

2019年度研究実施報告書

(対象期間 2019年11月13日～2020年3月31日)

担当課題：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

研究機関名：国立研究開発法人国立がん研究センター

研究責任者：浜本 隆二 rhamamot@ncc.go.jp

【1. 実施内容】

実施項目1

1-1 研究目的

動物ではなく、ヒトの情報から創薬ターゲット分子を探索するAIの開発実装を行うことを目的とする。対象疾患としては、特に肺がんに焦点を当てて行う。

1-2 研究概要・要旨

肺がんに対する臨床情報（電子カルテを始めとする診療情報＋患者検体のオミックスデータ）を収集・利用して創薬ターゲットを探索するAI手法の開発を行う。

1-3 実施内容

- ① オミックスデータの取得：国がんで行っているバイオバンク事業等と連携し、それらの事業で取得された肺がん検体からDNA/RNAを取得し、医薬健栄研に提供した。また、肺がん検体のChIP-seq解析を行った。
- ② 肺がん患者の臨床情報のデータベース化：①の肺がん検体と紐付いた臨床情報を収集し、AIの入力データとして利用するためにクレンジング等の作業を行った。
- ③ 統合データ解析のAI手法開発：①のオミックスデータとそれに紐付く②の臨床情報を用い、肺がんの本態解明、ひいては創薬ターゲット候補を見出すためのAI手法の開発を行う。

1-4 結果、成果等

- ① データ収集に関しては、全エクソーム解析及び診療情報に関しては521症例新たに肺がんデータベースに追加し、総計1569症例となり世界最大規模のデータベースとなった。その他にも全ゲノム解析は242症例（総計272症例）、DNAメチル化解析は160症例（総計430症例）、ChIP-seq解析は100症例（総計100症例）、トランスクリプトーム（RNA-seq）解析は654症例（総計1494症例）が肺がんデータベースに追加された。
- ② 呼吸器内科・呼吸器外科にある電子カルテ情報を、ゲノム・エピゲノム情報と紐づけた統合データベースを構築した。
- ③ 肺がんオミックス解析を目的とした新規アルゴリズムを4種類開発し、学会発表及び論文発表を行った。

【2. 外部発表、論文投稿等】

[学会発表]

1. 浜本 隆二. がんの統合的解明を目指したオミックス情報の階層的ネットワークに対する機械学習・深層学習技術の応用: 第39回医療情報学連合大会 (招待講演)
2. 浜本 隆二. 人工知能 (AI) 技術の医療応用への取組-医用画像解析を中心として-: 第37回日本脳腫瘍学会学術集会 (招待講演)
3. 小林 和馬, 三宅 基隆, 渡辺 裕一, 浜本 隆二. 臨床のリアル・ワールド・データを利用した人工知能技術の開発を加速するための統合的なデータ・プラットフォームの構築: 第2回日本メディカル AI 学会学術集会
4. 浅田 健, 小林 和馬, 高橋 慧, 高澤 建, 小松 正明, 金子 修三, 浜本 隆二. 肺がん予後予測を施行したマルチオミックス解析: 第2回日本メディカル AI 学会学術集会
5. 高澤 建, 浜本 隆二. 知識行列を用いたスパース正則化付き NMF による DNA メチル化データ解析手法の開発: 第2回日本メディカル AI 学会学術集会
6. Syuzo Kaneko, Ryuji Hamamoto. Toward the application of precision medicine: Multi-omics analysis including clinical ChIP-seq datasets may reveal pivotal regulation of gene expression in cancer: 第2回日本メディカル AI 学会学術集会

[論文投稿]

1. Kobayashi K, Murakami N, Takahashi K, Inaba K, Igaki H, Hamamoto R, Itami J. A Population-based Statistical Model for Investigating Heterogeneous Intraprostatic Sensitivity to Radiation Toxicity After ¹²⁵I Seed Implantation. *In Vivo*, 33:2103-2111 (2019)
2. Asada K, Kobayashi K, Joutard S, Tubaki M, Takahashi S, Takasawa K, Komatsu M, Kaneko S, Sese J, Hamamoto R. Uncovering Prognosis-related Genes and Pathways by Multi-omics Analysis in Lung Cancer. *Biomolecules*, 10:524 (2020)
3. Hamamoto R, Komatsu M, Takasawa K, Asada K, Kaneko S. Epigenetics Analysis and Integrated Analysis of Multiomics Data, Including Epigenetic Data, Using Artificial Intelligence in the Era of Precision Medicine. *Biomolecules*, 10:62 (2020)
4. Isago, H, Mitani A, Mikami Y, Horie M, Urushiyama H, Hamamoto R, Terasaki Y, Nagase T. Epithelial Expression of YAP and TAZ Is Sequentially Required in Lung Development. *Am J Respir Cell Mol Biol*, 62 :256-266 (2020)

【3. 知財化について】

該当なし