

がんゲノムデータ解析：臨床現場への実装

Introduction of the cancer genome data analysis service

毛利 涼* 岡村 容伸* 野原 祥夫* 谷嶋 成樹*

Ryo Mori, Yasunobu Okamura, Sachio Nohara, Shigeki Tanishima

近年のがん遺伝子研究の発展により治療法は化学療法薬からがん遺伝子やがん遺伝子変異をターゲットにした分子標的薬中心の時代に向けて急速に変化している。さらに最新の治療法である免疫チェックポイント阻害剤はがん細胞にみられる高頻度の遺伝子変異に対する免疫反応を利用している。そのため、がんの遺伝子変異状態を検査することで患者にこれまでにない最適な治療を提案できる可能性がでてきた。当社では、これらの臨床現場からのニーズに対応できる「がんゲノムデータ解析サービス」を構築した。このサービスは、これまで生命科学基礎研究分野の支援事業で培ったバイオインフォマティクスを最大限活用し、治療方針の決定に必要な遺伝子変異情報を医師に対してわかりやすく、すばやく提供することが可能である。また、このサービスではインターネットから分離したセキュアな環境で自動的に解析を行うことができ、安全かつすみやかに情報を提供することを可能にしている。本稿では当社で提供するがんゲノムデータ解析サービスについて紹介を行う。

Cancer genome analysis plays an increasingly significant role in cancer treatment as a result of spreading of molecular-targeted therapy. One of the new concepts in cancer treatment “Immune checkpoint blockade” helps immunity cells kill cancer cells with many genomic mutations. Examination of genomic mutation in cancer cells enables patients to receive the precise therapy. We built a new cancer genomic mutation analysis service to meet those needs. Our service is backed by our experiences of supporting the service for basic biology scientists using bioinformatics. Our service can provide the “Cancer driver gene-mutation reports” to medical doctors clearly, rapidly and securely because our service automatically analyzes the data in “secure” environment, which is isolated from the Internet. Here, we show our cancer genome data analysis service.

1. まえがき

厚生労働省の全国人口動態では、日本人の死因第1位として悪性新生物（＝がん）が報告され、その数は全体の29%に及んでいる⁽¹⁾。また一生のうち二人に一人ががんを発症するといわれており、国民病の一つとしてあげられる。がんの治療は外科手術や化学療法が主流であったが、近年の研究の発展により遺伝子をターゲットとした分子標的薬が利用できるようになりつつある。

がんとは、遺伝子に後天的変異が起き特定の細胞集団が異常な増殖を示す疾患である。がんにおいては後天的に以下のような段階で変異が蓄積することによって発症することが知られている（多段階発がん：Multistep carcinogenesis）⁽²⁾。以下にその段階を示す（図1）。

- (1) 特定の細胞集団でがんを抑制する遺伝子（がん抑制遺伝子）に変異が起り、細胞増殖の抑制機構が機能を喪失する。この状態は車のブレーキが効かない

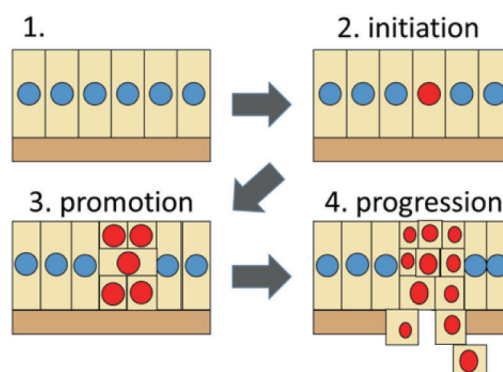


図1 多段階発がん

状態に例えられる。この段階ではアクセルは適切に踏まれている。

- (2) (1)で変異の起こった細胞集団でがん化を引き起こす遺伝子（がん原遺伝子）に変異が起こり、細胞増殖の活性化が起こる（initiation）。この状態は車のアクセルが踏まれた状態に例えられる。
- (3) 増殖が抑制されないため(2)の細胞集団が増殖する（promotion）。
- (4) 増殖した細胞集団が他の組織へ浸潤・転移する（progression）。

なお、(1)と(2)で起こったがん化に直接影響を与える遺伝子変異のことをドライバー変異と呼ぶ。希に遺伝的にドライバー変異を持っている場合は家族性のがんになる。

一般に、がんの種類はがんが最初に発生した部位（以下、原発巣と表記する）及びがん化した細胞の種類によって分類される。原発巣によってメジャーなドライバー変異が異なる。例えば脳腫瘍など神経膠腫ではIDH1/2遺伝子などの変異がドライバー変異として知られる一方、悪性黒色腫ではBRAF遺伝子の変異などがドライバー変異として知られる。

これまで、抗がん剤は、活発に増殖する細胞を阻害することでがん細胞の増殖を抑える化学療法が主であった。化学療法の対象である「増殖する細胞」にはがん細胞のみならず皮膚や毛根などの正常な状態で活発に増殖する細胞も標的になってしまうため副作用が大きいという問題があった。

近年、がん遺伝子についての研究成果を応用して、がん細胞を選択的に阻害する薬剤が盛んに開発されている。例えばImatinib（商品名グリベック）はがん由来の異常な染色体フィラデルフィア染色体から作られるBCR-ABLタンパク質を標的とする。がん細胞由来の異常な変異（＝治療目的変異）を持つ遺伝子から作られる異常タンパク質を薬剤ターゲットとする分子標的薬や免疫チェックポイント阻害剤Nivolumab（商品名オプジーボ）⁽⁴⁾がこれにあたる。以下にいくつかの分子標的薬を掲載した（表1）。

さらに、最新のがん遺伝子研究では、これまで知られていた原発巣とドライバー変異の関係に限らず、様々な原発巣から共通的なドライバー変異の存在が明らかになってきた。たとえば、乳がんのドライバー変異といわれてきたERBB2（HER2タンパク質）遺伝子変異を有する肺がんの患者に、乳がんの分子標的薬であるハーセプチンが薬効を示すという例が報告されている。このように、原発巣の臓器にとらわれず網羅的な遺伝子変異プロファイルに基づきドライバー変異を同定し、高い治療効果を望める薬物を選択することが臨床現場での喫緊の課

表1 分子標的薬とその対象遺伝子(抜粋)⁽³⁾

| 分子標的薬 | 対象遺伝子 | 日本承認年 |
|----------------------|-------------|-------|
| Gefitinib/Iressa | EGFR | 2002年 |
| Imatinib/Gleevec | フィラデルフィア染色体 | 2001年 |
| Nivolumab/Opdivo | PD-1 | 2014年 |
| Osimertinib/Tagrisso | EGFR T790M | 申請中 |
| Trametinib/Mekinist | MEK | 2016年 |
| Vemurafenib/Zelboraf | BRAF V600E | 2014年 |

題となっている。

それらの臨床現場の課題に対して、当社では2015年度に北海道大学病院様との共同研究により、網羅的ながん遺伝子変異解析技術を開発した。それは、高精度かつ短時間・低コストで遺伝子配列を解読できる次世代シーケンサによるターゲットシーケンス法を用いて、がんドライバー変異情報を解析する技術である。

この技術に基づき、当社は2016年度より医療機関内の検査システムを経由して得られたがん組織細胞から「がんゲノムデータ解析レポート」の生成までの情報解析プロセス自動化を実現した「がんゲノムデータ解析サービス」の提供を開始した。自動的に出力されるレポートには、がんドライバー変異の判定に必要とされる一塩基多型（SNV）、コピー数変異（CNV）、ゲノム不安定性の解析結果およびドライバー候補としての重要度に基づきランキングした情報、関連した薬剤、薬剤耐性や予後、治験の実施に関する解析結果が含まれる。

当社の「がんゲノムデータ解析サービス」は、がん遺伝子診断を医療機関内の迅速な検査システムとして運用するために必要な情報解析サービスである。このサービスの導入により医療機関は、がん遺伝子診断のチーム医療カンファレンスに必要となるゲノム解析情報を迅速に入手することが可能になる。

2. サービスの流れ

本章では当社が提供している「がんゲノムデータ解析サービス」の流れを示す。

全体のワークフロー

医療機関を含む全体のワークフローを診断・解析ワークフローに示す（図2）。

- (1) 患者は診断を受けるため、医療機関を受診する。その際に検査の説明を受け、診断を受ける意志があれば診断を開始する。
- (2) 医療機関は患者のがん組織および血液のサンプルを取得し、取得したサンプルをシーケンシングし、データを当社に送信する。

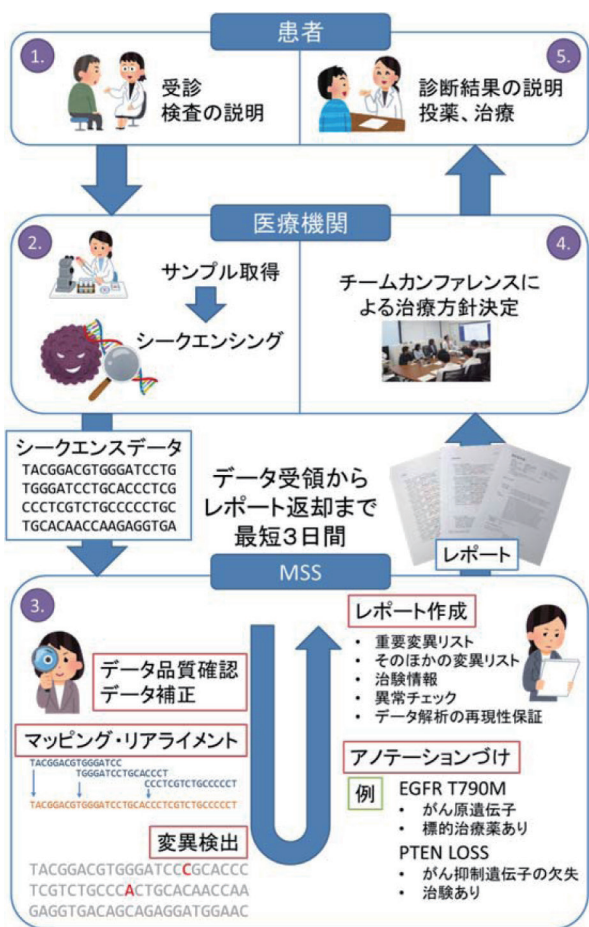


図2 診断・解析ワークフロー図

- (3) 当社は受け取ったデータの解析及びレポート作成を行う。解析結果の異常をチェックし、データ解析の再現性を保証したデータを返却する。
- (4) 医療機関は当社から返却されたデータを用いて医療機関内でチームカンファレンスを行い、患者への治療方針を決定する。その後患者に診断結果の説明を行う。
- (5) 患者は受診日から最短二週間でがん遺伝子診断の検査結果の説明及び検査結果から妥当とされる治療を受ける。

患者の受診から診断結果の説明までの所要時間は最短二週間である。短期間で返却することで、治療目的変異が見つかった場合速やかに患者の治療を開始することができる。

当社側の解析「がんゲノムデータ解析サービス」

「がんゲノムデータ解析サービス」として提供される当社側の解析について述べる。当社ではこれらの解析を自動で行うことにより最短3日で医療機関へ解析結果を返却する。

- (1) データ品質確認・補正、マッピング、リアライメント
医療機関から受領したデータを精度よく解析するため

には、機械的なエラーやサンプル固定法に由来するエラーを除く必要がある。そこで当社ではリードの品質確認をはじめとして精度の高いデータを得るための取組みを行っている。

シーケンサから出力されたリードについて、品質の低いリードはエラーの可能性が高いため取り除く。そのほかシーケンシングやサンプルの保管方法などが原因でないエラーの補正を行う。

補正を行ったデータをヒトゲノム配列にマッピングを行い、挿入欠損部分についてはリアライメントを行う。これらにより、変異検出を行うために十分な品質をもったマッピング情報を生成する。

(2) 変異検出

医療機関から受領したデータについて、正常な組織由来の検体と比較してがん組織由来の検体でのみ検出される変異を探索する。この作業により、がん由来の変異のみを確認することができる。

検出する変異の種類はがん化を引き起こす遺伝子変異 (SNV、INDEL) や異常増殖時に正しく分裂できなくなることで発生する遺伝子部位の異常な複製や欠損 (CNV)、がん細胞由来の遺伝子変異数の指標 (変異率やマイクロサテライト不安定性) を検出し統計的な有意性を判断する。これらの遺伝子変異はがん原遺伝子である可能性があるため、これらを調べることで分子標的薬など最適な治療が可能な遺伝子変異が見つかる可能性がある。

(3) アノテーションづけ

検出された変異について、その影響度を各種データベースから情報を提供する。この情報には、日本人でその変異が見られる割合や、がん細胞での観測される頻度、変異の影響度予測情報、関連する薬剤、治験情報が含まれる。当社内部でさまざまなデータを統合したデータベースを持っており、これらのデータベースを組み合わせることで精度の高い情報を提供する。

(4) レポート作成

アノテーション付けしたデータを医師が理解しやすいように分類し、数値の可視化を施したレポートを作成する。臨床情報とがん由来変異情報が揃っているもの、いずれか一方のみのものでランキングすることで、悪性度の高いものをピックアップしやすいようにしている。また変異影響度予測データベースの数値を可視化することで、薬剤、薬剤耐性や予後が揃っていない変異についても一定の評価ができる。

(5) GenomeJackの連携・自動実行

解析は当社製品のNGSゲノムブラウザGenomeJack⁽⁵⁾と連携して行っている。GenomeJackではGUIによるワークフローの作成ができ (図3)、作成したワークフロー

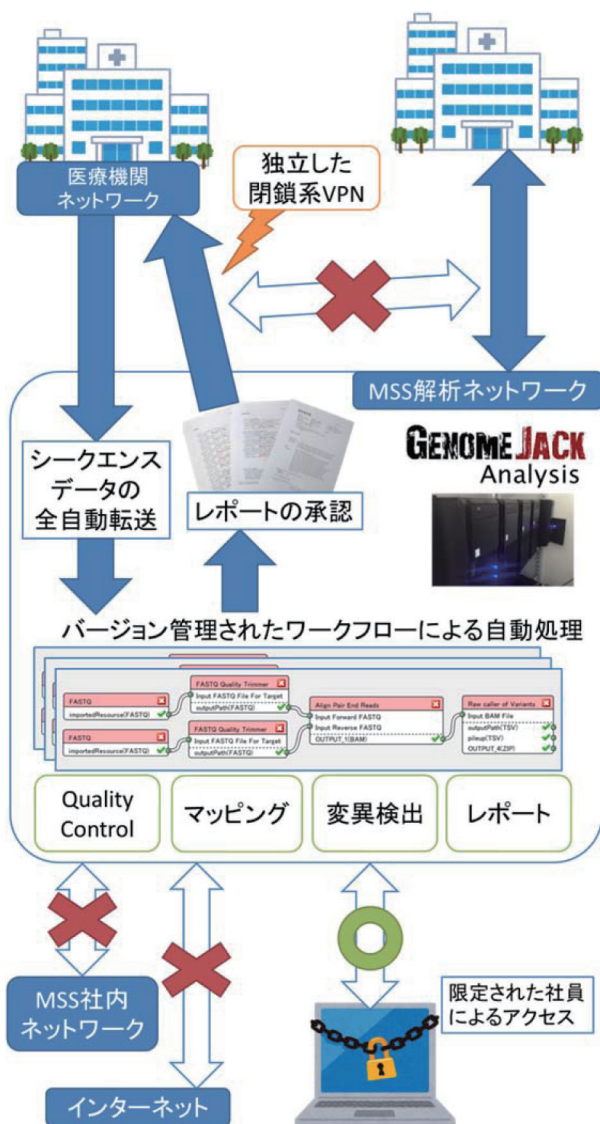


図3 当社解析サービスのセキュリティおよびGenome Jack Analysis を用いた自動実行フロー

一を複数管理することができる。

医療機関側で新しいシーケンスデータが作成されると自動で当社解析マシンへとデータが送信される。当社に送信されたデータはGenomeJackにより自動的に解析パイプラインにかけられる。予めサンプルの情報を入力することで、様々なパイプラインを自動で実行することができる。

これらの実装により実行時のヒューマンエラーを防ぐことができる。またブラウザソフトがデータの管理を行うことで、実行時に解析データをブラウザに登録してお

き後から容易に視覚的に確認することができる。さらにツールやデータベースのバージョン管理もGenomeJackを介して行うことができる。

(6) セキュリティ

当社－医療機関間のデータのやり取りには独立した閉鎖系VPNを使用しており、解析が行われるネットワーク（＝MSS解析ネットワーク）はインターネットからは分離されている（図3）。また、MSS解析ネットワークにアクセスできる社員は臨床シーケンシングにかかわる一部の社員に限定されている。これらによりデータのセキュリティを確保している。

3. むすび

2016年4月より開始された北海道大学病院 網羅的がん遺伝子検査である「クラーク検査」にて、今回の「がんゲノムデータ解析サービス」が用いられている。当該検査では2016年9月末現在までに約70件の検体に関するデータ解析を北海道大学病院様から委託され、解析結果を提供した。

今後は様々な医療機関へのサービス提供を目指すと共に、当社が今まで培ってきた遺伝子解析技術を最大限駆使して、機能不明遺伝子変異の解明やアノテーション情報の拡充などに取り組む予定である。

最後に、共同研究先である北海道大学医学部と北海道大学病院の皆様には、深くお礼を申し上げる。

引用文献

- (1) 厚生労働省, 人口動態統計平成27年度第7表, 日本: 厚生労働省, 2015
- (2) J. C. Barrett, "Mechanisms of Multistep Carcinogenesis and Carcinogen Risk Assessment," Environmental Health Perspectives Vol.100, pp. 9-20, U.S.A., 1993
- (3) 水上民夫, "がん分子標的薬開発状況に関する情報," 文部科学省. [オンライン]. Available <https://scads.jfcr.or.jp/db/table.html>
- (4) D. M. Pardoll, "The blockade of immune checkpoints in cancer immunotherapy," Nature Review Cancer, 2012
- (5) 三菱スペース・ソフトウェア, "次世代ゲノムブラウザ GenomeJackの開発," vol.22