

機関番号：34315

研究種目：若手研究（B）

研究期間：2009 ～ 2010

課題番号：21710212

研究課題名（和文） 環境ゲノムの機能多様性の解析

研究課題名（英文） Analysis for functional diversity of environmental genomes

研究代表者

奥田 修二郎 （OKUDA SHUJIRO）

立命館大学・生命科学部・助教

研究者番号：00512310

研究成果の概要（和文）：

本課題の目的としては、環境微生物のゲノム配列を利用して、環境の違いと微生物の機能との関係を明らかにすることである。さまざまな環境から得られる微生物の遺伝子配列を利用して比較解析を行った結果、代謝系のカテゴリにおける差は、大きくないことがわかったが、環境微生物では、より小さなサブネットワークが連結しやすいという傾向を発見した。このことにより、環境中の微生物は、代謝物質のやりとりにより相互作用している可能性が示唆された。

研究成果の概要（英文）：

This project aims to clarify the relationship between environments and microbes using environmental genomic data from metagenomics analyses. I compared various environmental metagenomics data and found that the metabolic categories of them have no large differences but sub-networks of metabolism were linked each other in environmental microbes. Therefore, it was implied that microbes have interactions through compounds in their environments.

交付決定額

（金額単位：円）

	直接経費	間接経費	合計
2009 年度	2,200,000	660,000	2,860,000
2010 年度	1,300,000	390,000	1,690,000
年度			
年度			
年度			
総計	3,500,000	1,050,000	4,550,000

研究分野：バイオインフォマティクス

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：メタゲノム、環境

1. 研究開始当初の背景

自然環境中では、生物は多くの他の生物と共に生きていることから、環境との相互作用を理解するためには、生物を生物群としてとらえる必要がある。最近になって、ある環境サンプル中の生物が持つゲノムの配列をすべてホールゲノムショットガン法によって決

めてしまうという方法が開発された。「メタゲノム解析」と呼ばれるその手法を利用すると、これまでの手法では同定困難であった生物種ゲノムの配列も、シーケンスすることが可能となり、環境ゲノム解析の新たな局面が始まったと言っても過言ではない。このようなメタゲノム解析は、まだ始まったばかりで、

利用できるサンプルは少ないものの、環境ゲノミクスを行う上で非常に有効なデータである。

2. 研究の目的

本課題の目的は、環境の違いが、環境ゲノムの機能とどのような関係にあるのかを明らかにすることである。そのためには、

(1) メタゲノム配列内の遺伝子コンテンツの機能を同定すること。

(2) パスウェイなどの高次機能ネットワーク上での構造を同定する。

(3) それらから環境ゲノムとしての機能を推定することが必要である。

(4) オペロンのようなゲノム構造の特徴と関連付けることが出来れば、ゲノム進化についての考察も可能となる。

遺伝子の機能から始まり、ゲノム構造、ネットワーク構造という高次機能を推定することにより、環境ゲノムとしての機能の理解を目指す。

3. 研究の方法

環境によって異なる微生物群の多様な機能の推定を次の行程で行った。

(1) 機能アノテーションの推定

メタゲノム配列をクエリーにし、KAAS 自動アノテーションシステムで遺伝子の機能推定を行った。KAAS は、KEGG データベースによって提供されている遺伝子オーソロジー (KO) に基づいたアノテーションを BLAST のベストヒット情報に基づいて行うものである。

(2) 代謝系ネットワークの推定

KEGG によって提供されるパスウェイ情報に、(1) によって得られた各サンプルの KO による機能アノテーションをクエリーにしてマッピングを行い、代謝系の再構築を行った。また、代謝系のグローバルネットワークを提供する KEGG Atlas システムを利用して、代謝系全体のネットワークの再構築も行った。このシステムの最大の特徴として、代謝系全体を俯瞰することができる点がある。得られた代謝ネットワークの違いをビジュアル化し、目で見て比較することが可能である。

(3) モジュール解析

KEGG では、パスウェイ情報をさらに細分化した MODULE というデータベースも提供されている。MODULE では、代謝パスウェイ中の連続

した代謝反応を一括りにしたモジュールという単位を定義している。(1) によって得られた KO のデータをこのモジュールにマッピングした。これにより、環境毎に微生物群が持つモジュールの有無を理解することが可能になる。

(4) ネットワークの連結性

モジュールは連続した単位であることから、その両端の化合物を基準に、モジュール同士の連結性について解析することが可能である。(3) から得られた各環境中でのモジュールの有無の情報を元に、モジュールの連結の程度を推測した。

4. 研究成果

(1) 代謝ネットワークの再構築

各環境メタゲノムデータにアノテーションされた KO は、そのメタゲノムデータのサイズに比例して、アサインされる数が増えて行く傾向を示した(図1)。また、各サンプルの KO を KEGG パスウェイマップにアサインし、各マップのカテゴリ毎に、その頻度を計算した。図1右はその一つを例として挙げたものである。代謝系のカテゴリ毎に、頻度の違いがあることがわかった。しかし、サンプル間で比較を行ったところ、その傾向に違いは少なく、どのサンプルでも同様の傾向を示した。

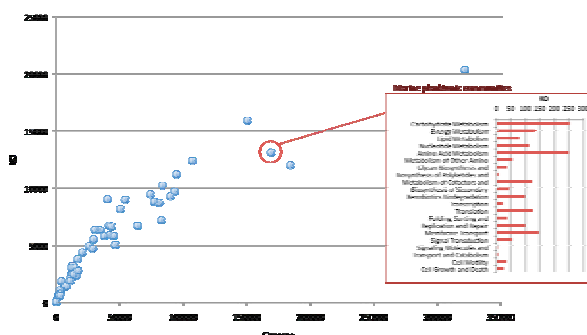


図1 ゲノムサイズと KO の数および代謝系の再構築

(2) モジュール解析

各サンプルで得られた KO から、モジュールの有無を計算した。各モジュールは複数の一連の化学反応を含んでいるが、この解析では、それらの全てのステップが存在するときに、そのモジュールを「有」とみなした。また、メタゲノムデータの対照として、ランダムなモデルをいくつか利用した。

ランダムに KO をサンプリングしたモデル

ゲノム1つをサンプリングしたモデル

ランダムに複数のゲノムをサンプリングしたモデル

これらのランダムなモデルと、環境中のサンプルであるメタゲノムで、モジュールの再構築を行い、比較した。図2は、各サンプルでのKOの数と、再構築されたモジュールの数をプロットしたものである。

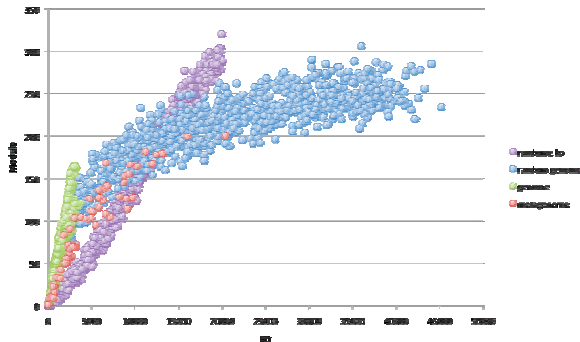


図2 モジュール再構築

KOをランダムにサンプリングすると、それに比例して再構築されるモジュールも増えていることがわかった。また、単一ゲノムだけでモジュールを再構築すると、得られるモジュールは、KOの数に比例した。また、ランダムなKOの場合に比べ、非常に効率よくモジュールを再構築していることがわかった。この点で、生命を構成するゲノムは、機能的な単位とも考えられるモジュールのようなものを非常に効率よく保持していると言える。さらに、その単一ゲノムを組み合わせたモデルは、単一ゲノムの時よりは、モジュールの再構築の程度は低い値を示した。このゲノムの組み合わせモデルは、仮想的にはメタゲノムを模しているものと想定できるが、実際のメタゲノムでの結果とは異なっていることがわかった。実際のメタゲノムデータでは、組み合わせモデルでの下限の辺りにプロットされており、モジュールの再構成においては、効率が悪い遺伝子組成であることが判明した。

(3) ネットワークの連結性

図2から、環境のゲノムでは、モジュールの再構成の効率が悪い可能性が示唆されたが、ここでは、モジュールの連結性を評価した。図3は、各サンプルで再構成されたモジュールの数に対するモジュール間の連結性を示

している。ランダムなモデルは、図2で用いたものと同じものを利用した。その結果、メタゲノムのサンプルは、モジュールの連結性が非常に高い値を示した。また、メタゲノムは、他のランダムなモデルが示す値の中で、最も高い値と似た値を示すことがわかった。これは、メタゲノム中では、隣接するモジュールの出現率が高いことを示している。実際の環境中では、様々な微生物種が混在しているが、それらがモジュールと言う単位で、物質の代謝を行っていることと、その代謝によって得られる物質を介して、他の微生物と相互作用している可能性が示唆された

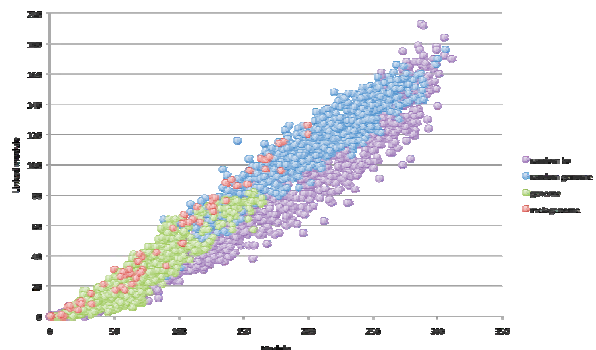


図3 モジュールの連結性

(4) データカバレッジ

メタゲノムデータは、完全にすべてのゲノム配列を決定しているわけではないことから、解析におけるカバレッジを決める閾値設定が有効となる。今回のモジュール解析では、すべての要素を持っているモジュールのみを解析対象にしたが、この解析で閾値を設定した場合、各サンプルがより多くのモジュールを持つと予測されることになる。このような閾値を設定した解析を行うための手法の開発を行った。その結果、KEGGにおけるPATHWAY、MODULE、BRITEなどの各データベースに対して、自動的に必要な情報を抽出することが可能なシステムを開発することができた。インターネット用のブラウザからそれらの情報を得ることが出来るシステムも同時に開発し、現在 in house で利用することができている。

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)

[雑誌論文](計1件)

著者名: Shujiro Okuda and Akiyasu C. Yoshizawa、論文表題: ODB: a database for operon organizations, 2011 update., 査読: 有、巻: 39、発行年: 2011、ページ: D552-D555

[学会発表](計1件)

発表者: 奥田修二郎、発表表題: 環境ゲノムのパスウェイ解析、学会名等: 日本微生物生態学会、発表年月日: 2010年11月21日、発表場所: 広島大学(広島県)

[図書](計0件)

[産業財産権]

出願状況(計0件)

名称:
発明者:
権利者:
種類:
番号:
出願年月日:
国内外の別:

取得状況(計0件)

名称:
発明者:
権利者:
種類:
番号:
取得年月日:
国内外の別:

[その他]

ホームページ等

6. 研究組織

(1)研究代表者

奥田 修二郎 (OKUDA SHUJIRO)

立命館大学・生命科学部・助教

研究者番号: 00512310

(2)研究分担者

()

研究者番号:

(3)連携研究者
()

研究者番号: