Habit Formation after Random Interval Training Is Associated with Increased Adenosine A_{2A} Receptor and Dopamine D_2 Receptor Heterodimers in the Striatum

Yan He [1†], Yan Li [2†], Mozi Chen [3], Zhilan Pu [1], Feiyang Zhang [1], Long Chen [1], Yang Ruan [1], Xinran Pan [2], Chaoxiang He [1], Xingjun Chen [1], Zhihui Li [1*] and Jiang-Fan Chen [1,4*]

[1] School of Optometry and Ophthalmology and Eye Hospital, Institute of Molecular Medicine, Wenzhou Medical University, Wenzhou, China, [2] Department of Geriatrics and Neurology, The 2nd Affiliated Hospital & Yuying Children's Hospital of Wenzhou Medical University, Wenzhou, China, [3] Ma'anshan Municipal Hospital Group and Municipal People's Hospital, Ma'anshan, China, [4] Department of Neurology, School of Medicine, Boston University, Boston, MA, USA

Striatal adenosine A_{2A} receptors (A_{2A}Rs) modulate striatal synaptic plasticity and instrumental learning, possibly by functional interaction with the dopamine D_2 receptors (D_2Rs) and metabotropic glutamate receptors 5 (mGluR5) through receptor-receptor heterodimers, but *in vivo* evidence for these interactions is lacking. Using *in situ* proximity ligation assay (PLA), we studied the subregional distribution of the A_{2A}R-D_2R and A_{2A}R-mGluR5 heterodimer complexes in the striatum and their adaptive changes over the random interval and random ratio training of instrumental learning. After confirming the specificity of the PLA detection of the A_{2A}R-D_2R heterodimers with the A_{2A}R knockout and D_2R knockout mice, we detected a heterogeneous distribution of the A_{2A}R-D_2R heterodimer complexes in the striatum, being more abundant in the dorsolateral than the dorsomedial striatum. Importantly, habit formation after the random interval training was associated with the increased formation of the A_{2A}R-D_2R heterodimer complexes, with prominant increase in the dorsomedial striatum. Conversely, goal-directed behavior after the random ratio schedule was not associated with the adaptive change in the A_{2A}R-D_2R heterodimer complexes. In contrast to the A_{2A}R-D_2R heterodimers, the A_{2A}R-mGluR5 heterodimers showed neither subregional variation in the striatum nor adaptive changes over either the random ratio (RR) or random interval (RI) training of instrumental learning. These findings suggest that development of habit formation is associated with increased formation of the A_{2A}R-D_2R heterodimer protein complexes which may lead to reduced dependence on D_2R signaling in the striatum.

Keywords: A_{2A} receptor, D_2 receptor, receptor-receptor heterodimers, goal-directed behavior, habit, striatum

INTRODUCTION

The adenosine A$_{2A}$ receptors (A$_{2A}$Rs) are highly enriched in the striatopallidal neurons of the striatum (Svenningsson et al., 1999) where A$_{2A}$Rs are co-localized with and form heterodimers with the dopamine D$_2$ receptors (D$_2$Rs) and metabotropic glutamate 5 receptors (mGluR5) (Tebano et al., 2009; Pinna et al., 2014; Taura et al., 2015). Possibly through the receptor-receptor heterodimerization, striatopallidal A$_{2A}$Rs interact antagonistically with D$_2$Rs (Canals et al., 2003; Trifilieff et al., 2011), and synergistically with mGluR5 (Ferré et al., 2002; Kachroo et al., 2005). By these functional interactions, striatopallidal A$_{2A}$Rs can modulate dopamine and glutamate signaling and striatal synaptic plasticity and cognitions including instrumental behaviors (Chen, 2014). Indeed, genetic inactivation of striatal A$_{2A}$Rs impaired habit formation (Yu et al., 2009) and pharmacological reduction of A$_{2A}$R-mediated cAMP-pCREB signaling in the dorsal medium striatum (DMS) enhanced goal-directed ethanol drinking (Nam et al., 2013) and reversed meth-amphetamine-induced facilitation of habit formation (Furlong et al., 2015). However, the mechanism underlying the A$_{2A}$R modulation of instrumental behaviors is not known.

Striatal long-term depression (LTD) that is restricted to striatopallidal neurons and requires activation of D$_2$Rs and mGluR5 (Kreitzer and Malenka, 2007; Lovinger, 2010) is the main form of plasticity of synaptic transmission in the dorsolateral striatum (DLS; Partridge et al., 2000; Yin and Knowlton, 2006; Lovinger, 2010). The loss of striatopallidal LTD is associated with a shift in behavioral control from goal-directed (Furlong et al., 2015) action to habitual responding (Nazzaro et al., 2012). Since activation of striatopallidal A$_{2A}$Rs can convert the striatopallidal synaptic plasticity from LTD to long-term potentiation (LTP; Shen et al., 2008), striatopallidal A$_{2A}$R signaling may interact with D$_2$R-/mGluR5-/endocannabinoids-mediated striatal LTD in striatopallidal neurons to modify instrumental learning. Thus, we postulated that striatopallidal A$_{2A}$Rs may exert their effects on D$_2$R- or mGluR5-mediated striatal synaptic plasticity and instrumental learning through the physical association of the A$_{2A}$R-D$_2$R and A$_{2A}$R-mGluR5 heterodimers in the striatopallidal neurons. Here, using two instrumental learning schedules coupled with *in situ* proximity ligation assay (PLA), we investigated the heterogeneous distribution of the A$_{2A}$R-D$_2$R and A$_{2A}$R-mGluR5 heterodimers in the DLS and DMS and their adaptive changes after the random interval (to promote habit) and random ratio (to promote goal-directed behavior) training schedules.

MATERIALS AND METHODS

Animals

All animals were handled in accordance with the protocols approved by the Institutional Ethics Committee for Animal Use in Research and Education at Wenzhou Medical University, China. Eighteen adult C57B6/J (n = 6/experimental group), three A$_{2A}$R knockout mice (from Chen's laboratory at Boston University School of Medicine) and three D$_2$R knockout mice (from The Jackson Laboratory, USA, Drd2^{tm1Low}, stock No. 003190) were used for the experiments.

Instrumental Behavior Training Schedules

Instrumental training and behavioral testing schedules were performed following the procedure by Rossi et al. (Rossi and Yin, 2012). In brief, mice first underwent a 5-day food deprivation schedule to reach 80–85% of their free-feeding weight before instrumental training sessions. Mice were then given one 30-min magazine training session during which one drop of 20 µl 20% sucrose solution as reward was delivered on a random time 60-s schedule. During continuous reinforcement (CRF) sessions, each lever press resulted in delivery of the sucrose reward. Sessions ended after 60 min or 50 rewards had been earned, whichever came first. After CRF, mice underwent a random interval (RI) schedule to promote habit formation or a random ratio (RR) schedule to promote goal-directed behavior. Mice underwent the RI schedule were trained for 2 days on random interval 30 s (RI30), with a 0.1 probability of reward availability every 3 s contingent upon lever pressing, followed by 4 days on RI60 schedule. Progressively leaner schedules of reinforcement were used for the RR training procedure: RR5 (each response was rewarded at a probability of 0.2 on average), RR10, RR20 each for 2 days.

Following the RI and RR training sessions, a 2-day devaluation test was conducted. A specific satiety procedure was applied to alter the current value of a specific reward. On each day the mice were allowed to have free access to home chows which were used for maintaining their weights (i.e., valued condition, the sucrose solution was still valued) or sucrose solution which was earned by their lever pressing in the training sessions (i.e., devalued condition, the sucrose solution was devalued) for at least an hour to achieve sensory-specific satiety. Immediately after the unlimited pre-feeding session, mice were given a 5-min extinction test during which the lever was inserted and pressing times was recorded but without reward delivery. The orders of the valued and devalued condition tests (day 1 or 2) were counterbalanced across each group. Mice insensitive to manipulation of outcome value, that is habit, would mildly change lever presses on the devalued condition compared to the valued condition, whereas goal-directed mice that performed sensitively to outcome value would significantly reduce their lever presses on the devalued condition. The control mice underwent food deprivation schedule and were handled exactly the same way every day just as the RR and RI training group before instrumental training. During the training sessions, the control mice were also placed in the operant chambers in which the sucrose reward was delivered in a random 30/60 s (corresponding to RI30/RI60) manner but without lever stretched. In the devaluation test, the control mice were also exposed to the operant chamber for 5 min with no lever stretched out and no reward available. In the present study, three groups (n = 6 for each group) were examined for the A$_{2A}$R-D$_2$R/A$_{2A}$R-mGLuR5 heterodimers in the striatum after instrumental learning: (a) mice without instrumental training as "control group," (b) mice underwent RI/RR training sessions as "RI group," and (c) mice underwent RR training sessions as "RR group."

Proximity-Ligation Assay (PLA)

After two additional RI60 or RR20 training sessions following devaluation test, mice were sacrificed for PLA detection of the $A_{2A}R$-D_2R and $A_{2A}R$-mGLuR5 heterodimers in the striatum. We performed PLA analysis according to the procedure described recently (Augusto et al., 2013; Pinna et al., 2014). Three sections from one brain (from anterior, middle, and posterior parts of the striatum, respectively) were rinsed in TBS at room temperature. The sections were incubated with 1% BSA and 0.5% Triton X-100 for 2 h at room temperature for blocking and permeabilization. The mouse anti-$A_{2A}R$ (1:300; millipore) and rabbit anti-D_2R (1:300; millipore)/rabbit anti-mGluR5 (1:300; millipore) were incubated with sections overnight at room temperature. Sections were then rinsed for four times (30 min each time) in TBS with 0.2% Triton X-100 following the manufacturer's protocol. Slices were then incubated at 37°C with the PLA secondary probes (1:5; Olink Bioscience) for 2 h. After rinsing with "Duolink II" Wash Buffer A, the slices were then incubated with the ligation-ligase solution for 30 min at 37°C followed by rinsing with Duolink II Wash Buffer A. Sections were ready for amplification with polymerase (1:40; Olink Bioscience). Then the sections were washed in decreasing concentrations SSC buffers (Olink Bioscience) and mounted on slides. Fluorescent mounting medium (containing DAPI) were applied on the sections. The fluorescence images (three non-overlapping and random microscopic fields respective from DMS and DLS of each brain section) were acquired by confocal microscope (**Figure 2A**). The quantit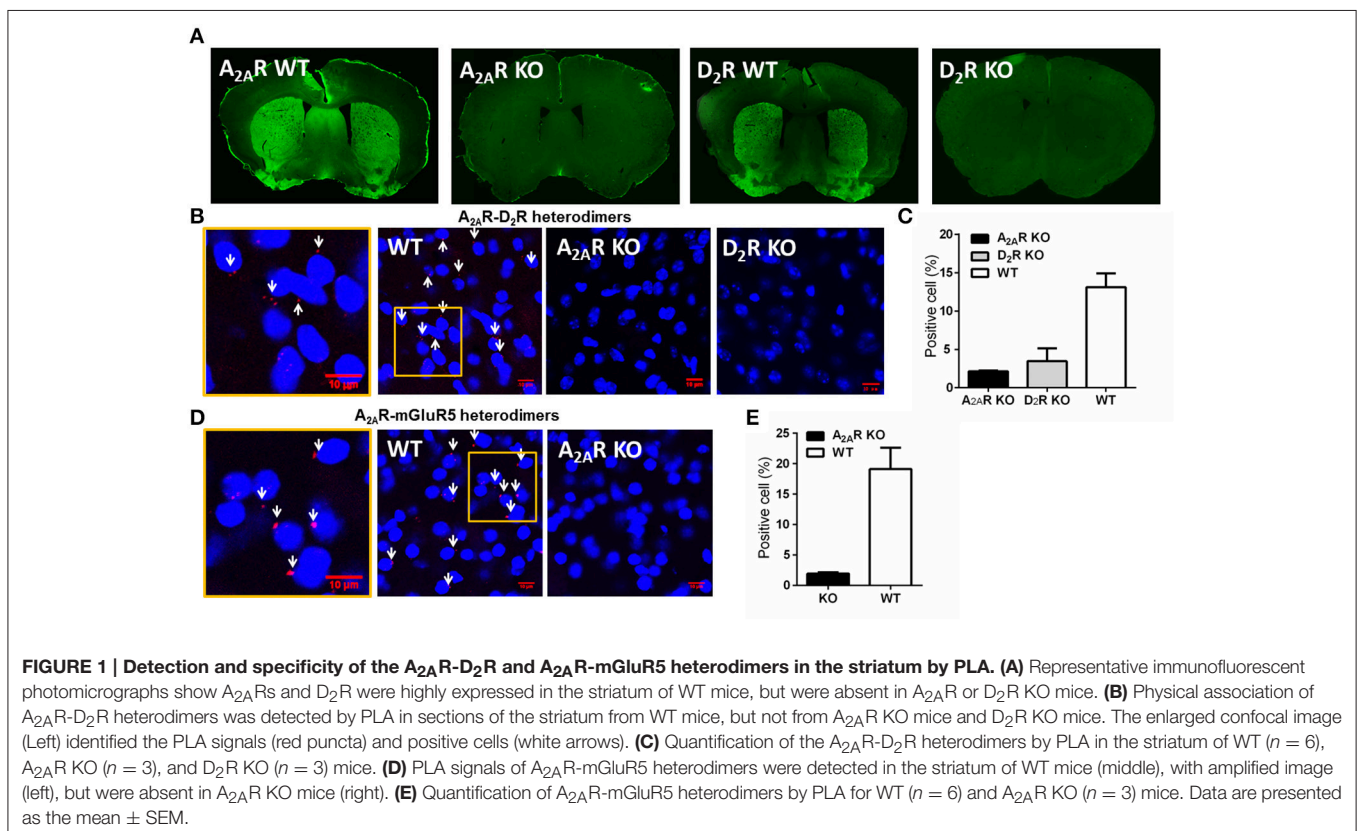ative analysis was done following the procedure by Bonaventura et al. (2014). The cells surrounded by the red puncta were defined as positive cells (white arrows in **Figure 1**). The cell number was counted by software "Image J." Each microscopic field image was quantified as "positive cell number/total cell number." The quantified value of the experimental mice was normalized to that of $A_{2A}R$ KO mice (as the background).
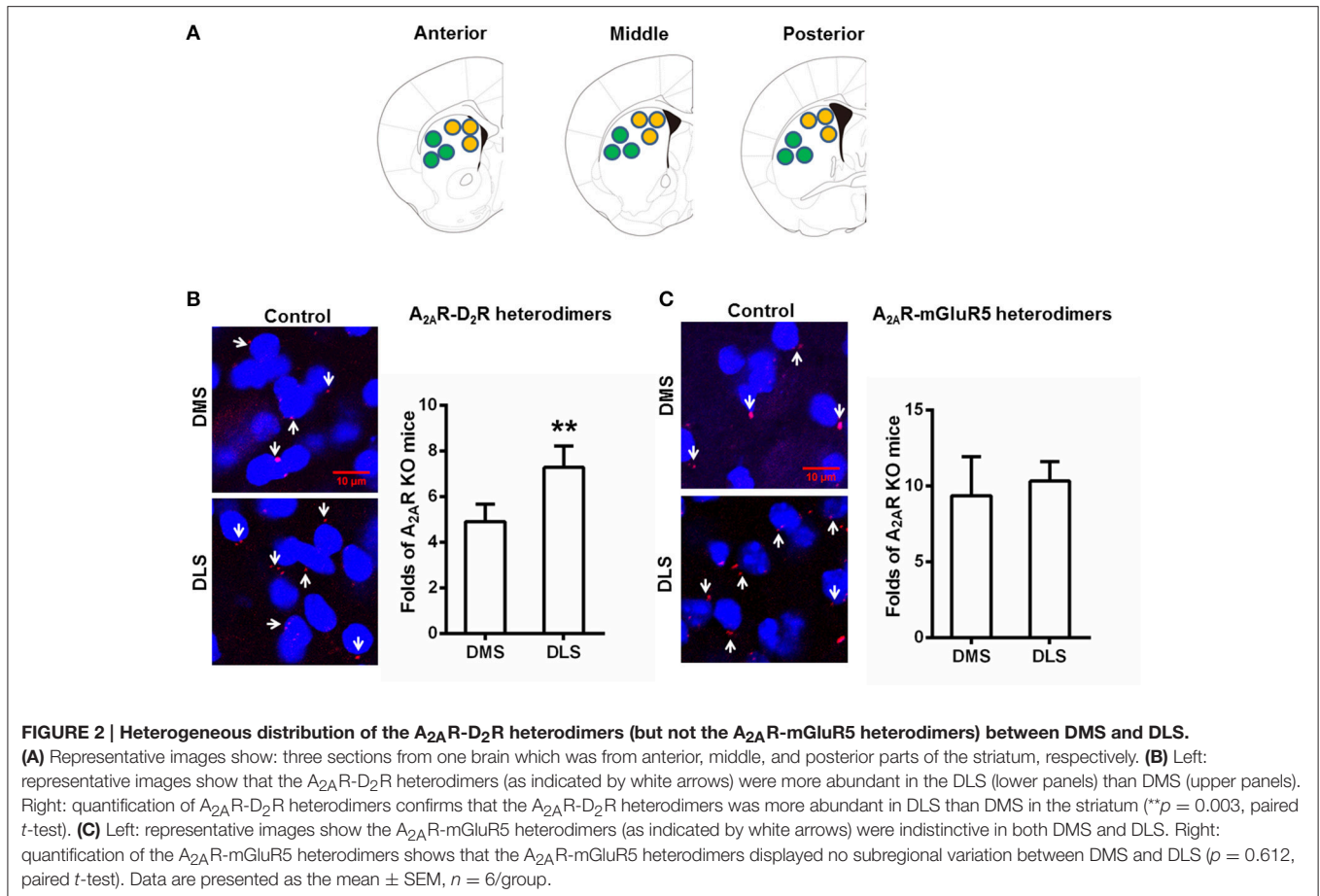
Immunofluorescence

Mice were deeply anesthetized and then transcardially perfused with 0.01 M PBS (pH = 7.4) followed by ice-cold 4% paraformaldehyde. Brains were post-fixed in 4% paraformaldehyde for 4–6 h at 4°C, and then allowed to equilibrate using gradient sucrose solution (10, 20, and 30%). Immunofluorescence were performed on 30 μm free-floating sections. Free-floating sections were washed in PBS and incubated for 60 min in 0.3% Triton X-100 and 10% normal donkey serum and then incubated with mouse anti-$A_{2A}R$ antibody (Millipore, 1:200) and rabbit anti-D_2R antibody (Millipore, 1:200) at 4°C overnight. Brain sections were incubated with Alexa 488-conjugated secondary antibodies (Invitrogen, 1:1000). The sections were washed and mounted. Fluorescent mounting medium were applied on the sections. Images were acquired by a fluorescence microscope.

Statistical Analysis

Instrumental behavior training and test processes were analyzed using two-way ANOVA for repeated measurements with the training or test sessions as within-subjects effect and the different

FIGURE 1 | Detection and specificity of the $A_{2A}R$-D_2R and $A_{2A}R$-mGluR5 heterodimers in the striatum by PLA. (A) Representative immunofluorescent photomicrographs show $A_{2A}R$s and D_2R were highly expressed in the striatum of WT mice, but were absent in $A_{2A}R$ or D_2R KO mice. **(B)** Physical association of $A_{2A}R$-D_2R heterodimers was detected by PLA in sections of the striatum from WT mice, but not from $A_{2A}R$ KO mice and D_2R KO mice. The enlarged confocal image (Left) identified the PLA signals (red puncta) and positive cells (white arrows). **(C)** Quantification of the $A_{2A}R$-D_2R heterodimers by PLA in the striatum of WT (n = 6), $A_{2A}R$ KO (n = 3), and D_2R KO (n = 3) mice. **(D)** PLA signals of $A_{2A}R$-mGluR5 heterodimers were detected in the striatum of WT mice (middle), with amplified image (left), but were absent in $A_{2A}R$ KO mice (right). **(E)** Quantification of $A_{2A}R$-mGluR5 heterodimers by PLA for WT (n = 6) and $A_{2A}R$ KO (n = 3) mice. Data are presented as the mean ± SEM.

FIGURE 2 | Heterogeneous distribution of the A$_{2A}$R-D$_2$R heterodimers (but not the A$_{2A}$R-mGluR5 heterodimers) between DMS and DLS.
(A) Representative images show: three sections from one brain which was from anterior, middle, and posterior parts of the striatum, respectively. **(B)** Left: representative images show that the A$_{2A}$R-D$_2$R heterodimers (as indicated by white arrows) were more abundant in the DLS (lower panels) than DMS (upper panels). Right: quantification of A$_{2A}$R-D$_2$R heterodimers confirms that the A$_{2A}$R-D$_2$R heterodimers was more abundant in DLS than DMS in the striatum (**$p = 0.003$, paired t-test). **(C)** Left: representative images show the A$_{2A}$R-mGluR5 heterodimers (as indicated by white arrows) were indistinctive in both DMS and DLS. Right: quantification of the A$_{2A}$R-mGluR5 heterodimers shows that the A$_{2A}$R-mGluR5 heterodimers displayed no subregional variation between DMS and DLS ($p = 0.612$, paired t-test). Data are presented as the mean ± SEM, $n = 6$/group.

training procedures as between-subjects effect. In the PLA assay, two-way ANOVA was used with striatal subregions and training procedures as main effects. Paired t-test was conducted to compare the distribution difference of A$_{2A}$R-D$_2$R and A$_{2A}$R-mGluR5 heterodimers between the DMS and DLS. One-way ANOVA with LSD *post-hoc* was used to compare the distribution variation of A$_{2A}$R-D$_2$R heterodimers on the anterior-posterior axes in both DMS and DLS after instrumental learning.

RESULT

Detection of the A$_{2A}$R-D$_2$R and A$_{2A}$R-mGluR5 Heterodimers by PLA in the Striatum

To detect the striatal A$_{2A}$R-D$_2$R and A$_{2A}$R-mGluR5 heterodimers by PLA assay, we first confirmed the specificity of the PLA detection of the A$_{2A}$R-D$_2$R and A$_{2A}$-mGluR5 heterodimers using the A$_{2A}$R KO and D$_2$R KO mice. The specificity of the A$_{2A}$R and D$_2$R antibody was evident with highly enriched expression pattern of the A$_{2A}$Rs or D$_2$Rs in the striatum by immunofluorescence, which was absent in the A$_{2A}$R or D$_2$R KO mice (**Figure 1A**). Moreover, the specific labeling of the A$_{2A}$R-D$_2$R heterodimer signals (cellular membrane) were detected in ~15% of striatal neurons of wild-type mice in

close association of DAPI (nuclei; as indicated by white arrow; **Figures 1B,C**). Importantly, these signals for the A$_{2A}$R-D$_2$R protein complexes were essentially absent in the striatum of the A$_{2A}$R KO or D$_2$R KO mice (**Figures 1B,C**), confirming the specificity of the PLA detection of the A$_{2A}$R-D$_2$R heterodimers. Similarly, the A$_{2A}$R-mGluR5 heterodimer signals were detected by PLA in the striatum of wild-type mice but not in A$_{2A}$R KO mice, supporting the specificity of PLA detection of the A$_{2A}$R-mGluR5 heterodimers (**Figures 1D,E**).

Heterogeneous Distribution of the A$_{2A}$R-D$_2$R Heterodimers (but Not the A$_{2A}$R-mGluR5 Heterodimers) in the DMS and DLS
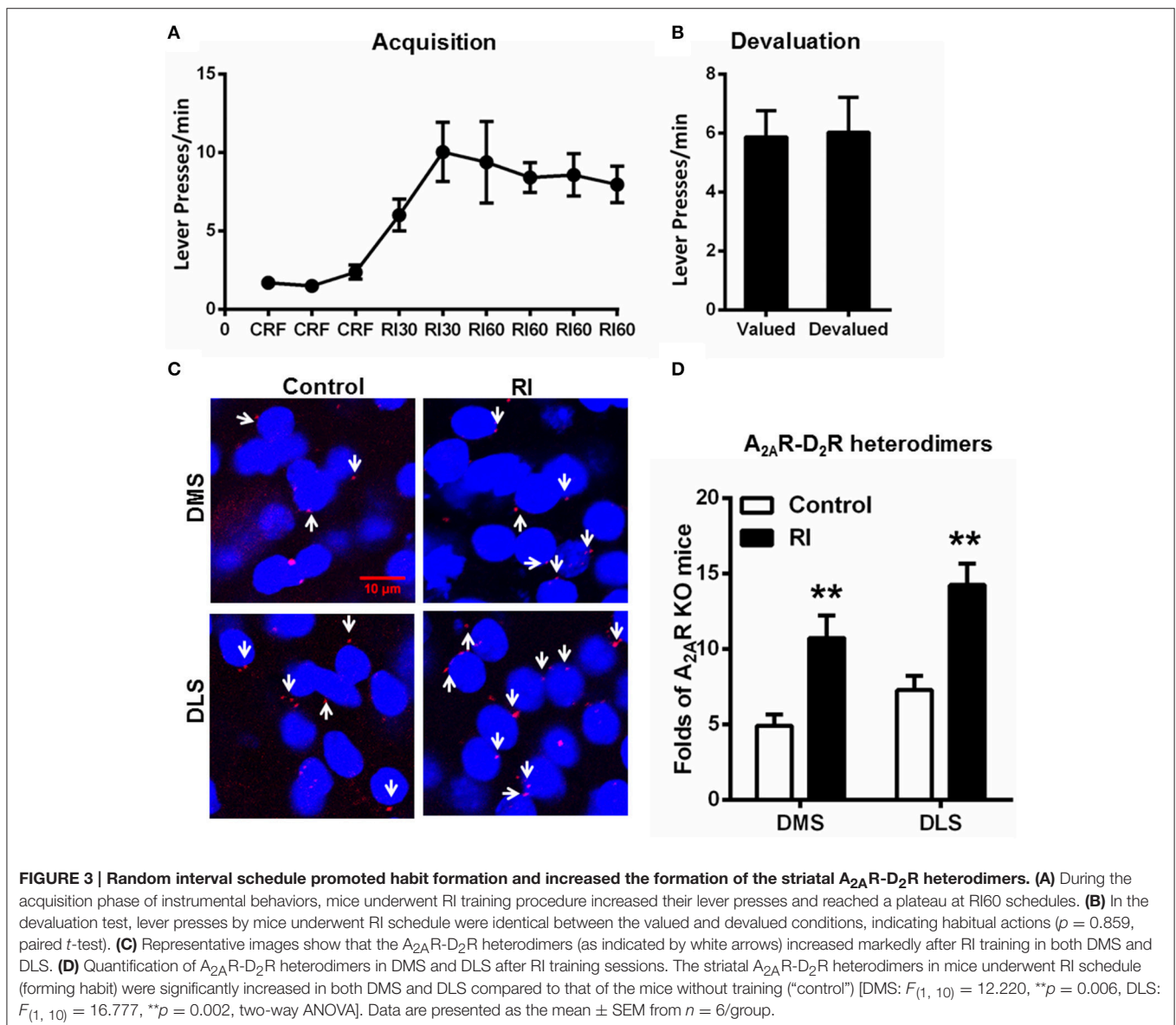
Following the confirmation of the specificity of the PLA detection of the A$_{2A}$R-D$_2$R and A$_{2A}$R-mGluR5 heterodimers, we examined the heterogeneous distribution of these heterodimers in the DMS and DLS in normal mice. PLA analysis showed that the A$_{2A}$R-D$_2$R complexes were more prominent in the DLS than DMS (**Figure 2B**). Quantitative analysis confirmed the heterogeneous distribution of A$_{2A}$R-D$_2$R heterodimers in the striatum (i.e., DLS > DMS; **Figure 2B**). By contrast, there was no heterogeneous distribution of the A$_{2A}$R-mGluR5 heterodimers in the DMS and DLS by PLA (**Figure 2C**).
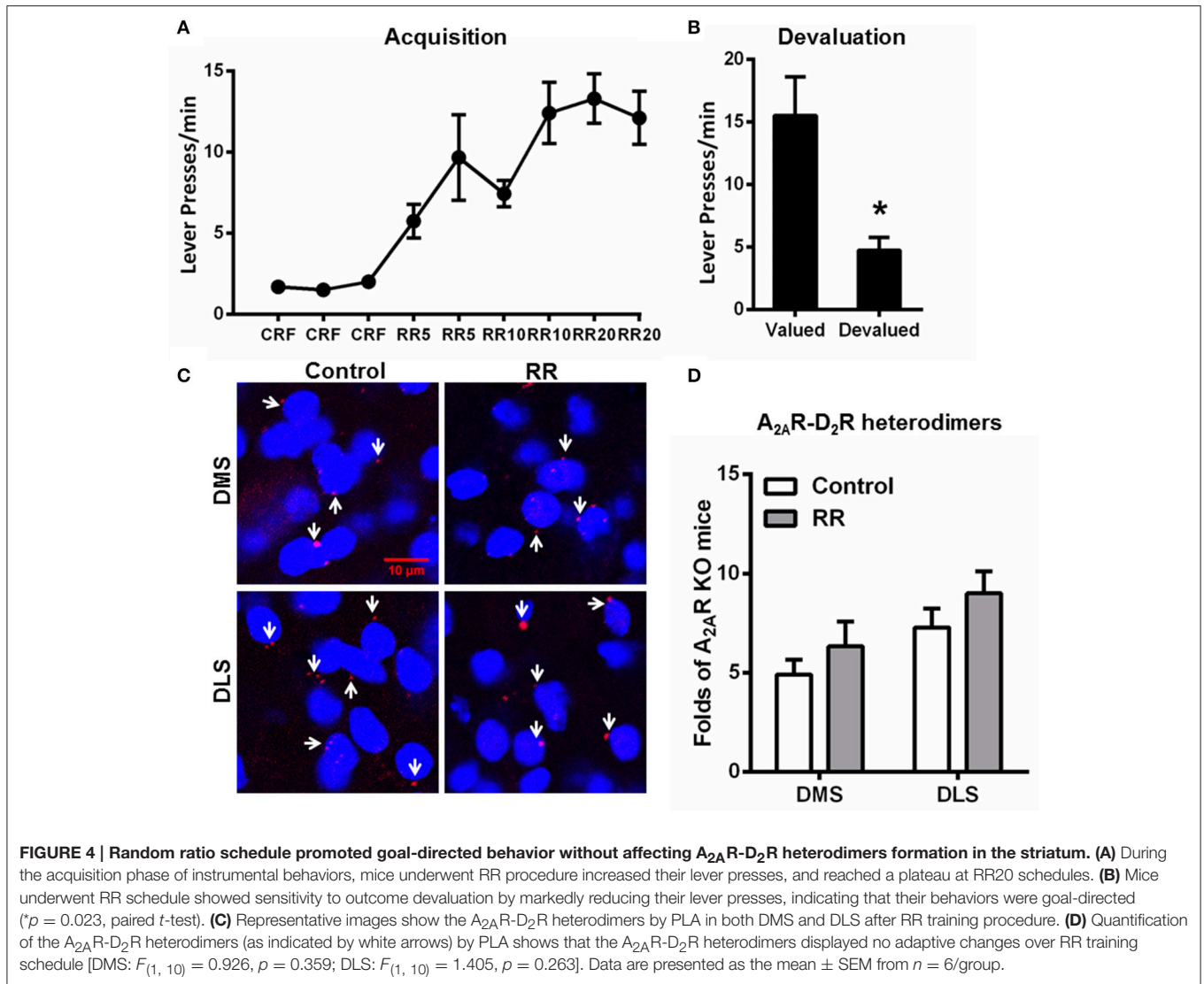
Random Interval Schedule Promoted Habit Formation and Increased the Formation of the Striatal $A_{2A}R$-D_2R Heterodimers

Following RI training sessions, mice gradually, and steadily increased their lever presses and reached plateau at RI60 schedule (**Figure 3A**). Consistent with the previous studies (Dickinson et al., 1983; Yin and Knowlton, 2006), devaluation test (**Figure 3B**) showed that mice trained by RI procedure showed insensitive to outcome devaluation, indicating their habitual action. In association with habit formation after the RI training sessions, we detected the increased formation of the $A_{2A}R$-D_2R heterodimers compared to mice without training ("control;" **Figures 3C,D**). Moreover, the $A_{2A}R$-D_2R heterodimers increased in both DMS and DLS, and accordingly, the heterogeneous pattern of this heterodimers in DMS and DLS persisted after the trainings.

Random Ratio Promoted Goal-Directed Behavior without Affecting the $A_{2A}R$-D_2R Heterodimer Formation in the Striatum

Over the RR training sessions, mice also gradually and steadily increased their lever presses and reached plateau at RR20 schedule (**Figure 4A**). Consistent with the previous studies (Dickinson et al., 1983; Yin and Knowlton, 2006), devaluation test showed that mice trained by RR schedule markedly reduced their lever presses in the devalued condition, indicating their goal-directed behavior (**Figure 4B**). In association with goal-directed behavior after RR training sessions, we did not detect any significant change in the $A_{2A}R$-D_2R heterodimers compared to mice without training ("control;" **Figures 4C,D**). Furthermore, the striatal $A_{2A}R$-D_2R heterodimers underwent RI schedule were significantly increased in both DMS and DLS compared to the RR group [DMS: $F_{(2, 17)} = 6.351$, $p = 0.010$, DLS: $F_{(2, 17)} = 9.605$,

FIGURE 3 | Random interval schedule promoted habit formation and increased the formation of the striatal $A_{2A}R$-D_2R heterodimers. (A) During the acquisition phase of instrumental behaviors, mice underwent RI training procedure increased their lever presses and reached a plateau at RI60 schedules. **(B)** In the devaluation test, lever presses by mice underwent RI schedule were identical between the valued and devalued conditions, indicating habitual actions ($p = 0.859$, paired t-test). **(C)** Representative images show that the $A_{2A}R$-D_2R heterodimers (as indicated by white arrows) increased markedly after RI training in both DMS and DLS. **(D)** Quantification of $A_{2A}R$-D_2R heterodimers in DMS and DLS after RI training sessions. The striatal $A_{2A}R$-D_2R heterodimers in mice underwent RI schedule (forming habit) were significantly increased in both DMS and DLS compared to that of the mice without training ("control") [DMS: $F_{(1, 10)} = 12.220$, **$p = 0.006$, DLS: $F_{(1, 10)} = 16.777$, **$p = 0.002$, two-way ANOVA]. Data are presented as the mean ± SEM from $n = 6$/group.

FIGURE 4 | Random ratio schedule promoted goal-directed behavior without affecting A_{2A}R-D_2R heterodimers formation in the striatum. (A) During the acquisition phase of instrumental behaviors, mice underwent RR procedure increased their lever presses, and reached a plateau at RR20 schedules. **(B)** Mice underwent RR schedule showed sensitivity to outcome devaluation by markedly reducing their lever presses, indicating that their behaviors were goal-directed (*$p = 0.023$, paired t-test). **(C)** Representative images show the A_{2A}R-D_2R heterodimers by PLA in both DMS and DLS after RR training procedure. **(D)** Quantification of the A_{2A}R-D_2R heterodimers (as indicated by white arrows) by PLA shows that the A_{2A}R-D_2R heterodimers displayed no adaptive changes over RR training schedule [DMS: $F_{(1, 10)} = 0.926$, $p = 0.359$; DLS: $F_{(1, 10)} = 1.405$, $p = 0.263$]. Data are presented as the mean ± SEM from $n = 6$/group.

$p = 0.002$; *$p < 0.05$, **$p < 0.01$; one-way ANOVA with LSD *post-hoc* test, $n = 6$/group]. Thus, the increased association of striatal A_{2A}R-D_2R heterodimers is selectively induced by the RI training schedule in association with habit formation, but not affected by RR training schedule which produced goal-directed behavior.

A_{2A}R-D_2R Heterodimers in the DMS Showed Prominent Increases after RI Training on Anterior-Posterior Axes

We have also performed detailed analysis of the A_{2A}R-D_2R hetrerodimers in three subregions of the DMS and DLS on anterior-posterior axes (i.e., anterior, middle, and posterior, **Figure 5**). In the DMS, the A_{2A}R-D_2R heterodimers in both the anterior and posterior parts were increased after RI training compared to the control or RR group (**Figure 5A**). In the DLS, the change in the A_{2A}R-D_2R heterodimers was relatively less pronounced such that the increase was observed

only in the middle DLS after RI training compared to the control (**Figure 5B**). There was no difference in the A_{2A}R-D_2R heterodimers in the anterior, middle and posterior part of the DMS and DLS between the RR or control groups, with exception of apparently a small increase in the middle part of the DMS in the RR group (**Figures 5A,B**).

The Striatal A_{2A}R-mGluR5 Heterodimers Display Neither Subregional Distribution Nor Adaptive Changes after the RI and RR Training Schedules

We also examined the subregional distribution and adaptive changes of the A_{2A}R-mGluR5 heterodimers after the RI and RR training schedules (**Figure 6**). The A_{2A}R-mGluR5 heterodimers displayed neither heterogeneous distribution between DMS and DLS nor adaptive changes after the RI training (leading to habit formation) or RR training (leading to goal-directed behavior) sessions (**Figures 6A,B**). These indicated
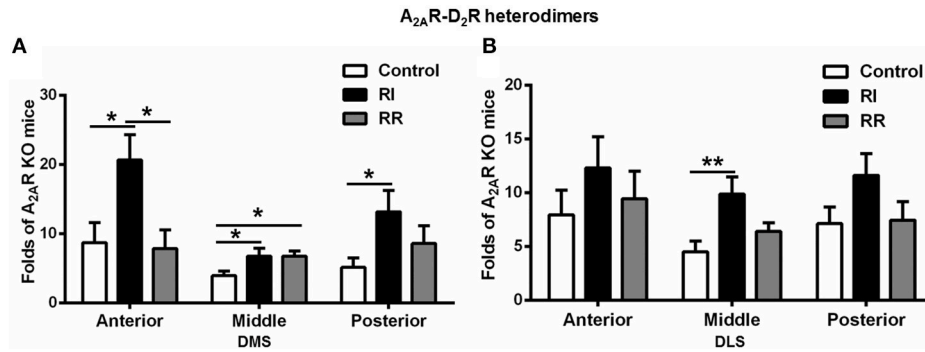
FIGURE 5 | A$_{2A}$R-D$_2$R heterodimers in the DMS showed more prominent increases after RI training on anterior-posterior axes. (A) In the DMS, the A$_{2A}$R-D$_2$R heterodimers in the anterior and posterior parts were significantly higher in the RI group compared to the control and RR groups [anterior: $F_{(2, 16)}$ =5.135, $p = 0.021$, LSD post-hoc: RI vs. control, *$p = 0.017$, RI vs. RR, *$p = 0.011$; posterior: $F_{(2, 16)}$ = 3.881, $p = 0.046$, LSD post-hoc: RI vs. control, *$p = 0.015$]. There was also a relatively small increase in the middle part of the DMS in the RI and RR group compared to the control group [middle: $F_{(2, 16)}$ = 3.889, $p = 0.045$, LSD post-hoc: RI vs. control, *$p = 0.024$, RR vs. control, *$p = 0.041$]. (B) In the DLS, the A$_{2A}$R-D$_2$R heterodimers in the middle part was higher in the RI group than the control and RR group [$F_{(2, 17)}$ = 5.223, *$p = 0.019$, LSD post-hoc: RI vs. control, **$p = 0.006$]. Data are presented as the mean ± SEM.
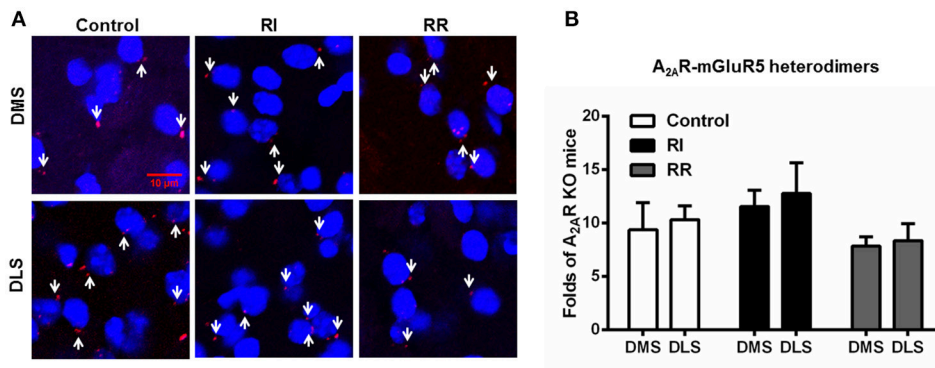
FIGURE 6 | The striatal A$_{2A}$R-mGluR5 heterodimers displayed neither subregional distribution variation nor adaptive changes over the RI and RR training schedules. (A) Representative images show the A$_{2A}$R-mGluR5 heterodimers by PLA in both DMS and DLS and after different training procedures. (B) Quantification of the A$_{2A}$R-mGluR5 heterodimers (as indicated by white arrows) by PLA shows that the A$_{2A}$R-mGluR5 heterodimers displayed neither subregional distribution variation between DMS and DLS nor adaptive changes over RI or RR training schedule [Subregion main effect: $F_{(1, 30)}$ = 0.325, $p = 0.573$; Training main effect: $F_{(2, 30)}$ = 2.245, $p = 0.123$]. Data are presented as the mean ± SEM from n = 6/group.

that such heterogeneous distribution (DLS vs. DMS) and adaptive changes over the instrumental learnings (i.e., increased formation of the heterodimers) were specific for the A$_{2A}$R-D$_2$R heterodimers.

DISCUSSION

The A$_{2A}$R-D$_2$R heterodimers have been studied extensively in cultured cells and brain tissues by fluorescence resonance energy transfer (FRET; Torvinen et al., 2005) and receptor binding with biochemical finger printing (Ciruela et al., 2006) and by blocking peptides targeting the presumed A$_{2A}$R-D$_2$R interaction site (Azdad et al., 2009) and by co-immunoprecipitation (Ciruela et al., 2006). Since heteromerization of A$_{2A}$Rs and other GPCRs (such as A$_{2A}$R-D$_2$R and A$_{2A}$R-mGluR5) has been demonstrated mostly in cultured cell lines with overexpressed recombinant receptors that may result in the creation of many

more heterodimers than naturally exist, it is essential to detect the normal distribution of the A$_{2A}$R-D$_2$R and A$_{2A}$R-mGluR5 heterodimers in the intact striatum in order to infer their possible physiological functions. However, the direct detection of the A$_{2A}$R-D$_2$R and A$_{2A}$R-mGluR5 heterodimers in intact animals and its physiological relevance has been proven to be difficult. Recently, PLA has been developed to detect the presence of the A$_{2A}$R-D$_2$R (Trifilieff et al., 2011) and A$_{2A}$R-CD73 heterodimers in the striatum (Augusto et al., 2013). For example, Bonaventura et al. showed that the A$_{2A}$R-D$_2$R heterodimers by PLA were reduced in dopamine-depleted caudate-putamen after chronic treatment with L-dopa in non-human primates (Bonaventura et al., 2014). The specificity of the A$_{2A}$R-D$_2$R and A$_{2A}$R-mGluR5 heterodimers using PLA was further validated here by demonstrating the detection of the A$_{2A}$R-D$_2$R and A$_{2A}$R-mGluR5 heterodimers in the striatum of WT but neither in A$_{2A}$R KO nor in D$_2$R KO mice.

Using PLA, we demonstrated the heterogeneous distribution of the A$_{2A}$R-D$_2$R heterodimers in DMS and DLS as well as their adaptive changes over the instrumental learning procedures. Given the critical role of the DLS in control of habit formation, the more abundant A$_{2A}$R-D$_2$R heterodimers in DLS than DMS under the basal condition may suggest that the A$_{2A}$R-D$_2$R heterodimers in DLS may contribute to habit formation. It should be noted that there was no subregional variation in the A$_{2A}$R-mGluR5 heterodimers between the DMS and DLS. Since D$_2$R-mediated striatal LTD is preferentially founded in the DLS striatopallidal neurons (Shen et al., 2008), the prominent DLS distribution of A$_{2A}$-D$_2$R heterodimers may suggest a possible role of the A$_{2A}$R and D$_2$R (rather than the A$_{2A}$R-mGLuR5) interaction in modulating LTD in the DLS.

Importantly, following instrumental learning, our analysis reveals that the formation of the A$_{2A}$R-D$_2$R heterodimers increased (by nearly two-folds) over the RI schedule which resulted in habitual behavior compared to their control level. The increased association of the A$_{2A}$R-D$_2$R heterodimers was seen selectively after the RI schedule (to promote habit), but was not seen in the mice trained by the RR schedule (to promote goal-directed behavior). Genetic and pharmacological studies have implicated several neurotransmitter and neuromodulator receptors take effects on development of goal-directed and habit behaviors, including D$_1$R, D$_2$R (Yin et al., 2009), CB$_1$R (Hilário et al., 2007), NMDAR (Yin et al., 2005), A$_{2A}$R (Yu et al., 2009; Li et al., 2016), and Gpr6 (Lobo et al., 2007), by alteration of instrumental behaviors. However, to the best of our knowledge, this is the first report on the molecular marker of habitual formation that is selectively induced by the RI (but not the RR) training schedule. This novel molecular correlate of the RI training and possibly habit formation would be useful in molecular dissecting and monitoring habit formation. Moreover, our detailed analysis of A$_{2A}$R-D$_2$R heterodimer changes on anterior-posterior axes indicated that the A$_{2A}$R-D$_2$R heterodimers in the DMS showed more dynamic changes after RI training in agreement with our recent finding that the DMS A$_{2A}$R signaling plays a major role in control of instrumental learning (Li et al., 2016).

Given the well-documented A$_{2A}$R-D$_2$R antagonistic interaction, striatopallidal A$_{2A}$Rs may affect animals' sensitivity to dopamine signaling through the increased A$_{2A}$R-D$_2$R heterodimers. Dopamine signaling apparently has more prominent role during the early stage of instrumental learning (goal-directed behavior) than the late stage (habitual behavior; Choi et al., 2005). As the RI training progresses, the formation of the striatal A$_{2A}$R-D$_2$R heterodimers increases, resulting in the increased inhibition of the A$_{2A}$R on the D$_2$R signaling in the striatopallidal neurons and consequently the less dependence of dopamine signaling at the late stage of instrumental learning. On the other hand, the lack of adaptive changes of the A$_{2A}$R-mGluR5 heterodimers after instrumental learning suggests that striatopallidal A$_{2A}$R activity may preferentially interact with dopamine signaling to modify instrumental learning process.

Since the LTD in striatopallidal neurons, which is modulated by the D$_2$R (Kreitzer and Malenka, 2007) and A$_{2A}$R activities (Shen et al., 2008), is associated with a shift in behavioral control from goal-directed action to habitual responding (Nazzaro et al., 2012). We speculated that the increased formation of the A$_{2A}$R-D$_2$R heterodimers after the RI learning may increase the inhibition effect of the A$_{2A}$Rs on the D$_2$Rs, and consequently reduce D$_2$R-mediated striatal LTD in striatopallidal neurons to modify instrumental learning. The prominent changes in the A$_{2A}$R-D$_2$R heterodimers in the DMS and its correlation with development of habitual behavior lend support for our interpretation that the increased formation of the A$_{2A}$R-D$_2$R heterodimers after RI training augments the inhibitory effect of the A$_{2A}$R on the D$_2$R activity and consequently on goal-directed behavior, manifesting as a habitual behavior. Thus, the A$_{2A}$R-D$_2$R heterodimers may partially account for recent demonstrations that optogenetic activation of the striatal A$_{2A}$Rs promotes habit (Li et al., 2016) and pharmacological blockade of the A$_{2A}$R promote goal-directed ethanol intake (Nam et al., 2013) and reverse meth-amphetamine-induced facilitation of habitual action (Furlong et al., 2015), albeit the A$_{2A}$R may control habit formation by distinct mechanism other than the A$_{2A}$R-D$_2$R heterodimerization. If the functional significance of the A$_{2A}$R-D$_2$R heterodimers can be demonstrated by future studies with direct manipulation of these heterodimers (such as the blocking peptide specifically targeting the A$_{2A}$R-D$_2$R heterodimers interface; Azdad et al., 2009) in intact animals, the A$_{2A}$R-D$_2$R heterodimers may represent a novel therapeutic target for controlling abnormal habit formation associated with obsessive compulsive disorders and relapse of drug addiction.

AUTHOR CONTRIBUTIONS

YL, YH, ZL, and JC designed the experiment. YL, YH, MC, ZP, FZ, LC, CH, and XC collected the data. YL, YH, MC, YR, XP, ZL, and JC analyzed the data. YL, YH, ZL, and JC wrote the manuscript.

FUNDING

REFERENCES

Augusto, E., Matos, M., Sevigny, J., El-Tayeb, A., Bynoe, M. S., Muller, C. E., et al. (2013). Ecto-5′-nucleotidase (CD73)-mediated formation of adenosine is critical for the striatal adenosine A$_{2A}$ receptor functions. *J. Neurosci.* 33, 11390–11399. doi: 10.1523/JNEUROSCI.5817-12.2013

Azdad, K., Gall, D., Woods, A. S., Ledent, C., Ferre, S., and Schiffmann, S. N. (2009). Dopamine D$_2$ and adenosine A$_{2A}$ receptors regulate NMDA-mediated excitation in accumbens neurons through A$_{2A}$-D$_2$ receptor heteromerization. *Neuropsychopharmacology* 34, 972–986. doi: 10.1038/npp.2008.144

Bonaventura, J., Rico, A. J., Moreno, E., Sierra, S., Sanchez, M., Luquin, N., et al. (2014). L-DOPA-treatment in primates disrupts the expression of A$_{2A}$ adenosine-CB$_1$ cannabinoid-D$_2$ dopamine receptor heteromers in the caudate nucleus. *Neuropharmacology* 79, 90–100. doi: 10.1016/j.neuropharm.2013.10.036

Canals, M., Marcellino, D., Fanelli, F., Ciruela, F., de Benedetti, P., Goldberg, S. R., et al. (2003). Adenosine A$_{2A}$-dopamine D$_2$ receptor-receptor heteromerization: qualitative and quantitative assessment by fluorescence and bioluminescence energy transfer. *J. Biol. Chem.* 278, 46741–46749. doi: 10.1074/jbc.M306451200

Chen, J. F. (2014). Adenosine receptor control of cognition in normal and disease. *Int. Rev. Neurobiol.* 119, 257–307. doi: 10.1016/B978-0-12-801022-8.00012-X

Choi, W. Y., Balsam, P. D., and Horvitz, J. C. (2005). Extended habit training reduces dopamine mediation of appetitive response expression. *J. Neurosci.* 25, 6729–6733. doi: 10.1523/JNEUROSCI.1498-05.2005

Ciruela, F., Casado, V., Rodrigues, R. J., Lujan, R., Burgueno, J., Canals, M., et al. (2006). Presynaptic control of striatal glutamatergic neurotransmission by adenosine A1-A$_{2A}$ receptor heteromers. *J. Neurosci.* 26, 2080–2087. doi: 10.1523/JNEUROSCI.3574-05.2006

Dickinson, A., Nicholas, D. J., and Adams, C. (1983). The effect of the instrumental training contingency on susceptibility to reinforcer devaluation. *Q. J. Exp. Psychol.* 35, 35–51. doi: 10.1080/14640748308400912

Ferré, S., Karcz-Kubicha, M., Hope, B. T., Popoli, P., Burgueno, J., Gutierrez, M. A., et al. (2002). Synergistic interaction between adenosine A$_{2A}$ and glutamate mGlu5 receptors: implications for striatal neuronal function. *Proc. Natl. Acad. Sci. U.S.A.* 99, 11940–11945. doi: 10.1073/pnas.172393799

Furlong, T. M., Supit, A. S., Corbit, L. H., Killcross, S., and Balleine, B. W. (2015). Pulling habits out of rats: adenosine 2A receptor antagonism in dorsomedial striatum rescues meth-amphetamine-induced deficits in goal-directed action. *Addiction Biol.* doi: 10.1111/adb.12316. [Epub ahead of print].

Hilário, M. R., Clouse, E., Yin, H. H., and Costa, R. M. (2007). Endocannabinoid signaling is critical for habit formation. *Front. Integr. Neurosci.* 1:6. doi: 10.3389/neuro.07.006.2007

Kachroo, A., Orlando, L. R., Grandy, D. K., Chen, J. F., Young, A. B., and Schwarzschild, M. A. (2005). Interactions between metabotropic glutamate 5 and adenosine A$_{2A}$ receptors in normal and parkinsonian mice. *J. Neurosci.* 25, 10414–10419. doi: 10.1523/JNEUROSCI.3660-05.2005

Kreitzer, A. C., and Malenka, R. C. (2007). Endocannabinoid-mediated rescue of striatal LTD and motor deficits in Parkinson's disease models. *Nature* 445, 643–647. doi: 10.1038/nature05506

Li, Y., He, Y., Chen, M., Pu, Z., Chen, L., Li, P., et al. (2016). Optogenetic activation of adenosine A$_{2A}$ receptor signaling in the dorsomedial striatopallidal neurons suppresses goal-directed behavior. *Neuropsychopharmacology* 41, 1003–1013. doi: 10.1038/npp.2015.227

Lobo, M. K., Cui, Y., Ostlund, S. B., Balleine, B. W., and Yang, X. W. (2007). Genetic control of instrumental conditioning by striatopallidal neuron-specific S1P receptor Gpr6. *Nat. Neurosci.* 10, 1395–1397. doi: 10.1038/nn1987

Lovinger, D. M. (2010). Neurotransmitter roles in synaptic modulation, plasticity and learning in the dorsal striatum. *Neuropharmacology* 58, 951–961. doi: 10.1016/j.neuropharm.2010.01.008

Nam, H. W., Hinton, D. J., Kang, N. Y., Kim, T., Lee, M. R., Oliveros, A., et al. (2013). Adenosine transporter ENT1 regulates the acquisition of goal-directed behavior and ethanol drinking through A$_{2A}$ receptor in the dorsomedial striatum. *J. Neurosci.* 33, 4329–4338. doi: 10.1523/JNEUROSCI.3094-12.2013

Nazzaro, C., Greco, B., Cerovic, M., Baxter, P., Rubino, T., Trusel, M., et al. (2012). SK channel modulation rescues striatal plasticity and control over habit in cannabinoid tolerance. *Nat. Neurosci.* 15, 284–293. doi: 10.1038/nn.3022

Partridge, J. G., Tang, K. C., and Lovinger, D. M. (2000). Regional and postnatal heterogeneity of activity-dependent long-term changes in synaptic efficacy in the dorsal striatum. *J. Neurophysiol.* 84, 1422–1429.

Pinna, A., Bonaventura, J., Farre, D., Sanchez, M., Simola, N., Mallol, J., et al. (2014). L-DOPA disrupts adenosine A$_{2A}$-cannabinoid CB$_1$-dopamine D$_2$ receptor heteromer cross-talk in the striatum of hemiparkinsonian rats: biochemical and behavioral studies. *Exp. Neurol.* 253, 180–191. doi: 10.1016/j.expneurol.2013.12.021

Rossi, M. A., and Yin, H. H. (2012). Methods for studying habitual behavior in mice. *Curr. Protoc. Neurosci.* Chapter 8, Unit 8.29. doi: 10.1002/0471142301.ns0829s60

Shen, W., Flajolet, M., Greengard, P., and Surmeier, D. J. (2008). Dichotomous dopaminergic control of striatal synaptic plasticity. *Science* 321, 848–851. doi: 10.1126/science.1160575

Svenningsson, P., Le Moine, C., Fisone, G., and Fredholm, B. B. (1999). Distribution, biochemistry and function of striatal adenosine A$_{2A}$ receptors. *Prog. Neurobiol.* 59, 355–396. doi: 10.1016/S0301-0082(99)00011-8

Taura, J. V., Fernández-Due-as, and Ciruela, F. (2015). Visualizing G protein-coupled receptor-receptor interactions in brain using proximity ligation *in situ* assay. *Curr. Protoc. Cell Biol.* 67, 17.17.1–17.17.16. doi: 10.1002/0471143030.cb1717s67

Tebano, M. T., Martire, A., Chiodi, V., Pepponi, R., Ferrante, A., Domenici, M. R., et al. (2009). Adenosine A$_{2A}$ receptors enable the synaptic effects of cannabinoid CB1 receptors in the rodent striatum. *J. Neurochem.* 110, 1921–1930. doi: 10.1111/j.1471-4159.2009.06282.x

Torvinen, M., Torri, C., Tombesi, A., Marcellino, D., Watson, S., Lluis, C., et al. (2005). Trafficking of adenosine A$_{2A}$ and dopamine D2 receptors. *J. Mol. Neurosci.* 25, 191–200. doi: 10.1385/JMN:25:2:191

Trifilieff, P., Rives, M. L., Urizar, E., Piskorowski, R. A., Vishwasrao, H. D., Castrillon, J., et al. (2011). Detection of antigen interactions *ex vivo* by proximity ligation assay: endogenous dopamine D2-adenosine A$_{2A}$ receptor complexes in the striatum. *Biotechniques* 51, 111–118. doi: 10.2144/000113719

Yin, H. H., and Knowlton, B. J. (2006). The role of the basal ganglia in habit formation. *Nat. Rev. Neurosci.* 7, 464–476. doi: 10.1038/nrn1919

Yin, H. H., Knowlton, B. J., and Balleine, B. W. (2005). Blockade of NMDA receptors in the dorsomedial striatum prevents action-outcome learning in instrumental conditioning. *Eur. J. Neurosci.* 22, 505–512. doi: 10.1111/j.1460-9568.2005.04219.x

Yin, H. H., Mulcare, S. P., Hilario, M. R., Clouse, E., Holloway, T., Davis, M. I., et al. (2009). Dynamic reorganization of striatal circuits during the acquisition and consolidation of a skill. *Nat. Neurosci.* 12, 333–341. doi: 10.1038/nn.2261

Yu, C., Gupta, J., Chen, J. F., and Yin, H. H. (2009). Genetic deletion of A$_{2A}$ adenosine receptors in the striatum selectively impairs habit formation. *J. Neurosci.* 29, 15100–15103. doi: 10.1523/JNEUROSCI.4215-09.2009