

平成 21 年 6 月 12 日現在

研究種目：若手研究 (B)

研究期間：2006～2008

課題番号：18700383

研究課題名 (和文) 強化学習仮説に基づく淡蒼球の役割について明らかにする研究

研究課題名 (英文) Analysis of basal ganglia based on the reinforcement learning theory

研究代表者

上田 康雅 (UEDA YASUMASA)

京都府立医科大学・医学研究科・助教

研究者番号：60332954

研究成果の概要：

大脳基底核線条体の投射細胞には、ドーパミン D1 受容体を持ち淡蒼球内節へ投射する直接路とドーパミン D2 受容体を持ち淡蒼球外節から視床下核を経て淡蒼球内節へ投射する間接路の異なる二つの経路が存在する。今回これらのドーパミン受容体をそれぞれの拮抗薬で阻害したときの動物の試行錯誤学習課題における影響を調べたところ、D1 受容体拮抗薬の注入によって報酬の履歴を反映した適切な行動選択が障害された。一方、D2 受容体の拮抗薬では D1 受容体拮抗薬の注入によって見られたような障害は現れなかった。このことから、淡蒼球内節には報酬履歴に基づいた行動選択に必要な情報が送られていることが示唆された。

交付額

(金額単位：円)

	直接経費	間接経費	合計
2006 年度	1,800,000	0	1,800,000
2007 年度	1,000,000	0	1,000,000
2008 年度	700,000	210,000	910,000
年度			
年度			
総計	3,500,000	210,000	3,710,000

研究分野：総合領域

科研費の分科・細目：神経科学 神経・筋肉生理学

キーワード：大脳基底核、強化学習、行動選択、線条体、淡蒼球、ドーパミン

1. 研究開始当初の背景

大脳基底核は、大脳皮質の広範な領域から線条体を介して入力を受け、その出力は淡蒼球内節から視床を介して再び大脳皮質へ再び入力するという、いわゆる大脳皮質基底核ループを形成している。大脳基底核は、パーキンソン病に代表される臨床的な知見から、運動制御に重要な役割を担う部位であると長い間考えら

れてきた。その一方で、近年線条体へ投射する黒質緻密部のニューロンを含めた中脳ドーパミンニューロンが、報酬に対する予測誤差(報酬予測誤差情報)をコードするということが報告されると、この報酬予測誤差情報を用いた大脳基底核の強化学習モデルが(Sutton, R. and Barto, A. G. 1998)提唱された。我々の研究室で

はこの仮説に基づいて大脳基底核の機能を明らかにすべく、平成14年度からATRの銅谷・鮫島のグループと共同で研究を行ってきた。その結果、サルに複数の選択肢の中から試行錯誤によって報酬確率

(価値)の高い選択肢を選ばせるという行動課題を行わせると、随意的な行動選択直前の待機期間中に、多くの被験および尾状核ニューロンが、選択肢の報酬確率に依存して放電頻度を増減させることが明らかになった。このことは、線条体のニューロンが報酬価値そのものを表現するのではなく、選択肢特異的な報酬確率、つまり行動価値を表現することを示している(Samejima et al., 2005)。次の大きな問題は、線条体ニューロンに表現されたこれら行動価値が、強化学習仮説に基づいてどのように行動選択に結びつくのかという疑問である。この線条体からの情報は投射細胞によって淡蒼球外節と淡蒼球内節に送られている。これらの経路は、線条体から淡蒼球内節を経て視床へ投射する直接路と、線条体から淡蒼球外節、視床下核を経て淡蒼球内節へ投射し視床へ投射する間接路に分かれている。直接路に含まれる線条体の細胞には、ドーパミンD1受容体が存在し、間接路に含まれる線条体の細胞にはドーパミンD2受容体が存在することが示唆されている。

2. 研究の目的

本研究の目的は、大脳基底核の強化学習モデルの生理学的検討と、検証である。強化学習仮説に基づいたモデルの研究から銅谷は、線条体が行動価値をコードし、黒質網様部と淡蒼球において報酬予測の競合が起こる結果、行動選択が行われるというモデルを提唱している(Doya 2000)。淡蒼球外節と内節では線条対の投射様式

が異なることから、これらの異なる経路の強化学習仮説に基づく役割の違いを明らかにすれば、それぞれの淡蒼球特異的な役割について有用な知見となると考えられる。

3. 研究の方法

サルに右倒しか左倒しかを、自由に選択させる試行錯誤学習課題(図1参照)を遂行させる。ただし、左右の行動選択後に、大きな報酬を得られる確率がそれぞれ異なっている。たとえば(50-10)ブロックでは、左倒しを選択した場合、その後大きな報酬をもらえる確率は50%で、逆に右倒しを選択すると、10%の確率でしか大きな報酬が得られない。ブロックの最初のトライアルでは、どちらの行動を選択すれば、より高確率で大きな報酬が得られるかはわからないが、試行錯誤を繰り返すうちに、左右どちらを選択すればより高確率で大きな報酬が得られるのかがわかり、その方向を繰り返し選択するようになる。

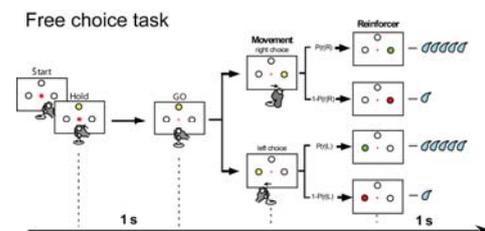


図1

このような課題を遂行中の動物の線条体に、ドーパミンD1受容体拮抗薬(SCH23390)、およびD2受容体拮抗薬(Eticlopride hydrochloride)を注入し、線条体から淡蒼球内節へ投射する回路と線条体から淡蒼球外節へ投射する回路特異的な、ドーパミンを介した情報の流れを阻害する。それによって起こる行動への影響を調べることで、淡蒼球内節および外節の役割について考察する。

4. 研究成果

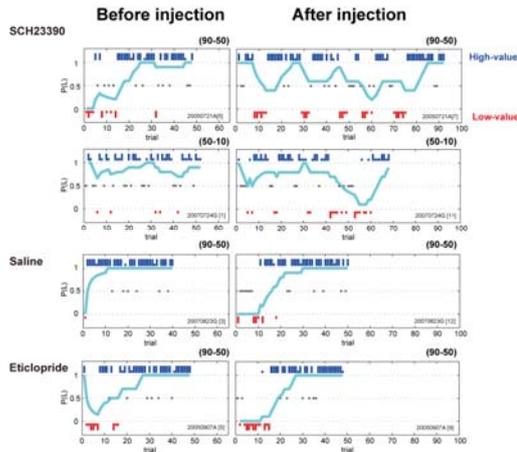


図 2

D1受容体拮抗薬の注入前、サルは報酬価値の高い選択枝を50トライアル以内に高い確率で選択するようになる。このことは試行錯誤を繰り返し、それぞれの行動価値を線条体において正しく表現できた結果、それらを利用した適切な行動選択ができていたと考えられる。D1受容体拮抗薬を注入すると、薬剤注入前に比べて価値の低い選択枝を選択する割合が増加する。しかし、このような障害は生理食塩水の注入、およびD2受容体拮抗薬の注入では観察されなかった（図2参照）。このD1受容体拮抗薬の影響を、サルが選択した行動とその結果得られる報酬の履歴から考察した。報酬が連続してもらえた場合では、D1受容体の拮抗薬の注入前および注入後においても同じ様にブロック内の価値の高い行動を選択するが（図3左）、行動に伴う大報酬が連続して欠落した場合、D1受容体拮抗薬注入後ではブロック内の価値の高い行動を選択する確率が注入前のその選択確率に比べて有意に低下した（図3右）。注入前には、連続して大報酬が欠落しても、長期の大報酬履歴を参考に行動選択を行うので、3回の大報酬欠落でも、行動選択はブロック内の

(50 - 10)

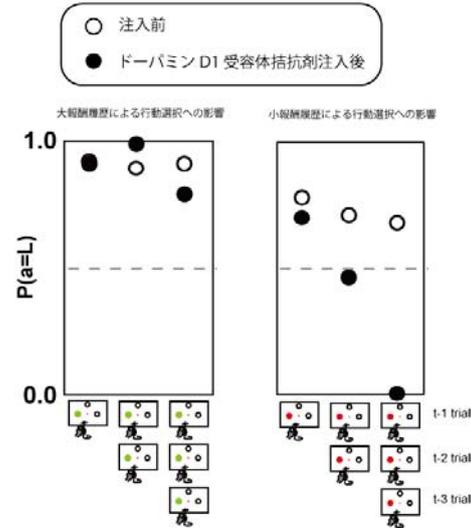


図 3

価値の高い方に偏っていると考えられる。この現象について強化学習仮説に基づいて考察してみる。適切な行動選択のためのそれぞれの行動価値は、強化学習仮説によれば、動物が過去に選択した行動とそれによって得られた報酬予測誤差情報によって行動毎に更新される（パネル1参照）。今回見られた障害は、おそら

左および右のそれぞれの行動価値の、報酬予測誤差を使った更新は、以下のように、左の行動価値は左を選択した時に起こり、右の行動価値は右を選択した時にそれぞれ起こる。

$$Qa(i) = Qa(i-1) + \alpha(i-1) [r(i-1) - Qa(i-1)] + \epsilon_0 \quad \text{if } a = a(i-1),$$

$$Qa(i) = Qa(i-1) + \epsilon_0, \quad \text{if } a \neq a(i-1)$$

$r(i)$: i 番目の試行で得られた報酬

—: 報酬予測誤差情報

ϵ_0 : 正規分布を持つノイズ

α : 学習率

左および右それぞれの行動価値を使用して、以下の式から i 試行目のトライアルで、左を選択する行動確率を計算する。

$$P_L(i) = \frac{1}{1 + \exp[-\beta \{Q_L(i) - Q_R(i)\}]}$$

パネル 1

くこのモデルにおいて、D1受容体を阻害することで、ドーパミンを介した報酬予測誤差情報を線条体内の行動価値を表現する細胞に適切に伝えられなくなったことにより、履歴に基づいた行動価値が正しく形成されなかったことによると考えられる。履歴に基づいた適切な行動価値が障害された動物は、より直近の報酬履

歴に基づいて行動選択をせざるを得なくなっていると考えられる。このことから、線条体から淡蒼球へ向かう直接路は、報酬履歴に基づく行動価値の形成と行動選択に重要な情報を送っていることが示唆された。一方間接路の情報を担うと考えられる経路の役割に関しては、今回の解析では大きな障害を観察することができなかった。今後は、直接路を介した行動価値の情報と間接路によって送られてくる情報が統合される淡蒼球内節の機能について、さらに詳細な解析が必要となる。このためには、この領域において同一課題を使用した神経活動の記録および解析が必要になるが、現在2頭の動物のトレーニングが終了し、複数細胞の同時記録システムも完成したので、今後このテーマを継続する予定である。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[学会発表] (計3件)

① Y. Ueda(発表者), K. Yamanaka, K. Samejima, K. Doya and M. Kimura
Adaptation of reward-based free-choices is impaired by D1 but not D2 dopamine receptor antagonist in the striatum
472.8/NN29 Society for Neuroscience 2008
Washington DC Mon. Nov. 17

② Y. Ueda(発表者), K. Samejima, K. Doya and M. Kimura
Selective impairment of reward-based adaptive choice of actions by intra-striatal injection of dopamine D1

receptor antagonist

P1-h09 第30回日本神経科学学会、2007年9月10日、横浜.

③ Y. Ueda(発表者), K. Samejima, K. Doya and M. Kimura
Representation of value of action, action and its outcome in sub-populations of striate neurons
Neurosci. Res. Suppl. 55:S230.PS3A-I151 第29回日本神経科学学会、2006年7月21日、東京.

[図書] (計1件)

① M. Kimura and Y. Ueda
Sensorimotor Learning and the Basal Ganglia. In: Encyclopedia of Neuroscience, Eds. Binder MD, Hirokawa N, Windhorst U. Springer(2008). 総ページ数 4399頁

6. 研究組織

(1) 研究代表者

上田 康雅 (UEDA YASUMASA)

京都府立医科大学・医学研究科・助教

研究者番号：60332954