

科学研究費助成事業 研究成果報告書

平成 28 年 6 月 9 日現在

機関番号：82706

研究種目：挑戦的萌芽研究

研究期間：2014～2015

課題番号：26550053

研究課題名(和文)メタゲノミクスと数理統計学の融合による新規な環境影響評価法の開発

研究課題名(英文)Development of new method for environmental impact assessment by combination of genomics and mathematical statistics

研究代表者

高見 英人(TAKAMI, Hideto)

国立研究開発法人海洋研究開発機構・海底資源研究開発センター・上席研究員

研究者番号：70359165

交付決定額(研究期間全体):(直接経費) 3,000,000円

研究成果の概要(和文):本研究は、あらゆる環境に生息する微生物のメタゲノム解析から得られる膨大なデータをもとに環境微生物が有する様々な生理代謝機能を可視化し、数理・統計物理学的な手法やネットワーク解析を駆使して環境への影響を鋭敏に評価する方法論の開発を目的とした。具体的には既知の海洋メタゲノム配列を研究代表者らが開発中の生理代謝機能評価システム(MAPLE)を用いて環境微生物が有する様々な生理代謝機能を機能の充足率として数値化し、その数値の過大/過少評価を補正するための指標(Q値)の計算手法を考案した。また、機能の存在度の計算手法も開発し、これらの数値をもとに環境間の生理代謝機能や生物種組成の違い等を明らかにした。

研究成果の概要(英文):Objective of this study is to develop a new method for environmental impact assessment based on the metagenomic data of microbial communities obtained from the targeted environments by combination of genomics and mathematical statistics. First of all, we quantified the metabolic and physiological potential based on completion ratio of the KEGG functional modules. Module completion ratio is easy-to-understand measure to evaluate functional potential but when the genes used for a module are shared with the other modules, the MCR does not necessarily reflect the working probability. Thus, we proposed the use of the Q-value for determining the significance of module completion. We also developed calculation method for abundance of each complete module to evaluate functional robustness. Since we could clarify the difference in functional potential between different sites using these methods, it is expected that our method is sufficiently applicable to actual environmental impact assessment.

研究分野：ゲノム科学

キーワード：環境メタゲノム 生理代謝機能 環境影響評価 機能モジュール MAPLEシステム

1. 研究開始当初の背景

地球の70%を占める海洋には様々な生物が生息するが、時に原油流出事故や地震・津波などの環境攪乱によって大きな影響を受ける。このような一過的イベントによる生物への影響は、大型生物を指標にした死滅や逃避などによる個体数減少の調査によりその現状把握が可能である。しかし、影響を受けた環境がどのようなプロセスを経て復元するのか、それにはどのような物理・化学的要因のみならず生物学的な作用が必要なのかを知る術はない。一方、海洋には海域や深度を問わず様々な微生物群集が生息しているが、その構成と機能は環境の物理化学的要因で決まる。微生物群集は、わずかな環境条件の違いに反応して増殖や生理状態を変えるが、従来の分離・培養法ではこれらの情報を網羅的に読み取ることができない。これを一変させたのが、環境中の全ての遺伝子(メタゲノム)情報を網羅的に解読し、環境が持つ生物学的機能の解読を可能とするメタゲノミクスである。つまり、普遍的に存在する微生物群集のメタゲノム配列を機能情報へと変換し、様々な機能の”充足度”、”冗長度”、機能を担う生物種の”数”、”割合”などを数値化し、定量的な解析を可能とする方法論が確立されれば、同じく数値化された物理化学的要因とを合わせたこれまでにない高感度の環境影響評価法と新たな環境シミュレーション技法の開発へと発展させることが可能となるのである。この着想に基づき本研究の提案に至った。

2. 研究の目的

原油流出事故、地震・津波による自然災害、大気中のCO₂濃度の上昇に伴う海洋酸性化などによって、海洋に生息する生物は様々な影響を受けている。これらの環境攪乱による生物への影響評価は、一般に大型生物を指標にした死滅や逃避などによる個体数減少を基準にしているが、その結果は環境破壊の最終段階で現れるため、早期の緩和策導入を決める基準にはならない。そこで、大型生物ではなく、環境中に普遍的に存在し、環境変化に敏感で増殖も速い微生物が持つ様々な環境適応能や代謝能などを基準とした全く新しい環境影響評価法を開発する。これは、これまで交わることがなかった異なる2分野からのアプローチ(最新の環境メタゲノミクスと数理統計学)を組み合わせた新しい方法論を構築することで実現可能となる。本研究の成果を、環境維持力、復元力を予測する新規な環境シミュレーション技法へと発展的に繋げることを目指す。

3. 研究の方法

本研究は大型生物の死滅や個体減少を基準とした海洋における環境影響評価法を一変し、普遍的に存在する微生物群集のメタゲノム配列情報に基づく環境影響評価を可能とするため、まず、配列情報から機能情報に翻訳し、その機能ポテンシャルを数値化したデータセットを作成する。それをもとに、「生物-環境」相互作用の抽出・定量化法と実環境の生態系をデザインする逆

生態学的手法を考案する。さらにこの二つの方法論を展開することで、環境パラメータと代謝ポテンシャルの関係を統計モデルで定式化し、代謝ネットワークを基盤とした生態系の構築原理を見いだす。更に、代謝ネットワーク理論を応用し、群集組成の変動に対する代謝ポテンシャルのロバスト(頑強)性を定量化する手法を構築する。この技法を種々の環境におけるメタゲノムデータに適用し、その有効性を評価する。

4. 研究成果

本研究課題“メタゲノミクスと数理統計学の融合による新規な環境影響評価法の開発”では、大型生物の死滅や個体減少を基準とした海洋における環境影響評価法を一変し、普遍的に存在する微生物群集のメタゲノム配列情報に基づく環境影響評価を可能とすることを目的としているが、まず、それに必要なメタゲノム配列情報に由来する生物の機能を数値化したデータセットの作成を行った。具体的には、Venter 研究所が中心となって行ったグローバル海洋サンプリングプロジェクトで採取された、サルガッソー海及びガラパゴス諸島近海、4 サイト(GS000c, GS000d, GS030, GS031)のメタゲノム配列データ(約120~140万配列)から研究代表者らが開発中の生理・代謝機能ポテンシャル評価システム、MAPLE を用いて、エネルギー代謝、中央代謝や環境応答、物質輸送などに関わる機能のポテンシャルを、各反応単位でモジュール化された機能の充足率(MCR %)という形で数値化した(表1)。

表1. MAPLEシステムによるMCRの計算結果

Module ID	MCR % GS000c	MCR % GS000d	MCR % GS030	MCR % GS031
M00165	100	100	90.9	100
M00166	100	100	100	100
M00167	100	100	100	100
M00168	100	100	100	100
M00169	100	100	100	100
M00172	100	75	75	75
M00171	85.7	28.6	28.6	28.6
M00170	50	50	50	50
M00173	90.9	90.9	81.8	81.8

また、数値化された MCR の統計的な偏りによる過大/過小評価を補正するための指標として、Q 値 (significance) の計算手法の開発を行った。

Q 値の計算は、

- (1) KEGG から公開されている機能モジュールの構成の特徴。
- (2) ホモロジーサーチ時に得られたスコア(類似度の優位性)。
- (3) モジュールを構成する遺伝子(KO)の存在度。

この3点を考慮して行った。KOの存在度は、KOがアサインされたメタゲノム中の配列数を各KOの平均長で標準化した数値を算出し、その値を各KOの存在度(アバンダンス)として用いた(図1)。

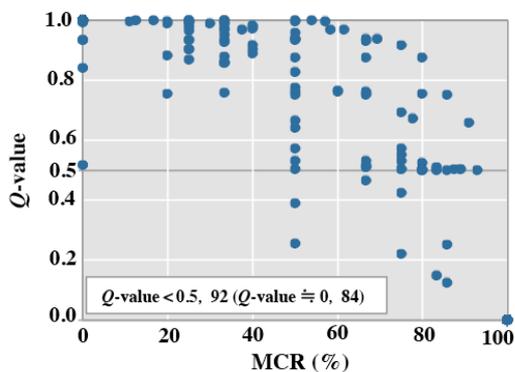


図1. GS000cのメタゲノムによるMCRとQ値の関係

これらの結果、充足率が100%に満たない機能モジュールでもQ値の優位性から、機能ポテンシャルの存在を優位に示唆できることがわかった。また、このQ値の優位性を示すには、対象とする環境にもよるが、300万配列の遺伝子セットがあれば問題ないことがわかった。一方、MCRが100%となり、モジュールが完成した場合は、KOのアバンダンスをもとに、各モジュールのアバンダンスを計算し、モジュールを完成した生物種数も算出した(図2)。

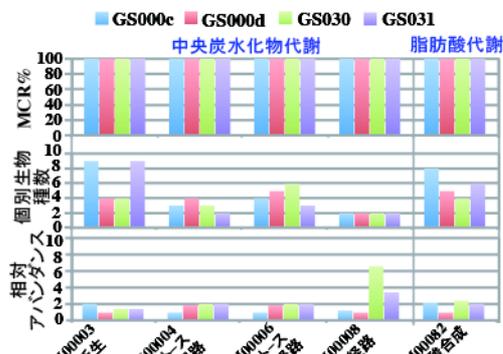


図2. 充足されたモジュールの相対アバンダンスと生物種数

次に、個別生物種ごとの各機能モジュールの充足率(MCR)の計算結果に基づき、近年、環境中での窒素や炭素循環への役割が注目されるようになったアーキアの機能分類をクラスタリング手法やPCA等の統計学的手法を用いて行った。その結果、同じ生理代謝機能を有する生物種は、系統学的分類パターンに必ずしも依存せずクラスタ化され、各機能モジュールの充足率に基づき機能分類が可能であることがわかった。一方、実環境にはゲノム既知生物種との類似性が低い生物種が数多く含まれているが、メタゲノム的手法を駆使すれば、未培養菌からのゲノム再構築が可能である。そこで、機能未知生物種の生理代謝機能の推測を機能未知の未培養菌ゲノム配列から計算されたMCRパターンに基づいて行ったところ、系統的には独立栄養型栄養生物に近いが、機能的には従属栄養生

物であることがわかった(図3)。

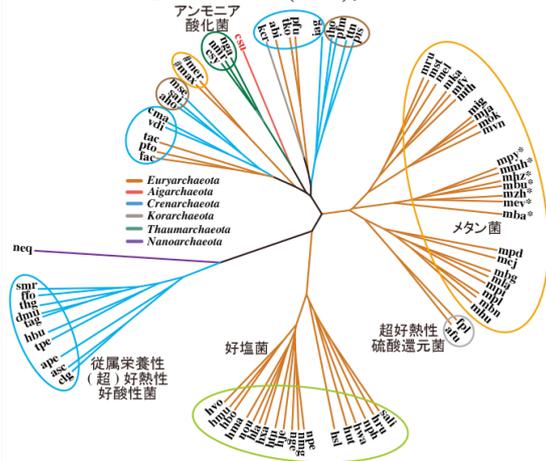


図3. MCRパターンに基づくアーキアの機能的分類

MCRは研究代表者らが開発中のMAPLEシステムで自動的に計算されるが、MCRが100%に満たないモジュールでも、Q値によって完成度が高いと示唆される場合がある。そこで、実際手作業でモジュールの再評価を行ったところ、完成型モジュールであることがわかった。このように、ゲノム情報が既知であれば未培養菌であってもMCRパターンに基づき機能分類が可能で、また未完成のモジュールであってもQ値によって高い完成度が示唆された場合には、再評価を行う価値があることが示された。

最後に、既知の海洋メタゲノム配列に潜む様々な生理代謝機能をMCR、モジュールの存在度(abundance)、モジュール機能を担う生物種の組成を数値化し、比較することによって、地理的に極めて類似する海域間の機能ポテンシャルにも大きな違いがあることを見出した。

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)

1. Hidetoshi Takami, Wataru Arai, Kazuhiro Takemoto, Ikuo Uchiyama, Takeaki Taniguchi. (2015) Functional classification of uncultured "Ca. Caldichaeum subterraneum" using the Maple system. *PLoS ONE*, vol. 10, e0132994. 査読あり. doi: 10.1371/journal.pone.0132994

[学会発表] (計12件)

1. 荒井 渉、谷口 丈晃、竹本和広、守屋勇樹、五斗進、高見英人、MAPLE 2.1による生理・代謝機能ポテンシャルの比較ゲノム解析、第10回日本ゲノム微生物学会、2016年3月4日(東京工業大学、東京都・目黒区)
2. 荒井 渉、谷口 丈晃、竹本和広、守屋勇樹、五斗進、高見英人、MAPLE 2.1による生理・代謝機能ポテンシャルの比較メタゲノム解析、第30回日本微生物生態学会、

- 2015年10月23日(土浦亀城プラザ、茨城県・土浦市)
3. 諏訪裕一、村上由夏、荒井渉、池田拓哉、勝山千恵、小椋義俊、林哲也、吉永郁夫、福原康平、黒岩恵、高見英人、メタゲノム解析による生理・代謝機能ポテンシャルから見た富栄養湖沼(北浦)の窒素循環像、第30回日本微生物生態学会、2015年10月23日(土浦亀城プラザ、茨城県・土浦市)
 4. Hideto Takami, Wataru Arai, Kazuhiro Takemoto, Ikuo Uchiyama, Takeaki Taniguchi. Functional classification of uncultured “*Ca. Caldiarchaeum subterraneum*” using the Maple system、第30回日本微生物生態学会、2015年10月23日(土浦亀城プラザ、茨城県・土浦市)
 5. Hideto Takami, Wataru Arai, Yoshihiro Takaki, Toyoda Atsushi, Takeaki Taniguchi, Shinro Nishi, Taishi Tsubouchi, Mikihiko Kaewai, Fumio Inagaki. Metagenomic characterization of carbon fixation pathways in deep seafloor microbial communities. ISME15, August 25th, 2015, Seoul-Korea
 6. Wataru Arai, Takeaki Taniguchi, Yuki Moriya, Susumu Goto, Hideto Takami. MAPLE system for evaluation of metabolic and physiological potential harbored in genome and metagenome. ISME15, August 25th, 2015, Seoul-Korea
 7. 高見英人、荒井渉、竹本和広、内山郁夫、谷口丈晃、MAPLE システムによるアーキアの機能分類と未培養アーキア “*Ca. Caldiarchaeum subterraneum*” の生理・代謝ポテンシャル解析、第28回日本アーキア研究会、2015年7月23日(愛媛大学、愛媛県・松山市)
 8. 荒井渉、谷口丈晃、五斗進、守屋勇樹、竹本和広、黄鶴、高見英人、メタゲノム中の生理代謝機能ポテンシャルを評価する新手法” MAPLE” システム、日本農芸化学会 2015年度大会、2015年3月27日(岡山大学、岡山県・岡山市)
 9. 高見英人、荒井渉、内山郁夫、谷口丈晃、未培養アーキアの MAPLE システムを用いた比較機能解析、第9回日本ゲノム微生物学会、2015年3月6日(神戸大学、兵庫県・神戸市)
 10. 高見英人、荒井渉、内山郁夫、谷口丈晃、MAPLE システムによる未培養アーキア *Ca. ‘Caldiarchaeum subterraneum’* の生理・代謝ポテンシャル解析、日本農芸化学会 2015年度大会、2015年3月27日(岡山大学、岡山県・岡山市)
 11. 高見英人、メタゲノムから見た海底生態系の機能ポテンシャル、環境微生物系合同大会、2014年10月22日(アクトシティー浜松、静岡県・浜松市) 招待講演
 12. 荒井渉、谷口丈晃、守屋勇樹、五斗進、磯部一夫、高見英人、ゲノム・メタゲノム中の潜在的機能評価システム” MAPLE” . 環境微生物系合同大会、2014年10月19日(アクトシティー、静岡県・浜松市)
- [図書] (計1件)
1. Hideto Takami et al., Springer, Encyclopedia of Metagenomics, 2015, 727 (525-539)
6. 研究組織
- (1) 研究代表者
高見 英人(TAKAMI, Hideto)
国立研究開発法人・海洋研究開発機構・海底資源研究開発センター・上席研究員
研究者番号:70359165
 - (2) 研究分担者
竹本 和広(TAKEMOTO, Kazuhiro)
九州工業大学・情報工学研究院・准教授
研究者番号:40512356