B770

枯草菌の転写因子結合配列と真性細菌のプロモータ領域との配列比較

○蒔田 由布子^{1,2}、中井 謙太² (¹名大・工・応用物理、²東大・医科研・HGC)

我々のグループでは以前に枯草菌 *Bacillus subtilis* における転写因子とそのプロモータ配列中の結合情報を集めたデータベース DBTBS (Ishii et al., 2001; <http://dbtbs.hgc.jp>) を作成した。今回はこれを出発点として、それらの既知転写因子結合配列 (シスエレメント) と、ゲノム決定のされている真性細菌全般のプロモータ配列との比較を行ったので報告する。具体的には、まず DBTBS のシスエレメント情報から転写因子ごとに重み行列を作成した。次に基生研の内山氏による細菌の比較ゲノムデータベース MBGD (<http://mbgd.genome.ad.jp>) により定義されたオルソログ遺伝子の上流プロモータ配列を、それらの重み行列を用いて探索し、同様のモチーフの存在を調べた。この結果を、各プロモータに対するオルソログ遺伝子の一覧表の形で表示している (新バージョンの DBTBS は近日公開を予定している)。以上のような機能を付加することにより、従来はゲノム比較をしても、トランスに働く転写因子がどう保存されているかだけが注目されてきたのに対して、さらにシスエレメントの保存状況の違いを網羅的に見ることが可能になった。実際、転写因子の保存とシスエレメントの保存パターンは必ずしも一致するものではないことを示唆するデータが得られてきている。これらの情報から細菌ゲノムにおける転写制御ネットワークの進化を明らかにすべく、現在さらに研究を進めている。

Y.Makita and K.Nakai : Sequence comparison between the transcription factor binding sites of *Bacillus subtilis* and other bacterial promoter regions.

B772

カラーコーディング法によるタンデムリピート検出の性能評価

○棟方 光¹、Abulimiti Aimaiti¹、種田 晃人¹、大澤 研二²、清水 俊夫¹ (¹弘前大・理工・電子情報システム工学、²群馬大・理工・ナノ材料システム工学)

現在、タンデムリピート (TR) の検出を行うプログラムとして、様々な手法が提案されている。しかしながら、自動化手法においては、一致度の低い TR や長い周期の TR を検出できない、あるいは検出数が少ないという問題がある。また、可視化手法の多くは、TR の検出に特化されたものではない。可視化手法の一つであるカラーコーディング (CC) 法が、一致度の低い長周期の TR を検出するのに有用な手法であることが示されている (1) が、その性能についてはやや不明な点がある。そこで本研究では、3 生物種 (*H. influenzae*, *D. radiodurans*, *Y. pestis*) を CC 法による TR 検出の対象とし、その特徴を把握する為に、以下の三つの点に注目した。第一に、CC 法で TR の存在しそうな領域を抽出した時に、実際にその領域に TR が存在しているかどうか (specificity)。第二に、異なる検索者により、同じゲノムを検索した際の結果の相違について。第三に、他手法である Tandem Repeats Finder (TRF) との比較である。その結果、CC 法の specificity は約 90 % であることが分かった。また、異なる検索者による同一ゲノムでの検索では、ゲノムの組成が偏っていると、個人差が大きくなることが分かった。これは組成の偏りにより TR の数が増加し、視認による判断を惑わす為と考えられる。CC 法と TRF の比較では、TRF で検出できる TR のほぼ全てを CC 法で検出可能であること、さらには CC 法がより低一致度、長周期の TR を検出することが可能であることが分かった。

(1) Yoshida et al., *J. Mol. Biol.*, **298**, 343-349, 2000

H. Munakata, A. Abulimiti, A. Taneda, K. Oosawa and T. Shimizu : Evaluation of the color-coding method for finding tandem repeats in prokaryotic genomes

B771

バクテリアにおけるタンデムリピートの検出とその解析

○Abulimiti Aimaiti¹、棟方 光¹、種田 晃人¹、大澤 研二²、清水 俊夫¹ (¹弘前大・理工・電子情報システム工学、²群馬大・理工・ナノ材料システム工学)

タンデムリピート (TR) の検出を行う手法の一つであるカラーコーディング (CC) 法は、他の手法では検出が困難な低一致度、長周期の TR の検出に適している (1)。本研究では、この手法を用いて、44 種のバクテリアゲノム中の TR の検出を行った。その結果、各ゲノムにおいて検出総数にかなりの差が生じていることが明らかとなった。これはゲノムサイズの差異が第一要因と考えられるが、その他にも Achaz らが指摘した組成の偏りや重複のメカニズムとの関係もあり得る (2)。そこである区間毎にゲノムを分割し、それらの区間の組成の偏りの指標として情報エントロピーと TR の数にどのような相関が見られるのかを解析した。その結果、幾つかの生物種において負の相関が見られた。また、遺伝子重複による新たな機能をもつタンパク質が生成される要因の一つとして、TR が考えられる。本研究では、これに関連する TR の同定を目的として、TR とそれが存在する領域とその周辺の遺伝子の配置をそれぞれの位置関係から 7 つに分類した。現在、これらの組み合わせから考えられる TR のゲノム構造に及ぼす影響や、TR と関連する重複遺伝子がどのような機能をもつのか、などの解析を進めているところである。

(1) Yoshida et al., *J. Mol. Biol.*, **298**, 343-349, 2000

(2) Achaz et al., *Nucleic Acids Res.*, **30**, 2987-2994, 2002

A. Abulimiti, H. Munakata, A. Taneda, K. Oosawa and T. Shimizu : Analysis of tandem repeats detected in prokaryotic genomes

B773

化学反応ネットワークとして見た生物の性質とその進化

○松浦 健太郎、金子 邦彦 (東大総文)

生物も所詮化学反応ネットワークである。しかしランダムにネットワークを組んでもうまくいかないことは未だかつて試験管の中で生物を作れたことがない事実からも明らかである。そこで生物らしさをもつネットワークの性質、条件とは何か? という動機から本研究は始まった。具体的には計算機上にいくつかの仮想細胞を作る。それらを特徴づける量は化学物質の成分量とする。その量は質量作用の法則をベースにおいた微分方程式に従って時間発展する。反応式の詳細によらない性質を議論したいのでそこにあまりこだわらない。そういう視点から生物を見ると力学系の考え方を適用できる。すなわち最終的な成分量の落ち着き先をアトラクターと呼び、そこは成分量の揺らぎに対して安定である。従ってアトラクターは仮想細胞の一つの状態と捉えることができるであろう。またアトラクターはいくつか存在し、初期値やパラメータ (ここでは餌の量とする) によって落ち着き先は変わる。そこで本研究はまずパラメータの操作に対して、そういったアトラクター間の移り変わる規則について調べた。すると状態 A から状態 B へ移る時にさまざまなレベルの不可逆性が見えることが分かった。第二にパラメータの種類と量の変化の仕方 (例えばランダムでもよい) を決め、生き残った仮想細胞に対して変異 (=新しい物質を合成する反応を付け加えたり、既存の反応が起こらないようにしたり) を起こさせた時に、どのようなネットワーク構造を持つ仮想細胞が進化で生じるかを調べた。これはアトラクター間の遷移規則とも関連する。以上のことについて議論したい。

K.Matsuura, K.Kaneko : The nature of living beings and its revolution from the viewpoint of a chemical network