Cerebellar granule cells encode the expectation of reward

Mark J Wagner[1,*], **Tony Hyun Kim**[1,2,*], **Joan Savall**[1], **Mark J Schnitzer**[1,3], and **Liqun Luo**[1]

[1]Department of Biology and Howard Hughes Medical Institute, Stanford University, Stanford, CA 94305, USA

[2]Department of Electrical Engineering, Stanford University, Stanford, CA 94305, USA

[3]Department of Applied Physics, Stanford University, Stanford, CA 94305, USA

Abstract

The human brain contains ~60 billion cerebellar granule cells[1], which outnumber all other neurons combined. Classical theories posit that a large, diverse population of granule cells allows for highly detailed representations of sensorimotor context, enabling downstream Purkinje cells to sense fine contextual changes[2–6]. Although evidence suggests a role for cerebellum in cognition[7–10], granule cells are known to encode only sensory[11–13] and motor[14] context. Using two-photon calcium imaging in behaving mice, here we show that granule cells convey information about the expectation of reward. Mice initiated voluntary forelimb movements for delayed water reward. Some granule cells responded preferentially to reward or reward omission, whereas others selectively encoded reward anticipation. Reward responses were not restricted to forelimb movement, as a Pavlovian task evoked similar responses. Compared to predictable rewards, unexpected rewards elicited markedly different granule cell activity despite identical stimuli and licking responses. In both tasks, reward signals were widespread throughout multiple cerebellar lobules. Tracking the same granule cells over several days of learning revealed that cells with reward-anticipating responses emerged from those that responded at the start of learning to reward delivery, whereas reward omission responses grew stronger as learning progressed. The discovery of predictive, non-sensorimotor encoding in granule cells is a major departure from current understanding of these neurons and dramatically enriches contextual information available to postsynaptic Purkinje cells, with important implications for cognitive processing in the cerebellum.

Mice voluntarily grasped the handle of a manipulandum (**Methods**) and pushed it forward ~8 mm for delayed receipt of a sucrose water reward (Fig. 1a). Highly trained mice made

many forelimb movements per session (191 ± 13 movements, mean ± s.e.m., across 20 experiments in 10 mice). To record neural activity, we used mice that expressed the genetically-encoded Ca^{2+} indicator GCaMP6f selectively in cerebellar granule cells (Fig. 1b, Extended Data Fig. 1a). We developed a chronic imaging preparation to visualize fluorescence responses in granule cell somas during behavior (Video S1; Fig. 1c,d; Extended Data Fig. 1b,c; Supplementary Note 1; $n = 43 \pm 4$ neurons per session). Mice began licking robustly during the delay period following a forelimb movement in anticipation of reward (Fig. 1e,f). Following reward delivery, the handle returned after a delay to permit the mouse to initiate the next movement.

The times of peak Ca^{2+} activity were heterogeneous and collectively spanned the task duration in highly trained mice (Fig. 1g). 85% of all recorded neurons exhibited significant task modulation ($n = 561$ total neurons from 6 mice). Some neurons exhibited maximal fluorescence during the forelimb movement (Fig. 1g example cells ~50–90; Extended Data Fig. 2a). Others were inhibited during movement (example cells ~1–40; Extended Data Fig. 2b). Consistent with the traditional role of sensorimotor representation in the cerebellum[15], neural response magnitude covaried significantly with peak movement velocity in 20% of granule cells (Extended Data Fig. 2c,d). Intriguingly, many other neurons exhibited response peaks during the delay period before the reward (example cells ~90–140) or during reward consumption (example cells ~140–170; Extended Data Fig. 2a).

Given the prominence of sensorimotor signals in the cerebellum, neural activity near the time of reward delivery could represent body movement or reward sensing. To discern its origins, we examined Ca^{2+} responses when omitting reward delivery on a randomly interspersed 1/6–1/4 of trials. We observed that some granule cells responded preferentially following reward delivery, as compared to instances of omitted reward (Fig. 2a top; Extended Data Fig. 3a–c). In principle, these could result from differences in overt motor output such as licking, which was substantially prolonged following reward compared to omitted reward (Fig. 2a; Extended Data Fig. 2e,f). We therefore compared rewarded trials with exceptionally high or low amounts of licking during reward consumption and found that reward-selective neurons were not modulated by licking (Fig. 2a bottom). Nevertheless, this does not exclude the possibility that reward-selective cells simply encode water-related sensory stimulus.

Surprisingly, many other granule cells exhibited larger responses following omitted reward than rewarded trials. Responses to omitted reward occurred without unique sensory input, and so cannot be a sensory response. We divided these responses in two types (**Methods**). The first type ("reward omission") became active following the omitted reward (Fig. 2b top; Extended Data Fig. 3d). The second type ("reward anticipation") became active before expected reward delivery and ceased to be active when the mouse received reward (Fig. 2c top, blue curve). But if expected reward was omitted, the neurons continued to be active for longer (Fig. 2c top, red curve; Extended Data Fig. 3e). Reward omission and reward anticipation neurons were also insensitive to licking during reward consumption (Fig. 2b,c second row). Thus, reward omission responses are not due to sensory input or reduced licking.

We hypothesized that reward anticipation neurons encoded a cognitive state of expectant waiting. As anticipatory licking is a behavioral readout of anticipation[16], we reasoned that it should influence the activity of reward anticipation neurons. Indeed, these neurons exhibited more anticipatory activity on trials with more anticipatory licking (Fig. 2c third row), and these quantities covaried on single trials (Fig. 2c bottom). On the other hand, when we omitted reward, mice stopped licking when they concluded no reward would be received, and therefore ceased anticipating. Therefore, activity of these neurons following omitted reward also covaried with the amount of licking following omitted reward (Fig. 2c bottom). By contrast, following reward delivery, licking exerted no effect on these neurons' responses (Fig. 2c bottom). Thus, reward anticipation cells track licking only when it represents anticipation, but not during reward consumption.

Three additional lines of evidence argue against body movement as a cause of reward-related responses. First, we leveraged natural variability in mouse body motion to determine its effect on reward signaling. Via video tracking, we identified sets of rewarded trials with body motion most similar or most dissimilar to body motion on omitted reward trials, and found that reward-related responses were similar on both sets of trials (Extended Data Fig. 4; Video S2). Second, inter-trial interval analyses revealed that reward omission cells do not encode preparation for the next trial (Extended Data Fig. 5). Third, to decouple movement and reward, we trained mice to alternate push-for-reward with pull-for-reward trials (Fig. 2d, black curves). Mice developed anticipatory licking in both conditions (Fig. 2d, colored curves). Reward anticipation neurons identified solely from activity on pushing trials (Fig. 2e top) exhibited highly conserved reward anticipation responses on pulling trials (Fig. 2e bottom). Thus, reward anticipation cells generalize across sensorimotor context. Both reward and reward omission responses were similarly generalized (Extended Data Fig. 6a,b). By contrast, pushing or pulling movement-encoding cells exhibited substantially different responses (Extended Data Fig. 6c,d). Although we cannot exclude the possibility that smaller covert motion unaccounted for by these analyses could contribute to apparent reward-related signaling, these results suggest that granule cells can signal reward expectation independent of body movement.

To quantify the prevalence of reward responses in all recorded cerebellar lobules (Fig. 2f), we computed each cell's response preference for reward vs. omitted reward and compared it to its response to high vs. low reward licking (Fig. 2g). We classified 5.5% of neurons as reward cells and 12.3% as reward omission cells, both with minimal sensitivity to licking. Reward anticipation cells contributed an additional 8.9% of neurons (Fig. 2h; Extended Data Fig. 3h). Consistent with the prominence of reward signals, granule cell ensembles linearly discriminated reward outcome on single-trials with $93 \pm 2\%$ accuracy (Extended Data Fig. 7a–e). In addition, linear decoding of granule cell ensembles accounted for $44 \pm 3\%$ of the fine moment-by-moment fluctuations in a behavioral estimate of reward anticipation (Extended Data Fig. 7f–h).

To examine whether cerebellar granule cells encode reward expectations in disparate reward contexts, we retrained 5 mice that had performed the operant task for a Pavlovian task in which reward was delivered at a fixed delay following a tone. Tone was separated from the prior trial's reward by a random delay. Among normal trials we randomly interspersed three

types of probe trials on which we omitted the reward after a tone, delivered a large reward after a tone, or delivered a reward without a preceding tone ($n = 241 \pm 3$ total trials per each of 11 sessions in 5 mice). After training on this task, mice also began licking before reward delivery as in the forelimb movement task (Fig. 3a). Reward-related Ca^{2+} responses in the Pavlovian task resembled those in the operant task: reward responding, omitted reward responding, and reward anticipation (Fig. 3b–d top; Extended Data Fig. 8a–c). These cells occurred in all imaged lobules in proportions similar to those seen in the forelimb movement task (Extended Data Fig. 8d; 5.1% reward, 9.3% reward omission, 5.6% reward anticipation). Reward anticipation neurons were again sensitive to anticipatory licking but not reward licking (Extended Data Fig. 8e), indicating signaling of expectation rather than licking.

Unexpected reward trials further supported that granule cells encode reward expectation. Sensory reward stimulus and licking response on these trials were the same as on normal trials (Fig. 3a; $p = 0.75$, $n = 11$ experiments, Wilcoxon rank-sum test for time of 50% decline in licking during reward consumption). Some reward cells were also found to encode expectation rather than only sensory input, as they exhibited larger responses to unexpected than expected reward (Fig. 3b bottom). Reward omission neurons did not distinguish expected from unexpected reward (Fig. 3c bottom; Extended Data Fig. 8b), suggesting a selective sensitivity to reward omission. Furthermore, the cognitive state of anticipation should be absent during unexpected reward, despite sensorimotor input identical to expected reward. Indeed, we found that reward anticipation neurons were silent following unexpected reward (Fig. 3d bottom; Extended Data Fig. 8c). Thus, these cells selectively encode anticipation but not reward or reward consumption. Comparing reward preference to unexpected reward preference across mice revealed that 12% of neurons preferred unexpected reward whereas 9% preferred expected reward (Fig. 3e, **Methods**). In addition, some neurons distinguished normal rewards from large rewards, with minimal sensitivity to licking (Extended Data Fig. 8g–i). The Pavlovian task thus confirmed the reward signaling observed during forelimb movements, while uncovering additional encoding of reward expectation and reward magnitude.

To investigate how reward anticipation signals develop during the training phase of our tasks, we tracked activity of the same granule cells each day while mice learned the forelimb pushing task (Fig. 4a, Extended Data Fig. 9a–g). Comparing population responses during the task late versus early in learning revealed a substantial decrease in neurons responding robustly to reward, and a substantial increase in neurons responding robustly during the delay period in anticipation of reward (Fig. 4b). Following the same neurons across days over the course of learning (Fig. 4c; Extended Data Fig. 9h), we found that neurons active during forelimb movement (example cells ~20–50) appeared to be more stable than neurons active around the reward period (example cells ~60–80). Comparing responses on the first and fifth day of exposure to omitted rewards, we observed many more neurons with reward omission responses (Fig. 4d; example cells ~60–90).

To quantify these observations, we performed retrospective analyses of neurons whose responses were strongest on the last day of imaging. Interestingly, neurons with strong anticipatory responses on day 6 primarily responded only after reward earlier in learning

(Fig. 4e top). For neurons with the strongest day-6 preference for omitted reward compared to reward (Fig. 4f top), responses to reward omission became stronger over days. By contrast, neurons with the strongest day-6 forelimb movement response also responded to forelimb movement on all previous days (Fig. 4g top). These differences were also evident when we quantified the responses across all recorded neurons (Fig. 4e–g, bottom). Over the same period, changes in licking and in forelimb motion were modest (Extended Data Fig. 9i,j) and were therefore unlikely to account for neural response changes. Thus, reward-related responses are highly dynamic during learning, with reward responses becoming progressively more anticipatory and omitted reward response preferences growing in magnitude over days. Given the importance of granule cell signaling in learning[17], the adaptive changes we observe are well placed to impact downstream cerebellar learning processes.

To our knowledge, this is the first *in vivo* recording of cerebellar granule cells during the execution and learning of goal-directed behavior. Besides movement-encoding granule cells as predicted from previous studies[18,19,14], we found that granule cells signal reward expectation in multiple contexts (Supplementary Table 1) and in all cerebellar lobules imaged. Reward omission cells substantially outnumbered reward cells, even though reward is a sensory stimulus that elicits a larger licking response. This discrepancy may be related to our finding that omitted reward responses increase while reward responses decrease during learning. The abundance of reward omission granule cells could relate to cerebellar signaling of unexpected events[20].

Reward signals have been best studied in the ventral tegmental area (VTA)[21,22] but also documented in other brain regions such as the ventral striatum[23], orbitofrontal cortex (OFC)[24], and dorsal raphe nucleus (DRN)[25]. Most VTA dopamine neurons respond selectively to unexpected rewards or reward-predicting stimuli and are suppressed by omitted rewards. Thus, reward anticipation granule cells do not resemble VTA responses. Rather, they are reminiscent of responses in striatum[23], OFC[24], and DRN[25] during goal-directed behavior. Reward omission signals are found mainly in anterior cingulate cortex and the lateral habenula[26,27]. Granule cell reward signals could thus arise from many places although unlikely from a direct VTA→cerebellum projection (Extended Data Fig. 10). Neocortex provides an especially large mossy fiber input[8] via the pons and thus merits further study.

An outstanding question is how reward context contributes to cerebellar function. Classical models posit that granule cells signal sensorimotor context. The incorporation of reward, reward omission, and reward anticipation signals should allow the cerebellar cortex to integrate sensorimotor information with signals reflecting internal brain state, drive, and affective status, and in so doing drastically expanding its function as a learning machine (Supplementary Note 2). Studying the causal role of these cells will require future technical advances to specifically manipulate reward-related granule cells without disrupting those essential for sensorimotor functions. Nevertheless, that granule cells can encode reward expectation clearly indicates that the contextual information available to downstream Purkinje cells is far richer than previously described, and provides a means for cerebellar involvement in a wide variety of cognitive computations.

Methods

Mice

To express the Ca^{2+} indicator GCaMP6f[28] in cerebellar granule cells, we used cre- and tTA-dependent GCaMP6f transgenic mouse line Ai93 (*TRE-lox-stop-lox-GCaMP6f*)[29]. We crossed the Ai93 mouse to a cre-dependent tTA mouse ztTA (*CAG-lox-stop-lox-tTA*)[30]. We then crossed Ai93/ztTA mice to *Math1-cre*[31] which in the cerebellum is expressed selectively in granule cell progenitors[32]. We used a total of ten Ai93/ztTA/*Math1-cre* triple transgenic mice (4 female and 6 male) for all experiments. Six contributed to the main pushing operant task data in Fig. 1–2 and Extended Data Fig. 1–3 and 7, five of those mice contributed to the Pavlovian task data in Fig. 3 and Extended Data Fig. 8, and three of them contributed to the operant learning data in Fig. 4 and Extended Data Fig. 9. The remaining four mice contributed to the push-pull operant task data in Fig. 2d,e and Extended Data Fig. 6, and three of those also contributed to the video tracking data in Extended Data Fig. 4. These sample sizes permitted acquisition of hundreds of cells per data set with hundreds of trials, sufficient to make the statistical claims in the study. Mice were aged 6–12 weeks at the start of procedures. For Extended Data Fig. 10, we used 4 Ai14 mice (*lox-stop-lox-tdTomato*)[33] and one *frt-stop-frt-lox-tdTomato* mouse (derived from Ai65, *frt-stop-frt-lox-stop-lox-tdTomato*[29] by crossing to germline-cre; kindly provided by Andrew Shuster). Stanford University's Administrative Panel on Laboratory Animal Care (APLAC) approved all procedures. All control conditions were internal to each animal and thus neither randomization nor blinding was performed.

Histology

We confirmed expression of GCaMP6f in cerebellar granule cells in fixed tissue from animals after performing experiments. We anesthetized mice using tribromethanol (Avertin) and transcardially perfused them with phosphate-buffered saline (PBS) followed by 4% paraformaldehyde (PFA). We extracted the brains into 4% PFA for 24 h of post-fixation, followed by at least 24 h in 30% sucrose solution. We cut 40 or 60 μm tissue sections on a cryotome (Leica). To label Purkinje neurons we used a monoclonal anti-calbindin mouse antibody at 1:1000 dilution in PBST (Sigma). To stain for GCaMP6f we used a polyclonal GFP chicken antibody at 1:2000 dilution in PBST (Aves Labs). We incubated both primary antibodies for 48 hours, followed by 3 hours in FITC donkey anti-chicken and Alexa-647 goat anti-mouse secondary antibodies (Jackson Immunoresearch), both at 1:500 dilution in PBST. We then stained for DAPI at 1:20,000 dilution for 20 minutes. We imaged the sections using a confocal microscope (Zeiss) and a 40× 1.4 NA objective (Fig. 1b) or a 20× 0.75 NA objective (Extended Data Fig. 1a). To stain for tyrosine hydroxylase (TH; Extended Data Fig. 10), we used a polyclonal rabbit anti-TH antibody (Millipore AB152) at 1:2000 dilution followed by donkey anti-rabbit secondaries conjugated either to Alexa-488 or Alexa-647 (Jackson Immunoresearch) at 1:500 dilution.

Surgical procedures

We anesthetized mice using isoflurane (1.25–2.5 % in 0.7–1.3 L/min of O_2) during surgeries. We removed hair from a small patch of skin, cleaned the skin, and made an incision and removed the patch of skin. We then peeled back connective tissue and muscle

and dried the skull. We then drilled a 3 mm diameter cranial window centered rostrocaudally over the post-lambda suture and centered 1.5 mm right of the midline. This positioned the window over cerebellar lobules VIA, VIB and simplex. To seal the skull opening, we affixed a #0 3 mm diameter glass cover slip (Warner Instruments) to the bottom of a 3 mm outer diameter, 2.7 mm inner diameter stainless steel tube (McMaster) cut to 1 mm height. We stereotaxically inserted the glass / tube combination into the opening in the skull at an angle of 45° from the vertical axis and 25° from the AP axis. We then fixed the window in place and sealed it using Metabond (Parkell). We next affixed a custom stainless steel head fixation plate to the skull using Metabond (Parkell) and dental cement (Coltene Whaledent). The 1.2 mm thickness fixation plate had a 5 mm opening to accommodate the stainless steel tube protruding from the window, and two lateral extensions to permit fixing the plate to stainless steel holding bars during imaging and behavior.

For viral surgeries (Extended Data Fig. 10), we drilled a small hole (~0.5 mm) in the cranium over the cerebellum, either over Lobule VI (−6.8 mm AP, 0.75 mm lateral, 0.35 mm below the brain surface; $n = 4$ mice) or over Lobule Crus I (−7.2 mm AP, 3 mm lateral, $n = 1$ mouse). We injected 500 nL of either *CAV2-cre* into Ai14 animals ($n = 4$ mice) or *AAVretro-EF1a-FLPo* into *frt-stop-frt-lox-tdTomato* mouse ($n = 1$ mouse). Animals were sacrificed 1 – 2 weeks after viral infection.

Behavior

For all behavior, mice were water restricted to 1 mL of water per day. Mice were monitored daily for signs of distress, coat quality, eye closing, hunching, or lethargy to assure adequate water intake. During behavioral training and imaging, mice generally received all water during daily training sessions. For each task, mice trained for 7–14 days for ~30–60 minutes daily, depending on satiety. In both tasks, we recorded licking at 200 Hz sampling rate using a capacitive sensor coupled to the metal water port which delivered ~6 μL 4% sucrose water reward near the animal's mouth. Raw binary lick traces were smoothed with a 2nd order Butterworth filter with 5 Hz cutoff frequency for all analyses except Extended Data Fig. 7f–h, which used instantaneous lick rate as described below. For all experiments mice were head-fixed with their bodies from the torso down in a custom printed plastic tube. For video tracking experiments this tube was printed from optically transparent material.

Forelimb movement task

Mice learned to voluntarily initiate pushing the handle of a manipulandum. We custom designed the manipulandum in a double SCARA mechanical configuration[34] to allow two-dimensional planar motion with minimal inertia. The robot was constructed from custom printed plastic parts and actuated by two motors (Maxon RE-max 21) and monitored by two encoders (Gurley Precision Instruments R120B). Robotic control relied on nested feedback loops in FPGA (10 kHz; National Instruments LX50) and a real-time operating system computer (1 kHz; National Instruments cRIO-9024) both in a National Instruments cRIO chassis, as well as a Windows PC (200 Hz). The controllers were all programmed in Labview and permitted precise robotic positioning and application of forces to the handle to restrict motion as needed (Wagner et al., manuscript in preparation). The device recorded the handle position with a 200 Hz sampling rate and encoder resolution of 0.003 mm. The

device permitted linear movements of maximum length 8 mm, after which the trial terminated. Following a delay (either 600 ms or 800 ms for 3 mice each), a solenoid released a drop of 4% sucrose water from a tube near the mouse's mouth. Following another delay (either 500 ms or 2 s for 3 mice each) the handle began to return to the home position. This process completed either 2 s or 3.5 s (for 3 mice each) after the previous reward delivery, any time after which the mouse could initiate the next movement. For studies of omitted reward response, on a randomly interspersed minority of 1/6 to 1/4 of trials no reward was delivered.

For the body motion tracking in Extended Data Fig. 4, we used two cameras (The Imaging Source) to visualize the mouse's right side directly, and the mouse's underside via a mirror (Video S2). Behavioral video frame acquisition was synchronized to the two-photon frame acquisition at 29.9 Hz. We manually annotated the videos to track the x–y motion of the right forepaw and the base of the tail from the side view, and of the two hind paws from the underside view. For analysis in Extended Data Fig. 4, we computed for each rewarded trial the time-varying Euclidean distance to the average omitted reward trial body trajectory across the 8 body coordinates (x and y motion of forepaw, tail base and two hind paws). We then took the mean square of this distance from 0.1 to 1.5 s relative to reward to quantify each trial's similarity to omitted reward body motion.

The alternating push-for-reward / pull-for-reward task followed a similar structure as above. After the mouse made a pushing movement and received reward, instead of returning to the home position, the robot released (following the same 3.5 s delay as above) to allow a pulling motion back to the prior home position. Mice were typically trained on this task for ~2 weeks beyond the initial training needed to learn the push-only task.

For learning experiments in Fig. 4, we began imaging studies when mice had achieved sufficient basic competency on the task to produce enough forelimb movements for statistically meaningful analyses (> ~30 movements in a session). Thus initial learning of basic task performance preceded the imaging study, and mice had experienced the forelimb movement task for 4 – 6 days prior to imaging.

Pavlovian tone task

A computer played a 500-ms 8 kHz pure tone, followed by a fixed delay (1.2 s) before reward delivery. A randomized 2–6 s inter-trial interval separated reward from the tone of the succeeding trial. In addition, during imaging, 1/10 of trials consisted of an unexpected reward delivered 2 s after the preceding reward, with no tone, 1/10 of trials consisted of a tone followed by omitted reward, and 1/10 of trials consisted of a tone followed by a larger reward ($2\times$ volume for 2 mice, $3\times$ volume for 3 mice). All mice imaged during the Pavlovian task were previously trained on the forelimb movement task.

Two-photon microscopy

We performed all Ca^{2+} imaging using a custom two-photon microscope with articulating objective arm[35]. We used a $40\times$ 0.8 NA objective (LUMPlanFLN-W, Olympus) for all experiments. 920 nm laser excitation was delivered to the sample from a Ti:sapphire laser (MaiTai, Spectra Physics) at powers of ~50–65 mW. We used ScanImage software[36] (Vidrio

Technologies) to control all image acquisition hardware. All data except Fig. 2d,e and Extended Data Fig. 4, 6 were acquired at 13.5 Hz and 150 μm field of view using galvanometer scanning mirrors. Those remaining data were collected were collected at 29.9 Hz and 320 μm field of view using resonant scanning mirrors. We focused into the tissue ~100–200 μm below the pia surface to reach the granule cell layer.

To ensure alignment of the articulating objective to the glass window on the brain, we performed a back-reflection procedure. We projected a low power visible red laser (CPS180, ThorLabs) co-aligned to the infrared beam onto the glass window. We then visualized the red back-reflection on an iris placed at the objective port. We positioned the mouse and objective angles to center the back-reflection into the iris aperture. This procedure was essential for tracking the same granule cells across days. Slight deviations in image angle result in a different two-photon sectioning angle and therefore a different set of granule cells, due to their extremely small size and high packing density. During image acquisition, we compensated slow axial drifts in real time by frequently comparing the acquired images to the initial image and using an objective z-piezo (P-725.4CD, Physik Instrumente).

To align imaging data to behavioral data, the behavioral computer acquired the microscope's frame clock signal simultaneously with the mouse's behavioral data.

For chronic imaging (Fig. 4, Extended Data Fig. 9), we recorded the coordinates of the field of view with respect to a landmark such as the intersection of blood vessels at the boundary between different lobules. We identified lobules based on vasculature patterns and confirmed the assignment in three mice by visualizing the entire cerebellum after extracting brains at the end of experiments.

Image preprocessing

We first corrected two-photon line scan artifacts to compensate for non-linear motion of the galvanometer mirror. We recorded the position feedback signal of the x (fast axis) scanning mirror and compared to the commanded waveform to determine deviations from the ideal scan pattern. We then inverted this scan error to assign pixels to their true location in the image and thereby compensated the resulting distortion from the nonlinear galvanometer motion. We then compensated rigid lateral brain motion using TurboReg[37].

Extraction of granule cell Ca²⁺ signals

We identified individual active cerebellar granule cells in our imaging videos using automated cell sorting based on principal and independent component analyses (PCA/ICA)[38]. Cells corresponded to a weighted sum of pixels forming a spatial filter. We used automated segmentation and thresholding to truncate these filters down to individual cell bodies by eliminating any spurious, disconnected components. We extracted each neuron's time varying fluorescence trace by applying the spatial filter to the processed videos. We then removed high-frequency noise by low-pass filtering the resulting traces with a 2nd-order butterworth filter (–3 dB frequency: 4 Hz). We removed slow drifts from each trace by subtracting a 10th-percentile-filtered (15 s sliding window) version of the signal. Finally, we z-scored each neuron's fluorescence trace to correct for differences in brightness between cells, and then reported all fluorescence values in the resulting s.d. units.

Aligning granule cells across days

We used TurboReg to align the mean image of each day to the final day, used as the reference. For each day independently, we performed the cell sorting procedure outlined above. In general, this produced an only partially overlapping set of cells between days. We then manually took the union of all unique and spatially non-overlapping cells identified in all 6 days to produce a much larger set of cell spatial filters which we then back-applied to the original imaging data from each day. Thus, neuron counts in these datasets exceeded the standard single-day cell sorting results by factors of ~2x.

Fluorescence response analysis

For Fig. 1 and Extended Data Fig. 2, we aligned data to the midpoint of each forelimb movement. For all other figures and analyses, we aligned data in both the forelimb movement task and tone task to the time of reward delivery. For omitted reward trials, we aligned data to the time at which reward would have been delivered following movement termination or tone onset. For each neuron we averaged the reward-aligned fluorescence response to produce the triggered averages used in all figures.

Definition of granule cell response types

We identified forelimb speed sensitive cells (Extended Data Fig. 2c,d) by averaging their fluorescence from −0.1 to +0.3 s relative to reach midpoint on each trial. We then took the Spearman correlation of the single-trial fluorescence with the peak forelimb movement speed. Cells with $p < 0.01$ (permutation test) were tabulated as significant forelimb speed cells.

We defined reward neurons in both the forelimb and Pavlovian tasks as those whose mean fluorescence averaged from 0.1 to 1 s was > 0.3 s.d. higher than following reward omission. Reward omission neurons conversely had responses > 0.3 s.d. greater than following reward delivery; however, we excluded reward anticipation neurons from this tally, as defined below. To verify that our classified reward outcome selective cells were statistically meaningful, we employed a shuffle test in which we scrambled the "rewarded" and "omitted reward" trial labels (or big reward or unexpected reward for Pavlovian task data) randomly 1,000 times. For each shuffle we computed the reward selectivity as described above. If < 50 of 1,000 shuffles ($p < 0.05$) yielded a larger reward or omitted reward preference than was observed, we concluded the reward preference was significant. Across all data sets in both operant and Pavlovian tasks, 97% of reward omission cells and 98% of reward cells, as defined by activity differences above, fulfilled this criterion. By contrast, the shuffle test alone was less stringent, classifying 1.9 and 2.2 times more reward and reward omission cells respectively at the p < 0.05 level. We defined cells using the more conservative and analytically simpler response difference metric for ease of presentation and consistency with all other analyses in the study.

To exclude cells whose reward selectivity was driven by sensitivity to licking, we further required minimal licking sensitivity defined as < 0.2 s.d. absolute difference between 25 highest and 25 lowest licking trials averaged from 0.1 to 1 s.

We similarly defined neurons significantly discriminating expected from unexpected reward or normal from large reward (in the Pavlovian task) by response differences > 0.3 s.d. averaged from 0.1 to 1 s. 97% of cells sensitive to reward magnitude and 97% of those sensitive to reward expectation defined in this way fulfilled the shuffle test described above, whereas the shuffle test alone less stringently classified 1.8 and 1.4 times as many reward expectation sensitive and reward magnitude sensitive cells respectively at the $p < 0.05$ level.

To identify reward anticipation cells in both the forelimb and Pavlovian tasks, we used two criteria. We required a substantial rise in fluorescence during the delay period (> 0.3 s.d. difference between the mean fluorescence from −0.25 to −0.05 s and the mean fluorescence from −1.3 to −1 s relative to reward), as well as greater fluorescence following omitted reward than reward (> 0.3 s.d. difference in mean fluorescence from 0.1 to 0.6 s).

To identify cells responsive to pushing or pulling movements (Extended Data Fig. 6c,d), we averaged the fluorescence from −1.3 to −1 s relative to reward on each trial and then averaged across pushing trials and pulling trials separately. Cells with a > 0.3 s.d. rise in fluorescence on pushing trials were tabulated as pushing cells, while those with a > 0.3 s.d. rise on pulling trials were pulling cells, compared to mean activity prior to reaching, −1.8 to −1.3 s.

To identify cells inhibited following tone onset (Extended Data Fig. 8f), we subtracted the average fluorescence following the tone (−0.8 to −0.5 s) from the average fluorescence prior to the tone (−1.8 to −1.3 s). We included all cells with a decrease > 0.5 s.d.

For selectivity scatter plots (Fig. 2g, 3e, Extended Data Fig. 8d,h), each point was computed from all trials, and thus has an associated standard error which we excluded for visual clarity but which typically ranged from ~0.1–0.15 s.d.

Population decoding analysis

To linearly discriminate reward outcome from ensemble granule cell activity, for each experiment we constructed a vector of true reward outcomes (0 for reward omission, 1 for rewarded trials). We further constructed a matrix of predictor variables from each cell's mean fluorescence between 0 to 1 s on each trial. We then determined the optimal weighting of all cells by fitting a lasso logistic regression from the ensemble activity matrix to the reward outcomes vector (MATLAB). The lasso performs a series of logistic regressions while varying a penalty that discourages non-zero weights on cells. With increasing penalty, the number of cells included in the regression decreases to the most informative set. For each penalty level, the regression computes the 10-fold cross-validated reward outcome classification accuracy (where $1/10^{th}$ of trials were left out of the fitting procedure to use for testing). This allowed us to determine the minimal cell ensemble size with the highest classification accuracy, which we reported in Extended Data Fig. 7a.

To linearly decode reward anticipation from ensemble granule cell activity, we first defined the time-varying reward anticipation state as the amount of licking (lick rate binned at 200 ms) from −1.5 s to +1.5 s relative to reward delivery. If reward was delivered, we defined anticipation to decline to zero at time +0.1 s following reward. If reward was withheld,

licking continued to indicate anticipation, and licking declined as mice concluded no reward was forthcoming (Extended Data Fig. 7g bottom). We then convolved this signal with a 200-ms exponential to simulate GCaMP6f Ca^{2+} unbinding kinetics[28]. Using this time-varying single trial metric of reward-anticipation, we then fit a lasso linear regression using the simultaneously acquired time-varying fluorescence traces of all neurons. This returned the weighted sum of neurons that optimally recapitulated the reward anticipation signal (Extended Data Fig. 7g top). We assessed the performance of this decoder with the 10-fold cross-validated fraction of variance accounted for by the decoder output (Extended Data Fig. 7f). For each lasso regularization penalty level, we recorded the 10-fold cross validated fraction of variance accounted for by the decoder output (Extended Data Fig. 7h).

Statistical analysis

We used MATLAB (Mathworks) for all statistical tests. We compared medians of two groups using the Wilcoxon rank-sum test. We probed the median difference between groups of paired samples using the Wilcoxon signed-rank test. We also compared the median of a distribution to zero using the Wilcoxon signed-rank test. These nonparametric tests do not assume the data follow a particular statistical distribution. Spearman correlation coefficient significance was determined by permutation test. Histogram error bars were computed from

counting statistics as $\sqrt{N\left(1-\frac{N}{N_{total}}\right)}$, where N = number per bin and N_{total} = total elements.

To determine whether the modulation of an individual cell by forelimb movement was significant in Extended Data Fig. 2a,b, we used an exact permutation test via simulated random datasets. Whereas the observed traces derived from averaging trials aligned to reach midpoint, the simulated random dataset was constructed by averaging the same number of "trials" aligned to random times during the 20–30 minute imaging session. We constructed 1,000 such random datasets. For each cell, on each randomization, we quantified the peak average fluorescence between −2 to 2 s relative to trial alignment. We then sorted all randomizations by peak average fluorescence and determined the $p < 0.01$ cutoff as the 10th largest of the 1,000 simulations. We then compared the observed peak average fluorescence to the $p = 0.01$ cutoff. Cells exceeding this cutoff were significant and tabulated in Extended Data Fig. 2a. We then performed the same analysis using the minimum average fluorescence in Extended Data Fig. 2b.

Data availability statement: Data and code are available from the author upon request.
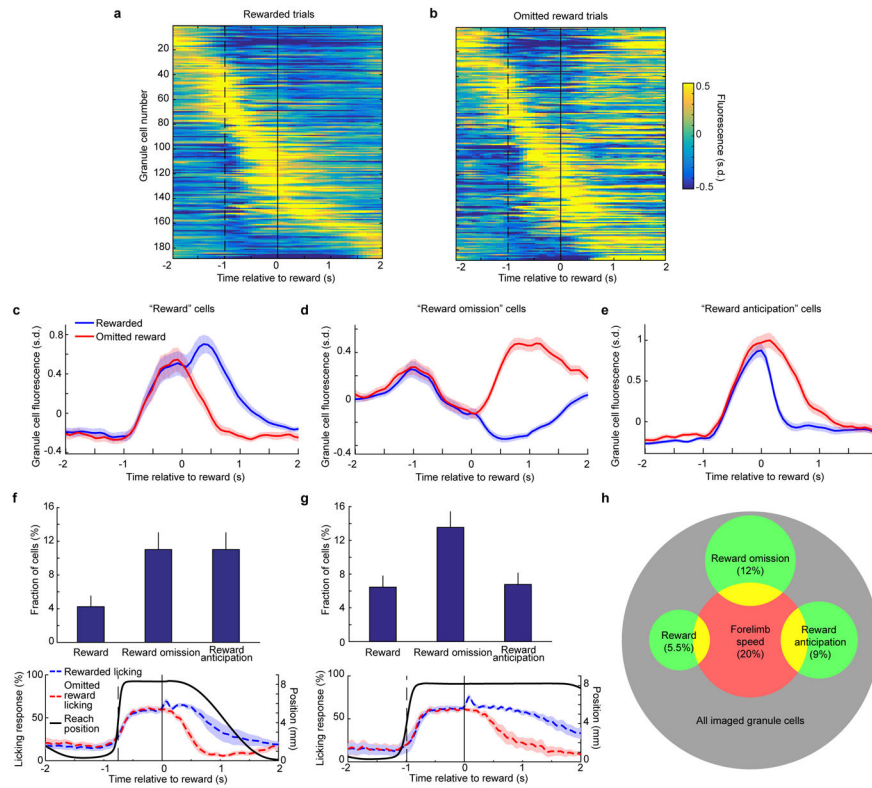
Extended Data
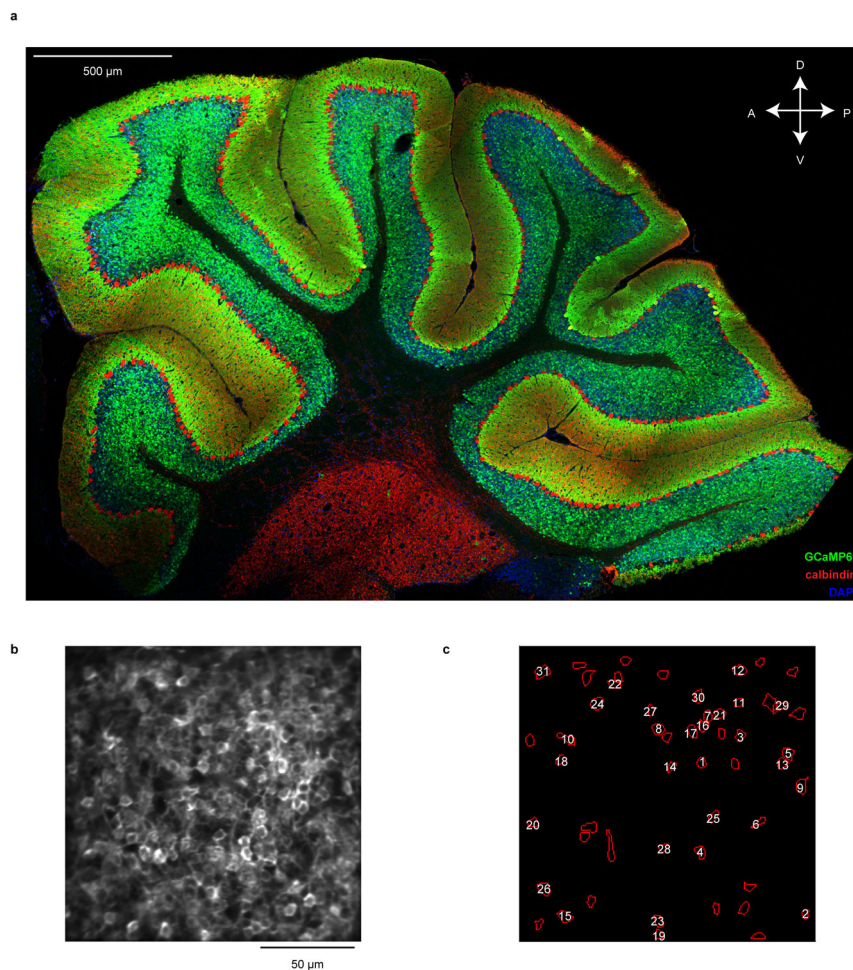
Extended Data Fig. 1. Ca^{2+} imaging in cerebellar granule cells
a, Parasagittal section of the cerebellum of a transgenic mouse (*Math1-Cre* / *CAG-lox-stop-lox-tTA* / *TRE-lox-stop-lox-GCaMP6f*) used for *in vivo* two-photon Ca^{2+} imaging. GCaMP6f expression (green) is widespread throughout most granule cells. GCaMP-expressing somas were not detected in the molecular layer, and only rarely coincided with Purkinje cells (red). For unknown reasons, granule cell expression is substantially reduced in lobules IX and X. A, anterior; P, posterior; D, dorsal; V, ventral. **b,** Mean two-photon fluorescence image for the session shown in Fig. 1c,d. c, Location of all identified active cerebellar granule cells in the field of view in **b** (*n* = 53 cells total). Numbered cells indicate the example cell traces shown in Fig. 1d, counting from the bottom to the top.
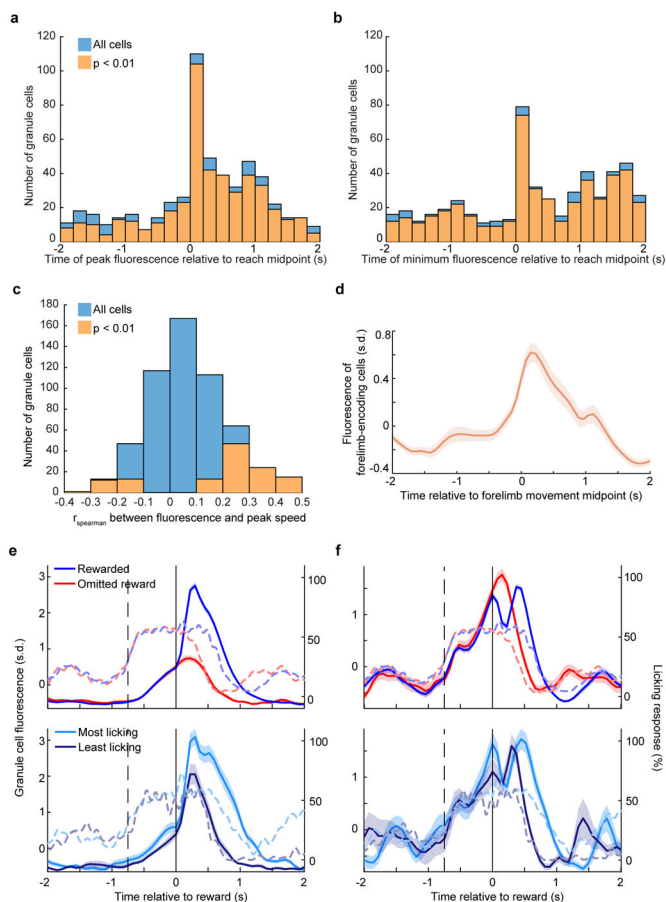
Extended Data Fig. 2. Granule cells encode movement in a forelimb movement operant task
a, b, Distribution of times of peak (**a**) or minimum (**b**) trial-averaged fluorescence response relative to reach midpoint (blue histograms, $n = 561$ total neurons from 6 mice). Orange histograms denote the subset of cells whose peak (**a**) or minimum (**b**) trial-averaged fluorescence modulation was significant. 85% of cells exhibited significant positive modulation, while 90% of cells exhibited significant negative modulation, at a point between −2 to 2 s relative to forelimb movement. To compute significance we compared observed peak and minimum fluorescence, to fluorescence for randomized datasets (**Methods**). **c,** For each cell we computed the Spearman correlation coefficient between single-trial fluorescence (mean from −0.1 to +0.3 s relative to movement midpoint) and peak movement velocity. Histogram denotes distribution of Spearman coefficients across neurons ($n = 561$ total neurons from 6 mice). Neurons correlated with $p < 0.01$ (permutation test) are shown in orange. **d,** Mean movement-aligned fluorescence of granule cells whose single-trial fluorescence correlated significantly with peak movement speed, shown in **c** ($n = 111$ neurons with $p < 0.01$ for correlation coefficients, shown in orange in **c**). **e, f,** Two example granule cells that encode licking. For these cells, response differences between reward outcomes (top row, examples) can be explained by the encoding of the licking response on rewarded trials (bottom row, 25 trials with the most and least licking from 0.1 to 1 s), $n = 209$ rewarded and 68 omitted reward trials. Dashed vertical lines denote average time of

forelimb movement midpoint, solid vertical line denotes time of reward. In this and all subsequent figures, shaded regions denote s.e.m.

Extended Data Fig. 3. Granule cell reward responses during the operant task
a, b, Fluorescence response of all granule cells recorded from three experiments in lobules VIa, VIb, and simplex from one example mouse on rewarded trials and omitted reward trials. Each row shows the trial-averaged response of a single neuron. Dashed vertical line denotes the average forelimb movement midpoint; solid vertical line denotes time of reward delivery. Many more neurons appear to respond preferentially following omitted reward than reward delivery ($n = 188$ neurons). **c–e,** Average reward-aligned fluorescence of all reward-preferring cells (**c**), omitted reward-preferring cells (**d**), and reward anticipation cells (**e**), from all mice and lobules during forelimb movements ($n = 31$ reward cells, 69 reward omission cells, 50 reward anticipation cells from 13 forelimb movement sessions in 6 mice). See **Methods** for cell identification criteria. **f, g,** Comparison of the cohort of mice that performed the operant task with briefer delay periods (**f**, $n = 6$ experiments in 3 mice with delay between the end of forelimb movement and reward delivery = 0.6 s and delay between reward delivery and manipulandum handle return = 2 s), or longer delay periods (**g**, $n = 7$ experiments in 3 mice with reward delay = 0.8 s and post-reward delay = 3.5 s). Top, prevalence of reward response types as fraction of total neurons (error bars denote counting error). Bottom, average movement and licking behavior across mice on each task version. Results did not differ substantially between the two task versions and thus all data were pooled for all analyses aside from these figure panels. Across all mice, 50% of peak licking rise from baseline was reached in anticipation 0.8 ± 0.04 s before reward. Licking was

prolonged following reward compared to omitted reward ($p = 4 \times 10^{-4}$ Wilcoxon rank sum test, $n = 6$ mice; licking declined to half of its anticipatory level by 1.4 ± 0.14 s following reward compared to 0.7 ± 0.08 s following omitted reward). **h,** Venn diagram illustrating multiplexed representations in granule cells. Relative areas are true to observed cell proportions. Corresponding counting errors for reward-related cell classifications are provided in Fig. 2h. For forelimb speed cells, counting error was 1.7%. The prevalence of multiple representations in a granule cell matched predictions of independent probabilities of each representation (1.1% of cells encode reward and forelimb speed, 2% encode reward omission and forelimb speed, and 2.3% encode reward anticipation and forelimb speed, compared to the independence null hypothesis of 1.1%, 2.4%, and 1.8%, respectively).
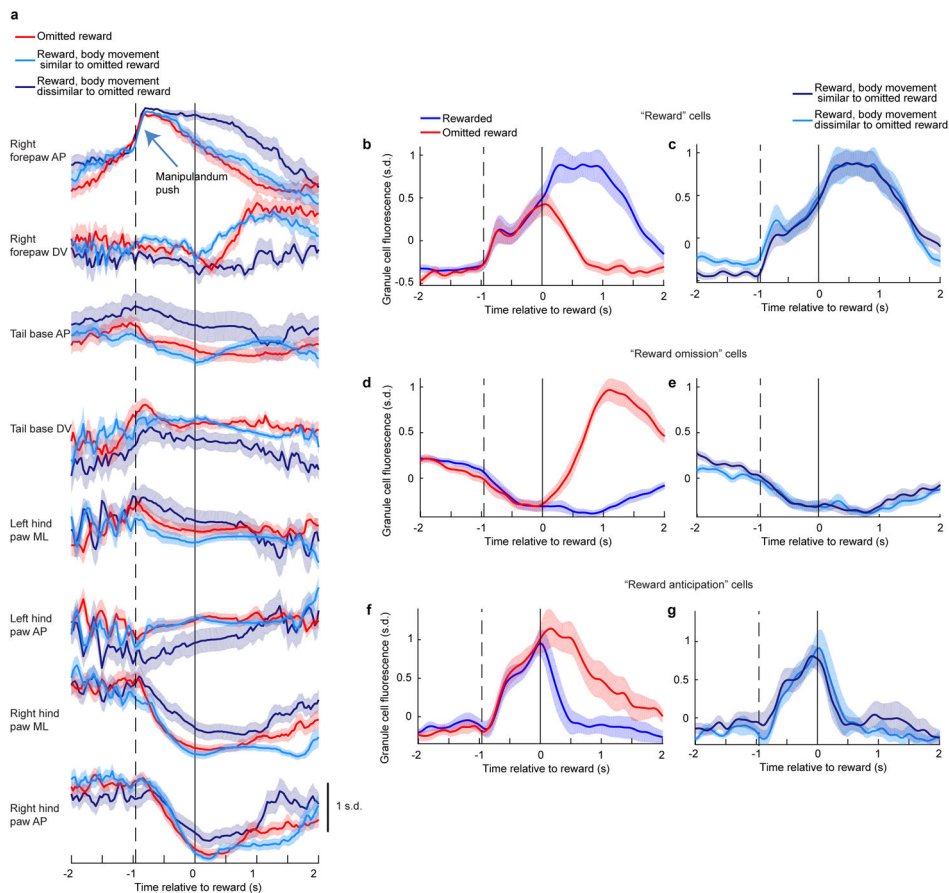

Extended Data Fig. 4. Body movement does not explain reward signaling in granule cells
We placed mice ($n = 3$) in a clear tube during imaging experiments and recorded video of their body movement from the right side and from underneath the animal (Video S2). a, For an example mouse, we computed the average body trajectory for each trial type: omitted reward, and the 25 trials most similar or most dissimilar to omitted reward body motion (**Methods**). AP, anterior-posterior, DV, dorsal-ventral, ML, medial-lateral. Motion on reward-similar-to-omitted-reward trials more closely matched motion on omitted reward trials than did motion on reward-dissimilar-to-omitted-reward trials. **b–g,** For reward cells, reward omission cells, and reward anticipation cells, despite robust signaling of reward

outcome (**b, d, f**), higher similarity of body trajectory on rewarded trials to that on omitted reward trials did not result in cellular responses more similar to those on omitted reward trials (**c, e, g**), $n = 21$ reward cells, 41 reward omission cells, 10 reward anticipation cells (from $n = 201$ total granule cells analyzed from 3 mice). Therefore body movement is unlikely to be the cause of granule cell reward signaling. Dashed vertical lines denote average time of forelimb movement midpoint.
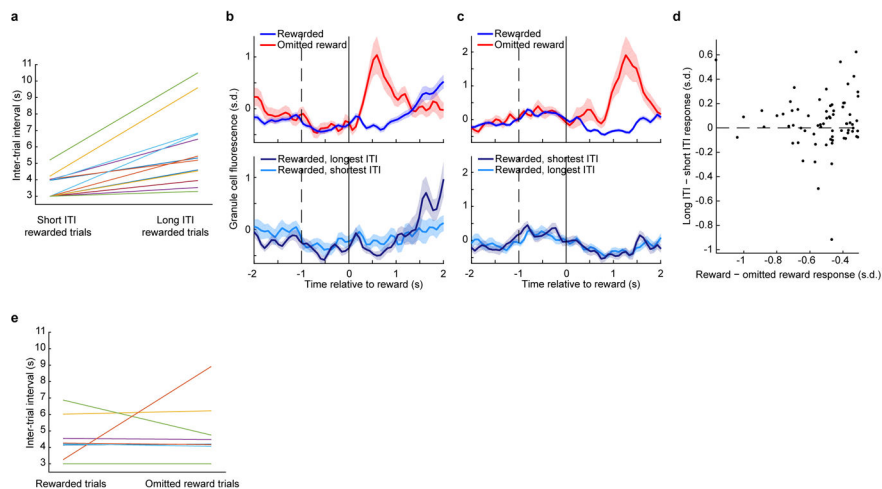
Extended Data Fig. 5. Inter-trial interval (ITI) analyses do not support that reward omission responses encode preparation for the next trial

One alternative explanation for the response of "reward omission" cells on omitted reward trials is that, following a trial in which the mouse does not receive a reward, the mouse is more anxious to begin the next trial and therefore quickly begins preparing for the next forelimb movement. If "reward omission" cells were actually just "next trial preparation cells," then these putative earlier motor preparations on omitted reward trials would elicit a larger response. That these cells exhibit on average no response following *rewarded* trials could reflect mice choosing to wait before preparing the next trial following reward delivery compared to omitted reward. We tested two predictions of this hypothesis. First, we reasoned that if, following a rewarded trial, mice choose to initiate the next trial very quickly, putative "next trial preparation cells" should exhibit increased response, as they do following omitted reward. By contrast, on rewarded trials after which mice wait before initiating the next trial, the lack of motor preparations should result in a smaller response in "next trial preparation cells." Second, if mice were substantially more anxious to initiate the next trial following omitted reward, ITIs following omitted reward trials should be shorter compared to ITIs following rewarded trials. **a–d,** To test the first prediction, we leveraged natural variability in mouse behavior to identify rewarded trials after which mice initiated the next movement very quickly and therefore had the shortest ITI (the earliest time that the robot returns to permit the mouse to initiate the next trial is 2 or 3.5 s following the previous reward, each in 3 mice). For each imaging session, we identified groups of 25 rewarded trials with the longest ITIs and those with the shortest. These two groups of rewarded trials had substantially different ITIs, indicating that their next-trial-preparatory movements varied substantially (mean ITI for the "short" group was 3.6 s, for the "long" group 5.8 s, $n = 13$ sessions). Each line in (**a**) represents one imaging session. Yet despite the large difference in

next-trial-preparations in these two groups of trials, reward omission cells remained silent in both cases, despite robust responses on omitted reward trials (two cells from two example mice in **b, c; b** is the example cell from Fig. 2B, $n = 97$ rewarded and 25 omitted reward trials; for **c**, $n = 129$ rewarded and 34 omitted reward trials). Across all 69 identified reward omission cells **(d)**, there was no tendency for a stronger response when mice initiated the next trial quickly compared to when they waited before doing so. Thus the prediction that putative "next trial preparation cells" respond to earlier next trial preparations was not borne out. **e,** To test the second prediction that mice were preparing the next trial more quickly following omitted reward trials, thereby leading to greater preparatory movements encoded by putative "next trial preparation cells," we grouped ITIs according to whether they followed rewarded or omitted reward trials within each imaging session (indicated by each line). We found no consistent difference in how long mice chose to wait before initiating the next trial following either reward or omitted reward trials ($p = 0.93$ Wilcoxon signed-rank test, $n = 13$ imaging sessions from 6 mice). Thus, the second prediction was also not borne out. Taken together, the selective response of reward omission cells to omitted reward trials is more likely to be related to reward than next-trial-preparations.
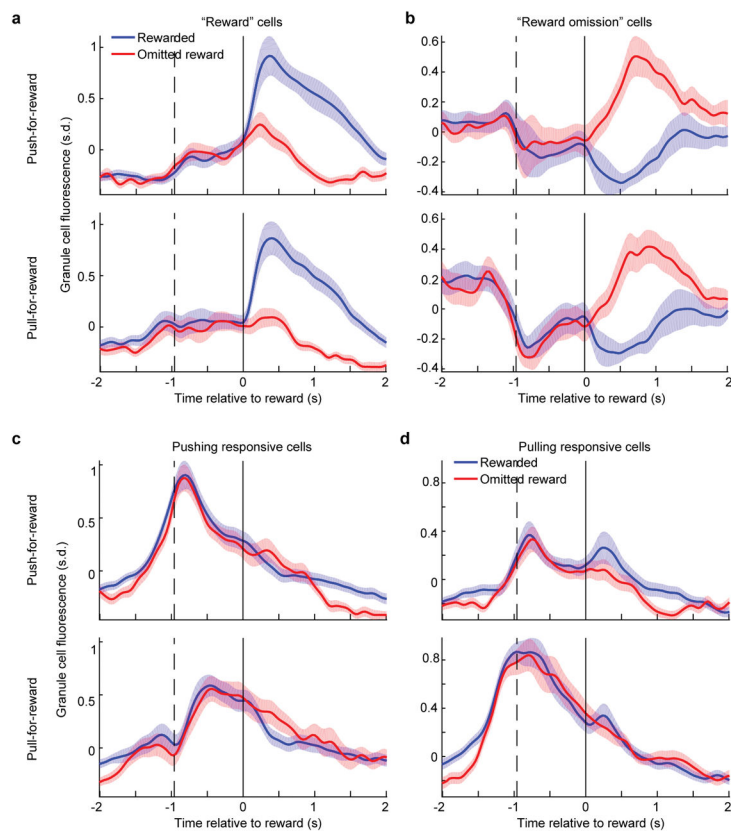
Extended Data Fig. 6. Granule cell responses in alternate push-for-reward and pull-for-reward trials

a,b, We identified reward **(a)** and reward omission cells **(b)** based only on push-for-reward trials and computed their average response (top). We then computed the average response of these same cells on pull-for-reward trials (bottom) and found they were highly preserved (*n*

= 23 reward omission and 30 reward cells from 4 mice). **c,d,** For comparison, we identified cells that responded to forelimb movement based only on push-for-reward trials ($n = 25$ pushing cells) and computed their average response (**c**, top). We then compared this to the average response of these cells on pull-for-reward trials (**c**, bottom) and found it was substantially weaker. Similarly, when we identified cells responsive to forelimb motion based only on pulling trials (**d,** bottom, $n = 42$ pulling cells) the response of these cells on pushing trials (top) was substantially weaker. This indicates that movement responses **(c,d)** are substantially less generalized across sensorimotor contexts than reward signaling **(a,b)**. Dashed vertical lines indicate average time of forelimb pushing or pulling movement midpoint, solid line denotes time of reward.
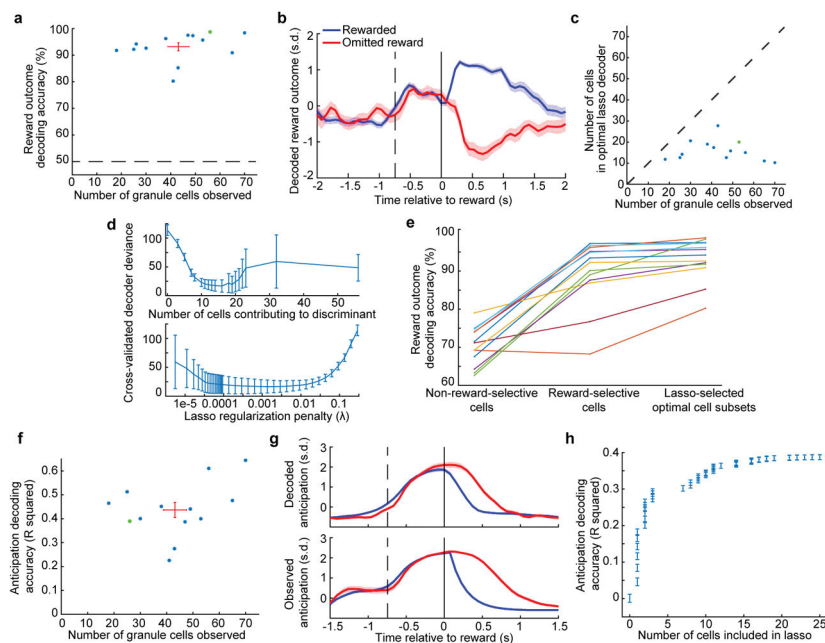
**Extended Data Fig. 7. Granule cell ensembles discriminate reward outcome and decode behavior**
a, We sought to discriminate reward from omitted reward trials by linearly decoding ensemble granule cell activity. We first used lasso logistic regression to identify the minimal set of neurons that achieve optimal decoding accuracy for each imaging session. For this minimal set, we fit a linear discriminant to the mean fluorescence from 0 to 1 s of each cell on each trial. We tabulated the discriminant's cross-validated accuracy for each imaging session (dots). Red bars denote mean ± s.e.m. across sessions ($n = 13$ experiments in 6 mice; **Methods**). Dashed line denotes chance accuracy. Green dot denotes example session used in (**b**) and (**d**). **b,** For an example imaging session, we applied the discriminant weighting to the time-varying cellular responses on each trial and averaged the output across all rewarded and omitted reward trials ($n = 56$ neurons, 64 rewarded trials, 19 omitted reward trials). The large separation following reward vs reward omission reflects accurate neural decoding. **c,** In general, the lasso determined that optimal cross-validated decoding was achieved with a minority of recorded cells. **d,** For the example session shown in **b**, we examined how cross-validated reward outcome decoding accuracy varied with the number of neurons included in the decoder, by varying the lasso penalty. We found that optimal performance was achieved

with a subset of cells, indicating that larger groups of cells resulted in some overfitting (**Methods**). Error bars indicate s.e.m. from cross-validation. **e,** To determine the importance of reward-selective cells in decoding, we fit linear discriminants while excluding reward-selective cells (> 0.2 s.d. absolute fluorescence difference between reward conditions averaged from 0.1 to 1 s), as well as discriminants using only reward selective cells. We compared these decoders' performance to the optimal subset determined from lasso regression, and found that reward-selective cells recover most of the optimal decoder performance. Each line represents one imaging session ($n = 13$ sessions). **f,** We reasoned that if granule cells can signal the mouse's reward anticipation, it should be possible to use neuronal activity to decode this anticipation on a moment-by-moment basis. We therefore defined the mouse's instantaneous anticipation state to be its lick rate (in 200 ms bins) until it received reward, in which case we defined anticipation to decline to zero (**Methods**). For each imaging session, we performed a linear regression to approximate the mouse's time-varying reward anticipation behavior by using the time-varying fluorescence of all cells. We quantified regression performance as the R^2 fraction of variance in reward anticipation that was accounted for by the regression output (using cross-validation). Each dot denotes a single imaging session. Red bars denote average decoder performance. Green dot denotes example session used in (**g, h**). **g,** For one example session, concurrence between decoded anticipation (top) and observed anticipation according to the definition in **f** (bottom), from a single imaging session averaged across all rewarded (blue) and omitted reward trials (red) ($n = 26$ neurons, 171 rewarded trials, 54 omitted reward trials). **h,** For the example session in **e**, we performed a lasso regression that penalizes non-zero weights on cells, to restrict the number of cells used for decoding. We varied the penalty from zero to maximum in order to determine how accuracy scales with the number cells (**Methods**). Reward anticipation decoding accuracy (using cross-validation) reached nearly asymptotic levels with typically ~10–20 included neurons. Error bars indicate s.e.m. from cross-validation.
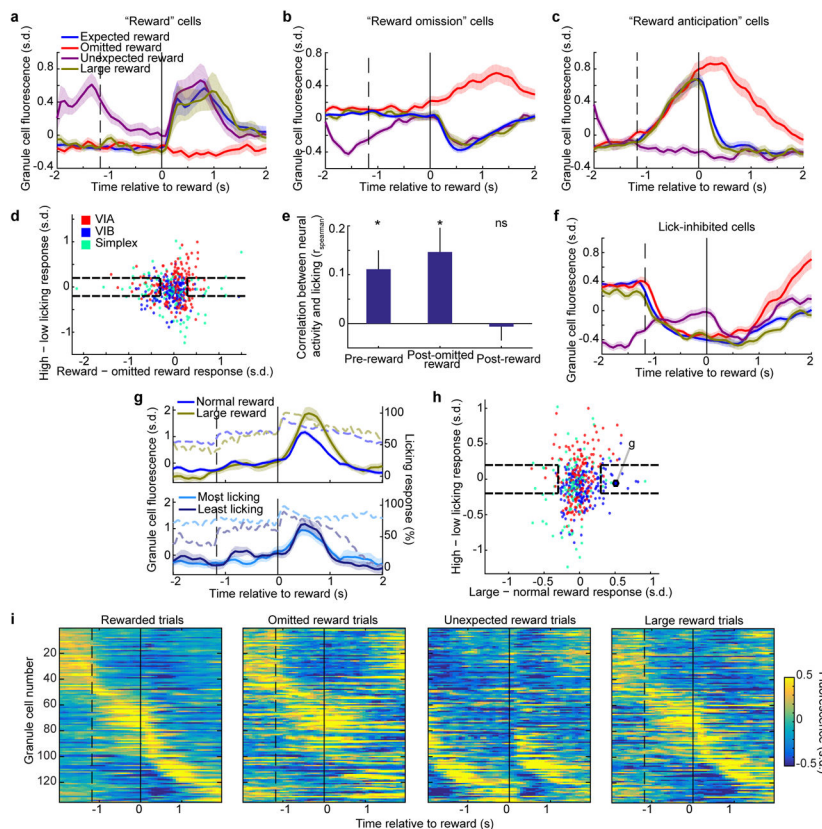
Extended Data Fig. 8. Granule cell reward responses during a Pavlovian tone–reward task
a–c, Average reward-aligned fluorescence of all reward preferring cells (**a**), reward omission cells (**b**), and reward anticipation cells (**c**), from all mice and lobules during the tone–reward task ($n = 23$ reward, 42 reward omission, and 25 reward anticipation cells from 11 experiments in 5 mice). On average, reward anticipation neurons were silent following unexpected reward ($p = 0.24$ Wilcoxon signed-rank test; mean fluorescence change of -0.05 ± 0.05 s.d. comparing $0 - 1$ s to -0.25 to -0.05 s relative to unexpected reward, $n = 25$ neurons). Reward omission cells did not distinguish expected from unexpected reward ($p = 0.48$ Wilcoxon signed-rank test comparing mean fluorescence from 0 to 1 s, $n = 42$ reward omission neurons). Dashed vertical lines indicate time of tone onset. **d,** Scatter of response properties of individual neurons (*colored dots*) showing reward preference (*x*-axis) versus licking sensitivity (*y*-axis) during the tone–reward task ($n = 450$ neurons). **e,** Single-trial correlation between licking and activity of each reward anticipation neuron either before reward delivery, after reward omission, or after reward delivery, averaged across all reward anticipation neurons during the Pavlovian task ($n = 25$ reward anticipation neurons from 11 experiments in 5 mice; $p = 0.02$ pre-reward, $p = 0.015$ post-omitted reward, $p = 0.72$ post-reward; Wilcoxon signed-rank test). As during forelimb movements, reward anticipation neurons correlate with licking only when licking represents anticipation. Following reward, when anticipation ceases, licking exerts no effect on activity. **f,** A subset of cells exhibited decreased fluorescence following the tone. To determine what these cells might be encoding, we identified all such neurons (**Methods**) and examined their response on the various trial types. We determined that these cells remain inhibited while the mouse is licking, beginning

with anticipatory licking through reward consumption ($n = 20$ cells from 5 mice). Importantly, on unexpected reward trials, these neurons are also inhibited. This is unlike reward anticipation cells in (**c**) that cease to be active following reward delivery and also remain silent on surprise reward trials. Thus cells inhibited by licking are more classically sensorimotor. **g,** First row compares trials with a normal sized reward to randomly interspersed trials with an larger reward. Second row compares normal reward trials with the most and least reward licking. **h,** Plot of each cell's response difference between normal and large rewards (*x*-axis) and preference for licking on normal reward trials (*y*-axis). Dashed boxes indicate reward magnitude sensitive neurons without substantial licking sensitivity. Example cell from **g** is outlined. **i,** Each row shows the trial-averaged Ca^{2+} response of a single neuron. Cells in each panel (trial types indicated above) are ordered identically based on their response on rewarded trials ($n = 135$ neurons from three sessions in lobules VIa, VIb, and simplex from an example mouse).
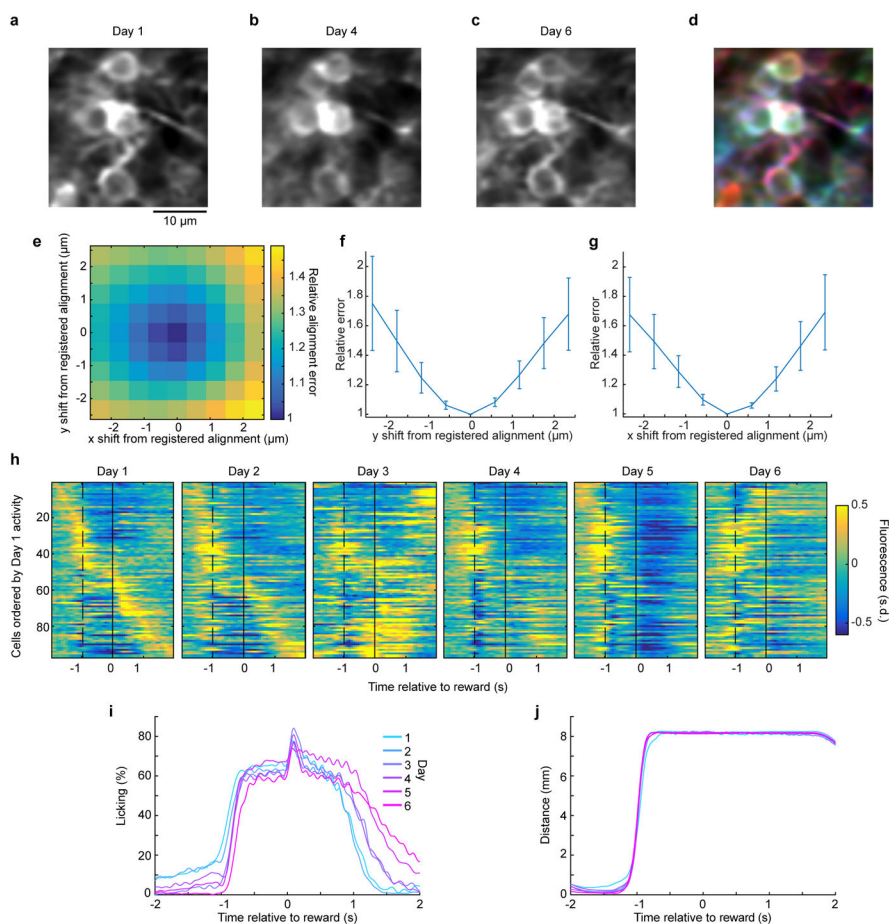
Extended Data Fig. 9. Chronic imaging cell tracking and registration
a–c, Magnified view of mean two-photon image from the regions shown in Fig. 4a on Day 1 (**a**), Day 4 (**b**), and Day 6 (**c**). **d**, Colorized overlay of the images in **a–c** in red, blue and green. We rigidly aligned the mean fluorescence image on each day to that of the final day using TurboReg[37], resulting in unambiguous alignment of visible morphological features of individual granule cells. **e**, To quantify any ambiguity in the image registration we offset our

images from optimal alignment by small amounts. For one example session, we quantified the image concordance of Day 1 and Day 6, as a function of displacing the Day 1 image in the x and y directions relative to the registered optimum at zero (sum squared pixel differences between days, normalized to the registered optimum). There is a clear trough in the alignment error at the optimum, demonstrating that even slight, submicron misalignments are easily detected by image registration. Thus, there is little appreciable ambiguity in the alignment procedure. **f, g,** Average alignment error as a function of image displacement from the registered optimum, as in **e**, here averaged across all sessions and mice ($n = 15$ alignments from 3 mice). Error bars denote s.e.m. across alignments. Even the smallest, submicron, single pixel displacements result in significantly higher alignment error than the registered optimum ($p = 4.4 \times 10^{-6}$ and 5.8×10^{-5} for one-pixel x and y misalignments respectively, Wilcoxon signed-rank test). **h,** Mean fluorescence response of all neurons for the example mouse shown in Fig. 4c, here ordered by their Day 1 activity peak response time ($n = 97$ neurons). **i–j,** Change over the 6 days of the imaging study in licking behavior (**i**) and forelimb movement behavior (**j**) for the mouse in (**h**). Gross changes in motor behavior were relatively modest over the days of the imaging study (**Methods**).
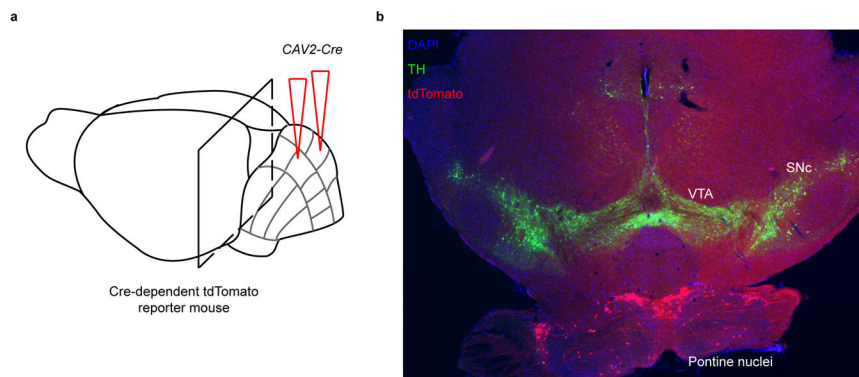
Extended Data Fig. 10. Granule cell reward responses unlikely result from a direct midbrain dopaminergic projection to the cerebellar cortex

Previous literature on the topic of dopamine in the cerebellum has been controversial, with some anatomical tracing studies suggesting a projection to cerebellar cortex from ventral tegmental area (VTA)[39,40], while others failed to find such a projection[41]. Some studies identified the presence of dopamine in the cerebellar cortex directly[42–44], yet a major confound arises due to the large noradrenergic projection to the cerebellum from the locus coeruleus, as dopamine is a precursor to norepinephrine[45]. To determine whether our widespread reward-related signals were likely to be driven by a direct dopaminergic projection, we traced the inputs to the cerebellar cortex using viral methods. **a,** Schematic. We injected *CAV2-cre*, cre recombinase expressed from canine adenovirus-2 known to robustly infect axons and their terminals in many neuronal types[46] including dopaminergic neurons specifically[47,48], into the cerebellar cortex of a highly sensitive cre-reporter Ai14 transgenic mouse. Thus any neuron in a region presynaptic to the cerebellar injection site infected by CAV2 will express tdTomato. We injected either the vermis of Lobule VI ($n = 3$ mice) or for comparison also the hemisphere lobule crus I (1 mouse). **b,** We stained serial coronal brain sections for tyrosine hydroxylase (TH, a marker for dopaminergic neurons)

and examined the distribution of input cells in the midbrain. In all 4 mice examined (sixty-four 40- or 60-micron sections encompassing all midbrain dopamine neurons), we did not find any VTA or substantia nigra pars compacta (SNc) dopamine neurons projecting to the cerebellar cortex. As a positive control, we noted that all mice exhibited robust tdTomato expression in known inputs to the cerebellum such as the pontine nuclei shown above. To exclude the unlikely possibility that putative VTA dopamine neurons that project to the cerebellum cannot take up CAV2 efficiently, we also performed an experiment where we injected *AAVretro-EF1a-FLPo*, a virus that robustly infects axonal terminals[49], into cerebellar lobule VI of a mouse that expresses FLP-dependent tdTomato, and again did not find tdTomato+ neurons in the VTA or SNc, but abundant tdTomato+ neurons in pontine nuclei (data not shown). Thus if a direct midbrain dopaminergic projection to the cerebellum exists, it must be very sparse, and therefore unlikely to drive the very large and widespread reward-related signals in our granule cell imaging data.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

1. Herculano-Houzel S. Coordinated Scaling of Cortical and Cerebellar Numbers of Neurons. Front Neuroanatom. 2010; 4:12.

2. Marr D. A theory of cerebellar cortex. J Physiol. 1969; 202:437–470. [PubMed: 5784296]

3. Albus JS. A theory of cerebellar function. Math Biosci. 1971; 10:25–61.

4. Fujita M. Adaptive filter model of the cerebellum. Biol Cybern. 1982; 45:195–206. [PubMed: 7171642]

5. Rancz EA, et al. High-fidelity transmission of sensory information by single cerebellar mossy fibre boutons. Nature. 2007; 450:1245–1248. [PubMed: 18097412]

6. Huang CC, et al. Convergence of pontine and proprioceptive streams onto multimodal cerebellar granule cells. eLife. 2013; 2:e00400. [PubMed: 23467508]

7. Ito M. Control of mental activities by internal models in the cerebellum. Nat Rev Neurosci. 2008; 9:304–313. [PubMed: 18319727]

8. Strick PL, Dum RP, Fiez JA. Cerebellum and nonmotor function. Annu Rev Neurosci. 2009; 32:413–434. [PubMed: 19555291]

9. Stoodley CJ, Valera EM, Schmahmann JD. Functional topography of the cerebellum for motor and cognitive tasks: an fMRI study. NeuroImage. 2012; 59:1560–1570. [PubMed: 21907811]

10. Tsai PT, et al. Autistic-like behaviour and cerebellar dysfunction in Purkinje cell Tsc1 mutant mice. Nature. 2012; 488:647–651. [PubMed: 22763451]

11. Bengtsson F, Jörntell H. Sensory transmission in cerebellar granule cells relies on similarly coded mossy fiber inputs. Proc Natl Acad Sci. 2009; 106:2389–2394. [PubMed: 19164536]

12. Bing YH, Zhang GJ, Sun L, Chu CP, Qiu DL. Dynamic properties of sensory stimulation evoked responses in mouse cerebellar granule cell layer and molecular layer. Neurosci Letters. 2015; 585:114–118.

13. Ishikawa T, Shimuta M, Häusser M. Multimodal sensory integration in single cerebellar granule cells in vivo. eLife. 2015; 4:e12916. [PubMed: 26714108]

14. Powell K, Mathy A, Duguid I, Häusser M. Synaptic representation of locomotion in single cerebellar granule cells. eLife. 2015; 4:e07290.

15. Coltz JD, Johnson MT, Ebner TJ. Cerebellar Purkinje cell simple spike discharge encodes movement velocity in primates during visuomotor arm tracking. J Neurosci. 1999; 19:1782–1803. [PubMed: 10024363]

16. Waelti P, Dickinson A, Schultz W. Dopamine responses comply with basic assumptions of formal learning theory. Nature. 2001; 412:43–48. [PubMed: 11452299]

17. Galliano E, et al. Silencing the Majority of Cerebellar Granule Cells Uncovers Their Essential Role in Motor Learning and Consolidation. Cell Reports. 2013; 3:1239–1251. [PubMed: 23583179]

18. Coltz JD, Johnson MTV, Ebner TJ. Cerebellar Purkinje Cell Simple Spike Discharge Encodes Movement Velocity in Primates during Visuomotor Arm Tracking. J Neurosci. 1999; 19:1782–1803. [PubMed: 10024363]

19. Medina JF, Lisberger SG. Links from complex spikes to local plasticity and motor learning in the cerebellum of awake-behaving monkeys. Nat Neurosci. 2008; 11:1185–1192. [PubMed: 18806784]

20. Brooks JX, Cullen KE. The primate cerebellum selectively encodes unexpected self-motion. Curr Biol. 2013; 23:947–955. [PubMed: 23684973]

21. Schultz W. Predictive Reward Signal of Dopamine Neurons. J Neurophysiol. 1998; 80:1–27. [PubMed: 9658025]

22. Cohen JY, Haesler S, Vong L, Lowell BB, Uchida N. Neuron-type-specific signals for reward and punishment in the ventral tegmental area. Nature. 2012; 482:85–88. [PubMed: 22258508]

23. Schultz W, Apicella P, Scarnati E, Ljungberg T. Neuronal activity in monkey ventral striatum related to the expectation of reward. J Neurosci. 1992; 12:4595–4610. [PubMed: 1464759]

24. Tremblay L, Schultz W. Reward-Related Neuronal Activity During Go-Nogo Task Performance in Primate Orbitofrontal Cortex. J Neurophysiol. 2000; 83:1864–1876. [PubMed: 10758098]

25. Miyazaki K, Miyazaki KW, Doya K. Activation of Dorsal Raphe Serotonin Neurons Underlies Waiting for Delayed Rewards. J Neurosci. 2011; 31:469–479. [PubMed: 21228157]

26. Matsumoto M, Hikosaka O. Representation of negative motivational value in the primate lateral habenula. Nature neuroscience. 2009; 12:77–84. [PubMed: 19043410]

27. Kawai T, Yamada H, Sato N, Takada M, Matsumoto M. Roles of the Lateral Habenula and Anterior Cingulate Cortex in Negative Outcome Monitoring and Behavioral Adjustment in Nonhuman Primates. Neuron. 2015; 88:792–804. [PubMed: 26481035]

28. Chen TW, et al. Ultrasensitive fluorescent proteins for imaging neuronal activity. Nature. 2013; 499:295–300. [PubMed: 23868258]

29. Madisen L, et al. Transgenic mice for intersectional targeting of neural sensors and effectors with high specificity and performance. Neuron. 2015; 85:942–958. [PubMed: 25741722]

30. Li L, et al. Visualizing the distribution of synapses from individual neurons in the mouse brain. PloS one. 2010; 5:e11503. [PubMed: 20634890]

31. Matei V, et al. Smaller inner ear sensory epithelia in Neurog1 null mice are related to earlier hair cell cycle exit. Dev Dynam. 2005; 234:633–650.

32. Ben-Arie N, et al. Math1 is essential for genesis of cerebellar granule neurons. Nature. 1997; 390:169–172. [PubMed: 9367153]

33. Madisen L, et al. A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. Nat Neurosci. 2010; 13:133–140. [PubMed: 20023653]

34. Figielski, A., Bonev, IA., Bigras, P. 2007 IEEE International Conference on Systems, Man and Cybernetics; p. 1562-1566.

35. Lecoq J, et al. Visualizing mammalian brain area interactions by dual-axis two-photon calcium imaging. Nat Neurosci. 2014; 17:1825–1829. [PubMed: 25402858]

36. Pologruto TA, Sabatini BL, Svoboda K. ScanImage: Flexible software for operating laser scanning microscopes. Biomed Eng Online. 2003; 2:1–9. [PubMed: 12605721]

37. Thevenaz P, Ruttimann UE, Unser M. A pyramid approach to subpixel registration based on intensity. IEEE T Image Process. 1998; 7:27–41.

38. Mukamel EA, Nimmerjahn A, Schnitzer MJ. Automated analysis of cellular signals from large-scale calcium imaging data. Neuron. 2009; 63:747–760. [PubMed: 19778505]

39. Simon H, Le Moal M, Calas A. Efferents and afferents of the ventral tegmental-A10 region studied after local injection of [3H]leucine and horseradish peroxidase. Brain Research. 1979; 178:17–40. [PubMed: 91413]

40. Ikai Y, Takada M, Shinonaga Y, Mizuno N. Dopaminergic and non-dopaminergic neurons in the ventral tegmental area of the rat project, respectively, to the cerebellar cortex and deep cerebellar nuclei. Neuroscience. 1992; 51:719–728. [PubMed: 1362601]

41. Swanson LW. The projections of the ventral tegmental area and adjacent regions: a combined fluorescent retrograde tracer and immunofluorescence study in the rat. Brain research bulletin. 1982; 9:321–353. [PubMed: 6816390]

42. Dahlström A, Fuxe K, Olson L, Ungerstedt U. Ascending Systems of Catecholamine Neurons from the Lower Brain Stem. Acta Physiologica Scandinavica. 1964; 62:485–486. [PubMed: 14252583]

43. Kizer JS, Palkovits M, Brownstein MJ. The projections of the A8, A9 and A10 dopaminergic cell bodies: evidence for a nigral-hypothalamic-median eminence dopaminergic pathway. Brain Research. 1976; 108:363–370. [PubMed: 1276901]

44. Panagopoulos NT, Papadopoulos GC, Matsokis NA. Dopaminergic innervation and binding in the rat cerebellum. Neurosci Lett. 1991; 130:208–212. [PubMed: 1795884]

45. Glaser PEA, et al. Cerebellar neurotransmission in attention-deficit/hyperactivity disorder: Does dopamine neurotransmission occur in the cerebellar vermis? Journal of Neuroscience Methods. 2006; 151:62–67. [PubMed: 16451810]

46. Schwarz LA, et al. Viral-genetic tracing of the input-output organization of a central noradrenaline circuit. Nature. 2015; 524:88–92. [PubMed: 26131933]

47. Hnasko TS, et al. Cre recombinase-mediated restoration of nigrostriatal dopamine in dopamine-deficient mice reverses hypophagia and bradykinesia. Proc Natl Acad Sci. 2006; 103:8858–8863. [PubMed: 16723393]

48. Beier KT, et al. Circuit Architecture of VTA Dopamine Neurons Revealed by Systematic Input-Output Mapping. Cell. 2015; 162:622–634. [PubMed: 26232228]

49. Tervo DG, et al. A Designer AAV Variant Permits Efficient Retrograde Access to Projection Neurons. Neuron. 2016; 92:372–382. [PubMed: 27720486]
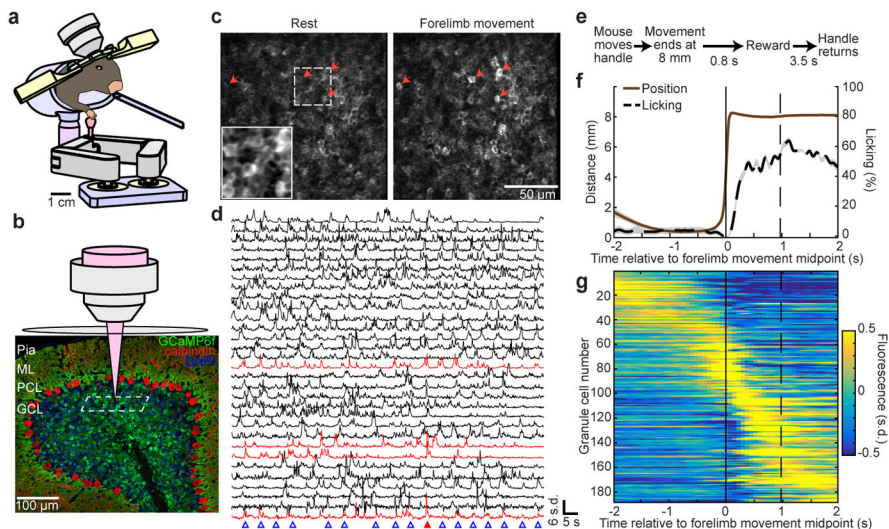
Figure 1. Two-photon Ca^{2+} imaging of cerebellar granule cells during an operant task
a, Mice voluntarily pushed a manipulandum forward for sucrose water reward. We performed Ca^{2+} imaging while recording the paw position and the mouse's licking. **b,** Confocal image of the cerebellar cortex of a transgenic mouse expressing GCaMP6f in granule cells. Calbindin immunostain for Purkinje cells in red. ML, molecular layer; PCL, Purkinje cell layer; GCL, granule cell layer. Two-photon imaging plane is schematized (dashed white box). **c,** Example *in vivo* two-photon images of cerebellar granule cells at rest and during a forelimb movement (500-ms average). Arrows denote example granule cells exhibiting fluorescence increases during this forelimb movement. Inset shows magnified view of mean fluorescence signals. **d,** Each row depicts the Ca^{2+} trace over time of one granule cell from the image in **c**. Blue triangles indicate forelimb movements. Red traces correspond to cells with red arrows in **c**. Red triangle denotes forelimb movement shown in **c**. Cells are ordered according to Extended Data Fig. 1c. e, Task structure. See Extended Data Fig. 3f for an alternative condition. **f,** Trial-averaged forelimb movement and licking (68 trials from an example mouse). Solid and dashed vertical lines denote midpoint of forelimb movement and average time of reward, respectively. **g,** Each row shows the trial-averaged Ca^{2+} response of a single neuron, with colors representing fluorescence signal in the unit of standard deviation (s.d.) from the mean (188 cells from three sessions in lobules VIa, VIb, and simplex from the mouse in **f**.). In this and all subsequent figures, shaded regions denote s.e.m.
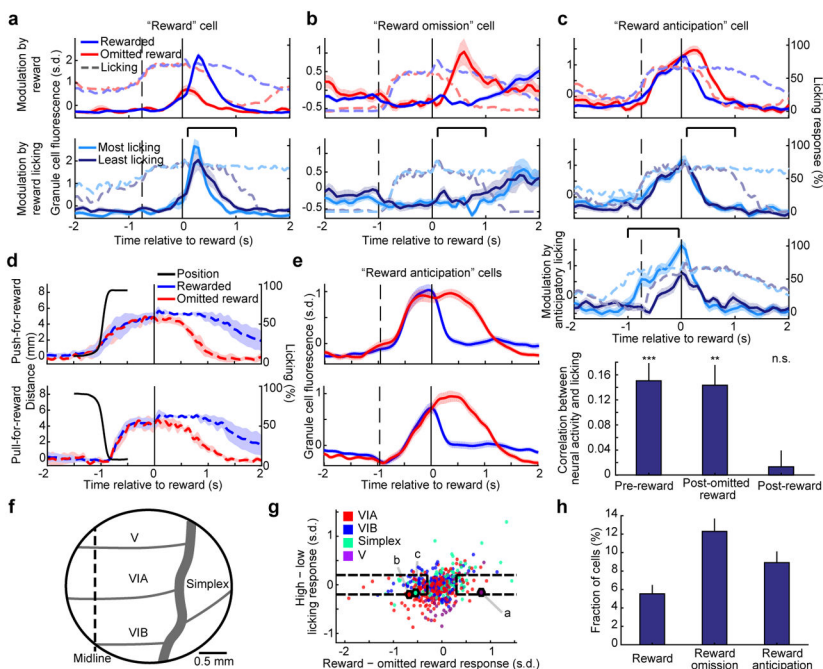
Figure 2. Granule cells encode reward context during a forelimb movement operant task
a–c, Trial averaged Ca^{2+} response (solid traces) of three example granule cells, superimposed on licking traces (dashed). Solid and dashed vertical lines denote reward onset and midpoint of forelimb movement, respectively. First row compares rewarded trials and omitted reward trials (trial numbers in **a–c**, 228, 97, 171 rewarded and 77, 25, 54 omitted reward, respectively). Second row compares rewarded trials with the most or least licking in response to reward delivery (25 of each in the bracketed period). **c,** Third row compares trials with the most or least anticipatory licking (25 of each in the bracketed period). Fourth row shows the relationship between licking and activity of all reward anticipation neurons. Bars denote the Spearman correlation between fluorescence response and licking either prior to reward delivery (−1 to −0.05 s), or following omitted reward or reward delivery (0.1 to 0.6 s). *** $p = 8 \times 10^{-6}$ pre-reward; ** $p = 5 \times 10^{-4}$ post-omitted reward; n.s. $p = 0.59$ post-reward (Wilcoxon signed-rank test; $n = 50$ reward anticipation neurons from 6 mice). **d,e,** In a modified task where mice alternated pushing-for-reward (top) with pulling-for-reward (bottom) trials, forelimb movement and licking responses are indicated as solid and dashed lines, respectively (**d**). Reward anticipation neurons classified on pushing trials (**e**, top) maintain similar responses on pulling trials (**e**, bottom), average of 41 neurons from 4 mice. **f,** Illustration of 3 mm cranial window. Grey lines represent cerebellar lobule boundaries. **g,** For each granule cell recorded during the (pushing only) operant task, we quantified the reward vs. reward omission response preference (*x*-axis; mean fluorescence response difference from 0.1 to 1 s), and the licking response preference (*y*-axis; mean response difference between trials with the most and least reward licking from 0.1 to 1 s; $n = 6$ mice, 561 cells). Colors denote lobule origin of the cells. Dashed boxes indicate neurons we classified as selective for reward or omitted reward, with minimal licking sensitivity. Example cells from **a–c** are outlined. **h,** Prevalence of reward, reward omission, and reward anticipation neurons. Reward omission excludes reward anticipation neurons.
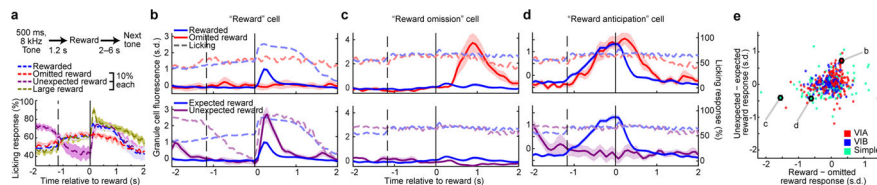
Figure 3. Granule cells encode reward context during a Pavlovian tone–reward task

a, Top, task illustration. Bottom, average licking response (11 sessions in 5 mice). **b–d,** Trial averaged response of three example granule cells (solid traces) superimposed on licking response (dashed). Dashed and solid vertical lines indicate the time of tone onset and reward delivery, respectively. First row compares rewarded trials and randomly interspersed omitted reward trials. Second row compares rewarded trials to interspersed unexpected rewards not preceded by a tone (trial numbers in **b–d**: 178, 163, 163 rewarded, 26, 24, 24 omitted reward, and 26, 24, 24 unexpected reward, respectively). **e,** Plot of each cell's response differences between rewarded and omitted reward trials (*x*-axis), and between unexpected and expected reward trials (*y*-axis). Colors denote lobule origin of the cells (450 cells). Example cells from **b–d** are outlined.
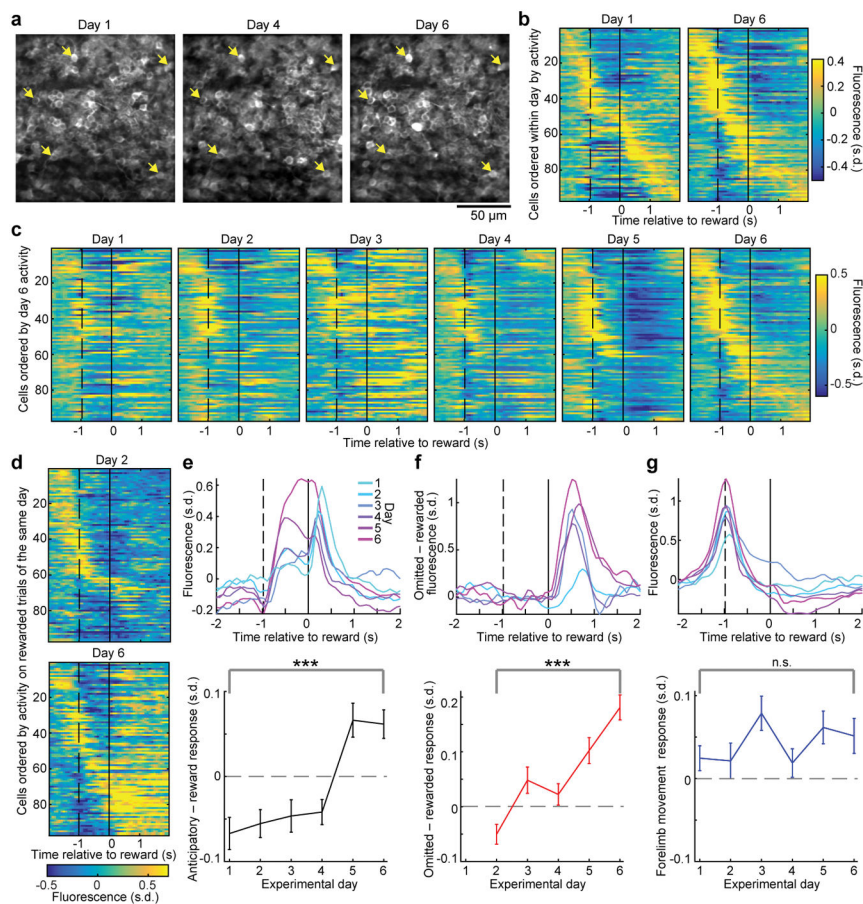
Figure 4. Emergence of reward expectation responses during forelimb movement task learning
a, Example *in vivo* two-photon mean fluorescence images of the same granule cells acquired on different days, registered to the final day (magnified in Extended Data Fig. 9). Arrows indicate example corresponding neurons across days. **b,** Average responses of all detected granule cells on rewarded trials on Day 1 and Day 6 of imaging, sorted separately for each day by time of peak response (97 neurons from an example mouse). **c,** Average response of all granule cells on rewarded trials on all six days, sorted by their Day 6 activity, for the mouse in **b. d,** Average response to omitted reward on Day 2 and Day 6, ordered by time of peak response on rewarded trials on the same days. **e–g,** *Top*, For each day, average fluorescence of the top 10% of cells across mice (24 neurons) ranked by their Day 6: **(e)** anticipatory rise in fluorescence (mean fluorescence difference between −0.25 to −0.05 s and −1.3 to −1 s), **(f)** response preference for omitted reward over reward (mean difference over 0.1 to 1 s), or **(g)** forelimb movement response (fluorescence rise during movement, −1.3 to −1 s, compared to pre-movement, −1.8 to −1.3 s). *Bottom*, summary across all neurons of changes in anticipatory responsiveness (**e**), omitted reward preference (**f**), or forelimb movement responsiveness (**g**). (***$p < 10^{-6}$; n.s. $p = 0.76$; $n = 233$ neurons from 3 mice, Wilcoxon signed-rank test).