

厚生労働記者会・厚生日比谷クラブ
文部科学記者会・科学記者会 同時発表

2022 年 5 月 30 日
横浜市立大学
国立研究開発法人国立がん研究センター
Genomedia 株式会社

病理組織画像からがんの遺伝子異常を予測する 新規スクリーニングシステム Virtual Sequencing を開発 ～病理学と AI の協働による次世代病理学の創生を目指して～

横浜市立大学大学院医学研究科分子病理学 藤井誠志 主任教授、国立がん研究センター東病院 吉野孝之 消化管内科長らの研究グループは、Genomedia 株式会社など企業との共同研究により、H&E（ヘマトキシリン・エオジン）染色^{*1}の大腸がんの病理組織画像から、がんの遺伝子異常を検出する高性能な検出器の作製に成功しました。

本研究では、病理医ががんにおける病理形態学的特徴を目視で抽出し、DL（ディープラーニング）^{*2}を使用して個別に学習させることにより、様々ながんの病理形態学的特徴を検出する検出器を作製しました。さらに、これらの検出器を用いて個別の病理形態学的特徴を有すると予測された病理組織画像の組み合わせに対して DL を使用することによって、高性能な遺伝子異常予測モデルを構築しました。これにより、病理組織画像からがんの遺伝子異常を予測できる新しいスクリーニングシステム「Virtual Sequencing」^{*3}の実用化が進むことが期待されます。がんの病理形態像は遺伝子異常、エピゲノム^{*4}変化等の合算から作られますが、病理形態像から遺伝子異常を検出することは一般に困難であり、次世代病理学に繋がる研究成果です。

本研究成果は、米医学術誌「Clinical Cancer Research」に掲載されました。（2022 年 4 月 25 日オンライン公開）

研究成果のポイント

- DL を用いて作製された病理組織学的特徴の検出器を使用し、これまでにない新規の方法で遺伝子異常の検出器を作製しました。
- 33 個の病理組織学的特徴のうち、12 個と 27 個の特徴に対して AUC^{*5} がそれぞれ >0.90 と >0.80 を達成しました。
- 2 つの病理形態学的特徴の組み合わせによる、大腸がんの遺伝子異常予測の 2 つの独立したコホートを用いた検証において、*BRAF*^{V600E}（AUC 0.851 と 0.859）と MSI-H^{*6}（AUC 0.923 と 0.862）に対して高い精度で遺伝子異常を検出しました。

研究背景

日常の病理診断では、生検または外科手術によって採取された病理標本はホルマリン固定された後にパラフィンに包埋されたブロック標本が作製されます。そしてそのブロック

標本から薄切された未染色切片が H&E 染色され、それを顕微鏡下で観察し、病理医が病理診断を行います。がんの個別化医療（precision medicine）では、がん細胞、がん組織を含むブロック標本から薄切された未染色切片から核酸を抽出し、がん遺伝子パネル検査を用いて遺伝子異常が検出されます。そして、その結果に基づいて分子標的治療の可否が決定されます。がんの病理形態像は遺伝子異常、エピゲノム変化等の合算から作られますが、がんの個別化医療の実践により、病理形態像と遺伝子異常・エピゲノム変化についての情報を紐づけた試料が蓄積され、これらの関係を検討する機会が得られるようになりました。

研究内容

産学連携全国がんゲノムスクリーニング「SCRUM-Japan」*7 の GI-SCREEN-Japan（現 MONSTAR-SCREEN）のコホート*8 から、*BRAF*^{V600E} 変異やマイクロサテライト不安定性 (MSI) に関する次世代シーケンス配列 (NGS) データを持つ大腸がん患者の 1,657 の病理組織画像（患者一人に対して一画像）を入手し、986 画像の探索的コホート、248・423 画像の検証コホートに分けました。

病理医がまず H&E 染色画像を目視で見比べて、どのような形態像が遺伝子異常に関連しているのかを注意深く観察して見出し、従来の病理組織学的診断に用いる病理組織学的特徴、これまで遺伝子異常との関連が示唆されている病理組織学的特徴、そして全く新規で従来の病理組織学的診断には用いない病理組織学的特徴を選抜しました。

次に、一人の病理医が病理組織画像に対して、33 個の病理組織学的特徴について高精度で、約 105 万に及ぶアノテーション（ラベル付け）を行い、第一段階のがんの病理形態学的特徴を予測するモデルを構築しました。

そして、この検出器を用いて画像を分類して得られた病理組織画像の組み合わせに対して、その特徴（遺伝子異常）との関係に関して DL を使用して遺伝子異常予測モデルを構築し、検出器を作製しました。その結果、病理画像から病理組織学的特徴を抽出し、遺伝子異常を高精度に検出するアルゴリズムが確立され、新しいスクリーニングシステムの開発に成功しました（図 1）。

この 2 段階での検出器作製のアプローチは、本研究で初めて用いられたこれまでにない新規の方法です。この方法では、病理形態学的特徴が予測された画像の組み合わせの中から、DL によって遺伝子異常との関係の存在が示唆される画像の組み合わせを探索することで、元来ブラックボックスに陥りかねない DL を用いながらも、作製した検出器が利用する病理形態学的特徴領域の組み合わせによって説明できるようになります。

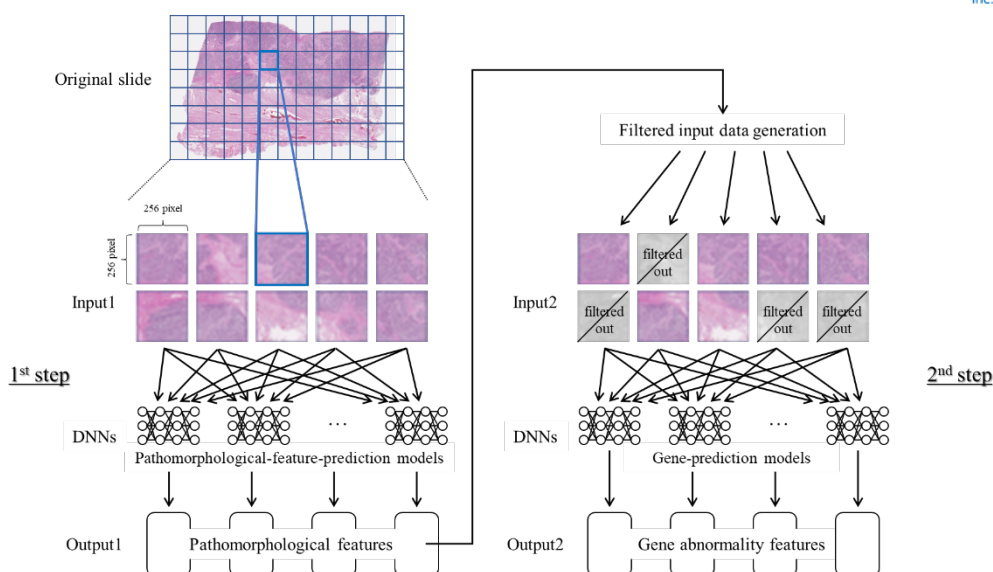


図1 遺伝子異常検出アルゴリズム作成のアプローチ (Clinical Cancer Research より改変)

本研究で対象とした **BRAF^{V600E}** および **MSI-H** は、**BRAF** 阻害剤または免疫チェックポイント阻害剤によるがん治療のために確認される、よく知られたバイオマーカーですが、本研究では、2つの病理組織学的特徴を使用する遺伝子異常検出アルゴリズムを用いた、**BRAF^{V600E}** および **MSI-H** に対する検出器により、2つの独立した検証コホートで高い AUC 値を再現しました (図 2)。特異的な遺伝子異常をもつ結腸・直腸がんの患者さんに、最も適した治療を選択する有用なスクリーニングツールとして役立つ可能性があります。

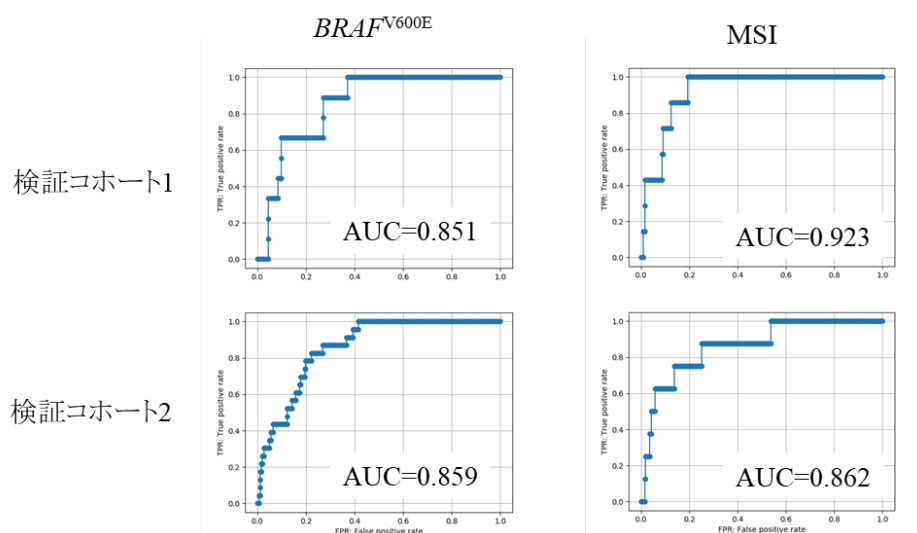


図2 **BRAF^{V600E}** および **MSI-H** の病理学的特徴と AUC の関係 (Clinical Cancer Research より改変)

病理画像から遺伝子異常を検出するこのスクリーニングシステムを本研究グループでは「Virtual sequencing (バーチャル シーケンシング)」と提唱し (図 3)、病理組織学的特徴数を増やすことによって検出器の性能向上をさせる等、実際の医療に用いることを意識したアプローチを行っています。

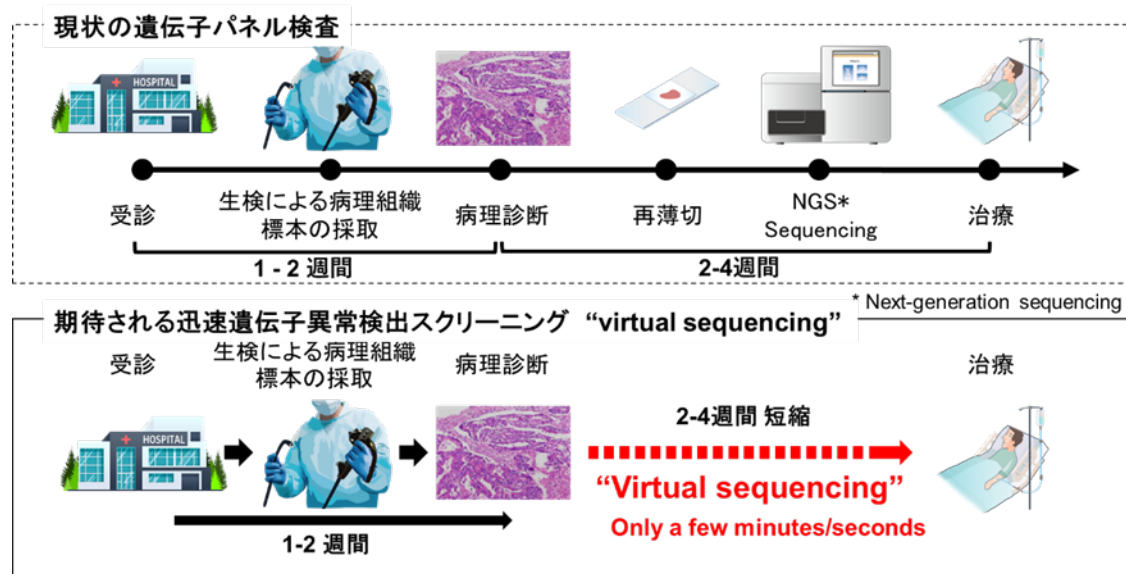


図3 Virtual sequencing の概要

今後の展開

遺伝子パネル検査などの従来の検査を行う前に、遺伝子異常を予測できるこの新しい方法は、近い将来の大腸がんなど、がんに対する治療戦略の迅速な決定に利活用できることが期待されます。病理学の面からは、病理医とAIが協働することによって可能になる“次世代病理学の創生”を目指し、このプロジェクトを引き続き遂行していきます。

論文情報

タイトル： Rapid screening using pathomorphological interpretation to detect *BRAF*^{V600E} mutation and microsatellite instability in colorectal cancer

著者： Satoshi Fujii*, Daisuke Kotani, Masahiro Hattori, Masato Nishihara, Toshihide Shikanai, Junji Hashimoto, Yuki Hama, Takuya Nishino, Mizuto Suzuki, Ayatoshi Yoshidumi, Makoto Ueno, Yoshito Komatsu, Toshiki Masuishi, Hiroki Hara, Taito Esaki, Yoshiaki Nakamura, Hideaki Bando, Tomoyuki Yamada, Takayuki Yoshino

*Corresponding author.

掲載雑誌： Clinical Cancer Research

DOI： 10.1158/1078-0432.CCR-21-4391

用語説明

*1 H&E (Hematoxylin & Eosin/ヘマトキシリン・エオジン) 染色：実際の病理診断に用いる病理標本に対する染色方法であり、それをスキャナーで取り込んだ病理組織画像のこと。

*2 DL（ディープラーニング）：経験からの学習により自動で改善するコンピュータアルゴリズムである機械学習技術の 1 つで、対象の全体像から細部までの各々の粒度の概念を階層構造として関連させて学習する手法。

*3 Virtual Sequencing（VSQ）：本研究を通じて提唱された、がんの病理形態学的な特徴から特定の遺伝子の異常を予測する手法。

*4 エピゲノム：ゲノム（DNA の塩基配列）に加えられた後天的な修飾で、主に DNA のメチル化やヒドロキシメチル化、ヒストンタンパク質の修飾（メチル化、アセチル化、リン酸化など）が知られている。

*5 AUC：ROC 曲線を作成した時に、グラフの曲線より下の部分の面積を AUC（Area Under the Curve）と言う。AUC は 0 から 1 までの値をとり、値が 1 に近いほど判別能が高いことを示す。

*6 MSI-H：高頻度マイクロサテライト不安定性（MSI-High）は、ゲノムの DNA からなる短い文字列が何度も繰り返す「マイクロサテライト」とよばれる部分の繰り返し回数に異常が起こった状態で、がんが発生しやすい状態と考えられている。

*7 SCRUM-Japan：2013 年に開始した肺がん患者さんを対象とした LC-SCRUM-Japan（現：LC-SCRUM-Asia）と、2014 年に開始した消化器がん患者さんを対象とした GI-SCREEN-Japan（現：MONSTAR-SCREEN）が統合した、産学連携がんゲノムスクリーニングプロジェクト。固形がん患者さんを対象に、がんの遺伝子異常を調べるプロジェクトであり、2015 年 2 月の設立以降、約 3 万例を超える進行固形がん患者さんが研究に参加。本プロジェクトの成果として、既に 11 品目の医薬品と 9 種類の体外診断薬の薬事承認を取得している。全国から 200 を超える医療機関と 17 社の製薬企業や検査会社が参画し、アカデミアと臨床現場、産業界が一体となって、日本のがん患者さんの遺伝子異常に合った治療薬や診断薬の開発を行っている。

（参考）SCRUM-Japan ホームページ <http://www.scrum-japan.ncc.go.jp/index.html>

*8 コホート：共通の因子を持つ観察対象集団のこと。