Deletion of the alternative splice variant KChIP4a in midbrain dopamine neurons selectively enhances extinction learning

Authors: Kauê Machado Costa[1,2], Jochen Roeper[1*]

Affiliations:

[1]Insitute of Neurophysiology, Goethe-University Frankfurt, Frankfurt am Main, Germany

[2] International Max Planck Research School for Neural Circuits, Max Planck Institute for Brain Research, Frankfurt am Main, Germany

*Correspondence to: roeper@em.uni-frankfurt.de

Abstract: Inhibition of midbrain dopamine neurons is thought to underlie the signaling of events that are less rewarding than expected and drive learning based on these negative prediction errors. It has recently been shown that Kv4.3 channels influence the integration of inhibitory inputs in specific subpopulations of dopamine neurons. The functional properties of Kv4.3 channels are themselves strongly determined by the binding of auxiliary β-subunits; among them KChIP4a stands-out for its unique combination of modulatory effects. These include decreasing surface membrane trafficking and slowing inactivation kinetics. Therefore, we hypothesized that KChIP4a expression in dopamine neurons could play a crucial role in behavior, in particular by affecting the computation of negative prediction errors. We developed a mouse line where the alternative exon that codes for the KChIP4a splice variant was selectively deleted in midbrain dopamine neurons. In a reward-based reinforcement learning task, we observed that dopamine neuron-specific KChIP4a deletion selectively accelerated the rate of extinction learning, without impacting the acquisition of conditioned responses. We further found that this effect was due to a faster decrease in the initiation rate of goal-directed behaviors, and not faster increases in action disengagement. Furthermore, computational fitting of the behavioral data with a Rescorla-Wagner model confirmed that the observed phenotype was attributable to a selective increase in the learning rate from negative prediction errors. Finally, KChIP4a deletion did not affect performance in other dopamine-sensitive behavioral tasks that did not involve learning from disappointing events, including an absence of effects on working memory, locomotion and novelty preference. Taken together, our results demonstrate that an exon- and midbrain dopamine neuron-specific deletion of an A-type K^+ channel β-subunit leads to a selective gain of function in extinction learning.

One Sentence Summary:

Exon- and midbrain dopamine neuron-specific deletion of the Kv4 channel β-subunit KChIP4a selectively accelerates extinction learning

Main text

Introduction

Midbrain dopamine (DA) neurons have been causally implicated in the regulation of many forms of emotional and cognitive functions, including motivation, learning and the control of voluntary movements. One of the ways in which DA neuron activity is thought to influence these neural processes is by the coding of reward prediction errors, i.e. the difference between expected and experienced stimuli in the value domain (*1*). It has been demonstrated, across multiple behavioral contingencies, experimental methods and organisms, that some DA neurons respond to unexpected rewards (positive prediction errors) with phasic increases in firing and DA release, which are attenuated when rewards are predicted by conditioned stimuli or operant actions (*1–3*). Conversely, when an expected reward is omitted or reduced (negative prediction errors) DA neurons respond with phasic decreases (or "pauses") in firing from their baseline. While there is controversy on whether DA neurons are capable of symmetrically signaling positive and negative prediction errors (*4–8*), and on how the dynamics of DA neuron firing pauses are related to negative prediction error signaling (*6, 9, 10*), a number of recent studies have demonstrated that transient inhibition of DA neuron firing is causally implicated in learning from negative prediction errors (*9, 11–13*).

But what factors control the integration of inhibitory input by these neurons? In addition to the circuit-determined timing of pre-synaptic release and the properties of inhibitory synapses onto DA neurons, it is also likely that cell-autonomous biophysical properties of DA neurons contribute to integration of inhibition. Previous studies have demonstrated that different subpopulations of midbrain DA neurons possess different biophysical properties; specifically, DA neurons in the ventral tegmental area (VTA) that project to the pre-frontal cortex, amygdala and the medial shell and core subregions of the nucleus accumbens (NAcc) display longer rebound delays from hyperpolarizing inhibition than DA neurons that project to the dorsal striatum and the lateral shell of the NAcc (*14*). This would indicate that the former group of DA neurons possesses an intrinsic biophysical mechanism that extends firing pauses beyond the inhibitory synaptic event. It was recently shown that the pharmacological blockade of A-type K^+ currents, which are mediated by Kv4.3 channel complexes (*15*), fully abolished the augmented rebound delay observed in mesolimbic DA neurons (*16*). Likewise, it was demonstrated that the

duration of rebound delays in midbrain DA neurons were inversely correlated with the inactivation kinetics of Kv4.3 channels, i.e. the slower the inactivation of A-type currents, the longer the rebound delay. Thus, modifiers of Kv4.3 channel function are candidates to control inhibition integration and consequently coding of negative prediction errors in DA neurons.

Kv4 channel function is strongly controlled by the binding of modulatory β-subunits, including KChIPs (K^+ channel interacting proteins), which can induce significant changes in channel properties such as surface trafficking and inactivation kinetics (*17–19*). There are four genes coding for distinct KChIP proteins (KCNIP 1-4), each with multiple splice variants. However, one of those, KChIP4a (a splice variant of KChIP4) stands out; it possesses a special K^+ channel inhibitory domain (KID), which has a unique effect of dramatically slowing down the inactivation of Kv4 channels (*17*, *20–22*). In addition, the KChIP4a KID suppresses membrane trafficking of Kv4 channels due to the presence on an ER-retention motif (*17*, *21*, *22*). The size of each effect depends on the relative ratio between KChIP4a and other KChIP isoforms within the native channel complex (*21*, *23*, *24*). Based on these previous findings, we tested whether the presence of KChIP4a in DA neurons affects learning from negative reward prediction errors.

Results

We developed a mouse line where the exon coding for the KID of KCNIP4 (i.e. exon 3) was floxed. Subsequently, we crossed this line with a DAT-cre KI line (see Material and Methods) for selective deletion of KChip4a in midbrain DA neurons (Ex3d, for "exon 3 deletion"). We subjected these mice (males and females, > 8 weeks old, N = 9) and their littermate DAT-cre KI controls (CTRL, N=20) to a battery of behavioral tests. To test of main hypothesis, we carried out a reinforcement learning task where water restricted mice learned that an auditory cue signaled the availability of sugar water (10% sucrose) in a reward port for 11 daily sessions (acquisition), followed by 6 daily sessions where reward was omitted (extinction) (*12*). We did not observe a significant difference between the genotypes in any of the measured behavioral variables during acquisition (Fig 1A-D). In contrast, we found that Ex3d and CTRL mice differed dramatically in terms of extinction learning dynamics. The Ex3d group displayed a significantly accelerated reduction in the time they spent in the reward port during cue

presentation (interaction effect, $F_{5, 135} = 3.2$, $P < 0.01$; Fig 1A) and a faster increase of response latency (interaction effect, $F_{5, 135} = 2.615$, $P < 0.05$; Fig 1B) across extinction sessions.

We then asked whether the effect on the time in port metric during extinction was due to reduced initiation of or faster disengagement from goal directed behaviors. We inferred this by quantifying the average number and duration of directed head entries into the reward port (initiated during cue presentation), which can respectively be interpreted as indicators of the probability of initiating a reward-seeking behavior and the average time to disengage from it. During extinction, we found that only the rate of change in the number of head entries was significantly faster in the Ex3d group in relation to controls (interaction effect, $F_{5, 135} = 3.45$, $P < 0.01$; genotype effect, CTRL > Ex3d, $F_{1, 27} = 4.82$, $P < 0.05$ (Fig 1C). Conversely, there was no genotype effect on head entry duration dynamics (Fig 1D). This indicated that Kchip4a deletion only affected the probability of initiating goal directed behaviors during extinction.

To verify whether we were observing a direct effect on learning (which would be compatible with an impact on reinforcement learning) or a reduction of responding in the absence of direct reinforcement (more compatible with an impact on motivational processes), we analyzed the trial-by-trial behavior measures in the first extinction session. If the Ex3d mice already displayed a reduced time in port on the first extinction trial (when the animals were first confronted with an unexpected reward omission), then the genotype effect would be more likely attributable to decreased motivation (*25*). However, we found that Ex3d and CTRL mice were statistically indistinguishable in their responses during the first extinction trial in relation to all measured variables (Fig 1A-D). Only in later trials of the first extinction session, genotype differences emerged in response latency (interaction effect, $_{F9, 243} = 2.01$, $P < 0.05$; genotype effect, Ex3d > DAT-cre KI, $F_{1, 27} = 7.51$, $P < 0.05$; Fig 1B) and in the number of head entries (genotype effect, CTRL > Ex3d, $F_{1, 27} = 5.98$, $P < 0.05$; Fig 1C). We also identified a strong trend for a gradual divergence between genotypes, starting on the second trial, for the time in port metric (genotype effect, CTRL > Ex3d, $F_{1, 27} = 3.51$, $P = 0.07$; Fig 1A). Taken together, these measures confirm that the Ex3d effect progressively developed as the animals learned that the CS no longer predicted US availability, strongly suggesting a selective effect on learning from expected reward omission, i.e. negative prediction errors.

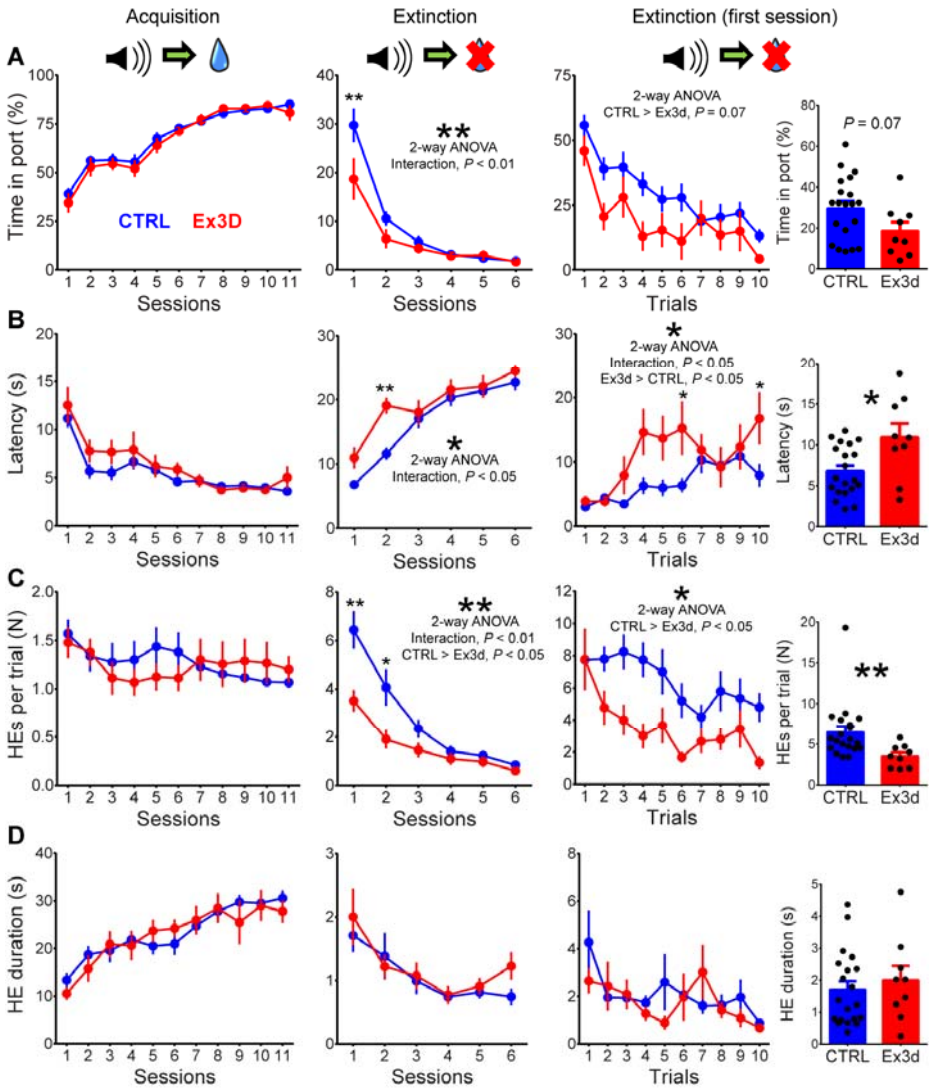

Fig 1. Deletion of KChIP4a in midbrain DA neurons selectively accelerates extinction learning.

Note that, in all measured variables, there was no significant difference between genotypes during acquisition or in the first trial of the first session of extinction. A: Time spent in port during CS presentation on a session-by-session basis during acquisition and extinction, as well as on a trial-by-trial and session average basis on the first extinction session. Note the faster response extinction for Ex3d mice in relation to controls, including the significant pair-wise difference during the first extinction session (Sidak's multiple comparisons test). Also note the strong trend towards faster trial-by-trial extinction in the first extinction session. B: Latency to response during CS presentation for all phases of the reinforcement learning task. Note the faster increase in this measure during extinction in Ex3d mice, including the significant pair-wise difference during the second extinction session. Importantly, there was also faster trial-by-trial learning according to this measure in the first extinction session, including significant pair-wise differences in the sixth and tenth trials of the session, as well as in the session average. C: Number of directed head entries into the reward port during CS presentation. Note the faster response extinction for Ex3d mice in relation to CTRL, including the significant pair-wise difference during the first and second extinction sessions. Also note that there was also a faster decrease in the number of head entries on a trial-by-trial basis in the first session of extinction for Ex3d mice. D: Duration of directed head entries into the reward port during CS presentation. Note the lack of a genotype effect during all phases of acquisition and extinction. *P < 0.05; **P < 0.01.

In order to formally test the hypothesis that only negative prediction errors were affected, we fitted the trial-by-trial time in port values for both Ex3d and CTRL groups with a modified Rescorla-Wagner model which assumed different learning rates for positive (α_P) and negative (α_N) prediction error-based learning, and that the time in port was a linear readout of the associative strength (V) (*13*, *26*, *27*). This model was able to successfully fit the behavioral data, including the different learning kinetics during extinction (Fig 2A). Comparisons of the learning rates for the best fits (lowest absolute residual) showed that indeed the Ex3d group had significantly higher α_N values, i.e. faster learning from negative prediction errors ($P < 0.05$), without any change in α_P (Fig 2B). Importantly, empirically derived model constraint parameters (V_0, V_{max} and V_{min}), as well as the goodness of fit (measured by the residuals between the experimental data and the model output), were similar for both genotypes (Fig 2C). These results support the interpretation that that the effect of deleting KChIP4a in DA neurons can be explained by a selective increase in learning from negative prediction errors.

In order to test the selectivity of the Ex3d effect, we tested the mice in additional DA-sensitive behavioral tasks, namely open field exploration (Fig 3), novel object preference (Fig 4), hole-board exploration (Fig 5) and spontaneous alternation in the plus maze (Fig 6). There was no difference between the genotypes in any variable determined from these tasks, indicating that deleting KChIP4a in DA neurons did neither impact locomotion, anxiety, spontaneous exploratory behavior, novelty preference, short-term object memory retention nor working memory, but rather selectively enhanced extinction learning in a reward-based reinforcement learning task.

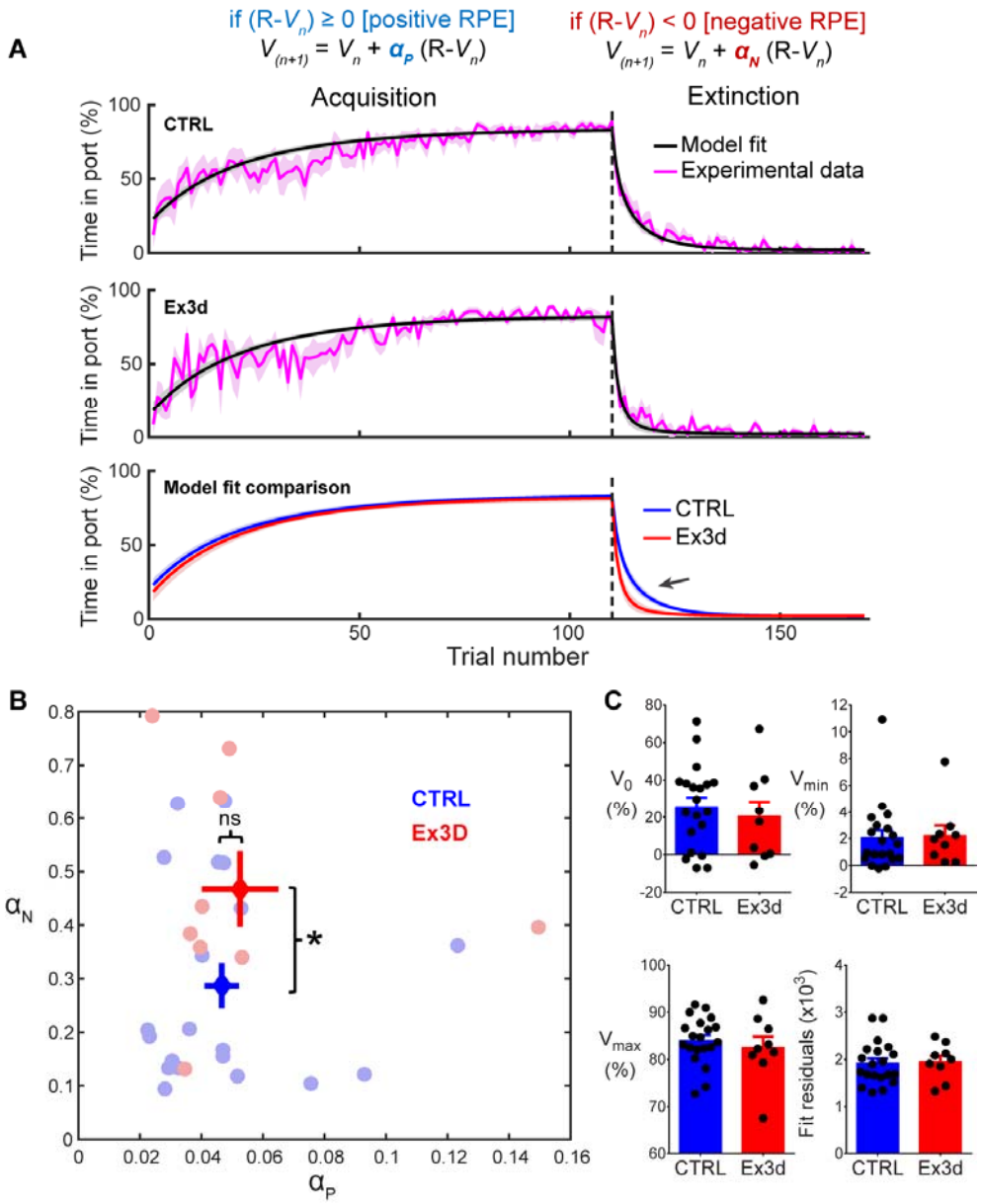

Fig 2. Model fitting reveals that the loss of KChIP4a in midbrain DA neurons induces a selective increase in learning from negative prediction errors.

A: Overlay of model fits and trial-by-trial experimental data for the Ex3d and CTRL groups. Note that the model tracks the behavioral data with relatively good accuracy in both groups, and that the model fits overlap almost completely during acquisition, but diverge in extinction, reflecting the faster extinction learning kinetics of the Ex3d mice in relation to controls (black arrow). B: Combination of best fit α_P and α_N values for Ex3d (red) and CTRL (blue) groups. Faded circles represent individual values. Large solid crosses indicate the mean ± SEM of each group. Note the selective increase in α_N (learning from negative prediction errors) in Ex3d in relation to littermate controls. C: V_0 (initial assumed associative strength), V_{min} (minimal associative strength, V_{max} (maximal associative strength), and the residuals of the best model fit for Ex3d and CTRL groups. Note that there was no significant difference observed between the compared groups (Mann Whitney tests, $P > 0.05$). *$P < 0.05$.

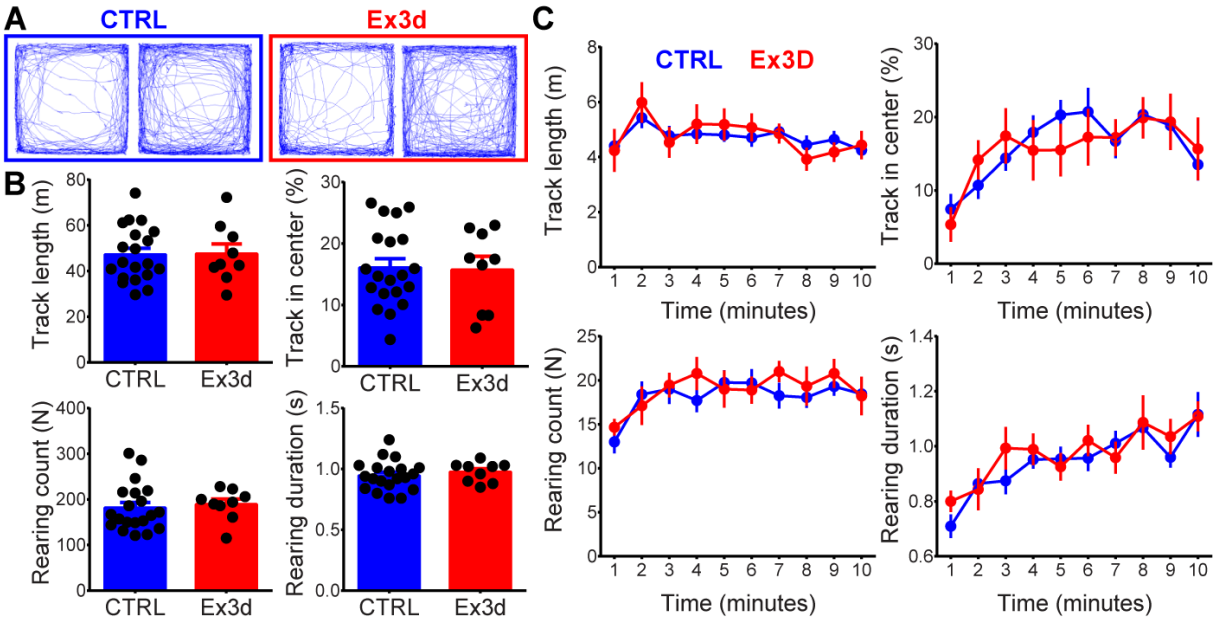

Fig 3. KChIP4a deletion in DA midbrain neurons does not affect locomotion and rearing in the open field.

A: representative tracks from mice of both genotypes. B: averages of the total track length, proportion of track length in the center of the open field, number and duration of rearings during the 10 minutes of the task. C: Same variables as in B, but averaged across every minute of the 10 minute task. Note the absence of a significant genotype difference for all variables (unpaired T tests and two-way repeated measures ANOVA, $P > 0.05$).

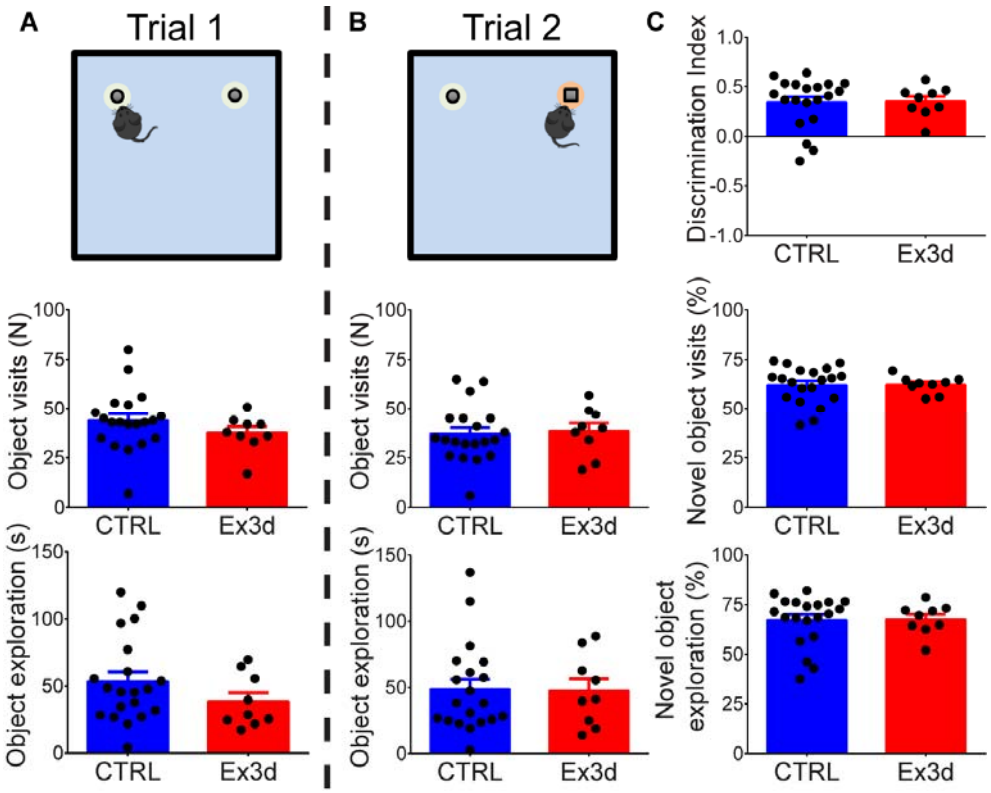

Fig 4. KChIP4a deletion in DA midbrain neurons does not affect performance in the novel object recognition task.

A: Cartoon representation of object disposition during trial 1 of the task (note that both objects are similar) and the total number of object visits and the summed duration of object exploration bouts. B: Cartoon representation of object disposition during trial 2 of the task (note that one of the objects has been replaced by another object with different shape) and the total number of object visits, the summed duration of object exploration bouts, the proportion of object visits and exploration time directed at the new object. Note the absence of a genotype effect on the measured variables (unpaired T tests, $P > 0.05$).

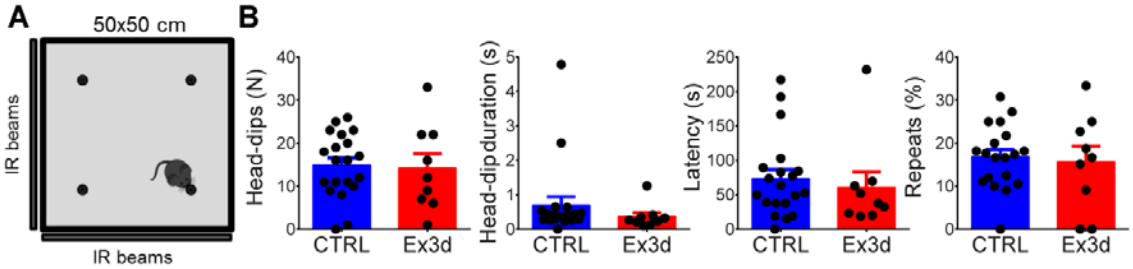

Fig 5. Loss of KChIP4a in DA midbrain neurons does not affect behavior in the hole board arena.

A: Cartoon schematic of the hole board arena. B: Total number of head dips, total summed head-dip durations, latency to first dip and of repeat head-dips. Note the absence of a genotype effect on these variables (Mann Whitney tests, P > 0.05).

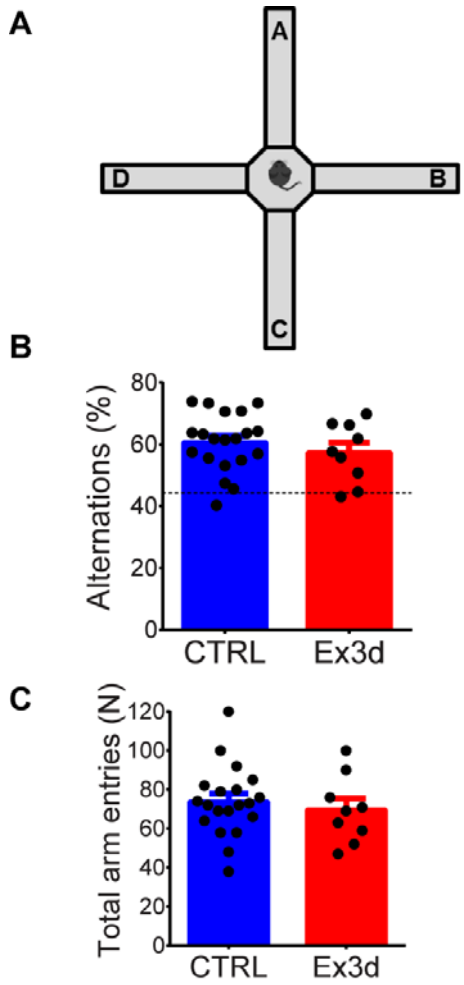

Fig 6. KChIP4a deletion in DA midbrain neurons does not impact working memory-guided spontaneous alternations in the plus maze.

A: Cartoon schematic of the plus maze. Spontaneous alternations were evaluated by tracking the order of entry into each arm. B: proportion of spontaneous alternations (in relation to the total possible alternations). Dotted line indicates chance performance level. Note that most mice in both groups performed well above chance level. C: Total arm entries performed during the task. Note the absence of a genotype effect (unpaired T tests, $P > 0.05$).

Discussion

We found that the selective removal of the splice variant KChIP4a from midbrain DA neurons had a highly specific effect on the learning rate during extinction of a conditioned response. It did not affect acquisition learning or other behavioral processes such as locomotion and working memory. Moreover, this effect on extinction learning was circumscript to changes in the rate with which animals initiated reward-seeking behaviors, without affecting their termination dynamics. Fitting of the behavior to a reinforcement learning model confirmed that this phenotype is compatible with a selective increase in the learning rate from negative prediction errors. Given the known functions of KChIP4a and the nature of reward prediction error signaling by DA neurons, it is likely that the behavioral phenotype of Ex3d mice is attributable to changes in the properties of Kv4.3 channel complexes and a consequent modulation of the integration of inhibitory inputs (*1*, *16*, *17*).

Several studies have suggested that longer pauses or stronger reductions in the firing rate of DA neurons, as well as stronger reductions in DA release or DA axon Ca^{++} transients in the ventral striatum, are correlated with larger negative prediction error signaling (*5*, *6*, *13*, *28*). Furthermore, lesions of the lateral habenula, whose neurons are excited by disappointing and aversive outcomes and disynaptically inhibit DA neurons (*29*), resulted in blunted firing rate decreases in DA neurons during negative outcomes and affect behavior in a manner that is congruent with a selective reduction in negative prediction error signaling (*13*). Detailed computational modelling of the post-synaptic effects of DA release on D1 and D2 receptors (D1R and D2R, respectively) in striatal projection neurons (SPNs) have predicted that the duration of DA neuron firing pauses results in a linear reduction in D2R occupancy, which would release indirect SPNs from D2R-mediated inhibition. This establishes a potential single-cell mechanism by which indirect pathway SPNs read out pauses in DA neuron firing during negative prediction error-based learning (*8*).

Within this framework, the most likely mechanism for the behavioral effects in Ex3D mice are a gain-of-function of Kv4.3 channels, likely to be mediated by increased surface expression. This would amplify synaptically-triggered inhibitory firing pauses in DA neurons, leading to larger reductions in DA release during the evaluation of negative outcomes, and a consequently greater activation of indirect pathway SPNs, ultimately resulting in an acceleration of learning from negative prediction errors.

The fact that the selective deletion of KChIP4a in DA midbrain neurons produces a robust behavioral phenotype indicates that manipulations of this protein are highly penetrant across multiple scales of biological organization (from cellular physiology to behavior). The specificity of the phenotype further suggests that these changes are restricted to a functionally distinct subpopulation of DA neurons. Perhaps this is due to both the unique nature of KChIP4a, and the fact that some subpopulations of DA neurons lack homeostatic mechanisms for buffering the effects of KChIP4a deletion. Given that rebound from hyperpolarization in DA midbrain neurons is typically controlled by the balance of A-type and HCN currents, the DA neurons subpopulations most likely to be affected by KChIP4a deletion would be those that lack HCN currents, including neurons that project to the medial shell and core of the NAcc (*14*, *16*).

An alternative interpretation of our results is that the DA-neuron selective deletion of KChIP4a results primarily in an acceleration of Kv4.3 channel inactivation, due to the absence of the KID. This would presumably decrease K charge transfer after transient hyperpolarizations, which would accelerate the rebound slope, effectively neutralizing the intrinsic inhibition-amplifying mechanism observed in atypical VTA DA neurons, and potentially reducing the duration of firing pauses in these cells *in vivo* (*14*, *16*). While this may seem at first completely at odds with our behavioral results, a very recent pair of studies, predicated on the precise optogenetic inhibition of DA neurons during reinforcement learning tasks, suggested that negative prediction errors are more effectively signaled by *multiple short pauses* rather than by *single long pauses* in DA neuron firing (*9*, *11*). If this is true, than perhaps the phenotype observed in the Ex3d mice could be due not to an increase in pause duration *per se*, but to a transformation of single inhibitory events into multiple shorter pauses.

In summary, we have identified KChIP4a in DA neurons as a selective molecular modulator of learning from negative prediction errors, singling out this subunit as a likely contributor to the computation of learning signals and a homeostatic controller of adaptive behaviors.

Materials and Methods

Animals

All animal procedures described in this thesis were approved by the German Regierungspräsidium Gießen and Darmstadt (license numbers: V54-19c 20/15 –FU/1100 & FU1203). Mice were bred and housed until 8 weeks of age at MFD diagnostics (Mainz, Germany). Animals were then maintained under a 12 h dark/light cycle and housed in groups of two to four animals with food (R/M-keeping, Ssniff, Germany) and tap water available *ad libitum*, except when they underwent water restriction. Nesting material and a red acrylic glass shelter (mouse house, Tecniplast, Germany) were used as enrichment.

The DA-neuron selective KChIP4a exon 3 deletion (Ex3d) mouse line

The mouse KChIP4 gene, KCNIP4, is located on chromosome 5 (B3, 48.39 – 49.52 Mb), is 1135 kb in length and has 14 exons, with ATG translation initiation codons in exons 1, 2, 3, 5, and 8, and the STOP codon located in exon 14 (*30*). KChIP4a (a.k.a. KChIP4.4) is a 229 amino acid protein that corresponds to the alternative splicing of exons 3 through 8b of KChIP4, with exon 3 coding for the KID (*20*, *31*).

A new transgenic mouse line was developed for assessing the specific function of the KChIP4a KID using the cre-lox system. Exon 3 of the KCNIP4 gene (which encodes the KID) was flanked with loxP sites using a vector designed to display two homology regions in a C57BL/6N genetic background, including a short homology region of 1795 bp and a long region of 5858 bp. This vector was electroporated into C57BL/6N embryonic stem cells and screened clones were used for the generation of mice with homozygotic floxing of KCNIP4 exon 3 ($KChIP4\text{-}Ex3^{lox/lox}$). These mice were then crossed with a line with heterozygous DAT-cre knock-in ($DAT\text{-}cre^{+/-}$ or *DAT-cre KI*) mice. The offspring of this crossing were then interbred, leading to the production of mice where exon 3 of the KCNIP4 gene was selectively excised from DAT-expressing neurons (Ex3d; $DAT\text{-}cre^{+/-}$ / $KChIP4\text{-}Ex3^{flox/flox}$) and DAT-cre KI control littermates (CTRL; $DAT\text{-}cre^{+/-}$ / $KChIP4\text{-}Ex^{WT/WT}$).

Genetic identity of all KChIP4-Ex3$^{flox/flox}$ mice was confirmed with polymerase chain reaction (PCR) genotyping throughout the breeding process. For detecting the conditional deletion allele, the following primers were used:

5-TAG TTA TGA CAA GAC AGG AGC TAG TAC CAC TAA GC-3

and

5-GAA CTG GAC TGA AGC AAA ACA AAA CAC G-3.

These primers target the flanking region of the FRT site and the PCR product lengths were 385 bp for the WT allele and 477 bp for the conditional deletion allele. The PCR protocols were run with 1 cycle at 94 °C for 120s (denaturing), 35 cycles of 94 °C for 30s (denaturing), 65 °C for 30s (annealing) and 68°C for 30s (extension), followed by 1 cycle at 68 °C for 480s (completion).

Reinforcement learning task

Male and female Ex3d (N = 10) and CTRL (N = 20) mice, over 8 weeks old, were genotyped and individually identified by the implantation of a subcutaneous transponder microchip (1.4 mm x 9 mm ISO FDX-B glass transponder, Planet ID, Germany). The experimenter was blind to the genotype of the mice and the order in which animals were tested was pseudo-randomized by a staff technician that was not directly involved in the design of the study. Behavioral comparisons were performed between Ex3d and DAT-cre KI mice. One Ex3d mouse showed erratic behavior within the conditioning box and was excluded from data analysis (final N of Ex3d group = 9).

Mice were motivated by water restriction (~85% of their initial body weight) and were rewarded with a solution of 10% sucrose in tap water. Daily water rations varied between 1 and 1.5 mL depending on the animal's weight on that day, with a set target of 85% of the initial body weight. Except during experimental sessions, water was always delivered in a cup placed in their home cage. Mice were also closely monitored in order to ensure that the water supply was consumed and not spilt over or contaminated and the health of the water restricted mice was also evaluated daily (*32*). This protocol was very successful in maintaining stable body weights

around the desired target for multiple weeks and never resulted in signs of overt dehydration or ill health (e.g. hunched posture, abnormal gait or ruffled fur).

The full experimental paradigm spanned 24 days. On the first four days, mice were submitted only to water restriction. In the following two days, they were tamed, i.e. gently handled (held up above their home cage on a spread palm, with no constriction or entrapment by the experimenter) until they no longer tried to escape from the experimenters hand, showed no overt signs of stress and anxiety and readily drank a portion of liquid reward (0.2 ml) given by the experimenter via a syringe while being held. The following day, the animals where placed inside the operant chamber with the reward port removed for 50 minutes in order to acclimate to the experimental conditions. After this time they were returned to their home cage and given their daily ration of water. The day after that, mice underwent shaping, i.e. they were placed in the operant chamber for another 50 minutes, now with the reward port present, and at semi-random time intervals (mean of 60 seconds, varying between 30 and 90 seconds), a reward portion (16.66 µL) was delivered at the port, with no cue to its delivery. The following day, mice were submitted to the conditioning task.

The task used in this study is based on the study by Steinberg et al. (*12*). In this paradigm, animals learned to associate a CS (sound tone pulsed at 3Hz – 0.1s on/0.2 s off – at 70 db) to the availability of reward in the reward port. Each session consisted of ten trials (mean ITI of 4 minutes, varying between 1.5 minutes and 6.5 minutes) in which the auditory cue was on for 30 seconds. Animals could trigger reward delivery to the port by entering it during CS presentation. Rewards were delivered in a cycle of 2 s reward delivery (16.66 µL) followed by a 3 s consumption interval. Delivery was continuous for as long as the animal kept its head in the port during CS presentation. This allowed for a maximum of 6 rewards per trial and a maximum of 60 rewards per session (a total of ~1 mL or reward in the task per day). Additional water supplementation, when needed to complete the animal's daily water ration, was provided in the home cage as described in the previous paragraph.

This acquisition phase lasted for 11 daily sessions. After acquisition, extinction of the conditioned response was tested by consistently omitting the reward during CS presentation for six daily sessions. After extinction, animals were returned to their home cage and received water ad libitum for at least three days before being submitted to other behavioral tasks. All animals

recovered their initial body weight in this period of time and showed no long-term adverse effects from water restriction.

Performance in this task was quantified by the total time the animals spent in the port during the cued trials and the latency to enter the port after cue onset. Time in port was normalized both as a percentage of total reward availability time and also by subtracting the amount of time the animal spend in the port 30 s before cue onset (*12*). This means that CS-US associations were quantified by the animal responding both quickly and selectively to the cue. In addition, the time stamps of every head entry and head retraction into the port were recorded, which allowed the quantification of the dynamics of head entries during cue presentation and during the ITI. Latency was quantified as the time between cue initiation and the first head entry into the reward port; if animals did not respond in a trial, the maximal possible latency value (30 s) was ascribed to that trial. The dynamics of head entries during extinction were used to infer how the animal adapts it's behavioral response in the absence of an expected reinforcement, i.e. whether it perform less head entries (a proxy of the initiation rate of reward-seeking behaviors), head entries that are just shorter in length (indicating a faster disengagement from reward-seeking behaviors) or a combination of both.

Fitting of behavioral data with a modified Rescorla-Wagner model

A modified Rescorla-Wagner model, which applied two different learning rates for positive (α_P) and negative (α_N) prediction errors, was used in order to formally quantify differences in learning between Ex3d and CTRL genotype groups (*13*, *26*, *27*). In detail:

$$if\ (R - V_n)\ \geq 0\ [positive\ RPE]$$

$$V_{n+1} = V_n + \boldsymbol{\alpha_P}(R - V_n)$$

$$if\ (R - V_n) < 0\ [negative\ RPE]$$

$$V_{n+1} = V_n + \boldsymbol{\alpha_N}(R - V_n)$$

where

V = associative strength between CS and US

α_P and α_N = learning rates from positive and negative prediction errors, respectively

R = maximal associative strength of the US (reward value)

n = trial index

All modelling and fitting was done with custom MatLab® code. The percentage of time in port metric was assumed to be a linear read-out of V. The limits of V in relation to the empirical data were bound by the initial value of V before the task (V_0), the lower and upper limits of associative strength readout (V_{min} and V_{max}, respectively). These parameters were derived empirically from each animal, with V_0 being set as the average response in the first three trials of acquisition, and V_{min} and V_{max} being set as the average responses in the last two sessions of acquisition and extinction, respectively. Fitting of learning rates was performed with least absolute residuals method by implementing the *fminsearch* MatLab® function (*13*). Goodness of fit was inferred from the residual values of the optimal fit for each animal.

Open field

Spontaneous locomotor and exploratory activity (track length, wall distance, time in center and number of rearings) were evaluated in an open field arena (a lidless box measuring 52 × 52 cm, under red illumination at 3 lux) using a video tracking system (Viewer II, Biobserve, Germany) as described in a previous study from our group (*33*). For a comprehensive phenotyping of mutant mice, the total track length, the time spent and track length within the center area of the arena (defined as a 30x30 cm square zone with all sides equidistant from the walls), as well as the number and duration of rearings (recorded via infra-red beam breaks at a height of 4.5 cm, and defined by being at least 200 ms long and two subsequent rearings had to be at least 80 ms apart), were evaluated. After ten minutes, the animals were returned to their home cages for at least two minutes, and then subjected to the novel object recognition test.

Novel object recognition

A novel object exploration and preference task was used to assess the mice's recognition, memory retention and preference for novel stimuli (*34*, *35*). In this task, animals were placed in the same open field arena described in the previous section, after the open field test (which also served as a habituation for the novel object recognition test), but within the arena two identical objects (stainless steel cylinders; 3 cm diameter x 6 cm height) were placed at equal lengths from each other and at 15 cm from the upper left and right corners. The animals were allowed to freely explore the arena and the objects for 10 minutes (trial 1), after which they were removed from the arena. Subsequently, one of the objects was replaced by a different, novel object (plastic coated rectangular prism; 3x3 cm base x 6 cm height), and the animals were again allowed to explore the arena and the objects (trial 2).

Object recognition was analyzed using Biobserve's Object Recognition plug-in, and object interaction events were defined as periods in which a mouse was directly facing the object (snout directed to the object within a 180° angle) at a distance shorter than 3 cm. Both the number and duration of these interaction events were quantified for both trials of the task. Exploration dynamics (number and duration of objected-directed exploration) were analyzed for both trials, as well as the differences between exploration of an already explored ("old") and novel ("new") object in trial 2, were analyzed. For the quantification of novel object discrimination and preference, a discrimination index was calculated by dividing the difference between the time spent exploring the novel object and the time spent exploring the familiar object by the sum of these two measures (*34*, *35*).

$$DI = \frac{T_{New} - T_{Old}}{T_{New} + T_{Old}}$$

where

DI = discrimination index

T_{New} = time spent exploring the novel object

T_{Old} = time spent exploring the familiar object

Hole board

Exploration of holes, a naturalistic behavior for mice, was evaluated with the hole board task (*36*). In this task, mice were put in an open field arena similar to the one used for the open field test, but with 4 to 16 circular holes (2 cm in diameter) on its floor (Actimot2, TSE systems). When confronted with such an arena, mice tend to spontaneously check the holes by dipping their heads in them. The performance of head dips was recorded with infra-red beam breaks placed immediately under the inferior surface of the floor board. The latency to the first head dip, as well as the total number and duration of head dips, were quantified. The percentage of repeated head dips (two dips performed sequentially into the same hole) as well as the percentage of dips into the preferred hole (the hole that was most explored during the session) were also quantified and compared between genotypes.

Spontaneous alternation in the plus maze

Working memory performance was quantified between genotypes using the spontaneous alternation in the plus maze (*37*). Mice were placed in a plastic plus maze (four equidistant 35 x 4.5 cm arms radiating at 90° angles from a circular central arena with 10 cm diameter; walls were 15 cm in height) and allowed free exploration for 12 minutes (*38*). An arm entry was defined as when the animal enters an arm with all its four paws. Each session was recorded with a video camera and quantification of arm entries was made through visual analysis of the videos. A spontaneous alternation was marked when the animal explored four different arms in five consecutive arm entries, and the proportion of spontaneous alternations was quantified as the number of real alternations divided by the total possible number of alternations (sum of all arm entries minus four); chance performance in this task is calculated to be 44% (*37*, *38*).

Statistical analyses

For all two-group comparisons, data were tested for normality with the Kolmogorov-Smirnov test. If both distributions were Gaussian, differences were analyzed using two-tailed T-tests; otherwise, the two-tailed Mann-Whitney test was used. Comparisons between groups over multiple trials or sessions were evaluated using two-way repeated measures

ANOVA with Sisak's multiple comparisons *post-hoc* tests. Statistical significance was set at $P < 0.05$ for all comparisons.

Acknowledgments:

The authors would like to thank Beatrice Fischer, Jasmine Sonntag and Felicia Müller-Braun for technical support. This research was supported by DFG CRC 1080 (A11). KMC received a one year fellowship and financial support from the Max Planck Society (International Max Planck Research School for Neural Circuits).

Additional Author notes:

Author contributions: KMC and JR designed the research project, interpreted the results and wrote the manuscript. KMC conducted the experiments and analyzed the data.

References

1. W. Schultz, P. Dayan, P. R. Montague, A Neural Substrate of Prediction and Reward. *Science (80-.).* **275**, 1593–1599 (1997).
2. J. Y. Cohen, S. Haesler, L. Vong, B. B. Lowell, N. Uchida, Neuron-type-specific signals for reward and punishment in the ventral tegmental area. *Nature*. **482**, 85–88 (2012).
3. W. Schultz, R. M. Carelli, R. M. Wightman, Phasic dopamine signals: from subjective reward value to formal economic utility. *Curr. Opin. Behav. Sci.* **5**, 147–154 (2015).
4. Y. Niv, M. O. Duff, P. Dayan, Dopamine, uncertainty and TD learning. *Behav. Brain Funct.* **1**, 6 (2005).
5. A. S. Hart, R. B. Rutledge, P. W. Glimcher, P. E. M. Phillips, Phasic dopamine release in the rat nucleus accumbens symmetrically encodes a reward prediction error term. *J. Neurosci.* **34**, 698–704 (2014).
6. H. M. Bayer, B. Lau, P. W. Glimcher, Statistics of midbrain dopamine neuron spike trains in the awake primate. *J. Neurophysiol.* **98**, 1428–1439 (2007).
7. H. M. Bayer *et al.*, Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron*. **47**, 129–41 (2005).
8. J. K. Dreyer, K. F. Herrik, R. W. Berg, J. D. Hounsgaard, Influence of phasic and tonic dopamine release on receptor activation. *J. Neurosci.* **30**, 14273–14283 (2010).
9. C. Y. Chang, M. Gardner, M. G. Di Tillio, G. Schoenbaum, Optogenetic Blockade of Dopamine Transients Prevents Learning Induced by Changes in Reward Features. *Curr. Biol.* **27**, 3480–3486.e3 (2017).
10. H. Matsumoto, J. Tian, N. Uchida, M. Watabe-Uchida, Midbrain dopamine neurons signal aversion in a reward-context-dependent manner. *Elife*. **5** (2016), doi:10.7554/eLife.17328.
11. C. Y. Chang *et al.*, Brief optogenetic inhibition of dopamine neurons mimics endogenous negative reward prediction errors. *Nat. Neurosci.* **19**, 1–8 (2015).
12. E. E. Steinberg *et al.*, A causal link between prediction errors, dopamine neurons and learning. *Nat Neurosci*. **16**, 966–973 (2013).
13. J. Tian, N. Uchida, Habenula Lesions Reveal that Multiple Mechanisms Underlie Dopamine Prediction Errors. *Neuron*. **87**, 1304–1316 (2015).
14. S. Lammel *et al.*, Unique Properties of Mesoprefrontal Neurons within a Dual Mesocorticolimbic Dopamine System. *Neuron*. **57**, 760–773 (2008).
15. B. Liss *et al.*, Tuning pacemaker frequency of individual dopaminergic neurons by Kv4.3L and KChip3.1 transcription. *EMBO J.* **20**, 5715–24 (2001).
16. R. A. Tarfa, R. C. Evans, Z. M. Khaliq, Enhanced sensitivity to hyperpolarizing inhibition in mesoaccumbal relative to nigrostriatal dopamine neuron subpopulations. *J. Neurosci.* (2017).
17. H. H. Jerng, P. J. Pfaffinger, Modulatory mechanisms and multiple functions of somatodendritic A-type K + channel auxiliary subunits. *Front. Cell. Neurosci.* **8**, 1–20 (2014).
18. H. H. Jerng, P. J. Pfaffinger, M. Covarrubias, Molecular physiology and modulation of somatodendritic A-type potassium channels. *Mol. Cell. Neurosci.* **27**, 343–369 (2004).
19. W. F. An *et al.*, Modulation of A-type potassium channels by a family of calcium sensors. *Nature*. **403**, 553–556 (2000).
20. M. H. Holmqvist *et al.*, Elimination of fast inactivation in Kv4 A-type potassium channels by an auxiliary subunit domain. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 1035–1040 (2002).

21. Y.-Q. Q. Tang *et al.*, Auxiliary KChIP4a suppresses A-type K+ current through Endoplasmic Reticulum (ER) retention and promoting closed-state inactivation of Kv4 channels. *J. Biol. Chem.* **288**, 14727–14741 (2013).
22. H. H. Jerng, P. J. Pfaffinger, Multiple Kv channel-interacting proteins contain an N-terminal transmembrane domain that regulates Kv4 channel trafficking and gating. *J. Biol. Chem.* **283**, 36046–36059 (2008).
23. M. Kitazawa, Y. Kubo, K. Nakajo, The stoichiometry and biophysical properties of the Kv4 potassium channel complex with K+ channel-interacting protein (KChIP) subunits are variable, depending on the relative expression level. *J. Biol. Chem.* **289**, 17597–609 (2014).
24. J. Zhou *et al.*, Different KChIPs Compete for Heteromultimeric Assembly with Pore-Forming Kv4 Subunits. *Biophys. J.* **108**, 2658–2669 (2015).
25. K. C. Berridge, From prediction error to incentive salience: mesolimbic computation of reward motivation. *Eur. J. Neurosci.* **35**, 1124–43 (2012).
26. K. A. Raczka *et al.*, Empirical support for an involvement of the mesostriatal dopamine system in human fear extinction. *Transl. Psychiatry*. **1**, e12 (2011).
27. Rescorla, A. Wagner, A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement (1972).
28. W. Menegas *et al.*, Opposite initialization to novel cues in dopamine signaling in ventral and posterior striatum in mice. *Elife*. **6**, 751–767 (2017).
29. S. Hong, T. C. Jhou, M. Smith, K. S. Saleem, O. Hikosaka, Negative reward signals from the lateral habenula to dopamine neurons are mediated by rostromedial tegmental nucleus in primates. *J Neurosci*. **31**, 11457–11471 (2011).
30. Bethesda (MD): National Library of Medicine (US); National Center for Biotechnology Information, National Center for Biotechnology Information (NCBI)[Internet] (2017), (available at https://www.ncbi.nlm.nih.gov/).
31. X.-Y. Deng *et al.*, Identification of the Alternative Promoters of the KChIP4 Subfamily. *Acta Biochim. Biophys. Sin. (Shanghai).* **37**, 241–247 (2005).
32. Z. V. Z. Guo *et al.*, Procedures for Behavioral Experiments in Head-Fixed Mice. *PLoS One*. **9**, e88678 (2014).
33. J. Schiemann *et al.*, K-ATP channels in dopamine substantia nigra neurons control bursting and novelty-induced exploration. *Nat. Neurosci.* **15**, 1272–1280 (2012).
34. M. Antunes, G. Biala, The novel object recognition memory: Neurobiology, test procedure, and its modifications. *Cogn. Process.* **13**, 93–110 (2012).
35. A. Ennaceur, J. Delacour, A new one-trial test for neurobiological studies of memory in rats. 1: Behavioral data. *Behav. Brain Res.* **31**, 47–59 (1988).
36. J. N. Crawley, *What's Wrong With My Mouse? Behavioural Phenotyping of Transgenic and Knockout Mice* (Wiley-Interscience, 2007).
37. M. E. Ragozzino, K. E. Unick, P. E. Gold, Hippocampal acetylcholine release during memory testing in rats: augmentation by glucose. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 4693–4698 (1996).
38. R. C. Lennartz, The role of extramaze cues in spontaneous alternation in a plus-maze. *Learn. Behav. a Psychon. Soc. Publ.* **36**, 138–144 (2008).