

次世代を超えたDNAシーケンス技術

大場 利治

はじめに

2000年に入ったころ、『生物学ハンドブック』に「核酸配列決定法とその利用」¹⁾をまとめてから15年以上経ち、その時にはなかった「次世代シーケンサー (NGS)」が実用現場で使われ始めている今日、再度、DNAシーケンスの歴史を振り返ってみるのに良い機会だと思い本稿の執筆を引き受けた。

筆者はタカラバイオ株式会社のドラゴンジェノミクスセンター(現:バイオメディカルセンター)にてDNAシーケンスの受託サービスに関わった。また、次世代シーケンサーに関しては、受託会社としてシーケンサーが出始めてからそれらを利用する立場から技術を比較、その特徴に違いから各社のシーケンサーを導入して、研究目的に応じたシーケンサーの使い分けの提案をしてきた。アカデミアやシーケンサーの開発・提供企業と異なり、ある程度客観的に利用する立場から公平にお話しができるかと考えた。

サンガー法とシーケンス

サンガー法(ジデオキシ法)によるシーケンスは現在も行われているが1970年代に提唱された原理である²⁾。詳細は『生物学ハンドブック』¹⁾をみていただければと考えるが簡単にその原理を述べる。この方法はDNAポリメラーゼによるDNA鎖の伸長反応をAGCTのそれぞれのジデオキシヌクレオチドを利用して配列特異的に

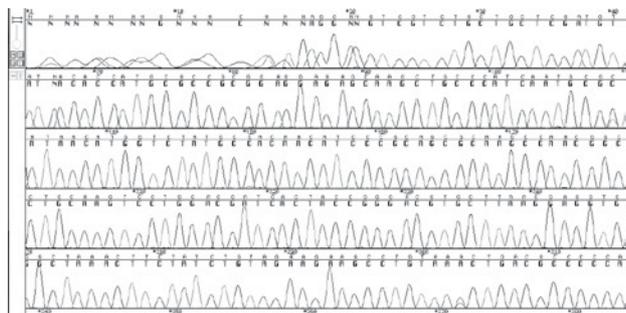


図1. 蛍光検出シーケンサーのデータと表示例。配列解析ソフト Sequencherによるデータの読み込みと表示例。

停止することのできる長さの異なるDNA断片を電気泳動することにより配列を決定する方法である。

筆者が研究生で名古屋大学にいた1990年前半には電気泳動の平板ゲルを自分で作り、その検出感度の問題からラジオアイソトープ (RI) を利用してバンドを検出するといったやり方が主流であり、グラジエントゲルを作り、いかに長くシーケンスできるか競ったものである。その後、電気泳動と検出にはDNAシーケンサーが登場し、電気泳動は平板ゲルからキャピラリー電気泳動法に、検出はRIから感度が改良された蛍光検出に移り、これは現在も完成した域で利用される技術になっている(図1)。次世代シーケンスが全盛の現在でも結果の検証のためにサンガー法を用いたシーケンスは頻繁に行われている。

次世代シーケンサーの始まり

2000年ごろヒトの全ゲノム解析やモデル生物の全ゲノム配列がサンガー法のシーケンサーを並べた大型の解析センターであいついで発表された(ヒト³⁾、シロイヌナズナ⁴⁾など)。このころから全ゲノム配列を利用したポストゲノム解析として全遺伝子の発現解析(トランスクリプトーム解析)や個人医療に向けたヒト遺伝子の一塩基多型の解析が注目された。このためには短時間で大量の配列解析が必要となり並列シーケンサーのアイデアが生まれた。

Lynx社によって開発されたMPSS (massively parallel signature sequencing) は細胞内で発現しているPolyA RNAをマイクロビーズ上でcDNA化して、マイクロビーズをフローセルに敷き詰め、ビーズ上で配列認識用のアダプターのライゲーションおよび16種類の蛍光プローブのハイブリダイゼーションを繰り返すことにより、並列的に配列決定をする技術であった⁵⁾。100万以上の配列を一度のRunで得られるこれまではまったく発想の異なる並列シーケンス法であった。その後、並列化の部分は、オイル中に存在する各水滴内に断片化した1分子の1本鎖DNAと1個のビーズが含む状態でPCR反応を行うエマルジョンPCR法、2種のアダプターのDNA配列を表面に持つフローセル上で鋳型DNAを

結合させ、ブリッジPCRにより1種類のDNAクラスターをフローセル上に合成するブリッジPCRの方法が登場し大幅に進展した。また、蛍光や発光検出のカメラやコンピュータ技術の進展により、今日ある次世代シーケンサーによる並列シーケンサーの確立につながっている。

次世代シーケンサーの比較と応用

次世代シーケンス (NGS) とよばれる一連のシーケンス技術は2005年ごろから実用的に使われるようになってきた。エマルジョンPCRの技術とパイロシーケンスの原理⁷⁾で検出の454社のシーケンサー（その後、ロシュ社に買収、すでに終売）が出始めた2005年くらいから、筆者らは受託サービス会社の立場で各社の次世代シーケンサーを利用してきた。

454社のシーケンサーでは、まず、エマルジョンPCRによってビーズ上に1種類のDNAを増幅する。シーケンサーでは、ビーズをシーケンサーのピコタイタプレート内の各ウェルに一つずつロードする。次に、4種のヌクレオチド (dTTP, dATP, dCTP, dGTP) を順次添加する。DNAポリメラーゼによる伸長反応により、dNTPがDNA鎖に取り込まれる際にピロリン酸が放出され、このピロリン酸を基質として sulfurylase により ATP を生成、ATP と ルシフェリン を基質として luciferase による発光反応を起こし、この発光シグナル強度を CCD カメラで検出する。次のヌクレオチドが添加される前に、組込まれなかった塩基はすべて Apyrase により分解する。これを1サイクルとしてシーケンス反応が完了するまでこの工程を繰り返す。そして発光の有無とシグナル強度から塩基配列を決定する⁷⁾。

2006年に、フローセル上でブリッジPCRを行い、DNA増幅する技術を利用した Solexa 社の Genome Analyzer が登場した。その後2007年に Solexa 社は イルミナ社に合併、HiSeq にその技術が受け継がれている。このシーケンサーでは、まず、ブリッジPCRにより1種類のDNAクラスターをフローセル上に大量に合成する。シーケンサーでは、3'末端に保護基のあるヌクレオチド (dTTP, dATP, dCTP, dGTP) にそれぞれを検出できるように異なる蛍光色素でラベルした可逆的ターミネーター (改変ヌクレオチド) が使われる。4種すべての改変ヌクレオチドを同時にフローセルに添加し、相補的な塩基を持つヌクレオチド取り込みの後、残りの塩基のものは洗い流し、蛍光検出する。その後、保護基、蛍光色素を除去する。これを1サイクルとしてシーケンス反応が完了するまでこの工程を繰り返す⁸⁾。この方法では一度に一つのヌクレオチドを取り込ませることになりパイロ

シーケンスの連続した塩基の検出における読み取りエラーの多い欠点を克服している。

その後、エマルジョンPCRとSBH (sequencing by hybridization) による検出の Applied Biosystems 社 (現在、Thermo Fisher Scientific 社) の SOLiD (すでに終売)、エマルジョンPCRとIon半導体による検出の Ion Torrent 社 (現在、Thermo Fisher Scientific 社) の Ion PGM や Ion Proton など各社の次世代シーケンサーを導入、利用してきた。現在もがん遺伝子検査によく用いられる Ion Torrent のシーケンス技術は、まず、エマルジョンPCRによってビーズ上に1種類のDNAを増幅する。シーケンサーでは、半導体マイクロチップ上の微小なウェルに一つずつロードされたビーズに、4種のヌクレオチド (dTTP, dATP, dCTP, dGTP) を順次添加される。DNAポリメラーゼによる伸長反応によりdNTPがDNA鎖に取り込まれ水素イオンが放出される。このイオンの持つ電荷によるpHの変化をイオンセンサーが検出しデジタル信号に変えることになる⁹⁾。蛍光や発光検出する従来の装置と異なりスキャナやカメラ、光源などを必要とせず、各塩基の取り込みを数秒で記録できることから、この装置はより安価で信頼性の高い診断検査の実現を期待されている。

それぞれのシーケンサーの原理から、その解析結果のリード数やリード長が異なり、一長一短があり、その適した用途が変わってくる。主要なシーケンサーの特徴概要を表1にまとめた。

たとえば、イルミナやIon Torrentのシーケンサーはリード長が100 base~250 baseと短いのが特徴である。大量のリード数が出せるのが特徴である。短いリード長だと新規のゲノム配列解析を行う用途には不向きである。なぜならば、生物のゲノムDNA中には少なからず繰り返し(リピート)配列があるからだ。新規のゲノム解析ではアセンブルという手法を用いてリード配列を並べて推定されるゲノム配列を組み立てていく。ゲノム配列中のリピー

表1. 主な次世代シーケンサーの特徴と用途比較

装置	454 GS FLX+ (※終売)	HiSeq 2500 (Rapid Mode)	Ion PGM (Ion 318 chip)
原理	Pyrosequencing	Sequencing by synthesis	Ion semiconductor sequencing
リード数	約100万	約2.5億	約400万
リード長	700 base	100/150/250 base	~200 base
用途	・新規ゲノム配列解析	・変異解析 ・遺伝子発現解析	・癌遺伝子の 変異解析

ト配列が数百baseを超えると、短いリード情報では、その並びがどうなっているかがわからない。そのような場合は長いリード長を出力できるシーケンサーが有利である。そのような用途では約700 baseのリード長を得ることのできた454社のシーケンサーは多く使われた。

逆に、遺伝子発現解析や変異解析においては大量のリード数が出力できるシーケンサーが有利である。このような解析には、すでにゲノム解析で決められたリファレンスとなるゲノム配列が用いられるので、リファレンスの配列に類似配列を探してマッピングするという手法ですむからである。

第3世代シーケンサー ～一分子シーケンサー～

新規のゲノム配列解析を行う用途に適した非常に長いリード長を得ることのできる一分子シーケンサーがPacBio社から発表された。DNAを1分子単位でPCRによる増幅無しでシーケンスできるまったく新しい原理のシーケンサーであることから第3世代シーケンサーとも言われている。

このシーケンサーでは、鋳型DNAのアダプター部分にプライマーとポリメラーゼを結合させSMRT (single molecule real time) cellにいれ反応する。SMRT cell上にはZMW (zero-mode waveguide) と呼ばれる無数の孔が開いており、このZMWの底面に1分子のDNA-ポリメラーゼ複合体が固定される。反応は4種類の異なる蛍光の付いた塩基が、ZMWの外から入り込み、DNA鋳型に取り込まれると蛍光が切り離されてZMWの穴の底で光るのを検出する¹⁰⁾。

このシーケンサーはリード長平均10 kbaseと非常に長いことが特徴として上げられる。このため新規ゲノム解析に幅広く使われ、微生物ゲノム解析においては454シーケンサーの座を奪う形となった。また、植物などの未知の大型ゲノム解析においても威力を発揮し、イルミナの短いリード解析と組み合わせることによって、迅速にゲノム配列を決定することができるようになった。たとえば、京都大学の安井らは高栄養源の穀物として注目されているキヌアのゲノム解析結果を発表した¹¹⁾。理研の持田らは多くの漢方薬の原料となっている生薬の甘草などのゲノム解析結果を発表した¹²⁾。

さらに、最近、USBメモリほどの大きさのコンパクトで持ち運び可能な一分子シーケンサーとしてOxford Nanopore社からナノポアDNAシーケンサー「MinION」が販売された¹³⁾。非常にコンパクトで場所を問わないことや長いリード長を得られることから、環境調査からベッドサイドまで広い分野での展開が考えられる。

実用の時代 ～精密医療や産業への利用～

DNAシーケンスの技術の応用は多岐な分野にわたる。NGSの進歩に伴い、広い分野に実用的に使われる時代になった。2017年5月に行われた第5回NGS現場の会でのポスター発表のカテゴリー別の演題数をみると図2のようになる。

これまでは基礎研究や技術開発の側面がまだ強かったのが、一気に実用化の時代に入ってきている。シーケンサーの技術進歩、変化は激しいが、今後はそれぞれの用途に応じたシーケンサーとなっていくと考えられる。イルミナ社の短いリード長のシーケンサーもMiSeqのような卓上型のシーケンサーが登場してきた。また、Ion Torrentのシーケンサーもコンパクトで目的の遺伝子の変異解析をみるためのCancer Panelとともに利用されている。

用途の一つの方向として精密医療の分野がある。遺伝子変異が原因のがんでは遺伝子変異を調べることにより分子標的薬がわかる場合がある。このような患者に有効な治療薬を届けることを目的の一つにSCRUM-Japanプロジェクトが行われている。この際の遺伝子検査にはThermo Fisher Scientific社のIon PGMシーケンサーおよびOncomine cancer research Panelを用いた変異解析が用いられる。このプロジェクトでは治験のデータベース化も行われており、適した治験があれば紹介できる。患者が少ないとなかなか製薬会社が手を出せない希少疾患に対しても医師主導型治験が行われる場合も増えている。たとえば、国立がんセンター東病棟の葉らは、RET遺伝子陽性の肺がん患者に対してRET分子標的治療薬・バンデタニブ（進行性の甲状腺髄様がんの治療薬、肺がんでは適用されていない）を投与する医師主導治験を実施したところ、約半数にがんの縮小効果が認められた

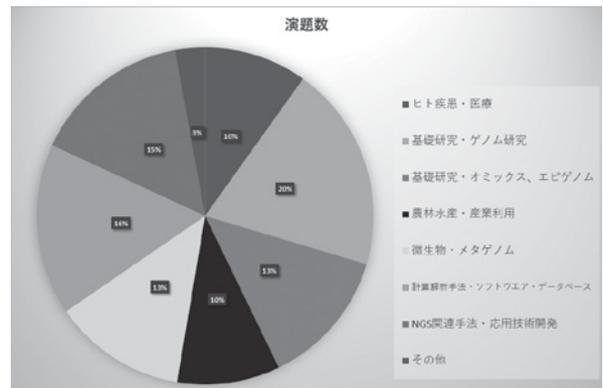


図2. 第5回NGS現場の会におけるカテゴリー別の演題数

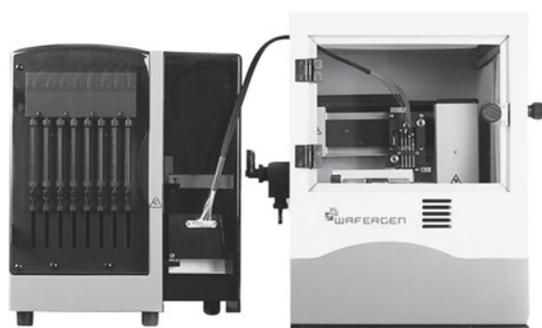


図3. NGS用途にシングルセルを分取する装置の例 (写真: ICCELL8 MultiSample NanoDispenser16, タカラバイオ株式会社提供).

14).

産業分野への次世代シーケンサーの応用も進んでいる。動植物のジェノタイピングを行う方法として広く行われているのはRAD-seqと呼ばれる方法である。制限酵素で切断したゲノムDNA断片の末端配列をNGSで決定することで、マーカーの探索や比較的lowコストでジェノタイピングを行うことができる。たとえば、信州大学の松村らは、ニガウリにおいてRAD-seq法によって見いだしたDNAマーカーのうち雌性型形質と連鎖する複数のDNAマーカーを見いだし報告している¹⁵⁾。

また、基礎研究分野では、効率よく1細胞（シングルセル）を分取しシーケンス解析に必要なDNA量を得るための装置がいくつか販売されるようになったため、シングルセル解析が盛んである（図3）。

シングルセルを分離して細胞ごとの遺伝子発現解析やゲノム解析を行うことで、バルク解析ではわからなかった細胞集団内の不均一性の情報を得ることができる。これによって新規な細胞の発見や分化過程での詳細な遺伝子発現変化を調べることができ、新規なメカニズムの発見につながるがあると考えられる。たとえば、細胞が分化する過程では細胞ごとに分化の速度が異なるため、シングルセルレベルで解析を行うことで細胞分化に関わる遺伝子発現の変化をより高解像度に捉えることが期待できる。Trapnellらは、筋芽細胞（myoblast）から筋細胞（myocyte）、筋間（myotube）へ至る過程をサンプリングし、バルク解析では“平均値”の中に埋もれていた主要調節因子の発現のスイッチ様の変化を発見している¹⁷⁾。

おわりに

このようにDNAシーケンス技術はその方法を劇的に発展させ、これまでにアプローチできなかったシングルセルレベルでの細胞の挙動にまで迫ることができるようになってきた。シングルセル解析が可能になった背景には超微量のDNAやRNAからのシーケンスの反応に持つていくための前処理のライブラリー作製に使われる試薬の進歩も忘れてはいけぬ。シングルセルのRNAからのライブラリー作製には、SMART-seqが使われることが多いが、逆転写酵素、アダプター配列付加の工夫¹⁸⁾、DNAポリメラーゼの配列バイアスの低減などの技術が盛り込まれており、ここには従来の生物工学の手法での酵素の改良などの多くの努力がされている。さらに基礎研究ばかりでなく、精密医療や効率の良い動植物の育種などの産業利用も実用的になされるようになってきた。今後はますますこれらの技術的な効率があがることでコストダウンがおき、産業利用が広がるであろう。

なお、本文中の社名、商品名は各社の商標または登録商標である。

文 献

- 1) 日本生物工学会 編：生物工学ハンドブック，p. 257，コロナ社 (2005).
- 2) Sanger, F. *et al.*: *Annu. Rev. Biochem.*, **57**, 1 (1988).
- 3) International Human Genome Sequencing Consortium: *Nature*, **409**, 860 (2001).
- 4) Arabidopsis Genome Initiative: *Nature*, **408**, 796 (2000).
- 5) Brenner, S. *et al.*: *Nat. Biotechnol.*, **18**, 630 (2000).
- 6) Santoro, S. W. *et al.*: *Proc. Natl. Acad. Sci. USA*, **94**, 4262 (1997).
- 7) Ronaghi, M. *et al.*: *Science*, **281**, 5375 (1998).
- 8) Bentley, D. R. *et al.*: *Nature*, **456**, 53 (2008).
- 9) Rothberg, J. M. *et al.*: *Nature*, **475**, 348 (2011).
- 10) Schadt, E. E. *et al.*: *Hum. Mol. Genet.*, **19**, R227 (2010).
- 11) Yasui, Y. *et al.*: *DNA Res.*, **23**, 535 (2016).
- 12) Mochida, K. *et al.*: *Plant J.*, **89**, 181 (2017).
- 13) Oxford Nanopore Technol. 社：https://nanoporetech.com/jp/ (2017/6/5)
- 14) Yoh, K. *et al.*: *Lancet Respir. Med.*, **5**, 42 (2017).
- 15) Matsumura, H. *et al.*: *PLoS ONE*, **9**, e87138 (2014).
- 16) タカラバイオ社：http://catalog.takara-bio.co.jp/ (2017/6/5)
- 17) Trapnell, C. *et al.*: *Nat. Biotechnol.*, **32**, 381 (2014).
- 18) Picelli, S. *et al.*: *Nat. Protoc.*, **9**, 171 (2014).