深圳シーケンス工場

The sequence factory

DAVID CYRANOSKI 2010年3月4日号 Vol. 464 (22-24)

大規模シーケンス施設 BGI が推進する野心的なプログラムにより、
中国はゲノム塩基配列解読で世界のリーダーとなるかもしれない。
こうしたやり方が、産業的拡大の時代に通用するのかどうか、検証する。

BGI は中国第一級のゲノムシーケンス施設である。かつては「北京ゲノミクス研究所 (Beijing Genomics Institute)」と称していたが、本部を深圳に移動した2007年より公式名称を「BGI」としている。そのBGIに、北京大学を後にした李英睿 (Li Yingrui) がやって来たのは、2006年のことだった。李は、現在23歳。そばかすのある顔には、まだ初々しさが残る。BGIの先輩によれば、李は大学時代、授業中にテレビゲームや居眠りをしていたという。これに対して李は、「講義中に寝てはいません。出席しなかっただけです」とおどける。

BGIで李は、130人のパイオインフォマティシャンチームを束ねているが、その中に彼よりも年長の者はほとんどいない。李のゲーム好きは、シーケンサーから毎日大量に出力されるデータの解析に大いに役立った。しかし李は、「ゲームよりもこっちのほうが断然おもしろいですよ。夢中になってしまいます」と話す。

BGIは、この数年間でいくつもの論文を一流の学術ジャーナルに発表し、ゲノム塩基配列解読の最前線に躍り出てきた。最近の成果としては、キュウリ¹やジャイアントパンダ²のゲノム、古代人の最初の完全ゲノム³、それに5770億塩基対の配列データからなる1000種を超える腸内細菌ゲノムの塩基配列⁴を発表している。

まるで示し合わせたかのようにスタッフが口をそろえるBGIのミッションは、一般の人にとってゲノミクスが重

BGIのシーケンサー室。生命のゲノムの系統樹を作り上げるのに数千件のプロジェクトが役立つだろう。

要なものであることを証明することだ。BGIの副所長で、コペンハーゲン大学の教授でもある王俊 (Wang Jun) は、「BGI全体がその重大な責任を感じています」と語る。その戦略は、塩基配列を解読 (シーケンス) することであり、その対象は、解読したいと思うありとあらゆるものであり、BGIだけでなく世界中の研究者たちが持ち込むものも増え続けている。BGIでは、進化の主要な枝をカバーできる生命のゲノム系統樹を作成する事業の一環として、微生物1万種のゲノムと、動植物1000種のゲノムプロジェクトを発足させた。このプロジェクトでは、イネなどの重要な種は100回以上、人間に至っては無制限にシーケンスが行われる。

そのミッションを果たすためBGIは、自らゲノミクス工場に変貌し、若いパイオインフォマティシャン軍団と高価な充実した装備をもって、高品質の塩基配列データを低コストで生み出そうとしている。こうした装備は絶えず更新され、2010年1月には、BGIは世界最新・最高速のシーケンサー「HiSeq 2000」をイルミナ社から128台購入すると発表した。このシーケンサーは、1台で1日当たり250億塩基対をシーケンスすることができる。全部がフル稼働すると、1年で1万人以上のゲノムを解読することになる。バイラー医科大学 (米国テキサス州ヒューストン) で分子生物学計算機リソースを統括するDavid Wheelerによれば、これは米国全体の

シーケンシング能力を凌駕する^{りょうが}という。「ゲノミクスの世界地図は完全に塗り替えられました」。

BGIが科学をやみくもな機械化におとしめた、という非難が深圳で怒りを買うことはほとんどない。王自身、BGIがプロジェクトに投入した知的な資本はほとんどないとおどけ、「我々は『筋肉』であって『脳』ではありません」という。しかし、そうしたコメントとは裏腹に、経験豊富な研究リーダーから新人に至るまで、BGIのすべてのスタッフには、自分たちがシーケンシング能力のバランスだけではなく、生物学、医学、農学にも影響を与えることができる、という自信が見え隠れしている。シーケンシング能力を拡張させるために抱え込んだ大きな借金を思えば、それは容易ではないだろう。科学と財務の板挟みになり、BGIをビジネスとするのか、非営利の研究施設とするのか、創設者も決めかねているようだ。世界のゲノム研究者たちは、BGIがどうバランスを取るのかに注目している。シンガポールゲノム研究所の幹部で国際ヒトゲノム機構を率いるEdison Liuは、「BGIが金だけのためにゲノムの解読を行うなら、見向きもされなくなるだろう」と警告する。

中央から遠く離れる

1990年代、世界はゲノム塩基配列解読に熱中し、それはヒトゲノムの解読へとつながった。そうした流れに中国は出遅れながらも、最終的には追いついたのだ。これは、強い意志をもち、カリスマ的で、時に人を不快にさせることもあるBGIのリーダー、楊煥明（Yang Huanming）に負うところが大きい。ヒトゲノムプロジェクトが完了に近づく中、楊たちは、なんとか中国を計画に参加させようと画策し、中国科学院の支援を得て、建設および創立用の資金として100万人民币元（約1400万円）を確保した。そして、1999年9月9日、9時9分9秒、王によりBGIの設立が発表された。「中国では『9』は『長く続くこと』を意味

する縁起のよい数字であり、BGIの末永い繁栄を願ったのです」と王は語る。同年11月、中国政府から交付された300万人民币元（約4100万円）の補助金のほとんどをBGIが受け取り、ヒトゲノムの1%の塩基配列を解読するのに使われた。中国は、この国際プロジェクトに参加した唯一の開発途上国で、担当する3000万塩基を1年足らずで片付けた。

2000年、資金不足の中、デンマークの研究機関のためにブタゲノムの「スキャン」を完了させたとき、楊は「もっと重要なこと」をやろうと決意したという。そして2001年BGIは、イネのゲノムプロジェクトを発足させた。このプロジェクトは杭州市から6000万人民币元（約8億2000万円）の補助金を受け、当時最新鋭のシーケンサー36台が購入された。2002年BGIは、インディカ種のイネのゲノム塩基配列を*Science*に発表した⁵。それは国際共同研究チームがジャポニカ種のゲノム塩基配列を発表する数か月前のことだった。

続いてBGIは、ニワトリ⁶やカイコガ⁷のゲノムDNAを解読し、2003年には重症急性呼吸器症候群（SARS）の原因であるコロナウイルスのゲノムRNAも解読して⁸、診断キットを世に送り出した。これには、胡錦濤中国国家主席もご満悦だった。BGIへのご褒美は、中国科学院への参画だった。この栄誉により資金提供を受けられるようになったが、制約も受けるようになった。中国科学院の研究施設では、スタッフの数は150人までとされている。しかし、当時BGIにはその2倍が在籍しており、さらなる増員も見込んでいた。楊は、BGIのスタッフの一部を中国科学院の正規雇用とし、それ以外の者は別契約とすることで、中国科学院からの予算を最大限に引き出したが、「誰も満足しませんでしたよ」と楊はいう。

深圳に移ることはガス抜きとなった。楊は、北京の監視から少し離れたたいと考えていたところだった。それに、初期費用1000万人民币元（約1億4000万

大量生産

過去10年、BGIはこれまでにない速度で、数百万～数十億塩基対のシーケンスを行ってきた。下に示すのは、その成果の一例である。

インディカ種のイネゲノム

4億6600万塩基対
カバー率：6回
作業期間：2000年4月～2001年10月

カイコガゲノム

4億8000万塩基対
カバー率：6回
作業期間：2003年6月～9月

初めてのアジア人ゲノム

30億塩基対
カバー率：36回
作業期間：2007年1月～10月

パンダゲノム

30億塩基対
カバー率：73回
作業期間：2008年3月～10月

古代グリーンランド人

30億塩基対
カバー率：20回
作業期間：2009年5月～12月

円)と年間補助金 2000 万人民币元(約 2 億 7000 万円)も魅力的だった。深圳は中国南部に広がる「世界の工場」の中心で、人口 1200 万人の多くが安価な衣料品と電気製品の製造に携わり、中国経済の奇跡に一役買っていた。楊は、自分が北京で衝突した人について多くを語らないが、「深圳では、山は高く、皇帝は遠い」が口癖だ。

拡大する BGI

BGIにとって深圳は心地よい場所である。その中で、BGIは成長し、雇用人数は全国で 1500 人(その 3 分の 2 以上が深圳で従事)になり、年内に 3500 人まで急増させる計画だ。中国開発銀行から 100 億人民币元(約 1400 億円)の融資を受け、新しいシーケンサーの購入に充てられた。BGIのシーケンシング能力が拡大するにつれ、コストも増大する。BGIのスタッフは、新しいシーケンサーにどれだけ支払ったのか語らないが、定価は 1 台約 340 万人民币元(約 4600 万円)だ。その購入の発表が、新型モデルの発売当日だったために、競合するほかのゲノミクス研究施設から不満の声が上がった。彼らによれば、イルミナ社が旧型モデルを他の施設に販売する一方で、BGIと秘密裏に取引を進めていたというのだ。しかしながら同社はそうした主張を否定し、アップグレードを希望する顧客向けの下取りプログラムの 1 つだと説明している。

新しいシーケンサーのうち 100 台は香港の新研究所に置き、国際共同研究に役立てる計画だ。しかし、香港の人件費はこれまでの BGI の相場よりも高く、40～50 人ほどの雇用となるだろう。さらに試薬類に毎年約 10 億人民币元(約 140 億円)かかり、電気代として別途 900 万人民币元(約 1 億 2000 万円)が必要だ。楊は、融資は完済できると強調する。しかし、シーケンスサービスが商品化され、どんどん価格が下がってきており、返済は不透明だ。

BGI がシーケンスサービスの市場を独占する、と決まったわけではない。

学術研究機関のほかに、新技術を使う民間企業も競合相手になりつつある。ヒトゲノムシーケンスサービスに特化したコンプリート・ゲノミクス社(米国カリフォルニア州マウンテンビュー)は、4 月にスタートしたが、2010 年だけで 5000 件のヒトゲノムの塩基配列解読を見込んでいる。

BGI の支払い能力は、ゲノムのシーケンスを、代金を支払っても、自分がするよりも迅速かつ確実にしてくれる BGI に委託する外部の研究者に依存している。しかし、多くのシーケンス施設と同様、BGI は単なるサービス提供者の域を越えようとしているようだ。王を育て、BGI とは縁深いワシントン大学(米国シアトル)のゲノミクス研究者 Maynard Olson は、そうでなければならない、と話す。「外注がうまくいくのは、発注側と受注側に何らかの科学的なつながりがある場合に限られます。BGI にはさまざまな解析技術がありますが、すべてを駆使するわけにはいかないので、どんなレベルのデータ解析でもどの技術で行うのかお互いが合意する必要がある、できあいの商品のように扱うことはできません」。

シーケンスにかかるコストの半分を共同研究者に負担してもらって、共同で成果を発表することが理想であるが、共同研究者に資金がなくてもおもしろいプロジェクトならば、BGI がそのほとんどを負担してもよい、と楊はいう。

コペンハーゲン大学の Eske Willerslev は、楊の発言は道理にかなっていると感じている。Willerslev は、永久凍土から発掘された髪の毛から、4000 年前の古代グリーンランド人のゲノムシーケンスを BGI と共同で行った。Willerslev 自身の研究室でも 1 週間で約 500 億塩基をシーケンスする能力があったが、彼は BGI の専門技術を求めた。「次世代のシーケンスプラットフォームの能力を十分に引き出せる BGI のスタッフを、私は尊敬しています」と Willerslev はいう。古代人のゲノムは 2 か月半で解読され、約 50 万ドル(約 4700 万円)の費用は、

Willerslev のスポンサーと BGI で折半された。

Willerslev によれば、BGI は、シーケンス作業だけではなく科学的な面でも不可欠な存在だったという。「王俊と私とで重要な科学的問題に関して折り合いがついたので、プロジェクト全体が進められることになったのです」。

科学のためか、サービス提供者としてか

BGI の「科学」が力任せのシーケンスサービスにとどまるものでないことを示すのは、容易でないかもしれない。華南理工大学(中国広東省広州市)の学生ながら、BGI に所属する弱冠 21 歳の羅鋭邦(Luo Ruibang)は、博士課程の教育を受けていない。彼は、国際学会で自分が最高レベルの研究ができることをほかの研究者に理解してもらうのに苦労したという。「本当に私に能力があるのかどうか、疑問に思う外国人が大勢います」。しかしながら、アジア人とアフリカ人のゲノムにはあって白人のゲノムにない大きな DNA 領域の発見を発表した論文⁹では、羅と李が共同第一著者だった。

李をはじめとする BGI の研究リーダーたちは、こうした若者たちでも塩基配列をつなぎ合わせ、きちんと検証できると確信している。「これは新しい分野なのです。ですからどのみち、多くの経験が蓄積されていないのです」と王はいう。しかし、データの解釈と実験の立案は別物である。BGI のスタッフも、実験の知識が足りないことを認めている。「我々は生物学について知らないことが多いのです」と李はいう。Liu は、BGI は生物学的側面の弱点を解消する必要があるといいつつも、BGI の研究に協力的だ。「BGI は基本的にシーケンス屋だが、強力な武器をもった賢いシーケンス屋です」と語る。

ある意味で、パンダゲノムの解読は、BGI の強さを世界に示す手段だった。パンダを研究対象とする生物学者はほとんどおらず、そのゲノム塩基配列が基礎および応用研究の両面で飛躍的な進

歩につながるとは考えにくい。しかしそれは BGI にとって、次世代シーケンシング技術の威力を示す機会となった。従来の技術では、1000 塩基対の DNA をシーケンスできるが、ある程度時間がかかる。一方、次世代シーケンシング技術では、シーケンスできる DNA こそ 1 つが 100 塩基対以下という短いものだが、そのスピードが非常に速い。ただ、王によれば、シーケンスできる DNA は数千倍に増えても、研究の進んでいないゲノムでは、最終的なゲノムの形に構築するのが難しいという。「BGI では無理だろうと考える人もいました」。今回ゲノムが解読されたパンダが北京オリンピックのマスコットだったことで、一般からも関心が寄せられ、地方政府の支援も得られた。「それに、パンダはかわいいですからね」と王はいう。

BGI は、外部の研究チームが主導する生物学的に重要性が高いプロジェクトでも強力な共同研究者となっている。例えば、昨年 10 月に発表されたカイコガ 40 系統のゲノムからは、交配と飼育化の歴史を示す遺伝子が約 300 個発見された¹⁰。最近では、チルー（チベットレイヨウ）のゲノムの解読が完了した。チルーは、海拔 4500 メートルを超える高地で走り回ることができ、高所への適応についての手がかりが得られるかもしれない。青海大学（中国青海省西寧市）で高所医学を研究する格日力（Ge Ri-Li）は、チルーのエネルギー輸送に関係する遺伝子を研究しようとしている。そして、高山病が少ないチベット族のゲノムを解読し、最終的には、人間が高所に適応する能力を高めたいと思っている。

研究だけでは、100 億人民元（約 1400 億円）の銀行融資を返済することはできない。確かに、BGI の収入のいくらかは、総シーケンス作業の 40% を占めている共同研究からのものである。しかしそれ以上に利幅が大きいのは、大学や種苗会社、製薬会社向けのシーケンスサービスで、仕事の 55% を占める（残りの 5% は BGI 独自のプロジェクト）。

ゲノミクスのニューフェース。左から、王俊副所長、羅鋭邦、李英睿。

こうした作業から得られる総収入は、2009 年は 3 億人民元（約 41 億円）だった。しかしまだ不十分だ、と BGI のマーケティング担当役員を務める梁鴻生はいう。2010 年、梁は 12 億人民元（約 160 億円）の収入を見込んでいる。

農業用の知的財産権から収入が得られる可能性もある。200 件を超える特許を保有する BGI は、河北でゲノミクスを利用したアワの育種に取り組んでおり、ラオスでも農業プロジェクトを進めている。海外サービスの拡大により、さらに収入が得られるかもしれない。3 年以内に、また、コペンハーゲンとサンフランシスコに事務所を構える計画だ。BGI は、2008 年に発足した、中国人 100 人のゲノムプロジェクト「炎黄」データベースへのアクセスに課金することも考えている。BGI では、収録データを数千人に増やしたいと考えている。楊によれば、アクセス料はコンピューターの代金と維持費、つまり「原価」をカバーするものであり、データの代金ではないという。

BGI が最先端を走り続けるためにはためめ技術革新が必要であり、実入りのよい共同研究を維持しなければならないのだ、と Wheeler は語る。Wheeler によれば、幅広いシーケンシング技術を試験できる大きな国家的施設が、米国には 3 か所あるという。「それらの施設どうしが、技術の向上をめざし、絶えず互いに刺激し合っています。協力しつつも、効率や分析方法を改善するため

に互いを批評し合っているのです」。単一の技術に立脚する単独の研究施設は、その技術が時代遅れになってくると、強みを失うだろう。

楊も技術進歩の不確定要素を認識している。今回、新型の次世代シーケンサーへの投資に二の足を踏まなかったのはなぜか、との問いに対し、「保有する設備はしかるべきときに取り換えるだけです」と説明する。しかし、あまりに早く買い換えときが来るようなことになれば、運が尽きたということだろう、とも話す。

「リスクは多くても、私はあえてリスクを取ることを評価します」と Olson はいう。共同研究と新技術を果敢に追求することにより、BGI の野心的な取り組みが世界から注目を集め続けることは間違いない。「要するに、BGI は刺激的なことをしているということです。これは、強力な科学の開発という課題に対する中国式の解決法であり、どれだけ機能するかは、いずれわかることでしょう」。

（翻訳：小林盛方）

David Cyranoski は *Nature* のアジア・パシフィック地域の特派員。

- Huang, S. et al. *Nature Genet.* **41**, 1275-1281 (2009).
- Li, R. et al. *Nature* **463**, 311-317 (2010).
- Rasmussen, M. et al. *Nature* **463**, 757-762 (2010).
- Qin, J. et al. *Nature* **464**, 59-65 (2010).
- Yu, J. et al. *Science* **296**, 79-92 (2002).
- Hillier, L. W. et al. *Nature* **432**, 695-716 (2004).
- Xia, Q. et al. *Science* **306**, 1937-1940 (2004).
- Qin, E. et al. *Chinese Science Bulletin* **48**, 941-948 (2003).
- Li, R. et al. *Nature Biotechnol.* **28**, 57-63 (2010).
- Xia, Q. et al. *Science* **326**, 433-436 (2009).