

Form I/I'型ルビスコに関連する新規タンパク質群の発見

番場 浩平 (筑波大学 生物学類) 指導教員: 稲垣 祐司 (筑波大学 計算科学研究センター)

【背景・目的】

ルビスコタンパク質は、一般に Calvin-Benson-Bassham (CBB) 回路で CO₂ を固定する酵素であり、陸上植物や藻類、シアノバクテリアなどの光合成生物に広く分布している。ルビスコタンパク質ファミリーは、陸上植物やシアノバクテリアにみられる Form I や、渦鞭毛藻と一部のプロテオバクテリアにみられる Form II、古細菌にみられる Form III、細菌と古細菌にみられカルボキシラーゼ活性に必要な残基を欠く Form IV に分類され、そのアミノ酸配列に大きな多様性が存在する。カルボキシラーゼ活性の有無から、Form I-III は true rubisco、Form IV は rubisco-like protein (RLP) とも呼ばれる。Form I ルビスコには large subunit (LSU) と small subunit (SSU) があり、LSU と SSU が 8 つずつ組み合わさってできる複合体が酵素として機能する。一方で他の Form のルビスコには SSU が存在せず、Form I の LSU に当たるドメインのみが存在し、この二量体が機能単位である。

Form IV の内部系統まで反映したルビスコタンパク質ファミリーの横断的な系統解析は、2012 年の Erb, T.らによる報告^①が最新のものである。そこから今に至るまでの十余年、シーケンス技術の向上によってゲノムデータの集積が進んでいる。本研究では最新のゲノムデータ中の細菌・古細菌のルビスコタンパク質を探索した。

【手法】

ルビスコ配列の収集

NCBI BLAST+にて、Erb, T.らの系統解析^①に用いられた 333 配列をクエリとして同源性検索を行った。対象データベースは nr, refseq_protein, env_nr の 3 つで、対象系統は細菌と古細菌である。得られた配列のうち、hsp 毎の query coverage が 50 以上かつ e-value が 1e-10 以下のものを以降の解析に使用した。

系統解析

Forms I-IV を対象とした系統解析では、前処理として CD-HIT にて収集したルビスコ配列のクラスタリングを行い、MAFFT で整列後、trimAl でギャップが 20 % 以上ある座位を削除した。最終的に 351 アミノ酸残基を含む 305 配列からなるアライメントから最尤系統樹を復元するには IQ-TREE を用いた。置換モデルには ModelFinder によって選択された LG+F+R10 を使用し、トポロジーの信頼性を 100 回の bootstrap 解析によって検証した。

Form I とその近縁配列の系統解析では BLAST で収集した配列のうち、(1) Form I、(2) 外群として Form III の一部、そして (3) 先行研究^②にて報告された Form I' 配列を使用した。前処理として (1) の配列のみを CD-HIT でクラスタリングし、MAFFT で整列、trimAl でギャップが 30 % 以上ある座位を削除した。最終的に 486 アミノ酸残基を含む 425 配列から構成されるアライメントを、IQ-TREE を使用して最尤法により解析した。置換モデルには ModelFinder によって選択された LG+R10 を使用し、トポロジーの信頼性を 100 回の bootstrap 解析によって検証した。

【結果・考察】

ルビスコタンパク質ファミリーの系統解析

True rubisco では Form I と Form II が単系統を形成し、Form III がその側系統となった。RLP では分岐順こそ先行研究^①と異なるものの、Aful や YkrW といった既知のグループが再現された。

この系統解析結果において注目すべき点は以下の 3 つである。一点目として、Form I の基部で分岐し、Form I と姉妹群を形成することが BP 値 91 % でサポートされている配列群の存在、二点目として、IV-GOS の姉妹群となる配列群の存在、三点目は IV-AMC と姉妹群を形成する配列群の存在である。本研究ではこのうち Form I の進化を考える上で重要になり得る Form I の基部から分岐する配列群についてより詳細な研究を行った。

Form I に関連する系統の調査

Form I の基部で分岐する配列群として、Form I' が報告されている^②。これは Anaerolineae 目の Chloroflexi 門細菌にみられるルビスコタンパク質で、LSU が八量体を形成する点は Form I と類似しているが、こちらは SSU を持たないと考えられている。本研究で検出した Form I 関連ルビスコタンパク質が Form I' に該当するか検討した結果、今回検出した Form I 関連ルビスコタンパク質群は Form I' よりもさらに基部で分岐する配列をふくむことがわかった。以降この新規タンパク質群を Form I'' とする。

Form I'' が Form I' と同様に SSU を必要としないかを推測するため、ゲノム中で Form I'' の LSU をコードする遺伝子の近傍遺伝子を調査した。一般的に Form I を用いる生物のゲノム中では、Form I の LSU をコードする *rbcL* 遺伝子と近接して SSU をコードする *rbcS* 遺伝子が存在する。しかし Form I'' の LSU をコードする遺伝子周辺に SSU をコードする遺伝子は見つからなかった。このことから、Form I'' には SSU が不要であると示唆された。また、Form I'' をコードするゲノム配列では、LSU 遺伝子が CBB 回路の構成要素をコードする遺伝子に囲まれているものと、Form III と同じ代謝経路の構成要素をコードする遺伝子と近接するものがあつた。これより Form I'' は Form III から Form I に進化する途上にある可能性が示唆された。

【展望】

Form I'' は、Form I・I' のクレードと Form III の中間的な系統的位置をとる。Form I'' が Form III のように二量体単位で機能するのか、Form I・I' のように八量体で機能するのかは現状不明である。今後 Form I の起源と進化を解明するためのカギとなる可能性がある Form I'' について、機能や高次構造の予測を行う。また、本研究で新たに発見された 2 つの RLP 配列群についても解析を進める。

【参考文献】

- (1) Erb, T., *et al.* (2012). *Nat Chem Biol*, **8**, 926-932.
- (2) Banda, *et al.* (2020). *Nat. Plants*, **6**, 1158-1166.