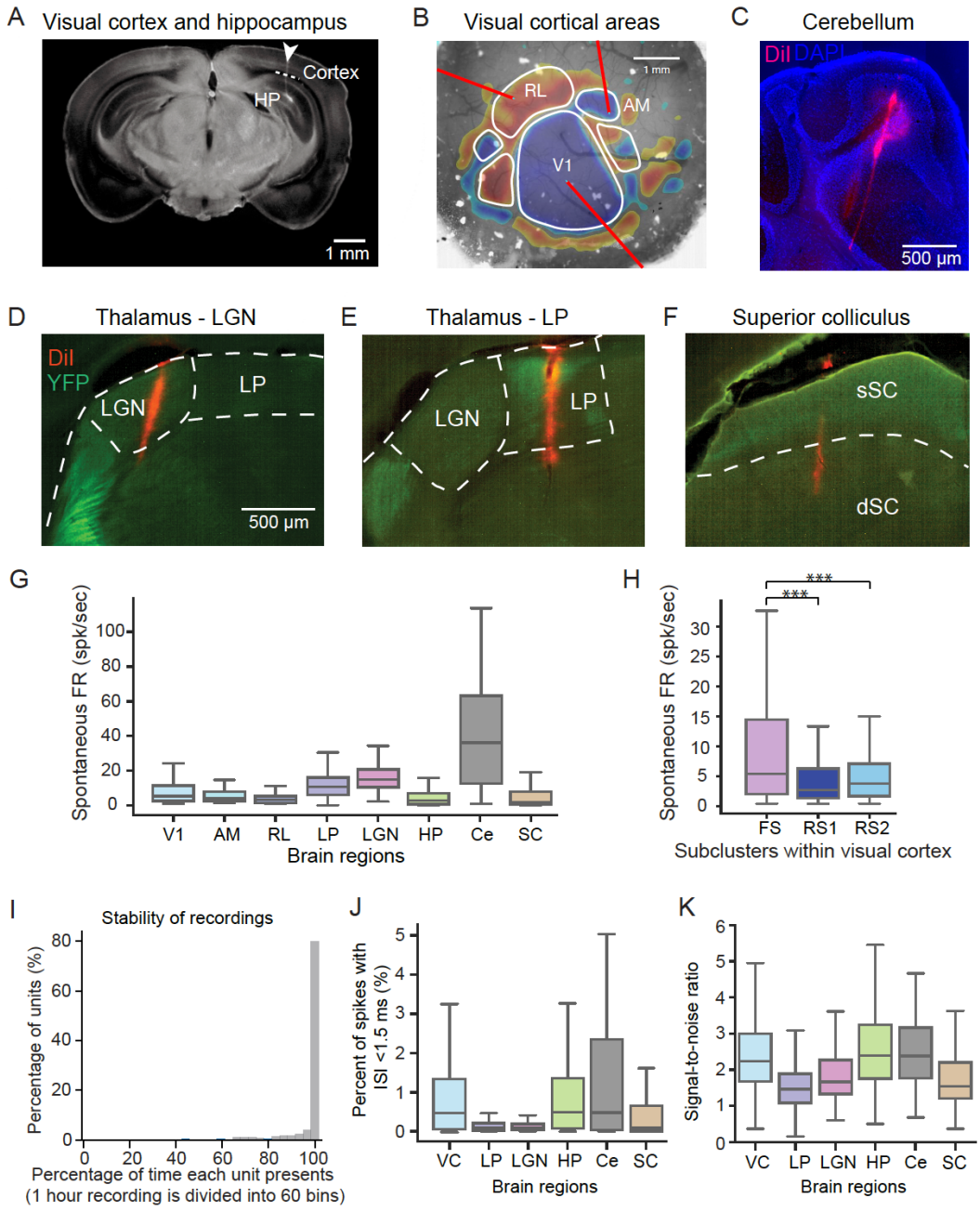
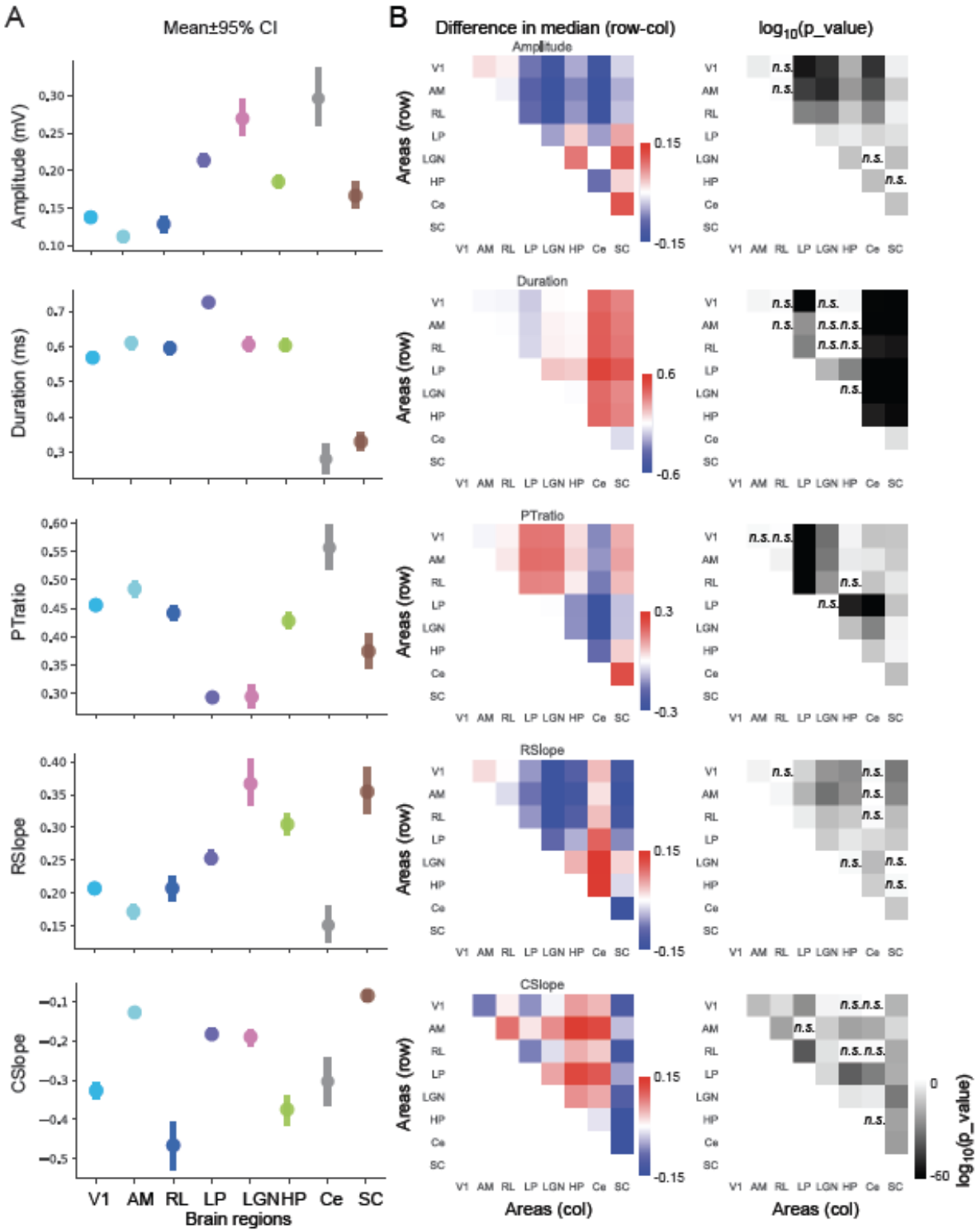
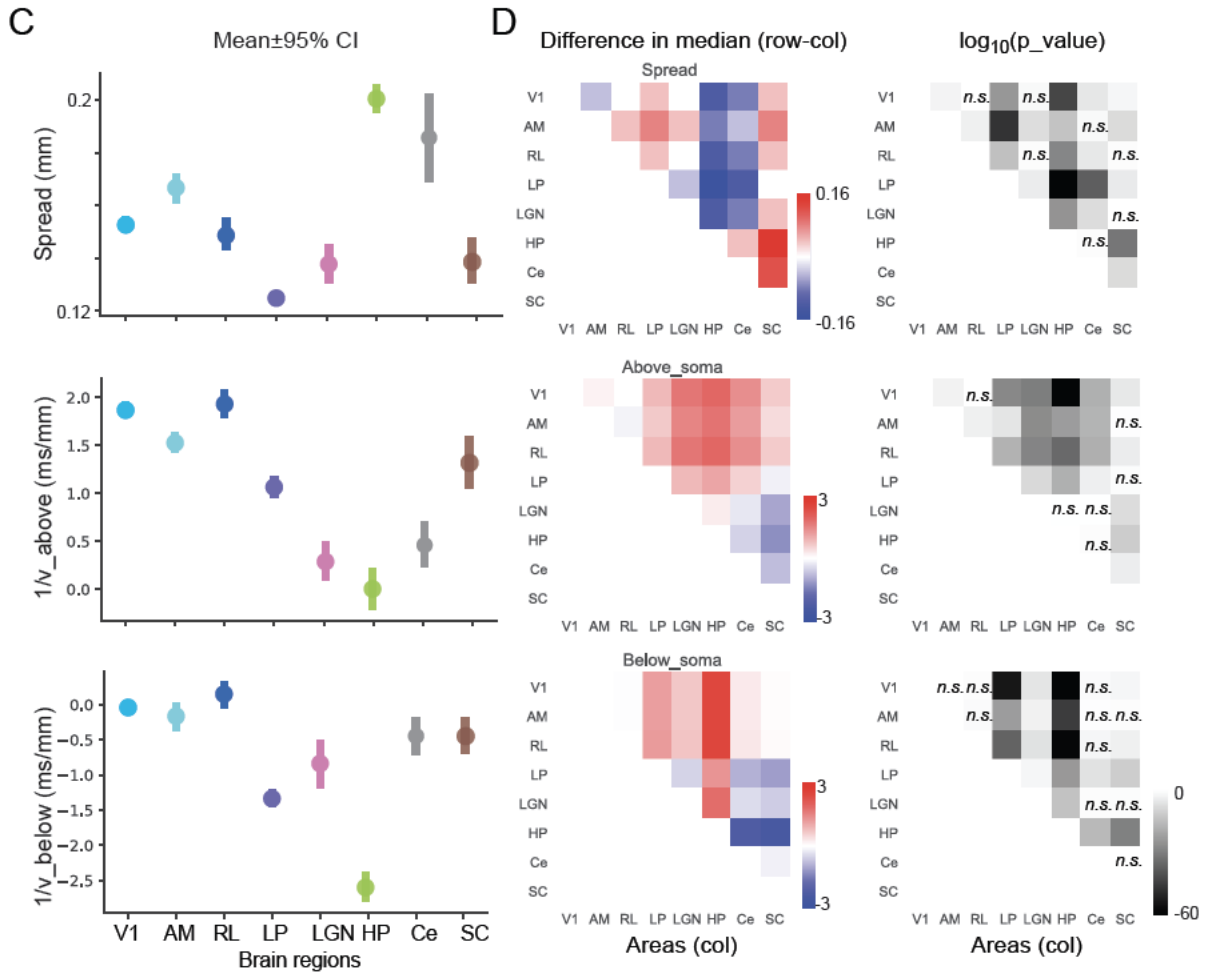


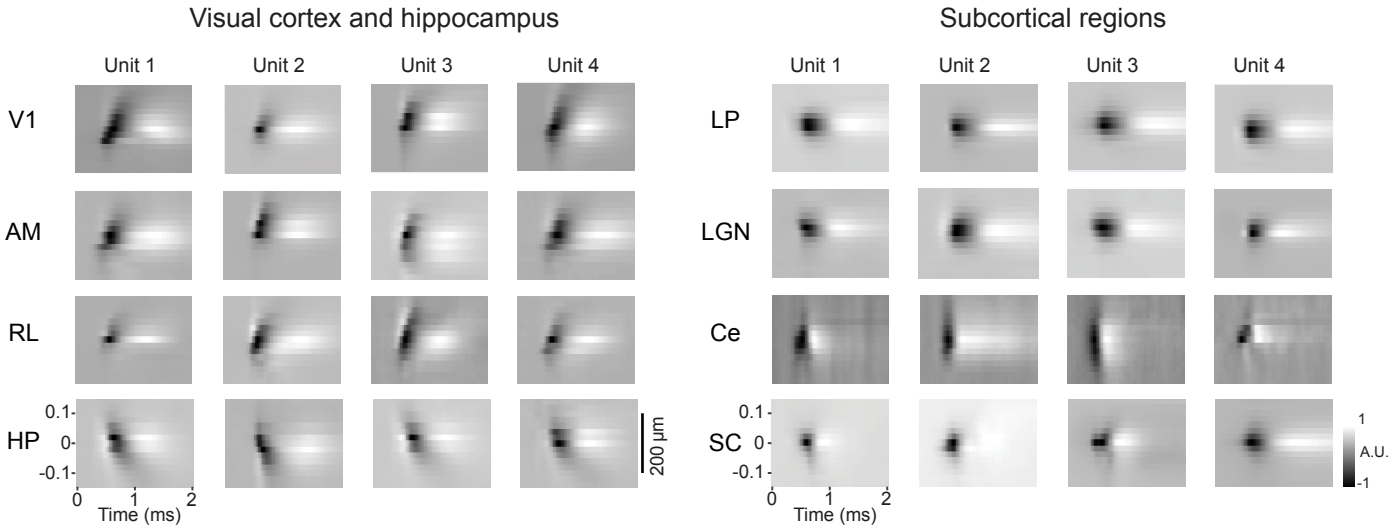
Supplementary Figure 1. Histological verification of recording location from different brain regions. **A)** Example probe track imaged and reconstructed using optical project tomography (indicated by white arrow). The white matter separates the visual cortex and hippocampus (illustrated with white dashed line). **B)** Overlay of retinotopic sign map and vasculature image to define subareas in visual cortex to guide probe insertion (red lines indicate probes). **C)** Example probe track in cerebellum (red). DAPI (blue) labels cell bodies and shows cytoarchitecture of cerebellum. **D)** Probe track labeled with DiI (red) from example LGN recording in a VGAT-ChR2-EYFP mouse (green labeling shows EYFP in inhibitory neurons of the thalamic reticular nucleus). **E,F)** Probe track verification for LP and SC recordings in NTSR1-GN209 x Ai32 mice. sSC: superficial layers of SC, dSC: deep layers of SC. **G)** Distribution of spontaneous firing rate in different cortical and subcortical areas. **H)** Spontaneous firing rate for subclasses of units in visual cortex (*** p_value<0.001). **I)** Stability of recordings for 1 hour sessions quantified as percent of time each unit is present during the session. Overall, 86±12% units are present for more than 90% of the total time. **J)** Refractory period violations of sorted units. We use a conservative value for refractory period of 1.5 ms consistently across areas, which may overestimate the true violation rate. Nonetheless the average percent of refractory period violations across all units is very low (0.88±0.03%). **K)** Signal-to-noise ratio of spike waveforms from different brain areas (Suner et al., 2005).



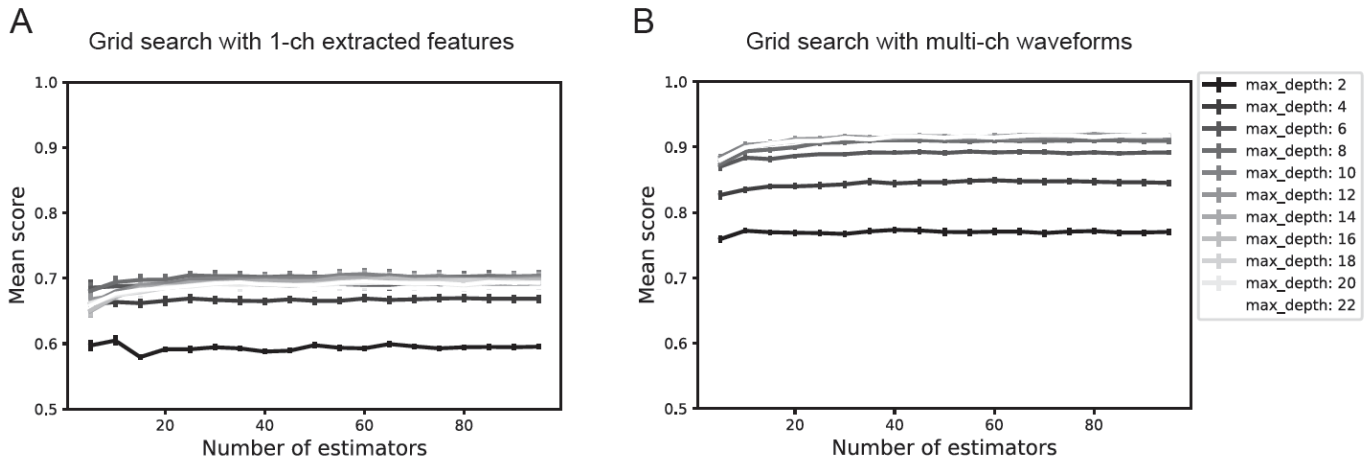
Supplementary Figure 2. Statistics of extracted waveform features. A) Mean and confidence interval (95% confidence interval) of features extracted from 1-ch waveforms from different brain regions. **B)** Pairwise statistics of extracted waveform features. Left panel: Difference of median of the distributions (row-column) for features extracted from 1-channel waveform. Right panel: paired t-test p_value shown in log10 scale, with darker color corresponds to smaller p-value. Significant comparisons were determined by p_value > 0.05 (Bonferroni corrected n=28). Non-significant comparisons are indicated by n.s.



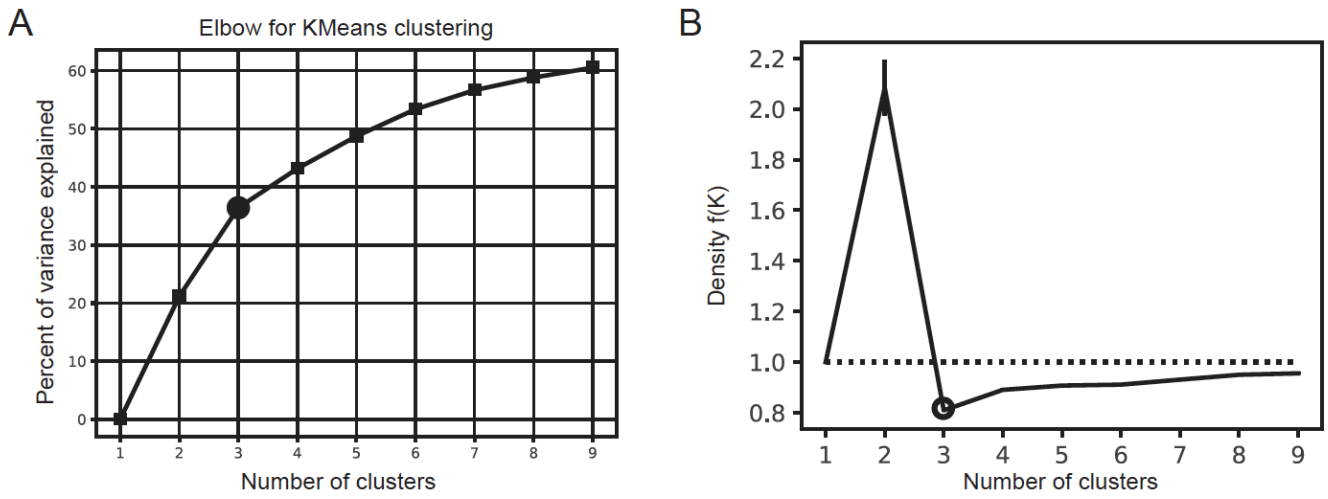
Supplementary Figure 2 (continued). Statistics of extracted waveform features. C) Mean and confidence intervals of features extracted from multi-channel waveforms from different brain areas. **D)** Pairwise comparisons for features extracted from multi-channel waveforms.



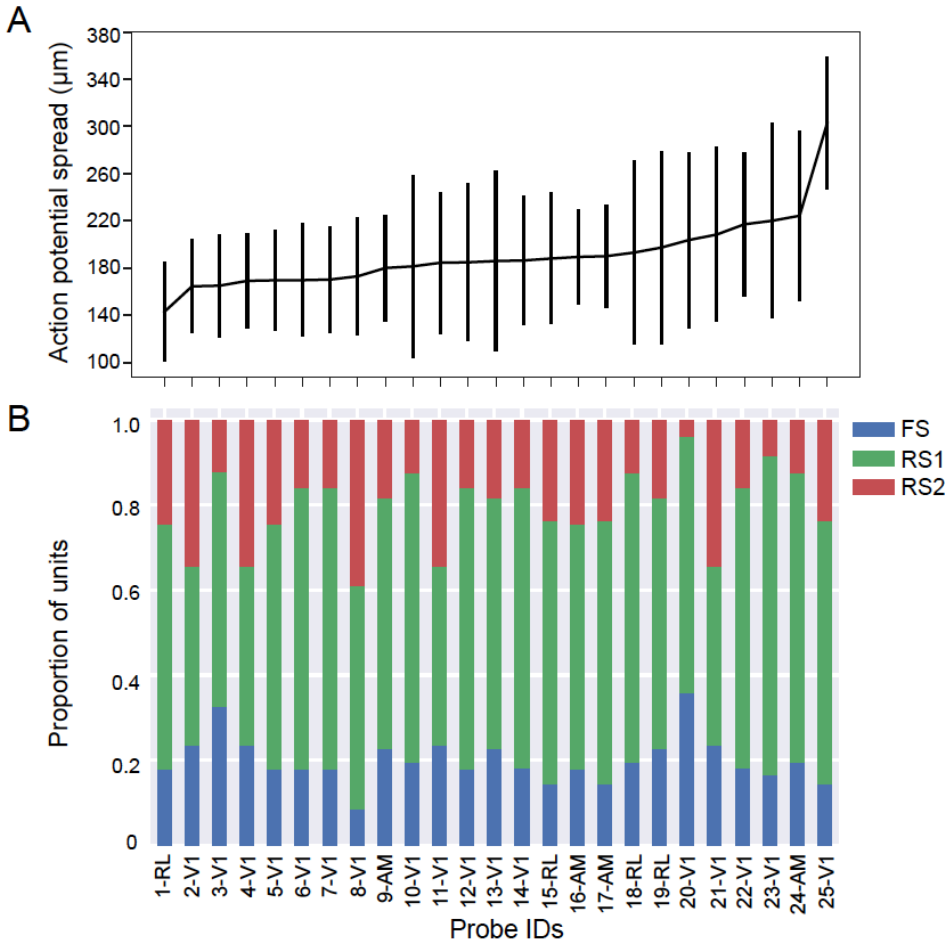
Supplementary Figure 3. Additional examples of multi-channel waveforms from different brain areas. Four example single unit waveforms from each of eight different brain regions. Each heat map shows amplitude of spike over time, measured on channels above and below the soma.



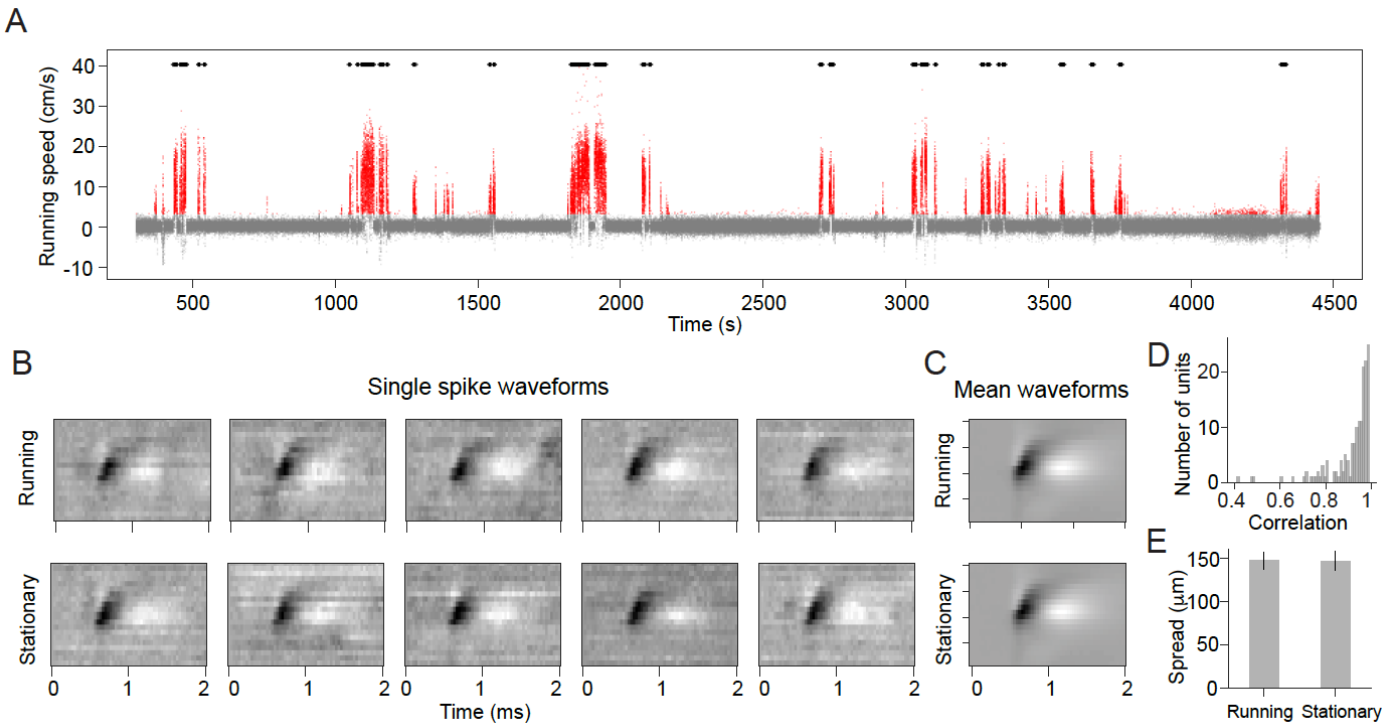
Supplementary Figure 4. Grid search with different hyperparameters for random forest classification. **A)** 5-fold cross-validation with features extracted from 1-channel waveforms (n_feature=4) as a function of number of estimators (decision trees). Standard error calculated across 5 repetitions. Color intensity of the lines corresponds to different tree depths (see legend). **B)** 5-fold cross-validation with multi-channel waveforms (n_feature=1800) as a function of number of estimators. This result indicates that performance of random forest classification is more sensitive to the maximum tree depth rather than the number of estimators for our dataset. Since performance plateaus above certain tree depth and number of estimators for different feature sets, there is no need to fine tune hyperparameters individually for different feature sets.



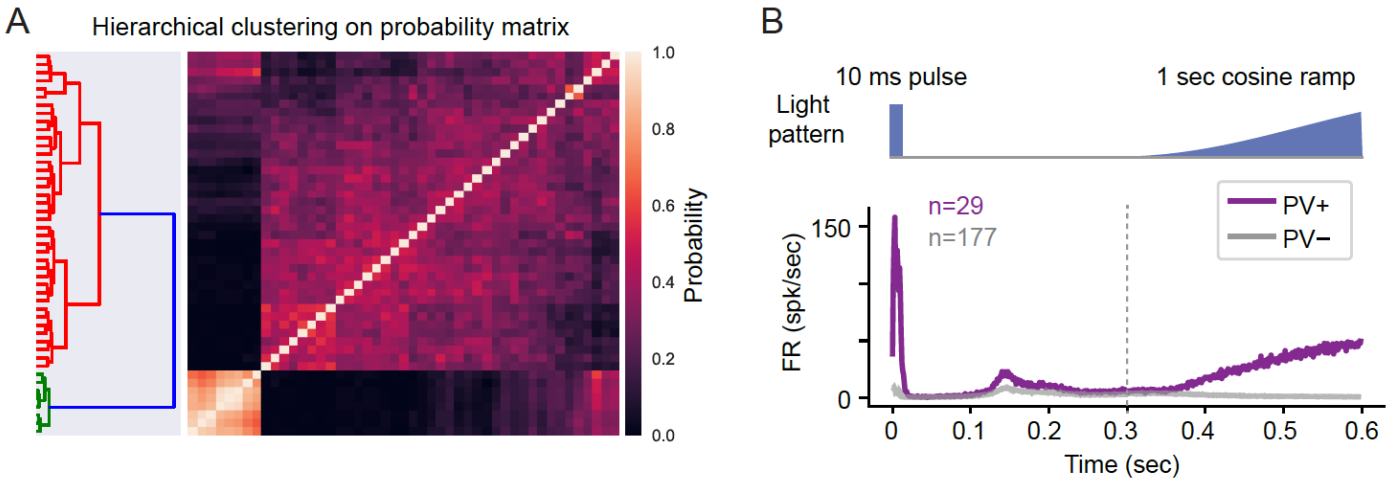
Supplementary Figure 5. Determining number of K clusters in visual cortex. A) Elbow method for finding number of K clusters. Percent of variance explained is calculated as a function of number of K. The turning point is highlighted by a black circle. **B)** Density function f(K) as a function of number of k (see Methods). Values lower than 1 indicate concentration of data given the value of K. The number of K clusters that gives rise to the minimum f(K) is determined as the optimal number of clusters (highlighted by a red circle). Both methods indicate K = 3 clusters.



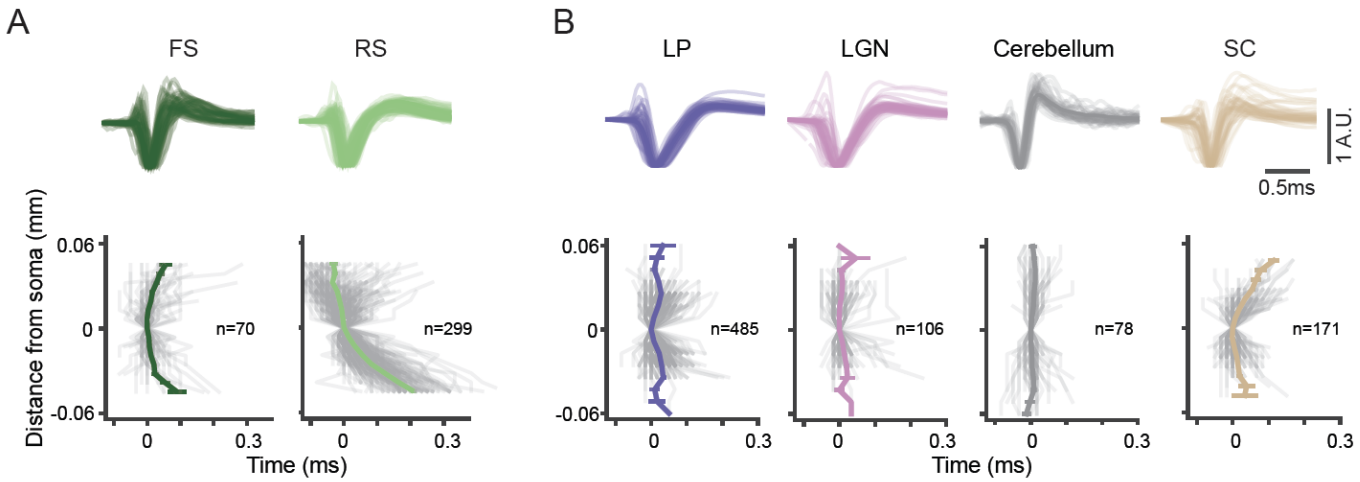
Supplemental Figure 6. Proportion of neuronal clusters in visual cortex for different probes. **A**) Sorted distribution of spread of RS1 units on each probe (n=25 probes from 11 mice). The spread of RS1 units show unidirectional BAPs, which could indicate the alignment of probes with the apical dendrites of individual neurons. **B**) Proportion of FS, RS1 and RS2 units on each probe arrange by the averaged spread of RS1 units of each probe. This result showed that there is no clear bias of RS1/RS2 distribution as a function of individual probes, cortical areas or the spread of BAPs.



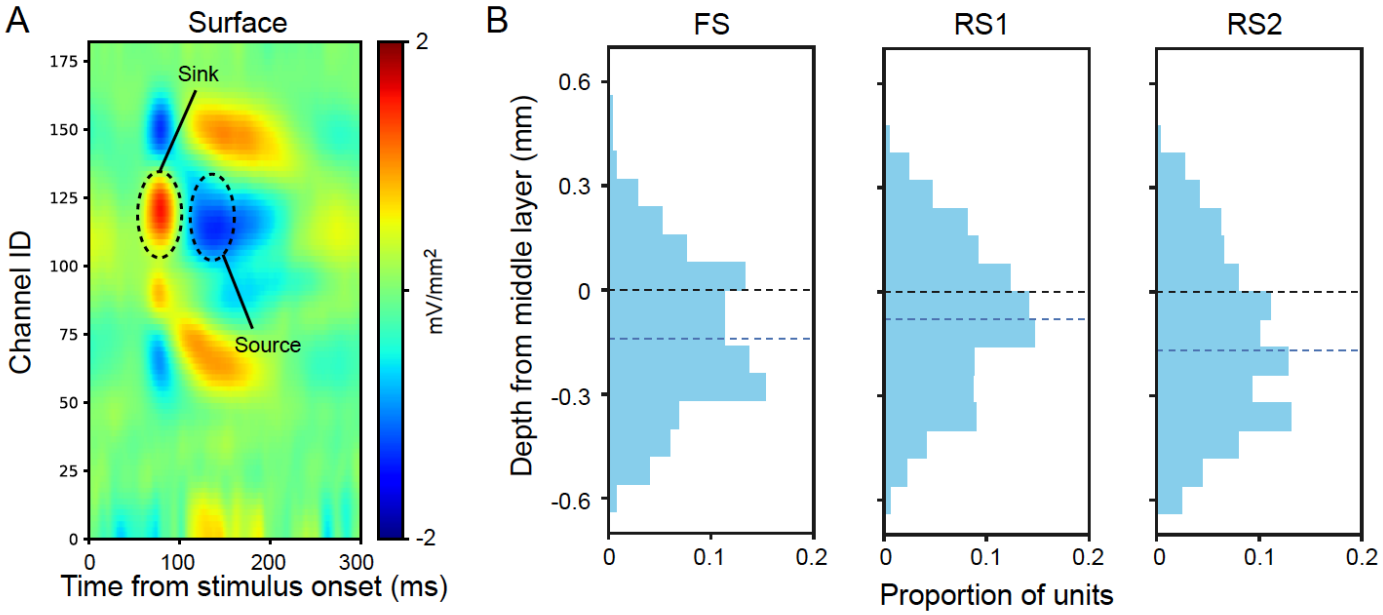
Supplementary Figure 7. Stability of backpropagating action potentials during running and stationary states. A) Running speed of an example mouse during a complete recording session. Gray dots indicate speed smaller than 3 cm/s, and red dots indicate speed larger than 3 cm/s. Running epochs are indicated by black bars at the top, which is defined by continuous running faster than 10cm/s for 2 seconds. **B**) Example single spike waveforms of cortical unit with BAPs during different epochs of running or stationary. **C**) Averaged waveform of an example unit during running and stationary states. **D**) Distribution of correlation coefficient between the averaged multi-channel waveforms during running and stationary states for all units in the example mouse (0.92±0.09, n=131 units). **E**) Averaged spread of BAPs during running and stationary states (t-test $p>0.05$). Error bar indicates standard deviation.



Supplementary Figure 8. Hierarchical clustering of neuronal response to photostimulation. **A)** Post-stimulus time histograms (bin size=1 ms) describing response to photostimulation is used as a feature vector for each unit and K-means clustering is applied 1000 times to the PCA components of the response matrix to construct a co-clustering matrix that reflects the probability of pairs of units belonging to the same cluster. Hierarchical clustering is applied to the probability matrix to find different clusters. This example shows the probability matrix (right) and the hierarchical clustering dendrogram (left) for an example mouse (PV-IRES-Cre; Ai32 (ChR2)) for units in visual cortex. The cluster of units with responses that tightly follow the light pattern is defined as the opto-tagged subpopulation. **B)** Summary plot of mean response to light pattern for all PV+ neurons (n=29 units) from two mice and mean response of the non-opto-tagged neurons (n=177 units).



Supplementary Figure 9. Spike propagation profiles for subcortical regions. **A)** Top: Example waveforms from FS and RS units in hippocampus. Bottom: waveform trough propagation trajectories for all hippocampal FS and RS units. **B)** Top: Example waveforms from LP, LGN, SC and cerebellum. Bottom: propagation trajectories for each region. Gray lines indicate individual units and the colored lines indicate the mean±sem.



Supplementary Figure 10. Laminar distribution of visual cortex waveform clusters. A) An example CSD from one insertion. Layer 4 is estimated by the structure of current sinks and sources. Center of mass of the first sink is defined as 0, indicating middle of layer 4. Data from 8 mice (n=1285 units) is aligned to middle layer of visual cortex. **B)** Proportion of units as a function of laminar depth. Dashed black line indicates depth 0. Dashed blue line indicate median of distribution. The distributions of RS1 and RS2 are significantly different (t-test, p=1.44E-5).