

令和 5 年 6 月 23 日現在

機関番号：82626

研究種目：基盤研究(B)（一般）

研究期間：2018～2022

課題番号：18H03367

研究課題名（和文）Microbiome mining: machine learning for discovery of genetic dark matter, metabolic pathways, and ecological processes from metagenomes

研究課題名（英文）Microbiome mining: machine learning for discovery of genetic dark matter, metabolic pathways, and ecological processes from metagenomes

研究代表者

延 優（NOBUMASARU・KONISHI）（NOBU, Masaru Konishi）

国立研究開発法人産業技術総合研究所・生命工学領域・主任研究員

研究者番号：40805644

交付決定額（研究期間全体）：（直接経費） 13,400,000円

研究成果の概要（和文）：生命は様々な生化学的な能力を持つが、そのすべての基盤が各細胞内に格納されているゲノム及び遺伝子に刻まれている。しかし、その遺伝子の半数以上が機能がまだ特定されていない未知遺伝子であり、我々はまだ生命が秘める能力のごく一部しか捉えられていない。本研究では未知遺伝子をもたらし新たな機能を紐解くことで、様々な生態系の物質循環プロセスを駆動する未知生物現象を発見し、見落とされてきた古代と現代の生命の進化の解明にも成功した。

研究成果の学術的意義や社会的意義

持続的な社会実現において地球上の様々な生態プロセスを理解し、社会の自然環境への影響を最小限にする必要がある。地球上の多様な生態系の物質循環、社会から自然界への廃棄物の流れ、どちらの中心にも微生物の特異的な能力による生化学反応がある。しかし、現代のデータサイエンスが駆動する学術界では、生命の研究は進めば進むほど未知な領域が拡大していく。本研究ではこの膨大な未知微生物・未知遺伝子データの解読を軸に、地球や社会にとって有用な微生物機能を発掘する手法の開発に貢献し、さらにこれまで見落とされてきた進化の道筋の解明にも成功した。

研究成果の概要（英文）：The basis for all of life's various biochemical capabilities is encoded in the genome and genes stored within each cell. However, more than half of the genes are unknown genes whose functions have not yet been identified, and we have only captured a small fraction of life's hidden capabilities. By unraveling the new functions of these unknown genes, we have discovered unknown biological phenomena that drive material-cycling processes in various ecosystems, and have also succeeded in elucidating the evolution of life in both ancient and modern times, which has been overlooked.

研究分野：Evolution

キーワード：Metagenomics

科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属します。

1. 研究開始当初の背景

土壌、堆積物、海底の地表から数ミリメートル、岩盤の地下数キロメートル、そして人間の体内でさえ、酸素のない生態系が存在し、そこでは「嫌気性」微生物が生態系と生物地球化学サイクルを支配している。しかし、嫌気性微生物は生来代謝が複雑で増殖が難しいため、その代謝ポテンシャルの多くは未知のままである。隠蔽された海洋硫黄循環、共生メタン酸化、炭化水素無機化など、多くの生物地球化学的に重要な微生物プロセスはまだ未解明である。我々は、これらの生態学的謎のメカニズムは、地球の嫌気性生態系全体に偏在する、多数の未解明遺伝子と微生物の中に隠されていると考えている。第一に、生態系には、培養も特性解析もされたことのない、非常に多様で豊富な生物が広く生息していることがわかってきた。第二に、培養生物と未培養生物のゲノムの両方が、未知の遺伝子の膨大なプールを保有している。培養微生物のゲノムでさえ、25~35%の遺伝子が機能未知である。未知の生物はどうだろうか？ハイスループット DNA シーケンスに基づく生態系レベルのゲノミクス（「メタゲノム」）により、環境サンプルから未培養生物の個々のゲノムを直接単離することが可能になり始めている。このようなゲノムから、未培養生物の遺伝子の 60~70%は機能が未確定であることがわかった。最後に、ほとんどの未培養微生物群（合計 100 部門：~30 部門が培養され、~70 部門が未解析）の代謝に関する知見はほとんどなく、彼らの異化能力は覆い隠されたままである。

2. 研究の目的

ゲノミクスの登場により、各微生物のゲノムに含まれる遺伝子の少なくとも 25~70%の機能は不明なままであることが明らかになった。さらに、まだ培養もゲノム解析もされていない微生物が驚くほど膨大に存在することから、この未知の遺伝子のプール（genetic dark matter; GDM）は、生化学的に検査された遺伝子の範囲（1000 倍以上）よりもはるかに広いことが示唆されている。ここでは、この謎めいた遺伝子の膨大なプールを「遺伝的暗黒物質」（GDM）と呼ぶ。このゲノムの時代において、GDM に光を当てることは、世界中の微生物が駆動する生物地球化学プロセスの理解を再定義し、実験的微生物学者や生態学者が築いた基盤の上に構築するための次のフロンティアである。我々は、GDM の生化学的役割とは何か、また GDM がどのように新規の経路レベル、微生物レベル、生態系レベルの機能に寄与しているのかを明らかにすることを目的とした。この深い知識ギャップを埋めるために、GDM の機能や新規代謝経路を de novo で予測するための機械学習とモデリングを採用する。生態系ゲノミクス（メタゲノミクス）により、注目される生息地から未培養生物のゲノムを分離し、機械学習を適用するための包括的な GDM データベースを公開する。これにより、既知および新規の遺伝子、パスウェイ、生物間の相互作用の外挿が可能になり、我々が「マイクロバイーム・マイニング」と呼ぶ新分野となる。

3. 研究の方法

メタゲノム解析による微生物ゲノムデータベースの拡充を目指し、様々な環境から未培養微生物のゲノムを取得した。対象となる環境は、海洋（海底堆積物など）、土壌（泥炭地など）、地下水（地下水など）、人工環境（嫌気性バイオリクターなど）など、注目度の高い無酸素生態系である。次世代 DNA シーケンス（Illumina HiSeq 2500 および PacBio Sequel System）と提案者が開発した高精度バイオインフォマティクス解析パイプラインを用いて、環境サンプルから新規ゲノムを直接取得する。これは、複数の最先端のアセンブリー（SPAdes、IDBA_UD、megahit など）およびゲノムクラスター化（MaxBin、MetaBAT、MyCC など）アルゴリズムを統合したもので、微生物群集の 0.1%程度の生物についても、数百の高品質ゲノムを正確に復元することができる。

最先端の遺伝子クラスタリング技術を用い GDM の分類、機能推定、及び GDM が構成する未知代謝経路の予測を行った。すなわち、(i) 同じゲノムに共起しやすい遺伝子、または (ii) 近傍に存在しやすい遺伝子は、生化学的に相互作用し、パスウェイを形成する可能性が高い。このロジックに基づき、遺伝子ネットワークから代謝経路の予測を最適化するために、代謝経路が十分に解明されている遺伝子や生物の共起ネットワーク解析と近傍ネットワーク解析を行った。また、新規代謝経路と既知代謝経路の相関を同定し、代謝カップリングをさらに同定した。このような遺伝子間および代謝経路間の解析から、代謝経路における機能的ギャップ、de novo タンパク質構造モデリング、文脈化（すなわち、宿主生物のニッチ、系統、生息地）、および他の代謝経路との関係に基づいて、GDM 遺伝子の機能を推測することができた。

未培養生物のゲノム、GDM を誘発する新規パスウェイ、GDM 機能推定により、嫌気性生息地全体で起こっている見過ごされている生態学的ニッチと生物地球化学のプロセスについて、これまでにない洞察を打ち出す。このような生態系におけるこれらのプロセスの影響と普及をさらに確認するために、生態系全体の遺伝子発現解析（「メタトランスクリプトミクス」）を通して、対象となる代謝経路の活性を間接的に測定する。この遺伝子発現解析は、固有の GDM、代謝経路、生物の潜在的な寄与を解明するだけでなく、これらのパスウェイや生物がどのように相互作用

して複雑な生物地球化学的マクロプロセスを形成しているのか、最初の手がかりを提供する。

4. 研究成果

生命は大きく分けて3種類に分類される：複雑な細胞構造を持つ我々真核生物と単純な細胞構造をもつ2種類の原核生物（真正細菌、古細菌）。主な研究成果として、本研究課題では、(1)古代生命の代謝機能を保持するレア未知微生物の発見、(2)古細菌の祖先の代謝機能の推定、(3)古細菌から真核生物への複雑化の進化の道筋の再構築に成功した。学术界にて広く認知されている学術誌 Nature, PNAS Nexus や微生物学分野最高峰の学術誌 ISME Journal に誌上発表できている。

古代生命の代謝機能を保持するレア未知微生物の発見

超苦鉄質岩の蛇紋岩化は、微生物の無機栄養代謝を支える水素分子 (H₂) を供給する一方で、高アルカリ性や限られた電子受容体の利用可能性など、極めて困難な条件をもたらす。2つの蛇紋岩化活性系を調査した結果、従来の H₂-/CO₂ 依存的なホモアセトジェネシスは、ピコモルレベルの CO₂ により原位置では熱力学的に不利であることが明らかになった。メタゲノミクスと熱力学を通して、生息環境に適応した代謝を行うユニークな分類群を発見した。その中には、新しい深部分岐門である "Ca. Lithacetigenota" と呼ばれる、蛇紋岩を擁する系にのみ生息し、H₂ を利用する代替的な岩石肥沃化様式をコードする遺伝子を保有している。これらの推定される代謝は、CO₂ ではなく、その場で検出された蛇紋岩由来と推定される還元炭素化合物、すなわちギ酸およびグリシンを利用する。前者は部分的なホモ酢酸生成経路を用い、後者は希少なセレノプロテインであるグリシン還元酵素を介する経路を用いる。微生物群を調査した結果、これまで検出されてこなかった古代のグリシン還元酵素は多様であり、蛇紋岩が棲息する環境ではほぼどこにでも存在することがわかった。Ca. Lithacetigenota グリシン還元酵素は、その基本的な系統を表しており、異化型グリシン還元は、高アルカリ性で CO₂ が乏しい条件下でも地起源の H₂ からエネルギーを得るための、テラバクテリアによる古くからのバクテリアの技術革新であることを示唆している。ここで示されたユニークな非 CO₂ 還元代謝は、蛇紋岩化に関連した環境における重要な障害を克服するために極限環境微生物が採用する可能性のある戦略、すなわち初期地球における原始的な無機栄養代謝に関連する可能性のある特徴に光を当てるものである。

古細菌の祖先の代謝機能の推定

遺伝子のメタン生成は太古の代謝として広く受け入れられているが、その正確な進化の軌跡については依然として激しい論争が続いている。メタン生成の発生時期、祖先の形態、同種の代謝との関係については、さまざまな説が存在する。本稿では、メタン生成の古さを示す新たな証拠となる、補酵素の生合成に関与する代謝関連タンパク質の系統樹を報告する。さらに、主要な異化に関与するタンパク質の系統を再検討した結果、最後の古細菌の共通祖先 (LACA) は、汎用性の高い H₂-、CO₂-、メタノールを利用するメタン生成を行うことが可能であったことが示唆された。メチル/アルキル-S-CoM 還元酵素ファミリーの系統学的解析に基づき、我々は、現在のパラダイムとは対照的に、基質特異的機能は、補酵素 F430 を用いた自己触媒実験から予測されるように、タンパク質を含まない反応に由来すると思われる非特異的な祖先まで遡る並行進化を通して出現したことを提案する。LACA の後、メタン生成リソオートトロフィーを中心とする継承/喪失/革新は、古代のライフスタイルの分岐と一致し、それは現存する古細菌のゲノム的に予測される生理学に明確に反映されている。このように、メタン生成は古細菌の特徴的な代謝であるだけでなく、祖先古細菌がとってきた謎めいたライフスタイルと、今日顕著な生理学に至る変遷を解明する鍵でもある。

古細菌から真核生物への複雑化の進化の道筋の再構築

真核生物の起源については、いまだに不明な点が多い。現在のデータでは、真核生物は「アスガルド」古細菌として知られる古細菌系統から出現した可能性が示唆されている。これらの古細菌には真核生物に類似したゲノム上の特徴が見られるにもかかわらず、古細菌から真核生物への進化的移行は、培養された代表的な古細菌やそれに対応する生理学的知見がないため、依然として不明である。本論文では、Lokiarchaeota に関連するアスガルド古細菌を、海洋深部の堆積物から 10 年かけて分離したことを報告する。この古細菌 'Candidatus Prometheoarchaeum syntrophicum' 株 MK-D1 は、嫌気性で非常に成長が遅く、小さな球菌（直径約 550nm）で、合成栄養によってアミノ酸を分解する。アスガルド古細菌には真核生物のような細胞内複合体が提案されているが、この分離株には目に見えるオルガネラのような構造はない。その代わりに、Ca. P. syntrophicum は真核生物様な未知遺伝子を多数保持し、他の古細菌では見られない形態学的に複雑で長く枝分かれする独特の突起を持っている。培養とゲノミクスから得られたデータ、真核生物様な未知遺伝子、既存の文献の理性的な解釈に基づき、我々は真核形成の仮説的モデルを提案した。

5. 主な発表論文等

〔雑誌論文〕 計7件（うち査読付論文 7件/うち国際共著 0件/うちオープンアクセス 1件）

1. 著者名 Kurth Julia M., Nobu Masaru K., Tamaki Hideyuki, de Jonge Nadieh, Berger Stefanie, Jetten Mike S. M., Yamamoto Kyosuke, Mayumi Daisuke, Sakata Susumu, Bai Liping, Cheng Lei, Nielsen Jeppe Lund, Kamagata Yoichi, Wagner Tristan, Welte Cornelia U.	4. 巻 15
2. 論文標題 Methanogenic archaea use a bacteria-like methyltransferase system to demethoxylate aromatic compounds	5. 発行年 2021年
3. 雑誌名 The ISME Journal	6. 最初と最後の頁 3549 ~ 3565
掲載論文のDOI (デジタルオブジェクト識別子) 10.1038/s41396-021-01025-6	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Koto Akiko, Nobu Masaru K., Miyazaki Ryo	4. 巻 11
2. 論文標題 Deep Sequencing Uncovers Caste-Associated Diversity of Symbionts in the Social Ant <i>Camponotus japonicus</i>	5. 発行年 2020年
3. 雑誌名 mBio	6. 最初と最後の頁 e00408-20
掲載論文のDOI (デジタルオブジェクト識別子) 10.1128/mBio.00408-20	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Preiner Martina, Igarashi Kensuke, Muchowska Kamila B., Yu Mingquan, Varma Sreejith J., Kleinermanns Karl, Nobu Masaru K., Kamagata Yoichi, Tuysuz Harun, Moran Joseph, Martin William F.	4. 巻 4
2. 論文標題 A hydrogen-dependent geochemical analogue of primordial carbon and energy metabolism	5. 発行年 2020年
3. 雑誌名 Nature Ecology & Evolution	6. 最初と最後の頁 534 ~ 542
掲載論文のDOI (デジタルオブジェクト識別子) 10.1038/s41559-020-1125-6	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Imachi H, Nobu MK, Nakahara N, Morono Y, Ogawara M, Takaki Y, Takano Y, Uematsu K, Ikuta T, Ito M, Matsui Y, Miyazaki M, Murata K, Saito Y, Sakai S, et al.	4. 巻 577
2. 論文標題 Isolation of an archaeon at the prokaryote-eukaryote interface	5. 発行年 2020年
3. 雑誌名 Nature	6. 最初と最後の頁 519 ~ 525
掲載論文のDOI (デジタルオブジェクト識別子) 10.1038/s41586-019-1916-6	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

1. 著者名 Suenami Shota, Nobu Masaru K, Miyazaki Ryo	4. 巻 9
2. 論文標題 Community analysis of gut microbiota in hornets, the largest eusocial wasps, <i>Vespa mandarinia</i> and <i>V. simillima</i>	5. 発行年 2019年
3. 雑誌名 Scientific Reports	6. 最初と最後の頁 na
掲載論文のDOI (デジタルオブジェクト識別子) 10.1038/s41598-019-46388-1	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Mei Ran, Kaneko Masanori, Imachi Hiroyuki, Nobu Masaru K.	4. 巻 2
2. 論文標題 The origin and evolution of methanogenesis and Archaea are intertwined	5. 発行年 2023年
3. 雑誌名 PNAS Nexus	6. 最初と最後の頁 na
掲載論文のDOI (デジタルオブジェクト識別子) 10.1093/pnasnexus/pgad023	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Nobu Masaru K., Nakai Ryosuke, Tamazawa Satoshi, Mori Hiroshi, Toyoda Atsushi, Ijiri Akira, Suzuki Shino, Kurokawa Ken, Kamagata Yoichi, Tamaki Hideyuki	4. 巻 17
2. 論文標題 Unique H ₂ -utilizing lithotrophy in serpentinite-hosted systems	5. 発行年 2022年
3. 雑誌名 The ISME Journal	6. 最初と最後の頁 95 ~ 104
掲載論文のDOI (デジタルオブジェクト識別子) 10.1038/s41396-022-01197-9	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計0件

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
---------------------------	-----------------------	----

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8 . 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------