

# Activation of Dorsal Raphe Serotonin Neurons Is Necessary for Waiting for Delayed Rewards

Kayoko W. Miyazaki,<sup>1</sup> Katsuhiko Miyazaki,<sup>1</sup> and Kenji Doya<sup>1,2</sup>

<sup>1</sup>Neural Computation Unit, Okinawa Institute of Science and Technology, Okinawa 904-0412, Japan, and <sup>2</sup>Computational Neuroscience Laboratories, Advanced Telecommunications Research Institute International, Kyoto 619-0288, Japan

The forebrain serotonergic system is a crucial component in the control of impulsive behaviors. We previously reported that the activity of serotonin neurons in the midbrain dorsal raphe nucleus increased when rats performed a task that required them to wait for delayed rewards. However, the causal relationship between serotonin neural activity and the tolerance for the delayed reward remained unclear. Here, we test whether the inhibition of serotonin neural activity by the local application of the 5-HT<sub>1A</sub> receptor agonist 8-hydroxy-2-(di-*n*-propylamino) tetralin in the dorsal raphe nucleus impairs rats' tolerance for delayed rewards. Rats performed a sequential food-water navigation task that required them to visit food and water sites alternately via a tone site to get rewards at both sites after delays. During the short (2 s) delayed reward condition, the inhibition of serotonin neural activity did not significantly influence the numbers of reward choice errors (nosepoke at an incorrect reward site following a conditioned reinforcer tone), reward wait errors (failure to wait for the delayed rewards), or total trials (sum of reward choice errors, reward wait errors, and acquired rewards). By contrast, during the long (7–11 s) delayed reward condition, the number of wait errors significantly increased while the numbers of total trials and choice errors did not significantly change. These results indicate that the activation of dorsal raphe serotonin neurons is necessary for waiting for long delayed rewards and suggest that elevated serotonin activity facilitates waiting behavior when there is the prospect of forthcoming rewards.

## Introduction

Several studies have shown that a reduced level of serotonin (5-hydroxytryptamine, 5-HT) in the CNS promotes impulsive behaviors (Evenden, 1999; Cardinal, 2006), including motor impulsivity (the failure to suppress inappropriate actions) and choice impulsivity (the choice of small immediate rewards over larger delayed rewards). The dorsal raphe nucleus (DRN) is the major origin of serotonergic projections to the forebrain. DRN serotonergic neurons also release serotonin in the DRN through their local axon collaterals. Recently, we showed that serotonin efflux in the DRN increases while rats perform a task that requires waiting for a delayed reward (Miyazaki et al., 2011a). We also found that DRN serotonergic neurons increase tonic firing while rats wait for delayed rewards and cease firing before rats give up waiting for long delayed rewards (Miyazaki et al., 2011b). These results demonstrated a correlation between dorsal raphe serotonin activation and waiting behavior for delayed rewards. However, whether the

activation of serotonergic neurons is causal and necessary for waiting for a delayed reward remains unclear.

The aim of the current research is to test any causal relationship between serotonergic activity and waiting behavior for delayed reward. To examine this causality, we applied a 5-HT<sub>1A</sub> receptor agonist locally in the DRN; this treatment is known to suppress 5-HT neural activity through autoreceptors. Systemic application of 5-HT<sub>1A</sub> receptor agonists has been commonly used to suppress 5-HT neural activity in the DRN (Carli and Samanin, 2000; Casanovas et al., 2000; Waterhouse et al., 2004; Miyazaki et al., 2011b). However, systemic application can affect 5-HT<sub>1A</sub> receptors in both the DRN and the projection sites, which can have different contributions to animal behaviors (Carli and Samanin, 2000; Assié et al., 2010). In this research, to achieve sustained reduction of 5-HT neural activity during rats' task performance, we applied the 5-HT<sub>1A</sub> receptor agonist 8-hydroxy-2-(di-*n*-propylamino) tetralin (8-OH-DPAT) locally into the DRN by reverse dialysis without affecting 5-HT<sub>1A</sub> receptors in the target areas (Adell et al., 1993; Casanovas et al., 2000). Using dual probe microdialysis, we assessed the effect of 5-HT<sub>1A</sub> receptor agonist application in the DRN by monitoring 5-HT efflux in the medial prefrontal cortex (mPFC) (Casanovas et al., 2000). Rats performed a sequential food-water navigation task in which they alternately visited food and water sites to acquire rewards after waiting periods (Fig. 1). The rats performed the task under two reward conditions: a short delayed reward (SDR) condition in which they had to wait for 2 s, and a long delayed reward (LDR) condition in which they had to wait for 7–11 s (delay length was different for each rat and fixed during the test session for individual rats). We found that the suppression of 5-HT neural activ-

Received Feb. 20, 2012; revised April 29, 2012; accepted May 17, 2012.

Author contributions: K.W.M., K.M., and K.D. designed research; K.W.M. performed research; K.W.M. and K.M. analyzed data; K.W.M., K.M., and K.D. wrote the paper.

We thank the members of the Neural Computation Unit for their helpful comments and discussion. We also thank Kiyoko Arakaki for her kind laboratory animal care. A part of this study was supported by "Integrated research on neuropsychiatric disorders" carried out under the Strategic Research Program for Brain Sciences by the Ministry of Education, Culture, Sports, Science and Technology of Japan, and Grant-in-Aid for Scientific Research on Innovative Areas: Prediction and Decision Making (23120007).

This article is freely available online through the *JNeurosci* Open Choice option.

Correspondence should be addressed to either of the following: Kayoko W. Miyazaki or Kenji Doya, Neural Computation Unit, Okinawa Institute of Science and Technology, 1919-1 Tancha, Onna, Okinawa 904-0412, Japan. E-mail: kmizayaki@oist.jp or doya@oist.jp.

DOI:10.1523/JNEUROSCI.0915-12.2012

Copyright © 2012 the authors 0270-6474/12/3210451-07\$15.00/0

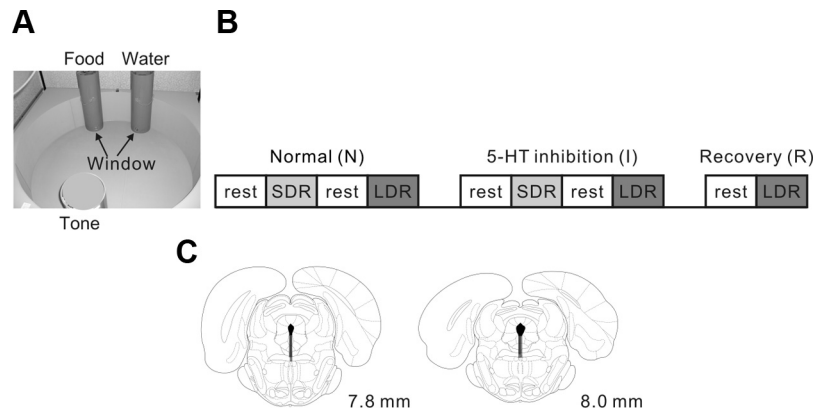
ity in the DRN increased premature exit from reward sites before the delivery of delayed rewards, which indicated impaired patience for delayed rewards.

## Materials and Methods

**Subjects.** All experimental procedures were performed in accordance with the guidelines determined by the Okinawa Institute of Science and Technology Experimental Animal Committee. Ten male Long–Evans rats (Japan SLC) weighing 300–340 g at the beginning of the behavioral training period were used in the present study. The animals were housed with one rat per cage at 24°C on a 12 h light/dark cycle (lights on 8:00 A.M. to 8:00 P.M.). Eight animals contributed to the data reported here. Two animals were excluded because of either the position of the microdialysis probe (one) or machine trouble during the microdialysis experiment (one). All training and test sessions were conducted during the light period 5–6 d per week. The rats were deprived of food and water in their home cage and received their daily food and water rations during the experimental sessions only (~15 g/d and 20 ml/d, respectively).

**Behavioral apparatus and training.** A free operant task that we designed as a sequential food–water navigation task was used. The rats were individually trained and tested in a cylindrical apparatus 1.5 m in diameter with a 45 cm high wall; three cylinders that were identical in appearance were used as the tone site, and the food site and the water site were fixed in an isosceles triangle (Fig. 1A). Each cylinder contained a rectangular window 4 cm long, 3 cm wide, and 2.3 cm deep that was positioned 5 cm above the floor. An infrared photo-beam crossed the entrances of all of the windows to detect the rats' nose poke responses positioned at a depth of 0.5 and 1 cm from the bottom of the window. The open field apparatus was surrounded by a soundproof box (2.5 m square and 2.4 m in height). Four 100 W lamps were set at the four corners of the box and indirectly illuminated the open field. One speaker and one camera were positioned above the open field. The position of the rat was monitored by a video tracking system (CV-2000; Keyence). When the rat poked its nose through the small windows in the three cylinders, the control infrared photo-beam was interrupted to detect the rat's responses. The tone site nose poke induced an 8 kHz tone (tone 1: 0.4 s, 70 dB) from the speaker. At the food site, a small food pellet (45 mg) was delivered into the window through a food dispenser. At the water site, a spout of water (length: 0.5 cm, duration: 2–4 s) was provided at the front wall of the window positioned 0.5 cm below the center.

The beginning of the sequential food–water navigation task was signaled by turning the four lamps on, and termination was indicated by only one lamp being on. The behavioral instrumental response in this task was for the rat to hold its nose in a fixed posture in either the tone site window, waiting for the conditioned reinforcement tone, or the reward site windows, waiting for the food or water rewards. This task required the rats to perform alternate visits and nose pokes to the food site and the water site via visiting the tone site and performing a nose poke there. The rats initiated a trial by nose poking in a fixed posture to achieve continuous interruption of the photo-beam at the tone site during a delay period until tone 1 was presented, signaling that a reward was available at one of the reward sites. After presentation of tone 1, the rat was required to continue nose poking at one of the reward sites during another delay period until the reward was delivered. To continue the task, the rats had to alternately visit two reward sites via the tone site. In this task, food was always rewarded first. We called the delay periods that preceded the tone and the rewards (food and water) the tone delay and the reward delay, respectively. Depending on whether tone 1 preceded the presentation of food or water, we referred to tone 1 as either the food tone or the water



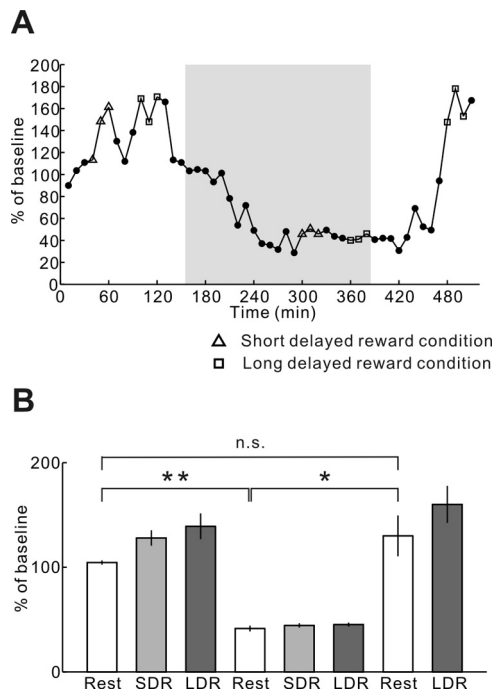
**Figure 1.** Experimental apparatus and design of the sequential food–water navigation task with variable delayed rewards. **A**, Open field, reward cylinders (food site, water site) and tone cylinder (tone site) for the task. Windows for nose pokes (reward locations) are indicated. The tone cylinder also contained a small window at the same position as the reward cylinders. **B**, Schematic of the movements required for the rats to receive rewards at both sites. The rats performed the task under two reward conditions. In the SDR condition, the rats obtained a reward after 2 s of nose poking at each reward site. In the LDR condition, the rats were required to poke their nose for 7–11 s before obtaining the reward. Each reward condition lasted for 30 min and was preceded by a rest period. **C**, Locations of the microdialysis probes in the DRN. Black bars represent the 2 mm length of the dialysis membranes. Numbers beside each plate correspond to millimeters from bregma. Coronal drawing modified from Paxinos and Watson (1998) with permission.

tone, respectively. The food and water delay periods were always set at the same length. During the initial training period, the tone delay was fixed at 0.2 s, and the reward delay was set to 0 s. The reward sites did not change until the rats were able to receive a reward at one of the sites.

Two types of error were present in this task: reward choice error and reward wait error. Reward choice error occurred when the rat nose poked at an incorrect reward site following tone 1. Food choice error entailed the rat nose poking at the water site when the correct nose poke location was the food site, whereas the reverse was the case for water choice errors. Reward wait error indicated that the rat failed to wait for the rewards during the delay period by keeping his nose in a fixed posture. Reward choice error was signaled by a 500 Hz tone (tone 2: 0.4 s, 72 dB) immediately after a nose poke at the wrong reward site. The occurrence of the two types of wait errors (food and water wait error) was not signaled. The rats could start the next trial at any time after reward consumption or after making a reward choice error or a wait error. The rats were trained daily for a period of 2 h. It took 3 weeks or less for the rats to learn the sequential food–water navigation task.

**Surgery and cannulation.** After the rats had mastered the sequential food–water navigation task, they were anesthetized with equithesin (3 ml/kg, i.p.), and a guide cannula (AG-8; Eicom) was stereotaxically implanted in the DRN (from bregma: posterior, –7.8 mm; lateral, 0 mm; ventral, –5.5 mm) and the mPFC (from bregma: posterior, 3.2 mm; lateral, 0.8 mm; ventral, –3.0 mm) according to the atlas produced by Paxinos and Watson (1998). Four rats were implanted with two guide cannulae at both the DRN and the mPFC. Six rats were implanted at the DRN only. The guide cannula was fixed onto the skull and anchored with dental acrylic and stainless steel screws. A dummy cannula (AD-8; Eicom) was inserted into the guide cannula and secured to the guide cannula with a cap nut (AC-1; Eicom) to prevent infection and occlusions. The animals were housed individually after surgery and were allowed at least 1 week to recover.

**Task sequence with two reward conditions.** After recovering from surgery, the rats were retrained on a task sequence (a 30 min task performance followed by a 30 min rest period) consecutively four or five times per day. After the rats had mastered the task sequence, we introduced two reward conditions: a short reward delay and a long reward delay. In the SDR condition, a rat could obtain a reward when it nose poked each reward site for 2 s. In the LDR condition, the rats were required to keep their nose in the reward site for 7–11 s before the reward was delivered. Removing the nose for >500 ms before the end of this delay period caused a wait error, in which no reward was presented. The rats were



**Figure 2.** Changes in serotonin efflux in the mPFC during a task sequence. **A**, A typical example of serotonin efflux in the mPFC during the task sequence. **B**, Average serotonin levels in the mPFC during all reward conditions ( $n = 3$ ;  $\pm$  SEM). Asterisks indicate significant differences, as assessed by a paired  $t$  test,  $*p < 0.05$ ,  $**p < 0.01$ .

trained in five SDR task sequences per day for five successive days. One day before a pharmacological test session, we examined the delay length for the LDR condition for the individual rats. The rats were tested for an extended reward delay condition in which the reward delay at both sites was increased gradually every 300 s (3, 5, 7, 9, 11, and 13 s). The rats performed the extended reward delay condition task four times, with a 30 min rest period between each task. The delay length before giving up waiting was defined as the delay period when the rats stopped performing the task under the extended reward delay condition. A maximal integer smaller than the mean give-up waiting time was used for the reward delay length of the LDR condition for the individual rats. Estimated delay length for individual rats was fixed within all LDR tests in the test session for each rat.

**Suppression of DRN serotonergic neural activity during task performance.** A dialysis probe inserted in the DRN was used for local infusion using reverse dialysis to deliver the 5-HT<sub>1A</sub> receptor agonist 8-OH-DPAT (300  $\mu$ M; Sigma), which has been shown to inhibit serotonin neuronal firing through the activation of 5-HT<sub>1A</sub> autoreceptors (Fig. 2). A dialysis probe in the mPFC was used to monitor serotonin efflux to examine the inhibitory effect of 8-OH-DPAT on serotonergic neural activity (Casanovas et al., 2000). On the day before the test session, a dialysis probe (A-I-8-03; length 3 mm, outer diameter 0.22 mm, 50,000 molecular weight cutoff; Eicom) was calibrated using a standard serotonin solution (20 mM NaH<sub>2</sub>PO<sub>4</sub>, 28 nM serotonin) before each experiment. The probe recovery (mean  $\pm$  SEM) for serotonin was  $13.1 \pm 4.2\%$  ( $n = 3$  for serotonin). The detection limit was  $\sim 30$  fg per sample. Following calibration, the probe was carefully inserted into the guide cannula of the mPFC. A dialysis probe (A-I-8-02; length 2 mm, outer diameter 0.22 mm, 50,000 molecular weight cutoff; Eicom) was also carefully inserted into the guide cannula of the DRN. Each probe was secured to the guide cannula with a screw. The inlet and outlet of each probe were connected to a five channel swivel (MCS/5A; Instech) through a freely moving tube (WT-20T; Eicom). The rats were placed in the open field overnight. The perfusion rate of the modified Ringer's solution was changed to 0.2  $\mu$ l/min. Eight rats were used for this experiment. Three of 8 rats were implanted with two dialysis probes (0.22 mm in diameter, 2 and 3 mm tip length for the DRN and the mPFC, respec-

tively) in the DRN and the mPFC simultaneously. Five rats were implanted with one dialysis probe into the DRN.

On the experimental day, the flow rate of the Ringer's solution was increased to 1  $\mu$ l/min. To augment the levels of serotonin in the dialysate of the mPFC probe, the perfusate of the mPFC probe contained a low concentration of citalopram (1  $\mu$ M) (Sigma). The perfusate of the DRN probe did not contain citalopram. It took  $\sim 60$  min to stabilize the serotonin baseline levels. The outflow of the mPFC probe was collected in a sample loop and injected automatically by an auto-injector (ESA-20; Eicom) once every 10 min into a high-performance liquid chromatography apparatus with electrochemical detection (HTEC-500; Eicom). Extracellular serotonin levels in the mPFC were measured by the high-performance liquid chromatography apparatus using electrochemical detection every 10 min. An Eicompact PP-ODS column (4.6 mm inner diameter  $\times$  30 mm, Eicom) was used to separate serotonin. The mobile phase contained 100 mM sodium phosphate buffer (pH 6.0, Wako Pure Chemicals), 2.0 mM sodium 1-decanesulfonate (Tokyo Chemical Industry), 0.1 mM disodium EDTA (Dojindo), and 1% (v/v) methanol (Wako Pure Chemicals). The flow rate was 500  $\mu$ l/min, and the system temperature was 25°C. The concentration of serotonin was measured by setting the working electrode at +400 mV against an Ag/AgCl reference electrode.

After a stable baseline level of serotonin was obtained, three baseline samples were collected. In the test session, the rats performed the following task sequence: rest (baseline) (30 min)—SDR (30 min)—rest (30 min)—LDR (30 min)—8-OH-DPAT injection to the DRN by reverse dialysis (90 min)—rest (30 min)—SDR (30 min)—rest (30 min)—LDR (30 min)—washout of 8-OH-DPAT (90 min)—rest (30 min)—LDR (30 min) (Fig. 1B). We measured the number of acquired rewards (food and water sites), reward choice errors (food and water sites), and reward wait errors (food and water sites). After the pharmacological experiment had been completed (day 1), the rats were placed in the open field overnight. The perfusion rate of modified Ringer's solution was changed to 0.2  $\mu$ l/min. Approximately 23 h after the previous day's experiment had started, the same task sequence performed on day 1 was conducted on experimental day 2.

**Histology.** The rats were deeply anesthetized with 100 mg/kg sodium pentobarbital intraperitoneally and then perfused with 0.9% NaCl followed by 10% formalin. Their brains were removed and stored in 10% formalin for a minimum of 24 h before being prepared into 60 mm coronal sections. Cresyl violet staining was used to help verify the placement of the probe tracts.

**Data analysis.** The behavioral data from 3 rats that were implanted with a dialysis probe in the mPFC and the DRN were omitted to avoid any possibility of the effect of the citalopram. The behavioral results were collected and analyzed from 5 rats that were implanted with a dialysis probe only in the DRN. The neurochemical data were transformed into percentage changes from baseline and were analyzed using a one-way repeated-measures ANOVA followed by Tukey's HSD test for multiple comparisons. Two-way ANOVA with task execution time intervals of 10 min (three levels; first, second, and last 10 min) and drug conditions (three levels; normal, drug, and recovery) as within-subjects factors were used for analysis of behavioral data for every 10 min period (Fig. 3). Two-way ANOVA using reward type (two levels; food and water) and drug conditions (three levels; normal, drug, and recovery) as within-subjects factors were used for analysis of behavioral data during the 30 min task periods (Figs. 4, 5). Statistical analyses were performed using the SPSS and Matlab (MathWorks) statistical packages.

**Error classification.** There were three possible patterns when reward choice error occurred. Water choice error occurred after (1) food reward success (from food site to food site), (2) water wait error (from water site to food site), or (3) water choice error (from food site to food site). Food choice error was preceded by (1) water reward success, (2) food wait error, or (3) food choice error. There were three possible patterns when reward wait error occurred. Water wait error occurred after (1) food reward success (from food site to water site), (2) water wait error (from water site to wait site), or (3) water choice error (from food site to water site). Food wait error was preceded by (1) water reward success, (2) food wait error, or (3) food choice error.

## Results

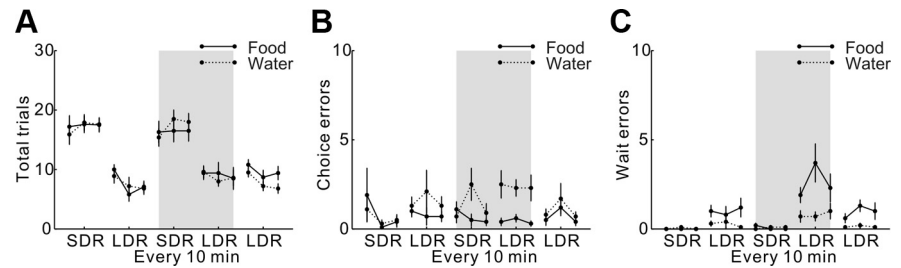
To examine the effect of reverse dialysis application of 8-OH-DPAT (300  $\mu\text{M}$ , 1  $\mu\text{l}/\text{min}$ ) in the DRN, we monitored serotonin efflux in the mPFC by implanting a second microdialysis probe containing citroplam (1  $\mu\text{M}$ , 1  $\mu\text{l}/\text{min}$ ) in the perfusate to enhance sensitivity (Fig. 2A). Local application of 8-OH-DPAT significantly reduced serotonin efflux in the mPFC compared with the control condition in which only the vehicle (Ringer's solution) was applied (paired  $t$  test,  $p = 0.0048$ ,  $n = 3$ ) (Fig. 2B). Dopamine efflux in the mPFC did not significantly change by local application of 8-OH-DPAT in the DRN (paired  $t$  test,  $p = 0.23$ ,  $n = 3$ ). Sixty minutes after the application of 8-OH-DPAT in the DRN, serotonin efflux in the mPFC was reduced to  $\sim 40\%$  of the baseline level and recovered to the baseline level within 90 min after the perfusate was switched to Ringer's solution (paired  $t$  test,  $p = 0.35$ ) (Fig. 2B).

### Rat behavior during the three drug conditions

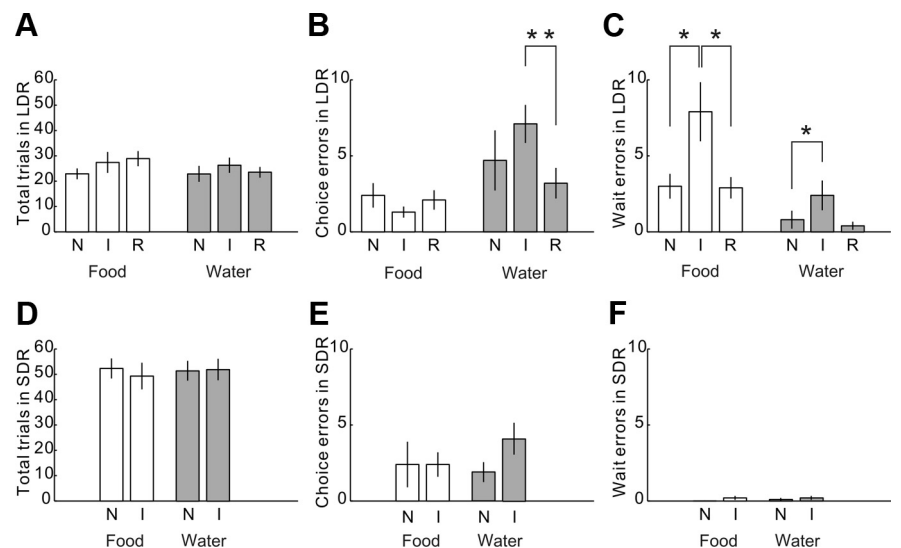
Behavioral data were collected from five rats implanted with a single dialysis probe in the DRN. The behavioral data from rats implanted with two dialysis probes were excluded from the analysis to avoid any effect of citroplam applied to the mPFC on the rats' task behavior. Probe replacements within the DRN are shown in Figure 1C. In both the SDR and LDR conditions, the rats continued to perform the task during the 30 min task period. Figure 3 shows the number of total trials (Fig. 3A), reward choice errors (Fig. 3B), and reward wait errors (Fig. 3C) every 10 min in the order of the task sequence; data are shown separately for the trials for the food site and the water site.

The number of total trials in every 10 min did not significantly differ between the normal and drug conditions in both food and water trials in both the SDR (main effect of the Drug Condition during the food trial,  $F_{(1,9)} = 1.03$ ,  $p = 0.34$ ; during the water trial,  $F_{(1,9)} = 0.10$ ,  $p = 0.76$ ) and LDR conditions (during the food trial,  $F_{(2,18)} = 1.48$ ,  $p = 0.25$ ; during the water trial,  $F_{(2,18)} = 0.74$ ,  $p = 0.49$ ). The number of total trials was largest during the second 10 min period in the water trials in the SDR condition, although there was no significant difference in the food trials (main effect of Time during the food trials,  $F_{(2,18)} = 0.09$ ,  $p = 0.92$ ; during the water trials,  $F_{(2,18)} = 8.38$ ,  $p = 0.003$ ). There was a decrease over time in the number of total trials shown by the significant main effect of Time in both food and water trials in the LDR condition (main effect of Time during the food trial,  $F_{(2,18)} = 4.23$ ,  $p = 0.03$ ; during the water trial,  $F_{(2,18)} = 8.64$ ,  $p = 0.002$ ) (Fig. 3A). There was no significant interaction between the Drug Condition and Time in both the SDR ( $F < 0.60$ ,  $p > 0.56$ ) and LDR conditions ( $F < 1.87$ ,  $p > 0.14$ ).

The number of reward choice errors did not significantly differ between the normal and drug conditions in either the food or water



**Figure 3.** Effect of inhibition of serotonergic neural activity on the sequential food-water navigation task with variable delayed rewards. **A–C**, Markers represent the means of total trials (**A**), choice errors (**B**), and wait errors (**C**) every 10 min during all reward conditions. Error bars represent  $\pm$  SEM ( $n = 10$ ). Straight line, Food-related trial; dotted line, water-related trial. Gray color indicates the number of each trial in the serotonin inhibition condition.



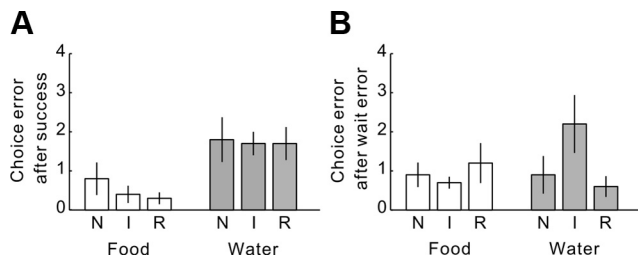
**Figure 4.** Number of behavioral events during the different reward and drug conditions. Behavioral events during normal (N), inhibition (I), and recovery (R) conditions are shown. **A–C**, Markers represent the means of the total trials (**A**), choice errors (**B**), and wait errors (**C**) during a 30 min period in the LDR condition. **D–F**, Markers represent the means of the total trials (**D**), choice errors (**E**), and wait errors (**F**) during a 30 min period in the SDR condition. Error bars represent  $\pm$  SEM ( $n = 10$ ). Asterisks indicate significant differences assessed by a paired  $t$  test, \* $p < 0.05$ , \*\* $p < 0.01$ . White bar, Food-related trial; gray bar, water-related trial.

trials in either the SDR (main effect of the Drug Condition, food choice error,  $F_{(1,9)} = 0.05$ ,  $p = 0.82$ ; water choice error,  $F_{(1,9)} = 2.34$ ,  $p = 0.16$ ) or LDR conditions (main effect of the Drug Condition, food choice error,  $F_{(2,18)} = 1.08$ ,  $p = 0.44$ ; water choice error,  $F_{(2,18)} = 3.48$ ,  $p = 0.053$ ). There were no significant effects of Time or interaction between Drug Condition and Time in either the SDR ( $F < 2.51$ ,  $p > 0.11$ ) or LDR conditions ( $F < 1.01$ ,  $p > 0.38$ ) (Fig. 3B).

Food and water wait errors rarely occurred in the SDR condition and did not significantly differ between the normal and drug conditions (main effect of the Drug Condition, food wait error,  $F_{(1,9)} = 2.25$ ,  $p = 0.17$ ; water wait error,  $F_{(1,9)} = 0.31$ ,  $p = 0.59$ ). In contrast, in the LDR condition, the number of food and water wait errors increased during the infusion of 8-OH-DPAT into the DRN (main effect of the Drug Condition, food wait error,  $F_{(2,18)} = 5.45$ ,  $p = 0.014$ ; water wait error,  $F_{(2,18)} = 4.18$ ,  $p = 0.032$ ) (Fig. 3C). Neither the main effect of Time nor the Drug Condition  $\times$  Time interaction was significant in either the SDR ( $F < 2.25$ ,  $p > 0.13$ ) or LDR conditions ( $F < 1.75$ ,  $p > 0.20$ ).

### Number of behavioral events during different reward and drug conditions

Figure 4 shows the number of total trials, reward choice errors, and reward wait errors during the 30 min task period of the LDR



**Figure 5.** Number of reward choice errors classified by previous trials. Reward choice errors during normal (N), inhibition (I), and recovery (R) conditions are shown. **A**, Markers represent the means of the food choice errors and water choice errors preceded by water success (left) and food success (right). **B**, Markers represent the means of the food choice errors and water choice errors that occurred after food wait errors (left) and water wait errors (right). Error bars represent  $\pm$  SEM. White bar, Food-related trial; gray bar, water-related trial.

(Fig. 4A–C) and SDR (Fig. 4D–F) conditions before, during, and after 8-OH-DPAT infusion.

In the LDR condition, the number of total trials did not significantly change (main effect of the Reward,  $F_{(1,9)} = 1.47$ ,  $p = 0.26$ ; main effect of the Drug Condition,  $F_{(2,18)} = 1.05$ ,  $p = 0.37$ ; Reward  $\times$  Drug Condition interaction,  $F_{(2,18)} = 1.63$ ,  $p = 0.22$ ) (Fig. 4A). The number of water choice errors in the LDR condition was larger than the number of food choice errors (main effect of the Reward,  $F_{(1,9)} = 5.58$ ,  $p = 0.042$ ) (Fig. 4B). Although there was no significant main effect of the drug condition (main effect of the Drug Condition,  $F_{(2,18)} = 2.51$ ,  $p = 0.11$ ), water choice errors significantly increased during the inhibition condition compared with the recovery condition (paired  $t$  test, normal vs inhibition  $p = 0.081$ ; inhibition vs recovery  $p = 0.0056$ ). There was no significant interaction between the reward and drug conditions (Reward  $\times$  Drug Condition interaction,  $F_{(2,18)} = 2.97$ ,  $p = 0.077$ ) (Fig. 4B). The food and water wait errors in the LDR condition increased during 8-OH-DPAT treatment, and the number of food wait errors was larger than that of water wait errors (main effect of the Reward,  $F_{(1,9)} = 10.29$ ,  $p = 0.011$ ; main effect of the Drug Condition,  $F_{(2,18)} = 12.89$ ,  $p = 0.0003$ ; Reward  $\times$  Drug Condition interaction,  $F_{(2,18)} = 1.42$ ,  $p = 0.27$ ) (Fig. 4C). Food wait errors significantly increased during the inhibition condition compared with the other two drug conditions (paired  $t$  test, inhibition vs normal  $p = 0.046$ ; inhibition vs recovery  $p = 0.024$ ). The number of water wait errors during the inhibition condition also increased compared with the other two drug conditions (paired  $t$  test, inhibition vs normal  $p = 0.048$ ; inhibition vs recovery  $p = 0.063$ ).

In the SDR condition, the number of total trials (Fig. 4D), choice errors (Fig. 4E), and wait errors (Fig. 4F) did not significantly differ for the reward or drug conditions (Fig. 4D, main effect of the Reward,  $F_{(1,9)} = 0.71$ ,  $p = 0.42$ ; main effect of the Drug Condition,  $F_{(1,9)} = 0.38$ ,  $p = 0.55$ ; Reward  $\times$  Drug Condition interaction,  $F_{(1,9)} = 2.04$ ,  $p = 0.19$ ) (Fig. 4E, main effect of the Reward,  $F_{(1,9)} = 0.33$ ,  $p = 0.58$ ; main effect of the Drug Condition,  $F_{(1,9)} = 1.91$ ,  $p = 0.20$ ; Reward  $\times$  Drug Condition interaction,  $F_{(1,9)} = 1.70$ ,  $p = 0.22$ ) (Fig. 4F, main effect of the Reward,  $F_{(1,9)} = 0.31$ ,  $p = 0.59$ ; main effect of the Drug Condition,  $F_{(1,9)} = 1.33$ ,  $p = 0.28$ ; Reward  $\times$  Drug Condition interaction,  $F_{(1,9)} = 0.31$ ,  $p = 0.59$ ).

### Reward choice errors classified on the basis of previous behavioral events

In the present task, after a wait error trial, the rat was supposed to visit the same site again to get a reward. If the rat goes to the other reward site after a wait error, a choice error is flagged. To examine

the possibility that the increase in water choice errors was a secondary effect of increased water wait errors (Fig. 4C), we classified the choice errors according to the performance in the previous trial under the LDR condition. After successful trials, water choice errors (heading erroneously to food) occurred more frequently than food choice errors (heading erroneously to water), suggesting that rats have a choice bias of food over water. There was no significant main effect of the Drug Condition (main effect of the Reward,  $F_{(1,9)} = 12.07$ ,  $p = 0.007$ ; main effect of the Drug Condition,  $F_{(2,18)} = 0.48$ ,  $p = 0.63$ ; Reward  $\times$  Drug Condition interaction,  $F_{(2,18)} = 0.16$ ,  $p = 0.86$ ) (Fig. 5A). After wait error trials, while food choice errors (following a food wait error) were not affected (one-way ANOVA,  $F_{(2,18)} = 0.50$ ,  $p = 0.61$ ), water choice errors (following a water wait error) increased during the drug infusion condition (one-way ANOVA,  $F_{(2,18)} = 3.73$ ,  $p = 0.04$ ) (Fig. 5B). Note that, because of the small number of samples, the difference did not reach significance in a two-way ANOVA (main effect of the Reward,  $F_{(1,9)} = 0.52$ ,  $p = 0.49$ ; main effect of the Drug Condition,  $F_{(2,18)} = 2.87$ ,  $p = 0.083$ ; Reward  $\times$  Drug Condition interaction,  $F_{(2,18)} = 2.34$ ,  $p = 0.13$ ) and paired  $t$  test (inhibition vs normal  $p = 0.057$ ; inhibition vs recovery  $p = 0.061$ ). These results suggest that the increase in the water choice errors in the LDR condition during the inhibition condition (Fig. 4B) was a secondary effect of the increase in water wait errors (Fig. 4C) after which the rats were supposed to return to the water site.

### Discussion

We examined the causal relationship between patience for delayed reward and 5-HT neural activity through the inhibition of 5-HT neural activity by the local application of 8-OH-DPAT in the DRN using the reverse dialysis method (Casanovas et al., 2000). We found that, in the long delayed reward condition, the numbers of food and water wait errors significantly increased during the inhibition of 5-HT neural activity while the numbers of total food and water trials were not significantly influenced. Although the number of water choice errors also significantly increased during the inhibition condition compared with the recovery condition, it was a secondary effect of an increase in wait errors. The numbers of total trials, reward choice errors, and reward wait errors in the SDR condition did not significantly change under the inhibition of 5-HT neural activity. In summary, the inhibition of DRN serotonergic neural activity selectively impaired the rats' waiting behavior for long delayed rewards without impairing their behavior for short delayed rewards of their general activity levels. Our results show that activation of dorsal raphe serotonergic neurons is necessary for rats to exhibit waiting behavior in anticipation of forthcoming rewards. The results from the present study support the hypothesis that increased activity of serotonin neurons in the DRN signals a longer time scale for reward evaluation (Doya, 2002) and facilitates the behaviors involved in the expectation of delayed rewards.

In this research, we showed that the inhibition of DRN 5-HT neural activity disrupted rats' patience to wait for long delayed reward without affecting cognitive or motor function. In the SDR condition, the inhibition of 5-HT neural activity did not significantly influence the numbers of total trials, reward choice errors, or reward wait errors. Even in the LDR condition, although water choice errors increased as a result of the increase in water wait errors, after reward wait errors, rats did not perseverate with reward choice errors but revisited the same reward site until they obtained the reward and then visited the other reward site in the next trial. These results also show that the inhibition of DRN

5-HT neural activity did not impair the working memory that the rats used to determine where to go to get a reward in a particular trial after obtaining a reward at a specific site in the previous trial.

In microdialysis measurement experiments using a similar task, we previously found that 5-HT efflux significantly increased in a delayed reward condition but not in a reward omission condition compared with an immediate reward condition (Miyazaki et al., 2011a). We also found that the DRN 5-HT neuronal firing increased specifically while waiting for a delayed reward at reward sites and that the increase in 5-HT neural firing ceased before the rats gave up waiting for long delayed rewards (Miyazaki et al., 2011b). The present results revealed a causal relationship between 5-HT neural activity and waiting behavior for delayed rewards. This series of studies suggest that the dorsal raphe 5-HT system promotes waiting for future rewards and contributes to the modulation of patience for the attainment of rewards.

The systemic application of 5-HT<sub>1A</sub> receptor agonists has been commonly used to suppress 5-HT neural activity in the DRN (Carli and Samanin, 2000; Casanovas et al., 2000; Waterhouse et al., 2004; Miyazaki et al., 2011b). However, because 5-HT<sub>1A</sub> receptors are expressed not only on 5-HT neurons as autoreceptors but also as postsynaptic receptors in the target areas such as the prefrontal cortex and the hippocampus, the effects could be a mixture of the central serotonin suppression and selective activation of 5-HT<sub>1A</sub> receptors in the target areas. Some of the 5-HT<sub>1A</sub> receptor-expressing neurons project back to DRN 5-HT neurons, making the interpretation of the results even more difficult (Sharp et al., 2007). For example, in the 5-choice serial reaction time task (5-CSRTT), Carli and Samanin (2000) combined a systemic dose of 8-OH-DPAT and dorsal raphe injection of the 5-HT<sub>1A</sub> antagonist WAY 100635 to show that some of the effects of 8-OH-DPAT (slowing of responding and increased errors of omission) depend on the stimulation of postsynaptic 5-HT<sub>1A</sub> receptors. Assié and colleagues compared the effects of the highly selective postsynaptic 5-HT<sub>1A</sub> receptor agonist F15599 and the nonselective agonist F13714 and showed different relative effects on antidepressant-like activity (immobility in the forced swimming test) and 5-HT syndrome (forepaw treading and flat body posture) (Newman-Tancredi et al., 2009; Assié et al., 2010). These results suggest that the local application of 5-HT<sub>1A</sub> receptor agonist is desirable to activate presynaptic 5-HT<sub>1A</sub> receptors without affecting postsynaptic 5-HT<sub>1A</sub> receptors.

Acute systemic administration of serotonin selective reuptake inhibitors (SSRIs) is also thought to suppress 5-HT neural activity in the DRN (Hajós et al., 1995; Artigas et al., 1996). However, the behavioral effects of the acute systemic administration of SSRIs on impulsive behavior were not consistent. For example, some studies showed the administration of SSRIs increased the selection rate of a large, delayed reward over a small, immediate reward, indicating a decrease in impulsive choice (Bizot et al., 1988, 1999). By contrast, a lack of effect has also been reported (Evenden and Ryan, 1996). Furthermore, opposite effects of different doses of citalopram in a probabilistic reversal learning have been reported (Bari et al., 2010). These inconsistent results may be due to opposing effects of SSRI on 5-HT neuron firing and postsynaptic 5-HT concentration. Low doses of SSRIs may mainly affect 5-HT transporters in the DRN, which causes the suppression of 5-HT neural activity by the activation of 5-HT<sub>1A</sub> receptors (Hajós et al., 1995; Artigas et al., 1996). High doses of the SSRIs may also affect 5-HT transporters in the projection sites and activate postsynaptic 5-HT receptors (Invernizzi et al., 1992). In this research, application of the 5-HT<sub>1A</sub> receptor agonist 8-OH-DPAT locally into the DRN by reverse dialysis allowed us to assess the effect of reduction of 5-HT neural activity without

affecting 5-HT<sub>1A</sub> receptors in the target areas (Adell et al., 1993; Casanovas et al., 2000).

Classic theories suggested that central 5-HT neurons are involved in the behavioral inhibition associated with the prediction of negative rewards or punishment (Soubrié, 1986; Daw et al., 2002; Dayan and Huys, 2009; Boureau and Dayan, 2011; Cools et al., 2011). We propose that the 5-HT system is involved in the decrease of behavioral activity not only to avoid aversive events with a prediction of punishment but also to obtain a reward with a prediction of reward (Miyazaki et al., 2012). To clarify the decrease in these two behavioral activities, we defined the animals' behavior as "waiting to obtain reward" when they decreased their activity to obtain a reward and used the term "waiting to avoid punishment" when animals suppressed their activity to prevent punishment (Miyazaki et al., 2012). In our task, maintaining nose-poking for a delayed reward characterizes "waiting to obtain reward". When the expected water reward was suddenly omitted for several consecutive trials, the duration of nose-poking gradually shortened (Miyazaki et al., 2011b). This result suggests that rats maintained their nose-poking behavior at reward sites to receive rewards when they predicted that a reward was forthcoming. Rats are not likely to continue nose poking with a prediction of punishment that they could not get rewards when they remove their nose from reward sites before the reward presentation. Depletion of forebrain 5-HT induced via the intraventricular administration of the selective neurotoxin 5,7-dihydroxytryptamine (5,7-DHT) (Harrison et al., 1997) and systemic application of 8-OH-DPAT (Carli and Samanin, 2000) increased the premature response in the 5-CSRTT. In the 5-CSRTT, the requirement that the rat withhold nose-poke responses in one of the five apertures until an internal stimulus light is briefly illuminated represents "waiting to obtain reward" because the light stimulus presented in one of the five apertures works as a conditioned reinforcer. In a symmetrically rewarded go/no-go conditional visual discrimination task, global 5-HT depletion by 5,7-DHT impaired the ability of previously trained rats to subsequently inhibit their responses to no-go signals (Harrison et al., 1999). Our results further support the role of the central serotonin in "waiting to obtain reward" and behavioral withholding to acquire future rewards but not in avoiding punishment.

An interesting question is whether the artificial activation of DRN serotonergic neurons causes rats to be more tolerant with respect to long delayed rewards. The effectiveness of 5-HT<sub>1A</sub> receptor antagonists introduced to the DRN by reverse dialysis in activating serotonergic neural activity is debated (Adell and Artigas, 1998; Hajós et al., 2001), and the electrical stimulation of the DRN has been shown to activate nonserotonergic neurons. Optogenetic methods (Zhang et al., 2009; Depuy et al., 2011) may enable us to control serotonergic neuronal activity with precise timing and assess its effects on behaviors involved in the expectation of rewards and punishments.

## References

- Adell A, Artigas F (1998) A microdialysis study of the in vivo release of 5-HT in the median raphe nucleus of the rat. *Br J Pharmacol* 125:1361–1367.
- Adell A, Carceller A, Artigas F (1993) In vivo brain dialysis study of the somatodendritic release of serotonin in the raphe nuclei of the rat: effects of 8-hydroxy-2-(Di-n-Propylamino)tetralin. *J Neurochem* 60:1673–1681.
- Artigas F, Romero L, de Montigny C, Blier P (1996) Acceleration of the effect of selected antidepressant drugs in major depression by 5-HT<sub>1A</sub> antagonists. *Trends Neurosci* 19:378–383.
- Assié MB, Bardin L, Auclair AL, Carilla-Durand E, Depoortère R, Koek W,

- Kleven MS, Colpaert F, Vacher B, Newman-Tancredi A (2010) F15599, a highly selective post-synaptic 5-HT<sub>1A</sub> receptor agonist: in-vivo profile in behavioral models of antidepressant and serotonergic activity. *Int J Neuropsychopharmacol* 13:1285–1298.
- Bari A, Theobald DE, Caprioli D, Mar AC, Aidoo-Micah A, Dalley JW, Robbins TW (2010) Serotonin modulates sensitivity to reward and negative feedback in a probabilistic reversal learning task in rats. *Neuropsychopharmacology* 35:1290–1301.
- Bizot JC, Thiébot MH, Le Bihan C, Soubrié P, Simon P (1988) Effects of imipramine-like drugs and serotonin uptake blockers on delay of reward in rats. Possible implication in the behavioral mechanism of action of antidepressants. *J Pharmacol Exp Ther* 246:1144–1151.
- Bizot J, Le Bihan C, Puech AJ, Hamon M, Thiébot M (1999) Serotonin and tolerance to delay of reward in rats. *Psychopharmacology (Berl)* 146:400–412.
- Boureau YL, Dayan P (2011) Opponency revisited: competition and cooperation between dopamine and serotonin. *Neuropsychopharmacology* 36:74–97.
- Cardinal RN (2006) Neural systems implicated in delayed and probabilistic reinforcement. *Neural Netw* 19:1277–1301.
- Carli M, Samanin R (2000) The 5-HT<sub>1A</sub> receptor agonist 8-OH-DPAT reduces rats' accuracy of attentional performance and enhances impulsive responding in a five-choice serial reaction time task: role of presynaptic 5-HT<sub>1A</sub> receptors. *Psychopharmacology* 149:259–268.
- Casanovas JM, Berton O, Celada P, Artigas F (2000) In vivo actions of the selective 5-HT<sub>1A</sub> receptor agonist BAY x 3702 on serotonergic firing and release. *Naunyn-Schmiedeberg's Arch Pharmacol* 362:248–254.
- Cools R, Nakamura K, Daw ND (2011) Serotonin and dopamine: unifying affective, motivational, and decision functions. *Neuropsychopharmacology* 36:98–113.
- Daw ND, Kakade S, Dayan P (2002) Opponent interactions between serotonin and dopamine. *Neural Netw* 15:603–616.
- Dayan P, Huys QJ (2009) Serotonin in affective control. *Annu Rev Neurosci* 32:95–126.
- Depuy SD, Kanbar R, Coates MB, Stornetta RL, Guyenet PG (2011) Control of breathing by raphe obscurus serotonergic neurons in mice. *J Neurosci* 31:1981–1990.
- Doya K (2002) Metalearning and neuromodulation. *Neural Netw* 15:495–506.
- Evenden JL (1999) Varieties of impulsivity. *Psychopharmacology (Berl)* 146:348–361.
- Evenden JL, Ryan CN (1996) The pharmacology of impulsive behaviour in rats: the effects of drugs on response choice with varying delays of reinforcement. *Psychopharmacology (Berl)* 128:161–170.
- Hajós M, Gartside SE, Sharp T (1995) Inhibition of median and dorsal raphe neurons following administration of the selective serotonin reuptake inhibitor paroxetine. *Naunyn-Schmiedeberg's Arch Pharmacol* 351:624–629.
- Hajós M, Hoffmann WE, Tetko IV, Hyland B, Sharp T, Villa AEP, (2001) Different tonic regulation of neuronal activity in the rat dorsal raphe and medial prefrontal cortex via 5-HT<sub>1A</sub> receptors. *Neurosci Lett* 304:129–132.
- Harrison AA, Everitt BJ, Robbins TW (1997) Central 5-HT depletion enhances impulsive responding without affecting the accuracy of attentional performance: interactions with dopaminergic mechanisms. *Psychopharmacology* 133:329–342.
- Harrison AA, Everitt BJ, Robbins TW (1999) Central serotonin depletion impairs both the acquisition and performance of a symmetrically reinforced go/no-go conditional visual discrimination. *Behav Brain Res* 100:99–112.
- Invernizzi R, Belli S, Samanin R (1992) Citalopram's ability to increase the extracellular concentrations of serotonin in the dorsal raphe prevents the drug's effect in the frontal cortex. *Brain Res* 584:322–324.
- Miyazaki KW, Miyazaki K, Doya K (2011a) Activation of central serotonergic system in response to delayed but not omitted rewards. *Eur J Neurosci* 33:153–160.
- Miyazaki K, Miyazaki KW, Doya K (2011b) Activation of dorsal raphe serotonin neurons underlies waiting for delayed rewards. *J Neurosci* 31:469–479.
- Miyazaki K, Miyazaki KW, Doya K (2012) The role of serotonin in the regulation of impulsivity and patience. *Mol Neurobiol* 45:213–224.
- Newman-Tancredi A, Martel JC, Assié MB, Buritova J, Laressergues E, Cusi C, Heusler P, Bruins Slot L, Colpaert FC, Vacher B, Cussac D (2009) Signal transduction and functional selectivity of F15599, a preferential post-synaptic 5-HT<sub>1A</sub> receptor agonist. *Br J Pharmacol* 156:338–353.
- Paxinos G, Watson C (1998) *The rat brain in stereotaxic coordinates*, Ed 4. San Diego: Academic.
- Sharp T, Boothman L, Raley J, Quéré P (2007) Important messages in the 'post': recent discoveries in 5-HT neurone feedback control. *Trends Pharmacol Sci* 28:629–636.
- Soubrié P (1986) Reconciling the role of central serotonin neurons in human and animal behavior. *Behav Brain Sci* 9:319–364.
- Waterhouse BD, Devilbiss D, Seiple S, Markowitz R (2004) Sensorimotor-related discharge of simultaneously recorded, single neurons in the dorsal raphe nucleus of the awake, unrestrained rat. *Brain Res* 1000:183–191.
- Zhang F, Gradinaru V, Adamantidis AR, Durand R, Airan RD, de Lecea L, Deisseroth K (2009) Optogenetic interrogation of neural circuits: technology for probing mammalian brain structures. *Nat Protocol* 5:439–456.