Behavioral and molecular underpinnings of performance variability in eyeblink conditioning in male and female mice

Maria Roa Oyaga[1], Sebastiaan K.E. Koekkoek[1], Aleksandra Badura[1,2]

[1] Department of Neuroscience, Erasmus MC, 3000 Rotterdam, the Netherlands

[2] Netherlands Institute of Neuroscience, Royal Dutch Academy for Arts and Sciences, Amsterdam, the Netherlands

Abstract

The functional and molecular sources of behavioral variability in mice are not fully understood. As a consequence, the predominant use of male mice has become a standard in animal reseach, under the assumption that males are less variable than females. Similarly, to homogenize genetic background, neuroscience studies have almost exclusively used the C57BL/6 (B6) strain. Here, we examined individual differences in performance in the context of associative learning. We performed delayed eyeblink conditioning while recording locomotor activity in mice from both sexes in two strains (B6 and B6CBAF1). Further, we used a C-FOS immunostaining approach to explore brain areas involved in eyeblink conditioning across subjects, and correlate them with behavioral performance. We found that B6 male and female mice show comparable variability in this task and that females reach higher learning scores. We found a strong positive correlation across sexes between learning scores and voluntary locomotion. C-FOS immunostainings revealed positive correlations between C-FOS positive cell density and learning in the cerebellar cortex as well as multiple, previously unreported extra-cerebellar areas. We found consistent and comparable correlations in eyeblink performance and *C-fos* expression in B6 and B6CBAF1 females and males. Taken together, we show that differences in motor behavior and activity across brain areas correlate with learning scores during eyeblink conditioning across strains and sexes.

Introduction

Individuality is continuously shaped throughout life by interactions between our genes and the environment. Our behaviors, emotions and cognitive functions are distinct and unique characteristics that differentiate ourselves from the rest of the population. In humans, these are known as personality traits and have been extensively studied in the field of psychology. Although some components of animal behavioral variability have been described, questions regarding other sources that could evoke differences in behavior are still unanswered (Pfaff, 2001; Bucán and Abel, 2002; Tye et al., 2011; Leung and Jia, 2016).

For example, female mice? are considerably under-investigated in neuroscience due to the presumption that hormonal fluctuations caused by the estrous cycle might introduce non-comparable variability across sexes (Meziane et al., 2007). Clinical studies have reported that multiple neurobiological disorders show sexual dimorphism, which indicates that understanding sex differences in neural function and behavior is essential for understanding and treating diseases (McCarthy et al., 2012; Shumer, Nokoff, 2017; Smith et al., 2019). Therefore, the underrepresentation of female rodents in research has gained attention in recent years. Meta-analysis in rats and mice show that females and males exhibit comparable variability across behavioral, morphological and physiological traits, and that for most traits, female estrous cycle does not need to be considered (Simpson and Kelly, 2012; Becker et al., 2016).

In addition, the development of inbred mouse strains was initially intended to homogenize the genetic background in order to increase comparability between animals and increase the power of studies (Festing, 1999). However, the current golden standard to keep mice exclusively on a C57BL/6 background (further reffered to as B6) limits the generalization of findings (Rivera and Tessarollo, 2008; Sittig et al., 2016). When it comes to inter-strain variability, behavioral differences have been reported, yet, it has not been systematically investigated across all commonly used paradigms (Faure et al., 2017; Arnold and Newland, 2018).

Together, the limited knowledge on the sources of behavioral heterogeneity might lead to assumptions, which introduce experimental design bias (Åhlgren and Voikar, 2019). Hence, gaining a deeper understanding of

the sources of behavioral variability could give us indications on how to interpret data and ensure better reproducibility across laboratories.

In order to study how sex and strain influence mouse behavior and brain activity, a reliable and controlled paradigm is needed. We investigated this in the context of delayed eyeblink conditioning. Eyeblink conditioning is a cerebellar-dependent associative learning paradigm, in which an initially neutral, conditoned stimulus (CS, a flashing light), becomes predictive of an unconditioned stimulus (US, an air-puff to the cornea), which elicits a blink. The paradigm consists of pairing the CS with the US, and over time an association is formed where blinking is triggered by the CS alone. The newly learned association is called conditioned response (CR) (Gormezano et al., 1962).(**Fig. 1A, B**).

There is limited knowledge on performance differences due to sex in eyeblink conditioning. Studies suggest that, in humans, children show differences in learning delayed eyeblink conditioning that persist throughout adulthood. Girls show more CRs in the first five days of learning compared to boys, and women show a continuous increase in CRs compared to men (Löwgren et al., 2017). In rodents, the literature presents contradictory findings and sex differences in eyeblink conditioning remain to be further studied. For example, in rats, stress seems to enhance eyeblink conditioning in males but hinders learning in females (Wood and Shors, 1998). However, in rabbits, males and females show similar conditioning profiles but females seem to adapt faster to stress (Schreurs et al., 2018). Finally, in mice, although females and males reach similar learning scores in a 10-day trace eyeblink conditioning paradigm, females show increased CRs compared to males in the first five days of learning (Rapp et al., 2021).

Studies have shown that this form of associative learning most likely relies on Purkinje cell capacity to precisely time the CS signals coming from the pons via parallel fibers, and the CS information from the inferior olive via climbing fibers (Heiney et al., 2014a; ten Brinke et al., 2015). The CR signal leaves the cerebellum via the interposed nucleus, which ultimately connects to the muscles controlling the eyeblink reflex (Gao et al., 2016; ten Brinke et al., 2017). Several cerebellar areas modulating eyeblink conditioning have been identified in mice; lobule VI in the vermal region and crus I and simplex in the hemispheric region (Heiney, Kim, et al., 2014; Gao et al., 2016). Inactivation of lobule VI and crus I during development causes deficits in learning, indicating a crucial role

in eyeblink conditioning (Badura et al., 2018). Beyond the pontocerebellar and olivocerebellar systems, little is known about the potential involvement of other brain areas in eyeblink conditioning (Boele et al., 2010; Ruigrok, 2011; D'Angelo et al., 2016; Kratochwil et al., 2017). The amygdala has been proposed to have a role in associative learning, given its implication in fear conditioning and arousal (Lee and Kim, 2004). Specifically, lesions in the amygdala during the first days of training highly impair learning, while lesions in later stages do not appear to affect learning (Lee and Kim, 2004).

Although learning is modulated by analogous brain circuits across mice, studies often do not show data of animals that do not reach profficient learning scores (Fonnesu and Kuczewski, 2019). Genetically homogeneous groups are expected to have similar learning rates, consequently, animals that deviate from their group mean are frequently classified as outliers and removed from analysis (Osborne and Overbay, 2004). Alteratively, data is gathered only from mice that reach a certain performace threshold (Rousselet and Pernet, 2012). This severly limits our understanding of behavioral variability. Possible mechanisms underlying learning differences could be arousal levels and locomotor activity, which both influence cortical function (McGinley et al., 2015; Vinck et al., 2015; Williamson et al., 2015), or stress levels which can affect neuronal firing in the deep cerebellar nuclei (DCN) and hippocampus (Joëls, 2009; Schneider et al., 2013). In the cerebellum, although locomotion modulates activity in the cortex, the relevance of this modulation is still not fully understood (Ozden et al., 2012; Hoogland et al., 2015; Powell et al., 2015). A recent study showed that imposed locomotor activity enhances eyeblink learning by increased activation of the mossy fiber pathway to the cerebellar cortex (Albergaria et al., 2018).

In this study we investigate the effect of sex in behavioral variability by employing eyeblink conditioning and voluntary locomotion to quantify behavioral differences in B6 mice. Additionally, we explore brain regions that may have a modulatory role in eyeblink conditioning by utilizing *C-fos* expression as a proxy for neural activity during learning. Finally, to study inter-strain variability, we also examine eyeblink learning profiles and *C-fos* expression in B6CBAF1 mice.

Materials and Methods

Animals

All experiments were performed in accordance with the European Communities Council Directive. All animal protocols were approved by the Dutch National Experimental Animal Committee (DEC). C57BL/6 mice were ordered from Charles River (n = 16 males; n = 16 females), and B6CBAF1 mice from Janvier (n = 7 males; n = 9 females). Mice were group-housed and kept on a 12-hour light-dark cycle with ad libitum food and water. All procedures were performed in male and female mice approximately 8-12 weeks of age.

Eyeblink pedestal placement surgery

Mice were anesthetized with isoflurane and oxygen (4% isoflurane for induction and 2-2.5% for maintenance). Body temperature was monitored during the procedure and maintained at 37°C. Animals were fixed in a stereotaxic device (Model 963, David Kopf Instruments, Tujunga CA, USA). The surgery followed previously described standard procedures for pedestal placement (Gao et al., 2016; ten Brinke et al., 2017). In short, the hair on top of the head was shaved, betadine and lidocaine was applied on the skin and an incision was done in the scalp to expose the skull. The tissue on top of the skull was removed and the skull was kept dry before applying Optibond™ prime adhesive (Kerr, Bioggio, Switzerland). A pedestal equipped with a magnet (weight ~1g), was placed on top with Charisma®, (Heraeus Kulzer, Armonk NY, USA) which was hardened with UV light. Rymadil was injected subcutaneously (5mg per kg). Mice were left under a heating lamp for recovery during at least 3 hours. Mice were given 3-4 resting days before starting experiments.

Eyeblink conditioning

Mice were habituated to the set-up (head fixed to a bar suspended over a cylindrical treadmill in a sound and light isolating chamber) for 5 days with increasing exposure (15, 15, 30, 45 and 60 min). Training started after two rest days. Twenty-eight 57BL/6 mice (n = 16 males; n = 16 females), and 16 B6CBAF1 mice (n = 7 males; n = 9 females) were trained using the standard eyeblink protocol (Brinke et al., 2015, Koekkoek et al., 2002). Ten CS-only trials of 30 ms with an inter-trial interval (ITI) of 10 ± 2 s were presented before the first training session

to acquire a baseline measurement. Mice were next trained for 5 consecutive days. Each session consisted of 20 blocks of 12 trails each (1 US only, 11 paired and 1 CS only) with an ITI of 10 ± 2 s. The CS was a 270 ms blue LED light (~ 450 nm) placed 7 cm in front of the mouse. The US was a 30 ms corneal air puff co-terminating with the CS. The puffer was controlled by a VHS P/P solenoid valve set at 30 psi (Lohm rate, 4750 Lohms; Internal volume, 30 µL, The Lee Company®, Westbrook, US) and delivered via a 27.5 mm gauge needle at 5 mm from the center of the left cornea. The inter-stimulus interval was 250 ms. Eyelid movements were recorded with a camera (Baseler aceA640) at 250 frames/s. 4 C57BL/6 mice (n = 2 males; n = 2 females), were trained using a pseudoconditioning protocol. Pseudoconditioning protocol consisted of 20 blocks of 12 trails each (1 puff only, 12 LED only) with an ITI of 10 ± 2 s. The puff and LED stimulus had the same characteristics as in the conditioning protocol. Data was analyzed with a custom written MATLAB code as previously described (Giovannucci et al., 2017; Badura et al., 2018). Traces were normalized within each session to the UR max amplitude. The CR detection window was set to 650-730 ms and CRs were only classified as such when the amplitude was equal or higher than 5% of the UR median. The CR percentage was calculated as the number of counted CRs (equal or higher than 5% of the UR median) divided by the total CS trials per session.

Locomotor activity

An infrared camera (ELP 1080P) (sampling frequency 60 frames/s) was placed in each of the eyeblink boxes and connected to an external computer (independent from the eyeblink system). The cameras were positioned at the right back corner of the chamber on top of a magnet tripod attached to a custom made metal block which allowed stable fixation. The recording angle was standardized by selecting the same reference in the field of view of each camera. Simultaneous video acquisition from the three cameras was performed in Ipi Recorder software (http://ipisoft.com/download/). Body movement recording was parallel to eyelid recording during the training sessions. The output videos (.avi format) from each mouse and session were approximately 35 min (corresponding to the length of an eyeblink session).

Locomotor activity analysis

We used DeepLabCut (DLC) to track body parts from videos (Mathis et al., 2018) (**Fig. 2**). To increase generalization, we extracted 40 frames of 4 different videos from two males and two females (total of 160 frames). Next, frames were manually labeled with 5 body parts (tail base, hip, knee, right back paw and nose). These frames were used for training the pre-trained deep neural network ResNet50 (He et al., 2016; Insafutdinov et al., 2016). Evaluation of the network was done to confirm a low error in pixels between labeled frames and predictions. Video analysis was done by using the trained network to get the locations of body parts from all mice and sessions (16 mice x 5 sessions = 80 videos). DLC output is a matrix with x and y positions in pixels and the likelihood of this position for each body part. We used this matrix to calculate distance covered and speed per body part with a custom written code (https://github.com/BaduraLab/DLC_analysis)

Tissue processing

Mice were anesthetized with 0.2 ml pentobarbital (60 mg/ml) and perfused with 0.9% NaCl followed by 4% paraformaldehyde (PFA). Given the peak time expression of *C-fos* (Chung, 2015), animals were perfused 90 min after finishing the last training session. Brains were dissected from the skull and stored in 4% PFA at room temperature (rT) for 1.5 hours. They were next changed to a 10% sucrose solution and left overnight at 4°C. Brains were embedded in 12% gelatin and 10% glucose and left in a solution with 30% sucrose and 4% PFA in PBS at rT for 1.5 hours. Next, they were transferred to a 30% sucrose solution in 0.1 PB and kept at 4°C. Whole brains were sliced at 50 μm with a microtome and slices were kept in 0.1 PB.

Immunostaining and Imaging

Sections were incubated in blocking solution (10% NHS, 0.5% Triton in PBS) for an hour at rT. After rinsing, sections were incubated for 48 hours at 4°C on a shaker in primary antibody solution with 2% NHS (1:2000 Rabbit anti-C-FOS, ab208942, Abcam; 1:1000 Rat anti-Ctip2, ab18465, Abcam; 1:1000 Rabbit anti-GABAalpha6, G5555, Sigma-Aldrich; 1:1000 Rabbit anti-Cux1, (Ellis et al., 2001)). After rinsing, sections were incubated for 2 hours at rT on a shaker with secondary antibody (1:500 Donkey anti-rabbit A594, 711-585-152, Jackson; 1:500 Donkey anti-Rabbit A488, 711-545-152, Jackson; 1:500 Donkey anti-rabbit Cy5, 711-175-152, Jackson; Donkey

anti-rat Cy3, 712-165-150, Jackson). Sections were counterstained with DAPI. Finally, sections were rinsed in 0.1 PB, placed with chroomulin on coverslips and mounted on slide glasses with Mowiol.

Sections were imaged with a Zeiss AxioImager 2 (Carl Zeiss, Jena, Germany) at 10x. A DsRed filter and an exposure time of 300 ms was used for the Alexa 595 channel (C-FOS). The DAPI channel was scanned at 20 ms or 30 ms exposure time. Tile scans were taken from whole brain slices. We processed half the sections obtained from slicing, hence, the distance between tile scan images was 100 µm. High resolution images were taken with a LSM 700 confocal microscope (Carl Zeiss, Jena, Germany).

Image analysis

We developed an image analysis workflow for brain region identification and quantification of C-FOS positive neurons following eyeblink conditioning (**Supp. Fig. 1**) (https://github.com/BaduraLab/cell-counting) . The workflow combines Fiji and a SHARP-Track, a software written in MATLAB initially developed to localize brain regions traversed by electrode tracks (Shamash et al., 2018) (https://github.com/cortex-lab/allenCCF/tree/master/SHARP-Track). Brain slices were preprocessed (rotating, cropping and scaling) with a custom written macro in Fiji (Schindelin et al., 2012). Next, slices were registered to the Allen Brain Atlas using the SHARP-Track user interface. Segmentation was performed on the registered slices in Fiji. Given the characteristic C-FOS staining pattern in the cerebellar granule layer (**Fig 3. A**), we used different thresholding algorithms for the cerebellum and for the rest of the brain. Following that, automated cell counting of C-FOS positive neurons was performed with a custom written macro in Fiji (cerebellum - circularity: 0.5-1, size: 0-20 pixels, rest of the brain - circularity: 0.7-1, size: 0-20 pixels) to get the X and Y coordinates of every detected cell. The output matrix of coordinates was used to create a ROI array (1 cell = 1 ROI) per slice in SHARP-Track. This step allows one to one matching between the ROI array and the previously registered slice. Finally, the reference-space locations and brain regions of each neuron were obtained by overlapping the registration array with the ROI array. ROI counts were normalized by brain region surface following the hierarchical structure of the Allen Brain Atlas. The surface of each brain areas was calculated per slice and cell density was defined as ROI counts/surface.

Statistics

Statistics were performed in MATLAB and GraphPad Prism 6. Data is reported as mean ± std or sem. Normality was tested and accepted for both eyeblink CR amplitudes and for speed of the right back paw. The corresponding statistical test for the *p* values reported are specified in Results. Time data (training sessions) was analyzed using two-way repeated measures ANOVA for sex and session. Sex effect is reported in Results, session effect is significant in all groups (indicating learning through time) and interaction is reported if significant. For Kendalls's correlation on C-FOS data, we report Tau and *p* values.

Results

B6 female and male mice show comparable variability in eyeblink conditioning and females reach higher learning scores

To study differences in learning profiles between sexes, we performed delayed eyeblink conditioning experiments with B6 females (n = 14) and males (n = 14). First, we habituated the animals to the set-up for increasing periods of time over five days to decrease anxiety levels and optimize training. Next, we subjected mice to a 5-day training paradigm in order to capture behavioral variability in the first stages of learning. This training length was selected considering that animals show the most variability within the first days of acquisition, start showing reliable CRs in day four/five and eventually plateau during the last 5 days of training (Heiney et al., 2014b; Giovannucci et al., 2017). Our aim was to compare variability between males and females at the beginning of learning and to use C-FOS immunostainings as a proxy for neural activity to identify possible eyeblink-related areas. Evidence suggests that C-FOS protein greatly increases after exposure to novel objects, sorroundings or stimuli while continuous, long-term, exposure to persistent stimuli returns *C-fos* expression to basal levels (Joo et al., 2015; Gallo et al., 2018; Bernstein et al., 2019). Therefore, we performed a 5-day training paradigm instead of the standard 10-day acquisition training.

The blue LED light (conditioned stimulus, CS) was triggered 250 ms prior to the puff to the cornea (unconditioned stimulus, US) in paired trials and the two stimuli co-terminated (**Fig. 1A**). Sessions consisted of 20 blocks of 12 trails each (1 US only, 11 paired and 1 CS only). Eyelid movements were recorded with a high-speed camera. Mice learned the association between the stimuli progressively and developed a gradually increasing conditioned response (CR) (**Fig 1A, 1B**). Males and females had comparable learning profiles, and the variances during training sessions were not significantly different between sexes (F-test for two sample variances in CR amplitude of paired trials, $F = 3.66$, $p = 0.11$). In paired trials (CS-US), both sexes followed similar learning profiles but started to diverge on the fourth session (**Fig. 1B**). In CS only trials, females showed a slight increase in CR percentage on session four that culminated with a 60% CR responses in session five opposed to 40% in males (two-way ANOVA for sex and sessions: sex effect: $F(1,26) = 1.461$, $p = 0.237$, interaction sex and session: $F(4,104) =$

4.01, p = 0.02, Cohen's d session five: 0.88) (**Fig. 1C**). The CR amplitude (measured as the response normalized to UR max amplitude = 1) during CS trials was significantly higher in females compared to males (two-way ANOVA for sex and sessions: sex effect: $F(1,26) = 6.109$, $p = 0.0203$, interaction sex and session: $F(4,104) = 5.12$, $p = 0.0008$, Cohen's d session five: 0.93) (**Fig. 1D**). On the last session of training, females reached an average amplitude of 0.55 while males reached an average of 0.33 (**Fig. 1E**). Overall, these results show that male and female mice show comparable variance in eyeblink conditioning but females reach higher learning scores in a 5 day training paradigm.

Learning scores correlate with spontaneous locomotor activity

We next asked whether mice show behavioral differences in voluntary locomotor activity during training. We investigated whether higher learning scores would correlate with higher spontaneous locomotor activity. For this purpose, we added infrared cameras to the eyeblink setups to record mice body movements during eyeblink sessions. The cameras were placed at the right back corner of the box, allowing a wide recording angle to capture whole body movements (**Fig. 2A**). We recorded videos of full training sessions for each mouse, which were later analyzed offline.

To track different body parts and get a meaningful movement output, we used DeepLabCut (DLC), a software for automated animal pose tracking (Mathis et al., 2018) (Materials and Methods; *Movement analysis*). This approach allows movement tracking without utilizing physical markers on the body that can hinder natural movement. We tracked 5 body parts: tail base, hip, knee, right back paw and nose (**Fig. 2A**). After training and refining the deep neural network, videos were analyzed to detect each body part in each video frame. The spatial coordinates of the different body parts during a training session were then used to calculate movement parameters.

After confirming normal distribution of the spatial coordinates per body part over training sessions, we performed the Grubbs's test for outlier removal to discard possible tracking errors. Animals were head-fixed on top of the wheel, hence, Y position was similar between both sexes (**Fig 2. B,C**). However, we observed differences in X position across the sexes; the right back paw and knee boxplots do not overlap with the hip and tail ones in

females, while in males they do. This indicates a tendency in males to rotate their body axis from left to right, which causes the hip and tail x position to fluctuate more compared to females (**Fig 2. B,C**). We selected speed of the right back paw as a proxy for general locomotion behavior on further analysis.

Animals increased their running speed on the wheel during training, and both females and males had comparable variances (F-test for two sample variances in speed: $F = 0.58$, $p = 0.305$). We found that females ran significantly faster than males during learning, reaching an average speed of 31 cm/s compared to males that reached 24 cm/s (two-way ANOVA for sex and sessions: sex effect: $F(1,26) = 12.17$, $p = 0.0017$, Cohen's d session five: 1.07) (**Fig. 2D,E**). We observed a similar increasing trend between CR amplitude and running speed across sexes. Thus, we performed a linear regression between speed of the right back paw and CR amplitude on the last session which shows a clear correlation between the variables ($R^2 = 0.75$, $p = 0.002$) (**Fig.2 F**). These results reveal that mice that spontaneously run faster on the wheel, reach higher learning scores in eyeblink conditioning.

Learning scores correlate with C-fos expression

To explore brain regions that could have a role in modulating associative learning during eyeblink conditioning, we performed C-FOS immunostainings following the last training session. *C-fos* is an immediate early expressed gene, a family of transcription factors that is expressed shortly after a neuron has depolarized. Because of its precise time window of expression, it is widely used as an activity marker (Chung, 2015). We imaged sections of whole brains with a fluorescent microscope and developed an image analysis workflow to quantify C-FOS positive neurons and identify their location (**Supp. Fig. 1**).

To ensure an appropriate control for the quantification of C-FOS positive cells, we included a pseudoconditioned mouse in each group (n = 2 males, 2 females). These mice went through the same experimental steps as the conditioned mice with the only exception that they were not trained with paired CS-US trials. Instead, we exposed them to a protocol with CS and US only trials, keeping the same structure and duration as the conditioned protocol. Pseudoconditioned mice did not acquire an association given that there was no substrate for learning. These animals showed marginal increase in the running speed over training sessions, and displayed lower

C-fos expression in most areas of the brain, excluding the visual cortex, where we found an abundance of C-FOS positive cells. (**Supp. Fig. 2**).

We quantified the density of C-FOS expressing cells in brain regions defined within the hierarchical structure of the Allen Brian Atlas (cerebrum, brainstem and cerebellum). In order to indentify regions potentially involved in associative learning, we selected mice that showed a CR amplitude of 0.4 or higher in CS only trials (n = 14; 5 males, 9 females). We performed a Kendall's correlation (non-parametric rank order regression) between density of C-FOS positive cells and CR amplitude on the last session. In the cerebellum, C-FOS labelling was clearly localized in the granule cell layer (**Fig 3A**). In the cerebellar hemispheric regions, crus 1 and simplex had a significant correlation between C-FOS cell density and CR amplitude (crus 1: tau = 0.42, p = 0.042, simplex: tau = 0.52, p = 0.009). In the cerebellar vermis, lobule VI also had a significant correlation and the highest Tau (lobule VI: tau = 0.8, p = 0.009) (**Fig. 3B**). In the midbrain and hindbrain, we found a significant correlation in the red nucleus, the racial nucleus, the inferior olive and the pontine nuclei (Red nucleus: tau = 0.43, p = 0.041, Facial nucleus: tau = 0.57, p = 0.006, inferior olive: tau = 0.76, p = 0.0008, pontine nuclei: tau = 0.48, p = 0.021) (**Fig. 3E**). Finally, we found a positive correlation in the visual, motor and somatosensory cortices and the amygdala (visual cortex: tau = 0.51, p = 0.013, motor cortex: tau = 0.69, p = 0.0003, somatosensory cortex: 0.54, p = 0.007, amygdala: tau = 0.63, p = 0.001) (**Fig. 3H**).

Together, these results confirm the previously reported areas associated with eyeblink conditioning within the olivo-cerebellar and ponto-cerebellar systems (Ruigrok, 2011; D'Angelo et al., 2016) and suggest that other areas might be involved in this learning task.

C-FOS positive cells indentity

To gain insight on the types of cells expressing *C-fos*, we performed several double immunostainings. To confirm that C-FOS positive cells observed in the granule cell layer were indeed granule cells, we co-stained for C-FOS and GABAα6, a granule cell specific marker (**Fig 4. A, B**). To understand whether there was a certain layer specificity in *C-fos* expression following eyeblink conditioning, we performed two double immunostaings with C-FOS; CUX1, a marker for upper cortical layers (II-IV) and CTIP2, for lower cortical layers (V-VI). Although we

detected C-FOS positive cells in all layers of the cortex, we observed a higher colocalization of CUX1 and C-FOS compared to CTIP2 and C-FOS, specially in the visual, somatosensory and motor cortices (**Fig 4. C-F**). This indicates an enrichment of *C-fos* expression in upper cortical layers, which suggest increased neural activity in these layers during associative learning.

Correlation between learning scores and C-fos expression is consistent in B6CBAF1 strain

Given the evidence of the possible unwanted effects of highly inbred mouse strains like B6 in replicability and reproducibility (Åhlgren and Voikar, 2019), we wanted to investigating inter-strain variability in associative learning. We asked whether the results obtained in B6 mice would be consistent in a different mouse strain. For this purpose, we performed eyeblink conditioning together with C-FOS immunostainings in B6CBAF1 mice, which are the F1 hybrids of B6 and CBA strains. Hybrid mice are used due to their hybrid vigor, the robustness and health gained from a high degree of hetereozygosity (Wolfer et al., 2002). B6CBAF1 mice have significantly less retinal degeneration and hearing loss compared to B6 mice, which makes them an appropriate candidate for visual and auditory experiments (Erway et al., 1996; Milon et al., 2018; Ohlemiller, 2019).

B6CBAF1 mice learned the association between the stimuli and gradually formed CRs. We observed a trend indicating similar sex differences between B6CBAF1 mice and B6. Females reached 53% CR percentage compared to 35% in males (two-way ANOVA for sex and sessions: sex effect: $F(1,14) = 2.55$, $p = 0.237$, interaction sex and session: $F(4,104) = 3.01$, $p = 0.021$, Cohen's d session five: 0.86) (**Fig 5. A**). When it comes to the amplitude of these responses, BFCBAF1 females showed a trend towards slightly higher CR amplitude over training sessions compared to males (two-way ANOVA for sex and sessions: sex effect: $F(1,14) = 2.55$, $p = 0.132$, Cohen's d session five: 0.84) (**Fig 5. B**).

We followed the same analysis pipeline to quantify *C-fos* expression in brain slices of B6CBAF1 mice after eyeblink conditioning. Next, we selected mice that showed a CR amplitude of 0.4 or higher in CS only trials (n =11; 3 males, 8 females) and performed a Kendall's correlation between density of C-FOS positive cells and CR amplitude on the last training session. The granule cell layer also contained C-FOS labelling, and crus 1, the simplex and lobule VI were found to have a significant positive correlation (crus 1: tau = 0.82, $p = 0.0001$, simplex: tau =

0.7, p = 0.005, lobule VI: tau = 0.75, p = 0.0007) (**Fig 5. C, D**). In the hindbrain, the correlation between C-FOS cells and learning was also significant in the inferior olive and the pontine nuclei (inferior olive: tau = 0.85, p = 0.0004, pontine nuclei: tau = 0.78, p = 0.0003) (**Fig 5. E, F**). Additionally, the visual, motor and somatosensory cortices showed significant positive correlations (visual cortex: tau = 0.64, p = 0.0057, motor cortex: tau = 0.75 p = 0.0008, somatosensory cortex: tau = 0.6, p = 0.009) (**Fig 5. G, H**).

Variability between sexes and strains

To further understand inter-strain and inter-sex variability in our data-set, we calcultated the coefficient of variance (CV) for each of the variables that we quantify in both B6 and B6CBAF1 mice. We selected the 14 *learners* (CR amplitude on session 5 > 0.4) B6 mice (n = 5 males, 9 females) and the 11 *learners* B6CBAF1 mice (n = 3 males, 8 females) and grouped them by strain and by strain and sex (**Fig. 6**). For each group, we calculated the CV for the common variables acquired and previously reported, which can be grouped in two main categories; eyeblink performance and *C-fos* expression. Eyeblink performances includes CR amplitude and percentage. *C-fos* expression includes the density of C-FOS positive cells in the areas of the brain where we have found a positive significant correlation accros both strains: crus 1, simplex, lobule VI, inferior olive, pontine nuclei, and the visual, motor and somatosensory cortices. When comparing strains, we observed that the variances for each variable were similar, with the exception of the somatosensory cortex, where B6CBAF1 mice seem to be more variable compared to B6 (**Fig. 6A**). B6 female and male mice had similar CVs for most of the variables, although in crus 1 and the somatosensory cortex males seem to have a slightly higher CV (**Fig. 6B**). However, this could be due to the difference in sample size. We observed something similar between B6CBAF1 female and male mice; males showed slightly higer CV in C-FOS density in the simplex (**Fig. 6C**). Additionally, B6CBAF1 mice had the highest CV in the somatosensory cortex.

Discussion

Understanding behavioral variability in the context of neuroscience is a challenge. We are still far from fully understanding how factors like sex and strain give rise to differences in behavior.

We tackled this question by making use of a well-known learning paradigm to study behavioral variability. We found that B6 female and male mice showed comparable variance in eyeblink conditioning and locomotion while being head-fixed on a rotating wheel. Although the variance within these behaviors was not different between sexes, females reached higher learning scores and running speeds within five days of training. Importantly, we found a robust correlation between learning scores and running speed which is consistent across sexes. In a similar way, we found that enriched *C-fos* expression across several brain areas positively correlates with learning, which suggests the involvement of these regions in eyeblink conditioning. Finally, we observed similar results in a hybrid mouse strain (B6CBAF1).

Sex, strain and behavior: comparable variability but difference in performance

Opposed to what is sometimes assumed in behavioral science, we observed that sexes show comparable variability. However, we found significant differences in performance during eyeblink conditioning and locomotion. Common behaviors like duration of running vary between female and male mice in the wild (Lightfoot et al., 2004; Goh & Ladiges, 2015). Behaviors widely assessed in research such as fear conditioning and navigation on the Morris water maze also show differences depending on sex (Roof & Stein, 1999; Keeley et al., 2013; Yang et al., 2013; Gruene, Flick, et al., 2015). In the context of cerebellar-dependent learning, evidence shows that estradiol increases the density of parallel fiber to Purkinje cell synapse and induces long-term potentiation, which improves memory formation (Andreescu et al., 2007). A similar mechanism could potentially explain the higher learning scores observed in females in eyeblink conditioning. In trace eyeblink conditioning, both sexes reach similar learning scores but females show significantly higher CR percentage compared to males in the first five days of learning, which is in line with our findings (Rapp et al., 2021). Considering that adapting motor reflexes is

a highly conserved behavior, it is logical that both sexes reach similar asymptotic learning scores. However, these findings together with our results suggest that females exhibit faster learning rates during the first stages of learning.

Using females could reduce the training time to achieve desired scores, which would be advantageous for certain experiments, particularly time-sensitive ones, such as calcium imaging or electrophysiological measurements. In addition, our results show that male and female mice have similar variability, which indicates that females can be included in studies without taking into account the estrous cycle phase. In general, utilizing both sexes would reduce the overall number of animals used in research and increase the relevance and generalization of scientific findings.

A common experimental setting in neuroscience involves head-fixing awake mice and placing them on a running wheel. A recent study investigated sex differences in head-fixed running behavior and found that female mice ran forward naturally within the first two days, while males took seven days to progressively learn to only run forward (Prawira, 2019). These results are consistent with our findings and, together, they provide evidence on the importance of considering sex as a controlled variable in experimental design. In our experiments, the differences in learning scores between sexes were strongly correlated to the changes in locomotor activity on the wheel. Our results show that the previously reported correlation between imposed locomotor activity and learning scores (Albergaria et al., 2018) persists when mice can initiate locomotion voluntarily. This suggests that spontaneous running activity might facilitate associative learning and could be predictive of learning scores.

There are mechanisms that could explain this relationship. For example, higher running speeds on the wheel could be an expression of higher arousal levels, which would involve threat-related signals from the amygdala and noradrenergic and cholinergic input from the cortex (Tye et al., 2011; Reimer et al., 2016; McCall et al., 2017). Consequently, the involvement of the cortex and the amygdala would facilitate acquisition of the association. In addition, locomotor signals coming from the body and CS signals could converge onto granule cells and, subsequently, convey enhanced CS information to Purkinje cells (Albergaria et al., 2018). This mechanism would also be plausible considering that multimodal signals can merge in mossy fibers to granule cells (Sawtell, 2010; Ishikawa et al., 2015; Wagner et al., 2017).

Finally, we have found that, besides running slower, males tended to have a tilted position on the wheel compared to females. These differences in body position could be partially caused by differences in stress levels that, at the same time, could affect learning rates. It is known that stress plays an important role in modulating neural activity in the hippocampus. Corticosterone - among other stress hormones – increases CA1/CA3 firing rates shortly after a stressful period and induces molecular cascades that enhance calcium influx, which disrupts hippocampal function (Joëls, 2009). Similar mechanisms have been described in the cerebellum; calcium-based excitability in the DCN is altered in animals with higher levels of corticosterone evoked by shipping stress (Schneider et al., 2013).

Associative learning networks

Our results show that *C-fos* expression upon eyeblink conditioning in the cerebellar cortex is localized in the granule cell layer. This is not surprising given that multiple forms of plasticity have been studied within the synapses in this layer. For example, the mossy fiber-granule cell synapse undergoes both long-term potentiation and long-term depression (Gao et al., 2012), and evidence has shown that granule cell activity adapts over time during eyeblink conditioning (Giovannucci et al.,2017), and other types of learning (Knogler et al., 2017; Wagner et al., 2017). In addition, induction of LTP by theta-burst stimulation in acute cerebellar slices activated cAMP-responsive element binding protein (CREB) cascade which, in turn, activated *C-fos* expression (Gandolfi et al., 2017). Our results are consistent with these findings and, overall, they provide evidence on how plasticity at the input level in the cerebellar cortex can evoke transcriptional processes that contribute to learning consolidation. The strongest correlations between the *C-fos* expression and CR amplitude within the cerebellum were observed in simplex, lobule VI, and crus 1, which is consistent with the "eyeblink region", but expands beyond the small area usually recorded using electrophysiological approaches (Heiney et al., 2014a; ten Brinke et al., 2015). Strong *C-fos* expression in crus 1 supports our previous findings showing importance of this lobule in eyeblink conditioning (Badura et al., 2018).

Outside the cerebellum, we identified several brain areas that play a role in eyeblink conditioning. At the brainstem level, we found a relation between high learning scores and *C-fos* expression in the pons, the inferior olive, the red nucleus and the facial nucleus. High activity in the red nucleus and facial nucleus is to be expected,

given that these two nuclei, together with the oculomotor nucleus, execute the blink. The inferior olive and the pontine nuclei relay the US and CS information to the cerebellar cortex, respectively. During early training sessions, the US is a highly aversive stimulus, which makes it comparetively more salient than the CS signal. Hence, one would expect increased activity in the pons relative to the inferior olive. However, in later learning stages (when animals have consolidated the association), the CS is predictive of the US, which would increase the activity in the inferior olive realive to the pons. We found a correlation between learning scores and C-FOS positive cells in both the pons and the inferior olive, which could indicate an intermediate stage of learning, where animals have learned the association but the US information is still relevant.

Moreover, we found higher *C-fos* expression in the visual, motor and somatosensory cortices in mice with higher learning scores. Processing in these cortices could facilitate the CS to become more salient and ultimately predict the US. The somatosensory cortex projects to the lateral amygdala which, in turn, projects to the central amygdala to ultimately contact the pons. The high *C-fos* expression found in the amygdala points towards a two stage conditioning model; where the amygdala would have an initial role with arousal as a salient feature and a second phase where the cerebellum would take over to form precisely-timed CRs (Boele et al., 2010). In the motor cortex, higher C-fos levels in high performing mice might be due to locomotor activity rather than learning itself. However, as mentioned above, this could play a role in learning either by directly affecting cerebellar input or indirectly as arousal.

Finally, we found an enrichment of *C-fos* expression in upper cortical layers (II, III and IV), specially in the visual and somatosensory cortices. The principal excitatory neurons in layers II/III have large axons that project to other telencephalic areas, such as the cortex itself and the striatum (Adesnik and Naka, 2018), while neurons in layer IV form loops within the layer and connect to layers II/III and VI (Scala et al., 2019). Layers VI and VII are thought to be the main outputs of the cerebral cortex, connecting to multiple subcortical areas and the thalamus, respectively (Harris and Shepherd, 2015). The higher C-FOS density in upper cortical layers indicates higher activity, which could reflect feedforward loops within neuronal populations and translaminar connectivity that could reinforce learning. However, research is needed to determine the identity of these neurons, for example, to confirm whether they are excitatory, inhibitory or both.

Together, these findings give us a better understanding of the networks underlying eyeblink conditioning and provide candidate brain areas to be further researched in the context of associative learning.

References

1. Adesnik H, Naka A (2018) Cracking the Function of Layers in the Sensory Cortex. Neuron 100:1028–1043 Available at: https://doi.org/10.1016/j.neuron.2018.10.032.
2. Åhlgren J, Voikar V (2019) Experiments done in Black-6 mice: what does it mean? Lab Anim (NY) 48:171–180.
3. Albergaria C, Silva NT, Pritchett DL, Carey MR (2018) Locomotor activity modulates associative learning in mouse cerebellum. Nat Neurosci 21:725–735 Available at: http://dx.doi.org/10.1038/s41593-018-0129-x.
4. Andreescu CE, Milojkovic BA, Haasdijk ED, Kramer P, De Jong FH, Krust A, De Zeeuw CI, De Jeu MTG (2007) Estradiol improves cerebellar memory formation by activating estrogen receptor β. J Neurosci 27:10832–10839.
5. Arnold MA, Newland MC (2018) Variable behavior and repeated learning in two mouse strains: Developmental and genetic contributions. Behav Processes 157:509–518 Available at: https://doi.org/10.1016/j.beproc.2018.06.007.
6. Badura A, Verpeut JL, Metzger JW, Pereira TD, Pisano TJ, Deverett B, Bakshinskaya DE, Wang SSH (2018) Normal cognitive and social development require posterior cerebellar activity. Elife 7:1–36.
7. Becker JB, Prendergast BJ, Liang JW (2016) Female rats are not more variable than male rats: A meta-analysis of neuroscience studies. Biol Sex Differ 7:1–7 Available at: http://dx.doi.org/10.1186/s13293-016-0087-5.
8. Bernstein HL, Lu YL, Botterill JJ, Scharfman HE (2019) Novelty and novel objects increase c-fos immunoreactivity in mossy cells in the mouse dentate gyrus. Neural Plast 2019.
9. Boele HJ, Koekkoek SKE, De Zeeuw CI (2010) Cerebellar and extracerebellar involvement in mouse eyeblink conditioning: The ACDC model. Front Cell Neurosci 3:1–13.
10. Bucán M, Abel T (2002) The mouse: genetics meets behavior. Nat Rev Genet Available at: https://www.nature.com/articles/nrg728?draft=marketing.
11. Chung L (2015) A Brief Introduction to the Transduction of Neural Activity into Fos Signal. Dev Reprod

19:61–67.

12. D'Angelo E, Galliano E, De Zeeuw CI (2016) Editorial: The olivo-cerebellar system. Front Neural Circuits 9:2015–2017.
13. Daniel E Shumer, Natalie J Nokoff NPS (2017) Sexual dimorphism in predisposition to Alzheimer's disease Daniel. Physiol Behav 176:139–148.
14. Ellis T, Gambardella L, Horcher M, Tschanz S, Capol J, Bertram P, Jochum W, Barrandon Y, Busslinger M (2001) The transcriptional repressor CDP (Cutl1) is essential for epithelial cell differentiation of the lung and the hair follicle. Genes Dev 15:2307–2319.
15. Erway LC, Shiau YW, Davis RR, Krieg EF (1996) Genetics of age-related hearing loss in mice. III. Susceptibility of inbred and F1 hybrid strains to noise-induced hearing loss. Hear Res 93:181–187.
16. Faure A, Pittaras E, Nosjean A, Chabout J, Cressant A, Granon S (2017) Social behaviors and acoustic vocalizations in different strains of mice. Behav Brain Res 320:383–390 Available at: http://dx.doi.org/10.1016/j.bbr.2016.11.003.
17. Festing MFW (1999) Warning: The use of heterogeneous mice may seriously damage your research. Neurobiol Aging 20:237–244.
18. Fonnesu M, Kuczewski N (2019) Doubts on the efficacy of outliers correction methods.
19. Gallo FT, Katche C, Morici JF, Medina JH, Weisstaub N V. (2018) Immediate early genes, memory and psychiatric disorders: Focus on c-Fos, Egr1 and Arc. Front Behav Neurosci 12:1–16.
20. Gandolfi D, Cerri S, Mapelli J, Polimeni M, Tritto S, Fuzzati-Armentero MT, Bigiani A, Blandini F, Mapelli L, D'Angelo E (2017) Activation of the CREB/c-Fos pathway during long-term synaptic plasticity in the cerebellum granular layer. Front Cell Neurosci 11:1–13.
21. Gao Z, Proietti-Onori M, Lin Z, ten Brinke MM, Boele HJ, Potters JW, Ruigrok TJH, Hoebeek FE, De Zeeuw CI (2016) Excitatory Cerebellar Nucleocortical Circuit Provides Internal Amplification during Associative Conditioning. Neuron 89:645–657.
22. Gao Z, Van Beugen BJ, De Zeeuw CI (2012) Distributed synergistic plasticity and cerebellar learning. Nat Rev Neurosci 13:619–635.

23. Giovannucci A, Badura A, Deverett B, Najafi F, Pereira TD, Gao Z, Ozden I, Kloth AD, Pnevmatikakis E, Paninski L, De Zeeuw CI, Medina JF, Wang SSH (2017) Cerebellar granule cells acquire a widespread predictive feedback signal during motor learning. Nat Neurosci 20:727–734.
24. Gormezano I, Schneiderman N, Deaux E, Fuentes I (1962) Nictitating membrane: Classical conditioning and extinction in the albino rabbit. Science (80-) 138:33–34.
25. Harris KD, Shepherd GMG (2015) The neocortical circuit: Themes and variations. Nat Neurosci 18:170–181.
26. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit 2016-Decem:770–778.
27. Heiney SA, Kim J, Augustine GJ, Medina JF (2014a) Precise control of movement kinematics by optogenetic inhibition of Purkinje cell activity. J Neurosci 34:2321–2330.
28. Heiney SA, Wohl MP, Chettih SN, Ruffolo LI, Medina JF (2014b) Cerebellar-dependent expression of motor learning during eyeblink conditioning in head-fixed mice. J Neurosci 34:14845–14853.
29. Hoogland TM, De Gruijl JR, Witter L, Canto CB, De Zeeuw CI (2015) Role of synchronous activation of cerebellar purkinje cell ensembles in multi-joint movement control. Curr Biol 25:1157–1165 Available at: http://dx.doi.org/10.1016/j.cub.2015.03.009.
30. Insafutdinov E, Pishchulin L, Andres B, Andriluka M, Schiele B (2016) Deepercut: A deeper, stronger, and faster multi-person pose estimation model. Lect Notes Comput Sci (including Subser Lect Notes Artif Intell Lect Notes Bioinformatics) 9910 LNCS:34–50.
31. Ishikawa T, Shimuta M, Häuser M (2015) Multimodal sensory integration in single cerebellar granule cells in vivo. Elife 4:1–10.
32. Joëls M (2009) Stress, the hippocampus, and epilepsy. Epilepsia 50:586–597.
33. Joo JY, Schaukowitch K, Farbiak L, Kilaru G, Kim TK (2015) Stimulus-specific combinatorial functionality of neuronal c-fos enhancers. Nat Neurosci 19:75–83.
34. Knogler LD, Markov DA, Dragomir EI, Štih V, Portugues R (2017) Sensorimotor Representations in Cerebellar Granule Cells in Larval Zebrafish Are Dense, Spatially Organized, and Non-temporally

Patterned. Curr Biol 27:1288–1302.

35. Kratochwil CF, Maheshwari U, Rijli FM (2017) The long journey of pontine nuclei neurons: From rhombic lip to cortico-ponto-cerebellar circuitry. Front Neural Circuits 11:1–19.
36. Leung C, Jia Z (2016) Mouse genetic models of human brain disorders. Front Genet 7:1–20.
37. Löwgren K, Bååth R, Rasmussen A, Boele HJ, Koekkoek SKE, De Zeeuw CI, Hesslow G (2017) Performance in eyeblink conditioning is age and sex dependent. PLoS One 12:1–15.
38. Mathis A, Mamidanna P, Cury KM, Abe T, Murthy VN, Mathis MW, Bethge M (2018) DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. Nat Neurosci 21:1281–1289 Available at: http://dx.doi.org/10.1038/s41593-018-0209-y.
39. McCall JG, Siuda ER, Bhatti DL, Lawson LA, McElligott ZA, Stuber GD, Bruchas MR (2017) Locus coeruleus to basolateral amygdala noradrenergic projections promote anxiety-like behavior. Elife 6:1–23.
40. McCarthy MM, Arnold AP, Ball GF, Blaustein JD, de Vries GJ (2012) Sex differences in the brain: The not so inconvenient truth. J Neurosci 32:2241–2247.
41. McGinley MJ, Vinck M, Reimer J, Batista-Brito R, Zagha E, Cadwell CR, Tolias AS, Cardin JA, McCormick DA (2015) Waking State: Rapid Variations Modulate Neural and Behavioral Responses. Neuron 87:1143–1161 Available at: http://dx.doi.org/10.1016/j.neuron.2015.09.012.
42. Meziane H, Ouagazzal AM, Aubert L, Wietrzych M, Krezel W (2007) Estrous cycle effects on behavior of C57BL/6J and BALB/cByJ female mice: Implications for phenotyping strategies. Genes, Brain Behav 6:192–200.
43. Milon B, Mitra S, Song Y, Margulies Z, Casserly R, Drake V, Mong JA, Depireux DA, Hertzano R (2018) The impact of biological sex on the response to noise and otoprotective therapies against acoustic injury in mice. Biol Sex Differ 9:1–14.
44. Ohlemiller KK (2019) Mouse methods and models for studies in hearing. J Acoust Soc Am 146:3668–3680.
45. Osborne JW, Overbay A (2004) The power of outliers (and why researchers should ALWAYS check for them). Pract Assessment, Res Eval 9.

46. Ozden I, Dombeck DA, Hoogland TM, Tank DW, Wang SSH (2012) Widespread state-dependent shifts in cerebellar activity in locomoting mice. PLoS One 7.
47. Pfaff D (2001) Precision in mouse behavior genetics. Proc Natl Acad Sci U S A 98:5957–5960.
48. Powell K, Mathy A, Duguid I, Häusser M (2015) Synaptic representation of locomotion in single cerebellar granule cells. Elife 4:1–18.
49. Prawira Y (2019) Sex differences in head-fixed running behavior. SSRN Electron J 5:1–19.
50. Rapp AP, Weiss C, Matthew Oh M, Disterhoft JF (2021) Intact female mice acquire trace eyeblink conditioning faster than male and ovariectomized female mice. eNeuro 8:1–9.
51. Reimer J, McGinley MJ, Liu Y, Rodenkirch C, Wang Q, McCormick DA, Tolias AS (2016) Pupil fluctuations track rapid changes in adrenergic and cholinergic activity in cortex. Nat Commun 7:1–7 Available at: http://dx.doi.org/10.1038/ncomms13289.
52. Rivera J, Tessarollo L (2008) Genetic Background and the Dilemma of Translating Mouse Studies to Humans. Immunity 28:1–4.
53. Rousselet GA, Pernet CR (2012) Improving standards in brain-behavior correlation analyses. Front Hum Neurosci 6.
54. Ruigrok TJH (2011) Ins and outs of cerebellar modules. Cerebellum 10:464–474.
55. Sawtell NB (2010) Multimodal integration in granule cells as a basis for associative plasticity and sensory prediction in a cerebellum-like circuit. Neuron 66:573–584 Available at: http://dx.doi.org/10.1016/j.neuron.2010.04.018.
56. Scala F, Kobak D, Shan S, Bernaerts Y, Laturnus S, Cadwell CR, Hartmanis L, Froudarakis E, Castro JR, Tan ZH, Papadopoulos S, Patel SS, Sandberg R, Berens P, Jiang X, Tolias AS (2019) Layer 4 of mouse neocortex differs in cell types and circuit organization between sensory areas. Nat Commun 10:1–12 Available at: http://dx.doi.org/10.1038/s41467-019-12058-z.
57. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B, Tinevez JY, White DJ, Hartenstein V, Eliceiri K, Tomancak P, Cardona A (2012) Fiji: An open-source platform for biological-image analysis. Nat Methods 9:676–682.

58. Schneider ER, Civillico EF, Wang SSH (2013) Calcium-based dendritic excitability and its regulation in the deep cerebellar nuclei. J Neurophysiol 109:2282–2292.
59. Schreurs BG, Smith-Bell C, Burhans LB (2018) Sex differences in a rabbit eyeblink conditioning model of PTSD. Neurobiol Learn Mem 155:519–527.
60. Shamash P, Carandini M, Harris K, Steinmetz N (2018) A tool for analyzing electrode tracks from slice histology. bioRxiv:447995 Available at: https://www.biorxiv.org/content/early/2018/10/19/447995%0Ahttp://dx.doi.org/10.1101/447995.
61. Simpson J, Kelly JP (2012) An investigation of whether there are sex differences in certain behavioural and neurochemical parameters in the rat. Behav Brain Res 229:289–300 Available at: http://dx.doi.org/10.1016/j.bbr.2011.12.036.
62. Sittig LJ, Carbonetto P, Engel KA, Krauss KS, Barrios-Camacho CM, Palmer AA (2016) Genetic Background Limits Generalizability of Genotype-Phenotype Relationships. Neuron 91:1253–1259 Available at: http://dx.doi.org/10.1016/j.neuron.2016.08.013.
63. Smith REW, Avery JA, Wallace GL, Kenworthy L, Gotts SJ, Martin A (2019) Sex differences in resting-state functional connectivity of the cerebellum in autism spectrum disorder. Front Hum Neurosci 13:1–13.
64. ten Brinke MM, Boele HJ, Spanke JK, Potters JW, Kornysheva K, Wulff P, IJpelaar ACHG, Koekkoek SKE, De Zeeuw CI (2015) Evolving Models of Pavlovian Conditioning: Cerebellar Cortical Dynamics in Awake Behaving Mice. Cell Rep 13:1977–1988 Available at: http://dx.doi.org/10.1016/j.celrep.2015.10.057.
65. ten Brinke MM, Heiney SA, Wang X, Proietti-Onori M, Boele HJ, Bakermans J, Medina JF, Gao Z, De Zeeuw CI (2017) Dynamic modulation of activity in cerebellar nuclei neurons during pavlovian eyeblink conditioning in mice. Elife 6:1–27.
66. Tuttle AH, Philip VM, Chesler EJ, Mogil JS (2018) Comparing phenotypic variation between inbred and outbred mice. Nat Methods 15:994–996 Available at: http://dx.doi.org/10.1038/s41592-018-0224-7.
67. Tye KM, Prakash R, Kim SY, Fenno LE, Grosenick L, Zarabi H, Thompson KR, Gradinaru V, Ramakrishnan C, Deisseroth K (2011) Amygdala circuitry mediating reversible and bidirectional control

of anxiety. Nature 471:358–362.

68. Vinck M, Batista-Brito R, Knoblich U, Cardin JA (2015) Arousal and Locomotion Make Distinct Contributions to Cortical Activity Patterns and Visual Encoding. Neuron 86:740–754 Available at: http://dx.doi.org/10.1016/j.neuron.2015.03.028.
69. Wagner MJ, Kim TH, Savall J, Schnitzer MJ, Luo L (2017) Cerebellar granule cells encode the expectation of reward. Nature 544:96–100 Available at: http://dx.doi.org/10.1038/nature21726.
70. Williamson RS, Hancock KE, Shinn-Cunningham BG, Polley DB (2015) Locomotion and Task Demands Differentially Modulate Thalamic Audiovisual Processing during Active Search. Curr Biol 25:1885–1891 Available at: http://dx.doi.org/10.1016/j.cub.2015.05.045.
71. Wolfer DP, Crusio WE, Lipp HP (2002) Knockout mice: Simple solutions to the problems of genetic background and flanking genes. Trends Neurosci 25:336–340.
72. Wood GE, Shors TJ (1998) Stress facilitates classical conditioning in males, but impairs classical conditioning in females through activational effects of ovarian hormones. Proc Natl Acad Sci U S A 95:4066–4071.

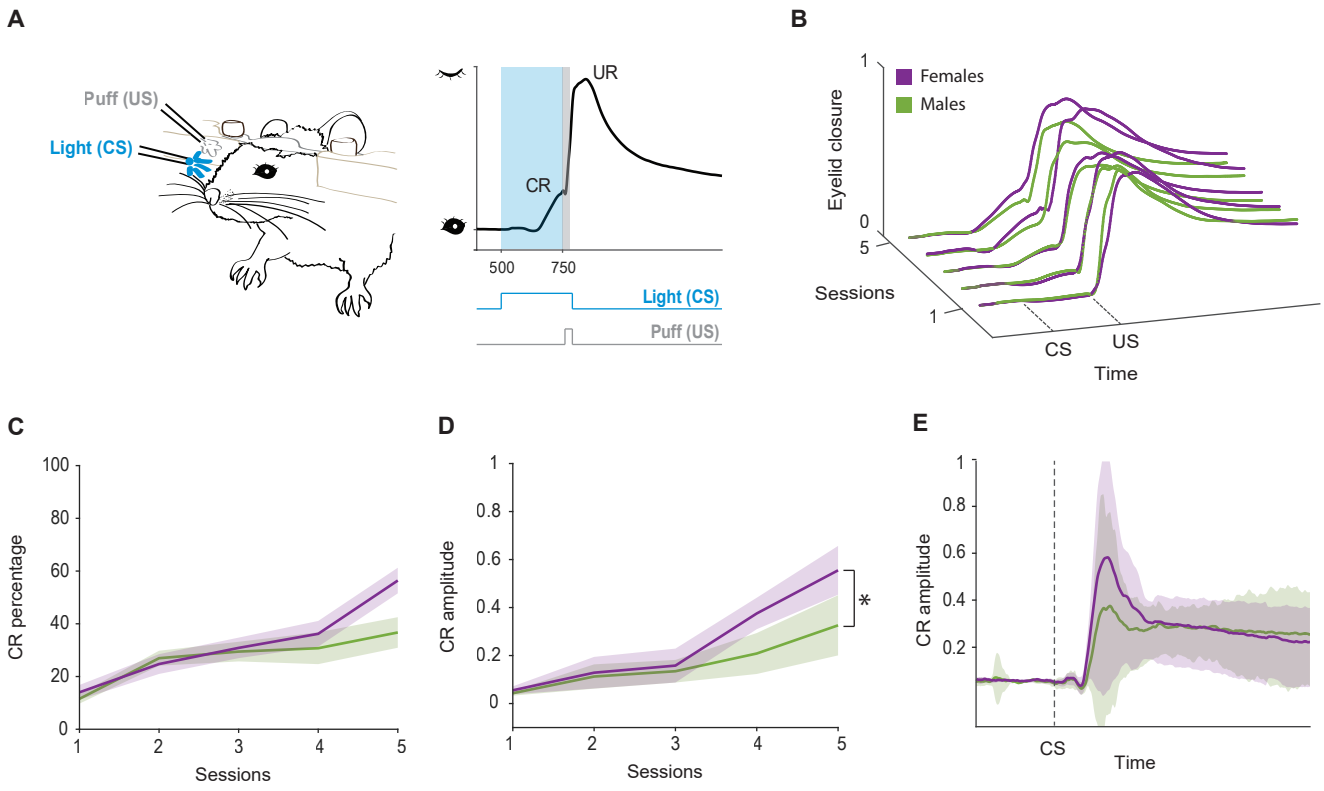

Figure 1: B6 female and male mice show comparable variability in eyeblink conditioning and females reach higher learning scores. A) Experimental setup. Mouse with implanted headplate is head-fixed on top of a freely rotating wheel. A blue light (conditioned stimulus, CS) is presented 250 ms before a puff (unconditioned stimulus, US) to the same eye. In a trained mouse, the CS produces an anticipatory eyelid closure (conditioned response, CR) followed by a blink reflex triggered by the US (unconditioned response, UR). B) Paired trails average traces in females and males over training sessions. The CR progressively develops due to the CS-US paring. C) CR percentage in CS only trials over training sessions. Purple: females, green: males. Shaded area: sem. D) CR amplitude in CS only trials over training sessions (two-way ANOVA for sex and sessions: sex effect: $F(1,26) = 6.109$, $p = 0.0203$) Shaded area: sem. E) Average response in CS only trials in the last session of training. Purple: females (n=14), green: males (n=14). Shaded area: std.

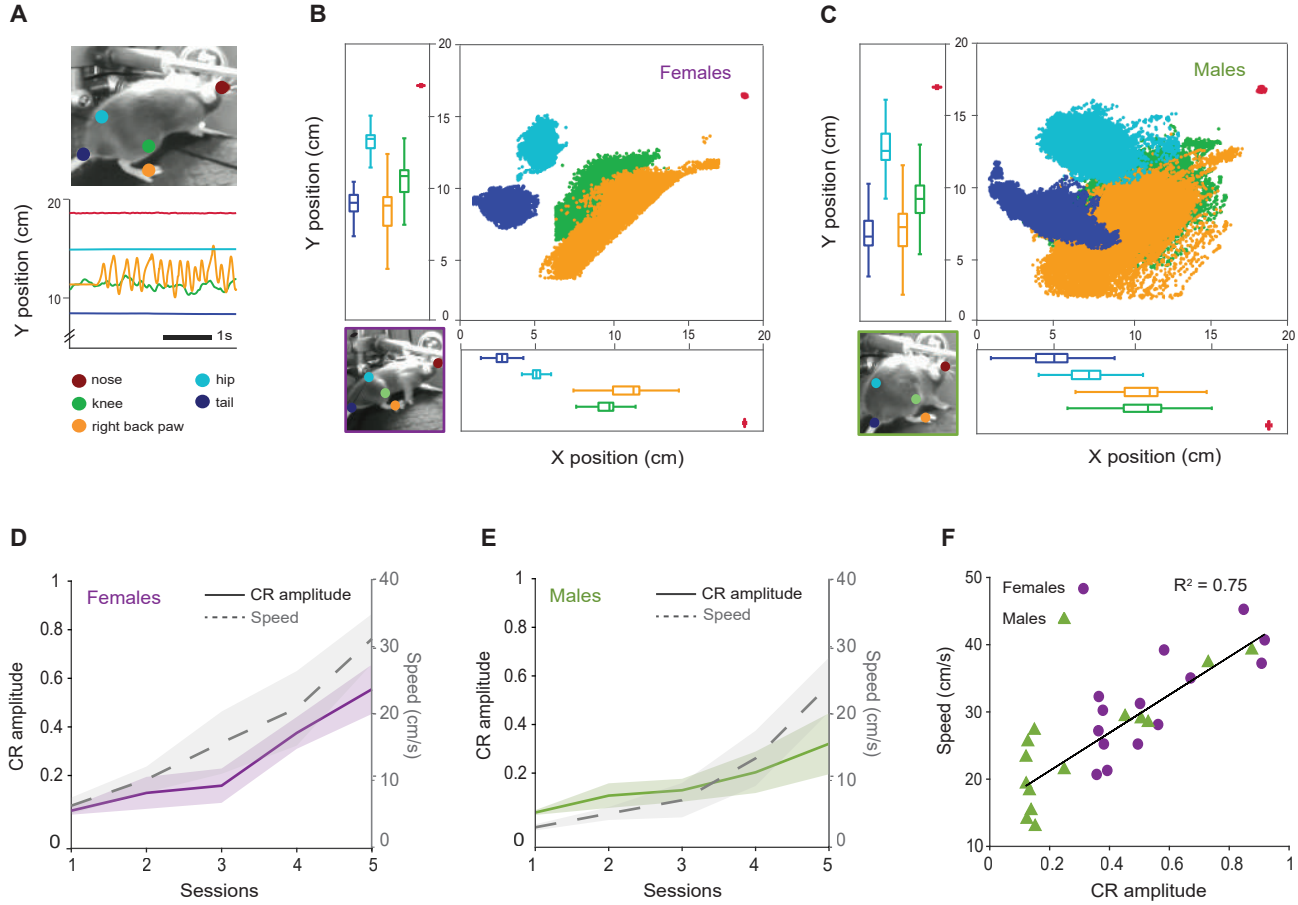

Figure 2: Learning scores correlate with spontaneous locomotor activity. A) Top: Example view of tracking with DeepLabCut. Bottom: Example tracking traces (Y position change). B) Scatter plot and boxplots of each body part (females, n = 14). C) Scatter plot and boxplots of each body part (males, n = 14). D and E) CR amplitude and speed of the right back paw over training sessions. Purple: females (n = 14), green: males (n = 14). Shaded area: sem. Speed: two-way ANOVA for sex and sessions: sex effect: $F(1,26) = 12.17$, $p = 0.0017$. F) Positive correlation between CR amplitude and speed of the right back paw on the last session of training (linear regression: $R2 = 0.7534$, $p = 0.002$).

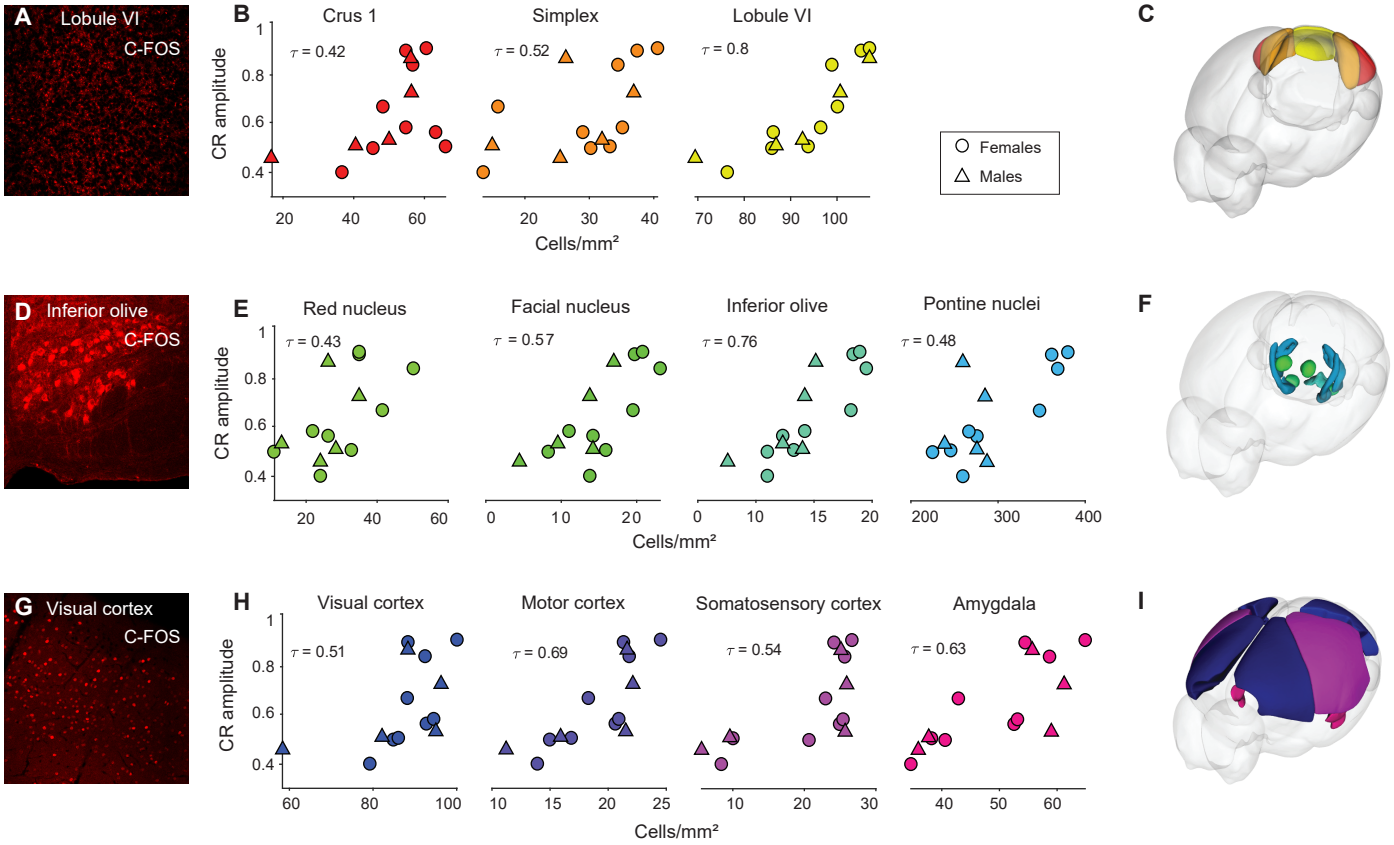

Figure 3: Learning scores correlate with C-fos expression. A) C-FOS positive granule cells in lobule VI in the cerebellum. B) Cerebellar areas with significant positive correlation between C-FOS positive cell density and CR amplitude (crus 1: tau = 0.42, p = 0.042, simplex: tau = 0.52, p = 0.009, lobule VI: tau = 0.8, p = 0.009). C) 3D model with significant areas highlighted. D) C-FOS positive cells in the inferior olive. E) Brainstem areas with significant positive correlation between C-FOS positive cell density and CR amplitude (Red nucleus: tau = 0.43, p = 0.041, Facial nucleus: tau =0 .57, p = 0.006, inferior olive: tau = 0.76, p = 0.0008, pontine nuclei: tau = 0.48, p = 0.021). F) 3D model with significant areas highlighted. G) C-FOS positive cells in the visual cortex. H) Cortical areas with significant positive correlation between C-FOS positive cell density and CR amplitude (visual cortex: tau = 0.51, p = 0.013, motor cortex: tau = 0.69, p = 0.0003, somatosensory cortex: 0.54, p = 0.007, amygdala: tau = 0.63, p: 0.001). I) 3D model with significant areas highlighted.

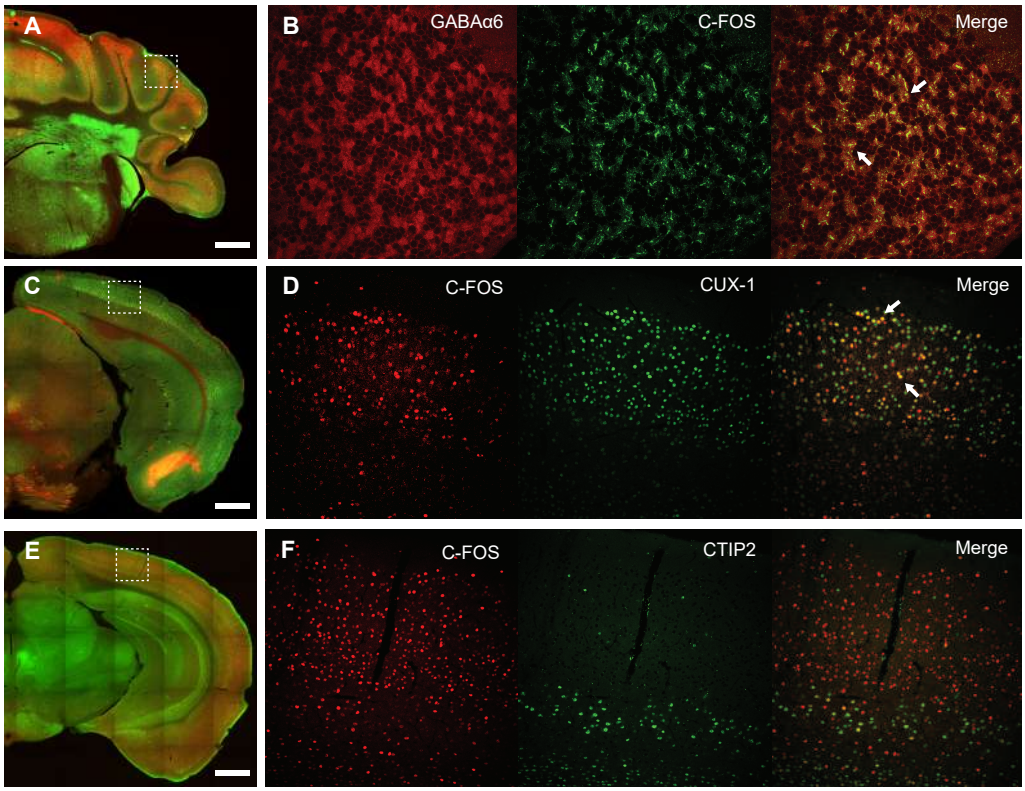

Figure 4: Co-localization of C-FOS and other neuronal markers. A) crus 1, example image used for quantification of C-FOS positive cells. Red: GABAα6, Green: 11C-FOS. Scale: 500 µm. B) Confocal image of the zoomed in area in A, 60x. C) visual cortex, example image used for quantification of C-FOS positive cells. Red: C-FOS, green: CUX-1. Scale: 500 µm. D) Confocal image of the zoomed in area in C, 40X. E) visual cortex, example image used for quantification of C-FOS positive cells. Red: C-FOS, green: CTIP2. Scale: 500 µm. F) Confocal image of the zoomed in areas in E, 40X.

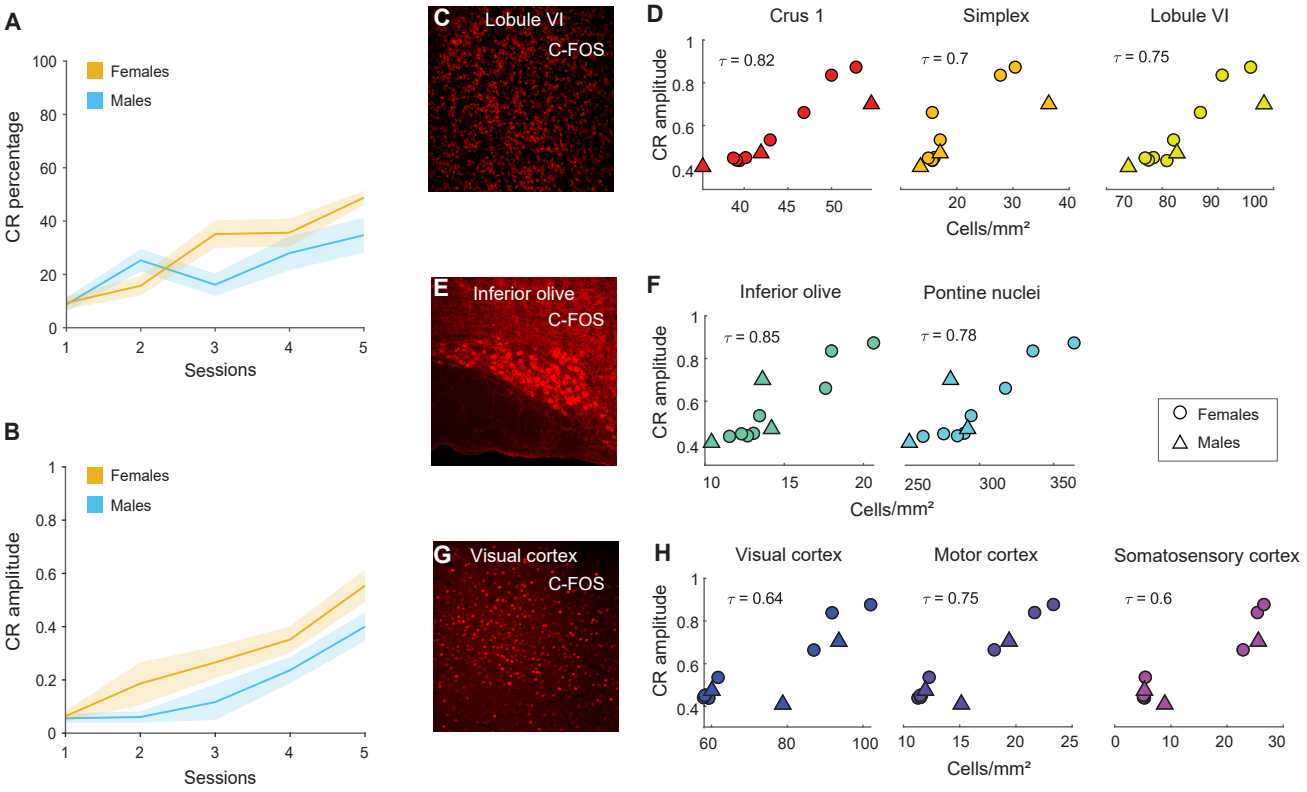

Figure 5: Correlation between learning scores and C-fos expression is consistent in B6CBAF1 mice. A) CR percentage in CS only trials over training sessions. Yellow: females (n = 9), cyan: males (n = 7). Shaded area: sem. B) CR amplitude in CS only trials over training sessions. Shaded area: sem C) C-FOS positive cells in the lobule VI. D) Cerebellar areas with significant positive correlation between C-FOS positive cell density and CR amplitude (crus 1: tau = 0.82, p = 0.0001, simplex: tau = 0.7, p = 0.005, lobule VI: tau = 0.75, p = 0.0007). E) C-FOS positive cells in the inferior olive. F) Brainstem areas with significant positive correlation between C-FOS positive cell density and CR amplitude (inferior olive: tau = 0.85, p = 0.0004, pontine nuclei: tau = 0.78, p = 0.0003). G) C-FOS positive cells in the visual cortex. H) Cortical areas with significant positive correlation between C-FOS positive cell density and CR amplitude (visual cortex: tau = 0.64, p = 0.0057, motor cortex: tau = 0.75, p = 0.0008, somatosensory cortex: tau = 0.6, p = 0.009).

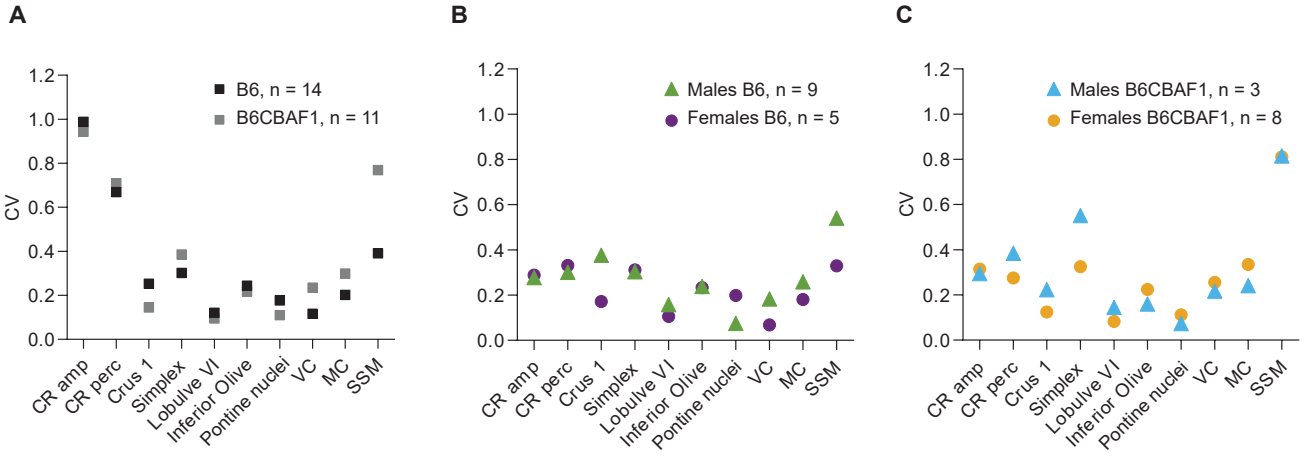

Figure 6: Strain and sex variability. Learners were selected if CR amplitude on session 5 > 0.4. CV = STD/mean. A) For strains: B6, n = 14, B6CBAF1, n = 11. B) For B6 mice: males B6, n = 9, females B6, n = 5. C) For B6CBAF1 mice: males, n = 3, females, n = 8. VC: visual cortex; MC: motor cortex; SSM: somato-sensory cortex; CV: Coefficients of variation.