35:15555–15567.

Cauli B, Audinat E, Lambolez B, Angulo MC, Ropert N, Tsuzuki K, Hestrin S, Rossier J. 1997. Molecular and physiological diversity of cortical nonpyramidal cells. *J Neurosci*. 17: 3894–3906.

Chaudhuri R, Knoblauch K, Gariel MA, Kennedy H, Wang XJ. 2015. A large-scale circuit mechanism for hierarchical dynamical processing in the primate cortex. *Neuron*. 88(2): 419–431.

Conde F, Lund JS, Jacobowitz DM, Baimbridge KG, Lewis DA. 1994. Local circuit neurons immunoreactive for calretinin, calbindin D-28k or parvalbumin in monkey prefrontal cortex—distribution and morphology. *J Comp Neurol*. 341: 95–116.

Constantinidis C, Goldman-Rakic PS. 2002. Correlated discharges among putative pyramidal neurons and interneurons in the primate prefrontal cortex. *J Neurophysiol*. 88:3487–3490.

Constantinidis C, Funahashi S, Lee D, Murray JD, Qi XL, Wang M, Arnsten AFT. 2018. Persistent spiking activity underlies working memory. *J Neurosci*. 38(32):7020–7028.

DeFelipe J. 1997. Types of neurons, synaptic connections and chemical characteristics of cells immunoreactive for calbindin-D28K, parvalbumin and calretinin in the neocortex. *J Chem Neuroanat*. 14(1):1–19.

DeFelipe J, Fariñas I. 1992. The pyramidal neuron of the cerebral cortex: morphological and chemical characteristics of the synaptic inputs. *Prog Neurobiol*. 39(6):563–607.

Dombrowski SM, Hilgetag CC, Barbas H. 2001. Quantitative architecture distinguishes prefrontal cortical systems in the rhesus monkey. *Cereb Cortex*. 11:975–988.

Elston GN. 2000. Pyramidal cells of the frontal lobe: all the more spinous to think with. *J Neurosci*. 20:RC95.

Elston GN, Rosa MGP. 1997. The occipitoparietal pathway of the macaque monkey: comparison of pyramidal cell morphology in layer III of functionally related cortical visual areas. *Cereb Cortex*. 7(5):432–452.

Elston GN, Rosa MGP. 2000. Pyramidal cells, patches, and cortical columns: a comparative study of infragranular neurons in TEO, TE, and the superior temporal polysensory area of the macaque monkey. *J Neurosci*. 20:RC117.

Elston GN. 2003. Cortex, cognition and the cell: new insights into the pyramidal neuron and prefrontal function. *Cereb Cortex*. 13(11):1124–1138.

Elston GN, Benavides-Piccione R, DeFelipe J. 2005. A study of pyramidal cell structure in the cingulate cortex of the macaque monkey with comparative notes on inferotemporal and primary visual cortex. *Cereb Cortex*. 15:64–73.

Elston GN, Benavides-Piccione R, Elston A, Manger P, DeFelipe J. 2011. Pyramidal cells in prefrontal cortex of primates: marked differences in neuronal structure among species. *Front Neuroanat*. 5:2.

Freeman JB, Dale R. 2013. Assessing bimodality to detect the presence of a dual cognitive process. *Behav Res Ther*. 45:83–97.

Freedman DJ, Assad JA. 2016. Neuronal mechanisms of visual categorization: an abstract view on decision making. *Annu Rev Neurosci*. 39:129–147.

Fuster JM. 2008. *The prefrontal cortex*. 4th ed. London: Elsevier.

Gabbott PLA, Bacon SJ. 1996. Local circuit neurons in the medial prefrontal cortex (areas 24a, b, c, 25 and 32) in the monkey. Quantitative areal and laminar distributions. *J Comp Neurol*. 364:609–636.

Ghaderi P, Marateb HR, Safari MS. 2018. Electrophysiological profiling of neocortical neural subtypes: a semi-supervised method applied to in vivo whole-cell patch-clamp data. *Front Neurosci*. 12:823.

Galarreta M, Hestrin SA. 2002. Electrical and chemical synapses among parvalbumin fast-spiking GABAergic interneurons in adult mouse neocortex. *Proc Natl Acad Sci USA*. 99:12438–12443.

Goldman-Rakic PS. 1995. Cellular basis of working memory. *Neuron*. 14(3):477–485.

Goldman MS, Compte A, Wang XJ. 2009. Neural integrator models. *Encyclopedia Neurosci*. 6:165–178.

Goldman-Rakic PS. 1990. Cellular and circuit basis of working memory in prefrontal cortex of nonhuman primates. *Prog Brain Res*. 85:325–336.

González-Burgos G, Barrionuevo G, Lewis DA. 2000. Horizontal synaptic connections in monkey prefrontal cortex: an in vitro electrophysiological study. *Cereb Cortex*. 10(1):82–92.

González-Burgos G, Krimer LS, Povysheva NV, Barrionuevo G, Lewis DA. 2005. Functional properties of fast spiking interneurons and their synaptic connections with pyramidal cells in primate dorsolateral prefrontal cortex. *J Neurophysiol*. 93(2):942–953.

González-Burgos G, Miyamae T, Krimer Y, Gulchina Y, Pafundo DE, Krimer O, Bazmi H, Arion D, Enwright JF, Fish KN, et al. 2019. Distinct properties of layer 3 pyramidal neurons from prefrontal and parietal areas of the monkey neocortex. *J Neurosci*. 39(37):7277–7290.

Guadaño-Ferraz A, Viñuela A, Oeding G, Bernal J, Rausell E. 2005. RC3/neurogranin is expressed in pyramidal neurons of motor and somatosensory cortex in normal and denervated monkeys. *J Comp Neurol*. 493(4):554–570.

Hartigan JA, Hartigan PM. 1985. The dip test of unimodality. *Ann Stat*. 13:70–84.

Henze DA, Borhegyi Z, Csicsvari J, Mamiya A, Harris KD, Buzsáki G. 2000. Intracellular features predicted by extracellular recordings in the hippocampus in vivo. *J Neurophysiol*. 84(1):390–400.

Hodge RD, Bakken TE, Miller JA, Smith KA, Barkan ER, Graybuck LT, Close JL, Long B, Johansen N, Penn O, et al. 2019. Conserved cell types with divergent features in human versus mouse cortex. *Nature*. 573(7772):61–68.

Hussar CR, Pasternak T. 2009. Flexibility of sensory representations in prefrontal cortex depends on cell type. *Neuron*. 64(5):730–743.

Hwang J, Romanski LM. 2015. Prefrontal neuronal responses during audiovisual mnemonic processing. *J Neurosci*. 35:960–971.

Kawaguchi YM, Katsumaru H, Kosaka T, Heizmann CW, Hama K. 1987. Fast spiking cells in rat hippocampus (CA1 region) contain the calcium-binding protein parvalbumin. *Brain Res*. 416:369–374.

Kennedy H, Van Essen DC, Christen Y, Callaway E. 2016. Inhibitory cell types, circuits and receptive fields in mouse visual cortex. *Micro- meso-, and macro-connectomics of the brain*. Cham (CH): Springer.

Khayat PS, Niebergall R, Martinez-Trujillo JC. 2010. Frequency-dependent attentional modulation of local field potentials in macaque area MT. *J Neurosci*. 30:7037–7048.

Kim Y, Yang GR, Pradhan K, Venkataraju KU, Bota M, García Del Molino LC, Fitzgerald G, Ram K, He M, Levine JM, et al. 2017. Brain-wide maps reveal stereotyped cell-type-based cortical architecture and subcortical sexual dimorphism. *Cell*. 171(2):456–469.

Krystal JH, Karper LP, Seibyl JP, Freeman GK, Delaney R, Bremner JD, Heninger GR, Bowers MB, Charney DS. 1994. Subanesthetic effects of the noncompetitive NMDA antagonist, ketamine, in humans. Psychotomimetic, perceptual, cognitive, and neuroendocrine responses. *Arch Gen Psychiatry*. 51:199–214.

Leavitt M, Mendoza-Halliday D, Martinez-Trujillo JC. 2017. Persistent activities encoding working memories: not fully distributed. *Trends Neurosci*. 40:328–346.

Lennert T, Martinez-Trujillo JC. 2011. Strength of response suppression to distracter stimuli determines attentional-filtering performance in primate prefrontal neurons. *Neuron*. 70:141–152.

Lim S, Goldman MS. 2013. Balanced cortical microcircuitry for maintaining information in working memory. *Nat Neurosci*. 16:1306–1314.

Luebke JI. 2017. Pyramidal Neurons Are Not Generalizable Building Blocks of Cortical Networks. *Front Neuroanat*. 11:11.

Lundqvist M, Herman P, Miller EK. 2018. Working memory: delay activity, yes! Persistent activity? Maybe not. *J Neurosci*. 38(32):7013–7019.

Markram H, Toledo-Rodriguez M, Wang Y, Gupta A, Silberberg G, Wu C. 2004. Interneurons of the neocortical inhibitory system. *Nat Rev Neurosci*. 5:793–807.

McCormick DA, Connors BW, Lighthall JW, Prince DA. 1985. Comparative electrophysiology of pyramidal and sparsely spiny stellate neurons of the neocortex. *J Neurophysiol*. 54:782–806.

Mendoza-Halliday D, Torres S, Martinez-Trujillo JC. 2014. Sharp emergence of feature-selective sustained activity along the dorsal visual pathway. *Nat Neurosci*. 17:1255–1262.

Mendoza-Halliday D, Martinez-Trujillo JC. 2017. Neuronal population coding of perceived and memorized visual features in the lateral prefrontal cortex. *Nat Commun*. 8:1–13.

Meskenaite V. 1997. Calretinin-immunoreactive local circuit neurons in area 17 of the cynomolgus monkey, Macaca fascicularis. *J Comp Neurol*. 379(1):113–132.

Mitchell JF, Sundberg KA, Reynolds JH. 2007. Differential attention-dependent response modulation across cell classes in macaque visual area V4. *Neuron*. 55:131–141.

Murray JD, Bernacchia A, Freedman DJ, Romo R, Wallis JD, Cai X, Padoa-Schioppa C, Pasternak T, Seo H, Lee D, et al. 2014. A hierarchy of intrinsic timescales across primate cortex. *Nat Neurosci*. 17:1661–1663.

Nowak LG, Azouz R, Sanchez-Vives MV, Gray CM, McCormick DA. 2003. Electrophysiological classes of cat primary visual cortical neurons in vivo as revealed by quantitative analyses. *J Neurophysiol*. 89:1541–1566.

Paxinos G, Huang XW, Toga AW. 2000. *The rhesus monkey brain in stereotaxic coordinates*. San Diego, USA: Academic Press.

Petrides M. 2005. Lateral prefrontal cortex: architectonic and functional organization. *Philos Trans R Soc Lond B Biol Sci*. 360(1456):781–795.

Pollock E, Everest M, Brown A, Poulter MO. 2014. Metalloproteinase inhibition prevents inhibitory synapse reorganization and seizure genesis. *Neurobiol Dis*. 70:21–31.

Quirk MC, Sosulski DL, Feierstein CE, Uchida N, Mainen ZF. 2009. A defined network of fast-spiking interneurons in orbitofrontal cortex: responses to behavioral contingencies and ketamine administration. *Front Syst Neurosci*. 3:13.

Rao SG, Williams GV, Goldman-Rakic PS. 1999. Isodirectional tuning of adjacent interneurons and pyramidal cells during working memory: evidence for microcolumnar organization in PFC. *J Neurophysiol*. 81(4):1903–1916.

Raghanti MA, Spocter MA, Butti C, Hof PR, Sherwood CC. 2010. A comparative perspective on minicolumns and inhibitory GABAergic interneurons in the neocortex. *Front Neuroanat*. 4:3.

Roux L, Buzsaki G. 2015. Tasks for inhibitory interneurons in intact brain circuits. *Neuropharmacology*. 88:10–23.

Runyan CA, Schummers J, VanWart A, Kuhlman SJ, Wilson NR, Huang ZJ, Sur M. 2010. Response features of Parvalbumin-expressing interneurons suggest precise roles for subtypes of inhibition in visual cortex. *Neuron*. 67:847–857.

Sclar G, Maunsell JH, Lennie P. 1990. Coding of image contrast in central visual pathways of the macaque monkey. *Vision Res*. 46:457–461.

Sharpee T. 2014. Toward functional classification of neuronal types. *Neuron*. 83:1329–1334.

Singec I, Knoth R, Ditter M, Volk B, Frotscher M. 2004. Neurogranin is expressed by pyramidal cells but not interneurons in the rodent and monkey neocortex and hippocampus. *J Comp Neurol*. 479(1):30–42.

Trainito C, von Nicolai C, Miller EK, Siegel M. 2019. Extracellular spike waveform dissociates four functionally distinct cell classes in primate cortex. *Curr Biol*. 29(18):2973.e5–2982.e5. doi: 10.1016/j.cub.2019.07.051.

Tremblay R, Lee S, Rudy B. 2016. GABAergic interneurons in the Neocortex: from cellular properties to circuits. *Neuron*. 91(2):260–292.

Treue S, Martinez-Trujillo JC. 1999. Feature-based attention influences motion processing gain in macaque visual cortex. *Nature*. 399:575–579.

Vergara J, Rivera N, Rossi-Pool R, Romo R. 2016. A neural parametric code for storing information of more than one sensory modality in working memory. *Neuron*. 89:54–62.

Vigneswaran G, Kraskov A, Lemon RN. 2011. Large identified pyramidal cells in macaque motor and premotor cortex exhibit "thin spikes": implications for cell type classification. *J Neurosci*. 31(40):14235–14242.

Wang B, Yin L, Zou X, Ye M, Liu Y, He T, Deng S, Jiang Y, Zheng R, Wang Y, et al. 2015. A subtype of inhibitory interneuron with intrinsic persistent activity in human and monkey neocortex. *Cell Rep*. 10:1450–1458.

Wang XJ. 1999. Synaptic basis of cortical persistent activity: the importance of NMDA receptors to working memory. *J Neurosci*. 19:9587–p603.

Wang XJ. 2001. Synaptic reverberation underlying mnemonic persistent activity. *Trends Neurosci*. 24:455–463.

Wang XJ. 2006. A microcircuit model of prefrontal functions: ying and yang of reverberatory neurodynamics in cognition. *The Prefrontal Lobes: Development, Function and Pathology*. 92–127.

Wang XJ. 2013. The prefrontal cortex as a quintessential "cognitive-type" neural circuit: working memory and decision-making. *Principles of frontal lobe function*. Cambridge (UK): Cambridge University Press, pp. 226–248.

Wang XJ, Tegner J, Constantinidis C, Goldman-Rakic PS. 2004. Division of labor among distinct subtypes of inhibitory neurons in a cortical microcircuit of working memory. *Proc Natl Acad Sci USA*. 101:1368–1373.

Wang XJ, Yang GR. 2018. A disinhibitory circuit motif and flexible information routing in the brain. *Curr Opin Neurol*. 49: 75–83.

Wang M, Yang Y, Wang CJ, Gamo NJ, Jin LE, Mazer JA, Morrison JH, Wang XJ, Arnsten AF. 2013. NMDA receptors subserve persistent neuronal firing during working memory in dorsolateral prefrontal cortex. *Neuron*. 77(4): 736–749.

Whissell PD, Cajanding JD, Fogel N, Kim JC. 2015. Comparative density of CCK- and PV-GABA cells within the cortex and hippocampus. *Front Neuronanat*. 9:124.

Williams SM, Goldman-Rakic PS, Leranth C. 1992. The synaptology of parvalbumin−immunoreactive neurons in the primate prefrontal cortex. *J Comp Neurol*. 320(3):353–369.

Wilson FA, O'Scalaidhe SP, Goldman-Rakic PS. 1994. Functional synergism between putative gamma-aminobutyrate-containing neurons and pyramidal neurons in prefrontal cortex. *Proc Natl Acad Sci USA*. 91:4009–4013.

Zaitsev AV, Gonzales-Burgos G, Povysheva NV, Kroner S, Lewis DA, Krimer LS. 2004. Localization of calcium-binding proteins in physiologically and morphologically characterized interneurons of monkey dorsolateral prefrontal cortex. *Cereb Cortex*. 15:1178–1186.

Zaitsev AV, Gonzalez-Burgos G, Povysheva NV, Kröner S, Lewis DA, Krimer LS. 2005. Localization of calcium-binding proteins in physiologically and morphologically characterized interneurons of monkey dorsolateral prefrontal cortex. *Cereb Cortex*. 15(8):1178–1186.

Zaitsev AV, Povysheva NV, Gonzalez-Burgos G, Rotaru D, Fish KN, Krimer LS, Lewis DA. 2009. Interneuron diversity in layers 2-3 of monkey prefrontal cortex. *Cereb Cortex*. 19(7): 1597–1615.

Zhu Y, Qiao W, Liu K, Zhong H, Yao H. 2015. Control of response reliability by parvalbumin-expressing interneurons in visual cortex. *Nat Commun*. 6:1–11.

Zaksas D, Pasternak T. 2006. Directional signals in the prefrontal cortex and in area MT during a working memory for visual motion task. *J Neurosci*. 26(45):11726–11742.

Zylberberg J, Strowbridge BW. 2017. Mechanisms of persistent activity in cortical circuits: possible neural substrates for working memory. *Annu Rev Neurosci*. 40:603–627.

Xiangmin Xu, Keith DR, Edward MC. 2010. Immunochemical characterization of inhibitory mouse cortical neurons: Three chemically distinct classes of inhibitory cells. *J Comp Neurol. J Comp Neurol*. 518(3):389–404.

Luebke JI. 2017. Pyramidal Neurons Are Not Generalizable Building Blocks of Cortical Networks. *Front Neuroanat*. 11:11.

Elston GN. 2003. Cortex, cognition and the cell: new insights into the pyramidal neuron and prefrontal function. *Cereb Cortex*. 13 (11):1124–38. Review.