

※注意※

本研究に関する論文は、プレプリントでピアレビュー（査読）前です。

PRESS RELEASE (2021/09/27)

人工知能であらゆる疾患の治療薬を見つける方法を開発 — 新型コロナウイルス（含デルタ株）治療薬候補の発見に成功 —

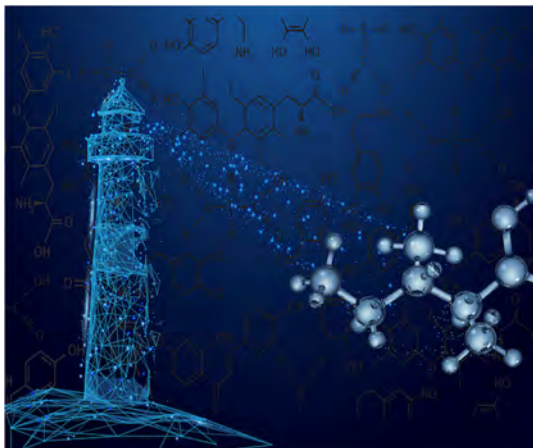
九州大学生体防御医学研究所の中山 敬一 主幹教授、米国ハーバードメディカルスクール・システム生物学部門の清水秀幸リサーチフェロー、北海道大学人獣共通感染症国際共同研究所の澤洋文 教授の研究グループは、疾患の原因となるタンパク質のアミノ酸配列（※1）のみから、そのタンパク質を狙った治療薬を見つけ出す方法を開発しました。

生体のさまざまな現象は遺伝子の設計図をもとに作られたタンパク質が担っています。疾患の原因となるタンパク質の立体構造が分かれば、そのタンパク質に対する治療薬を探すこともできますが、現在でもタンパク質の多くはその構造が分かっておらず、そのことが創薬における一つの壁になっていました。

本研究グループはこの問題を解決するため、タンパク質の立体構造を全く使わずに、より容易に入手できるタンパク質のアミノ酸配列のみから治療薬候補を高速に見つけ出す人工知能、LIGHTHOUSE（“灯台”の意）を開発しました。がんや感染症、生活習慣病といったさまざまな疾患の治療薬をLIGHTHOUSEに予測させ、その予測を実験で検証したところ、新たな抗がん剤や抗菌薬を見つけることができました。さらに新型コロナウイルスについても、国内で深刻な感染状況を引き起こしているデルタ株を含め多くの変異株の治療に有望な化合物を見出しました。

今回の成果は、これまでの創薬研究の進め方を大きく変える可能性を秘めているものであり、より迅速に薬を開発することができるようになると期待されます。

本研究成果は、2021年9月27日（月）（日本時間）にbioRxivで公開されました。なお、研究の詳細および用語解説は別紙を参照。



（イメージ図）

自然界には膨大な数の化合物が存在しますが、LIGHTHOUSEは灯台のように目的に合う化合物を「照らし出す」ことで創薬を大幅に加速させます。

研究者からひとつこと：大量の化合物候補の中から治療薬を照らし出してくれるLIGHTHOUSEは、まさに「灯台」のように、創薬開発の進むべき方向を示してくれる大きな指針となります。創薬にかかる時間やコスト、労力を大きく減らし、患者さんに薬という形で成果を還元できる日が1日でも早く来るよう、今後も研究を続けます。



中山主幹教授

【お問い合わせ】 生体防御医学研究所 主幹教授 中山 敬一（なかやま けいいち）
TEL : 092-642-6815 FAX : 092-642-6819
Mail : nakayak1@bioreg.kyushu-u.ac.jp

人工知能であらゆる疾患の治療薬を見つける方法を開発 — 新型コロナウイルス（含デルタ株）治療薬候補の発見に成功 —

<研究の背景と経緯>

医学研究の成果を患者さんに還元できるようになるまで時間がかかる大きな原因の一つは、有望な治療標的タンパク質に対する薬をつくるのが現在の技術でも簡単ではないことがあげられます。分子量500以下の小分子化合物（※2）に限定しても、 10^{60} もの化合物が存在すると言われており、その中から薬を見つけ出すのは時間・コスト・労力が大きくかかってしまいます。スーパーコンピューター（スパコン）を使ったドッキングシミュレーション（※3）など、コンピューターによる予測法も提案されていますが、たくさんの計算リソースが必要な上、ドッキングシミュレーションの前提となるタンパク質の立体構造はその多くが未知のままです。近年盛んに研究されている人工知能を使う創薬研究も発表されていましたが、それらのほとんどはコンピューターシミュレーションのみの解析であり、実際に新しい薬を見つけ出したわけではありませんでした。

そこで本研究グループは、さまざまな病気の治療薬を見つけ出すことができる汎用的な人工知能の開発と、実際の治療薬の発見を目的に研究を開始しました。

<研究の内容>

大規模な国際プロジェクトから得られた STITCH データ（100 万以上の化合物-タンパク質ペアのデータ）を人工知能の訓練に使いました。人工知能に学習させるには、化合物もタンパク質も、何らかの数値ベクトルに変換しなくてはなりません。化合物は、その結合様式も加味した上で、グラフとして表現した後に MPNN（※4）にて数値ベクトルに変換しました。タンパク質側は、アミノ酸配列をそれぞれ特徴が異なる3つの方法（CNN, AAC, Transformer, ※5）で数値ベクトルに変換しました。化合物とタンパク質の数値ベクトルを足し合わせ、さらに一連の演算をすることで、最終的にその化合物が「どれくらい薬らしいか」を表す数値が得られるような仕組みです。我々はこの手法を Lead Identification with Graph-ensemble network for arbitrary Targets by Harnessing Only Underlying primary SEquence、省略して LIGHTHOUSE（“灯台”の意）と命名しました。

次に、LIGHTHOUSE を使ってがんの悪性化に関わる酵素 PPAT（※6）と呼ばれるタンパク質を抑制する化合物を探索しました。PPAT をノックダウン（※7）するとさまざまながんの進行を食い止めることができることは知られていますが、PPAT の立体構造は未だ解明されておらず、PPAT の阻害剤もありません。そこで我々は、ZINC データセット（※8）に登録されている10億近い化合物を LIGHTHOUSE を用いて探索し、発見した最も有望な化合物を調べることで、世界で初めて PPAT 阻害活性がある化合物を見つけることに成功しました。PPAT はあらゆるがんの悪性化に関わっていることを考えると、この化合物は多くのがん患者さんに有効である可能性があります。

さらに、我々は新型コロナウイルス感染症（COVID-19）の治療に有望な化合物を LIGHTHOUSE で予測し、エトキシゾラミドというすでに緑内障治療薬や利尿薬（※9）などとして本邦で承認されている化合物を見出しました。ヒト培養細胞を用いた感染実験において、エトキシゾラミドは新型コロナウイルスの感染を抑え、元々の新型コロナウイルスだけでなくデルタ株を含めさまざまな変異ウイルスから細胞を保護する働きがあることを確かめました。

<今後の展開と治療応用への期待>

新型コロナウイルスに対しては、LIGHTHOUSE が予測したエトキシゾラミドをリード化合物（※10）とし、その派生化合物を探索することで、より効能の高い治療薬の開発につながると示唆されます。また LIGHTHOUSE はアミノ酸配列さえあれば治療薬候補を見つけることができるため、さまざまな魅力的な治療標的タンパク質に対する創薬を大きく加速させることが期待されます。

<参考図>

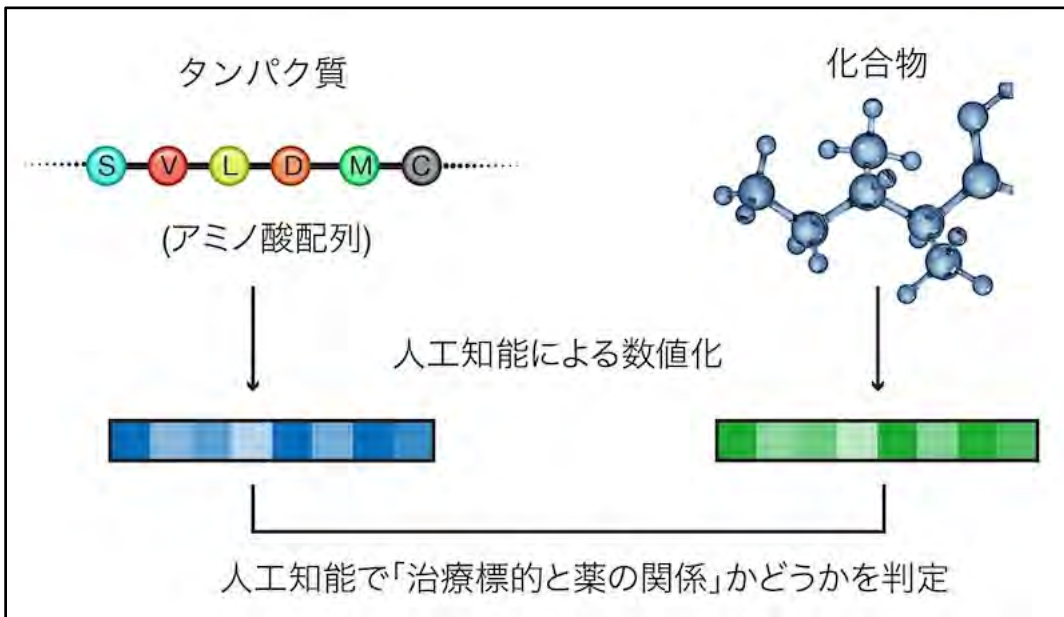


図1 LIGHTHOUSEの模式図

人工知能 (AI) である LIGHTHOUSE は、タンパク質のアミノ酸配列 (左) と化合物 (右) をそれぞれ数値の集まり (数値ベクトル) に変換した後、これらが標的タンパクと薬の関係にあるかをスコアで表現します。このスコアが高いほど、化合物が「薬」になる確率が高いと言えます。LIGHTHOUSE は従来の技術よりも非常に高速で、従来技術の数千倍のスピードで薬を探索することができます。

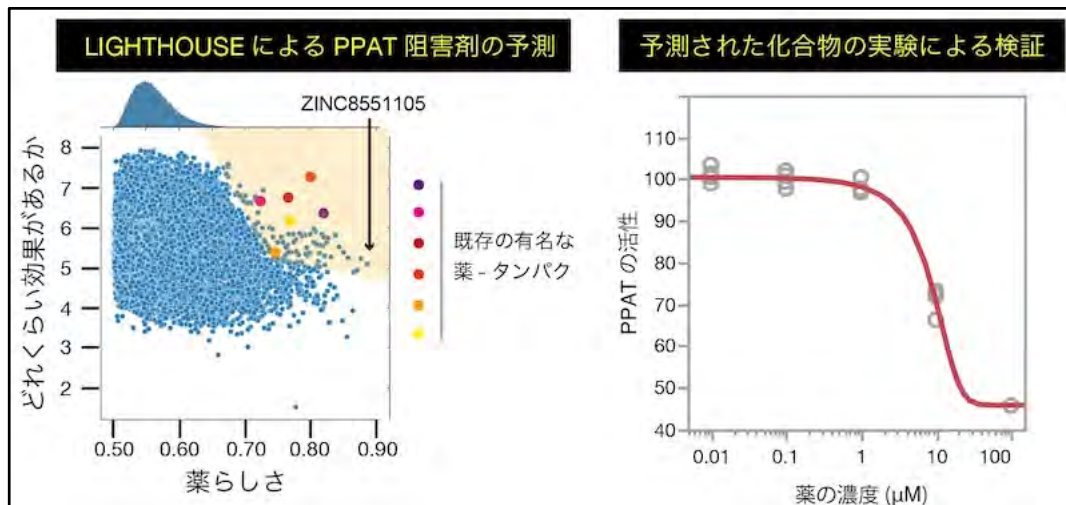


図2 LIGHTHOUSEによるPPAT阻害剤の予測と検証

さまざまながんの悪性化に強く関わっている PPAT を抑制する化合物を LIGHTHOUSE によって予測し、既存の他の薬-タンパクと同等以上のスコアを持つ化合物を候補として抽出しました (左図、網掛け)。そのトップヒットを実験的に検証したところ、確かに PPAT の抑制効果があることが実証されました (右図)。

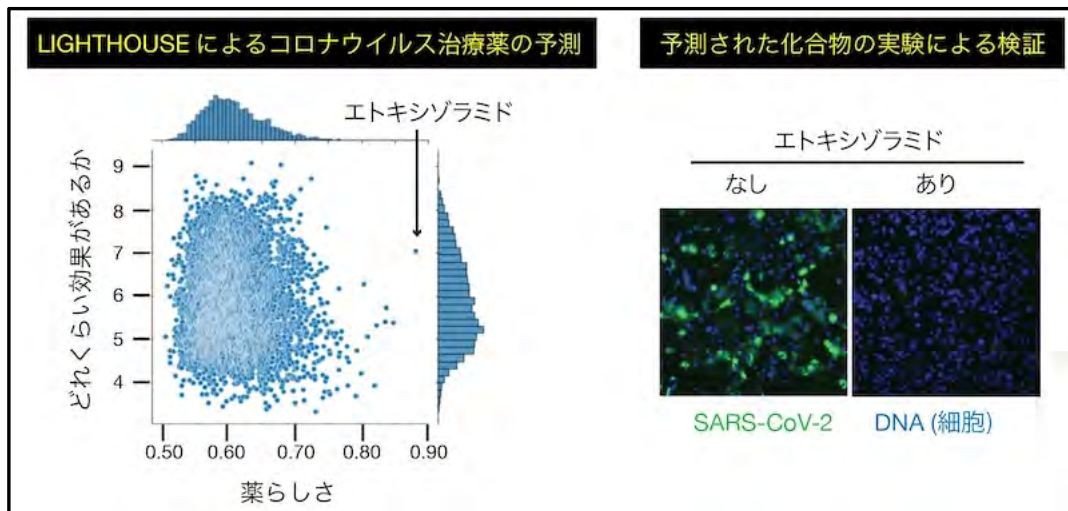


図3 LIGHTHOUSE によるコロナウイルス治療薬の発見

LIGHTHOUSE を使って、COVID-19 を引き起こす SARS-CoV-2 ウイルス治療薬候補を探索し、すでに緑内障などの治療薬として承認されているエトキシゾラミドを発見しました（左図）。この予測された化合物は、細胞にかけた SARS-CoV-2 の増殖を抑えることが実験で示されました（右図、緑色が SARS-CoV-2、青色が細胞）。

<用語解説>

- (※1) タンパク質のアミノ酸配列: タンパク質は、20種類のアミノ酸の組み合わせでできています。そのアミノ酸を端から順に並べたものを、タンパク質のアミノ酸配列といいます。これに対して、生体内ではタンパク質は複雑に折りたたまれており、その構造のことを立体構造といいます。アミノ酸配列に比べて立体構造は解明することがずっと難しく、多くのタンパク質の立体構造は未知です。
- (※2) 小分子化合物: 小さな分子量 (目安として 500 未満) の化合物を小分子化合物といい、現在流通している薬の多くはこの小分子化合物のカテゴリーに分類されます。
- (※3) ドッキングシミュレーション: タンパク質の立体構造をもとに、コンピューター上で仮想の化合物がそのタンパクに結合するかをシミュレーションすることです。
- (※4) MPNN: Message Passing Neural Networks の略で、ここでは化合物の元素をノード (頂点)、化学結合をエッジ (頂点間の線) とみなした図形 (グラフ) とし、そのグラフの情報を処理し統合しています。
- (※5) CNN, AAC, Transformer: タンパク質を数値ベクトルに変換するための手法で、それぞれ「近隣のアミノ酸」、「物理化学的性質」、「遠くのアミノ酸」の情報をうまく反映させた数値ベクトルが得られます。
- (※6) PPAT: phosphoribosyl pyrophosphate amidotransferase という名前の酵素で、細胞増殖時に必要な DNA の材料の合成に関わっています。PPAT 高発現 (たくさんある状態) だと、DNA もたくさん作られ、がん細胞の増殖にとって有利になります。
- (※7) ノックダウン: 目的の遺伝子の量を実験的に減らす技術です。ノックダウン法はとても有用な方法ですが、あくまでも研究目的であり、患者さんに使うことはできません。そのため、ノックダウンと同じ効果がある、PPAT 阻害剤が熱望されていましたが、立体構造が不明だということもあって、既存の方法ではこれまで開発できていませんでした。
- (※8) ZINC データセット: 全世界のさまざまな試薬メーカーが取り扱う、購入可能な化合物を集めたデータセットです。
- (※9) 緑内障治療薬や利尿薬: 緑内障は眼の中の圧力 (眼圧) が高くなって起こる病気です。利尿薬は腎臓に働いて尿の排出量を増やす薬です。
- (※10) リード化合物: 改良の余地があるもののそれなりに良い活性を持つ化合物のことです。創薬においては最初から素晴らしい化合物を見つけ出すことは稀で、ほとんどのケースではリード化合物をなんとか見つけ出し、それが見つかった後にリード化合物に関連した化合物を調べて、より効能の高い薬にしていくという手順をとります。

<論文情報>

タイトル: "LIGHTHOUSE illuminates therapeutics for a variety of diseases including COVID-19"
(LIGHTHOUSE は COVID-19 をはじめとする様々な疾患の治療薬を照らし出す)

掲載誌: bioRxiv

本成果は、以下の事業・研究領域・研究課題によって得られました。

科学研究費補助金・特別推進研究

研究課題名: 「幹細胞における細胞周期の制御と代謝系との連関に関する総合的研究」

研究代表者: 中山 敬一 (九州大学 生体防御医学研究所 主幹教授)

研究期間: 平成 30 年 4 月 ~ 令和 5 年 3 月

科学研究費補助金・若手研究

研究課題名: 「胃がん・大腸がんのユニバーサル予後予測法の開発とその治療応用」

研究代表者: 清水 秀幸 (九州大学 生体防御医学研究所 学術研究員)

研究期間: 平成 31 年 4 月 ~ 令和 3 年 3 月

<お問い合わせ>

【研究内容に関すること】

中山 敬一（なかやま けいいち）

九州大学 生体防御医学研究所 主幹教授

TEL : 092-642-6815 FAX : 092-642-6819 携帯電話 : 090-3608-4654 (24 時間対応可能)

Mail : nakayak1@bioreg.kyushu-u.ac.jp

【報道に関すること】

九州大学広報室

TEL : 092-802-2130 FAX : 092-802-2139

Mail : koho@jimui.kyushu-u.ac.jp