

食塩嗜好性変化に関わる大脳基底核神経回路の解析と再構成

疋田 貴俊¹, 内田 裕輝², 山下 祐一²

¹ 大阪大学蛋白質研究所, ² 国立精神・神経医療研究センター

概要 ヒトを含めて動物は、低濃度の食塩を含む食物を好み、高濃度の食塩を含む食物を嫌う。しかしながら、極端な減塩を行うと高濃度の食塩を好むようになることが知られている。一方、食塩の過剰摂取が継続した場合に、どのような食嗜好の変化が起きるかは定まっていない。これらの食塩の濃度依存的な嗜好性と、体内環境による嗜好性変化についての脳内基盤は明らかになっていない。

大脳基底核は運動のバランスの制御部位であるのみならず、好ましいものを求める報酬行動、不快なものを避ける忌避行動、報酬か忌避かを選択する意思決定を支配する必須の脳部位である。大脳基底核の主要な神経回路である線条体および側坐核から黒質への投射は、直接路と間接路の2つの伝達経路からなる。直接路は報酬行動及び薬物依存を支配する伝達経路であり、一方間接路は忌避行動及び行動柔軟性を支配する伝達経路である。また直接路と間接路のスイッチングがドーパミンシグナル伝達による可塑性変化によることが示されている。

食塩の濃度依存的な嗜好性の脳内基盤を調べることを目的として、前年度までに、マウスに対して、水、低濃度の食塩水、高濃度の食塩水をランダムに提示する行動課題を確立し、食塩の濃度依存的な嗜好性は体内環境によって変化することを示した。このことは体内環境を一定に保とうとする恒常性機構として捉えられる。

そこで、本年度は、恒常性維持プロセスの計算モデルを用いた解析を行い、食塩の濃度依存的な嗜好性の理論的再構成を目的とした。

まず恒常性維持プロセスの計算モデルとして、ホメオスタシス強化学習を使用し、マウスの食塩嗜好性の計算理論として、ホメオスタシス強化学習が妥当であることを示した。次に、口腔感覚を内部状態変化の予測として扱うことで、胃への直接の食塩水注入が脱塩マウスの報酬にはならない、という食塩嗜好性の説明が可能であることを示した。さらに、多次元内部状態のホメオスタシス強化学習による二瓶選択実験の再現を行なった。

今後の課題として、恒常性維持プロセスの計算モデルを用いた解析を進め、これまでに示した、食塩の濃度依存的な嗜好性に関する側坐核のドーパミン変動の理論的再構成を目指す。

1. 研究目的

ヒトを含めて動物は、低濃度の食塩を含む食物を好み、高濃度の食塩を含む食物を嫌う。しかしながら、極端な減塩を行うと高濃度の食塩を好むようになることが知られている。一方、食塩の過剰摂取が継続した場合に、どのような食嗜好の変化が起きるかは定まっていない。これらの食塩の濃度依存的な嗜好性と、体内環境による嗜好性変化についての脳内基盤は明らかになっていない。

大脳基底核は運動のバランスの制御部位であるのみならず、好ましいものを求める報酬行動、不快なものを避ける忌避行動、報酬か忌避かを選択する意思決定を支配する必須の脳部位である⁽¹⁾。大脳基底核の神経回路網は大脳皮質—線条体および側坐核—黒質網様部—視床の神経経路で形成され、さらに海馬体や扁桃体から記憶や情動の情報入力を受ける。大脳基底核の主要な神経回路である線条体および側坐核から黒質への投射は、直接路と間

接路の2つの伝達経路からなる。直接路は報酬行動及び薬物依存を支配する伝達経路であり、一方間接路は忌避行動及び行動柔軟性を支配する伝達経路である⁽¹⁾。また直接路と間接路のスイッチングがドーパミンシグナル伝達による可塑性変化によることが示されている⁽²⁾。そこで、食塩の濃度依存的な嗜好性の脳内基盤を調べることを目的として、前年度までにマウスに対して、水、低濃度の食塩水、高濃度の食塩水をランダムに提示する行動課題を確立し、食塩の濃度依存的な嗜好性は体内環境によって変化することを示した。このことは体内環境を一定に保とうとする恒常性機構として捉えられる。そこで本年度は、恒常性維持プロセスの計算モデルを用いた解析を行い⁽³⁾、食塩の濃度依存的な嗜好性の理論的再構成を目的とした。

2. 研究方法

2.1 ホメオスタシス強化学習(HRL)

恒常性維持プロセスの計算モデルとして、ホメオスタシス強化学習(HRL)を使用した⁽⁴⁾⁽⁵⁾。HRLにおいては、ホメオスタシスの偏りの最小化が、報酬の総和を最大化するとし、ホメオスタシスを強化学習プロセスとする。ドライブ関数 $D(H)$ を、入力変数を内部状態 H 、出力変数をドライブ D として(式1)に定義した。

$$D(H) = \sqrt[m]{\sum_{i=1}^N |H_i^* - H_{i,t}|^n} \quad (1)$$

内部状態 H は、例えば体温、血中グルコース濃度や水分状態を表す。各内部状態 H_i は、理想的な値(ホメオスタシスが最も安定した状態) H_i^* をもつ。ドライブ関数 $D(H)$ は、時点 t における各内部状態 $H_{i,t}$ と、理想的な内部状態 H_i^* との距離として定義される。すべての内部状態が理想的であるとき、ドライブは最小値をとることになる。(式1)において、 m, n はそうした距離を定義するためのハイパーパラメータである。

時点 t から $t+1$ にかけて、内部状態 $H_{i,t}$ が理想状態 H_i^* に近づけば、ドライブの値が減少し、離れればドライブの値が増加する。ドライブ関数を定義したのち、報酬 R が定義された。報酬 R は、時点 t のドライブと時点 $t+1$ のドライブの差として定義される。内部状態が何らかの栄養状

態であり、時点 t での行動選択が栄養摂取だった場合、内部状態は摂取量 K_t だけ増加する(式2)：

$$\begin{aligned} R(H_t, K_t) &= D(H_t) - D(H_{t+1}) \\ &= D(H_t) - D(H_t + K_t) \end{aligned} \quad (2)$$

後で述べるように、HRLの特徴である「実際の摂取量と、内部状態増加の予測としての味覚」を実装した場合、報酬 R は(式3)のように定義される：

$$R(H_t, \widehat{K}_t) = D(H_t) - D(H_t + \widehat{K}_t) \quad (3)$$

本研究では、内部状態が時定数 τ に依存して緩やかに減少することを想定した(式4)：

$$H_{t+1} = \left(1 - \frac{1}{\tau}\right) H_t + K_t \quad (4)$$

これは、実際の動物において、乾燥や代謝によって体内の水分状態やナトリウム状態(後述)が減少してゆくことに相当する。

選択可能なすべての行動に対して、行動価値 Q を定めた。時点 t において、行動価値 Q が大きな行動ほど選択されやすくなる。 Q は、各行動によって得られた報酬を元に、毎時点、更新がなされる(式5)：

$$Q_{t+1}(a) = Q_t(a) + \alpha_{t,q}(r_t - Q_t(a)) \quad (5)$$

(式5)において、報酬と、現在の時点の Q 値は、それぞれ実際に得られた報酬と、報酬の期待値であると考えることができる。報酬と現在の時点の Q 値の差は報酬予測誤差であり、これを元に Q 値の更新が行われると捉えることができる。なお、 α は1回の報酬予測誤差を Q 値の更新に反映する度合いを制御するパラメータ(学習率)を表す。

2.2 内部状態増加の予測としての口腔感覚

ホメオスタシスを考えるにあたって、動物が内部状態を監視することだけでなく、内部状態変化の予測と行動の関連学習を行うことが重要である。先行研究では、摂取によって得られた口腔感覚が、摂取によって内部状態がどれだけ増加するかを予測する値として設定されていた⁽⁴⁾。食塩摂取行動のアルゴリズムを理解するために、実際の摂取量と口腔感覚を区別することが重要であるか検証するため、さらに、将来的に味覚障害が栄養状態に与える影響について検討するための基盤を形成するため、マウス

の胃内注入実験⁽⁶⁾を、内部状態増加の予測としての口腔感覚を含む HRL (Fig. 2A)を用いて再現し、実際の行動実験の結果と一致するか検証した。計算モデルは 3 グループ用意した。ナトリウム不足 ($H_{t=0} = 0$) から開始し、通常の食塩水摂取が可能な control 群、内部状態が半分満ちた状態 ($H_{t=0} = H^*/2$) から計測を開始した NaCl intragastric infusion (IG) 群、どの行動をとっても常時塩味を感覚しており、ナトリウム不足状態 ($H_{t=0} = 0$) から計測を開始した NaCl oral 群である (Fig. 2B)。Fig. 1 と同様、エージェントがとりうる状態は一つ、行動は、何もしない行動、食塩水摂取行動の 2 通りを設定した (Fig. 1C)。

再現した動物実験データは、(Lee et al., 2019 Nature; Fig. 2C) を用いた⁽⁶⁾。これは 3 グループの比較から成る。ナトリウム の再吸収を阻害する利尿薬により、ナトリウム不足状態になった control 群、ナトリウム不足状態であり、計測前に胃内に直接食塩水を注入した NaCl IG 群、ナトリウム不足状態であり、口腔内で常に塩味を感覚している NaCl oral 群である。著者らは、これらマウスの食塩水摂取の回数 (licking 回数) を 10 分間計測した。動物実験の結果 (Lee et al., 2019; Fig. 2C) では、control 群と IG infusion 群の食塩水摂取回数に有意な差は無かった。一方で、Oral 群の結果は、Control 群よりも有意に低下していた。

今回、モデルと動物実験でそれぞれの摂取回数を比較し、似た傾向を示すか否かを検証することで、モデルの妥当性について考察した。報酬の定義式として、(式 6) を用いた:

$$R(H_t, \widehat{K}_t) = D(H_t) - D\left(\left(1 - \frac{1}{\tau}\right)H_t + \widehat{K}_t\right) \quad (6)$$

モデルの 1 時点を動物実験の 6 秒と考えた。これは、本研究において一貫している。

2.3 内多次元 HRL

従来の HRL 研究では、水分状態、体温など、1 種類の内部状態に着目してモデルが実装されてきた^{(4) (5)}。ところが、実際の生体の内部状態は多次元であり、HRL の理論を一般化するためには、複数の内部状態を考慮することが望まれている。そこで本研究では、HRL の内部状態を、水分状態とナトリウム状態の 2 通りに設定し、水と食塩水に対する嗜好性を検討したモデルを作成した。モデルの挙動が実際のマウスの振る舞いと一致するか否か確認する

ための動物実験として、二瓶選択実験 (Two-bottle preference test) を設定した。二瓶選択実験は、二つのボトルに別の溶液を用意し、ケージの中の動物がそれらを摂取した量を比較することで、二つの溶液に対する嗜好性を検討する動物実験である (Fig. 3B)。再現した動物実験データは、公開データ⁽⁷⁾を用いた。著者らは、蒸留水と食塩水の入ったボトルを二瓶として設定した。計算モデルの構築にあたって、エージェントが取りうる状態は 1 つ、行動は 3 つに設定した。 $a0$: 何もしない行動, $a1$: 水摂取, $a2$: 食塩水摂取である (Fig. 3A)。HRL は、4 通りの初期状態で挙動を比較された。水とナトリウム の内部状態が充足した状態 ($H_{w,t=0} = Hw^*$, $H_{s,t=0} = Hs^*$) から開始した control 群、水欠乏状態 ($H_{w,t=0} = 0$, $H_{s,t=0} = Hs^*$) から開始した water-depleted 群、ナトリウム欠乏状態 ($H_{w,t=0} = Hw^*$, $H_{s,t=0} = 0$) から開始した sodium-depleted 群、水とナトリウムが両方不足した状態 ($H_{w,t=0} = 0$, $H_{s,t=0} = 0$) から開始した water/sodium-depleted 群である。動物実験において、マウスは 4 つのグループに分けられていた。不足状態のない control 群、水のみ不足した water-depleted 群、ナトリウムのみ不足した sodium-depleted 群、水とナトリウムの不足した water/sodium-depleted 群である。動物実験の結果が (Fig. 3C-E) の左のグラフである⁽⁷⁾。水もナトリウムも充足した control 群では、水摂取と食塩水摂取の量はどちらもわずかだったが、水もナトリウムも欠乏した群では、水摂取と食塩水摂取の両方が有意に増加していた (Fig. 3C)。ナトリウムのみ不足していた群では、食塩水摂取のみが有意に上昇していた (Fig. 3D)。水のみ不足していた群では、水摂取が顕著に増加し、食塩水摂取も有意に増加していた (Fig. 3E)。

計算モデルと動物実験において、対応する群の、水と食塩水の摂取量をそれぞれ比較し、HRL が内部状態を 2 次元とする食塩嗜好性を考察するにあたって妥当か否かを、また、HRL が多次元内部状態を表現しうるか否かを検証した。

3. 研究結果

3.1 HRL はホメオスタシス様の挙動を示す

HRL を用いて、食塩嗜好性とホメオスタシス様の内部状態の変化を再現できるか否かを検討するために、まずマウスが食塩水の入ったボトルから食塩水を摂取する行動の再現を試みた (Fig. 1A, B)。想定する動物実験におい

て、時点 t で、マウスは二つの行動から一つの行動を選択することができる設定とした。選択肢は、食塩摂取行動と何もしない行動である (Fig. 1C)。マウスが塩分不足の際には食塩嗜好性を増加させ、ナトリウム充足状態では食塩嗜好性を低下させる、という挙動が、モデルでも見られるか否か検証した (Fig. 1B)。内部状態はナトリウム状態の 1 次元とし、実験開始時にマウスはナトリウム欠乏状態であると仮定した。

時点 $t = 0$ から $t = 80$ まで、ナトリウム内部状態が不足しているため、摂取行動が選択された際、大きな報酬が得られていた (Fig. 1D)。これにより、摂取行動の価値 (赤線) は急激に上昇し、大きな正の値を維持していた。この間、

摂取行動が選択される確率は 1 に近い値を継続していた。ナトリウム濃度が理想状態 ($H^* = 50$) を上回ると、摂取によって罰が得られることとなり、摂取の行動価値は急激に低下した。内部状態が時定数 τ によって減少することで内部状態が理想状態に近づくため、何もしない行動によって報酬が得られることとなり、 $t = 100$ 付近で、何もしない行動の行動価値 (黒線) が摂取の行動価値を上回った。これにより、何もしない行動が選択される頻度が上昇し、内部状態は時定数 τ に依存して緩やかに減少した。内部状態が理想状態を下回ると、再び摂取行動の価値が上昇し、内部状態が増加した。

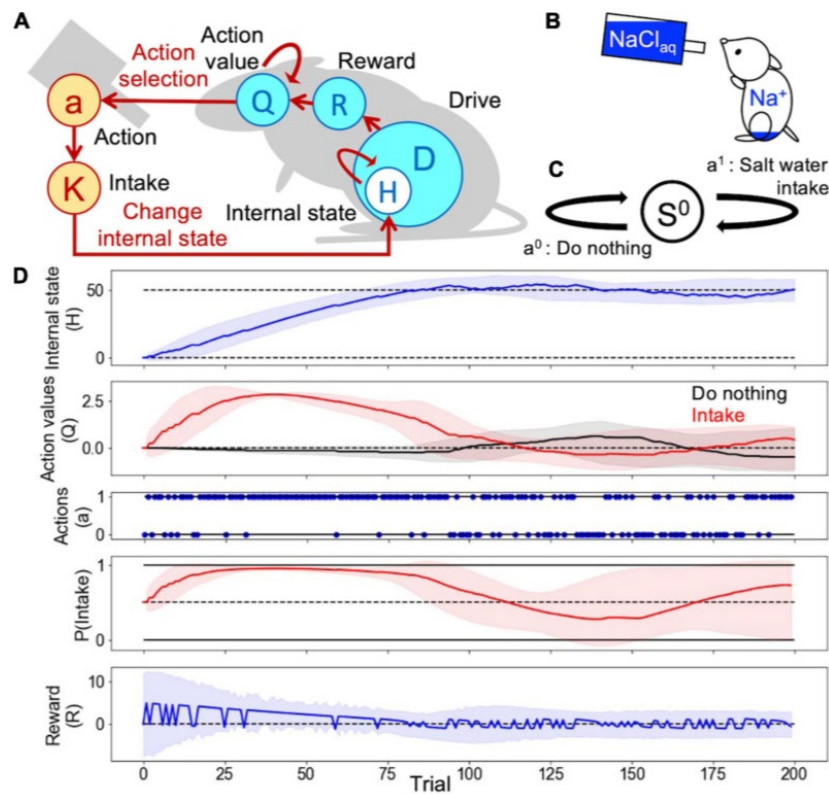


Fig. 1 Homeostatic behavior according to the homeostatic reinforcement learning (HRL) model.

(A) Schematic drawing of the computational process of the HRL model. (B) In an assumed animal behavior, sodium-depleted mice were able to lick saltwater. (C) Definition of a state and two actions in Simulations 1 and 2. (D) Example of homeostatic behavior. Changes in internal sodium state (H), the value of each action (Q-value), selected actions (a), probability of sodium intake [P(Intake)], and magnitude of reward (R) over time are plotted. The solid lines represent the results of a trial, and the light-colored error ranges represent the mean \pm 2 SD of 100 trials. The dotted line in the panel of the internal state indicates the ideal point ($H^* = 50$) of sodium taste. In the panel related to actions (a), action 1 represents “Intake behavior,” and action 0 indicates “do nothing.” At the beginning of the simulation, the internal sodium state and Q-values for each action were set to 0. After several random selections of action, the Q-value of sodium intake was increased, and the internal sodium state quickly reached the ideal point, maintaining homeostatic regulation of behavior.

3. 2 摂取の口腔感覚と実際の摂取量の分割

HRL の振る舞いにおいて、control 群と IG infusion 群に大きな差は見られなかった。一方、control 群よりも oral 群の結果は大幅に低下していた (Fig. 2D)。ナトリウム不足から開始した control 群では、食塩水摂取の行動価値が大きく上昇していた (Fig. 2E)。IG infusion 群は、ナトリウムが半ば充足 ($H_0 = H^*/2$) した状態を初期状態として、食塩水摂取の価値を大きく上昇させた (Fig. 2F)。Oral stimulation 群では、何もしない行動と食塩水摂取の両方の行動で口腔感覚 (内部状態増加の期待値) が上昇し、報酬が得られたため、何もしない行動と食塩水摂取の行動価値に差がつかなかった (Fig. 2G)。

3. 3 多次元内部状態の HRL による two-bottle preference test の再現

HRL の結果として、すべての群において、動物実験と同様の結果が得られた (Fig. 3C, D, E)。HRL の内部状態の推移、と行動価値の推移を示した (Fig. 3F-I)。水/ナトリウム不足群では、水摂取とナトリウム摂取の両方の行動価値が上昇していた (Fig. 3F)。ナトリウム不足群では、食塩水摂取の価値だけが上昇していた (Fig. 3G)。水不足群では、水摂取とナトリウム摂取の両方の価値が上昇していた (Fig. 3L)。不足状態のない control 群では、両方の摂取行動の価値が低下していた (Fig. 3I)。すべての群の内部状態は、最終的に安定していた (Fig. 3F-I)。

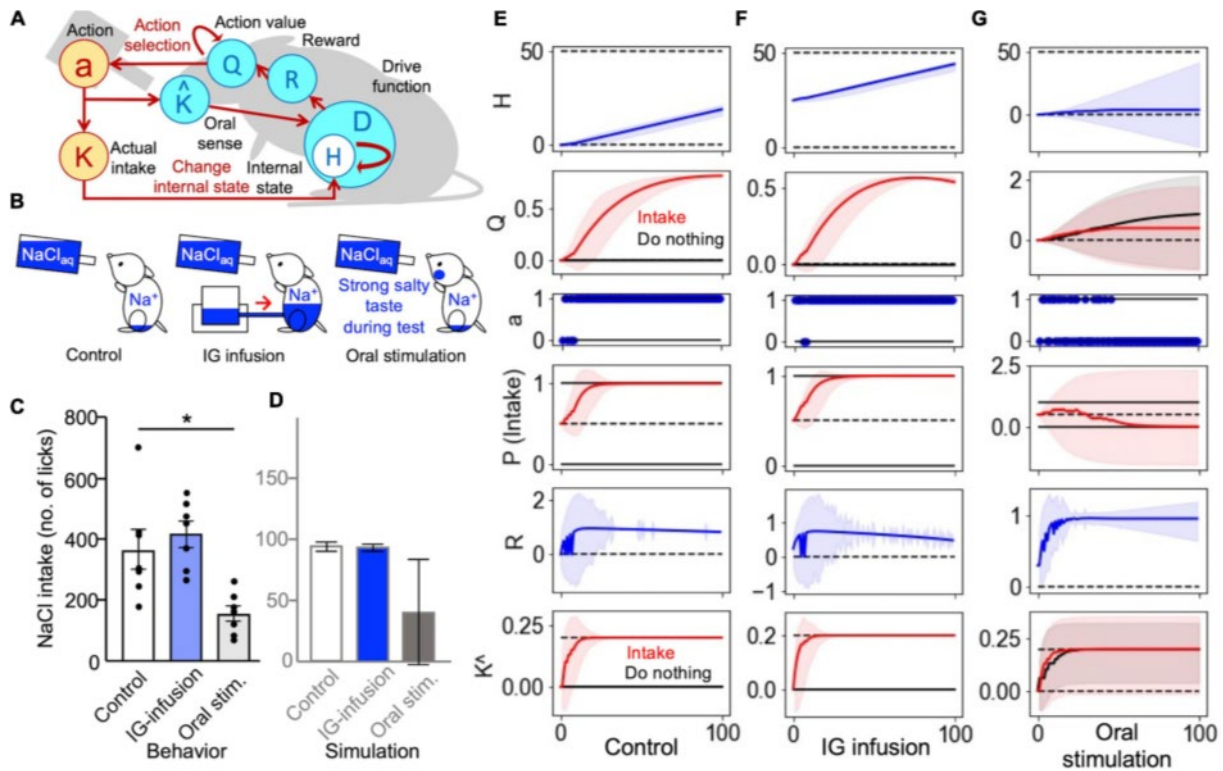


Fig. 2 Homeostatic behavior according to the homeostatic reinforcement learning (HRL) model.

(A) Schematic illustration of Simulations 2 and 3: Taste perception (K^{\wedge}) was a predictor of an increase in the internal state. (B) Three groups for the behavioral tests. (C) The results of the behavioral experiment (6). * $P < 0.05$. (D) The results of the computational model. Intra-gastric infusion did not change the level of sodium intake, while oral stimulation decreased intake. (E-G) Transitions of the simulation. (E) The control group exhibited an increased Q-value for intake. (F) A congenial increase in the Q-value was observed in the IG-infusion group. (G) The values of “do nothing” and “intake” were reinforced in the oral-stimulation group. (E-G) The solid lines represent the results of a trial, and the light-colored error ranges represent the mean \pm 2 SD of 100 simulated agents.

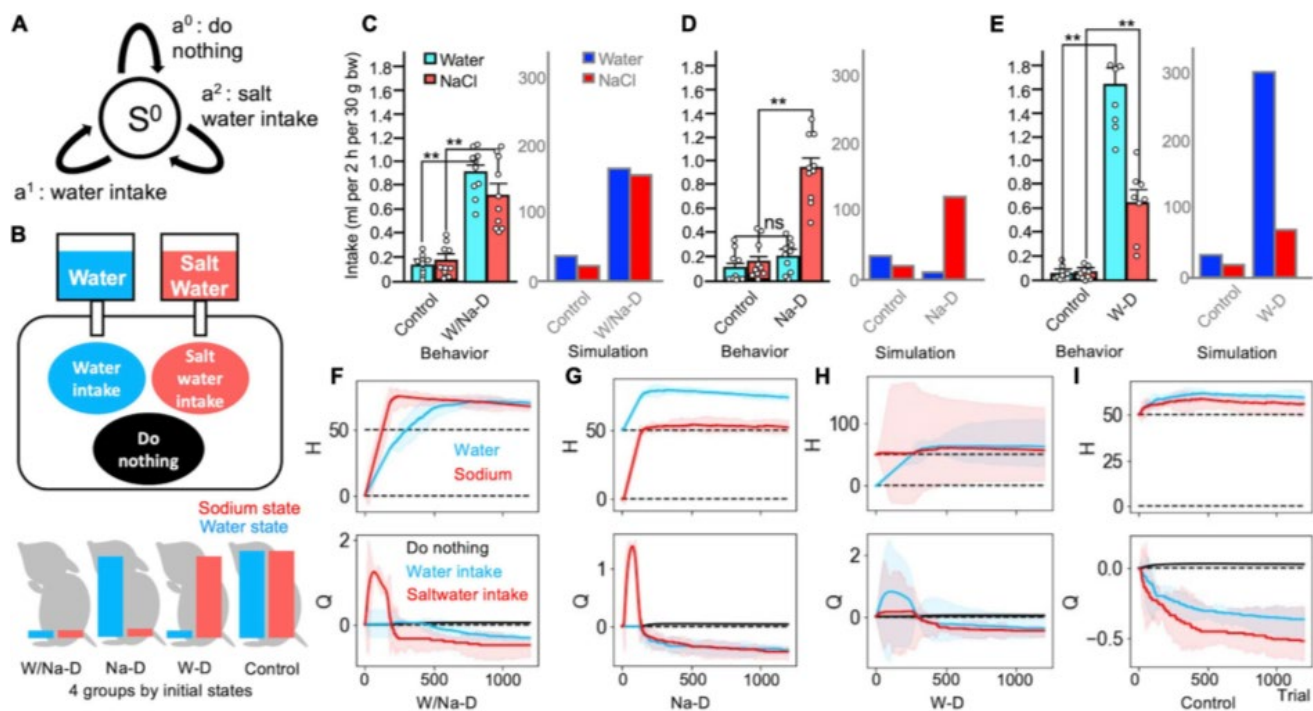


Fig. 3 Multi-dimensional homeostatic reinforcement learning (HRL) model as a suitable explanation of findings in the two-bottle preference test.

(A) Definition of a state and three actions in Simulations 3 and 4. (B) Schematic drawing of Simulation 3 (two-bottle preference test). (C) Comparison of behavioral data (7) and simulated data in the control (all-satisfied) and water/sodium-depleted groups. Control groups exhibited minor intake, and both depleted groups demonstrated copious volumes of intake in both sets of data. (C–E) Data averaged over 100 simulated agents. $**P < 0.01$. (D) In the sodium-depleted group, water intake was slight, while saltwater intake was increased. (E) The water-deficient groups exhibited abundant water intake and non-negligible saltwater intake. (F–I) Transitions of the HRL models. (F) The water/sodium-deficient HRL model exhibited strong increases in the values of water and saltwater intake. (G) The action value for sodium intake soared in the sodium-depleted model. (H) The water-depleted model exhibited increased values for water and saltwater intake. (I) The control model refused both intakes. (F–I) The solid lines represent the results of a trial, and the light-colored error ranges represent the mean ± 2 SD of 100 simulated agents.

4. 考察

本研究で検討した問題は、(i) マウスの食塩嗜好性の計算理論として、HRL が妥当であるか？(ii) 口腔感覚を内部状態変化の予測として扱うことで、食塩嗜好性の説明がされうるか？(iii) HRL が多次元内部状態に拡張可能であるか？というものであった。これらに対し、(Fig. 1)では、HRL がホメオスタシスの維持と相同な振る舞いをすることを示した。また、(Fig. 2)では、胃への直接の食塩水注入が脱塩マウスの報酬にはならないという性質を HRL が再現しうることを確かめた。さらに、(Fig. 3)では、HRL の振る舞いが、二瓶選択実験と定量的に類似することと共に、多次元 HRL を実装することが可能であることを示した。

5. 今後の課題

本研究により、食塩嗜好性のモチベーション、学習を介さない食塩嗜好性と報酬として強化される食塩嗜好性の差異、食塩嗜好性変化のタイムスケール、食塩嗜好性が

増加・減少する速度と加速度の制御機構、水分欲求・体液量の関与、神経回路操作の強度といった、食塩嗜好性が内包する興味深い機能が、定量的に説明・理解される可能性がある。こうした機能の中には、実験による観察を行うことが困難なものもあり、その説明に理論研究が貢献することが期待される。

これまでに、食塩の濃度依存的な嗜好性に側坐核のドーパミン変動と間接路の神経活動が関与していることを示した。今後、この神経ネットワーク機構を恒常性維持プロセスの計算モデルを用いた解析を進め、食塩の濃度依存的な嗜好性の理論的再構成を目指す。

6. 文献

- Hikida T. et al. Distinct roles of synaptic transmission in direct and indirect striatal pathways to reward and aversive behavior. *Neuron*, 66, 896-907 (2010).
- Hikida, T. et al. Pathway-specific modulation of nucleus

- accumbens in reward and aversive behavior via selective transmitter receptors. *Proc. Natl. Acad. Sci. U.S.A.*, 110, 342-347 (2013).
3. Uchida, Y., Hikida, T., Yamashita, Y. Computational mechanisms of osmoregulation: a reinforcement learning model for sodium appetite. *Front. Neurosci.* 16, 857009 (2022).
 4. Keramati, M. and Gutkin, B. Homeostatic reinforcement learning for integrating reward collection and physiological stability. *Elife*, 3, 1–26 (2014).
 5. Hulme, O. J., Morville, T., and Gutkin, B. Neurocomputational theories of homeostatic control. *Phys. Life Rev.*, 31, 214–232 (2019).
 6. Lee, S. et al., Chemosensory modulation of neural circuits for sodium appetite. *Nature*, 568, 93–97. (2019).
 7. Matsuda, T. et al., Distinct neural mechanisms for the control of thirst and salt appetite in the subfornical organ. *Nat. Neurosci.*, 20, 230–241 (2017)..

Analysis and Reconstruction of Basal Ganglia Circuit in Salt Preference

Takatoshi Hikida¹, Yuuki Uchida², Yuichi Yamashita²

¹Institute for Protein Research, Osaka University,

²National Institute of Neuroscience, National Center of Neurology and Psychiatry

Summary

Animals prefer low concentrations of salt and dislike high concentrations of salt. However, it is known that extreme salt reduction leads to a preference for high salt concentrations. The brain basis for these salt concentration-dependent changes in preference has not been clarified.

The basal ganglia control not only motor balance but also reward and aversive behavior, and decision-making. The projection from the striatum and the nucleus accumbens (NAc) to the substantia nigra consists of two main pathways, the direct and the indirect pathway, in the basal ganglia circuit. The direct pathway controls reward behavior, while the indirect pathway controls aversive behavior. In addition, it has been shown that switching between the direct and indirect pathways is due to plasticity induced by dopamine signaling. We have established the licking operant task for mice and showed that the concentration-dependent preference for salt was shown to depending on the internal environment. This can be regarded as a homeostatic mechanism that tries to maintain a constant internal environment. Therefore, this year, we conducted a mathematical model analysis using homeostatic reinforcement learning to theoretically reconstruct the concentration-dependent preference for salt.

First, we used homeostatic reinforcement learning, and showed that homeostatic reinforcement learning is a valid computational theory for salt craving behavior in mice. Next, by treating oral sensation as a predictor of internal state change, we showed that the direct injection of saline solution into the stomach is not a reward for desalinized mice, which can explain the salt craving behavior. In addition, we reproduced the two-bottle choice experiment by using homeostatic reinforcement learning of multidimensional internal states.

As a future work, we will continue the analysis using mathematical models based on homeostatic reinforcement learning and aim to theoretically reconstruct the dopamine fluctuations in the nucleus accumbens involved in the salt concentration-dependent preference.