

令和 3 年 5 月 31 日現在

機関番号：13101

研究種目：基盤研究(A) (一般)

研究期間：2018～2020

課題番号：18H04123

研究課題名(和文)がんゲノム医療実現化に向けた高次元データ解析基盤構築

研究課題名(英文)High-dimensional data analysis platform realizing cancer genome medicine

研究代表者

奥田 修二郎 (Okuda, Shujiro)

新潟大学・医歯学系・教授

研究者番号：00512310

交付決定額(研究期間全体)：(直接経費) 35,200,000円

研究成果の概要(和文)：がんという病気は細胞中のDNAに損傷(変異)が起きること、その変異が長年の間修復されずに積み重なることにより発症する。現在では超ハイスループットなDNAシーケンサーの発達により比較的安価に変異同定が可能になった。そこで重要となるのが、がんゲノム情報のデータ統合化とデータ解析基盤の構築である。これまでに大腸がん、胃がん、肺がん、乳がん等700症例以上の固形がんにおける遺伝子検査結果のデータベース化を実施した。さらに臨床情報とゲノム変異の情報を合わせてあらゆる条件でデータ解析を実施できるシステムの開発を実施した。さらにこれの人工知能への応用として、病理画像を用いた病態の判定技術を開発した。

研究成果の学術的意義や社会的意義

ある程度以上の症例数が含まれるコホートなどになるとがんゲノムデータは非常に膨大な変異の情報となる。従来は、表計算ソフトなどを駆使し手作業でさまざまな解析を実施してきたが、このような大規模データになると手作業での解析は非常に困難であり、また、エラーを誘発する危険がある。我々が開発してきたデータ解析基盤は、この問題を解決し、がんゲノムデータ解析を効率化する事が可能になる。精密医療としてのがんゲノム医療が実際の患者に適用されているが、臨床データの解析においてエラーがなく信頼性の高い解析結果に基づいたエビデンスが得られる環境は、学術に関してのみならず社会的にも非常に大きな意義を持っていると考える。

研究成果の概要(英文)：Cancer is a disease caused by damage (mutation) to the DNA in a cell and the accumulation of mutations that are not repaired for many years. The development of very high-throughput DNA sequencers has made it possible to identify mutations at a relatively low cost. Therefore, it is important to integrate the data of cancer genome information and build a data analysis infrastructure. To date, we have created a database of genetic test results for more than 700 solid tumors, including colorectal cancer, gastric cancer, lung cancer, and breast cancer. In addition, we have developed a system that can analyze data under any conditions by combining clinical information and genome mutation information. Furthermore, as an application of this system to artificial intelligence, we have developed a technique for determining pathological conditions using pathological images.

研究分野：バイオインフォマティクス

キーワード：がんゲノム データベース 精密医療

1. 研究開始当初の背景

(1) がんという病気は細胞中の DNA に損傷 (変異) が起きること、その変異が長年の間修復されずに積み重なることにより発症する。これまで個人のがん組織における変異を全て同定するには膨大な費用がかかったが、現在では次世代シーケンサー (NGS) と呼ばれる超ハイスループット DNA シーケンサーの発達により比較的安価に変異同定が可能になった。米国では The Cancer Genome Atlas (TCGA) などのがんゲノムを網羅的に調べる大型プロジェクトが立ち上がり、数百・千人規模でのがん変異の情報が蓄積されている (<http://cancergenome.nih.gov/>)。これらの成果から、がん細胞が持つ変異は特定の遺伝子に偏りがあることがわかってきた。がん発症に関連する数百の遺伝子 (ターゲット遺伝子) の変異を調べれば、概ねそのがん組織における原因変異を同定することが可能である。米国を始め世界中でこのターゲット遺伝子のシーケンスに基づいたがんゲノム研究が進展している。

(2) がん組織は最初にごん化した細胞と、それとは異なる変異パターンをもつサブクローンとの集合体である (腫瘍内不均一性)。この腫瘍内不均一性が理由で、変異同定は非常に難しい計算処理を必要とする。しかし、ターゲット遺伝子のみの配列を非常に深くシーケンスすることにより、情報量を嵩上げすることで変異同定の精度が上がる。現在のターゲットシーケンスでは 500 倍程の高深度シーケンスが主流である。そこから得られた変異情報を元に、分子標的薬との適合性を評価する。最適な分子標的薬が存在する場合には、その薬剤での治療が最も効果を示すことが期待できる。このように最適な治療選択を精密に実施する Precision medicine を実現するには、実際の変異情報と分子標的薬との適合性を確かめる必要があるが、現在網羅的にこの適合性を確認した実験結果は存在しない。しかしながら、たとえば分子標的薬の 1 つ Trastuzumab (ERBB2 阻害薬) では、ERBB2 が高発現 (Her2 陽性) であれば適用可能と判断され顕著に奏功する場合もあるが、40%程度は効果が認められない。この薬剤は ERBB2 を標的にするため下流の遺伝子にも変異がある場合 (約 60%) には、その効果が期待できないためである。分子標的薬の標的タンパク質が変異を受けても、その他の遺伝子も変異している場合、あるいは、腫瘍不均一性により他の変異パターンのクローンが多勢の場合は、その効果の判定が非常に難しくなることがある。

2. 研究の目的

(1) いわゆる医療ビッグデータの中には、現状では明確ではない変異遺伝子と分子標的薬の効果の関連について推測できる情報が埋まっていることが期待できる。そこで、一人の人間では扱いきれない程の大規模情報とがんゲノム情報から実臨床に意味のある関連情報を取り出すことを目的としている。

(2) Precision medicine としてのクリニカルシーケンシングでは、遺伝子変異と最適治療薬の推薦が必要であるが、その治療法選択が実際にどの程度効果を示すか、また、その効果判定の精度はどの程度か、ということに答えを見つけない限り実臨床への応用は不可能であり、これががんゲノム医療実現における最も高いハードルである。つまり、ある遺伝子変異が起きる時、細胞の中ではその他の遺伝子も含め何が起きているのかについて、実データに基づいた確度の高い予測モデルを確立する。

(3) がんゲノム医療として実際に提案される最適治療法の根拠になっているデータ処理の中で、なんらかのエラーが存在する場合、致命的な過誤を生じる可能性がある。人の手作業という最もエラーの発生する可能性の高い処理を無くし、研究用データ解析および実臨床応用においてエラーが生じないシステムを構築する。

(4) 大型計算機を駆使した人工知能や方法論としての深層学習は、次代の日本を担う技術の根幹と目されているが、これを医療に応用する場合に意味のあるモデルを作成する。

3. 研究の方法

(1) 申請者らはすでに大腸がん 201 症例、胃がん 207 症例、肺がん 167 症例、乳がん 53 症例を含め 700 症例以上の固形がんにおける遺伝子検査を実施してきた。これらのゲノム変異情報に加え関連する臨床情報をデータベース化する。年齢・性別、ステージ、TNM 分類、全生存期間、無増悪生存期間等の情報に加え、術前・後の薬剤使用履歴など様々な情報を収集する。

(2) がん化の原因の多くは、関連するシグナル伝達系の何処かに遺伝子変異が起きることによって細胞増殖のサイクルが狂うことにある。シグナル伝達系を中心にがん関連パスウェイの情報を文献より収集し、データベース化する。

(3) 遺伝子変異のパターンは人それぞれ異なるが、申請者らは分子標的薬との関連が示唆される特異的な遺伝子変異を持つ患者群をクラスター化する方法を開発してきた。分子標的薬の効果の評価するにはクラスター毎に臨床情報と合わせた統計解析が必須である。本研究課題により得られる臨床情報に基づく医療統計解析を自動・半自動的に実施できるシステムを開発する。

(4) 深層学習の応用として画像判別は非常に有用であり、その手法を病理画像の判定に応用す

る。深層学習に必要となる学習用画像セットを簡便に作成するためのシステムの構築に加え、それをシームレスに深層学習システムに接続するための仕組みの構築も実施する。

4. 研究成果

(1) 大腸がん、胃がん、肺がん、乳がん等の固形がんにおけるがんゲノムデータとその臨床情報をデータベース化してきた。また、比較解析のために TCGA における関連のがんゲノムデータをダウンロードし、同様のフォーマットに変換することで、同じデータベース内で管理できる体制を整えた。日本人側のデータは遺伝子パネル検査から得られたパネル遺伝子に関するものが多いため、TCGA のデータでも同様にパネル遺伝子での変異頻度等が得られるように遺伝子のダウンサンプリングも実施した。これらのデータにより、さまざまな臨床情報と遺伝子変異との関連の解析が外科学分野を始め多くの臨床分野の研究者らとの共同研究という形で実施してきた。大腸癌の場合、大腸における左右差などの部位の違いと特定の遺伝子に関して変異頻度の有意な違いが観察されることを見つけ報告した。また、さらにこの研究は、BRAF 遺伝子のような特定のホットスポットに遺伝子変異が入りやすい遺伝子について、ホットスポット変異とそれ以外の臨床的意義について調査し、予後に関与する因子の可能性があることについても報告した。このように本申請において開発してきた統合解析プラットフォームにがんゲノムの情報と臨床情報を入力し、簡便に解析ができる環境が得られたことにより、上記のようなデータ解析において結果が得られるまでの時間をかなり短縮することに成功している。本システムの開発そのものに関する実績以外にもそれらを応用した共同研究などですでに 20 報以上の論文が公開されており、がんゲノムデータ解析を促進することに成功している。このシステムをさらに広範囲に応用して利用できるように現在でも改良に努めている。

(2)他にも、遺伝子変異の量(Tumor mutation burden)が非常に多く蓄積するタイプの症例(TMBが高い)に関する大腸癌や乳がんにおける臨床的意義についての研究も進展した。乳がんにおいてはこの TMB が高いグループにおいて臨床背景として喫煙が関連している事を発見した。また、大腸癌において TMB が高い症例は、同時にマイクロサテライト不安定性の状態であることがよく観察されるが、これらの症例では腫瘍浸潤リンパ球が有意に増大していることを病理画像からも確認しており、免疫チェックポイント阻害剤の効果との関連が示唆された。このような TMB が非常に高い症例における病理画像には湯用浸潤リンパ球のような病理医が目でその状態について判定できる部分が含まれているが、すべての TMB が高い症例でその状態を示すわけではないことも知られている。そこで、この TMB の高い症例を人工知能的に判定できるモデルの作成を実施した。その結果、非常に高精度に TMB の高い症例の病理画像を区別することができる人工知能の開発に成功し報告した。また、この技術は他の癌にも応用できるため、現在、臓器横断的にこの人工知能の技術を応用し、医療用人工知能の開発を実施している。

<引用文献>

Eifert, C., et al. Clinical application of a cancer genomic profiling assay to guide precision medicine decisions. *Personalized Medicine* 14:4 2017.

Nagahashi, M., et al. Genomic landscape of colorectal cancer in Japan: clinical implications of comprehensive genomic sequencing for precision medicine. *Genome Medicine* 8:136 2016

凌一 韋ほか、「変異頻度に基づくがんゲノムデータの新規クラスタリング手法とその臨床応用」、医療情報学-春季学術大会論文、38(5):305-312 (2018)

5. 主な発表論文等

〔雑誌論文〕 計15件（うち査読付論文 14件／うち国際共著 0件／うちオープンアクセス 6件）

1. 著者名 Nagahashi M, Shimada Y, Ichikawa H, Kameyama H, Takabe K, Okuda S, Wakai T.	4. 巻 110
2. 論文標題 Next generation sequencing-based gene panel tests for the management of solid tumors.	5. 発行年 2018年
3. 雑誌名 Cancer Sci.	6. 最初と最後の頁 6-15
掲載論文のDOI（デジタルオブジェクト識別子） 10.1111/cas.13837	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -
1. 著者名 Ling Y, Okuda S.	4. 巻 38
2. 論文標題 Novel gene fusions found in cervical cancer.	5. 発行年 2018年
3. 雑誌名 EBioMedicine	6. 最初と最後の頁 13-14
掲載論文のDOI（デジタルオブジェクト識別子） 10.1016/j.ebiom.2018.11.013	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -
1. 著者名 Takeuchi S, Okuda S.	4. 巻 24
2. 論文標題 Knowledge base toward understanding actionable alterations and realizing precision oncology.	5. 発行年 2019年
3. 雑誌名 Int J Clin Oncol.	6. 最初と最後の頁 123-130
掲載論文のDOI（デジタルオブジェクト識別子） 10.1007/s10147-018-1378-0.	査読の有無 無
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -
1. 著者名 Shimada Yoshifumi, Okuda Shujiro, Watanabe Yu, Tajima Yosuke, Nagahashi Masayuki, Ichikawa Hiroshi, Nakano Masato, Sakata Jun, Takii Yasumasa, Kawasaki Takashi, Homma Kei-ichi, Kamori Tomohiro, Oki Eiji, Ling Yiwei, Takeuchi Shiho, Wakai Toshifumi	4. 巻 56
2. 論文標題 Histopathological characteristics and artificial intelligence for predicting tumor mutational burden-high colorectal cancer	5. 発行年 2021年
3. 雑誌名 Journal of Gastroenterology	6. 最初と最後の頁 547 ~ 559
掲載論文のDOI（デジタルオブジェクト識別子） 10.1007/s00535-021-01789-w	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Ling Yiwei, Watanabe Yu, Nagahashi Masayuki, Shimada Yoshifumi, Ichikawa Hiroshi, Wakai Toshifumi, Okuda Shujiro	4. 巻 18
2. 論文標題 Genetic profiling for diffuse type and genomically stable subtypes in gastric cancer	5. 発行年 2020年
3. 雑誌名 Computational and Structural Biotechnology Journal	6. 最初と最後の頁 3301 ~ 3308
掲載論文のDOI (デジタルオブジェクト識別子) 10.1016/j.csbj.2020.10.021	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

[学会発表] 計13件 (うち招待講演 7件 / うち国際学会 3件)

1. 発表者名 Takeuchi, S., Ling, Y., Watanabe, Y., Shimada, Y., Wakai, T., Okuda, S.
2. 発表標題 Mutational signatures and clinicopathological relationship extracted from clinical sequence data of 201 cases with colorectal cancer.
3. 学会等名 ASHG2019 (国際学会)
4. 発表年 2019年

1. 発表者名 Ling, Y., Yoshizaki, H., Okuda, S.
2. 発表標題 Genome-wide analysis of mutational impact on post translational modification.
3. 学会等名 ISMB2019 (国際学会)
4. 発表年 2019年

1. 発表者名 若井俊文、島田能史、永橋昌幸、市川寛、油座築、根本万里子、中野麻恵、廣瀬雄己、滝沢一泰、坂田純、亀山仁史、小林隆、栗田学、吉原弘祐、奥田修二郎
2. 発表標題 固形癌における包括的ゲノム解析に基づくPrecision Medicine
3. 学会等名 第57回日本癌治療学会学術集会
4. 発表年 2019年

1. 発表者名 永橋昌幸、若井俊文、島田能史、市川寛、羽入隆晃、滝沢一泰、石川卓、坂田純、小林隆、亀山仁史、竹内志穂、奥田修二郎
2. 発表標題 次世代シーケンサーを用いた遺伝子検査とゲノム解析データベース構築
3. 学会等名 第78回日本癌学会学術総会
4. 発表年 2019年

1. 発表者名 奥田修二郎
2. 発表標題 Precision Medicine実現に向けたバイオインフォマティクスの役割
3. 学会等名 第57回日本癌治療学会学術集会（招待講演）
4. 発表年 2019年

1. 発表者名 奥田修二郎
2. 発表標題 Precision medicine実現に向けたがんゲノムインフォマティクス
3. 学会等名 第8回生命医薬情報学連合大会
4. 発表年 2019年

1. 発表者名 奥田修二郎
2. 発表標題 プレジジョンメディスン実現に向けたがんゲノムインフォマティクスとマイクロバイオームへの応用
3. 学会等名 第61回日本婦人科腫瘍学会・第18回日本婦人科がん分子標的研究会（招待講演）
4. 発表年 2019年

1. 発表者名 奥田修二郎
2. 発表標題 Precision medicine を目指すがんゲノムインフォマティクス
3. 学会等名 JPrOS2019/JES2019 (招待講演)
4. 発表年 2019年

1. 発表者名 竹内志穂、凌一葦、渡邊由、若井俊文、奥田修二郎
2. 発表標題 がんプレジジョンメディシン実現に向けたパイオインフォマティクス：遺伝子変異と治療薬の知識ベース
3. 学会等名 第41回日本分子生物学会年会
4. 発表年 2018年

1. 発表者名 Shujiro Okuda
2. 発表標題 Computational approaches towards understanding human genome, proteome, and proteogenome data
3. 学会等名 2018 Taiwan-Japan Joint Conference for Genomics and Proteogenomics and Annual Retreat of Taiwan Genomics and Genetics Society (招待講演) (国際学会)
4. 発表年 2018年

1. 発表者名 奥田修二郎
2. 発表標題 新潟大学におけるがんゲノムプロジェクト
3. 学会等名 第22回医療情報学会春季学術集会 (招待講演)
4. 発表年 2018年

1. 発表者名 奥田修二郎、渡辺由、凌一葦、竹内志穂
2. 発表標題 がんゲノム医療を加速するデータ解析システム
3. 学会等名 第78回新潟癌治療研究会（招待講演）
4. 発表年 2018年

1. 発表者名 奥田修二郎
2. 発表標題 病理画像から遺伝子変異を予測する
3. 学会等名 第9回生命医薬情報学連合大会（招待講演）
4. 発表年 2020年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究分担者	赤澤 宏平 (Akazawa Kohei) (10175771)	新潟大学・医歯学総合病院・教授 (13101)	
研究分担者	島田 能史 (Shimada Yoshifumi) (20706460)	新潟大学・医歯学系・助教 (13101)	
研究分担者	永橋 昌幸 (Nagahashi Masayuki) (30743918)	新潟大学・医歯学総合病院・研究准教授 (13101)	

6. 研究組織（つづき）

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究分担者	若井 俊文 (Wakai Toshifumi) (50372470)	新潟大学・医歯学系・教授 (13101)	
研究分担者	市川 寛 (Ichikawa Hiroshi) (50721875)	新潟大学・医歯学系・助教 (13101)	

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関