

【 技術 】

2022年3月7日

富士通株式会社

国立大学法人東京医科歯科大学

スーパーコンピュータ「富岳」と「発見するAI」で、がんの薬剤耐性に 関わる未知の因果メカニズムを高速に発見する新技術を開発

1,000兆通りの可能性から患者の遺伝子の特徴を一日以内で絞り込み、効率的な薬剤開発に寄与

富士通株式会社(注1)(以下、富士通)と国立大学法人東京医科歯科大学(注2)(以下、東京医科歯科大学)は、富士通が開発した現場のデータから新たな発見の手掛かりを提示する技術「発見するAI」をスーパーコンピュータ「富岳」(注3)上に実装し、従来は実行困難であった2万変数のデータを一日以内で超高速計算することが可能で、1,000兆通りの可能性から未知の因果を発見できる技術(以下、本技術)を開発しました。

両者が、がん医療と創薬の現場課題である抗がん剤の薬剤耐性(注4)を分析するために、がんの細胞株から得られた遺伝子発現量(注5)データに本技術を適用した結果、これまでの研究成果では知られていない、肺がん治療薬の耐性の原因を示唆する遺伝子の新たな因果メカニズムを抽出することに成功しました。これにより、患者一人ひとりに対応した効果的な抗がん剤創薬の実現に向けて本技術の活用が期待できます。

両者は今後、さらに時間軸や位置データなどの多くのデータを組み合わせた多層的、総合的な分析を実施し、薬効メカニズムやがんの起源の解明などの重要課題における発見の手掛かりを提示する研究を加速するとともに、医療や創薬分野だけでなく、複雑に交錯する因子を発見して意思決定を行うようなマーケティングやシステム運用、生産現場など、活用分野の拡大に取り組んでいきます。

【 開発背景と課題 】

がんの分子標的薬(注6)は、患者への投与を続けると、その薬剤に対して耐性を獲得したがん細胞が増殖し、再発することがあります。このように、複数のドライバー変異(注7)を獲得した細胞群が変幻自在に異常な振る舞いをするがん耐性獲得のメカニズム解明には、精緻なデータと解析技術が不可欠です。薬剤開発や既存薬の再配置(注8)の臨床治験では、薬剤効果が期待される患者のおおまかな同定が重要ですが、個人や臓器における遺伝子やその発現量などにより薬剤効果が異なり、複数の遺伝子の発現量を組み合わせたパターン数は1,000兆通りを超える(注9)ほど膨大なため、医療や創薬の進歩の鍵を握る効率性が高い探索技術への期待が高まっています。

富士通の開発技術「発見するAI」は、判断根拠の説明や知識発見が可能なAI技術「Wide Learning」(注10)を用いて、特徴的な因果関係を持つ条件を網羅的に抽出する技術として開発(注11)されましたが、ヒトの全2万個の遺伝子を対象とした網羅的な探索には通常の計算機で4,000年以上かかる試算であったため、処理の高速化が課題となっていました。

【 開発した技術 】

両者は、プロセッサ間とプロセッサ内の並列性を特長とするスーパーコンピュータ「富岳」上にヒトの全遺伝子を実用的な時間で分析できるよう、条件探索と因果探索を行うアルゴリズムを並列化し、実装することで、計算性能を最大限引き出しました。さらに、「発見するAI」を活用し、統計情報に基づき薬剤耐性を生み出す条件となりうる有望な遺伝子の組み合わせを抽出することで、一日以内で網羅的な探索を実現する技術を開発しました。

【 効果 】

スーパーコンピュータ「富岳」上で本技術を用いて DepMap(注12)の公開データを実行した結果、ヒトの全遺伝子に対して条件と因果関係を一日以内で網羅的に探索し、肺がんの治療薬に耐性を持つ原因となる遺伝子を特定することに成功

しました。肺がんの治療薬などで用いられるゲフィチニブは、EGFR 経路(注 13)の活性化を押さえてがんの進行を抑える分子標的薬であり、耐性が発生する現象が知られています。DepMap のがん細胞株約 300 種の遺伝子発現量データとゲフィチニブの感受性、耐性データを解析し、ゲフィチニブが効かないがん細胞株の条件とメカニズムを網羅的に探索した結果、*ZNF516*, *E2F6*, *EMX1* の 3 つの転写因子(注 14)の発現量が低いという条件を探り当て、その条件を満たす肺がん細胞株では、転写因子 *SP7* と *PRRX1* をトリガーとするメカニズムが耐性を生み出す因果として新たに示唆されました。

| 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) |

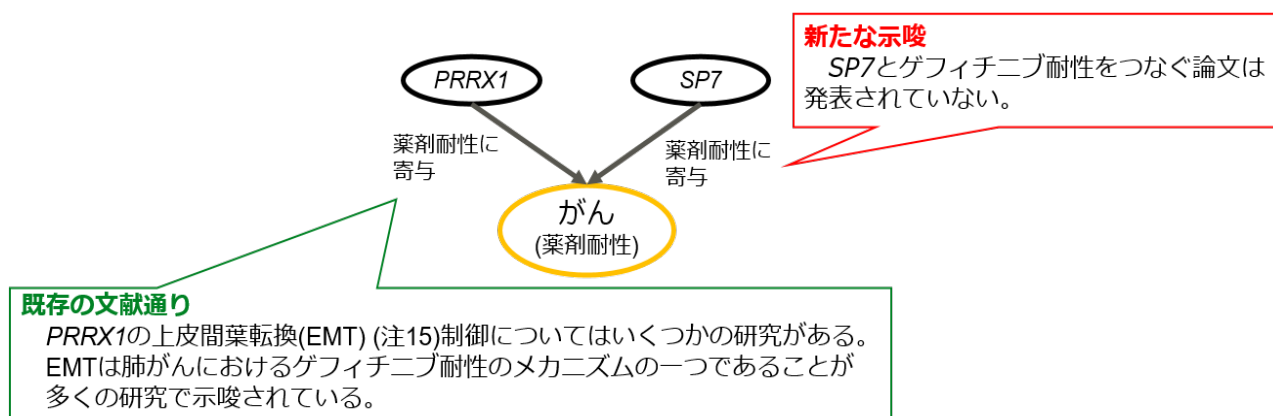


図: 本技術により示唆されたゲフィチニブ耐性の条件と因果関係

なお、本技術開発は、スーパーコンピュータ「富岳」成果創出加速プログラム(注 16)において、東京医科歯科大学と京都大学、富士通などが参画する課題「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」解決に向けた取り組みの一環として実施しました。

【京都大学 医学研究科 教授 小川誠司氏のコメント】

「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。新規薬剤開発の成功の鍵は、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、本課題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。「富岳」に実装されている点も期待感を大きくしています。

【今後】

両者は今後、遺伝子の発現量や変異のデータに加え、時間軸や位置データを組み合わせた多層的、総合的な分析を実施し、薬効メカニズムやがんの起源の解明などの重要課題における発見の手掛かりを提示する研究を加速するとともに

に、創薬、医学分野における実験研究の現場と連携していきます。東京医科歯科大学は、本研究で開発した技術を用いて、がんをはじめとする難病の攻略法の研究を推進します。さらに、富士通は、マーケティングやシステム運用、生産現場において複雑に交錯する因子を発見し、意思決定を支援する取り組みを進めます。

【謝辞】

本研究は、文部科学省のスーパーコンピュータ「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」(JPMXP1020200102)の一環として実施されたものです。また、本研究の一部は、スーパーコンピュータ「富岳」の計算資源の提供を受け、実施しました(課題番号:hp200138, hp210167)。

【商標について】

記載されている製品名などの固有名称は、各社の商標または登録商標です。

【注釈】

- (注1) 富士通株式会社:本社 東京都港区、代表取締役社長 時田 隆仁。
- (注2) 国立大学法人東京医科歯科大学:所在地 東京都文京区、学長 田中雄二郎。
- (注3) スーパーコンピュータ「富岳」:スーパーコンピュータ「京」の後継機として理化学研究所に設置された計算機。2020年6月から2021年11月にかけて世界のスーパーコンピュータに関するランキングの主要4部門で4期連続で1位を獲得するなど、世界トップの性能を持つ。2021年3月9日から共用開始。
- (注4) 薬剤耐性:薬剤の投与を続けているうちに、その薬剤の効果が弱まってしまう現象。
- (注5) 遺伝子発現量:DNAから複製されたRNA(転写により一部のDNA配列を鋳型として合成されたDNAと同じ核酸)の量。
- (注6) 分子標的薬:病気の原因となっている分子(タンパク質や遺伝子等)だけに作用するように設計された治療薬。
- (注7) ドライバー変異:がんの発生や悪化の直接的な原因となる遺伝子変異。
- (注8) 既存薬の再配置:ある疾患の治療薬として開発、認可された既存薬を、別の疾患治療用に適用すること。
- (注9) 1,000兆通りを超える:がんに関係することが判明している主な50個の遺伝子の組み合わせに限定し、各遺伝子の発現量を2分類(遺伝子の発現の「高い」「低い」など)とした場合でも、条件数は2の50乗となり、1,000兆通りを超える。
- (注10) 「Wide Learning」:公式サイト「Hello, Wide Learning!」(<https://widelearning.labs.fujitsu.com/>)
- (注11) 特徴的な因果関係を持つ条件を網羅的に抽出する技術として開発:「ヒトやモノなどのデータの一つひとつが持つ特徴的な因果関係を発見する技術を開発」(2020年12月17日プレスリリース)(<https://pr.fujitsu.com/jp/news/2020/12/17.html>)
- (注12) DepMap:米Broad Instituteが提供する約600種の間違ったがん細胞株に対する約4500の薬剤の感受性、耐性のデータ。がん細胞株の変異情報と全遺伝子の発現データ等が含まれる。(<https://depmap.org/portal/>)
- (注13) EGFR経路:上皮成長因子(EGF)という生体分子の刺激を受け取って、細胞増殖を活性化するメカニズム。
- (注14) 転写因子:遺伝子の転写(RNAの合成)を制御する遺伝子。
- (注15) 上皮間葉転換(EMT):上皮系の細胞から遊走性を獲得した間葉系の細胞に転換する現象で、がんの転移に関与していることが示唆されている。
- (注16) スーパーコンピュータ「富岳」成果創出加速プログラム:2020年5月より開始した、文部科学省が「富岳」を用いて成果を早期に創出することを目的としたプログラム。

【関連リンク】

・東京医科歯科大学と富士通研究所、「富岳」を用いてがんの遺伝子ネットワーク分析を1日以内に実現(2020年11月10日プレスリリース)(<https://pr.fujitsu.com/jp/news/2020/11/10.html>)

以上

《本件に関するお問い合わせ》

富士通株式会社

研究本部 人工知能研究所

電話:044-754-2328 (直通)

E-mail: fj-widelearning@dl.jp.fujitsu.com

東京医科歯科大学

総務部総務秘書課広報係

〒113-8510 東京都文京区湯島1-5-45

電話 03-5803-5833

Fax 03-5803-0272

メール kouhou.adm@tmd.ac.jp