

大規模データ解析と人工知能技術による  
がんの起源と多様性の解明

文部科学省「富岳」成果創出加速プログラム



analysis and artificial intelligence technology

Unravelling origin of cancer and diversity by

# Newsletter

No. 7

2023年3月31日

# 「富岳」成果創出加速プログラム 大規模データ解析と人工知能技術によるがん の起源と多様性の解明

## 人馬一体の研究チームの成しとげたこと

課題代表者 東京医科歯科大学M&Dデータ科学センター 特任教授 **宮野 悟**



がんの起源と多様性の解明は世界のがん研究の中心になっている課題です。4年前、この課題申請をしたとき、具体的目的を次のA)-C)に設定しました。

- A) **発がんの初期のクローン選択はどのようにして生ずるのか**
- B) **どのようにして多様性が生じ、転移や治療抵抗性の獲得、再発に至るのか**
- C) **それらはどのようにがんの臨床的な振舞いと関連するのか**

これらの課題を、ゲノムシーケンスデータ解析パイプラインGenomon、ネットワーク推定・解析アプリパッケージSiGN、富士通研究所が開発したネットワークを入力とするDeep Tensorという3つのアプリと「富岳」などのスパコンで解決することを目標としました。

研究の流れが図1です。

ベースには小川が独自に開発した単細胞

シーケンス技術、大規模高度サンプリング技術があります。

小川のところに導入されている最先端シーケンサーをもちいてDNA, RNA等のシーケンスデータをだしました。

これらをGenomonで解析し、がんの変異の全貌を明らかにしました。

また定量化したデータからSiGNというネットワーク推定アプリにより出力されたネットワークを入力として、機械学習システムDeep Tensor解析を合流させました。

SiGNは、2006年来、次世代スパコンプロジェクト、及びその後のプロジェクトで開発してきたもので、富岳に実装しました。Deep Tensor は連携機関である富士通研究所が開発し、本課題で説明可能AIへと発展させ、富士通研究所が実装しました。

図2は研究実施体制です。東京医科歯科大学が代表機関となり2020年4月に新

設されたM&Dデータ科学センターの面々、特に、Heewon Park教授、及び愛知県がんセンターの山口類先生が活躍が情報系として活躍しました。そして、京都大学大学院医学研究科の小川誠司教授のグループから多数、特に、そしてプロジェクトの途中で東大医科学研究所に異動した南谷泰仁教授（京大を兼ねており特定教授）、越智陽太郎助教が活躍しました。また、東大医科学研究所ヒトゲノム解析センターの片山琴絵先生には、計算リソースのマネージメントでご支援いただきました。富士通研究所も連携機関としてはいり、AIの導入でご支援いただきました。小川教授はスウェーデンのカロリンスカ研究所をつなぎました。

このチームは、モータースポーツの頂点であるF1にたとえると、ドライバーが小川教授、コンストラクターは情報系のメンバー達、そしてエンジンはスパコンです。これらが人馬一体（図3）となって戦い、チャンピオンを達成しました。

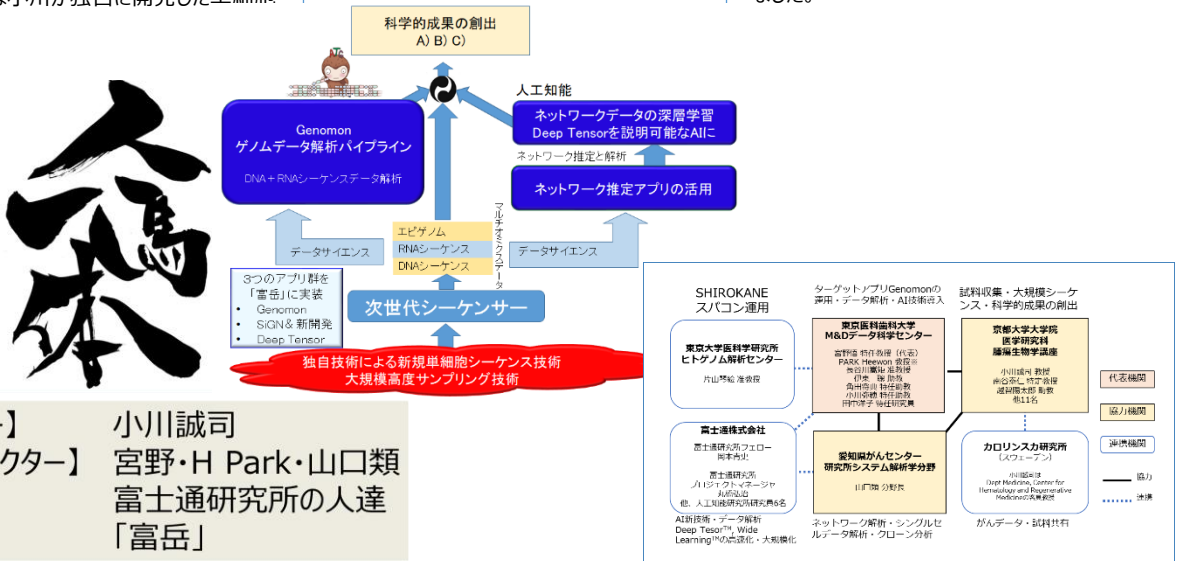


図3：人馬一体の研究

図1：研究の流れ

図2：研究体制

本課題はプロジェクトの総合的推進も含め次の3つのサブテーマからなっております。

1. サブテーマA：大規模データ解析アプリケーションの実装と実行
2. サブテーマB：大規模データ創出によるがんの起源と多様性の解明
3. サブテーマC：プロジェクトの総合的推進

サブテーマAの計画は、Genomon、ネットワーク推定・解析アプリ、Deep Tensor + TRIP、そして新たにWide Learningの「富岳」実装と実行でした。その実績をまとめると以下ようになります。

- ・ シーケンスデータ解析パイプライン Genomonを計算のボトルネックになっているところを見出しプログラムを改良し「富岳」に実装

使用ノード数をファイルシステムとの関連で最適化チューニングしました。GenomonはサブテーマBの科学的成果の創出に貢献しました。

- ・ 3000検体以上のがん全ゲノム（がんゲノム、正常ゲノムのペア）シーケンスのGenomonでの解析を実施

この開発状況、チューニング状況については達成状況の説明の中で述べます。RNAの解析も実施しました。様々な困難がありましたが、サブテーマBの科学的成果の創出に貢献しました。

- ・ NetworkProfilerなどのSiGNパッケージアプリを「富岳」に実装

大規模な遺伝子発現データセットからネットワークを推定し説明可能AIの入力としました。同時に、がんの多様性をとらえるための新たなネットワーク推定技術を開発・実装し、転移や治療抵抗性に関する科学的成果へつなげました。クローンを単位としたクローンネットワークの変化をネットワークとして捉える方式も開発しました。「京」やほかのスパコンでは「待てられない」ほどの時間が「富岳」ではサクッと計算できました。

- ・ AIアプリDeep Tensorを「富岳」に実装

Deep Tensorを説明可能アプリにするために新たな補助アプリTRIPを開発し、上皮間葉転換に関する科学的成果を得ました。「富岳」とTRIPがなければ到達できなかった世界を体験しました。

- ・ AIアプリWide Learningを「富岳」に実装

薬剤耐性メカニズムの科学的理解に貢献創薬の強い支援アプリになることを確信しました。「富岳」は私たちが“impossible”の世界へ解き放ちました（プレス発表）。

サブテーマBの計画は、大規模データ創出によるとがんの起源と多様性の解明です。サブテーマAと人馬一体の研究体制によりA), B), C)の各問題を世界トップレベルで解決しました。その実績をまとめると以下のようになります。

- ・ A-1) クローン拡大履歴の解析

クローン拡大履歴の解析のために、同一幹細胞由来領域(単一クリプト)、あるいは単一細胞由来培養のコロニーの全ゲノムシーケンスを行ない、分子クロックを決めました。

- ・ A-2) 膵がん、大腸がん、胃がん、および乳がんの初期病変の全ゲノムシーケンスによるクローン履歴の推定

A-1で決定した分子クロックと組み合わせることで、がんの初期病変の全ゲノムシーケンスによるクローン履歴の推定法を開発し、膵がん、大腸がん、胃がん、および乳がんの初期病変の全ゲノムシーケンスによるクローン履歴を推定しました。

- ・ A-3) 単一細胞シーケンスによる高齢健康人のクローン性造血の解析

高齢健康人のクローン性造血の解析のために、多数の単一細胞で同時に変異と発現を解析可能なプラットフォームを用いてクローン性造血の解析を行いました。

- ・ B-1) 単一細胞シーケンスを用いたがんのクローン進展の解明

新規シングルセル解析プラットフォームを用いることにより、同一患者内で同一環境内に存在する複数クローンを区別して比較することで、新たな変異を獲得することにより引き起こされる形質変化や、特定の変異が選択される環境変化の改正を行いました。

- ・ B-2) 全ゲノム解析を用いた新規の発がん・抵抗性獲得メカニズムの解明

全ゲノム解析を用いた新規の発がん・抵抗性獲得メカニズムの解明のために、骨髄系腫瘍500例について、全ゲノムシーケンスと、長鎖リードシーケンスないし疑似長鎖リードを補完的に使い、がんゲノムの構造異常を解

明しました。

- ・ C-1) 造血器腫瘍の網羅的なドライバー遺伝子異常の解析

大規模な検体集積を行ってきた骨髄系腫瘍4000検体、小児急性リンパ性白血病1000検体、悪性リンパ腫500検体の大規模な試料を用いた網羅的なドライバー遺伝子異常の解析を行いました。

- ・ C-2) 大腸がんの網羅的なドライバー遺伝子異常の解析

大規模な検体集積を行ってきた大腸がんについて3,000例を超える試料を用いた網羅的なドライバー遺伝子異常の解析を行いました。

こうして、研究目的のA)-C)を世界トップレベルで解決しました。

サブテーマCは、プロジェクトの総合的推進です。COVID-19禍のなかで、Zoomによる本課題に関する「富岳」ミーティング2回/月実施、公開シンポジウム「がんはどのようにはじまり、そしてどのように強敵にばけるのか」

(2022年12月18(日))の開催、電子版ニュースレターを7 issueを発行、成果のプレス発表、ホームページの開設と運用、メディアインタビュー（中央公論など）によるアウトリーチ活動、研究成果は論文として発表しました。

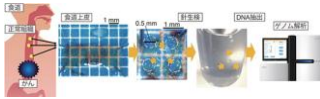
以上の実績と成果をまとめると、

**独創的新技術（実験技術、大規模検体収集、AI導入、大規模計算手法、効率化実装の技芸等）**により、スパコンも加えた人馬一体体制で**問題A), B), C)に関する世界トップの研究成果を創出したこと**です。特に、現在投稿中の「**乳がんのライフヒストリーの解明**」は思春期前後にクローンが発生し（がんではない）30年ほどの年月を経て乳がんを発症するという乳がん発症のメカニズムの解明を解明したことです。これは**乳がん発症に新たな示唆を与える極めて科学的・社会的インパクトのある**ものです。さらに、がんの多様性という観点で行った**説明可能AIx富岳の研究は、生命科学・創薬研究の未来を見通したチャレンジ**と位置づけています。

**独自性・優位性を示す科学的成果**に10項目を挙げました。以下の成果は、本課題の実施期間(2020年4月~2023年3月)中に発表されたものですが、長年取り組んでいた研究がこの期間に表出したもので、このプログラム開始前夜から全世界が注目する成果を出し続けておりました。ポスト「京」プログラムでは、2019年1月に次の成果を発表しておりました。

**Age-related remodelling of oesophageal epithelia by mutated cancer drivers.**

Yokoyama A, Kakiuchi N, Yoshizato T, Nannya Y, Suzuki H, Takeuchi Y, Shiozawa Y, Sato Y, Aoki K, Kim SK, Fujii Y, Yoshida K, Kataoka K, Nakagawa MM, Inoue Y, Hirano T, Shiraishi Y, Chiba K, Tanaka H, Sanada M, Nishikawa Y, Amanuma Y, Ohashi S, Aoyama I, Horimatsu T, Miyamoto S, Tsunoda S, Sakai Y, Narahara M, Brown JB, Sato Y, Sawada G, Mimori K, Minamiguchi S, Haga H, Seno H, Miyano S, Makishima H, Muto M, Ogawa S. *Nature*. 2019 Jan;565(7739):312-317



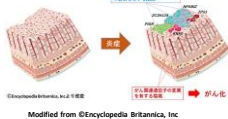
- **食道がんにおいて遺伝子変異が乳児期から獲得され、加齢とともに増加し70歳以上では全食道面積の40~80%が、がん遺伝子の変異をもった細胞で置き換わる。**
- **がんがなぜ高齢者に生じ、「飲酒」や「喫煙」によって促進されるのかを解明。**
- **がんが生ずる初期のメカニズムの解明の突破口を創った。**
- **がんの早期診断、予防、がんの死亡率の低減に資する大きな社会的意義。**

また、2020年1月には、潰瘍性大腸炎とがんの関係について2の成果が出ていました。

**Frequent mutations that converge on the NFKB1Z pathway in ulcerative colitis.**

Kakiuchi N, Yoshida K, Uchino M, Kihara T, Akaki K, Inoue Y, Kawada K, Nagayama S, Yokoyama A, Yamamoto S, Matsuura M, Horimatsu T, Hirano T, Goto N, Takeuchi Y, Ochi Y, Shiozawa Y, Kogure Y, Watatani Y, Fujii Y, Kim SK, Kon A, Kataoka K, Yoshizato T, Nakagawa MM, Yoda A, Nanya Y, Makishima H, Shiraishi Y, Chiba K, Tanaka H, Sanada M, Sugihara E, Sato TA, Maruyama T, Miyoshi H, Taketo MM, Oishi J, Inagaki R, Ueda Y, Okamoto S, Okajima H, Sakai Y, Sakurai T, Haga H, Hirota S, Ikeuchi H, Nakase H, Marusawa H, Chiba T, Takeuchi O, Miyano S, Seno H, Ogawa S. *Nature*. 2020 Jan; 577(7789): 260-265

- **損傷と修復を繰り返す「潰瘍性大腸炎」。その大腸上皮を総なめゲノム解析して、**がんに移行する人・移行しない人のメカニズム**を解明。**
- **潰瘍性大腸炎による上皮再構築メカニズムと発がんとの関係を解明**
- **大腸炎の長期罹患患者の大腸上皮は、大腸がんで認められる遺伝子変異の他に、炎症に関わるIL-17シグナル経路の遺伝子変異を獲得した細胞が増加し、直腸では全体の50~80%の面積がこれら**の遺伝子変異を有する細胞**によって置き換わっている。**
- **NFKB1Z (IκBζ) やZC3H12Aの変異を獲得した上皮細胞は発がんしがた**い。



**① 乳がんのライフストーリーの解明**

正常組織においてすでに、がん遺伝子変異を有するクローンの拡大が生じていることが近年報告され、がんの起源に関する知見が集まりつつあります。しかし、正常組織におけるこのようなクローンがどのような過程をへて臨床的にがんと診断されるにいたるかについては分かっていません。私共は、まず、単一乳腺細胞由来のオルガノイドの全ゲノム解析により、正常の乳腺細胞に蓄積する変異の速度を測定しました。そして、乳がんの外科切除標本から採取した多数の微小試料の全ゲノムシーケンスにもとづいて系統樹解析を行って、最初のドライバー変異の獲得から乳がんの診断にいたるまでの課程の全貌を明らかにすることに成功しました。ここに示す代表例の解析(図4)では、思春期前後に独立なder(1;16)転座を獲得した二つの乳腺上皮細胞に由来し、これらの起源の細胞の子孫が、30年の年月を経て、乳腺内に乳腺のほぼ全域に拡大し、この中から、複数のがんが生じていることが明らかになりました。この結果は、乳がんの発症メカニズムに新たな科学的知見を与えるとともに、超早期診断、予防法の確立に寄与すると考えております。

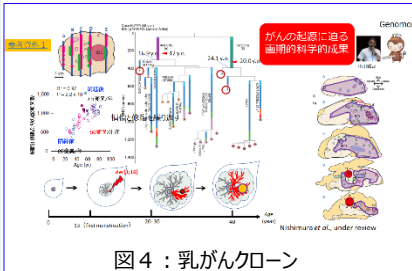


図4：乳がんクローン

**② クローン性造血の臨床予後への影響を解明—遺伝子変異とコピー数異常の統合的な知見—**

血液がんの前がん病変である**60歳を過ぎるころから皆に起こる血液の遺伝子変異「クローン性造血」**は、血液がんのみならず心血管疾患のリスク上昇とも関連し、注目が集まっています。過去の研究で、クローン性造血は遺伝子変異とコピー数異常(Copy number alterations: CNA)という2種類のゲノム異常で引き起こされることが報告されました。しかし、両方の異常を同時に調べた研究はなく、その関係性はよく分かっていませんでした。今回、約1万人の被験者についてクローン性造血の統合解析を実施しました。その結果、

➤ **クローン性造血では遺伝子変異とCNAが高頻度に共存すること、**

➤ **両者が共存すると血液腫瘍・心血管疾患のリスクが有意に上昇すること、**

などを解明しました。この研究結果は、クローン性造血に基づく予後予測の重要な指標となる事が期待されます。

**Combined landscape of single-nucleotide variants and copy number alterations in clonal hematopoiesis.**  
Saiki R, Momozawa Y, Nannya Y, Nakagawa MM, Ochi Y, Yoshizato T, Terao C, Kuroda Y, Shiraishi Y, Chiba K, Tanaka H, Nuida A, Imoto S, Matsuda K, Morisaki T, Murakami Y, Kamatani Y, Matsuda S, Kubo M, Miyano S, Makishima H, Ogawa S. *Nature Medicine*. 2021 Jul; 27(7):1239-1249  
DNMT3A, TET2, JAK2, TP53などの遺伝子では遺伝子変異とコピー数異常が頻繁に共存し、遺伝子の両方異常を来していると考えられました。これらの遺伝子の両方異常は血液がんでは既知の所見ですが、今回の研究によりクローン性造血の段階から存在することが判明しました。

**③ 上部尿路上皮がんの分子分類と新規分子診断—尿中遺伝子変異の検出で高精度の診断が可能**

上部尿路上皮がんは腎盂や尿管の尿路上皮に発生する予後不良ながんであり、その希少さから遺伝学的な背景が十分に解明されていませんでした。また、上部尿路上皮がんは診断が困難なため治療の遅れにつながることもあり、新たな診断方法の開発が待たれていました。上部尿路上皮がんの腫瘍検体及び術前に採取した尿を用いて大規模なゲノム解析を行い、

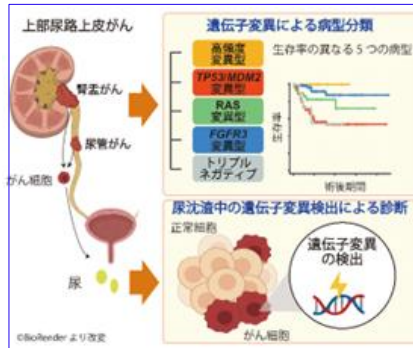
➤ **上部尿路上皮がんは遺伝子変異に基づき、異なる生存率を示す5つの分子病型に分類できること、**

➤ **術前の尿中にはがん組織と同一の遺伝子異常が認められ、上部尿路上皮がんの精度の高い診断が可能となること、**

を証明しました。今回の研究結果により、199例の上部尿路上皮がんの各病型のゲノム異常に基づく治療選択や、尿を用いた簡便かつ精度の高い診断が可能となり、治療の向上が期待されます。

**Molecular classification and diagnostics of upper urinary tract urothelial carcinoma.**

Fujii Y, Sato Y, Suzuki H, Kakiuchi N, Yoshizato T, Lenis AT, Maekawa S, Yokoyama A, Takeuchi Y, Inoue Y, Ochi Y, Shiozawa Y, Aoki K, Yoshida K, Kataoka K, Nakagawa MM, Nannya Y, Makishima H, Miyakawa J, Kawai T, Morikawa T, Shiraishi Y, Chiba K, Tanaka H, Nagae G, Sanada M, Sugihara E, Sato TA, Nakagawa T, Fukuyama M, Ushiku T, Aburatani H, Miyano S, Coleman JA, Homma Y, Solit DB, Kume H, Ogawa S. *Cancer Cell*. 2021 Jun 14;39(6):793-809.e8



上部尿路上皮がんにおいて、5.5%の症例は他の症例と比べて著しく遺伝子変異の多い状態でした。これらの多くが大腸がん等で頻りにみられる、ミスマッチ修復遺伝子の異常に起因するものでした。また、どのような遺伝子に高頻度に変異が起きているかを解析し、26個の重要なドライバー遺伝子を同定しました。

次に、膀胱がんの既存データを用いて腎盂、尿管、膀胱の各部位のがんにおける遺伝子異常の比較を行ったところ、一部の遺伝子異常において部位毎に頻度の差があることがわかりました。例えば、RAS(HRAS, KRAS, NRAS)の変異は腎盂がんに多く、ERBB2の変異は膀胱がんに多く、KDM6A、TERTプロモーターの変異は膀胱がん及び腎盂がんで多く認められました。一方、尿管がんではKMT2Dの変異が高頻度変異の症例が多く認められました。これらのことから、同じ尿路上皮という細胞に発生するがんで、解剖学的部位により発がんに関わる遺伝子が異なる可能性が示唆されました。

④ コヒーシン遺伝子変異による白血病発症の機序を解明—染色体の3次元構造の異常が発がん—

急性骨髄性白血病や骨髄異形成症候群は血液のがんで、現在でも根治の難しい難治性の病気です。正常な血液細胞が遺伝子変異を獲得することで白血病を発症しますが、近年、本研究グループより新規コヒーシン遺伝子変異（コヒーシンは、複製された染色体を娘細胞に均等に分離するために必須な過程に中心的な役割を果たすタンパク質複合体です）が10-20%の白血病に認められることを報告しました。しかし、コヒーシン変異がどのように白血病を引き起こすかは、十分に分かっていませんでした。さらに、がん細胞にしばしば複数の遺伝子変異が蓄積していることも分かってきましたが、複数の変異が発がんに関わる機序についても、不明な点が多く残されていました。3,000症例以上の白血病の大規模なゲノム解析、ノックアウトマウスの解析、更に次世代シーケンサーを活用した統合的エピジェネティック解析（DNAの化学修飾を調べる解析）を実施しました。その結果、

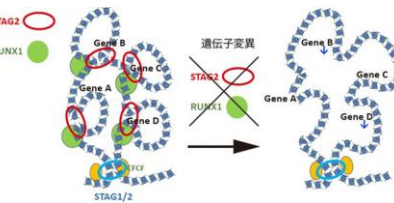
- コヒーシンの一つであるSTAG2遺伝子変異が、RUNX1遺伝子変異と高頻度に共存すること、
- これらの遺伝子の両方に異常が生じることで、骨髄異形成症候群を発症すること、
- コヒーシン変異が染色体の3次元構造の変化などのエピジェネティックな異常を

引き起こし、更にRUNX1変異が加わることでその異常が加速し、発がんに至ること、

などを解明しました。GPUも用いて解析しました。

Combined Cohesin-RUNX1 Deficiency Synergistically Perturbs Chromatin Looping and Causes Myelodysplastic Syndromes.

Ochi Y, Kon A, Sakata T, Nakagawa MM, Nakazawa N, Kakuta M, Kataoka K, Koseki H, Nakayama M, Morishita D, Tsuruyama T, Saiki R, Yoda A, Okuda R, Yoshizato T, Yoshida K, Shiozawa Y, Nannya Y, Kotani S, Kogure Y, Kakiuchi N, Nishimura T, Makishima H, Malcovati L, Yokoyama A, Takeuchi K, Sugihara E, Sato TA, Sanada M, Takaori-Kondo A, Cazzola M, Kengaku M, Miyano S, Shirahige K, Suzuki HI, Ogawa S. *Cancer Discov.* 2020 Jun;10(6):836-853.



- STAG2/RUNX1変異による染色体ループ破綻および特定の遺伝子発現低下の模式図
- Stag2とRunx1両遺伝子ノックアウトマウスは骨髄異形成症候群を発症する
- コヒーシン遺伝子に変異が生じると、染色体の3次元構造の変化などのエピジェネティックな異常が生じる。

⑤ ゲノム解析から急性赤白血病の変異プロファイルと治療標的を解明—特定の遺伝子変異群の組み合わせと、特徴となる遺伝子の増幅が鍵—

急性骨髄性白血病と一言にいっても組織学的にも様々で予後も均一ではありません。この研究で急性赤白血病がどのような遺伝子異常を獲得して発症するかが判明するとともにその中でも予後が極めて不良な一群が存在し、それに対し治療標的を見つけマウスモデルでその治療標的の阻害薬が奏効することが明らかとなりました。

124例の急性赤白血病（AEL）に対し次世代シーケンサーを用いた包括的なゲノム解析を行い、変異プロファイルからAELは4つのサブグループに分類されることを明らかにしました。そのなかで予後不良のTP53変異群ではJAK2、EPOR、ERG、ETS2を含むDNA増幅が他急性骨髄性白血病（AML）と比較しAELで高頻度に認めました。他のサブグループでもAELに特徴的な変異を認め、単独の変異ではなく特定の変異の組み合わせがAELを特徴づけると考えました。

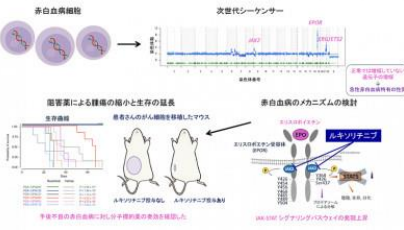
JAK2、EPORのDNA増幅は発現上昇に

寄与し共にJAK-STATシグナリング経路に関与することから治療標的になりうると考え、TP53変異とJAK2、EPOR増幅をもつ検体を用いて患者由来のマウスモデルを作成し、JAK阻害薬（ルキソリチニブ）により腫瘍細胞の増殖が抑制され、生存を延長されることを明らかにしました。

AELの変異プロファイルと表現形の関連を明らかにするものであり、また、治療のさらなる最適化に資するものであると考えます。

Amplified EPOR/JAK2 Genes Define a Unique Subtype of Acute Erythroid Leukemia.

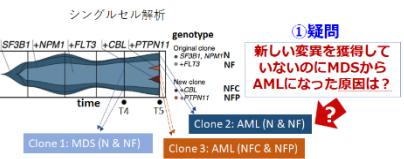
Takeda J, Yoshida K, Nakagawa MM, Nannya Y, Yoda A, Saiki R, Ochi Y, Zhao L, Okuda R, Qi X, Mori T, Kon A, Chiba K, Tanaka H, Shiraiishi Y, Kuo MC, Kerr CM, Nagata Y, Morishita D, Hiramoto N, Hangaiishi A, Nakazawa H, Ishiyama K, Miyano S, Chiba S, Miyazaki Y, Kitano T, Usuki K, Sezaki N, Tsurumi H, Miyawaki S, Maciejewski JP, Ishikawa T, Ohyashiki K, Ganser A, Heuser M, Thol F, Shih LY, Takaori-Kondo A, Makishima H, Ogawa S. *Blood Cancer Discov.* 2022 Sep 6;3(5):410-427.



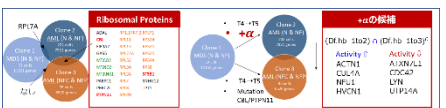
⑥ 骨髄異形成症候群(MDS)から急性骨髄性白血病(AML)へ至るクローン変化の謎

シングルセル解析で、骨髄異形成症候群（MDS）から急性骨髄性白血病(AML)に至るクローン変化の謎に迫りました。

新しい変異を獲得していないのにMDSからAMLになったことが疑問点です。



そこでクローンごとにネットワークを推定し、ハブ遺伝子に注目してその重なりをみるといくつかのサジェスションが浮かびあがってきました。



しかし、MDSからAML移行の原因探査には各クローンの違いを特徴付けるネットワーク探査技術が必要と考え、データ解析の方法論開発をしました。



条件における遺伝子名と発現量	各遺伝子に関する既存の文献による指摘	条件に関して、既存の文献による指摘から考えられる可能性
ZNF516が低発現	EGFRの発現を抑えることに寄与。EGFR経路を抑制することで細胞増殖を抑える機能がある。	ZNF516の発現量が低いことは、ゲフィチニブが標的としているEGFR経路が活性化している可能性あり。
E2F6が低発現	E2F1とE2F2は細胞周期経路を回して細胞増殖を促進。E2F6はその機能を阻害。	E2F6の発現量が低いことは、細胞増殖を促進している可能性あり。(EGFR経路の対象外)
EMX1が低発現	Wnt経路(細胞周期経路を促進)を抑制することで知られる。	EMX1の発現量が低いことは、Wnt経路が抑制されず、細胞増殖を促進している可能性あり。(EGFR経路の対象外)

「Wide Learning」を用いた「発見するAI」は、薬剤開発の際の重要な興味であるバイオマーカー探索に資する技術となるでしょう。新規薬剤開発の鍵は、効果が期待される患者さんを同定して臨床試験を行うことです。たとえばイマチニブを慢性骨髄性白血病ではなく大腸がんに対して治験を行えば必ず失敗します。そこで、薬剤の効果が期待できる人を予測するマーカー（遺伝子変異や遺伝子発現情報）があれば、その人だけを対象として臨床試験を行うことで、臨床試験の経費を大きく削減し、その成功確率を高めることができます。この点から、この技術は製薬メーカー等が大いに興味を持つことが期待されます。

プレスリリース: <https://pr.fujitsu.com/jp/news/2022/03/7.html>  
<https://www.tmd.ac.jp/press-release/20220307-1/>  
 \*Wide Learning技術: 複数の網羅的な列挙により、判断根拠の説明や知識発見が可能なAI技術。  
 正解が少ないデータでも高精度に学習するAIの新技術「Wide Learning」を開発(2018年9月19日プレスリリース)  
<https://pr.fujitsu.com/jp/news/2018/09/19-2.html> Home - Hello, Wide Learning! <https://widelearning.labs.fujitsu.com/ja/index.html>

## 終わりに

2019年、「がん研究にスパコンは不可欠」ということが常識となっていた。さらに「大規模

データ解析」と「AIの活用」が、がんの理解と社会への還元には不可欠という認識にいたり本課題を始めました。「超高齢社会のなかで、がんの起源(卵)は何か、いつ生まれているのか、どのようにがんは発症するのか、多様性を獲得したがんはどのように転移・治療抵抗性を獲得していくのか、病態との関連は」。このがんの中心課題に取り組み、人馬一体となって世界トップの成果を創出できたと考えています。今後、医学・生命系の研究者等が、より親しみやすく大規模な計算リソース(計算ノード、ストレージ、ネットワーク等)にアクセスし、もしくは無意識に利用する環境が到来するものと思います。Society 5.0の時代にスパコンがインソレイトされてはならない。

## 1. 代表機関

国立大学法人東京医科歯科大学  
 研究課題開発責任者  
 M&Dデータ科学センター 特任教授 宮野悟

## 2. 協力機関

国立大学法人京都大学  
 協力機関代表者  
 大学院医学研究科 教授 小川誠司  
 愛知県がんセンター  
 協力機関代表者  
 研究所システム解析分野 分野長 山口類

## 3. 連携機関

(株)富士通研究所  
 連携機関連絡担当者  
 富士通株式会社フェロー 岡本青史  
 国立大学法人東京大学  
 連携機関連絡担当者  
 医科学研究所 准教授 片山琴絵

Karolinska Institutet  
 連携機関連絡担当者  
 Department of Medicine/Center for Hematology and Regenerative Medicine  
 Visiting Professor 小川誠司

## 4. 事業参加者

事業参加者(代表機関)  
 東京医科歯科大学M&Dデータ科学センター  
 統合解析分野 特任教授 宮野悟  
 准教授 長谷川嵩矩  
 助教授 伊東聡  
 特任助教 角田将典  
 特任助教 小川弥穂  
 特任研究員 田中洋子  
 Sungshing Womenn's University  
 (M&Dデータ科学センター客員教授)  
 Heewon Park

事業協力者(協力機関)  
 京都大学大学院医学研究科  
 腫瘍生物学講座 教授 小川誠司  
 特定教授 南谷泰仁  
 助教 越智陽太郎  
 他11名  
 愛知県がんセンター研究所  
 システム解析分野 分野長 山口類

連携参加者(連携機関)  
 富士通株式会社フェロー 岡本青史  
 富士通株式会社 富士通研究所  
 プロジェクトマネージャ 丸橋弘治  
 東京大学医科学研究所  
 ヒゲム解析センター 准教授 片山琴絵  
 先端医療研究センター 教授 南谷泰仁  
 (2023年3月1日時点)



2023年3月16日小川研究室にて(京大医) 田中洋子・伊東聡・小川誠司・垣内伸之・宮野悟

スーパーコンピュータ「富岳」成果創出加速プログラムについて  
 (理化学研究所計算科学研究センターより抜粋)  
<https://www.r-ccs.riken.jp/jp/fugaku/promoting-researches>

スーパーコンピュータ「富岳」成果創出加速プログラムは、「富岳」を用いた成果を早期に創出することを目的として文部科学省が設置しました。①人類の普遍的課題への挑戦と未来開拓、②国民の生命・財産を守る取組の強化、③産業競争力の強化、④研究基盤の4領域について課題の選定が行われ、19課題が選定されています。選定された課題は、「富岳」の計算資源を優先的に無償で使用し、速やかな成果創出を目指します(2020~2022年度)。

課題名: 大規模データ解析と人工知能技術によるがんの起源と多様性の解明  
 課題番号: JPMXP1020200102

課題ID  
 2022年度: hp220163  
 2021年度: hp210167  
 2020年度: hp200138

# Information



文部科学省「富岳」成果創出加速プログラム  
課題名:大規模データ解析と人工知能技術によるがんの起源と多様性の解明  
ニュースレター No. 6  
発行日★2023年2月28日  
課題代表者★宮野 悟

- 東京医科歯科大学M&Dデータ科学センター 統合解析分野
- 郵便物宛先: 〒113-8510 東京都文京区湯島1-5-45
- 宅配便宛先: 〒101-0062 東京都千代田区神田駿河台2-3-10  
駿河台キャンパス22号館5階

- TEL: 03-5280-8620 FAX: 03-5280-8632
- E-mail: [mdsc.dsc@tmd.ac.jp](mailto:mdsc.dsc@tmd.ac.jp)
- HP: <https://odcla.mddsc.jp>