

厚生労働科学研究費補助金
(政策科学総合研究事業 (臨床研究等 ICT 基盤構築・人工知能実装研究事業))
分担研究年度終了報告書

課題名 : 肺がん統合データベース構築及び AI 技術を用いたオミックス解析
研究分担者名 : 浜本 隆二
国立がん研究センター研究所医療 AI 研究開発分野・分野長

研究要旨

次世代シーケンサーや遺伝子編集技術に象徴されるバイオテクノロジーの加速的な進化の一方で、医薬品開発を巡る近年の研究開発の生産性は急激に低下し、「イノベーションの欠乏」が叫ばれて久しい。10 年前後の期間にわたって開発される新薬にかかる費用はますます増大し、一つの薬を上市するために必要な開発コストは 1000 億円以上に上る。また、候補化合物から医薬品として世に出されるものは 2 万から 3 万に一つであり、その成功確率は極めて低い。特に、新薬候補物質の効果を始めて患者で示す”Phase II 試験 “での開発中止 (Phase II attrition) が顕著であり、開発した薬剤の多くがマウスでは効くものの、人間ではその有効性が証明されないという、マウス実験を基盤とする現状の医薬品開発パイプラインの構造的な課題がある。本研究課題申請者らが対象疾患とする肺がんにおいても、患者の遺伝子を調べることで最適な治療薬を選択する「がんゲノム医療」が本邦でも開始されたものの、遺伝子の異常が検出されながらも適合する薬剤が存在しないケースが多数出現することが判明しており、社会問題となっている。以上の事から、新薬開発プロセスにおける構造的な課題を克服し、疾患治療で現状打破を得るためのイノベーションが現在期待されている。そのための鍵となるものが、ヒトの薬剤投与時や疾患罹患時における網羅的分子プロファイルに現れる生体変化の統合的理解と、実際の医療から得られるリアルワールド・データの解析に基づいた薬剤のヒトにおける有効性や毒性予測、更にこうした解析を可能にする最新の計算機アルゴリズムであると考えられる。

近年、「50 年来の技術的ブレークスルー」とも言われる深層学習と最新の機械学習手法の組み合わせが、生命情報処理に対するアルゴリズム革命をもたらすものと期待されている。特に、生物の脳神経系にヒントを得た情報処理メカニズムである人工ニューラルネットワークを多層化した深層学習は、社会や産業の形を変え得る画期的な情報技術として大きな注目を集めている。深層学習を既存の機械学習手法と比して特徴付けるものとしては、多種のデータを入力として取り扱えるマルチモーダル学習、複数の異なるタスクをモデルの一部として共有できるマルチタスク学習、少数の教師ありデータから汎化性能を得ることのできる半教師あり学習や教師無し学習、階層的な特徴を自動的に獲得できる表現学習などが挙げられる。近年のコンピュータ処理能力の飛躍的向上と加速的に蓄積されるビッグデータを基盤として、深層学習は画像や音声といった比較的高度な知的タスクで、既に人間を上回る性能を示すに至り、生命情報処理ならびに創薬の分野においてもその応用に関する世界的な研究開発競争が激化している。そこで、本邦における分子標的治療薬の創生及び生命情報処理に関わる第一線の研究者が一つのチームとして協力し合うことで、がんの生体時空間にわたるシステムの統合理解に基づいた創薬候補因子の探索と、がん治療におけるあるべき不均一性を包含した、臨床のリアルワールド・データに基づいたヒトにおける有効性や毒性の予測を可能にすると捉え、医薬品開発における「イノベーションの欠乏」を解決するための新しい技術開発に取り組んだ。その結果 AI を活用した複数の解析プラットフォームの開発に成功した。また AI 解析において最重要な事項の一つとして、質の高いデータベースの構築が挙げられるが、我々は臨床データを効率的に収集するプラットフォームの構築に成功した。その結果 AI 解析を志向した世界最大規模の肺がん統合データベースを、本研究課題で構築することに成功した。本データベースは質の高いデータが保存されていることから、本邦の財産と考えられる貴重な研究成果であると判断している。

A. 研究目的

医薬品開発において、近年国内外を問わず創薬ターゲットの枯渇が問題となっている。現在残されているのは高難易度の創薬ターゲットのみであるがために、新薬の研究開発には多額の費用が必要となっており、これが高薬価、ひいては医療費の高騰の要因となっている。

更に、臨床試験段階で期待していた薬効が得られず開発が中断する例が増えていることも問題点として挙げられる。特に医薬品開発の70~80%がPhase2で中止となっており、この約60%が、薬効が得られなかったことが原因との報告がある。つまり、「動物では効くが、ヒトでは効かなかった」という事案が多発している。これは現在の創薬研究開発スキームの限界であると考えられる。

このような現状を打開する解決策として、人工知能 (AI; Artificial intelligence) が注目されている。AIのパフォーマンスと可能性に創薬・医療・ヘルスケア分野が大きな期待を寄せており、今後国際競争が激化することが必至である。

これらの現状を背景に、本事業では、「創薬ターゲットの枯渇問題」を克服すべく、動物からではなくヒトの情報から創薬ターゲット分子を探索するAIの開発実装を目的とする。つまり、臨床情報 (=電子カルテを始めとする診療情報+オミックスデータ) を収集・利用して創薬ターゲットを探索するAI手法の開発をおこなう。本事業では、対象疾患として部位別がん死亡者数1位である肺がんを選択し、これらの臨床情報収集とそれを支援する基盤構築、異種かつ大量のデータを統合して創薬ターゲット候補となる生体分子群を自動的に抽出するAI手法の開発を行う。また、本事業で作成される肺がんの疾患統合データベース、機能分子を特定するためのAI及び知識ベース等を多くの研究者等に利用してもらうための環境 (オープンプラットフォーム) の構築を目指す。

今年度は、i) データ収集：世界最大規模の肺がん統合データベースの拡充、ii) 肺がんデータプラットフォームの構築、iii) 解析アルゴリズムの開発を目標とする。

B. 研究方法

- ① 肺がんオミックス統合データベースの拡充：特にエピゲノム解析 (ChIP-seq 解析、DNA メチル化解析) を優先して解析を進める。
- ② 肺がんデータプラットフォームの構築：AI 解析を志向した効率的なデータ及び臨床情報収集システムを構築する。
- ③ 解析アルゴリズムの開発：AI を用いてマルチオミックスデータ解析システムを開発する。

(倫理面への配慮)

本研究課題は、国立がん研究センター・研究倫理審査委員会承認を受け (研究開発課題：2019-108、“新薬創出を加速する人工知能の開発”)、実施している。

C. 研究結果

1.

概要: AI の学術研究を行う上で最重要事項の一つとして、質の高いデータベースを構築することが挙げられる。我々は今年度も引き続き肺がんオミックスデータベースの拡充に取り組み、臨床情報 1,714 症例・RNA-seq 解析データ 1,682 症例・全エクソーム解析データ 1,556 症例、及び DNA メチル化解析データ 403 症例・全ゲノム解析データ 413 症例・ChIP-seq(H3K27Ac)解析データ 222 症例まで統合データベースを拡大した。この症例数は、TCGA データベースに保存されている肺がん症例数である 1,020 を超え、世界最大規模であり、我が国における貴重な財産であると判断している。

2.

概要: AI 研究により優れた研究成果を出していくためには、質の高い構造化データを準備する事は非常に重要である。しかし、多くの医療データは不均質で構造化されておらず、価値創造の機会を逸しているのが現状である。我々は富士フイルム株式会社と共同で、AI 技術と最先端の ICT 技術を組み合わせ、効率的にデータを構造化するプラットフォームの構築に成功し、2022 年4月5日に「Synapse Creative Space」という名称で製品化した。(本プラットフォームは医療 AI 研究を発展させるうえで重要なイノベーションであり、貴重な成果である。)

3.

概要: がんは複雑な疾患である為、その本態解明を行い創薬に発展させる為には、臨床検体から得られた膨大な

オミックスデータを、詳細な診療情報と共に効率的に解析する技術を開発することが必須であると考えている。そこで、我々は機械学習・深層学習技術を活用して、がんに関する医療ビッグデータを解析する手法の開発に取り組んだ。その結果令和3年度は、DNA メチロームデータ解析に応用した新手法、methPLIER を開発し、Briefings in Bioinformatics 誌に投稿し、peer-review に回った(現在 revision 中)。

4.

概要:肺がん統合データベース構築においては、ロボットも活用した RCRA ChIP-seq 法を用いて、世界に先駆けて大規模な ChIP-seq 解析を施行した結果、pan-negative 肺がんにおいて新規治療標的となる遺伝子の発現調整領域に、super-enhancer が形成されていることを突き止めた。この成果は、ゲノム解析のみでは発見できず、様々なオミックスデータをマルチモーダルに解析して初めて得られる結果であるため、大変意義があると判断している。現在論文発表の準備を進めるとともに、臨床試験に向けた取り組みも進めている。

D. 研究発表

1. 論文発表

- 1) Kim K, Ryu TY, Jung E, Lee J, Han TS, Kim SK, Roh YN, Lee MS, Jung CR, Lim JH, Hamamoto R, Lee HW, Hur K, Son MY, Kim DS, Cho HS. SMYD2 is a key regulator of lung cancer metastasis by controlling SMAD3 expression. *Exp Mol Med. in press*
- 2) Shirasawa T, Yoshida T, Shiraishi K, Takigami A, Takayanagi D, Imabayashi T, Matsumoto Y, Masuda K, Shinno Y, Okuma Y, Goto Y, Horinouchi H, Yotsukura M, Yoshida Y, Nakagawa K, Tsuchida T, Hamamoto R, Yamamoto N, Motoi N, Kohno T, Watanabe S, Ohe Y. Identification of inflamed-phenotype of small cell lung cancer leading to the efficacy of anti-PD-L1 antibody and chemotherapy. *Lung Cancer. in press*
- 3) Hamamoto R, Takasawa K, Shinkai N, Machino H, Kouno N, Asada K, Komatsu M, Kaneko S. Analysis of super-enhancer using machine learning and its application to medical biology. *Brief Bioinform. bbad107. Online ahead of print.*
- 4) Ito T, Takayanagi D, Sekine S, Hashimoto T, Shimada Y, Matsuda M, Yamada M, Hamamoto R, Kato T, Shida D, Kanemitsu Y, Boku N, Kohno T, Takashima A, Shiraishi K. Comparison of clinicopathological and genomic profiles in anal squamous cell carcinoma between Japanese and Caucasian cohorts. *Sci Rep. 2023 Mar 3;13(1):3587.*
- 5) Asami Y, Kobayashi Kato M, Hiranuma K, Matsuda M, Shimada Y, Ishikawa M, Koyama T, Komatsu M, Hamamoto R, Nagashima M, Terao Y, Itakura A, Kohno T, Sekizawa A, Matsumoto K, Kato T, Shiraishi K, Yoshida H. Utility of molecular subtypes and genetic alterations for evaluating clinical outcomes in 1029 patients with endometrial cancer. *Br J Cancer. 2023 Apr;128(8):1582-1591.*
- 6) Dozen A, Shozu K, Shinkai N, Ikawa N, Aoyama R, Machino H, Asada K, Yoshida H, Kato T, Hamamoto R, Kaneko S, Komatsu M. Tumor Suppressive Role of the PRELP Gene in Ovarian Clear Cell Carcinoma. *J Pers Med. 2022 Dec 2;12(12):1999.*
- 7) Shozu K, Kaneko S, Shinkai N, Dozen A, Kosuge H, Nakakido M, Machino H, Takasawa K, Asada K, Komatsu M, Tsumoto K, Ohnuma SI, Hamamoto R. Repression of the PRELP gene is relieved by histone deacetylase inhibitors through acetylation of histone H2B lysine 5 in bladder cancer. *Clin Epigenetics. 2022 Nov 12;14(1):147.*
- 8) Kukita A, Sone K, Kaneko S, Kawakami E, Oki S, Kojima M, Wada M, Toyohara Y, Takahashi Y, Inoue F, Tanimoto S, Taguchi A, Fukuda T, Miyamoto Y, Tanikawa M, Mori-Uchino M, Tsuruga T, Iriyama T, Matsumoto Y, Nagasaka K, Wada-Hiraike O, Oda K, Hamamoto R, Osuga Y. The Histone Methyltransferase SETD8 Regulates the Expression of Tumor Suppressor Genes via H4K20 Methylation and the p53 Signaling Pathway in Endometrial Cancer Cells. *Cancers (Basel). 2022 Oct 31;14(21):5367.*
- 9) Hossain E, Abdelrahim M, Tanasescu A, Yamada M, Kondo H, Yamada S, Hamamoto R, Marugame A, Saito Y, Bhandari P. Performance of a novel computer-aided diagnosis system in the characterization of colorectal polyps, and its role in meeting Preservation and Incorporation of Valuable Endoscopic Innovations standards set by the American Society of Gastrointestinal Endoscopy. *DEN Open. 2022 Oct 28;3(1):e178.*

- 10) Hamamoto R, Koyama T, Kouno N, Yasuda T, Yui S, Sudo K, Hirata M, Sunami K, Kubo T, Takasawa K, Takahashi S, Machino H, Kobayashi K, Asada K, Komatsu M, Kaneko S, Yatabe Y, Yamamoto N. Introducing AI to the molecular tumor board: one direction toward the establishment of precision medicine using large-scale cancer clinical and biological information. *Exp Hematol Oncol*. 2022 Oct 31;11(1):82.
- 11) Hopkins J, Asada K, Leung A, Papadaki V, Davaapil H, Morrison M, Orita T, Sekido R, Kosuge H, Reddy MA, Kimura K, Mitani A, Tsumoto K, Hamamoto R, Sagoo MS, Ohnuma SI. PRELP Regulates Cell-Cell Adhesion and EMT and Inhibits Retinoblastoma Progression. *Cancers (Basel)*. 2022 Oct 8;14(19):4926.
- 12) Yamada M, Shino R, Kondo H, Yamada S, Takamaru H, Sakamoto T, Bhandari P, Imaoka H, Kuchiba A, Shibata T, Saito Y, Hamamoto R. Robust automated prediction of the revised Vienna Classification in colonoscopy using deep learning: development and initial external validation. *J Gastroenterol*. 2022 Nov;57(11):879-889.
- 13) Hamamoto R, Takasawa K, Machino H, Kobayashi K, Takahashi S, Bolatkan A, Shinkai N, Sakai A, Aoyama R, Yamada M, Asada K, Komatsu M, Okamoto K, Kameoka H, Kaneko S. Application of non-negative matrix factorization in oncology: one approach for establishing precision medicine. *Brief Bioinform*. 2022 Jul 18;23(4):bbac246.
- 14) Hashimoto T, Takayanagi D, Yonemaru J, Naka T, Nagashima K, Yatabe Y, Shida D, Hamamoto R, Kleeman SO, Leedham SJ, Maughan T, Takashima A, Shiraishi K, Sekine S. Clinicopathological and molecular characteristics of RSP0 fusion-positive colorectal cancer. *Br J Cancer*. 2022 Oct;127(6):1043-1050.
- 15) Ono S, Komatsu M, Sakai A, Arima H, Ochida M, Aoyama R, Yasutomi S, Asada K, Kaneko S, Sasano T, Hamamoto R. Automated Endocardial Border Detection and Left Ventricular Functional Assessment in Echocardiography Using Deep Learning. *Biomedicines*. 2022 May 6;10(5):1082.

2. 学会発表

- 1) 浜本 隆二、今後の医学領域における AI 研究及び社会実装の展望、第 40 回日本脳腫瘍学会学術集会・教育講演：鴨川グランドホテル（鴨川）、2022 年 12 月 4 日
- 2) 浜本 隆二、実臨床応用を志向した医療 AI 研究：ビッグデータが拓く新しい医学、第 5 回日本眼科アレルギー学会学術集会・特別講演：米子コンベンションセンター（米子）、2022 年 11 月 5 日
- 3) 浜本 隆二、臨床応用を目的とした医療 AI 研究開発：研究の立案から社会実装まで、第 60 回日本癌治療学会学術集会・基調講演：神戸コンベンションセンター（神戸）、2022 年 10 月 20 日
- 4) Ryuji Hamamoto, Social implementation of AI-based medical device programs、第 81 回日本癌学会学術総会・シンポジウム：パシフィコ横浜（横浜）、2022 年 9 月 29 日
- 5) 浜本 隆二、日本のメディカル AI の現況の総論、第 67 回日本透析医学会学術集会・総会・特別講演：パシフィコ横浜（横浜）、2022 年 7 月 3 日
- 6) 浜本 隆二、AI のがん研究への導入と医療への実装、第 26 回日本がん分子標的治療学会学術集会・Year in Review2：石川県立音楽堂（金沢）、2022 年 6 月 30 日
- 7) 浜本 隆二、PRISM の成果と今後の戦略～肺がんプロジェクトを中心として～、第 4 回日本メディカル AI 学会学術集会・シンポジウム：トークネットホール仙台（仙台）、2022 年 6 月 10 日
- 8) 浜本 隆二、ゲノム医学における AI の活用、第 119 回日本内科学会講演会・シンポジウム：ロームシアター京都（京都）、2022 年 4 月 16 日
- 9) 浜本 隆二、実臨床応用を目的とした医療 AI 研究開発、第 122 回日本外科学会定期学術集会・「外科学の未来を拓く」：熊本城ホール（熊本）、2022 年 4 月 14 日

E. 知的財産権の出願・登録状況（予定を含む。）

1. 特許取得

なし

2. 実用新案登録

なし

3. その他
なし