

大規模配列決定の新技术：Pyrosequencing を中心として

ヒトゲノム配列の完全解読を見るまでもなく、いまや全ゲノム配列の解読は日常的な研究と化している。しかしながらその膨大なコストのため、ゲノムの読まれる真核生物種は興味の対象のごく一部であり、モデル生物の他は、経済効果の大きな病原生物や家畜に大きく偏っている。このため、コストの安い配列解読技術の進歩が待たれるのは言うまでもない。と同時にゲノム解読計画が配列決定の技術を飛躍的に進歩させてきたのもまた事実である。現在の塩基配列決定の主役はサンガー法であり、また、キャピラリーシーケンシングである。サンガー法の詳細な説明は多くの教科書で記載されているのを参照してもらいたい。簡単に言うとサンガー法では、塩基のアナログを標識し、そのDNA合成に伴う取り込みを見ることで配列を決定する。かつてはRIによって標識し、ゲルで泳動するのが一般的だったが、最近では、蛍光物質で標識し、細いキャピラリーにポリマーを充填した中を泳動して検出するのが一般的である。しかし、この手法では、DNAが長くなるにつれて泳動の際の移動度の違いが小さくなるため、一度に読める長さには限界がある。現在の手法では、99%以上の精度で700bp程度を読むのが限界である。また、一度に読めるDNAの数はキャピラリーの数によって制限されるため、ゲノム解読のような大量のDNAを扱う際には、多数のシーケンサーが必要となってくる。これらのことから、新しい技術革新が期待されていたが、既に一部実用化され普及が待たれる手法の一つがPyrosequencingと呼ばれる手法である。

Pyrosequencingでは、DNA合成において塩基が取り込まれる際に放出されるピロリン酸(Pyrophosphate)を検出する。Pyrosequencingの原理と問題点はRonaghi(2001)に詳しい。dNTPがDNAに取り込まれる際には、 $\text{DNA}(n)+\text{dNTP}\rightarrow\text{DNA}(n+1)+\text{PPi}$ の反応が起こり、ヌクレオチドがDNAに結合すると同時にピロリン酸(PPi)を放出する。ピロリン酸は2つのリン酸が結合した化合物で高いエネルギーを持つ。それぞれのdNTPを個別に反応液に加えた際に放出されるピロリン酸を検出することで配列を決定しようというのがPyrosequencingの原理である。実際には、ATP sulfurylaseによりピロリン酸とAdenosine 5'phosphosulfate (APS、アデニン-リボース-リン酸基-硫酸基の結合した化合物)から $\text{ATP}(+\text{SO}_4^{2-})$ を合成し、このATPとルシフェリンを混ぜてルシフェラーゼによって発光させる。この光を検出するのである。反応を終了させてから検出を行なうサンガー法と異なり、Pyrosequencingでは、反応と検出を同時に行なう。そのためには、それぞれのヌクレオシド3リン酸(dATP, dCTP, dTTP, dGTP)を別々に与えなければならない。また、反応後に残ったヌクレオシド3リン酸及びATPを確実に取り除かなければならない。そのために2種類の方法が開発さ

れている。一つは、物理的に洗浄によって取り除く方法、もう一つは、酵素反応的に取り除く方法である。後者では、ApyraseというdNTPやATPを分解する（ $dNTP \rightarrow dNMP + 2P_i$, $ATP \rightarrow AMP + 2P_i$ ）酵素を用いる。ここでは、リン酸基はひとつずつ取り除かれるためピロリン酸は合成されない。物理的な洗浄よりも酵素による分解の方が取り扱いは簡単だが、分解産物が残ることと活性が衰えることが問題である。

Pyrosequencingでは、配列はそれぞれのヌクレオチドを加えた際の光の強さとして得られる。これはPyrogramと呼ばれる。DNA合成で次に付加されるべきヌクレオチドが反応液に加えられたときにはヌクレオチドがDNAに取り込まれピロリン酸が出て最終的に光が検出される。違うヌクレオチドが加えられたときには発光しない。AAAのように同じ塩基が連続するときにはそれだけ大量のピロリン酸が放出されるので光の強度が高くなる。ところがあまり長く同じ塩基が連続する場合には、時間内にDNA合成が終わらないものが出てくるため後ろのDNA合成が不均一になり、配列が読めなくなってしまう。5から6個同じ塩基が続く場合にこのようなことが起こる。これはPyrosequencingの大きな問題点の一つである。

MarguliesらはPyrosequencingに*in vitro*のDNA増幅を組み合わせることで大量の配列を短時間で得る手法を開発した(Margulies et al. 2005)。彼らは、まず、ビーズにDNAを結合させ、そのビーズとPCRの反応液をエマルジョンの形で包み込んだ。断片ビーズに1分子のDNAだけを最初に結合させることができれば、全く同じ配列を持ったDNA断片がPCRによって増幅され、ビーズに結合する。DNA断片の増幅後、ビーズを洗浄し、光ファイバーで作ったプレートのウェルに入れる。この際、ウェル当たり1つのビーズだけが入るように最適化する。ウェルの大きさは、直径44 μ m、深さ55 μ m、容積は75 μ lで、ビーズの大きさは直径28 μ mなので大抵は1つのビーズだけがウェルに入ることになる。酵素の結合したビーズを加えた上でATGCそれぞれのヌクレオチドを加えて反応させる。後はPyrosequencingの手法の通りである。

Marguliesらの手法の最大のメリットはシーケンスの前過程のサブクローニングの大幅な省略が可能となることであろう。通常、DNA断片をプラスミドに挿入したライブラリーを構築し、大腸菌にトランスフォーメーションしてプラスミドを増やした上でPCRによってDNA断片を増幅する（あるいはプラスミドを精製する）という面倒な過程が必要となる。これを小液量でのPCR反応で代替することで時間と費用が節約できる。また、スケールが下がったことにより大量のシーケンスが一度に決定可能である。1回4時間のランで25Mbの配列が得られる。これは、キャピラリーを使用したサンガー法の場合の約100倍の量である。一方で問題点も提起された。彼らは*Mycoplasma genitalium*のゲノムシー

ケンスを行ない、公表されている配列と比較した。その結果、配列の精度は96%とサンガー法の99.4%に比べてかなり低く、一つのDNA断片からは110bp程度しか読めていない。精度の低さは、一回の作業で100倍の量のシーケンスが得られることで補うことができる。おそらく問題となるのは得られるDNA配列の長さであろう。80bp程度の長さの反復配列が存在すると、配列をアセンブルすることは不可能になってしまう。従来の方法ならば700bp程度の長さが得られるのでこの違いは大きい。現在の技術では、反復配列の多い真核生物のゲノムを読むことは不可能であるが、反復配列の少ない原核生物のゲノムシーケンスや、組織ごとの発現カタログ作成のためのESTなどには現時点で使用可能である。尤も、得られる配列の長さは技術の進歩によっておそらく急速に改善されるであろうから悲観するには当たらない。

Pyrosequencingには専用の機械が必要になる。一方、Shendureらは研究室に普及している機械を使って低価格で配列決定する手法を考案した(Shendure et al. 2005)。この手法はなんと蛍光顕微鏡で配列決定するというものである。DNAの配列決定には様々な目的があるが、Shendureらはその中でResequencingに特化した手法を考えた。Resequencingとは既にわかっている配列にごく近縁な配列を読み直すことで、バクテリアの株間の違いや変異体でどのような置換が入っているのかを知るために行なわれる。この場合、わかっているゲノム配列のどの領域と一致するのかがわかれば良いため、長い配列を得ることは重要ではない。

Shendureらはまず、制限酵素とローリングサークル型のDNA複製酵素を利用して、2つの17~18bpのゲノム由来の配列が、共通する3つの配列には含まれたDNA断片を増幅した。この2つのゲノム由来断片は最初にゲノムを切断した際のDNAの両末端の領域に由来する。すなわち、タグを付けたDNA断片の両末端17~18bpを残して、中央部を特定の配列に入れ替えたことになる。このDNA断片をビーズに結合し、アクリルアミドゲルで固めたスライド(アレイ)を作成した。配列決定は様々な方法を試したが、Degenerate ligation法が適していたらしい。共通配列部分に相補的なオリゴDNA(アンカープライマー)をまず、ビーズ上の配列に結合させる。そののち、NNNNANNNN-Cy5、NNNNGNNNN-Cy3、NNNNCNNNN-TexasRed、NNNNTNNNN-FRETの4種類の、中央の塩基だけを固定しそれぞれの3'末端に異なる蛍光標識を付けた縮重オリゴDNA(縮重9mer)をハイブリダイゼーションさせる。Nはランダムな塩基である。配列が100%相補する場合だけ縮重9merとアンカープライマーがリガーゼによって結合される。洗浄後に励起光を当て蛍光顕微鏡で写真を撮り、画像解析によってそれぞれのビーズにおける蛍光を識別することで中央の塩基を決定する。これをアンカープライマーを変えて繰り返すことで一つのDNA断片につき26bpの配列を決定した。1サイクルは135分と非常に長い、1視野に含

まれる全てのビーズを同時に観測できるので、秒当たり 140 bp の速度で配列決定できる計算になる。と同時にビーズを密に配置できれば、それだけで配列決定の速度が向上するという利点もある。配列の間違いは 100 万塩基に 1 個程度であった。また、コストは従来の 9 分の 1 程度まで下げることができた。

Margulies らの手法と Shendure らの手法はどちらも短い配列を大量に得ることに特徴がある。また、*in vitro* での DNA 増幅技術であるエマルジョン PCR など、小液量で DNA を増幅する手法を用いることで、サブクローニングやトランスフォーメーションなどの過程を省略し、時間を短縮すると同時にスケールダウンに成功している。スケールダウンはコストダウンに加えて配列決定の速度を加速する。これらはキャピラリーでの電気泳動では実現できない利点である。その一方で配列の精度や一度に読める長さは従来の方法に劣っている。目的に合わせていくつかの配列決定技術を併用していく時代が訪れているのかもしれない。

Ronaghi M.

Pyrosequencing sheds light on DNA sequencing.

Genome Res. 2001 Jan;11(1):3-11.

Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, Berka J, Braverman MS, Chen YJ, Chen Z, Dewell SB, Du L, Fierro JM, Gomes XV, Godwin BC, He W, Helgesen S, Ho CH, Irzyk GP, Jando SC, Alenquer ML, Jarvie TP, Jirage KB, Kim JB, Knight JR, Lanza JR, Leamon JH, Lefkowitz SM, Lei M, Li J, Lohman KL, Lu H, Makhijani VB, McDade KE, McKenna MP, Myers EW, Nickerson E, Nobile JR, Plant R, Puc BP, Ronan MT, Roth GT, Sarkis GJ, Simons JF, Simpson JW, Srinivasan M, Tartaro KR, Tomasz A, Vogt KA, Volkmer GA, Wang SH, Wang Y, Weiner MP, Yu P, Begley RF, Rothberg JM.

Genome sequencing in microfabricated high-density picolitre reactors.

Nature. 2005 Sep 15;437(7057):376-80.

Shendure J, Porreca GJ, Reppas NB, Lin X, McCutcheon JP, Rosenbaum AM, Wang MD, Zhang K, Mitra RD, Church GM.

Accurate multiplex polony sequencing of an evolved bacterial genome.

Science. 2005 Sep 9;309(5741):1728-32.

2005/12/26

小島 健司 著
禁 無断複写転載