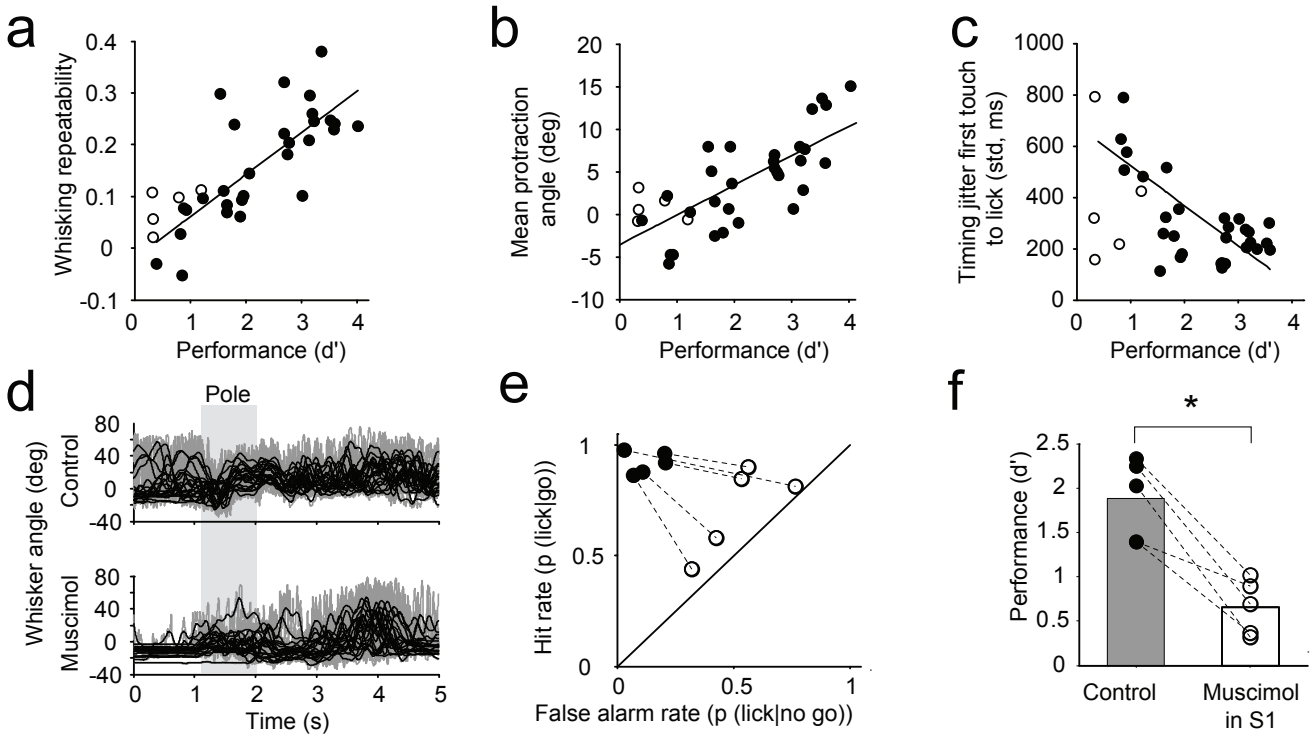
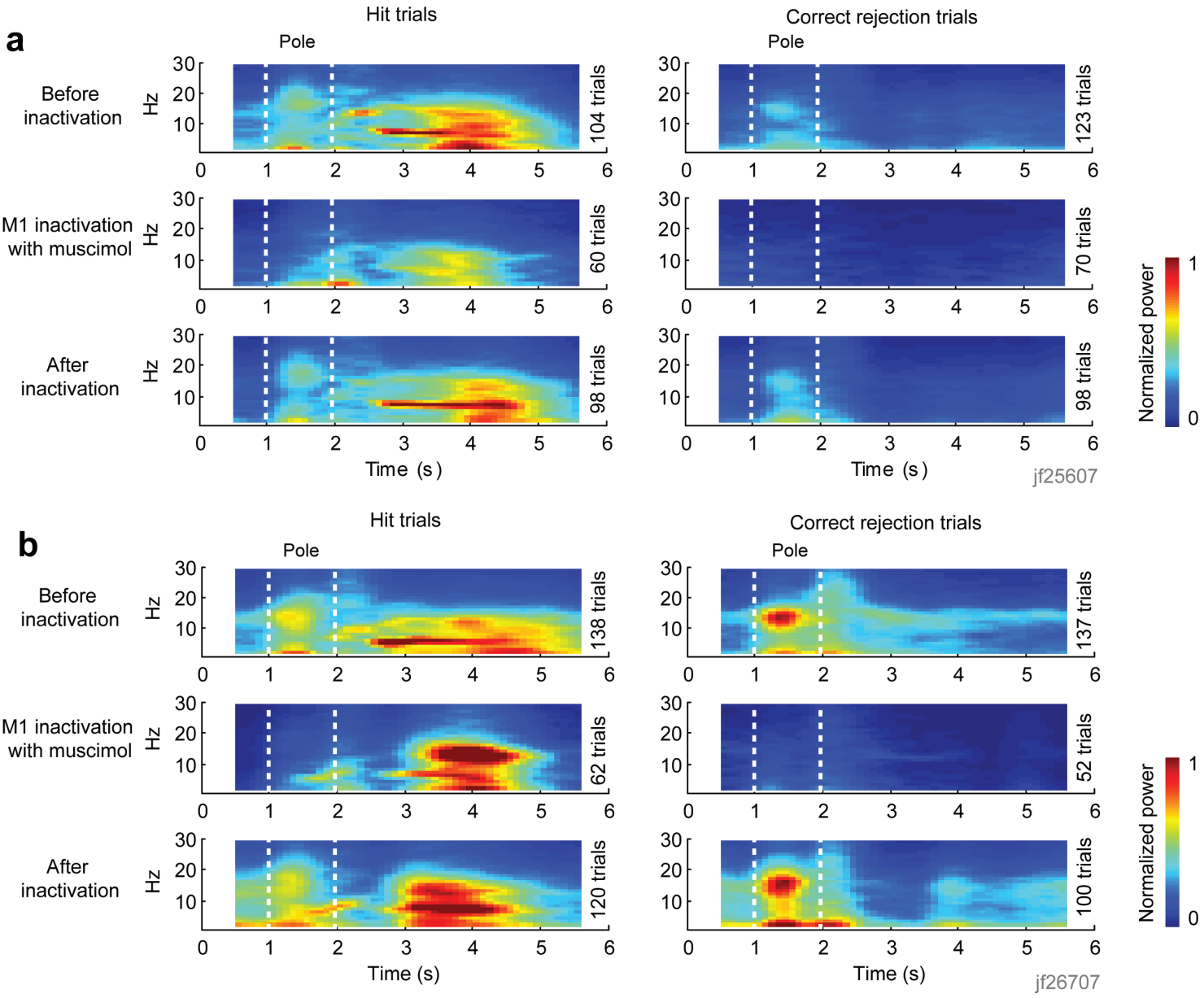
doi:10.1038/nature11039



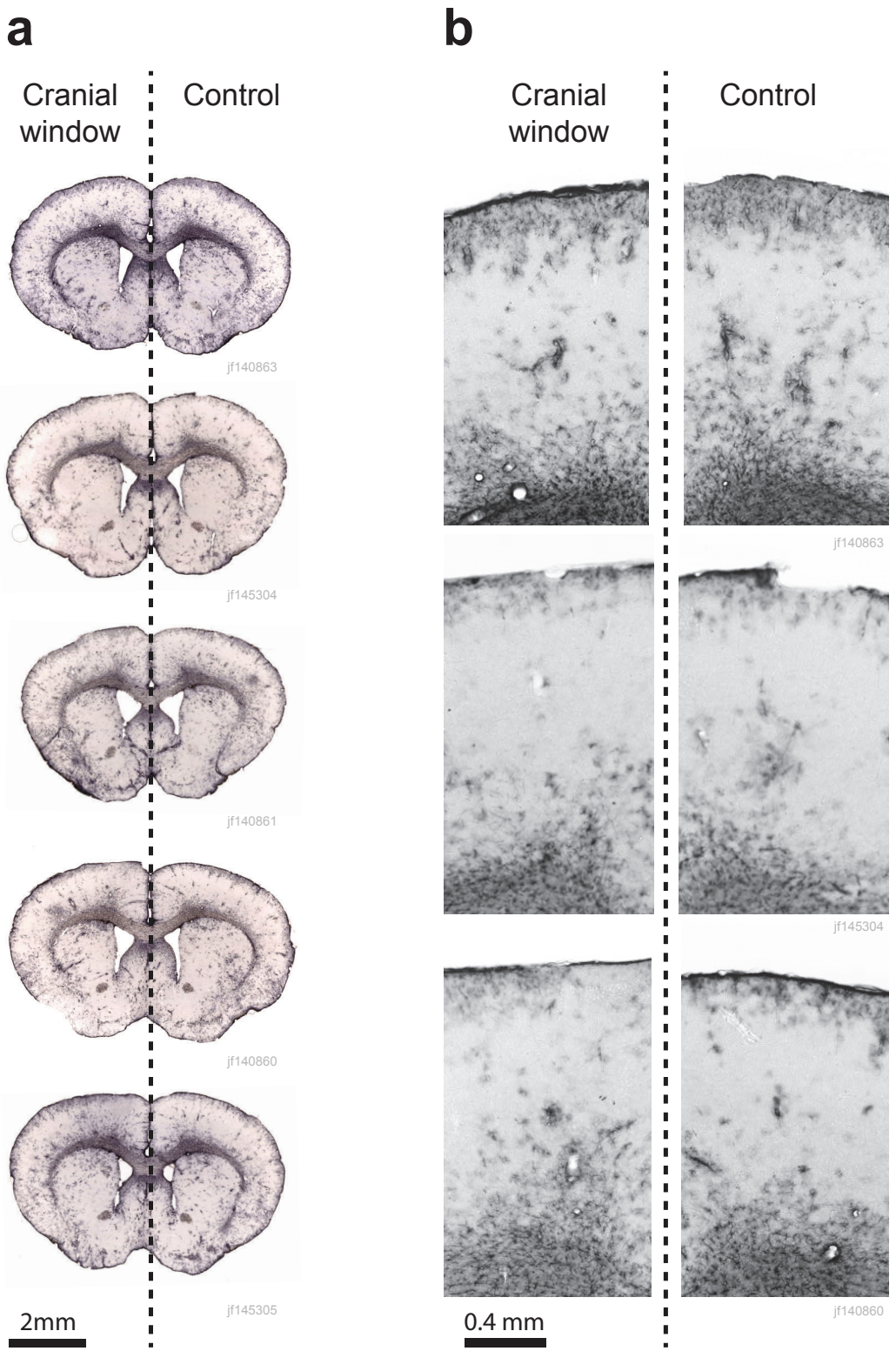
Supplementary Figure 1 | Learning related changes in behavior.

a-c) Changes in motor parameters as a function of performance. **a)** Whisking repeatability, defined as the amplitude of the first peak (at one trial length) of the autocorrelation of whisking amplitude (all trials in one session were concatenated). **b)** Mean whisker protraction angle during the sampling period. **c)** Timing jitter (std) of the first lick relative to first touch. Each circle represents one behavioral session (6 sessions/mouse, 5 mice). Open circles indicate sessions with inactivated vM1. **d-e)** vM1 inactivation in expert mice. **d)** Whisker angle (gray) and setpoint (black) in Hit trials under control conditions (top) and after muscimol inactivation (bottom). **e)** Hit rates decrease and false alarm rates increase after inactivation of vM1 (control, solid circles; muscimol, open circles). **f)** Behavioral performance drops after inactivation of S1 (n = 5 mice, control, solid circles; muscimol, open circles).



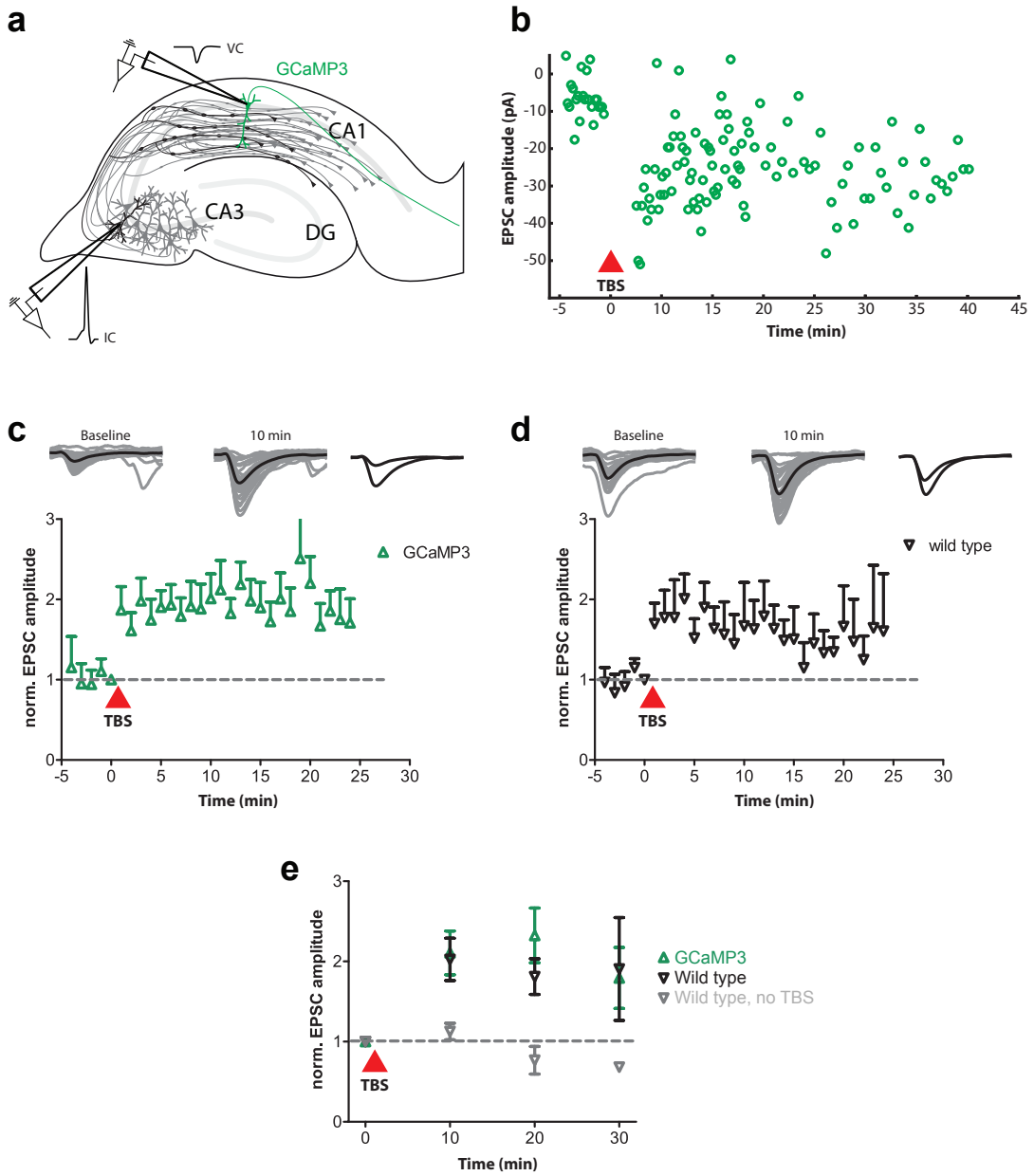
Supplementary Figure 2 | Effects of vM1 inactivation on whisking pattern.

Whisking and effect of muscimol inactivation for 2 representative animals (animal 1 & 3, Supplementary Table T1). The panels represent averages of the spectrogram (1-30Hz, 1s sliding window) of the whisker angle across Hit trials (left) and Correct Rejection trials (right) before, during and after muscimol injection into vM1 (Methods). White dotted lines delineate the sampling period.



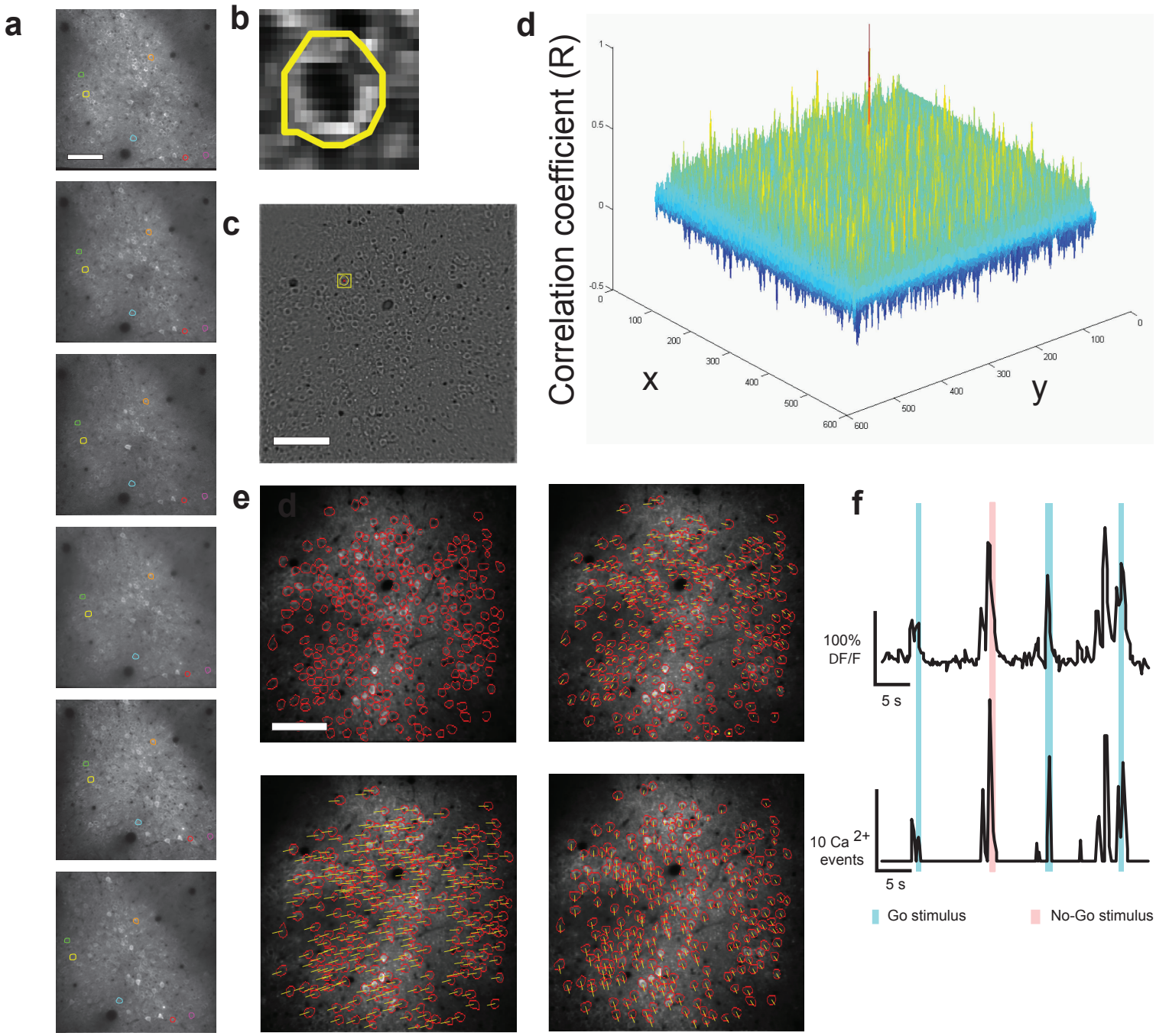
Supplementary Figure 3 | Immunohistochemical staining for glial fibrillary acidic protein (GFAP) after chronic imaging.

a-b) Coronal sections through motor cortex of 5 mice at the level of the craniotomy and the site for chronic imaging (left hemisphere; 90 days after surgery). **a)** low magnification. **b)** high magnification. Immunoreactivity under the imaging window was identical to the control (right) hemisphere.



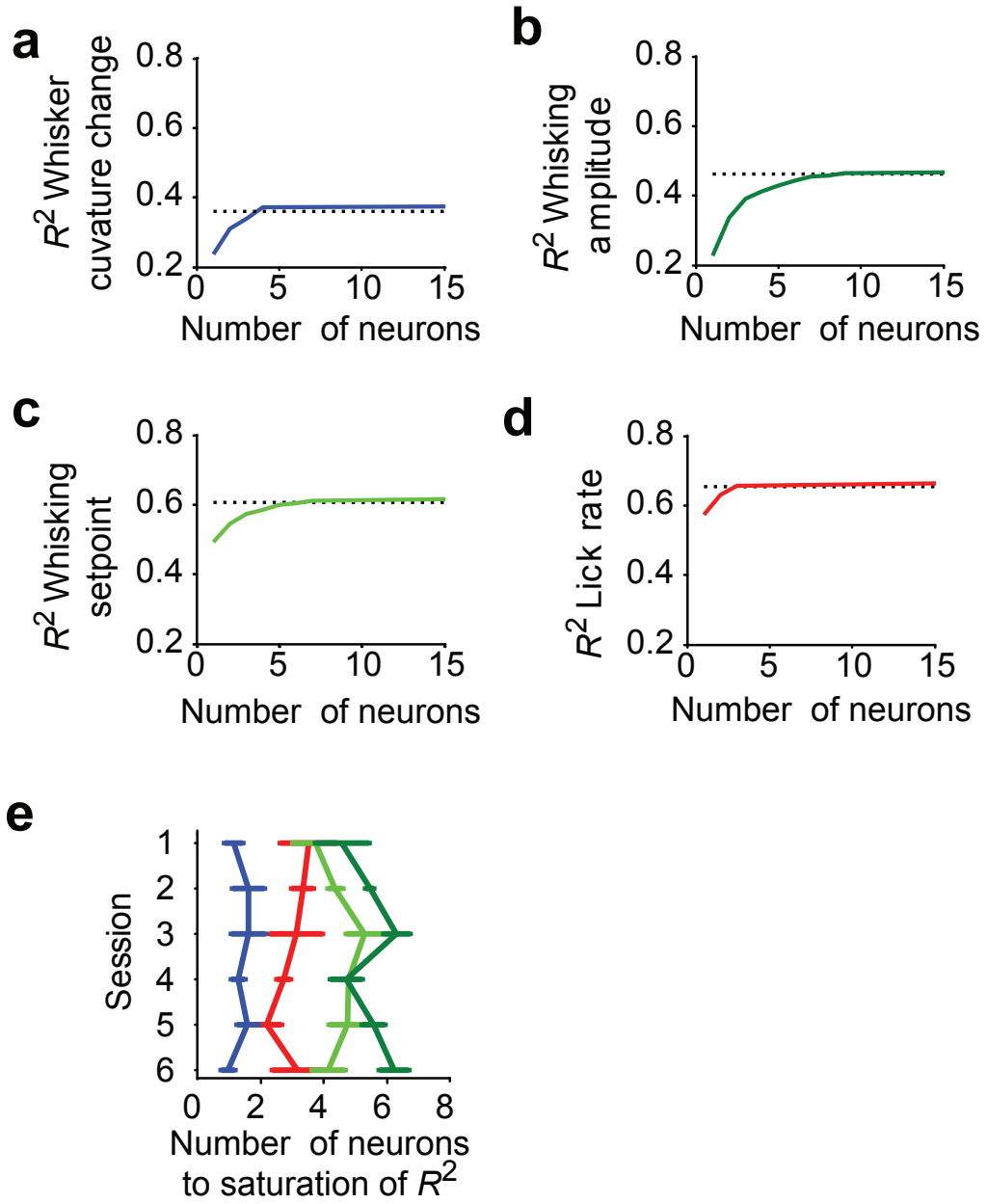
Supplementary Figure 4 | Normal synaptic plasticity in GCaMP3-expressing neurons.

a) Schematic of the experiment. Recordings were made from pairs of CA3-CA1 pyramidal neurons in cultured hippocampal slices. The postsynaptic CA1 cells were expressing CaMP3 and cerulean (volume marker). Parallel experiments were done with untransfected CA1 pyramidal neurons (wild-type). To probe excitatory postsynaptic currents (EPSCs) in CA1 neurons single action potentials were elicited in CA3 cells. **b)** Example LTP experiment in a GCaMP3-expressing CA1 pyramidal neuron. Single action potentials were evoked every 10 seconds for 5 min to establish baseline synaptic transmission. LTP was induced with a theta-burst protocol (TBS; 10 x 5 action potential bursts at 100 Hz; repeated 4x at 5 Hz) paired with temporally matched postsynaptic action potentials. After the pairing protocol, synaptic transmission was probed as during the baseline period.
c-d) Average LTP in GCaMP3-expressing (n = 15) and wild-type (n = 10) CA1 pyramidal cells. Top traces: gray, single responses; black, mean response. **e)** Comparison of LTP between GCaMP3-expressing neurons, wild type CA1 neurons, and wild- type neurons without LTP induction.



Supplementary Figure 5 | Alignment of ROIs across multiple imaging sessions.

a) Field of view over six imaging sessions. A few ROIs were marked with colored ROIs. **b)** Close-up view of an individual ROI. **c)** Field of view after luminance normalization. **d)** A 50 x 50 pixels square around the ROI of interest (yellow box in c) was used to search for the best match in the next day's average image using a normalized cross-correlation. The center of the new ROI position was moved to the location with the highest correlation value (red peak). **e)** Displacement vectors (yellow) of individual ROIs (red) for four consecutive imaging sessions. Scale bars, 100 μm. **f)** A fluorescence trace corresponding to a single ROI (top) and deconvolved event rate (bottom; see Methods).

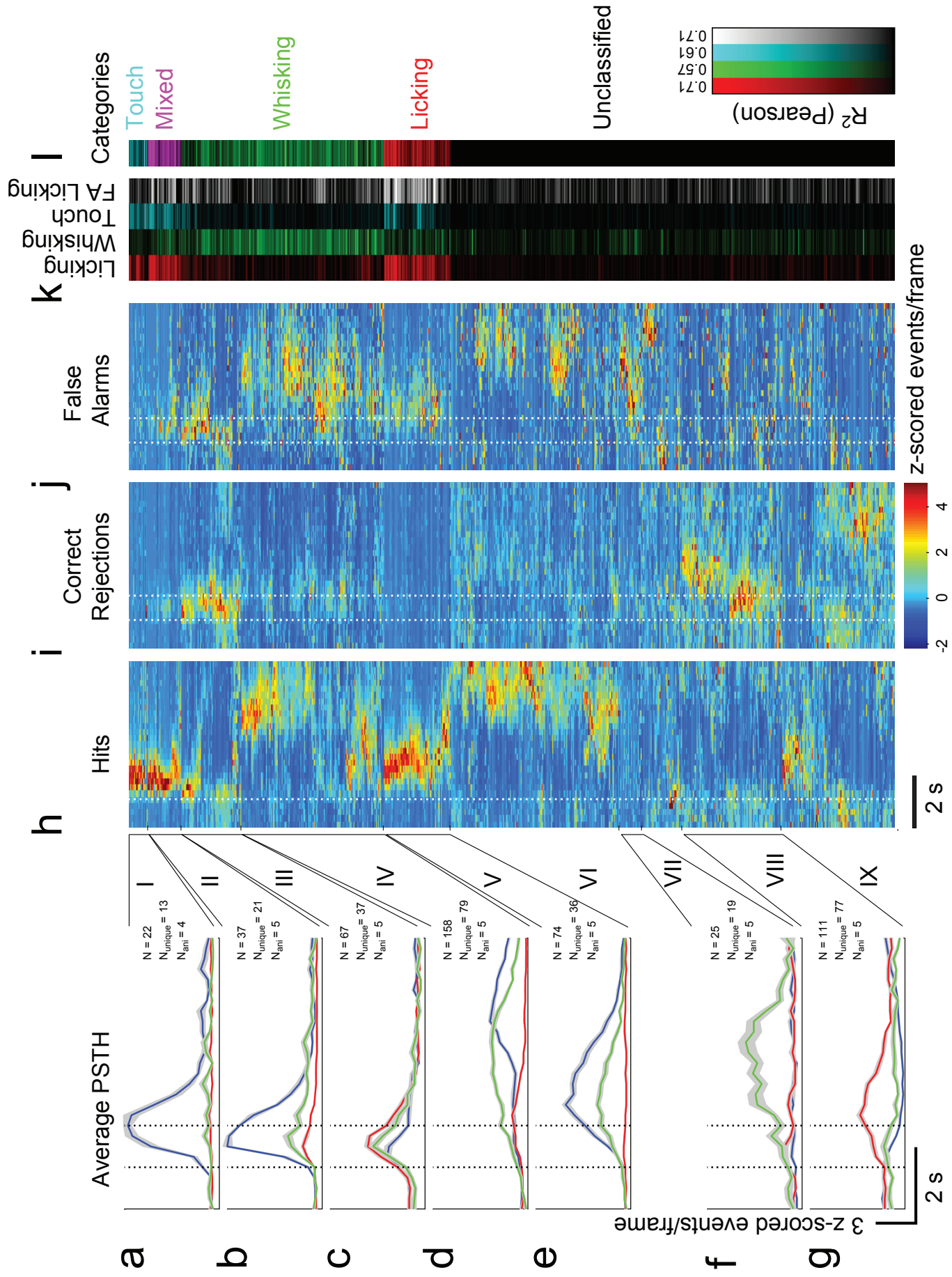


Supplementary Figure 6 | Decoding accuracy as a function of population size

a-e) Decoding accuracy (R^2) as a function of population size (same data as in **Fig. 3a-e**). Starting with the neuron with the highest Ri^2 we added neurons until saturation (95% of the Ri^2 for the full population; black dotted line). **a)** Whisker curvature change. **b)** Whisking amplitude. **c)** Whisking setpoint. **d)** Lick rate. **e)** Number of neurons required to achieve 95% of the full population R^2 for each behavioral variable in function of the session days. The minimum number of neurons required to reach saturation of the population R^2 did not depend on learning sessions (Two-way repeated measures ANOVA; $p>0.6$; $F_{5,20}=0.73$) but it did depend on the decoded behavioral variable ($p<10^{-5}$; $F_{3,12}=42$). Error bars represent sem. Each behavioral feature required a different number of neurons to reach saturation (licking, 3.0 ± 0.3; touch, 1.4 ± 0.2; whisking setpoint, 4.5 ± 0.2; whisking amplitude; 5.5 ± 0.2; mean ± sem; n=30). These numbers are tiny compared to 30,000 L2/3 neurons in vM1; 10^5 neurons per mm^2 of cortical area; vM1 area, 3 mm^2; fractional thickness of L2/3, 0.1, ref 9), suggesting that coding is highly redundant.

doi:10.1038/nature11039

Supplementary Figure S7



Supplementary Figure 7 | Representations of individual neurons.

All highly active neurons were classified across all sessions by how well they represented specific types of behavioral features (852 neurons across all 5 sessions and all 6 animals; 286 unique neurons).

a-g) PSTHs averaged across neurons with shared activity patterns. (Hits, blue; Correct Rejections, red; False Alarms, green; Misses were rare and are thus not shown. N is the number of total observations grouped in one category, N_{unique} is the number of individual neurons in which this response has been observed neurons, N_{ani} is the number of animals in which these neurons were observed). The vertical dashed lines mark the sampling period. Shaded curve represent sem.

a) Touch cells (I) are active only during interactions with the pole during Hit trials (blue) and in Miss trials (relatively rare, not shown). They can be distinguished from licking neurons (V) because they are not active during False Alarm trials (green).

b) Mixed cells (II) are strongly correlated with licking during Hit trials, but also decode additional features such as touch or whisking. In some sessions in highly performing animals there were too few False Alarm trials to determine if a neuron decoded touch or licking better, and these neurons were classified as mixed.

c) Whisking cells, early (III) correlate to whisking before and during the sampling period. We focused on Correct Rejection trials to distinguish between neurons coding for licking or whisking, since licking is usually associated with large whisking amplitudes and changes in Setpoint in Hit trials. Several neurons correlated with whisking only during a specific phase of the task (e.g. the example cell in **Fig. 2g**, Cell A).

d) Whisking cells, late (IV) correlate to the whisking features after the sampling period.

e) Licking cells (V) correlate with licking and therefore show high activity during the response period in Hit and False Alarm trials. Lick neurons tend to fire after touch neurons become active (**Fig. 5**). Unclassified neurons (VI) are active mainly after rewarded (Hits) or unrewarded licking (False Alarms).

f) A group of unclassified neurons (VII) show activity uniquely after a false alarm responses. They could signal an error signal related to the lack of an expected reward.

g) Another group of unclassified neurons (VIII) increase activity during correct rejection trials.

Several neurons in the unclassified category IX, show activity that predicting trial outcome. These neurons show higher firing before correct choices (Hits & Correct Rejections) compared to incorrect choices (False Alarms) before the start of the task. These activity patterns might be related to the animals' attentional state.

h-j) Task-aligned activity for all neuron-sessions averaged over trials, clustered by activity pattern (see note on clustering, below). Activity was normalized by computing the z-score per neuron. The white dotted lines mark the sampling period.

k) Strength of correlation between the Random Forest model and different behavioral features (R^2).

l) Classification based on heuristic rules (see above and also the main text).

Notes on clustering for panels h-j: We divided neurons into five categories (see caption of and main text): Touch, Whisking, Licking, Mixed and Unclassified. Within each category we computed a similarity index between pairs of neurons based on the trial-averaged responses in Hit, Correct Rejection and False Alarms trials. For each neuron

we concatenated their trial-average response for these tree trial-types. The similarity index was the pairwise correlation between the concatenated trial-average responses between every neuronal pair. We computed a weighted correlation to take into account that the number of False Alarm trials is typically lower than either Hit or Correct Rejection trials.

Given the two vectors $\bar{y}$ and $\bar{x}$ and the weight vector $\bar{w}$ (by the relative number of trials in each trial type):

$$weighted\ mean:\ \mu(\bar{x};\bar{w}) = \frac{\sum_i w_i x_i}{\sum_i w_i}$$

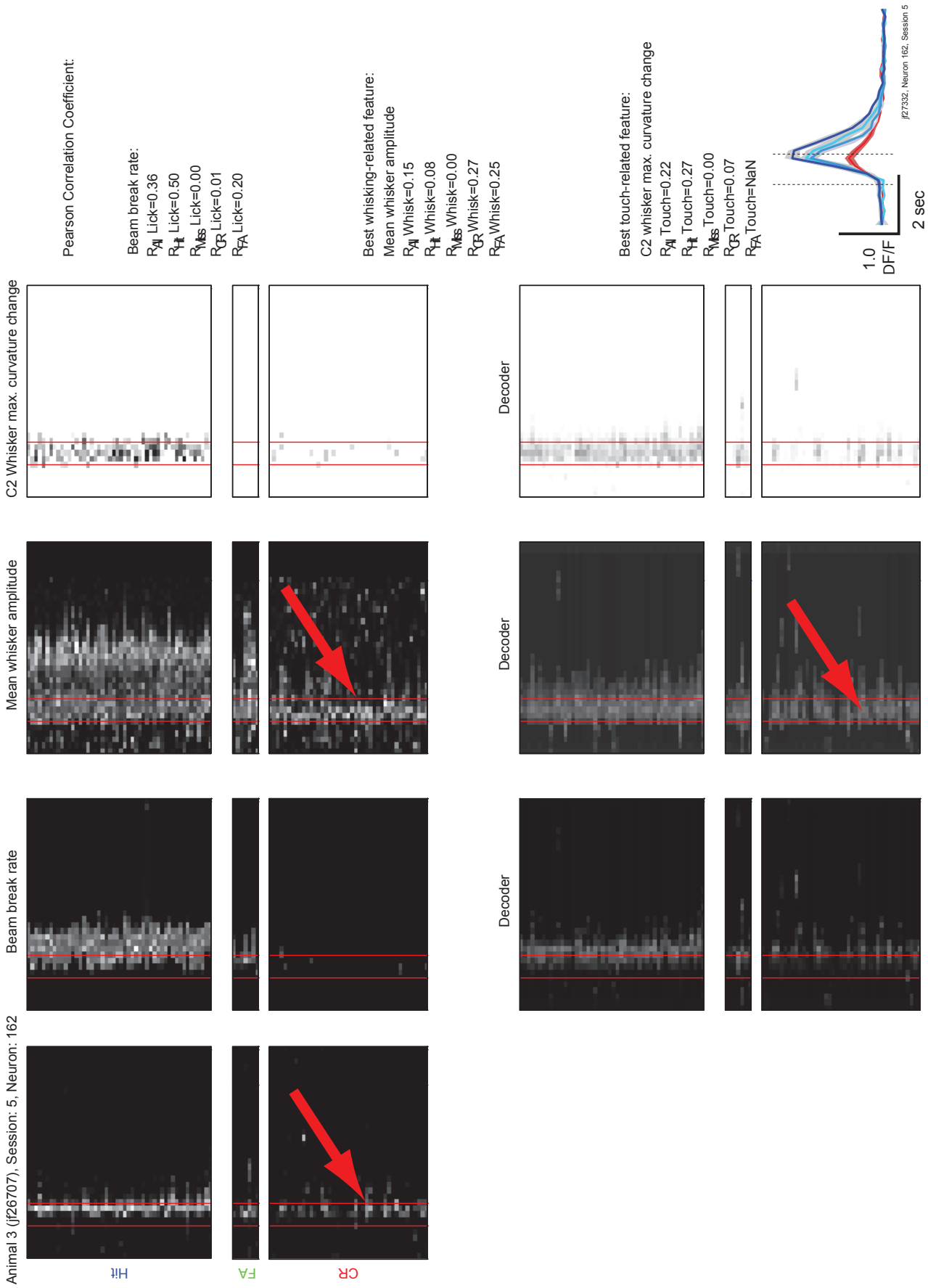
$$weighted\ covariance: cov(\bar{x},\bar{y};\bar{w}) = \frac{\sum_i w_i\big(x_i - \mu(\bar{x};\bar{w})\big)\big(y_i - \mu(\bar{y};\bar{w})\big)}{\sum_i w_i}$$

$$weighted\ correlation: corr(\bar{x},\bar{y};\bar{w}) = \frac{cov(\bar{x},\bar{y};\bar{w})}{\sqrt{cov(\bar{x},\bar{x};\bar{w})cov(\bar{y},\bar{y};\bar{w})}}$$

Neurons were re-ordered using hierarchical clustering for each category separately (using *linkage* and *dendrogram* from Matlab with unweighted average distances between clusters).

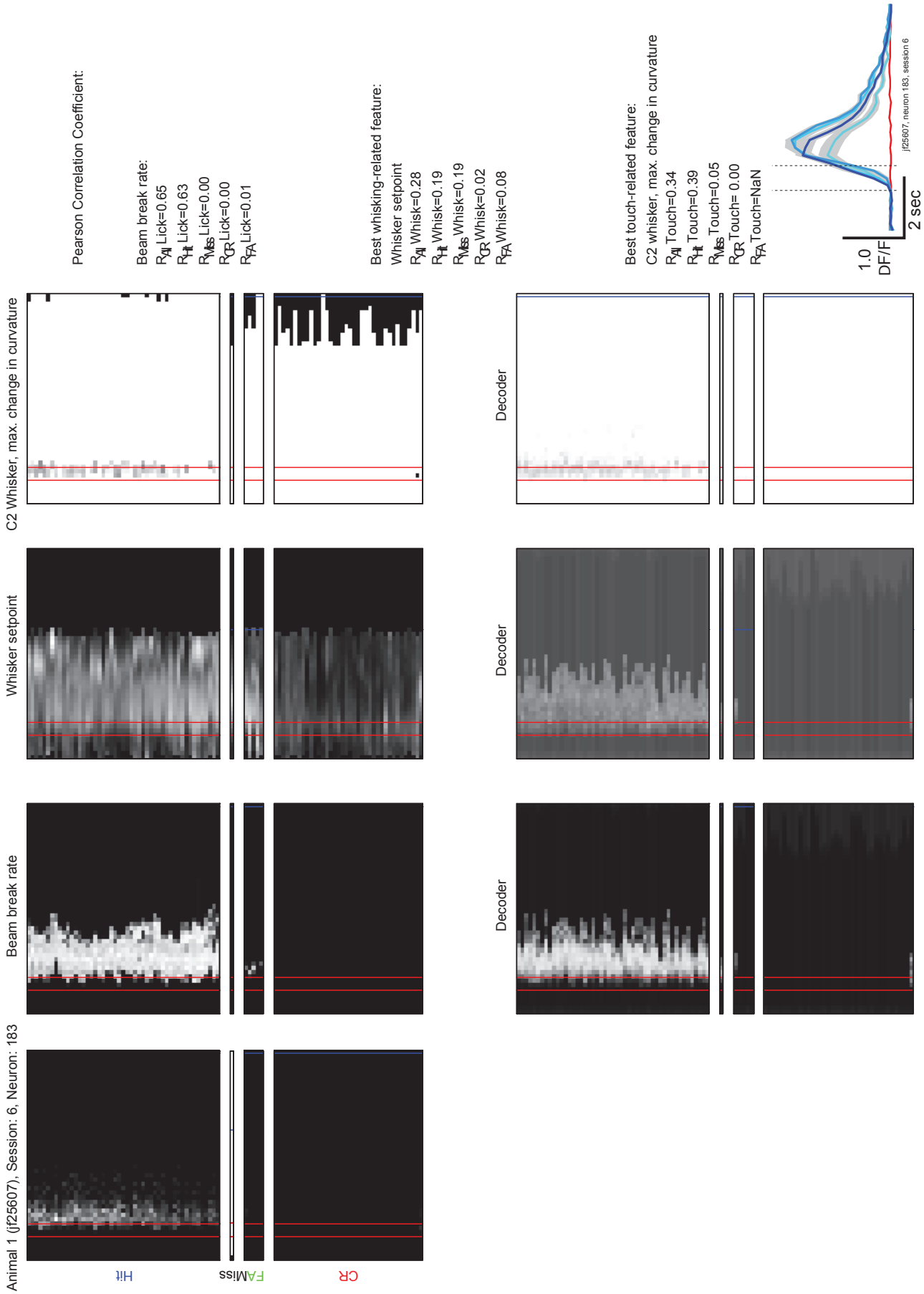
doi:10.1038/nature11039

Supplementary Figure S8



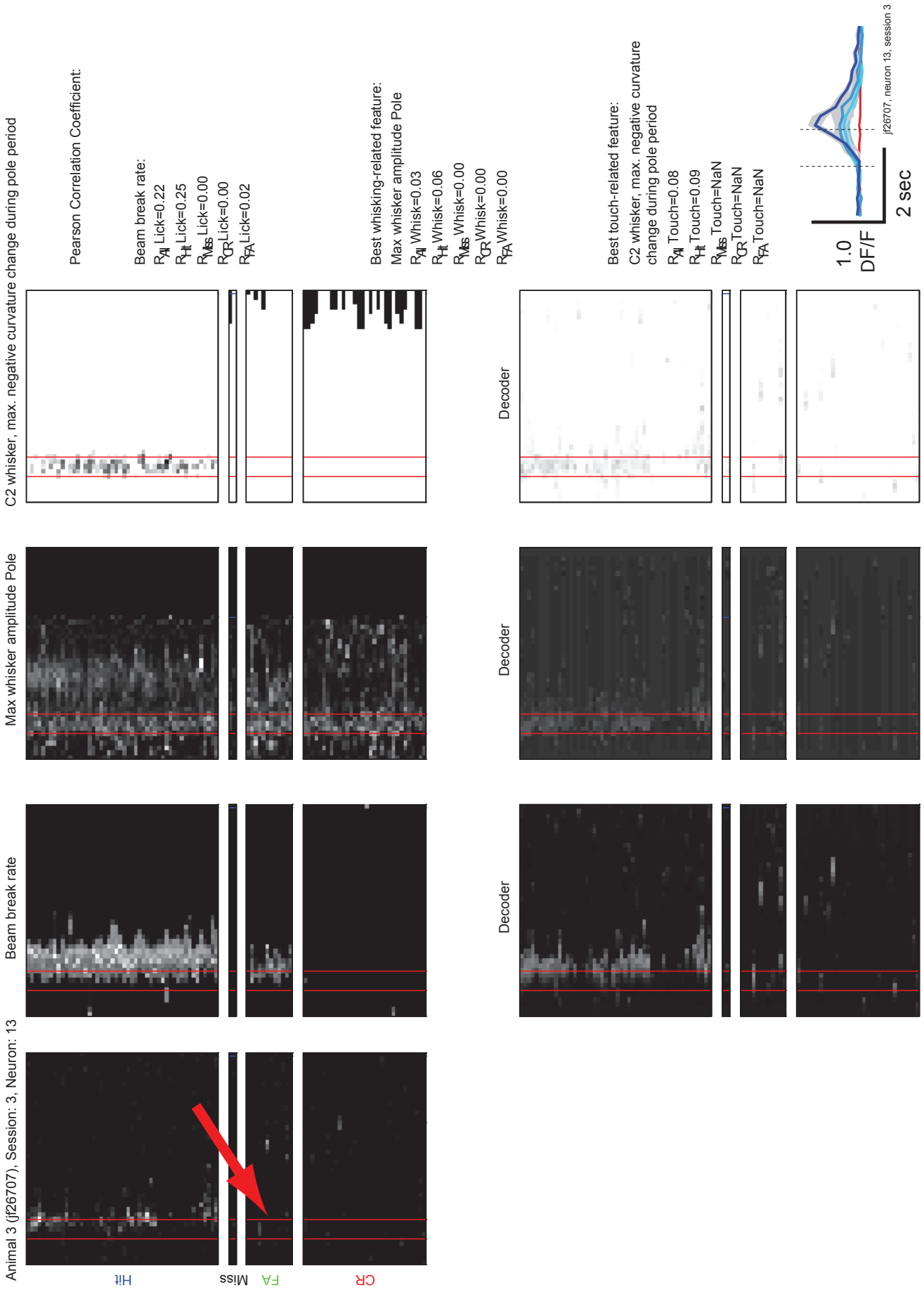
doi:10.1038/nature11039

Supplementary Figure S9



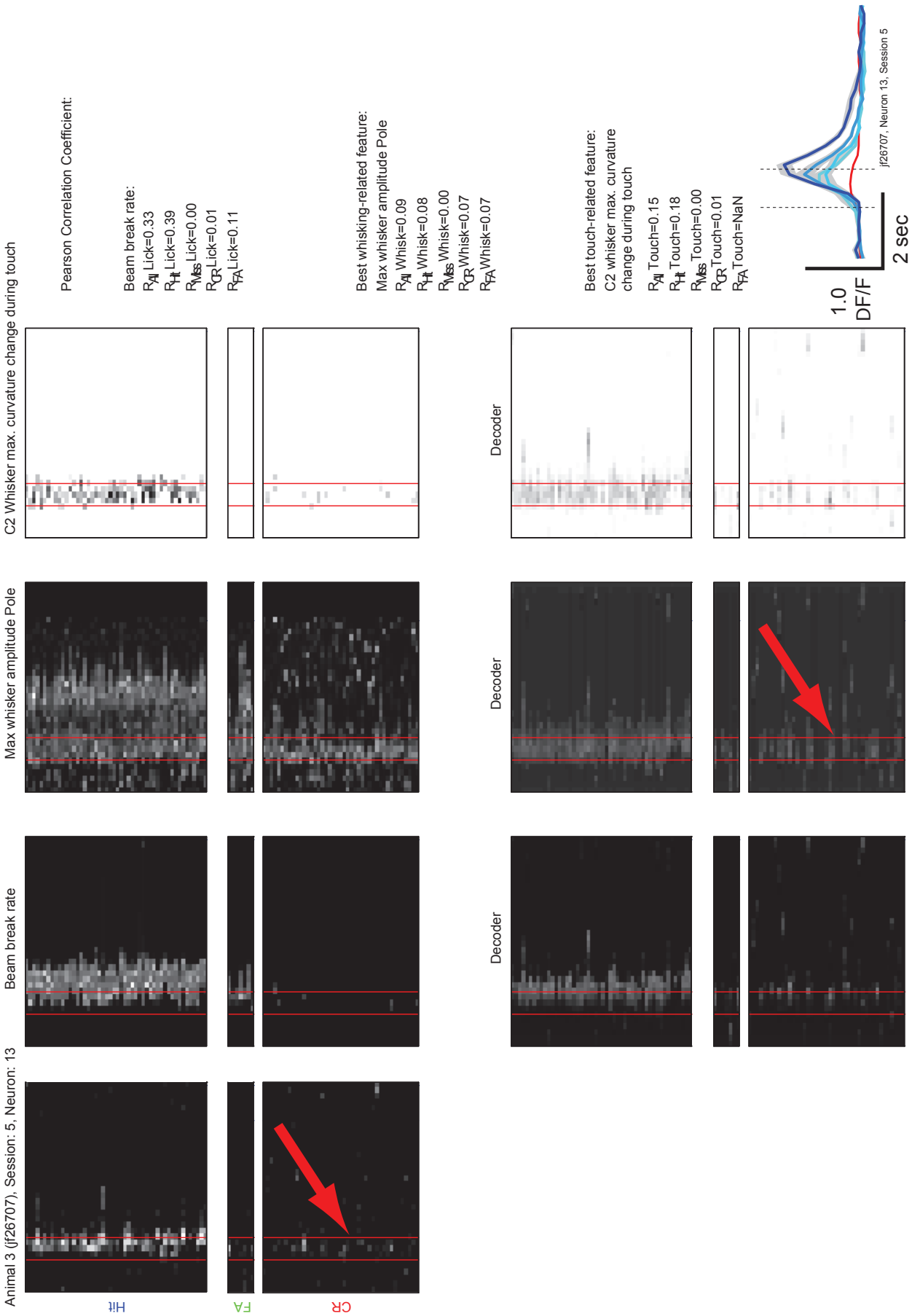
doi:10.1038/nature11039

Supplementary Figure S10



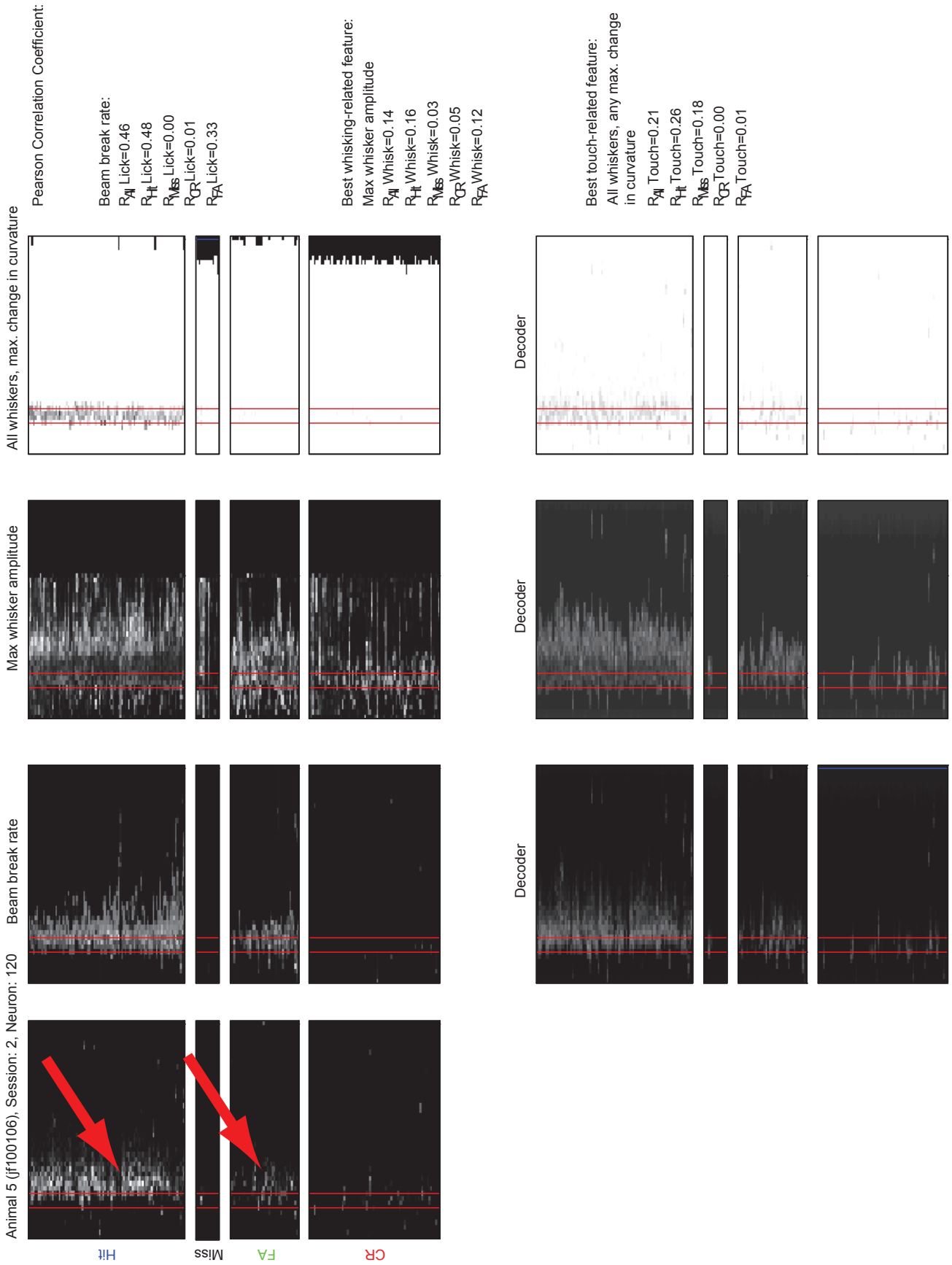
doi:10.1038/nature11039

Supplementary Figure S11



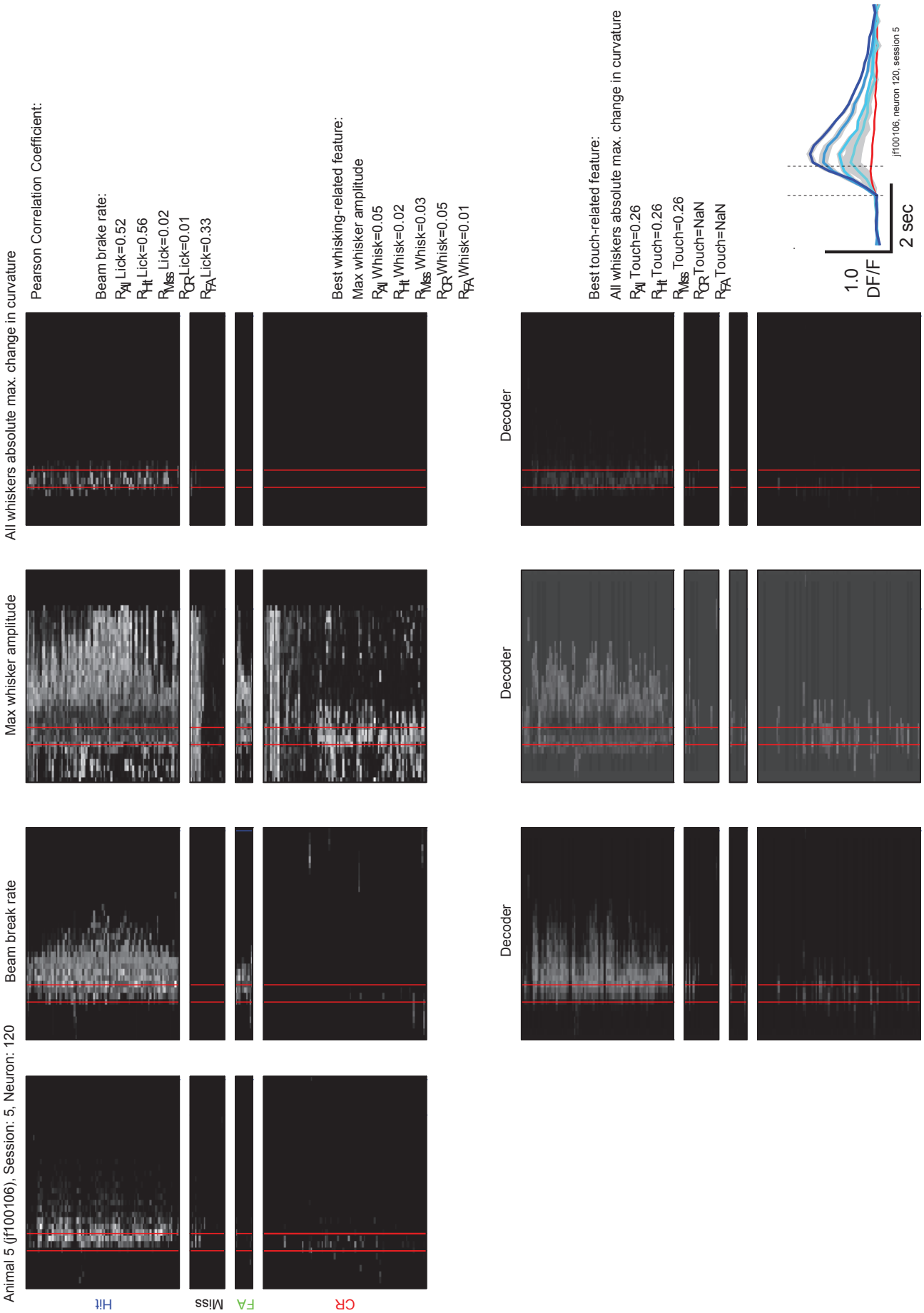
doi:10.1038/nature11039

Supplementary Figure S12



doi:10.1038/nature11039

Supplementary Figure S13



Supplementary Figure 8-13 | Example neurons to illustrate neuron classification.

The upper panels represent from left to right: event rate (events / frame) of an individual neuron, licking (beam break rate), whisking related feature (based on whisker setpoint or amplitude), touch related feature (based on curvature change). Each line represents one trial, grouped vertically by trial type. The red lines mark the sampling period. The lower row of panels shows the model of the best feature within a behavioral category (see Methods) based on the activity of the single neuron, derived from Random Forests. The numbers on the right are the Pearson correlation coefficients between the behavioral feature and the prediction from the model. Only the whisking and touch related features with the highest correlation coefficients are shown. The panel on the right lower corner is a PSTH of the neuron's responses to multiple 'Go'-positions (different hues of blue, light blue is most posterior, see Supplementary **Fig. S14**) and Correction Rejections (red).

8| Mixed neuron with good prediction for touch (overall R^2 value for touch is 0.22), but also whisking-related signals in Correction Rejection trials (red arrows).

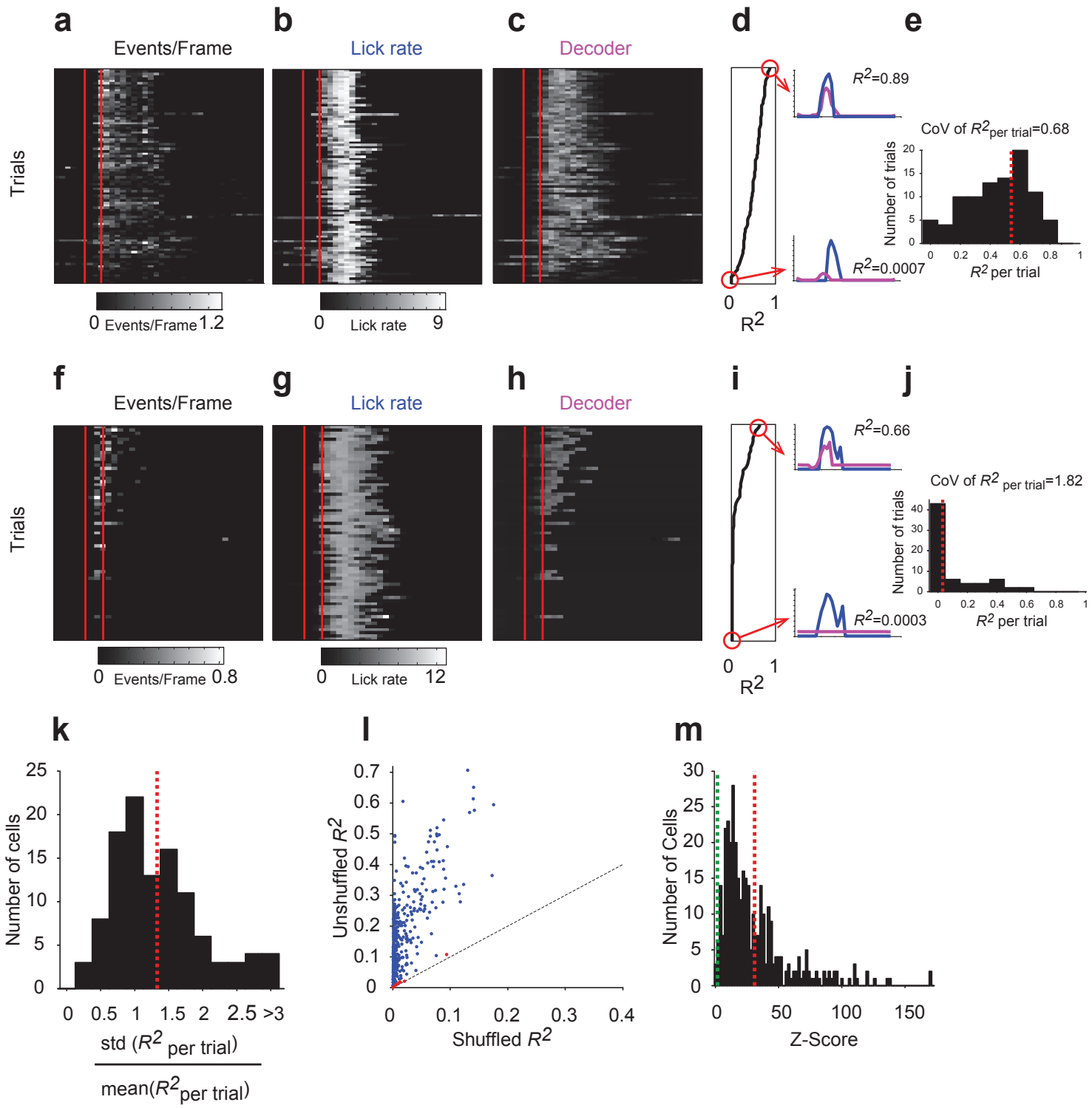
9| Mixed neuron which decodes licking with high accuracy, but is position-dependent.

10-12| Changes in classification from touch to mixed. **10)** Touch neuron could be disambiguated from licking by the lack of activity during False Alarm trials (red arrow). In addition, this neuron shows position dependence. **11)** The same neuron was classified as a mixed type two sessions later. The neuron now shows activity during the sampling period on Correct Rejection and False Alarm trials (red arrow), permitting accurate decoding of whisking during Correct Rejectiontrials. The neuron remained position-dependent.

12-13| Changes in classification from licking neuron to mixed type. **12)** Licking neuron with showing activity during the Hits and False Alarms (red arrows). **13)** Same neuron 2 sessions later classified as mixed type, since it shows less activity during False Alarms, but strong position dependence (inset lower right).

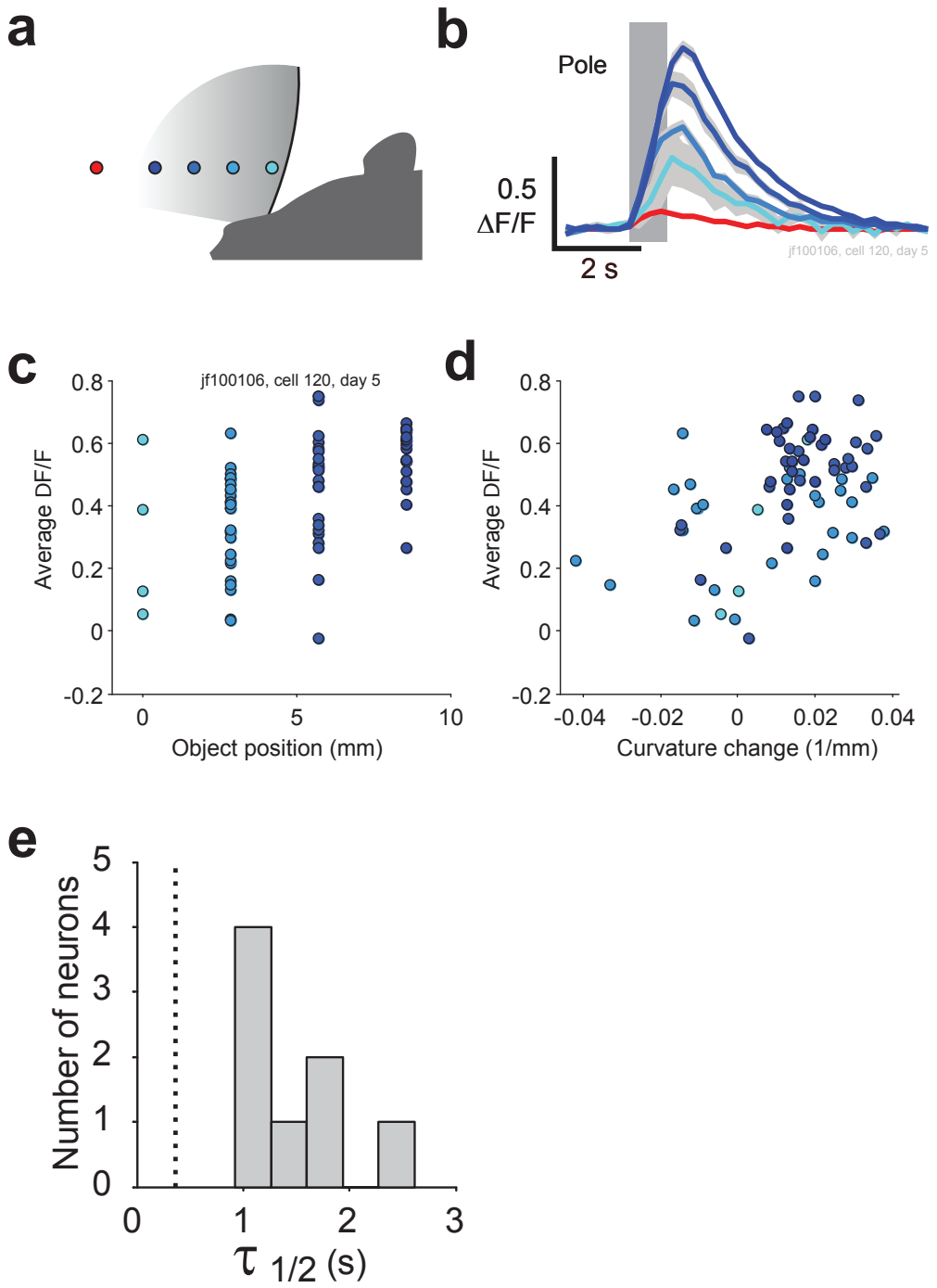
doi:10.1038/nature11039

Supplementary Figure S14



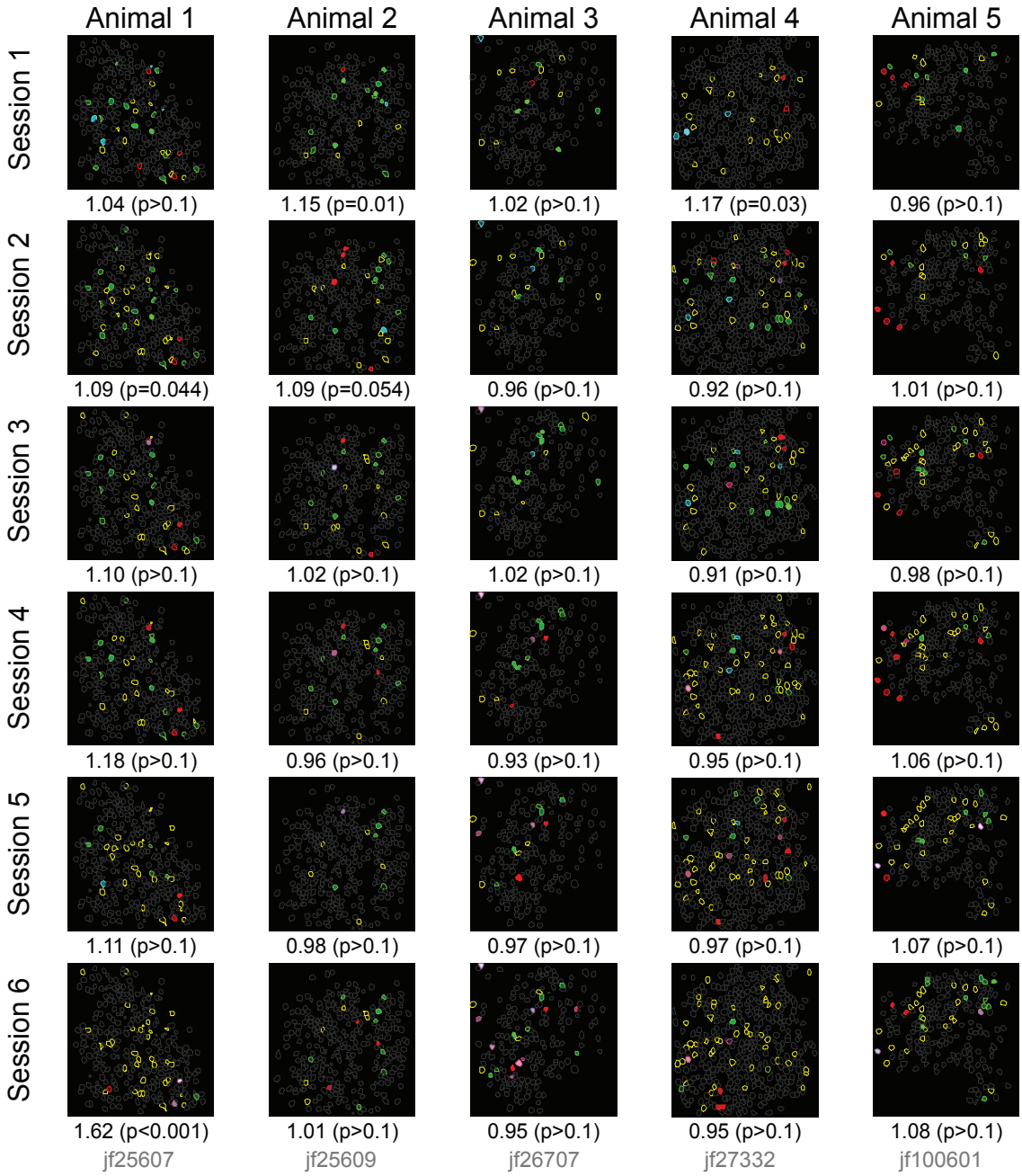
Supplementary Figure 14 | Single neuron decoding algorithms capture the trial-by-trial dynamics of motor and sensory variables.

a) Calcium events in hit trials for a cell classified as a licking neuron in a highly performing mouse (d'>2). Red bars represents the period where the pole was up. **b)** Lick-rate. **c)** Decoded Lick-rate. **d)** R^2 between lick-rate and the decoded lick-rate in each trial ($R^2_{perTrial}$). We sorted trials from the highest to the lowest R^2. The upper and lower panels show two examples of trials with high and low $R^2_{perTrial}$ respectively. **e)** Distribution of $R^2_{perTrial}$ for all trials. The dotted red line represents the median of $R^2_{perTrial}$. To summarize how reliable the cell was able to decode licking trial-by-trial we computed the coefficient of variation of $R^2_{perTrial}$ (COV = standard deviation of $R^2_{perTrial}$ divided by its mean value). **f-j)** Same as (a-e) but for an 'unreliable' lick cell in a highly performing mouse (d'>2). **k)** Distribution of reliability coefficients (i.e. CoV of $R^2_{perTrial}$) for all the classified neurons in highly performing mouse (d'>2) computed only for hit trials (N=111). The median CoV was 1.33 (dotted red) and ranged from 0.39 (most reliable) to 3.3 (least reliable). **l)** Trial-shuffled bootstrap test of the individual neuron's decoder. To assess whether the correlations between measured behavioral variables and the decoded variables appeared by chance, we shuffled trials (1000 times) and recomputed the Pearson correlation coefficient between data and model (R^2). The p-value was calculated as the proportion of trial-shuffled bootstrap replicates that were higher than the unshuffled R^2. This test indicates whether the decoder extracts information beyond a given trial-averaged behavioral variable. We performed this test in all the classified neurons (N=358) in the hit trials for the behavioral variable that the neuron codes for. Only 7 neurons (~2%) were non-significant (P>0.05; red dots). **m)** Distribution of the z-score converted R^2 values: ($[R^2_{unshuffled} - mean(R^2_{shuffled})]/std(R^2_{shuffled})$).The dotted red line represents the median z-score and the green line represents Z=1.64 that corresponds to a significance value of 0.05.



Supplementary Figure 15 | Neural activity modulated by object location.

a) Schematic of object locations. **b)** Activity as a function of object location (same color code as in a, lines are mean, shading standard error of the mean). The grey region indicates the sampling period. Note that object location dependent activity outlasts the presence of the pole by several seconds. **c)** The activity of an example neuron represented as a function of object object location. **d)** Activity as a function of whisker curvature change. **e)** Persistent object location-dependent activity. Activity (as in b) was deconvolved with the GCaMP3 decay time (ref 34) to correct for the dynamics of calcium and GCaMP3. Activity was then normalized to the value at the time of last touch and the decay time of the signal was computed. The histogram indicates the decay time of persistent activity after last touch. The dashed vertical line indicates $\tau_{1/2}$ (384ms) expected for GCaMP3 fluorescence without persistent activity (ref 34).

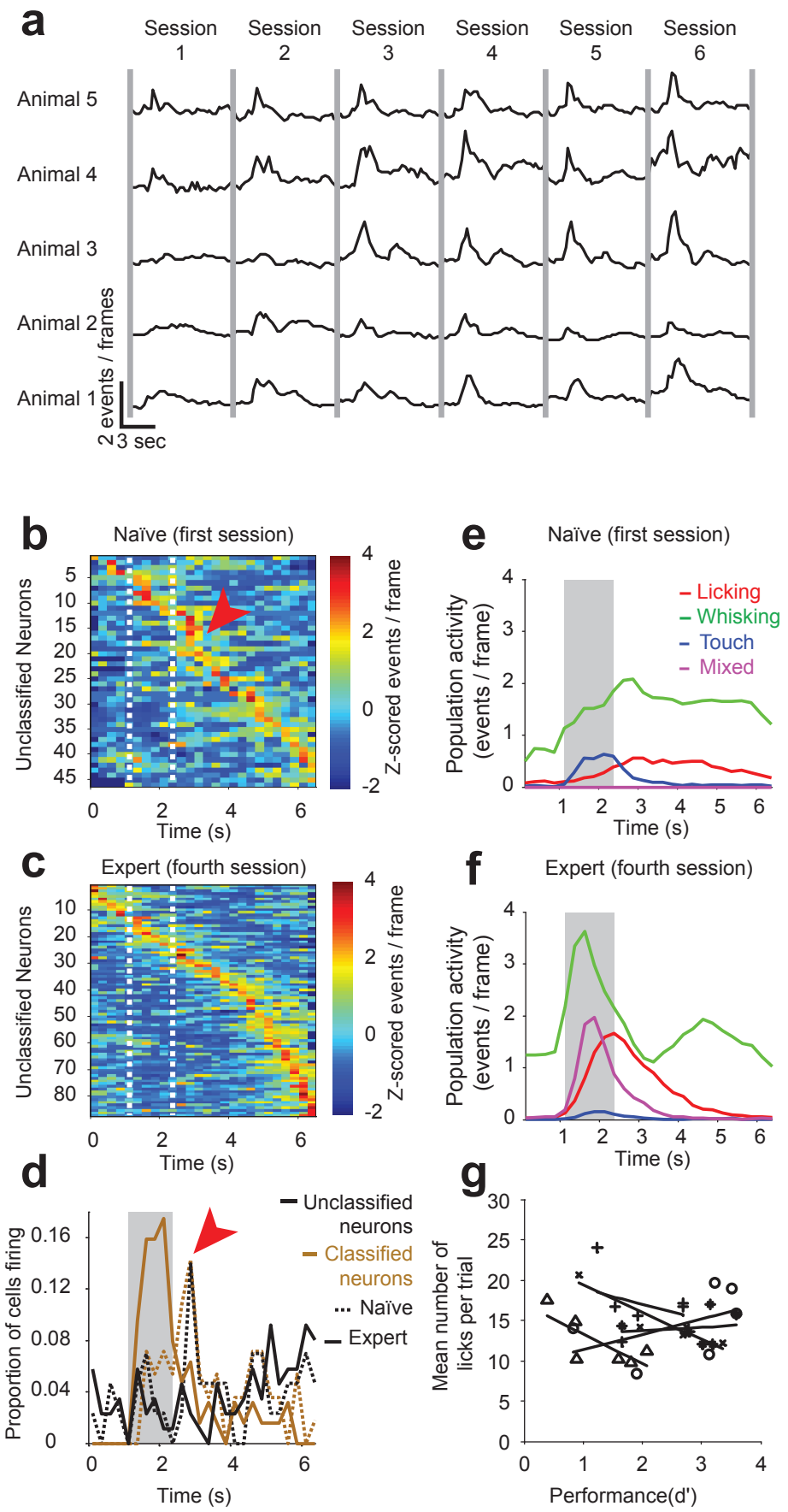


Supplementary Figure 16 | Spatially intermingled representations.

Spatial locations of classified neurons. Individual regions of interest (ROIs) colored according to the scale in a. ROIs with yellow boundaries correspond to non-classified, task-related neurons. Numbers below images denote the spatial clustering index (SCI; > 1, clustering; 1, unclustered) and the p-value of clustering. We measured the center coordinates of each active cell (including active but unclassified cells). SCI is the ratio of the mean distance between neuronal pairs belonging to different representations divided by the mean distance between neuronal pairs belonging to the same representations exclusively. Under the null hypothesis that different representations are intermingled, SCI ~ 1. SCI > 1 indicates spatial clustering. To test for statistical significance we shuffled the cell's labels and computed SCI on the shuffled data 1000 times; the p-value is the fraction of shuffled SCI higher than the unshuffled SCI. Significant clustering was not detected.

doi:10.1038/nature11039

Supplementary Figure S17



Supplementary Figure 17 | Temporal structure of neuronal activity.

a) Average event rate. Each line represents the mean number of events per frame across all trial types and summed across all neurons. **b-c)** Trial averages of unclassified neurons ordered by their peak activity (for classified neurons see **Fig. 5**). **b)** Neurons in naïve animals (first session). The arrowhead indicates an enhanced activation of neurons during the reward period in naive animals. **c)** Neurons in expert animals (fourth session). **d)** Fraction of neurons with peak firing rate during the trial. Dotted lines indicate first day, the solid lines the fourth day in classified (brown) and unclassified neurons (black). In naïve mice there was an over-representation of neurons active during the reward period (red arrowhead). In expert mice these effects disappeared for classified and unclassified neurons (see **Fig. 5**). These neuronal dynamics are consistent with an unexpected reward delivery in naïve animals and a predicted future reward delivery after pole contact in trained animals. Note also that unclassified neurons have peak activity spread out throughout the entire trial. These neurons could serve to assign the appropriate temporal credit to the performed motor actions based on actual and/or expected reward delivery. Alternatively, these neurons could play a role in orchestrating the appropriate sequence of motor actions during learning. **e-f)** Temporal structure of the sum of activity of different classified neuron types (licking, red; touch, blue; mixed, magenta; whisking, green). **e)** Naive animals. **f)** Expert animals (fourth session). Whisking cells are active early, followed by touch (or mixed neurons conveying contact information) and then lick neurons. Whisking neurons exhibit a second peak that is correlated with high-amplitude post-reward whisking (see **Fig. 2g**). Skilled performance in a sensorimotor task is linked to a reorganization of neuronal firing of specific response-types in an appropriate sequence.**g)** Average number of licks per trial as a function of behavioral performance (each symbol corresponds to an animal; lines are linear fits). d' is a measure of behavioral performance (expert > 1.75, corresponding to approximately 80 %).

Days post-injection	Present whiskers	Pole position
Active neurons (total neurons)	Performance in d' (trials)	Event rate (events/sec/roi)

Animal	Session 1			Session 2			Session 3			Session 4			Session 5			Session 6		
#1 (jf 25607)	14	C1,C2, C3	Single go	15	C1,C2, C3	Single go	16	C1,C2, C3	Single go	17	C1,C2, C3	Multi go	21	C1,C2, C3	Multi go	23	C1,C2, C3	Multi go
	32 (290)	d'(322)=0.83	$3.4*10^{-4}$	40 (290)	d'(389)=1.90	$3.9*10^{-4}$	31 (290)	d'(329)=3.13	$4.8*10^{-4}$	32 (290)	d'(277)=3.59	$6.4*10^{-4}$	33 (290)	d'(232)=3.52	$4.5*10^{-4}$	39 (290)	d'(215)=3.22	$6.8*10^{-4}$
#2 (jf 25609)	14	C1,C2, C3	Single go	15	C1,C2, C3	Single go	16	C1,C2, C3	Single go	17	C1,C2, C3	Multi go	21	C1,C2, C3	Multi go	23	C1,C2, C3	Multi go
	18 (285)	d'(361)=1.66	$3.6*10^{-4}$	25 (285)	d'(270)=3.02	$6.8*10^{-4}$	22 (285)	d'(373)=3.19	$6.4*10^{-4}$	17 (285)	d'(247)=2.77	$5.5*10^{-4}$	11 (285)	d'(210)=3.16	$10*10^{-4}$	17 (285)	d'(143)=3.58	$6.6*10^{-4}$
# 3 (jf 26707)	14	C1,C2, C3	Single go	15	C1,C2, C3	Multi go	16	C1,C2, C3	Multi go	18	C1,C2, C3	Multi go	20	C1,C2, C3	Multi go	22	C1,C2, C3	Multi go
	14 (184)	d'(224)=0.93	$12*10^{-4}$	15 (184)	d'(288)=1.96	$11*10^{-4}$	17 (184)	d'(278)=2.76	$12*10^{-4}$	14 (184)	d'(209)=2.74	$17*10^{-4}$	14 (184)	d'(327)=3.35	$11*10^{-4}$	21 (184)	d'(300)=2.69	$15*10^{-4}$
#4 (jf 27332)	14	C1,C2, C3	Single go	15	C1,C2, C3	Single go	16	C1,C2, C3	Single go	18	C1,C2, C3	Multi go	20	C1,C2	Multi go	22	C1,C2	Multi go
	23 (380)	d'(267)=0.39	$0.2*10^{-4}$	39 (380)	d'(261)=0.86	$3.5*10^{-4}$	37 (380)	d'(272)=0.89	$2.9*10^{-4}$	52 (380)	d'(405)=1.80	$4*10^{-4}$	52 (380)	d'(259)=1.60	$4.2*10^{-4}$	55 (380)	d'(201)=2.06	$1.6*10^{-4}$
# 5 (jf 100106)	14	C1,C2, C3	Single go	15	C1,C2	Single go	16	C1,C2	Multi go	17	C2	Multi go	18	C2	Multi go	19	C2	Multi go
	16 (197)	d'(384)=1.23	$9.4*10^{-4}$	21 (197)	d'(226)=1.93	$1.6*10^{-4}$	37 (197)	d'(426)=1.66	$6.1*10^{-4}$	35 (197)	d'(364)=1.54	$7.9*10^{-4}$	36 (197)	d'(228)=2.69	$7.7*10^{-4}$	37 (197)	d'(166)=2.69	$10*10^{-4}$

Supplementary Table 1 | Summary data for all 5 animals and 6 behavioral sessions.

Top row, left to right: Number of days after window surgery; remaining whiskers; pole position (single go position, multiple go positions). Bottom row, left to right: Number of active neurons (total number of cells); behavioral performance expressed as d'; mean event rate across all measured ROIs (Hz).

doi:10.1038/nature11039

		Day(n+1)					
		Silent	Lick	Touch	Whisk	Mixed	Unclassified
Day(n)	Silent	507	23	2	29	7	179
	Lick	18	30	1	0	7	7
	Touch	7	0	7	0	6	2
	Wisk	33	1	0	114	0	46
	Mixed	2	2	3	1	14	1
	Unclassified	114	7	0	44	3	213

Supplementary Table 2 | Dynamics of classified neurons.

Matrix of changes in response classification across subsequent days. Only neurons that were at least one day categorized as Touch, Licking, Whisking, Mixed or Unclassified are shown (n=286 unique neurons, across 5 animals and 6 sessions).