

//// Cutting Edge ////

機械学習とバーチャルスクリーニングを通じた SARS-CoV-2 RNA 依存性 RNA ポリメラーゼの阻害剤予測

株式会社 Elix Nazim Medzhidov、結城 伸哉

1. はじめに

重症急性呼吸器症候群コロナウイルス2 (SARS-CoV-2) の世界的な拡散により、COVID-19パンデミックが発生した。SARS-CoV-2は、一本鎖プラス鎖RNAゲノム [(+)ssRNA] を有し、ニドウイルス目コロナウイルス科ベータコロナウイルス属に属する[1]。COVID-19は、非定型肺炎を特徴とする致死性の呼吸器疾患である[2]。COVID-19治療薬は現時点で存在せず、計算機を使用した解析をはじめ、様々な手法でドラッグリポジショニングの取り組みが行われてきた[3-5]。本研究では、複数の機械学習モデルを使用し、それらを分子ドッキングアプローチと組み合わせ、ウイルス複製機構の重要な構成要素であるSARS-CoV-2 RNA依存性RNAポリメラーゼ (RdRp) に対し阻害活性を有しうる抗ウイルス薬および抗炎症薬のスクリーニングを行った。

2. コロナウイルスと SARS-Cov-2

2.1 概要

コロナウイルスは、様々な動物に感染する多様なウイルス群であり、ヒトでは軽度から重度の呼吸器感染症を引き起こす可能性がある。人獣共通感染源の高病原性コロナウイルスである重症急性呼吸器症候群コロナウイルス (SARS-CoV) と中東呼吸器症候群コロナウイルス (MERS-CoV) は、2002年と2012年にそれぞれヒトに感染し、致命的な呼吸器疾患を引き起こした。こうしてコロナウイルスは、21世紀の公衆衛生上の懸念事項として認識されるようになった[6]。2019年末、高い感染力を有するSARS-CoV-2が中国の武漢市で報道され、その後世界でもウイルス性肺炎の感染が広がりを見せた。COVID-19と呼ばれるこの病気は、感染者数と流行地域数の両方で過去に発生したSARSやMERSを上回っており、世界の公衆衛生に並々ならぬ脅威をもたらしている。

2.2 SARS-CoV-2の起源と拡散

2019年12月下旬、中国・武漢の複数の保健施設から原因不明の肺炎患者のクラスターが報告された[7]。患者の多くは、発熱、咳、胸部不快感などのウイルス性肺炎の症状を呈し、重症の場合呼吸困難や両側肺浸潤など、SARSやMERSの患者と同様の症状が見られた[7, 8]。患者の多くは、武漢市の中心部にある華南海鮮卸売市場と何かしらの関連を持っていたとされる[9, 10]。12月31日、武漢市衛生委員会は原因不明の肺炎発生を世界保健機関(WHO)に通知した[11]。

2.3 Genomic Characterization of SARS-CoV-2 / SARS-CoV-2 の遺伝的特徴

SARS-CoV-2 は SARS-CoV と比して 79%、MERS-CoV と比して 50% のゲノム配列同一性を有する[12]。SARS-CoV-2 は他のベータコロナウイルスと共通のゲノム構造を持つ。SARS-CoV-2 のゲノム構造は他のベータコロナウイルスと共通しており、レプリカーゼ (ORF1a/ORF1b)、スパイク (S)、エンベロープ (E)、メンブレン (M)、ヌクレオカプシド (N) の 6 つのオープンリーディングフレーム (ORF1a/ORF1b) が 5' から 3' の順に配列されている。また、遺伝子には、付属タンパク質をコードする 7 つの ORF が散在している[1]。SARS-CoV-2 がコードする多くのタンパク質の長さは、SARS-CoV の対応するタンパク質と近い。4 つの構造遺伝子 (上記 S、E、M、N) のうち、SARS-CoV-2 は S 遺伝子を除いて SARS-CoV と 90% 以上のアミノ酸同一性を共有している[12]。レプリカーゼ遺伝子は 5' ゲノムの 3 分の 2 をカバーし、大きなポリタンパク質 (pp1ab) をコードしており、ウイルス複製に関与する 16 の非構造タンパク質をコードする。SARS-CoV-2 非構造タンパク質のほとんどは、SARS-CoV の構成アミノ酸と 85% 以上同じアミノ酸で構成されている[1]。

2.4 SARS-CoV-2 RdRp を対象とした理由

RdRps は、RNA テンプレートから RNA 合成を触媒するマルチドメインタンパク質で、ウイルスのゲノム複製と転写プロセスを担う[13]。コロナウイルス属が属するニドウイルス目は、ウイルスの複製と転写を促進すべく ORF1a および ORF1b ウイルスポリプロテインの開裂産物として産生される非構造タンパク質 (NSP) によって作動する RNA 合成専用の複雑な機構を持つという特徴を有する[14]。

ウイルスの宿主細胞への侵入は、感染開始の重要な過程である。SARS-CoV-2 は、ウイルススパイク (S) タンパク質の受容体結合ドメインを利用して宿主細胞上のアンジオテンシン変換酵素 2 (ACE2) に結合し、エンドソームを介して細胞内に侵入する[15]。細胞内侵入の後、ウイルスゲノム RNA が放出され、ウイルスタンパク質の翻訳やウイルスゲノム複製のための鋳型として機能する。RNA 依存性 RNA ポリメラーゼ (RdRp) は、非構造タンパク質 12 (nsp12) としても知られており、新しいウイルスのゲノム RNA の産生を担うウイルスレプリカーゼ複合体の重要な構成要素である[16]。

ほとんどすべての RNA ウイルスは、そのゲノムにコードされた RdRp を有する。配列の類似性が比較的低いにもかかわらず、さまざまな RNA ウイルス由来の RdRp は、「右手モデル」として説明される 3 つの構成ドメイン(「親指ドメイン」、「手のひらドメイン」、「指ドメイン」)と同様の構造的類似性を有しており、これは逆転写酵素とも類似している。種々のウイルスから得られた RdRp の配列解析から、活性部位と手のひらドメインに重要な残基が保存されていることが明らかになっている[13, 17-19]。また、SARS-CoV-2 RdRp の構造解析の結果、RdRp の一般的な構造、保存された A-G モチーフの存在、主要な触媒アミノ酸の保存、C 型肝炎ウイルス (HCV) やポリオウイルス RdRp との構造的な類似性が確認された[20]。RdRp はウイルスの複製サイクルにおいて重要な役割を果たしつつ人体に対応するものが存在しないため、治療中に望ましくない副作用が生じるリスクを軽減できる

重要な治療標的と見なされている。

3. ドラッグリポジショニングにおける人工知能の利用

3.1 分子ドッキングと機械学習によるアプローチの比較

SARS-CoV-2 に有効な治療薬の開発が急務となっており、これまで多くの研究が行われてきた。最近では、SARS-CoV-2 の異なるタンパク質を標的とした様々なドラッグリパーパシングの候補を提案する *in silico* 創薬報告が数多く発表されている[5, 21-24]。これらの研究のほとんどは従来の分子ドッキング解析を利用しており、標的やリガンドの立体構造に関する情報を必要とする。この情報が得られれば、標的タンパク質上の指定された領域に対してドッキングシミュレーションを行い、タンパク質とリガンドの結合エネルギーを予測することができる。しかし、標的タンパク質の立体構造取得は困難で、費用も時間もかかる。さらに、結合エネルギーが良好であっても、リガンドによる阻害活性が高いとは限らない。

3.2 SARS-CoV-2 ドラッグリパーパシングにおける人工知能の利用

創薬における機械学習活用の技術開発は、従来の創薬プロセスを短縮し、新薬創出にかかるコストを削減できる可能性がある[25, 26]。最近では、立体構造に関する情報が得られない場合でも比較的短期間で潜在的な薬剤候補を見つけることができる代替手段として、機械学習を用いたアプローチがいくつか検討されている。

英国に拠点を置く BenevolentAI 社は、種々の科学文献から得られる情報を統合した AI 由来のナレッジグラフを多く活用している[4]。同グループは、宿主タンパク質 AAK1 の阻害を標的とし、関節リウマチの治療薬として承認されているバリシチニブを同定した。バリシチニブは、クラスリンを介したエンドサイトーシスを阻害することで、細胞のウイルス感染を抑制すると考えられている。バリシチニブが有する抗炎症性は、COVID-19 患者に多く見られるサイトカインレベルの上昇に有効である可能性が高い[27]。同様に、Beck らは、Molecule Transformer-Drug Target Interaction (MT-DTI) と呼ばれる深層学習の技術を活用した薬物-標的相互作用モデルを応用し、SARS-COV-2 関連のプロテアーゼとヘリカーゼを標的とする可能性を有する市販の抗ウイルス薬を予測した論文を発表している[3]。このモデルは分子を1次元の文字列表す SMILES (Simplified molecular-input line-entry system) とアミノ酸 (AA) 配列を使用しているため、実験で3次元結晶構造が確認されていない標的タンパク質を扱うことが可能である。また、米国に拠点を置く Atomwise 社は、新たな抗ウイルス剤の開発を目指し、複数のコロナウイルスに高度に共通して保存されている SARS-CoV-2 タンパク質結合部位をターゲットにした研究に注力している。具体的には、AtomNet の深層畳み込みニューラルネットワーク技術[28]を使い、数百万の仮想化合物をスクリーニングしているほか、学術研究者と15のパートナーシップを結んで、モデルで予測された化合物を *in vitro* アッセイで確認している[29]。

複数のウイルスの RdRp において構造上の類似性および活性部位での主要アミノ酸の保

存が認められ、また、広範な抗 RdRp 薬であるレムデシビルが同定されたことで、有効な RdRp 阻害薬の化学構造に類似パターンが存在する可能性が示された。このことから、潜在的な RdRp 阻害剤を同定するため、教師あり機械学習アルゴリズムを実装することとした。

4. Discovery of SARS-CoV-2 RdRp Inhibitors / SARS-CoV-2 RdRp 阻害剤の探索

4.1 データセット収集及びモデルの学習

まずは、HCV、デングウイルス、ポリオウイルス、インフルエンザウイルスの RdRp に対し、過去の実験で既に活性値が確認されている低分子を含むデータセットを構築した。データセットは PubChem [30]と ChEMBL [31]のバイオアッセイから取得した。分類モデルの学習を目的とし、既の実験で活性値 (IC₅₀/EC₅₀) が知られているエントリを選択し、活性値と閾値から二値の活性ラベル (活性あり・なし) を割り当てた。学習に使用する活性のカットオフ閾値は 5 μM とした。最終的に構築されたデータセットは、活性ラベル付きの化合物 1,356 (656 種類が不活性、700 種類が活性あり) が含まれている。バリデーションデータセットは、データセット全体から 20%の化合物 (活性化合物と不活性化合物) を無作為に選択したものを使用した。学習データセットは、残りの 80%の化合物を使用した。学習データセット中の分子は HCV、ポリオウイルス、デングウイルス、インフルエンザウイルスの RdRp に対して阻害活性を示した。これらの阻害剤を機械学習モデルに学習させることで、機械学習のモデルに有効な RdRp 阻害剤の化学的特徴の学習が可能になると考えられる。次に、当該モデルが既知の前臨床および臨床 RdRp 阻害剤を同定する能力を評価した。最後に、RdRp に対する阻害活性を潜在的に有する物質の候補を特定するため、FDA 承認の抗ウイルス薬および抗炎症薬を、当該モデルでスクリーニングした。また、SARS-CoV-2 の RdRp タンパク質に対する抗ウイルス薬や抗炎症薬の分子ドッキング解析も併せて実施した。今回の機械学習モデルでは、RdKit [32]を用いて化合物を分子フィンガープリントに変換したものを入力特徴量として使用した。フィンガープリントは、Morgan フィンガープリント[33]とトポロジカルフィンガープリントを使用しており、トポロジカルフィンガープリントに関しては、結合数が最小 1 から最大 7 までの化合物のすべてのサブグラフを抽出して計算した。モデルの実装は、主に scikit-learn ライブラリ[34]を使用した。

受信者操作時特性曲線下の面積 (以下、ROC-AUC、又は AUROC とする) 0.8 を超えたモデルは複数ある。具体的には、グラフ畳み込みネットワーク、メッセージパッシングネットワーク、ランダムフォレスト分類器 (円形フィンガープリント、トポロジカルフィンガープリント)、リッジ分類器 (円形フィンガープリント)、ラッソ分類器 (トポロジカルフィンガープリント)、3 層マルチレイヤーパーセプトロン (円形フィンガープリント)、XGBoost 分類器 (トポロジカルフィンガープリント) が、0.8 を超えていた。そのうちの 1 つである円形フィンガープリントのランダムフォレスト分類器は ROC-AUC 値が 0.9 を超えていた。精度 (ACC) は、ランダムフォレスト分類器が 84%と最も良い数値を示した。

4.2 前臨床 RdRp 阻害剤のテストセットを用いたモデル評価

結果検証には、ROC-AUC に基づく上位 3 つのモデルを使用し、既知の前臨床 RdRp 阻害剤のテストセットを対象とした予測を行った。ROC-AUC と精度に加え、真陽性 (TP)、真陰性 (TN)、偽陽性 (FP)、偽陰性 (FN) の割合も表 1.1 に示す。

モデル	AUROC	ACC	Confidence Interval (alpha=0.05)	TP	TN	FP	FN
GraphConv	0.700	0.700	[0.558, 0.842]	0.65	0.75	0.25	0.35
RandomForest (C)	0.725	0.725	[0.587, 0.863]	0.50	0.95	0.05	0.50
3-layer MLP (C)	0.625	0.625	[0.475, 0.775]	0.50	0.75	0.25	0.50

表 1.1 テストデータに対するモデルのパフォーマンス

ランダムフォレスト分類器は陰性例の検出に非常に優れていたが (真陰性率 95%)、陽性例について見てみると、検出可能な活性分子は半数に留まった (真陽性率 50%)。グラフ畳み込みモデル等の他のモデルは、より多くの活性分子を検出できるが (真陽性率は 65%)、真陰性率は 75%に低下し、より多くの偽陽性を検出してしまう。更に、異なるモデルの出力間の相関はあまり高くない。これらの事情を勘案し、アンサンブルモデルを使用することとした。

今回の実験では、RBF カーネルを用いた単純な Support Vector Machine が最も良く機能した。このモデルは、最も良い結果を示した上位 10 個のモデルの出力値を入力特徴量として使用している。最初に、元の学習セットを用いて個々のモデルの学習を行った。次に、バリデーションセットを 2 つの同じサイズのサブセットに分割した。そのうちの 1 つをアンサンブルモデルの学習に使用し、もう 1 つは検証セットに充てた。アンサンブルモデルでは、テストセットを用いた他のすべてのモデルの性能をわずかに上回った (表 1.2)。

Dataset	AUROC	ACC	Confidence Interval (alpha=0.05)	TP	TN	FP	FN
Validation	0.875	0.875	[0.819, 0.931]	0.871	0.879	0.129	0.121
Test	0.750	0.750	[0.616, 0.884]	0.600	0.900	0.100	0.400

表 1.2 アンサンブルモデルの結果

モデルの評価は、バリデーションセットと既知の前臨床 RdRp 阻害剤のテストセットのパフォーマンスに基づき実施した。最も良い性能を示した 3 つのモデルと (表 1.1 参照)、アンサンブルモデルの予測結果を分析した。モデルの学習はリガンドの情報のみを使用し、標的である SARS-CoV-2 RdRp の立体構造は使用していない。そのため、標的タンパク質とリガンドの立体構造情報に基づく従来の分子ドッキングアプローチと比較することが望ましいと考えられる。そこで、SARS-CoV-2 RdRp (PDB ID: 6m71) の活性部位に対する抗ウイルス・抗炎症データセットのバーチャルスクリーニングを AutoDock Vina を用いて行った[35]。

5. 機械学習とドッキングシミュレーションの結果

5.1 抗ウイルスデータセットの結果

解析の結果、抗ウイルス薬と抗炎症薬の両方のデータセットから標的候補を複数同定した。本章では具体的に紹介する。

我々のモデルは抗ウイルスデータセットからレムデシビルを同定した。レムデシビルは SARS-CoV-2 RdRp を標的とすることが確認されたヌクレオシドアナログであり、COVID-19 患者の治療薬として米国 FDA によって承認されている。レムデシビルは陽性対照としてテストセットに含まれていた。なお、上述のモデルで同定されたバロキサビル・マルボキシル、TMC-310911 (ASC09)、およびユミフェノビル (アルビドール) は、COVID-19 を対象とした臨床試験でも研究が進んでいる。臨床試験登録番号と識別番号は、バロキサビル・マルボキシルが ChiCTR2000029544、ASC09 が NCT04261907、ユミフェノビルが NCT04350684 である。バロキサビル・マルボキシルはインフルエンザウイルスの RdRp に作用し[36]、TMC-310911 は HIV-1 に対して開発されたプロテアーゼ阻害剤であり[37]、ユミフェノビルはヘマグルチニン (HA) 糖タンパク質を標的にしてウイルスの細胞内への侵入を抑制する抗インフルエンザ薬[38]である。

これらの薬剤候補に加え、最も良い性能を示した 4 つのモデル (表 1.1 の 3 つのモデル及びアンサンブルモデル、以下「最良モデル」とする) が、RdRp 阻害剤の可能性として、抗ウイルスデータセットからベクラブビルとアスナプレビルを同定した。ベクラブビルは、HCV RdRp (NS5B) の非ヌクレオシド系阻害剤である[39]。SARS-CoV-2 と同様、HCV も一本鎖エンベロープ型ポジティブセンス RNA ウイルスである。SARS-CoV-2 と HCV RdRp の活性部位は構造的類似性を示しており、両者とも触媒部位の保存アミノ酸を保有している[20]。ベクラブビルの SARS-CoV-2 RdRp に対する結合エネルギーは -9.2 kcal/mol となり、本候補の阻害性を示唆している。別の抗 HCV 薬であるアスナプレビルは、HCV のプロテアーゼを標的とすることが知られている。全ての最良モデルが潜在的な抗 RdRp 候補としてアスナプレビルを同定し、AutoDock Vina を用いた結合エネルギーは、-7.5 kcal/mol であった。他の候補としては、パリタプレビル、ファルダプレビル、シメプレビル、ベドロプレビル (HCV プロテアーゼ阻害剤)、レディパスビル、オダラスビル、およびベルパタスビル (HCV NS5A 阻害剤) が挙げられる (表 1.3 参照、4 つの最良モデルのうち少なくともいずれか 2 つで予測され、SARS-CoV-2 RdRp に対する結合エネルギーも比較的低いため)。我々のモデルは RdRp 阻害剤のみを対象としていたが、HCV のプロテアーゼまたは NS5A タンパク質を標的とするいくつかの抗 HCV 薬も RdRp 阻害剤として分類された。なお、これらの候補分子は、分子ドッキング解析に基づく SARS-CoV-2 の RdRp に対する結合エネルギーが良好であると予測された。これらの分子が実際に RdRp に作用するかどうか確認するためには、実験を通じて更なる検証が必要である。

物質名	対象物質を予測した 予測モデルの数	SARS-CoV-2 RdRp との 結合エネルギー (kcal/mol)
Beclabuvir	4	-9.2
Asunaprevir	4	-7.5
Paritaprevir	3	-10.5
Faldaprevir	3	-9.6
Odalasvir	3	-8.8
Simeprevir	3	-8.7
Vedroprevir	3	-8.6
Velpatasvir	3	-8.6
Telaprevir	3	-8.3
Dolutegravir	3	-8.0
Sofosbuvir	3	-6.9
Uprifosbuvir	3	-6.8
Entecavir	3	-6.6
Lobucavir	3	-6.6
Trifluridine	3	-6.3
Nevirapine	3	-6.1
Ledipasvir	2	-9.2
Ruzasvir	2	-8.1
Baloxavir marboxil	2	-8.0
TMC-310911(ASC09)	2	-7.9
Adafosbuvir	2	-7.8
Remdesivir	2	-7.5
Saquinavir	2	-7.2
Abacavir	2	-7.1
Maribavir	2	-7.1
Elvitegravir	2	-6.6
Vidarabine	2	-6.5
Efavirenz	2	-6.3
Valganciclovir	2	-6.2
Valomaciclovir	2	-6.2
Sorivudine	2	-6.1
Ibacitabine	2	-6.1
Idoxuridine	2	-5.9
Fialuridine	2	-5.9
Didanosine	2	-5.8
Umifenovir	2	-5.8

表 1.3 RdRp に作用する可能性のある抗ウイルス薬一覧
及び結合エネルギー（計算は AutoDock Vina を使用）

5.2 抗炎症性データセットの結果

抗ウイルス薬は、患者の負荷軽減には役立つが、ウイルス誘発性肺炎そのものに直接対処するものではない。新型コロナウイルスによる肺炎は、SARS-CoV-2 が肺で引き起こす炎症の結果生じるものである[40]。したがって、肺炎を発症した COVID-19 の患者には、肺の炎症を抑えるための追加治療が必要となる可能性がある。上述のモデルは数多の抗ウイルス薬から RdRp 阻害剤を同定する能力を有するだけでなく、抗炎症薬に対しても同様の効果

を発揮しうることから、SARS-CoV-2 RdRp に対する RdRp 阻害剤の特性と結合エネルギーの予測に焦点を当て、解析を行った。抗炎症性データセットの分析の結果、全ての最良モデルが、天然物質であるベツリン酸とルピオールが抗 RdRp 活性を有していると予測することが明らかになった。他にも、リフィテグラスト、アントラフェニン、ウルソール酸、酢酸デキサメタゾン、リン酸プレドニゾロンは、最良モデルのうち少なくとも2つのモデルによって予測されており、また、-7.5 から-9.5 kcal/mol の範囲の結合エネルギーで SARS-CoV-2 RdRp の活性部位に結合することが予測された (表 1.4)。なお、ベツリン酸とウルソール酸は、HIV に対する抗ウイルス活性が報告されている五環式トリテルペノイドである[41]。

物質名	対象物質を予測した 予測モデルの数	SARS-CoV-2 RdRp との 結合エネルギー (kcal/mol)
Betulinic Acid	4	-7.4
Lupeol	4	-7.2
Lifitegrast	3	-9.5
Antrafenine	3	-8.7
Ursolic acid	3	-8.0
Floctafenine	3	-7.1
Cimicoxib	3	-7.0
Acemetacin	3	-6.8
Morniflumate	3	-6.8
Loteprednol	3	-6.8
Polmacoxib	3	-6.8
Andrographolide	3	-6.7
Dexamethasone acetate	2	-7.6
Prednisolone phosphate	2	-7.5
Cortisone acetate	2	-7.3
Mometasone furoate	2	-7.3
Prednicarbate	2	-7.1
Deflazacort	2	-7.1
Clobetasone	2	-6.8
Rimexolone	2	-6.8
Robenacoxib	2	-6.8
Hydrocortisone probutate	2	-6.8
Mometasone	2	-6.6
Diflunisal	2	-6.5
Lumiracoxib	2	-6.5
Etoricoxib	2	-6.5
Clobetasol	2	-6.5
Apremilast	2	-6.5
Bisindolylmaleimide I	2	-6.5
Talniflumate	2	-6.3
NS-398	2	-6.2
Firocoxib	2	-5.6
Dimethyl sulfone	2	-3.0

表 1.4 RdRp に作用する可能性のある抗炎症性薬一覧
及び結合エネルギー (計算は AutoDock Vina を使用)

6. 結論

COVID-19 の世界的パンデミックの社会的・経済的影響は、世界中の数多くの人々の生活に深刻な影を落としている。多様な生態系において、人命を奪う可能性のあるウイルスを保有する動物が存在し、また、グローバルに人々が移動する現代社会においては、病原体の世界的な伝播が容易になっていることから、将来、新たなパンデミックの出現によってより脆弱な環境にさらされ得る。有効な治療薬の発見のために多大な努力がなされているが、これまでのところ、COVID-19 の患者の治療に役立つ薬剤特定は未だ発展途上であり、多くの死亡者が報告され続けている。これは、創薬のプロセス全体を迅速化できるよう、現在の創薬の方法を慎重に、しかし新しい方法で見直していく必要があることを示唆している。本論文では、機械学習アルゴリズムに基づくスクリーニングを通じ、SARS-CoV-2 RdRp 阻害剤の候補を予測した。今後は、当該分子が SARS-CoV-2 RdRp に対し活性を有するかどうか、モデルだけでなく実際の実験を通じて検証することが望ましいと考える。

7. 謝辞

本論文の作成に当たり、多くの人々のご協力を頂きました。特に、丁寧な原稿の修正と翻訳に協力して頂いた加藤嵩侑氏、伊東美絵氏（ともに株式会社 Elix）に感謝の意を表します。今回のパンデミックは編集現時点でも、多くの人々に多大なる影響を及ぼしています。人々の安全で健康な生活のために、命がけで COVID-19 の患者のケアをして下さっている医療従事者の皆様にも感謝の意を表し、謝辞とさせていただきます。

参考文献

- [1] Chan, J. F.-W. et al.: Genomic characterization of the 2019 novel human-pathogenic coronavirus isolated from a patient with atypical pneumonia after visiting Wuhan. *Emerging Microbes & Infections*, **9**, 221–236 (2020). <http://doi:10.1080/22221751.2020.1719902>.
- [2] Zhou, P. et al.: A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*, **579**, 270–273 (2020). <http://doi:10.1038/s41586-020-2012-7>.
- [3] Beck, B. R. et al.: Predicting commercially available antiviral drugs that may act on the novel coronavirus (SARS-CoV-2) through a drug-target interaction deep learning model. *Computational and Structural Biotechnology Journal*, **18**, 784–790 (2020). <http://doi:10.1016/j.csbj.2020.03.025>.
- [4] Richardson, P. et al.: Baricitinib as potential treatment for 2019-nCoV acute respiratory disease. *The Lancet*, **395**, e30–e31 (2020). [http://doi:10.1016/s0140-6736\(20\)30304-4](http://doi:10.1016/s0140-6736(20)30304-4).
- [5] Kandeel, M.; Al-Nazawi, M.: Virtual screening and repurposing of FDA approved drugs against COVID-19 main protease. *Life Sciences*, **251**, 117627 (2020). <http://doi:10.1016/j.lfs.2020.117627>.

- [6] Cui, J. et al.: Origin and evolution of pathogenic coronaviruses. *Nature Reviews Microbiology*, **17**, 181–192 (2019). <http://doi:10.1038/s41579-018-0118-9>.
- [7] Zhu, N. et al.: A novel coronavirus from patients with pneumonia in China, 2019. *New England Journal of Medicine*, **382**, 727–733 (2020). <http://doi:10.1056/nejmoa2001017>.
- [8] Gralinski, L. E.; Menachery, V. D.: Return of the coronavirus: 2019-nCoV. *Viruses*, **12**, 135 (2020). <http://doi:10.3390/v12020135>.
- [9] Deng, S.-Q.; Peng, H.-J.: Characteristics of and public health responses to the coronavirus disease 2019 outbreak in China. *Journal of Clinical Medicine*, **9**, 575 (2020). <http://doi:10.3390/jcm9020575>.
- [10] Jiang, S. et al.: An emerging coronavirus causing pneumonia outbreak in Wuhan, China: calling for developing therapeutic and prophylactic strategies. *Emerging Microbes & Infections*, **9**, 275–277 (2020). <http://doi:10.1080/22221751.2020.1723441>.
- [11] Wu, Z.; McGoogan, J. M.: Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China. *JAMA*, **323**, 1239–1242 (2020). <http://doi:10.1001/jama.2020.2648>.
- [12] Lu, R. et al.: Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *The Lancet*, **395**, 565–574 (2020). [http://doi:10.1016/s0140-6736\(20\)30251-8](http://doi:10.1016/s0140-6736(20)30251-8).
- [13] Venkataraman, S. et al.: RNA dependent RNA polymerases: Insights from structure, function and evolution. *Viruses*, **10**, 76 (2018). <http://doi:10.3390/v10020076>.
- [14] Snijder, E. J. et al.: The nonstructural proteins directing coronavirus RNA synthesis and processing. *Advances in Virus Research*, **96**, 59–126 (2016). <http://doi:10.1016/bs.aivir.2016.08.008>.
- [15] Lan, J. et al.: Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature*, **581**, 215–220 (2020). <http://doi:10.1038/s41586-020-2180-5>.
- [16] Masters, P. S.: The Molecular biology of coronaviruses. *Advances in Virus Research*, **66**, 193–292 (2006). [http://doi:10.1016/s0065-3527\(06\)66005-3](http://doi:10.1016/s0065-3527(06)66005-3).
- [17] Bruenn, J. A.: A structural and primary sequence comparison of the viral RNA-dependent RNA polymerases. *Nucleic Acids Research*, **31**, 1821–1829 (2003). <http://doi:10.1093/nar/gkg277>.
- [18] Xu, X. et al.: Molecular model of SARS coronavirus polymerase: implications for biochemical functions and drug design. *Nucleic Acids Research*, **31**, 7117–7130 (2003). <http://doi:10.1093/nar/gkg916>.
- [19] Černý, J. et al.: Evolution of tertiary structure of viral RNA dependent polymerases. *PLoS ONE*, **9**, e96070 (2014). <http://doi:10.1371/journal.pone.0096070>.
- [20] Gao, Y. et al.: Structure of the RNA-dependent RNA polymerase from COVID-19 virus. *Science*, **368**, 779–782 (2020). <http://doi:10.1126/science.abb7498>.

- [21] Hage-Melim, L. I. da S. et al.: Virtual screening, ADME/Tox predictions and the drug repurposing concept for future use of old drugs against the COVID-19. *Life Sciences*, **256**, 117963 (2020). <http://doi:10.1016/j.lfs.2020.117963>.
- [22] Quimque, M. T. J. et al.: Virtual screening-driven drug discovery of SARS-CoV2 enzyme inhibitors targeting viral attachment, replication, post-translational modification and host immunity evasion infection mechanisms. *Journal of Biomolecular Structure and Dynamics*, , 1–23 (2020). <http://doi:10.1080/07391102.2020.1776639>.
- [23] Naik, B. et al.: High throughput virtual screening reveals SARS-CoV-2 multi-target binding natural compounds to lead instant therapy for COVID-19 treatment. *International Journal of Biological Macromolecules*, **160**, 1–17 (2020). <http://doi:10.1016/j.ijbiomac.2020.05.184>.
- [24] Jiménez-Alberto, A. et al.: Virtual screening of approved drugs as potential SARS-CoV-2 main protease inhibitors. *Computational Biology and Chemistry*, **88**, 107325 (2020). <http://doi:10.1016/j.compbiolchem.2020.107325>.
- [25] Zhavoronkov, A. et al.: Deep learning enables rapid identification of potent DDR1 kinase inhibitors. *Nature Biotechnology*, **37**, 1038–1040 (2019). <http://doi:10.1038/s41587-019-0224-x>.
- [26] Fleming, N.: How artificial intelligence is changing drug discovery. *Nature*, **557**, S55–S57 (2018). <http://doi:10.1038/d41586-018-05267-x>.
- [27] Stebbing, J. et al.: COVID-19: combining antiviral and anti-inflammatory treatments. *The Lancet Infectious Diseases*, **20**, 400–402 (2020). [http://doi:10.1016/s1473-3099\(20\)30132-8](http://doi:10.1016/s1473-3099(20)30132-8).
- [28] Wallach, I. et al.: AtomNet: A deep convolutional neural network for bioactivity prediction in structure-based drug discovery. *arXiv*, (2015).
- [29] <https://www.businesswire.com/news/home/20200521005238/en/Atomwise-Partners-with-Global-Research-Teams-to-Pursue-Broad-Spectrum-Treatments-Against-COVID-19-and-Future-Coronavirus-Outbreaks>
- [30] Kim, S. et al.: PubChem 2019 update: improved access to chemical data. *Nucleic Acids Research*, **47**, D1102–D1109 (2018). <http://doi:10.1093/nar/gky1033>.
- [31] Gaulton, A. et al.: ChEMBL: a large-scale bioactivity database for drug discovery. *Nucleic Acids Research*, **40**, D1100–D1107 (2012). <http://doi:10.1093/nar/gkr777>.
- [32] Landrum, G.: RDKit: Open-Source Cheminformatics Software. *null*, (2016).
- [33] Rogers, D.; Hahn, M.: Extended-connectivity fingerprints. *Journal of Chemical Information and Modeling*, **50**, 742–754 (2010). <http://doi:10.1021/ci100050t>.
- [34] Pedregosa, F. et al.: Scikit-Learn: Machine learning in Python. *Journal of Machine Learning Research*, **12**, 2825–2830 (2011).
- [35] Trott, O.; Olson, A. J.: AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of Computational Chemistry*, **31**, 455–461 (2010). <http://doi:10.1002/jcc.21334>.

- [36] Giacchello, I. et al.: Insights into RNA-dependent RNA polymerase inhibitors as anti-influenza virus agents. *Current Medicinal Chemistry*, **27** (2020).
<http://doi:10.2174/0929867327666200114115632>.
- [37] Dierynck, I. et al.: TMC310911, A novel human immunodeficiency virus type 1 protease inhibitor, shows in vitro an improved resistance profile and higher genetic barrier to resistance compared with current protease inhibitors. *Antimicrobial Agents and Chemotherapy*, **55**, 5723–5731 (2011).
<http://doi:10.1128/aac.00748-11>.
- [38] Gasparini, R. et al.: Compounds with anti-influenza activity: present and future of strategies for the optimal treatment and management of influenza. Part I: Influenza life-cycle and currently available drugs. *Journal of preventive medicine and hygiene*, **55**, 69–85 (2014).
- [39] Gentles, R. G.: Discovery of Beclabuvir: A potent allosteric inhibitor of the hepatitis C virus polymerase. *Topics in Medicinal Chemistry*, **31**, 193–228 (2019).
http://doi:10.1007/7355_2018_38.
- [40] Tay, M. Z. et al.: The trinity of COVID-19: immunity, inflammation and intervention. *Nature Reviews Immunology*, **20**, 363–374 (2020). <http://doi:10.1038/s41577-020-0311-8>.
- [41] Xiao, S. et al.: Recent progress in the antiviral activity and mechanism study of pentacyclic triterpenoids and their derivatives. *Medicinal Research Reviews*, **38**, 951–976 (2018).
<http://doi:10.1002/med.21484>.