

機関番号：32639
 研究種目：基盤研究（A）
 研究期間：2008～2010
 課題番号：20240040
 研究課題名（和文） 行動選択の価値のアップデートと記憶におけるドーパミン系と眼窩前頭皮質の役割
 研究課題名（英文） Involvement of dopamine system and orbito frontal cortex in update and memory of value signals
 研究代表者
 木村 實（KIMURA MINORU）
 玉川大学・脳科学研究所・教授
 研究者番号：40118451

研究成果の概要（和文）：ドーパミン細胞が長期的な報酬予測を表現するかどうかを調べた。3択によって3回の報酬を得る課題を日本ザルに行わせた。予測的な舌の運動から、動物は各試行での報酬価値ではなく、長期的報酬（強化学習の価値関数）を予測していることが判明した。ドーパミン細胞は、動物の行動と同様に長期的報酬予測を表現することが分かった。将来のゴールに向けて長期的な予測と誤差を表現し、線条体などの標的部位での価値のアップデートや意志決定に必須の役割を担うと考えられる。

研究成果の概要（英文）：We examined whether the midbrain dopamine neurons encode signals of long-term reward prediction and its error. Three Japanese monkeys performed multi-step choices for 3 rewards. Anticipatory licking during the task performance showed that monkey's expectation did not reflect reward probability of individual trials but precisely reflected reward value (probability x volume) of expected total rewards with future rewards discounted, value function of reinforcement learning. Remarkably, dopamine neuron activity encoded the long-term reward prediction rather than reward value of individual trials. The results suggested that the dopamine signals of long-term reward prediction may play critical roles in update of value signals for directing future goals in the target structures such as the striatum.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2008年度	12,300,000	3,690,000	15,990,000
2009年度	11,300,000	3,390,000	14,690,000
2010年度	15,000,000	4,500,000	19,500,000
年度			
年度			
総計	38,600,000	11,580,000	50,180,000

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：ドーパミンニューロン、線条体、価値判断、報酬予測、報酬予測誤差、試行錯誤、目標・ゴール、教科学習

1. 研究開始当初の背景

中脳ドーパミン細胞は大脳基底核線条体と大脳辺縁系、大脳前頭葉皮質に投射し、それぞれの脳部位でのシグナル伝達の可塑性調節をしていること、黒質緻密部のドーパミン細胞の脱落によってパーキンソン病症状を呈することが知られている。近年、種々の外界の事象に関連するドーパミン細胞の放電頻度が、その事象の持つ報酬価値（得られる確率、量）や報酬予測の誤差 (Schultz et al., 1997; Satoh とKimura ら 2003; Bayer et al., 2005)、動機づけレベル (Satoh とKimura ら 2003; Redgrave et al., 1998) を表現することがSchultz らの研究と木村らの研究によって明らかになった。このことによって、大脳皮質からの運動・認知情報と中脳ドーパミン細胞から報酬予測誤差情報を受け取る線条体の情報処理様式が強化学習のアルゴリズムと矛盾しないことが判明し、従来から謎であった大脳基底核の機能理解が大きく進んだ。木村のグループは、複数の選択肢の中から試行錯誤をしながら自分の最も望ましい選択肢の行動を選択する際に、線条体の投射細胞が過去の行動選択とその結果得た報酬（有無や大小）に基づいて、各々の選択肢の報酬価値の大小を反映するように放電頻度を変えることを明らかにし (Samejima とKimura らScience, 2005)、大脳基底核の強化学習モデル (Doya 2000; Houk and Barto, 1996; O' Doherty et al., 2004) を実験的に支持する明確な証拠を提供した。

試行錯誤によって一連の選択行動をおこなひ、最終的に一定量の報酬を獲得する際に、

一連の複数の予測される行動の報酬価値を表現するという可能性である。このことによって、現時点では最も望ましい選択肢の選択が不可能であっても将来にわたってトータルの報酬を最大にする選択スケジュールをたてることによって、目標に到達する高級な行動計画が可能になる。木村らはドーパミン系がこの機能に重要な役割を担うであろうと考えている。また、外界の事象の持つ報酬価値を表現する神経細胞が線条体だけでなく、大脳皮質の広い部位で確認されている。

2. 研究の目的

このような背景において、大脳基底核の機能、ドーパミン細胞の役割を明らかにする上で、ドーパミン細胞や線条体細胞に報酬の価値判断の情報を伝達している脳部位はどこか、大脳基底核と大脳皮質による選択肢の価値のアップデート、長期的な保存のためのしくみを明らかにすることが大変重要になっている。それは、「ドーパミン細胞や線条体の細胞が担う報酬価値の情報は、強化学習仮説が示すとおり、試行錯誤によって感覚事象や行動の選択肢のもつ報酬価値をアップデートする上で特に重要であり、学習が完成したあとではドーパミン細胞は活動せず、長期的な記憶として大脳皮質に保存される。一旦特定の事象での報酬価値が確定すると、価値判断が求められる状況では大脳皮質の保存されている情報が線条体やドーパミン細胞に伝達されて使用される」という作業仮説が導かれる。海馬がエピソードの学習と一時的な記憶の座であり長期的な記憶は大脳皮質に貯蔵されるという知見 (Ji, Wilson, 2007; Siapas, Lubenov,

Wilson, 2005; Morris et al., 2007) と同様なメカニズムで、望ましいアクションの選択に関する学習が脳基底核—ドーパミン系によってなされ、選択肢の価値情報は脳皮質、とりわけ眼窩前頭皮質に保存されるという仮説で、勿論まだ検証されていない。そこで、大脳眼窩前頭皮質、ドーパミン細胞と線条体に焦点を絞り、木村らが最近の数年間で調べてきた複数の選択肢の中から報酬価値の高い選択肢を試行錯誤によって選択するルールをサルに学習・記憶させ、将来選択する複数ステップの報酬価値をドーパミン細胞、大脳眼窩前頭皮質と線条体の細胞がどのように表現するかを調べた上で、これらの脳部位をGABA_A受容体アゴニストであるMuscimol によって不活性化させて、ドーパミン細胞の神経活動がどのように変化するかを明らかにする。同時にサルの行動選択行動にどのような変化が生じるかを詳細に調べる。

3. 研究の方法

木村らの作業仮説（上記）を検証するために、20年度には以下の研究計画を推進する。

- (1)サルを実験動物に、複数回の試行錯誤による価値判断と行動選択によって報酬に繋がる選択肢を選択する課題を学習させる。この選択課題は木村らが従来から用いている課題を基に、どのように被験者が複数ステップ先までの行動選択での報酬予測をおこなっているかを確認することができるように工夫する。
- (2)この課題を行っているサルの中脳と眼窩前頭皮質に微小電極を刺入し、単一ドーパミン細胞と皮質細胞の神経放電を記録し、複数ステップ先までの行動選択での報酬予測と予測誤差を表現するか、2つの脳部位の表現する神経情報の類似点と相違点は何かを、強化学習のアルゴリズムでの計算論的な推定を含めて詳細に調べる。

4. 研究成果

木村のグループは、従来の研究成果に基づいて、「2. 研究の目的」欄に記載の作業仮説を検証するための、実験と理論による共同研究を行った。研究課題の提案時点では、眼窩前頭皮質の神経活動の遮断による行動課題遂行とドーパミンニューロンの放電活動への影響を調べることを検討していたが、ドーパミンニューロンの新しい活動特性を見出したので、この点に焦点を当てて精力的に研究を推進し、成果を得た。

ドーパミン細胞が長期的な報酬の予測とその誤差の情報を担うかどうかを、3 択の行動選択課題を実行中の3頭のニホンザルのドーパミン細胞の活動を記録することによって調べた。3 択課題では、第一選択 (N1) での平均正解 (報酬を得る) 確率は 33%、第二選択 (N2) では 50%、そして第三選択 (N3) では 85-90%であった。正解試行の場合には、次の1回または2回の試行でも同じ選択 (R1 または R2) をすれば正解で再び報酬を得ることができる。当然ながら正答率はほぼ 100%であった。したがって、試行錯誤期に1回、続く繰り返し期に2回、合計3回の報酬を得ることができる。まず、サルが各試行で、その試行での報酬期待 (確率) を予想しているのか、それとも、長期的な報酬予測をしているかどうかを調べた。つまり、試行錯誤期には全3回分の報酬の中の1回目の報酬確率と、第2、第3回分の報酬の和に基づいて期待し、繰り返し期に2回目の報酬を得ると (R1) 残りの試行 (R2) の報酬確率がほぼ 100%であるものの残りは1回分、つまり総報酬の3分の1しか残っていないので、R1 試行での期待値よりも低くなってしまふ。図1には試行タイプごとに、報酬を期待する水舐め行動の時間を棒グラフで示すが、サルは各試行の報酬を期待している (緑の丸) のではないことが判

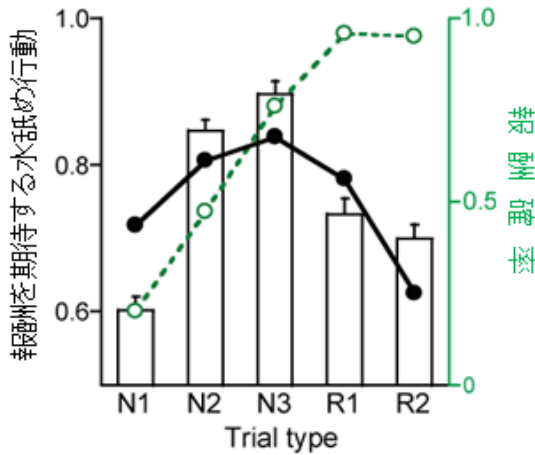


図1. 長期的な報酬予測に基づく水舐め行動

明した。そこで、長期的報酬予測を試行錯誤期から繰り返し期までの各試行での3回の報酬を得るまでの総報酬（確率 \times 量、式 (1)）として求めた（図1、黒の丸）。強化学習の価値関数で用いられるように、将来の報酬は割り引いて計算している。図1に示す通り、サルは各試行での報酬でなく、長期的報酬予測を行っていることが明らかになった。

$$V(s_t) = E\left\{\sum_{k=0}^{\infty} \gamma^k r_{t+k+1} \mid s_t = s\right\} \quad (1)$$

r_i : future rewards
 γ : discount factor (割引定数)
 k : an index for future steps

そこで、ドーパミン細胞 93 個の放電を微小電極によって記録して、各試行での課題スタート刺激に対する応答の強さを調べた。サルが試行錯誤で1回の報酬、繰り返し期に2回の報酬を得られることを学習する前には、ドーパミン応答は各試行によってほとんど変化がなかった（図2、学習初期）。しかし、課題を十分学習すると、繰り返し期の R1、R2 試行での応答が著しく減少し、強化学習の価値関数で定量的に良く推定できることが分かった（図2、習熟期）。念のために、長期的な予測が不可能な古典的条件付け課題での条件刺激に対するドーパミン細胞の応答

を調べた。すると、条件付けに用いられる視覚条件刺激のもつ報酬確率、すなわち目先の報酬価値（確率）を良く反映する活動を示した。この結果は、ドーパミン細胞は長期的な報酬予測が可能な状況では、将来の報酬を割り引いた上で、ゴール（この実験では3回の報酬を得る）に到達するまでの各段階での長期的報酬予測を正確に反映することが明らかになった（Enomoto et al., 論文投稿・改訂中）。ドーパミン細胞の信号は、線条体を中心に脳皮質前頭葉、脳辺縁系などの価値判断や意志決定に重要な役割を担うと考えられている脳部位に伝達され、その領域でのシナプス伝達の可塑性を介する調節に必須の役割を担うことが知られているので、報酬や懲罰に基づいて将来期待されるゴールへの到達に向けて、意志決定や行動選択に必須の役割を担うと考えられる。

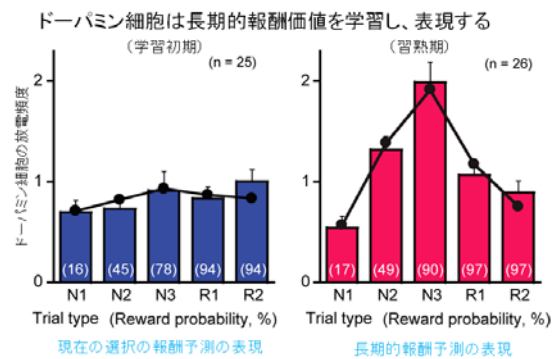


図2. 中脳ドーパミン細胞による長期的報酬予測の表現

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)

[雑誌論文] (計 5 件)

- ①Muranishi M, Inokawa H, Yamada H, Ueda Y, Matsumoto N, Nakagawa M, Kimura M. Inactivation of the putamen selectively impairs reward history-based action selection. *Exp Brain Res*. 209(2):235-246 (2011) 査読有
<http://www.springerlink.com/content/9206161690133t47/>
- ② Inokawa H, Yamada H, Matsumoto N, Muranishi M, Kimura M. Juxtacellular labeling of tonically active neurons and phasically active neurons in the rat striatum. *Neuroscience* 168(2):395-404 (2010) 査読有
<http://www.sciencedirect.com/science/article/pii/S0306452210004975>
- ③Tsytarev V, Fukuyama H, Pope D, Pumbo E. and Kimura M. Optical imaging of interaural time difference representation in rat auditory cortex. *Front. Neuroeng.* 2:2(2009) 査読有
<http://www.frontiersin.org/neuroengineering/10.3389/neuro.16.002.2009/abstract>
- ④Minamimoto T, Hori Y. and Kimura M. Roles of the thalamic CM-PF complex-Basal ganglia circuit in externally driven rebias of action. *Brain Res. Bull.* 78: 75-79(2009) 査読有
<http://www.sciencedirect.com/science/article/pii/S0361923008002906>
- ⑤Hori Y, Minamimoto T, Kimura M. Neuronal encoding of reward value and direction

of actions in the primate putamen. *J Neurophysiol* 102: 3530-3543(2009) 査読有

<http://jn.physiology.org/content/102/6/3530.full>

[学会発表] (計 12 件)

- ①Kimura M, Enomoto K, and Yamanaka K. Value signals for decision and action selection in the basal ganglia. *Janelia Conference Neural Circuits of Decision-Making* 2011年3月6日 Janelia Farm Research Campus, VA, USA
- ②Enomoto K, Matsumoto N, Haruno M, Kimura M. Differential encoding of reward value and its prediction error by dopamine neurons located in ventromedial and dorsolateral parts of the midbrain. The 40th annual meeting of Society for Neuroscience. 2010年11月17日 San Diego, USA
- ③Haruno M, Frith C, Sakagami M, Kimura M. Top-down signal from the frontal cortex regulates social value orientation. 第33回日本神経科学大会 2010年9月4日 神戸国際会議場, 兵庫県
- ④Enomoto K, Matsumoto N, Haruno M, Kimura M. Dopamine neurons in ventromedial and dorsolateral part of the midbrain differentially encode reward value and its error. 33回日本神経科学大会 2010年9月3日 神戸国際会議場, 兵庫県
- ⑤木村實 ご褒美と罰によって私達の行動を司る脳のしくみ. 包括的脳科学研究・教育センター設立記念シンポジウム 2010年1月16日 東北大学, 宮城県
- ⑥ Kimura M. Neuronal basis of action valuation and selection in the basal

ganglia. International Symposium : New Perspectives on Neural Mechanisms of Cognition and Action. 2009年11月13日 玉川大学, 東京都

- ⑦村西学, 井之川仁, 山田洋, 木村實. 被殻の機能脱落は状態価値ではなく行動価値を障害する. 第32回日本神経科学大会 2009年9月17日 名古屋国際会議場, 愛知県
- ⑧山中航, 堀由紀子, 上田康雅, 木村實. サルの視床CM核細胞の報酬予告と行動に対する応答. 第32回日本神経科学大会 2009年9月16日 名古屋国際会議場, 愛知県
- ⑨榎本一紀, 松本直幸, 木村實. ドーパミン細胞は複数のステップで得られる総報酬量に基づく報酬価値とその誤差を表現する. 第32回日本神経科学学会大会 2009年9月16日 名古屋国際会議場, 愛知県
- ⑩Kimura M, Yamada H, Enomoto K, Matsumoto N, Hori Y, Ueda Y. and Yamanaka K. Involvement of the basal ganglia in valuation and selection of actions. The 36th Congress of the International Union of Physiological Sciences (IUPS2009) 2009年7月29日 国立京都国際会館, 京都府
- ⑪Ueda Y, Samejima K, Yamanaka K, Doya K, Kimura M. Adaptation of reward-based free-choices is impaired by D1 but not D2 dopamine receptor antagonist in the striatum. The 38th annual meeting of Society for Neuroscience 2008年11月16日 Washington DC, USA
- ⑫村西学, 井之川仁, 山田洋, 木村實. 被殻へのムシモル注入による報酬価値に基づく行動選択の障害. 第31回日本神経科学大会 2008年7月10日 東京国際フォーラム,

東京都

[図書] (計2件)

- ① Kimura M. and Ueda Y. Sensorimotor Learning and the Basal Ganglia In: Encyclopedia of Neuroscience. Springer 3650-3653(2009)
- ② Doya K. and Kimura M. The basal ganglia and the encoding of value. In: Neuroeconomics: Decision Making and the Brain. Academic Press 407-416(2008)

6. 研究組織

(1) 研究代表者

木村 實 (KIMURA MINORU)
玉川大学・脳科学研究所・教授
研究者番号 : 4 0 1 1 8 4 5 1

(2) 研究分担者

松本 直幸 (MATSUMOTO NAOYUKI)
京都府立医科大学・医学研究科・講師
研究者番号 : 0 0 2 5 2 7 2 6