完全ゲノムのこだわりとドラフト解析の使い方

宮崎大学フロンティア科学実験総合センターの林 哲也 センター長をインタビューしました。

先生のご研究の目的について教えてください。

基本は病原菌を対象とした研究ですが、病原細菌を中心とした様々な細菌のゲノム多様性の実態がどのようなものか、そしてどういう形でそうした多様性が生まれてくるかの2つを大きな研究目的としています。

病原菌の場合は、実際の臨床の場では様々な株が分離されてきます。なんとか菌というと1種類というイメージしかないかもしれませんが、例えば同じ大腸菌、あるいは同じヘリコバクター菌でも、バクテリアの株は実に多様です。ただ、そういった多様性が菌種によってはあまりはっきりしないケースもあるし、すごく明確なケースもあります。同じ菌種でも、全く病原性がないものと病原性をもつもの。病原性があっても全く違う病気を起こすタイプ。異なる病気を起こすというような違いはなかったとしても、非常に病原性が強かったり、弱かったりする株など。おそらく、ほとんどの病原菌種でそういう違いはあると考えているのです。そしてそれはゲノムの違いに由来するだろう、と。細菌のゲノムはそれだけ多様化しているので、変化のメカニズムや実際に菌株間でゲノムがどう違うかをまず見ないと始まらないですよね。

また、あるクローンが、時間的・地理的にどう広がっていくのか、それに伴ってどのような多様化が生じるのかを探索するのも僕らの目標でもあります。ある変異を持った株がドミナントに広がる場合もあります。僕はもともと病原性のメカニズムやタンパク毒素から研究をスタートとしているので、もちろんある特別な病原遺伝子、その遺伝子産物と宿主とインターラクション、そしてその結果何が起きるかというような病原性解析にも当然興味があります。ただ、そういった研究には多くの人が取り組んでいるので、うちではそういった個別研究よりは、もっとゲノムから病原菌を見てみよう、ゲノムがどう広がっていくかを考えようとしています。実際は、海外にはそういうことをしているグループはたくさんいるので、非常に競争が激しいのですが、国内にはそういう目で研究を進めている人がそれほどいないので、最低うちはやるぞと思っています。

僕らが最初に手を出したのは大腸菌で、その一番目が腸管出血性大腸菌O157でした。みなさん、たぶん名前聞いたことがあると思いますけど、去年富山で起きたようなO111とか、他にも似たような大腸菌がいるわけですよね。それでは、その違いはどうなんだ、と調べるわけです。同じ大腸菌でも、違う病気を起こすものが最低7～8種類もあります。ただ、全部をうちがやるわけにはいかないので、メインは腸管出血性大腸菌に絞っています。

O157は腸管出血性大腸菌で一番主要なのですが、その次がO26、そしてO111、O113、O145、O121など。一番大事な最初の4つはうちのラボでゲノムを決めたました。また、最初にゲノムを決めたO157については、様々なO157株のバリエーションを調べて、ゲノムがどう違うか、そしてなんで違いが生まれたかをはっきりさせるための一連の仕事もしています。

林 哲也 センター長
（宮崎大学フロンティア科学実験総合センター）

次の段階としては、地域とか年代を経てどのようにクローンが変化するか、という点を調べたいなと思っています。こうしたことをずっとやりたくていたのですが、サンガーの知り合いに先を越されて、サイエンスやネイチャーなどに論文が出て、私は去年がっかりでした。ちょうど株を集めてスタートしようと思っていた時だったので。

僕は完全ゲノムっていうものにも、非常にこだわっています。ドラフト配列は簡単に取得できるようになってきていて、コアゲノムの情報は簡単に抽出できるのですけれども、リピート部分とか、似た様なファージが複数あったら、そういった情報についてはドラフトではお手上げなんですよね。例えばO157の中でどこがどう変化するかっていうのは、ドラフトの情報では一部しかわからない。O26で菌株間の比較をしましょうって言ったときに、完全配列がなければリファレンスにならない。だからリファレンスになるべき高精度の完全ゲノムの配列は、絶対必要だと考えているわけなのです。例えば、O157のゲノム上には18種類のプロファージがあるというようなことは、完全ゲノムがないと見えてこない。ところが、ドラフトから完全ゲノムに仕上げる研究チームの数はどんどん減ってるんですよ。なので、逆にうちはそこにこだわろうかなと思っているわけです。

「僕は完全ゲノムっていうのに、非常にこだわっています。」

MiSeqはどのような用途で使う予定ですか？

まずは、もともと持っている454のデータから完全ゲノムを出したときにどうしてもエラーが出てきますから、1回はMiSeqをランしてエラーを修正する、という使い方。次に、例えば全国からある菌種のある系統を数百株集めて、その多様性を調べるときにはすべて完全ゲノムをするのは無理なので、MiSeqを使おうと。そのほうが一気に情報が取れるし、その場合はドラフトでいい。そういう意味での完全ゲノムとドラフトの使い分けを僕はしたいなと思っています。

あとは、フォスミドなどをある程度の数集めて、今まではキャピラリーで個別に解析していたようなことをMiSeqでまとめてやろうと思ってます。いくつまでできるか、ちょっとトライしないといけないのですが。例えば真核生物なら、それこそトランスポゾンというかレトロトランスポゾンの類が出てきてしまうし、またバクテリアの種類によってはリピートの種類や数にはいろいろなので、いくつまでできるかはトライしてみないと分からないですが。これは最初454を考えて準備もしていたのですが、実はライブラリー作りに時間がかかってしまうので、いまではMiSeqにしようかなって思ってます。

MiSeqを選んだ理由は？

まず、いままで私たちが使っていた454と違うケミストリーのものが必要だったこと。それから、ハンドリングの簡便性とコスト。この3つが一番大きいですね。MiSeqは圧倒的に低コストが魅力的。なので、この後250bpまでリードがのびて、さらにメイトペアが使えるようになると実は454でないとダメというようなことが少なくなるかもなって思ってもいるのですけど（笑）。MiSeqのデータを見せてもらう限りは、まあ、自分たちのデータ見てもそうなんですけど、やっぱりクオリティーが圧倒的によいと思います。ちょうど私たちはバクテリア以外にも真核生物も手をだしているので、クオリティーはやはり重要です。あと、プレパレーションの時間、速さ、簡単さ。ハンドリングは圧倒的に楽だと思います。

大型のHiSeqシステムはあればあったでいいと思うけど、僕はコストパフォーマンスとそこから得られるデータを考えたときに、うちの大学だけで考えたらMiSeqの方が絶対いいと考えました。うちの大学では需要を考えても、大型のプロジェクトはなかなか無いんじゃないかなと思うので。ただ、トラブった時にはMiSeqが2台あるといいかな、と（笑）。なので安定稼働できればいい。

データ処理の負担は今のところ大丈夫です。遺伝研のデータ解析サービスや共同研究者と一緒に*de novo* などを行っているので、我々は最低限のことをやっています。ただし、本当のことを言うと、単純な配列処理以上に、例えば本当の比較ゲノム解析などのほうがはるかに難しいのですよね。だからそちらに関しては、バイオインフォのプロに助けてもらって、一緒に解析を進めています

「まず、いままで私たちが使っていた454と違うケミストリーのものが必要だったこと。それからハンドリングの簡便性とコスト。この3つが一番大きいですかね」

MiSeqシステムの詳細については以下まで。
www.illuminakk.co.jp/miseq

イルミナ株式会社

〒108-0014
東京都港区芝5-36-7　三田ベルジュビル22階
Tel (03)4578-2800　Fax (03)4578-2810
www.illuminakk.co.jp

代理店

本製品の使用目的は研究に限定されます。

© 2013 Illumina, Inc. All rights reserved.
Illumina, illumina*Dx*, BaseSpace, BeadArray, BeadXpress, cBot, CSPro, DASL, DesignStudio, Eco, GAIIx, Genetic Energy, Genome Analyzer, GenomeStudio, GoldenGate, HiScan, HiSeq, Infinium, iSelect, MiSeq, Nextera, NuPCR, SeqMonitor, Solexa, TruSeq, TruSight, VeraCode, the pumpkin orange color, the Genetic Energy streaming bases design は Illumina, Inc の商標または登録商標です。
その他の会社名や商品名は、各社の商標または登録商標です。予告なしに仕様を変更する場合があります。

Pub. No. 770-2011-J014 11MAY12