

## Come fly with us

# ゲノミクスの使徒となったショウジョウバエ

Ewan Birney



ショウジョウバエ 12 種のゲノム比較解析が行われたことは、我々にとっても意味がある。ハエからヒトまでの進化の間に選択圧に耐えて残ってきた塩基配列は、機能的に重要であるに違いないと考えられるからである。

Nature Vol.450 (184-185) / 8 November 2007

キイロショウジョウバエ (*Drosophila melanogaster*) は、遺伝学や分子生物学の分野では常に人気者である。ありふれた生き物でありながら、多細胞生物の生物学を解明する手がかりを提供し続けてくれるからである。しかし、1990 年代には線虫 (*Caenorhabditis elegans*) がゲノミクスの主役となり、我々ヒトとともに「主賓」扱いだったショウジョウバエは、「主賓リスト」から外されてしまいそうになった。

こうした状況が変化し始めたきっかけは、Celera Genomics 社<sup>1</sup> が、大型生物種のゲノム解読に取り組む前にキイロショウジョウバエのゲノム塩基配列を解読したことだった。進化ゲノミクスの時代に入り、今回、ショウジョウバエ属の 10 種の塩基配列解読と 12 種のゲノム比較解析について、*Nature* 2007 年 11 月 8 日号<sup>2,3</sup> とほかのジャーナルで 40 以上の関連論文が報告された。これで、ショウジョウバエは研究対象としてほかの生物種に追いつき、ゲノミクスの分野でも人気者の地位を得た。

1 つの生物のいかなる側面も、進化の過程で出現して維持され続けてきたものである。したがって、ゲノムを読み解くため、なかでも生物種間で保存されているタンパク質をコードする遺伝子を見つけ出すために、進化解析が常に用いられてきた。しかし、進化の

過程は、1 つのゲノムに含まれる遺伝子を分類整理することだけによらず、効率的に研究できる。具体的にいうと、研究者は進化の相補的な 2 つの側面、すなわち負の選択と正の選択を研究することができるのである。Stark たち<sup>2</sup> は、負の選択、すなわちゲノム内の機能要素の存在を調べた。負の選択では、ランダムな変異を多数こらむっても機能が変化しない (図 1a)。これに対して、*Drosophila* 12 ゲノムコンソーシアム (Clark たち)<sup>3</sup> は、正の選択、すなわち異なる種で新たに獲得された機能を調べた (図 1b)。

ショウジョウバエ属は種間に著しい多様性がみられ、こうした比較解析にうってつけの研究素材である。結果的に今回の解析では、*D. simulans* (右の写真) と *D. sechellia* のように非常に近縁な種の組み合わせ (ヒトと、ヒトに非常に近い霊長類との遺伝学的な距離に相当) に加えて、*D. grimshawi* のように類縁関係の離れたショウジョウバエも調べられた。*D. grimshawi* は多数あるハワイ産外来種の 1 つで、実験用の普通のショウジョウバエに比べて体の大きさが 100 倍もあり、両者の遺伝学的な距離はヒトとトカゲの距離に相当する。

Stark たち<sup>2</sup> は、機能要素を新たに見つけ出したり既知の要素について解明を進めたりするために、機能要素に関する現在活用できるほとんどの知識を駆使

し、なおかつ自然に存在する範囲での変異や選択を利用した。Stark たちは、よく解明されているタンパク質コード遺伝子から、遺伝子発現を調節するあまり解明されていないモチーフまで、知られているすべての種類の機能要素について検討している。これらの解析によって、キイロショウジョウバエの特異的なゲノム配列のために不正確だった生物学的情報を特定することができた。

Stark たちは、タンパク質をコードする DNA の連続した塩基配列内に埋もれている、進化的に保存された要素をいくつか発見した。それらの中には、終止コドン（1つのタンパク質の配列終了を知らせる塩基3個の配列）や、フレームシフト変異（タンパク質コード配列の読み取り枠がずれる）などがある。例えばメッセジャー RNA へ転写された終止コドンがタンパク質翻訳の段階で無視されるような、こうした遺伝子構造は、正常な翻訳ルールと両立できるとはちょっと考えにくい。そのため、これらの知見からは、翻訳の前段階に何か未知の mRNA プロセッシング機構が存在しているか、もしくは、別の翻訳様式が存在していることがうかがわれる。

マイクロ RNA は、自然に存在する短い一本鎖 RNA で、遺伝子発現を調節する。Stark たちは次に、マイクロ RNA などの非コード（タンパク質をコードしない）RNA 配列の遺伝子を調べ、新しいマイクロ RNA 配列を発見した。これによりキイロショウジョウバエのマイクロ RNA のリストは、74個から101個へと拡大した。

Stark たちは、調節モチーフという、もう1種類の機能要素も調べた。そして、調節モチーフの非常に詳しい「辞典」を作成し、この種のモチーフが機能を果たしていると思われる一群の例（1つの生物のゲノム規模では初めて）も示した。ゲノミクスを用いて調節モチーフの活性例を見いだすというのは、実に画期的な新手法である。Stark たちは「branch length score (分枝の長さ指数)」とよばれる方法を用いた。この方法は、実際のデータでよくあるアラインメント（塩基配列の整列化）や配列決定の際の誤りに配慮したもので、1つの系統樹全体に適用できる。Stark たちは、自分たちの手法をさまざまな統計学的側面から慎重に査定し、ショウジョウバエを研究する分子生物学者や遺伝子調節に取り組む研究グループが安心して使える「機能要素の金山」を提供してくれたのである。

正の選択の側面に関する Clark たちの知見<sup>3</sup>は、Stark たちの研究ほど、キイロショウジョウバエの分子

生物学研究への直接利用ができない。しかし Clark たちの成果は、進化の最中に生物がどのように生じてくるかをゲノム規模で知るための包括的な手がかりを、初めてもたらしてくれた。Clark たちの解析は、統計学的な意味では Stark たちのものほど威力はないが、それは当然である。Clark たちの目的は、正の選択を解明することにあるからだ。正の選択は、ショウジョウバエのさまざまな系統を越えて非連続的に起こるが、Stark たちの調べた負の選択は、それに比べて一定しており、全データ群にわたって容易に集計できる（図1）。

それでも、Clark たちの成果はショウジョウバエの種の進化を解明する糸口となる。例えば、ショウジョウ



*D. simulans*



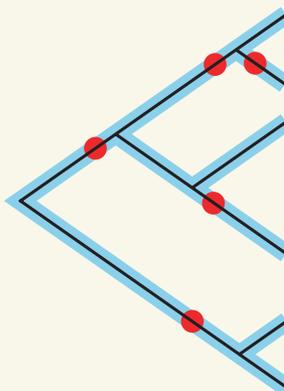
今回新たにゲノムが解読された10種と、既に解読済みの2種を合わせた12種のショウジョウバエ（麻酔をかけた状態）。

## a 負の選択

1 ATGGGGTTG---GTTTCC 種A  
 ATGGGTTTTGTGGTTTCC 種B  
 ATGGGATTAGTGGTCTCC 種C

2 ATAGAGTTGG--GTCTAC  
 ACATTTTT--GTGGAATAC  
 AGGCAATTAG--GTGCC

3 ATGAGGTTG---GTTTCC  
 AGGT--TGGTGGTTCTCC  
 ATGGGATTAGGTTGCTCC



## b 正の選択

4 ATGGGGTTG---GTTTCC  
 ATTTGTTTTGTGGTTTCC  
 ATAGCAATTAGTGGTCTCC

5 ATCGAGTTGG--GTCTGC  
 ACATTTTT--GTGGAATTC  
 AGGCAATTAG--GTGCC

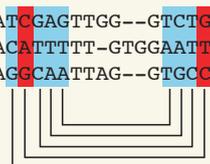


図1 進化上の2種類の選択。a, 進化の解析として、Stark たち<sup>2</sup>は負の選択(青色)を研究した。負の選択では、特異的な塩基(A、T、C、Gの4種類)が全系統のゲノムを通じてほぼ一定に維持され、ゲノムの機能要素の保存を確実なものにしている。このような解析では大きく3つの方法を用いる。タンパク質をコードする保存された塩基配列を見つけ出すこと(1)、タンパク質をコードしない非コードRNA遺伝子で、保存された対合塩基を見つけ出すこと(2)、アラインメントで、保存された特異的モチーフを見つけ出すこと(3)である。b, それに対してClark たち<sup>3</sup>は、正の選択(赤色)の例を探した。正の選択では、異なる種において特異的な塩基の修飾が起こり、新しい機能の獲得につながる。正の選択を調べるには主に2つの手法が使われる。負の選択を受けるコドン群に埋もれた急速進化するコドンの特定(4)と、非コードRNA構造の背景の中で急速に進化する塩基対合の探索(5)である。図の中央にある系統樹は種の系統発生を示しており(すべての枝を図示してはいない)、負の選択は相対的に持続的だが、正の選択は断続的であることがよくわかる。2と5にある黒色の線は、二次構造での塩基対合を表している。そのため、線をつないだ塩基の配列上の位置関係は青色の各カラム上では(一次元では)保存されていないかもしれないが、線をつないだ位置の塩基対合は、二次構造において適切な位置で塩基対合を保持している。

ウバエ属の12種のゲノム(そのうちの10種は今回配列解読がなされて報告されたもの)を比較すると、ゲノム再編成が大規模スケールでも小規模スケールでも著しく高頻度で見られることが明らかになった。また、遺伝子のうち約3分の1が、少なくとも1個のアミノ酸の位置に影響する変異を通じて正の選択を受けたこともわかった。これは、1つのゲノムにおいて多数の遺伝子で正の選択が起こったことを意味している。

1つの生物種において、1種類のアミノ酸を指定する全コドンのうち特定のコドンが選択的に使われることを、コドン使用頻度の偏り(codon-usage bias)といい、生物が異なれば、この偏りも異なっている。Clark たちは、ショウジョウバエ属の*D. willistoni*種をほかの複数の種と比較し、この種ではコドン使用頻度の偏りがかなり小さいことを発見した。

Clark たちは、さらに、嗅覚や免疫にかかわるタンパク質類をコードする遺伝子(タンパク質コード遺伝子の中でも正の選択が働くと考えられている)が、ゲノムのほかの部分より速く進化してきたことも明らかにした。殺虫剤耐性といった、ショウジョウバエの生理の特異的な部分を制御する遺伝子にも、高速進化がみられた。ショウジョウバエは人類との長いつきあいの中で、殺虫剤の導入から人類移住に伴う種の移動まで、さまざまな環境の激変に耐えてきた。したがって、ショウジョウバエの進化史に起こったこうした出来事とゲノムの変化とを相関づけるような、さまざまな興味深い研究がおそらく期待できるだろう。

今回の2つの研究<sup>2,3</sup>の知見は、特にヒトゲノムのさらなる研究に、どのような広い意味をもつのだろうか。負の選択の場合、Stark たちが使った進化ゲノミクス

の手法によって、1つのクレード（近縁な生物からなるひとまとまりの系統分類群）全体で保存されている機能要素について、目覚ましい洞察がはっきりと得られる。現時点でかなり進行している哺乳類ゲノムプロジェクト案<sup>4</sup>は、Starkたちのデータセットとほぼ同じくらい統計学的に威力があるようである。これによって、哺乳類の大規模スケールのゲノム解析に適用可能で、実験的手法を補完するような一連の非常に有効な探索手法が手に入ることになるだろう。これらの手法を使って研究成果を出すには、理論上、ショウジョウバエ系統でも哺乳類系統でも質的に違いはないはずである。とはいえ、哺乳類ではゲノムのサイズがショウジョウバエよりも大きく、また、1種のゲノム全体でも異種のゲノム間でも中立進化の速度の揺らぎがショウジョウバエの場合より大きい可能性があり、いくつかの興味深い問題を乗り越えねばならないかもしれない。

研究者が懸念しているのは、進化ゲノミクスの手法を使って得られたデータが、ChIP-chip法やChIPseq法などの実験的手法で得られたデータと完全に重ならないのではないかということだ。ChIP-chip法やChIPseq法は、転写因子結合部位やその他の機能的なDNA要素を、*in vivo*で包括的にマッピングする手法である。特に、これらの実験的手法では、進化ゲノミクスの標準感度では進化的に保存されたものと判定されていない要素群を定義づけることが多い。以前の研究報告<sup>5</sup>やStarkたちによって論じられているように、このミスマッチはどうやら、種を越えて、またさまざまな研究室で行われた解析を越えて一貫してみられるようである。そのため、これはおそらく、系統特異的な要素が大量に存在するせいでも、あるいは使われた手法に欠陥があるせいでもなく、選択が働く前に、一見すると中立の進化過程がどうやって生化学活性をもつ新規の要素を生じるのかについて、我々の理解が不足していることを反映しているのだろう。

Clarkたち<sup>3</sup>による正の選択の解析は、多細胞生物のクレードで行われたものとして、間違いなく最も広範で最も詳細な研究である。彼らの研究により、種間の差異、つまり進化がどのように適応的变化につながるのかを解明するには、用いる手法を向上させねばならず、より大規模なデータやより広範な種を考察しなければならないことがよくわかる。こうした論拠は、もっと多くの種のゲノム塩基配列解読を行う（現在は経費の安い塩基配列解読技術が登場したおかげで実現

可能である）ことや、各クレードのほかの構成種について実験研究を行うことを後押しする理由づけとなる。こうした研究は、塩基配列の変化と機能要素の変化とを相関づけたり、開発された新しい手法を検証したりするために不可欠である。

ショウジョウバエ属についていえば、次の段階では、さらに多くのショウジョウバエ種や双翅目のほかの種のゲノム塩基配列解読を行い、それらを、今回論じているショウジョウバエ属や遠縁にあたるカの仲間の配列解読済みゲノム<sup>6,7</sup>に加えるべきである。さらに、個体群レベルでもっと多数のゲノム塩基配列を調べるべきである。つまり、同一種の数個体の塩基配列を解読して、進化解析データとの比較用に使えるような古典的な集団遺伝学解析の材料を集めるべきである。こうした解析用リソースを作り出そうとする試みは、ショウジョウバエ属のいくつかの種で順調に進んでいる。最後に、進化解析によって得られたデータと比較するには、実験に大きく貢献してきたキイロショウジョウバエを超えるようなほかの種で、新しい実験結果を得ようとする協調的な努力が必要である。

Clarkたちの知見からすると、ヒトに固有のものも含めた霊長類間の興味深い適応的变化を解明するには、個体群レベルのデータと高精度のゲノム塩基配列の双方を得るのに最適な塩基配列解読戦略を用いて、あらゆる現生霊長類（そして可能な場合には、DNA復元の可能な絶滅霊長類も）のゲノムの塩基配列解読を行う必要がおそらくあるだろう。機能要素の変化と塩基配列の変化とを相関づけるには、厳選した霊長類種から樹立した細胞株を用いた基礎的な分子生物学研究も必要になるだろう。

今回の成果に話を戻そう。Starkたち<sup>2</sup>とClarkたち<sup>3</sup>が提示し解析したデータは、進化ゲノミクスの威力を示した特筆すべき初めての例である。この分野は、今後10年間の中心的な研究テーマの1つとなるだろう。これで、ゲノムミクス研究者は、ショウジョウバエのファンクラブで、同僚の遺伝学者や分子生物学者とようやく肩を組めるようになったわけである。 ■

Ewan Birney、欧州バイオインフォマティクス研究所（英）

1. Adam, M. D. *et al. Science* **287**, 2185–2195 (2000).
2. Stark, A. *et al. Nature* **450**, 219–232 (2007).
3. *Drosophila* 12 Genomes Consortium *Nature* **450**, 203–218 (2007).
4. [www.broad.mit.edu/mammals](http://www.broad.mit.edu/mammals)
5. The ENCODE Project Consortium *Nature* **447**, 799–816 (2007).
6. Holt, R. A. *et al. Science* **298**, 129–149 (2002).
7. Nene, V. *et al. Science* **316**, 1718–1723 (2007).