Predictive and reactive reward signals conveyed by climbing fiber inputs to cerebellar Purkinje cells

Dimitar Kostadinov[1], Maxime Beau[1], Marta Blanco Pozo[1,2], and Michael Häusser[1]

1. Wolfson Institute for Biomedical Research and Department of Neuroscience, Physiology and Pharmacology, University College London, London, UK

2. Present address: Department of Experimental Psychology, University of Oxford, Oxford, UK

Corresponding authors:

Michael Häusser (m.hausser@ucl.ac.uk) or Dimitar Kostadinov (dimvladkost@gmail.com)

ABSTRACT

There is increasing evidence for a cerebellar contribution to cognitive processing, but the specific input pathways conveying this information remain unclear. We probed the role of climbing fiber inputs to Purkinje cells in generating and evaluating predictions about associations between motor actions, sensory stimuli, and reward. We trained mice to perform a visuomotor integration task to receive a reward and interleaved cued and random rewards between task trials. Using two-photon calcium imaging and Neuropixels probe recordings of Purkinje cell activity, we show that climbing fibers signal reward expectation, delivery, and omission. These signals map onto cerebellar microzones, with reward delivery activating some microzones and suppressing others, and reward omission activating both reward-activated and reward-suppressed microzones. Moreover, responses to predictable rewards are progressively suppressed during learning. Our findings elucidate a specific input pathway for cerebellar contributions to reward signaling and provide a mechanistic link between cerebellar activity and the creation and evaluation of predictions.

INTRODUCTION

The cerebellum is thought to facilitate smooth behavioral execution and learning by generating expectations about the sensory consequences of actions and using sensory input to inform future motor output – i.e. forming internal models of how we interact with the world (1, 2). Purkinje cells, the output neurons of the cerebellar cortex, are crucial for the construction and updating of internal models (3). These neurons receive thousands of inputs from parallel fibers carrying contextual sensory and motor information, and a single but exceptionally strong input from a climbing fiber. These climbing fibers, which generate complex spikes in Purkinje cells, carry supervisory instructive signals and modify the synaptic weights of parallel fiber inputs to Purkinje cells (4-6).

Climbing fiber activation triggers complex spikes in Purkinje cells at low rates (~0.5-2 Hz) and yet exerts a powerful influence on cerebellar function at the level of Purkinje cell populations. This is due to the anatomical and functional relationships between inferior olive neurons, the source of climbing fibers, and Purkinje cells (7, 8). Olivary neurons are gap-junction coupled and exhibit subthreshold oscillations (9, 10), and neighboring olivary neurons, which innervate neighboring Purkinje cells, fire action potentials synchronously and consequently trigger synchronous complex spikes in neighboring Purkinje cells. In this way, functional clusters of Purkinje cells, known as microzones, experience climbing fiber activation coherently and coordinate cerebellar output via synchronous output to neurons of the cerebellar nuclei (11). Population recording methods have recently made it possible to address how olivary neurons engage Purkinje cell populations during behavior (12-20).

In well-studied tasks that engage the cerebellum, the instructive signals conveyed by climbing fibers are usually considered as error signals in an extrinsic framework, e.g. retinal slip during visual tracking (21, 22). However, there is increasing evidence for the cerebellum's involvement in a higher-order processing (23) including spatial navigation (24), language processing (25), and, notably for our study, reward (15, 26). We therefore examined whether the climbing fiber inputs to Purkinje cells may carry internally generated instructive signals and tested this possibility directly by studying how reward context is represented by climbing fiber inputs to Purkinje cell populations.

We demonstrate topographically organized encoding of reward context in the complex spiking patterns of Purkinje cell populations in lobule simplex of the cerebellar cortex, a region traditionally thought to modulate forelimb movements (15, 27). We recorded dendritic calcium signals (a proxy for climbing fiber input and complex spikes) using two-photon microscopy and made direct recordings of complex spikes using Neuropixels probes while mice received rewards with varying degrees of predictability: after performing a trained motor action, after a

tone-cue that preceded reward by a fixed delay, and randomly without prompting. Population activity of Purkinje cells represented reward context in a diverse but predictable manner organized spatially into microzones: some microzones exhibited elevated activity at reward delivery (‘reward-activated microzones’), while other microzones were inhibited (‘reward-suppressed microzones’). Some of these microzones also exhibited an elevated rate of complex spiking in anticipation of upcoming reward, with this behavior preferentially expressed in reward-suppressed microzones. When rewards were omitted on motor trials, both reward-activated and reward-suppressed microzones exhibited omission-related feedback error signals. Omitting tone-cued rewards also triggered feedback error signals, and these signals occurred just after the time of expected reward. Finally, the degree of reward predictability modulated reward-related sensory responses in a graded fashion: the more predictable the reward, the smaller the sensory response it triggered. Combined with the recent demonstration that cerebellar granule cells also encode reward context (26), our data demonstrate that the cerebellar cortex has access to the information streams necessary to create and evaluate expectations about higher-order variables.

RESULTS

Population Purkinje cell complex spike recordings during a sensorimotor task

We trained mice to perform a visually guided sensorimotor integration task to study the variety of climbing fiber signals conveyed to Purkinje cells during behavior. Mice were head-fixed in front of an array of monitors and trained to use a steering wheel placed in front of their forepaws to control a virtual object (**Figure 1A, left**). The object appeared at an eccentric visual position ~45^{o} from the visual midline, and mice had to move it into the center of the environment to receive a delayed water reward. After ~10 days of pre-training aimed to establish an association between steering wheel turns and virtual object movement, mice were transitioned to a more difficult version of the task. Here, the object always appeared on the same side (left) at a fixed position, and the mice were required to make a single, continuous turn of the steering wheel to translate the object/wheel to a visible target region ±15^{o} from the visual midline (**Figure 1A, right**). Each trial was initiated by the appearance of the object, and well-trained mice initiated movements as soon as the object appeared. After the mouse initiated a wheel movement, trial outcome was assessed by the position of the object when the mouse had, for the first time since trial start, stopped moving the wheel for 100 ms continuously. If the object was positioned within the central target region, a delayed reward was given 500 ms after the object was stopped (400 ms from trial evaluation). Mice performed 214 ± 8 trials per session (mean ± s.e.m., n = 61 sessions from 6 mice), and there were 3 possible outcomes of each trial: undershoots, correct (rewarded) trials, and overshoots (**Figure 1B**). Behavioral performance plateaued after <1 week on this final task version and resulted in the following breakdown of performance: 28 ± 4% undershoots, 59 ± 2% correct trials, and 13 ± 3% overshoots (mean ± s.e.m., n = 6 mice, averaged within mouse for sessions ≥ 5 of final task version; **Figure 1C**). By comparison, performance on days 1 and 2 of the final task version was significantly lower (46 ± 3% correct trials, mean ± s.e.m., n = 6 mice, $p < .05$).

To image Purkinje cell populations during our task, we expressed Cre-dependent GCaMP6f virus in Pcp2(L7)-cre mice. Injections were targeted to the left lobule simplex and adjacent vermis (**Figure 1D**), regions known to be involved in forelimb movements (15, 27). Purkinje cell population activity was recorded using resonant scanning two-photon microscopy to measure dendritic calcium signals, faithful indicators of climbing fiber input and complex spiking in Purkinje cells (28-31). Our fields of view (FOVs) yielded 219 ± 27 (mean ± s.e.m., n = 13 different fields from 9 mice) distinct dendritic regions of interest (ROIs) corresponding to individual Purkinje cells (**Figure 1E**). Individual dendritic ROIs exhibited fast calcium transients indicative of complex spikes, and we extracted the size and timing of these events for each recorded neuron (**Figure 1F**).

As a first step towards understanding how Purkinje cell activity is related to different aspects of the behavior, we aligned the calcium responses of our neurons to two important time points in the task across trial outcomes: (1) movement initiation and (2) movement termination and reward delivery (which occurred with a fixed time interval), and sorted neurons by their response during these epochs on correct trials. Subsets of Purkinje cells exhibited elevated dendritic calcium signals in the interval immediately prior to movement initiation and during the movement itself (**Figure 1G**). Because animals usually initiated wheel movements immediately upon object appearance, these signals may reflect either object appearance (a sensory signal) or the wheel movement (a motor signal). To distinguish these (not mutually exclusive) possibilities, we compared activity in trials where mice reacted rapidly from those in which they did not, and wheel movement within trials to those made during inter-trial intervals (**Supplementary Figure 1**). We found that while the object appearance itself could evoke responses in our recorded neurons, movement-aligned activity was similar for trials in which animals reacted quickly or slowly, and also similar for wheel movements initiated within trials and outside of trials. Furthermore, trial-by-trial analysis of population activity as a function of reaction time showed a tighter linkage between activity and movement onset than object appearance. Thus, movement onset-aligned activity is preferentially related to movement.

Many Purkinje cells also exhibited elevated calcium signals in the interval between the end of wheel movements and the reward, and at the time of the reward delivery itself (**Figure 1H**). Overall, we found that movement onset-related activity was not predictive of trial outcome (**Figure 1I**), while reward delivery on correct trials modulated our recorded populations potently (**Figure 1J**). In subsequent experiments, we explored how reward-related signals were organized in Purkinje cell populations and how they could be modulated by reward context.

Topographic organization of reward-related signals

Microzones constitute a fundamental unit of cerebellar processing and are defined by the relationship between Purkinje cells and the climbing fibers that innervate them (7, 8, 28, 30, 32). We asked whether the functional segregation of Purkinje cells activated by reward delivery maps onto microzones. To begin, instead of sorting ROIs based on response magnitude, we sorted them based on anatomy: orthogonally to the parasagittal axis of Purkinje cell dendrites. This sorting revealed groups of reward delivery-activated and reward delivery-suppressed Purkinje cells (**Figure 2A**). To classify individual Purkinje cells into microzones systematically in all our recorded FOVs, we used principal component analysis to reduce the dimensionality of each dataset and performed k-means clustering followed by a series of validation steps to identify microzonal clusters (see Methods and **Supplementary Figure 2**). The clustering results for our example FOV are plotted for the first 3 components in **Figure**

2B. These functionally defined clusters mapped onto anatomically clustered Purkinje cell populations, revealing almost perfect mediolateral segregation into microzones (**Figure 2C**, **Supplementary Figure 3**). This method yielded similar results to previously established correlation-based methods for identifying microzones, as evident from the block-diagonal correlation matrix structure of our sorted Purkinje cells (see Methods and **Supplementary Figure 3**). Fields of view (670 μm x 670 μm) contained 5.3 ± 0.3 microzones (mean ± s.e.m., n = 6 fields from 6 mice). Microzones were 170 ± 10 μm wide (~17 dendrites wide) and contained 34 ± 3 dendrites (n = 1101 dendrites, 32 microzones), consistent with reported microzonal widths on the order of 100-200 μm (15, 28, 30) (**Supplementary Figure 3**).

We used our microzonal groupings to ask how functionally related groups of Purkinje cells encoded reward-related activity in their complex spiking patterns. Most Purkinje cells within a given microzone exhibited similar patterns of reward-related activity (**Figure 2D**) and microzones segregated into two groups – those that increased their activity upon reward delivery ('reward-activated', **Figure 2D**, Clusters 3, 5, and 6; **Figure 2F**) and those that decreased their activity upon reward-delivery ('reward-suppressed', **Figure 2D**, Clusters 1, 2, and 4; **Figure 2G**). Across our 6 FOVs, we found an equal proportion of reward-activated and reward-suppressed microzones (16 of each). These reward-related groupings were not strictly related to movement onset-related activity, with both reward-activated (11 of 16) and reward-suppressed (7 of 16) microzones showing significant activation at the time of movement onset (assessed across trials for intervals 300-0 ms prior to movement onset and compared to baseline firing rates; Wilcoxon signed-rank test). We also aligned a subset of our FOVs (4 of 6) to coarser anatomical maps of the cerebellar surface. These coarse maps showed a gross level of stereotypy between animals with alternating groups of reward-activated and reward-suppressed neurons that could contain multiple functionally identified microzones (**Supplementary Figure 4**).

We next asked whether complex spikes in Purkinje cells, at the level of microzones, may encode upcoming reward predictively and whether they may signal lack of reward on incorrect trials. We found that both reward-activated and reward-suppressed microzones could exhibit elevated activity in the delay period between movement offset and reward (assessed across trials for delay intervals 0 - 200 ms and 200 - 400 ms after movement offset, and compared to interval 500 prior to movement offset; Wilcoxon signed-rank test, $p < 0.025$ with Bonferroni correction), with a higher fraction of reward-suppressed microzones showing significant modulation (**Figure 2H**). Activity during the delay period was similarly elevated on correct and incorrect trials (**Figure 2E-G**), suggesting that mice used reward delivery as the ultimate signal reflecting trial outcome. Indeed, mice that licked predictively during the delay period (3 of 6 mice) did so similarly for correct and incorrect trials. To rule out the possibility that delay period

activation was simply a reflection of this licking motor program, we correlated the level of activation in reward-predictive microzones to the level of predictive licking that each animal exhibited and found no relationship (**Supplementary Figure 5**). Thus, predictive licking cannot explain the elevated activity observed during the delay period.

Subsets of activated and suppressed microzones also exhibited elevated activity in the period after expected reward time on incorrect trials (assessed across trials for post-reward intervals 100 – 300 ms and 300 - 500 ms after expected reward time compared to delay period 500 - 0 ms prior to reward time; Wilcoxon signed rank test, $p < 0.025$ with Bonferroni correction), with a higher proportion of reward-suppressed microzones showing significant modulation (**Figure 2H**). Thus, complex spikes in Purkinje cell populations collectively encode reward-related information in our task, including putatively predictive signals, bidirectionally-modulated reactive signals, and error-like signals associated with lack of reward on incorrect trials.

Predictability modulates reward-related sensory responses in trained mice

In some experiments, mice were occasionally provided with random rewards during inter-trial intervals of the motor task to maintain their motivation. When we analyzed these experiments and compared reward-related responses within the task to those given randomly during inter-trial intervals, we noticed that random rewards triggered significantly larger responses than rewards earned during correct trials of the task (**Supplementary Figure 6**). We reasoned that this difference may reflect an expectation-dependent modulation of the reward-related sensory cue (solenoid sound), similarly to the suppression of climbing fiber responses to predicted periocular air puffs during eye-blink conditioning (33).

To test this directly, most of the mice in our study (5 of 6 mice from Figure 2) were trained to perform the motor task with interleaved random or tone-cued rewards on a subset (10% each) of inter-trial intervals (**Figure 3A**). Thus, we could compare how climbing fiber inputs to Purkinje cells convey information about random (not predictable), operant, and tone-cued (fully predictable) rewards (**Figure 3B**). Consistent with tone-cued rewards being more predictable (and carrying a greater degree of expectation) than operant rewards, all mice exhibited greater predictive licking during the delay between the tone cue and reward than during the delay between a correctly executed operant trial and reward (**Supplementary Figure 7**). Predictive lick was, by definition, not present in the random reward condition. The level of reward predictability had a clear influence on reward-related sensory responses: random reward evoked the largest signals, operant rewards evoked signals of intermediate size, and tone-cued rewards exhibited strong suppression of the sensory response typically associated with reward delivery (**Figure 3C-D**). To further validate that reward predictability exerted a suppressive effect on reward-related sensory signals, we also analyzed data from

2 mice that were trained on an easier version of our task, where all vigorous wheel movements towards the midline produced a correct trial and were rewarded (see Methods). In these mice, we found that reward responses were suppressed even more than those mice trained on our normal task (**Supplementary Figure 8**).

How does the predictability of reward alter the patterns of activity displayed by populations of Purkinje neurons? To answer this question, we computed correlations between the mean activity response vectors in each Purkinje cell over the interval 0 to 500 ms post-reward in our three reward conditions. Random and operant rewards triggered highly correlated activity patterns, confirming that similar subsets of Purkinje cell dendrites were activated in these two reward conditions. In contrast, the correlation between activity patterns recruited by either random or operant rewards with those recruited by tone-cued rewards was lower than between random and operant rewards (**Figure 3E**), demonstrating that reward predictability modulated these responses in a similar manner. To test whether representations of reward predictability varied continuously or whether they were categorically different across our reward conditions, we performed trial-by-trial analysis of the reward responses in individual neurons for trials with different amounts of predictive licking. This analysis did not show any obvious trend of greater suppression of the reward response in trials with stronger predictive licking for either tone-cued or operant rewards (**Supplementary Figure 7**). Thus, while predictive licking was categorically different across our different reward conditions, it was not sufficient to explain the differences in reward responses across different reward categories.

We also analyzed responses to random and cued rewards separately for Purkinje cells in reward-activated and reward-suppressed microzones (defined within the operant motor task; **Figure 3F**). We found that, on average, Purkinje cells from both groups were activated by the predictive tone cue and exhibited little modulation at the time of reward. In contrast, random reward delivery could activate not only those Purkinje cells that were activated by the reward cue in the operant task, but also Purkinje cells that were suppressed by operant rewards. Thus, the level of predictability exerts a bidirectional influence on reward-related activity across Purkinje cell populations, modulating responses when there is ambiguity in the outcome and remaining neutral when there is no ambiguity.

To validate that the reward-related modulation of Purkinje cell dendritic calcium signals do indeed reflect modulation in complex spiking, and not some other process (e.g. modulation of dendritic calcium signals by molecular layer interneurons (34)), we complemented our imaging experiments with direct electrophysiological recordings of complex spikes in Purkinje cells (**Figure 4A**) using Neuropixels probes. We performed these experiments in a minimal behavioral task in which we presented mice with tone-cued and random rewards (without a motor task), and processed the electrophysiological recordings using automated spike sorting

methods combined with post hoc manual curation (**Supplementary Figure 9**; see Methods). Recordings from Purkinje cells were readily identifiable by a range of criteria, including the presence of complex spikes (**Figure 4B, left**) and high-frequency simple spikes (**Figure 4B, right**), which exhibited characteristic pauses in firing after complex spikes (**Figure 4C**). Most recorded neurons (56 / 61 cells, n = 3 mice) exhibited an increase in complex spikes upon delivery of random rewards (**Figure 4D**). In agreement with our imaging experiments, the response to tone-cued rewards in these Purkinje cells was significantly suppressed (**Figure 4E-F, top**). Furthermore, the minority of Purkinje cells in our recordings that exhibited suppressed complex spike response to random rewards also were activated by the tone cues and exhibited minimal modulation at reward time when rewards were cued (**Figure 4E-F**). Thus, the results of our imaging experiments are highly consistent with those observed using direct electrophysiological recordings of complex spikes.

Modulation of reward-related responses develops with training

Reward expectation must by definition be associated with the development of trained behavior, and expectation signals should be absent in naïve mice. To test this, we analyzed recordings from the first day of training, when mice could begin to form associations between rewards and the tone-cues, wheel turns, or solenoid clicks that preceded rewards. Naïve mice learned to lick to rewards over the course of this first session, but the majority of 'within-trial' rewards on this first day were given as auto-rewards (see Methods). In naïve mice, reward delivery evoked dendritic calcium events in Purkinje cells that were similar across all conditions (**Figure 5A-C**). Interestingly, we observed the development of suppression of tone-cued rewards even within this first day of training. While random and within-trial rewards evoke similar responses, the response to tone-cued rewards was slightly reduced when averaging all tone-cued rewards (on average 10) given on this first day. However, when we analyzed only the first 3 tone-cued rewards given within each session, we saw no difference in the response when we compared them to random rewards (**Figure 5D**). Thus, the suppression of responses for predictable rewards was learned and could develop rapidly during training. To further support the idea that mice learned to associate task parameters and reward, we compared the latency to the first lick for random rewards in naïve and trained mice. The lick latency in trained mice was significantly shorter than in naïve mice, consistent with a learned association that developed with training (**Figure 5E**).

Fictive reward on operant trials triggers error signals across microzones

We next tested how omission of reward could produce an error response, similar to that observed at reward time on incorrect motor trials, and to the reward-omission response recently reported by Heffley and colleagues (15). We took advantage of the fact that the

solenoid valve-associated sensory cue that was audible at reward time represents the most immediate signal that reward would be delivered across reward conditions in our task. We introduced perturbation trials on 10% of correct trials in our motor task in which we triggered an identical solenoid valve to the one that normally delivered our reward but was not coupled to reward: i.e. we gave a fictive reward (**Figure 6A**). In 5 of our 6 mice, we recorded from FOVs that showed reward-related activity (**Figure 6B**, compared to **Figure 2C**) and measured the differences in neural activity and behavior between real and fictive reward presentation (**Figure 6C-D**). Responses on trials with real and fictive reward were similar during the pre-reward delay period and the immediate post-reward period (**Figures 6C-G**), demonstrating that mice could not distinguish between the sound of real and fictive reward. However, Purkinje cells exhibited strong activation in the later post-reward period (+100 to +200 ms) (15), presumably when mice realized the lack of reward delivery (**Figure 6C-F**). This reward-related error signal was present across our two groups of Purkinje cells (**Figure 6G**). Thus, reward-related error signals transcend microzone boundaries: both reward-activated and suppressed microzones can convey these signals.

Feedback error signals caused by omission of tone-cued reward

Modulation of reward-related activity on tone-cued rewards is drastically reduced in trained animals. Given the cerebellum's crucial role in motor timing (3, 12, 33, 35), we reasoned that mice may learn the delay interval between the cue and reward for tone-cued rewards, and wondered how violations of this expectation would be represented in the climbing fiber input to Purkinje cells. To test this directly, we introduced tone cues in our task that were not followed by rewards (**Figure 7A**). To obtain enough repetitions for each condition, we altered the likelihood of the reward event types during inter-trial intervals of our operant task, such that 30% of intervals contained a cued reward and 10% of intervals contained a cue but no reward (3:1 reward to omission ratio) and recorded from the same 5 mice subjected to fictive reward (**Figure 7B**). In two of these mice, we also recorded video of orofacial movements during these experiments (**Supplementary Figure 10**). Cued omission of reward evoked responses in many Purkinje cells (**Figure 7C-D**) at the time of expected reward (computed over 0 – 200 ms post expected reward, **Figure 7E-F**). These error signals were related to the expectation based on the tone cue, because we omitted any sensory signal at the time of the reward itself. Importantly, responses to the tone-cue were similar for rewarded and unrewarded cues (**Figure 7E-F**). We again asked if Purkinje cells defined as reward-activated and reward-suppressed in our motor task encoded this reward omission differently, and found that, as for our analysis of incorrect trials within the motor task, these error responses were expressed more strongly (but not exclusively) by neurons in reward-suppressed microzones (**Figure 7G**). We also validated that these omission responses were present in our

electrophysiological recordings of Purkinje cells by omitting cued rewards in our simple conditioning paradigm. We found neurons with significant increases in complex spike rates at expected reward time (**Figure 7H**), confirming that omission-related error signals identified in our imaging experiments are reflected in the underlying complex spike patterns of Purkinje cells.

DISCUSSION

The nature and variety of signals conveyed to Purkinje cell populations by climbing fibers has been vigorously debated. This debate has centered on whether climbing fibers carry feedback error signals or timing signals to sculpt ongoing and future actions (36). Here we show that when mice learn to associate multiple parameters – operant wheel movements, tone-cues, and solenoid clicks – with reward, this reward context is encoded in climbing fiber input to Purkinje cells. Specifically, climbing fiber signals encode parameters related to internally generated expectations, namely those relating to reward expectation, delivery, and evaluation. In this way, the cerebellum can utilize all relevant signals – be they self-generated or sensed – to make predictions about the future, evaluate these predictions, and relay them to the rest of the brain.

Microzonal organization of reward signals in Purkinje cells

Our results demonstrate that climbing fiber inputs signal reward bidirectionally, as activation and suppression, via distinct but adjacent groups of microzones (7, 32). Purkinje cells in microzones that were suppressed by reward delivery were more likely to exhibit reward-predictive activity, while Purkinje cells that exhibited reward-related sensory responses exhibited expectation-dependent modulation of these responses. However, when reward was expected but not delivered, both groups could exhibit error signals in response to this violated expectation. Notably, these error signals were strongest (i.e. strongly engaged in both reward-activated and reward-suppressed microzones) in our fictive reward condition, when mice both made the correct action and were provided with the reward-associated sensory signal, and less prominent on incorrect motor trials and when tone-cued rewards were omitted. Mice also made larger, less stereotyped orofacial movements upon omission of expected reward, presumably in search of the reward they were expecting. The generality of these error signals, which manifest both on a neural and behavioral level, imply that when expectations are violated, climbing fibers may be activated in a heterogeneous manner to destroy previously created associations, since the outcomes of these expectations were not fulfilled.

Learned, temporally specific suppression of sensory responses to predictable rewards

The degree of predictability of upcoming reward exhibited a profound influence on reward-related signals in our trained mice. The greater the likelihood of upcoming reward, the greater the suppression of responses to reward. Reward delivery elicited large climbing fiber responses in Purkinje cells when reward was delivered randomly, moderate responses when reward was delivered in a motor trial context in which success was not guaranteed, and virtually no responses when reward was cued with a fully predictive tone. These reward-

related expectations developed with training: responses to reward were similar at the very beginning of training, and suppression of predictively reward-related signals developed rapidly (within the first training session for fully predictable rewards). The mechanism of this suppression is unclear, but a potential source may be the cerebellum itself, whose output could exert either an indirect excitatory or direct inhibitory influence over the inferior olive (37). These expectation signals were also temporally specific: omission of reward on tone-cued trials evoked omission-related activation of Purkinje cells specifically at the time of expected reward.

Relationship between reward-related signals in the two input streams to Purkinje cells

Cerebellar granule cells have recently also been shown to encode reward (26), presumably driven by mossy fiber input of unknown origin. Assessing the similarity of these granule cell signals with reward-related climbing fiber signals will require a careful comparison of the reward contingencies of these signals, ideally using the same behavioral task. Specifically, a spatial organization of reward-related signals (**Figure 2A-D**) and activity suppressed by reward (**Figure 2G**) have not yet been observed in granule cells. If the granule cell and climbing fiber-mediated reward signals indeed exhibit similar behavioral contingencies, it will be interesting to examine whether these signals converge on the same Purkinje cells, as might be expected from microzonal functional organization (7). Simultaneous encoding of reward-related signals by granule cell and climbing fiber inputs to Purkinje cells parallels the acquisition of predictive signals in these inputs during delayed eye-blink conditioning (33, 38). Robust representation of reward signals in these two input pathways, which can drive plasticity mechanisms in Purkinje cells, may be crucial for the role of the cerebellar cortex in guiding learned behavior.

Relationship with reward signals elsewhere in the brain

Our data highlight the diversity of information about reward expectation and delivery provided by climbing fiber inputs to Purkinje cells. Reward-related complex spike responses are inversely scaled by reward predictability in both reward-activated and reward-suppressed microzones, consistent with temporal-difference (TD) prediction error models (39) invoked in studies of the midbrain dopaminergic system (40, 41) and for Purkinje cells during eye-blink conditioning (33). In this framework, unexpected stimuli should evoke stronger responses than predictable ones. However, in contrast to the predictions of TD models, in which neurons activated by reward delivery would be suppressed by omission of reward (and vice-versa), we observed that reward omission was signaled as an increase in the climbing fiber input in both reward-activated and reward-suppressed Purkinje cells (15).

The ramp-like increase in climbing fiber activity observed in some Purkinje cells in anticipation of reward (**Figure 2D-H**) represents a non-canonical mode of firing for climbing fibers, which typically have been reported to exhibit brief changes in firing rates locked to sensory and motor events. The mechanism of this steady activation is not clear, but it may reflect a change in excitability of olivary neurons triggered by descending inputs from the cerebellum itself (37, 42-44). These patterns of activation are similar to those of GABAergic neurons in the ventral tegmental area (41) and serotonergic neurons in the dorsal raphe nucleus (45), which progressively increase their activity in anticipation of upcoming reward.

Understanding how cerebellar circuits engage with processing of reward in other parts of the brain is an important avenue for future research. The afferent inputs to the inferior olive arise from a variety of cortical and subcortical sources (37, 46). Cerebellar outputs target the midbrain dopaminergic system (47) and can influence both premotor (48, 49) and basal ganglia (50) circuits via the thalamus. Thus, the olivo-cerebellar system may interact with the canonical reward circuitry of the brain through these reciprocal connections. Overall, our findings lend further support to the idea that the cerebellum coordinates with the rest of the brain to process a range of cognitive functions.

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding authors upon reasonable request.

CODE AVAILABILITY

The custom analysis code used in this study is available from the corresponding authors upon reasonable request.

ACKNOWLEDGEMENTS

We are grateful to P. Dayan, M. Fisek, S. Tsutsumi, C. Buetfering, B. Clark, Y. Chung, and the members of the Häusser lab for discussions and comments on the manuscript. We would like to thank N. Steinmetz for help with Neuropixels recordings, M. Pachitariu for generously providing us access to Kilosort2 prior to general distribution, and N. Smith for illustrations. This work was supported by the Wellcome Trust (M.H., PRF 201225), ERC (M.H. , AdG 695709), and EMBO (D.K., ALTF 914-2015).

AUTHOR CONTRIBUTIONS

D.K. and M.H. conceived the project and wrote the manuscript. D.K., M.B., and M.B.P. performed experiments and analysis.

COMPETING INTERESTS

The authors declare no competing financial interests.

REPORTING SUMMARY

Further information on research design is available in the Life Sciences Reporting Summary linked to this article.

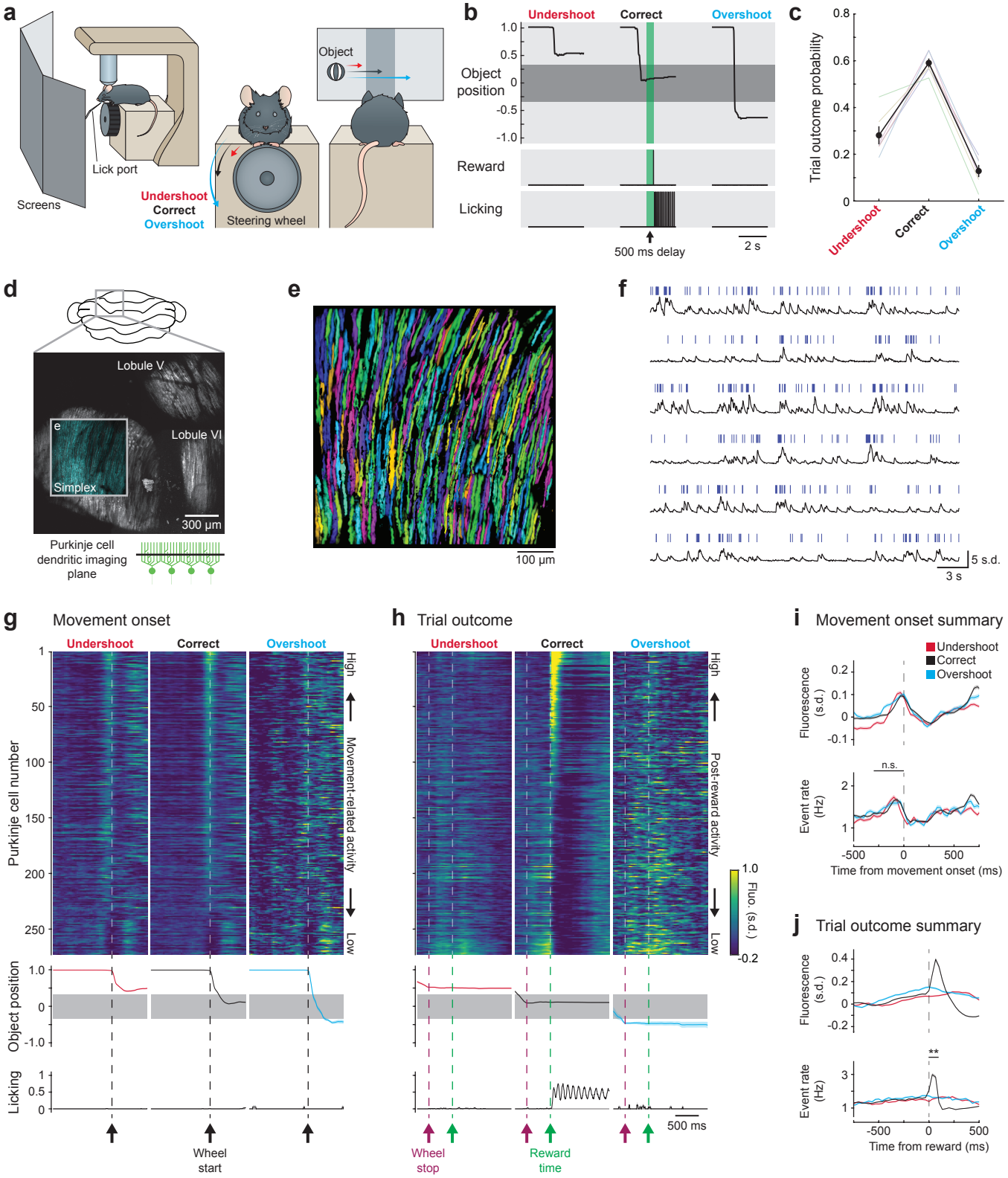


Figure 1: Population Purkinje cell complex spike imaging during a sensorimotor task

Figure 1: Population Purkinje cell complex spike imaging during a sensorimotor task

a. Behavioral setup: mice were head-fixed in front of three monitors and trained to use a steering wheel to translate a virtual object from an eccentric visual position (45° left of midline) to the midline (±15° target) to obtain a delayed reward.

b. Example behavioral trials: single undershoot (left), correct (middle) and overshoot (right) wheel trajectories, along with reward time and licking behavior.

c. Behavioral performance in well-trained mice. Colored lines represent performance of individual mice (averaged across sessions), and thick black line represents average performance across mice. Data are shown as mean ± s.e.m. (n = 6 mice).

d. GCaMP6f-labeled Purkinje cells in lobule simplex and adjacent vermis. A field of view (FOV) in a typical lobule simplex recording location is shown in cyan.

e. Extracted Purkinje cell dendritic regions of interest (ROIs) from field highlighted in **d**.

f. Six example Purkinje cell dendrite fluorescence traces (black) and extracted dendritic events (blue). The thickness of the blue line denotes event amplitude.

g. Top: trial-averaged Ca^{2+} responses in Purkinje cell population aligned to wheel movement onset for undershoot, correct and overshoot trials. Cells are sorted by the first coefficient of principal component analysis (PCA) performed over the interval ±500 ms from movement onset on correct trials. Middle: trial-averaged steering wheel position. Bottom: trial-averaged licking. Position and licking traces are shown as mean ± s.e.m. (n = 97 undershoots, 156 corrects and 18 overshoots).

h. Same as **g** but aligned to reward delivery and sorted by over the interval ±500 ms. Purkinje cell dendritic responses were sorted independently in **g** and **h**.

i. Mean time course of fluorescence responses (top) and detected events (bottom) aligned to movement onset (vertical dashed line). Mean response for statistical comparisons was computed on detected events over an interval of −300 to 0 ms from movement onset (bar above traces). Data are shown as mean ± s.e.m. (n = 1,101 neurons from 6 FOVs in 6 mice). No group was significantly different from any other group (NS; Kruskal–Wallis test, H = 4.6, d.f. = 2, P = 0.1).

j. Same as **i** but aligned to reward time (500 ms after wheel stop). Mean response for statistical comparisons was computed on detected events over the interval of 0 to +100 ms post-reward (bar above traces). Data are shown as mean ± s.e.m. (n = 1,101 neurons from 6 FOVs in 6 mice). Response on correct trials was significantly different from response on undershoot and overshoot trials (Kruskal–Wallis test, H = 22.7, d.f. = 2, $P = 1 \times 10^{-5}$, significance values for Bonferroni-corrected individual comparisons: correct versus undershoot trials, P = 0.004; correct versus overshoot trials, $P = 1 \times 10^{-5}$; undershoot versus overshoot, P = 0.5).

Statistics summary: n.s. = not significant, $^{**}P < 0.01$.

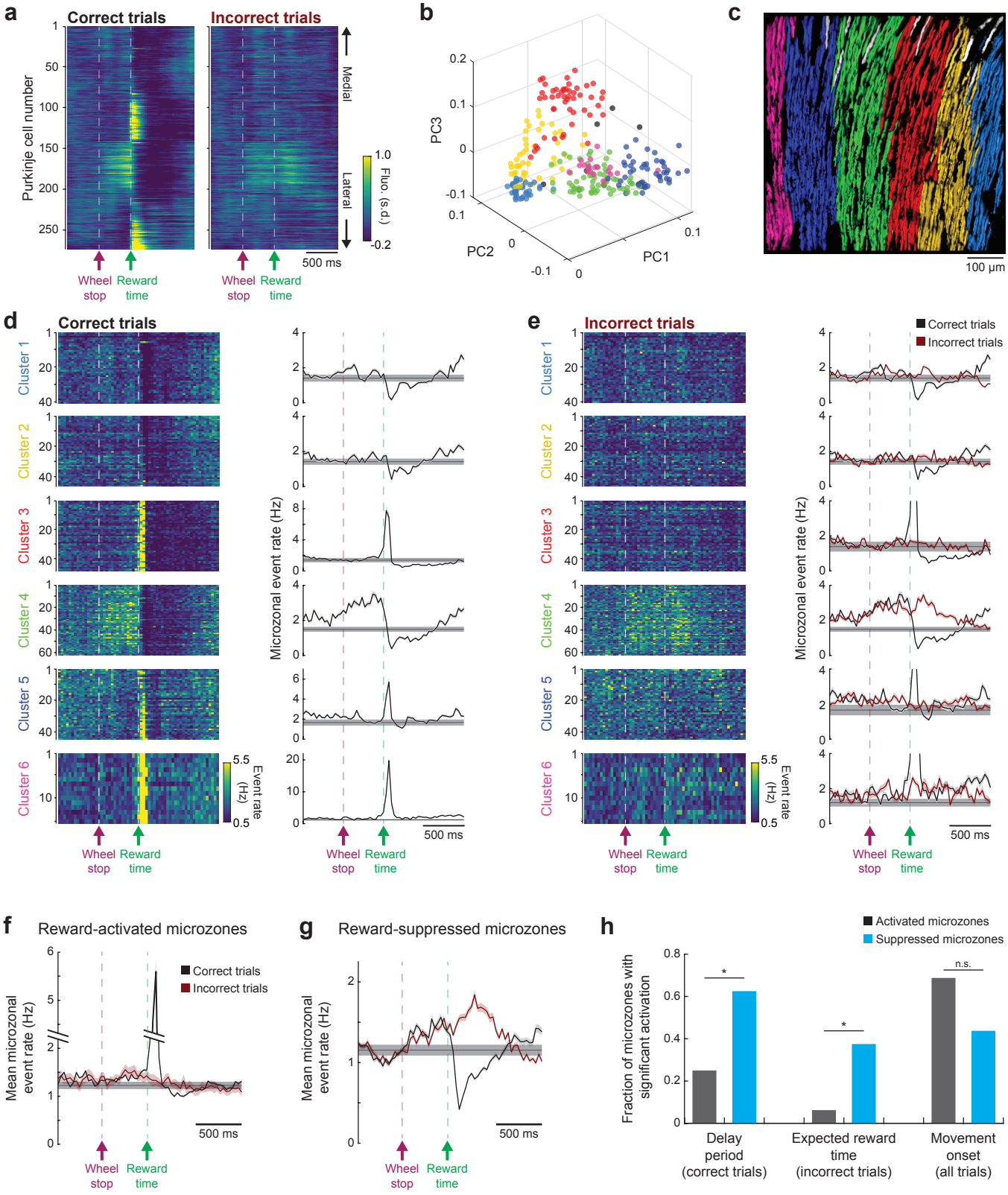


Figure 2: Reward-activated and suppressed Purkinje cells segregate to distinct cerebellar microzones

Figure 2: Reward-activated and suppressed Purkinje cells segregate to distinct cerebellar microzones

a. Trial-averaged Ca^{2+} responses in example FOV aligned to reward delivery on correct trials (left) or reward delivery time on incorrect trials (right), sorted anatomically from medial to lateral and warped by the local curvature of our recorded Purkinje cell ROIs.

b. PCA projection of *z*-scored Purkinje cell dendritic Ca^{2+} activity (spontaneous activity only) onto first three components. Individual ROIs are colored on the basis of *k*-means clustering ($k = 6$) of neuronal projections onto first six principal components ($p = 6$). Outlier ROIs shown in gray; $n = 273$ neurons.

c. Anatomical mapping of functionally identified clusters. Colors correspond to those in **b** (outlier ROIs shown in gray).

d. Trial-averaged dendritic events plotted separately for neurons within a cluster (left) and as the average microzone response (right) on correct trials. Data in right panels are shown as mean ± s.e.m. across trials. Group correspondence is denoted by color of *y* axis labels.

e. Same as **d** for incorrect trials. Note that correct trial responses are cropped to better illustrate responses on incorrect trials.

f. Time course of mean microzonal event rates on correct trials (black) and incorrect trials (red) for reward-activated microzones. Data are shown as mean ± s.e.m. across microzones ($n = 16$ microzones, 6 mice).

g. Same as **f** for reward-suppressed microzones ($n = 16$ microzones, 6 mice).

h. Fraction of reward-activated (gray) and reward-suppressed (cyan) microzones that show elevated activity during the delay period on correct trials (assessed using two-sided Wilcoxon signed-rank test with Bonferroni correction), at the time of expected reward on incorrect trials (assessed using two-sided Wilcoxon signed-rank test with Bonferroni correction) and during the movement onset period on all trials (assessed using two-sided Wilcoxon signed-rank test). Statistical significance between proportions was assessed using a Chi-squared test ($n = 16$ reward-activated and 16 reward-suppressed microzones, significance values for individual comparisons: delay period (correct trials), $P = 0.03$; expected reward time (incorrect trials), $P = 0.03$; movement onset (all trials), $P = 0.15$).

Note that in **d**-**g**, gray bars indicate mean ± 2 s.d. of baseline event rate. Statistics summary: n.s. = not significant, $^*P < 0.05$.

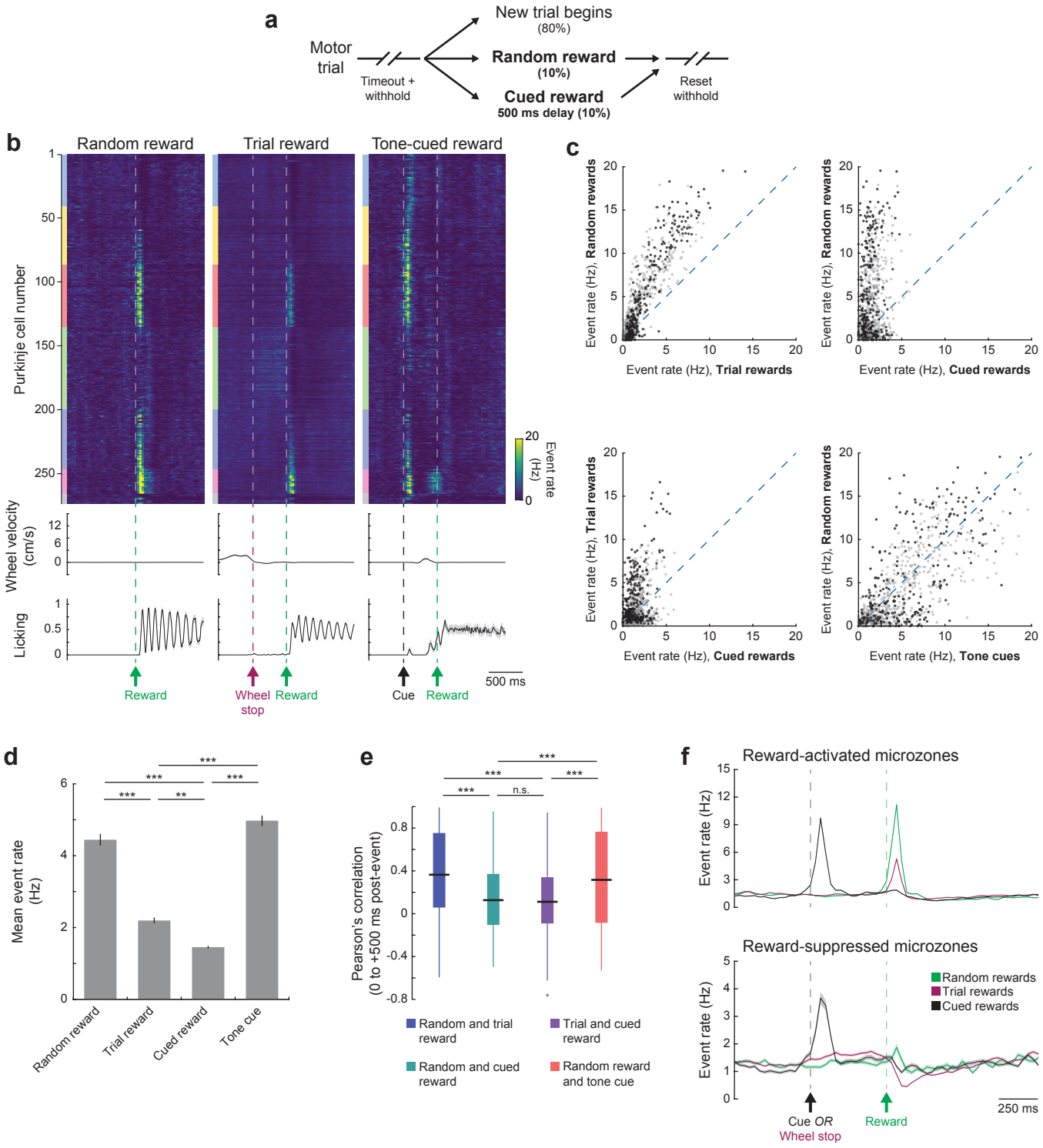


Figure 3: Predictability modulates reward responses in trained mice

Figure 3: Predictability modulates reward responses in trained mice

a. Schematic of reward perturbation experiments: during each behavioral session, we randomly interspersed random rewards (10% of inter-trial intervals) or tone-cued rewards (also 10% of inter-trial intervals; 500 ms delay between cue onset and reward).

b. Top: trial-averaged population response of a representative FOV (same as Fig. 2) to random, operant and tone-cued rewards. ROIs are sorted first by mediolateral position of identified microzones, then mediolaterally within each microzone. Color blocks adjacent to each heatmap denote microzonal designation, following the color scheme of Fig. 2 (gray, unclustered). Middle: trial-averaged steering wheel velocity. Bottom: trial-averaged licking. Velocity and licking are shown as mean ± s.e.m. across trials (n = 30 random rewards, 156 trial rewards and 30 tone-cued rewards).

c. Scatter plots showing pairwise comparisons of response amplitude (computed as mean over 0 to +100 ms after each event) across different reward conditions; $n = 891$ neurons from 5 FOVs in 5 mice. Data points from representative FOV (**b**) are shown in darker gray.

d. Cell-wise average of Purkinje cell dendritic response to each reward-related event. Data are shown as mean ± s.e.m. ($n = 891$ neurons from 5 FOVs in 5 mice, Kruskal–Wallis test, $H = 460$, d.f. = 3, $P = 2 \times 10^{-99}$, significance values for Bonferroni-corrected individual comparisons: random versus trial reward, $P = 2 \times 10^{-18}$; random versus cued reward, $P = 3 \times 10^{-33}$; trial versus cued reward, $P = 0.009$; trial reward versus tone cue, $P = 1 \times 10^{-57}$; cued reward versus tone cue, $P = 5 \times 10^{-82}$).

e. Summary of Pearson's correlations between pairs of reward-related events. Data are shown as box plots: center line, median; box edges, interquartile range; whiskers, range without outliers; gray points, outliers ($n = 891$ neurons from 5 FOVs in 5 mice, Kruskal–Wallis test, $H = 237$, d.f. = 3, $P = 5 \times 10^{-51}$, significance values for Bonferroni-corrected individual comparisons: random and trial reward versus random and cued reward, $P = 7 \times 10^{-32}$; random and trial reward versus trial and cued reward, $P = 3 \times 10^{-35}$; random and cued reward versus trial and cued reward, $P > 0.9$; random and cued reward versus random reward and tone cue, $P = 1 \times 10^{-17}$; trial and cued rewards versus random reward and tone cue, $P = 4 \times 10^{-20}$).

f. Time course of mean responses across reward conditions for Purkinje cells in reward-activated microzones (top, $n = 361$ neurons) and reward-suppressed microzones (bottom, $n = 470$ neurons). Note that 60 neurons were not clustered into a microzone and excluded from this analysis. Data are shown as mean ± s.e.m.

Statistics summary: n.s. = not significant, $^{**}P < 0.01$, $^{***}P < 0.001$.

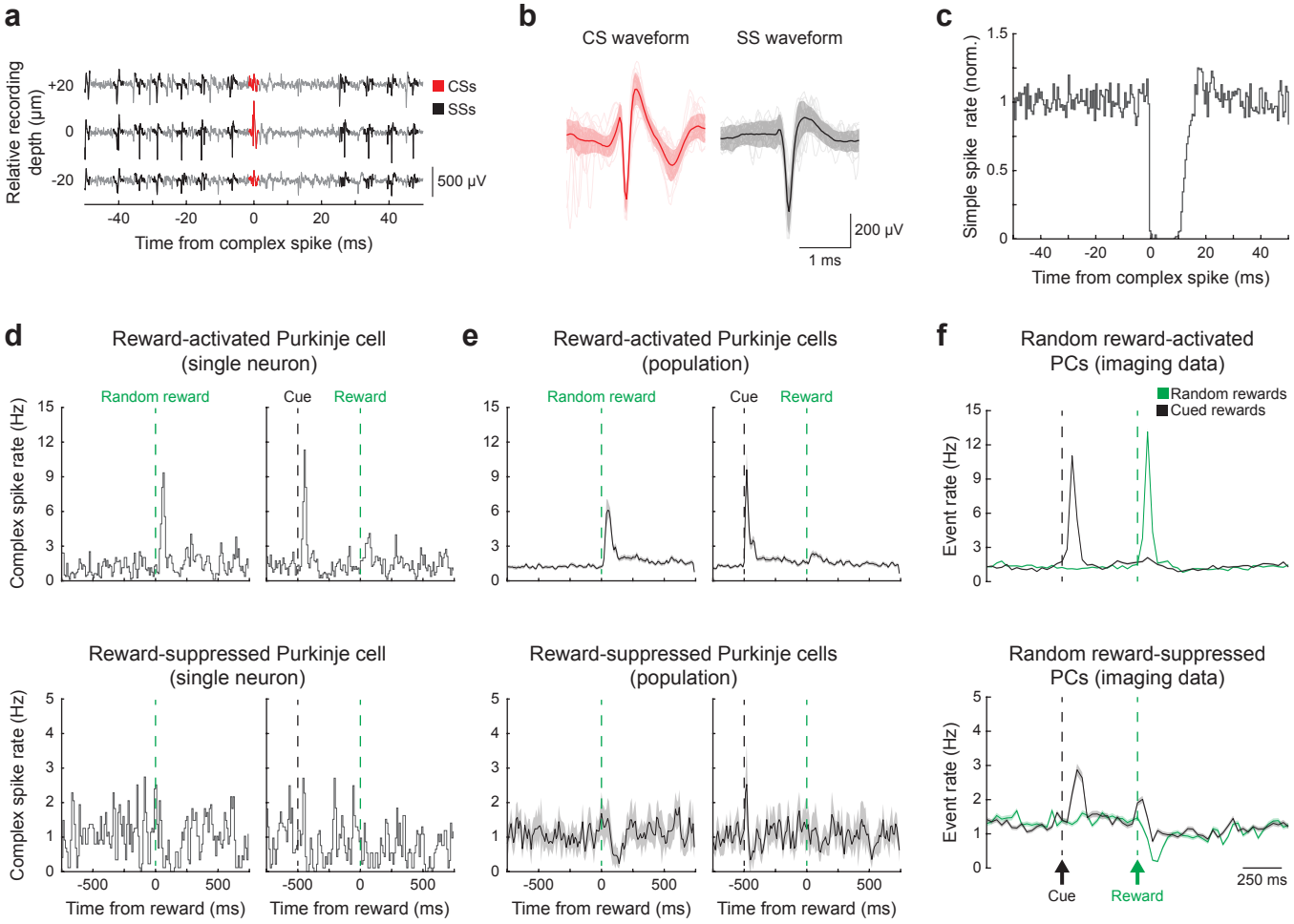


Figure 4: Electrophysiological recordings of complex spikes during cued and random reward presentation

Figure 4: Electrophysiological recordings of complex spikes during cued and random reward presentation

a. Example raw traces (gray) recorded on three adjacent vertically consecutive sites (20 μm vertical separation) of a Neuropixels probe within a Purkinje cell layer. The simple spikes and complex spikes of a single Purkinje cell are highlighted in black and red, respectively. Several other Purkinje cells were identified in this recording but are not highlighted.

b. Examples of waveforms of complex spikes (CS waveform, left) and simple spikes (SS waveform, right) recording using Neuropixels probes (same recording as in **a**). Each panel shows detected spike waveform (mean ± s.d.) and 20 overlaid raw traces.

c. Normalized histogram of simple spike firing rate (same neuron as **a** and **b**) aligned to time of complex spikes, demonstrating the characteristic post-CS pause.

d. Peristimulus time histogram (bin size = 10 ms) of complex spikes in example units that were activated (top) and suppressed (bottom) by random reward delivery on random reward trials (left) and cued reward trials (right); $n = 146$ random rewards and 154 cued rewards.

e. Same as **d** but for all recorded units that showed activation (top, $n = 56$ neurons from three mice) and suppression (bottom, $n = 5$ neurons from three mice) to random reward delivery. Data are shown as mean ± s.e.m.

f. Random and tone-cued reward responses (imaging data) in Purkinje cells (PCs) activated by random reward (top, $n = 280$ neurons, 236 of 361 from trial reward-activated microzones and 44 of 470 from trial reward-suppressed microzones) and Purkinje cells suppressed by random reward (bottom, $n = 273$ neurons, 28 of 361 from trial reward-activated microzones and 245 of 470 from trial reward-suppressed microzones). Modulation of individual Purkinje cells was assessed by comparison of response in post-reward period (33–133 ms post-reward) to pre-reward withhold period (1 s). Data are shown as mean ± s.e.m.

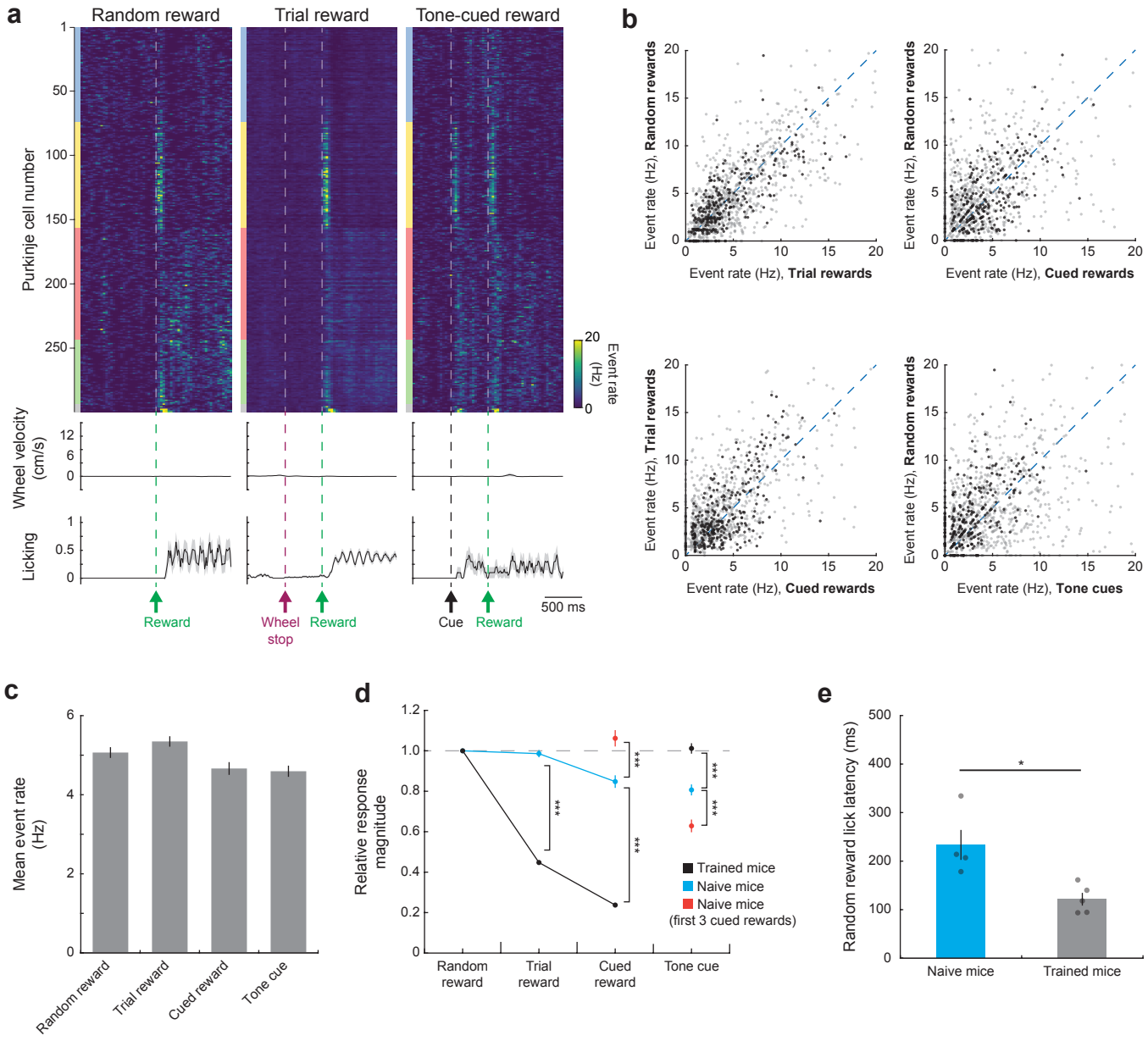


Figure 5: Modulation of reward-related responses develops with training

Figure 5: Modulation of reward-related responses develops with training

a. Top: trial-averaged population response of a representative FOV (same as Fig. 2) to random, operant and tone-cued rewards in naïve mice (first training session). ROIs are sorted first by mediolateral position of identified microzones, then mediolaterally within each identified microzone. Color blocks adjacent to each heatmap denote microzonal designation, following the color scheme of Fig. 2 (gray, unclustered). Middle: trial-averaged steering wheel velocity. Bottom: trial-averaged licking. Velocity and licking are shown as mean ± s.e.m. across trials; n = 8 random rewards, 70 trial rewards and 10 tone-cued rewards.

b. Scatter plots showing pairwise comparisons of response amplitude (computed as mean over 0 to + 100 ms after each event) across different reward conditions; $n = 1{,}187$ neurons from 5 FOVs in 5 mice. Data points from a representative FOV (**a**) are shown in darker gray.

c. Cell-wise average of Purkinje cell dendritic response to each reward-related event, pooled over the same 1,187 cells in 5 mice. Data are shown as mean ± s.e.m.

d. Relative response magnitude in neurons responsive to random reward (mean response over 0–100 ms after random reward >2 s.d. above baseline) in trained (black), naïve mice (whole first session, cyan) and naïve mice (first three trials only, red). Data are shown as mean ± s.e.m.; $n = 400$ neurons (of 891) in trained mice and $n = 710$ neurons (of 1,187) in naïve mice (Kruskal–Wallis test, $H = 1857$, d.f. = 11, $P < 1 \times 10^{-99}$, significance values for Bonferroni-corrected individual comparisons: trained versus naïve mice (trial reward), $P = 1 \times 10^{-73}$; trained versus naïve mice (cued reward), $P = 1 \times 10^{-59}$; trained versus naïve mice (tone cue), $P = 3 \times 10^{-15}$; naïve mice (all trials) versus naïve mice (first three trials) (cued reward), $P = 2 \times 10^{-91}$; naïve mice (all trials) versus naïve mice (first three trials) (tone cue), $P = 7 \times 10^{-43}$).

e. Comparison of latency to first lick in trained mice (gray) and naïve mice (cyan). For naïve mice, trials in which mice did not produce a lick to reward delivery (typically the first 5–10 rewards) were excluded ($n = 5$ trained mice and 4 naïve mice; licks were not registered for one naïve mouse). Data are shown as mean ± s.e.m., $P = 0.02$ (two-side Wilcoxon rank-sum test). Statistics summary: $*P < 0.05$, $***P < 0.001$.

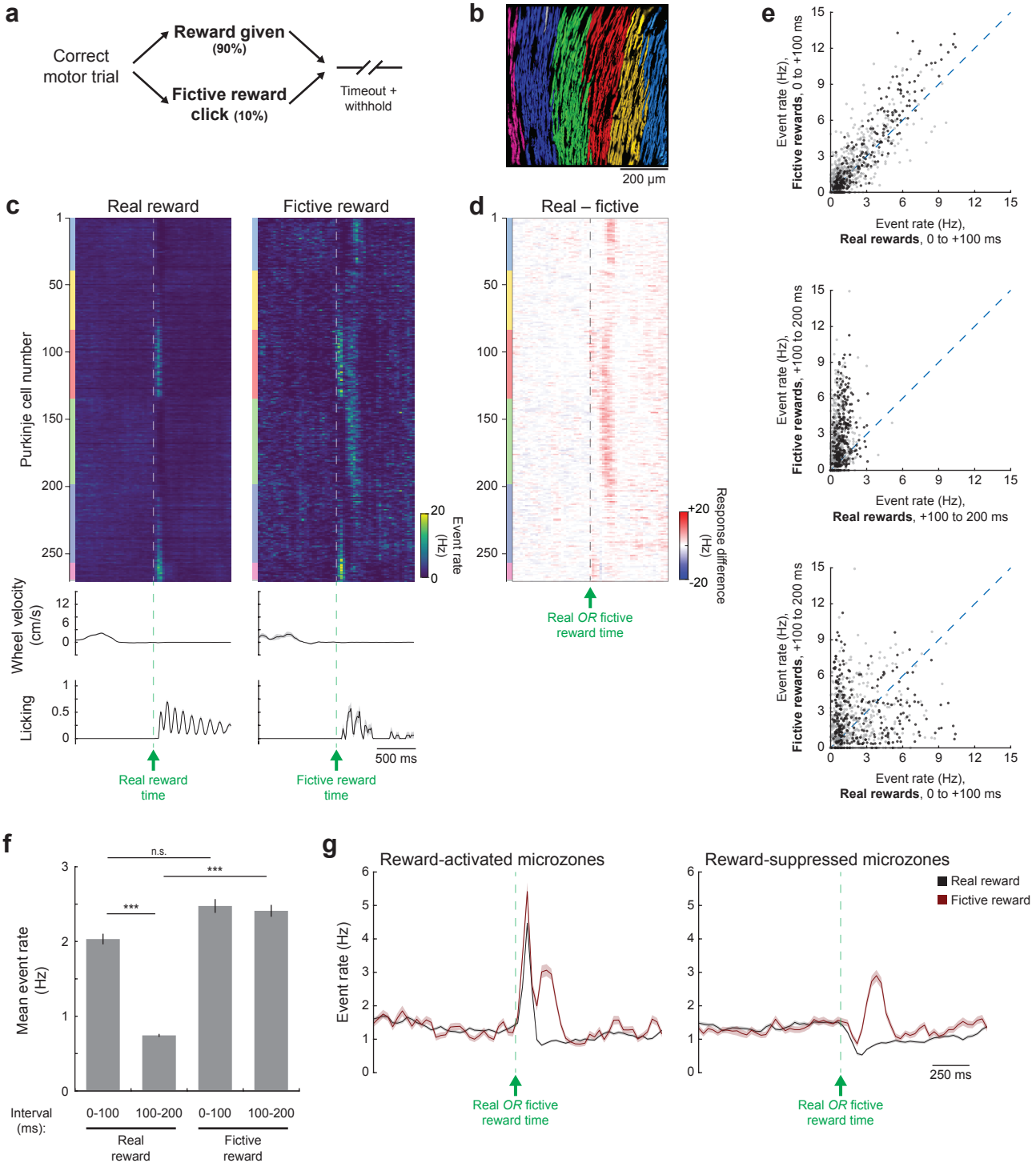


Figure 6: Fictive rewards on operant trials trigger error signals across microzones

Figure 6: Fictive rewards on operant trials trigger error signals across microzones

a. Schematic of fictive reward experiments: on 10% of correct motor trials, correct trials triggered a second solenoid that mimicked the reward sound.

b. Anatomical mapping of functionally identified microzones from an example FOV (same as Fig. 2 on different recording day). Outlier ROIs shown in gray.

c. Top: population response heatmap (trial-averaged events) of FOV from **b** to real (left) and fictive (right) rewards. ROIs are sorted first by mediolateral position of identified microzones, then mediolaterally within each identified microzone. Color blocks adjacent to each heatmap denote microzonal designation, following the color scheme in **b**. Middle: trial-averaged steering wheel velocity. Bottom: trial-averaged licking. Velocity and licking are shown as mean ± s.e.m. across trials; $n = 141$ real rewards and 16 fictive rewards.

d. Mean difference image (smoothed over three frames) comparing responses to real and fictive rewards.

e. Pairwise comparisons of reward-related responses at different time intervals after delivery of real and fictive rewards. Data pooled from 832 Purkinje cell dendritic ROIs from 5 FOVs in 5 mice (1 FOV per mouse). Data points from a representative FOV (**b**) are shown in darker gray.

f. Cell-wise average of Purkinje cell dendritic response to each reward-related event. Data are shown as mean ± s.e.m. ($n = 832$ neurons from 5 FOVs in 5 mice, Kruskal–Wallis test, $H = 333$, d.f. = 3, $P = 8 \times 10^{-72}$, significance values for Bonferroni-corrected individual comparisons: real reward (0–100 ms) versus real reward (100–200 ms), $P = 3 \times 10^{-40}$; real reward (0–100 ms) versus fictive reward (0–100 ms), $P = 0.7$; real reward (100–200 ms) versus fictive reward (100–200 ms), $P = 4 \times 10^{-56}$).

g. Time course of mean responses on real reward trials (black) and fictive reward trials (red) for Purkinje cells in reward-activated microzones (left, $n = 368$ neurons) and reward-suppressed microzones (right, $n = 405$ neurons). Note that 59 neurons were not clustered into a microzone and excluded from this analysis. Data are shown as mean ± s.e.m.

Statistics summary: n.s. = not significant, ***$P < 0.001$.

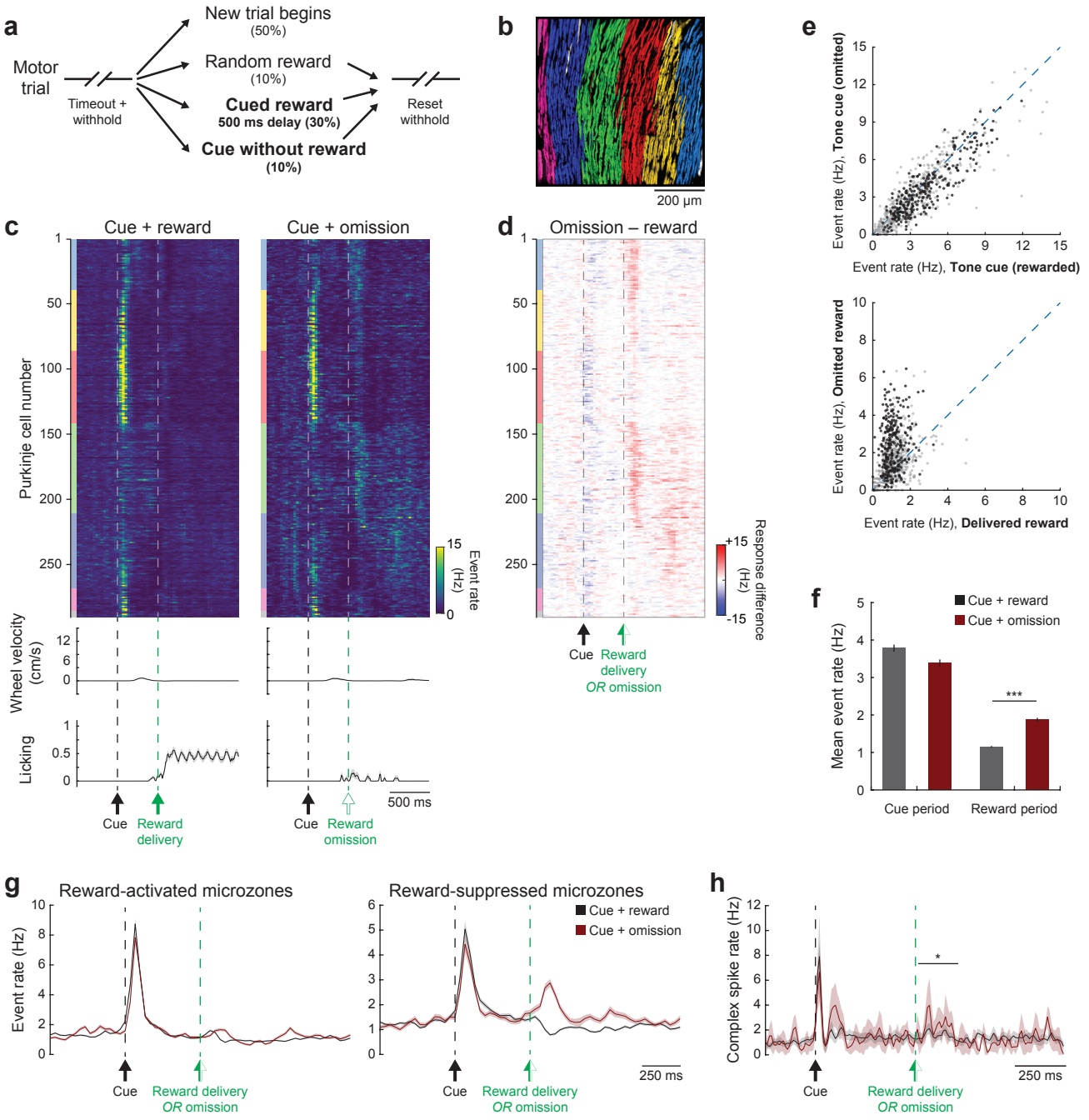


Figure 7: Omission of cued rewards triggers feedback error signals

Figure 7: Omission of cued rewards triggers feedback error signals

a. Schematic of cued reward omission: during each behavioral session, we randomly interspersed random rewards as in previous experiments (10% of inter-trial intervals), tone-cued rewards (30% of inter-trial intervals) or tone cues with reward omitted (10% of inter-trial intervals).

b. Anatomical mapping of functionally identified microzones from an example FOV (same as Fig. 2 but on a different recording day). Outlier ROIs shown in gray.

c. Top: trial-averaged population response of a representative FOV to tone-cued rewards and omissions. ROIs are sorted first by mediolateral position of identified microzones, then mediolaterally within each identified microzone. Color blocks adjacent to each heatmap denote microzonal designation. Middle: trial-averaged steering wheel velocity. Bottom: trial-averaged licking. Velocity and licking are shown as mean ± s.e.m. across trials; $n = 59$ cued rewards and 20 cued omissions.

d. Mean difference image (smoothed over three frames) comparing responses to real and fictive rewards.

e. Pairwise comparisons of reward-related responses at different time intervals after delivery of real and fictive rewards. Data pooled from 765 Purkinje cell dendritic ROIs from 4 FOVs in 4 mice (1 FOV per mouse). Data points from a representative FOV (**b**) are shown in darker gray.

f. Cell-wise average of Purkinje cell dendritic response to each reward-related event measured over interval 0–200 ms after each event. Data are shown as mean ± s.e.m. and statistical significance between cued rewards and cued omissions was assessed using the two-sided Wilcoxon signed-rank test ($n = 765$ neurons from 4 FOVs in 4 mice, $P = 2 \times 10^{-42}$).

g. Time course of mean event rates (from imaging experiments) on real reward trials (black) and fictive reward trials (red) for Purkinje cells in reward-activated microzones (left, $n = 349$ neurons) and reward-suppressed microzones (right, $n = 362$ neurons). Note that 54 neurons were not clustered into a microzone and excluded from this analysis.

h. Time course of mean complex spike rates (from electrophysiology experiments) on real reward trials (black) and fictive reward trials (red) ($n = 7$ neurons from 3 mice). Electrophysiological complex spike recordings were acquired without a motor task. Data are shown as mean ± s.e.m. ($P = 0.02$, two-sided Wilcoxon signed-rank test).

Statistics summary: n.s. = not significant, *P <0.05, ***P < 0.001.

REFERENCES

1. Wolpert DM, Miall RC, Kawato M. Internal models in the cerebellum. Trends in Cognitive Sciences. 1998;2(9):338-47.

2. Kawato M, Furukawa K, Suzuki R. A hierarchical neural-network model for control and learning of voluntary movement. Biological Cybernetics. 1987;57(3):169-85.

3. Medina JF. The multiple roles of Purkinje cells in sensori-motor calibration: to predict, teach and command. Curr Opin Neurobiol. 2011;21(4):616-22.

4. Marr D. A theory of cerebellar cortex. Journal of Physiology. 1969;202:437-70.

5. Albus JA. A theory of cerebellar function. Mathematical Biosciences. 1971;10:25-61.

6. Ito M. Cerebellar long-term depression: characterization, signal transduction, and functional roles. Physiological Reviews. 2001;81(3):1143-95.

7. Apps R, Garwicz M. Anatomical and physiological foundations of cerebellar information processing. Nat Rev Neurosci. 2005;6(4):297-311.

8. Sugihara I, Shinoda Y. Molecular, topographic, and functional organization of the cerebellar cortex: a study with combined aldolase C and olivocerebellar labeling. J Neurosci. 2004;24(40):8771-85.

9. Mathy A, Ho SS, Davie JT, Duguid IC, Clark BA, Hausser M. Encoding of oscillations by axonal bursts in inferior olive neurons. Neuron. 2009;62(3):388-99.

10. Llinas R, Baker R, Sotelo C. Electrotoniccoupling between neurons in cat inferior olive. Journal of Neurophysiology. 1981;37(3):560-71.

11. Tang T, Blenkinsop TA, Lang EJ. Complex spike synchrony dependent modulation of rat deep cerebellar nuclear activity. Elife. 2019;8.

12. Welsh JP, Lang EJ, Sugihara I, Llinas R. Dynamic organization of motor control within the olivocerebellar system. Nature. 1995;374:453-7.

13. Ozden I, Dombeck DA, Hoogland TM, Tank DW, Wang SS. Widespread state-dependent shifts in cerebellar activity in locomoting mice. PLoS One. 2012;7(8):e42650.

14. Ghosh KK, Burns LD, Cocker ED, Nimmerjahn A, Ziv Y, Gamal AE, et al. Miniaturized integration of a fluorescence microscope. Nat Methods. 2011;8(10):871-8.

15. Heffley W, Song EY, Xu Z, Taylor BN, Hughes MA, McKinney A, et al. Coordinated cerebellar climbing fiber activity signals learned sensorimotor predictions. Nat Neurosci. 2018;21(10):1431-41.

16. De Gruijl JR, Hoogland TM, De Zeeuw CI. Behavioral correlates of complex spike synchrony in cerebellar microzones. J Neurosci. 2014;34(27):8937-47.

17. Hoogland TM, De Gruijl JR, Witter L, Canto CB, De Zeeuw CI. Role of Synchronous Activation of Cerebellar Purkinje Cell Ensembles in Multi-joint Movement Control. Curr Biol. 2015;25(9):1157-65.

18. Najafi F, Giovannucci A, Wang SS, Medina JF. Coding of stimulus strength via analog calcium signals in Purkinje cell dendrites of awake mice. Elife. 2014;3:e03663.

19. Mukamel EA, Nimmerjahn A, Schnitzer MJ. Automated analysis of cellular signals from large-scale calcium imaging data. Neuron. 2009;63(6):747-60.

20. Deverett B, Koay SA, Oostland M, Wang SS. Cerebellar involvement in an evidence-accumulation decision-making task. Elife. 2018;7.

21. Medina JF, Lisberger SG. Links from complex spikes to local plasticity and motor learning in the cerebellum of awake-behaving monkeys. Nat Neurosci. 2008;11(10):1185-92.

22. Yang Y, Lisberger SG. Purkinje-cell plasticity and cerebellar motor learning are graded by complex-spike duration. Nature. 2014;510(7506):529-32.

23. Strick PL, Dum RP, Fiez JA. Cerebellum and nonmotor function. Annu Rev Neurosci. 2009;32:413-34.

24. Rochefort C, Arabo A, Andre M, Poucet B, Save E, Rondi-Reig L. Cerebellum shapes hippocampal spatial code. Science. 2011;334(6054):385-9.

25. Stoodley CJ, Schmahmann JD. Functional topography in the human cerebellum: a meta-analysis of neuroimaging studies. Neuroimage. 2009;44(2):489-501.

26. Wagner MJ, Kim TH, Savall J, Schnitzer MJ, Luo L. Cerebellar granule cells encode the expectation of reward. Nature. 2017;544(7648):96-100.

27. Lee KH, Mathews PJ, Reeves AM, Choe KY, Jami SA, Serrano RE, et al. Circuit mechanisms underlying motor memory formation in the cerebellum. Neuron. 2015;86(2):529-40.

28. Ozden I, Sullivan MR, Lee HM, Wang SS. Reliable coding emerges from coactivation of climbing fibers in microbands of cerebellar Purkinje neurons. J Neurosci. 2009;29(34):10463-73.

29. Kitamura K, Hausser M. Dendritic calcium signaling triggered by spontaneous and sensory-evoked climbing fiber input to cerebellar Purkinje cells in vivo. J Neurosci. 2011;31(30):10847-58.

30. Schultz SR, Kitamura K, Post-Uiterweer A, Krupic J, Hausser M. Spatial pattern coding of sensory information by climbing fiber-evoked calcium signals in networks of neighboring cerebellar Purkinje cells. J Neurosci. 2009;29(25):8005-15.

31. Gaffield MA, Christie JM. Conversion of graded presynaptic climbing fiber activity into graded postsynaptic Ca signals by Purkinje cell dendrites. BioRxiv. 2018.

32. Oscarsson O. Functional units of the cerebellum - sagittal zones and microzones. Trends in Neurosciences. 1979;2:143-5.

33. Ohmae S, Medina JF. Climbing fibers encode a temporal-difference prediction error during cerebellar learning in mice. Nat Neurosci. 2015;18(12):1798-803.

34. Rowan MJM, Bonnan A, Zhang K, Amat SB, Kikuchi C, Taniguchi H, et al. Graded Control of Climbing-Fiber-Mediated Plasticity and Learning by Inhibition in the Cerebellum. Neuron. 2018;99(5):999-1015 e6.

35. Ivry RB, Keele SW. Timing functions of the cerebellum. Journal of Cognitive Neuroscience. 1989;1(2):136-52.

36. Lang EJ, Apps R, Bengtsson F, Cerminara NL, De Zeeuw CI, Ebner TJ, et al. The Roles of the Olivocerebellar Pathway in Motor Learning and Motor Control. A Consensus Paper. Cerebellum. 2017;16(1):230-52.

37. Ten Brinke MM, Boele HJ, De Zeeuw CI. Conditioned climbing fiber responses in cerebellar cortex and nuclei. Neurosci Lett. 2019;688:26-36.

38. Giovannucci A, Badura A, Deverett B, Najafi F, Pereira TD, Gao Z, et al. Cerebellar granule cells acquire a widespread predictive feedback signal during motor learning. Nat Neurosci. 2017;20(5):727-34.

39. Sutton RS. Learning to predict by methods of temporal differences. Machine Learning. 1988;3:9-44.

40. Schultz W, Dayan P, Montague PR. A neural substrate of prediction and reward. Science. 1997;275:1593-9.

41. Watabe-Uchida M, Eshel N, Uchida N. Neural Circuitry of Reward Prediction Error. Annu Rev Neurosci. 2017;40:373-94.

42. Turecek J, Yuen GS, Han VZ, Zeng XH, Bayer KU, Welsh JP. NMDA receptor activation strengthens weak electrical coupling in mammalian brain. Neuron. 2014;81(6):1375-88.

43. Mathy A, Clark BA, Hausser M. Synaptically induced long-term modulation of electrical coupling in the inferior olive. Neuron. 2014;81(6):1290-6.

44. Lefler Y, Yarom Y, Uusisaari MY. Cerebellar inhibitory input to the inferior olive decreases electrical coupling and blocks subthreshold oscillations. Neuron. 2014;81(6):1389-400.

45. Miyazaki K, Miyazaki KW, Doya K. Activation of dorsal raphe serotonin neurons underlies waiting for delayed rewards. J Neurosci. 2011;31(2):469-79.

46. Garden DLF, Rinaldi A, Nolan MF. Active integration of glutamatergic input to the inferior olive generates bidirectional postsynaptic potentials. J Physiol. 2017;595(4):1239-51.

47. Carta I, Chen CH, Schott AL, Dorizan S, Khodakhah K. Cerebellar modulation of the reward circuitry and social behavior. Science. 2019;363(6424).

48. Gao Z, Davis C, Thomas AM, Economo MN, Abrego AM, Svoboda K, et al. A cortico-cerebellar loop for motor planning. Nature. 2018;563(7729):113-6.

49. Chabrol FP, Blot A, Mrsic-Flogel TD. Cerebellar contribution to preparatory activity in motor neocortex. BioRxiv. 2018.

50. Chen CH, Fremont R, Arteaga-Bracho EE, Khodakhah K. Short latency cerebellar modulation of the basal ganglia. Nat Neurosci. 2014;17(12):1767-75.

ONLINE METHODS

Animals

All animal procedures were approved by the local Animal Welfare and Ethical Review Board at University College London and performed under license from the UK Home Office in accordance with the Animals (Scientific Procedures) Act 1986. We used male Pcp2(L7)-Cre mice (line Jdhu – B6.Cg-Tg(Pcp2-Cre)3555Jdhu/J) (51) aged between 3 and 6 months. Male mice were preferred in our task because they were larger and more willing to initiate wheel movements at the beginning of training, facilitating more rapid learning in our task. Mice were group housed prior to surgery, single-housed after surgery, and maintained on a 12:12 day-night cycle. In total, data from 12 mice (9 imaging and 3 electrophysiology) were used in this study.

Headplating, virus injection, and chronic window installation

A minimum of 2 hours before surgery, mice were injected with dexamethasone to reduce swelling during surgery. A single procedure, during which mice were maintained under 1.5-2% isoflurane anesthesia, was performed on each mouse lasting approximately 2 hours to install a headplate over the cerebellar cortex, infect Purkinje cells with GCaMP6f, and install a chronic window for chronic imaging experiments. Buprenorphine (1 mg/kg, subcutaneous, Vetergesic) was administered peri-operatively for analgesia. Once mice were anesthetized, custom headplates with an oval inner opening 7 mm long and 9 mm wide were installed over the forelimb regions of the cerebellar cortex on the left side of each mouse (lobule simplex and adjacent paravermis lobules V and VI) and secured with dental cement (Super-Bond C&B, Sun-Medical). This corresponded to the posterior tip of the interparietal bone, 1.8 mm displaced from the midline (approximately 6 mm caudal and 1.8 mm lateral from bregma). Mice to be used for imaging experiments were next injected with virus and implanted with a cranial window, while mice used for electrophysiology experiments were allowed to recover at this point.

For mice used in imaging experiments, we performed a 3 mm craniotomy, centered in the middle of the headplate hole, to expose the cerebellar cortex for virus injection and window installation. We then injected Cre-dependent GCaMP6f (52) virus (AAV1.CAG.Flex.GCaMP6f.WPRE.SV40) diluted 1:12 from stock titer in 3 locations spanning paravermis and intermediate lobule simplex. At each location, ~100 nl of virus solution was pressure-injected at depths of 500, 375, and 250 μm below the cerebellar surface at 2 minute intervals. We waited ~5 minutes after the final of set of 3 injections before retracting the injection pipette. In total, ~1 μl of diluted virus was injected per mouse. Finally, a 3 mm single-paned coverslip was press-fit in to the craniotomy, sealed to the skull by a thin layer of

cyanoacrylate (VetBond) and fixed in place by dental cement. The conical portion of a nitrile rubber seal (RS Components, Stock no. 749-581) was then glued to the headplate with dental cement and filled with Kwik-Cast to protect the window preparation during recovery and between recording sessions. Mice were allowed to recover for a minimum of 7 days before beginning water restriction, during which time they were given post-operative analgesia as needed.

After mice had recovered from surgery, they were placed under water restriction for at least 5 days during which time they were acclimated to the recording setup and expression-checked. All mice were maintained at 80-85% of their initial weight over the course of recording experiments. Trained mice typically received all their water for the day from rewards during the behavioral task, while naïve mice were supplemented to 1 g water per day with Hydrogel.

On the day of electrophysiology experiments, a small craniotomy (<1mm diameter) was performed over the proximal part of lobule simplex under brief anesthesia (< 20 minutes), a nitrile rubber seal was affixed to the headplate to act as a recording chamber, and the chamber was filled with Kwik-Cast. Mice were allowed to recover for >2 hours before experiments began.

Behavior

Motor task training protocol

Mice were head-fixed in front of an array of 3 monitors with screens arranged at 135^{o} relative to each other and the central screen directly in front of the mouse (creating 3 sides of an octagon). Below their forepaws was Lego rubber tire that could be rotated left and right, and whose angle was measured using a rotary encoder coupled to the wheel's axle. We used the MATLAB-based software ViRMEn (53) to construct and operate the virtual reality (VR) environment. The rotation of the steering wheel translated the virtual object (a revolving black and white beach ball) displayed on the screens during each operant motor trial.

Mice were initially trained to translate the virtual object, which appeared in the middle of either the left or right screens (at +45^{o} or -45^{o}), towards the visual midline to receive a reward (inspired by the visual decision-making task of Burgess and colleagues (54)), at a high wheel gain (9 degrees/mm wheel rotation). On the first few days of training, the virtual object drifted towards the midline and triggered an auto-reward after a long delay (60 – 180 seconds). These auto-reward sessions were useful to allow mice to make the initial associations necessary to perform the more difficult versions of the task. The data from naïve mice shown in Figure 5 come from the first day of these auto-reward sessions. After several days (~1 week of training), mice learned to make wheel turns on their own accord to receive rewards. At this point, they were switched to a unilateral version of the task (left trials only) and increased the difficulty in

multiple steps. We decreased the gain to 6 degrees/mm and rewarded all trials in which the mice moved the object past the visual midline. This simplified task version facilitated training mice to react rapidly to object appearance and to make vigorous movements. The data shown in Supplementary Figure 8 come from this task version. We then made the task slightly harder by decreasing the gain to 2.25 degrees/mm, such that mice had to make more than one movement (typically two) to get the wheel to the target region (±15^{o} from the visual midline) and only reward trials in which the object was left unmoved in the target region for 500 ms. After mice learned to do this consistently (on >70% of trials), we analyzed the wheel movements for each mouse and identified a gain for each mouse that was most likely to produce a correct trial in a single movement – defined as one where the wheel is stopped in the target region for 100 ms. The mean gain across the mice used in this study was 3.3 degrees/mm corresponding to a 13.6 mm translation of the wheel to hit the center of the target (range 2.5-4 degrees/mm). On the first day that mice were trained on this final task version, their performance was 30-50% and plateaued at ~60% after about 1 week of training on this final task version. All recordings in 'trained' mice were performed after behavior had plateaued.

Rewards on correct trials consisted of ~3 µl of a sugar water solution (5% sucrose) and were delivered through a solenoid valve (NResearch, Part number 225PNC1-21) whose click was audible to the mouse. Reward delivery on correct motor trials was delivered 400 ms after trial evaluation (500 ms after the wheel stopped moving) and were followed by short (0-2 s) timeout, while incorrect motor trials were followed by a long (5-7 s) timeout. After completion of the timeout, a variable withhold period (1.5-2.5 s) was enforced, in which time mice were obligated to not lick or turn the steering wheel. Lick were detected using an electrical lick circuit (55). As indicated in the main text, random or tone-cued reward were administered upon the completion of these withhold periods. Tone cues for reward trials consisted of a 100 ms long, 4 kHz tone followed by 400 ms of silence before reward delivery. The timing of these cued reward was designed to mimic those of the operant motor rewards, which required the wheel to be stopped for 100 ms to trigger a reward 400 ms later (same 500 ms total delay). Random rewards were given immediately upon the completion of the withhold period. In all mice used for the analyses in this study, random and tone-cued reward were included throughout training with 10% probability of each extra reward type being given on any single inter-trial interval, except in perturbation experiments as indicated. Behavioral parameters and task-related triggers were fed back to the VR system through an Arduino and National Instruments DAQ card (NI USB-6212).

Pavlovian conditioning task

Mice used for Neuropixels electrophysiology experiments were trained on a Pavlovian conditioning paradigm consisting of an equal mixture of cued and random rewards during

training (9 training sessions). The same tone cues, timing intervals, and solenoid valves were used for these experiments as for the tone-cued and random reward imaging experiments. On the day of recording (session 10), mice were presented with 50 baseline trials of cued and random rewards (equal probability), after which 20% of rewards were randomly omitted.

Data acquisition

Two-photon calcium imaging

Imaging experiments were performed through a 16x/0.8 NA objective (Nikon) using a Sutter MOM microscope equipped with the Resonant Scan box module. A Ti:Sapphire laser tuned to 930 nm (Mai Tai, Spectra Physics) was raster scanned using a resonant scanning galvanometer (8 kHz, Cambridge Technologies) and images were collected at 512x512 pixel resolution over fields of view of 670 μm x 670 μm at 30 Hz. Sample plane power used for recordings ranged from 30-70 mW and recordings were performed midway between the pial surface and the Purkinje cell body layer, at depths of ~75 μm. The microscope was controlled using ScanImage (Version 2015, Vidrio Technologies) and tilted to ~10 degrees such that the objective was orthogonal to the surface of the brain and coverglass. Blood vessel landmarks were used to approximately find the same fields of view across imaging sessions, and fine scale adjustments were made to maximize day-to day overlap by taking short imaging movies (10 seconds) and aligning them to the previous day's recordings.

Electrophysiological recordings

Electrophysiological recordings were made using Neuropixels ("Phase 3A") electrode arrays (56) mounted on a custom 3D-printed plastic piece and affixed to a 3 axis micromanipulator with one axis tilted to be perpendicular to stereotaxic coordinates. This manipulator axis was used to lower the probe into the cerebellum at ~8 μm/s to a final depth of ~3 mm. Electrodes were allowed to settle for a minimum of 20 minutes before beginning experiments. Signals were recorded from the distal 384 channels (covering ~3.84 mm of linear distance). Recordings were made in external reference mode with gain of 500 for the action potential band (300 Hz high-pass filter) and acquired at 30 kHz using SpikeGLX software (http://billkarsh.github.io/SpikeGLX/). Electrodes were coated with a lipophilic dye (DiI) to facilitate histological identification of electrode tracks.

Video analysis of orofacial movements

Frontal video of mice on omission trials was recorded at 100 Hz using an Allied Vision Mako U-130B camera. To analyze orofacial movements, the brightness of a region of interest surrounding each mouse's mouth (~4 x 8 mm) was averaged, baseline subtracted (8^{th} percentile of a 2 second rolling average surrounding each data time point) and aligned to

behavior. Because the mice tongues appeared bright in these videos, we could use the brightness value at each time point as a proxy for tongue movements.

Anatomical mapping and histology

The anatomical maps shown in Supplementary Figure 4 were made by taking tiled z-stacks of the exposed portions of the cerebellum (in live mice) and stitching them to create a panoramic image of the cerebellar cortical surface. Imaging fields of view were manually aligned to these reference images.

For histological experiments requiring post-mortem histology, mice were deeply anesthetized with ketamine/xylazine then transcardially perfused with PBS then 4% paraformaldehyde (PFA) in PBS. Brains were removed and post-fixed overnight in 4% PFA in PBS.

Data analysis

Extraction of Purkinje cell dendritic ROIs and identification of putative complex spikes

ROIs corresponding to single Purkinje cells were extracted using a combination of Suite2p software in MATLAB (57) for initial source extraction and custom-written software to merge over-segmented dendrites. For each recorded field of view, we identified individual dendrites using the following protocol:

(1) After running initial segmentation using Suite2p, all dendritic segments corresponding to a fluorescent portion of a Purkinje cell dendrite in the mean fluorescence image were selected for further processing using Suite2p's built-in user interface.
(2) Correlations of the baseline-subtracted (8^{th} percentile of a 2 second rolling average surrounding each data time point) fluorescence traces of all selected dendritic segments were computed. Dendritic segments that did not exhibit correlations above 0.5 with any other dendritic segments were classified as unique Purkinje cell ROIs.
(3) The dendritic segments that did exhibit correlations above 0.5 with any other segment were classified in to non-redundant groups. These groups were then visualized in a custom written MATLAB graphical user interface that displayed each segment in a different color and overlaid it and its correlated partners on the mean fluorescence image. Segments that originated from the same single dendrite (i.e. had highly correlated fluorescence traces and were aligned in the axis of Purkinje cell dendrites) were merged. A weighted average of the fluorescence trace of each group of merged dendritic segments was computed based on the number of pixels in each segment.

An event detection algorithm, MLspike (58), was used to identify fast dendritic calcium transients, faithful indicators of complex spiking activity in Purkinje cells (28-30, 59-61), in each

dendritic ROI. As input to MLspike, we used baseline subtracted fluorescence traces (ΔF) to which we added the maximum value of each trace (input values near zero are problematic for MLspike). The baseline fluorescence parameter (F_0) was set as the 25th percentile of each fluorescence trace, the sampling rate (*dt*) was set at 1/30 (30 Hz), and the indicator decay parameter (tau) was set to 0.15. The output of MLspike is an event time, as well as an amplitude (an integer multiple of the unitary event size detected of each trace). Events detected in consecutive bins, which are very likely to reflect a large dendritic event corresponding to a single complex spike rather than multiple separate complex spikes at our imaging rates (30 Hz), were summed and binned in to the first time point of each sequence. Event amplitudes for each ROI were normalized by the mean amplitude of detected events for that ROI. The absolute event rate across all recorded Purkinje cell dendritic ROIs in this study was 1.4 ± 0.4 Hz (mean ± s.d., n = 2854 ROIs from 13 fields of view in 9 mice), consistent with previously reported rates of complex spiking during behavior (62). Event amplitudes were converted into rates by multiplying by the imaging frequency (30 Hz), creating a complex spike firing rate weighted by event amplitude. Treating all detected complex spike events the same (i.e. setting their magnitude equal to 1) produced very similar results (**Supplementary Figure 11**).

Synchronization of behavior and recordings

All behavioral parameters – trial onset and offset triggers, wheel translation, reward deliveries, tone cues, VR frame update times, 2-photon imaging frame times, and video frame times – were acquired simultaneously and digitized at 5 kHz using a National Instruments (NI USB-6212) and saved using PackIO software (63). Subsequent analysis was performed off-line using custom written scripts in MATLAB (Version 2017a or 2018a).

Recorded dendritic fluorescence traces and extracted events (complex spikes) were aligned to different behavioral events of interest at the first frame whose acquisition began after each event and averaged across occurrences of each of these behavioral events. Wheel movement initiation was defined as the first time on each motor trial that wheel velocity exceeded 1 mm/s. Binary licking traces, whose value was one when the mouse's tongue contacted the lickport and was zero otherwise, were averaged in their raw format in all plots and quantifications.

Identification of Purkinje cell microzones

Initial spatial sorting of Purkinje cells was by (1) fitting a line through the pixels comprising each ROI and using this line to create a local direction vector for each ROI, (2) binning these ROI vectors at a density of 32x32 pixels – creating a 16x16 grid from our 512x512x pixel images with 1 mean vector per square, (3) fitting local contour lines to this grid using MATLAB's 'streamline' function, and (4) grouping ROIs by their closest local contour line,

sorting ROIs orthogonally to this contour line, and concatenating groups closest to each contour line. ROIs were organized and indexed from medial to lateral, by convention. This analysis demonstrated clearly that parasagittal clusters of Purkinje cells exhibited uniformity in their responses to reward.

To systematically identify functional clusters of Purkinje cells (microzones) in our recordings, we devised the following analysis pipeline:

(1) We normalized our recordings by z-scoring our baseline-normalized fluorescence data matrices and performed principal component analysis on both our whole baseline data matrix ('all data') as well as just spontaneous activity obtained by concatenating the withhold periods before the start of each trial ('spontaneous only').

(2) To determine the relevant principal component subspace in which to cluster our data, we performed 1000 shuffles of our 'all data' matrix where each neuron's activity was jittered in time over the interval ± 400 ms (± 12 imaging frames). We computed a mean and standard deviation of the variance explained by the principal components of these temporally jittered data and took the first n principal components of our real data whose variance explained exceeded the mean + 2 s.d. of the shuffled data. The number of principal components with significant information varied between 4 and 7, depending on the field of view.

(3) The coefficients associated with this number of principal components (p) were used for k-means clustering of the 'all data' matrix' and 'spontaneous only' matrix. The number of clusters in this p-dimensional subspace of our data was chosen programmatically using silhouette criterion values to identify the optimal number from a range [1:12]. To optimize clustering, centroid positions were re-seeded 1000 times and the solution yielding the lowest within-cluster distances was used for further analysis.

(4) Identified clusters were mapped on to anatomy and all further analysis was performed using the 'spontaneous only' matrix to align with the original conception of microzones, but results from the 'all data' matrix were used as comparison.

(5) The following criteria were applied sequentially to refine identified microzones:
    a. Clusters with fewer than 5 members were merged with their closest neighbor.
    b. Cluster with clear multipeaked spatial distributions in the mediolateral axis were split into separate clusters. To identify multipeaked distributions, the mediolateral coordinates of ROI centroids were binned at ~80 μm/bin (64 pixels/bin) and normalized to the peak bin. Secondary peaks were defined as those containing counts greater than 40% of the largest bin of the histogram.

c. ROIs that were spatial outliers along the mediolateral axis of a given cluster were excluded from further analysis. These outliers were defined as having a mediolateral centroid position greater than three scaled median absolute deviations from the median mediolateral centroid of the cluster.

(6) We sorted Purkinje cells within each cluster's spatial as described above and also sorted microzones relative to each other based on median ROI position. Thus, we sorted ROIs within each microzone based on their mediolateral position, and also sorted microzones relative to each other based on their mediolateral position.

Electrophysiological analysis

Data from Neuropixels recordings were automatically spike sorted with Kilosort 2 (https://github.com/MouseLand/Kilosort2) (64) and manually curated using the 'Phy' GUI (https://github.com/kwikteam/phy). Given the foliation of the cerebellar cortex, recordings typically yielded multiple crossings of the Purkinje cell layer, and we were usually able to isolate 4-5 Purkinje cell units from each layer. Purkinje cells were identified by their characteristic electrophysiological signature (65, 66), including the presence of complex spikes and simple spikes. The rate of complex spikes was 1.4 ± 0.5 Hz (mean ± s.d., n = 61 units, 3 recordings from 3 mice). As shown in **Supplementary Figure 9**, complex spikes exhibited either a narrow waveform followed by spikelets if the recording site was perisomatic, or a broader waveform when the recording site was dendritic and in the molecular layer (66). All recording sites were confirmed by post-hoc histology, in which recording tracks (labelled with DiI coating the recording electrode) were identified in 100 µm coronal cerebellar sections in brains fixed after recording and counterstained using Neurotrace 435/455 (**Supplementary Figure 9a**). Recording tracks were aligned to the Allen Mouse Common Coordinate Framework (CCF; (67, 68)) using 'Allen CCF tools', a custom GUI for 3D alignment of electrode tracks to histology (69). Spike sorting analysis and complex spike identification were performed with the experimenter blind to task conditions. After these sorting procedures, units were aligned to behavior and grouped into reward-activated and reward-suppressed categories based on responses to random reward. Units activated at reward omission were identified by inspection.

Statistical analysis

No statistical methods were used to pre-determine sample sizes, but our sample sizes are similar to those reported in previous publications (15, 20, 26). No randomization of experimental subjects was necessary as all mice were trained and recorded under the same conditions. Behavioral events within each training session were randomized on a trial-by-trial basis within the temporal ranges and incidence rates described in the text. Data collection and

analysis were not performed blind to the conditions of the experiment, but analysis relied on code that was standardized for all experimental conditions.

Categorical comparisons between proportions were made using the Chi-squared test. Data distributions were not assumed to be normally distributed and all statistical comparisons between groups of continuous variables were performed using non-parametric tests – the Wilcoxon rank-sum test and sign test were used to study differences between two groups of unpaired and paired data, respectively, and the Kruskal-Wallis test was used when more than 2 groups were compared. Bonferroni correction was applied for multiple comparisons. In general, 95% confidence intervals ($p < 0.05$) were used to define statistical significance.

METHODS-ONLY REFERENCES

15. Heffley W, Song EY, Xu Z, Taylor BN, Hughes MA, McKinney A, et al. Coordinated cerebellar climbing fiber activity signals learned sensorimotor predictions. Nat Neurosci. 2018;21(10):1431-41.

20. Deverett B, Koay SA, Oostland M, Wang SS. Cerebellar involvement in an evidence-accumulation decision-making task. Elife. 2018;7.

26. Wagner MJ, Kim TH, Savall J, Schnitzer MJ, Luo L. Cerebellar granule cells encode the expectation of reward. Nature. 2017;544(7648):96-100.

28. Ozden I, Sullivan MR, Lee HM, Wang SS. Reliable coding emerges from coactivation of climbing fibers in microbands of cerebellar Purkinje neurons. J Neurosci. 2009;29(34):10463-73.

29. Kitamura K, Hausser M. Dendritic calcium signaling triggered by spontaneous and sensory-evoked climbing fiber input to cerebellar Purkinje cells in vivo. J Neurosci. 2011;31(30):10847-58.

30. Schultz SR, Kitamura K, Post-Uiterweer A, Krupic J, Hausser M. Spatial pattern coding of sensory information by climbing fiber-evoked calcium signals in networks of neighboring cerebellar Purkinje cells. J Neurosci. 2009;29(25):8005-15.

51. Zhang XM, Ng AH, Tanner JA, Wu WT, Copeland NG, Jenkins NA, et al. Highly restricted expression of Cre recombinase in cerebellar Purkinje cells. Genesis. 2004;40(1):45-51.

52. Chen TW, Wardill TJ, Sun Y, Pulver SR, Renninger SL, Baohan A, et al. Ultrasensitive fluorescent proteins for imaging neuronal activity. Nature. 2013;499(7458):295-300.

53. Aronov D, Tank DW. Engagement of neural circuits underlying 2D spatial navigation in a rodent virtual reality system. Neuron. 2014;84(2):442-56.

54. Burgess CP, Lak A, Steinmetz NA, Zatka-Haas P, Bai Reddy C, Jacobs EAK, et al. High-Yield Methods for Accurate Two-Alternative Visual Psychophysics in Head-Fixed Mice. Cell Rep. 2017;20(10):2513-24.

55. Slotnick B. A simple 2-transistor touch or lick detector circuit. J Exp Anal Behav. 2009;91(2):253-5.

56. Jun JJ, Steinmetz NA, Siegle JH, Denman DJ, Bauza M, Barbarits B, et al. Fully integrated silicon probes for high-density recording of neural activity. Nature. 2017;551(7679):232-6.

57. Pachitariu M, Stringer C, Dipoppa M, Schröder S, Rossi LF, Dalgleish H, et al. Suite2p: beyond 10,000 neurons with standard two-photon microscopy. bioRxiv. 2017.

58. Deneux T, Kaszas A, Szalay G, Katona G, Lakner T, Grinvald A, et al. Accurate spike estimation from noisy calcium signals for ultrafast three-dimensional imaging of large neuronal populations in vivo. Nat Commun. 2016;7:12190.

59. Ozden I, Lee HM, Sullivan MR, Wang SS. Identification and clustering of event patterns from in vivo multiphoton optical recordings of neuronal ensembles. J Neurophysiol. 2008;100(1):495-503.

60. Tsutsumi S, Yamazaki M, Miyazaki T, Watanabe M, Sakimura K, Kano M, et al. Structure-function relationships between aldolase C/zebrin II expression and complex spike synchrony in the cerebellum. J Neurosci. 2015;35(2):843-52.

61. Ramirez JE, Stell BM. Calcium Imaging Reveals Coordinated Simple Spike Pauses in Populations of Cerebellar Purkinje Cells. Cell Rep. 2016;17(12):3125-32.

62. Streng ML, Popa LS, Ebner TJ. Climbing Fibers Control Purkinje Cell Representations of Behavior. J Neurosci. 2017;37(8):1997-2009.

63. Watson BO, Yuste R, Packer AM. PackIO and EphysViewer: software tools for acquisition and analysis of neuroscience data. bioRxiv. 2016.

64. Pachitariu M, Steinmetz NA, Kadir SN, Carandini M, Harris KD. Fast and accurate spike sorting of high-channel count probes with Kilosort. Advances in Neural Information Processing Systems. 2016:4448-56.

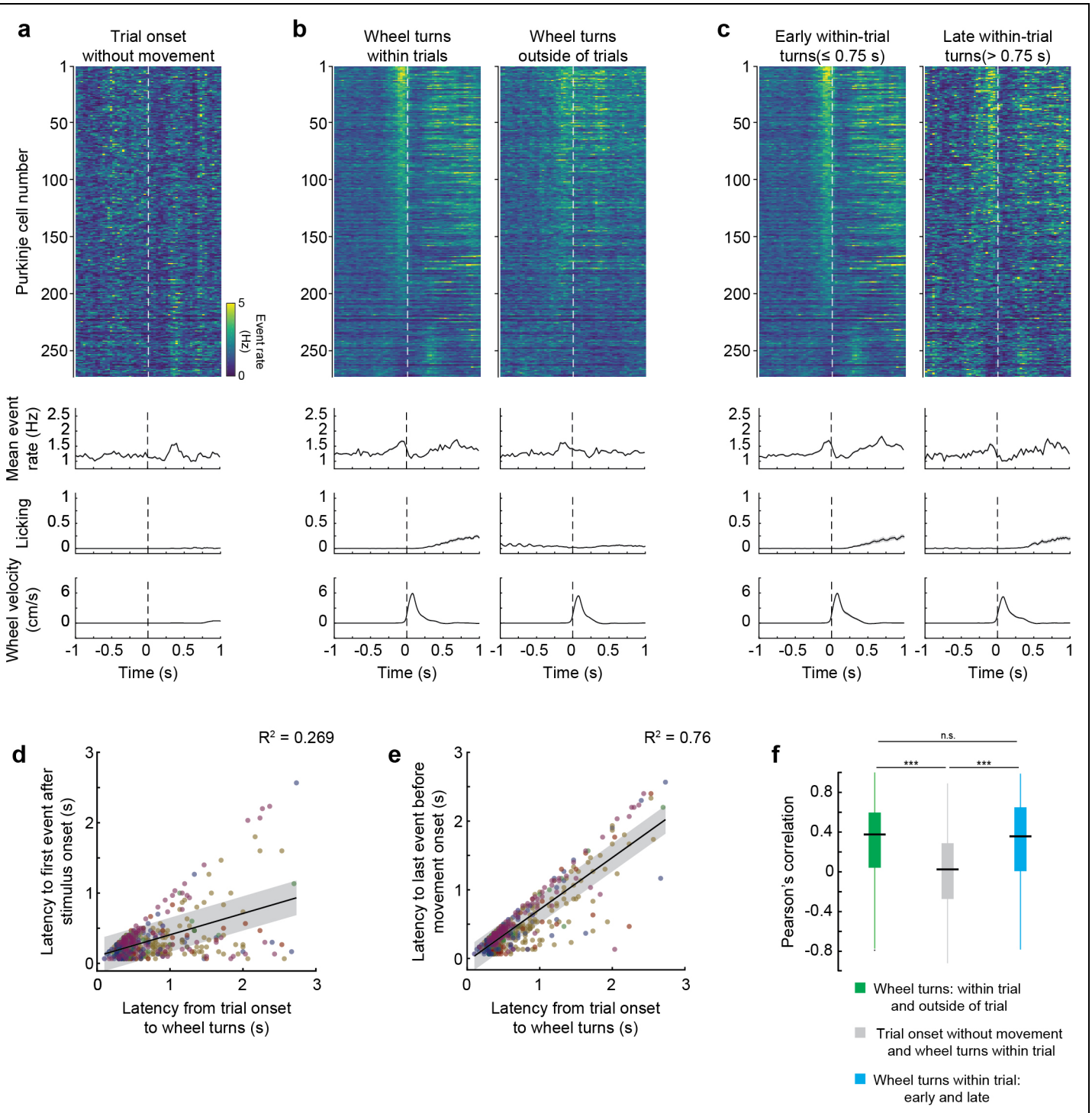
65. Armstrong DM, Rawson JA. Activity patterns of cerebellar cortical neurones and climbing fibre afferents in the awake cat. Journal of Physiology. 1979;289:425-48.

66. Gao H, Solages C, Lena C. Tetrode recordings in the cerebellar cortex. J Physiol Paris. 2012;106(3-4):128-36.

67. Dong H-W. The Allen reference atlas: A digital color brain atlas of the C57Bl/6J male mouse. Hoboken, NJ: John Wiley & Sons Inc; 2008.

68. Oh SW, Harris JA, Ng L, Winslow B, Cain N, Mihalas S, et al. A mesoscale connectome of the mouse brain. Nature. 2014;508(7495):207-14.

69. Shamash P, Carandini M, Harris K, Steinmetz N. A tool for analyzing electrode tracks from slice histology. bioRxiv. 2018.



Supplementary Figure 1

Activity aligned to movement onset is preferentially related to movement and not to object appearance

a. Top: Events extracted from population two-photon calcium imaging of Purkinje cells during the behavioral task, expressed as a heatmap. Activity in an example session is aligned to object appearance (dashed line) on trials in which the mouse did not immediately

turn the wheel (turn latency > 0.75 s). Bottom: summary of activity and behavior for all sessions. N = 1101 neurons, 6 mice, 6 sessions (1 session per mouse). Data are shown as mean ± s.e.m.

b. Top: Heatmap of activity in an example session aligned to wheel turns (dashed line) that occurred within a trial (left) or outside of a trial (right). Bottom: summary of activity and behavior for all sessions. N = 1101 neurons, 6 mice, 6 sessions (1 session per mouse). Data are shown as mean ± s.e.m.

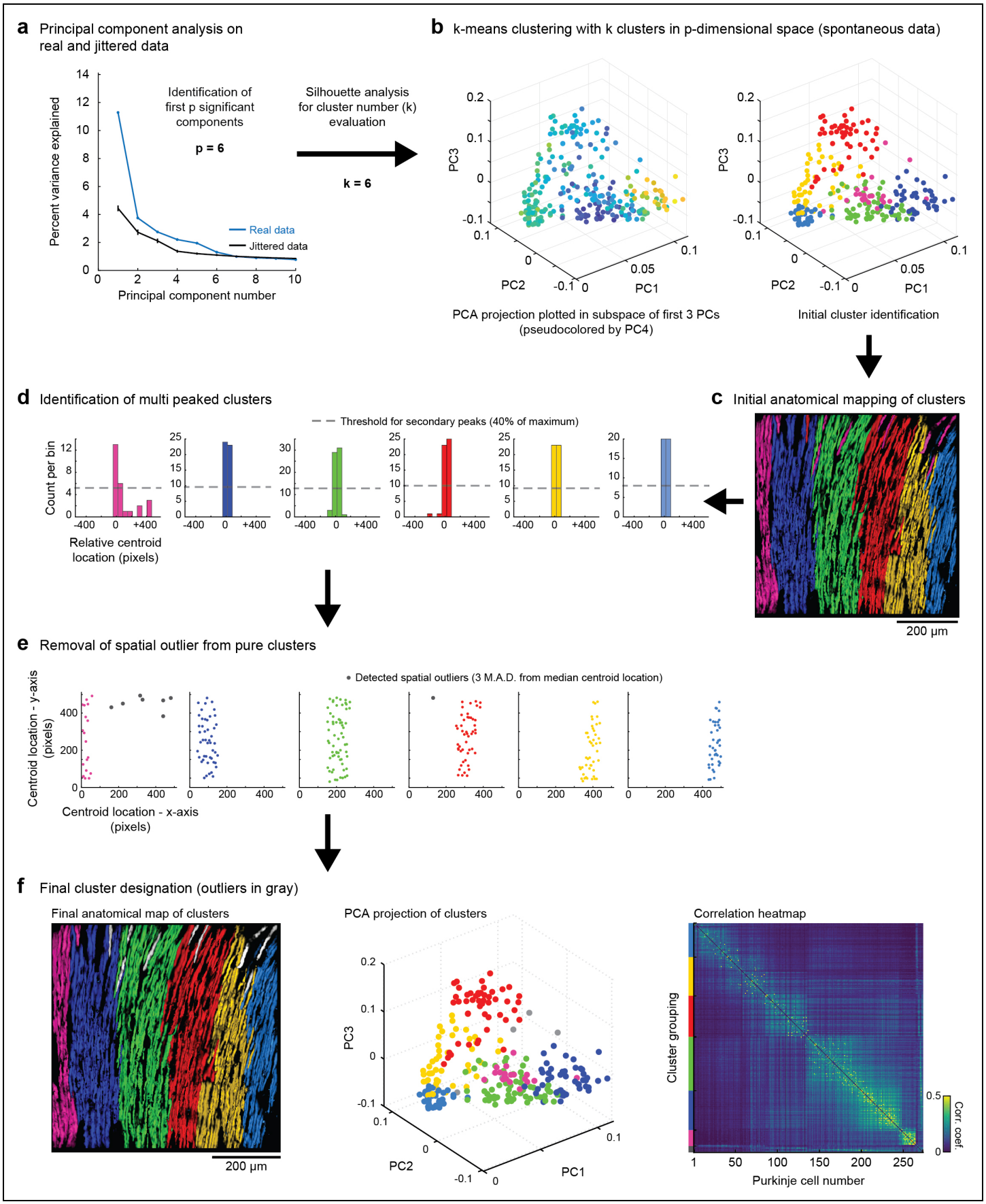
c. Top: Heatmap of activity in an example session aligned to wheel turns (dashed line) that occurred early within a trial (turn latency ≤ 0.75 s) or late in a trial (turn latency > 0.75 s). Bottom: summary of activity and behavior for all sessions. N = 859 neurons, 5 mice, 5 sessions (1 session per mouse). Note that one session was excluded from these analyses because it only had 4 trials in which it initiated a movement with turn latency > 0.75 s. Data are shown as mean ± s.e.m.

d. Comparison of first significant population-wide event after object appearance (defined as the time points when the mean event rate for a field of view exceeded the mean + 2 s.d. of activity during the withhold period on each trial) to the reaction time of the mouse on each trial. Dots represent individual trials and dots of the same color represent individual sessions (one session per mouse). Black line and shaded bar represent linear fit and 95% confidence interval through all data points (linear regression, $p = 3x10^{-42}$). N = 594 trials, 6 mice, 6 sessions (1 session per mouse).

e. Same as panel **d** but for the last population-wide event before movement onset (linear regression, $p = 8x10^{-190}$).

f. Summary of Pearson's correlation coefficients of activity traces for individual neurons between wheel turns within trial and outside of trial (green bar, interval -300 to +300 ms relative to wheel movement, N = 1101 neurons, 6 mice, 6 sessions), between trial onsets without movement and wheel turns within a trial (gray bar, interval +100 to 700 ms for trial onset and -300 to +300 ms for wheel turns in order to align peaks of activity, N = 1101 neurons, 6 mice, 6 sessions), and between wheel turns that occurred early within a trial and movements that occurred late within a trial (blue bar, interval -300 to +300 ms relative to wheel movement, N = 859 neurons, 5 mice, 5 sessions). Intervals for correlation analysis were chosen to align the peaks of the responses for each condition. Data are shown as box plots: center line, median; box edges, interquartile ranges; whiskers, range without outliers (1.5 times the interquartile range from box edges); black points, outliers (Kruskal-Wallis test, H = 359, d.f. = 2, $p = 9x10^{-79}$, significance values for Bonferroni-corrected individual comparisons: wheel turns: within trial and outside of trial vs trial onset without movement and wheel turns within trial, $p = 7x10^{-63}$; wheel turns: within trial and outside of trial vs wheel turns within trial: early and late, $p > 0.9$; trial onset without movement and wheel turns within trial vs wheel turns within trial: early and late, $p = 1x10^{-54}$).

Statistics summary: n.s. = not significant, ***$p < 0.001$.


a Principal component analysis on real and jittered data
Percent variance explained
Principal component number
Identification of first p significant components
p = 6
Silhouette analysis for cluster number (k) evaluation
k = 6
Real data
Jittered data
b k-means clustering with k clusters in p-dimensional space (spontaneous data)
PC1
PC2
PC3
PCA projection plotted in subspace of first 3 PCs (pseudocolored by PC4)
Initial cluster identification
c Initial anatomical mapping of clusters
200 µm
d Identification of multi peaked clusters
Threshold for secondary peaks (40% of maximum)
Count per bin
Relative centroid location (pixels)
e Removal of spatial outlier from pure clusters
Detected spatial outliers (3 M.A.D. from median centroid location)
Centroid location - y-axis (pixels)
Centroid location - x-axis (pixels)
f Final cluster designation (outliers in gray)
Final anatomical map of clusters
200 µm
PCA projection of clusters
Correlation heatmap
Cluster grouping
Purkinje cell number
Corr. coef.

Supplementary Figure 2

PCA-based clustering reveals correlated groups of Purkinje cell dendrites

Workflow for microzone identification is shown:

a. Significant principal components were identified by performing 1000 iterations of PCA on cell-wise jittered data (interval ± 400 ms, whole recording) and choosing first p principal components that explained significantly more variance than the jittered components.

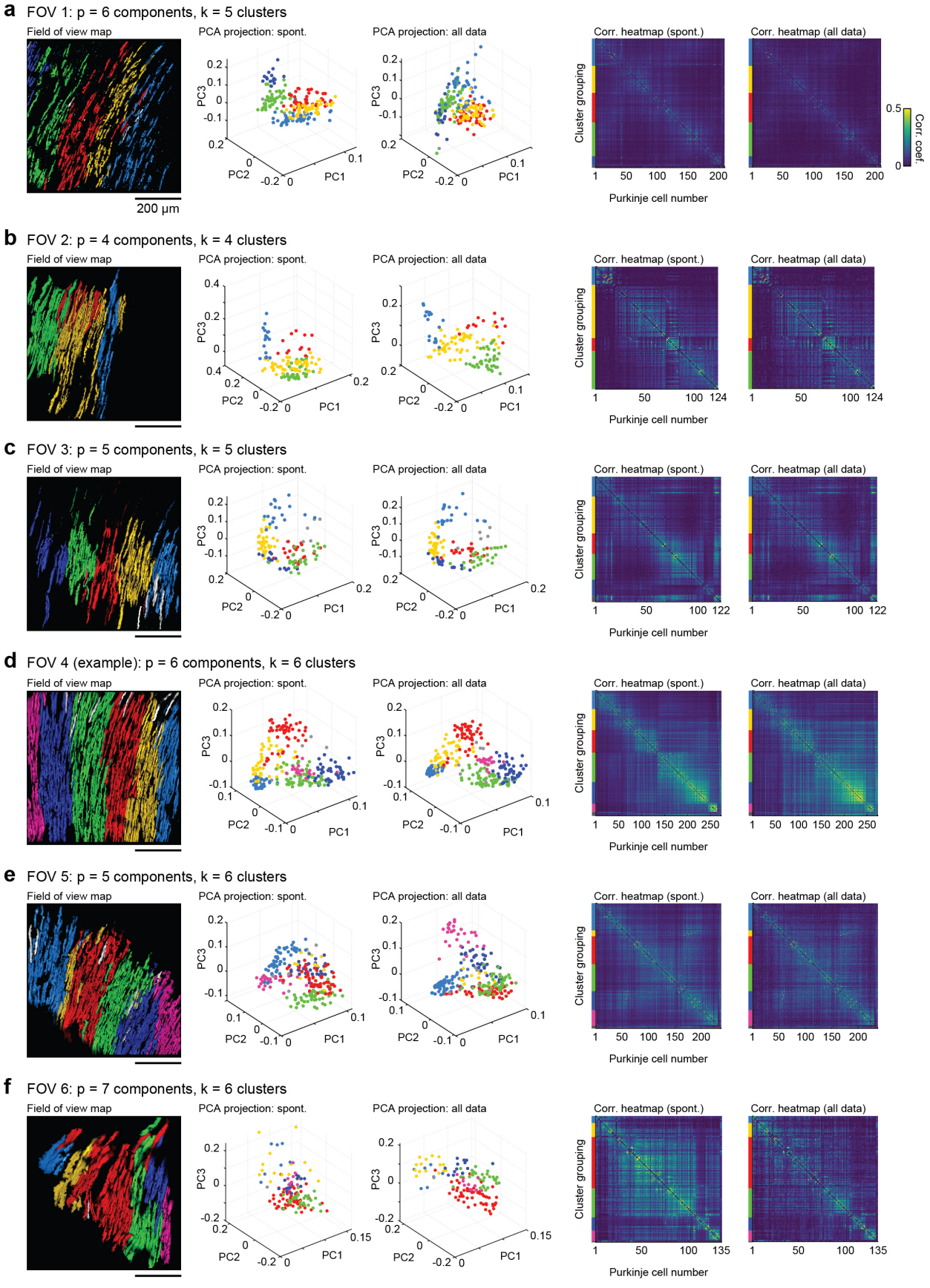
b. The appropriate number of clusters for each data set (spontaneous data only) were chosen from the interval [1:12] based on silhouette analysis. PCA projection in 3 dimensions shown before clustering (left, pseudocolored by coefficients of 4th component) and after initial clustering (right, colored by cluster).

c. The initial clusters for each recording were mapped on anatomy and exclusion criteria were applied to obtain pure clusters.

d. ROI centroids were projected on to the mediolateral anatomical axis and binned at ~80 um (64 pixels). Secondary peaks were identified with a threshold of 40% of the maximum bin. If secondary peaks were found, individual ROIs were assigned to the closest peak (none found in example dataset).

e. The median position for each cluster was computed and ROIs that were more than 3 median absolute deviations away from median were excluded as outliers.

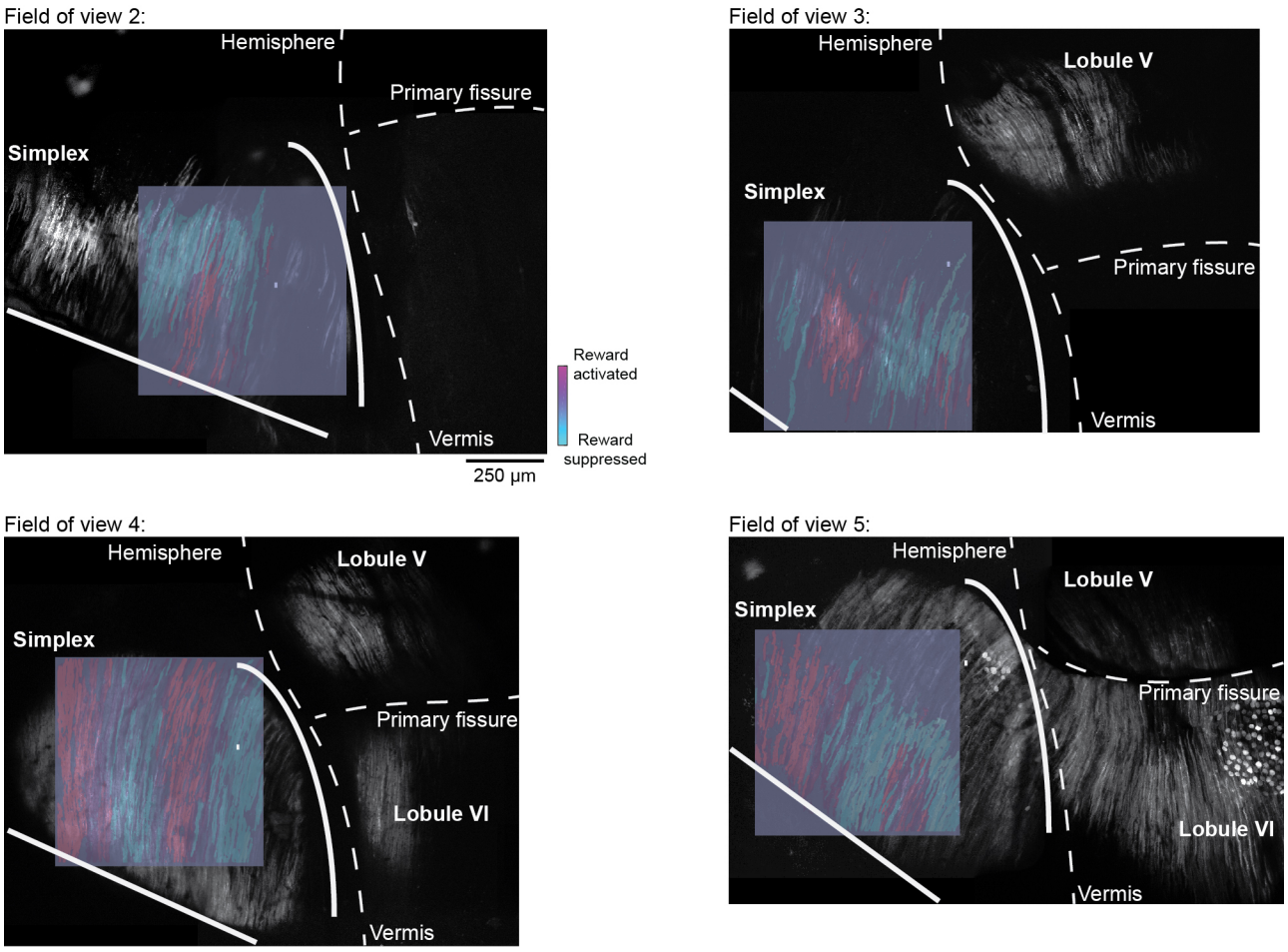
f. Final cluster designations were assigned, ROIs were organized spatially within each cluster and clusters were organized spatially relative to each other.


a
FOV 1: p = 6 components, k = 5 clusters
Field of view map
200 µm
PCA projection: spont.
PCA projection: all data
PC1
PC2
PC3
Corr. heatmap (spont.)
Corr. heatmap (all data)
Cluster grouping
Purkinje cell number
Corr. coef.
b
FOV 2: p = 4 components, k = 4 clusters
c
FOV 3: p = 5 components, k = 5 clusters
d
FOV 4 (example): p = 6 components, k = 6 clusters
e
FOV 5: p = 5 components, k = 6 clusters
f
FOV 6: p = 7 components, k = 6 clusters

Supplementary Figure 3

Microzonal identification for each recorded dataset

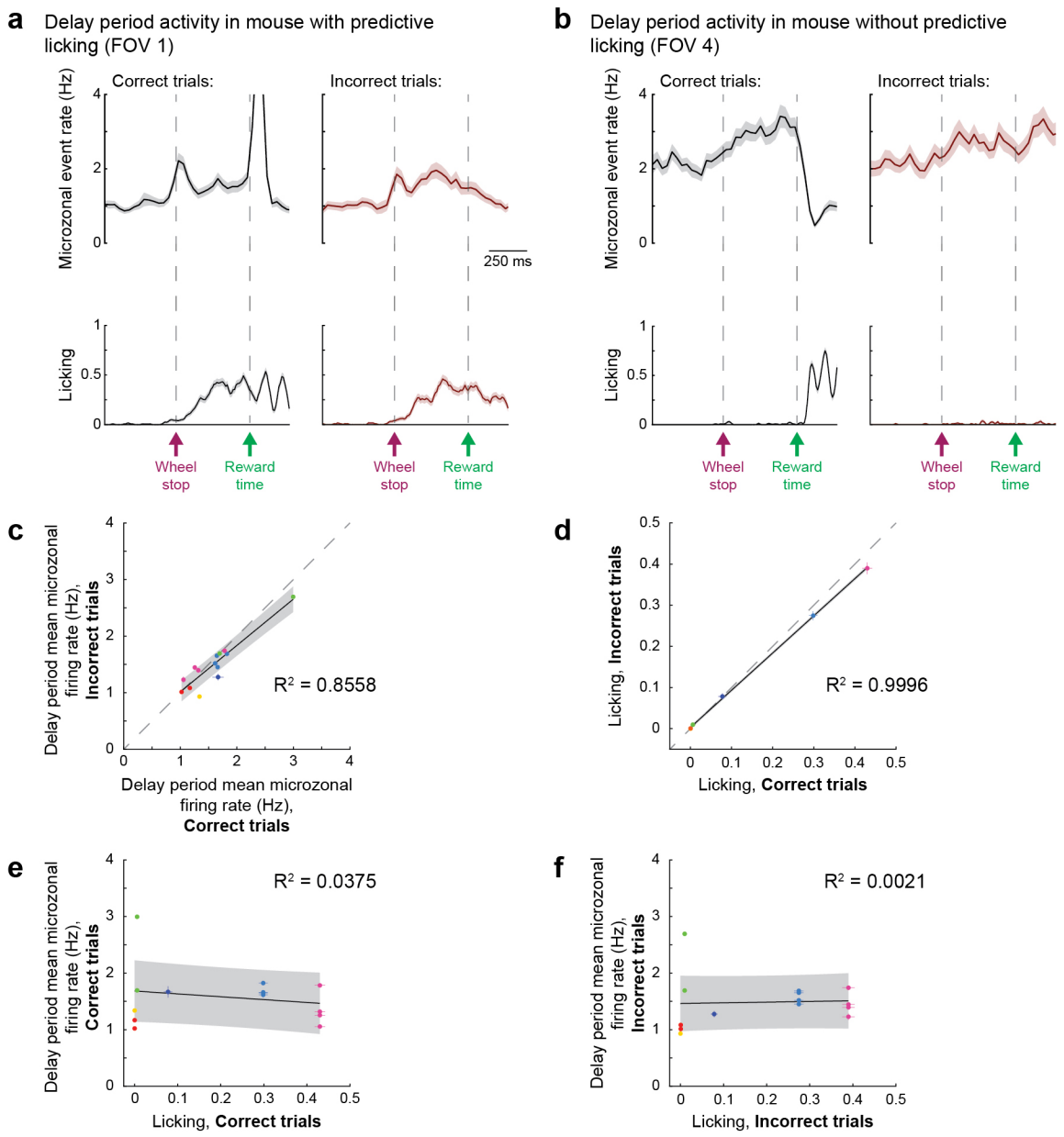
Anatomical maps (subpanel 1), clustering (subpanels 2 and 3), and correlation heatmaps (subpanels 4 and 5) for fields of view used in Figures 1 – 2. Analysis was done independently for spontaneous activity acquired during withhold periods (spont.) and for whole recording session (all data). Numbers of neurons for fields of view in panels **a**-**f** = 210, 124, 122, 273, 237, and 135, respectively.



Supplementary Figure 4

Mapping recorded fields of view onto cerebellar anatomy

Coarse anatomical maps with landmarks are shown in grayscale and reward response is shown in color. Field of view number corresponds to numbers given in Supplementary Figure 3. Purkinje cell dendrites with high reward activity (’reward-activated’) are shown in magenta, and Purkinje cell dendrites with low reward activity (’reward-suppressed’) are shown in cyan. Number of trial rewards per field of view: 120 (Field of view 2), 143 (Field of view 3), 156 (Field of view 4), and 52 (Field of view 5).



Supplementary Figure 5

Relationship between delay period activity and predictive licking on operant trials

a. Example of microzone from animal that exhibited predictive licking showing delay period activity on both correct (left, black) and incorrect (right, red) trials. N = 162 correct trials and 124 incorrect trials.

b. Same as panel **a** but for animal that did not exhibit predictive licking. N = 156 correct trials and 115 incorrect trials.

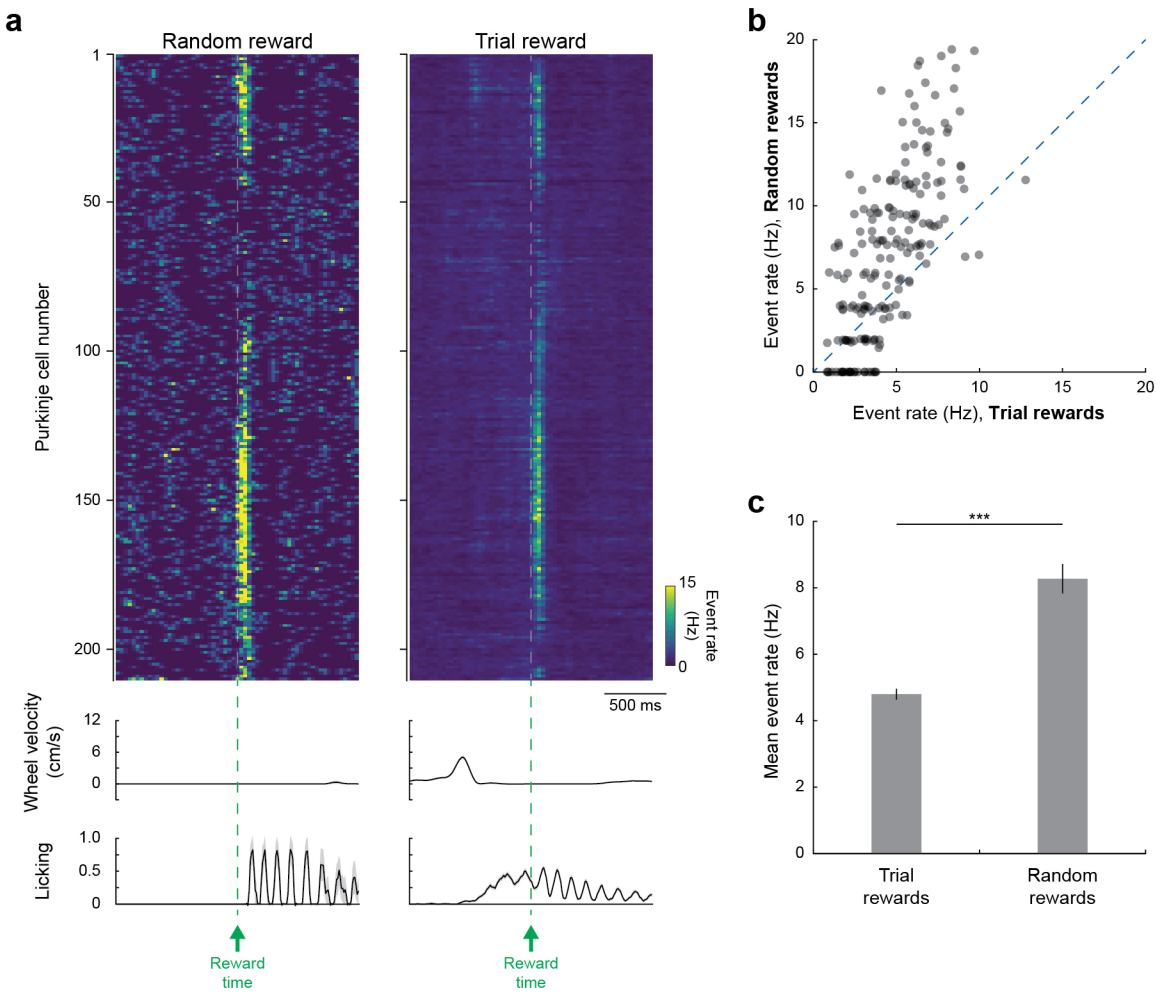
c. Comparison of mean microzonal activity during delay period (-500 to 0 ms relative to reward) in significantly activated microzones on correct and incorrect trials. Dots of the same color represent microzones from the same recording. N = 14 microzones from 6 mice.

d. Comparison of degree of predictive licking on correct and incorrect trials. N = 6 sessions from 6 mice.

e. Relationship between level of delay period activation and degree of predictive licking on correct trials. N = 14 microzones from 6 mice.

f. Same as panel **e** but for incorrect trials.

Traces in panels **a** and **b** and data points in panels **d**-**f** are shown as mean ± s.e.m. Black line and shaded bars in panels **d**-**f** represent linear fit and 95% confidence interval through all data points (linear regression; panel **c**, $p = 2x10^{-6}$; panel **d**, $p = 5x10^{-8}$; panel **e**, $p = 0.5$; panel **f**, $p = 0.9$).



Supplementary Figure 6

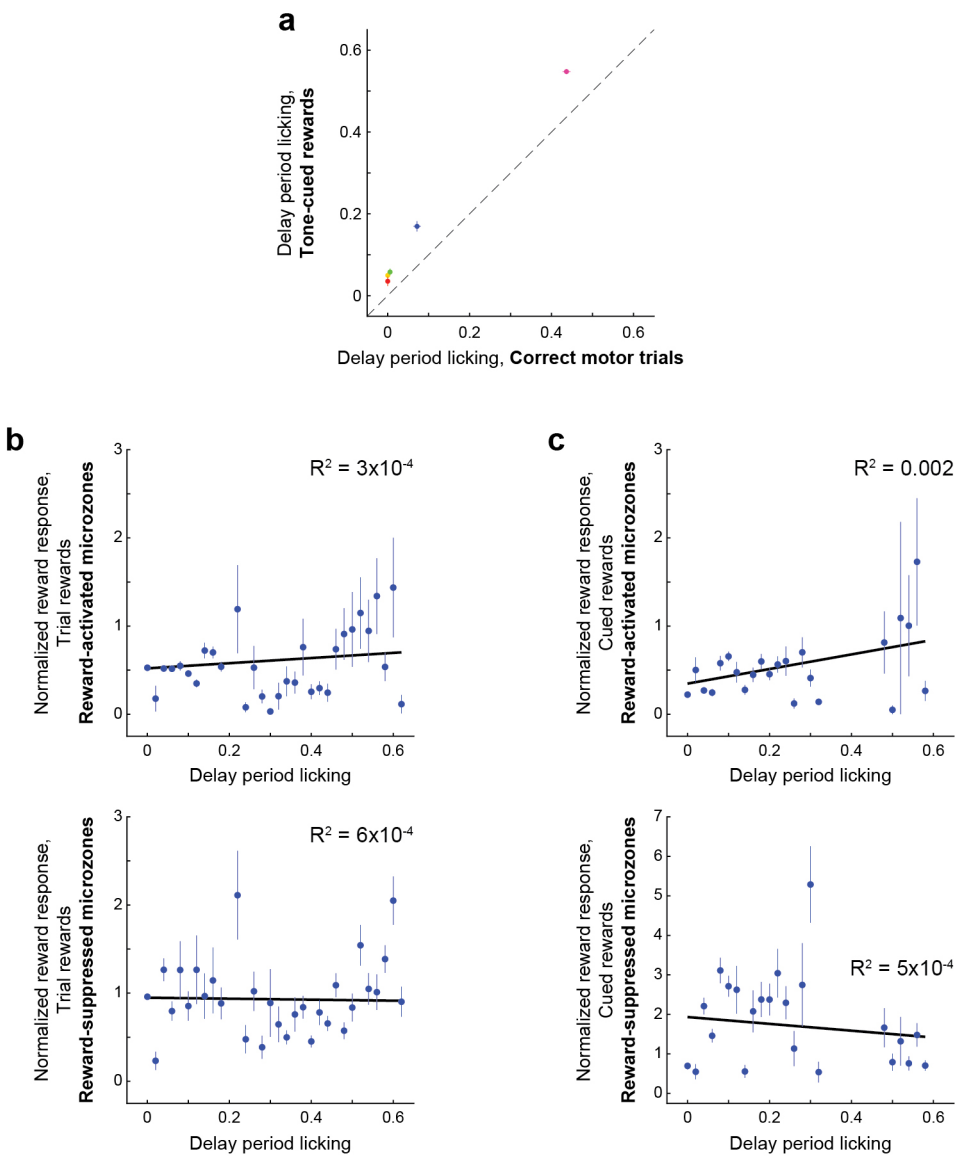
Initial observation of reward response modulation

a. Top: Trial-averaged population response in the first mouse in which we observed reward response modulation (Supplemental Figure 3, FOV 1). Middle: Trial-averaged steering wheel velocity. Bottom: Trial-averaged licking response. Data are shown as mean ± s.e.m. across trials (n = 162 trial rewards and 5 random rewards). In this experiment, the mouse was not subjected to all of the potential reward contingencies, only to random and trial-dependent (operant) rewards, but still exhibited the characteristic reward response modulation.

b. Scatter plot comparing reward-related responses over interval 0 to +100 ms post reward for random rewards and rewards earned on correct motor trials.

c. Cell-wise average of Purkinje cell dendritic response to trial rewards and random rewards taken over interval 0 to +100 ms post reward. Data are shown as mean ± s.e.m. across cells (n = 210 cells, two-sided Wilcoxon signed-rank test, $p = 3x10^{-18}$).

Statistics summary: ***p < 0.001.



Supplementary Figure 7

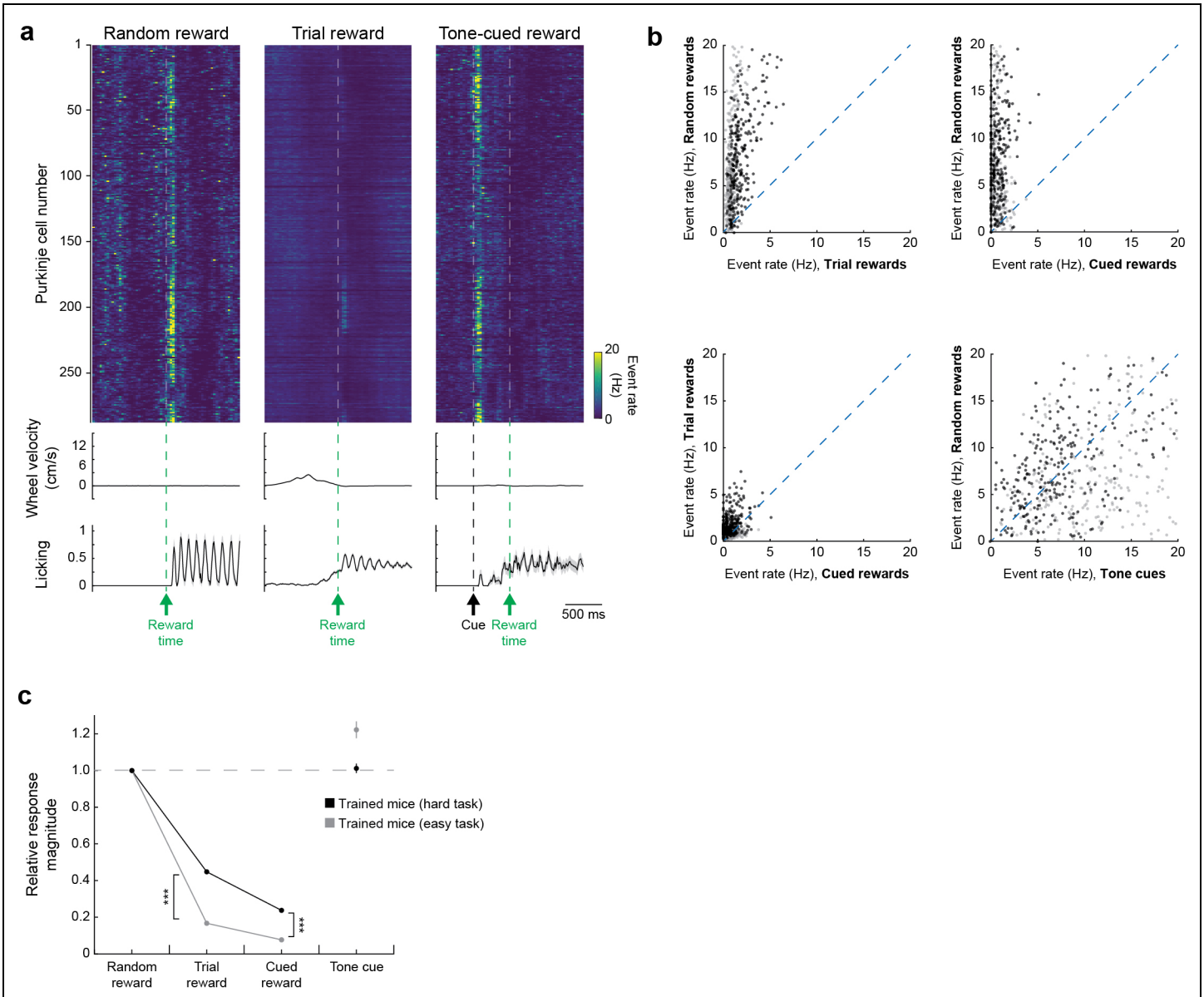
Quantification of licking responses across reward conditions

a. Comparison of delay period (reward-predictive) licking for tone-cued rewards and correct motor trials (-500 to 0 ms relative to reward in both conditions, which had the same delay interval). Mean licking responses in individual mice (n = 5) are as single colored dots with error bars showing across trial s.e.m. All mice exhibited higher predictive licking for tone-cued rewards than on correct motor trials ($p = 3x10^{-22}$ (Mouse 1, yellow), $p = 5x10^{-17}$ (Mouse 2, red), $p = 1x10^{-16}$ (Mouse 3, green), $p = 1x10^{-7}$ (Mouse 4, blue), $p = 1x10^{-6}$ (Mouse 5, magenta), two-sided Wilcoxon rank sum test across trials).

b. Normalized response to trial rewards (mean activity in interval 0 to +100 ms after reward in Purkinje cells from reward-activated microzones (top) and reward-suppressed microzones (bottom) as a function of degree of predictive licking. Reward-related responses and predictive licking were quantified on individual trials, after which responses in individual neurons were normalized to the mean response to random reward per neuron and all neuron-trial pairs were binned for plotting according level of predictive licking in each trial.

c. Same as panel a but for cued rewards.

Data are shown as mean ± s.e.m. In panels **b** and **c**, N = 38337 neuron-condition pairs (trial rewards, activated), 51888 pairs (trial rewards, suppressed), 11584 pairs (cued rewards, activated), and 11314 pairs (cued rewards, suppressed). The number of Purkinje cells = 361 (reward-activated) and 450 (reward-suppressed) pooled from 5 mice. Black line represents linear fit through all data points (not binned).



Supplementary Figure 8

Stronger suppression of reward response on operant trials with an easier task

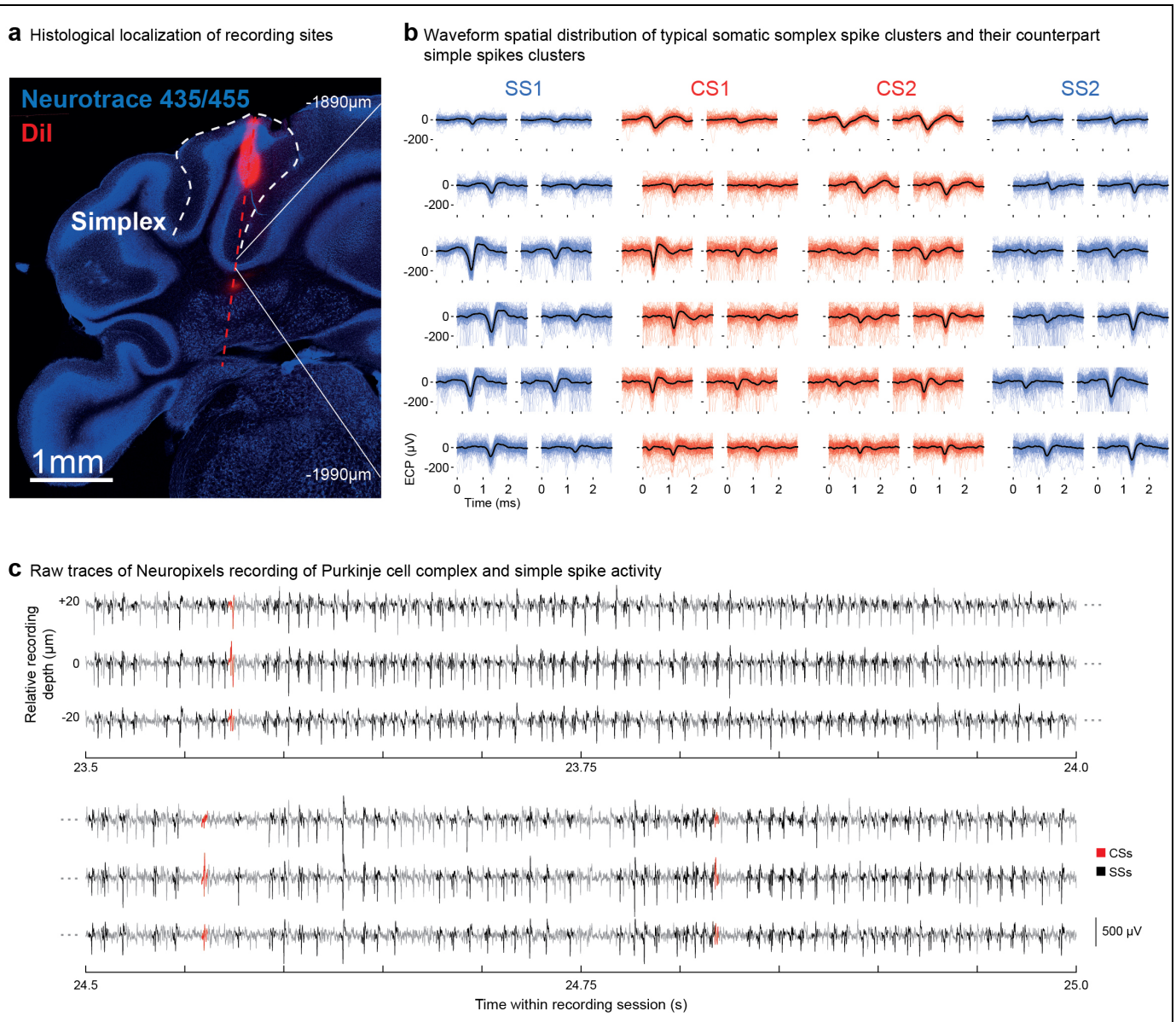
a. Top: Trial-averaged population response of a representative field of view to random, operant, and tone-cued rewards on a different task version in which all wheel movements towards the midline were rewarded. ROIs are by mediolateral position within the field of view. Middle: Trial-averaged steering wheel velocity. Bottom: Trial-averaged licking response. Velocity and licking are shown as mean ± s.e.m. across trials (n = 11 random rewards, 210 trial rewards, and 20 tone-cued rewards).

b. Scatter plots showing pairwise comparisons of response amplitude (computed as mean over 0 to +100 ms after each reward-related event) across different reward conditions. N = 556 neurons from 2 FOVs in 2 mice. Data points from representative field of view (panel **a**) are shown in darker gray.

g. Relative response magnitude in neurons responsive to random reward (mean response over 0 to 100 ms after random reward > 2 s.d. above baseline) in mice trained on the task used throughout the majority of this study (black) and the easier task version (gray). Data are shown as mean ± s.e.m., and n = 400 neurons (of 891) for mice trained on the hard task and n = 417 neurons (of 556) for mice trained

on the easy task (Kruskal-Wallis test, $H = 2543$, d.f. = 7, $p < 1x10^{-99}$, significance values for Bonferroni-corrected individual comparisons: hard task vs easy task (trial reward), $p = 5x10^{-25}$; hard task vs easy task (cued reward), $p = 9x10^{-15}$).

Statistics summary: ***$p < 0.001$.



Supplementary Figure 9

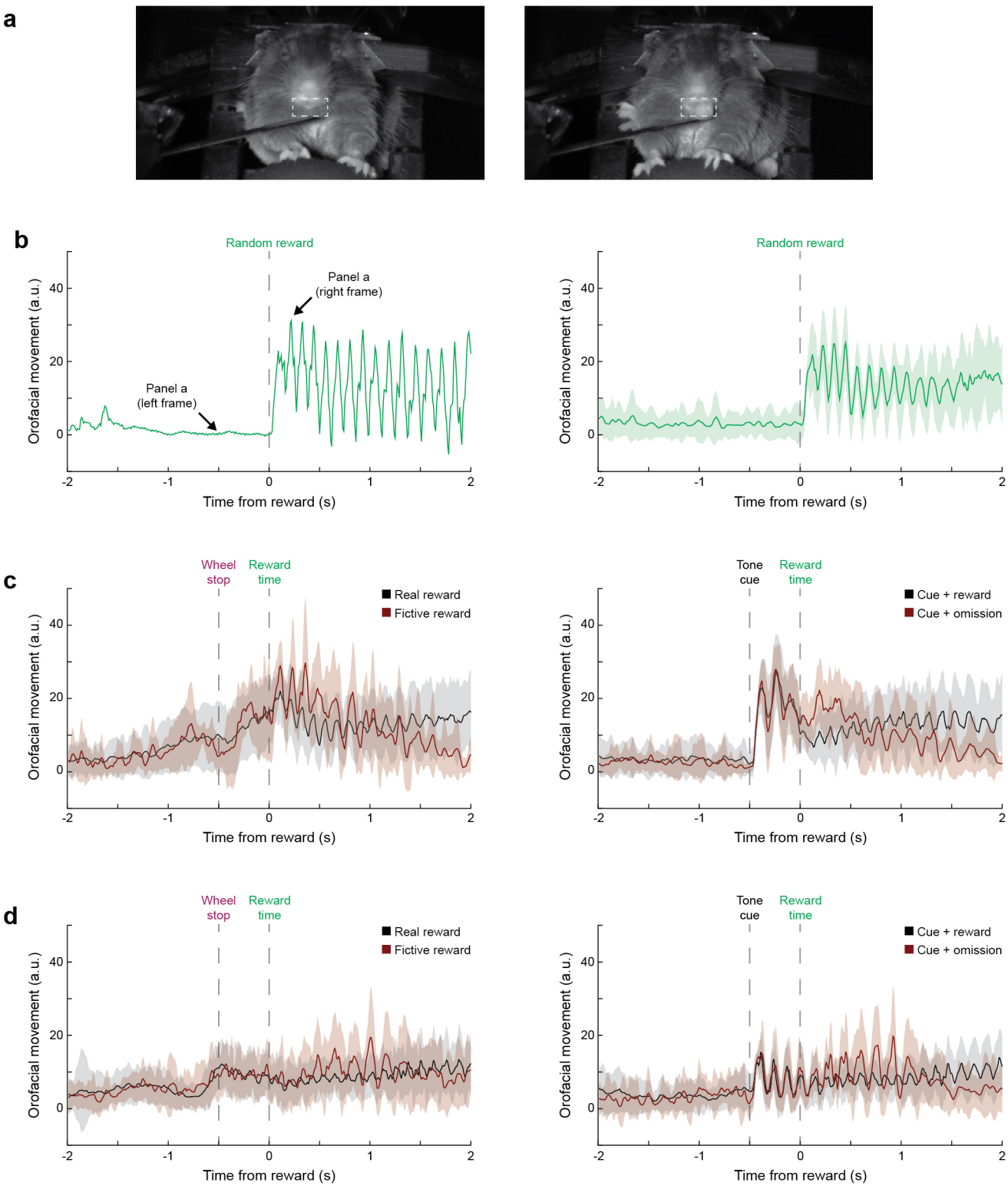
Recordings from Purkinje cells using Neuropixels probes

a. Histological localization of the Neuropixels recording track in a coronal section of cerebellum (6.2mm caudal from bregma). Blue: Neurotrace, Red: DiI (highlighting the electrode track). Dashed red line shows the projection of the electrode track onto this anatomical section.

b. Examples of complex spikes (CSs) and simple spikes (SSs) recorded using Neuropixels probes. Each panel shows 100 overlaid raw traces (with mean and s.d. shown as black line and shaded area, respectively) on adjacent recording sites (20 µm vertical and horizontal separation), situated from -1890 µm to -1990 µm below the surface of the brain shown in panel **a**. Red: CS recordings showing channels located in the molecular (top channels) and Purkinje (middle channels) layers of the cerebellar cortex. Blue: corresponding SS recordings from the same putative Purkinje cells exhibiting post-CS pauses, as in Figure 4C (SS1 and SS2 associated with CS1 and CS2, respectively).

c. Example raw traces (gray) recorded on three adjacent vertically consecutive channels (20 µm separation) of a Neuropixels probe within the Purkinje cell layer. The spikes of the simple spike cluster (black) and complex spike cluster (red) shown in Figure 4A-C are

highlighted. Note that several other simple and complex spike clusters were identified within this Purkinje cell layer but are not highlighted for clarity.



Supplementary Figure 10

Video analysis of orofacial movements on omission trials

To better understand how complex spiking responses upon reward omission might be reflected behaviorally, we analyzed videos of orofacial movements in a subset of mice presented with real and fictive trial reward and with delivered and omitted cued rewards. We observed that upon violations of reward expectations, mice exhibited larger orofacial movements, consistent with them searching for an expected reward.

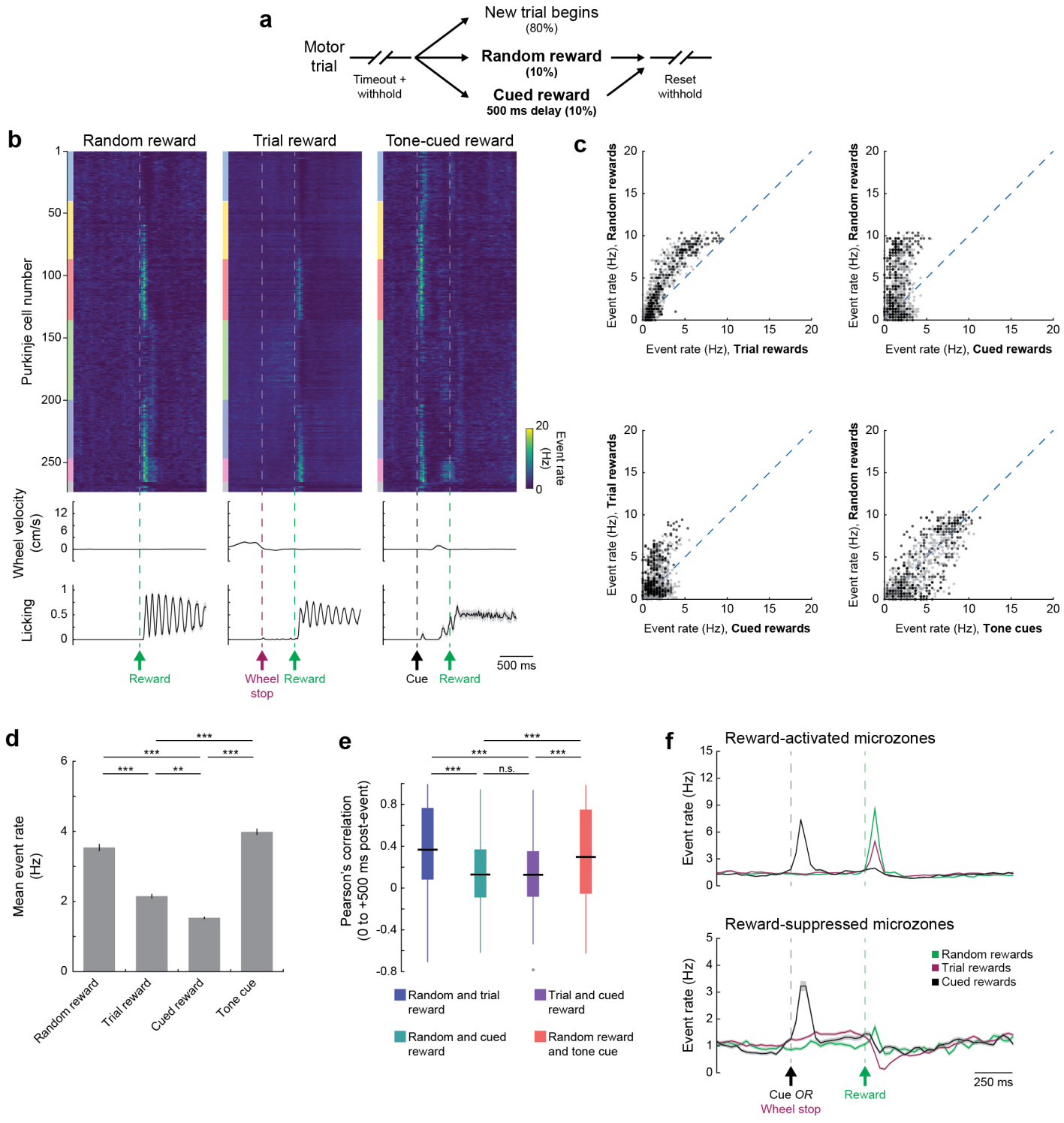
a. Two video frames (taken at 100 Hz) show mouse 500 ms before a random reward was delivered (left) and 220 ms after the random reward was delivered (right). White dashed boxes indicate regions of analysis (~4 x 8 mm).

b. Orofacial movement signal recorded for a single random reward (left, corresponding to video frames in panel **a**) and the mean orofacial movement signal (right) across all random rewards in a session (n = 30).

c. Mean orofacial movement signals from example mouse (same session as panels **a** and **b**) on operant trials (left) with real reward (black, n = 52) and fictive reward (red, n = 8), and signals on cued reward trials (right) on which reward was given (black, n = 40) and omitted (red, n = 14).

d. Same as panel **c** but for a different mouse. N = 54 operant trials with real reward, 14 operant trials with fictive reward, 53 cue + reward trials, and 12 cue + omission trials.

Data in panels **b** - **d** are shown as mean ± s.d. across trials.



Supplementary Figure 11

Modulation of reward responses calculated using unweighted event detection

Data used in this figure are the same as in Figure 3 of the main text, but all detected event amplitudes are set to a value of 1.

a. Schematic of reward perturbation experiments: during each behavioral session, we randomly interspersed random rewards (10% of inter-trial intervals) or tone-cued rewards (also 10% of inter-trial intervals; 500 ms delay between cue onset and reward).

b. Top: Trial-averaged population response of a representative field of view (same as Figure 2) to random, operant, and tone-cued rewards. Dendritic calcium event matrix was binarized for this analysis, setting all event amplitudes to a value of 1. ROIs are sort first by medio-lateral position of identified microzones, then mediolaterally within each identified microzone. Color blocks adjacent to each heatmap denote microzonal designation, following the color scheme of Figure 2 (gray = unclustered). Middle: Trial-averaged steering wheel velocity. Bottom: Trial-averaged licking response. Velocity and licking are shown as mean ± s.e.m. across trials. N = 30 random rewards, 156 trial rewards, and 30 tone-cued rewards.

c. Scatter plots showing pairwise comparisons of response amplitude (computed as mean over 0 to +100 ms after each reward-related event) across different reward conditions. N = 891 neurons from 5 FOVs in 5 mice. Data points from representative field of view (panel **b**) are shown in darker gray.

d. Cell-wise average of Purkinje cell dendritic response to each reward-related event. Data are shown as mean ± s.e.m. (n = 891 neurons from 5 FOVs in 5 mice, Kruskal-Wallis test, H = 428, d.f. = 3, $p = 2x10^{-92}$, significance values for Bonferroni-corrected individual comparisons: random vs trial reward, $p = 4x10^{-17}$; random vs cued reward, $p = 1x10^{-32}$; trial vs cued reward, p = 0.004; trial reward vs tone cue, $p = 1x10^{-51}$; cued reward vs tone cue, $p = 4x10^{-77}$).

e. Cell-wise comparison of Pearson's correlations between pairs of reward-related events computed over 0 to +500 ms after each event. Data are shown as box plots: center line, median; box edges, interquartile ranges; whiskers, range without outliers (1.5 times the interquartile range from box edges); gray points, outliers (n = 891 neurons from 5 FOVs in 5 mice, Kruskal-Wallis test, H = 245, d.f. = 3, $p = 8x10^{-53}$, significance values for Bonferroni-corrected individual comparisons: random and trial reward vs random and cued reward, $p = 1x10^{-34}$; random and trial reward vs trial and cued reward, $p = 6x10^{-37}$; random and cued reward vs trial and cued reward, p > 0.9; random and cued reward vs random reward and tone cue, $p = 5x10^{-17}$; trial and cued rewards vs random reward and tone cue, $p = 1x10^{-18}$).

f. Time course of mean responses across reward conditions for Purkinje cells in reward-activated microzones (top, n = 361 neurons) and reward-suppressed microzones (bottom, n = 470 neurons). Note that 60 neurons were not clustered into a microzone and so excluded from this analysis. Data are shown as mean ± s.e.m.

Statistics summary: n.s. = not significant, **p < 0.01, ***p < 0.001.