

厚生労働科学研究費補助金
難治性疾患等克服研究事業（難治性疾患克服研究事業）
遺伝学的検査の実施拠点の在り方に関する研究
テーマ別分担研究報告書

次世代シーケンサー解析拠点の役割・必要性・意義について

研究分担

松本直通（横浜市立大学大学院医学研究科・遺伝学・教授）
松田 文彦（京都大学大学院医学系研究科附属ゲノム医学センター・教授）
松原 洋一（国立成育医療センター研究所・所長）
梅澤 明弘（立成育医療センター研究所再生医療センター・センター長）
辻 省次（東京大学医学部附属病院・神経内科・教授）

要旨：大規模ゲノム解析では、従来の医学研究で経験したことがないような大規模化が研究の前提条件となる。最近になり、疾患発症との関連で注目される低頻度アレルの多くは、集団毎に固有の低頻度の機能障害性アレルが集積したことが示されており、日本人を対象とした解析拠点が必要になる。拠点整備においては、十分な規模の次世代シーケンサーを整備し、解析能力の飛躍的向上、高品質化、コスト削減の実現が重要である。さらに、ゲノム解析に最適化された計算機資源の集中整備が必要である。診断を確定するために行う遺伝子検査は、これまでのように特定の遺伝子の解析を行うのではなく、網羅的な遺伝子解析が必要になってきており、次世代シーケンサーを用いたゲノム解析を、クリニカルシーケンシングに応用することが必要になってきており、次世代シーケンサー解析拠点の役割・必要性・意義が格段に高まってきている。

次世代シーケンサーを用いたゲノム解析，解析拠点の必要性

拠点化することにより、次世代シーケンサーによるゲノム解析能力の大規模化、ゲノムインフォマティクスの大規模化・高度化が実現できる。このことにより、ゲノム配列解析の精度の向上、スループットの向上、スケールメリットによる試薬のコスト削減が可能になる。ゲノム解析はスケールメリットの大きい分野であり、拠点化することの意義は大きい。

高度なゲノム解析技術を担う役割

疾患毎に、解析対象に含めるべき遺伝子数は飛躍的に増大してきている。例えば、心筋症では 50-70 遺伝子、てんかんでは 53-130 遺伝子が解析対象となる (HL Rehm Nature Review Genetics (2013))。このように、クリニカルシーケンシングにおいて、網羅的な遺伝子解析が必須となってきており、次世代シーケンサーの持つ役割は大きくなっている。現在の次世代シーケンサーは一定頻度で error reads が含まれ variants を call する条件についても解析の目的に応じた最適化していく必要がある。また、見出された variants については、クリニカルシーケンシングにおいては、必ず、Sanger 法など

の標準的な塩基配列解析方法による確認作業が必要となる。また、想定される変異の種類（構造多型、リピート伸長など）によっては、array CGH, repeat-primed PCR など、他の解析を併用する必要があり、統合的かつ高度なゲノム解析能力が必要となり、クリニカルシーケンシングの技術開発研究も含め、次世代シーケンサー解析拠点の持つ役割は大きい。

ゲノムインフォマティクスの重要性

次世代シーケンサーを用いた大規模ゲノム解析においては、ゲノムインフォマティクスの持つ役割が飛躍的に増大してきており、今後は、大規模臨床情報を含めたビッグデータの研究分野として発展する。次世代シーケンサーから産生される膨大な情報から、配列情報の抽出、ヒトゲノム参照配列への整列、variants の call などの一連の解析、さらに、病的意義を有する変異の探索、評価においては、さまざまなデータベースを参照すること、機能解析の予測、あるいは、機能解析を加えるなど、重層的な解析が求められる。また、必要に応じて、新しいアルゴリズムの開発研究も求められる。これらの解析を実施するためには、大規模並列処理ができる計算サーバーを備え、インフォマティクスの研究者が担当する研究体制が必要であり、このような役割を次世代シーケンサー解析拠点が担う必要がある。また、医学系の情報と、大規模ゲノム情報を同時に扱い、OJT (on the job training) が可能な環境が、人材不足が指摘されているメディカルゲノムインフォマティクスの分野の人材育成という点できわめて重要となり、このような環境の構築・提供が次世代シーケンサー解析拠点の役割となる。特に重要な点は、新しい解析アルゴリズムの解析など、研究面で高いレベルを目指すことが大切であり、そのような研究の場を提供し、ゲノムインフォマティクスに造詣の深い研究者を育成する機能が、次世代シーケンサー解析拠点の重要な役割となる。

医療におけるクリニカルシーケンシングの必要性

診断確定のために、遺伝子検査、特に、複数の遺伝子を対象としたクリニカルシーケンシングの必要性が医療現場で高まってきている。個別の研究室で遺伝子検査ができるところもある一方、網羅的な解析・解釈ができる研究室は限られており、このような状況において、次世代シーケンサー解析拠点が果たす役割が大きい。依頼側から見ると、どのような要件を満たせば依頼できるのかなど、依頼から結果の受け取りまでの流れが明快に示されていれば、そのようなフレームワークを利用してクリニカルシーケンシングを依頼できるメリットは大きいと考えられる。診断確定のための必要なステップとしてクリニカルシーケンシングを利用できる体制の整備と、依頼側から見て満足できる turnaround time (受け取りから結果返却までの期間) で、ゲノム解析結果を返却できるだけの体制を拠点側で整備し、透明性を確保することが重要である。また、解析結果については、依頼側が診療や臨床研究に活用できるようにすると共に、一定の猶予期間後に、データベース登録を義務づけるなどの仕組みを整備して、疾患変異データベースを充実させていくことも大切であると考えられる。近い将来、クリニカルシーケンシングが医療の中で重要な役割を果たすようになると予測され、検査の標準化、品質管理、倫理面の課題などを含め、医療制度の中でクリニカルシーケンシングを実装していく上でも、貢献することが期待される。

解析結果の解釈についての重要性

網羅的なゲノム配列解析を行うことにより、膨大な数のゲノム上の variants が見出される。例えば、全エクソン配列解析を実施すると、200-300 個の新規（データベースに登録されていないもの）の非同義置換（アミノ酸置換を伴う変異）が見出され、これらの variants について、どのように解釈を与えるかが重要となる。この解釈にあたっては、1. 大規模の日本人健常者のゲノム多様性のデータベースを用いた分析（新規の変異であるか、既知の変異であればそのアレル頻度などの分析）、2. 対象とする疾患に関して、HGMDをはじめとする疾患関連変異のデータベース、特に日本人の変異データベースの分析、3. 疾患の表現型の多様性を考慮に入れた分析（臨床側で想定していなかった遺伝子の変異であることが判明することもある）、4. 病原性変異の判断基準に基づく検討（一定の基準に基づく、病原性変異としてのランクづけ）、5. 種間の保存性などに基づく、deleterious mutation の推定のスコアリング（ランクづけ）、6. 必要に応じて、連鎖解析など家系分析を追加しての検討（候補遺伝子領域の絞り込み）、次世代シーケンサー以外の解析方法の適用などが必要になる。このように豊富な分析能力を整備するには、拠点がその開発を含めて担うことが必要であり、それを、医療コミュニティ（医師、遺伝カウンセラー、研究者など、幅広い職種）で活用していくことが望まれる。

研究面での次世代シーケンサー解析拠点の役割

次世代シーケンサー解析拠点においては、クリニカルシーケンシングの提供を行うとともに、未だ発症機構が未解明な疾患について、遺伝性疾患の病因遺伝子の解明、孤発性疾患（complex trait）の疾患感受性遺伝子の解明に対して貢献することが求められる。遺伝性疾患については、現在病因遺伝子が未解明の疾患の多くは、小家系だったり、きわめて稀であったりと、病因遺伝子の解明が困難な疾患が数多く残されており、その点で、連鎖解析、エクソーム・全ゲノムシーケンシングなど、高度かつ統合的な解析技術を投入する必要がある。また、complex trait の疾患については、サンプル数の大規模化が成果をあげる上で必須であるので、難病研究班との連携を含め、臨床側と次世代シーケンサー解析拠点で、適切な多施設共同研究体制を構築してそれぞれに役割分担をすることも求められる。このような研究は疾患ゲノムコホート研究として位置づけられ、大規模研究プロジェクトとして推進する必要があり、次世代シーケンサー解析拠点が、多施設共同研究体制のもとにこのような研究を担当する役割を果たすことが望ましい。また、その解析結果などを研究者コミュニティが広く活用でき、わが国全体のゲノム医学研究が発展するような仕組みを整備することも求められる。また、個別の研究室において行われる疾患遺伝子探索研究も数多くあり、このような研究に対しても、次世代シーケンサー拠点が必要に応じて一定枠の支援的な役割を果たすことがわが国の研究全体の発展にも貢献する。

データベースの構築・維持

上述したように、クリニカルシーケンシングにおいては、観察されたゲノム variants に対する解釈が最も重要となる。この解釈においては、健常者集団のゲノム多様性についてのデータベース、疾患毎の変異データベースが重要なリソースとなる。特に、ゲノム多様性、疾患関連変異は、民族毎に固有のもの（ethnicity-specific）が少なくないことがわかってきており、日本人集団のデータベースの持つ役割が重要な位置を占める。これまでこのような日本人に特化したデータベースは存在してい

なかったが、先に公開された 1,208 名の健常者日本人集団のゲノム多様性のデータベース (<http://www.genome.med.kyoto-u.ac.jp/SnpDB/>) は、厚労科研でサポートされた次世代シーケンサー 5 拠点 が協力して構築されたものであり、次世代シーケンサー解析拠点の持つ重要な役割といえる。このようなデータベースは広く研究者がさまざまな研究、診療に活用することができ、わが国の研究の基盤的な役割を果たす。今後、生活習慣病など頻度の高い疾患の発症に対する影響度の高いゲノム variants の検索を目的に exome-association study が発展すると予測されるが、そのような研究の基盤として、多施設共同研究体制に基づいて、さらにサンプルサイズを拡大することが重要である。疾患に関連する変異データベースは、個別の疾患についていくつかの小規模データベースはわが国でも存在するが、幅広く多くの疾患を網羅したデータベースは存在していない。次世代シーケンサーを用いたクリニカルシーケンシングの拠点化は、このような変異データベースの構築に対しても大きく寄与すると考えられる。また、疾患関連の変異データベースの構築・維持については、拠点のみならず、個別研究で見出される疾患関連変異についても、多くの研究者の協力に基づきデータベースへの deposit を推進することも重要である。

国際的には、Human Variome Project という大きな動きがあり、全世界で、ヒトゲノムの多様性情報のデータベースの構築していくことが進められている。日本では、まだ、Japan Node が設置されておらず、Japan Node を設置し、国際的な Human Variome Project の推進に貢献していくためにも、このような次世代シーケンサー解析拠点が重要な役割を担うことが求められる。

国際的な動向

ゲノム医学分野で、国際的に activity の高い次世代シーケンサー解析拠点としては、英国 Sanger 研究所、米国 Broad 研究所、Baylor College of Medicine などあげることができる。また、ゲノム医学に特化していないが、全ての生物を対象としたゲノムシーケンシングの分野では、中国の BGI が群を抜いて活発な研究を進めている。いずれの研究所においても、資源を集中し、次世代シーケンサー、ゲノムインフォマティクスの大規模解析拠点を構築している。最近になり、米国 Broad 研究所は CLIA の認証を取得し、クリニカルシーケンシングの分野に乗り入れている、Baylor College of Medicine は早い時期から、次世代シーケンサーをクリニカルシーケンシングの分野に取り入れ、大きな成果をあげている。英国 Sanger 研究所は、ヨーロッパのゲノム解析拠点としても機能しており、多くの研究者が自らのサンプルを持って、研究を行う仕組みも整備している。このように国際的な動向は、資源の集中化と高度化、さらにその拠点を研究者コミュニティ全体が活用する仕組みも整備してきている。