

## 質疑応答

Q. ARODの祖先型設計の事例を示していただいておりますが、共通祖先の配列を探っていくことで、高い選択性と耐熱性を有する酵素にいきつくのはなぜでしょうか？

進化系統樹上で、進化するほど選択性が下がり、耐熱性が下がるといったことが逆に生じるという理解になるのでしょうか？

Ans. これは回答が極めて難しいものの一つで、今でも議論が進められているものになります。一説として共通祖先の生物は、現存する生物よりもより少ない遺伝子数で代謝を回さないといけないため、祖先の酵素・タンパク質は単体で多様な機能(例えば広選択性)を発揮する必要があったのではと考えられています(Jensen RA, Annu Rev Microbiol 1974, 30:409-425, Khersonsky, O. et al., Curr Opin Chem Biol, 2006, 10, 498-508をご参照ください)。耐熱性について、太古の昔の環境は現在に比べてかなり高温だったことが提唱されています。これは祖先型酵素・タンパク質を復元し、その時間と耐熱性を調べた研究からも示唆されています(Gaucher, E. A. et al., Nature, 2008, 451, 704-7のFig. 3が分かりやすいかと思います)。このことから、祖先型設計を実行し、高温環境時の時間軸にのるような祖先型酵素・タンパク質配列を復元できれば、耐熱性を獲得させることが可能だと考えられています。一方で耐熱性獲得がコンセンサス効果によってもたらされる、いわゆるアーティファクトではないか、という報告もあります(Trudeau, D. L et al., Mol. Evol. Biol., 2016, 33, 2633-41)。

2つめのご質問に関してですが、ターゲットとなるタンパク質によっては祖先型配列にすると耐熱性が下がることも起こります(設計に用いたタンパク質配列が適切でないなどの理由もありますが)。選択性を高くするよう進化してきた酵素・タンパク質も現存するはずです。ですので進化系統樹上での選択性・耐熱性の変化は酵素・タンパク質の種類に強く依存するのでは、と考えております。例えばすべての生物の生存に必須な酵素・タンパク質は上記仮説に従うような耐熱性・選択性の変化を生じる(祖先型配列に近づくほど耐熱性と選択性が向上する)と思いますが、その他の酵素・タンパク質は例外的な機能変化を起こす可能性もあるのでは、と考えております。

Q. ここで用いられている計算は、どのような計算機を使用し、どのくらいの計算 時間がかかるのでしょうか？

Ans. 計算機については通常の市販されているSpecのもの(10-15万円程度のデスクトップPC)を使っています。計算時間は機械的に配列分類するのであれば、1週間程度で可能です(配列ライブラリが大きくなると計算時間が必要になります)。

Q. 示されていた例では、解析対象にしたタンパク質の中に一つのモチーフがあるように見えたが、複数のモチーフが見つかることもありえると思います。モチーフごとに異なる分類になる場合はどのように処理するのでしょうか？

Ans.モチーフの候補残基には、それがどのくらいモチーフとして妥当か評価するスコアが計算されます。そのスコアが高いものから順に選抜していくことになります。一方でご指摘の問題点があることも判明しています。これを解決するため、異なるモチーフで分類したデータを最終的に統合するツールを開発中です。

Q. 各特徴を持ったサブファミリーを集めてきてモチーフを探すのではなく、ランダムな配列集団から4-8個程度のアミノ酸モチーフを持つクラスターに分類する手法で、それぞれのクラスターが何らかの特徴を持っているはずだ、という理解でよろしいでしょうか？

また、iAngler法はINTMSAlignをより短い領域で行っているというイメージでしょうか？INSTMSAlign, iAngler法共

に、公開の予定がありましたら教えていただけますでしょうか。

Ans.ご質問内容の解釈で間違いありません。iAngler法は、ランダムな配列集団からアミノ酸モチーフを基に各クラスターに分類する手法です。各クラスターは何等かの特徴(機能)を有していると推察していますが、それを合理的に予測することはできておりません。

iAngler法はINTMSAlignを細かく繰り返し行うことで、各アミノ酸残基における出現頻度のperturbationを解析・モチーフを同定することを行っています。ですのでINTMSAlignを短い領域で行うのではなく、細かく沢山行うというイメージです。INTMSAlignおよびiAngler法は現在、Web版含めて開発中です。開発を完了させ、公開につなげたいと考えております。

Q. PpLAAOがLAAOだろうと推定するのは何か指標があるのでしょうか？

まず、パラログと推定するには配列類似性から予想するのでしょうか？それともUniprotあるいはCATHなどのモチーフから機能を予想して選抜するのでしょうか？それとも先生が開発した4アミノ酸を持つパラログを選抜？

Ans. 先に設計した祖先型L-Arginine Oxidase (AncAROD)の基質選択性が広がったことから、AncARODとパラログの関係にある配列はLAAO活性を有すると予測しました。ただPpLAAOの基質選択性の広さまで配列から予測することはできておりません。なお配列類似性からPpLAAOがパラログであると予測しております (BlastpのE値が $1.0E-25$ 以下の配列を探索したところ、PpLAAOがヒットしました)。

Q. AncLAAOの配列設計

可溶性発現させるためにPpLAAO配列に基づき祖先型を設計していると思いますが、その際にHisol法なども適用しながら設計するのでしょうか？発現量、kcat/Kmを満足するものが取れる保証はあるのでしょうか？あるいは、それも狙ってAncLAAOの配列設計をするのでしょうか？(サブファミリー概念の適用を導入？)

Ans. 可溶性発現させる際にHiSol法は適用していません。発現量、酵素学的パラメータが良好なものがとれることは保証されておらず、設計してみないと分かりません。ただサブファミリー配列分類後に祖先型設計で作成したタンパク質は可溶化・高機能化するものが多い印象を受けています。今後も本手法を異なるタンパク質群に適用、解析を進めることで、設計した配列が有用な酵素機能を持つか否か判別できるスコアを作れるのではと考えております。

▶【北日本支部】2020年度オンラインシンポジウム「情報科学を駆使して生命分子を見る・知る・使う」