

戦略的創造研究推進事業 CREST  
研究領域「イノベーション創発に資する  
人工知能基盤技術の創出と統合化」  
研究課題「人工知能を用いた統合的ながん医療  
システムの開発」

## 研究終了報告書

研究期間 2016年11月～2019年3月

研究代表者：浜本 隆二  
(国立がん研究センター研究所  
分野長)

## § 1 研究実施の概要

### (1) 実施概要

我が国において現在がんは死亡原因の第1位で、国民の2人に1人は人生に一度はがんにかかるといわれる報告ががん対策情報センターから発表されており、“がん”という疾患はもはや国民病の様相を呈している。近年分子医学の進歩と集学的治療の進歩で、がんという疾患に対する恐怖は以前と比較すると減弱しつつあるが、依然人間にとって生命の脅威となる疾患であり、がん研究の進展及び社会への還元は、常に国民から期待されている。一方、我が国においては世界でもトップレベルの質の高い、がんの基礎研究・臨床研究・疫学研究が長い間継続的に行われてきており、蓄積されたデータは膨大な量になる。これまでは蓄積された膨大なデータを、統合的に解析する手段が無かったが、近年の人工知能(AI)技術の発展により、ビッグデータの解析が可能な時代となっている。そこで、本研究プロジェクトにおいては、これまで長年蓄積されてきた我が国のがんに関する医療ビッグデータを、本邦におけるがん対策拠点である国立がん研究センターが主導で統合的に解析し、日本人研究者による日本人の為のがん医療システムを開発することを目標としてきた。研究チームとしては国立がん研究センターの浜本グループが、本研究の中心的役割を担い、統合的推進を図ってきた。機械学習・深層学習技術を用いた解析に関しては、株式会社 Preferred Networks の岡野原グループ及び産業技術総合研究所の瀬々グループが担当し、内視鏡画像解析に関しては日本電気株式会社 (NEC) の協力も得て行ってきた。

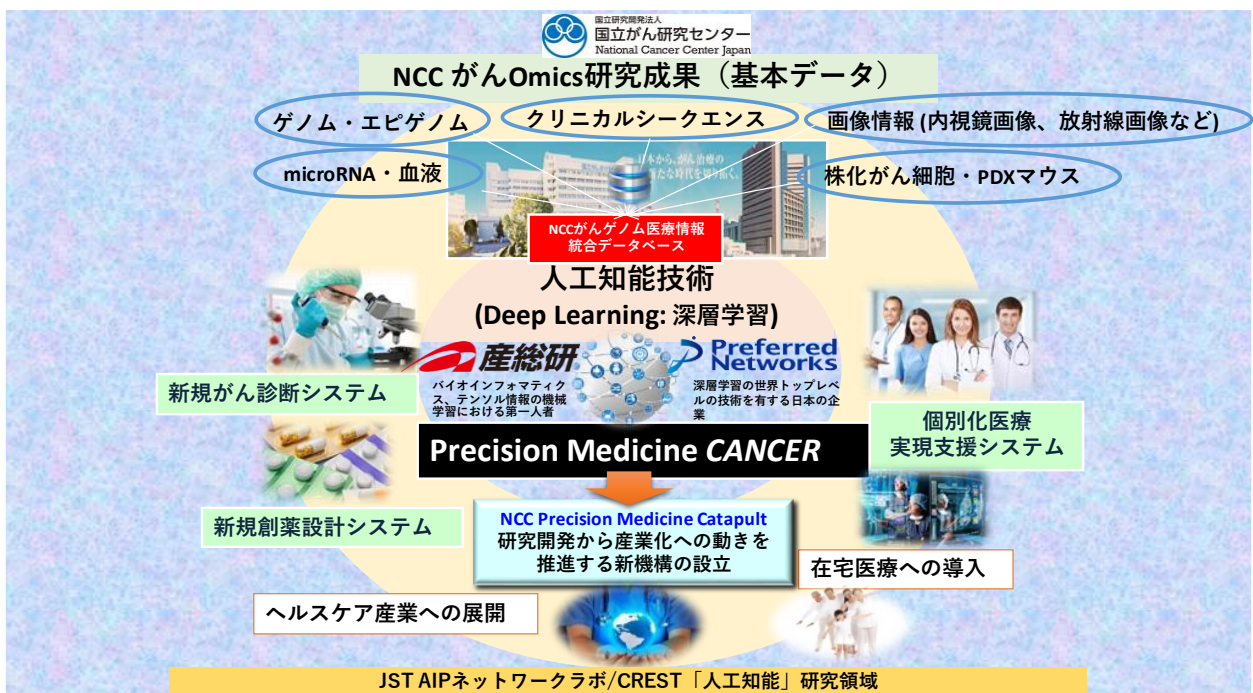
本プロジェクトに関しては主に3つの研究テーマに取り組んできた。一つ目は AI 技術を活用したマルチオミクス解析を志向した、エピジェネティクス解析手法の確立である。現在のがん医療分野において Precision Medicine は世界的な潮流となっているが、その中核と判断され現在推進されているのは、次世代シーケンサーを用いたクリニカルシーケンスである。“がん”という疾患の病態を考えた場合、その発生や進展に遺伝子の変異が重要な働きをしていることは、長年の研究成果から明らかとなっている。また抗がん剤の感受性予測や予後予測にも遺伝子変異がバイオマーカーとして重要な働きをしていることは、現在既にコンセンサスが得られている。しかし同時に遺伝子の変異に関する情報だけではがんの本態に関して説明することはできず、抗がん剤の感受性予測や予後予測、さらに副作用予測も遺伝子変異の情報だけでは説明しきれないことも明らかになってきている。Precision Medicine が目指す個々に最適化された治療法を確立する為には、各患者さんが保持する heterogeneity な性質を適切に解析する為には必要があり、遺伝子変異以外の情報を加味し統合的に解析することは今後 Precision Medicine を推進する上で、必須になると考えている。その観点から我々は、FFPE sample や細胞数の少ない細胞からも ChIP-sea 解析が必要な、次世代型 ChIP-seq 法の開発に取り組み、開発に成功した。本解析法には LabDroid も活用し、データ解析の段階で機械学習を取り入れる努力も行った。今後は多数の臨床検体 (FFPE サンプルも含む) を用いて、ヒストン修飾情報及びクロマチン構造情報を取得し、ゲノム情報など他のオミクスデータと統合化し、マルチモーダルな解析を行うことを予定している。

二つ目として AI 技術を用いた医用画像解析を重点課題として取り組んできた。本課題で用いた医用画像は主に内視鏡画像・放射線画像及び皮膚腫瘍画像である。内視鏡画像に関しては、深層学習を活用した大腸がん及び前がん病変発見の為のリアルタイム内視鏡診断サポートシステムの開発に取り組み、世界トップレベルの精度を示すシステムの開発に成功した。本システムに関しては現在論文投稿を進めると共に、早期社会実装を目指し、医薬品医療機器総合機構 (PMDA) と既に面談を繰り返している。今後は全国の医療施設への導入を目指し、クラウド化構想にも取り組んでいく。放射線画像解析に関しては、MRI 脳腫瘍画像に対する Radiomics を展開し、個別化・精密化された医療体系の実現を目指した。その結果自動化された Radiomics 解析系により、単純な腫瘍サイズの変化では捉えきれない画像特徴量の経時的変化を網羅的かつ定量的に解析することが可能となった。またセグメンテーションに用いた 3D-Unet の最終出力層を fine tuning することによって、腫瘍領域に対して選択的に精度を

向上させた検出器を構築した。さらに放射線治療後のフォローアップの頭部 MRI に対して、深層学習による腫瘍領域の検出器とレジストレーション Radiomics 解析系を組み合わせることで、治療効果の自動的・定量的判定を可能にするシステムを構築した。今後 Radiomics プロジェクトに関しては、脳腫瘍外科コンソーシアムと密接に連携することで、大規模のコホートを用いて学習器の性能を brush up すると共に、社会実装を目指して企業と連携していく予定である。皮膚腫瘍のプロジェクトに関しては、深層学習を用いてほくろと悪性黒色腫を識別する学習器の開発に取り組み、高い精度を示す学習器の構築に成功した。

三つ目の重点課題として、我々は AI 技術を用いた解析を志向した、がん統合マルチオミックスデータベースの構築に取り組んだ。最近の機械学習・深層学習用アルゴリズムの進歩は非常に早く、最先端のアルゴリズムを活用することにより、様々な成果を出すことが可能な時代となっている。一方、アルゴリズムがいくら発達しても良好な結果を出すためには、質の高いデータを用いることが重要であることは、AI 技術を活用した医学研究全般について当てはまる。例えば本邦は OECD 諸国の中で、CT や MRI などの画像診断装置を人口あたり最も多く有しており、実際国立がん研究センターにおいても 4 億枚以上という、世界でもトップクラスの医用放射線画像を保有している。しかし現状ではその多くが不均質で構造化されておらず、価値創造の機会が大きく損なわれてきた。こうした医用画像を単なる「写真」の集まりとして保管するのではなく、コンピュータによる解析、学習、推論の対象となるようなビッグデータとして蘇らせるための取り組みが急務である。そこで我々は独自にアノテーションツールを開発し、また vendor neutral archive (VNA) システムも導入することで、AI 技術を用いた解析を志向したがん統合マルチオミックスデータベースシステムを構築した。本システムの導入に伴い、中央病院と新研究棟 4 階に設置された GPGPU クラスタ（本クラスタも CREST 予算で導入された）を専用線で結ぶ工事も始めており、大規模がんオミックスデータが効率的に AI 技術を用いて解析が可能なシステムの構築を行っている。将来的には新研究棟で AI 技術を用いた解析結果が、中央病院に還元されるシステムとして還元していく予定である。

## 研究課題「人工知能を用いた統合的ながん医療システムの開発」の概念図



(国立研究開発法人国立がん研究センター研究倫理審査委員会[許可]研究課題番号: 2016-496)

## (2) 顕著な成果

### <優れた基礎研究としての成果>

1.

概要:我々は次世代型 ChIP-seq 解析法の開発に成功した。その結果、ホルマリン固定パラフィン包埋 (FFPE) 標本からや、細胞数が少ないサンプルから ChIP-seq が可能になり、今後ヒストン修飾解析やクロマチン解析など幅広い分野で活用されることが期待される。特に FFPE 標本から転写因子結合部位を ChIP-seq 解析を用いて同定する技術は世界初の報告であり、新規性・有用性が高い技術であると判断している。本研究成果は現在原著論文として報告する準備をしている (Nature Methods, *in preparation*)。

2.

概要:我々は深層学習を活用した大腸がん及び前がん病変発見の為のリアルタイム内視鏡診断サポートの開発に成功した。深層学習を活用したリアルタイム内視鏡診断システムとしては、世界トップレベルの精度を示し、現在早期社会実装を目指し PMDA と面談を繰り返している。本成果は基礎医学研究としても意義があり、現在原著論文として発表準備を行っている (Scientific Reports, *in revision*)。

3.

概要:我々が構築した Radiomics 解析系により、単純な腫瘍サイズの変化では捉えきれない画像特徴量の経時変化を、網羅的かつ定量的に解析することが可能となった。また放射線治療後のフォローアップの頭部 MRI に対して、深層学習による腫瘍領域の検出器とレジストレーション、Radiomics 解析系を組み合わせることで、治療効果の自動的・定量的を可能にするシステムを構築した。さらに大規模な Glioblastoma MRI 画像を解析することで、画像から遺伝子型を高精度に予測することに成功した。これらの成果は全て原著論文として投稿する準備を現在している。

### <科学技術イノベーションに大きく寄与する成果>

1.

概要:我々は今回次世代 ChIP-seq 解析法を確立した事により、FFPE 臨床検体からも精度高くヒストン修飾解析及び転写因子結合領域の解析を行うことが可能になった。本技術は Precision Medicine を推進するうえにおいて重要で、今後エクソーム解析やトランスクリプトーム解析と同様に、大規模な臨床検体を網羅的に解析できる可能性を示唆している (特許取得準備中)。実際現在産業技術総合研究所・人工知能研究センター及びロボティック・バイオロジーインスティテュート (RBI) 株式会社と共同で、早期社会実装を目指した共同研究を開始している。

2.

概要:我々は深層学習を活用した大腸がん及び前がん病変発見の為のリアルタイム内視鏡サポートシステムの開発に成功したが、本システムは大腸内視鏡検査における病変発見率の向上、医師の技術格差の解消及び病変見逃しの予防に大変有用である。現在早期社会実装を目指しており、2020 年度までの製品化及びグローバル普及を目指している。

3.

概要:我々は放射線治療後のフォローアップの頭部 MRI に対して、深層学習による腫瘍領域の検出器とレジストレーション、Radiomics 解析系を組み合わせることで、治療効果の自動的・定量的判定を可能にするシステムを構築した。また MRI 画像から Glioblastoma の遺伝子型を高精度に予測する学習器の開発に成功した。本研究に関しては今後企業と共同で早期社会実装を目指す予定で、話し合いを開始している。

### <代表的な論文3報>

1. Vassiliki Saloura, Theodore Vougiouklakis, Makda Zewde, Takehiro Suzuki, Naoshi Dohmae, Xiaolan Deng, Kazuma Kiyotani, Jae-Hyun Park, Yo Matsuo, Mark Lingen, Ryuji Hamamoto, Yusuke Nakamura: WHSC1L1-mediated EGFR mono-methylation enhances the cytoplasmic and nuclear oncogenic activity of EGFR. **Scientific Reports**, 7, 40664 (2017)
2. Chul-Hwan Lee, Marlene Holder, Daniel Grau, Ricardo Saldaña-Meyer, Jia-Ray Yu, Rais Ahmad Ganai, Jenny Zhang, Miao Wang, Gary LeRoy, Marc-Werner Dobenecker, Danny Reinberg, Karim-Jean Armache: Allosteric Activation Dictates PRC2 Activity Independent of Its Recruitment to Chromatin. **Molecular Cell**, 40, 422-434 (2018)
3. Yoshinobu Shigekawa, Shinya Hayami, Masaki Ueno, Atsushi Miyamoto, Norihiko Suzuki, Manabu Kawai, Seiko Hirono, Ken-ichi Okada, **Ryuji Hamamoto**, Hiroki Yamaue: Overexpression of KDM5B/JARID1B is associated with poor prognosis in hepatocellular carcinoma. **Oncotarget**, 9, 34320-34335 (2018)

## § 2 研究実施体制

### (1) 研究チームの体制について

#### (1) 浜本グループ

- ① 研究代表者: 浜本 隆二 (国立がん研究センター研究所 分野長)  
研究参加者: 落谷 孝広 (国立がん研究センター研究所 主任分野長)  
研究参加者: 金子 修三 (国立がん研究センター研究所 ユニット長)  
研究参加者: 山田 真善 (国立がん研究センター中央病院 医員)  
研究参加者: 小林 和馬 (国立がん研究センター研究所 連携大学院生)  
研究参加者: 三宅 基隆 (国立がん研究センター中央病院 医員)  
研究参加者: 高橋 雅道 (国立がん研究センター中央病院 医員)  
研究参加者: 陣内 駿一 (国立がん研究センター中央病院 がん専門修練医)  
研究参加者: 三原 直樹 (国立がん研究センター中央病院 部長)  
研究参加者: 田代 志門 (国立がん研究センター研究支援センター 部長)  
研究参加者: 上條 憲一 (国立がん研究センター研究所 外来研究員 [本務・NEC 医療ソリューション事業部])  
研究参加者: 今岡 仁 (国立がん研究センター研究所・外来研究員 [本務・NEC データサイエンス研究所])

#### ② 研究題目

- ・本研究の中心的役割を担い、統合的推進を図る。

#### (2) 岡野原グループ

- ① 主たる共同研究者: 岡野原 大輔 ((株)Preferred Networks 取締役副社長)  
研究参加者: 大田 信行 (Preferred Networks America 最高執行責任者)  
研究参加者: 菅原 洋平 ((株)Preferred Networks エンジニア)  
研究参加者: 平野 湧一郎 ((株)Preferred Networks エンジニア)

#### ② 研究題目

- ・深層学習技術を用いた統合的な解析を行う。特に microRNA データ解析及び医用画像データ解析を中心に行う。

### (3) 瀬々グループ

① 主たる共同研究者:瀬々 潤 (産業技術総合研究所人工知能研究センター 研究チーム長)

研究参加者:河口 理紗 (産業技術総合研究所人工知能研究センター 特別研究員)

#### ② 研究題目

・機械学習全般の解析手法を用いた統合的解析を行う。特にエピジェネティクス解析及び医用画像データ解析を中心に行う。

#### (2) 国内外の研究者や産業界等との連携によるネットワーク形成の状況について

本研究課題においては、1. AI 技術を用いたエピジェネティクス解析及び創薬・診断への応用、2. AI 技術による microRNA を対象とした早期がん診断システムの開発、3. AI 技術による高精度の画像診断システムの開発、4. AI 技術を用いたクリニカルシーケンスデータ解析/Precision Medicine を指向したがん治療の最適化という、4つのテーマについて研究を推進させてきた。それぞれのテーマに関して国内外の研究者や産業界等との連携によるネットワーク形成の状況を下記に示す。

##### [1. AI 技術を用いたエピジェネティクス解析及び創薬・診断への応用]

本研究テーマに関して、LabDroid”まほろ“の活用に関しては、RBI 株式会社の夏目徹博士、産業技術総合研究所人工知能研究センターの光山統泰博士及び東京医科歯科大学の浅原弘嗣博士と共同で行った。また臨床検体の解析に関しては、東京大学附属病院産婦人科の曾根献文博士及び和歌山県立医科大学附属病院第二外科の速水晋也博士の協力を得て本研究に使用した。さらに ChIP-seq データベースの構築に関しては、九州大学大学院医学研究院の沖真弥博士と連携して行った。

##### [2. AI 技術による microRNA を対象とした早期がん診断システムの開発]

本テーマに関しては、国立がん研究センター中央病院の各科の先生から協力を得て行った。

##### [3. AI 技術による高精度の画像診断システムの開発]

本テーマに関して、深層学習技術を活用した大腸がん及び前がん病変発見の為のリアルタイム内視鏡診断サポートシステムの開発に関しては、国立がん研究センター内視鏡科及び NEC と連携を組み行った。

放射線画像解析に関しては、国立がん研究センター中央病院放射線治療科・放射線診断科・脳脊髄腫瘍科、産業技術総合研究所人工知能研究センター及び国内 16 施設の脳神経外科と連携を組み行った。

皮膚腫瘍解析に関しては、国立がん研究センター中央病院皮膚腫瘍科と連携を組み行った。

##### [4. AI 技術を用いたクリニカルシーケンスデータ解析/Precision Medicine を志向したがん治療の最適化]

肺がん統合データベース構築に関しては、国立がん研究センター研究所ゲノム生物学研究分野、国立がん研究センター中央病院医療情報部・放射線診断科・放射線治療科と連携を組み行った。