

機関番号：82706

研究種目：若手研究 (B)

研究期間：2009～2010

課題番号：21710209

研究課題名 (和文) ゲノム多型検出法の開発：可変部位・多型部位の検出からゲノム構築論理の理解へ

研究課題名 (英文) Detection of genome rearrangement through genome comparison

研究代表者

河合 幹彦 (KAWAI MIKIHICO)

独立行政法人海洋研究開発機構・海洋・極限環境生物圏領域・ポストドクトラル研究員

研究者番号：80451904

研究成果の概要 (和文)：

相同性検索結果から挿入欠失・置換といった多型を選び出す、データベース検索対応型のゲノム多型検索手法を作成した。相同領域をブロックとして検出し、そのブロック同士の位置関係の組み合わせで、多型のパターンを判定した。多型領域に機能既知遺伝子と対をなしている機能未知遺伝子を探索した。この検索手法により、完全ゲノム配列以外の配列情報に加えメタゲノム配列データも対象とした多型検索や、特定の近縁ゲノムの組み合わせ(多対多)にとどまらない多対配列データベースレベルの比較ができるようになった。

研究成果の概要 (英文)：

Homology-based methods widely used to infer a function of a gene cannot be applied to genes lacking any detectable homology with functionally or structurally characterized genes. In the present study with the aim of identifying such genes, we developed a method that does not stand on homology of a gene but instead stands on homology of its surrounding nucleotide sequences.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2009年度	2,600,000	780,000	3,380,000
2010年度	1,100,000	330,000	1,430,000
年度			
年度			
年度			
総計	3,700,000	1,110,000	4,810,000

研究分野：ゲノム科学

科研費の分科・細目：ゲノム科学・ゲノム情報科学

キーワード：機能予測・ゲノム多型・ゲノム進化・ゲノム比較・系統プロファイル

1. 研究開始当初の背景

ゲノム科学の進展は、ゲノムが遺伝子の種類、数や配置が大きく変化していく流動的なものであることを明らかにした。

近年では同じ属・同じ種の複数のゲノムが決定されることが普通になってきている。こういった、近縁のゲノム配列を比較することで、そのゲノムに特異的な領域を同定し、逆位・転移といったゲノム構造の経てきた変化

を検出し、時にはそれらがどのような分子機構で生じたのかを推測できる。近年、ゲノム配列決定が格段に加速・低コスト化し、多数のゲノム配列が明らかになる時代になってきた。それらを活用するためのゲノム比較のインフラマティクスが要請されている。

遺伝子の機能推定は、通常共通祖先由来の相同遺伝子 (オーソログ) をもとに行う。すなわち機能未知遺伝子と機能既知遺伝子と

の配列類似性をもとにした検索である。しかし、遺伝子配列の多様性が明らかになるにつれ、機能的に同等な遺伝子間に配列の類似性がない場合があることがわかってきた。例えば、制限酵素遺伝子では、全く配列相同性のない異なった立体構造骨格を持つ遺伝子が知られている。

機能的に関係する遺伝子は、ゲノム再編過程でゲノム上で近接した状態で維持されることが多い。すなわち相同領域に挟まれた非相同領域上の遺伝子群は機能的に関連している場合が多く、これをもとに機能推定できる可能性を示唆する。

ゲノム進化機構を探るには注目する事象が最近起きていることが望ましい。すなわち近縁ゲノム同士を比較することが有用である。ゲノム同士の対応する遺伝子の分布パターンから機能的に関係する遺伝子を抽出する系統プロファイル法は、遺伝子の有無に注目する。そのため、近縁ゲノム比較に適用しても得られる知見が限られていた。しかし、部分的に欠失した遺伝子など、軽微な違いを検出して系統プロファイル法で扱えるように数値化すれば、系統プロファイル法をより近縁なゲノムの比較にも適用できるようになると考えられる。

2. 研究の目的

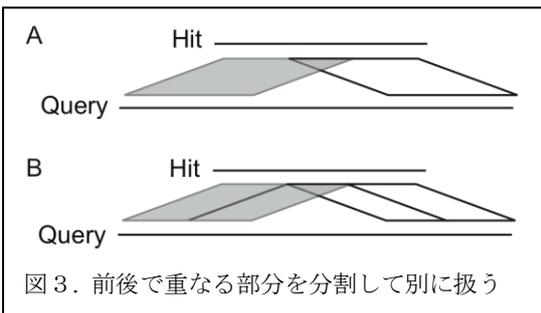
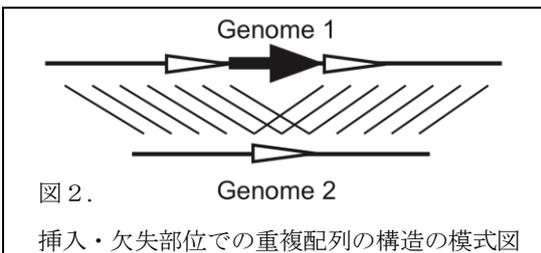
(1)ゲノム配列比較によって中・大規模ゲノム多型を検出する。ペアワイズのゲノム比較でなく、ゲノムのある領域の塩基配列を問い合わせ配列に、配列データベースに対して相同性検索を行い、挿入欠失・置換といったゲノム多型をスクリーニングする手法を開発する。本手法を遺伝子自体の配列の相同性からは検出できない新規遺伝子探索に適用する。

(2)相同遺伝子の数や偽遺伝子などの特徴を数値化した遺伝子属性を用いた系統プロファイル法を開発する。その手法によって、近縁ゲノム間の比較から系統特異的な進化パターンを検出する。

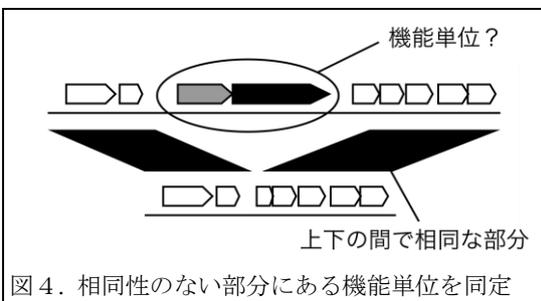
3. 研究の方法

(1)対象配列の塩基配列とその両隣の隣接配列を問い合わせ配列にして配列データベースを検索する。検索結果をもとに、塩基配列レベルのオーソログ領域の検出を行う。オーソログ判定には、ゲノム比較ソフトCGATのデータ構築プログラムを改良したものを用いた。元々のプログラムでは領域Aと比較対象の領域Bがオーソログ領域かどうかを判定する。このときオーソログ領域をブロックとして返せるように改良した(図1)。近縁ゲノム間でも挿入・欠失・置換などのゲノムの多型はごく普通に起きている(Kawai et al. 2006, Tsuru, Kawai et al. 2006)。挿入・

欠失は数塩基から数十塩基の順方向重複配列の間で起きることが多い(図2)。こういった場合、今までのプログラムでは前半と後半を別のブロックとして検出する(図3A)が、全体を一つのオーソログなブロックとして検出するのが適当である。そこで、前後それぞれのアラインメントを前後で重なる部分と重ならない部分に分割するという前処理を組み込むことで、一つのオーソログ領域として検出できるように改良した(図3B)。こうして抽出したブロック同士の位置関係の組み合わせで、多型のパターンを判定して拾うべき多型を絞り込む。



本手法を新規遺伝子探索に適用した。2つ以上の遺伝子からなる機能単位であって、1つが配列類似性から検出できる場合、それを足場にして隣接する配列から機能が推定できない遺伝子の機能を予測できる。



具体例として、制限酵素修飾酵素遺伝子を用いた。制限修飾系は、修飾酵素でメチル化されたDNAを自己とみなして保護する一方、メチル化されていないDNAを非自己と見なし制限酵素で切断する。そのため、ほとんど

の場合、2つの遺伝子はゲノム上で隣り合っ
て存在している。修飾酵素遺伝子は配列類似
性からの検出が容易であるのに対し、制限酵
素遺伝子は配列の違いが大きく、配列類似性
からは検出できないものが多く知られてい
る。

まず、修飾酵素遺伝子を足場に、既知制限
酵素遺伝子がどの程度検出できるか検討し
た。次に、隣接する制限酵素遺伝子が知られ
ていない修飾酵素遺伝子に本手法を適用し
て、機能未知遺伝子から制限酵素遺伝子酵素
を探索した。

(2) 数値属性を用いた系統プロファイル法。
与えられた複数のゲノムの遺伝子セットを、
オーソログに分類する。オーソログ解析では、
複数の生物種を比較し、共通祖先由来の相同
遺伝子（オーソログ）遺伝子を1つの遺伝子
グループ（オーソログクラスタ）とする。所属
研究室で開発したユーザ自前のゲノムを加
えたオーソログ解析が行える RECOG
(Research Environment for Comparative
Genomics)を使用した。次に、それぞれのオー
ソログクラスタごとに、相同遺伝子の数の
違いや遺伝子が分断しているかどうかとい
う遺伝子の属性を数値化する（図5）。得ら
れる数値ベクトルを、注目する生物種を分け
るテンプレートパターンと比較して、パター
ンが関連するものを抽出する。

Organism	1	2	3	4	5	6	7	8
Presence/absence	+	+	+	+	-	-	-	-
Paralog number	3	3	3	3	1	1	1	1
Split	+	+	+	+	-	-	-	-
Partial deletion	del	del	del	del	+	+	+	+
Gap number	gap	gap	gap	gap	-	-	-	-

図5. 遺伝子属性を用いた系統プロファイル

4. 研究成果

(1) 通常修飾酵素遺伝子が制限酵素遺伝子と
対をなすことから、この多型検索プログラム
を使い、多型領域に修飾酵素遺伝子と対をな
している機能未知遺伝子から制限酵素遺伝
子を探索した。修飾酵素遺伝子を足場に、既
知制限酵素遺伝子がどの程度検出できるか
検討したところ、現在知られる5種類の制限
酵素立体構造骨格をすべて検出することが
できた。

次に、隣接する制限酵素遺伝子が知られ
ていない修飾酵素遺伝子に本手法を適用し
て、機能未知遺伝子から制限酵素遺伝子酵
素を探索した。5千の既知修飾酵素遺伝子のうち、
制限酵素と対になっていない3千から、非相
同領域に修飾酵素遺伝子と機能未知遺伝子
が隣接する例を300程度見いだした。相同
遺伝子で同様の遺伝子構成になっているか

どうかを調べることで最終的に数十程度に
絞り込めた。立体構造予測を行い、いくつか
の遺伝子は、通常の相同性検索では検出で
きない程度に異なっているものの、既知制限
酵素と立体構造レベルでは同じ骨格を持つだ
ろうことを明らかにした。

共同研究先の東京大学新領域創成科学研
究科小林一三研究室において、本研究手法を
適用して得られた候補の制限酵素活性の有
無が検討された。その結果、実際に制限酵
素活性を持つ新規遺伝子が実験的に同定さ
れた(Khan, Furuta, Kawai et al., *Nucleic
Acids Research*, 38(9), 3019-30, 2010)。
この制限酵素ファミリーは既知の遺伝子と
有意な配列類似性を持たないことから、新し
い立体構造骨格を持つ可能性が示唆された。
DNA切断酵素の構造機能関連の多様性の理解
に貢献すると考えられる。DNA認識配列の切
断パターンも、遺伝子配列が既知の制限酵
素からは見いだされていない新しいパター
ンであり、新しい認識配列の制限酵素を探索
するには、本手法のように、配列相同制では
検出できない遺伝子に注目することが有効で
あると考えられる。

本手法を用いることで、数十塩基から数十
キロ塩基単位での塩基配列の挿入欠失や置
換が検出できる。この手法を適用することで、
多細胞進化のモデル生物である多細胞植物
プランクトンのオオヒゲマワリ (*Volvox*) の
初期発生時に特異的に発現するトランスポ
ゾンの挿入場所を検出した。

本検索手法により、完全ゲノム配列以外の
配列情報に加えメタゲノム配列データも対
象とした多型検索や、特定の近縁ゲノムの組
み合わせ(多対多)にとどまらない多対配列
データベースレベルの比較が可能になった。
新規遺伝子を探索する上で遺伝子自体が既
知遺伝子と相同性を示さない場合に有用で
あることを示した。

(2) ある遺伝子が他のゲノムの相同遺伝子で
は幾つに分断されているかを数値化した数
値属性を使うことで、偽遺伝子候補を検出で
きた。同属別種のゲノムがそれぞれ複数決ま
っている分類群を用いて検討したところ、遺
伝子の有無だけでは検出できない違いを検
出できた。二成分制御系やトランスポーター
の遺伝子群が検出された。近縁種間で環境適
応戦略の違いを反映した遺伝子レパートリ
ーの多様化が現在進行形で進んでいる可能
性が考えられる。

東京大学新領域創成科学研究科小林一三
研究室との共同研究において、本研究で開発
した数値属性を用いた系統プロファイル法を
ピロリ菌のサブタイプのゲノム比較に適
用した。分断遺伝子に注目した系統プロフ
ァイルの結果、東アジア型のピロリ菌では酸化
還元に関わる特定の経路(モリブデン補酵素

を利用する経路)が崩壊していることを明らかにした(Kawai et al., BMC Microbiology, 11(1), 104 (2011))。

パラログの個数に注目した系統プロファイルの結果、東アジア型ピロリ菌で系統群特異的に増加している外膜タンパク質遺伝子、減少している外膜タンパク質遺伝子を見いだした。いずれも、ゲノムの逆位に伴う遺伝子重複という共通のゲノム進化機構によって説明できることを示した(Furuta*, Kawai* et al. (*: equal contribution), Proceedings of the National Academy of Sciences, U. S. A., 108, 1501-1506 (2011))。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[学会発表] (計3件)

- ① Mikihiko Kawai, Phylogenetic Profiling Based on Gene Property: Identification of Decaying Pathways in a Subset of Closely Related Organisms, APBC2011 The 9th Asia Pacific Bioinformatics Conference, 2011/01/12-14, Songdo Convensia, Incheon, Korea
- ② Mikihiko Kawai, Phylogenetic Profiling Based on Gene Properties: Identifying Gene Families with Characteristic Distribution among Closely Related Organisms, Asian Young Researchers Conference on Computational and Omics Biology AYRCOB 2010, 2010/3/10-12, National Cheng Kung University, Taiwan
- ③ 河合 幹彦, ゲノム比較インフォマティクスによるゲノム進化の方向性の解明、ゲノム微生物学会若手の会、2009/11/19、東京・八王子セミナーハウス

6. 研究組織

(1) 研究代表者

河合 幹彦 (KAWAI MIKIHICO)

独立行政法人海洋研究開発機構・海洋・極限環境生物圏領域・ポストドクトラル研究員

研究者番号：80451904