

時系列蛍光データを用いた表現型データベースの構築

●園池 公毅

東京大学大学院新領域創成科学研究科

<研究の目的と進め方>

塩基配列情報は、バイオインフォマティクスの手法とっていわば「扱いやすい」情報である。これは、どのような生物であっても塩基配列情報は4種の文字の一次元配列として扱うことができ、結果として塩基配列情報解析は、多くの生物に同一の手法を適用して行なうことができることに起因している。しかしながら、ゲノム配列の決定を中心としたゲノム研究が一段落し、ゲノム上の遺伝子の機能解析が主要な問題の一つであるポストゲノムの時代に入り、そのような研究の普遍性は失われつつあるように見える。遺伝子の機能解析に不可欠な変異株の表現型情報は、その多様性と定性的な性質のため一般的に大量解析になじまない。さらには、ある特定の表現型に対して、ゲノムワイドに表現型情報を蓄積したとしても、別の表現型については、全く別のフォーマットの情報を蓄積せざるを得ず、研究対象の生物種、あるいは研究対象の生物現象によっても、その表現型に関するデータのフォーマットはばらばらにならざるを得ない。このことが、いったんは塩基配列という共通言語を手に入れたと思われた生物学が、再びバベルの塔の状況を経験している大きな原因だと思われる。

本研究においては、このような状況を部分的にでも解決するため、シアノバクテリアというモデル生物において、多くの異なる表現型を、単純なデータフォーマットに基づく単一の表現型によって記述することを目的とする。具体的には、単細胞原核光合成生物であるシアノバクテリアを材料とした場合には、光合成反応をいわば道具として使うことにより、広い範囲の遺伝子変異株の表現型情報をクロロフィル蛍光強度の一次元の数値データとして、統一されたフォーマット上に蓄積することが可能であることを利用する。実際に、我々は、蛍光強度の時系列データという極めて単純な単一の表現型を使いながら、ゲノム上の全遺伝子の5割程度について、遺伝子破壊の影響を解析することが少なくとも潜在的には可能であるシステムを開発した。この新しく開発された方法に基づいて、シアノバクテリアのゲノム上の全遺伝子の破壊株を対象に、遺伝子破壊株の機能的クラスタリングを行ない、各遺伝子の機能推定につなげることが本研究の最終的な目的である。

<2007年度の研究の当初計画>

2007年度当初までの研究により、以下の4点の成果が得られ、それに基づき今年度の研究計画を立案した。

1. シアノバクテリア *Synechocystis* sp. PCC 6803 のゲノムの15%をカバーする500株からなる遺伝子破壊株コレクションを作成した。

2. 上記の遺伝子破壊株コレクションについて、弱光培養条件、強光培養条件の二種の条件下でクロロフィル蛍光の経時変化のデータを集積した。

3. 上記のクロロフィル蛍光のデータを比較することにより、光化学系量比調節に関する遺伝子をスクリーニングすることができることを示した。

4. 任意の二つの変異株のクロロフィル蛍光時系列データがどれだけ似ているかを偏差二乗和に基づいて定量化することに成功した。

実際に、このクロロフィル蛍光の定量化データを用いることによって、光合成の光化学系量比の調節に関わる遺伝子の候補株を簡便に抽出することに成功している。しかしながら、光化学系量比は、光合成の調節メカニズムの一つであり、光合成に関連する遺伝子の変異を、光合成色素であるクロロフィル蛍光によって検出しただけであれば、本方法により広い範囲の表現型を解析できるという証明にはならない。特定の代謝系の複数の遺伝子が類似したクロロフィル蛍光挙動を示すことは確認しているが、今のところ、機能が未知であった遺伝子を本方法により、特定の代謝系に帰属させることはできていない。そこで、今年度はまず、光合成とは直接関係のない代謝系の遺伝子の変異について、その変異株とクロロフィル蛍光挙動が類似している変異株を抽出し、その代謝系に異常があるかどうかを調べることにより、本研究手法の有用性を証明することとした。この際、特定の代謝系の遺伝子の数が少ないと解析が困難であることから、いくつかの代謝系を選び、その構成要素の遺伝子の変異株を集中的に作成して解析することを考慮することとした。

もう一つの課題として、クロロフィル蛍光挙動の定量化の方法の確立が上げられる。クロロフィル蛍光挙動データベースを公開にするにあたり、多くの人の興味は、自分が興味を持つ遺伝子と表現型が似ている遺伝子にはどのようなものがあるのか、という点であると考えられる。そのような要請に答えられるかどうかは、どこまでクロロフィル蛍光挙動の定量化を適切に行なえるかに依存する。現在は、各時点での蛍光強度自体を定量化の基礎にしているが、各時点での蛍光強度の微分値を用いて定量化方法に変更することにより、クロロフィル蛍光挙動の「形」をより適切に表現する方法を確立することができるのではないかと考えた。

<2007年度の成果>

昨年度までの研究においては、クロロフィル蛍光挙動による遺伝子の機能解析の可能性を探るため、まず、光化学系量比調節に関わる因子の探索を行った。光合成においては、光化学系IIと光化学系Iが協調して働くため、光環境に応じて2つの光化学系は適切な量比に調節される必要がある。これまでに、この光化学系量比調節に強光下で欠損を示す遺伝子破壊株として pmgA 破壊株と sll1961 破壊株が報告されており、これらの破壊株は強光下でお互いによく似たクロロフィル蛍光挙動を示す。そこで、類似の機能に欠損のある遺伝子破壊株は類似のクロロフィル蛍光挙動を

示すと仮定し、pmgA 破壊株と sll1961 破壊株に共通してみられるクロロフィル蛍光挙動（強光培養時に、励起光の照射から約 0.5 秒後に現れる初期ピークが野生株よりも低く、45 秒後の蛍光レベルがそれらの初期ピークよりも高い）を指標に破壊株コレクションの中からクロロフィル蛍光挙動の類似した破壊株 6 種（sll1961, pmgA, ccmK2, slr1916, ctaEI, ctaCI, slr0645）を候補株として選抜した。そして実際に、これらの候補株の細胞の低温蛍光スペクトルを測定し、光化学系量比の指標となる光化学系 I(F725) と光化学系 II(F695) の蛍光強度の比を計算したところ、弱光に順化した各候補株は野生株と同様に F725/F695 が 2 程度であったが、強光に順化した野生株では光化学系 I の減少に伴い F725/F695 が 1 程度になるのに対して候補株では野生株よりも大きな値を示した。従ってこれらの候補株は、実際に強光下での光化学系量比に異常を持つことが明らかとなった。このことは、クロロフィル蛍光挙動を単純に比較する方法によって効率良く光化学系量比に異常がある変異株をスクリーニングできることを意味する。

そこで、今年度は、クロロフィル蛍光挙動の類似性による選抜を目で見て行なうのではなく、偏差二乗和に基づく定量的な類似性に基づいて行なうこととした。その結果、各時点におけるクロロフィル蛍光挙動の絶対値の偏差二乗和を類似距離とした場合、上記の 6 種の変異株が類似距離の小さな順に並べたリストの上位に来ることを確かめた。このリストの上位に来るその他の変異株について光化学系の量比を測定すると、全てではないものの、いくつかの変異株で野生型とは異なる光化学系量比を示すことが明らかとなり、偏差二乗和に基づく定量的な解析が有効であることが示された。

さらに、クロロフィル蛍光挙動の類似性の定量化をより効率的に行なうため、他の定量化手法についても検討した。各時点におけるクロロフィル蛍光挙動の絶対値を用いる代わりに、隣り合う 2 点間の差分を用いる微分類似距離は、絶対値を用いた類似距離よりもより有効な指標となることが明らかとなった。これは、対象となるクロロフィル蛍光挙動の時間変化の絶対値が、時間が後になるに連れて前の状態の影響を受けるため情報量が減るのに対して、時間変化の微分値、すなわち蛍光の時間変化の傾きを用いた場合は、時間が後になっても含まれる情報量が減らないことによるものと考えられる。

さらにこの他、1) 光化学系量比の調節に関わる変異株に特徴的なクロロフィル蛍光挙動を示す時間に重みをつける変異株特異的類似距離、2) クロロフィル蛍光挙動の時間変化の各ポイント 126 点における変異体と野生株の差を 126 次元空間にベクトルとして示した場合の任意の 2 つの変異株のベクトルのなす傾きを類似距離とした場合の類似距離、の 2 つの方法についても検討した。前者については、比較的有効な指標となることは確かめられたが、注目する表現型によって指標が異なるという欠点があり、これについてはそれ以上検討しないこととした。一方、後者の方法では、他の方法では見つけづらい表現型の強さが異なる変異株についても情報が得られることがわかった。これは、クロロフィル蛍光挙動の時間変化をベクトルとして考えた場合、ベクトルの長さが表現型の強さに対応し、ベクトルのなす傾きが表現型の種類に対応することによるためと考えられる。今後は、この方法により、遺伝子の機能クラスタリングに取り組む予定としている。

また、収集したクロロフィル蛍光のデータについては、ようやく準備が整い、Fluorome - The Cyanobacterial Chlorophyll

Fluorescence Database として公開するに至った。現在のところアクセス数は多いとは言えないが、日本とアメリカを中心に 25 カ国から利用が見られる。

現在公開している約 500 の変異株のクロロフィル蛍光共同データベースの拡充に関しては、変異株の作成とクロロフィル蛍光の測定を並行して進行中である。ランダムに変異を入れたものについては約 100 の変異株を作成し、そのうち 50 についてクロロフィル蛍光の測定が終了している。今後はこれらも随時データベースに追加、公開していく予定である。

<達成できなかったこと、予想外の困難、その理由>

代謝系の中での相互作用に的を絞る、各代謝系を構成する遺伝子の破壊株の表現型の間の相互作用を見ることによって、その代謝系のネットワークを明らかにすることもねらっていたが、こちらはまだ成功していない。ゲノムワイドな議論、および、生理学的な解析を絡めた研究が進行しているのにもかかわらず、特定の代謝系に的を絞った研究が進んでいない理由は、純粹に、調べた遺伝子の数がゲノムの 15% に過ぎず、単一の代謝系に注目してしまうと、その中に入る遺伝子の数が極めて少なくなってしまうことにより、議論が難しいためである。特定の代謝系に絞っての変異株の集中作成も進行中であるが、まだクロロフィル蛍光データを十分に蓄積できるところまでは至っていない。

<今後の課題>

上記のように、クロロフィル蛍光の類似性の定量化については一定の方向性が得られたので、今後は、最善の手法を決定し、以後はその決定に基づいて解析を進める予定である。

一方で、その基盤となるデータについては、変異体の作成が律速段階となるため、短期間に大きな進展は予想しづらい。しかしながら、今後半年間に約 200 の変異体についてのクロロフィル蛍光挙動を新規に取得できる見込みであり、そのうちの 50-100 は、いくつかの代謝系に的を絞った変異体である。1 つの代謝系について数十の変異体情報が得られるようになれば、クロロフィル蛍光挙動からの遺伝子機能解析がどこまで可能かを明らかにできると考えている。これに関しては、今年度の阻害剤を用いた予備的な実験から、呼吸系に関わる遺伝子について特徴的なクロロフィル蛍光挙動を特定できる見込みであり、これを突破口に、窒素代謝系など、その他の代謝系でも研究を進展させていきたいと考えている。

<成果公表リスト>

1) 論文

- 0704301554
Ozaki, H., Ikeuchi, M., Ogawa, T., Fukuzawa, H. and Sonoike, K., Large Scale Analysis of Chlorophyll Fluorescence Kinetics in *Synechocystis* sp. PCC 6803: Identification of the Factors Involved in the Modulation of Photosystem Stoichiometry., *Plant Cell Physiology*, 48(3), 451-458 (2007)

2) データベース

- 0708080823
Fluorome - The Cyanobacterial Chlorophyll Fluorescence Database
<http://sunlight.k.u-tokyo.ac.jp/fluorome/>