nature neuroscience

Technical Report

https://doi.org/10.1038/s41593-023-01267-x

Tracking neural activity from the same cells during the entire adult life of mice

In the format provided by the authors and unedited

This PDF file includes:

Supplementary Text

Supplementary Figures 1 to 5

Supplementary Tables 1 to 5

Supplementary Methods

Supplementary Text

Distribution of firing rate and related ISI violations

To address the potential issue of inadequate reporting of inter-spike interval (ISI) violations using fractions for neurons with low firing rate, we took several precautions: First, we excluded neurons with firing rate less than 0.1 spikes/s[1]. Then, we carefully inspected the firing rate of all recorded neurons and found that 2.7% of neurons had firing rate between 0.1 and 1 spike/s (Supplementary Figure 1a, dark blue). These neurons were recorded by mesh electronics with sparsely distributed electrode arrays during spontaneous activity recording. Next, we calculated the ISI violations using a metric that reports the relative firing rate of hypothetical neurons generating these violations[2]. Results demonstrated that 96.4% of neurons (Supplementary Figure 1b, 0.0539 ± 0.1240, median ± SD, n = 111 neurons from 6 independent mice) had ISI violations[2] less than the threshold defined by the previous studies[3]. We also included the ISI violations and firing rate of all stable neurons in Supplementary Table 5. Together, these results demonstrate the well-isolated, single-unit recording by mesh electronics.

ML-based recording stability validation and analysis

We benchmarked the stability of the recording in an unbiased manner using a machine-learning-based algorithm to examine the entire spike waveform rather than manually selected features. Specifically, we built a classification-autoencoder by combining the conventional autoencoder and supervised classification. A conventional autoencoder is a self-supervised artificial neural network comprising encoder and decoder subnetworks that project input spike waveform into a low-dimensional (2D in our analysis) representation and back to the input data space. Besides the

design of the original autoencoder, we further added a classification subnetwork and built a classification-autoencoder to classify the sorted neuron identity of the spike waveform (Supplementary Fig. 2a). The classification subnetwork took both the 2D representation of the spike waveform and the corresponding electrode information as input and assigned the spike with a sorted neuron identity. As a result, each classification-autoencoder can i) project spike waveforms into a 2D representation space, ii) classify each 2D representation as a specific neuron in the training data, and iii) reconstruct the embeddings back to the input data space (Supplementary Fig. 2a). The projection, classification, and the reconstruction were simultaneously optimized during the training of the autoencoder. We then verified the stability of the waveforms by examining the performance of the classification-autoencoder from three aspects: classification accuracy, reconstruction accuracy, and representation space consistency. Notably, unlike the aforementioned waveform similarity analysis where we validated the stability of the averaged spike waveform within each neuron, the classification-autoencoder, because of the computational efficiency, can validate the stability of every individual spike. We envisioned that if a recording was stable enough, the recorded spike waveform from testing data should be indistinguishable from training data. Then, the trained classification autoencoder can i) consistently project the spike waveforms to the same manifold space, ii) accurately reconstruct the spike waveforms, and iii) correctly classify the spikes with high accuracy. On the contrary, for unstable recordings, the accuracy of these results will be low.

We trained the classification-autoencoder using data from the first 6-month recording and tested the performance using the data from the following 8-month recording. Importantly, using the noise gathered from electrodes without spikes, we constructed an abnormal dataset as the anomaly control data for the classification-autoencoder. The abnormal dataset only contained highly

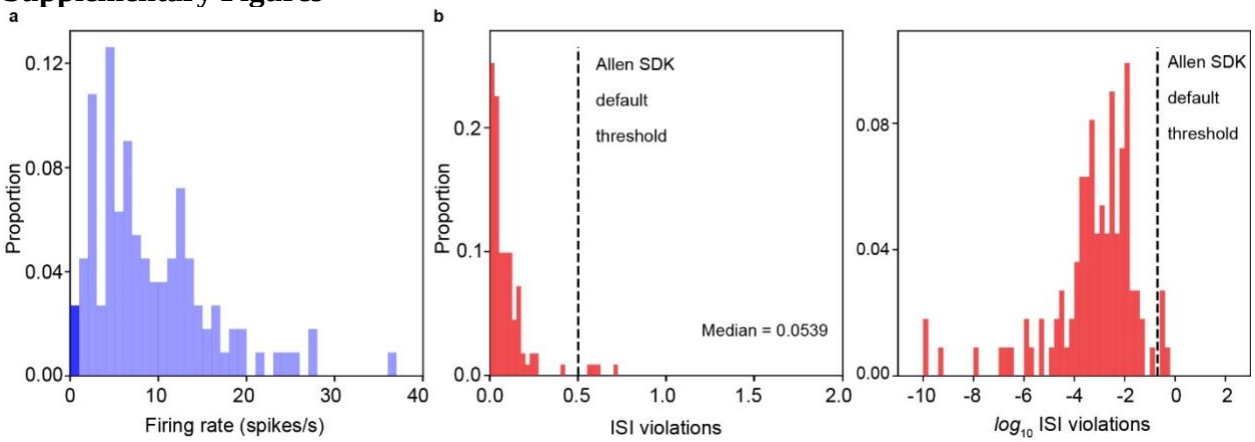
distorted waveforms from the background noise and thus can be used to verify whether the classification-autoencoder could distinguish the real neuron spike from the noise (Supplementary Fig. 2a, b). After training, we first verified the stability of the spike waveforms in the recording by examining the classification and reconstruction accuracy. The results showed that the trained classification-autoencoder successfully classified and reconstructed the testing data with high accuracy (classification accuracy = 94%, n = 144 neurons from 6 mice, Supplementary Fig. 2c and Supplementary Fig. 3b) and low reconstruction mean square error (MSE, individual MSE = 117.33 ± 0.06, mean ± SEM, Supplementary Fig. 2b, d, and Supplementary Fig. 3a). In contrast, the reconstruction of the anomaly control dataset showed a significantly higher MSE (individual MSE = 297.91 ± 1.81, mean ± SEM, Supplementary 2d). In addition, a threshold was automatically identified to distinguish the testing and anomaly control spikes for each mouse, which kept most spikes from the testing dataset (88 ± 2%, mean ± SEM, Supplementary Fig. 3c) and eliminated most spikes from the anomaly control dataset (80 ± 7%, mean ± SEM, Supplementary Fig. 3c). This result proved that the recording data from the remaining eight months was sufficiently stable to be accurately classified and reconstructed. Then, we verified the stability of the spike waveforms by examining the consistency of the autoencoders' representation space in the bottleneck layer (Supplementary Fig. 2a). We constructed a manifold convex hull from the representation embeddings of training data to quantify the within-manifold subset of testing dataset spikes (Supplementary Fig. 2e). The results showed that the autoencoder projected most of the testing spikes (94 ± 0.6%, mean ± SEM) to the correct training manifold (Supplementary Fig. 2f). Moreover, the manifold of the training and testing data was significantly more separable than the manifold of the anomaly data (Supplementary Fig. 3d). Collectively, we used the classification accuracy, reconstruction accuracy, and within-manifold percentage of the classification-

autoencoder to confirm the stability of the spikes detected from the same neurons over time. The results showed that most of the neurons (86%, $n = 144$ neurons from 6 mice) were stable over time.

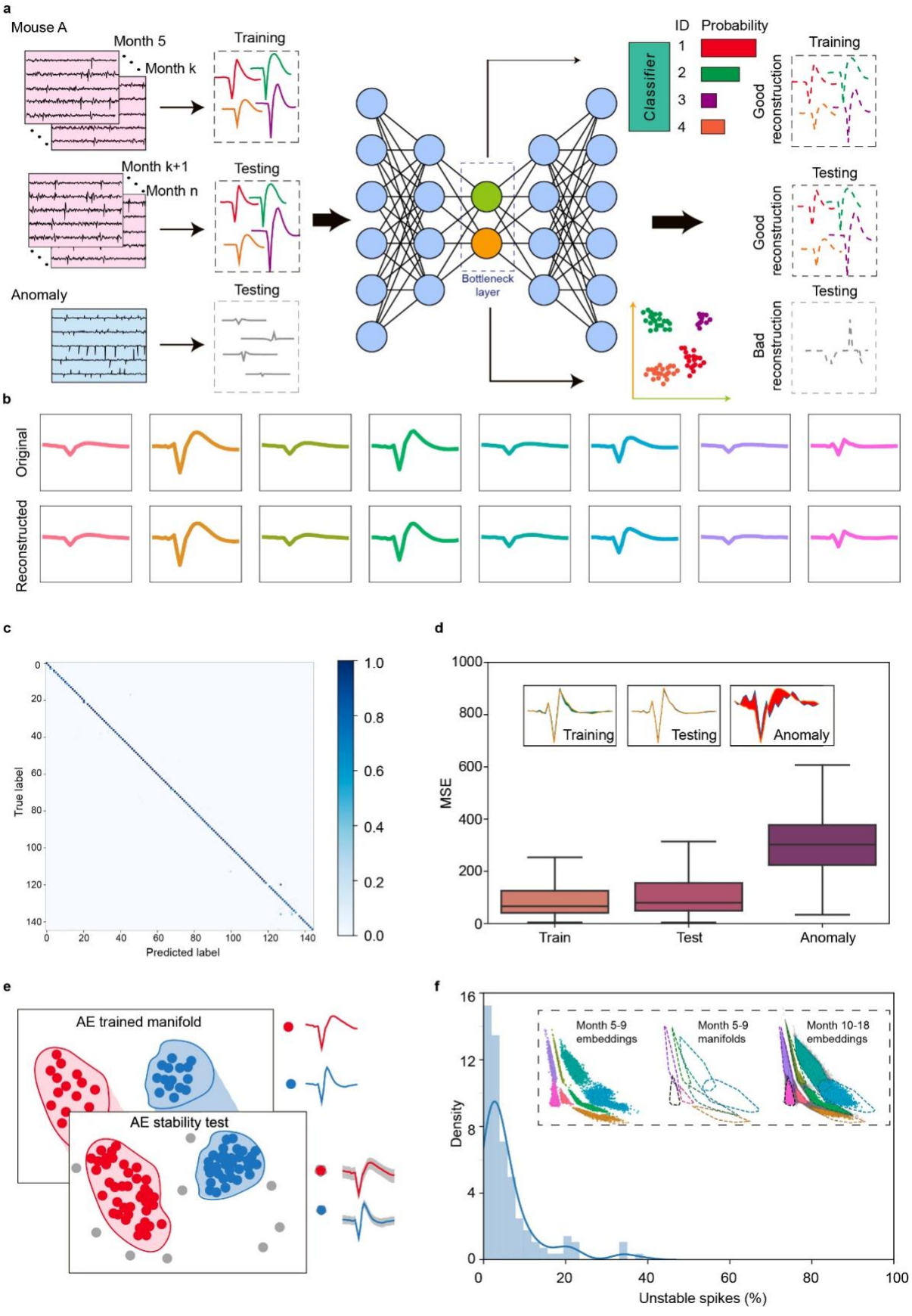
Comparison of signal between sparsely distributed and tetrode-like mesh electrode arrays

As a result, both types of mesh electronics showed high levels of waveform similarity (sparsely distributed electrodes: 0.98 ± 0.03 versus tetrode-like electrodes 0.97 ± 0.04, mean ± SD, Supplementary Fig. 4a, b), low L-ratio (sparsely distributed electrodes: 0.002 ± 0.004 versus tetrode-like electrodes: 0.006 ± 0.006, mean ± SD, Supplementary Fig. 4c), high Silhouette score (sparsely distributed electrodes: 0.73 ± 0.09 versus tetrode-like electrodes: 0.84 ± 0.06, mean ± SD, Supplementary Fig. 4d), and <10% fraction of spike violating refractory periods (sparsely distributed electrodes: 0.17% ± 0.33% versus tetrode-like electronics 2.05% ± 1.49%, mean ± SD, Supplementary Fig. 4e). There were no significant differences in the change rates of the five waveform features between the two types of electronics ($p > 0.05$ for all comparisons, two-tailed *t* test, Supplementary Fig. 4f).

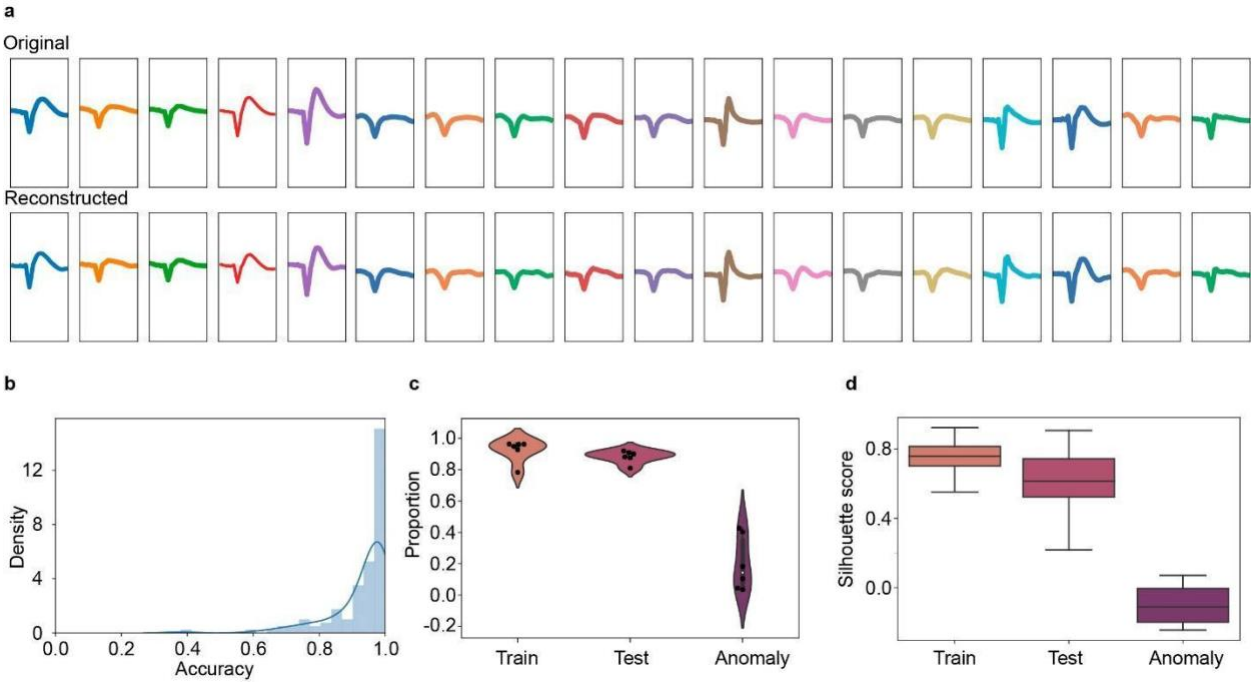
Supplementary Figures



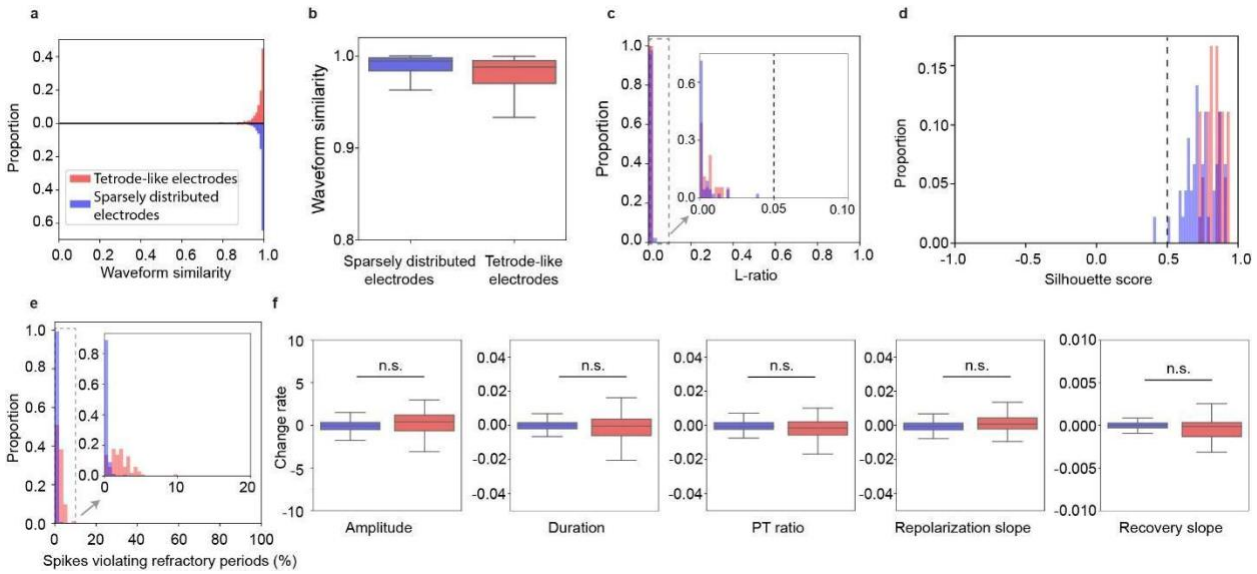
Supplementary Fig. 1 | Distribution of firing rate and ISI violations. a, Distribution of firing rate across 111 neurons from 6 independent mice. Dark blue indicated the neurons with firing rate between 0.1 and 1 spike/s. **b**, Distribution of ISI violations on a linear (left) and logarithmic (right) scale. We added 10^{-10} to the ISI violation values (right), because the log of 0 is an undefined value. Default AllenSDK thresholds are indicated by dotted lines. n = 111 neurons from 6 independent mice.


a
Mouse A
Month 5
Month k
Training
Month k+1
Month n
Testing
Anomaly
Testing
Bottleneck layer
ID
Probability
1
2
3
4
Classifier
Training
Good reconstruction
Testing
Good reconstruction
Testing
Bad reconstruction
b
Original
Reconstructed
c
True label
Predicted label
d
MSE
Training
Testing
Anomaly
Train
Test
Anomaly
e
AE trained manifold
AE stability test
f
Month 5-9 embeddings
Month 5-9 manifolds
Month 10-18 embeddings
Density
Unstable spikes (%)

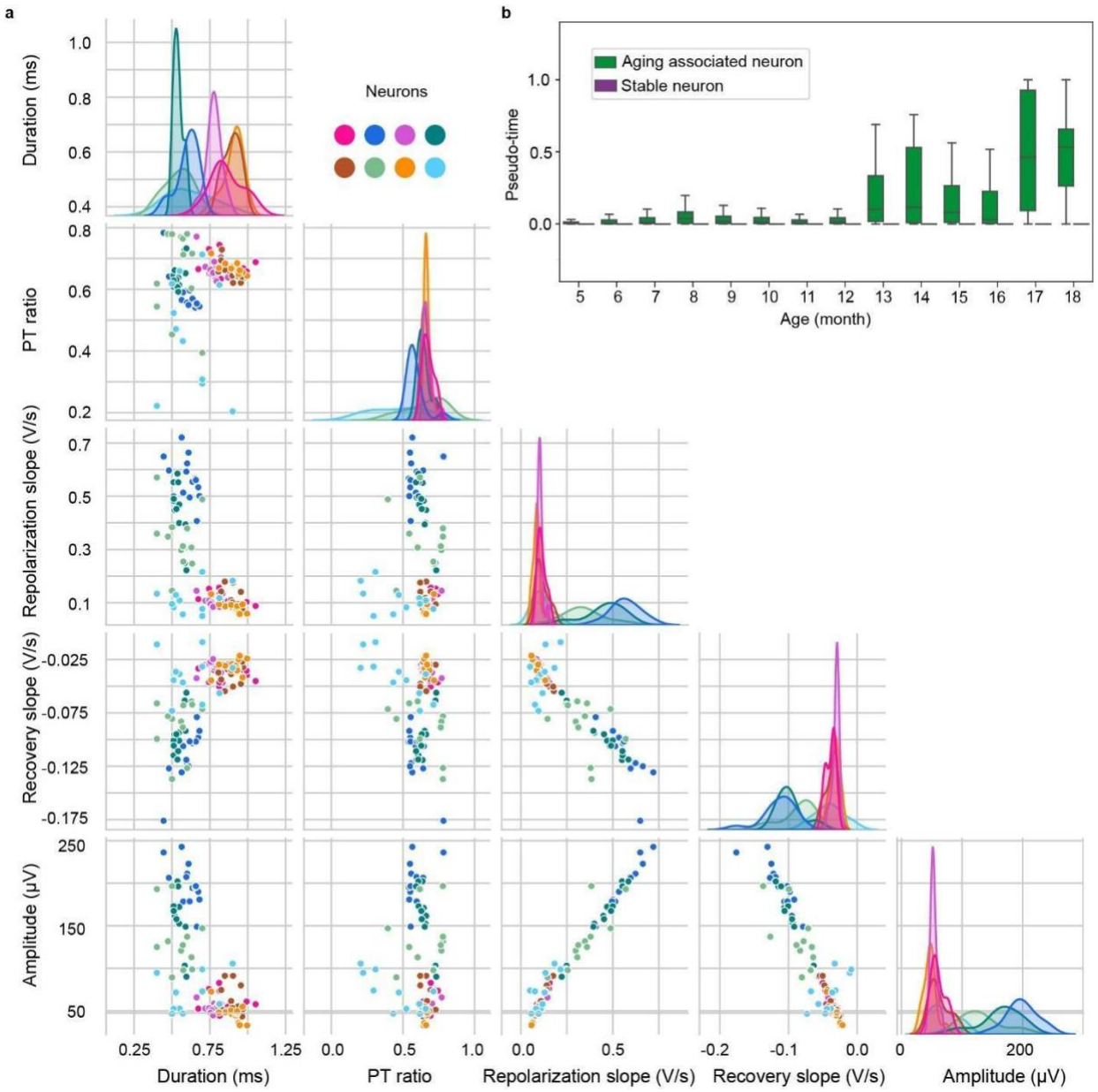
Supplementary Fig. 2 | Machine learning-based stability test. a, Schematics of an autoencoder-based model which consists of an encoder, a classification head, and a decoder. Single-unit action potential waveforms are first extracted from the recording voltage traces and encoded with electrode number to a lower-dimensional latent space. The latent representation is then used for classification and decoding. Spikes from the first 6 recording months were used for training. The subsequent 8 months of data were used for testing. Spikes from background noise were used as anomaly control data. **b**, Original and reconstructed average representative single-neuron waveforms from the testing data. Original waveforms are colored by neuron labels. Reconstructed waveforms are colored by the classification output. **c**, Confusion matrix comparing testing dataset ground-truth neuron labels with autoencoder classification output across six mice. **d**, Boxplot of median and quartile range of mean-squared error (MSE) distribution for each dataset training, testing, and anomaly dataset. n = 363,303 spikes for the training data, n = 2,948,280 spikes for the testing data, n = 3,826 spikes for the anomaly data. All spikes from 6 independent mice. Box, 75% and 25% quantiles. Line, median. Whisker, the maxima/minima or to the median ± 1.5× IQR. Inset: representative average waveforms and their corresponding reconstructions by the trained mouse-specific autoencoder. The colored areas correspond to the difference between the reconstructed and original waveform. **e**, Schematics showing that the autoencoder can be used to quantify the stability of the recording. The trained manifold (highlighted in red and blue) was used to test the stability of the recording by quantifying the percentage of the neuron single-unit action potentials from the testing data falling inside the training manifold. **f**, Distribution plot illustrating the percentage of the testing dataset for each neuron falling outside of the training manifold as per the process illustrated in (**e**). Low levels of unstable spikes suggest recording stability through train-test waveform resemblance. Inset: applying noise rejection process to representative mouse data, in order: training embeddings, training manifold boundaries, testing embeddings with out-of-manifold spikes colored in grey.



Supplementary Fig. 3 | Machine learning-based stability analysis. a, Original and reconstructed average neuron waveforms from an additional representative mouse. **b**, Distribution plot of classification accuracy calculated for each neuron label class. **c**, Proportion of spikes kept for each dataset when using MSE-based thresholds for stability verification. **d**, Silhouette score calculated using autoencoder embeddings of training, testing, and anomaly control datasets with associated true neuron labels. The training and testing scores reflect the emergent latent space cluster separability for observed neurons while the anomalous scores highlight the poor separability of previously unseen embedded neuron waveforms. Box, 75% and 25% quantiles. Line, median. Whisker, the maxima/minima or to the median ± 1.5× IQR, n = 6 independent mice.



Supplementary Fig. 4 | Comparison of signals between sparsely distributed and tetrode-like mesh electrode arrays. a-b, Distribution (**a**) and boxplot (**b**) of waveform similarity within each neuron. Red: tetrode-like mesh electrode arrays, 0.97 ± 0.04, mean ± SD, *n* = 92 neurons from 5 mice. Blue: sparsely distributed mesh electrodes, 0.98 ± 0.03, mean ± SD, *n* = 111 neurons from 6 independent mice. Boxplots, 75% and 25% quantiles. Line, median. Whisker, the maxima/minima or to the median ± 1.5× interquartile range (IQR). **c-d,** Distribution of L-ratio and silhouette score for signals between the two types of electronics. Dashed lines indicate L-ratio of 0.05 and silhouette score of 0.5, commonly taken as a threshold of high cluster quality. **e,** Refractory periods violations (defined as an ISI <1.5 ms) for sparsely distributed electrodes (0.17% ± 0.33%, mean ± SD, *n* = 144 neurons) and tetrode-like (2.05% ± 1.49%, mean ± SD, *n* = 102 neurons) electrode arrays. **f**, Comparison of the change rate of five waveform features (amplitude, $p = 0.0585$; duration, $p = 0.3492$, PT ratio, $p = 0.0633$; repolarization slope, $p = 0.1254$; recovery slope, $p = 0.0525$) across all recording periods between sparsely distributed and tetrode-like mesh electrode arrays, *n.s.*, not significant, two-tailed *t* test, *n* same as **a-b**. Boxplots, 75% and 25% quantiles. Line, median. Whisker, the maxima/minima or to the median ± 1.5× IQR.



Supplementary Fig. 5 | Feature selection and electrophysiology pseudotime for potential aging-associated analysis. a, Pair plot of five waveform features. Dots represent mean values of paired features calculated over neuron clusters over the entire recording period. Diagonal subplots show neuron features in univariate distributions using kernel density estimators. **b,** Box plot of the electrophysiology pseudotime from two representative neurons recorded from the same electrode over the adult life of mice. Box, 75% and 25% quantiles. Line, median. Whisker, the maxima/minima or to the median ± 1.5× IQR. *n* = 11,254 spikes for the aging-associated neuron, *n* = 15,672 spikes for the stable neuron.

Supplementary Tables

Distance (μm)	2-week					
	Normalized neuron intensity		Normalized astrocytes intensity		Normalized microglia intensity	
	p_{mesh}	p_{film}	p_{mesh}	p_{film}	p_{mesh}	p_{film}
0	0.0748	4.1×10^{-7}	0.0097	3.5×10^{-5}	0.0039	0.1461
25	0.071	0.0126	0.0052	6.2×10^{-5}	0.14	0.0025
50	0.0844	0.02	0.006	4.2×10^{-6}	0.0128	0.0002
75	0.3455	0.3612	0.0112	8.1×10^{-5}	0.0535	2.5×10^{-5}
100	0.8307	0.0226	0.0091	0.0064	0.1518	0.0028
125	0.7089	0.4417	0.2494	0.0309	0.1589	0.0185
150	0.7319	0.8177	0.33	0.001	0.3535	0.2976
175	0.5906	0.1598	0.2512	0.097	0.3851	0.7243
200	0.3489	0.7509	0.5831	0.0659	0.188	0.588
225	0.8024	0.362	0.2836	0.2095	0.2507	0.6383
250	0.9098	0.2575	0.1751	0.4317	0.2443	0.7389
275	0.6163	0.8205	0.3835	0.7257	0.1136	0.5492
300	0.4576	0.7631	0.3397	0.5322	0.4621	0.7483

Supplementary Table 1 | *p* value of the histology studies (Figure 2f) at 2-week post-implantation.

Distance (μm)	6-week					
	Normalized neuron intensity		Normalized astrocytes intensity		Normalized microglia intensity	
	p_{mesh}	p_{film}	p_{mesh}	p_{film}	p_{mesh}	p_{film}
0	0.0559	2.4×10^{-5}	0.2244	5.8×10^{-6}	0.062	0.0409
25	0.1531	0.0098	0.2095	4.1×10^{-5}	0.0866	0.0016
50	0.4855	0.0951	0.6203	0.0006	0.0521	0.0007
75	0.4092	0.1522	0.5489	0.0069	0.5509	0.0033
100	0.523	0.5089	0.9624	0.0422	0.0954	0.0662
125	0.4831	0.5955	0.735	0.0002	0.6815	0.1206
150	0.7924	0.5036	0.6851	0.4195	0.5978	0.1063
175	0.2126	0.4505	0.6826	0.4081	0.863	0.4837
200	0.8181	0.6738	0.7687	0.934	0.7431	0.3627
225	0.2111	0.9956	0.6758	0.9337	0.4206	0.3682
250	0.986	0.6202	0.6	0.0935	0.6819	0.5
275	0.4736	0.9057	0.9991	0.5149	0.9739	0.0968
300	0.9316	0.8039	0.6977	0.9831	0.8934	0.0521

Supplementary Table 2 | *p* value of the histology studies (Figure 2f) at 6-week post-implantation.

	12-week					
Distance (μm)	Normalized neuron intensity		Normalized astrocytes intensity		Normalized microglia intensity	
	p_{mesh}	p_{film}	p_{mesh}	p_{film}	p_{mesh}	p_{film}
0	0.2854	0.0011	0.944	0.0015	0.0526	0.1487
25	0.48	0.0115	0.5352	4.1×10^{-6}	0.2146	0.0195
50	0.9012	0.3262	0.6492	0.0002	0.5722	0.0023
75	0.5711	0.65	0.2974	0.0014	0.9284	0.0031
100	0.8269	0.3028	0.0799	0.0023	0.7632	0.0251
125	0.6988	0.8403	0.531	0.0021	0.8044	0.1159
150	0.4222	0.4709	0.9141	0.0027	0.9596	0.191
175	0.7566	0.5817	0.9417	0.0944	0.7625	0.0809
200	0.2994	0.337	0.8323	0.1699	0.4819	0.1024
225	0.5969	0.7654	0.1973	0.5074	0.9778	0.139
250	0.8477	0.8035	0.9466	0.758	0.8857	0.1291
275	0.8384	0.8877	0.5879	0.3684	0.8438	0.1638
300	0.8619	0.6785	0.9593	0.8285	0.7151	0.3723

Supplementary Table 3 | *p-value* of the histology studies (Figure 2f) at 12-week post-implantation.

	1-year					
Distance (μm)	Normalized neuron intensity		Normalized astrocytes intensity		Normalized microglia intensity	
	p_{mesh}	p_{film}	p_{mesh}	p_{film}	p_{mesh}	p_{film}
0	0.5433	0.002	0.6919	0.2444	0.2809	0.0186
25	0.4228	0.0114	0.7056	8.2×10^{-6}	0.2818	0.0171
50	0.859	0.1517	0.529	0.0008	0.0571	0.0144
75	0.3164	0.4687	0.9861	0.0176	0.7057	0.0622
100	0.7132	0.9337	0.9574	0.0149	0.6415	0.1222
125	0.5212	0.671	0.8465	0.0004	0.2474	0.0992
150	0.5361	0.7934	0.2338	0.0049	0.7298	0.3726
175	0.8664	0.4237	0.5769	0.0317	0.5972	0.3555
200	0.5057	0.278	0.7005	0.0082	0.894	0.7377
225	0.6613	0.8738	0.5275	0.0288	0.297	0.7804
250	0.3133	0.9908	0.9346	0.0076	0.9463	0.6111
275	0.7049	0.5446	0.6987	0.1506	0.5632	0.3745
300	0.4409	0.9227	0.3381	0.4451	0.0534	0.6584

Supplementary Table 4 | *p-value* of the histology studies (Figure 2f) at 1-year post-implantation.

IDs	ISI violations	Firing rate	IDs	ISI violations	Firing rate	IDs	ISI violations	Firing rate
1	0.1052	12.35	38	0.1375	9.76	75	0.0769	10.66
2	0.1234	4.46	39	0.0224	13.55	76	0.6126	0.79
3	0.1081	18.75	40	0.0656	4.62	77	0.0959	0.91
4	0.0534	2.44	41	0.1162	7.47	78	0.023	4.29
5	0.0539	2.1	42	0.0388	8.81	79	0.2167	8.53
6	0.0808	1.98	43	0.0015	36.44	80	0.029	2.03
7	0.0774	3.36	44	0.1704	6.39	81	0.0261	2.13
8	0.1183	16.07	45	0.0393	17.02	82	0.116	2.08
9	0.0338	5.83	46	8.6×10^{-5}	12.37	83	0.0271	6.52
10	0.1518	6.44	47	0.1362	14.38	84	0.0106	2.95
11	0.0647	3.78	48	0.0297	8.17	85	0.1547	0.83
12	0.0434	8	49	0.0496	6.34	86	0.0239	6.05
13	0.0011	6.83	50	0.0477	1.75	87	0.0705	4.29
14	0.2563	6.89	51	0.0396	7.03	88	0.0172	7.73
15	0.1127	6.64	52	0.008	13.27	89	0.4028	4.05
16	0.0173	12.53	53	0.0283	4.32	90	0.7132	3.62
17	0.0116	5.44	54	0.0806	4.69	91	0.1589	19.84
18	0.021	12.76	55	0.1626	12.16	92	0.0111	17
19	0.087	13.12	56	0.061	10.46	93	0.1577	24.58
20	0.0822	15.39	57	0.1182	7.13	94	0.1772	27.64
21	0.0347	2.96	58	0.0214	11.04	95	0.0029	23.68
22	0.2613	5.24	59	0.2424	13.43	96	0.0197	14.65
23	0.0263	11.28	60	0.0084	10.49	97	0.0548	19.53
24	0.0466	5.46	61	0.032	8.98	98	0.0764	7.31
25	0	1.77	62	0.1123	14.91	99	0.5953	15.76
26	0.1945	4.43	63	0.0262	21.67	100	0.0004	25.66
27	0.0555	1.54	64	0.0296	4.12	101	0.0384	7.56
28	0.0036	5.72	65	0.0837	6.78	102	0.001	27.23
29	0.1047	4.72	66	0.0028	4.83	103	0.0349	9.53
30	0.1549	4.1	67	0.2259	9.1	104	0	10.27
31	0.0313	2.83	68	0.0803	11.01	105	0.0146	13.67
32	0.0409	2.05	69	0.0401	11.39	106	0.1564	16.03
33	0.0603	6.18	70	0.0357	9.33	107	0.0786	11.36
34	0.0527	2.7	71	0.1021	12.44	108	0.1332	4.57
35	0.1391	12.87	72	0.0053	4.17	109	0.0048	18.05
36	0.0194	5.61	73	0.1491	12.47	110	0.558	1.83
37	0.009	2.97	74	0.0714	5.09	111	0.029	2.22

Supplementary Table 5 | Summary of the firing rate and ISI violations of individual sorted neurons.

Supplementary Methods

Bending stiffness calculations

We estimated and compared the bending stiffness values (D) of mesh and thin-film electronics using a beam model. The bending stiffness of mesh electronics with a three-layer polymer/mesh/polymer structure can be calculated as

$$D = \frac{E_S}{12}(wh^3 - w_m h_m^3) + \frac{E_m}{12} w_m h_m^3 \quad (1)$$

where E_s and E_m are the young's moduli of SU8 and gold (2 and 79 GPa, respectively), w is the width of SU-8 scaffolds, w_m is the width of gold interconnects, h_{mesh} = 0.91 μm and h_{film} = 0.92 μm is the measured total thickness of SU-8, h_{metal} = 80 nm is the thickness of Cr/Au interconnects. The calculated bending stiffness of mesh and thin-film electronics were 1.26×10^{-15} N·m^2 and 3.98×10^{-14} N·m^2, respectively.

Autoencoder-based spike waveform stability verification

All classification autoencoders were built using TensorFlow (https://www.tensorflow.org). The encoder consisted of two fully connected (FC) layers before the bottleneck 2D layer. Inputs to the network were concatenations of individual spike waveforms and their corresponding one-hot encoded electrode. From the bottleneck, the classification head, which consists of a single FC layer with softmax activation, outputs probabilities to a particular neuron class. The decoder's symmetric structure with respect to the encoder allowed for the reconstruction of the original waveform. Activations for encoder and decoder layer were set as LeakyReLU with *alpha* = 0.3 as well as an added L_2 regularizer on the encoder's last layer, enforcing lower latent embedding values. Adam optimizer with default parameters was used. For a given dataset, $X \in R^{nX30}, n \in \aleph$

is the number of the spikes, each composed of 30 data points of waveforms. We denote the input batch to the autoencoder $\hat{X} \in R^{pX(30+c)}$, where p is the number of the mini-batch and c is the one-hot embedding of the channel. The autoencoder considered here was composed of an encoder, f_e, decoder f_d and classifying head f_c parametrized by $\theta = (\theta_e, \theta_d, \theta_c)$. We used θ as the network weights in the downstream description. The network's loss function over a batch of spikes is thus:

$$L(\hat{X}, \theta) = \sum_{i=1}^{p} [\alpha | |x_i - f_d(f_e(x_i))| |_2^2 + \beta \, CE(f_c(f_e(x_i)))] \quad (2)$$

where CE is the categorical cross-entropy loss used for multi-class classification problems. The loss function was defined by two components: reconstruction and classification, each with their own weight to regulate their overall contribution to the loss. Training the network consisted of learning the correct parametrization θ which minimized the loss function over the training dataset. Early stopping was applied on an evaluation dataset, which was a subset split from the training data, to avoid overfitting of the training data. Each autoencoder was independently trained for one mouse using the first few months' recording data. After training, we used three criteria to verify the stability of the spike waveforms in the recording. First, we examined how well the spike can be reconstructed by the classification autoencoder by using mean square error (MSE). We used a threshold to identify the unstable spikes in the testing data. The threshold was automatically calculated by summing the average value and standard deviation of the MSE of the training data.

Brain aging analysis at the single-neuron level

We quantitatively assessed the multivariate (*i.e.*, x- and y-axis) spread of cluster centroid positions by comparing average position shifts between consecutive cluster centroid positions to average cluster distribution spreads. Average cluster centroid position shifts were compared to average

cluster distribution spread. $X \in R^{dx2}$ is a vector of cluster centroid positions over d days in each 2D (x, y) embedding. For each identified neuron, we calculated:

$$\sqrt{(\frac{u_x}{\sigma^{av}{}_x})^2 + (\frac{u_y}{\sigma^{av}{}_y})^2} \quad (3)$$

with $u_x = \frac{1}{d-1}\sum_{i=1}^{d-1} |x_{i+1} - x_i|$, $u_y = \frac{1}{d-1}\sum_{i=1}^{d-1} |y_{i+1} - y_i|$the average successive cluster centroid absolute position shifts along each axis, and $\sigma^{av}{}_x = \frac{1}{d}\sum_{i=1}^{d} \sigma^{x}{}_i$, $\sigma^{av}{}_y = \frac{1}{d}\sum_{i=1}^{d} \sigma^{y}{}_i$ the average cluster distribution SD along each axis.

Furthermore, these UMAP embeddings were used to compute a trajectory graph with the corresponding "pseudo temporal" scale of evolution, with Monocle3 (https://cole-trapnell-lab.github.io/monocle3). The convex hull delimiting regions spanned by individual neuron time-evolution was calculated using Scipy. PCA was used as another form of dimensionality reduction to view the evolution of spike waveforms per cluster. Clusters from the same electrodes were used to obtain cluster centroids for individual electrodes and recording months along with 2 SD confidence ellipses of the corresponding covariance. Confidence ellipses were centered on cluster means while radii were calculated without explicit computation of eigenvalues by using Pearson correlation coefficients and equations:

$$\sqrt{\lambda_1} = \sqrt{1 + p} \quad (4)$$

$$\sqrt{\lambda_2} = \sqrt{1 - p} \quad (5)$$

where λ_1 and λ_2 are the eigenvalues of the covariance matrix and p is the Pearson correlation coefficient.

Feature trajectory analysis was performed by fitting trajectories, from feature embeddings in a selected 3D feature space of mean feature values per cluster over real recording months. The features chosen for the representation were repolarization slope, peak-trough ratio, and duration. These were selected based on prior correlation analysis to make sure the trajectory was not a degenerate representation of feature evolution caused by highly correlated features. Trajectory fittings were constructed using a B-spline interpolation on previously discussed points.

References

1. Steinmetz, N.A., et al. Neuropixels 2.0: A miniaturized high-density probe for stable, long-term brain recordings. *Science* **372**, eabf4588 (2021).

2. Hill, D.N., Mehta, S.B. & Kleinfeld, D. Quality metrics to accompany spike sorting of extracellular signals. *J. Neurosci.* **31**, 8699-8705 (2011).

3. Siegle, J.H., et al. Survey of spiking in the mouse visual system reveals functional hierarchy. *Nature* **592**, 86-92 (2021).