

ポリケチド合成酵素ファミリーの包括的解析

*Comprehensive analysis of polyketide synthase families*

京都大学化学研究所バイオインフォマティクスセンター化学生命科学 清水 祐吾

背景と目的

ポリケチドはバクテリアや菌類、植物等によって産生される天然化合物(二次代謝物質)であり、ポリフェノール、マクロライド、ポリエン、エンジン、ポリエーテルといった多様な化学構造を取り得ることや、自己防衛、情報伝達、毒性因子、色素生成、摂食阻害、UV 防御などの様々な生物学的機能が知られている。また、エリスロマイシン、ラパマイシン、ロバスタチンといった臨床的に重要な多くの薬品を含む、抗生物質、抗癌剤、抗寄生虫剤、免疫抑制剤、コレステロール低下剤等の薬理活性を持つ多様な化合物が含まれていることが知られているタンパク質ファミリーである。ポリケチドは生物内においてポリケチド合成酵素(PKS)によって触媒されることで合成されている。その過程は脂肪酸合成酵素(FAS)による脂肪酸合成経路と類似しており、アセチル CoA やマロニル CoA などの単純なスターター基質にマロニル CoA やその誘導体の伸長鎖基質を繰り返し縮合、修飾を行うというものである。縮合の回数、各伸長鎖における修飾の種類と有無、そして他の酵素による反応後修飾によってポリケチドの広範な多様性が生み出されている。

現在、ゲノム決定や配列解析によってポリケチドと推定されるタンパク質の遺伝子配列が大量に発見されているが、その機能は未知のものも多く、タンパク質の配列から計算機的に生成物ポリケチドの化学構造及びその性質を予測するという試みが行われている。PKS はその反応様式及び配列ドメイン構成によって大きく3種類に分類されているが、従来の分類ではとらえられないような PKS も見つかっており、生成物予測に利用できるような包括的な分類体系を整えることが必要となっている。そこで本研究では、PKS ファミリーの候補となるタンパク質を包括的に探索し、網羅的に分類を行った。さらに PKS の中でも特に未知の部分の多いタイプ III の PKS に着目してより詳細な分類を行った。

検討内容

ゲノムが解明された生物種の遺伝子・タンパク質データベースである KEGG GENES に対し、既知の PKS である pikAIII(タイプ I)、actI/ORF1,2,3(タイプ II)、CHS1(タイプ III)のそれぞれをシードとし配列類似度検索法である PSI-BLAST を用いて PKS、PKS 類似タンパクの遺伝子の抽出を行った。PSI-BLAST の繰り返し回数は 10 回まで行い、繰り返し毎に  $e\text{-value}10^{-3}$  の閾値にて配列を取得した。これらの遺伝子のタンパク質配列のドメイン情報をタンパク質のドメインデータベースである PFAM と隠れマルコフモデルを用いたドメイン検索ソフトウェアである HMMER を用いて抽出(閾値  $e\text{-value}10^{-3}$ )し、得られたデータを必須ドメイン(KS、ACP、AT)の組み合わせによって網羅的に分類をすることでその特徴付けを行った。また、PFAM における KS ドメインの定義の違いを利用することによりタイプ III の PKS 候補を抽出し、その配列をマルチプルアラインメントし、Neighbor-Joining 法を用いて系統樹を作成した。

結果と考察

各シードから得られた PSI-BLAST の結果を比較した結果、タイプ I とタイプ II のシードから得られた結果は

一致性が高くなった。また、タイプ III シードから得られた結果とそれ以外のタイプのシードから得られた結果には繰り返し回数が 4 以下では重複がほとんど見られなかったが 5 回目以降では重複が見られた。そこで、繰り返し回数 4 の PSI-BLAST 結果を採用し、これに加えてドメイン情報を使用することでタイプ I~III を区別することが可能となった。しかしながら、FAS のように PKS と近縁のファミリーのタンパク質で、類似の反応系を持ったものは PKS と同様のドメインを持っていることが知られており、この分類中にも FAS やその他類似の非 PKS 配列が大量に含まれる結果となった。系統樹を作成した結果(図 1、ここではタイプ III のもののみを示す)、この FAS を含む非 PKS 配列を PKS 配列と区別することが可能になり、タイプ III 候補配列では 939 配列中 348、すなわち全体のおよそ 4 割弱が非 PKS 配列であることがわかった。KEGG GENES に含まれる既知 PKS 配列に加え、全ゲノムが決定されていない生物における既知 PKS 配列を UniProt データベースより取得し系統樹に挿入(図 1 青色部、ここではバクテリアタイプのもののみを示す)することで既知 PKS 配列を含むクラスターを分類することが可能になった。また、近い既知 PKS 配列のないクラスターに関しては KEGG OC や KEGG ORTHOLOGY の情報を付加することによってある程度の分類を行うことが可能となった。

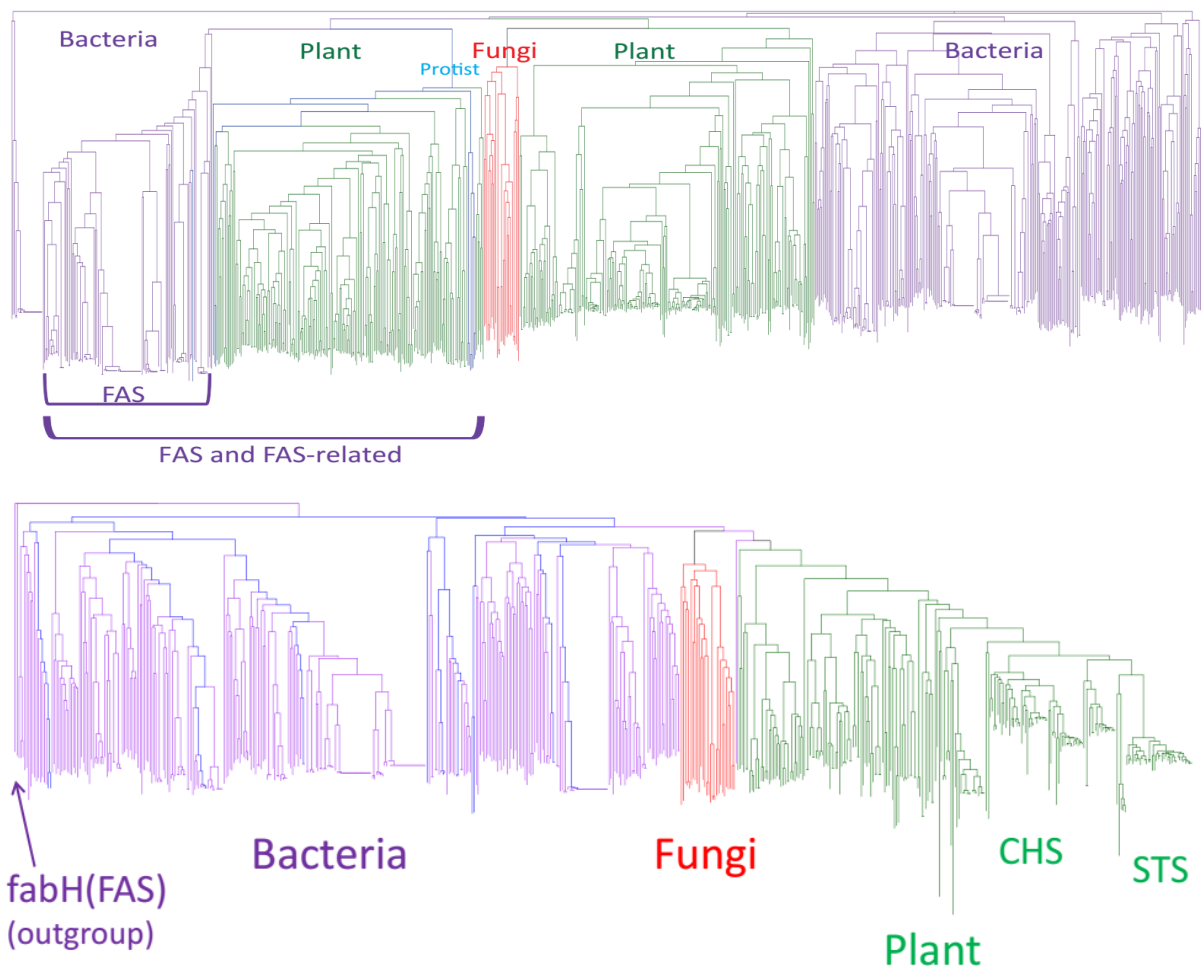


図 1. タイプ III の PKS 候補配列で作成した系統樹。上は FAS 等の非 PKS 配列を含む状態、下は非 PKS 配列を除去した状態(外群として使用した FAS を除く)。紫色、赤色、緑色、水色の部分はそれぞれバクテリア、菌類、植物、原生生物の配列。青色の部分は既知のバクテリアタイプ III PKS の配列。