

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第5795579号

(P5795579)

(45) 発行日 平成27年10月14日 (2015. 10. 14)

(24) 登録日 平成27年8月21日 (2015. 8. 21)

(51) Int. Cl.

F I

C 1 2 Q 1/68 (2006. 01)

C 1 2 Q 1/68 A

C 1 2 M 1/00 (2006. 01)

C 1 2 M 1/00 A

C 1 2 N 15/09 (2006. 01)

C 1 2 N 15/00 Z N A F

C 1 2 Q 1/06 (2006. 01)

C 1 2 N 15/00 A

G O 1 N 33/53 (2006. 01)

C 1 2 Q 1/06

請求項の数 6 (全 169 頁) 最終頁に続く

(21) 出願番号 特願2012-517804 (P2012-517804)
 (86) (22) 出願日 平成22年6月25日 (2010. 6. 25)
 (65) 公表番号 特表2012-531211 (P2012-531211A)
 (43) 公表日 平成24年12月10日 (2012. 12. 10)
 (86) 国際出願番号 PCT/US2010/040106
 (87) 国際公開番号 W02010/151842
 (87) 国際公開日 平成22年12月29日 (2010. 12. 29)
 審査請求日 平成25年6月20日 (2013. 6. 20)
 (31) 優先権主張番号 61/317, 644
 (32) 優先日 平成22年3月25日 (2010. 3. 25)
 (33) 優先権主張国 米国 (US)
 (31) 優先権主張番号 61/259, 565
 (32) 優先日 平成21年11月9日 (2009. 11. 9)
 (33) 優先権主張国 米国 (US)

(73) 特許権者 506115514
 ザ リージェンツ オブ ザ ユニバーシ
 ティ オブ カリフォルニア
 アメリカ合衆国 カリフォルニア州 94
 607 オークランド フランクリン ス
 トリート 1111 トウエルフス フロ
 ア
 (74) 代理人 100108453
 弁理士 村山 靖彦
 (74) 代理人 100064908
 弁理士 志賀 正武
 (74) 代理人 100089037
 弁理士 渡邊 隆
 (74) 代理人 100110364
 弁理士 実広 信哉

最終頁に続く

(54) 【発明の名称】 系統発生分析のための方法およびシステム

(57) 【特許請求の範囲】

【請求項 1】

1つのアッセイにおける試料中の複数の操作的分類単位 (OTU) の存在を判定するための方法であって、

a) 標識された核酸を含む試料を準備する工程；

b) 複数のOTUのためのパーフェクトマッチ (PM) プローブとミスマッチ (MM) プローブとの対と前記核酸を接触させる工程であって、PMプロローブは、OTUに対応するポリヌクレオチドに相補性を有するオリゴヌクレオチドであり、MMプロローブは、一つ以上のヌクレオチド位置で対応するPMプロローブとは異なるオリゴヌクレオチドである、工程；

c) 前記プロローブのそれぞれについてのハイブリダイゼーションシグナル強度を測定する工程；並びに

d) 前記核酸が複数のOTUのOTUから由来するか否かを判定する工程であって、前記判定は、

(i) PMプロローブとMMプロローブのシグナル強度を比較する工程；

(ii) 試料に添加された既知の核酸を相補する陽性対照プロローブについてのシグナル強度の分布を測定する工程；

(iii) PMプロローブによって相補される何れの核酸も相補しない陰性対照プロローブについてのシグナル強度の分布を測定する工程；

(iv) 前記工程 (i) ~ (iii) のシグナル強度及び分布に基づいて、少なくとも一対のPM及びMMプロローブが、試料中のOTUの存在の指標となる可能性を計算する工程

10

20

についてのコンピュータ判読可能な指示を含むコンピュータシステムを使用することを含む、工程を含む、方法。

【請求項 2】

(A) 工程(iv)が、少なくとも一対のPM及びMMプローブの応答スコアを計算することを含み、応答スコアが、特定のプローブ対が標的配列に結合しており、バックグラウンドシグナルでない前記可能性の測定値であり、及び/または

(B) 工程(d)が、対応するPMプローブの推定上のクロスハイブリダイゼーション能に基づいて前記可能性についてペナルティを課すことを更に含む、請求項 1 に記載の方法。

【請求項 3】

前記PM及びMMプローブのシグナル強度に基づいて、OTU存在量を計算する工程を更に含む、請求項 1 に記載の方法。

【請求項 4】

少なくとも 2 つの試料の間のOTUの存在または存在量を比較する工程を更に含む、請求項 1 に記載の方法。

【請求項 5】

前記少なくとも 2 つの試料の一方が不健康な個体から得た試料であり、前記少なくとも 2 つの試料の他方が全体として健康な集団から由来するものであり、さらに不健康な状態または条件のバイオシグネチャーの指標を同定する工程を更に含む、請求項 4 に記載の方法。

【請求項 6】

コンピュータによって実行された際に、1 つのアッセイ由来の複数のOTUについてのパーフェクトマッチ(PM)プローブとミスマッチ(MM)プローブの対を含むマイクロアレイから得たハイブリダイゼーションシグナル強度を測定し、核酸が複数のOTUのOTUから由来するか否かを判定する方法を実施する指示を含む一過的ではないコンピュータ判読可能な媒体であって、

(a) PMプローブは、OTUに対応するポリヌクレオチドに相補性を有するオリゴヌクレオチドであり、MMプローブは、一つ以上のヌクレオチド位置で対応するPMプローブとは異なるオリゴヌクレオチドであり；

(b) 前記アレイは標識された核酸を含む試料によって接触され；

(c) 前記判定は、

(i) PMプローブとMMプローブのシグナル強度を比較する工程；

(ii) 試料に添加された既知の核酸を相補する陽性対照プローブについてのシグナル強度の分布を測定する工程；

(iii) PMプローブによって相補される何れの核酸も相補しない陰性対照プローブについてのシグナル強度の分布を測定する工程；

(iv) 前記工程(i)～(iii)のシグナル強度及び分布に基づいて、少なくとも一対のPM及びMMプローブが、試料中のOTUの存在の指標となる可能性を計算する工程によって実施される、一過的ではないコンピュータ判読可能な媒体。

【発明の詳細な説明】

【技術分野】

【0001】

相互参照

本願は、以下の同時係属の米国仮特許出願、2009年6月26日に出願された米国出願第61/220,937号[代理人整理番号IB-2733P]、2009年11月9日に出願された米国出願第61/259,565号[代理人整理番号IB-2733P1]、2010年3月25日に出願された米国出願第61/317,644号[代理人整理番号IB-2733P2]、2010年5月24日に出願された米国出願第61/347,817号[代理人整理番号IB-2733P3]に関し、これらの優先権を主張し、そのそれぞれは、参照により本明細書に組み込まれている。

【0002】

連邦政府支援の研究に関する陳述

本発明は、エネルギー省によって授与された契約番号DE-AC02-05CH11231の下での政府支援、国土安全保障省からの補助金、およびカリフォルニア州水質委員会(State of California Water Quality Board)からの契約番号07-576-550-0を用いて行われた。政府は、本発明において、ある特定の権利を有する。

【背景技術】

【0003】

10^{30} もの多くの微生物ゲノムを包括的に用いて、複数の異なる環境条件および宿主条件にわたって、マイクロバイーム内およびマイクロバイーム間の多様性は、よく認識されている(Huseら、(2008)、PLoS Genetics 4(11): e1000255)。この多様性の結果として、マイクロバイームの内容を特徴づけることは、現在の手法にとっての課題である。第一に、標準的な培養技術は、自然における微生物のほんの小さい部分を維持することに成功している。配列決定などのより直接的なプロファイリングの手段は、2つのさらなる課題に直面している。所与の試料中の非常に多くの異なるゲノム、およびメンバー同士の間の相動性の程度の両方は、すでに労力を要する手順について複雑な問題を提示する。

【0004】

核酸およびタンパク質などのバイオポリマーは、疾患を診断し、または生物を同定するために、有用な遺伝子の探索において同定されることが多い。しばしば、ハイブリダイゼーションまたは別の結合反応が、同定ステップの一部として使用される。可能な標的の数が試料中で増加するにつれて、様々なハイブリダイゼーション反応を検出するためのシステムの設計は、結合またはハイブリダイゼーションデータの分析とともに、困難性が増大する。マイクロバイームを含む個々の種または群が、高度に保存されたポリヌクレオチドに基づいて1つのアッセイで検出または定量化される場合と同様に、多くの類似標的が試料中に存在する場合、設計問題および分析問題が切実になる。例えば、ヒトの腸において見出される細菌の約98%は、わずか4つの細菌の区分に属する一方で、これは、系統レベルで約36,000の異なるファイトタイプを含み、99%超の配列同一性を有する(Hattoriら、(2009)、DNA Res. 16:1~12頁)。ある特定の重複する分類群をおそらく含有する一方で、他の宿主の腸によって提示される様々な環境が、様々なマイクロバイームを支持することが予期される。様々な源によって潜在的に汚染された水源などの、複数のサブ環境からの寄与が組み合わされる状況では、数千の分類群を正確に同定することは、現在の検出方法にとって重要な課題である。

【先行技術文献】

【特許文献】

【0005】

【特許文献1】米国特許第5,644,048号

【特許文献2】米国特許第5,386,023号

【特許文献3】米国特許第5,637,684号

【特許文献4】米国特許第5,602,240号

【特許文献5】米国特許第5,216,141号

【特許文献6】米国特許第4,469,863号

【特許文献7】米国特許第5,235,033号

【特許文献8】米国特許第5,034,506号

【特許文献9】米国特許第6,524,793号

【特許文献10】米国特許第5,539,083号

【特許文献11】米国特許第5,578,832号

【特許文献12】米国特許第5,556,752号

【特許文献13】米国特許第5,510,270号

【特許文献14】米国特許第6,449,562号

【非特許文献】

【0006】

- 【非特許文献 1】Huseら、(2008)、PLoS Genetics 4(11): e1000255)
- 【非特許文献 2】Hattoriら、(2009)、DNA Res. 16:1~12頁
- 【非特許文献 3】Beaucageら、Tetrahedron、49(10):1925頁(1993)
- 【非特許文献 4】Letsinger、J. Org. Chem.、35:3800頁(1970)
- 【非特許文献 5】Sprinzlら、Eur. J. Biochem.、81:579頁(1977)
- 【非特許文献 6】Letsingerら、Nucl. Acids Res.、14:3487頁(1986)
- 【非特許文献 7】Sawaiら、Chem. Lett.、805頁(1984)
- 【非特許文献 8】Letsingerら、J. Am. Chem. Soc.、110:4470頁(1988)
- 【非特許文献 9】Pauwelsら、Chemica Scripta、26:141頁(1986)
- 【非特許文献 10】Magら、Nucleic Acids Res.、19:1437頁(1991) 10
- 【非特許文献 11】Briuら、J. Am. Chem. Soc.、111:2321頁(1989)
- 【非特許文献 12】Eckstein、Oligonucleotides and Analogues: A Practical Approach、Oxford University Press
- 【非特許文献 13】Egholm、J. Am. Chem. Soc.、114:1895頁(1992)
- 【非特許文献 14】Meierら、Chem. Int. Ed. Engl.、31:1008頁(1992)
- 【非特許文献 15】Nielsen、Nature、365:566頁(1993)
- 【非特許文献 16】Carlssonら、Nature、380:207頁(1996)
- 【非特許文献 17】Denpcyら、Proc Natl. Acad. Sci. USA、92:6097頁(1995)
- 【非特許文献 18】Kiedrowshiら、Angew. Chem. Intl. Ed. English、30:423頁(1991)
- 【非特許文献 19】Letsingerら、Nucleosides & Nucleotides、13:1597頁(1994) 20
- 【非特許文献 20】2章および3章、ASC Symposium Series 580、「Carbohydrate Modifications in Antisense Research」、Y. S. SanghuiおよびP. Dan Cook編
- 【非特許文献 21】Mesmaekerら、Bioorganic & Medicinal Chem. Lett.、4:395頁(1994)
- 【非特許文献 22】Jeffsら、J. Biomolecular NMR、34:17頁(1994)
- 【非特許文献 23】Tetrahedron Lett.、37:743頁(1996))
- 【非特許文献 24】6章および7章、ASC Symposium Series 580、「Carbohydrate Modifications in Antisense Research」、Y. S. SanghuiおよびP. Dan Cook編
- 【非特許文献 25】Jenkinsら、Chem. Soc Rev.、(1995)169~176頁
- 【非特許文献 26】Rawls、C & E News、1997年6月2日、35頁
- 【非特許文献 27】Ngら A spatially addressable bead-based biosensor for simple and rapid DNA detection. Biosensors & Bioelectronics、23:803~810頁、2008 30
- 【非特許文献 28】Geiss Gら Direct multiplexed measurement of gene expression with color-coded probe pairs. Nature Biotech. 26:317~325頁、2008
- 【非特許文献 29】DeSantisら、Nucleic Acids Res. (2006) 34:W394~399
- 【非特許文献 30】Thompsonら、Nucleic Acids Res (1994) 22:4673~4680頁
- 【非特許文献 31】Edgar、Nucleic Acids Res. (2004) 32:1792~1797頁
- 【非特許文献 32】Zhouら(Appl. Environ. Microbiol. (1996) 62:316~322頁)
- 【非特許文献 33】Wuら、Appl. Environ. Microbiol. (2006) 72、4931~4941頁
- 【非特許文献 34】Hurtら、Appl. Environ. Microbiol. (2001) 67:4495~4503頁
- 【非特許文献 35】McGrathら、J. Microbiol. Methods (2008)75:172~176頁 40
- 【非特許文献 36】Gaoら、Appl. Environ. Microbiol. (2007) 73:563~571頁
- 【非特許文献 37】DeSantisら、Microbial Ecology、53(3):371~383頁、2007
- 【非特許文献 38】Froehlerら、Nucleic Acid Res. 14:5399~5407頁(1986)
- 【非特許文献 39】McBrideら、Tetrahedron Lett. 24:246~248頁(1983)
- 【非特許文献 40】Egholmら、Nature 363:566~568頁(1993)
- 【非特許文献 41】Schenら、Science 270:467~470頁(1995)
- 【非特許文献 42】Fodorら、1991、Science 251:767~773頁
- 【非特許文献 43】Peaseら、1994、Proc. Natl. Acad. Sci. U.S.A. 91:5022~5026頁
- 【非特許文献 44】Lockhartら、1996、Nature Biotechnology 14:1675頁
- 【非特許文献 45】Blanchardら、Biosensors & Bioelectronics 11:687~690頁 50

- 【非特許文献 4 6】MaskosおよびSouthern、1992、Nuc. Acids. Res. 20: 1679 ~ 11684頁
- 【非特許文献 4 7】Rheeら、Appl. Environ. Microbiol. (2004) 70:4303 ~ 4317頁
- 【非特許文献 4 8】Cheeら、Science (1996) 274:610 ~ 614頁
- 【非特許文献 4 9】Wuら、Appl. Environ. Microbiol. (2006) 72:4931 ~ 4941頁
- 【非特許文献 5 0】Zhang L.ら、A model of molecular interactions on short oligonucleotide analysis systems、Nat. Biotechnol. 2003、21(7):818 ~ 821頁
- 【非特許文献 5 1】Yergeauら、Environmental microarray analyses of Antarctic soil microbial communities. ISME J. 3:340 ~ 351頁、2009
- 【非特許文献 5 2】Lozuponeら、UniFrac-an online tool for comparing microbial community diversity in a phylogenetic context、BMC Bioinformatics、7、371頁; 2006 10
- 【非特許文献 5 3】de Hoonら、Open source clustering software、Bioinformatics、20、1453 ~ 1454頁; 2004
- 【非特許文献 5 4】Brodieら、Application of a High Density Oligonucleotide Analysis system Approach to Study Bacterial Population Dynamics during Uranium Reduction and Reoxidation. Applied Environ Microbio. 72:6288 ~ 6298頁、2006
- 【非特許文献 5 5】Team RCD(2008) R: A language and environment for statistical computing
- 【非特許文献 5 6】Goslee SC & Urban DL (2007) The ecodist package for dissimilarity-based analysis of ecological data. J Stat Softw 22(7): 1 ~ 19頁
- 【非特許文献 5 7】Evans J、Sheneman L、& Foster JA (2006) Relaxed neighbor-joining: a fast distance-based phylogenetic tree. Construction method. J Mol Evol 62:785 ~ 792頁 20
- 【非特許文献 5 8】Lozupone C & Knight R (2005) UniFrac: a new phylogenetic method for comparing microbial communities. Applied and Environmental Microbiology 71(12):8228 ~ 8235頁
- 【非特許文献 5 9】DeSantis TZら、(2006) Greengenes、a chimera-checked 16S rRNA gene database and workbench compatible with ARB. Applied and Environmental Microbiology 72(7):5069 ~ 5072頁
- 【非特許文献 6 0】Juteau P、Tremblay D、Villemur R、Bisailon JG、& Beaudet R (2005) Analysis of the bacterial community inhabiting an aerobic thermophilic sequencing batch reactor (AT-SBR) treating swine waste Applied Microbiology and Biotechnology 66:115 ~ 122頁 30
- 【非特許文献 6 1】Anderson RTら、(2003) Stimulating the in situ activity of Geobacter species to remove uranium from the groundwater of a uranium-contaminated aquifer. AEM 69: 5884 ~ 5891頁
- 【非特許文献 6 2】Hugenholz 2002、Genome Biol. 3(2): 1 ~ 8頁
- 【非特許文献 6 3】Jacobsonら、1995、Journal of Combinatorial Designs 4: 405頁
- 【発明の概要】
- 【発明が解決しようとする課題】
- 【0007】 40
- マイクロバイオームの研究は、環境変化、疾患、免疫学的機能、および生理的機能の起源への新しい洞察を提供することができるので、特定の生物、または分類群を認識することができる核酸、タンパク質、または他のプローブを設計するための改善された方法が必要とされている。同様に、高い信頼度レベルで微生物群集のメンバーの検出および定量化を可能にするデータ分析のための改善された方法も必要とされている。
- 【課題を解決するための手段】
- 【0008】
- いくつかの実施形態は、1つのアッセイで、少なくとも10,000の異なるOTUの存在、非存在、相対存在量、および/または量を判定することができる複数のプローブを備えるシステムを提供する。いくつかの実施形態では、システムは、糞便汚染を示すバイオシグネチ 50

ヤーを生成するように構成される。いくつかの実施形態では、プローブは、1つまたは複数の高度に保存されたポリヌクレオチドに選択的にハイブリダイズし、これは、16S rRNA遺伝子、23S rRNA遺伝子、5S rRNA遺伝子、5.8S rRNA遺伝子、12S rRNA遺伝子、18S rRNA遺伝子、28S rRNA遺伝子、gyrB遺伝子、rpoB遺伝子、fusA遺伝子、recA遺伝子、cox1遺伝子、nifl3遺伝子、これらに由来するRNA分子、またはこれらの組合せを含むことができる。いくつかの実施形態では、保存されたポリヌクレオチドは単位複製配列である。いくつかの実施形態では、プローブは、基質に結合することができる。いくつかの実施形態では、プローブは、アレイを形成することができる。いくつかの実施形態では、基質は、ビーズ、微小球、ガラス、プラスチック、またはシリコンを含む。いくつかの実施形態では、システムは、OTUのそれぞれの同じ高度に保存された領域に対して、シーケンシング反応を実施することができる。いくつかの実施形態では、システムは、1つまたは複数の種特異的なプローブをさらに含む。いくつかの実施形態では、OTUのそれぞれは、細菌、古細菌、または真菌である。

10

【0009】

いくつかの実施形態では、システムは、複数の陽性対照プローブをさらに含む。いくつかの実施形態では、システムは、複数の陰性対照プローブをさらに含む。いくつかの実施形態では、陰性対照プローブは、高度に保存されたポリヌクレオチド中に見出される配列と相補性でない配列を含む。いくつかの実施形態では、陽性対照プローブは、配列番号51~100から選択されるポリヌクレオチドと相補的である配列を含む。いくつかの実施形態では、陽性対照プローブは、配列番号51~100から選択される1つまたは複数の配列を含む。

20

【0010】

いくつかの実施形態では、システムは、前記OTUの存在、非存在、相対存在量、および/または量についての最終コールを行う前に、前記インターロゲーションプローブ(interrogation probe)の少なくとも1つのサブセットからデータを除去する。いくつかの実施形態では、データは、インターロゲーションプローブのクロスハイブリダイゼーション能に基づいて除去される。

【0011】

いくつかの実施形態は、95%超の信頼度レベルおよび95%超の感度レベルを伴って、1つのアッセイで、存在する全核酸の 1×10^{-3} 以下を含む1つまたは複数の第1の核酸配列を検出することができるシステムであって、1つまたは複数の第1の核酸配列と残りの標的核酸のセットは、少なくとも95%相同であるシステムを提供する。いくつかの実施形態では、システムは、糞便汚染を示すバイオシグネチャーを生じるように構成される。いくつかの実施形態では、核酸配列の1つまたは複数の、16S rRNA遺伝子、23S rRNA遺伝子、5S rRNA遺伝子、5.8S rRNA遺伝子、12S rRNA遺伝子、18S rRNA遺伝子、28S rRNA遺伝子、gyrB遺伝子、rpoB遺伝子、fusA遺伝子、recA遺伝子、cox1遺伝子、nifl3遺伝子、これらに由来するRNA分子、またはこれらの組合せである。いくつかの実施形態では、核酸は単位複製配列を含む。

30

【0012】

いくつかの実施形態は、1つのアッセイで、複数の異なるOTUの存在、非存在、相対存在量、および/または量を判定するためのシステムであって、複数のポリヌクレオチドインターロゲーションプローブ、複数のポリヌクレオチド陽性対照プローブ、および複数のポリヌクレオチド陰性対照プローブを含むシステムを提供する。いくつかの実施形態では、システムは、糞便汚染を示すバイオシグネチャーを生じるように構成される。いくつかの実施形態では、システムは、前記微生物の存在、非存在、相対存在量、および/または量についての最終コールを行う前に、前記インターロゲーションプローブの少なくとも1つのサブセットからデータを除去する。いくつかの実施形態では、データは、インターロゲーションプローブのクロスハイブリダイゼーション能に基づいて除去される。

40

【0013】

いくつかの実施形態は、95%超の信頼度を伴って、1つのアッセイで、1つのドメインの1

50

0,000を超える異なるOTU(例えば、細菌、古細菌、または真菌)の存在、非存在、相対存在量、および/または量を検出することができるシステムを提供する。いくつかの実施形態では、システムは、糞便汚染を示すバイオシグネチャーを生じるように構成される。いくつかの実施形態では、システムは、前記OTUのそれぞれにおける同じ高度に保存された領域に選択的にハイブリダイズする複数のプローブを備える。いくつかの実施形態では、システムは、前記OTUのそれぞれの同じ高度に保存された領域に対して、シーケンシング反応を実施することができる。いくつかの実施形態では、システムは、前記高度に保存された配列にハイブリダイズしない種特異的なプローブをさらに備える。いくつかの実施形態では、システムは、100の種-特異的なプローブを備える。

【0014】

10

いくつかの実施形態は、試料から1つまたは複数の微生物の存在、非存在、相対存在量、および/または量を判定するためのシステムであって、複数のOTUを備え、10TU当たりのプローブの中央値数は、26未満であるシステムを提供する。いくつかの実施形態は、試料から1つまたは複数の微生物の存在、非存在、相対存在量、および/または量を判定するためのシステムであって、複数のOTUを備え、1プローブ当たりのクロスハイブリダイゼーションの中央値数は、20未満であるシステムを提供する。いくつかの実施形態では、システムは、糞便汚染を示すバイオシグネチャーを生じるように構成される。

【0015】

いくつかの実施形態は、試料の状態を判定するための方法であって、a)前記試料を複数の異なるプローブと接触させるステップと、b)前記プローブのそれぞれについてのハイブリダイゼーションシグナル強度を測定するステップであって、前記測定は、前記試料のバイオシグネチャーを確立するステップと、c)前記試料のバイオシグネチャーを糞便汚染のバイオシグネチャーと比較するステップとを含む方法を提供する。

20

【0016】

本発明の一態様では、試料中の微生物の存在、相対存在量、および/または量の確率を求めるための方法であって、a)微生物中の高度に保存されたポリヌクレオチドと特異的にハイブリダイズしない陰性対照プローブのハイブリダイゼーションシグナル強度分布を求めるステップと、b)陽性対照プローブのハイブリダイゼーションシグナル強度分布を求めるステップと、c)複数の異なるインターロゲーションプローブについてのハイブリダイゼーションシグナル強度を求めるステップであって、インターロゲーションプローブのそれぞれは、高度に保存されたポリヌクレオチド内のセクションと相補的であるステップと、d)陰性および陽性プローブのハイブリダイゼーションシグナル強度を使用することによって、異なるインターロゲーションプローブについてのハイブリダイゼーションシグナルが、微生物の存在、相対存在量、および/または量を表す確率を求めるステップとを含む方法が提供される。いくつかの実施形態では、陰性および陽性プローブのハイブリダイゼーションシグナル強度が使用されることによって、インターロゲーションプローブのハイブリダイゼーションデータが正規化またはフィッティングされる。さらなる実施形態では、インターロゲーションプローブのハイブリダイゼーションデータの正規化またはフィッティングは、陰性および陽性対照プローブのA+T含量、または正規分布およびガンマ分布を利用する。他の実施形態では、陰性対照プローブおよび/または陽性対照プローブは、パーフェクトマッチおよびミスマッチプローブを含む。さらなる実施形態では、陰性および陽性対照プローブの正規分布およびガンマ分布は、前記プローブについてのペアディファレンス(pair difference)スコアを計算することを伴う。他の実施形態では、複数の異なるインターロゲーションプローブについてのハイブリダイゼーションシグナル強度は、各プローブのG+C含量に基づいて減衰される。

30

40

【0017】

一態様では、試料中のユニークなポリヌクレオチドまたは微生物の存在または量の確率を求めるための方法であって、a)試料を、複数の異なるプローブと接触させるステップと、b)プローブのそれぞれに対する、試料ポリヌクレオチドについてのハイブリダイゼーションシグナル強度を求めるステップと、c)ハイブリダイゼーションシグナル強度データに

50

基づく可能なリストから1つのOTU/分類群を分析から除去または希薄化し、それによって、残りのハイブリダイゼーションシグナル強度データの信頼度レベルを増大させるステップとを含む方法が提供される。いくつかの実施形態では、除去または希薄化するステップは、そのようなOTU内である特定の閾値強度をクリアする一定割合のプロープを有するOTUのみに対して実施される。いくつかの実施形態では、ある特定の閾値をクリアするOTUのみがさらに分析される。なおさらなる実施形態では、除去または希薄化するステップは、OTUに由来するプロープの、他のOTUに由来するポリヌクレオチドとの潜在的なクロスハイブリダイゼーションに基づいて、試料中に存在するOTUにペナルティを課すことによって実施される。いくつかの実施形態では、ペナリゼーションは、他のOTUとのクロスハイブリダイゼーションに関する潜在性と正に相関する。他の実施形態では、クロスハイブリダイゼーションに基づくペナリゼーションは、最低レベルから開始して、系統樹の各レベルで実施される。さらなる実施形態では、ハイブリダイゼーションシグナル強度閾値を超えるスコアを得ているペナルティを課されたOTUのみがさらに分析される。さらに他の実施形態では、OTUを含む系統樹の一部のみが分析される。

【0018】

本発明のさらなる態様では、試料中の複数の異なる生物の存在または量を判定するための方法であって、各プロープのGC含量を求めるステップと、陽性対照プロープ強度および陰性対照プロープ強度に対する各プロープ強度を比較することによって、前記プロープの量を求めるステップとを含む方法が提供される。

【0019】

本発明の別の態様では、一連の異なる生物からの1つまたは複数の生物が試料中に存在する確率を求めるためのコンピュータ実行可能ロジックであって、a)陰性対照プロープおよび陽性対照プロープの強度との比較に基づいて、個々のインターロゲーションプロープ強度が正確である尤度(likelihood)を求めるためのアルゴリズムと、b)第1の変位値閾値(quantile threshold)をクリアするOTUからのインターロゲーションプロープの強度に基づいて、個々のOTUが存在する尤度を求めるアルゴリズムと、c)前記OTUの配列を分析するプロープの、他のOTUからの配列とのクロスハイブリダイゼーションについての潜在性に基づいて、第1の変位値閾値をクリアした1つまたは複数のOTUにペナルティを課すためのアルゴリズムとを含むコンピュータ実行可能ロジックが提供される。

【0020】

さらなる態様では、試料中の1つまたは複数の微生物の存在、および任意選択により量を判定するためのコンピュータ実行可能ロジックであって、少なくとも10,000、20,000、30,000、40,000、50,000、60,000、70,000、80,000、90,000、または100,000の高度に保存されたポリヌクレオチドのそれぞれに選択的に結合する一連のプロープからの強度を分析し、前記試料中に存在するすべての種のうちの少なくとも90%、95%、97%、またはそれ以上の存在を判定するためのロジックを含むコンピュータ実行可能ロジックが提供される。判定は、少なくとも90%、95%、98%、99%、または99.5%の信頼度レベルを伴って行うことができる。

【0021】

別の態様では、試料中の1つまたは複数の微生物の存在を判定するためのコンピュータ実行可能ロジックであって、一連の少なくとも1000の異なるインターロゲーション完全プロープ(interrogation perfect probe)を分析するためのロジックと、前記判定を行うプロセスにおいて、前記インターロゲーション完全プロープの少なくとも10%から情報を廃棄するためのロジックとを含むコンピュータ実行可能ロジックが提供される。

【0022】

一態様では、プロープ選択のための方法であって、a)一連の高度に保存されたポリヌクレオチドを選択するステップと、b)前記複数のポリヌクレオチドを複数の標準ポリヌクレオチドと比較することによって、キメラ配列を同定するステップと、c)比較ステップにおいて同定されたキメラ配列を除去するステップと、d)残りのポリヌクレオチドと相補的なプロープを選択するステップとを含む方法が提供される。いくつかの実施形態では、少な

くとも500,000の高度に保存されたポリヌクレオチドが選択される。他の実施形態では、複数のポリヌクレオチドのメンバーは、これが複数の標準ポリヌクレオチドのメンバーと95%超の類似性を共有する場合、キメラ配列でないとみなされる。さらに他の実施形態では、複数のポリヌクレオチドが、これら自体と比較されることによってキメラ配列が同定される。他の実施形態では、高度に保存されたポリヌクレオチドは、16S RNA遺伝子、23S RNA遺伝子、5S RNA遺伝子、5.8S rRNA遺伝子、12S rRNA遺伝子、18S rRNA遺伝子、28S rRNA遺伝子、gyrB遺伝子、rpoB遺伝子、fusA遺伝子、recA遺伝子、cox1遺伝子、nifD遺伝子、またはこれらの組合せに由来する配列を含む。

【0023】

別の態様では、プローブ選択の方法であって、a)複数の核酸配列を選択するステップと、b)複数の核酸配列を、複数の標準核酸配列と整列させることによって、複数の核酸配列のそれぞれにおける挿入点を同定するステップと、c)少なくとも10、20、30、40、50、もしくはそれ以上の挿入点を有するか、または長さが少なくとも100核酸である挿入を有する配列を除去するステップと、d)残りの核酸と相補的なプローブを選択するステップとを含む方法が提供される。

10

【0024】

さらなる態様では、プローブ選択の方法であって、a)複数の核酸配列を選択するステップと、b)複数の核酸配列を選別するステップと、c)残りの核酸配列に対して階層的クラスタリングを実施することによってガイドツリーを生成するステップと、d)前記ガイドツリーにおける各ノードと相補的なプローブを選択するステップとを含む方法が提供される。いくつかの実施形態では、複数の核酸配列を選別するステップは、PCRプライマー人工産物を含むと同定された配列を除去するステップ、挿入を含むと同定された配列を除去するステップ、キメラと同定された配列を除去するステップ、またはこれらの任意の組合せを含む。

20

【0025】

一態様では、1つの状態を示すマイクロバイームシグネチャーを同定するための方法であって、a)前記状態を伴わない対照試料および前記状態を伴う参照試料における少なくとも1,000の異なるOTUの存在および任意選択により存在量を比較するステップと、b)前記状態と関連する1つまたは複数のOTUを同定するステップとを含む方法が提供される。いくつかの実施形態では、状態は油流出である。いくつかの実施形態では、前記対照試料に対する、前記参照試料中の前記OTUの存在および任意選択により存在量における類似性の増大は、前記状態のレメディエーションを示す。いくつかの実施形態では、前記対照試料に対する、前記参照試料中の前記OTUの存在および任意選択により存在量における類似性の程度の変化は、前記状態のレメディエーションの尺度として提供される。いくつかの実施形態では、この方法は、前記状態のレメディエーションの所定レベルに到達するまでの時間を予測するステップをさらに含む。

30

【0026】

一態様では、試料中の状態をアッセイするためのプローブを選択するための方法であって、a)前記状態を有する1つまたは複数の試験試料を、細菌、古細菌、真菌などの、1つのドメインの少なくとも10,000のOTU、または1つのドメインの各既知のOTUの存在または非存在の確率について同時にアッセイする検出システムに適用するステップと、b)前記状態を有さない1つまたは複数の対照試料を、前記検出システムに適用することによって、前記対照試料中の前記OTUの存在または非存在の確率を求めるステップと、c)対照試料と関連していない、試験試料と関連したOTUのパターンを求めるステップと、d)試験試料と関連したOTUを選択的に検出するプローブを同定するステップとを含む方法が提供される。いくつかの実施形態では、1つまたは複数の同定されたプローブが、低密度プローブシステムにおいて使用するために選択される。いくつかの実施形態では、パターンは、最大200の異なるOTUからなる。他の実施形態では、試料は、水試料であり、状態は、糞便汚染、毒性藻類ブルーム汚染、魚類病原菌の存在、点源汚染、非点源汚染、またはこれらの組合せである。いくつかの実施形態では、汚染のあるタイプのユニークなバイオシグネチャー

40

50

が、汚染源を求めるために使用される。いくつかの実施形態では、試料はヒトまたは動物試料である。いくつかの実施形態では、試料は、腸、呼吸器系、口腔、洞、鼻孔、尿生殖路、皮膚、糞便、乳房、またはこれらの組合せから得られる。いくつかの実施形態では、その試料において特徴づけられる(例えば、診断または予知される)状態は、クローン病、過敏性腸症候群、癌、鼻炎、胃潰瘍、大腸炎、アトピー、喘息、新生仔の壊死性全腸炎、アクネ、食物アレルギー、胃食道逆流症、肥満症、または歯周病である。いくつかの実施形態では、試料は、食物、水、土壌、または空気試料である。いくつかの実施形態では、試料は、森林、工芸作物、または他の植物からのものである。

【0027】

一態様では、状態についての少なくとも1つの新しいインジケータ種を同定するための方法であって、a)前記状態を伴わない対照試料を、1つの実験においてアッセイすることによって、すべての既知の細菌、古細菌、または真菌の各OTUの存在または非存在を判定するステップと、b)前記状態を伴う試験試料を、1つの実験においてアッセイすることによって、すべての既知の細菌、古細菌、または真菌の各OTUの存在または非存在を判定するステップと、c)(a)および(b)からの結果を比較することによって、存在量が、状態の変化に応じて所定の割合で変化する少なくとも1つの微生物を同定するステップであって、同定される微生物種は、前記状態についての前記新しいインジケータ種を表すステップとを含む方法が提供される。いくつかの実施形態では、同定される微生物は、状態の存在下で存在量が減少する一方で、他の実施形態では、同定される微生物は、存在量が増加する。いくつかの実施形態では、所定の割合は、存在量の少なくとも2倍の変化である。いくつかの実施形態では、所定の割合は、存在量の統計的に有意な変化である。

【0028】

別の態様では、微生物または微生物の選択された群が試料中に存在する確率を求めるためのシステムであって、開示されたアルゴリズムによって同定された2つ以上のプローブを備えるシステムが提供される。いくつかの実施形態では、システムは、95%、99%、または99.5%超の信頼度レベルで確率を求める。他の実施形態では、判定は、同時に、または1つのアッセイを使用して実施される。

【0029】

一態様では、1つのアッセイで、複数のプローブの、区別される各生物から単離された、高度に保存された核酸との選択的ハイブリダイゼーションに基づいて、95%、99%、または99.5%以上の精度/信頼度で、系統樹上の2つのOTUを区別することができるシステムが提供される。

【0030】

別の態様では、95%超の精度および/または信頼度レベルを伴って、1つのアッセイで、ある環境からの少なくとも10,000のOTUを含むマイクロバイームシグネチャーを生成することができるシステムが提供される。いくつかの実施形態では、プローブは、検出されている各生物に由来する核酸と選択的にハイブリダイズする。

【0031】

一態様では、微生物汚染源を検出するための方法であって、1つのアッセイで、前記試料中に天然に存在しない少なくとも20、50、100、またはそれ以上の微生物のOTUの存在および量を判定するステップと、OTUの存在および量のパターンを使用して汚染源を特定するステップとを含む方法が提供される。

【0032】

別の態様では、1つのアッセイで、少なくとも50の異なる糞便の分類群の存在および量を検出することができるシステムが提供される。いくつかの実施形態では、検出は、複数のプローブの、検出される各生物から単離された、高度に保存された核酸との選択的ハイブリダイゼーションに基づく。いくつかの実施形態では、検出は、表4に列挙された生物または分類群を同定する複数のプローブの選択的ハイブリダイゼーションに基づく。いくつかの実施形態では、検出は、清潔な水の分類群を示す核酸と選択的にハイブリダイズする1つまたは複数のプローブのハイブリダイゼーションを検出するステップを含み、前記

プローブは、表11に列挙された生物または分類群を同定する複数のプローブから選択される。

【0033】

さらなる態様では、水中の糞便汚染を検出するための方法であって、表4に列挙された分類群を検出する1つまたは複数のポリヌクレオチドの水試料中での存在または非存在を検出するステップを含む方法が提供される。いくつかの実施形態では、この方法は、表11中に列挙された清潔な水の分類群を示すポリヌクレオチドと選択的にハイブリダイズする1つまたは複数のプローブのハイブリダイゼーションを検出するステップをさらに含む。

【0034】

別の態様では、水試料を試験するための方法であって、前記水中の桿菌(Bacilli)、バクテロイデス(Bacteroidetes)、およびクロストリジウム(Clostridia)(BBC)種と、-プロテオバクテリア(A)種との比を計算するステップであって、1.0を超える値は、糞便汚染を示すステップを含む方法が提供される。いくつかの実施形態では、比を計算するステップは、培養、直接計数、PCRクローニング、配列決定、または遺伝子発現アレイの使用を利用しない。いくつかの実施形態では、桿菌、バクテロイデス、クロストリジウム、および-プロテオバクテリア種は、表4に列挙された種を含む。いくつかの実施形態では、BBC種とA種の比を計算するステップは、水試料を複数のプローブと接触させるステップを含む。いくつかの実施形態では、複数のプローブは、高度に保存された遺伝子と相補性である。

【0035】

さらなる態様では、毒性藻類ブルームの尤度を予測するための方法であって、a)水試料を、表6から選択されるラン藻類(Cyanobacteria)に由来する核酸に選択的に結合する複数のプローブと接触させるステップと、b)得られたハイブリダイゼーションデータを使用することによって、水試料中のラン藻類の量および組成を求めるステップと、c)環境条件を測定するステップと、d)ラン藻類の量および組成、ならびに環境条件に基づいて毒性藻類ブルームの尤度を予測するステップとを含む方法が提供される。いくつかの実施形態では、ラン藻類の核酸に対するプローブは、本方法を使用して選択され、表6に列挙された属を検出する。いくつかの実施形態では、環境条件は、水温、濁り、窒素濃度、酸素濃度、炭素濃度、リン酸濃度および/または太陽光レベルを含む。さらなる実施形態では、水の管理の決定は、毒性藻類ブルームの尤度に基づいて行われる。

【0036】

一態様では、対象の状態または対象のための療法を判定するための方法であって、前記対象に由来する試料に対して1つの核酸アッセイを実施することによって、少なくとも1000のOTUの存在および/または量を判定するステップを含む方法が提供される。

【0037】

別の態様では、試料の状態を予測するための方法であって、a)試料中の微生物の少なくとも1,000のOTUの存在または非存在の確率として、微生物個体数データを求めるステップと、b)試料中の前記微生物の1つまたは複数の遺伝子の遺伝子発現データを求めるステップと、c)発現データおよび個体数データを使用することによって、試料の状態を予測するステップとを含む方法が提供される。いくつかの実施形態では、試料は、水または土壌試料である。

【0038】

別の態様では、本発明は、油流出によって引き起こされた被害を評価するための方法を提供する。いくつかの実施形態では、この方法は、(a)前記油流出によって影響されていない1つまたは複数の位置からの1つまたは複数の試料中の少なくとも1,000のOTUの存在、非存在、および/または存在量を求め、それによって無影響のバイオシグネチャーを確立するステップと、(b)前記油流出によって影響された位置からの1つまたは複数の試料中の少なくとも1,000のOTUの存在、非存在、および/または存在量を求め、それによって油流出に影響されたバイオシグネチャーを確立するステップと、(c)前記無影響のバイオシグネチャーを、前記油流出に影響されたバイオシグネチャーと比較するステップであって、

前記バイオシグネチャーの差異は、前記油流出によって影響された前記位置のマイクロバイオームに対する影響を示すステップとを含む。いくつかの実施形態では、ステップ(b)は、第1の時間および第2の時間で実施される。いくつかの実施形態では、前記第1の時間と前記第2の時間の間の前記バイオシグネチャーの前記差異の変化を使用することによって、油流出被害のレメディエーションの進行を追跡する。

【0039】

参照による組み込み

本明細書に述べられるすべての刊行物、特許、および特許出願は、各個々の刊行物、特許、または特許出願が、参照により組み込まれるように具体的に、かつ個々に示されているのと同じ程度に、参照により本明細書に組み込まれている。

10

【0040】

本発明の新規な特徴は、添付の特許請求の範囲に具体的に記載されている。本発明の原理が利用される、例示的な実施形態を記載する以下の詳細な説明を参照することにより、本発明の特徴および利点はよりよく理解されることになり、その添付図面は以下の通りである。

【図面の簡単な説明】

【0041】

【図1】適切なコンピュータシステム環境の例を示す図である。

【図2】本発明の方法により得られるデータの遠隔取得または分析のためのネットワークシステムを示す図である。

20

【図3】プローブ選択プロセスのフローチャートを示す図である。

【図4】定量的標準(QS)プローブおよび陰性対照(NC)プローブからの観察されたペアディファレンススコア、dの分布を明示する図である。

【図5】79のアレイにわたるスケールの変動を示す図である。

【図6】計算負荷分散のためのプレパーティションプロセスを示す図である。

【図7】汚染源を決定するために3種の潜在的な汚染源、すなわち、下水、腐敗層汚泥および畜産廃棄物と比した汚染水試料中の微生物群集組成を比較するベクトル図である。

【図8】プールされたクリーンルーム試料(x軸)ごとにPhyloChipにより検出されたOTUの数(y軸)を示す対数棒グラフである。孢子カウントにより検出された孢子の数が示されている。

30

【図9】各プールされたクリーンルーム試料において検出された共通な科とユニークな科のネットワークを示す図表示である。

【図10A】プールされたクリーンルーム試料についてのPhyloChip上でのプローブのペア拡散スコア頻度のグラフである。

【図10B】関係ネットワークとしてPCR陰性プールされたクリーンルーム試料においてPhyloChipにより検出された共通に検出される門(phyla)を示す図表示である。

【図11】2つの異なるPhyloChip実験で観察された糞便細菌(*Faecalibacterium*)OTU 36742に対するプローブ応答を比較している図表示である。

【図12】2つの異なるPhyloChip実験で観察されたルミノコッカス(*Ruminococcus*)OTU 38712に対するプローブ応答を比較している図表示である。

40

【図13】陰性対照プローブのd観察を示す密度プロットである。

【図14】16S単位複製配列濃度対PhyloChip応答を示すグラフである。

【図15】ペア「応答スコア」r分布(新規)に基づく検出アルゴリズム対陽性フラクション計算(G2 PhyloChipを用いて以前使用された)のボックスプロット比較である。

【図16】rスコアメトリック対受信者動作特性(R.O.C)プロットによるpfの比較を示す2つのグラフである。

【図17】類似する生物群集形態序列(ordination)クラスターからのPhyloChip結果を示すグラフである。

【図18】類似する生物群集形態序列クラスターからのPhyloChip結果を示すグラフである。

50

【図 19】4つのサンプリング地点がまったく異なっており、生物学的複製が極めて高いレベルの類似性を示していることを示すNMS分析を示す図である。

【図 20】サンプリング場所間の全体的群集構造の観察された違いをもっとも著しく明確にする約50ばかりの微生物OTUを同定するための収縮セントロイド(Shrunken Centroids)法と呼ばれる分析のヒートプロット概要である。

【図 21】気候の変化に応答する群集組成の変化の程度が異なる表示である。

【図 22】a)2005年8月に北緯度、中緯度および南緯度の地点から収集された新鮮な試料、ならびにb)同時期同地点からの新鮮な試料および移植対照試料(移植の1年後)のPhyloChip細菌OTUのNMS序列を示す2つのグラフである。両グラフに描かれている新鮮な試料は同一試料である。バーは3複製の1標準偏差を表している。

【図 23】移植された1年後に収集された相互に移植された試料と移植された対照のPhyloChip細菌OTUのNMS序列を示す4つのグラフである。矢印は、起源対照の組成から離れた移植された試料の組成の変化の軌道を示している。

【図 24】a)2007年9月に北緯度、中緯度および南緯度の地点から収集された新鮮な試料、ならびにb)同時期同地点からの新鮮な試料および移植対照試料(移植の3年後)のPhyloChip細菌OTUのNMS序列を示す2つのグラフである。両グラフに描かれている新鮮な試料は同一試料である。バーは3複製の1標準偏差を表している。

【図 25】移植された3年後に収集された相互に移植された試料と移植された対照のPhyloChip細菌OTUのNMS序列を示す4つのグラフである。矢印は、起源対照の組成から離れた移植された試料の組成の変化の軌道を示している。

【図 26】動物の種および類型(反芻動物および草食動物、鰭脚類、鳥類)による糞便試料中の検出された細菌分類群のクラスター分析を示す図表である。

【図 27】種の類型ごとのインジケータOTUの数を示す棒グラフである。

【図 28】インジケータ群集が発生源識別のために汚染水試料と比較されたことを示す序列グラフである。

【図 29】FIBと強い相関関係を有する下水分類群を示す棒グラフである。

【図 30】群集組成の比較を示したクラスター分析の結果を示す図表である。群集は、受水域における時間、発生源および受水域の類型に従ってクラスター化することが可能である。

【図 31】糞便微生物群集に対する受水域における時間の効果を示す棒グラフである。

【図 32】廃棄物微生物群集に対するクリーク対湾水の効果を示す棒グラフである。

【図 33】油プルーム(plume)による細菌分類群の濃縮を示す図である。

【発明を実施するための形態】

【0042】

本明細書で使用する場合、用語「オリゴヌクレオチド」は、通常一本鎖のポリヌクレオチドを指し、これは、合成ポリヌクレオチドまたは天然に存在するポリヌクレオチドである。オリゴヌクレオチドの長さは一般に、その特定の役割、例えば、プローブ、プライマーなどによって支配される。オリゴヌクレオチドを調製するための様々な技法、例えば、生物学的合成または化学合成を使用することができる。本発明の核酸は一般に、ホスホジエステル結合を含有することになるが、いくつかの場合では、以下に略述されるように、例えば、ホスホルアミド(Beaucageら、Tetrahedron、49(10):1925頁(1993)およびその中の参考文献; Letsinger、J. Org. Chem.、35:3800頁(1970); Sprinzlら、Eur. J. Biochem.、81:579頁(1977); Letsingerら、Nucl. Acids Res.、14:3487頁(1986); Sawaiら、Chem. Lett.、805頁(1984); Letsingerら、J. Am. Chem. Soc.、110:4470頁(1988); および Pauwelsら、Chemica Scripta、26:141頁(1986)); ホスホロチオエート(Magら、Nucleic Acids Res.、19:1437頁(1991); および米国特許第5,644,048号); ホスホロジチオエート(Briuら、J. Am. Chem. Soc.、111:2321頁(1989)); 0-メチルホスホロアミダイト(methylphosphoroamidite)連結(Eckstein、Oligonucleotides and Analogues: A Practical Approach、Oxford University Pressを参照); およびペプチド核酸主鎖および連結(そのすべてが参照により組み込まれている、Egholm、J. Am. Chem. Soc.、114:1895頁(1992); Meierら、Chem. Int

10

20

30

40

50

. Ed. Engl., 31:1008頁(1992); Nielsen、Nature、365:566頁(1993); Carlssonら、Nature、380:207頁(1996)を参照))を含む代替の主鎖を有することができる核酸類似体が含まれる。他の類似体核酸には、ポジティブ主鎖(Denpcyら、Proc Natl. Acad. Sci. USA、92:6097頁(1995));非イオン性主鎖(米国特許第5,386,023号;同第5,637,684号;同第5,602,240号;同第5,216,141号;および同第4,469,863号; Kiedrowskiら、Angew. Chem. Intl. Ed. English、30:423頁(1991); Letsingerら、J. Am. Chem. Soc.、110:4470頁(1988); Letsingerら、Nucleosides & Nucleotides、13:1597頁(1994);2章および3章、ASC Symposium Series 580、「Carbohydrate Modifications in Antisense Research」、Y. S. Sanghui およびP. Dan Cook編; Mesmaekerら、Bioorganic & Medicinal Chem. Lett.、4:395頁(1994); Jeffsら、J. Biomolecular NMR、34:17頁(1994); Tetrahedron Lett.、37:743頁(1996));ならびに米国特許第5,235,033号および同第5,034,506号、ならびに6章および7章、ASC Symposium Series 580、「Carbohydrate Modifications in Antisense Research」、Y. S. SanghuiおよびP. Dan Cook編に記載されたものを含む非リボース主鎖を有するものが含まれる。1つまたは複数の炭素環式糖を含有する核酸も、核酸の定義内に含まれる(Jenkinsら、Chem. Soc. Rev.、(1995)169~176頁を参照)。いくつかの核酸類似体は、Rawls、C & E News、1997年6月2日、35頁に記載されている。これらの参考文献のすべては、参照により明白に本明細書に組み込まれている。

【0043】

核酸は、DNA、RNA、またはハイブリッドであってもよく、デオキシリボヌクレオチドおよびリボヌクレオチドの任意の組合せ、ならびにウラシル、アデニン、チミン、シトシン、グアニン、イノシン、キサンチン、ヒポキサンチン、イソシトシン、イソグアニンと、ニトロピロールおよびニトロインドールなどの塩基類似体を含め塩基の任意の組合せを含有することができる。オリゴヌクレオチドは、市販の自動核酸シンセサイザーにおいて使用されるものなどの標準的な方法によって合成し、後にアレイ、ビーズ、または他の適当な表面に結合させることができる。あるいは、オリゴヌクレオチドは、フォトリソグラフィ技法または他の技法を使用して、アッセイ表面上に直接合成することができる。いくつかの実施形態では、リンカーが、アレイ表面またはビーズにオリゴヌクレオチドを結合させるのに使用される。

【0044】

本明細書で使用する場合、用語「核酸分子」または「ポリヌクレオチド」は、ポリマーのヌクレオチドまたは核酸ポリマーである化合物または組成物を指す。核酸分子は、天然化合物であっても、合成化合物であってもよい。核酸分子は、約2~5,000,000、またはそれ以上のヌクレオチドを有することができる。より大きい核酸分子は一般に、天然の状態で見出される。単離された状態では、核酸分子は、約10~50,000、またはそれ以上のヌクレオチド、通常、約100~20,000のヌクレオチドを有することができる。したがって、天然状態からの核酸分子の単離は、断片化をもたらすことが多いことが明白である。分子内構造との競合を低減するために、ハイブリダイゼーションの前に、より長い標的核酸分子、特にRNAを断片化することは有用となり得る。断片化は、化学的または酵素的に実現することができる。一般に、試料がDNAを含有する場合、デオキシリボヌクレアーゼ(DNase)などのヌクレアーゼが、ホスホジエステル連結を切断するのに使用される。核酸分子およびその断片として、それだけに限らないが、微生物、例えば、細菌、酵母、ファージ、染色体、ウイルス、ウイロイド、カビ、真菌、または他のより高等な生物、例えば、植物、魚、トリ、動物、ヒトなどのtRNA、mRNA、rRNA、ミトコンドリアDNAおよびRNA、葉緑体DNAおよびRNA、DNA/RNAハイブリッド、これらの生体物質または混合物を含めた精製または未精製形態のDNA(dsDNAおよびssDNA)およびRNA、遺伝子、染色体、プラスミド、コスミド、ゲノムが挙げられる。ポリヌクレオチドは、生体試料などの複雑な混合物のマイナーな画分のみとすることができる。

【0045】

本明細書で使用する場合、用語「ハイブリダイズする」は、ポリヌクレオチドの一本鎖が、成分塩基同士間の水素結合を通じて二本鎖構造を形成するプロセスを指す。2つのポ

10

20

30

40

50

リヌクレオチドの互いにハイブリダイズする能力は、2つのポリヌクレオチドの相補性の程度に基づき、これはひいては、マッチした相補的なヌクレオチド対の画分に基づく。別のポリヌクレオチドに相補的である所与のポリヌクレオチド中のヌクレオチドが多いほど、ハイブリダイゼーションのための条件がよりストリンジェントとなる場合があり、2つのポリヌクレオチドの間の結合がより特異的になる。ストリンジェンシーの増大は、温度を上昇させ、共溶媒の比を増大させ、塩濃度を低下させ、およびこれらの組合せによって実現することができる。

【 0 0 4 6 】

本明細書で使用する場合、用語「相補的な」、「補完する」、および「相補的核酸配列」は、ワトソン-クリック塩基対合則によって、別の核酸鎖中の塩基配列と関連づけられる核酸鎖を指す。一般に、2つのポリヌクレオチドは、一方のポリヌクレオチドが、逆平行センスで別のポリヌクレオチドに結合することができる場合、相補的であり、この場合、各ポリヌクレオチドの3'末端が他方のポリヌクレオチドの5'末端に結合し、次いで一方のポリヌクレオチドの各A、T(U)、G、およびCがそれぞれ、他方のポリヌクレオチドのT(U)、A、C、およびGと整列される。RNA塩基を含むポリヌクレオチドは、相補的なG/UまたはU/G塩基対も含むことができる。

10

【 0 0 4 7 】

本明細書で使用する場合、用語「クラスタリングツリー(clustering tree)」は、観察物、例えば、生物、遺伝子、およびポリヌクレオチドなどが、1つまたは複数のクラスターに分離される階層ツリー構造を指す。クラスタリングツリーのルートノードは、すべての観察物を含む1つのクラスターからなり、リーフノードは、個々の観察物に対応する。クラスタリングツリーは、観察物の様々な特性、例えば、遺伝子の配列および生物の形態学的特徴などに基づいて構築することができる。当技術分野で公知の多くの技法、例えば、階層的クラスタリング分析を、クラスタリングツリーを構築するのに使用することができる。クラスタリングツリーの非限定例は、系統発生的、分類学的、または進化的ツリーである。

20

【 0 0 4 8 】

本明細書で使用する場合、用語「操作的分類単位(operational taxon unit)」、「OTU」、「分類群」、「階層的クラスター」および「クラスター」は、互換的に使用される。操作的分類単位(OTU)は、クラスタリングツリーのノードを構成する1つまたは複数の生物の群を指す。クラスターのレベルは、その階層順序によって決定される。一実施形態では、OTUは、系統発生分析の目的のために、妥当な分類群であると暫定的に仮定された群である。別の実施形態では、OTUは、研究下で現存する分類学的単位のいずれかである。さらに別の実施形態では、OTUは、名称およびランクを与えられる。例えば、OTUは、ドメイン、サブドメイン、界、亜界、門、亜門、綱、亜綱、目、亜目、科、亜科、属、亜属、または種を表すことができる。いくつかの実施形態では、OTUは、階層順序の任意のレベルでの、真正細菌、原生生物、または真菌の界に由来する1つまたは複数の生物を表すことができる。いくつかの実施形態では、OTUは、原核生物目または真菌目を表す。

30

【 0 0 4 9 】

本明細書で使用する場合、用語「kmer」は、長さkのポリヌクレオチドを指す。いくつかの実施形態では、kは、1~1000の整数である。いくつかの実施形態では、kは、1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50、55、60、65、70、75、80、85、90、95、100、125、150、175、200、250、300、400、500、600、700、800、900、または1000である。

40

【 0 0 5 0 】

本明細書で使用する場合、用語「パーフェクトマッチプローブ」(PMプローブ)は、高度に保存された標的遺伝子またはポリヌクレオチドの少なくとも一部と100%相補的なkmerを指す。完全な相補性は通常、プローブの長さ全体にわたって存在する。しかし、完全プローブは、標的遺伝子またはポリヌクレオチドと相補性を欠いている先導配列または後続(t

50

railing) 配列に隣接する完全な相補性の1つまたは複数のセグメントを有することができる。

【0051】

本明細書で使用する場合、用語「ミスマッチプローブ」(MMプローブ)は、PMプローブの1、2、3、4、5、6、7、8、9、または10ヌクレオチドを除いて、すべての位置で対応するPMプローブと同一である対照プローブを指す。一般に、同一でない1つまたは複数の位置は、PMプローブの中央または中央付近に位置される。いくつかの実施形態では、ミスマッチプローブは、ユニバーサルミスマッチプローブ、例えば、陽性プローブと比較して、設定された数以下のヌクレオチド変異または置換を有するミスマッチプローブのコレクションである。例えば、ユニバーサルミスマッチプローブは、PMプローブのセット中の任意の1つのPMプローブと比較して、5つ以下のヌクレオチドがヌクレオチド配列において異なることができる。いくつかの実施形態では、MMプローブは、アレイ中で、各試験プローブ、例えば、細菌性の16S rRNA配列を標的にするPMプローブに隣接して使用される。

10

【0052】

本明細書で使用する場合、用語「プローブ対」は、PMプローブおよびその対応するMMプローブを指す。いくつかの実施形態では、PMプローブおよびMMプローブは、データ処理および統計的な分析の間に互いに関連してスコアをつけられる。本明細書で使用する場合、用語「OTUに関連したプローブ対」は、OTU特異的なPMプローブおよびその対応するMMプローブからなるプローブの対として定義される。

【0053】

20

本明細書で使用する場合、「試料」は、任意の源からのものであり、それだけに限らないが、気体試料、流体試料、固体試料、またはこれらの任意の混合物を含む。

【0054】

本明細書で使用する場合、「微生物」または「生物」は、それだけに限らないが、ウイルス、ウィロイド、細菌、古細菌、真菌、原生動物などが含まれる。

【0055】

用語「感度」は、そのようなものとして正確に識別される、実際の陽性のもの(actual positives)の割合の尺度である。

【0056】

用語「特異性」は、そのようなものとして正確に識別される、実際の陰性のもの(actual negatives)の割合の尺度である。

30

【0057】

用語「信頼度レベル」は、試験の結果が事実で繰り返し可能であり、ランダムでないという、百分率として表された尤度を指す。信頼度レベルは、推定の信頼性を示すのに使用され、様々な方法によって計算することができる。

【0058】

本発明は、大まかには汚染、より具体的には水中の汚染を検出するためのシステムおよび方法に関する。「汚染」は、本明細書で使用する場合、分析される組成物中の任意の望ましくない成分または物質(「混入物」)の存在を指す。いくつかの実施形態では、分析される組成物は水である。さらなる実施形態では、混入物は微生物である。汚染は、閾値レベルを超える、1つまたは複数の混入物の存在から生じ得る。

40

【0059】

一態様では、本発明は、OTUのバイオシグネチャーを利用する。本明細書で使用する場合、用語「バイオシグネチャー」は、特定の状態との、1つまたは複数のOTUの1つまたは複数のメンバーのレベルの関連を指す。一実施形態では、バイオシグネチャーは、1つのアッセイを使用した、試料中の少なくとも5、10、20、50、100、250、500、1000、5000、10,000、20,000、30,000、40,000、50,000、60,000、70,000、80,000、90,000、100,000、250,000、500,000、または1,000,000のOTUの存在、非存在、および/または量の判定を含む。いくつかの実施形態では、バイオシグネチャーは、少なくとも1、2、3、4、5、6、7、8、9、10、15、20、30、40、50、75、100、125、150、175、200、250、300、またはそ

50

れ以上のOTUの存在、またはレベルの変化を含む。

【0060】

一実施形態では、バイオシグネチャーは、1つの状態、例えば、単一源による汚染と関連する。別の実施形態では、バイオシグネチャーは、状態の組合せ、例えば、2つ以上の源による汚染、例えば、2、3、4、5、6、7、8、9、10、またはそれ以上などの源による汚染と関連する。バイオシグネチャーは、それだけに限らないが、新鮮な水、飲料水、海水、再生水、処理済み水、脱塩水、汚水、湖、河川、小川、海洋、地表水、地下水、流出する水、廃水、帯水層、他の天然または非天然の水域、および既知の混入物を含めて、任意の試料について得ることができる。バイオシグネチャーは、純粋な試料、既知の混入物、またはこれらの組合せについて求めることができる。いくつかの実施形態では、試験試料のバイオシグネチャーは、既知のバイオシグネチャーと比較され、シグネチャーが同じである尤度に関して判定が行われる。さらなる実施形態では、試料のバイオシグネチャーは、汚染源からのバイオシグネチャーと比較される。試験試料のバイオシグネチャーが比較されるバイオシグネチャーは、試験試料のバイオシグネチャーの前、後、または実質的に同じ時間に求めることができる。バイオシグネチャーは、特定の源からの1つまたは複数の試料の1つまたは複数の分析の結果であってもよい。シグネチャーを分析することができる汚染源の例として、それだけに限らないが、ヒトからの糞便物質；渡り鳥および非渡り鳥を含めたトリ源からの糞便物質；ヘラジカ、雌ウシ、シカ、ヒツジ、ウマ、ブタ、およびヤギを含めた家畜(cattle)および家畜(livestock)からの糞便物質；ならびにアシカ、アザラシ、およびカワウソを含めた水生動物からの糞便物質が挙げられる。本明細書で検出される水汚染は、腐敗物(例えば、植物もしくは動物の腐敗)、油流出、工業廃棄物または副生成物、およびOTUバイオシグネチャーを相関させることができる任意の他の混入物であってもよい。

【0061】

いくつかの実施形態では、試験試料のバイオシグネチャーは、2つ以上の独立したシグネチャー、2、3、4、5、6、7、8、9、10、またはそれ以上などの独立したシグネチャーの組合せである。好適な実施形態では、試料中に含まれる2つ以上のバイオシグネチャーのそれぞれは、同時にアッセイされる。さらなる実施形態では、バイオシグネチャーのサブセットを、10、25、50、100、250、500、1000、2000、または5000以下のOTUの存在、非存在、および/またはレベルの判定を含む、低密度検出システムを使用して評価することができる。

【0062】

一態様では、本発明は、試料中の複数の生体分子および生物を検出および同定するための方法、システム、および組成物を提供する。本発明は、個々の生物またはOTUを区別する能力を利用する。一態様では、個々の生物またはOTUは、生物特異的および/またはOTU特異的なプローブ、例えば、オリゴヌクレオチドプローブを使用して同定される。より具体的には、いくつかの実施形態は、試料中の生体分子および生物を検出および同定するのに有用な生物特異的および/またはOTU特異的なオリゴヌクレオチドプローブの選択に関する。いくつかの実施形態では、オリゴヌクレオチドプローブは、複数の生物中の標的オリゴヌクレオチドおよびその相同体内の領域へのオリゴヌクレオチドプローブのクロスハイブリダイゼーションパターンに基づいて選択される。相同体は、少なくとも80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、または99.5%同一であるヌクレオチド配列を有することができる。そのようなオリゴヌクレオチドは、全体またはその一部としての遺伝子、または遺伝子間配列とすることができる。オリゴヌクレオチドは、長さが10~10,000超の範囲となり得る。いくつかの他の実施形態では、OTU特異的なオリゴヌクレオチドプローブの、他の生物またはOTUに特異的なプローブとのクロスハイブリダイゼーションに少なくともある程度基づいて、試料中のOTUの存在を検出するための方法が提供される。いくつかの実施形態では、試料バイオシグネチャーが比較されるバイオシグネチャーは、1つまたは複数のプローブについての標的の存在に関して陽性の結果を含む。

【0063】

一態様では、本発明は、試料のバイオシグネチャーを判定または評価をするための診断システムを提供する。一実施形態では、診断システムは、少なくとも1、2、3、4、5、6、7、8、9、10、15、20、30、40、50、75、100、125、150、175、200、250、300、またはそれ以上のプローブを備える。別の実施形態では、診断システムは、最大1、2、3、4、5、6、7、8、9、10、15、20、30、40、50、75、100、125、150、175、200、250、300、またはそれ以上のプローブを備える。

【0064】

大容量システム

本発明の一態様では、存在する微生物、および総個体数のこれらの百分率組成の観点から試料の総微生物個体数を評価することによって、試料のバイオシグネチャーを判定するための大容量システムが提供される。このシステムは、1つのアッセイで、少なくとも10,000、20,000、30,000、40,000、50,000、60,000、またはそれ以上の異なるOTUの存在または量を判定することができる複数のプローブを備える。一般に、プローブは、高度に保存されたポリヌクレオチドと選択的にハイブリダイズする。通常、プローブは、同じ高度に保存されたポリヌクレオチドまたはその一部内とハイブリダイズする。一般に、高度に保存されたポリヌクレオチドまたはその断片は、遺伝子またはその断片を含む。例示的な高度に保存されたポリヌクレオチドは、16S rRNA遺伝子、23S rRNA遺伝子、5S rRNA遺伝子、5.8S rRNA遺伝子、12S rRNA遺伝子、18S rRNA遺伝子、28S rRNA遺伝子、gyrB遺伝子、rpoB遺伝子、fusA遺伝子、recA遺伝子、cox1遺伝子、およびnifD遺伝子に見出されるヌクレオチド配列を含む。他の実施形態では、2以上、3以上、4以上、5以上、6以上、7以上、8以上、9以上、10以上、15以上、20以上、25以上、50以上のプローブのコレクションが使用され、そのそれぞれは、異なる高度に保存されたポリヌクレオチドと特異的にハイブリダイズする。例えば、1つのコレクションのプローブは、16S rRNA遺伝子の同じ領域に結合する一方で、第2のコレクションのプローブは、23S rRNA遺伝子の同じ領域に結合する。各コレクションが、異なる、別個の高度に保存されたポリヌクレオチドを認識する2つ以上のコレクションのプローブを使用することにより、より多くのプローブを生成および試験することが可能になり、その使用は、種またはOTUの間のより優れた識別をもたらすことができる。

【0065】

高度に保存されたポリヌクレオチドは通常、それぞれ、ドメイン、界、門、綱、目、科、または属にわたって、少なくとも80%、85%、90%、92%、94%、95%、または97%の相同性を示す。これらのポリヌクレオチドの配列は、進化系統を判定し、または系統発生的判定を行うために使用することができ、系統発生的マーカーとしても知られる。いくつかの実施形態では、バイオシグネチャーは、系統発生的マーカーの組合せの存在、非存在、および/または存在量を構成する。本明細書に開示されるプローブによって検出されるOTUは、起源において、細菌、古細菌、真菌、または真核生物であってもよい。さらに、本明細書に開示される方法は、細菌、古細菌、真菌、または真核生物であるOTUを定量化するのに使用することができる。様々なプローブセットを組み合わせることによって、細菌、古細菌、真菌、真核生物、またはこれらの組合せを検出するためのシステムを設計することができる。1つのアッセイとして行われるそのようなユニバーサル微生物試験は、様々な試料、環境、状態、および混入物についてのバイオシグネチャーのキャラクタリゼーションを含めて、多数の環境の組成および生態を評価および理解するために、大きな利益をもたらすことができる。

【0066】

本発明の別の態様では、1つのアッセイで、1つのドメインの少なくとも10,000、20,000、30,000、40,000、50,000、または60,000の異なるOTUの存在および任意選択により量の確率を求めることができるシステムが提供される。そのようなシステムは、90%、91%、92%、93%、94%、95%、99%、または99.5%を超える信頼度レベルで確率判定を行う。いくつかの実施形態では、バイオシグネチャーは、各確率判定の組み合わせされた結果を含むことができる。

【0067】

いくつかの実施形態は、クラスタリングツリー中のノードに特異的なオリゴヌクレオチドプローブを選択する方法を提供する。いくつかの実施形態では、この方法は、複数の生物についての高度に保存された標的ポリヌクレオチドおよびその相同体を選択するステップと、複数の生物のポリヌクレオチドおよび相同体を、クラスタリングツリーにクラスター化するステップと、クラスタリングツリー上の各ノードに対する第1のポリヌクレオチドとハイブリダイズする候補オリゴヌクレオチドプローブのクロスハイブリダイゼーションパターンを求めるステップとを含む。この判定は、プローブがその標的相補配列の相同体とクロスハイブリダイズする尤度を求めるために実施される(例えば、in silico)。候補オリゴヌクレオチドプローブは、複数の生物のうちの1つにおける高度に保存された標的ポリヌクレオチド、高度に保存された標的の断片、またはその相同体の1つに相補的である場合がある。いくつかの実施形態では、クラスタリングツリー上の各ノードに対する候補オリゴヌクレオチドプローブの変異体のクロスハイブリダイゼーションパターンの判定であって、変異体は、候補オリゴヌクレオチドプローブに対応するが、少なくとも1つのヌクレオチドミスマッチを含む判定のため、ならびに候補オリゴヌクレオチドプローブのクロスハイブリダイゼーションパターンおよび変異体のクロスハイブリダイゼーションパターンに基づいて、候補オリゴヌクレオチドプローブを選択または却下するための方法が提供される。いくつかの実施形態では、ノードは操作的分類単位(OTU)である。いくつかの実施形態では、ノードは1種の生物である。

10

【0068】

20

いくつかの実施形態は、試料中の複数の生物を検出するのに使用するためのOTU特異的オリゴヌクレオチドプローブを選択する方法を提供する。いくつかの実施形態では、この方法は、複数の生物から高度に保存された標的ポリヌクレオチドおよびその相同体を選択するステップと、複数の生物からの標的遺伝子およびその相同体のポリヌクレオチドを、1つまたは複数の操作的分類単位(OTU)にクラスター化するステップであって、各OTUは、1つまたは複数の群の類似ヌクレオチド配列を含むステップと、候補OTU特異的オリゴヌクレオチドプローブのOTUとのクロスハイブリダイゼーションパターンを求めるステップであって、候補OTU特異的オリゴヌクレオチドプローブは、複数の生物の1つに由来する標的遺伝子またはその相同体の断片に対応するステップと、候補OTU特異的オリゴヌクレオチドプローブの変異体のOTUとのクロスハイブリダイゼーションパターンを求めるステップであって、変異体は、候補OTU特異的オリゴヌクレオチドプローブから少なくとも1ヌクレオチドミスマッチを含むステップと、候補OTU特異的オリゴヌクレオチドプローブのクロスハイブリダイゼーションパターンおよび変異体のクロスハイブリダイゼーションパターンに基づいて、候補OTU特異的オリゴヌクレオチドプローブを選択または却下するステップとを含む。いくつかの実施形態では、候補OTU特異的オリゴヌクレオチドプローブは、候補OTU特異的オリゴヌクレオチドプローブが、他のOTUからのプローブと相補的である任意のポリヌクレオチドとクロスハイブリダイズしない場合、選択される。さらなる実施形態では、候補OTU特異的オリゴヌクレオチドプローブは、候補OTU特異的オリゴヌクレオチドプローブが、1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、20、30、40、50、100、200、500、または1000以下の他のOTU群中のポリヌクレオチドとクロスハイブリダイズする場合、選択される。

30

40

【0069】

いくつかの実施形態は、試料中の複数の生物を検出するのに使用するのための一連の生物特異的ポリヌクレオチドプローブを選択する方法を提供する。いくつかの実施形態では、この方法は、複数の生物における高度に保存された標的ポリヌクレオチドおよびその相同体を同定するステップと、候補生物特異的オリゴヌクレオチドプローブの、複数の生物における高度に保存された標的ポリヌクレオチドおよびその相同体の配列とのクロスハイブリダイゼーションパターンを求めるステップであって、候補オリゴヌクレオチドプローブは、複数の生物の1つからの標的配列またはその相同体の断片に対応するステップと、候補生物特異的オリゴヌクレオチドプローブの変異体の、複数の生物における高度に保存さ

50

れた標的配列およびその相同体の配列とのクロスハイブリダイゼーションパターンを求めるステップであって、変異体は、候補生物特異的オリゴヌクレオチドプローブから少なくとも1ヌクレオチドミスマッチを含むステップと、候補生物特異的オリゴヌクレオチドプローブのクロスハイブリダイゼーションパターンおよび候補生物特異的オリゴヌクレオチドプローブの変異体のクロスハイブリダイゼーションパターンに基づいて、候補生物特異的オリゴヌクレオチドプローブを選択または却下するステップとを含む。

【0070】

いくつかの実施形態では、OTU特異的オリゴヌクレオチドプローブは、他のOTUからのプローブと相補的であるいずれのポリヌクレオチドともクロスハイブリダイズしない。他の実施形態では、OTU特異的オリゴヌクレオチドプローブは、1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、20、30、40、50、100、200、500、または1000以下の他のOTU群中のポリヌクレオチドとクロスハイブリダイズする。いくつかの実施形態は、試料中の複数の生物を検出するのに使用するために、一連の生物特異的オリゴヌクレオチドプローブを利用する。さらなる実施形態では、候補生物特異的オリゴヌクレオチドプローブは、候補生物特異的オリゴヌクレオチドプローブが、複数の生物中の1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、20、30、40、50以下のユニークな生物の標的核酸分子とのみハイブリダイズする場合選択される。他の実施形態では、プロセスは反復性であり、複数の候補特異的な-特異的オリゴヌクレオチドプローブが選択される。しばしば、選択された生物特異的オリゴヌクレオチドプローブは、クラスター化および整列されて類似の配列の群にされ、これらの群は、10TU当たり1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、35、40、50、または60以下の生物特異的オリゴヌクレオチドプローブのマッチに基づく高い信頼度で生物の検出を可能にする。一般に、生物特異的オリゴヌクレオチドプローブが検出する候補生物は、少なくとも1つの系統発生的、系統学的、進化的、または分類学的ツリーのリーフまたはノードに対応する。生物特異的オリゴヌクレオチドプローブによって検出される候補生物がツリー上に占める位置の知見は、そのドメイン、門、綱、亜綱、目、科、亜科、または属の他のメンバーに、生物の関係のある情報を提供する。

【0071】

いくつかの実施形態では、本明細書に開示される方法は、複数の生物を検出および区別するのに使用することができる、階層的なセットのオリゴヌクレオチドプローブである、一連の生物特異的オリゴヌクレオチドプローブを選択および/または利用する。いくつかの実施形態では、この方法は、1つの分析で、すべての既知の細菌または古細菌分類群の少なくとも80%、85%、90%、95%、99%、または100%についての包括的なスクリーニングを可能にする、生物特異的またはOTU特異的オリゴヌクレオチドプローブを選択および/または利用し、したがって異なる所望の分類群の検出を増強する。いくつかの実施形態では、すべての既知の細菌または古細菌分類群の素性は、オリゴヌクレオチド特異的プローブ、PCRクローニング、および配列決定方法を使用することによって以前に同定された分類群を含む。いくつかの実施形態は、混合された標的核酸分子を、その適切な操作的分類単位(OTU)の指定に正確に分類することができる一連のオリゴヌクレオチドプローブを選択および/または利用する方法を提供する。そのような方法は、包括的な原核生物または真核生物の同定、したがって、包括的なバイオシグネチャーキャラクタリゼーションをもたらすことができる。

【0072】

いくつかの実施形態では、選択されたOTU特異的オリゴヌクレオチドプローブは、分類学的同定の異なるレベルでの特定のOTUに属する1つまたは複数の生物の相対存在量を計算するのに使用される。いくつかの実施形態では、本明細書に開示される方法によって選択される少なくとも1つの生物特異的またはOTU特異的オリゴヌクレオチドプローブを含む微粒子のアレイまたはコレクションが、特定の微生物群集活性を推測するために提供される。例えば、嫌気性環境、例えば、沼地からの微生物コンソーシアム中の個々の分類群の素性は、その相対存在量に沿って求めることができる。コンソーシアムがブタノール発酵す

ることができる微生物を宿している疑いのある場合、嫌気性環境中に適当な供給原料を供給した後、ブタノールの産生が認められる場合、ブタノール発酵に関与する分類群は、豊富な量の16S rRNAを有する微生物によって推測することができる。本発明は、沼地または他の嫌気性環境から得られる試料から同定することができる、ブタノールの嫌気性発酵をすることができる直接標識された16S rRNAの検出に基づいて分類群の存在量を測定するための方法を提供する。

【0073】

いくつかの実施形態は、特定の生物またはOTUの同定の信頼度レベルおよび/または感度レベルを増大させるために、複数のプローブを選択する。複数のプローブの使用により、特定の生物に対するマッチの信頼度レベルを大いに増大させることができる。いくつかの実施形態では、選択された生物特異的オリゴヌクレオチドプローブは、生物の検出が、1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、35、またはそれ以上のオリゴヌクレオチドプローブのマッチに基づくように、クラスター化および整列されて類似の配列の群にされる。いくつかの実施形態では、オリゴヌクレオチドプローブは、1つの種に特異的である。他の実施形態では、オリゴヌクレオチドプローブは、同じ亜属、属、亜科、科、亜目、目、亜綱、綱、亜門、門、亜界、または界における生物などの関連生物を認識する。

【0074】

パーフェクトマッチ(PM)プローブは、標的ポリヌクレオチド、例えば、特定の生物を識別する配列と完全に相補的である。いくつかの実施形態では、本発明のシステムは、ミスマッチ(MM)対照プローブを備える。通常、MMプローブは、1つまたは複数のヌクレオチドが異なるが、その他ではPMプローブと同一である。1つまたは複数のミスマッチを有するプローブを使用することによって、標的配列との非特異的な結合および可能性のあるノンマッチを示すことができる。いくつかの実施形態では、MMプローブは、プローブの中央、例えば、25merプローブの13位に位置された1つのミスマッチを有する。MMプローブは、「プローブ対」として、その対応するPMプローブに関連してスコアをつけられる。MMプローブを使用することによって、バックグラウンドハイブリダイゼーションを推定し、それによって、多くの現在の検出システムにとって重要な問題である、非特異的なハイブリダイゼーションによる偽陽性結果の出現を低減することができる。Affymetrix高密度プローブアレイまたはIlluminaビーズアレイなどのアレイが使用される場合、MMプローブは、アレイ上のその対応するPMプローブに隣接して、またはPMプローブ付近に位置されることが理想的である。

【0075】

いくつかの実施形態は、相対的に高い信頼度レベルを伴って、複数の原核生物分類群を同時に同定することを可能にする、一連のオリゴヌクレオチドプローブを選択および/または利用する方法に関する。一般に、同定の信頼度レベルは、少なくとも90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、または99.5%である。OTUは、高度に保存された領域において、平均で少なくとも85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、または99.5%の配列相同性を共有する高度に関連した種の個々の種または群を指す。複数のMMプローブを利用することによって、定量化、および測定の信頼度を増強することができる。いくつかの実施形態では、複数のインターロゲーションプローブの各インターロゲーションプローブは、約1~約20の対応するミスマッチ対照プローブを有する。さらなる実施形態では、各インターロゲーションプローブは、約1~約10、約1~約5、約1~4、1~3、2、または1の対応するミスマッチプローブを有する。これらのインターロゲーションプローブは、標的核酸配列、例えば、16S rRNA遺伝子内のユニーク領域を標的にし、少なくとも約10、20、50、100、500、1,000、2,000、5,000、10,000、20,000、30,000、40,000、50,000、60,000、70,000、80,000、90,000、100,000、250,000、500,000、または1,000,000の分類群を同定するための手段を提供する。いくつかの実施形態では、高密度オリゴヌクレオチドプローブシステムを通じて、1つのアッセイで、複数の標的を同時にアッセイまたは検出することができる。特定の原核生物分類群を同定するために、すべて

の標的ハイブリダイゼーションの合計が使用される。結果は、培養できない、または未知の生物を同定する、より効率的で、あまり時間のかからない方法である。本発明はまた、以前は実現することができなかった結果を提供する、例えば、他の方法が数日を必要とする場合に、数時間で結果を提供することができる。いくつかの実施形態では、マイクロバイオーム(すなわち、試料)をアッセイすることによって、20、19、18、17、16、15、14、13、12、11、10、9、8、7、6、5、4、3、2、または1時間未満で、その成分の微生物の素性および存在量を求めることができる。

【0076】

いくつかの実施形態では、OTU特異的オリゴヌクレオチドプローブのセットは、各分類群について約1~約500のプローブを含む。いくつかの実施形態では、プローブは、抗体を含めたタンパク質、またはオリゴヌクレオチドもしくはその断片を含めた核酸分子である。いくつかの実施形態では、オリゴヌクレオチドプローブは、標的核酸分子のヌクレオチド断片に対応する。いくつかの実施形態では、約1~約500、約2~約200、約5~約150、約8~約100、約10~約35、または約12~約30のオリゴヌクレオチドプローブを、各分類学的分類について設計することができる。他の実施形態では、分類群は、少なくとも1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、40、またはそれ以上のプローブを有することができる。いくつかの実施形態では、様々な分類群は、異なる数のプローブを有することができる一方で、他の実施形態では、すべての分類群は、1群当たり固定された数のプローブを有する。1つの分類群における複数のプローブは、OTUが存在するかどうかに関して、「コールを行う」としても知られる、判定を行うために使用することができる、さらなるデータを提供することができる。複数のプローブは、不十分なシグナル強度、クロスハイブリダイゼーション、または他の異常に基づいて、分析から1つまたは複数のプローブを除去することも可能にする。プローブを除去することにより、結果の信頼度レベルを増大させ、豊富でない微生物の検出をさらに可能にすることができる。オリゴヌクレオチドプローブはそれぞれ、約5~約100ヌクレオチド、約10~約50ヌクレオチド、約15~約35ヌクレオチド、または約20~約30ヌクレオチドとすることができる。いくつかの実施形態では、プローブは、少なくとも5mer、6mer、7mer、8mer、9mer、10mer、11mer、12mer、13mer、14mer、15mer、16mer、17mer、18mer、19mer、20mer、21mer、22mer、23mer、24mer、25mer、26mer、27mer、28mer、29mer、30mer、31mer、32mer、33mer、34mer、35mer、36mer、37mer、38mer、39mer、40mer、41mer、42mer、43mer、44mer、45mer、46mer、47mer、48mer、49mer、50mer、51mer、52mer、53mer、54mer、55mer、56mer、57mer、58mer、59mer、60mer、61mer、62mer、63mer、64mer、65mer、66mer、67mer、68mer、69mer、70mer、71mer、72mer、73mer、74mer、75mer、76mer、77mer、78mer、79mer、80mer、81mer、82mer、83mer、84mer、85mer、86mer、87mer、88mer、89mer、90mer、91mer、92mer、93mer、94mer、95mer、96mer、97mer、98mer、99mer、100mer、またはこれらの組合せである。

【0077】

いくつかの実施形態は、検出の信頼度を増大させるために、複数の、確認の、生物特異的またはOTU特異的プローブを選択する方法を提供する。いくつかの実施形態では、この方法は、非標的領域によるクロスハイブリダイゼーションの効果を最小限にするために、各パーフェクトマッチ(PM)プローブについて1つまたは複数のミスマッチ(MM)プローブも選択する。本明細書に開示される方法によって選択される生物特異的およびOTU特異的オリゴヌクレオチドプローブは、環境試料中に存在する数千の分類群を同時に同定し、対象とする群集における微生物およびその系統発生的関係の正確な同定を可能にすることができる。本明細書に開示される方法によって選択される生物特異的およびOTU特異的オリゴヌクレオチドプローブを使用するシステム、および本明細書に開示されるコンピュータ分析は、rRNA遺伝子配列決定技法に対して多数の利点を有する。そのような利点として、1回のマイクロバイオーム分析当たりの費用が低減すること、物理的分析およびコンピュータ分析の両観点から、1つの試料またはマイクロバイオーム当たりの処理速度が増大する

こと、分析手順は、キメラによって悪影響されず、人工のファイロタイプを作り出さず、バーコードPCRのバイアスによって変わらないことが挙げられる。さらに、定量的基準を、本発明のマイクロバイーム試料とともに実施することができ、ピロシーケンスにとって可能でないものである。

【0078】

いくつかの実施形態は、試料中の複数の生物を同時に検出するための分析システムまたはビーズ多重システムにおいて使用するために、一連のOTUまたは生物特異的オリゴヌクレオチドプローブを選択および/または利用するための方法を提供する。この方法は、微生物群集組成を求め、パイオシグネチャーを確立するために、標的核酸分子内の既知の多様性を標的にする。標的核酸分子は一般に、高度に保存されたポリヌクレオチドである。いくつかの実施形態では、高度に保存されたポリヌクレオチドは、高度に保存された遺伝子に由来し、一方、他の実施形態では、ポリヌクレオチドは、中等度の、または大きな配列変異を有する遺伝子の高度に保存した領域に由来する。さらなる実施形態では、高度に保存された領域は、イントロン、エクソン、または2つの遺伝子を分離する核酸の連結セクションである場合がある。いくつかの実施形態では、高度に保存されたポリヌクレオチドは、「系統発生的」遺伝子に由来する。系統発生的遺伝子として、それだけに限らないが、5.8S rRNA遺伝子、12S rRNA遺伝子、16S rRNA遺伝子-原核生物、16S rRNA遺伝子-ミトコンドリア、18S rRNA遺伝子、23S rRNA遺伝子、28S rRNA遺伝子、gyrB遺伝子、rpoB遺伝子、fusA遺伝子、recA遺伝子、cox1遺伝子、およびnifD遺伝子が挙げられる。真核生物では、rRNA遺伝子は、核、ミトコンドリア、または両方のものであってもよい。いくつかの実施形態では、他のrRNA遺伝子を使用して、または使用せずに、密接に関連した分類群を区別するために、16S-23S rRNA遺伝子の内部転写スペーサー(ITS)を使用することができる。例えば、rRNA、例えば、16Sまたは23S rRNAは、タンパク質に翻訳されるその遺伝子コードを有するのではなく、機能的分子としてタンパク質アセンブリー機構内に直接作用する。16S rRNAの構造的制約のために、遺伝子全体にわたる特定の領域は、高度に保存されたポリヌクレオチド配列を有するが、非構造的セグメントは、高度の変異性を有する場合がある。高い変異性の領域をプローブするステップを使用することによって、1つの種レベルを表すOTUを同定することができ、一方、変異性のより少ない領域を使用することによって、亜属、属、亜科、科、亜目、目、亜綱、綱、亜門、門、亜界、または界を表すOTUを同定することができる。本明細書に開示される方法を使用することによって、特定の生物、特定の生物を表すOTU、または生物の特定の分類群を表すOTUを同定するための、高レベルの特異性を提供する生物特異的およびOTU特異的オリゴヌクレオチドプローブを選択することができる。本明細書に開示されるシステムおよび方法は、バックグラウンドまたは密接に関連した生物のプールから密接に関連した微生物およびOTUを同定するのに特に有用である。

【0079】

本発明の方法によって選択および/または利用されるプローブは、OTU内に組織化することができ、これは、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、または99%を超える感度および/または特異性を伴ったアッセイを提供する。いくつかの実施形態では、感度および特異性は、ハイブリダイゼーションシグナル強度、OTU中のプローブの数、潜在的なクロスハイブリダイゼーション反応の数、存在する場合、ミスマッチプローブのシグナル強度、バックグラウンドノイズ、またはこれらの組合せに依存する。いくつかの実施形態では、1つのプローブを含むOTUが、少なくとも90%の感度および特異性を伴ったアッセイをもたらすことができる一方で、別のOTUは、少なくとも90%の感度および特異性を伴ったアッセイをもたらすのに少なくとも20のプローブを必要とする場合がある。

【0080】

いくつかの実施形態は、試料中の複数の生体分子ならびに生物を検出ならびに同定するのに使用するための、系統発生分析システムの設計ならびにシグナル処理および解釈のための方法に関する。より具体的には、いくつかの実施形態は、高い信頼度レベルを伴って、試料中の複数の生物を検出するのに使用するために、一連の生物特異的オリゴヌクレオ

チドプローブを選択する方法に関する。いくつかの実施形態は、高い信頼度レベルを伴って、試料中の複数の生物を検出するのに使用するために、一連のOTU特異的オリゴヌクレオチドプローブを選択する方法に関する。

【0081】

関連微生物種間での識別用のプローブによって標的にされる任意の15-bp ~ 30bpの領域にわたって、わずか1 ~ 数個のヌクレオチドの配列変異性を有することができる、16S rRNAのような高度に保存されたポリヌクレオチドの場合では、アッセイシステムにおいてプローブ-標的配列特性を最大にすることが有利である。本発明のいくつかの実施形態は、クロスハイブリダイゼーションの影響を有効に最小限にする生物特異的オリゴヌクレオチドプローブを選択する方法を提供する。一実施形態では、この方法は、(a)複数の生物に対応する標的核酸分子の配列を同定するステップと、(b)候補生物特異的オリゴヌクレオチドプローブの、複数の生物に由来する標的核酸分子とのクロスハイブリダイゼーションパターンを求めるステップであって、候補オリゴヌクレオチドプローブは、複数の生物に由来する標的核酸分子の配列断片に対応するステップと、(c)候補生物特異的オリゴヌクレオチドプローブの変異体の、複数の生物に由来する標的核酸分子とのクロスハイブリダイゼーションパターンを求めるステップであって、候補生物特異的オリゴヌクレオチドプローブの変異体は、候補生物特異的オリゴヌクレオチドプローブと比較して少なくとも1つのヌクレオチドミスマッチを含むステップと、(d)候補生物特異的オリゴヌクレオチドプローブのクロスハイブリダイゼーションパターンおよび候補生物特異的オリゴヌクレオチドプローブの変異体のクロスハイブリダイゼーションパターンに基づいて、候補生物特異的オリゴヌクレオチドプローブを選択または却下するステップとを含む。いくつかの実施形態では、試料中の複数の生物を検出するのに使用するための、一連のOTU特異的オリゴヌクレオチドプローブを選択する方法が提供される。いくつかの実施形態では、この方法は、(a)複数の生物に対応する標的核酸分子の配列を同定するステップと、(b)複数の生物に由来する標的核酸分子の配列を、1つまたは複数の操作的分類単位(OTU)にクラスター化するステップであって、各OTUは、1つまたは複数の群の類似配列を含むステップと、(c)候補OTU特異的オリゴヌクレオチドプローブのOTUとのクロスハイブリダイゼーションパターンを求めるステップであって、候補OTU特異的オリゴヌクレオチドプローブは、複数の生物の1つに由来する標的核酸分子の配列断片に対応するステップと、(d)候補OTU物特異的オリゴヌクレオチドプローブの変異体のOTUとのクロスハイブリダイゼーションパターンを求めるステップであって、候補OTU特異的オリゴヌクレオチドプローブの変異体は、候補OTU特異的オリゴヌクレオチドプローブと比較して少なくとも1つのヌクレオチドミスマッチを含むステップと、(e)候補OTU特異的オリゴヌクレオチドプローブのOTUとのクロスハイブリダイゼーションパターンおよび候補OTU特異的オリゴヌクレオチドプローブの変異体のOTUとのクロスハイブリダイゼーションパターンに基づいて、候補OTU特異的オリゴヌクレオチドプローブを選択または却下するステップとを含む。いくつかの実施形態では、候補OTU特異的オリゴヌクレオチドプローブは、候補OTU特異的オリゴヌクレオチドプローブまたはその変異体が他の標的配列とクロスハイブリダイズすると予測される場合、却下される。いくつかの実施形態では、所定量の予測されるクロスハイブリダイゼーションは許容される。

【0082】

いくつかの実施形態では、選択されたオリゴヌクレオチドプローブは、当技術分野で公知の任意の妥当な方法によって合成される。適当な方法のいくつかの例には、ファインポイントピンを用いたガラススライド上へのプリンティング、予め作製されたマスクを使用したフォトリソグラフィー、ダイナミックマイクロミラーデバイスを使用したフォトリソグラフィー、インクジェットプリンティング、または電気化学が含まれる。一例では、フォトリソグラフィー法を使用することによって、表面上に選ばれたオリゴヌクレオチドプローブを直接合成することができる。表面の適当な例には、ガラス、プラスチック、シリコン、および当技術分野で利用可能な任意の他の表面が含まれる。ある特定の例では、オリゴヌクレオチドプローブは、 $1\mu\text{m}^2$ 当たり約1,000プローブ ~ $1\mu\text{m}^2$ 当たり約100,000

10

20

30

40

50

プローブ、好ましくは、 $1\mu\text{m}^2$ 当たり約2000プローブ~ $1\mu\text{m}^2$ 当たり約50,000プローブ、より好ましくは、 $1\mu\text{m}^2$ 当たり約5000プローブ~ $1\mu\text{m}^2$ 当たり約20,000プローブのおおよその密度でガラス表面上に合成することができる。一例では、プローブの密度は、 $1\mu\text{m}^2$ 当たり約10,000プローブである。アレイ上のプローブの数は、かなり大きく、例えば、1アレイ当たり少なくとも 10^5 、 10^6 、 10^7 、 10^8 、または 10^9 プローブとすることができる。通常、大きいアレイについては、所与の長さのプローブの総数の比較的小さい割合(すなわち、約1%未満、0.1%、0.01%、0.001%、0.00001%、0.000001%、または0.0000001%)のみが、個々のOTUを標的にする。しばしば、下限のアレイは、10、25、50、100、500、1,000、5,000、または10,000、25,000、50,000、100,000、または250,000以下のプローブを有する。

10

【0083】

一般に、アレイまたは微粒子は、1つまたは複数の高度に保存されたポリヌクレオチドに対するプローブを有する。アレイまたは微粒子は、機能的に発現された遺伝子にハイブリダイズするさらなるプローブ(例えば、確認プローブ)を有することができ、それによって、分類群の同定の根拠を置くための代替または確認のシグナルをもたらす。例えば、アレイは、ペスト菌(*Yersinia pestis*)およびコレラ菌(*Vibrio cholerae*)に由来する16S rRNA遺伝子配列に対するプローブ、およびまた、ペスト菌caf1毒性遺伝子またはコレラ菌閉鎖毒素(zot)遺伝子に対する確認プローブを含むことができる。確認プローブに基づくハイブリダイゼーションシグナルの検出とカップリングした、特定のOTUに関連する16S rRNAポリヌクレオチドに結合しているプローブに基づくハイブリダイゼーションシグナル

20

【0084】

様々な長さのプローブを、アレイまたは微粒子上に使用することができる。上述したように、プローブは、1つの相補的セグメントからもっぱらなっているとしてもよく、または隣接、後続および/または介在性セグメントと並べられた1つまたは複数の相補的セグメントを有することができる。後者の状況において、相補的セグメントの全長は、プローブの長さより重要となり得る。機能的な条件において、MMプローブと比較して、PMプローブが標的ポリヌクレオチド、例えば、16S rRNAにより強くハイブリダイズすることを可能にするために、PMプローブの相補的セグメントは、十分に長いべきである。PMプローブは通常、完全な相補性を示す、少なくとも15ヌクレオチド、より通常には少なくとも16、17、18、19、20、21、22、23、24、25、または30塩基の長さを有する1つの相補的セグメントを有する。

30

【0085】

いくつかのアレイ、または微粒子のロットでは、すべてのプローブは同じ長さである。他のアレイ、または微粒子のロットでは、プローブ長は、定量化標準(quantification standard)(QS)プローブ、陰性対照(NC)プローブ、プローブ対、プローブセット(OTU)、およびこれらの組合せの間で変化する。例えば、いくつかのアレイは、すべて25merであるプローブ対を含むOTUまたはプローブセットの他の群と一緒に、すべて23merであるプローブ対を含むOTUの群を有することができる。他の長さのプローブ対のさらなる群を加えることができる。したがって、いくつかのアレイは、15mer、16mer、17mer、18mer、19mer、20mer、21mer、22mer、23mer、24mer、25mer、26mer、27mer、28mer、29mer、30mer、31mer、32mer、33mer、34mer、35mer、36mer、37mer、38mer、39mer、40mer、またはこれらの組合せのサイズを有するプローブ対を含むことができる。他のアレイは、同じ群、OTU、またはプローブセット内で異なるサイズのプローブを有することができる。これらのアレイにおいて、所与のOTUまたはプローブセット内のプローブは、互いに独立して長さが変化してもよい。異なる長さのプローブを有することを使用することによって、反応のpH、

40

50

温度、およびイオン状態でのオリゴヌクレオチドプローブのハイブリダイゼーション安定性に応じて、プローブからのハイブリダイゼーションシグナルを同等にすることができる。

【0086】

本発明の別の態様では、1つのアッセイで、複数の異なるOTUの存在または量を判定するためのシステムであって、複数のポリヌクレオチドインターロゲーションプローブ、複数のポリヌクレオチド陽性対照プローブ、および複数のポリヌクレオチド陰性対照プローブを備えるシステムが提供される。いくつかの実施形態では、システムは、1つのアッセイを使用して、試料中の少なくとも5、10、20、50、100、250、500、1000、5000、10,000、20,000、30,000、40,000、50,000、60,000、70,000、80,000、90,000、100,000、250,000、500,000、または1,000,000のOTUの存在、非存在、相対存在量、および/または量を検出することができる。いくつかの実施形態では、ポリヌクレオチド陽性対照プローブは、1)断片化の前に、規定量で標的核酸配列中にスパイクされた原核生物または真核生物の代謝関連遺伝子の配列を標的にするプローブ、または2)断片化および標識化の後に、ハイブリダイゼーションミックス中に加えられた予め標識されたオリゴヌクレオチドと相補的なプローブを含む。断片化の前に加えられた対照は、システムの断片化、ピオチン化、ハイブリダイゼーション、染色、および走査効率を一括して試験する。これはまた、2つ以上のアレイが1つの実験において使用される場合、または2つの別個の実験からのデータが組み合わせられる場合など、1つの実験または組み合わせられた実験において使用される複数の分析成分にわたって、全体的な蛍光強度が正規化されることを可能にする。第2の対照は、システムのハイブリダイゼーション、染色、および走査を直接アッセイする。両タイプの対照を、1つの実験で使用することができる。

【0087】

いくつかの実施形態では、QS基準(陽性対照)は、PMプローブである。他の実施形態では、QS基準は、PMおよびMMプローブの対である。さらなる実施形態では、QS基準は、PMおよびMMプローブの対と、対応するMMプローブを伴わないPMプローブとの組み合わせを含む。別の実施形態では、QS基準は、各対応するPMプローブについて、少なくとも1、2、3、4、5、6、7、8、9、10、またはそれ以上のMMプローブを含む。さらなる実施形態では、QS基準は、各対応するMMプローブについて、少なくとも1、2、3、4、5、6、7、8、9、10、またはそれ以上のPMプローブを含む。システムは、各1、10、100、または1000の異なるインターロゲーションプローブについて少なくとも1つの陽性対照プローブを備えることができる。

【0088】

いくつかの場合では、陽性対照プローブに相補的なスパイクインされたオリゴヌクレオチドは、G+C含量、ウラシル含量、濃度、またはこれらの組合せの点で異なる。いくつかの実施形態では、G+C%は、約30%~約70%、約35%~約65%、または約40%~約60%の範囲である。QS基準は、ウラシル取り込み頻度に基づいて選ぶこともできる。QS基準は、100中約1~100中約60、100中約4~100中約50、または100中約10~100中約50の範囲でウラシルを取り込むことができる。いくつかの場合では、これらの加えられるオリゴヌクレオチドの濃度は、1、2、3、4、5、6、または7桁にわたる範囲となる。約 $10^5 \sim 10^{14}$ 、 $10^6 \sim 10^{13}$ 、 $10^7 \sim 10^{12}$ 、 $10^7 \sim 10^{11}$ 、 $10^8 \sim 10^{11}$ 、および $10^8 \sim 10^{10}$ の濃度範囲を使用することができ、一般に、この範囲にわたる線形ハイブリダイゼーションシグナル応答を特徴とする。いくつかの実施形態では、本明細書に開示される方法を行うための陽性対照プローブは、表1に示された陽性対照配列と相補的なポリヌクレオチドを含む。陽性対照の標的として使用することができる他の遺伝子には、構造タンパク質、成長、細胞周期または生殖の調節を制御するタンパク質をコードする遺伝子、およびハウスキーピング遺伝子が含まれる。さらに、高度に保存された遺伝子または他の高度に保存されたポリヌクレオチドに基づく合成遺伝子を、試料に加えることができる。合成遺伝子を設計することができる有用な高度に保存された遺伝子には、16S rRNA遺伝子、18S rRNA遺伝子、23SrRNA遺伝子が含まれる。例示的な対照プローブは、配列番号51~100として与えられている。

【 0 0 8 9 】

【表 1】

表 1 陽性対照配列

陽性対照 ID	説明
AFFX-BioB-5_at	大腸菌(E.coli)ビオチン合成酵素
AFFX-BioB-M_at	大腸菌ビオチン合成酵素
AFFX-BioC-5_at	大腸菌 bioC タンパク質
AFFX-BioC-3_at	大腸菌 bioC タンパク質
AFFX-BioDn-3_at	大腸菌デチオビオチン合成酵素
AFFX-CreX-5_at	バクテリオファージ P1 cre リコンビナーゼタンパク質
AFFX-DapX-5_at	B.スプチリス(B. subtilis)dapB、ジヒドロジピコリン酸レダクターゼ
AFFX-DapX-M_at	B.スプチリス dapB、ジヒドロジピコリン酸レダクターゼ
YFL039C	サッカロミセス(Saccharomyces)、アクチン(Act 1p)タンパク質についての遺伝子
YER022W	サッカロミセス、RNA ポリメラーゼ II メディエーター複合体サブユニット(SRB4p)
YER148W	サッカロミセス、TATA-結合タンパク質、基本転写因子(SPT15)
YEL002C	サッカロミセス、オリゴサッカリルトランスフェラーゼ(OST)糖タンパク質複合体(WBP1)のペータサブユニット
YEL024W	サッカロミセス、ユビキノール-チトクロム-c レダクターゼ(RIP1)
合成 16S rRNA 対照	
SYNM neurolyt_st	マイコプラズマ・ニューロリティクム (Mycoplasma neurolyticum)16S rRNA 遺伝子の合成誘導体
SYNLc.oenos_st	リューコノストック・エノス(Leuconostoc oenos)16S rRNA 遺伝子の合成誘導体
SYNCau.cres8_st	カウロバクター・クレセニウス(Caulobacter crescenius)16S rRNA 遺伝子の合成誘導体
SYNFer.nodosm_st	フェルビドバクテリウム・ノドサム(Fervidobacterium nodosum)16S rRNA 遺伝子の合成誘導体
SYNSap.grandi_st	サプロスピラ・グランディス(Saprospira grandis)16S rRNA 遺伝子の合成誘導体

【 0 0 9 0 】

いくつかの実施形態では、陰性対照は、PMおよびMMプローブの対を含む。さらなる実施形態では、陰性対照は、PMおよびMMプローブの対と、対応するMMプローブを伴わないPMプローブとの組合せを含む。他の実施形態では、陰性対照プローブは、各対応する陰性対照PMプローブについて、少なくとも1、2、3、4、5、6、7、8、9、10、またはそれ以上のMMプローブを含む。システムは、各1、10、100、または1000の異なるインターロゲーションプローブ(PM)について少なくとも1つの陰性対照プローブを備えることができる。

【 0 0 9 1 】

一般に、陰性対照プローブは、16S rRNA遺伝子または他の高度に保存された遺伝子標的と、仮にあったとしても、弱くハイブリダイズする。陰性対照プローブは、原核生物または真核生物起源の代謝関連遺伝子と相補的となり得る。一般に、陰性対照プローブでは、標的物質は、試料中にまったくスパイクされない。いくつかの実施形態では、陰性対照プローブは、陽性対照にも使用されるプローブの同じコレクションからのものであるが、陽性対照プローブ方法と対照的に、陰性対照プローブと相補的な物質は、試料中にまったくスパイクされない。本質的に、対照プローブは、ユニバーサル対照プローブであり、システムの設計に応じて陽性または陰性対照プローブの役割を果たす。当業者は、ユニバーサル対照プローブは、高度に保存された配列分析システムに限定されず、本明細書に開示される本実施形態以外に用途を有することを理解するであろう。

【0092】

さらなる実施形態では、種特異的同定、または高度に保存されたポリヌクレオチドに対するプローブを用いて実現された結果の確認をもたらすために、高度に保存されていないポリヌクレオチドに対するプローブがシステムに加えられる。通常、これらの「確認」プローブは、パーフェクトマッチプローブによって認識される高度に保存されたポリヌクレオチドと、仮にあったとしても、非常に弱くクロスハイブリダイズする。有用な種特異的遺伝子には、代謝関連遺伝子、構造タンパク質、成長、細胞周期、もしくは生殖の調節を制御するタンパク質をコードする遺伝子、ハウスキーピング遺伝子、または毒性、毒素、もしくは他の病原因子をコードする遺伝子が含まれる。いくつかの実施形態では、システムは、少なくとも1、5、10、20、30、40、50、60、70、80、90、100、150、200、250、300、400、500、600、700、800、900、1000、5,000、または10,000の種特異的プローブを備える。

10

【0093】

いくつかの実施形態では、本発明のシステムは、アレイを備える。アレイの非限定例には、マイクロアレイ、ビーズアレイ、貫通孔アレイ、ウェルアレイ、および標的に対するハイブリダイジングプローブにおいて使用するのに適した、当技術分野で公知の他のアレイが含まれる。アレイは、任意の適切な構成、例えば、行と列のグリッドなどに配置することができる。アレイの一部の範囲は、OTU検出プローブを含み、一方、他の範囲は、画像方向、標準化対照、シグナルスケールリング、雑音低減処理、または他の分析のために使用することができる。対照プローブは、アレイの周囲に沿って、アレイにわたって対角線状に、セクションを交互にして、ランダムに、を含めて、アレイ内の任意の位置に配置することができる。いくつかの実施形態では、アレイ上の対照プローブは、PMおよびMMプローブのプローブ対を含む。対照プローブの数は変更することができるが、一般に、アレイ上の対照プローブの数は、1～約500,000の範囲である。いくつかの実施形態では、少なくとも10、100、500、1,000、5,000、10,000、25,000、50,000、100,000、250,000、または500,000の対照プローブが存在する。対照プローブ対が使用される場合、プローブ対は、1～約250,000対の範囲となる。いくつかの実施形態では、少なくとも5、50、250、500、2,500、5,000、12,500、25,000、50,000、125,000、または250,000の対照プローブ対が存在する。アレイは、プローブに加えて他の成分、例えば、プローブを支持体に結合させるリンカーなどを有することができる。いくつかの実施形態では、アレイを作製するための材料は、Affymetrix (Santa Clara, California)、GE Healthcare (Little Chalfont, Buckinghamshire, United Kingdom)、またはAgilent Technologies (Palo Alto, California.)から得ることができる。

20

30

【0094】

プローブがアレイの基質に結合されるアレイに加えて、多数の他の技術を、本発明の方法を実施するために開示されたシステムにおいて使用することができる。一実施形態では、プローブは、Ngら (Ngら A spatially addressable bead-based biosensor for simple and rapid DNA detection. Biosensors & Bioelectronics, 23:803～810頁、2008)に開示されているように、ビーズに結合され、次いでアレイ上に配置される。

【0095】

40

別の実施形態では、Luminex多重アッセイシステムによって例示されるように、プローブがビーズまたは微小球に結合され、ハイブリダイゼーション反応が溶液中で実施され、次いで、ビーズがフローサイトメトリーによって分析される。この分析システムでは、米国特許第6,524,793号に開示されているように、それぞれが、複数の同一のプローブでタグを付けられ、または標識されたビーズを有する、均質なビーズのサブセットが組み合わされることによって、プールされたビーズセットが作成され、これは、試料とハイブリダイズされ、次いでフローサイトメトリーを用いてリアルタイムで分析される。ビーズのサブセットは、例えば、レーザー励起可能な色素含量の変量を使用して、タグまたは標識の変化によって互いに区別することができる。

【0096】

50

さらなる実施形態では、プローブは、Illumina Veracode多重アッセイシステムによって例示されるように、円柱状ガラスマイクロビーズに結合される。ここでは、同一のデジタルホログラフィックエレメントを埋め込まれたマイクロビーズのサブセットが使用されることによって、プローブ標識されたマイクロビーズのユニークなサブセットが作り出される。ハイブリダイゼーションの後、マイクロビーズがレーザー光によって励起され、マイクロビーズのコードおよびプローブ標識が、リアルタイム多重アッセイで読まれる。

【0097】

別の実施形態では、NanoString nCounter Analysis System (Geiss Gら Direct multiplexed measurement of gene expression with color-coded probe pairs. Nature Biotech. 26:317~325頁、2008)によって例示されるように、溶液ベースのアッセイシステムが使用される。この方法では、試料は、レポータープローブの溶液と混合され、このレポータープローブは、ユニークな配列を認識し、試料中の核酸と、データ収集のために固体表面上に固定されるレポータープローブとの間で複合体を形成させるプローブを捕捉する。各レポータープローブはカラーコードされ、蛍光を通じて検出される。

【0098】

さらなる実施形態では、Panomics QuantiGene Plex 2.0アッセイシステムによって例示されるような分岐DNA技術が使用される。分岐DNA技術は、RNA検出および定量化のためのサンドイッチ核酸ハイブリダイゼーションアッセイを含み、これは、配列ではなく、レポーターシグナルを増幅する。試料源においてRNAを測定することによって、このアッセイは、標的ポリヌクレオチドの抽出および増幅に固有の変化または誤差を回避する。QuantiGene Plex技術は、上述したLuminexシステムなどの多重ビーズベースアッセイシステムと組み合わせることによって、全細胞または精製RNA調製物から直接、複数のRNA標的を同時に定量化することを可能にすることができる。

【0099】

プローブおよびその選択

複数の微生物を同時検出するのに使用するための標的プローブを設計するための例示的プロセス300が、図3に例示されている。簡単に言えば、配列が、状態301で、データベースから抽出される。一般に、データベースは、系統発生的配列、または他の高度に保存された配列もしくは相同配列を含む。配列は、状態302でキメラについて分析され、これは、さらなる考察から除去される。キメラ配列は、2つ以上の無関係の配列の結合、一般に、異なる遺伝子から生じる。任意選択により、配列は、状態303で、ヘアピンループ形成の傾向などの構造異常についてさらに分析することができ、同定された配列は、さらなる考察から引き続いて除去される。次に、状態304で、複数の配列アライメントが、データセット中の残りの配列に対して実施される。次いで、整列された配列は、状態305で、PCRプライマー配列などの実験室の人工産物について調査され、同定された配列は、さらなる考察から除去される。残りの配列は、状態306でクラスター化され、クラスター化された配列のセクションと完全な相補性を有するパーフェクトマッチ(PM)プローブが、状態307で選択される。任意選択により、配列包括度ヒューリスティクスが状態308で実施され、その後、対応するPMプローブについて、状態309で、ミスマッチ(MM)プローブを選択することによって、プローブ対が作られる。最後に、複数のプローブ対を含むプローブセットによって表されるOTUを状態310でアセンブルすることによって、階層的分類を構築する。

【0100】

一般に、プローブ選択のために使用される配列を抽出するためのデータベースは、対象とする特定の保存された遺伝子または高度に相同な配列、データベース中の配列の総数、データベースに列挙された配列内の配列全体の長さまたは高度に保存された領域の長さ、およびデータベース中の配列の質に基づいて選ばれる。一般に、等しい配列数であるが異なる配列長の2つのデータベースの間では、高度に保存された配列のより長い標的領域を有するデータベースが一般に、比較することができるより多い総数の可能な配列を含むことになる。いくつかの実施形態では、配列は、少なくとも300、400、500、600、700、800

、900、1,000、1,200、1,400、1,600、1,800、2,000、4,000、8,000、16,000、または24,000ヌクレオチドの長さである。一般に、より大きい数の総配列を有するデータベースが、比較するためのより多い材料を提供する。さらなる実施形態では、データベースは、少なくとも10,000、20,000、30,000、40,000、50,000、60,000、70,000、80,000、100,000、200,000、500,000、1,000,000、または2,000,000の配列項目を含む。プローブ構成のために特に対象とする遺伝子は、16S rDNA (16S rRNA遺伝子)である。他の保存された遺伝子として、18S rDNA、23S rDNA、gyrA、gyrB遺伝子、groEL、rpoB遺伝子、fusA遺伝子、recA遺伝子、sodA、cox1遺伝子、およびnifD遺伝子が挙げられる。さらなる実施形態では、2つの遺伝子の高度に保存されたセグメント間のスペーサー領域を使用することができる。例えば、16Sおよび23S rDNA遺伝子の間のスペーサー領域を、16Sおよび23S rDNAの保存されたセクションとともに使用することができる。

10

【0101】

いくつかの実施形態では、バイオシグネチャーの検出は、1つまたは複数のOTU内の既知の、または発見された標的とハイブリダイズするように設計されたプローブの使用を含む。いくつかの実施形態では、標的は、データベース中などの既知の標的のコレクションから選択される。本発明のいくつかの実施形態では、プローブを選択するのに使用されるデータベースは、対象とする生物、例えば、対象とする細菌、古細菌、真菌、真核生物、微生物、または原核生物の既知の配列の、少なくとも50%、60%、70%、80%、90%、95%、96%、97%、98%、99%、または最大100%を含む。データベース中の各個々の生物についての配列は、生物のゲノム、またはその非重複領域の20%超、30%、40%、50%、60%、70%、80%、90%、または95%超を含むことができる。いくつかの実施形態では、データベースは、配列決定されたものが中に含まれている生物のゲノム、またはその非重複配列の最大100%を含む。長さが1250ヌクレオチドより長い、ほぼ40,000の整列された16S rDNA配列のリストは、Lawrence Berkeley National Laboratoryによって実施されている公的にアクセス可能なデータベースである、Greengenesウェブアプリケーション上で見出すことができる。他の公的にアクセス可能なデータベースには、GenBank、Michigan State Universityのリボソームデータベースプロジェクト、Max Planck Institute for Marine MicrobiologyのSilvaデータベース、および国立保健研究所のNCBIが含まれる。2つ以上の私的なデータベースおよび/または公共のデータベースの内容を融合することによって作り出された独自の配列データベースまたは組合せも、本発明の方法を実施するのに使用することができる。いくつかの実施形態では、試料は、1つまたは複数の選ばれたデータベース中のすべての標的について同時にアッセイされる。他の実施形態では、試料は、1つまたは複数のデータベース中で同定される標的のサブセットについて同時にアッセイされる。いくつかの実施形態では、バイオシグネチャーは、1つまたは複数の選ばれたデータベース中の一部またはすべての標的について試料をアッセイする結果を含む。他の実施形態では、バイオシグネチャーは、1つまたは複数の選ばれたデータベース中の一部またはすべての標的について試料をアッセイする結果のサブセットを含む。

20

30

【0102】

状態302でキメラを検出及び除去するための、データベースから選択された配列の分析は一般に、重複断片を生成し、これらの断片を互いに比較することによって実施される。断片は、これらが少なくとも60%、70%、80%、90%、95%、または99%の配列同一性を有する場合、保持することができる。選択される配列の配列多様性が低い場合があるために、上記プロセスは潜在的にキメラを見逃すことが認識された。多様なキメラを含まない配列のコアセットに対して断片を比較することによって、より多くのキメラを同定し、配列セットから除去することができる。1つまたは複数の配列が、不明瞭なキメラ、例えば、キメラの親を含むキメラとして同定される場合では、そのキメラは除去され、親キメラは断片化され、第2の比較サイクルが実施される。データセットからの配列は、greengenes.lbl.govのGreengenesウェブサイトから入手可能なBellerophon3などの独自のソフトウェアプログラムを使用して、キメラについてスクリーニングすることができる。

40

【0103】

50

次いで、保持された非キメラ配列のデータセットは、既知の配列のコアセットに対して保持された配列を整列させることによって、状態303で構造異常についてスクリーニングすることができる。コアセットと比較した場合に、そのアライメントにおいて、少なくとも25、30、35、40、45、50、60、70、もしくは80のギャップを有するか、またはコアセットと比較した場合に50、60、70、80、90、100、110、120、130、140、150、160、170、180、190、200、250、300、もしくは400を超える塩基対の挿入を有する、保持されたデータセット中の配列は、配列異常を有するとしてタグを付けられ、データセットから除去される。

【0104】

次いで、スクリーニングされた配列は、既知の、キメラを含まないコアセットと比較するために、状態304で整列されて、多重配列アライメント(MSA)にされる。集中的なアライメント計算を実施するための1つのアライメントツールは、NAST(Nearest Alignment Space Termination)ウェブツール(DeSantisら、Nucleic Acids Res. (2006) 34:W394~399)である。MSAをコンパイルするために、任意の適切なアライメントツール、例えば、clustalw (Thompsonら、Nucleic Acids Res (1994) 22:4673~4680頁)およびMUSCLE (Edgar、Nucleic Acids Res. (2004) 32:1792~1797頁)を使用することができる。

【0105】

整列された配列は、状態305で、PCRプライマー配列を宿している配列について探索され、任意のそのように同定された配列は、データセットから除去される。

【0106】

次いで整列された配列は、状態306でクラスター化することによって、いわゆる「ガイドツリー」を作り出すことができる。最初に、配列は、kmerのリストに変換される。kmerのリストの対での比較が実施され、閾値類似性が見出される場合のみ、共通のkmerの百分率が疎行列中に記録される。疎行列は、例えば、完全連結を使用してクラスター化される。クラスタリングには、凝集型「ボトムアップ」または分岐型「トップダウン」階層的クラスタリング、距離「分割」クラスタリングおよびアライメントクラスタリングが含まれる。各クラスターから、最大の情報内容を有する配列が、代表として選ばれる。通常、ゲノム配列決定プロジェクトから得られた配列は、クラスター作成において優先権が与えられ、その理由は、これらがキメラである、または他の配列異常を有する可能性がより低いためである。以前のサイクルからの代表のみを使用して、周期的プロセスが繰り返される。各新しいサイクルについて、疎行列中に記録するための閾値が低減される。最終段階で、ルートノードがマルチ分岐ツリー(multifurcated tree)中で最終的な代表配列に連結される。各サイクル中に見出される代表配列は、得られるガイドツリー中でノードを表す。すべてのノードは、セルフリファレンシャル表(self-referential table)介して、そのクラスタリング結果に基づいて連結され、ガイドツリー中の任意の階層点への迅速なアクセスを可能にする。いくつかの実施形態では、結果は、データベース形式、例えば、構造化クエリー言語(SQL)準拠形式で記憶される。得られるガイドツリーにおいて、各リーフノードは、個々の生物を表し、ガイドツリーの最低レベルの上の各ノードは、候補OTUを表す。

【0107】

約 2×10^5 の配列から構築される一般的な距離行列は、400億の交点を必要とする場合があり、これは、ディスクにコード化される場合、約40ギガバイトのデータ空間を必要とする。40億に配列の量を2倍にすると、4倍のファイルサイズ(約160GB)を必要とする。疎行列を使用する、ここで例示されるクラスタリング方法は、大きいファイルの必要および計算時間の予期される増加を回避する。したがって、この方法は、従来の配列クラスタリング方法より効率的に実施することができる。さらに、配列アライメント(例えば、DNAアライメント)から作り出される距離行列では、1つのミスアライメントが、多くの距離値に影響し得る。対照的に、本明細書に例示されるクラスタリング方法は、kmerのアライメントに基づき、したがって、クラスタリング値に対するミスアライメントの効果は、有意に低減される。

【0108】

ガイドツリー構成の後に、今度は「選別された配列データセット」と呼ばれる、残りの配列のデータセットが、候補プローブ、例えば、PMプローブを選択するのに使用される。最初に、未対応の配列多型が同定され、上記で生成されたガイドツリーを使用するプレクラスタリングプロセスを使用して、選別された配列データセットから除去されることによって、最小類似性を越え、最大サイズ以下のクラスターが作り出される。一般に、クラスター化された配列は、少なくとも80%、85%、90%、95%、97%、または99%類似である。通常、クラスターは、1,000、500、200、100、80、60、50、40、30、20、または10以下の配列を有する。このプロセスにより、配列データ異常値を、近隣内の比較によって検出し、選別された配列データセットから除去することが可能になる。

10

【0109】

次に、残りの配列は、候補標的プローブを生成するために、所望のサイズに断片化される。一般に、断片は、約10mer~100mer、15mer~約50mer、約20mer~約40mer、約20mer~約30merの範囲である。通常、断片は、サイズが少なくとも15mer、20mer、25mer、30mer、40mer、50mer、または100merである。各候補標的プローブは、少なくとも1つのプレクラスタの閾値画分内に見出されることを必要とする。一般に、少なくとも80%、90%、または95%の閾値画分が使用される。

【0110】

次いで、少なくとも1つのプレクラスタの閾値画分内であるすべての候補PMプローブは、様々な生物物理学的パラメータ、例えば、融解温度(61~80)、G+C含量(35~70%)、-4kcal/molを超えるヘアピンエネルギー、自己二量体化の可能性(>35)などについて評価される。生物物理学的パラメータの設定境界に入らない候補PMプローブは、データセットから排除される。任意選択により、プローブは、フォトリソグラフィ合成を容易にするために、さらに選別することができる。

20

【0111】

各PM候補プローブの、各非標的インプット16s rRNA遺伝子配列とのクロスハイブリダイゼーションの尤度が求められる。各PM候補プローブについてのクロスハイブリダイゼーションパターンが記録される。

【0112】

次いで、状態308で実施される配列包括度ヒューリスティックスが、許容できる生物物理学的パラメータを有する候補PMプローブに適用される。

30

【0113】

各候補PMプローブについて、対応するMMプローブを、状態309で生成することができる。各MMプローブは、少なくとも1つのヌクレオチドによってその対応するPMプローブと異なる。いくつかの実施形態では、MMプローブは、1、2、3、4、5、6、7、8、9、または10のヌクレオチドによってその対応するPMプローブと異なる。MMプローブ内で、ミスマッチした1つまたは複数のヌクレオチドは、PMプローブ中の同じ1つまたは複数の位置において見出されない3つの中心塩基のいずれかを含むことができる。例えば、13位、すなわち、中心ヌクレオチドでグアニンを有する25merのPMプローブでは、MMプローブは、13位でアデニン、チミン、ウラシル、またはシトシンを有するプローブを含む。同様に、3'方向から読まれる場合に、12番目のヌクレオチド位置でアデニンおよび13番目のヌクレオチド位置でグアニンを有する25merのPMプローブでは、可能なMMプローブは、12番目のヌクレオチドでグアニンおよび13番目のヌクレオチド位置でアデニン、チミン、またはシトシンを有するプローブ;12番目のヌクレオチド位置でシトシンおよび13番目のヌクレオチド位置でアデニン、チミン、またはシトシンを有するプローブ;ならびに12番目のヌクレオチド位置でチミンおよび13番目のヌクレオチド位置でアデニン、チミン、またはシトシンを有するプローブを含む。いくつかの実施形態では、ミスマッチした1つまたは複数のヌクレオチドは、対応するPMプローブ中のヌクレオチドの任意の1つまたは複数を含む。MMプローブおよび/または代表されるミスマッチ位置の数を増やすステップを使用することによって、定量化、精度、および信頼度を増強することができる。

40

50

【 0 1 1 4 】

PMプローブについて上述したように、各候補MMプローブは、1つまたは複数の生物物理学的パラメータ、例えば、融解温度、G+C含量、ヘアピンエネルギー、自己二量体、およびフォトリソグラフィー合成ステップなどの設定された境界を満たすことが必要とされる。一般に、これらのパラメータは、PMプローブの生物物理学的パラメータと同一または実質的に類似である。

【 0 1 1 5 】

生物物理学的パラメータ、および任意選択により、上記フォトリソグラフィーパラメータを満たす候補MMプローブは、次いで、標的配列とのクロスハイブリダイゼーションの尤度についてスクリーニングされる。通常、中心のkmer長が評価される。25merの候補MMについて、候補MMからの中心のkmer、一般に15mer、16mer、17mer、18mer、または19merが標的配列に対して比較される。標的配列と同一である中心kmerを含む候補MMプローブが排除される。次に、適当な候補MMプローブを同定することができない候補PMプローブも排除される。

【 0 1 1 6 】

各候補OTUを評価することによって、OTU以外の配列とハイブリダイズすることができないPMプローブの数を求めることができる。

【 0 1 1 7 】

一実施形態では、プレパーティション(pre-partition)プロセスが実施される。プレパーティションは、最大パーティションサイズを超えない最大の可能なクレード(ノード_id)である。図6を参照。一般に、有用なパーティションサイズは、約1,000~約8,000ノードの範囲である。所定のサイズ範囲内にある任意のプレパーティションは、フルパーティションになる。最小パーティションサイズ未満のプレパーティションは、可能な場合、姉妹ノードをアセンブルすることによってパーティションに組み合わされる。例えば、パーティションが1000~2000メンバーのサイズの範囲にされることを仮定されたい。ノードAが1500遺伝子を表し、その親であるノードBが2500遺伝子を表す場合、ノードAは、プレパーティションとみなされる。ノードCがノードAの兄弟であり、ノードCが50遺伝子のみを表す場合、ノードCもプレパーティションであり、その理由は、ノードCをその親であるノードBへと移動することは、最大パーティションサイズの2000メンバーを超えて包含するためである。

【 0 1 1 8 】

候補配列クラスターを作り出すために、過渡的な配列クラスターが、対でのユニークな候補標的の計数、または対での共通の候補標的の計数に基づいて2つの距離行列の変化する閾値を使用して同定される。候補配列クラスター中の配列の大きい画分で、例えば、クラスター中の配列の90%以上に行き渡っているプローブが、PMを含む配列の計数、および所与のPMの遺伝子座についてあいまいでないデータを有する配列の計数を使用して同定される。各行き渡っているプローブについて、クラスター外でのクロスハイブリダイゼーションの可能性も試験される。クラスター-PMセットに関するすべての情報が記録される。クロスハイブリダイジングプローブのみが同定されるクラスターとして定義される効果のないクラスターは、データセットから除去される。

【 0 1 1 9 】

必要な場合、ある程度のクロスハイブリダイゼーションを示すことが予期されるプローブを選択することができる。潜在的にハイブリダイゼーションしがちなプローブが抑制されることによって、クラスター外の配列が、クラスター特異的PMプローブの多くとハイブリダイズすることができる確率が低減される。分布アルゴリズムを使用することによって、プローブ-配列相互接続(エッジ)のグラフを検査し、重複エッジを最小限にするプローブのセットを支持することができる。

【 0 1 2 0 】

すべてのパーティションからの解決を完了した後、パーティションにわたって設定された解の包括的な調和が実施される。配列クラスターは、OTUとしてロックされ、各クラス

10

20

30

40

50

ターのPMプローブセットは、他の残りのPMプローブセットに対する包括的なクロスハイブリダイゼーションについて試験される。プローブは、包括的なクロスハイブリダイゼーションパターンに基づく有用性についてランク付けされる。

【0121】

OTUは、アセンブルおよびアノテートされる。一般に、各OTUは、ドメイン、界、門、亜門、綱、亜綱、目、および科からの各ランクの1つの用語を使用して、分類学的にアノテートされる。結果として、分類学的な学名を伴わずに提示され、「環境試料」または「未分類」としてアノテートされたすべての16S rRNA配列は、分類学的なアノテーションに割り当てられる。

【0122】

NCBIによって認識された各属レベルの名称は、判読および記録される。各系列の分類学的用語について、重複した隣接する用語は除去される。ドメインレベルの用語は、直接的なパターンマッチによって見出され、門レベルの用語は、ドメインのすぐ下のランクとして見出される。目レベルの用語は、-alesの接尾辞によって見出され、科レベルの用語は、-eaeの接尾辞によって見出される。科レベルの用語は利用可能でないが、属が同定される場合(例えば、容認されたリストへのマッチによって)、属レベルの用語が、科レベルの用語を得るために使用される。認識された用語の間で見出されるすべての未認識の用語は、利用可能なランクに適合される(新しいランクは、追加の用語について作り出されない)。空のランクは、下位の用語からルート用語を引き出し、所定の接尾辞を加えることによって満たされる。最後に、OTUの科は、配列の科の割り当てから投票によって決定される。同点は、優先配列によって解消される(例えば、ゲノム配列決定プロジェクトに由来する配列は、最高の優先権が与えられる)。亜科内のすべてのOTUは、配列の中でのkmer距離によって比較され、OTUは、閾値類似性が観察される場合はいつでも亜科にリンクされる。各候補OTUは、候補OTUの配列全体にわたって行き渡り、OTU外の配列とハイブリダイズすることが予期されない標的の総数を求めるために評価される。

【0123】

開示されるアルゴリズムを使用して生成される例示的な25merのPMおよびMMプローブは、配列番号1~50として示されている。上記プロセスは、長さが少なくとも15ヌクレオチド~少なくとも200ヌクレオチドのサイズ範囲のプローブを選択するのに適用可能であり、連結配列を含めた共通または無関係の配列によって、片側または両側に隣接されたプローブを含むことが留意されるべきである。さらに、このプロセスによって選択されたプローブは、さらに処理することによって、元の選択されたプローブより小さい、または大きいプローブを得ることができる。例えば、配列番号1~50として列挙されたプローブは、3'末端、5'末端、または両方から配列を除去することによりさらに処理することによって、25merの配列の少なくとも一部と同一である、より小さい配列を生成することができる。他の実施形態では、より大きいプローブを、開示されるアルゴリズムによって同定されるプローブの配列を組み込むことによって生成することができ、すなわち、25merのプローブを、30mer以上、35mer以上、40mer以上、45mer以上、50mer以上、55mer以上、60mer以上、65mer以上、70mer以上、75mer以上、80mer以上、85mer以上、または90mer以上のプローブに組み込むことができる。さらに、配列番号1~50として列挙されたプローブは、一端で短くし、他端で長くすることによって、10mer~200merの範囲のプローブを得ることができる。

【0124】

上記プロセスによって選択されたプローブはまた、1つもしくは複数の塩基置換、例えば、チミンの代わりにウラシルを含むプローブ、ニトロピロールおよびニトロインドールなどの1つもしくは複数の塩基類似体を組み込んでいるプローブ、1つもしくは複数の糖置換、例えば、デオキシリボースの代わりにリボースを含むプローブ、またはこれらの任意の組合せを含む。同様に、本発明の方法によって選択されるプローブは、例えば、ホスホルアミドからなる代替のバックボーン化学的性質をさらに含むことができる。

【0125】

本発明の方法によって生成される推定上のプローブのコレクションのサイズは、16S rRNA遺伝子などのより小さい高度に保存された配列より多数の相同配列を可能にする23S rRNA遺伝子のもののように、より長い配列を有する特定の高度に保存された配列の長さにある程度依存する。いくつかの実施形態では、高度に保存された配列の長さは、少なくとも100bp、250bp、500bp、1,000bp、2,000bp、4,000bp、8,000bp、10,000bp、または20,000bpである。さらに、本発明の方法によって生成される推定上のプローブのコレクションのサイズは、配列がプローブを分析および生成するために選択される1つまたは複数のデータベース中の相同配列のコレクションのサイズにも依存する。相同配列のコレクションがより大きいほど、分析することができる配列のより大きいプールをもたらすことによって、より多くの推定上のプローブの生成を可能にする。いくつかの実施形態では、1つまたは複数のデータベース中の相同配列の開始コレクションは、少なくとも100,000、250,000、500,000、1,000,000、2,000,000、5,000,000、または10,000,000の配列を含む。推定上のプローブのコレクションのサイズは、所望のプローブの長さにさらに依存し、その理由は、ユニークな配列に結合するプローブの数が増加するにつれてプローブ長が減少するためである。特定の高度に保存された配列、データベースのサイズ、および所望のプローブの長さに応じて、少なくとも100、1,000、10,000、25,000、50,000、100,000、250,000、500,000、1,000,000、2,000,000、5,000,000、または10,000,000のプローブの推定上のプローブのコレクションを生成することができる。

【0126】

上記方法によって生成される推定上のプローブから、検出システムを構築することができる。検出システムは、任意の数のプローブを有することができ、この方法によって選択される1つのプローブ～すべてのプローブの範囲となり得る。いくつかの実施形態では、検出システムは、少なくとも1、2、3、4、5、6、7、8、9、10、15、20、25、30、35、36、40、45、50、55、60、65、70、80、90、100、125、150、200、300、400、500、1000、2,000、5,000、10,000、20,000、40,000、50,000、100,000、200,000、500,000、1,000,000、または2,000,000のプローブを含む。多数のプローブを有するシステムを使用することによって、試料、例えば、環境試料もしくは臨床試料中の妥当な微生物を同定し、かつ/またはバイオシグネチャーを生成することができる。別の実施形態では、妥当な微生物が分かった後、低度(例えば、1～10,000)～中程度の(例えば、10,000～100,000)数のプローブを有する検出システムを、1つまたは複数の特定のバイオシグネチャーを求めることなどの特殊目的用途のために設計することができる。いくつかの実施形態では、妥当な微生物の素性の知見を使用することによって、これらの微生物に対するさらなるプローブを選択することができる。例えば、第1のセットのプローブ中の5つの25merのプローブが、妥当な微生物とハイブリダイズする場合、これらの5つのプローブの変異体を生成し、その結合特性および生物物理学的特性について試験することができる(例えば、in silico)。あるいは、妥当な微生物の同定は、微生物を同定するのに最初に使用されたプローブと異なる新しいプローブの生成を導くことができる。例えば、新規の微生物が同定された後、特定の用途のために抗体を生成することができる。

【0127】

OTU特異的プローブ、例えば、階層ノード内に含まれる生物に特異的なオリゴヌクレオチドプローブを選択するために、1つを超える子ノードを有する各階層ノードについて、さらなるPMプローブを選ぶことができる。ある特定のノードへの選択についての標的を適格にするために、PMセットにマッチするノード内の配列の閾値細分が実施される。閾値細分の例には、0.2%、0.5%、1%、2%、3%、4%、5%、6%、7%、8%、9%、および10%が含まれる。直接のサブノード(子)をカバーすることも実施される。例えば、各標的は、少なくとも1つのサブノードの少なくとも25%を表すべきである。

【0128】

本明細書に開示される方法によって選択されるプローブの特異性は、いくつかの方法において実験的にバリデートすることができる。例えば、標的配列の存在下でのプローブのハイブリダイゼーションシグナルは、測定し、バックグラウンドシグナルに対して比較す

ることができる。標的配列は、1つまたは複数の純粋な培養物、または標的配列を含むことが知られている環境試料もしくは臨床試料から得ることができる。特定の分類群は、アレイ上のプローブの大部分(約70%~約100%、約80%~約100%、または約90%~約100%)が、バックグラウンドのハイブリダイゼーションシグナルより少なくとも約50倍、100倍、150倍、200倍、250倍、300倍、350倍、400倍、450倍、500倍、1,000倍大きいハイブリダイゼーションシグナルを有する場合、試料中に存在すると同定される。また、プローブのハイブリダイゼーションシグナルは、そのミスマッチプローブの1つまたは複数のハイブリダイゼーションシグナルと比較することができる。少なくとも1.05、1.10、1.15、1.20、1.25、1.30、1.40、1.45、または1.50のPM:MM比は、PMプローブが、その標的配列と選択的にハイブリダイズすることができることを示すことができる。プローブがその標的に選択的にハイブリダイズする能力を試験するためのさらなる方法は、以下にさらに説明される、ペアディファレンス(pair difference)スコア(d)を計算することである。1.0を超えるペアディファレンススコアは、そのミスマッチプローブの1つと比較して標的と選択的にハイブリダイズすることができることを示す。

【0129】

本明細書に開示される方法を使用することによって、タンパク質、DNA、RNA、DNA、またはRNA単位複製配列、および標的核酸分子に由来する天然rRNAなどの生体分子に対する生物特異的および/またはOTU特異的オリゴヌクレオチドプローブを選択および/または利用することができる。いくつかの実施形態では、プローブは、PCR増幅からのバイアスをまったく伴わずに、試料中の活発に代謝する生物を同定するために、試料に由来するrRNAをアレイ上に配置することができるように、天然のrRNAに対してアンチセンスであるように設計される。活発に代謝する生物は、無活動の、または死んだ生物と比較して、タンパク質の産生に使用される著しく高い数のリボソームを有する。したがって、いくつかの実施形態では、特定の時点でタンパク質を作製する1つまたは複数の生物の能力を測定することができる。このようにして、本実施形態のアレイシステムは、多様な群集内の代謝する生物を直接同定するために使用することができる。

【0130】

試料の調製

いくつかの実施形態では、使用される試料は、任意の源からの環境試料、例えば、天然に存在する、もしくは人工の雰囲気、水の系および源、土壌、または対象とする他の試料とすることができる。いくつかの実施形態では、環境試料は、例えば、屋内もしくは屋外空気または大気粒子収集システム;屋内の表面、および装置、デバイスまたは機器の表面から得ることができる。いくつかの実施形態では、生態系が試料採取される。生態系は、陸生であってもよく、それだけに限らないが、土壌、表面および表面の上の環境を含めたすべての既知の陸生環境を含む。生態系は、食糧農業機関の土地被覆分類システム(LCCS)および米国森林局によって開発されたForest-Range Environmental Study Ecosystems(FRES)において分類されるものを含む。例示的な生態系として、熱帯雨林、温帯雨林、温帯硬葉樹林、北方林、タイガおよび山地の針葉樹林など;サバンナおよびステップを含めた草原;砂漠;沼地(marsh)、沼地(swamp)、沼地(bog)、河口、および泥沼を含めた湿地帯;河岸生態系、高山およびツンドラ生態系が挙げられる。生態系には、水の環境、例えば、湖、小川、泉、サンゴ礁、海浜、河口、海山、海溝、および潮間帯などに関連するものがさらに含まれる。生態系は、土壌、腐植土、鉱質土壌、および帯水層も含む。生態系は、地下環境、例えば、鉱山、油田、洞窟、断層帯および破碎帯、地熱ゾーン、ならびに帯水層などをさらに包含する。生態系には、植物、動物、およびヒトに関連するマイクロバイオーームがさらに含まれる。例示的な植物に関連するマイクロバイオーームには、根、樹皮、幹、葉、および花の中または付近に見出されるものが含まれる。動物およびヒトに関連するマイクロバイオーームには、胃腸管、呼吸器系、鼻孔、尿生殖路、乳腺、口腔、耳道、糞便、尿、および皮膚中に見出されるものが含まれる。

【0131】

他の実施形態では、試料は、任意の種類の臨床的または医学的試料とすることができる

。例えば、哺乳動物の血液、尿、糞便、鼻孔、肺、または腸に由来する試料を、アレイシステムを使用してアッセイすることができる。また、本明細書に開示される方法および本実施形態のアレイシステムによって選択されるプローブは、動物の血液中の感染を同定するのに使用することができる。本明細書に開示される方法および本実施形態のアレイシステムによって選択されるプローブはまた、体の外側に直接または間接的に暴露される医学的試料、例えば、動物の肺、耳、鼻部、咽頭、消化器系または皮膚の全体などをアッセイするのに使用することができる。病院は現在、これらの範囲に存在する複雑な微生物群集を同定するための資源を欠いている。

【0132】

環境試料または臨床試料などの試料中の複数の生物から遺伝子配列を得るための技法およびシステムは、当業者によって周知である。例えば、Zhouら (Appl. Environ. Microbiol. (1996) 62:316~322頁) には、ロバストな核酸抽出および精製が提供されている。このプロトコールは、実験の目的および環境試料のタイプ、例えば、土壌、沈澱物、および地下水などに応じて改変することもできる。多くの市販のDNA抽出および精製キットも使用することができる。2pg未満の精製DNAを有する試料は、増幅を必要とする場合があり、これは、全群集ゲノム増幅(whole community genome amplification)(WCGA)方法(Wuら、Appl. Environ. Microbiol. (2006) 72、4931~4941頁)など、当技術分野で公知の従来の技法を使用して実施することができる。いくつかの実施形態では、高度に保存された配列、例えば、16S rRNA遺伝子、23S rRNA遺伝子、5S rRNA遺伝子、5.8S rRNA遺伝子、12S rRNA遺伝子、18S rRNA遺伝子、28S rRNA遺伝子、gyrB遺伝子、rpoB遺伝子、fusA遺伝子、recA遺伝子、cox1遺伝子、およびnifD遺伝子中に見出されるものなどが増幅される。通常、増幅は、PCRを使用して実施されるが、他のタイプの核酸増幅も使用することができる。一般に、増幅は、高度に保存された配列に特異的なユニバーサルプライマーの1つの対を使用して実施される。冗長度または総単位複製配列濃度の増加量について、異なる高度に保存された配列に特異的な2つ以上のユニバーサルプローブ対を使用することができる。代表的なPCRプライマーには、細菌プライマー27Fおよび1492Rが含まれる。

【0133】

環境試料から精製RNAを得るための技法およびシステムも、当業者に周知である。例えば、Hurtら (Appl. Environ. Microbiol. (2001) 67:4495~4503頁) によって記載された手法を使用することができる。この方法は、同じ試料内で同時にDNAおよびRNAを単離することができる。ゲル電気泳動方法も、群衆RNA(community RNA)を単離するのに使用することができる (McGrathら、J. Microbiol. Methods(2008)75:172~176頁)。5pg未満の精製RNAを有する試料は、増幅を必要とする場合があり、これは、cDNAを得るために、全群衆RNA増幅手法(WCRA)(Gaoら、Appl. Environ. Microbiol. (2007) 73:563~571頁)などの当技術分野で公知の従来の技法を使用して実施することができる。いくつかの実施形態では、環境試料採取およびDNA抽出は、先に述べたように行われる (DeSantisら、Microbial Ecology, 53(3):371~383頁、2007)。他の実施形態では、16S rRNAまたは23S rRNAは、直接標識され、任意の増幅を用いずに使用される。

【0134】

プローブの調製

分析システム、ビーズ、または他のシステムで使用されるオリゴヌクレオチドプローブを生成するための技法および手段は、当業者に周知である。例えば、オリゴヌクレオチドプローブは、例えば、N-ホスホネートまたはホスホラミジット化学反応を使用して、合成ポリヌクレオチドまたはオリゴヌクレオチドを合成することによって生成することができる (Froehlerら、Nucleic Acid Res. 14:5399~5407頁(1986); McBrideら、Tetrahedron Lett. 24:246~248頁(1983))。合成の配列は一般に、長さが約10~約500塩基の間、より一般的には、長さが約15~約100塩基の間、最も好ましくは、長さが約20~約40塩基の間である。いくつかの実施形態では、合成核酸には、イノシンなどの非天然塩基が含まれるが、決してイノシンに限定されない。適当な核酸類似体の例は、ペプチド核酸である (例えば、Egholmら、Nature 363:566~568頁(1993); 米国特許第5,539,083号を参照)。いくつか

の実施形態では、少なくとも10、25、50、100、500、1,000、5,000、10,000、20,000、40,000、50,000、60,000、70,000、80,000、90,000、100,000、200,000、500,000、1,000,000、または2,000,000のプローブが、アレイ上に含まれる。さらなる実施形態では、各PMプローブは、アレイ上に存在する、対応するMMプローブを有する。一般に、各プローブ対は、OTUに関連する。いくつかの実施形態では、少なくとも10、25、50、100、500、1,000、5,000、10,000、20,000、40,000、50,000、60,000、70,000、80,000、90,000、100,000、200,000、または500,000のプローブ対が、アレイ上に配置される。一般に、プローブ対のセットは、少なくとも1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、または35の存在するプローブ対を有する。

10

【0135】

いくつかの実施形態では、標的配列(例えば、16S rRNA遺伝子)中の特定の配列に相補的な陽性対照プローブは、内部定量化標準(QS)として使用され、システム中に含まれる。他の実施形態では、内部DNA定量化標準(QS)プローブとしても知られる陽性対照プローブは、スパイクインされた核酸配列標的とハイブリダイズするプローブである。通常、配列は、代謝関連遺伝子に由来する。いくつかの実施形態では、陰性対照(NC)プローブ、例えば、標的配列(例えば、16S rRNA遺伝子)中の配列と相補的でない、または認め得るほどにハイブリダイズしないプローブが、アレイ上に含まれる。QSプローブと異なり、試料処理の前に、NCプローブのために、標的物質は試料ミックス中にスパイクされない。

【0136】

20

ハイブリダイゼーションプラットフォームの製作

いくつかの実施形態では、プローブは、別個に合成され、次いで固体支持体または表面に結合され、固体支持体または表面は、例えば、ガラス、ラテックス、プラスチック(例えば、ポリプロピレン、ナイロン、ポリスチレン)、ポリアクリルアミド、ニトロセルロース、ゲル、シリコン、または他の多孔質もしくは非多孔質物質から作製することができる。いくつかの実施形態では、表面は、マイクロビーズまたはロッドの場合のように球状または円柱状である。他の実施形態では、表面は、アレイまたはマイクロアレイの場合のように平面である。例えば、Schenら、Science 270:467~470頁(1995)によって一般に記載された方法を、ガラスプレート上にプリントすることによって表面に核酸を結合させるために使用することができる。高密度オリゴヌクレオチドアレイを作製するために一般に使用される他の実施形態では、規定された配列と相補的な数千のオリゴヌクレオチドが、フォトリソグラフィ技法(例えば、Fodorら、1991、Science 251:767~773頁; Peaseら、1994、Proc. Natl. Acad. Sci. U.S.A. 91:5022~5026頁; Lockhartら、1996、Nature Biotechnology 14:1675頁; 米国特許第5,578,832号、同第5,556,752号、および同第5,510,270号を参照)、または規定されたオリゴヌクレオチドの迅速な合成および沈着のための他の方法(例えば、Blanchardら、Biosensors & Bioelectronics 11:687~690頁)によって、表面上の規定された位置にin situで合成される。これらの方法のいくつかでは、既知の配列のオリゴヌクレオチド(例えば、25mer)が、誘導体化ガラススライドなどの表面上に直接合成される。分析システムを作製するための他の方法も、例えば、マスキング(MaskosおよびSouthern、1992、Nuc. Acids. Res. 20: 1679~1684頁)によって利用可能である。本発明の実施形態は、任意のタイプのアレイ、例えば、ビーズベースのアレイ、ガラスプレートまたは上述したような誘導体化ガラススライド上のアレイ、およびナイロンハイブリダイゼーション膜上のドットプロットに適用可能である。

30

40

【0137】

本発明の実施形態は、それだけに限らないが、ビーズもしくは溶液多重反応プラットフォームを含めて、または複数のプラットフォーム、例えば、Affymetrix GeneChip(登録商標)Arrays、Illumina BeadChip(登録商標)Arrays、Luminex xMAP(登録商標)Technology、Agilent Two-Channel Arrays、MAGIChips(ゲル固定化された化合物の分析システム)、もしくはNanoString nCounter Analysis Systemにわたる、任意の分析システムにおいて使用するために適用可能である。Affymetrix(Santa Clara, CA, USA)プラットフォームDNA

50

アレイは、 $1\mu\text{m}^2$ 当たり10,000分子のおおよその密度で、フォトリソグラフィー法によってガラス表面上に直接合成されたオリゴヌクレオチドプローブ(約25mer)を有することができる(Cheeら、Science (1996) 274:610~614頁)。スポットされたDNAアレイは、既定の濃度で個々に合成されるオリゴヌクレオチドを使用し、化学的に活性化されたガラス表面に適用される。一般に、オリゴヌクレオチドの長さは、長さが数ヌクレオチド~数百の塩基の範囲とすることができるが、一般に、長さが約10mer~50mer、約15mer~40mer、または約20mer~約30merである。

【0138】

微粒子システム

当技術分野で知られる技法を使用して生成されるオリゴヌクレオチドは、微小球、ビーズ、マイクロビーズ、ロッド、またはアレイ中に使用するための他の微視的粒子、フローサイトメトリー、および他の多重アッセイシステムに構築および/または結合することができる。多数の微粒子が、直径約0.01~100マイクロメートルで市販されている。一般に、約0.1~50 μm 、約1~20 μm 、または約3~10 μm の微粒子が好適である。微粒子のサイズおよび形状は均一であってもよく、これらは変化してもよい。いくつかの実施形態では、様々なサイズ、形状、または両方のサブロットが、標識された微粒子の最終的な混合ロットを作製するためにサブロットを合わせる前に、プローブにコンジュゲートされる。したがって、個々のサブロットは、そのサイズおよび形状に基づいて区別および分類することができる。微粒子のサイズは、いわゆる前方散乱光または小角散乱光によって、実質的に任意のフローサイトメトリー装置で測定することができる。粒子の形状も、フローサイトメトリー、例えば、高分解能スリット走査法によって区別することができる。

【0139】

微粒子は、ガラス、ガラスコンポジット、金属、セラミック、またはポリマーを含めた任意の固体または半固体物質から作製することができる。しばしば、微粒子は、ポリスチレンまたはラテックス物質であるが、任意のタイプのポリマー物質が許容され、それだけに限らないが、臭素化ポリスチレン、ポリアクリル酸、ポリアクリロニトリル、ポリアクリルアミド、ポリアクロレイン、ポリブタジエン、ポリジメチルシロキサン、ポリイソブレン、ポリウレタン、ポリビニルアセテート、ポリ塩化ビニル、ポリビニルピリジン、ポリビニルベンジルクロリド、ポリビニルトルエン、ポリ塩化ビニリデン、ポリジビニルベンゼン、ポリメチルメタクリレート、またはこれらの組合せを含む。微粒子は、磁性であっても非磁性であってもよく、また、蛍光色素、量子ドット、または微粒子構造中に組み込まれ、もしくは微粒子の表面に結合された他のインジケータ物質を有することができる。しばしば、微粒子は、1~30%の架橋剤、例えば、ジビニルベンゼン、エチレングリコールジメタクリレート、トリメチロールプロパントリメタクリレート、またはN,N'メチレン-ビス-アクリルアミド、または当技術分野で公知の他の機能的に等価な作用剤なども含むことができる。

【0140】

標的の標識化

一実施形態では、核酸標的は、光の特定の波長に調整されたレーザースキャナが、特定のDNAプローブとハイブリダイズした蛍光分子の数を測定できるように標識される。アレイについて、核酸標的は一般に、長さで15~100ヌクレオチドの間に断片化され、ビオチン化ヌクレオチドが、末端DNAトランスフェラーゼによって断片の末端に付加される。後の段階で、オリゴヌクレオチドプローブにハイブリダイズするビオチン化断片は、サンドイッチ(ストレプトアビジン)法によって複数のフィコエリトリンフルオロフォアを添加するための基質として使用される。AGILENTまたはNIMBLEGENによって作製されたものなどのいくつかのアレイについて、精製群集DNAを、DNAポリメラーゼのクレノウ断片を使用するランダムプライミングによって蛍光標識することができ、1つを超える蛍光部分を使用することができる(例えば、1つの分析システムとのハイブリダイゼーションによって直接比較するために、対照をCy3で標識し、実験試料をCy5で標識することができる)。いくつかの標識化方法では、標的の複数の標識コピーを生成するために、増幅または酵

素ステップの間に分子標識を標的中に組み込む。

【0141】

いくつかの実施形態では、検出システムは、PCR増幅を用いずに、その結果として、PCR増幅に付随する固有のバイアスを伴わずに、複雑な群集の微生物の多様性を測定することができる。活発に代謝する細胞は一般に、リボソームのコピーをほとんど有さない無活動細胞または死細胞と比較して、タンパク質アセンブリーのためにその細胞内に約20,000以上のリボソームのコピーを有する。いくつかの実施形態では、rRNAは、環境試料から直接精製し、増幅ステップを用いずに処理することができ、それによって他の配列に対するいくつかの配列の優先的な増幅によって引き起こされるいずれのバイアスも回避する。したがって、いくつかの実施形態では、分析システムからのシグナルは、試料中に存在するrRNA分子の真の数を反映することができる。これは、各細胞内のrRNAコピーの数を乗じた細胞の数として表現することができる。次いで、試料中の細胞の数は、いくつかの異なる方法、例えば、定量的リアルタイムPCR、またはFISH(蛍光in situハイブリダイゼーション)などによって推測することができる。次いで、各細胞内のリボソームの平均数を計算することができる。

10

【0142】

ハイブリダイゼーション

ハイブリダイゼーションは、当業者に周知の条件下で実施することができる。Rheeら(Appl. Environ. Microbiol. (2004) 70:4303~4317頁)、およびWuら(Appl. Environ. Microbiol. (2006) 72:4931~4941頁)を参照。温度を変更することによって、ストリンジエンシーを低減または増大させ、程度の差はあるが多岐にわたる配列の検出を可能にすることができる。ロボットのハイブリダイゼーションおよびストリンジエンシー洗浄ステーションを使用することによって、より一貫した結果を与え、処理時間を低減することができる。いくつかの実施形態では、ハイブリダイゼーションおよび洗浄プロセスは、約30分未満、1時間、2時間、3時間、4時間、5時間、6時間、7時間、8時間、9時間、10時間、11時間、12時間、14時間、16時間、18時間、20時間、または24時間で実現することができる。一般に、ハイブリダイゼーションおよび洗浄時間は、プローブの標的分子へのより大きなアクセスのしやすさのために、微粒子ベース検出システムについて低減される。一般に、ハイブリダイゼーション時間は、複雑性の低いアッセイおよび/または過剰の標的分析物が存在するアッセイについて低減することができる。

20

30

【0143】

シグナルの定量化

ハイブリダイゼーションの後、アレイは、適当な走査デバイスを使用して走査することができる。従来のマイクロアレイスキャナの非限定例として、GeneChip Scanner 3000またはGeneArray Scanner、(Affymetrix、Santa Clara、CA)、およびProScan Array(Perkin Elmer、Boston、MA)が挙げられ、これらは、10pmまたはより微細な分解能を有するレーザーを備えることができる。走査された画像表示は、ピクセル画像として取り込まれ、保存し、画像定量化ソフトウェア(例えば、GeneChip Analysis system Analysis Suite、バージョン5.1 Affymetrix、Santa Clara、CA;およびImaGene 6.0、Biodiscovery Inc. Los Angeles、CA、USA)を使用してアレイ上の各スポットのピクセル密度(強度)を定量化することによって分析することができる。各プローブについて、個々のシグナル値は、画像構文解析(imaging parsing)およびxy座標への変換を通じて得ることができる。各特徴についての強度の概要を作成することができ、特徴を含むピクセルの中の分散の推定値を計算することができる。

40

【0144】

フローサイトメトリーベースの検出システムを用いて、微粒子の各サブロット中の微粒子の代表的な画分を検査することができる。サブセットとしても知られる個々のサブロットは、サブロット内の微粒子が、比較的均質であるが、任意の他のサブロット中の微粒子と少なくとも1つの特徴的な特徴が異なるように調製することができる。したがって、微粒子が属するサブロットは、米国特許第6,449,562号に記載されているように、従来の

50

フローサイトメトリー技法を使用して、様々なサブロットから容易に求めることができる。一般に、レーザーが個々の微粒子上に照らされ、少なくとも3つの公知の分類パラメータ値、すなわち、一般にサイズおよび屈折率と関連する前方光散乱(C_1)、一般にサイズと関連する側方光散乱(C_2)、および一般に標識された標的配列中に組み込まれた蛍光色素の存在から生じる、少なくとも1つの波長での蛍光発光(C_3)が測定される。異なるサブセットからの微粒子は、上記に列挙された分類パラメータのうちの少なくとも1つが異なり、各サブセットについての分類パラメータは分かっているので、微粒子のサブロットの素性は、1つのアッセイステップで、リアルタイムで、微粒子のプールのフローサイトメトリー分析の間に確認することができる。特定のプローブを表す微粒子の各サブロットについて、ハイブリダイゼーションシグナルの強度は、バックグラウンドの減算を実施した後、シグナル分散の推定値とともに計算することができる。

10

【0145】

データ処理および統計分析

高レベルの信頼度を伴った、少なくとも500、1,000、5,000、10,000、20,000、30,000、40,000、50,000、60,000、またはそれ以上の分類群の同時検出は、多数のプローブセットのシグナル強度をデコンボリューションする技法を、確率推定に組み入れることができる。いくつかの実施形態では、本発明の方法、組成物、およびシステムは、微生物が微生物の総個体数の0.05%未満を構成する場合、環境試料または臨床試料などの微生物の群集における微生物の存在または非存在を1つのアッセイで検出することを可能にする。いくつかの実施形態では、検出は、微生物の量、例えば、総微生物個体数中の微生物の百分率を求めるステップを含む。デコンボリューション技法は、NCプローブ対を分析システム中に組み入れること、およびQSプローブ対からのハイブリダイゼーションシグナルを、NCプローブ対のハイブリダイゼーション分布にフィッティングするためのデータの使用を含む。

20

【0146】

デコンボリューション技法は、試料中の核酸の検出および定量化、ならびに推定によって、試料中の微生物の検出、および定量化を可能にすることができる。本発明の一態様では、試料中の微生物の存在または量を判定するためのシステムであって、試料を複数のプローブと接触させるステップと、プローブを用いて、試料核酸のハイブリダイゼーションシグナルを検出するステップと、シグナルをデコンボリューションすることによって、特定の核酸が総核酸個体数の0.01%未満存在する場合に、核酸の集団中に存在する特定の核酸の存在、非存在および/または量を判定するステップとを含むシステムが提供される。いくつかの実施形態では、特定の核酸は、集団中の他の核酸と、少なくとも80%、85%、90%、91%、92%、93%、94%、95%、96%、または97%相同である。

30

【0147】

いくつかの実施形態では、画像化および走査された試料からのデータ出力は、以下の方法を使用してデコンボリューションおよび分析される。例示的な例としてアレイを使用して、ハイブリダイゼーションシグナルは、市販のソフトウェアを使用してピクセルについて生成された強度の概要および分散の推定値を有するxy座標に変換される。データは、CELファイル(Affymetrix)、またはFeature Reportファイル(NimbleGen)のような標準的なデータ形式を使用して出力される。

40

【0148】

ハイブリダイゼーションシグナルは、バックグラウンドが減算される。一般に、バックグラウンド強度は、象限中の最小強度2%のプローブの平均シグナル強度として、各象限について独立して計算される。他の閾値値、例えば、0.5%、1%、3%、4%、5%、または10%も使用することができる。次いでバックグラウンド強度は、さらなる計算が実施される前に、象限中のすべてのプローブから減算される。このノイズ除去手順は、象限ごとに、またはアレイ全体にわたって行うことができる。

【0149】

いくつかの実施形態では、異なる実験で実現された結果の比較、または反復される実験

50

の比較を可能にするために、アレイシグナルが正規化される。正規化は、いくつかの方法で実現することができる。一実施形態では、同じ標的に対する異なるプローブ間の再現性が、Zhang L.ら、A model of molecular interactions on short oligonucleotide analysis systems、Nat. Biotechnol. 2003、21(7):818~821頁に記載されているような位置依存最近傍(Position Dependent Nearest Neighbor)(PDNN)を使用して評価される。PDNNモデルは、配列特異的ノイズシグナルおよび非特異的バックグラウンドシグナルの推定を可能にし、したがってプローブの真の強度の推定を可能にする。

【 0 1 5 0 】

他の実施形態では、PMおよびMMプローブ対と、内部DNA定量化標準(QS)プローブ対との比較から観察される応答を用いたシグナル分布およびバックグラウンド分布のパーアレイモデル(per-array model)が作り出される。一実施形態では、各プローブ対が「陽性」である確率は、各プローブ対について、差異スコアdを計算することによって求められ、dは

10

【 0 1 5 1 】

【数 1】

$$d = 1 - \left(\frac{PM - MM}{PM + MM} \right) \quad \text{式 1}$$

【 0 1 5 2 】

20

として定義することができ、式中、

PM = パーフェクトマッチプローブのスケールされた強度、

MM = ミスマッチプローブのスケールされた強度、

d = ペアディファレンススコア

である。

dの値は、0~2の範囲となり得る。PM>>MMである場合、dの値は0に近づき、PM=MMである場合、d=1であり、PM<<MMである場合、dの値は2に近づく。

【 0 1 5 3 】

いくつかの実施形態では、内部DNA定量化標準(QS)および陰性対照(NC)プローブ対は、プローブの特性によってピン化および分別される。本発明の実施形態において使用することができるプローブの特性の例として、それだけに限らないが、結合エネルギー;A+T数、G+C数、およびT数を含めた塩基組成;配列の複雑性;クロスハイブリダイゼーション結合エネルギー;二次構造;ヘアピン形成能;融解温度;ならびにプローブの長さが挙げられる。プローブのこれらの特性は、プローブのハイブリダイゼーション性に影響する場合があり、例えば、A+T数は、プローブの水素結合に影響する場合があり、T数は、DNaseの使用によって生成される断片の長さおよび塩基組成に影響する場合がある。他の酵素システムを用いた断片化は、他の塩基の組成によって影響され得る。

30

【 0 1 5 4 】

一実施形態では、QSおよびNCプローブ対は、個々のプローブのA+T数およびT数に基づいてピン化および分別される。各ピン(A+T数、T数)について、陰性対照プローブからのd値が、正規分布にフィッティングされることによって、スケール(平均)および形状(標準偏差)が導出される。次いで、QSからのd値が、ガンマ分布にフィッティングされることによって、スケールおよび形状が導出される。各アレイについて、複数の密度プロットがこのプロセスによって生成される。同じアレイ内の2つの異なるプローブピンから生成された密度プロットの2つの例を、図4A~Bに示す。AT数は、両図に表されたプローブについて14である。T数は、図4A中のプローブについて9であり、一方、T数は、図4Bに表されたプローブについて10である。これらのグラフは、図4Bに示されるように、1つの追加のTでさえ、プローブガンマスケールパラメータの感知できる差異をもたらし得ることを実証する。79のアレイにわたるガンマスケールの変化は、図5に示されている。

40

【 0 1 5 5 】

50

ガンマ分布および正規分布から導出されるパラメータは、各プローブ対について、対応答スコア、 r を導出するのに使用され、 r は、プローブ対が陽性である確率、すなわち、プローブ対が標的配列に対して応答性である確率のインジケータであり、 r は、

【 0 1 5 6 】

【 数 2 】

$$r = \left(\frac{pdf_{\gamma}(X = d)}{pdf_{\gamma}(X = d) + pdf_{norm}(X = d)} \right)$$

式 2

【 0 1 5 7 】

として定義することができ、式中、

r = 特定のプローブ対が標的配列に結合しており、バックグラウンドシグナルでない可能性、すなわち、プローブ対が特定の標的配列に対して陽性である確率を見積もるための応答スコア

$pdf_{\gamma}(X = d)$ = d を、標的クラスATx Tyについて推定されたガンマ分布から引き出すことができる確率；

$pdf_{norm}(X = d)$ = d を、標的クラスATx Tyについて推定された正規分布から引き出すことができる確率

である。

r は、0～1の範囲となり得る。PM>>MMである場合、 r は1に近づき、PM<<MMである場合、 r は0に近づく。

【 0 1 5 8 】

インターロゲーションプローブ対の各セット、例えば、OTUは、対応答スコア、クロスハイブリダイゼーション関係、または両方に基づいてスコアを付けることができる。いくつかの実施形態では、システムは、前記微生物の存在または量に対して最終コールを行う前に、プローブ対セットの少なくとも1つのサブセットからデータを除去する。一実施形態では、データは、インターロゲーションプローブのクロスハイブリダイゼーション能に基づいて除去される。一実施形態では、プローブ対のスコアリングは、以下に記載されるような2段階プロセスによって実施される。

【 0 1 5 9 】

例えば、第1段階をクリアするプローブのみが次段階で分析される、2段階分析を実施することができる。第1段階では、プローブ対の各セットにわたる r の分布、 R が求められる。OTUに関連するプローブ対の各セットについて、すべてのプローブ対の r 値がセット内でランク付けされ、1つまたは複数の閾値 r 値を満たすプローブ対の百分率が求められる。しばしば、3つの閾値判定が、ランク付けされたプローブ対の全範囲にわたる25%の増分で行われる(四分位間のQ1、Q2、およびQ3)が、任意の数の閾値判定または百分率増分を使用することができる。例えば、判定は、プローブ対が80%の閾値をクリアしなければならない、70%での1つの増分を使用することができる。

【 0 1 6 0 】

一般に、ノイズからシグナルを区別するために、OTUは、このOTUに関連するプローブ対のセットのQ1、Q2、およびQ3が、それぞれ $Q1_{min}$ 、 $Q2_{min}$ 、および $Q3_{min}$ という閾値を越える場合、段階1をクリアするとみなされる。すなわち、段階1をクリアするOTUについて、そのOTUに関連するプローブ対のセット中のプローブ対の75%の r 値は、少なくとも $Q1_{min}$ でなければならず、プローブ対のそのセット中のプローブ対の50%の r 値は、少なくとも $Q2_{min}$ でなければならず、プローブ対のそのセット中のプローブ対の25%の r 値は、少なくとも $Q3_{min}$ でなければならず。 $Q1_{min}$ は、少なくとも約0.5、約0.55、約0.6、約0.65、約0.7、約0.75、約0.8、約0.82、約0.84、約0.86、約0.88、約0.90、約0.91、約0.92、約0.93、約0.94、約0.95、約0.96、約0.97、約0.98、または約0.99である。 $Q2_{min}$ は、少なくとも約0.5、約0.55、約0.6、約0.65、約0.7、約0.75、約0.8、約0.82、約0.84、約0.86、約0.88、約0.90、約0.91、約0.92、約0.93、約0.94、約0.95、約0.96、約0.97、約0.98、また

10

20

30

40

50

は約0.99である。Q3_{min}は、少なくとも約0.5、約0.55、約0.6、約0.65、約0.7、約0.75、約0.8、約0.82、約0.84、約0.86、約0.88、約0.90、約0.91、約0.92、約0.93、約0.94、約0.95、約0.96、約0.97、約0.98、約0.99、約0.992、約0.994、約0.996、約0.998、または約0.999である。いくつかの実施形態では、Q1_{min}、Q2_{min}、およびQ3_{min}は、スパイクイン実験から経験的に求められる。例えば、Q1_{min}、Q2_{min}、およびQ3_{min}は、2pMの単位複製配列濃度をクリアさせるように選ばれる。一実施形態では、Q1_{min}、Q2_{min}、およびQ3_{min}は、それぞれ、0.98、0.97、および0.82である。これらの閾値数値は、試料配列を断片化するためにDNaseを使用して経験的に導出された。DNaseは、T-バイアスを有するので、他の酵素の使用は、閾値数値のシフトを必要とする場合があり、経験的に導出することができる。

10

【0161】

第2段階では、第1段階をクリアするOTUのみが、クロスハイブリダイゼーションの可能な源とみなされる。いくつかの実施形態では、各OTUについて、 $r > 0.5$ を有するプローブ対(これらは、標的配列に応答性である可能性があるものとみなされたプローブ対である)のみがさらに分析される。他の場合では、 $r > 0.6$ 、0.7、0.8、または0.9を有するプローブのみが応答性であるとみなされ、さらに分析される。応答性である見込みがない(すなわち、 $r < 0.5$)プローブ対は、そのセットのRが全体的に応答性であっても、さらに分析されない。R_{0.5}は、すべてのプローブ対が $r > 0.5$ を有する、プローブ対のサブセットを表す。一般に、段階1で選ばれた四分位間のQ1、Q2、およびQ3値に基づいて、段階1をクリアするOTU中のプローブ対のほとんどが分析される。他の実施形態では、 $r > 0.55$ 、0.60、0.65、0.70、0.75、0.80、0.85、または0.90を有するプローブ対のみがさらに分析される。

20

【0162】

R_{0.5}サブセット中の各プローブ対について、推定的にクロスハイブリダイズするOTUの数(すなわち、プローブ対がクロスハイブリダイズすることができるOTUの数)が求められる。このプロセスにおいて、段階1をクリアしたOTUのみが、クロスハイブリダイゼーションの可能な源とみなされる。R_{0.5}サブセット中の各プローブ対は、その r 値を推定的にクロスハイブリダイズするOTUの数で除することによりペナルティが課されることによって、その陽性であるという修正された可能性が求められる。プローブ対についての陽性であるという修正された可能性は、 r_x 値によって表すことができる。 r_x は、

30

【0163】

【数3】

$$r_x = \frac{r}{\text{スカラー } S_{1x}} \quad \text{式3}$$

【0164】

として定義することができ、式中、

S_1 = 段階1クリアするOTUのセット、および

S_{1x} = 所与のプローブ対に対してクロスハイブリダイゼーション能を有する段階1をクリアするOTUのセットである。

40

【0165】

r_x は、第1段階の間に観察される群集を考慮すると、プローブ対の応答およびプローブ対の特異性に比例する。 r_x 値は、0~1の範囲となり得る。OTUと関連するプローブ対の各セットについて、 r_x が各プローブ対について計算され、セット内でランク付けされる。プローブ対の各セットにおける r_x 値の分布についての四分位間のQ1、Q2、Q3値が求められる。OTUによって表される分類群は、Q1がQ_{x1}より大きく、Q2がQ_{x2}より大きく、またはQ3がQ_{x3}より大きい場合存在するとみなされる。Q_{x1}は、少なくとも約0.5、少なくとも約0.55、少なくとも約0.6、少なくとも約0.65、少なくとも約0.7、少なくとも約0.75、少なくとも約0.8、少なくとも約0.85、少なくとも約0.90、少なくとも約0.95、または少なくとも約0.97である。Q_{x2}は、少なくとも約0.5、少なくとも約0.55、少なくとも約0.6、少なくとも

50

約0.65、少なくとも約0.7、少なくとも約0.75、少なくとも約0.8、少なくとも約0.85、少なくとも約0.90、少なくとも約0.95、または少なくとも約0.97である。 Q_{x3} は、少なくとも約0.5、少なくとも約0.55、少なくとも約0.6、少なくとも約0.65、少なくとも約0.7、少なくとも約0.75、少なくとも約0.8、少なくとも約0.85、少なくとも約0.90、少なくとも約0.95、または少なくとも約0.97である。一実施形態では、 Q_{x1} は、少なくとも0.66であり、すなわち、プローブ対のセット中のプローブ対の75%は、少なくとも0.66である r_x 値を有する。

【0166】

2段階ハイブリダイゼーションシグナル分析手順を、開示された方法を使用して選択されるプローブの任意の組合せを使用して生成されるデータを含めて、任意のアレイまたは微粒子が生成するデータセットからのハイブリダイゼーションシグナルに対して実施することができる。いくつかの実施形態では、手順の第2段階では、クロスハイブリダイゼーションの数、クロスハイブリダイゼーションシグナルの強度、またはこれらの2つの組合せに基づいてプローブにペナルティが課される。

【0167】

本明細書に開示される方法は、階層的プローブセットのスコア付けに有用である。OTUは、クラスタリングツリー上の任意の階層的レベルでのノードに存在し得る。本明細書で使用する場合、OTUは、1つまたは複数の生物の群、例えば、ドメイン、サブドメイン、界、亜界、門、亜門、綱、亜綱、目、亜目、科、亜科、属、亜属、種、または任意のクラスターなどである。いくつかの実施形態では、 R_{05} は、系統樹上の各ノードについて収集され、下位の R_{05} セットからのすべてのユニークなプローブからなる。例えば、「目」を表すOTUについての R_{05} セット中のプローブ対についての r_x 値を計算するために、クロスハイブリダイゼーション能を有する少なくとも1つの配列を含む、推定的にクロスハイブリダイズする、等しくランク付けされた分類群(すなわち、「目」ノード)の数が、式3における分母として使用される。

【0168】

いくつかの実施形態では、リーフレベルのOTU(例えば、種、亜属、または属)が最初に分析される。次いで、クラスタリングツリー中のノードの各次のレベルが分析される。一実施形態では、分析は、最高でドメインレベルまで実施される。別の実施形態では、分析は、最高で門レベルまで実施される。さらに別の実施形態では、分析は、最高で界レベルまで実施される。式3中のクロスハイブリダイゼーションについてのペナリゼーションは、同じ分類学レベル上のプローブについてのみ実施される。すべての存在する分類群は、セットRの最高値および最低値を廃棄した後、平均スケールのPMプローブ強度を使用して定量化される(HybScore)。いくつかの実施形態では、第1のレベルで存在する分類群のみがさらに分析される。

【0169】

いくつかの実施形態では、概略の存在量スコアが求められる。補正存在量スコアは、G+C含量およびウラシルの組み込みに基づいて作り出される。一般に、より高いG+C含量を有するプローブは、より高いハイブリダイゼーションシグナルを生じ、これは一般に、存在量スコアを補正するために相殺される。

【0170】

各分類学的ノードについての検出の確率は、末端ノード検出、およびクロスハイブリダイゼーション関係の幅を要約することによって求められる。階層的プローブは、クラスター分析に基づいて、新規生物の証拠についてスコアを付けられる。

【0171】

いくつかの実施形態では、システムは、プローブハイブリダイゼーションシグナル強度の分析から得られるものとともに、他のデータを分析することができる。いくつかの実施形態では、システムは、ハイスループット配列決定技法を用いて得られるデータを含めた、シーケンシング反応データを分析することができる。いくつかの実施形態では、配列決定データは、プローブを使用して本明細書に開示される方法によって分析される同じ高

10

20

30

40

50

度に保存された配列の同じ領域からのものである。

【0172】

大容量分析システムアプリケーション

多数の自然のヒトが作り出した環境は、環境のマイクロバイーム組成を求めるために試料採取およびアッセイすることができる。少なくとも10,000、20,000、30,000、40,000、50,000、60,000、70,000、80,000、90,000、100,000、200,000、500,000、または1,000,000の細菌または古細菌の分類群の存在または量を1つのアッセイで検出することができるアッセイシステムを有することによって、原核生物の生態系の完全なイメージを、迅速に、かつ比較的低費用で実現することができ、科学的な、健康管理の、または規制上の興味対象の多数の環境を検査する能力を提供する。

10

【0173】

生態系、物理的環境、作物、動物、ヒト、臓器系などに関連する特定のマイクロバイームの解明により、本明細書で互換的に使用される用語である、試料採取される特定の環境の「シグネチャー」、「バイオシグネチャー」または「フィンガープリント」の生成が可能になる。バイオシグネチャーが、正常もしくは健康な系もしくは個体からのもの、または物理的環境に生息し、もしくは物理的環境で生成されるアイテムを使用する個体の健康状態の維持に関連する物理的環境からのものである場合、正常または健康な場所のバイオシグネチャーを、異常または不健康な状態または条件に関連する変化をモニターするために、同じ環境からの後の試料を比較するための参照として使用することができる。例えば、水源の後のバイオシグネチャーが、マイクロバイームが、飲料水に関連するマイクロバイームからシフトしていることを示す場合、例えば、混入物および/または汚染源を特定し、これを処理および/または除去するためのステップをとることによって、継続したシフトを予防するために先制の対策をとることができる。さらなる例として、果樹園の後のフィンガープリントが、マイクロバイームが、健康な樹木および高い生産性と関連するマイクロバイームからシフトしていることを示す場合、先制の対策をとることによって、健康な微生物の増殖および維持に好都合である栄養分を適用することができ、または代わりに、堆肥茶を適用することによって、健康な微生物の数を増大させることができる。

20

【0174】

同様に、ある環境のバイオシグネチャーを、その環境の集団またはコレクションについての平均の、または正常なバイオシグネチャーを表す試料のプールから生成されるバイオシグネチャーと比較することができる。例えば、不健康な個体からの試料をアッセイすることができ、微生物のバイオシグネチャーを、健康な集団において一般に見られるバイオシグネチャーと比較することができる。一般的な集団において見られないか、または同じ蔓延率で見られない1つまたは複数の微生物が不健康な個体において検出される場合、治療的対策をとることによって、不健康な状態に関連する微生物を選択的に排除し、またはこの微生物の数を低減することができる。例えば、胃腸管のマイクロフローラを、アレルギーを患っている小児と健康な小児との間で比較することができる。アレルギー患者が、他の小児と比較して、患者の胃腸管中に1つまたは複数の支配的な微生物を有することが示される場合、蔓延している異常な微生物を特異的に標的にする利用可能な薬物および/または食事療法を投与することができる。代わりに、またはさらに、プロバイオティック食品またはサプリメントなどを通じて、健康な小児と関連する多数の微生物を導入することによって、アレルギー患者における胃腸の集団をシフトすることができる。同様に、アレルギー患者に、健康な微生物の増殖を促進する栄養補給剤を与えることができ、または小児の親に、小児の飲食物を、不健康な微生物のものに対して、健康な微生物の増殖に好都合である食物に変更するよう指示することができる。特定の微生物または微生物の群の蔓延と、疾患状態との間の関係が分かった後、疾患の進行または処置の応答も、本システムおよび方法を使用してモニターすることができる。

30

40

【0175】

腸、呼吸器系、尿生殖路、乳腺、皮膚、口腔、耳道、および皮膚を含めて、動物または

50

ヒトの多数のマイクロバイオーームを、本システムおよび方法を用いて分析することができる。臨床試料、例えば、血液、痰、鼻孔、糞便、および尿などを、本方法とともに使用することができる。正常な個体および疾患または状態を患っている個体の分析から、フィンガープリントまたはバイオシグネチャーの大きなデータベースを構築することができる。健康な状態と疾患に関連する状態との間のバイオシグネチャーを比較することによって、マイクロバイオーームの個々の成分の影響および重要性に関して関連付けを行うことができる。

【0176】

これらの関連付けが行われた後、処置を設計および試験することによって、疾患状態において見られる微生物叢の組成を変更することができる。さらに、疾患した個体における患部器官系の微生物組成を定期的にモニターすることによって、疾患進行または療法に対する応答を観察することができ、必要な場合、さらなる治療的対策をとることによって、マイクロバイオーーム組成を、健康な集団において見られる組成のより代表的なものに変更することができる。

【0177】

健康管理、水質および食物の安全性において大きな重要性を有する細菌の興味深い性質は、クオラムセンシングである。多くの細菌は、その種または関連種の他のメンバーの存在を感知することができ、特定の密度に到達すると、細菌は、様々な毒性または病原性因子を産生し始める。言い換えれば、細菌の遺伝子発現は、群として協調される。例えば、いくつかの細菌は、「粘液層」として知られるエキソポリサッカライドを産生する。エキソポリサッカライドの分泌は、白血球の微生物を食べる能力を減少させ、微生物を治療剤または清浄剤に対してより耐性にする場合がある。従来の方法は、クオラムセンシングおよび他の集団によって誘発される効果を検出または試験するために、特定の遺伝子発現の検出を必要とする。本システムおよび方法を使用することによって、生物膜形成または毒性産生などの所与の効果と関連したマイクロバイオーーム中で起こる変化を理解することができる。クオラムセンシングに至る条件を探し、判定するために、本システムおよび方法を伴うプロトコルを開発することができる。例えば、様々な時点で、様々な条件下で試料を試験することにより、毒性または病原性因子の発現を誘発された集団にどのように、かついつ介入し、またはこの集団を逆戻りさせるかを決定することができる。

【0178】

例えば、衛星および他の宇宙船のコンポーネントをアセンブルするのに使用されるクリーンルームを、本システムおよび方法を用いて調査することによって、どの微生物の群集が存在するかを理解し、他の惑星への陸生微生物の導入もしくはその試料の導入を防止するためのより良好な汚染除去および洗浄技法を開発し、または推定上の地球外微生物によって生成されたデータを、混入した陸生微生物によって生成されたデータと区別するための方法を開発することができる。

【0179】

例えば、食物調製箇所、集中治療施設、手術室などのクリーンルーム環境、薬物製造施設、医療用デバイス製造施設などを、本システムおよび方法を用いて調査することによって、局所的な微生物の群集の組成、および微生物の群集を含む個々の分類群の量を確認することができる。そのような試験は、製造プロセスにおける汚染、およびその後の汚染された消費者製品のリコール、または感染および疾患の蔓延を防止するための手段となり得る。

【0180】

一実施形態では、本システムおよび方法を用いて、環境状態または健康状態についての新しいインジケータ種を同定するための方法が提供される。状態は、正常または健康な状態のものであってもよい。あるいは、インジケータ種は、不健康または異常な状態についてののものであってもよい。新しいインジケータ種を同定するために、正常な試料が同時にアッセイされることによって、すべての既知の細菌、古細菌、または真菌と関連する各OTUの存在または量が判定され、この試験結果は、すべての既知の細菌、古細菌、または真

10

20

30

40

50

菌と関連する各OTUの存在または量が判定された状態の環境からの試料の同時アッセイで実現された結果と比較される。存在量の増加であっても、存在量の減少であっても、存在量が少なくとも、2倍、3倍、4倍、5倍、10倍、20倍、50倍、または100倍変化する微生物は、状態についての推定上のインジケータ種を表す。

【0181】

いくつかの実施形態では、根の成長、ならびに土壌中の炭素基質の利用可能性の増大、および重金属もしくはウランの存在などの土壌組成の変化、土壌pHの変化、ならびに沈殿物の量およびパターンの変化を含めた環境変化と関連するインジケータ種を同定するための方法が提供される。他の実施形態では、サンゴのストレス、ならびにサンゴ白化現象、または他の海洋および他の水の環境の変化と関連するインジケータ種を同定するための、本システムおよび方法を使用する方法が提供される。

10

【0182】

他の実施形態では、CF患者における疾患の進行、およびCOPDの憎悪を含めた、疾患状態、疾患の進行、治療レジメン、生菌投与と関連したインジケータ種を同定するための方法が提供される。他の実施形態では、群集中に1つまたは複数の新しい微生物を導入することに関連した環境状態または健康状態の変化をモニターするための方法が提供される。例えば、ヨーグルトまたはカセイ菌(*L. casei*)を含有する食物サプリメントを人に与えることなどの、個体中の腸マイクロバイオームの特定の微生物の百分率を増大させるための対策は、本方法およびシステムを使用してモニターすることができる。

20

【0183】

複合分析

試料中の微生物を同定および定量化する能力を、機能的遺伝子アレイなどの遺伝子発現技術と組み合わせることによって、集団を、観察される遺伝子発現と関連させることができる。同様に、マイクロバイオーム組成分析を、酵素、毒素、薬物、抗生物質、または他の試料成分を含めた化学物質、タンパク質の存在と関連させることができる。例えば、土壌試料から単離された核酸を分析することによって、マイクロバイオーム組成(例えば、バイオシグネチャー)を解明し、また、発現された遺伝子を同定することができる。南極上の露出した栄養分の乏しい土壌では、この分析は、キチナーゼおよびマンナーゼ発現をバクテロイデスと関連付け、CH₄-関連遺伝子をアルファプロテオバクテリア(*Alphaproteobacteria*)と関連付けた。(Yergeauら、*Environmental microarray analyses of Antarctic soil microbial communities*. ISME J. 3:340~351頁、2009)。有意な相関は、分類群の存在量と、C-およびN-サイクル遺伝子の存在量との間にも見出された。このデータから、ある特定の生物または生物の群は、予期または観察される酵素または分解プロセスの大部分を必要とし、または占めることを予測することができる。例えば、バクテロイデス門のメンバーは、試料現場で、昆虫および節足動物の外骨格ならびにまた真菌細胞壁の主要成分である、環境キチンの大部分をおそらく分解する。

30

【0184】

この方法を使用することによって、新しい抗生物質を産生する生物、培養できない生物でさえ同定することができる。例えば、土壌抽出物を、抗生物質活性について試験することができる。陽性の抽出物が見出される場合、抗生物質について一部が抽出された土壌の試料は、微生物の組成およびおそらく遺伝子発現について分析することができる。マイクロバイオームの主要な成分は、抗生物質活性と関連させることができ、相関は遺伝子発現データ通じて強化され、特定の生物または生物の群が、観察された抗生物質活性に関与することを予測することを可能にする。

40

【0185】

一態様では、本発明は、試料中の状態を判定するための方法を提供する。一実施形態では、この方法は、a)前記試料を複数の異なるプローブと接触させるステップと、b)前記プローブのそれぞれについてのハイブリダイゼーションシグナル強度を測定するステップとあって、前記測定は、前記試料についてのバイオシグネチャーを確立するステップと、c)前記試料のバイオシグネチャーを、糞便汚染についてのバイオシグネチャーと比較するス

50

テップとを含む。いくつかの実施形態では、試料についての予測を行うための方法であって、a)前記試料中の微生物の少なくとも100のOTUの存在または非存在の確率として、微生物の個体数データを求めるステップと、b)前記試料中の前記微生物による1つまたは複数の遺伝子の遺伝子発現データを求めるステップと、c)前記発現データおよび個体数データを使用することによって、前記試料についての予測を行うステップとを含む方法が提供される。いくつかの実施形態では、予測は、土壌または局所的環境中で観察される特徴または状態に關与する微生物の素性を必要とする。

【0186】

他の複合分析方法には、水試料中の微生物を保持する一方で、水試料の1つまたは複数の成分またはパラメータが変更される拡散チャンバーの使用が含まれる。例えば、水の塩分またはpHを、経時的に急激または徐々に変更することができる。拡散チャンバーは、例えば、未処理の汚水が配置される受入れ水(receiving water)の状態を模倣するのに有用である。特定の時間間隔の後、拡散チャンバー中の水試料のマイクロバイームを求めることができる。新しい環境状態に耐えることができない微生物は、死に、芳しくない状態もしくは捕食のために数が減少した状態になり、またはその数が変わらないままである。対照的に、新しい状態に耐えることができる微生物は、その数を少なくとも維持し、またはおそらく支配的な集団なりながら繁殖する。少なくとも10,000のOTUの存在または量を検出することができるシステムと結合した拡散チャンバーを使用することにより、新しい環境中に置かれた場合に、消滅し、または繁殖することができない微生物の同定を可能にすることができる。そのような微生物は「一過性」と呼ばれ、マイクロバイームの百分率組成は、急速に変化することを意味する。一過性の微生物の同定は、これらが環境中に導入された時間および/または場所を確認するために使用することができる。例えば、約4時間の半減期を有する、汚染水と関連する感知できる量の一過性の微生物が水試料中で同定されることは、その微生物は、過去1日以内(6半減期)に水体中におそらく導入されたことを示す。異なる一過性の微生物は、特定の状態について異なる半減期を有する場合がある。汚染水と関連する様々な一過性の微生物の受入れ水中での半減期の知見を備えることで、流出、例えば、汚水排出の時間経過を構築することができる。時間経過の使用は、排出源を特定するのに使用することができ、例えば、クルーズ船または貨物船による不法な排出の場合、違反者の識別および出頭命令を可能にする。

【0187】

拡散チャンバーは、半透性のカプセル、管、ロッド、もしくは球、または他の固体もしくは半固体物体の形態もとることができる。マイクロバイームまたは細菌の選択群は、カプセル内部に配置することができ、次いでこれは密閉され、ある環境中に指定された時間の期間投入される。取り出されたのち、カプセルが開けられ、マイクロバイームまたは細菌の選択群が試料採取されることによって、個々の成分の存在または量の変化が確認される。例えば、拡散チャンバー中に未処理汚水の試料を入れるのではなく、未処理汚水は、半透性カプセル中に入れられ、次いでこれは、ある量の受入れ水、または実際の受入れ水体中に配置される。カプセルは、マイクロバイームを試料採取するために、その量の受入れ水または水体から1回または定期的に取り出すことができる。あるいは、同一量のマイクロバイームを有する複数の使い捨てカプセルを使用することができ、各カプセルは、異なる時間に取り出し、試料採取することができる。カプセルまたは他の半透性容器に入れられたマイクロバイームは、特定の器官または系の環境に関連するマイクロバイーム組成の変化を測定するために、通常開口部を通じて生体中に導入することができる。例えば、マイクロバイームを含む半透性カプセルまたは管を、口または肛門を通じて胃腸系内に投入することができる。健康な個体からのマイクロバイームを、不健康な固体、例えば、クローン病または過敏性腸症候群を患っている患者中にこのようにして投入することによって、正常な、健康な個体に関連するマイクロバイームに対する不健康な状態の効果を確認することができる。このようにして、薬物有効性および治療プロトコルの効力も、既知のマイクロバイームに対する腸環境の効果に基づいて評価することができる。

10

20

30

40

50

【 0 1 8 8 】

低密度特殊目的の検出システム

いくつかの実施形態では、プローブは、アレイまたは微粒子を伴うものを含めた特殊目的のシステムを構築するために選択される。一般に、特殊目的「低密度」システムは、通常、特定の環境において、または特定の用途のために使用するために設計され、特定のバイオシグネチャーに関連するなど、特定の環境中に存在することが知られ、または予期される生物に特異的である、低減された数のプローブ、すなわち「ダウン選択された(down-selected)」プローブを特徴とする。いくつかの場合では、バイオシグネチャーは糞便汚染である。一般に、低密度システムは、10、20、50、100、200、500、1,000、2,000、5,000、もしくは10,000以下のダウン選択されたプローブ、または5、10、25、50、100、250、500、1,000、2,500、もしくは5,000のダウン選択されたプローブのプローブ対(PMおよびMMプローブ)を備える。いくつかの実施形態では、わずか1、2、3、4、5、6、7、8、9、または10のプローブが、10TU当たり使用される。さらなる実施形態では、PMプローブのみが使用される。一般に、これらのダウン選択されたプローブは、ロバストなハイブリダイゼーションシグナルを有し、ほとんど、またはまったくクロスハイブリダイゼーションを有さない。いくつかの実施形態では、ダウン選択されたプローブのコレクションは、1プローブ当たり20、15、10、8、7、6、5、4、3、2、または1未満の中央値クロスハイブリダイゼーション能数を有する。しばしば、ダウン選択されたプローブは、低減された数のプローブを有するOTUに属する。いくつかの実施形態では、ダウン選択プローブコレクションのOTUは、10TU当たり25、20、15、14、13、12、11、10、9、8、7、6、5、4、3、または2未満のプローブの中央値数を有する。一般に、低密度システムは、10、25、50、100、250、500、1,000、2,000、または5,000以下の分類群を認識するプローブを特徴とする。設定された数のプローブについて、いくつかの設計ストラテジーを、低密度システムのために使用することができる。1つの手法は、例えば、ミスマッチプローブをまったく用いずに、10TU当たり1つのプローブを使用して、同定されるOTUの数を最大にすることである。別の手法は、所望の信頼度レベルに基づいてプローブを選択することである。ここでは、各OTUの存在および量について少なくとも95%の信頼度レベルを実現するために、対応するミスマッチプローブとともに、各OTUについて複数のプローブを必要とする場合がある。特定の低密度用途のためのプローブは、適切な環境からの試料を、高密度分析システム、例えば、少なくとも10,000、20,000、30,000、40,000、50,000、60,000、70,000、80,000、90,000、100,000、250,000、500,000、もしくは1,000,000の、1つのドメイン、例えば、細菌、古細菌、もしくは真菌などのOTUの、または代わりに、1つのドメインの各既知のOTUについての存在または量の確率を1つのアッセイで求めることができる検出システムに適用することによって選択することができる。蔓延しているOTUに関連するプローブを、低密度システムのために選択することができる。あるいは、対象とする試料中に見られるOTUは、対照試料と比較し、差し引かれたOTUを、低密度システムのために選択された残りのOTUに関連するプローブと共有することができる。さらに、プローブは、対象とする環境と対照環境との間のOTUの蔓延の変化に基づいて選択することができる。例えば、対照試料と比較して、対象とする試料中で少なくとも2倍、5倍、10倍、100倍、または1,000倍超豊富なOTUが、ダウン選択されたプローブセット中に含まれる。この情報を使用して、ダウン選択されたアレイ、ビーズ多重システム、または他の低密度アッセイシステムが設計される。

【 0 1 8 9 】

「低密度」アッセイシステムは、選択微生物を同定し、互いに関連した様々な選択微生物の百分率組成を求めるのに使用することができる。低密度アッセイシステムは、開示された方法によって選択されたプローブを使用して構築することができる。これらの低密度システムは、少なくとも1、2、3、4、5、6、7、8、9、10、15、20、30、40、50、60、70、80、90、100、200、500、1000、またはそれ以上微生物を同定することができる。同定または定量化される代表的な微生物は、表2に列挙されている。

【 0 1 9 0 】

【表 2】

表 2 低密度アッセイシステムによって認識される代表的な微生物

種	用途
リステリア・モノサイトゲネス(<i>Listeria monocytogenes</i>)	食品加工プラントの食物安全性、環境調査
サルモネラ・エンテリカ(<i>Salmonella enterica</i>)亜種エンテリカ血清型エンテリティディス(<i>Enteritidis</i>)	食品加工プラントの食物安全性、環境調査
緑膿菌 (<i>Pseudomonas aeruginosa</i>)	肺の健康

10

【0191】

低密度アッセイシステムは、多数の環境用途および臨床用途に有用である。例示的な用途は、表2に列挙されている。これらの用途として、糞便または他の汚染の水質試験、動物またはヒト病原体の試験、水汚染源の特定、再生(reclaimed)または再生(recycled)水の試験、海洋排出ブルームを含めた汚水排出ストリームの試験、病原体に関する水産養殖のモニタリング、海浜、水泳領域、または他の水関連レクリエーション施設のモニタリング、および毒性藻類ブルームの予測が挙げられる。他の用途には、試験またはモニタリングの結果に基づいて水管理または治療決定を行うことが含まれる。

20

【0192】

食物モニタリング用途には、食品加工プラントでの生産ラインの定期的検査、食肉処理場の調査、食品媒介病原体、例えば、大腸菌系統0157:H7もしくは0111:B4、リステリア・モノサイトゲネス、またはサルモネラ・エンテリカ亜種エンテリカ血清型エンテリティディスなどについての、レストラン、病院、学校、矯正施設、および他の機関のキッチンおよび食物貯蔵エリアの検査が含まれる。貝および貝生産水は、麻痺性貝毒作用、神経毒性貝毒作用、下痢性貝毒作用、および健忘性貝毒作用に関与する藻類について調査することができる。さらに、輸入食糧は、食糧安全保障を確保するために、解放する前に税関内の間にスクリーニングすることができる。

【0193】

植物病原体モニタリング用途には、園芸および苗床モニタリング、例えば、突然のオーク枯れに関与する微生物である、フィトフトラ・ラモルム(*Phytophthora ramorum*)についてのモニタリング、作物病原体の監視および疾患管理、ならびに森林病原体の監視および疾患管理が含まれる。

30

【0194】

低密度システムを用いて得られるデータに基づいて同定、診断、予知、追跡、または治療することができる医学的状態として、それだけに限らないが、嚢胞性線維症、慢性閉塞性肺疾患、クローン病、過敏性腸症候群、癌、鼻炎、胃潰瘍、大腸炎、アトピー、喘息、新生仔の壊死性全腸炎、肥満症、歯周病、および微生物の存在、非存在、または個体数変化によって引き起こされ、悪化し、またはこれらに関連する任意の疾患または障害が挙げられる。システムに含められるOTUの慎重な選択を通じて、システムは、非常に低い割合の偽陽性または偽陰性の読みを生じる高レベルの信頼度を伴って1つまたは複数の状態または疾患を診断することができる診断デバイスとなる。

40

【0195】

医薬品、医療用デバイス、および微生物の汚染が主要な安全性の懸念事項である他の消耗品または極めて重要なコンポーネントのための製造環境は、緑膿菌、または黄色ブドウ球菌(*Staphylococcus aureus*)のような特定の病原体の存在、ヒトに付随するより一般的な微生物、水の存在に関連する微生物、またはその特定の環境もしくは類似の環境において以前に同定された汚染微生物数を表す他の微生物の存在について調査することができる。

50

【0196】

同様に、宇宙船を含めた繊細な装置の構築および組立領域は、生息することが知られ、またはそのような環境中に最も一般に導入される、以前に同定された微生物についてモニターすることができる。

【0197】

国家安全保障用途には、野兎病菌(*Francisella tularensis*)または炭疽菌(*Bacillus anthracis*)などの既知の生物テロリスト脅威についての、空気、水、および建物のモニタリングが含まれる。他の用途には、怪しい小包または郵便物の検査が含まれる。

【0198】

エネルギー安全保障は、ガスおよび油探索方法の改善を通じて、ならびに微生物増進石油回収(MEOR)によって向上させることができる。油層およびガス層は、メタン、エタン、プロパン、およびブタンを含めた、蓄積された炭化水素の低分子量成分を漏出することが多い。これらの炭化水素は、様々な微生物の食物源として機能を果たすことができる。炭化水素蓄積物の上に重なっている微生物群集を試料採取し、マイクロバイームを、炭化水素を欠いている類似の環境において観察されるマイクロバイームと比較することによって、次に油およびガス探索のための新しい領域を特定するのに使用することができる、インジケータ種を発見することができる。有望な油田中のグリッドアレイから、かつ各炭化水素インジケータ微生物の存在量に基づいて土壌試料を収集することができ、輪郭を示した表面マップを構築し、炭化水素ブルームの位置を正確に概説することができる。

【0199】

ほとんどの従来の石油回収プロセスでは、貯留層中の利用可能な油の15~50%しか回収することができない。三次石油回収は一般に、より高価な方法の抽出技法、例えば、熱回収、ケミカルフラッディング、または埋蔵量の最後の画分を抽出するための混合置換(ガス注入)などを伴う。微生物は、単純で安価な栄養分を使用して、in situでバイオサーファクタントまたはガスを産生することができるので、MEORは、より低コストの三次回収法をもたらす。さらに、ある特定の微生物は、長鎖炭化水素を代謝することによって、くみ出すのがより容易な、より小さい、粘性のより低い炭化水素を作ることができる(バイオクラッキング(biocracking))。油田の全マイクロバイームを測定またはモニターする能力は、より生産的なフィールドと関連する微生物の同定および単離を可能にすることができる。さらに、全マイクロバイーム手法は、マイクロバイームを観察し、油抽出に最適な群集組成を誘導または維持するために栄養分レベルを調整することによって生産を最適化するためのMEORフィールドのモニタリングを可能にする。

【0200】

法医学的科学的には、殺人調査における死亡時刻など、起こった事象を判定するための信頼できるシステムを必要とする。昆虫のコレクションおよび分類が現在使用されているが、微生物集団の変化は、死亡時刻および状況を判定するための別の道を提供することができる。

【0201】

バイオレメディエーションの成功には、所望の種がバイオレメディエーションプロジェクトの開始時に存在し、必須の、または好適な栄養分を、おそらく適時に補充することによって、その数が適切に維持されていることを保証するために、微生物集団の能動的なモニタリングおよび管理を必要とする場合がある。

【0202】

いくつかの実施形態では、低密度システムは、特定の生物において発現される遺伝子または配列に特異的な(相補的な)確認プローブも特徴とする。例えば、ペスト菌(*Yersinia pestis*)のcaf1毒性遺伝子、およびコレラ菌(*Vibrio cholerae*)の閉鎖毒素(zot)遺伝子も、ペスト菌またはコレラ菌に対する確認プローブである。

【0203】

キット

本明細書で使用する場合、「キット」は、本発明の方法を実施するために物質または試

10

20

30

40

50

薬を送達するための任意の送達システムを指す。アッセイとの関連で、そのような送達システムには、プローブを有するアレイまたはビーズ、反応試薬(例えば、適切な容器中のプローブ、酵素など)、および/または支持物質(例えば、アッセイを実施するための使用説明書に書かれた緩衝液など)の、1つの位置から別の位置への貯蔵、輸送、または送達を可能にするシステムが含まれる。例えば、キットは、本発明のアッセイのための妥当な反応試薬および/または支持物質を含む1つまたは複数のエンクロージャー(例えば、箱)を含む。

【0204】

本発明の一態様では、核酸標的を分析するためのキットが提供される。一実施形態によれば、キットは、1つのアッセイで、10、20、50、100、200、500、1,000、2,000、5,000、10,000、20,000、30,000、40,000、50,000、または60,000の異なるOTUに対する存在または量を判定することができる複数のプローブを含む。そのようなプローブは、例えば、アレイ、または複数のマイクロビーズに結合させることができる。いくつかの態様では、キットは、開示された方法を使用して、かつ/または1つもしくは複数の試料のバイオシグネチャーの同定および/もしくは比較において使用するために選択された、少なくとも5、10、15、20、50、100、200、500、1,000、2,000、5,000、10,000、20,000、50,000、100,000、200,000、500,000、1,000,000、または2,000,000のインターロゲーションプローブを含む。

10

【0205】

キットは、試料処理のための試薬も含むことができる。いくつかの実施形態では、試薬は、16S rRNA遺伝子の領域などの高度に保存された配列の領域を増幅するためのプライマーを含めた、試料核酸のPCR増幅のための試薬を含む。さらに他の実施形態では、試薬は、rRNAを直接標識化するための試薬を含む。さらなる実施形態では、キットは、キットを使用するための使用説明書を含む。他の実施形態では、キットは、遠隔のデータ分析およびソフトウェアプログラムの操作に電子的にアクセスするためのパスワードまたは他の許可を含む。そのようなキットは、環境モニタリング、疾患の診断、疾患の進行もしくは治療に対する応答のモニタリング、ならびに汚染源および/または1つもしくは複数の混入物の存在、非存在、もしくは量の特定を含めた、様々な用途を有することになる。

20

【0206】

コンピュータで実行される方法

30

図1は、適当なコンピュータシステム環境または構造の例を例示し、この中で、コンピュータサブシステムは、プローブ選択、試料の分析、およびリモートネットワークキングを含めた、本発明のソフトウェア実施形態を実行するための処理機能性をもたらしことができる。本明細書に開示される方法またはシステムはまた、パーソナルコンピュータ、サーバーコンピュータ、手持ち式またはラップトップデバイス、マルチプロセッサシステムなどを含めた、多数の他の汎用または特殊目的用コンピュータシステムを用いて機能を果たすことができる。

【0207】

本方法またはシステムは、コンピュータによって実行されるプログラムモジュールなどのコンピュータ実行可能な指示の一般的な脈絡で説明することができる。本方法またはシステムはまた、分散コンピューティング環境で実施することができ、この場合、タスクは、通信ネットワークを通じてリンクされている遠隔処理デバイスによって実施される。

40

【0208】

図1を参照すると、本方法またはシステムを実施するための例示的なシステムは、コンピュータ102の形態での汎用コンピュータデバイスを含む。

【0209】

コンピュータ102のコンポーネントには、それだけに限らないが、プロセッシングユニット104、システムメモリー106、およびシステムメモリーからプロセッシングユニット104を含めた、様々なシステムコンポーネントを結合するシステムバス108が含まれる。

【0210】

50

コンピュータ102は一般に、様々なコンピュータ判読可能な媒体を含む。コンピュータ判読可能な媒体には、揮発性および不揮発性媒体、リムーバブルおよびノンリムーバブル媒体の両方が含まれ、これは、コンピュータ記憶媒体を含むことができる。コンピュータ記憶媒体として、それだけに限らないが、RAM、ROM、EEPROM、フラッシュメモリ、または他のメモリ技術、CD-ROM、デジタル多用途ディスク(DVD)、または他の光学ディスク記憶装置、磁気カセット、磁気テープ、磁気ディスク記憶装置、または他の磁気記憶装置デバイスが挙げられる。

【0211】

システムメモリ106は、リードオンリーメモリ(ROM)110およびランダムアクセスメモリ(RAM)112などの揮発性および/または不揮発性メモリの形態でコンピュータ記憶媒体を含む。始動の間などに、コンピュータ102内のエレメント間で情報を移動させるのに役立つ基本ルーチンを含む基本入出力システム114(BIOS)が、ROM 110中に一般に記憶されている。RAM 112は一般に、プロセッシングユニット104に直ちにアクセス可能であり、かつ/またはほとんどなくプロセッシングユニット104によって作動されるデータおよび/またはプログラムモジュールを含む。図1は、オペレーティングシステム132、アプリケーションプログラム134、例えば、配列分析、プローブ選択、シグナル分析、およびクロスハイブリダイゼーション分析プログラムなど、他のプログラムモジュール136、ならびにプログラムデータ138を例示する。

【0212】

コンピュータ102は、他のリムーバブル/ノンリムーバブル、揮発性/不揮発性コンピュータ記憶媒体も含むことができる。例としてのみ、図1は、ノンリムーバブル不揮発性磁気媒体に読み書きするハードディスクドライブ116、リムーバブル不揮発性磁気ディスク120に読み書きする磁気ディスクドライブ118、およびリムーバブル不揮発性光学ディスク124、例えば、CD ROMまたは他の光学媒体などに読み書きする光学ディスクドライブ122を例示する。例示的な動作環境において使用することができる他のリムーバブル/ノンリムーバブル、揮発性/不揮発性コンピュータ記憶媒体として、磁気テープカセット、フラッシュメモリカード、デジタル多用途ディスク、デジタルビデオテープ、固体状態RAM、固体状態ROMなどが挙げられる。ハードディスクドライブ116は一般に、インターフェース126などのノンリムーバブルメモリインターフェースを通じてシステムバス108に接続され、磁気ディスクドライブ118および光学ディスクドライブ122は一般に、インターフェース128または130などのリムーバブルメモリインターフェースによってシステムバス108に接続される。

【0213】

上記に論じられ、図1に例示されたドライブおよびこれらの関連するコンピュータ記憶媒体は、コンピュータ判読可能な指示、データ構造、プログラムモジュール、およびコンピュータ102のための他のデータの記憶装置を提供する。例えば、図1では、ハードディスクドライブ116は、オペレーティングシステム132、アプリケーションプログラム134、他のプログラムモジュール136、およびプログラムデータ138を記憶するものとして例示されている。ユーザーは、入力デバイス、例えば、キーボード140、およびマウス、トラックボール、またはタッチパッド142などを通じてコンピュータ102にコマンドおよび情報を入力することができる。これらの、および他の入力デバイスは、システムバスに結合されたユーザー入力インターフェース144を通じてプロセッシングユニット104に接続されることが多いが、他のインターフェースおよびバス構造、例えば、パラレルポート、またはユニバーサルシリアルバス(USB)などによって接続することができる。モニター158または他のタイプの表示デバイスも、ビデオインターフェースまたはグラフィックス表示インターフェース156などのインターフェースを介してシステムバス108に接続される。モニター158に加えて、コンピュータは、他の周辺出力デバイス、例えば、スピーカー(図示せず)およびプリンター(図示せず)なども含むことができ、これらは、出力周辺インターフェース(図示せず)を通じて接続することができる。

【0214】

コンピュータ102は、本明細書に記載されるマイクロアレイまたは他のプローブシステムなどの分析システム中に組み込むことができる。あるいは、分析システムによって生成されるデータは、当技術分野で公知の様々な手段を使用してコンピュータシステムにインポートすることができる。

【0215】

コンピュータ102は、1つまたは複数のリモートコンピュータまたは分析システムへの論理結合を使用してネットワーク環境下で動作することができる。リモートコンピュータは、パーソナルコンピュータ、サーバ、ルータ、ネットワークPC、ピアデバイス、または他の一般的なネットワークノードであってもよく、一般に、コンピュータ102と比べて、上述したエレメントの多く、またはすべてを含む。図1に表された論理結合は、ローカルエリアネットワーク(LAN)148および広域ネットワーク(WAN)150を含むが、他のネットワークも含むことができる。そのようなネットワーク環境は、オフィス、企業規模のコンピュータネットワーク、イントラネット、およびインターネットにおいてありふれている。LANネットワーク環境下で使用される場合、コンピュータ102は、ネットワークインターフェースまたはアダプター152を通じてLAN 148に接続される。WANネットワーク環境下で使用される場合、コンピュータ102は一般に、モデム154、またはインターネットなどのWAN 150に対して通信を確立するための他の手段を含む。内部にあっても外部にあってもよいモデム154は、ユーザー入力インターフェース144、または他の適切な機構を介してシステムバス108に接続することができる。ネットワーク環境において、コンピュータ102に関連して表されるプログラムモジュール、またはその一部は、遠隔メモリー記憶デバイス中に記憶することができる。

【0216】

本発明のさらなる態様では、1つのアッセイで、20、50、100、200、500、1,000、2,000、5,000、10,000、20,000、30,000、40,000、50,000、または60,000を超える異なるOTUの存在または量を分析するための、コンピュータで実行される方法が提供される。一実施形態では、試料中の1つまたは複数の微生物の存在または量を判定するためのコンピュータ実行可能なロジックであって、少なくとも20、50、100、200、500、1,000、2,000、5,000、10,000、20,000、30,000、40,000、50,000、または60,000のユニークで高度に保存されたポリヌクレオチドのそれぞれに選択的に結合する一連のプローブからの強度を分析し、少なくとも90%、95%、96%、97%、98%、99%、または99.5%の信頼度レベルを伴って、前記試料中に存在するすべての種の少なくとも97%の存在を判定するためのロジックを含むコンピュータ実行可能なロジックが提供される。

【0217】

一実施形態では、一連の異なる生物から、1つまたは複数の生物が試料中に存在する確率を求めるためのコンピュータ実行可能なロジックが提供される。コンピュータロジックは、陰性対照プローブおよび陽性対照プローブの強度との比較に基づいて、個々のインターロゲーションプローブ強度が正確である尤度を求めるためのプロセスまたは指示と、第1の変位値閾値をクリアするOTUからのインターロゲーションプローブの強度に基づいて、個々のOTUが存在する尤度を求めるためのプロセスまたは指示と、第1の変位値閾値をクリアした1つまたは複数のOTUに、やはり第1の変位値閾値をクリアした他のプローブとクロスハイブリダイズするその潜在性に基づいてペナルティを課すためのプロセスまたは指示とを含む。

【0218】

さらなる実施形態では、試料中の1つまたは複数の微生物の存在を判定するためのコンピュータ実行可能なロジックが提供される。このロジックにより、少なくとも1000の異なるインターロゲーション完全プローブの分析が可能になる。このロジックは、判定を行うプロセスにおいて、インターロゲーションパーフェクトマッチプローブの少なくとも10%から情報を廃棄するステップをさらに提供する。いくつかの実施形態では、コンピュータ実行可能なロジックは、コンピュータ判読可能な媒体上に記憶され、コンピュータソフトウェア製品を代表する。

【0219】

他の実施形態では、本発明の態様を具現するコンピュータ実行可能なロジックが、ハードドライブまたは光学ドライブのようなコンピュータ媒体上に記憶されるコンピュータソフトウェア製品が提供される。一実施形態では、コンピュータソフトウェア製品は、実行される際、候補プローブを判定するための本明細書に記載される方法を実施する指示を含む。

【0220】

さらなる実施形態では、本発明の方法を実施することができるコンピュータシステムが提供される。いくつかの実施形態では、コンピュータシステムは、フローサイトメーターまたはマイクロアレイ画像法デバイスのような分析システムに一体化され、このシステムの一部である。他の実施形態では、コンピュータシステムは、分析システムに接続され、または移植されている。いくつかの実施形態では、コンピュータシステムは、ネットワーク接続によって分析システムに接続される。図2は、図1に例示されたコンピュータシステムを利用する遠隔データ取得または分析のためのネットワークシステムの一実施形態を例示する。この実施例では、試料は、市販のイメージングシステムおよびソフトウェアを使用して画像化される。データは、CELファイル(AFFYMETRIX(登録商標))、またはFeature Reportファイル(NIMBLEGEN(登録商標))のような標準的なデータ形式を使用して出力される。次いでデータは、本発明の方法を使用して分析するために、遠隔または中央の位置に送られる。いくつかの実施形態では、標準化された分析が実施され、シグナルの正規化、OTUの定量化、および視覚的分析をもたらす。他の実施形態では、カスタマイズされた分析が、ユーザーの特定の必要性のために設計された、固定されたプロトコルを使用して実施される。さらに他の実施形態では、ユーザーが設定できる分析が使用され、ユーザーが、各分析を実施する前に少なくとも1つの変数を調整することを可能にするプロトコルを含む。

【0221】

処理した後、結果は、後に使用または共有するために、交換可能なバイナリー形式で記憶される。さらに、ハイブリダイゼーションスコアおよびOTU確率値を、検出された試料群集をさらに統計分析するために、タブ区切りファイルに、またはUniFrac(Lozuponeら、UniFrac-an online tool for comparing microbial community diversity in a phylogenetic context、BMC Bioinformatics、7、371頁；2006)と適合した形式でエクスポートすることができる。

【0222】

いくつかの実施形態では、分類学的ツリー、ヒートマップ、階層的クラスタリング、平行座標(時系列)、バープロット、および多次元スケール分散プロットを含めた、データの複数のインタラクティブビュー(interactive view)が利用可能である。いくつかの実施形態では、分類学ツリーは、各検出されたOTUについての平均強度を示し、試料のヒートマップとしてツリーのリーフを示す。ツリーは、ある特定の強度または確率閾値未満のOTUを選別することによって、動的に剪定することができる。さらに、ツリーは、門から亜科までの任意のレベルで要約することができる。他の実施形態では、ユーザーは、組み込まれたC Clustering Library(de Hoonら、Open source clustering software、Bioinformatics、20、1453~1454頁；2004)からの標準的な距離法および連結法のいずれか、ならびに二次ヒートマップウィンドウに示される、得られるデンドログラムを使用して、OTUおよび試料の両方を階層的にクラスター化することができる。いくつかの実施形態では、試料の対での比較を促進するために、示差的なOTU強度のインタラクティブなバープロットを示す第3のウィンドウが提供される。任意の2つの試料について、差異バーの高さは、OTU間の平均強度の絶対的または相対的差異を示す。バーは、同定および比較を容易にするために、任意の分類学的ランクによって、水平軸に沿ってグループ化および分類することができる。同調した選択および選別は、データの複数のビューの間で途切れなくナビゲートするユニークな能力をユーザーに与える。例えば、ユーザーは、階層的クラスタリングウィンドウ中のクラスターを選択し、分類学ツリー中の選択された生物を同時に見ることが

でき、その系統発生的関係および環境の関係を直ちに明らかにする。さらなる実施形態では、分析システムからのデータ、すなわち、分析システムまたはフローサイトメーターは、ハイスループット配列決定データを用いて同時分析および表示することができる。いくつかの実施形態では、試料中に存在すると同定された各生物について、ユーザーは、特定の生物が見出される他の環境のリストを見ることができる。

【0223】

いくつかの実施形態では、スクリーン表示は、動的で同調していることによって、すべての他のビューにおいて同時に反映される、任意のビューに対する変化を用いたOTUの選択または選別を可能にする。さらに、16S rRNA遺伝子、18S rRNA遺伝子、または23S rRNA遺伝子の配列決定によって確認されたOTUは、すべてのビューにおいて同時表示することができる。

10

【0224】

ビジネス方法

本発明のいくつかの態様では、クライアントがアレイを画像化し、または多数の微粒子を走査し、分析のために、データを含むファイルをサービスプロバイダーに送るビジネス方法が提供される。サービスプロバイダーは、データを分析し、金銭的な補償と引き換えにユーザーにレポートを提供する。いくつかの実施形態では、ユーザーは、サービスプロバイダーの分析システムへのアクセスを有し、分析パラメータ、または結果の表示を操作および調整することができる。

【0225】

20

本発明の別の態様では、クライアントが、処理、画像化、または走査されるべき試料を送り、生物の存在または量についてデータが分析されるビジネス方法が提供される。サービスプロバイダーは、金銭的な補償と引き換えにクライアントにレポートを送る。本発明のいくつかの実施形態では、クライアントは、さらなる分析、およびデータの観察のための、一式のデータ分析および表示プログラムへのアクセスを有する。さらなる実施形態では、サービスプロバイダーは最初に、システムまたはキットをクライアントに供給する。キットは、存在するマイクロバイオームの大部分もしくは全部をアッセイするためのシステムを含むことができ、またはシステムは、特定の用途のために設計された「ダウン選択された」プローブを含むことができる。試料を処理および画像化した後、クライアントは、サービスプロバイダーによる分析のためにデータを送る。本発明のいくつかの実施形態では、クライアントのレポートは電子的である。他の実施形態では、クライアントは、データのさらなる観察、操作、比較、および分析のための一式のデータ分析および表示プログラムへのアクセスが提供される。いくつかの実施形態では、クライアントは、結果を比較するための独自のデータベースへのアクセスが提供される。他の実施形態では、クライアントは、結果を比較するために、1つまたは複数の公共データベース、または私的データベースと公共データベースの組合せへのアクセスが提供される。いくつかの実施形態では、独自のデータベースは、正常な試料についてのプールされた結果(フィンガープリント、バイオシグネチャー)、または疾患状態などの特定の異常な状況からのプールされた結果を含む。いくつかの実施形態では、バイオシグネチャーは、新しい試料分析を受け取ると、持続的かつ自動的にアップデートされる。

30

40

【0226】

いくつかの実施形態では、データベースは、高度に保存された配列表をさらに含む。いくつかの実施形態では、データベースは、新しい配列情報が、例えば、国立衛生研究所のHuman Microbiome Projectから利用可能になる際に、自動的にアップデートされる。さらなる実施形態では、プローブは、新しい配列情報に基づいて自動的にアップデートされる。配列情報の持続的なアップグレードおよびプローブセットの洗練により、マイクロバイオームの組成およびその個々の成分の量の判定において精度および分解能を増大させることが可能なる。いくつかの実施形態では、システムは、より早いマイクロバイオームバイオシグネチャーを、同じまたは実質的に類似の環境からのより遅いマイクロバイオームバイオシグネチャーと比較し、プローブセットの組成、ならびに関連OTU間の識別、マイク

50

ロバイオーム成分の同定および定量化を改善もしくは洗練し、または判定の精度を増大させるのに有用な情報についてのハイブリダイゼーションシグナル分析パラメータの変化を分析する。

【0227】

いくつかの実施形態では、データベースは、宿主の年齢、性別、および全体的な健康状態、宿主の地理的位置、宿主の飲食物(すなわち、西洋、アジア、または菜食主義者)、水源、宿主の職業または社会的状態、宿主の住宅状態を含めた、特定のマイクロバイオーム、例えば、健康な、および不健康なヒトの腸のマイクロフローラに関連する微生物叢についての情報をまとめる。

【0228】

いくつかの実施形態では、男性および女性の成人、ならびに小児についての参照の健康/正常なシグネチャーをベンチマークとして使用することによって、前駆症状および症候性の疾患状態、治療/療法に対する応答、感染、および/または疾患に関連する二次感染を特定することができる。

【0229】

いくつかの実施形態では、クライアントは、クライアントの試料マイクロバイオームと、1つまたは複数の参照マイクロバイオームとの比較に基づいて診断または治療勧告が与えられる。

【0230】

いくつかの実施形態では、データベースは、日常の食品加工プラントまたは食肉処理場の微生物検査からの集合した結果で維持される。日常または緊急の検査からの1つまたは複数の試料に由来するマイクロバイオームのフィンガープリントが、「清潔なプラント」、「汚れたプラント」、または特定の微生物の汚染問題を経験したことが知られているプラントの複合フィンガープリントと比較される。次いで比較結果は、提出実体(submitting entity)に送られる。

【0231】

他の実施形態では、養魚場は、植物プランクトンの存在量または毒性藻類ブルームの非存在の予測に基づいて管理され、そのような予測は、養魚場の現在のフィンガープリントを、十分に管理された養魚場、衰退しているか、または毒性藻類ブルームの出現が知られている養魚場の複合フィンガープリントと比較することから得られる。他の実施形態では、水産養殖設備は、マイクロバイオームのフィンガープリントを、健康な水産養殖設備のフィンガープリント、および同定された、または疑いのある病原体のアウトブレイクの間の水産養殖設備のフィンガープリントのデータベースと比較することによってモニターまたは管理される。

【0232】

さらに他の実施形態では、分水界からの水試料のマイクロバイオームが、分水界全体から集められたデータと比較されることによって、水質を最適にし、魚の集団を支持し、毒性藻類ブルームまたはデッドゾーンを最小限にする管理およびレメディエーションの実行が通知される。いくつかの実施形態では、水の検査は、汚染を低減し、規制上の基準を満たすことにおける処理施設の有効性を判定するために、処理施設を建設する前および後に実施される。さらに他の実施形態では、試料採取プログラムが設定され、このプログラムにおいて、試料が定期的に分析され、レクリエーション用水または家庭内消費のために使用される水源中で、微生物レベルがある特定の閾値を超える場合、自動警告システムが、地方、州、または連邦の機関に通報する。

【0233】

集合したフィンガープリントコレクションのさらなる例には、紙パルプ工場、石油精製所、製革所、製糖所、化学プラントを含めた製造、加工、または精錬所施設からの工業流出液体および流出物、ならびに糞便汚染のバイオシグネチャーが含まれる。

【0234】

(実施例)

10

20

30

40

50

以下の実施例は、本発明の様々な実施形態を説明する目的のものであり、いかなる形であれ本発明を限定することを意図しない。本実施例は、本明細書に記載される方法と併せて、好ましい実施形態を現在表しており、例となるものであり、本発明の範囲を限定するものとして意図しない。特許請求の範囲により限定される本発明の精神内に包含されるその変化および他の使用は、当業者であれば想到されよう。

【 0 2 3 5 】

(実施例1)

PhyloChipアレイ分析

標準技法を使用した試料調製、適用、インキュベーションおよび洗浄に続いて、PhyloChip G3アレイはAffymetrix製のGeneArray Scannerを使用して走査された。前記走査は、データをプローブごとにテキストコードされた表の個々の列に変える標準AFFYMETRIX(登録商標)ソフトウェア(GCOS v1.6 using parameter: Percentile v6.0)を使用してピクセル画像として取り込まれた。表3を参照されたい。

【 0 2 3 6 】

【表 3】

表3 アレイデータの例となる表示

[強度]			
細胞数=506944			
4 セルヘッダー=XY	平均	STDV	
NPIXELS			
0	0	167.0	
47.9	25		
1	0	4293.0	
1060.2	25		
2	0	179.3	
43.7	36		
3	0	4437.0	
681.5	25		

【 0 2 3 7 】

各分析システムは、細胞あたり1プローブ配列で約1,016,000細胞を有していた。前記分析システムスキャナは、細胞あたり30～45ピクセルの規則的格子内で0～65,000任意単位(a.u)の幅があるアレイ全体でシグナル強度を記録した。隣接する細胞間に2ピクセルマージンを使用し、使用可能なシグナルのプローブあたり約25～40ピクセルを残した。これらのピクセルから、AFFYMETRIX(登録商標)ソフトウェアは75パーセンタイル平均ピクセル強度(「平均」と示される)、約25～40ピクセル間のシグナル強度の標準偏差(「STDV」と示される)および細胞あたりの使用されるピクセル数(「NPIXELS」と示される)を算出した。シグナル強度が3標準偏差離れているピクセルを有する細胞はどれも異常値として分類された。

【 0 2 3 8 】

分析システムは、使用者定義数の水平割りと縦割りに分割された。デフォルト設定では、4つの水平割りと4つの縦割りが作成され、独立したバックグラウンド減算のための16の規則的に間隔が開けられたセクターを生じた。バックグラウンド強度は象限ごとに独立して、その象限のプローブの最小強度2%(デフォルト設定では)の平均シグナル強度として算出された。次に、バックグラウンド強度はさらなる算出前にすべてのプローブから減算された。

【 0 2 3 9 】

ノイズ値は、AFFYMETRIX(登録商標)GeneChip User Guide v3.3の推奨に従って評価された。ノイズ(N)は、スキャナがアレイ表面を読み取るときに観察されるピクセル強度シグ

ナルの変動のせいであり、その細胞を含むピクセルの数の平方根で割った同定されたそれぞれのバックグラウンド細胞内のピクセル強度の標準偏差として計算された。こうして得られた商の平均は下に記載される計算におけるNに使用された。

【 0 2 4 0 】

【数 4】

$$N = \frac{\sum_{i \in B} \frac{S_i}{\sqrt{\text{pix}_i}}}{\text{スカラー-B}}$$

10

【 0 2 4 1 】

Bはバックグラウンド細胞

S_i はBにおけるピクセル間の標準偏差 pix_i はBにおけるピクセルの総数

スカラー-Bはすべてのバックグラウンド細胞の総数、累積的

【 0 2 4 2 】

次に、すべてのプローブの強度は、スパイクインプローブの平均の観察されたシグナル強度が既定のシグナル強度を有するようにスケーリングされた。これは、対応するPMプローブの平均応答を下の方程式を使用して標的平均に強制するための倍率(S_f)を見つけることにより実現された。

【 0 2 4 3 】

20

【数 5】

$$S_f = \bar{e}_i / \frac{\sum_{i \in K_{pm}} e_i}{\text{スカラー-K}_{pm}}$$

【 0 2 4 4 】

\bar{e}_i =標的平均強度(デフォルト:2500)

スカラー- K_{pm} =どんなスパイクインも補完するプローブの総数

S_f =倍率

30

【 0 2 4 5 】

典型的には、既定のシグナル強度は約0～約65,000の範囲であった。倍率が導き出されると、すべての細胞強度に倍率を掛けあわせた。

【 0 2 4 6 】

ノイズ(N)は同じ係数によりスケールを変更され: $N_s = N \times S_f$ で、 N_s =スケーリングされたノイズ、 N =スケーリングされていないノイズ、および S_f =倍率であった。

【 0 2 4 7 】

代わりのまたは選択のステップとして、高ハイブリダイゼーションシグナル応答のMMプローブが同定され、プローブペアは除去された。

【 0 2 4 8 】

40

【数 6】

$$\left[\left(\frac{MM}{PM} > srt_r \right) \wedge (MM - PM > N_s \times sdtm_r) \right] \vee [PM \in O] \vee [MM \in O]$$

【 0 2 4 9 】

PM=パーフェクトマッチプローブのスケーリングされた強度

MM=ミスマッチプローブのスケーリングされた強度

srt_r =逆標準比率閾値(デフォルト:1.3)

$sdtm_r$ =逆標準差閾値乗数(デフォルト:130)

50

Ns=スケーリングされたノイズ

0=異常値セット

【 0 2 5 0 】

残りのプローブペアは、

【 0 2 5 1 】

【 数 7 】

$$\left(\frac{PM}{MM} > srt \right) \wedge (PM - MM > Ns^2 \times sdtm)$$

10

【 0 2 5 2 】

PM=パーフェクトマッチプローブのスケーリングされた強度

MM=ミスマッチプローブのスケーリングされた強度

srt=標準率閾値(デフォルト:1.3)

sdtm=標準差閾値乗数(デフォルト:130)

Ns=スケーリングされたノイズ

によりスコア化された。

【 0 2 5 3 】

OTUを「プレゼント(present)」として分類した後、プレゼントコールは、その下位のOTUのうちの少なくとも1つがプレゼントであれば、どんなノード(亜科、科、目など)も「プレゼント」と見なすことにより分類階層を通して上方に伝搬させた。

20

【 0 2 5 4 】

ハイブリダイゼーション強度はOTU存在量の尺度であり、所定のプローブセットにおけるプローブペアにわたるPMマイナスMM強度差のトリム平均(平均化する前に最大値と最小値が取り除かれる)としてプローブセットごとに任意単位で計算された。

【 0 2 5 5 】

(実施例2)

水質検査 糞便汚染アッセイ

上昇した糞便インジケータ細菌濃度およびヒト糞便汚染に関連する場所である、カリフォルニア州サンタバーバラの低地ミッションクリークとラグナ流域の日照り流水から本発明のアレイを用いてサンプリングすることになる。前記目的は、都市化されたクリークにおける全細菌群集組成および生物地理学的パターンを特徴付け、2)分子的方法により検出される分類群を従来の糞便インジケータ細菌と比較し、3)培養非依存群集ベースの糞便汚染モニタリングにおいて使用される細菌分類群の信頼性のある群(糞便汚染のインジケータ種)を解明することである。

30

【 0 2 5 6 】

前記流水域はサンタバーバラ都心部の都市化地域を流れている。サンプリングする場所には、雨水排水管、流れるクリークの区間、渦(M2、M4)および海が挙げられる。さらに、地点には旧ミッションクリーク支流がミッションクリークに注ぎ込む場所が含まれる。乾いたクリーク流は、上流領域における地下泉、灌漑と水洗に関連する都市流出水、地下水浸出、排水だめまたは地下ポンプ、および潜在的に違法な地下水接続を含む多くの発生源を有する。サンプリングは、サンプリング前の少なくとも48時間またはサンプリング中に雨が降っていない期間中に行われることになる。流域試料に加えて、ヒト糞便および下水がサンプリングされることになる。

40

【 0 2 5 7 】

材料と方法

試料説明、収集および抽出

水試料は日照り期間中の流水域から3~5日にわたり収集される。さらに、ヒト糞便下水流入を含む糞便試料が収集される。溶存酸素(DO)、pH、温度および塩分が各サンプリングとともに測定される。水試料は実験室において0.22µmフィルターで濾過し、UltraClean W

50

ater DNA kit (MoBio Laboratories)を使用してDNAを抽出し、-20 で保存する。全大腸菌群、大腸菌(*E. coli*)およびエンテロコッカス種(*Enterococcus* spp.)の濃度(IDEXXによる)の他にもヒト特異的バクテロイデス(*Bacteroides*)マーカー(HBM)の定量的PCR(qPCR)測定も実施される。

【0258】

分析システム解析のための16S rRNA遺伝子増幅

16S rDNAは、非変性細菌プライマー27F.jgiおよび1492Rを使用してgDNAから増幅される。ポリメラーゼ連鎖反応(PCR)は、TaKaRa Ex Taqシステム(Takara Bio Inc.、Japan)を使用して実施される。増幅プロトコールはすでに記載されている(Brodieら、Application of a High Density Oligonucleotide Analysis system Approach to Study Bacterial Population Dynamics during Uranium Reduction and Reoxidation. *Applied Environ Microbio*. 72:6288 ~ 6298頁、2006)。

10

【0259】

分析システム処理および画像データ解析

分析システム解析は、高密度系統発生分析システム(PhyloChip)を使用して実施される。そのプロトコールはすでに記載されている(Brodieら、2006)。手短に言えば、単位複製配列はイソプロパノール沈殿により40 µl未満の容積まで濃縮される。前記DNA単位複製配列は、デオキシリボヌクレアーゼを用いて断片化され、ビオチン標識され、変性され、48で一晩(>16時間)DNA分析システムにハイブリダイズされる。アレイは続いて洗浄され染色される。アレイはGeneArray Scanner (Affymetrix、Santa Clara、CA、USA)を使用して走査される。各プローブ(PM、MMおよび対照プローブ)の蛍光強度に関する情報を生み出すAffymetrixソフトウェアから得られるCELファイルは、Todd DeSantisにより設計されたCEL解析ソフトウェア(LBNL、Berkeley、USA)を使用して解析される。

20

【0260】

PhyloChipデータ正規化

すべての統計的解析はR(Team RCD(2008) R: A language and environment for statistical computing)において実施される。単位複製配列標的の定量化に伴う変動(定量化変動)ならびに標的断片化、標識化、ハイブリダイゼーション、洗浄、染色およびスキニングに伴う下流変動(分析システム技術変動)について補正するために、二段階正規化手法が開発されている。すなわち、まずPhyloChip実験ごとに、乗法誤差モデル下でスパイク対照プローブの強度をもっともよく説明する倍率が、最尤手順を使用して見積もられる。各実験における強度にはそれに対応する最適倍率が掛けられる。さらに、実験ごとの強度は、細菌と古細菌については別々に強度をそれに対応する全アレイ強度で割ることにより、全アレイ強度の変動について補正される。

30

【0261】

統計的解析

すべての統計的解析はRにおいて実施された。ブレイカーチス距離(Bray-Curtis distances)は、ecodistパッケージにおけるbcdist関数を用いて正規化蛍光強度を使用して計算された(Goslee SC & Urban DL (2007) The ecodist package for dissimilarity-based analysis of ecological data. *J Stat Softw* 22(7): 1 ~ 19頁)。群集データのブレイカーチス距離行列のマンテル相関、地理的距離および環境変数は、veganパッケージにおけるマンテル関数を使用して計算される。ピアソンの相関は、モンテカルロ(無作為化)テストの1000順列を用いて計算される。非計量的多次元尺度法(NMDS)はveganパッケージのmetaMDS関数を使用して実施される。緩和近接結合系統樹はClearcutを使用して作成される(Evans J、Sheneman L、& Foster JA (2006) Relaxed neighbor-joining: a fast distance-based phylogenetic tree. Construction method. *J Mol Evol* 62:785 ~ 792頁)。別個のクリアカット系統樹は、地点ごとに「常在」および「一過性」群集について作製される。重みなしUniFrac距離(Lozupone C & Knight R (2005) UniFrac: a new phylogenetic method for comparing microbial communities. *Applied and Environmental Microbiology* 71(12):8228 ~ 8235頁)は地点ごとに計算される。

40

50

【0262】

PhyloChip由来パラメータ

糞便分類群

3つの糞便試料すべてにおよび27の水試料すべてに存在する分類群は別々に集計される。「糞便分類群」の一覧表は、3つの糞便試料すべてに存在する分類群から水試料すべてに見出される分類群を取り除くことにより導き出される。

【0263】

一時的および常在亜集団

サンプリング期間にわたり各地点由来の少なくとも1つの試料に存在する分類群が集計され、その分類群の蛍光強度の分散が作製される。上位十分位数における分類群は「一時的」亜集団として定義され、下位十分位数における分類群は「常在」亜集団として定義された。

10

【0264】

BBC:A

バチルス(Bacilli)、バクテロイデス(Bacteroidetes)、クロストリジウム(Clostridia)および -プロテオバクテリア(-proteobacteria)の綱における分類群の数が集計される。その比率は以下の式、

【0265】

【数8】

$$BBC:A = \frac{Bac + Bct + Cls}{A}$$

20

【0266】

を使用して、計算される。

【0267】

前記綱それぞれのユニークな分類群の総数は、分析システムにより検出される各綱における全分類群で割ることにより正規化される。

【0268】

公表された研究由来の整列された配列は、Greengenesからダウンロードされ(DeSantis TZら、(2006) Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. Applied and Environmental Microbiology 72(7):5069~5072頁)、PhyloChip分類法を使用して再分類される。ユニークな分類群の総数は細菌綱ごとに集計される。BBC:Aは上記の式を使用して計算される。1つの綱で分類群が検出されない場合は、その綱の総数は0.5に設定される。

30

【0269】

生息地間の群集の違いを解決する

ミッションクリーク試料は3つの生息地類型、すなわち、海洋、河口湾ならびに淡水(クリークおよび雨水排水管路流水)に描写されている。流域試料と3つの糞便試料(2つの下水と1つのヒト糞便)のプレイカーチス距離が計算される。最初の2つの軸の非計量的多次元尺度法(NMDS)序列およびプロットングを使用して、試料間の距離を表示する。細菌群集は生息地類型により明確に分離される。排水管路試料は糞便試料にもっとも類似している。潟試料は海洋試料にもっとも類似している。

40

【0270】

生息地間で観察される細菌群集の違いの大多数を占めるシグネチャー分類群は、すべての生息地類型間の綱レベルでの検出された分類群を比較することにより同定される。各生息地類型における分類群の数は、パーセント検出を得るために試料類型ごとに検出された合計で割られる。糞便試料を市街地の上流で採取された試料または潟もしくは海洋由来の試料と比較することにより、低い割合の -プロテオバクテリアと高い割合のバチルスおよびクロストリジウムが存在することが明らかにされている。さらに、糞便試料中では5つ

50

の綱、ソリバクター(*Solibacteres*)、未分類アシドバクテリウム(*Acidobacteria*)、クロロフレクサス-4(*Chloroflexi-4*)、コプロサーモバクテリウム(*Coprothermobacteria*)およびフソバクテリウム(*Fusobacteria*)のみが検出されている。クロロフレクサス綱-3(*Chloroflexi-3*)はクリーク試料のみで検出されており、サーモミクロビウム(*Thermomicrobia*)、未分類ターマイトグループ1(*Termite group 1*)および未分類クロロフレクサスは海洋試料のみで検出されている。4つの生息地にわたりもっとも高い標準偏差を有する上位10綱は(降順に)、クロストリジウム、 γ -プロテオバクテリア、バチルス、 α -プロテオバクテリア(α -proteobacteria)、 β -プロテオバクテリア(β -proteobacteria)、放線菌(*Actinobacteria*)、フラボバクテリア(*Flavobacteria*)、バクテロイデス、シアノバクテリア(*Cyanobacteria*)、および ϵ -プロテオバクテリア(ϵ -proteobacteria)である。これらの綱のうち、クロストリジウム、バチルスおよびバクテロイデス割合は高いが、 α -プロテオバクテリア割合は低かった。これらの4つの分類群は、糞便汚染のインジケータとして使用することが可能である。

【0271】

「一過性」と「常在性」亜集団

サンプリング間でもっとも変動する分類群の亜集団が同定される。これらは「一過性」集団と呼ばれる。サンプリング期間安定したままである集団は「常在性」集団と呼ばれる。「一過性」と「常在性」亜集団において見られる分類群の比較により、地点による群集組成の違いが説明される。糞便分類群の6つの主要目(腸内細菌(*Enterobacteriales*)、ラクトバチルス(*Lactobacillales*)、アクチノミセス(*Actinomycetales*)、バクテロイデス(*Bacteroidales*)、クロストリジウム(*Clostridiales*)およびバチルス(*Bacillales*))は、時間をかけて糞便細菌の分布をさらに細かく分けるために比較される。いくつかの地点由来の試料中の一過性腸内細菌目の数は、その他の地点と比べると極端に高いが、バチルスの高い常在性亜集団を有する地点もある。偏在性であり、PhyloChipにより測定されるような、測定される環境変数の変化に影響されない細菌が同定されている。流域および糞便試料全体で類似する数の分類群を有する細菌綱には、ヴェルコミクロビア(*Verrucomicrobiae*)、プランクトミケス(*Planctomycetacia*)、 γ -プロテオバクテリア、アナエロリネア(*Anaerolineae*)、アシドバクテリウム(*Acidobacteria*)、スフンゴバクテリウム(*Sphingobacteria*)およびスピロヘータ(*Spirochaetes*)が挙げられた。

【0272】

バチルス、バクテロイデスおよびクロストリジウム対 α -プロテオバクテリアの比

4つの細菌綱、バチルス、バクテロイデス、クロストリジウムおよび β -プロテオバクテリアは、生息地類型間にもっとも高い分散を有するとして同定され、糞便インジケータとしてさらに展開される。

【0273】

バチルス、バクテロイデスおよびクロストリジウムの組合せ比率は、糞便試料中で検出された全綱の約20~35%に相当し、クリーク、潟および海洋などの予想されるもっとも清浄な水を有する地点でのその比率は10~15%未満である。クリーク水、潟および海洋試料において検出される分類群の少なくとも45%は β -プロテオバクテリアである。これらの微生物は、糞便試料中に見出されるプロテオバクテリアの比率が約35~45%で著しく低いので、浄水分類群(*Clean Water Taxa*)(表11)として分類された。糞便試料での β -プロテオバクテリアに対するバチルス、バクテロイデスおよびクロストリジウムの比(BBC:A)は、他の生息地類型において見出される比の約3~5倍の高さである。BBC:A比は地点ごとに計算され、すべての地点にわたり糞便分類群総数と同じパターンを示し、海洋水は約0.75~0.90のもっとも低いBBC:Aを有し、糞便汚染の観察地点に近い試料は約1.50~約1.90であった。

【0274】

この比は、非大腸菌型関連細菌を含有し、環境内の大腸菌型の増殖に起因する偽陽性糞便検出の可能性を回避する。バクテロイデスおよびクロストリジウムは周知の糞便関連嫌気性細菌である。バチルスは基本的には糞便関連ではないが、好気性好熱性ブタ排水バイ

10

20

30

40

50

オリアクターにおいて見出されている(Juteau P、Tremblay D、Villemur R、Bisaillon J G、& Beaudet R (2005) Analysis of the bacterial community inhabiting an aerobic thermophilic sequencing batch reactor (AT-SBR) treating swine waste *Applied Microbiology and Biotechnology* 66:115~122頁)。したがって、バチルス、バクテロイデスおよびクロストリジウムの存在は、排水汚染、廃棄物処理汚染およびヒト由来糞便汚染の優れた兆候である。-プロテオバクテリアは、環境において豊富な大部分が光合成細菌であり、地球規模の炭素、硫黄および窒素循環において重要な役割を果たしている。多くの-プロテオバクテリアは低栄養状態で繁殖し、非汚染水質環境において見出される非糞便細菌の良好な代理となるであろう。

【0275】

10

その結果は、マウス盲腸、ウシ結腸、下水汚染された地下水、ヒト結腸および二次下水を用いてPhyloChipにより分析される他の糞便関連発生源において見出されるBBC:Aによく匹敵する。これらの発生源は約1.2のBBC:Aを有している。これとは対照的に、嫌気性下水は0.80~0.99のBBC:Aを有する。

【0276】

糞便汚染を検出するためのBBC:A比の値を確認するために、シーケンシングにより得られた細菌群集の公表された研究が解析される。哺乳動物の腸、嫌気性消化槽汚泥、海洋、南極湖の氷および飲料水由来の比は、糞便試料と非糞便試料間には差が存在することも示している。哺乳動物の腸試料は約10~約260の範囲のBBC:Aを有する。嫌気性消化槽汚泥試料は少なくとも1~約10のBBC:Aを有する。これらの結果は、下水処理における嫌氣的に消化された廃棄物活性汚泥中の高度に選択された群集を反映しているのであろう。非糞便試料は、0~0.94のBBC:Aを有する。シーケンシング結果は、1.0のBBC:A閾値は、汚染水であることを示す1およびそれ以上の値を有する水における糞便汚染を同定するためのカットオフとして使用することが可能である。BBC:A値を計算する方法は、培養が必要ではないので、速さ、現在培養できない微生物を検出することが可能であり、PCRクローニングおよび高スループットシーケンシングに伴う経費と技術問題を回避するより大きな検出能力を含む数々の利点を提供する。

20

【0277】

その数は通常、糞便汚染の発生源により近い地点から得られる試料においては増加するために、前記BBC:A比を使用して糞便汚染の発生源を追うことが可能である。

30

【0278】

(実施例3)

糞便試料関連分類群

3つの糞便試料(サンタバーバラのヒト糞便、およびカリフォルニア州サンタバーバラ、El Estero下水処理施設の流入水由来の2つの生水)が収集された。9つの場所由来の水ラム試料も、カリフォルニア州サンタバーバラ郡のミッションクリークおよびラグナ流域内で収集された。3つの糞便試料すべてにおいて、および27の水試料すべてにおいて、PhyloChipアッセイを使用する分析により示されるように、分類群が存在していた。その結果は別個に集計された。

【0279】

40

503の分類群の一覧表は表4に示されており、3つの糞便試料すべてに存在していた分類群から27の水試料すべてにおいて見出された分類群を取り除くことにより導かれた。これらの503の分類群は、分析されたヒト糞便と下水試料に共通であるが、バックグラウンド環境には見出されない細菌を潜在的に表すことができるであろう。糞便関連亜集団に対する全細菌群集組成の類似性は、糞便汚染の兆候として有用である。

【0280】

【表 4 A】

表 4 糞便分類群

細菌界;OD1;OP11-5;未分類;未分類;sf_1;515	
細菌界;NC10;NC10-1;未分類;未分類;sf_1;452	
細菌界;アシドバクテリウム門;アシドバクテリウム綱-6;未分類;未分類;sf_1;897	
細菌界;バクテロイデス門;バクテロイデス綱;バクテロイデス目;未分類;sf_15;6233	
細菌界;バクテロイデス門;バクテロイデス綱;バクテロイデス目;プレボテラ科 (Prevotellaceae);sf_1;6011	
細菌界;バクテロイデス門;バクテロイデス綱;バクテロイデス目;ポルフィロモナス科 (Porphyromonadaceae);sf_1;5460	10
細菌界;バクテロイデス門;バクテロイデス綱;バクテロイデス目;プレボテラ科;sf_1;6047	
細菌界;バクテロイデス門;フラボバクテリア綱(Flavobacteria);フラボバクテリア目 (Flavobacteriales);フラボバクテリア科(Flavobacteriaceae);sf_1;5942	
細菌界;バクテロイデス門;フラボバクテリア綱;フラボバクテリア目;フラボバクテリア科 ;sf_1;5589	
細菌界;バクテロイデス門;未分類;未分類;未分類;sf_4;5703	
細菌界;バクテロイデス門;スフンゴバクテリア綱(Sphingobacteria);スフンゴバクテリア目 (Sphingobacteriales);スフンゴバクテリア科(Sphingobacteriaceae);sf_1;5459	
細菌界;バクテロイデス門;スフンゴバクテリア綱;スフンゴバクテリア目;スフンゴバクテ リア科;sf_1;5492	20
細菌界;バクテロイデス門;バクテロイデス綱;バクテロイデス目;バクテロイデス科 (Bacteroidaceae);sf_6;5792	
細菌界;バクテロイデス門;スフンゴバクテリア綱;スフンゴバクテリア目;毛状菌科 (Crenotrichaceae);sf_11;5619	
細菌界;バクテロイデス門;スフンゴバクテリア綱;スフンゴバクテリア目;毛状菌科 ;sf_11;6123	
細菌界;バクテロイデス門;スフンゴバクテリア綱;スフンゴバクテリア目;フレキシバクテ ー科(Flexibacteraceae);sf_19;5667	30
細菌界;クラミジア門(Chlamydiae);クラミジア綱;クラミジア目(Chlamydiales);クラミジア科 (Chlamydiaceae);sf_1;4820	
細菌界;シアノバクテリア門(Cyanobacteria);シアノバクテリア綱;クロロプラスト目 (Chloroplasts);クロロプラスト科;sf_11;5123	
細菌界;海洋グループ A;mgA-1;未分類;未分類;sf_1;6408	
細菌界;スピロヘータ門(Spirochaetes);スピロヘータ綱;スピロヘータ目(Spirochaetales);スピ ロヘータ科(Spirochaetaceae);sf_1;6502	
細菌界;スピロヘータ門;スピロヘータ綱;スピロヘータ目;スピロヘータ科;sf_1;6494	
細菌界;スピロヘータ門;スピロヘータ綱;スピロヘータ目;スピロヘータ科;sf_1;6583	40
細菌界;スピロヘータ門;スピロヘータ綱;スピロヘータ目;スピロヘータ科;sf_1;6476	
細菌界;スピロヘータ門;スピロヘータ綱;スピロヘータ目;スピロヘータ科;sf_1;6490	
細菌界;スピロヘータ門;スピロヘータ綱;スピロヘータ目;スピロヘータ科;sf_1;6506	
細菌界;スピロヘータ門;スピロヘータ綱;スピロヘータ目;スピロヘータ科;sf_1;6571	

【表 4 B】

細菌界;プロテオバクテリア門(Proteobacteria);アルファプロテオバクテリア綱(Alphaproteobacteria);アセトバクター目(Acetobacterales);ロゼコッカス科(Roseococcaceae);sf_1;7500	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;アセトバクター目;アセトバクター科(Acetobacteraceae);sf_1;7600	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;アゾスピリラム目(Azospirillales);アゾスピリラム科(Azospirillaceae);sf_1;6959	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;未分類;未分類;sf_6;7312	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;未分類;未分類;sf_2;6697	10
細菌界;プロテオバクテリア門;ベータプロテオバクテリア綱(Betaproteobacteria);ナイセリア目(Neisseriales);ナイセリア科(Neisseriaceae);sf_1;7675	
細菌界;プロテオバクテリア門;ベータプロテオバクテリア綱;MND1 クローニンググループ;未分類;sf_1;7808	
細菌界;プロテオバクテリア門;ベータプロテオバクテリア綱;メチロフィルス目(Methylophilales);メチロフィルス科(Methylophilaceae);sf_1;8137	
細菌界;プロテオバクテリア門;ベータプロテオバクテリア綱;ロドシクルス目(Rhodocyclales);ロドシクルス科(Rhodocyclaceae);sf_1;7817	
細菌界;プロテオバクテリア門;ベータプロテオバクテリア綱;未分類;未分類;sf_3;8036	
細菌界;プロテオバクテリア門;ベータプロテオバクテリア綱;バークホルデリア目(Burkholderiales);アルカリゲネス科(Alcaligenaceae);sf_1;7768	20
細菌界;プロテオバクテリア門;ベータプロテオバクテリア綱;バークホルデリア目;コマモナス科(Comamonadaceae);sf_1;7942	
細菌界;プロテオバクテリア門;ベータプロテオバクテリア綱;バークホルデリア目;コマモナス科;sf_1;7847	
細菌界;プロテオバクテリア門;ベータプロテオバクテリア綱;バークホルデリア目;コマモナス科;sf_1;7941	
細菌界;プロテオバクテリア門;ベータプロテオバクテリア綱;未分類;未分類;sf_3;8045	
細菌界;プロテオバクテリア門;ベータプロテオバクテリア綱;バークホルデリア目;コマモナス科;sf_1;7745	30
細菌界;プロテオバクテリア門;ベータプロテオバクテリア綱;バークホルデリア目;ラルストニア科(Ralstoniaceae);sf_1;7778	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱(Gammaproteobacteria);GAO クラスター;未分類;sf_1;9059	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;チオトリクス目(Thiotrichales);チオトリクス科(Thiotrichaceae);sf_3;8741	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;ウラン廃棄物クローン(uraniumwasteclones);未分類;sf_1;8231	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;オセアノスピリルム目(Oceanospirillales);オセアノスピリルム科(Oceanospirillaceae);sf_1;8596	40
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;レジオネラ目(Legionellales);コキシエラ科(Coxiellaceae);sf_3;9444	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;オセアノスピリルム目;アルカニボラクス科(Alcanivoraceae);sf_1;9658	

【表 4 C】

細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;シュードモナス目(Pseudomonadales);シュードモナス科(Pseudomonadaceae);sf_1;8601	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;未分類;未分類;sf_3;8959	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;アルテロモナス目(Alteromonadales);アルテロモナス科(Alteromonadaceae);sf_1;9486	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;アルテロモナス目;アルテロモナス科;sf_1;8863	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;アルテロモナス目;アルテロモナス科;sf_1;9501	10
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;アルテロモナス目;シュワネラ科(Shewanellaceae);sf_1;8581	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;パスツレラ目(Pasteurellales);パスツレラ科(Pasteurellaceae);sf_1;9237	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;エンテロバクター目(Enterobacteriales);エンテロバクター科(Enterobacteriaceae);sf_1;8554	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;エンテロバクター目;エンテロバクター科;sf_1;8885	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;エンテロバクター目;エンテロバクター科;sf_1;8700	20
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;エンテロバクター目;エンテロバクター科;sf_1;8529	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;エンテロバクター目;エンテロバクター科;sf_1;8770	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;エンテロバクター目;エンテロバクター科;sf_1;8225	
細菌界;未分類;未分類;未分類;未分類;sf_160;10012	
細菌界;プロテオバクテリア門;デルタプロテオバクテリア綱(Deltaproteobacteria);デスルフォビブリオ目(Desulfovibrionales);デスルフォビブリオ科(Desulfovibrionaceae);sf_1;10189	30
細菌界;プロテオバクテリア門;デルタプロテオバクテリア綱;シントロフォバクター目(Syntrophobacteriales);シントロファス科(Syntrophaceae);sf_3;9665	
細菌界;プロテオバクテリア門;イプシロンプロテオバクテリア綱(Epsilonproteobacteria);カンピロバクター目(Campylobacteriales);ヘリコバクター科(Helicobacteraceae);sf_23;10443	
細菌界;プロテオバクテリア門;イプシロンプロテオバクテリア綱;カンピロバクター目;ヘリコバクター科;sf_3;10576	
細菌界;プロテオバクテリア門;イプシロンプロテオバクテリア綱;カンピロバクター目;未分類;sf_1;10407	
細菌界;ゲマティモナス門(Gemmatimonadetes);未分類;未分類;未分類;sf_5;317	
細菌界;アクチノバクテリア門(Actinobacteria);アクチノバクテリア綱;ルプロバクター目(Rubrobacteriales);ルプロバクター科(Rubrobacteraceae);sf_1;1551	40
細菌界;アクチノバクテリア門;アクチノバクテリア綱;アシドミクロビウム目(Acidimicrobiales);未分類;sf_1;1666	
細菌界;アクチノバクテリア門;BD2-10 グループ;未分類;未分類;sf_2;1652	

【表 4 D】

細菌界;アクチノバクテリア門;アクチノバクテリア綱;アクチノミセス目(Actinomycetales);未分類;sf_3;2045	
細菌界;アクチノバクテリア門;アクチノバクテリア綱;アクチノミセス目;セルロモナダ科(Cellulomonadaceae);sf_1;1748	
細菌界;アクチノバクテリア門;アクチノバクテリア綱;アクチノミセス目;アクチノミセス科(Actinomycetaceae);sf_1;1684	
細菌界;アクチノバクテリア門;アクチノバクテリア綱;ビフィドバクテリウム目(Bifidobacteriales);ビフィドバクテリウム科(Bifidobacteriaceae);sf_1;1444	
細菌界;アクチノバクテリア門;アクチノバクテリア綱;アクチノミセス目;キネオスポリア科(Kineosporiaceae);sf_1;1598	10
細菌界;アクチノバクテリア門;アクチノバクテリア綱;アクチノミセス目;キネオスポリア科;sf_1;1961	
細菌界;アクチノバクテリア門;アクチノバクテリア綱;アクチノミセス目;コリネバクテリウム科(Corynebacteriaceae);sf_1;1517	
細菌界;アクチノバクテリア門;アクチノバクテリア綱;アクチノミセス目;コリネバクテリウム科;sf_1;1803	
細菌界;アクチノバクテリア門;アクチノバクテリア綱;アクチノミセス目;ディエジア科(Dietziaceae);sf_1;1970	
細菌界;ファーミキューテス門(Firmicutes);未分類;未分類;未分類;sf_8;2433	20
細菌界;ファーミキューテス門;クロストリジウム綱(Clostridia);未分類;未分類;sf_4;2398	
細菌界;クロロフレクサス門(Chloroflexi);デハロココイデス綱(Dehalococcoidetes);未分類;未分類;sf_1;2339	
細菌界;クロロフレクサス門;デハロココイデス綱;未分類;未分類;sf_1;2497	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目(Clostridiales);ペプトコッカス(Peptococc)/アシダミノコッカス(Acidaminococc);sf_11;709	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ペプトコッカス/アシダミノコッカス;sf_11;242	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ラクノスピラ科(Lachnospiraceae);sf_5;3042	30
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ラクノスピラ科;sf_5;3076	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ラクノスピラ科;sf_5;3171	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ラクノスピラ科;sf_5;2681	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ペプトストレプトコッカス科(Peptostreptococcaceae);sf_5;2721	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ペプトストレプトコッカス科;sf_5;2796	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;クロストリジウム科(Clostridiaceae);sf_12;2915	40
細菌界;TM7;未分類;未分類;未分類;sf_1;3025	
細菌界;ファーミキューテス門;バチルス綱;バチルス目(Bacillales);パエニバチルス科(Paenibacillaceae);sf_1;3299	

【表 4 E】

細菌界;ファーミキューテス門;パチルス綱;パチルス目;ハロパチルス科 (Halobacillaceae);sf_1;3344	
細菌界;ファーミキューテス門;パチルス綱;ラクトパチルス目(Lactobacillales);ストレプトコ ッカス科(Streptococcaceae);sf_1;3722	
細菌界;ファーミキューテス門;モリクテス綱(Mollicutes);アコレプラズマ目 (Acholeplasmatales);アコレプラズマ科(Acholeplasmataceae);sf_1;3976	
細菌界;アシドバクテリア門;未分類;未分類;未分類;sf_1;4222	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;クロストリジウ ム科;sf_12;4406	10
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ラクノスピラ科 ;sf_5;4212	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;クロストリジウ ム科;sf_12;4359	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;クロストリジウ ム科;sf_12;4475	
細菌界;未分類;未分類;未分類;未分類;sf_160;4410	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;クロストリジウ ム科;sf_12;4306	20
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;クロストリジウ ム科;sf_12;4427	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;クロストリジウ ム科;sf_12;4296	
細菌界;ウェルコミクロビウム門(Verrucomicrobia);ウェルコミクロビウム綱 (Verrucomicrobiae);ウェルコミクロビウム目(Verrucomicrobiales);ウェルコミクロビウム亜 門(Verrucomicrobiasubdivision)7;sf_1;559	
細菌界;バクテロイデス門;フラボバクテリア綱;フラボバクテリア目;フラボバクテリア科 ;sf_1;6200	
細菌界;プロテオバクテリア門;ベータプロテオバクテリア綱;パークホルデリア目;コマモ ナス科;sf_1;7971	30
細菌界;ウェルコミクロビウム門;ウェルコミクロビウム綱;ウェルコミクロビウム目;ウェ ルコミクロビウム亜門 5;sf_1;533	
細菌界;ウェルコミクロビウム門;未分類;未分類;未分類;sf_4;288	
細菌界;バクテロイデス門;バクテロイデス綱;バクテロイデス目;バクテロイデス科 ;sf_12;5320	
細菌界;バクテロイデス門;バクテロイデス綱;バクテロイデス目;バクテロイデス科 ;sf_12;5950	
細菌界;バクテロイデス門;バクテロイデス綱;バクテロイデス目;プレボテラ科;sf_1;5905	
細菌界;シアノバクテリア門;シアノバクテリア綱;ネンジュモ目(Nostocales);未分類 ;sf_1;5047	40
細菌界;シアノバクテリア門;シアノバクテリア綱;ネンジュモ目;未分類;sf_1;5072	
細菌界;シアノバクテリア門;シアノバクテリア綱;ネンジュモ目;未分類;sf_1;5191	
細菌界;シアノバクテリア門;シアノバクテリア綱;ネンジュモ目;未分類;sf_1;5199	

【表 4 F】

細菌界;BRC1;未分類;未分類;未分類;sf_1;5051	
細菌界;シアノバクテリア門;シアノバクテリア綱;クロロプラスト目;クロロプラスト科;sf_5;5130	
細菌界;未分類;未分類;未分類;未分類;sf_160;6337	
細菌界;未分類;未分類;未分類;未分類;sf_160;6360	
細菌界;アシドバクテリウム門;アシドバクテリウム綱;アシドバクテリウム目(Acidobacteriales);アシドバクテリウム科(Acidobacteriaceae);sf_14;6425	
細菌界;スピロヘータ門;スピロヘータ綱;スピロヘータ目;スピロヘータ科;sf_1;6529	10
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;アセトバクター目;アセトバクター科;sf_1;7529	
細菌界;プロテオバクテリア門;ベータプロテオバクテリア綱;未分類;未分類;sf_3;8007	
細菌界;プロテオバクテリア門;ベータプロテオバクテリア綱;MND1 クローングループ;未分類;sf_1;7993	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;アルテロモナス目;アルテロモナス科;sf_1;9491	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;アルテロモナス目;シュワネラ科;sf_1;8201	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;パスツレラ目;パスツレラ科;sf_1;8409	20
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;エンテロバクター目;エンテロバクター科;sf_1;9363	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;エンテロバクター目;エンテロバクター科;sf_1;8934	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;エンテロバクター目;エンテロバクター科;sf_1;8467	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;エンテロバクター目;エンテロバクター科;sf_1;8530	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;エンテロバクター目;エンテロバクター科;sf_1;9390	30
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;エンテロバクター目;エンテロバクター科;sf_1;8251	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;エンテロバクター目;エンテロバクター科;sf_1;8890	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;エンテロバクター目;エンテロバクター科;sf_1;8362	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;エンテロバクター目;エンテロバクター科;sf_1;8510	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;エンテロバクター目;エンテロバクター科;sf_1;8711	40
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;エンテロバクター目;エンテロバクター科;sf_1;8712	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;エンテロバクター目;エンテロバクター科;sf_1;8739	

【表 4 G】

細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;エンテロバクター目;エンテロバクター科;sf_1;9417	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;エンテロバクター目;エンテロバクター科;sf_1;8473	
細菌界;プロテオバクテリア門;デルタプロテオバクテリア綱;ミキソコッカス目(Myxococcales);ポリアンギウム科(Polyangiaceae);sf_3;10082	
細菌界;プロテオバクテリア門;デルタプロテオバクテリア綱;シントロフォバクター目;シントロフォバクター科(Syntrophobacteraceae);sf_1;9864	
細菌界;プロテオバクテリア門;デルタプロテオバクテリア綱;シントロフォバクター目;シントロフォバクター科;sf_1;9731	10
細菌界;アクチノバクテリア門;アクチノバクテリア綱;アクチノミセス目;フランキア科(Frankiaceae);sf_1;1286	
細菌界;アクチノバクテリア門;アクチノバクテリア綱;アクチノミセス目;ディエジア科;sf_1;1872	
細菌界;クロロフレクサス門;デハロココイデス綱;未分類;未分類;sf_1;2397	
細菌界;クロロフレクサス門;未分類;未分類;未分類;sf_1;2534	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ペプトコッカス/アシダミノコッカス;sf_11;710	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ペプトコッカス/アシダミノコッカス;sf_11;300	20
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ラクノスピラ科;sf_5;3218	
細菌界;ファーミキューテス門;カタバクター綱(Catabacter);未分類;未分類;sf_4;2716	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ペプトストレプトコッカス科;sf_5;2679	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ペプトストレプトコッカス科;sf_5;2714	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ペプトストレプトコッカス科;sf_5;2722	30
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ペプトストレプトコッカス科;sf_5;2993	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;クロストリジウム科;sf_12;3021	
細菌界;ファーミキューテス門;バチルス綱;ラクトバチルス目;カルノバクテリウム科(Carnobacteriaceae);sf_1;3536	
細菌界;ファーミキューテス門;バチルス綱;ラクトバチルス目;ストレプトコッカス科;sf_1;3869	
細菌界;ファーミキューテス門;バチルス綱;ラクトバチルス目;ストレプトコッカス科;sf_1;3588	
細菌界;ファーミキューテス門;モリクテス綱;アナエロプラズマ目(Anaeroplasmatales);エリシペロトリクス科(Erysipelotrichaceae);sf_3;3981	40
細菌界;ファーミキューテス門;カタバクター綱;未分類;未分類;sf_1;4261	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ラクノスピラ科;sf_5;4571	

【表 4 H】

細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ラクノスピラ科;sf_5;4623	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;クロストリジウム科;sf_12;4589	
細菌界;バクテロイデス門;バクテロイデス綱;バクテロイデス目;未分類;sf_15;5511	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;エンテロバクター目;エンテロバクター科;sf_1;8286	
細菌界;プロテオバクテリア門;デルタプロテオバクテリア綱;デスルフォバクター目(Desulfobacterales);デスルフォバクター科(Desulfobacteraceae);sf_5;10136	10
細菌界;アクウィフェクス門(Aquificae);アクウィフェクス綱;アクウィフェクス目(Aquificales);未分類;sf_1;2364	
細菌界;ウェルコミクロビウム門;ウェルコミクロビウム綱;ウェルコミクロビウム目;ウェルコミクロビウム科(Verrucomicrobiaceae);sf_6;871	
細菌界;ウェルコミクロビウム門;ウェルコミクロビウム綱;ウェルコミクロビウム目;ウェルコミクロビウム科;sf_1;1024	
細菌界;バクテロイデス門;スフンゴバクテリア綱;スフンゴバクテリア目;毛状菌科;sf_11;5334	
細菌界;クロロフレクサス門;アナエロリネア綱(Anaerolineae);未分類;未分類;sf_9;72	
細菌界;シアノバクテリア門;シアノバクテリア綱;ネンジュモ目;未分類;sf_1;5004	20
細菌界;アシドバクテリア門;ソリバクター綱(Solibacteres);未分類;未分類;sf_1;6426	
細菌界;スピロヘータ門;スピロヘータ綱;スピロヘータ目;スピロヘータ科;sf_1;6507	
細菌界;スピロヘータ門;スピロヘータ綱;スピロヘータ目;スピロヘータ科;sf_1;6460	
細菌界;スピロヘータ門;スピロヘータ綱;スピロヘータ目;スピロヘータ科;sf_1;6579	
細菌界;スピロヘータ門;スピロヘータ綱;スピロヘータ目;レプトスピラ科(Leptospiraceae);sf_3;6470	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;未分類;未分類;sf_6;7647	
細菌界;プロテオバクテリア門;ベータプロテオバクテリア綱;ニトロソモナス目(Nitrosomonadales);ニトロソモナス科(Nitrosomonadaceae);sf_1;8145	30
細菌界;プロテオバクテリア門;ベータプロテオバクテリア綱;パークホルデル目;コマモナス科;sf_1;7822	
細菌界;プロテオバクテリア門;ベータプロテオバクテリア綱;未分類;未分類;sf_3;7954	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;チオトリクス目;チオトリクス科;sf_3;8321	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;チオトリクス目;フランシセラ科(Francisellaceae);sf_1;9554	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;キサントモナス目(Xanthomonadales);キサントモナス科(Xanthomonadaceae);sf_3;8983	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;レジオネラ目;コクシエラ科;sf_3;8969	40
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;オセアノスピリルム目;ハロモナス科(Halomonadaceae);sf_1;8598	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;アルテロモナス目;アルテロモナス科;sf_1;9236	

【表 4 I】

細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;エンテロバクター目;エンテロバクター科;sf_1;8742	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;エンテロバクター目;エンテロバクター科;sf_1;9135	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;エンテロバクター目;エンテロバクター科;sf_1;9496	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;エンテロバクター目;エンテロバクター科;sf_1;8886	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;エンテロバクター目;エンテロバクター科;sf_1;9651	10
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;エンテロバクター目;エンテロバクター科;sf_1;8379	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;エンテロバクター目;エンテロバクター科;sf_1;9142	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;エンテロバクター目;エンテロバクター科;sf_1;9345	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;エンテロバクター目;エンテロバクター科;sf_1;8282	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;エンテロバクター目;未分類;sf_1;8430	20
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;エンテロバクター目;エンテロバクター科;sf_1;8505	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;エンテロバクター目;エンテロバクター科;sf_1;8528	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;エンテロバクター目;エンテロバクター科;sf_1;8936	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;エンテロバクター目;エンテロバクター科;sf_1;9060	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;エンテロバクター目;エンテロバクター科;sf_1;9274	30
細菌界;プロテオバクテリア門;デルタプロテオバクテリア綱;デスルフォビブリオ目;デスルフォビブリオ科;sf_1;10212	
細菌界;プロテオバクテリア門;デルタプロテオバクテリア綱;EB1021 グループ;未分類;sf_4;10024	
細菌界;プロテオバクテリア門;イプシロンプロテオバクテリア綱;カンピロバクター目;カンピロバクター科(Campylobacteraceae);sf_3;10397	
細菌界;アクチノバクテリア門;アクチノバクテリア綱;アシドマイクロビウム目;マイクロスリックス科(Microthrixineae);sf_1;1576	
細菌界;アクチノバクテリア門;アクチノバクテリア綱;アクチノミセス目;シュードノカルジア科(Pseudonocardiaceae);sf_1;1863	40
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;クロストリジウム科;sf_12;252	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ラクノスピラ科;sf_5;2709	

【表 4 J】

細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ラクノスピラ科;sf_5;3060	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ペプトストレプトコッカス科;sf_5;2729	
細菌界;ファーミキューテス門;バチルス綱;バチルス目;バチルス科(Bacillaceae);sf_1;234	
細菌界;ファーミキューテス門;バチルス綱;バチルス目;バチルス科;sf_1;3460	
細菌界;ファーミキューテス門;バチルス綱;バチルス目;ハロバチルス科;sf_1;3769	
細菌界;ファーミキューテス門;バチルス綱;バチルス目;バチルス科;sf_1;3900	10
細菌界;ファーミキューテス門;バチルス綱;バチルス目;カリオフアノン科(Caryophanaceae);sf_1;3285	
細菌界;ファーミキューテス門;バチルス綱;ラクトバチルス目;ラクトバチルス科(Lactobacillaceae);sf_1;3768	
細菌界;ファーミキューテス門;モリクテス綱;アコレプラズマ目;アコレプラズマ科;sf_1;4044	
細菌界;ファーミキューテス門;モリクテス綱;アコレプラズマ目;アコレプラズマ科;sf_1;4045	
細菌界;ファーミキューテス門;モリクテス綱;アナエロプラズマ目;エリシペロトリクス科;sf_3;3965	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;クロストリジウム科;sf_12;4614	20
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;クロストリジウム科;sf_12;4415	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;クロストリジウム科;sf_12;4548	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;クロストリジウム科;sf_12;4555	
細菌界;ニトリスピラ門(Nitrospira);ニトリスピラ綱;ニトリスピラ目(Nitrospirales);ニトリスピラ科(Nitrospiraceae);sf_2;542	
細菌界;ニトリスピラ門;ニトリスピラ綱;ニトリスピラ目;ニトリスピラ科;sf_2;697	30
細菌界;ナトロノアナエロビウム(Natronoanaerobium);未分類;未分類;未分類;sf_1;769	
細菌界;アシドバクテリア門;アシドバクテリア綱-4;エリン(Ellin)6075/11-25;未分類;sf_1;435	
細菌界;バクテロイデス門;バクテロイデス綱;バクテロイデス目;プレボテラ科;sf_1;5484	
細菌界;シアノバクテリア門;シアノバクテリア綱;シュードアナベナ目(Pseudanabaena);未分類;sf_1;5008	
細菌界;海洋グループ A;mgA-1;未分類;未分類;sf_1;6454	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ベロルホドスピラ目(Verorhodospirilla);未分類;sf_1;7109	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ブラジリゾビウム目(Bradyrhizobiales);ペイエリンク(Beijerinck)/ロードプラン(Rhodoplan)/メチロシスト(Methylocyst);sf_3;7401	40
細菌界;プロテオバクテリア門;ベータプロテオバクテリア綱;ロドシクルス目;ロドシクルス科;sf_1;7951	

【表 4 K】

細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;チオトリクス目;チオトリクス科;sf_3;9321	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;オセアノスピリルム目;ハロモナス科;sf_1;8317	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;シュードモナス目;モラクサラ科(Moraxellaceae);sf_3;9359	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;アルテロモナス目;アルテロモナス科;sf_1;8533	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;エンテロバクター目;エンテロバクター科;sf_1;9358	10
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;エンテロバクター目;エンテロバクター科;sf_1;9302	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;エンテロバクター目;エンテロバクター科;sf_1;8603	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;エンテロバクター目;エンテロバクター科;sf_1;9265	
細菌界;プロテオバクテリア門;デルタプロテオバクテリア綱;ミキソコッカス目;未分類;sf_1;10259	
細菌界;プロテオバクテリア門;デルタプロテオバクテリア綱;未分類;未分類;sf_7;10048	20
細菌界;プロテオバクテリア門;デルタプロテオバクテリア綱;EB1021 グループ;未分類;sf_4;9741	
細菌界;クロロフレクサス門;クロロフレクサス門-4;未分類;未分類;sf_2;2344	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ペプトコッカス/アシダミノコッカス;sf_11;39	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ラクノスピラ科;sf_5;3036	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ラクノスピラ科;sf_5;2825	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ペプトストレプトコッカス科;sf_5;58	30
細菌界;ファーミキューテス門;バチルス綱;ラクトバチルス目;ラクトバチルス科;sf_1;3566	
細菌界;ファーミキューテス門;バチルス綱;ラクトバチルス目;ストレプトコッカス科;sf_1;3251	
細菌界;ファーミキューテス門;モリクテス綱;アナエロプラズマ目;エリシペロトリクス科;sf_3;768	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;クロストリジウム科;sf_12;4297	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;クロストリジウム科;sf_12;4299	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;クロストリジウム科;sf_12;4502	40
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;クロストリジウム科;sf_12;4554	

【表 4 L】

細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;クロストリジウム科;sf_12;4157	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;クロストリジウム科;sf_12;4267	
細菌界;バクテロイデス門;バクテロイデス綱;バクテロイデス目;ポルフィロモナス科;sf_1;5961	
細菌界;バクテロイデス門;バクテロイデス綱;バクテロイデス目;プレボテラ科;sf_1;5916	
細菌界;バクテロイデス門;フラボバクテリア綱;フラボバクテリア目;フラボバクテリア科;sf_1;5473	10
細菌界;シアノバクテリア門;シアノバクテリア綱;ネンジュモ目;未分類;sf_1;5028	
細菌界;シアノバクテリア門;シアノバクテリア綱;ネンジュモ目;未分類;sf_1;5174	
細菌界;シアノバクテリア門;シアノバクテリア綱;ネンジュモ目;未分類;sf_1;5175	
細菌界;TM7;未分類;未分類;未分類;sf_1;5061	
細菌界;プロテオバクテリア門;ベータプロテオバクテリア綱;バークホルデリア目;バークホルデリア科(Burkholderiaceae);sf_1;7782	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;クロマチウム目(Chromatiales);未分類;sf_1;9282	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;オセアノスピリルム目;ハロモナス科;sf_1;8854	20
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;シュードモナス目;シュードモナス科;sf_1;8209	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;エンテロバクター目;エンテロバクター科;sf_6;8783	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ペプトコッカス/アシダミノコッカス;sf_11;304	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ペプトコッカス/アシダミノコッカス;sf_11;131	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;クロストリジウム科;sf_12;206	30
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ラクノスピラ科;sf_5;2834	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ラクノスピラ科;sf_5;2844	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ペプトストレプトコッカス科;sf_5;2694	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ペプトストレプトコッカス科;sf_5;3080	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ペプトストレプトコッカス科;sf_5;3182	40
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ペプトストレプトコッカス科;sf_5;619	
細菌界;ファーミキューテス門;バチルス綱;バチルス目;バチルス科;sf_1;305	
細菌界;ファーミキューテス門;バチルス綱;バチルス目;バチルス科;sf_1;3836	

【表 4 M】

細菌界;ファーミキューテス門;バチルス綱;バチルス目;バチルス科;sf_1;462	
細菌界;ファーミキューテス門;バチルス綱;バチルス目;バチルス科;sf_1;3831	
細菌界;ファーミキューテス門;バチルス綱;ラクトバチルス目;エンテロコッカス科(Enterococcaceae);sf_1;3288	
細菌界;ファーミキューテス門;バチルス綱;ラクトバチルス目;エンテロコッカス科;sf_1;3598	
細菌界;ファーミキューテス門;モリクテス綱;アコレプラズマ目;アコレプラズマ科;sf_1;3961	
細菌界;ファーミキューテス門;モリクテス綱;アコレプラズマ目;アコレプラズマ科;sf_1;3975	10
細菌界;ファーミキューテス門;モリクテス綱;アナエロプラズマ目;エリシペロトリクス科;sf_3;3952	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;クロストリジウム科;sf_12;4584	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;クロストリジウム科;sf_12;4459	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ラクノスピラ科;sf_5;4533	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ラクノスピラ科;sf_5;4539	20
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;クロストリジウム科;sf_12;4637	
細菌界;ファーミキューテス門;カタバクター綱;未分類;未分類;sf_4;4526	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;クロストリジウム科;sf_12;4560	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;クロストリジウム科;sf_12;4310	
細菌界;プロテオバクテリア門;ベータプロテオバクテリア綱;パークホルデリア目;未分類;sf_1;7879	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;キサントモナス目;キサントモナス科;sf_3;9211	30
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;アルテロモナス目;シュードアルテロモナス科(Pseudoalteromonadaceae);sf_1;9339	
細菌界;プロテオバクテリア門;デルタプロテオバクテリア綱;ミキソコッカス目;ポリアンギウム科;sf_3;10065	
細菌界;プロテオバクテリア門;デルタプロテオバクテリア綱;未分類;未分類;sf_9;9738	
細菌界;プロテオバクテリア門;イプシロンプロテオバクテリア綱;カンピロバクター目;ヘリコバクター科;sf_3;10572	
細菌界;アクチノバクテリア門;アクチノバクテリア綱;アクチノミセス目;ノカルジオブシス科(Nocardiopsaceae);sf_1;1385	40
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ペプトコッカス/アシダミノコッカス;sf_11;71	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ラクノスピラ科;sf_5;3059	

【 0 2 9 3 】

【表 4 N】

細菌界;TM7;TM7-3;未分類;未分類;sf_1;2697	
細菌界;ファーミキューテス門;バチルス綱;バチルス目;パエニバチルス科;sf_1;3630	
細菌界;ファーミキューテス門;バチルス綱;バチルス目;バチルス科;sf_1;3424	
細菌界;ファーミキューテス門;バチルス綱;バチルス目;バチルス科;sf_1;3661	
細菌界;ファーミキューテス門;バチルス綱;バチルス目;バチルス科;sf_1;283	
細菌界;ファーミキューテス門;バチルス綱;バチルス目;バチルス科;sf_1;829	
細菌界;ファーミキューテス門;バチルス綱;バチルス目;バチルス科;sf_1;3675	
細菌界;ファーミキューテス門;モリクテス綱;エントモプラズマ目(Entomoplasmales);エントモプラズマ科(Entomoplasmataceae);sf_1;4074	10
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;クロストリジウム科;sf_12;4156	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;クロストリジウム科;sf_12;4575	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;エンテロバクター目;エンテロバクター科;sf_1;8631	
細菌界;プロテオバクテリア門;イプシロンプロテオバクテリア綱;カンピロバクター目;ヘリコバクター科;sf_3;10534	
細菌界;ニトリスピラ門;ニトリスピラ綱;ニトリスピラ目;ニトリスピラ科;sf_1;179	20
細菌界;ウェルコミクロビウム門;ウェルコミクロビウム綱;ウェルコミクロビウム目;ウェルコミクロビウム亜門 7;sf_1;446	
細菌界;バクテロイデス門;スフンゴバクテリア綱;スフンゴバクテリア目;スフンゴバクテリア科;sf_1;6272	
細菌界;スピロヘータ門;スピロヘータ綱;スピロヘータ目;スピロヘータ科;sf_1;6487	
細菌界;スピロヘータ門;スピロヘータ綱;スピロヘータ目;スピロヘータ科;sf_1;6554	
細菌界;プロテオバクテリア門;ベータプロテオバクテリア綱;ニトロソモナス目;ニトロソモナス科;sf_1;7931	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ラクノスピラ科;sf_5;3111	30
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ラクノスピラ科;sf_5;2693	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ペプトストレプトコッカス科;sf_5;2913	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ペプトストレプトコッカス科;sf_5;1037	
細菌界;ファーミキューテス門;バチルス綱;バチルス目;スポロラクトバチルス科(Sporolactobacillaceae);sf_1;3365	
細菌界;ファーミキューテス門;バチルス綱;バチルス目;バチルス科;sf_1;3419	
細菌界;ファーミキューテス門;バチルス綱;バチルス目;ハロバチルス科;sf_1;3756	40
細菌界;ファーミキューテス門;バチルス綱;バチルス目;ハロバチルス科;sf_1;3849	
細菌界;ファーミキューテス門;バチルス綱;ラクトバチルス目;エンテロコッカス科;sf_1;3881	
細菌界;ファーミキューテス門;バチルス綱;ラクトバチルス目;ストレプトコッカス科;sf_1;3629	

【表 40】

細菌界;ファーミキューテス門;モリクテス綱;アナエロプラズマ目;エリシペロトリクス科;sf_3;144	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;クロストリジウム科;sf_12;4632	
細菌界;バクテロイデス門;フラボバクテリア綱;フラボバクテリア目;フラボバクテリア科;sf_1;5509	
細菌界;プロテオバクテリア門;デルタプロテオバクテリア綱;ミキソコッカス目;ポリアンギウム科;sf_3;9912	
細菌界;NC10;NC10-1;未分類;未分類;sf_1;536	10
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;エンテロバクター目;エンテロバクター科;sf_1;8640	
細菌界;アクチノバクテリア門;アクチノバクテリア綱;アクチノミセス目;未分類;sf_4;1337	
細菌界;アクチノバクテリア門;アクチノバクテリア綱;アクチノミセス目;キネオスポリア科;sf_1;1087	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;クロストリジウム科;sf_12;2786	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ユーバクテリウム科(Eubacteriaceae);sf_1;28	
細菌界;ファーミキューテス門;バチルス綱;バチルス目;バチルス科;sf_1;3540	20
細菌界;ファーミキューテス門;バチルス綱;バチルス目;バチルス科;sf_1;3827	
細菌界;ファーミキューテス門;バチルス綱;ラクトバチルス目;ラクトバチルス科;sf_1;3703	
細菌界;ファーミキューテス門;腸クロングループ;未分類;未分類;sf_1;4298	
細菌界;ファーミキューテス門;カタバクター綱;未分類;未分類;sf_4;4325	
細菌界;未分類;未分類;未分類;未分類;sf_92;9999	
細菌界;バクテロイデス門;バクテロイデス綱;バクテロイデス目;バクテロイデス科;sf_12;5621	
細菌界;BRC1;未分類;未分類;未分類;sf_1;5143	
細菌界;プロテオバクテリア門;ベータプロテオバクテリア綱;ロドシクルス目;ロドシクルス科;sf_1;8052	30
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;アルテロモナス目;アルテロモナス科;sf_1;8904	
細菌界;プロテオバクテリア門;デルタプロテオバクテリア綱;ミキソコッカス目;ポリアンギウム科;sf_3;10353	
細菌界;ファーミキューテス門;バチルス綱;バチルス目;バチルス科;sf_1;3283	
細菌界;ファーミキューテス門;バチルス綱;バチルス目;スタフィロコッカス科(Staphylococcaceae);sf_1;3258	
細菌界;ファーミキューテス門;バチルス綱;バチルス目;スタフィロコッカス科;sf_1;3605	
細菌界;ファーミキューテス門;バチルス綱;ラクトバチルス目;ロイコノストック科(Leuconostocaceae);sf_1;3497	40
細菌界;ファーミキューテス門;バチルス綱;ラクトバチルス目;ストレプトコッカス科;sf_1;3290	

【表 4 P】

細菌界;ファーミキューテス門;未分類;未分類;未分類;sf_8;4536	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ラクノスピラ科;sf_5;4155	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;クロストリジウム科;sf_12;4378	
細菌界; ウェルコミクロビウム門;ウェルコミクロビウム綱;ウェルコミクロビウム目;未分類;sf_3;11	
細菌界;アシドバクテリア門;アシドバクテリア綱;アシドバクテリア目;アシドバクテリア科;sf_14;208	10
細菌界;バクテロイデス門;バクテロイデス綱;バクテロイデス目;バクテロイデス科;sf_12;5275	
細菌界;バクテロイデス門;フラボバクテリア綱;フラボバクテリア目;フラボバクテリア科;sf_1;5423	
細菌界;プロテオバクテリア門;ベータプロテオバクテリア綱;ニトロソモナス目;ニトロソモナス科;sf_1;7805	
細菌界;プロテオバクテリア門;ベータプロテオバクテリア綱;ニトロソモナス目;ニトロソモナス科;sf_1;7858	
細菌界;アクチノバクテリア門;アクチノバクテリア綱;アクチノミセス目;フランキア科;sf_1;1105	20
細菌界;ゲマティモナス門;未分類;未分類;未分類;sf_5;1565	
細菌界;アクチノバクテリア門;アクチノバクテリア綱;アクチノミセス目;マイクロコッカス科(Micrococcaceae);sf_1;1213	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ラクノスピラ科;sf_5;2804	
細菌界;ファーミキューテス門;バチルス綱;バチルス目;スタフィロコッカス科;sf_1;3284	
細菌界;ファーミキューテス門;バチルス綱;バチルス目;スタフィロコッカス科;sf_1;3628	
細菌界;ファーミキューテス門;バチルス綱;ラクトバチルス目;ラクトバチルス科;sf_1;3547	
細菌界;ファーミキューテス門;バチルス綱;ラクトバチルス目;ラクトバチルス科;sf_1;3634	
細菌界;ファーミキューテス門;バチルス綱;ラクトバチルス目;エンテロコッカス科;sf_1;3261	30
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;クロストリジウム科;sf_12;4638	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;クロストリジウム科;sf_12;4275	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ペプトコッカス/アシダミノコッカス;sf_11;489	
細菌界;プロテオバクテリア門;ベータプロテオバクテリア綱;未分類;未分類;sf_3;7765	
細菌界;NC10;NC10-2;未分類;未分類;sf_1;10254	
細菌界;アクチノバクテリア門;アクチノバクテリア綱;アクチノミセス目;マイコバクテリア科(Mycobacteriaceae);sf_1;1365	40
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ペプトストレプトコッカス科;sf_5;3112	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;クロストリジウム科;sf_12;3219	

【表 4 Q】

細菌界;ファーミキューテス門;バチルス綱;バチルス目;バチルス科;sf_1;385	
細菌界;ファーミキューテス門;バチルス綱;バチルス目;バチルス科;sf_1;571	
細菌界;ファーミキューテス門;バチルス綱;バチルス目;スタフィロコッカス科;sf_1;3684	
細菌界;ファーミキューテス門;バチルス綱;ラクトバチルス目;エンテロコッカス科;sf_1;3433	
細菌界;プロテオバクテリア門;ベータプロテオバクテリア綱;ニトロソモナス目;ニトロソモナス科;sf_1;7831	
細菌界;ファーミキューテス門;バチルス綱;ラクトバチルス目;ラクトバチルス科;sf_1;3330	10
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;GAO クラスター;未分類;sf_1;8980	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ラクノスピラ科;sf_5;2756	
細菌界;ファーミキューテス門;バチルス綱;バチルス目;スタフィロコッカス科;sf_1;3545	
細菌界;ファーミキューテス門;バチルス綱;ラクトバチルス目;エンテロコッカス科;sf_1;3298	
細菌界;ニトリスピラ門;ニトリスピラ綱;ニトリスピラ目;ニトリスピラ科;sf_3;833	
細菌界;バクテロイデス門;バクテロイデス綱;バクテロイデス目;バクテロイデス科;sf_12;5474	20
細菌界;バクテロイデス門;未分類;未分類;未分類;sf_1;5745	
細菌界;バクテロイデス門;バクテロイデス綱;バクテロイデス目;プレボテラ科;sf_1;5769	
細菌界;シアノバクテリア門;シアノバクテリア綱;ユレモ目(Oscillatoriales);未分類;sf_1;5184	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;GAO クラスター;未分類;sf_1;9468	
細菌界;プロテオバクテリア門;デルタプロテオバクテリア綱;デススルフロモナス目(Desulfuromonadales);ゲオバクター科(Geobacteraceae);sf_1;9956	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ラクノスピラ科;sf_5;3066	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ラクノスピラ科;sf_5;3088	30
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ラクノスピラ科;sf_5;3075	
細菌界;ファーミキューテス門;バチルス綱;バチルス目;バチルス科;sf_1;3688	
細菌界;ファーミキューテス門;バチルス綱;バチルス目;スタフィロコッカス科;sf_1;3822	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ラクノスピラ科;sf_5;4167	
細菌界;ファーミキューテス門;バチルス綱;バチルス目;サーモアクチノミセス科(Thermoactinomycetaceae);sf_1;3539	
細菌界;バクテロイデス門;バクテロイデス綱;バクテロイデス目;リケネラ科(Rikenellaceae);sf_5;5889	40
細菌界;バクテロイデス門;バクテロイデス綱;バクテロイデス目;ポルフィロモナス科;sf_1;5932	

【 0 2 9 7 】

【表 4 R】

細菌界;バクテロイデス門;バクテロイデス綱;バクテロイデス目;プレボテラ科;sf_1;5437	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ラクノスピラ科;sf_5;3089	
細菌界;ファーミキューテス門;バチルス綱;バチルス目;スタフィロコッカス科;sf_1;3569	
細菌界;ファーミキューテス門;バチルス綱;ラクトバチルス目;ラクトバチルス科;sf_1;3767	
細菌界;ファーミキューテス門;バチルス綱;ラクトバチルス目;エンテロコッカス科;sf_1;3713	
細菌界;ファーミキューテス門;モリクテス綱;未分類;未分類;sf_6;149	
細菌界;クロロフレクサス門;デハロココイデス綱;未分類;未分類;sf_1;2487	10
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ラクノスピラ科;sf_5;2784	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ラクノスピラ科;sf_5;2937	
細菌界;ファーミキューテス門;バチルス綱;バチルス目;スタフィロコッカス科;sf_1;3794	
細菌界;ファーミキューテス門;バチルス綱;ラクトバチルス目;エンテロコッカス科;sf_1;3382	
細菌界;ファーミキューテス門;バチルス綱;ラクトバチルス目;エンテロコッカス科;sf_1;3318	
細菌界;ファーミキューテス門;バチルス綱;ラクトバチルス目;ストレプトコッカス科;sf_1;3397	20
細菌界;ファーミキューテス門;バチルス綱;ラクトバチルス目;ストレプトコッカス科;sf_1;3446	
細菌界;バクテロイデス門;バクテロイデス綱;バクテロイデス目;プレボテラ科;sf_1;5946	
細菌界;アクチノバクテリア門;アクチノバクテリア綱;アクチノミセス目;サーモモノスポラ科;sf_1;1406	
細菌界;アクチノバクテリア門;アクチノバクテリア綱;アクチノミセス目;コリネバクテリウム科;sf_1;1428	
細菌界;ファーミキューテス門;バチルス綱;ラクトバチルス目;エンテロコッカス科;sf_1;3392	30
細菌界;ファーミキューテス門;バチルス綱;ラクトバチルス目;エンテロコッカス科;sf_1;3680	
細菌界;ファーミキューテス門;モリクテス綱;アナエロプラズマ目;エリシペロトリクス科;sf_3;3943	
細菌界;バクテロイデス門;フラボバクテリア綱;フラボバクテリア目;フラボバクテリア科;sf_1;5339	
細菌界;シアノバクテリア門;シアノバクテリア綱;ユレモ目;未分類;sf_1;5215	
細菌界;アクチノバクテリア門;アクチノバクテリア綱;アクチノミセス目;シュードノカルジア科;sf_1;1402	
細菌界;ファーミキューテス門;バチルス綱;ラクトバチルス目;ラクトバチルス科;sf_1;3521	40
細菌界;ファーミキューテス門;バチルス綱;ラクトバチルス目;ラクトバチルス科;sf_1;3885	
細菌界;ファーミキューテス門;バチルス綱;ラクトバチルス目;ストレプトコッカス科;sf_1;3250	

【 0 2 9 8 】

【表 4 S】

細菌界;ファーミキューテス門;バチルス綱;ラクトバチルス目;ストレプトコッカス科;sf_1;3906	
細菌界;ファーミキューテス門;バチルス綱;ラクトバチルス目;ストレプトコッカス科;sf_1;3287	
細菌界;ファーミキューテス門;バチルス綱;ラクトバチルス目;未分類;sf_1;3481	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;クロストリジウム科;sf_12;4173	
細菌界;プロテオバクテリア門;デルタプロテオバクテリア綱;デスルフォバクター目;デスルフォバクター科;sf_5;10275	10
細菌界;バクテロイデス門;バクテロイデス綱;バクテロイデス目;プレボテラ科;sf_1;5940	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ラクノスピラ科;sf_5;3087	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ラクノスピラ科;sf_5;2991	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ラクノスピラ科;sf_5;4381	
細菌界;ファーミキューテス門;バチルス綱;ラクトバチルス目;アエロコッカス科(Aerococcaceae);sf_1;3504	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ラクノスピラ科;sf_5;4443	20
細菌界;バクテロイデス門;バクテロイデス綱;バクテロイデス目;プレボテラ科;sf_1;5398	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ラクノスピラ科;sf_5;2849	
細菌界;プロテオバクテリア門;ベータプロテオバクテリア綱;パークホルデリア目;コマモナス科;sf_1;7834	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;パスツレラ目;パスツレラ科;sf_1;9263	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ラクノスピラ科;sf_5;2726	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;未分類;sf_17;2683	30
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ラクノスピラ科;sf_5;3107	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ラクノスピラ科;sf_5;3033	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;クロストリジウム科;sf_12;2736	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ラクノスピラ科;sf_5;4538	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ラクノスピラ科;sf_5;2808	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ラクノスピラ科;sf_5;2733	40

【表 4 T】

細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;クロストリジウム科;sf_12;3019	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ラクノスピラ科;sf_5;2747	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ラクノスピラ科;sf_5;2793	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ラクノスピラ科;sf_5;4563	
細菌界;フソバクテリウム門(Fusobacteria);フソバクテリウム門;フソバクテリウム目(Fusobacteriales);フソバクテリウム科(Fusobacteriaceae);sf_1;488	10
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ラクノスピラ科;sf_5;3149	
細菌界;ファーミキューテス門;モリクテス綱;アナエロプラズマ目;エリシペロトリクス科;sf_3;3956	
細菌界;バクテロイデス門;バクテロイデス綱;バクテロイデス目;リケネラ科;sf_5;6032	
細菌界;バクテロイデス門;バクテロイデス綱;バクテロイデス目;バクテロイデス科;sf_12;5285	
細菌界;バクテロイデス門;バクテロイデス綱;バクテロイデス目;未分類;sf_15;5299	
細菌界;バクテロイデス門;バクテロイデス綱;バクテロイデス目;バクテロイデス科;sf_12;5424	20
細菌界;バクテロイデス門;バクテロイデス綱;バクテロイデス目;バクテロイデス科;sf_12;5551	
細菌界;バクテロイデス門;バクテロイデス綱;バクテロイデス目;バクテロイデス科;sf_12;5979	
細菌界;バクテロイデス門;バクテロイデス綱;バクテロイデス目;バクテロイデス科;sf_12;6064	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;パスツレラ目;パスツレラ科;sf_1;9360	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;パスツレラ目;パスツレラ科;sf_1;8228	30
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;パスツレラ目;パスツレラ科;sf_1;8861	
細菌界;フソバクテリウム門;フソバクテリウム門;フソバクテリウム目;フソバクテリウム科;sf_3;558	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ペプトコッカス/アシダミノコッカス;sf_11;181	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ラクノスピラ科;sf_5;2731	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ラクノスピラ科;sf_5;3032	40
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;未分類;sf_17;2730	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ラクノスピラ科;sf_5;2769	

【 0 3 0 0 】

【表 4 U】

細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ラクノスピラ科;sf_5;2928	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ラクノスピラ科;sf_5;2753	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;クロストリジウム科;sf_12;2898	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ラクノスピラ科;sf_5;2965	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ラクノスピラ科;sf_5;2737	10
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ラクノスピラ科;sf_5;3016	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ラクノスピラ科;sf_5;3185	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;未分類;sf_17;2912	
細菌界;ファーミキューテス門;パチルス綱;ラクトパチルス目;ストレプトコッカス科;sf_1;3253	
細菌界;ファーミキューテス門;モリクテス綱;未分類;未分類;sf_6;196	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ラクノスピラ科;sf_5;4500	20
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;クロストリジウム科;sf_12;4570	

【 0 3 0 1 】

ファーミキューテス門、プロテオバクテリア門、バクテロイデス門およびアクチノバクテリア門からなる糞便分類群が見出された。ファーミキューテス門のうち、もっとも多いのは、ラクノスピラ科、ペプトストレプトコッカス科、アシダミノコッカス科およびクロストリジウム科を含むクロストリジウム目由来であり、パチルス科、ハロパチルス科、スタフィロコッカス科を含むパチルス目由来の比較的わずかな割合の細菌の他にも、ラクトパチルス科、エンテロコッカス科およびストレプトコッカス科を含む類似の割合のラクトパチルス目が存在する。プロテオバクテリア門では、約3分の1がエンテロバクター科を含むエンテロバクター目由来であり、わずかな割合のアルテロモナス目がアルテロモナス科およびシュワネラ科を含む。他の比較的小さな成分集団には、パークホルデリア科、コマモナス科、アルカリゲネス科、オキサロバクター科(Oxalobacteraceae)およびラルストニア科を含むパークホルデリア目由来の分類群が含まれる。

【 0 3 0 2 】

いくつかの実施形態では、1回のアッセイにおいて表4から選択される少なくとも10、25、50、100、200、300、400または500の異なる糞便分類群の存在または量を検出するためのシステムが提供される。追加の実施形態では、前記システムは少なくとも10、25、50、100、200、300、400または500の異なる糞便分類群のそれぞれに選択的にハイブリダイズするプローブを含む。他の実施形態では、水試料中の表4に列挙されている糞便分類群の16S RNA配列からなる群から選択される1つまたは複数の核酸配列の存在または量を検出することを含む水の糞便汚染を検出するための方法が提供される。追加の実施形態では、前記検出法は、表11に列挙されている清浄水分類群の1つまたは複数の16S RNA配列を検出することに依拠している。さらに追加の実施形態では、前記水試料は、前記1つまたは複数の清浄水分類群に選択的にハイブリダイズする複数のプローブに接触させる。有用なプローブには、表11に列挙されている生物または分類群を同定するために使用することが可能なプローブが挙げられる。

【 0 3 0 3 】

10

20

30

40

50

(実施例4)

水質試験、糞便汚染およびフローサイトメトリー

水質は、微粒子ベースの多重システムを使用して試験される。糞便汚染水に付随する中核となる微生物(バチルス、バクテロイデスおよびクロストリジウム目)のコレクションを認識する複数のプローブは表4から選択される。清浄水に付随する中核となる微生物(- プロテオバクテリア)のコレクションを認識するさらに複数のプローブは、表11に列挙されている生物または分類群を同定する複数のプローブからも選択される。複数のプローブの前記2つのコレクション内のプローブごとに標識された微粒子のサブロットが作製される。前記プローブはカルボジイミドカップリングにより3.0マイクロメートルのラテックス微小球(Interfacial Dynamicsにより製造)にカップリングされる。カップリング後に前記サブロットは組み合わせられる。次に、陰性対照プローブ結合微粒子と陽性対照プローブ結合微粒子は添加されて標識微粒子の最終ロットが作製される。

10

【0304】

水試料は0.22 μm フィルター上で濾過され、UltraClean Water DNAキット(MoBio Laboratories)を使用してDNAを抽出される。16S rRNA遺伝子は、ユニバーサル細菌プライマー27Fと1492Rを使用してPCR増幅される。潜在的PCR増幅バイアスを最小限に抑えるために、試料ごとに温度勾配(48~58)にわたる8つの複製反応が実施される。各試料のプールされた単位複製配列(250ng)は、内部QS標準でスパイクされて、アッセイハイブリダイゼーションシグナルの正規化を可能にする。この混合物は、断片化され、ビオチン標識され、2.3 \times SSC緩衝液中約25 で、約40ピコモルのビーズ付着オリゴを約2倍量のビオチン標識単位複製配列と結合させることにより前記微粒子にハイブリダイズさせる。この混合物は室温で2時間インキュベートし、続いて洗浄し、300マイクロリットルの生理食塩水pH7.3で希釈し、「FAGSCAN」(Becton-Dickinson Immunocytometry Systemsにより製造)上で分析する。結果は、水試料が排泄物で汚染されていることを示す1.05のBBC:A比を示している。

20

【0305】

(実施例5)

糞便汚染関連分類群

本発明のシステムと方法を使用して糞便汚染が疑われる試料を分析することにより、糞便汚染についてバイオシグネチャーを決定することが可能である。DNAは標準法を使用して試料から抽出される。次に、16S rDNAは実施例2に記載される通りに増幅され、プロセシングされ、分析される。プローブハイブリダイゼーションによる分析は、実施例2に記載される通りにアレイを使用して、またはビーズに結合しているプローブを用いて実施例4における方法に類似するフローサイトメトリー法を使用することにより、行うことが可能である。評価されるプローブごとにおよび/または評価されるプローブにより表わされるOTUごとに、存在、非存在および/またはレベルを点数化することが可能である。次に、このデータのコレクション、またはそのサブセットは、糞便汚染による汚染についてのバイオシグネチャーとしての働きをすることができ、このバイオシグネチャーと試験試料のバイオシグネチャーを比較することが可能である。

30

【0306】

レクリエーション海岸近くで採取された水試料は、容認しがたい高レベルの糞便汚染を有すると同定される。一連の水試料が、海岸近くおよび近隣のクリークの流域沿いで収集される。前記水試料は処理され、次に低密度水質アレイ上でアッセイされる。撮像およびシグナル処理後、試料ごとにBBC:A比が計算される。前記BBC:Aシグナルは、海岸近くで約1.05~1.10であり、流域沿いで増加し、その後突然清浄水を示す0.95より下にさがる。もっとも高いBBC:A読みを有する位置を取り巻く地点が捜され、破裂した下水管ラインが見つけれられる。下水管ラインの修復により、クリーク流域の水質が上昇する。

40

【0307】

ファームキューテス門、プロテオバクテリア門、バクテロイデス門およびアクチノバクテリア門からなる糞便分類群が見つかりしている。ファームキューテス門のうち、もっとも

50

多いのは、ラクノスピラ科、ペプトストレプトコッカス科、アシダミノコッカス科およびクロストリジウム科を含むクロストリジウム目由来であり、パチルス科、ハロパチルス科、スタフィロコッカス科を含む比較的わずかな割合のパチルス目由来の細菌の他にもラクトパチルス科、エンテロコッカス科およびストレプトコッカス科を含む類似の割合のラクトパチルス目が存在する。プロテアバクテリア門では、約3分の1がエンテロバクター科を含むエンテロバクター目由来であり、わずかな割合のアルテロモナス目がアルテロモナス科およびシュワネラ科を含む。他の比較的少ない成分集団には、パークホルデリア科、コマモナス科、アルカリゲネス科、オキサロバクテリア科およびラルストニア科を含むパークホルデリア目由来の分類群が含まれる。

【0308】

いくつかの実施形態では、1回の水質試験アッセイにおいて表4から選択される少なくとも10、25、50、100、200、300、400または500の異なる糞便分類群の存在または量を検出するためのシステムが提供される。追加の実施形態では、前記システムは少なくとも10、25、50、100、200、300、400または500の異なる糞便分類群のそれぞれに選択的にハイブリダイズするプローブを含む。他の実施形態では、水試料中の表4に列挙されている糞便分類群の16S RNA配列からなる群から選択される1つまたは複数の核酸配列の存在または量を検出することを含む水の糞便汚染を検出するための方法が提供される。追加の実施形態では、前記検出法は、表11に列挙されている清浄水分類群の1つまたは複数の16S RNA配列を検出することに依拠している。さらに追加の実施形態では、前記水試料は、前記1つまたは複数の清浄水分類群に選択的にハイブリダイズする複数のプローブに接触させる。有用なプローブには、表11に列挙されている生物または分類群を同定するために使用することが可能なプローブが挙げられる。

【0309】

(実施例6)

毒性藻類ブルーム

シアノバクテリアは、藍藻類としても知られるが、水性マイクロバイオーム(microbiomes)の主成分を表す。適切な条件、通常は栄養素の十分な利用能下で、その数は急速に増加することがあり、藻類ブルームを生じる。栄養素が使い果たされると、ブルームは死滅しその後、利用可能な溶存酸素をすべて消費することが可能な細菌性分解を受け、肉眼で見える生物がいない死水域を生じる。これらのシアノバクテリアが有する他のシアノバクテリアまたは細菌の存在を感知し(クオラムセンシング)、特定の密度で神経毒を産生する能力も厄介である。前記シアノバクテリアもしくはその神経毒を含有する水または海産物、特に毒性藻類ブルームを有する水域由来の貝を摂取すると重傷または死をもたらすことがある。公衆衛生を守り飲料水および海産物の安全性を確保するためには、毒性藻類ブルームを含む藻類ブルームの可能性を予測するための方法が必要である。

【0310】

いくつかの実施形態では、a)表6から選択されるシアノバクテリア由来の核酸に選択的に結合する複数のプローブに水試料を接触させる、b)ハイブリダイゼーションデータを使用して前記水試料中のシアノバクテリアの量と組成を決定する、c)環境条件を測定する、およびd)シアノバクテリアの量と組成および環境条件に基づいて毒性藻類ブルームの可能性を予測することを含む毒性藻類ブルームの可能性を予測するための方法が提供される。追加の実施形態では、プローブは、表6に列挙されている属を検出するための上記の方法により選択される。モニターする有用な環境条件には、水温、濁り、窒素濃度、リン酸濃度もしくは鉄濃度、または太陽光の強さが挙げられる。追加の実施形態では、他の微生物、特に細菌性生物の存在または量が決定される。高い頻度で、毒性ブルーム産生シアノバクテリアは、クオラムセンシングを使用するある種の細菌と共生的に生息している。シアノバクテリアは、細菌のクオラムセンシングを読むまたは乗っ取ることができる可能性があり、したがって、共生的細菌の量についての知識は毒素発現にとり重要でありうる(たとえば、毒素レベルに影響を及ぼす、毒素レベルを触媒するまたは制御しうる)。クオラムセンシングができる細菌およびシアノバクテリアならびにこの現象が起きる密度につい

10

20

30

40

50

ての知識を含む、水性マイクロバイームに存在する集団の関係についての知識があれば、いつ毒性藻類ブルームが発生するのかを予測することが可能になる。この予知力で武装していれば、毒性藻類ブルームの可能性に基づいて、ブルームに見舞われる可能性のある水域での遊泳もしくは貝集めを禁止するまたは上水道を井戸水のような代替の水源に切り替えることを含む、水の管理の決定を下すことが可能になる。

【 0 3 1 1 】

【表 5】

表 6:毒性藻類ブルームシアノバクテリア属

属
ミクロキスティス(Microcystis)
アナベナ(Anabaena)
プランクトスリックス(Planktothrix)(オシラトリア(Oscillatoria))
ノストック(Nostoc)
ハパロシフォン(Hapalosiphon)
アナベノプシス(Anabaenopsis)
ノドゥラリア(Nodularia)
アフアニゾメノン(Aphanizomenon)
リングビア(Lyngbya)
シゾトリックス(Schizothrix)
シリンドロスペルモプシス(Cylindrospermopsis)
アフアニゾメノン(Aphanizomenon)
ウメザキア(Umezakia)

10

20

【 0 3 1 2 】

地方の湖のレクリエーション水域由来の水試料は、毒性藻類ブルームに関連するシアノバクテリアの100 OTUから核酸を検出するために上記の通りに選択されたプローブを有するダウン選択された系統発生アレイに適用される。水mlあたり50,000を超えるシアノバクテリアのシアノバクテリア密度と相関がある3つのシアノバクテリアOTUが検出され定量化される。前記水温は70 ° F (約21 ° C)、透明度板が見えるのは14インチ(35.6cm)の深さまでで透明度は不十分であり、次の5日間日照が予想され、周囲屋外の昼間の温度が90 ° F 台(30 ° C 台)まで上昇すると予想される。毒性藻類ブルームの可能性は90%超である。レクリエーション水域での遊泳区域を閉鎖する準備が行われ、水試料中でのシアノバクテリアの検出に基づいて2日後に地表水から井戸水に切り替えるよう上水道の管理者に通知される。

30

【 0 3 1 3 】

(実施例7)

PhyloChipアレイ

アレイシステム「PhyloChip」は、本明細書に記載される方法により選択される生物特異的およびOTU特異的16s rRNAプローブの一部を用いて作製された。前記PhyloChipアレイは、1,008の行と列の格子として配置された1,016,064プローブフィチャーからなっていた。これらのフィチャーのうち、90%は、16s rRNA遺伝子に対してそれぞれ正確なまたは不正確な相補性を有するオリゴヌクレオチドPMまたはMMであった。各プローブはミスマッチ対照プローブと対合して、標的特異的ハイブリダイゼーションをバックグラウンドと非標的クロスハイブリダイゼーションから区別する。残りのプローブは画像方向、標準化制御のためにまたは染色体の追加の標的領域を使用する病原体特異的シグネチャー単位複製配列検出のために使用された。(1)断片化の直前に定義された量で16s rRNA遺伝子単位複製配列混合物にスパイクされた原核生物代謝関連遺伝子の単位複製配列を標的にし、(2)前記ハイブリダイゼーション混合物に添加される前標識されたオリゴヌクレオチドに相補的である追加のプローブを用いて、各高密度16s rRNA遺伝子マイクロアレイが設計された。

40

50

第1の対照は、標的断片化、ビオチン化による標識化、アレイハイブリダイゼーションおよび染色/スキャニング効率を全体として試験した。第1の対照は、実験におけるアレイのすべてにわたり全体的蛍光強度を正規化することも可能にした。第2の対照は、ハイブリダイゼーション、染色およびスキャニングを直接アッセイした。

【0314】

前記プローブ配列に対する相補的標的は前記アレイにハイブリダイズし、蛍光シグナルは、データをプローブごとに個々のシグナル値に変え、典型的には人に解読可能なCELファイルとしてエクスポートする標準AFFYMETRIX(登録商標)ソフトウェア(GeneChip Microarray Analysis Suite, version 5.1)を使用してピクセル画像として取り込まれた。バックグラウンドプローブは、全強度のうちのもっとも低い2%の強度を生み出すプローブとしてCELファイルから同定された。バックグラウンドプローブの平均強度を全プローブの蛍光強度から減算した。ノイズ値(N)は、スキャナがアレイ表面を読み取る際にスキャナにより観察されるピクセル強度シグナルの変動であった。同定されたバックグラウンドプローブ強度それぞれ内のピクセル強度の標準偏差は、そのフィチャーを含むピクセルの数の平方根で割られた。得られた商の平均は下に記載される計算においてNに使用された。

【0315】

前の方法を使用して、陽性として点数化されたプローブペアは、2つの基準、すなわち(i)完全にマッチするプローブ(PM)からの蛍光強度の大きさは、ミスマッチ対照(MM)からの強度の少なくとも1.3倍である、および(ii)強度差、つまりPM引くMMは二乗されたノイズ値の少なくとも130倍である($>130 N^2$)、を満たすプローブペアである。陽性フラクション(PosFrac)は、プローブセットにおけるプローブペアの総数で割った陽性プローブペアの数としてプローブセットごとに計算された。OTUは、対応するプローブセットについてのそのPosFracが0.92よりも大きい場合には「プレゼント」と見なされた(クローンライブラリー分析からの経験的データに基づいて)。複製アレイは、それぞれにPosFrac閾値を超えることを要求することにより各OTUの存在を決定するのに集合的に使用することができた。プレゼントコールは、そのサブ序列OTUのうちの少なくとも1つが存在する場合にはどんなノード(亜科、科、目など)も「プレゼント」と見なすことにより分類階層を通して上方に伝播された。

【0316】

ハイブリダイゼーション強度は、OTU量の尺度であり、所定のプローブセットにおいてプローブペアにわたるPM引くMM強度差のトリム平均(平均化前に最大値と最小値が取り除かれる)としてプローブセットごとに任意単位で計算された。1未満の強度はすべて、それに続く対数変換でのエラーを回避するために1に変換された。

【0317】

実施例1に記載される分析法は、現在記載されているPhyloChip G3アレイに適用された試料にも適用することが可能である。

【0318】

ラテン方格検証(Latin Square Validation)は、PhyloChip G3アレイ上で実施された。既知の細菌および古細菌分類群ごとに複数のプローブを含有する新規のPhyloChipマイクロアレイ(G3)が製造された。前記アレイは、ラテン方格実験計画(Latin Square experimental design)を使用して、数桁にまたがる既知ではあるが無作為に割り当てられた濃度で組み合わせられた26の生物の三重混合物を用いてチャレンジされた。プローブ標的複合体は蛍光強度により定量化された。環境内で群集動態をモニターするために、水試料は点源汚水流出に続く2つの時点でサンフランシスコ湾(CA)から採取された。全16S rRNA遺伝子単位複製配列プール(約1000億分子/時点)が前記アレイを用いて評価された。3つの複製物が、78のラテン方格チップと1つの定量基準唯一対照を用いて異なる日に試験された。単位複製配列濃度範囲は $4.5 \log_{10}$ 超であった。標的濃度は0.25pMから477.79pMであり、ステップあたり37%プラス0 pM増加した(26の異なる濃度)。各チップは26の標的すべてを含有し、それぞれのチップが243ngトータルスパイクについてそれぞれ異なる濃度0~66ngを有していた。ラテン方格行列は示されていない。

【0319】

図14は、16S単位複製配列の濃度対PhyloChip応答を示すグラフである。濃度はPhyloChipハイブリダイゼーションチャンバー内の2を底とした対数ピコモル濃度として表示されている。y軸は、プローブセットにおける複数のパーフェクトマッチプローブの平均である。垂直なエラーバーは、3つの複製物試験の標準偏差を示す。0.98を超える r^2 乗値は、PhyloChip G3アレイが濃度の変化を追跡するその能力において定量的であることを示している。

【0320】

図15と16は、モデルベースの検出が、プローブセットの陽性フラクシオン検出の改良であることを示している。低濃度(2pMまで下がる)はラテン方格においてバックグラウンドから区別される。

10

【0321】

図15は、ペア「応答スコア」 r 分布(新規)に基づく検出アルゴリズム対陽性フラクシオン計算(G2 PhyloChipで以前使用された)のボックスプロット比較である。両プロットにおいて、x軸はスパイクイン16S単位複製配列の濃度である(矢印は2ピコモルで始まり500ピコモルまで伸びている)。y軸は、両方のプロットで0から1の間の範囲である。上のプロットのy軸はプローブセット内のすべてのプローブの中央値 r スコアを表示し、下のプロットのy軸は同じデータセットから陽性フラクシオンを表示している。低濃度の0.25pMでは、両方のプロットがスコアの広範な分布を示し(長いウィスカ参照)、2pMでは上のボックスプロットは、種々の細菌と古細菌種を使用する複数の測定がすべて非常に類似する中央値 r スコアを有していることを示す短いウィスカを有している。陽性フラクシオングラフ上の対応する濃度は広範な陽性フラクシオンスコアを有する。ほぼすべての濃度で、 r スコアは陽性フラクシオンよりも優れている。

20

【0322】

図16は、前記 r スコアメトリック対前記受信者動作特性(R.O.C)プロットによるpfの比較を示す2つのグラフである。下の曲線と比べた上の曲線の急勾配は、前記 r スコアメトリックのほうが偽陽性から真の陽性を前記pfメトリックよりも効率的に区別することが可能であることを実証している。灰色スケールバーは、曲線に沿った各点でのカットオフ値(r スコアまたはpfのどちらについても)を示している。

【0323】

30

前記検証は、前記新規のPhyloChip G3アレイが以前のG2アレイよりも試料における優れた生物検出および定量化が可能であることを示している。

【0324】

(実施例8)

水質試験-同定された汚染発生源

水試料は、本発明のシステムと方法を使用して、前記水試料についてのバイオシグネチャーを得て、それを実施例5に記載されるバイオシグネチャーなどの糞便汚染についてのバイオシグネチャーと比較することにより、糞便汚染による汚染についてアッセイすることが可能である。DNAは、実施例2におけるように、前記試料から抽出し、増幅し、プロセッシングし、分析することが可能である。プローブハイブリダイゼーションによる分析は、実施例2に記載される通りにアレイを使用して、またはビーズに結合しているプローブを用いて実施例4における方法に類似するフローサイトメトリー法を使用することにより実行することが可能である。評価されるプローブごとに、および/または評価されるプローブにより表わされるOTUごとにその存在、非存在および/またはレベルを点数化することが可能である。次に、このデータは、糞便汚染を含む1つまたは複数の汚染菌について1つまたは複数のバイオシグネチャーと比較することが可能である。試験試料のバイオシグネチャーと糞便汚染のバイオシグネチャー間の類似の程度が高い場合は、前記試料は糞便汚染を含有すると判定される。前記バイオシグネチャー間の類似の程度が低い場合は、前記試料は糞便汚染を含有していないと判定される。

40

【0325】

50

現実世界のシナリオでは、3つの発生源のうちのどれが汚染の一因となる可能性がもっとも高いかを決定するために、PhyloChipを使用して3種の潜在的汚染源、すなわち、下水、腐敗層汚泥および畜産廃棄物と比べて汚染水試料中の微生物群集組成を比較した。図7は各PhyloChip結果を二次元空間にプロットしている。前記プロットは、前記汚染水試料(高エンテロコッカス)が発生源群集、この場合は下水に向かうベクトルに沿って下がっていることを明らかにした。

【0326】

この実施例は、発生源が他には知られていない場合には、PhyloChipを使用して公共の水路におけるエンテロコッカス過剰の原因を同定する群集分析の力を例証している。

【0327】

別の例では、2009年2月のサウスリートマリンシティ衛生管理区域からの一次処理下水の764,000ガロン下水流出の地点近くのリチャード湾において2つの水試料が収集された。1つの試料(#3)は流出が始まった24時間後に施設のすぐ近くで収集され、エンテロコッカス、全大腸菌群および大腸菌についての培養ベース糞便インジケータ試験(IDEXX)のための水質基準を大幅に超えていた。第2の試料(#26)は流出が始まった72時間後に150m沖合で収集され、含有するすべての糞便インジケータ細菌の数は無視できる(検出限界未満)程度であった。表層水の試料は1リットルの無菌ビンで収集され、ローレンスバークレイ国立研究所での濾過(収集から5時間以内)まで4 で保存された。750mlの試料はWhatman Anodiscメンブランフィルター(47mm径、0.2 µm孔径)を通して真空濾過され、DNA抽出まで-80 で直ちに保存された。

【0328】

ゲノムDNAは、ビーズ破砕とフェノール/クロロホルム抽出法を使用してフィルターから抽出された。16SリボソームRNA遺伝子は、細菌にはユニバーサルプライマー27F(5'-AGAGTTGATCTCTGGCTCAG-3')と1492R(5'-GGTTACCTTGTTACGACTT-3')を、古細菌には4Fa(5'-TCCGGT TGATCTGCCRG-3')と1492Rを使用してPCRにより増幅された。各PCR反応は1×Ex Taq緩衝液(Takara Bio Inc., Japan)、0.125ユニット/µl Ex Taqポリメラーゼ、0.8mM dNTP混合物、1.0 µg/µl BSA、および300nM 各プライマーおよび0.5 µl 鋳型を含有していた。PCR条件は、95 (3分)、続いて30サイクル95 (30秒)、48~58 (25秒)、72 (2分)、続いて最終伸長72 (10分)であった。各DNA抽出物は、アニーリング温度48~58 の範囲にまたがる8複製25 µl 反応液中で増幅された。アニーリング温度が異なるPCR産物は試料ごとに組み合わされ、Microcon YM-100フィルター(Millipore)を使用して濃縮された。

【0329】

ゲル定量化に続いて、500ngの細菌16S rRNA遺伝子単位複製配列および50ngの古細菌単位複製配列はPhyloChip分析のためにプロセッシングされた。PCR産物は、原核生物および真核生物代謝関連遺伝子ならびに合成16S様遺伝子由来の対照単位複製配列を用いてスパイクされた。この混合物は、20 で10分間および98 で10分間インキュベートすることにより、デオキシリボヌクレアーゼI(0.02 U/µg DNA; Invitrogen)とOne-Phor-All緩衝液を使用して50~200bpまで断片化された。断片の末端標識化は、製造業者の使用説明書によりGeneChip WT Double Stranded DNA Terminal Labelingキット(Affymetrix # 900812)を使用して実現された。断片化された試料は、37 で60分間、続いて10分間の70 ステップでインキュベートすることにより、末端デオキシヌクレオチドトランスフェラーゼとAffymetrix DNA Labeling Reagentを使用して標識された。アレイへのハイブリダイゼーションは、GeneChip Hybridization, Wash, and Stainキット(Affymetrix #900720)を使用して実施された。標識DNA(42 µl)は、Control Oligonucleotide B2 (Affymetrix #900301)、DMSO(最終濃度15.7%)およびMES緩衝液と組み合わされて最終容積130 µlとし、99 で5分間、続いて48 で5分間変性された。次に全反応混合液はPhyloChipに添加され、48 、60rpmで一晩(16時間超)インキュベートされた。前記PhyloChipは続いて洗浄され、GeneChip Fluidics Station 450を使用してAffymetrixプロトコールにより染色され、GeneArray Scannerを使用して走査された。前記走査は、プローブごとに前記データを個々のシグナル値に変える標準Affymetrixソフトウェア(GeneChip Operating Software, version 1.0)

を使用してピクセル画像として取り込まれた。ここに記載される分析アルゴリズムを使用して、多数の分類群がどちらかの試料中に存在すると同定された。さらに、流出から最初の24時間以内に下水流出のすぐ近くの水試料に特有である多くの特徴的な分類群が見つかり、流出の72時間後に150メートル沖合で採取された推定非汚染試料において多くの異なる特徴的な分類群が同定された。下水流出地点(試料3)から同定される分類群の他にもその関連するプローブは、関連する受水域における糞便汚染の同定のための根拠として使用することが可能である。

【0330】

糞便細菌プローブセットおよびこのプローブセットの個々のプローブは、実施例1の方法を使用して作業の各ステップで分析された。前記2つの試料のそれぞれにおいて陽性として同定された全プローブセットおよび異なるものの要約的統計値が決定された(示されていない)。

【0331】

拡散チャンバー試験を用いたPhyloChipの使用により、所定の受水域における動物等の腸内細菌などの所定のマイクロバイオームの運命に関する重要な情報を得ることが可能になる。適切なインジケータ生物の選択を推進するために、拡散チャンバーを使用して異なる受水域などの第2の環境におけるマイクロバイオームのメンバーの生存率を調べることによる。塩水と淡水間のマイクロバイオーム生存プロファイルには大きな差異が存在する。選択された生物の異なる生存率を比較することにより、流出の年齢、たとえば、進行中对数日経っているを突き止めることもできる場合がある。拡散チャンバー試験において数種の生物を使用することはすでに周知であるが、全マイクロバイオーム分析を実施するPhyloChipの能力は以前には得られなかった結果をもたらすことになる。

【0332】

上記の下水試料は、拡散チャンバーを使用する拡散試験にも供された。受水域における生物からの捕食の効果を説明することができるよう、下水マイクロバイオームは受水域と混合された下水マイクロバイオームとともにそれぞれが試験された。

【0333】

(実施例9)

OTUの存在または非存在を判定するためにプローブペア応答のセットを評価する

水の汚水漏れ後の2つの時点で2つの湾の水試料が採取された。各試料由来のDNAが抽出され、PCR増幅され、消化され、標識されPhyloChipにハイブリダイズされた。2つの選択されたヒト糞便OTUに対するプローブセットからの応答パターンは、実例として慎重に調べられた。

溢水3-流出開始の24時間後、施設のすぐ前のくるぶしの深さ

溢水26-流出開始の72時間後、500フィート(約150メートル)沖合

OTU:36742

ss_id:2036742 細菌;ファーミキューテス門; クロストリジウム_SP; クロストリジウム_CL; クロストリジウム; 糞便細菌_FM; sfA; OTU:36742

このOTU中1配列:

DQ805677.1 gg_id:185502ヒト糞便クローンRL306aa189f12

OTU:38712

ss_id: 2038712 細菌: ファーミキューテス門; クロストリジウム_SP; クロストリジウム_CL; クロストリジウム; ルミノコッカス_FM;s fA; OTU:38712

このOTU中1配列:

DQ797288.1 gg_id:188731ヒト糞便クローンRL248_aai97d06

【0334】

図11と12では、プローブ応答が表されており、これらのOTU両方について閾値の使用が示されている。前記PhyloChipは、前記OTU内の特定のDNA標的に相補的な複数のDNAプローブを含有するように設計されている。これらの標的のそれぞれは、異なるA+T含量、異なるT含量を有する場合があり、他のOTUへの推定クロスハイブリダイゼーション可能性を有

している場合がある。これら3つの要因は、OTUの存在または非存在コールへのプローブ強度測定値のデコンボリューション(de-convolution)のために利用される。

【0335】

走査が収集された後、プローブ強度はバックグラウンド減算され、スパイクインに合わせてスケールリングされた。

【0336】

図11は、2つの異なるPhyloChip実験において観察される糞便細菌OTU 36742へのプローブ応答を比較している。前記「強度」棒グラフは、各PMおよびMMプローブからの強度を、ペアとしてグループ化されて、それぞれ青および赤で表示している。OTU 36742は30のプローブペアを有する。強度測定値は5.7から30334.3a.u.(任意単位)までの範囲である。次に、前記PMとMM強度を比較することによりプローブペアごとにペアディファレンススコアdを計算する。たとえば、ペア#6は溢水3についてPM強度9941.4およびMM強度903.4を報告した。

【0337】

【数9】

$$d = 1 - \left(\frac{PM - MM}{PM + MM} \right) = 1 - \left(\frac{9941.4 - 903.4}{9941.4 + 903.4} \right) = 0.166$$

【0338】

この変換を実施することにより、PMプローブとMMプローブ間の差を1つの数字を用いて表すことが可能になる。dの考えうる範囲は、0~2であり、dはPM>MMの場合は0に近づき、PM=MMの場合はd=1であり、PM<MMの場合はdは2に近づく。したがって、ペア#6は溢水3において配列特異的相互作用を表示したが、0.166が0に近いからである。d値は、それぞれのプローブペアのすぐ下に「d」と標識されて棒グラフ上にプロットされる。溢水26における同じプローブペア(#6)は0.870のd値を出したことに注目されたい。これは、0.870が0.166よりも0から離れているためにPM値とMM値間の隔たりがより少ないことを示している。前記プローブ組成はまったく同じであるために(異なる実験条件下で見ればそれは同じプローブペアである)、異なるチップにわたる同じプローブペアからのd値を比較するのは公平である。

【0339】

次のステップでは、プローブペアと様々なヌクレオチド組成との比較を可能にするために、dスコアは正規化される。この変換の目的は、ペアについてのd値が陰性対照(NC、どんな16S rRNA配列への潜在的クロスハイブリダイゼーションもないプローブペア)由来のd値のほうに類似しているのか、定量基準(QS、実験にスパイクされる非16S rRNA遺伝子にマッチするPMを有するプローブペア)である陽性対照由来のd値のほうに類似しているのかがどうかを決定することである。d_{QS}値はその標的のA+T総数とT総数に依存しているために、QSペアはこれらの属性によりクラスにグループ化され、d_{QS}値の別の分布が各々で見出される。d_{NC}値は同じようにグループ化される。各クラス内の応答に変動があるために、分布は前記観察値から評価される。溢水3については例が下に示されている。正規分布をしている陰性対照プローブのd観察を示しているオレンジ色の密度プロットの異なる形に注目されたい。図13に示されているように、クラス「9T 14AT」では、平均d_{NC}はクラス「4T 11AT」よりも大きく、分散もクラス「9T 14AT」の方が大きい。緑色の密度プロット(分布に従うと推定されている)を比較すると、クラス「4T 11AT」についての定量基準はほぼ常にゼロに近いdスコアを出し、クラス「9T 14AT」はより高いdスコアのより多くの観察値を含有する(PMとMM間では差異が少ない)。この例では、クラス「9T 14AT」はNCとQSの両方により共有されるより大きな範囲のdスコアを有することが見てとれる(図13)。

【0340】

次に、OTUプローブセット由来の各d値は、同じクラス由来のd_{QS}とd_{NC}の分布と比較される。たとえば、OTU 36742では、プローブ#6は9個のチミン塩基とチミンかアデニンのどち

らかである14個の塩基を有する(表7)。溢水3では、このペアは0.166のd値を達成した。

【 0 3 4 1 】

【表 6】

表 7 OTU 36742 についての PM 標的およびその T と A+T の総数

ペア#	PM 標的配列	T 総数	A+T 総数
1	TGATTACCTAGGTGTTGGAGGATTG	9	14
2	CAATCCTCCAACACCTAGGTAATCA	5	14
3	ACGCCGCGTAGAGGAAGAAGGTCTT	4	11
4	AAGACCTTCTTCCTCTACGCGGCGT	7	11
5	ATCCTGCGACGCACATAGAAATATG	5	14
6	CATATTTCTATGTGCGTCGCAGGAT	9	14
7	GACACGGCCCAGATTCTTACGGGAG	4	10
8	CTCCCGTAAGAATCTGGGCCGTGTC	6	10
9	TTTTCCTGGTAGTGCAGAGGTAGGC	8	12
10	GCCTACCTCTGCACTACCAGGAAAA	4	12
11	ACCAACTGACGCTGAGGCTTGAAAG	4	12
12	CTTCAAGCCTCAGCGTCAGTTGGT	8	12
13	TTGCTTCCTCCATCTAGTGGACAAC	8	13
14	GTTGTCCACTAGATGGAGGAAGCAA	5	13
15	GAAACAACGTCCCAGTTTGGACTGC	5	12
16	GCAGTCCAAACTGGGACGTTGTTTC	7	12
17	TGTTTCTTTCGGGACGCAGAGACAG	7	12
18	CTGTCTCTGCGTCCCGAAAGAAACA	5	12
19	GGCCCAGATTCTTACGGGAGGCAGC	4	9
20	GCTGCCTCCCGTAAGAATCTGGGCC	5	9
21	CTAATACCGCATTAGAGCCCACAGG	4	12
22	CCTGTGGGCTCTAATGCGGTATTAG	8	12
23	AGGCTTGAAAGTGTGGGTAGCAAAC	5	13
24	GTTTGCTACCCACACTTTCAAGCCT	8	13
25	AGTGGAACAACGGGTGAGTAACACAT	4	13
26	ATGTGTTACTCACCCGTTGTCCACT	9	13
27	GATTACCTAGGTGTTGGAGGATTGA	8	14
28	TCAATCCTCCAACACCTAGGTAATC	6	14
29	ACATGAGGAACCTGCCACATACAGG	3	12
30	CCTGTATGTGGCAGGTTCTCATGT	9	12

【 0 3 4 2 】

プローブ#6についての応答スコア、rを決定するために、d=0.166のプローブがNCの正規分布間(下の密度プロットにおけるオレンジ色)で見出される確率を見つけ、次にd=0.166のプローブペアがQSの分布間で見出される確率を見つけ、次に以下の方程式に従って比率を応答スコアrとして最終的に記録する。

【 0 3 4 3 】

【数 1 0】

$$r = \left(\frac{\text{pdf}_{\text{nc}}(X = d)}{\text{pdf}_{\text{nc}}(X = d) + \text{pdf}_{\text{norm}}(X = d)} \right)$$

【 0 3 4 4 】

r=プローブペアが標的に応答しバックグラウンドには応答していない可能性を測定するための応答スコア

pdf (X=d) =dを標的クラスATx Tyについて評価される 分布から引き出すことができ

る確率

$pdf_{norm}(X=d)$ = dを標的クラスATx Tyについて評価される正規分布から引き出すことができる確率

【 0 3 4 5 】

応答スコア、rは0~1の範囲があり、1はプローブペアは曖昧ではない陽性応答を有することが観察されたことを示している。r=0.5の場合、プローブペア応答はNCとQSに等しく似ており、したがって、応答は曖昧であると見なすことが可能である。OTU 36742由来の例プローブ#6を続けると、

【 0 3 4 6 】

【 数 1 1 】

$$pdf_r(X=0.166)=0.951$$

$$pdf_{norm}(X=0.166)=0.058$$

$$r = \left(\frac{0.951}{0.951 + 0.058} \right) = 0.943$$

【 0 3 4 7 】

OTU 36742 #6についてのr値は、両実験におけるその応答に従って図11にプロットされる。溢水26では、このプローブペアは溢水3ほど「陽性」ではなかった。

【 0 3 4 8 】

ある種のプローブペアについてはrスコア化には選択肢が存在する。第1のより厳密な選択肢は、QSまたはNCからの十分な観察値(使用者定義閾値)が記録されて上に記載される分布を評価してはじめて、rを計算する。この第1の選択肢は、rについてのプロットに関して図12 OTU 38717に示されている。赤色で囲まれたプローブペアは、下に記載される rQ_1 、 rQ_2 および rQ_3 を見つけるのには使用されなかった。第2の選択肢は、クラス特異的モデルが決定されていない場合は必ず、一般的 d_{QS} と d_{NC} モデルを使用して(そのクラスとは関係なくすべてのQSとNCペアを使用して)すべてのプローブペアについてrスコアを計算する。この選択肢は図12に示されていない。第2の選択肢の利点は、分析において使用されるプローブペアの数を増やすことである。第3の選択肢は、所定のアレイについてペア特異的クラスモデルが決定されていない場合に、もっとも近いクラスモデルを使用することを可能にする。たとえば、PhyloChipの実験走査が「異常値」プローブペアを隠し、このせいでクラス「4T 12AT」についてのQSまたはNCに対するペア総数が不十分になるならば、このクラスのペアを「5T 12AT」モデルと比較することができる。2つの選択肢のこのような混成により、多数のペアを観察することが可能になるし、近いクラス特異的応答スコアリングも可能になる。この選択肢も図12には示されていない。

【 0 3 4 9 】

次に、プローブセットについてのすべてのrスコアは、「ステージ1」プローブセット存在/非存在スコアリングにおいて集団的に検討される。OTU 36742についての30のプローブペアのうち、前記rスコアの多くが溢水3では1に近いが、溢水26においては1に近いrスコアはほとんどない(図11)ことに注目されたい。これらの分布を定量的に区別するためには、rスコアはランク付けされ、区切り点(四分位数)、 rQ_1 、 rQ_2 および rQ_3 が、ランク付けされた観察値を4つのサイズの等しいビン(bin)に分割することにより見つけられる。2つの実験にわたる2つのOTUについての計算された四分位数は表8に示されている。この表はプローブセットパフォーマンスを記載している。溢水3 OTU 36742 $rQ_2=0.934$ は、「OTU 36742を標的にするプローブペアのセットのうち、半分が0.934を超えるrスコアを出した」として読み取ることができる。

【 0 3 5 0 】

10

20

30

40

【表 7】

表 8 2つの実験にわたって比較された2つの OTU についての「ステージ1」結果

実験	OTU	rQ_1	rQ_2	rQ_3
溢水 3	36742	0.207	0.934	0.983
溢水 26	36742	0.015	0.172	0.763
溢水 3	38712	0.738	0.953	0.991
溢水 26	38712	0.789	0.985	0.996

【0351】

四分位数は、図11と12において応答スコア(r)の各プロット上で緑色の線として図示されている。OTUが「ステージ1」を通過するためには、以下の基準、すなわち rQ_1 0.200、 rQ_2 0.920、および rQ_3 0.977の3つすべてを満たさなければならない。これらの基準は、ラテン方格データ(本文書には示されていない)から学習された。表8から、溢水26におけるOTU 36742を除いて4つのOTUすべてがステージ1を通過している。

10

【0352】

ステージ1を通過しているOTUのみが、ステージ2スコア化において検討される。ステージ2の目的は、同じアレイ上でステージ1を通過するOTUの群集を検討する際に各応答性プロープペア($r > 0.5$)の特異性を評価することである。これは、その推定クロスハイブリダイゼーション可能性に従って各rスコアにペナルティを課すことにより実現される。ステージ1を通過する多くのOTUに対して推定クロスハイブリダイゼーション可能性を有するプロープペアは、ステージ1を通過する数少ないOTUに対して推定クロスハイブリダイゼーションを有するプロープペアよりも大きな係数によりペナルティを課せられることになる。ペナルティを課せられたスコア、 r_x は、

20

【0353】

【数12】

$$r_x = \frac{r_i}{\text{スカラー } (O_{s1} \cap O_{hi})}$$

【0354】

O_{s1} =ステージ1を通過するOTUのセット

30

O_{hi} =PMプロープにハイブリダイズする推定能力を有するOTUのセット

スカラー($O_{s1} \cap O_{hi}$)=ハイブリダイゼーション可能性を有しステージ1を通過するOTUの総数

として計算される。

【0355】

図12におけるプロープペア10(pp10)はこの効果を例証している。溢水3ではpp10は高rスコア(0.997)を達成した。pp10のPMは11の異なるOTUの配列に潜在的にハイブリダイズすることができ、この11うちの7つはステージ1を通過した(「ペナルティ」と標識された数字の列参照)。したがって、rスコアを7で割れば $r_x=0.142$ が得られる。図11と12で下方を指す矢印は、プロープペアごとのペナルティの大きさを示している。すべてのペナルティを検討した後、 r_x 値はランク付けされ、上記のように四分位数が見つけられる(r_xQ_1 、 r_xQ_2 、 r_xQ_3)。例は表9に示されている。

40

【0356】

【表 8】

表 9 2つの実験にわたって比較された2つの OTU についての「ステージ 1」と「ステージ 2」の結果

実験	OTU	r_{Q_1}	r_{Q_2}	r_{Q_3}	r_{xQ_1}	r_{xQ_2}	r_{xQ_3}
溢水 3	36742	0.207	0.934	0.983	0.200	0.529	0.947
溢水 26	36742	0.015	0.172	0.763	該当なし	該当なし	該当なし
溢水 3	38712	0.738	0.953	0.991	0.080	0.142	0.214
溢水 26	38712	0.789	0.985	0.996	0.158	0.496	0.864

10

【0357】

ここに記載される特定の例では、応答性のみに基づいて、溢水3には糞便細菌OTU 36742が存在したが溢水26には存在しなかったと結論付けることが可能である。溢水3においてのみ糞便細菌OTU 36742はステージ1を通過する。逆に、ルミノコッカスOTU 38712についてのプローブセットは、両溢水についてステージ1分析では応答性であったが、ステージ2における追加の自動分析精密化の後、前記プローブセットは溢水26においてのみ存在すると決定された。ステージ2のカットオフ値: r_{xQ_1} 0.100、 r_{xQ_2} 0.200、および r_{xQ_3} 0.300であり、ラテン方格データ(本文書には示されていない)から経験的に決定された。表9に示されるように、OTU 38712は溢水3においてこれらのカットオフ値を満たさなかった。

【0358】

20

(実施例10)

清浄な海洋水および処理された排水のマイクロバイームシグネチャーは流出水および海洋関連分類群を提供する

モンテシト衛生管理区域排水処理施設(サンタバーバラ、カリフォルニア州)から収集された脱塩流出水の試料および清浄海水(1000m沖合、サンタバーバラ、カリフォルニア州)の試料が1年の期間をかけて収集された。脱塩流出水は、清浄海水試料と同じように、処理と分析前に組み合わされた。試料処理および分析は実施例2に記載されている通りに実施された。脱塩流出水と清浄海水のマイクロバイームシグネチャーが比較された。流出水マイクロバイームには、清浄海水マイクロバイームでは見つからなかった266の分類群(表10)が含まれていた。清浄海水マイクロバイームには、流出水試料では見つからなかった231の分類群(表11)が含まれていた。

30

【0359】

同定された分類群は、それぞれ処理された流出水と清浄海水についての「シグネチャー分類群」を表す。シグネチャー分類群は、生下水、健康な、病気のまたは罹患した患者、食品安全性検査を繰り返して合格している食品加工工場、および日常的に召喚を受けている食品加工工場などの数多くの環境から同定することが可能である。シグネチャー分類群には多くの用途がある。たとえば、新鮮な水試料から生じるマイクロバイーム中の異なる生下水シグネチャー分類群の存在または特定の量は、雨水排水管を介して大量の水が水処理施設に送られるときに起こりうることであり、上流の水処理施設における処理が不十分であることを示すことが可能である。新鮮な水マイクロバイーム中の生下水分類群の存在または量は、下水管の漏水、維持が不適切な腐敗物置き場からの浸み出し、または不法排出も示すことが可能である。

40

【0360】

【表 9 A】

表 10 流出水マイクロバイオーーム

分類群	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目 (Rhodobacterales);ロドバクター科(Rhodobacteraceae);sf_1;6848	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバク ター科;sf_1;7602	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバク ター科;sf_1;6883	10
細菌界;バクテロイデス門;フラボバクテリア綱;フラボバクテリア目;フラボバクテリア科 ;sf_1;5671	
細菌界;バクテロイデス門;フラボバクテリア綱;フラボバクテリア目;フラボバクテリア科 ;sf_1;5695	
細菌界;バクテロイデス門;フラボバクテリア綱;フラボバクテリア目;フラボバクテリア科 ;sf_1;5896	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバク ター科;sf_1;7596	20
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバク ター科;sf_1;6982	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;未分類;未分類;sf_6;7252	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバク ター科;sf_1;7050	
細菌界;バクテロイデス門;フラボバクテリア綱;フラボバクテリア目;フラボバクテリア科 ;sf_1;5919	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバク ター科;sf_1;7288	30
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバク ター科;sf_1;7432	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバク ター科;sf_1;6664	
細菌界;シアノバクテリア門;シアノバクテリア綱;プロクロロン目(Prochlorales);未分類 ;sf_1;5076	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;未分類;未分類;sf_6;7196	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;アルテロモナス目;アルテロ モナス科;sf_1;8517	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;SAR86;未分類;sf_1;9648	40
細菌界;プロテオバクテリア門;未分類;未分類;未分類;sf_20;7365	
細菌界;アクチノバクテリア門;BD2-10 グループ;未分類;未分類;sf_1;1675	
細菌界;シアノバクテリア門;シアノバクテリア綱;クロコプラスト目;クロコプラスト科 ;sf_5;5007	

【表 9 B】

細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;未分類;未分類;sf_6;7510	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;SAR86;未分類;sf_1;9620	
細菌界;未分類;未分類;未分類;未分類;sf_148;5235	
細菌界;シアノバクテリア門;シアノバクテリア綱;サーモシネコッカス目(サーモシネコッカス目);未分類;sf_1;5012	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;クロマチウム目;エクトチオロドスピラ科(Ectothiorhodospiraceae);sf_1;9387	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;未分類;未分類;sf_3;8647	10
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;7054	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;7233	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;7045	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;6960	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;7405	20
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;7329	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;オセアノスピリルム目;アルカニボラクス科;sf_1;9043	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;未分類;未分類;sf_6;7520	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;7499	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;SUP05;未分類;sf_1;8953	30
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;7649	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ブラジリゾビウム目;未分類;sf_1;7143	
細菌界;アクチノバクテリア門;BD2-10 グループ;未分類;未分類;sf_1;1732	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;未分類;未分類;sf_3;9016	
細菌界;プランクトミセス門(Planctomycetes);プランクトミセス綱(Planctomycetacia);プランクトミセス目(Planctomycetales);プランクトミセス科(Planctomycetaceae);sf_3;4654	
細菌界;未分類;未分類;未分類;未分類;sf_148;4970	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;7429	40
細菌界;アクチノバクテリア門;アクチノバクテリア綱;アクチノミセス目;アシドサーマス科(Acidothermaceae);sf_1;1399	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;6894	

【表 9 C】

細菌界;アクチノバクテリア門;アクチノバクテリア綱;アシドミクロビウム目;アシドミクロビウム科(Acidimicrobiaceae);sf_1;1282	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;7033	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;7140	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;7085	10
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;未分類;未分類;sf_6;7421	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;未分類;未分類;sf_6;6858	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;未分類;未分類;sf_3;8333	
細菌界;プロテオバクテリア門;未分類;未分類;未分類;sf_20;7541	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;未分類;未分類;sf_3;9061	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;6796	
細菌界;ファーミキューテス門;クロストリジウム綱;ハロアナエロビウム目(Halanaerobiales);ハロバクテロイデス科;sf_1;887	20
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;6714	
細菌界;バクテロイデス門;スフンゴバクテリア綱;スフンゴバクテリア目;未分類;sf_3;5799	
細菌界;プランクトミケス門;プランクトミケス綱;プランクトミケス目;ピレルラ科(Pirellulaceae);sf_3;4801	
細菌界;バクテロイデス門;バクテロイデス綱;バクテロイデス目;リケネラ科;sf_5;5889	
細菌界;バクテロイデス門;フラボバクテリア綱;フラボバクテリア目;未分類;sf_3;5900	
細菌界;シアノバクテリア門;シアノバクテリア綱;クロコプラスト目;クロコプラスト科;sf_5;4983	30
細菌界;シアノバクテリア門;シアノバクテリア綱;クロコプラスト目;クロコプラスト科;sf_5;5111	
細菌界;シアノバクテリア門;シアノバクテリア綱;クロコプラスト目;クロコプラスト科;sf_5;5156	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;未分類;未分類;sf_3;8805	
細菌界;バクテロイデス門;スフンゴバクテリア綱;スフンゴバクテリア目;フレキシバクター科;sf_19;5404	
細菌界;アクウィフェクス門;アクウィフェクス綱;アクウィフェクス目;ハイドロゲノサーマス科(Hydrogenothermaceae);sf_1;737	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;コンシスティエール目(Consistiales);未分類;sf_5;7504	40
細菌界;バクテロイデス門;バクテロイデス綱;バクテロイデス目;リケネラ科;sf_5;5945	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;7224	

【表 9 D】

細菌界;プロテオバクテリア門;ベータプロテオバクテリア綱;未分類;未分類;sf_3;7923	
細菌界;バクテロイデス門;未分類;未分類;未分類;sf_4;6190	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;コンシスティエール目;未分類;sf_5;7203	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;コンシスティエール目;SAR11;sf_1;7376	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;7590	10
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;リゾビウム目(Rhizobiales);未分類;sf_1;7012	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;未分類;未分類;sf_3;8933	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;6866	
細菌界;シアノバクテリア門;シアノバクテリア綱;クロロプラスト目;クロロプラスト科;sf_5;5166	
細菌界;バクテロイデス門;フラボバクテリア綱;フラボバクテリア目;フラボバクテリア科;sf_1;6104	20
細菌界;シアノバクテリア門;シアノバクテリア綱;クロロプラスト目;クロロプラスト科;sf_5;5221	
細菌界;シアノバクテリア門;シアノバクテリア綱;クロロプラスト目;クロロプラスト科;sf_5;5120	
細菌界;バクテロイデス門;バクテロイデス綱;バクテロイデス目;リケネラ科;sf_5;5947	
細菌界;バクテロイデス門;バクテロイデス綱;バクテロイデス目;未分類;sf_15;6078	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;オセアノスピリルム目;未分類;sf_3;8961	
細菌界;バクテロイデス門;フラボバクテリア綱;フラボバクテリア目;フラボバクテリア科;sf_1;5641	30
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;メチロコックス目(Methylococcales);メチロコックス科(Methylococcaceae);sf_1;8821	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;アシドチオバチルス目;アシドチオバチルス科;sf_1;8913	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;未分類;未分類;sf_3;9456	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ハイフオモナス科(Hyphomonadaceae);sf_1;7584	
細菌界;シアノバクテリア門;未分類;未分類;未分類;sf_5;4993	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;オセアノスピリルム目;ハロモナス科;sf_1;9141	40
細菌界;シアノバクテリア門;シアノバクテリア綱;ゲイトレリネマ目(Geitlerinema);未分類;sf_1;4999	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;6771	

【表 9 E】

細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;オセアノスピリルム目;未分類;sf_3;9010	
細菌界;アシドバクテリウム門;アシドバクテリウム綱-9;未分類;未分類;sf_1;704	
細菌界;OP10;未分類;未分類;未分類;sf_4;728	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;7508	
細菌界;バクテロイデス門;フラボバクテリア綱;フラボバクテリア目;フラボバクテリア科;sf_1;5559	10
細菌界;バクテロイデス門;バクテロイデス綱;バクテロイデス目;未分類;sf_15;5998	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;クロマチウム目;クロマチウム科(Chromatiaceae);sf_1;8407	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;アルテロモナス目;アルテロモナス科;sf_1;9442	
細菌界;ゲマティモナス門;未分類;未分類;未分類;sf_6;2554	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ブラジリゾビウム目;未分類;sf_1;7255	
細菌界;バクテロイデス門;バクテロイデス綱;バクテロイデス目;リケネラ科;sf_5;6317	20
細菌界;アクチノバクテリア門;アクチノバクテリア綱;アクチノミセス目;ミクロコッカス科;sf_1;1266	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;7049	
細菌界;プロテオバクテリア門;イプシロンプロテオバクテリア綱;カンピロバクター目;ヘリコバクター科;sf_3;10534	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;7362	
細菌界;バクテロイデス門;フラボバクテリア綱;フラボバクテリア目;フラボバクテリア科;sf_1;5955	30
細菌界;プロテオバクテリア門;未分類;未分類;未分類;sf_21;8509	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;7373	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;GAO クラスター;未分類;sf_1;9008	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;7032	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;6661	40
細菌界;バクテロイデス門;未分類;未分類;未分類;sf_4;5637	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;エンテロバクター目;エンテロバクター科;sf_1;9309	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;6979	

【表 9 F】

細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;アルテロモナス目;アルテロモナス科;sf_1;9236	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;リゾビウム目;フィロバクテリア科(Phyllobacteriaceae);sf_1;7009	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;アルテロモナス目;アルテロモナス科;sf_1;9486	
細菌界;シアノバクテリア門;シアノバクテリア綱;ネンジュモ目;未分類;sf_1;5174	10
細菌界;シアノバクテリア門;シアノバクテリア綱;ネンジュモ目;未分類;sf_1;5028	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;未分類;未分類;sf_3;8883	
細菌界;クロロフレクサス門;アナエロリネア綱;未分類;未分類;sf_9;94	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;7523	
細菌界;バクテロイデス門;フラボバクテリア綱;フラボバクテリア目;フラボバクテリア科;sf_1;5490	
細菌界;シアノバクテリア門;シアノバクテリア綱;ネンジュモ目;未分類;sf_1;5175	20
細菌界;ウェルコミクロビウム門;ウェルコミクロビウム綱;ウェルコミクロビウム目;ウェルコミクロビウム亜門 7;sf_1;760	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;コンシスティエール目;SAR11;sf_2;7043	
細菌界;クロロフレクサス門;アナエロリネア綱;クロロフレクサス門-1f; 未分類;sf_1;765	
細菌界;プロテオバクテリア門;未分類;未分類;未分類;sf_28;10091	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱 GAO クラスター;未分類;sf_1;8980	
細菌界;アクウィフェクス門;アクウィフェクス綱;アクウィフェクス目;ハイドロゲノサーマス科;sf_1;220	30
細菌界;バクテロイデス門;スフンゴバクテリア綱;スフンゴバクテリア目;スフンゴバクテリア科;sf_1;5492	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;アルテロモナス目;アルテロモナス科;sf_1;8863	
細菌界;シアノバクテリア門;シアノバクテリア綱;スピルリナ目(Spirulina);未分類;sf_1;5034	
細菌界;バクテロイデス門;フラボバクテリア綱;フラボバクテリア目;フラボバクテリア科;sf_1;5499	
細菌界;ゲマティモナス門;未分類;未分類;未分類;sf_5;227	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;スフィンゴモナス目;スフィンゴモナス科;sf_1;7110	40
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;7125	
細菌界;シアノバクテリア門;シアノバクテリア綱;クロロプラスト目;クロロプラスト科;sf_5;5130	

【表 9 G】

細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;7536	
細菌界;未分類;未分類;未分類;未分類;sf_92;9999	
細菌界;プロテオバクテリア門;デルタプロテオバクテリア綱;未分類;未分類;sf_9;9993	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;6805	
細菌界;未分類;未分類;未分類;未分類;sf_148;5022	
細菌界;バクテロイデス門;バクテロイデス綱;バクテロイデス目;バクテロイデス科;sf_12;5950	10
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;7493	
細菌界;ウェルコミクロビウム門;ウェルコミクロビウム綱;ウェルコミクロビウム目;ウェルコミクロビウム亜門;sf_1;533	
細菌界;プロテオバクテリア門;デルタプロテオバクテリア綱;デスルフォバクター目;デスルフォバクター科;sf_5;9777	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;未分類;未分類;sf_6;6986	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;6679	20
細菌界;シアノバクテリア門;シアノバクテリア綱;ネンジュモ目;未分類;sf_1;5072	
細菌界;シアノバクテリア門;シアノバクテリア綱;ネンジュモ目;未分類;sf_1;5199	
細菌界;シアノバクテリア門;シアノバクテリア綱;ネンジュモ目;未分類;sf_1;5191	
細菌界;シアノバクテリア門;シアノバクテリア綱;ネンジュモ目;未分類;sf_1;5047	
細菌界;バクテロイデス門;フラボバクテリア綱;フラボバクテリア目;フラボバクテリア科;sf_1;5509	
細菌界;バクテロイデス門;フラボバクテリア綱;フラボバクテリア目;クライオモルファ科(Cryomorphaceae);sf_1;5400	30
細菌界;バクテロイデス門;フラボバクテリア綱;フラボバクテリア目;フラボバクテリア科;sf_1;5301	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;フルビマリナ目(Fulvimarina);未分類;sf_1;7281	
細菌界;プロテオバクテリア門;イプシロンプロテオバクテリア綱;カンピロバクター目;ヘリコバクター科;sf_3;10614	
細菌界;ファーミキューテス門;モリクテス綱;マイコプラズマ目(Mycoplasmatales);マイコプラズマ科(Mycoplasmataceae);sf_1;4102	
細菌界;ディクチオグロムス門(Dictyoglomi);ディクチオグロムス綱;ディクチオグロムス目(Dictyoglomales);ディクチオグロムス科(Dictyoglomaceae);sf_9;7579	40
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;アルテロモナス目;アルテロモナス科;sf_1;9586	
細菌界;シアノバクテリア門;シアノバクテリア綱;ネンジュモ目;未分類;sf_1;5004	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;7383	

【表 9 H】

細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;アルテロモナス目;アルテロモナス科;sf_1;8533	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;アルテロモナス目;アルテロモナス科;sf_1;9247	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;アルテロモナス目;アルテロモナス科;sf_1;8600	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;312	10
細菌界;ウェルコミクロビウム門;ウェルコミクロビウム綱;ウェルコミクロビウム目;ウェルコミクロビウム科;sf_6;203	
細菌界;アクチノバクテリア門;アクチノバクテリア綱;アクチノミセス目;マイクロバクテリウム科(Microbacteriaceae);sf_1;1135	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ユーバクテリウム科;sf_1;28	
細菌界;シアノバクテリア門;シアノバクテリア綱;シュードアナベナ目;未分類;sf_1;5008	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;6955	20
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;7084	
細菌界;バクテロイデス門;スフンゴバクテリア綱;スフンゴバクテリア目;スフンゴバクテリア科;sf_1;6250	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;7560	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;7211	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;6784	30
細菌界;バクテロイデス門;スフンゴバクテリア綱;スフンゴバクテリア目;フレキシバクテリウム科;sf_19;6261	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;6827	
細菌界;シアノバクテリア門;シアノバクテリア綱;クロロプラスト目;クロロプラスト科;sf_5;5060	
細菌界;OD1;OP11-5;未分類;未分類;sf_1;515	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;未分類;未分類;sf_6;7107	
細菌界;プロテオバクテリア門;デルタプロテオバクテリア綱;ミキソコッカス目;ポリアンギウム科;sf_3;10298	40
細菌界;アクチノバクテリア門;アクチノバクテリア綱;未分類;未分類;sf_1;1370	
細菌界;クロロフレキサス門;テルモミクロビウム綱(Thermomicrobia);未分類;未分類;sf_2;652	

【表 9 I】

細菌界;バクテロイデス門;バクテロイデス綱;バクテロイデス目;プレボテラ科;sf_1;6152	
細菌界;スピロヘータ門;スピロヘータ綱;スピロヘータ目;スピロヘータ科;sf_1;6458	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;7262	
細菌界;ウェルコミクロビウム門;ウェルコミクロビウム綱;ウェルコミクロビウム目;ウェルコミクロビウム科;sf_6;871	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;アルテロモナス目;アルテロモナス科;sf_1;9491	10
細菌界;バクテロイデス門;スフンゴバクテリア綱;スフンゴバクテリア目;フレキシバクター科;sf_19;5728	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;スフィンゴモナス目;スフィンゴモナス科;sf_1;7576	
細菌界;ウェルコミクロビウム門;ウェルコミクロビウム綱;ウェルコミクロビウム目;ウェルコミクロビウム科;sf_7;29	
細菌界;クロロビウム門;未分類;未分類;未分類;sf_6;5294	
細菌界;シアノバクテリア門;シアノバクテリア綱;クロコプラスト目;クロコプラスト科;sf_5;5039	20
細菌界;バクテロイデス門;フラボバクテリア綱;フラボバクテリア目;フラボバクテリア科;sf_1;5758	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;オセアノスピリルム目;ハロモナス科;sf_1;9446	
細菌界;ゲマティモナス門;未分類;未分類;未分類;sf_5;1127	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;クロストリジウム科;sf_12;4156	
細菌界;バクテロイデス門;スフンゴバクテリア綱;スフンゴバクテリア目;毛状菌科;sf_11;5463	
細菌界;シアノバクテリア門;シアノバクテリア綱;プレクトネマ目;未分類;sf_1;5010	30
細菌界;バクテロイデス門;スフンゴバクテリア綱;スフンゴバクテリア目;フレキシバクター科;sf_19;5994	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;エンテロバクター目;エンテロバクター科;sf_1;8173	
細菌界;TM7;未分類;未分類;未分類;sf_1;3025	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;リゾビウム目;未分類;sf_1;7339	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;オセアノスピリルム目;ハロモナス科;sf_1;8598	40
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;リゾビウム目;ブラジリゾビウム科;sf_1;7096	
細菌界;シアノバクテリア門;シアノバクテリア綱;クロコプラスト目;クロコプラスト科;sf_5;5006	

【表 9 J】

細菌界;未分類;未分類;未分類;未分類;sf_132;9820	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;6862	
細菌界;シアノバクテリア門;シアノバクテリア綱;プレクトネマ目;未分類;sf_1;5210	
細菌界;ゲマティモナス門;未分類;未分類;未分類;sf_5;2047	
細菌界;アクチノバクテリア門;アクチノバクテリア綱;アクチノミセス目;マイクロバクテリウム科;sf_1;1186	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;7364	10
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;7453	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;リゾビウム目;ブルセラ科;sf_1;6757	
細菌界;プロテオバクテリア門;デルタプロテオバクテリア綱;デススルフロモナス目;ゲオバクター科;sf_1;482	
細菌界;プロテオバクテリア門;デルタプロテオバクテリア綱;デスルフォバクター目;デスルフォバクター科;sf_5;10136	20
細菌界;シアノバクテリア門;シアノバクテリア綱;クロオコッカス目;未分類;sf_1;5219	
細菌界;クロロビウム門;クロロビウム綱;クロロビウム目;クロロビウム科;sf_1;995	
細菌界;アシドバクテリウム門;アシドバクテリウム綱;アシドバクテリウム目;アシドバクテリウム科;sf_16;6414	
細菌界;ウェルコミクロビウム門;ウェルコミクロビウム綱;ウェルコミクロビウム目;ウェルコミクロビウム科 sf_6;613	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;コンシスティエール目;未分類;sf_5;7592	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;リゾビウム目;未分類;sf_1;6726	30
細菌界;シアノバクテリア門;シアノバクテリア綱;クロロプラスト目;クロロプラスト科;sf_5;5182	
細菌界;アクチノバクテリア門;アクチノバクテリア綱;アクチノミセス目;キネオスポリア科;sf_1;1598	

【 0 3 7 0 】

【表 10 A】

表 11 清浄海水マイクロビオーム

分類群	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;6848	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;7602	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;6883	10
細菌界;バクテロイデス門;フラボバクテリア綱;フラボバクテリア目;フラボバクテリア科;sf_1;5671	
細菌界;バクテロイデス門;フラボバクテリア綱;フラボバクテリア目;フラボバクテリア科;sf_1;5695	
細菌界;バクテロイデス門;フラボバクテリア綱;フラボバクテリア目;フラボバクテリア科;sf_1;5896	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;7596	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;6982	20
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;未分類;未分類;sf_6;7252	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;7050	
細菌界;バクテロイデス門;フラボバクテリア綱;フラボバクテリア目;フラボバクテリア科;sf_1;5919	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;7288	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;7432	30
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;6664	
細菌界;シアノバクテリア門;シアノバクテリア綱;プロクロロン目;未分類;sf_1;5076	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;未分類;未分類;sf_6;7196	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;アルテロモナス目;アルテロモナス科;sf_1;8517	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;SAR86;未分類;sf_1;9648	
細菌界;プロテオバクテリア門;未分類;未分類;未分類;sf_20;7365	40
細菌界;アクチノバクテリア門;BD2-10 グループ;未分類;未分類;sf_1;1675	
細菌界;シアノバクテリア門;シアノバクテリア綱;クロコプラスト目;クロコプラスト科;sf_5;5007	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;未分類;未分類;sf_6;7510	

【 0 3 7 1 】

【表 10B】

細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;SAR86;未分類;sf_1;9620	
細菌界;未分類;未分類;未分類;未分類;sf_148;5235	
細菌界;シアノバクテリア門;シアノバクテリア綱;サーモシネコッカス目;未分類;sf_1;5012	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;クロマチウム目;エクトチオロドスピラ科;sf_1;9387	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;未分類;未分類;sf_3;8647	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;7054	10
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;7233	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;7045	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;6960	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;7405	20
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;7329	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;オセアノスピリルム目;アルカニボラクス科;sf_1;9043	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;未分類;未分類;sf_6;7520	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;7499	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;SUP05;未分類;sf_1;8953	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;7649	30
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ブラジリゾビウム目;未分類;sf_1;7143	
細菌界;アクチノバクテリア門;BD2-10 グループ;未分類;未分類;sf_1;1732	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;未分類;未分類;sf_3;9016	
細菌界;プランクトミケス門;プランクトミケス綱;プランクトミケス綱;プランクトミケス科;sf_3;4654	
細菌界;未分類;未分類;未分類;未分類;sf_148;4970	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;7429	40
細菌界;アクチノバクテリア門;アクチノバクテリア綱;アクチノミセス目;アシドサーマス科;sf_1;1399	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;6894	

【 0 3 7 2 】

【表 10C】

細菌界;アクチノバクテリア門;アクチノバクテリア綱;アシドミクロビウム目;アシドミクロビウム科;sf_1;1282	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;7033	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;7140	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;7085	10
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;未分類;未分類;sf_6;7421	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;未分類;未分類;sf_6;6858	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;未分類;未分類;sf_3;8333	
細菌界;プロテオバクテリア門;未分類;未分類;未分類;sf_20;7541	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;未分類;未分類;sf_3;9061	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;6796	
細菌界;ファーミキューテス門;クロストリジウム綱;ハロアナエロビウム目;ハロバクテロイデス科;sf_1;887	20
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;6714	
細菌界;バクテロイデス門;スフンゴバクテリア綱;スフンゴバクテリア目;未分類;sf_3;5799	
細菌界;プランクトミケス門;プランクトミケス綱;プランクトミケス綱;ピレルラ科;sf_3;4801	
細菌界;バクテロイデス門;バクテロイデス綱;バクテロイデス目;リケネラ科;sf_5;5889	
細菌界;バクテロイデス門;フラボバクテリア綱;フラボバクテリア目;未分類;sf_3;5900	
細菌界;シアノバクテリア門;シアノバクテリア綱;クロコプラスト目;クロコプラスト科;sf_5;4983	30
細菌界;シアノバクテリア門;シアノバクテリア綱;クロコプラスト目;クロコプラスト科;sf_5;5111	
細菌界;シアノバクテリア門;シアノバクテリア綱;クロコプラスト目;クロコプラスト科;sf_5;5156	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;未分類;未分類;sf_3;8805	
細菌界;バクテロイデス門;スフンゴバクテリア綱;スフンゴバクテリア目;フレキシバクター科;sf_19;5404	
細菌界;アクウィフェクス門;アクウィフェクス綱;アクウィフェクス目;ハイドロゲノサーマス科;sf_1;737	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;コンシスティエール目;未分類;sf_5;7504	40
細菌界;バクテロイデス門;バクテロイデス綱;バクテロイデス目;リケネラ科;sf_5;5945	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;7224	

【 0 3 7 3 】

【表 10D】

細菌界;プロテオバクテリア門;ベータプロテオバクテリア綱;未分類;未分類;sf_3;7923	
細菌界;バクテロイデス門;未分類;未分類;未分類;sf_4;6190	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;コンシスティエール目;未分類;sf_5;7203	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;コンシスティエール目;SAR11;sf_1;7376	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;7590	10
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;リゾビウム目;未分類;sf_1;7012	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;未分類;未分類;sf_3;8933	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;6866	
細菌界;シアノバクテリア門;シアノバクテリア綱;クロコプラスト目;クロコプラスト科;sf_5;5166	
細菌界;バクテロイデス門;フラボバクテリア綱;フラボバクテリア目;フラボバクテリア科;sf_1;6104	20
細菌界;シアノバクテリア門;シアノバクテリア綱;クロコプラスト目;クロコプラスト科;sf_5;5221	
細菌界;シアノバクテリア門;シアノバクテリア綱;クロコプラスト目;クロコプラスト科;sf_5;5120	
細菌界;バクテロイデス門;バクテロイデス綱;バクテロイデス目;リケネラ科;sf_5;5947	
細菌界;バクテロイデス門;バクテロイデス綱;バクテロイデス目;未分類;sf_15;6078	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;オセアノスピリルム目;未分類;sf_3;8961	
細菌界;バクテロイデス門;フラボバクテリア綱;フラボバクテリア目;フラボバクテリア科;sf_1;5641	30
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;メチロコックス目;メチロコックス科;sf_1;8821	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;アシドチオバチルス目;アシドチオバチルス科;sf_1;8913	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;未分類;未分類;sf_3;9456	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ハイフオモナス科;sf_1;7584	
細菌界;シアノバクテリア門;未分類;未分類;未分類;sf_5;4993	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;オセアノスピリルム目;ハロモナス科;sf_1;9141	40
細菌界;シアノバクテリア門;シアノバクテリア綱;ゲイトレリネマ目;未分類;sf_1;4999	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;6771	

【 0 3 7 4 】

【表 10E】

細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;オセアノスピリルム目;未分類;sf_3;9010	
細菌界;アシドバクテリウム門;アシドバクテリウム綱-9;未分類;未分類;sf_1;704	
細菌界;OP10;未分類;未分類;未分類;sf_4;728	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;7508	
細菌界;バクテロイデス門;フラボバクテリア綱;フラボバクテリア目;フラボバクテリア科;sf_1;5559	10
細菌界;バクテロイデス門;バクテロイデス綱;バクテロイデス目;未分類;sf_15;5998	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;クロマチウム目;クロマチウム科;sf_1;8407	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;アルテロモナス目;アルテロモナス科;sf_1;9442	
細菌界;ゲマティモナス門;未分類;未分類;未分類;sf_6;2554	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ブラジリゾビウム目;未分類;sf_1;7255	
細菌界;バクテロイデス門;バクテロイデス綱;バクテロイデス目;リケネラ科;sf_5;6317	20
細菌界;アクチノバクテリア門;アクチノバクテリア綱;アクチノミセス目;マイクロコッカス科;sf_1;1266	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;7049	
細菌界;プロテオバクテリア門;イブシロンプロテオバクテリア綱;カンピロバクター目;ヘリコバクター科;sf_3;10534	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;7362	
細菌界;バクテロイデス門;フラボバクテリア綱;フラボバクテリア目;フラボバクテリア科;sf_1;5955	30
細菌界;プロテオバクテリア門;未分類;未分類;未分類;sf_21;8509	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;7373	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;GAO クラスター;未分類;sf_1;9008	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;7032	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;6661	40
細菌界;バクテロイデス門;未分類;未分類;未分類;sf_4;5637	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;エンテロバクター目;エンテロバクター科;sf_1;9309	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;6979	

【 0 3 7 5 】

【表 10F】

細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;アルテロモナス目;アルテロモナス科;sf_1;9236	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;リゾビウム目;フィロバクテリア科;sf_1;7009	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;アルテロモナス目;アルテロモナス科;sf_1;9486	
細菌界;シアノバクテリア門;シアノバクテリア綱;ネンジュモ目;未分類;sf_1;5174	
細菌界;シアノバクテリア門;シアノバクテリア綱;ネンジュモ目;未分類;sf_1;5028	10
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;未分類;未分類;sf_3;8883	
細菌界;クロロフレクサス門;アナエロリネア綱;未分類;未分類;sf_9;94	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;7523	
細菌界;バクテロイデス門;フラボバクテリア綱;フラボバクテリア目;フラボバクテリア科;sf_1;5490	
細菌界;シアノバクテリア門;シアノバクテリア綱;ネンジュモ目;未分類;sf_1;5175	
細菌界; ウェルコミクロビウム門;ウェルコミクロビウム綱;ウェルコミクロビウム目;ウェルコミクロビウム亜門 7;sf_1;760	20
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;コンシスティエール目;SAR11;sf_2;7043	
細菌界;クロロフレクサス門;アナエロリネア綱;クロロフレクサス門-1f;未分類;sf_1;765	
細菌界;プロテオバクテリア門;未分類;未分類;未分類;sf_28;10091	
細菌界; プロテオバクテリア門;ガンマプロテオバクテリア綱;GAO クラスター;未分類;sf_1;8980	
細菌界;アクウィフェクス門;アクウィフェクス綱;アクウィフェクス目;ハイドロゲノサーマス科;sf_1;220	
細菌界;バクテロイデス門;スフンゴバクテリア綱;スフンゴバクテリア目;スフンゴバクテリア科;sf_1;5492	30
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;アルテロモナス目;アルテロモナス科;sf_1;8863	
細菌界;シアノバクテリア門;シアノバクテリア綱;スピルリナ目;未分類;sf_1;5034	
細菌界;バクテロイデス門;フラボバクテリア綱;フラボバクテリア目;フラボバクテリア科;sf_1;5499	
細菌界;グマティモナス門;未分類;未分類;未分類;sf_5;227	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;スフィンゴモナス目;スフィンゴモナス科;sf_1;7110	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;7125	40
細菌界;シアノバクテリア門;シアノバクテリア綱;クロコプラスト目;クロコプラスト科;sf_5;5130	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;7536	

【 0376 】

【表 1 0 G】

細菌界;未分類;未分類;未分類;未分類;sf_92;9999	
細菌界;プロテオバクテリア門;デルタプロテオバクテリア綱;未分類;未分類;sf_9;9993	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;6805	
細菌界;未分類;未分類;未分類;未分類;sf_148;5022	
細菌界;バクテロイデス門;バクテロイデス綱;バクテロイデス目;バクテロイデス科;sf_12;5950	10
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;7493	
細菌界;;ウェルコミクロビウム門;ウェルコミクロビウム綱;ウェルコミクロビウム目;ウェルコミクロビウム亜門 5;sf_1;533	
細菌界;プロテオバクテリア門;デルタプロテオバクテリア綱;デスルフォバクター目;デスルフォバクター科;sf_5;9777	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;未分類;未分類;sf_6;6986	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;6679	
細菌界;シアノバクテリア門;シアノバクテリア綱;ネンジュモ目;未分類;sf_1;5072	20
細菌界;シアノバクテリア門;シアノバクテリア綱;ネンジュモ目;未分類;sf_1;5199	
細菌界;シアノバクテリア門;シアノバクテリア綱;ネンジュモ目;未分類;sf_1;5191	
細菌界;シアノバクテリア門;シアノバクテリア綱;ネンジュモ目;未分類;sf_1;5047	
細菌界;バクテロイデス門;フラボバクテリア綱;フラボバクテリア目;フラボバクテリア科;sf_1;5509	
細菌界;バクテロイデス門;フラボバクテリア綱;フラボバクテリア目;クライオモルファ科;sf_1;5400	
細菌界;バクテロイデス門;フラボバクテリア綱;フラボバクテリア目;フラボバクテリア科;sf_1;5301	30
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;フルビマリナ目;未分類;sf_1;7281	
細菌界;プロテオバクテリア門;イプシロンプロテオバクテリア綱;カンピロバクター目;ヘリコバクター科;sf_3;10614	
細菌界;ファーミキューテス門;モリクテス綱;マイコプラズマ目;マイコプラズマ科;sf_1;4102	
細菌界;ディクチオグロムス門;ディクチオグロムス綱;ディクチオグロムス目;ディクチオグロムス科;sf_9;7579	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;アルテロモナス目;アルテロモナス科;sf_1;9586	40
細菌界;シアノバクテリア門;シアノバクテリア綱;ネンジュモ目;未分類;sf_1;5004	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;7383	

【 0 3 7 7 】

【表 10H】

細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;アルテロモナス目;アルテロモナス科;sf_1;8533	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;アルテロモナス目;アルテロモナス科;sf_1;9247	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;アルテロモナス目;アルテロモナス科;sf_1;8600	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;312	10
細菌界; ウェルコミクロビウム門;ウェルコミクロビウム綱;ウェルコミクロビウム目;ウェルコミクロビウム科;sf_6;203	
細菌界;アクチノバクテリア門;アクチノバクテリア綱;アクチノミセス目;マイクロバクテリウム科;sf_1;1135	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;ユーバクテリウム科;sf_1;28	
細菌界;シアノバクテリア門;シアノバクテリア綱;シュードアナベナ目;未分類;sf_1;5008	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;6955	20
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;7084	
細菌界;バクテロイデス門;スフンゴバクテリア綱;スフンゴバクテリア目;スフンゴバクテリア科;sf_1;6250	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;7560	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;7211	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;6784	30
細菌界;バクテロイデス門;スフンゴバクテリア綱;スフンゴバクテリア目;フレキシバクター科;sf_19;6261	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;6827	
細菌界;シアノバクテリア門;シアノバクテリア綱;クロロプラスト目;クロロプラスト科;sf_5;5060	
細菌界;OD1;OP11-5;未分類;未分類;sf_1;515	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;未分類;未分類;sf_6;7107	
細菌界;プロテオバクテリア門;デルタプロテオバクテリア綱;ミキソコッカス目;ポリアンギウム科;sf_3;10298	40
細菌界;アクチノバクテリア門;アクチノバクテリア綱;未分類;未分類;sf_1;1370	
細菌界;クロロフレクサス門;テルモミクロビウム綱;未分類;未分類;sf_2;652	

【 0 3 7 8 】

【表 10 I】

細菌界;バクテロイデス門;バクテロイデス綱;バクテロイデス目;プレボテラ科;sf_1;6152	
細菌界;スピロヘータ門;スピロヘータ綱;スピロヘータ目;スピロヘータ科;sf_1;6458	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;7262	
細菌界; ウェルコミクロビウム門;ウェルコミクロビウム綱;ウェルコミクロビウム目;ウェルコミクロビウム科;sf_6;871	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;アルテロモナス目;アルテロモナス科;sf_1;9491	10
細菌界;バクテロイデス門;スフンゴバクテリア綱;スフンゴバクテリア目;フレキシバクター科;sf_19;5728	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;スフィンゴモナス目;スフィンゴモナス科;sf_1;7576	
細菌界; ウェルコミクロビウム門;ウェルコミクロビウム綱;ウェルコミクロビウム目;ウェルコミクロビウム科;sf_7;29	
細菌界;クロロビウム門;未分類;未分類;未分類;sf_6;5294	
細菌界;シアノバクテリア門;シアノバクテリア綱;クロコプラスト目;クロコプラスト科;sf_5;5039	20
細菌界;バクテロイデス門;フラボバクテリア綱;フラボバクテリア目;フラボバクテリア科;sf_1;5758	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;オセアノスピリルム目;ハロモナス科;sf_1;9446	
細菌界;ゲマティモナス門;未分類;未分類;未分類;sf_5;1127	
細菌界;ファーミキューテス門;クロストリジウム綱;クロストリジウム目;クロストリジウム科;sf_12;4156	
細菌界;バクテロイデス門;スフンゴバクテリア綱;スフンゴバクテリア目;毛状菌科;sf_11;5463	
細菌界;シアノバクテリア門;シアノバクテリア綱;プレクトネマ目;未分類;sf_1;5010	30
細菌界;バクテロイデス門;スフンゴバクテリア綱;スフンゴバクテリア目;フレキシバクター科;sf_19;5994	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;エンテロバクター目;エンテロバクター科;sf_1;8173	
細菌界;TM7;未分類;未分類;未分類;sf_1;3025	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;リゾビウム目;未分類;sf_1;7339	
細菌界;プロテオバクテリア門;ガンマプロテオバクテリア綱;オセアノスピリルム目;ハロモナス科;sf_1;8598	40
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱; リゾビウム目;ブラジリソビウム科;sf_1;7096	
細菌界;シアノバクテリア門;シアノバクテリア綱;クロコプラスト目;クロコプラスト科;sf_5;5006	

【 0 3 7 9 】

【表 10 J】

細菌界;未分類;未分類;未分類;未分類;sf_132;9820	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;6862	
細菌界;シアノバクテリア門;シアノバクテリア綱;プレクトネマ目;未分類;sf_1;5210	
細菌界;グマティモナス門;未分類;未分類;未分類;sf_5;2047	
細菌界;アクチノバクテリア門;アクチノバクテリア綱;アクチノミセス目;マイクロバクテリウム科;sf_1;1186	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;7364	10
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;ロドバクター目;ロドバクター科;sf_1;7453	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;リゾビウム目;ブルセラ科;sf_1;6757	
細菌界;プロテオバクテリア門;デルタプロテオバクテリア綱;デススフロモナス目;ゲオバクター科;sf_1;482	
細菌界;プロテオバクテリア門;デルタプロテオバクテリア綱;デスルフォバクター目;デスルフォバクター科;sf_5;10136	20
細菌界;シアノバクテリア門;シアノバクテリア綱;クロオコッカス目;未分類;sf_1;5219	
細菌界;クロロビウム門;クロロビウム綱;クロロビウム目;クロロビウム科;sf_1;995	
細菌界;アシドバクテリウム門;アシドバクテリウム綱;アシドバクテリウム目;アシドバクテリウム科;sf_16;6414	
細菌界;ウェルコミクロビウム門;ウェルコミクロビウム綱;ウェルコミクロビウム目;ウェルコミクロビウム科;sf_6;613	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;コンシスティエール目;未分類;sf_5;7592	
細菌界;プロテオバクテリア門;アルファプロテオバクテリア綱;リゾビウム目;未分類;sf_1;6726	30
細菌界;シアノバクテリア門;シアノバクテリア綱;クロコプラスト目;クロコプラスト科;sf_5;5182	
細菌界;アクチノバクテリア門;アクチノバクテリア綱;アクチノミセス目;キネオスポリア科;sf_1;1598	

【 0 3 8 0 】

(実施例11)

クリーンルーム質試験

従来のクリーンルーム試験は、拭取り法とペトリ皿におけるいかなる孢子増殖であれ観察することに頼っている。拭取り法を使用して検出される微生物群集と実施例7のPhyloChipにより検出される微生物群集の比較が示されている。

【 0 3 8 1 】

様々なクリーンルームおよび衛星または宇宙船表面およびその試料に適用された合計で125のワイプが比較された。各試料はそれぞれ約250mLあり、したがって、試料を濃縮するのは困難である。前記試料は0.45 μmフィルター続いて0.2 μmフィルターを使用して濾過された。こうして得られた10mL液体はアミコンフィルターを使用して濃縮された。DNAはMaxwell抽出器を使用して抽出された。

【 0 3 8 2 】

【表 1 1】

表 12 孢子カウントによる 14 のプールされた試料

孢子総数	プールされた試料の数および ID 番号
孢子のない試料セット:7セット	GI-15, 16, 17, 25, 26, 27, 28
孢子総数: 1:	GI-18 (10 プールされた試料)
孢子総数 2 から 4:	GI-19 (15 プールされた試料)
孢子総数 5 から 9:	GI-20 (5 プールされた試料)
孢子総数 10 から 11:	GI-21 (4 プールされた試料)
孢子総数 32:	GI-22 (1 試料)
孢子総数 59:	GI-23 (1 試料)
孢子総数 151:	GI-24 (1 試料)

10

【 0 3 8 3 】

ここで図8を参照すると、ペトリ皿法では、前記PhyloChipにより見つけれられる微生物群集の多様性を予測していない。孢子カウント法により検出される孢子がたとえゼロの場合でも、PhyloChipはOTUを検出する。示されているように、実施例1および2に記載されている試験と分析の方法を使用して最大650のOTUが検出された。孢子総数とPhyloChip OTU総数の間には何の関係も観察されていない。

【 0 3 8 4 】

図9を参照すると、PhyloChipは、試料が共有しているどんな微生物科でもまたはユニークであるどんな微生物科でも検出することができる。図9は、共通のまたはユニークな科を示す試料の図画ネットワークを示している。暗い点は試料であり、明るい点は検出された科である。2つの科、すなわちシュードモナスとラルストニアは、大半の試料において見つかった。1つの試料に連結している科はユニークであり、多くの試料に連結している科は、前記試料が見つけれられた他の類似する環境間でおそらく普遍種である科であることを示している。

20

【 0 3 8 5 】

図10Aと10Bを参照すると、実施例7のPhyloChip上のペアディファレンススコア応答は、前記PhyloChipのほうがPCR法よりも16S単位複製配列に敏感であり感度がよいことを示している。図10Aでは、ペアードディファレンススコア応答は16S PCR産物に敏感である。すべてのプローブペアの頻度が示されている。示されているように、スコアがゼロに近づくに従って、プローブはそれだけ陽性であると判定される。PCR増幅ができない試料は、我々のPhyloChip検出結果とよく相関しており、応答性のプローブペアがほとんどないことを示していた。逆に、前記PCR試料が陽性であれば、それだけ多数のプローブペアが陽性的に応答した。図10Bでは、4種の門がPhyloChipにより検出された。プロテオバクテリア門、ファーミキューテス門、バクテロイデス門およびアクチノバクテリア門は、PCR産物が検出されなかった場合でも検出された。

30

【 0 3 8 6 】

(実施例12)

Rifle IFRCにおける微生物群集動態： フィールドにおける酢酸添加の影響

40

Rifle, CO Integrated Field Research Challenge地点の微生物群集特徴付けはほぼ10年前に始まった。初期の方法論には、クローンライブラリーアプローチを使用する地下水および堆積物の分析を含み、ジオバクター(Geobacter)様配列の濃縮を示していた。Rifleにおける最近の研究の試みは、3つのその後のフィールド規模酢酸アmendメント実験(Winchester [2007]、Big Rusty [2008]、およびBuckskin [2009])に、ならびに自然な生物現象領域の特徴付け(La Quinta (2009))に集中してきた。これらのフィールドアmendメントすべてが以前の実験から得られた結果を再現しており、生物刺激中地下水ではウランが還元されていた。しかし、PLFA、qPCR、TRFLPおよびマイクロアレイ分析(Akonni and Affymetrix-based LBNL G3 PhyloChip)を含む追加の分子的分的アプローチが用いられて細菌群集を特徴付けるようになった。定量的PCRは、フィールドアmendメント中のジオバクタ

50

一種の著しい変化を示した。TRFLPプロファイリングも、ジオバクター様配列が酢酸アmendメントの初期段階では地下水の細菌群集のほぼ50%に相当していたが、時間の経過とともにアシネトバクテリア(*Acinetobacteria*)およびデスルフォバクター(*Desulfobacter*)と遠い関係にある細菌に置き換わっていることを示した。Akonniマイクロアレイは、Winchester (2007)のデクロロモナス(*Dechloromonas*)およびデクロロソーマ(*Dechlorosoma*)に加えて、ジオバクター、ペロバクター(*Pelobacter*)およびゲオトリックス(*Geothrix*)のシグナルを検出した。さらに、2007年のプロファイルは2008年とは異なっていて、これはPLFAおよびqPCRデータにより支持されており、残りのバイオマス/刺激された群集がBig Rusty実験に入っていることを示している。G3 PhyloChipは、酢酸刺激地下水試料がバックグラウンド堆積物試料とどのように異なっているかを、大量のジオバクター種とそれよりも程度は劣るがデスルフォバクテリア科により記載した。両アレイは、優勢な鉄還元条件が優勢な硫黄還元条件に移行するに従ってアmendメント中のジオバクター種が減少することを示していた。

【0387】

G3 PhyloChipにより探索された後の試料は、デスルフォバクター科デスルフォビブリオ目、デスルフィトバクテリウム(*Desulfitobacterium*)、およびデスルホトマクルム(*Desulfotomaculum*)を含む大量の硫黄還元分類群細菌を含有していた。Rifle IFRCにおける活性な細菌を突き止めるために、Winchester実験中に地下水と堆積物において安定なアイソトーププロービング法が用いられた。具体的には、 ^{13}C 酢酸を使用して、24日時間枠にわたり堆積物の3つの粒群(粗砂、細粒土[8~約150ミクロン]、地下水[0.2~8ミクロン])について活性な微生物を評価した。結果は、プランクトンおよび粒子関連相においては活性な細菌間には違いがあることを示していて、ジオバクター様グループ(187、210、212bp)は地下水相において活性であり、アルファプロテオバクテリウム(166bp)は細粒土/砂上で増殖し、アシネトバクテリア(*Acinetobacter*)種(277bp)は地下水相でも粒子関連相でも多くの ^{13}C 酢酸を利用していた。自然に還元される堆積物(La Quinta)における微生物群集の分析により、ゲオバクター科が自然バックグラウンド群集の20%を構成し、Rifle IFRC地点から収集されたより酸化された堆積物の4倍であることが示された。La Quinta堆積物が酢酸と一緒にインキュベートされると、ゲオバクター科は決して優勢にはならず、La Quintaで見出されたゲオバクター科は、この科に属する他の生物とは異なる機能を果たしている可能性があることを示唆していた。

【0388】

(実施例13)

鉄、硫黄およびウラン還元中の生物刺激を受けた堆積物と地下水群集における複雑性および不均一性

生物刺激を受けた鉄および硫黄還元細菌(SRB)群集の系統発生的マイクロアレイ調査により、堆積物と地下水画分間の思いがけない類似性、重要な官能基の可変性、および潜在的に重要な低含量の生物についての洞察が明らかになった。Rifle、C0におけるU(VI)バイオレメディエーション実験に関連する一定範囲の酢酸改良無刺激試料由来の細菌群集は、極めて低含量の何万もの生物からでもDNAを検出することができる新たに開発されたLBNL PhyloChipを使用して比較された。これとは対照的に、もっと従来の技法(たとえば、クローンライブラリー)は、低含量群集メンバーを過小に評価する傾向がある。

【0389】

Rifle地下水への酢酸の添加は常在性の微生物群集を刺激してFe(III)および硫黄を連続的に、U(VI)を付随的に還元させた。おそらく、初期段階生物刺激中は豊富なジオバクター種がFe(III)およびU(VI)還元の原因であり、硫黄は主としてデスルフォバクター科により還元された。データによれば、非酢酸酸化SBRグループ、すなわちペプトコッカス科(*Petrococcaceae*)および以前は検出されなかったデスルフォビブリオ属(*Desulfovibrio*) (Anderson RTら、(2003) Stimulating the in situ activity of *Geobacter* species to remove uranium from the groundwater of a uranium-contaminated aquifer. AEM 69: 5884~5891頁)のわずかな濃縮が、硫黄還元中の酵素的U(VI)還元を目指す残りのジオバクター

種に対する潜在的競争相手を含んでいたことも示唆される。ジオバクターが、カラム実験(c)においては珪砂よりもFe(III)担持Rifle堆積物の方に付着するほうを強く好むことを示していること以外は、特定の試料処理(酢酸改良:[a, b]表面下地下水/堆積物、および[c]実験室または[d]井戸内フィールドカラム堆積物/石英; [e]自然に還元される表面下堆積物)内では、群集は高度に類似していた。奇妙なことに、硫黄還元堆積物のサブセット(d)は、他の硫黄還元堆積物(b~e)に対するよりもFe(III)および硫黄還元地下水群集に対するほうに大きな類似性を示した。これはおそらく、一つには上昇したジオバクターとデスルフォバクター科およびデスルフォピブリオ目との広い重複のせいであり、選択的堆積物に限定されているペプトコッカス科の差次的増加のせいである(c, e)。

【0390】

10

(実施例14)

表面下における天然微生物ホスファターゼ活性によるウランバイオミネラリゼーション

本実施例の目的は、不溶性ウランリン酸塩鉱物の作製による固定化ウランの促進を目的に、天然に存在する表面下微生物における微生物ホスホヒドロラーゼの役割を調べることである。我々の先のNABIR-ERSP(SBR)プロジェクトの結果は、DOE Oak Ridge Field Research Center (ORFRC)における放射性核種および金属汚染された土壌から単離された表面下微生物が、酸耐性であり鉛を含む数多くの有害重金属に耐性であることを示している。さらに、これらの鉛耐性分離菌の多くは、有機リン酸化合物上での増殖中に無機リン酸を遊離し、これに付随して金属リン酸塩沈殿物を生成することにより金属毒性を寛解することができるホスファターゼ表現型(すなわち、特にリン酸非抑制性であると推測される表現型)を示す。グリセロール3-リン酸由来の遊離リン酸は、pH範囲5~7でターミナル電子受容体として溶存酸素または硝酸塩のどちらかを含有する合成地下水中の低可溶性ウランリン酸塩鉱物としてU(VI)の95%も沈殿させるのに十分であった。本実施例では、表面下微生物群集における天然に存在する微生物ホスファターゼの活性が、汚染土壌中のリン酸塩鉱物の形成を介してウランを固定化するのかどうかを判定するための実験アプローチを我々は開発した。

20

【0391】

U(VI)および NO_3^- 汚染ORFRCエリア2およびエリア3土壌の表面下微生物群集応答の他にも、有酸素および無酸素増殖条件下での外因性有機リン塩添加、土壌スラリー、ならびにpH 5.5および7.0で行われたフロースルーリアクター実験に対する微生物集団応答についての特徴付けを実施中である。土壌スラリーおよびフロースルーリアクター実験は36日間および80日間25℃で、それぞれ唯一のC、PおよびN供給源として10mM G_2P および15mM NO_3^- を用いて行われた。有酸素増殖条件下では、スラリーインキュベーションの最後には4mMを超える可溶性 PO_4^{3-} が測定され、 NO_2^- は検出されなかった。無酸素土壌スラリーインキュベーションについて得られた予備データによれば、1mMを超える PO_4^{3-} の蓄積の他にも NO_2^- の蓄積とその後の除去が示された。3通りのインキュベーションに続いて、16S rDNA多様性のスラリーが、高密度16Sオリゴヌクレオチドマイクロアレイ(PhyloChip)を介して分析された。予備の結果によれば、有酸素条件下では、微生物群集構造は、改良されていない土壌の多様性と比べるとプロテオバクテリア分類群において低pHで濃縮されていることが示唆される。無酸素条件下でインキュベートされたスラリーの分析は、有酸素と無酸素両環境下で有機リン酸塩加水分解ができる細菌分類群を同定するために進行中である。3.7の開始pHを有する土壌のフロースルーリアクター研究では、5.5の間隙水pHに到達すると強い微生物活性が示された。pH調整の2日以内に脱窒素作用と有機リン酸塩加水分解が測定された。我々の土壌スラリーとカラム研究により、ORFRC汚染表面下土壌に住む微生物群集によるU(VI)の有機リン酸塩媒介隔離の潜在的有効性が実証されている。

30

40

【0392】

(実施例15)

表面下金属汚染物の加速されたレメディエーションに応答する微生物群集軌跡

DOE地点での表面下金属汚染物のレメディエーションは、微生物群集の生態学により大部分制御されている酸化/還元または錯体化の微生物機構を必要とする。群集構造、機能

50

および重要な環境要因間の関係を認識し定量化することにより、レメディエーション戦略に関する将来的決定を告げることが可能な定量的理解が得られる可能性がある。U生物還元および低水中U濃度の維持は、有機炭素(OC)供給速度に強く依存していることを我々はすでに発見している。我々の結果により、OC供給速度は微生物群集構造に著しい効果を有しており、2つの異なるOCタイプの効果が実験期間中は二次的であることが明らかになった。OC供給速度が異なればカラム内の環境条件も異なるという事実にも関わらず、異なるOC供給速度に起因する群集間の差異は時間とともに減少した。全体として、これらのデータは、バイオレメディエーションのために刺激された微生物群集は予測可能な軌跡に従う可能性があることを示している。

【0393】

我々の事前の研究、および微生物群集は制御し予測することが可能であるという前提の下で操作することの他にもこうして得られたレメディエーション能力に基づいて、我々の現在の計画の目的は、(1)OCアmendメントに続く微生物群集の構造、組成および機能の軌跡を重要な環境決定要因と関係付けることが可能かどうか、および環境決定要因により予測することが可能かどうかを判定すること、ならびに(2)微生物群集の機能的応答およびバイオレメディエーション能力の主要決定因子としての常在性微生物群集、堆積物、地下水およびOC供給速度という特徴の相対的重要性を評価することである。我々は、3つの堆積物(Oak Ridge, TN; Rifle, CO; Hanford, WA)およびその微生物群集を相互移植実験計画法を使用して分析している。前記3つの堆積物の最初の特徴付けによって、これらの堆積物は、鉱物特性、粒子サイズ分布、容積密度、塩基陽イオン、CEC、SAR、鉄、マンガ
ン、リン、および硫黄濃度、有機および無機炭素濃度、間隙水化学、ならびに微生物群集
サイズおよび組成が異なることが明らかにされている。フロースルーリアクターは、2つ
のOC供給速度で刺激を受けた地下水を受けるが、18ヶ月間の期間にわたり破壊的にサンプ
リングされている。現存のアレイ上に含まれていない機能的能力を探求するために、群集
DNA(全体構造)およびRNA(活性なメンバー)の16S PhyloChip分析、群集DNA(機能的な潜在
力)および群集RNA(活性な機能)のGeoChip機能的分析ならびにメタトランスクリプトーム
分析を使用して微生物群集軌跡は追跡されている。リアクター流出水および堆積物の地球
科学的特徴を使用して、微生物群集構造および機能的軌跡に影響を与える要因のモデル
を作製している。これらの分析により、DOE地点における表面下金属レメディエーション
の基底をなす微生物群集生態学の枠組みが提供されることになる。

【0394】

(実施例16)

序列における定量的分析支援

表面下堆積物は、Oak Ridge, TN; Hanford, WA; Rifle, COでの金属汚染DOE地点から収集された。1~3g堆積物を使用する複数(n=13~15)のgDNA抽出は各地点から実施された。抽出物は定量され、次に10ngのgDNAは8温度勾配16S PCRにより増幅された。温度プールから、500ngがG3 PhyloChipにハイブリダイズされた。OTUごとのハイブリダイゼーション強度は、プローブペアのOTUセットごとにPM-MM差のトリム平均として決定された。NMDS序列は、相対的量ではブレイカーチス距離を、存在/非存在データではSorensenを使用してRにおいて作成された。

【0395】

図17は、類似の生物群集形態序列クラスターからのPhyloChip結果を示す図である。OTUは、3つの異なる場所由来の表面下堆積物から採取された試料に存在するまたは存在しないと呼ばれた。試料間の距離行列は、Sorensen距離に基づいて作成された。距離行列はNMDSを使用して並べられ、試料位置により着色された。アノシム(Anosim)分析により、グループ内の試料のほうが異なるグループ由来の試料よりも組成が類似していることが明らかにされている。

【0396】

図18は、類似の生物群集形態序列クラスターからのPhyloChip結果を示す図である。OTUは、3つの異なる場所由来の表面下堆積物から採取された試料から定量化された。試料間

の距離行列は、ブレイカーチス距離に基づいて作成された。距離行列はNMDSを使用して並べられ、試料位置により着色された。アノシム分析により、グループ内の試料のほうが異なるグループ由来の試料よりも組成が類似していることが明らかにされている。R値は前のプロットと比べて大きく、類似の試料類型間の関係が定量的PhyloChipデータを利用した場合のほうが近いことを示している。

【0397】

(実施例17)

汚泥バイオリアクターにおける定量的分析

活性汚泥バイオリアクターは下水から有機物および栄養素を除去するのに広く使用されている。しかし、活性汚泥微生物群集を構造化する際のイミグレーションの役割はほとんど理解されていない。本格的な下水処理施設での1年を通じた一連の毎週試料からの証拠の集束線は、曝気槽流入水 NO_2^- と活性汚泥微生物群集構造の変化間の強い関連を示していた。この関連をさらに調べるため、この施設内のトランセクトに沿った4つの位置、すなわち、1)施設流入水、2)散水濾床バイオフィーム、3)散水濾床流出水、および4)活性汚泥バイオリアクターからサンプリングした。ここで、多相アプローチを介して、流入 NO_2^- は、上流バイオフィームベースの散水濾床から活性汚泥バイオリアクターへの微生物イミグレーションのシグネチャーであることを我々は明らかにしている。高密度系統発生マイクロアレイ(PhyloChip)分析により、散水濾床における過剰なメタン生成菌および硫黄還元細菌が明らかにされ、散水濾床流出水を介した活性汚泥への微生物輸送が示唆された。さらに、アンモニア酸化細菌(AOB)amoAコピー数が施設流入水と散水濾床流出水間で一桁増加し、散水濾床におけるAOBの蓄積および活性汚泥装置への著しいイミグレーションが示唆された。クローンライブラリーにより確認された分子フィンガープリンティング(T-RFLP)分析により、ニトロソモナスユーロピア(*Nitrosomonas europaea*)が散水濾床を支配し、「ニトロソモナス様」系統が活性汚泥では優位であることが明らかにされた。*N. ユーロピア*が上昇した流入水 NO_2^- 事象中活性汚泥において優位であることが以前明らかにされており、活性汚泥AOB群集動態は、一部、上流散水濾床からの糊崩れ(sloughing)を介したイミグレーションにより推進されることを示唆している。

【0398】

図19と20は、PhyloChip G3アレイを使用して実施された分析を説明している。図19は、4つのサンプリング地点がまったく異なっており、生物学的複製が極めて高度なレベルの類似性を示すことを実証するNMS分析を示している。図20は、縮んだセントロイドの方法と呼ばれる分析のヒートプロット要約である。この分析の基本概念は、サンプリング位置間の全体的群集構造の観察された差異をもっとも顕著に明確にする約50くらいの微生物OTUを同定することである。我々が仮説を立てたように、嫌気性細菌(特にメタン生成菌)は、50微生物類型のこのセットによく表されており、我々はこれらの微生物のサンプリング位置(すなわち、散水濾床および活性汚泥曝気槽)間の輸送の証拠とみている。さらに、ニトロスピラ属(亜硝酸酸化菌)も、この「最小の」データセットにかなり良く表されている。とりわけ、PhyloChip結果が基本的にニトロスピラ属の非存在、およびニトロスピラ属がはるかに豊富である下流活性汚泥装置における亜硝酸蓄積が基本的にないことを示しているサンプリング位置、すなわち散水濾床バイオフィームの1つにおけるわずかなレベルの亜硝酸蓄積を我々は見ている。

【0399】

まとめると、我々の結果は、連結した処理装置間のイミグレーションが活性汚泥微生物群集構造に著しい影響を与えることが可能である有力な証拠を提供する。

【0400】

(実施例18)

アメリカスギの森への気候変化の影響に関するPhyloChip G3分析

この計画では、沿岸のアメリカスギの森における土壌微生物群集の組成に対する気候変化の潜在的影響を調べた。気候変化に対する微生物群集の応答を理解することは、生態系サービスの変化を予測するのに重要であり、生態系執事にとっては興味深い。

【0401】

三方向相互移植実験は、沿岸のアメリカスギの森の緯度勾配にわたり行われた。試料は移植の1年後および3年後に収集された。細菌群集組成は、高密度16S rDNAマイクロアレイ(PhyloChip)を使用して分析された。気候変数および土壌変数(降雨量、土壌湿度、土壌温度、土壌CおよびN利用可能性、pH、土性)が測定された。群集組成の変化は、非計量的多次元尺度法(全群集では)およびANOVA(個々の分類群では)を用いて評価された。細菌群集組成と気候および土壌変数間の関係はマンテル(Mantel)検定を用いて調べられた。

【0402】

気候の変化は、細菌群集組成への強力な影響をもたらす媒介物を有していた。降雨量および土壌湿度へのその影響は群集組成と最も強力でもっとも重要な相関物であった。さらに、気候の変化に応答する細菌種の数1年目から3年目へと増加した。

10

【0403】

その結果によれば、気候変化は局地的規模での細菌群集組成への強力な影響をもたらす媒介物を有していた。降雨量は細菌群集組成と最も重要な相関関係を有していた。種の相互作用などの他の要因または他の確率プロセスも時間をかけて群集組成の変化に大きな影響を与えることがあるが、気候変化の影響に応答する種数は時間とともに増加し、3年では微生物群集組成への気候変化の長期的影響を評価できるほど長くはないと思われる。

【0404】

表13は、試料が移植された1年後と3年後両方の、移植された試料および対照の細菌群集組成と環境変数間の関係を表す重要な標準マンテル統計(r)を示している。

20

【0405】

【表12】

表 13

環境変数	軸1 マンテル検定 r	マンテル検定 p 値
移植の1年後		
年間降雨量	0.19	0.013
晩春の雨	0.19	0.019
全環境変数	0.19	0.015
移植の3年後		
年間降雨量	0.17	0.009
夏の雨	0.19	0.007
重量測定含水量	0.18	0.040
温度	0.13	0.047
最高温度	0.12	0.034
年間降雨量+温度	0.17	0.011

30

【0406】

【表 13】

表 14 1年後および3年後の移植に応答する細菌(OTU)

門(動物)/門(植物)	綱	1年後(no.*)	3年後(no.*)
アシドバクテリウム門		1	17
アクチノバクテリア門		0	38
バクテロイデス門		1	8
クロロフレクサス門		0	14
ファーミキューテス門		16	39
プランクトミケス門		0	9
プロテオバクテリア門	アルファ-	11	104
	ベータ-	21	15
	デルタ-	0	11
	ガンマ-	20	13
スピロヘータ門		1	18
他		3(2 門から)	39(18 門にわたり)
合計		74	325

【0407】

P<0.10でのANOVAによる処理(起源インキュベーション組合せ)間で相対存在量(OTU強度)に差があるOTUの数。

【0408】

図21は、気候の変化に応答する群集組成の変化の程度が異なることを示す図である。白四角は、序列における南緯度地点の位置を表し、黒四角は北緯度地点の位置を表す。白三角は、南緯度気候を経験した北緯度地点の群集を表す。矢印の長さは変化の程度を示す。

【0409】

図22は、a)2005年8月に北緯度、中間緯度および南緯度地点から収集された新鮮な試料、ならびにb)同時間(移植の1年後)に同地点由来の新鮮な試料および移植対照試料のNMS序列を示す2つのグラフである。両グラフに描かれている新鮮な試料は同じ試料である。バーは3複製物の1標準偏差を表す。

【0410】

図23は、移植1年後に収集された相互に移植された試料および移植された対照のNMS序列を示す4つのグラフである。矢印は、元の地点の対照の組成から離れた移植された試料の組成変化の軌跡を示す。

【0411】

図24は、a)2007年9月に北緯度、中間緯度および南緯度地点から収集された新鮮な試料、ならびにb)同時間(移植の3年後)同地点由来の新鮮な試料および移植対照試料のNMS序列を示す2つのグラフを示している。両グラフに描かれている新鮮な試料は同じ試料である。バーは3複製物の1標準偏差を表す。

【0412】

図25は、移植の3年後に収集された相互移植試料および移植された対照のNMS序列を示す4つのグラフである。矢印は、元の地点の対照の組成から離れた移植された試料の組成変化の軌跡を示す。

【0413】

(実施例19)

沿岸部カリフォルニアにおける哺乳動物および鳥類起源の糞便汚染の微生物群集分析

沿岸地域に生息する野生動物および家畜は、水質に影響を与える糞便微生物を堆積する

。沿岸水域が様々なヒトおよび動物起源の糞便汚染により損なわれる程度を、単一のバイオマーカーおよび低解像度プロファイリング法を用いて判定するのは困難である。腸微生物群集のハイスループット配列分析は、糞便源を確実に同定し、異論の多い水質問題を解決する潜在能力を有している。本研究では、我々は汚染源を区別する分類群を同定するために、多種多様な動物糞便およびヒト廃棄物から細菌群集を特徴付けた。次に我々は、水汚染事故中にこれらの発見の有用性を試験した。

【0414】

新鮮な糞便試料は、カモメ、ガチョウ、ひれ足動物(アシカおよびトド)、ウシ、ウマおよびヘラジカそれぞれの少なくとも4つの地理的に異なる集団から収集された。ヒト下水および腐敗廃棄物は複数の場所から集められた。我々は、各試料中の全標的プールの16S rRNA遺伝子コピーを検出することにより希少な細菌分類群と豊富な細菌分類群両方の相対存在量の違いを定量化することができるPhyloChipマイクロアレイを使用して細菌16S rRNA遺伝子組成を分析した。

10

【0415】

周囲の水試料は、北部カリフォルニアの9ヶ所のレクリエーション海岸で毎週2年にわたり、およびサンフランシスコ湾の大きな下水流出中に収集された。水試料は、一般的な糞便インジケータ試験により測定され、発生源同定のためにPhyloChipを使用して分析された。

【0416】

糞便細菌群集は動物種/類型により強くクラスター化した。我々は、ヒト廃棄物を動物糞便から、および異なる動物を互いに区別する何千もの細菌分類群を同定した。ヒト廃棄物試料は、処理の規模および類型に違いがあるにもかかわらず一つにクラスター化した。ウシおよびヘラジカの細菌群集はほとんど区別がつかず、これらの反芻動物の異なる集団間にはほとんど変動がなかった。これとは対照的に、鳥類の細菌群集は集団間で、同一種内でもはるかに可変性であった。ウマ集団は他の草食動物と一緒にクラスター化した。反芻動物とは組成が異なっていた。汚染事故中の水試料の分析により、我々の発生源特徴付けから展開させた特有の分類群のライブラリーが汚染の原因を首尾よく同定するまたは排除することが可能であることが実証された。

20

【0417】

糞便試料および清潔な水試料における検出された細菌分類群のクラスター分析が実施され、PhyloChip G3アレイが糞便試料中の3513の異なる細菌亜科(ステージ1分析を通過した)を検出したことを明らかにした。動物の種および類型(反芻動物および草食動物、ひれ足類、鳥類)による強いクラスタリングが示され図26に表示された。PhyloChip G3アレイを使用すると、ヒト発生源(浄化槽、下水)は動物および野生生物およびバックグラウンド水域とは異なっている。発生源識別子群集は発生源ごとに定義された。検出されたOTU(ステージ1通過)は、バックグラウンド水域よりも有意に高いアレイ強度を有していた(t-検定および平均アレイ強度差、 >2000)(図27)。図28では、インジケータ群集は発生源識別のために汚染された水試料と比較された。

30

【0418】

FIBに強い相関関係のある下水分類群は図29に示されている。下水中で見つかった4,625の異なる分類群の量は糞便インジケータと強く相関していた($r>0.9$)。もっとも相関している分類群はバクテロイデス目とクロストリジウム綱であった。

40

【0419】

トマレス(Tomales)湾拡散チャンバー実験において同定された潜在的インジケータ分類群の系統樹は示されていない。潜在的インジケータ分類群とは、特定の廃棄物にユニークであり受水域には存在しないOTUのことである。酪農場の廃棄物では165の潜在的インジケータ分類群が識別され、浄化槽廃棄物では119のインジケータ分類群が識別された。合計で13,341の異なる分類群が、G3チップを用いて廃棄物および受水域試料中で検出された。

【0420】

図30は、群集組成の比較を示したクラスター分析の結果を示す。群集は受水域における

50

時間、発生源および受水域の類型に従ってクラスター化することが可能である。

【0421】

図31は、糞便微生物群集に対する受水域における時間の効果を示す棒グラフである。4日間の液浸により、ウシおよび腐敗群集における類似の変化を有する分類学的群間の残留性の違いを示している。大半のプロテオバクテリア門は時間とともに相対存在量が減少する。クロストリジウム綱は時間とともに相対存在量が増加する。

【0422】

図32は、廃棄物微生物群集に対するクリーク対湾水の効果を示す棒グラフである。異なる水類型に対するウシおよび腐敗群集の類似する応答が例示されている。クロストリジウム綱、プロテオバクテリア、大腸菌群はクリークにおいて有利であり、プロテオバク

10

【0423】

したがって、食餌、場所または処理に起因する内発生源の変異性にも関わらず、動物が異なれば発生源の追跡のために利用することが可能なはっきりと異なる糞便微生物群集を宿す。

【0424】

(実施例20)

海洋マイクロバイームに対する油流出効果および徐染の評価

本発明の方法、組成物およびシステムを適用すれば、環境を支えその環境により支えられるマイクロバイームに対する環境の変化の効果を評価することが可能になる。本実施例では、本発明のアレイを使用して、健康な海洋環境のマイクロバイームのための基本線を確立し、次にこの基本線を使用してマイクロバイームに対する油流出の影響を評価する他にも修復努力の進行を評価する。

20

【0425】

微生物DNAは、油流出により影響を受ける多様な生態系を表す約150試料からの他にも類似の影響を受けていない生態系由来の約100試料から単離される。試料は、代表的な範囲の深海、商業上およびレクリエーション漁場、ならびに沿岸地域、たとえば、浜辺と湿地の表層水、入江および潟から収集される。理想的には、最初地点あたりで、およびそれぞれの年4回の再サンプリングで複数の試料(5~10)が収集される。実施例2のように、DNAは前記試料から抽出され、増幅され、プロセッシングされ、分析される。プローブハイブリダイゼーションによる分析は、実施例2および7において記載されるようなアレイを使用して行われる。評価されるプローブごとに/または評価されるプローブにより表わされるOTUごとに、存在、非存在および/もしくはレベルが点数化される。結果は、影響を受けていない海洋環境のバイオシグネチャーおよび油流出により影響を受ける海洋環境のバイオシグネチャーである。前記結果の解析およびバイオインフォマティクスデータマイニングは、各地点での微生物集団の状態に関する報告の他にも、被害を受けていない類似の海洋生態系と比べた微生物生態系サービスへの損害の範囲を示す解釈報告をもたらす。

30

【0426】

その後、試料は四半期ごとに各モニタリング地点から収集され、油流出に影響を受けた地域の最初のバイオシグネチャーからの変化の他にも影響を受けていない類似の生態系サンプリング地点と比べた連続する生態系被害が評価される。修復作業の相対的成功は、流出に影響を受けたバイオシグネチャーと影響を受けなかったバイオシグネチャー間の類似性の改善の程度の点から測定されるが、将来の油流出災害の封じ込めまたは分散のためのもっとも適切な行動を知らせるのに使用することができる。評価される健康な海洋微生物生態系ごとのプロファイルは、サンプリングの3四半期~5四半期間で確立され、特定の微生物種の相対存在量と多様性における正常な季節的変動を考慮に入れる。信頼度および確率情報を含む、改善された地点由来の微生物バイオシグネチャーに影響を受けなかった地点と比較することにより、地点特異的修復が追跡される。これらのパラメータが確立されると、油流出被害からのレメディエーションおよび健康で機能的海洋生態系の修復へ向か

40

50

う前進が計画され認定される。修復の程度には、影響を受けなかった海洋環境と影響を受けた海洋環境のバイオシグネチャー間の類似性の割合を表す修復スコアが割り当てられる。影響を受けなかった領域の微生物集団に対する影響を受け処理された領域の高類似性は、流出領域が回復しており健康な海洋生物を支えることができるという証拠を与える。影響を受けなかった海洋環境と影響を受けた海洋環境のバイオシグネチャー間の類似性の増加を追跡することの他にも、類似性の閾値に到達するとレメディエーション作業が停止されるレメディエーション作業の終点を定義することは、影響を受けない状態への回復に時間の投影を与える。閾値は約80%、85%、90%、95%、97.5%、98%、98.5%、99%、99.1%、99.2%、99.3%、99.4%、99.5%、99.6%、99.7%、99.8%、99.9%、99.95%よりも高い、またはそれよりも高い類似性でありうる。

10

【0427】

(実施例21)

細菌群集への深海石油ブルームの効果

メキシコ湾におけるディープウォーターホライズン流出由来の石油はこの生態系への莫大な炭素入力を表し、石油中の炭化水素成分は海水中に存在する微生物のための炭素基質として潜在的に働くことになるであろう。微生物群集および炭化水素分解に対するその潜在力へのブルームの影響が評価された。本研究は、2010年5月25日から6月2日まで2隻の船のクルージングで19のサンプリング地点にわたる。

【0428】

試料収集

20

有色溶解有機物(CDOM) WETstar蛍光光度計(WET Labs、Philomath、OR)をCTDサンプリングロゼット(Sea-Bird Electronics Inc.、Bellevue、WA)に取り付けて、これを使用して水面と海底間の深度プロファイルに沿って油の存在を検出した。蛍光光度計の結果は、その後実験室炭化水素分析を用いて確認された。合計で17の試料が10か所から分析された。

【0429】

CTDロゼットに取り付けられたNiskinピンを使用して、内海域と外海域の様々な深度で、検出された炭化水素と一緒に水試料を捕獲した。各試料から800~2000mLの海水を、0.22μm孔径を有する47mm径ポリエチルサルフォン膜を含有する無菌濾過装置(MO BIO Laboratories、Inc.、Carlsbad、CA)を通して濾過し、次に直ちに凍結して-20℃で保存した。フィルターはドライアイス上に載せ、DNAとリン脂質脂肪酸(PLFA)抽出まで-80℃で保存した。

30

【0430】

100mLの海水をシリンジ(syringe)濾過し、事前に空気を抜いた25mLの血清ピンに注入し、厚ブチルゴム栓でふたをした。100mLの海水は、栄養素分析のために125mL HDPEピン中で凍結させた。AODCのために、36mLの海水を4%ホルムアルデヒド(最終濃度)中で保存した。

【0431】

DNA抽出

各フィルターの4分の1を小片に切断し、Lysing Marix Eチューブ(MP Biomedicals、Solon、OH)に入れた。300μLのミラーリン酸緩衝液と300μLのミラーSDS溶解緩衝液を添加し混ぜ合わせた。次に600μLフェノール:クロロホルム:イソアミルアルコール(25:24:1)を添加し、チューブはFastPrep装置において5.5m/sで45秒間ビーズビートした。前記チューブは4℃で5分間、16,000×gで回転させた。540μLの上清を2mLのチューブに移し、等量のクロロホルムを添加した。チューブは混合させ、次に10,000×gで5分間回転させた。400μLの水相を別のチューブに移し、2容積のSolution S3(MoBio、Carlsbad、CA)を添加して反転させて混ぜ合わせた。残りの精製手順はMoBio Soil DNA抽出キットの使用説明書に従った。試料は60μL Solution S5に戻し-20℃で保存した。

40

【0432】

PCR増幅

前記16S rRNA遺伝子は、細菌ではプライマー27Fと1492Rを、古細菌では4Faと1492Rを用

50

いてPCRを使用して増幅させた。各PCR反応は、1×Ex Taq緩衝液(Takara Bio Inc., Japan)、0.025ユニット/μl Ex Taqポリメラーゼ、0.8mM dNTP混合液、1.0μg/μl BSA、および200pM各プライマーおよび鋳型として0.15~0.5ngゲノムDNAを含有していた。PhyloChipアッセイ(PhyloTech Inc., San Francisco, CA)分析では、各試料は、一定範囲のアニーリング温度にまたがる4複製25μl反応で増幅された。PCR条件は、95 (3分)に続いて30サイクルの95 (30秒)、46~56 (25秒)、72 (2分)に続いて最終伸長72 (10分)であった。各反応からの単位複製配列は試料ごとにプールし、QIAquick PCR精製キット(Qiagen, Valencia, CA)を用いて精製し、20μL溶出緩衝液に溶出させた。

【0433】

Phylochipアッセイデザイン

PhyloChipマイクロアレイプローブデザインを、少なくとも1,300ヌクレオチドを含有する全既知高品質16S rRNA遺伝子配列に適用した。配列(大腸菌塩基対47位から1473位)は、16S rRNA遺伝子データベース(greengenes.lbl.gov)から入手可能なNASt多重配列アライメントから抽出された。この領域は、2~4プライマーのみを使用して細菌または古細菌ゲノム材料を増幅させるためのPCRプライミング部位として使用することが可能な普遍的に保存されたセグメントに隣接しているという理由で選択された。1つのまたは両方の推定親に $\geq 90\%$ レーンマスク同一性を有するベレロボン(Bellerophon)発散比 ≥ 1.1 に遭遇する推定キメラ配列が同定され取り除かれた。3もしくはそれ以上のホモオクタマーまたはさらに長い配列、あるいは $\geq 0.3\%$ 多義ベースコールを有する配列も除外された。サブアライメントから、G+C含量が35~75%、RNAfoldにより計算される二次構造自由エネルギー(G) $\geq -4\text{kcal/mol}$ (17)、相補的融解温度が61 ~ 80、およびThermalignにより計算される自己二量体融解温度 < 35 の推定25mer標的が選択された。

【0434】

濾過されたrRNA遺伝子配列は、クラスターの各配列を表す完全相補性プローブの選択を可能にするようにクラスター化された。クラスターに配列同一性を有する17merを含有する推定単位複製配列は、そのクラスターに含められた。こうして得られた59,959クラスターは、それぞれが平均で0.5%の配列分散を包含していたが、操作的分類単位(OTU)と見なされた。前記OTUは、古細菌と細菌内で境界を画定された2界、147門、1,123綱、および1,219目を表した。各OTUは、Philip Hugenholtzにより主張された分類学的概要におけるそのメンバー生物の配置に従って1,464科のうちの1つに割り当てられた(Hugenholtz 2002, Genome Biol. 3(2): 1~8頁)。各科を含むOTUは、72%共通の7量体の遷移的(単純連結)配列同一性により亜科にクラスター化された。全体で、10,993亜科が見つかった。各OTUの分類学的位置の他にも各OTUを含む配列の付随的NCBI受託番号は、files_sequences_by_OTU_G3.gz, taxonomy_by_OTU_G3.gzにおいて入手可能である。

【0435】

OTUごとに、所定のOTUのしかし前記所定のOTUの外側の配列とは似ていないメンバーにおける普及について複数の特異的25mer標的が探求された。特定のOTUためのプローブ選択の第1のステップでは、前記OTUにおける配列それぞれが、重複する25mer、すなわち潜在的標的に分離された。次に、各潜在的標的は、できる限り多くの前記OTUの配列にマッチされた。Greengenesにより与えられる多重配列アライメントを使用して、各潜在的プローブ部位でグループサイズの個別の測定値が与えられた。たとえば、7配列を含有するOTUが、1つのメンバーが欠損データであるプローブ部位を有するならば、部位特異的OTUサイズは6にすぎなかった。考えうる標的をランク付けする際に、そのOTUの全メンバーのデータを有する標的は、OTUメンバーの一部分のみで見出される標的よりも好まれた。第2のステップでは、普及している標的のサブセットが選択され、プローブ配向は、意図されない単位複製配列へのハイブリダイゼーションを最小限に抑えるために逆相補体に反転させた。潜在的に問題を含むと推定されるプローブは、1を超えるOTUにおける配列にマッチする中央の17merを含有する25merであった。したがって、4つの隣接する塩基のうちの1つにおける特有の塩基のせいのみでOTUにユニークであるプローブは避けられた。その上、根底近くに共通のツリーノードを有するプローブは、末端枝近くに共通ノードを有するプロ

ープよりも好まれた。製作のために選択された標的配列に相補的なプローブはパーフェクトマッチ(PM)プローブと呼ばれる。各PMプローブが選ばれると、13番目の塩基を除いてすべての位で同一の対照25mer(ミスマッチプローブ、MM)と対合された。前記MMプローブは、任意のOTUにおける配列に相補的な中央17merを含有しなかった。前記PMおよびMMプローブは一緒に分析されるプローブペアを構成する。各OTUに割り当てられるプローブペアの平均数は37であった(標準偏差9.6)。

【0436】

選ばれたオリゴヌクレオチドは、Affymetrix Inc.(Santa Clara, CA)でフォトリソグラフィ法により、 μm^2 当たり10,000分子のおおよその密度で直接ガラス表面上に合成され、「ミディ100フォーマット」ハイブリダイゼーションカートリッジに置かれた。11,016, 064プローブフィチャーの全アレイは、1,008行と列の格子として配置された。品質管理、処理制御、画像の向き、標準化制御、階層的分類学的同定、病原体特異的シグネチャー検出および染色体のいくつかのインプリメント追加の標的領域のための追加のプローブ。さらに、分析の系統発生的範囲を広げることができるように、その配列が公共の貯蔵所に明白に入っていると確認される場合は、信頼度の低い16S配列に相補的なプローブが含まれた。PhyloChipアッセイ計画は、前分析、処理、前標識されたハイブリダイゼーション対照および陰性対照のための対照プローブを含む。前分析対照およびハイブリダイゼーション対照は、バックグラウンドシグナル強度の解釈において、および試料間の比較のための全体的蛍光強度の正規化を支持するためにも使用することが可能である。

【0437】

PhyloChipアッセイのための試料調製

ディープホライズン核酸から、500ngの細菌PCR産物と25ngの古細菌PCR産物をハイブリダイゼーションのために調製した。PCR産物はアガロースゲルにより検証される50~200bpの範囲まで断片化された。DNA調製のためには、市販のキット、Affymetrix (Santa Clara, CA) WT Double Stranded DNA Terminal Labelingが利用され、分析にはAffymetrix GeneChip Hybridization、Wash、and Stainキットが使用された。手短に言えば、断片化16S単位複製配列および非16S定量的単位複製配列基準対照は、8 μL の5 \times TDE緩衝液、40ユニットのTDE、3.32ナノモルのGeneChip標識試薬を含有する40 μL 反応液中でピオチンを用いて標識された。37 で60分のインキュベーション後、2 μL の0.5M EDTAが添加されて反応を終わらせた。標識DNAは65 μL の2 \times MESハイブリダイゼーション緩衝液、20.4 μL のDMSO、2 μL のAffymetrix対照オリゴB2、および0.4 μL 無ヌクレアーゼ水と組み合わせた。各反応混合液はアレイカートリッジのハイブリダイゼーションチャンバーに注入され、Affymetrixハイブリダイゼーションオープンにおいて48 、60RPMで16時間インキュベートされた。ハイブリダイゼーション溶液は取り除かれ、マイクロアレイは製造業者の使用説明書に従って染色され走査された。

【0438】

PhyloChipアッセイ分析

蛍光画像はGeneChip Scanner 3000 7G (Affymetrix, Santa Clara, CA)を用いて撮像された。個々のアレイフィチャーは、表面上の単一プローブ25merに応じて画像ファイル中ほぼ8 \times 8ピクセルを占めた。中央の9ピクセルは強度によりランク付けされ、その75%パーセンタイルはフィチャーの要約強度として使用された。プローブ強度はバックグラウンド減算され、定量的標準にスケールリングされ(非16Sスパイクイン)、異常値は前に記載された通りに同定された(DeSantisら、2007、Microb. Ecol. 53: 371頁)。OTUのハイブリダイゼーションスコア(HybScore)は、最大および最小を除いてパーフェクトマッチプローブの平均強度として計算された。

【0439】

前記PMおよび対応するMM強度の比較は、上に記載されるペアディファレンススコア、dとして要約される。dスコアは、プローブペアと様々なヌクレオチド組成の比較ができるように標準化される。この変換の目的は、ペアのd値が陰性対照(NC、どんな16S rRNA配列にも定量的標準にも潜在的クロスハイブリダイゼーションしないプローブペア)に由来す

るd値により類似するのか、陽性対照、定量的標準(QS、実験にスパイクされている非16S rRNA遺伝子とマッチするPMを有するプローブペア)に由来するd値により類似するのかを判定することである。d_{QS}値はその標的のA+T総数とT総数に依存しているために、QSペアはこれらの属性により網に分類され、それぞれにd_{QS}値の別の分布が見つかる。d_{NC}値は同じように分類される。観察結果から網ごとに分布が評価される。OTUプローブセット由来の各d値は、ペア応答スコア、r(上に記載されている)を作成するために同じ網由来のd_{QS}とd_{NC}の分布と比較される。OTUに相補的なプローブペアのセットのrスコアは、ステージ1プローブセット存在/非存在スコアリングにおいて集団的に検討される。最小では、18プローブペアが検討される。rスコアがランク付けされ、四分位数、 r_{Q_1} 、 r_{Q_2} および r_{Q_3} が見つかっている。OTUがステージ1を通過するためには、以下の基準、 r_{Q_1} 0.70、 r_{Q_2} 0.95、および r_{Q_3} 0.98の3つすべてを満たさなければならない。ステージ1を通過するOTUは、亜科検出のためにステージ2スコアリングにおいて検討される。このステージでは、クロスハイブリダイゼーション調整応答スコア、 r_x は上に記載される全応答性プローブ($r > 0.5$)について計算される。すべてのペナルティが検討された後、 r_x 値はランク付けされ四分位数が上記の通りに見つけられる(r_{xQ_1} 、 r_{xQ_2} 、 r_{xQ_3})。 r_{xQ_3} 値 ≥ 0.48 を有する亜科はプレゼントと見なされた。

【0440】

ブルーム内で著しく濃縮されたOTUは、スチューデントt検定を用いてlog2でp値 < 0.05 (HybScores)、9ブルーム試料のうち ≥ 4 のステージ1プレゼントコール、ならびに > 1000 ユニットおよび $> 35\%$ のバックグラウンド(ブルーム試料の外側)と比べて平均HybScoresの増加を達成したOTUと定義された。

【0441】

PhyloChipアッセイ性能

異なる種由来の26の16S rDNA混合物は、JacobsonとMathewsにより記載された半無作為化ラテン方格構造を使用してモック群集として調製された(Jacobsonら、1995, Journal of Combinatorial Designs 4: 405頁)。段階的機能は、各連続する生物が以前の生物よりも37%大きな最終濃度で添加されるように使用された。各試験生物は、各考えうる濃度段階の全混合物で表された。26のDNA混合物は、異なる日に3通りにハイブリダイズされた。その上、対照として、1ハイブリダイゼーションが定量的基準対照のみを使用して実施された。基準対照について0.5を超える応答スコア、rを生じたすべての16Sプローブペアは、それに続く分析からマスクされた。

【0442】

3,548の異なる25mer組合せを表す12,202の複製プローブからバックグラウンド減算されたプローブ強度を使用して、アッセイごとに変動係数(CV)を決定した。全体として、平均CV=0.097を生じる変動は少数であった。さらに、ラテン方格における各遺伝子の濃度と対応するHybScore生成の間には著しい相関が見られ、平均相関係数、 $r = 0.941$ であった。

【0443】

ハイブリダイゼーション混合物内の単位複製配列を検出し分類する能力は、受信者動作特性(ROC)曲線を使用して評価された。 r_{Q_1} 、 r_{Q_2} および r_{Q_3} プローブセット要約は、すべてのラテン方格結果由来の考えられるOTUのそれぞれから収集された。ROC曲線はプロットされて、存在を判定するための単一閾値を選ぶ効果を評価した。y軸、すなわち予想される陽性率は存在すると予想されプレゼントと呼ばれたOTUの割合である。x軸、すなわち予想されない陽性率は存在すると予想されずプレゼントと呼ばれたOTUの割合である。四分位数ごとの存在/非存在閾値は、もっとも厳密でない0からもっとも厳密な1まで変化した。たとえば、 r_{Q_1} プロットでは、0.5の閾値は予想される検出事象の97.5%を通過させる。存在を判定するための単一閾値に頼る代わりに、プローブセットの3つの四分位数すべてが、応答スコアの分布が確実に1に向けて歪められるように調べられる。まとめると、0.020の予想されない陽性OTU率を有する > 2 および < 348 pMの単位複製配列について0.961の予想される陽性OTU率を達成するためには、 r_{Q_1} 0.70、 r_{Q_2} 0.95、および r_{Q_3} 0.98が必要であった。ステージ2では、0.48の r_{xQ_3} 亜科閾値セットは、同じ濃度範囲にわたりラテン方格

10

20

30

40

50

データに適用された場合は、対応する0.019の予想されない陽性垂科率とともに0.969の予想される陽性垂科率を可能にした。

【0444】

ハイブリダイゼーション結果は、多変量統計に有用なフォーマットで各PhyloChipアッセイからの群集プロファイルに変えられた。ステージ2を通過する垂科内のステージ1を通過するOUTが前記群集プロファイルを構成した。ラテン方格モック群集の複製群集プロファイルは、序列により比較された。プロファイル間距離は、ブレイカーチス法または荷重Unifrac法のどちらかを用いて計算され、得られた距離行列は、非計量的多次元尺度法(NMDS)を用いて正しく並べられた。26のモック群集それぞれ由来のプロファイルは、どちらの距離法を使用してもはっきりと区別可能であった。どちらの距離行列(Adonis)を使用する分散分析も、モック群集間の有意差($p < 0.005$)を結論付けた。

10

【0445】

結果

前記ブルームは、微生物群集系統発生的組成と構造を著しく変えた。系統発生的マイクロアレイ(PhyloChipアッセイ)を使用して、検出可能な細菌の豊富さの40%減退および微生物群集組成の著しい変化が見出された。系統発生的マイクロアレイ分析により判定される群集組成の序列は、試料の2つのはっきり異なるクラスター、すなわち検出された油を有する試料で完全に構成された1つクラスターと油が検出されていない試料を有するもう1つのクラスターを明らかにした。これらの群間には、炭化水素以外の他の物理的または化学的要因には有意差はなく、微生物が分散油の存在に直接的に応答していることを示している。

20

【0446】

プロテオバクテリア綱の細菌のみがブルーム試料中で著しく濃縮されていた(表15)。ブルーム試料中、62門中951の異なる細菌分類群が検出されたが、gプロテオバクテリアとしてすべて分類された16の異なる分類群のみがブルームの外側の深海と比べてブルームにより著しく濃縮されていた(表15、図33)。濃縮された分類群のほぼすべてが炭化水素を分解させることが知られている、または寒冷環境では油の存在に刺激される(表15)。ブルーム濃縮細菌には、寒冷海水、海氷および周極生息地から知られている多くの好冷性および耐寒性種が含まれる。前記結果によれば、これらのプロテオバクテリアは深海ブルームにおいて微生物群集を支配していることが示される。細胞密度は高いが、分類学的豊かさは低く、濃縮された細菌の多様性はこれらの少数のプロテオバクテリアに制限されている。プロテオバクテリアのオセアノスピリルム目はPhyloChipアッセイにより分析された9つの油ブルーム試料すべてにおいて検出され、油のないバックグラウンド深海と比べて著しく濃縮されていた。

30

【0447】

【表 14】

表 15 油ブルームにより濃縮されたγプロテオバクテリア分類群。既知の炭化水素分解菌を含むまたは寒冷水域では炭化水素に応答して濃縮されることが以前明らかにされている分類群が示されている。

綱	科	炭化水素分解菌*	原油により濃縮される*	代表的配列	
エロモナス綱 (Aeromonadaceae)	エロモナス科 (Aeromonadaceae)	+	+	DQ816633.1 ゼブラフィッシュ腸クローン	10
アルテロモナス綱	コルウェリア科 (Colwelliaceae)	ND	+	EU491914.1 東太平洋海膨深海クローン	
アルテロモナス綱	シュードアルテロモナス科 (Pseudoalteromonadaceae)	+	+	AY646431.1 シュードアルテロモナス種 (Pseudoalteromonas)	
Arctic96B-1	未分類	ND	+	EU544859.1 北極海水クローン	
BPC036	未分類	ND	+	DQ925906.1 グアイマ ス海盆クローン	20
ハロモナス綱	ハロモナス科	+	+	DQ270747.1 ハロモナス種	
マリノバクター綱	マリノバクター科	+	+	DQ157009.2 マリノバ クターハロテリジェナ (Marinobacter haloterrigenus)	
マリノスピリラム綱 (Marinospirillum)	マリノスピリラム科	+	+	AF275713.1 マリノス ピリラムアルカリフィ ラム	
モラクサラ綱	モラクサラ科	+	+	AF200213.1 好冷性海 洋分離菌	30
オセアノスピリラム綱	マリノバクテリウム科	+	+	AY549003.2 海洋ボー ンクローン	
オセアノスピリラム綱	マリノモナス科 (Marinomonas)	+	+	EF673290.1 マリノモ ナス種	
オセアノスピリラム綱	未分類	+	+	AM747817.1 オセアニ セルペンティラハリオ チジス (Oceaniserpentilla haliotidis)	
シュードモナス綱	シュードモナス科	+	+	AM111047.1 シュード モナス種	40
シュワネラ綱	シュワネラ科	+	+	DQ665797.1 シュワネ ラフリギディマリナ (Shewanella frigidimarina)	
未分類	未分類_sfB	ND	ND	EU491790.1 東太平洋 海膨海底クローン	
未分類	未分類_sfC	ND	ND	EU652559.1 エル海洋 堆積物クローン	

ND=データなし

【0448】

図33は、油ブルームによる選抜細菌分類群の濃縮の図解を提供している。系統発生マイクロアレイ分析を使用して、ブルームと非ブルーム試料間の推定濃度の平均差を計算した。平均差は、濃縮された分類群亜科における代表的OTUについての非ブルーム濃度の割合として示されている(表15)。

【0449】

本発明の好ましい実施形態が本明細書において明らかにされ記載されてきたが、そのような実施形態は例としてのみ提供されていることは当業者には明白であろう。本発明から逸脱することなく、当業者であれば今や数多くの変形、変化および置き換えを思いつくで

10

20

30

40

50

あろう。本明細書に記載されている本発明の実施形態に対する様々な代替手段が本発明を実行する際に用いられることは理解されるべきである。以下の特許請求の範囲は本発明の範囲を限定しており、これらの特許請求の範囲内の方法および構造ならびにその同等物はそれにより含まれていることが意図されている。

【符号の説明】

【 0 4 5 0 】

- 106 システムメモリー
- 104 プロセシングユニット
- 156 図表示インターフェース
- 108 システムバス
- 126 ハードディスクドライブインターフェース
- 128 磁気ドライブインターフェース
- 130 光ドライブインターフェース
- 152 ネットワークインターフェース
- 144 ユーザー入力インターフェース
- 132 オペレーティングシステム
- 134 アプリケーションプログラム
- 136 他のプログラムモジュール
- 138 プログラムデータ

【図 1】

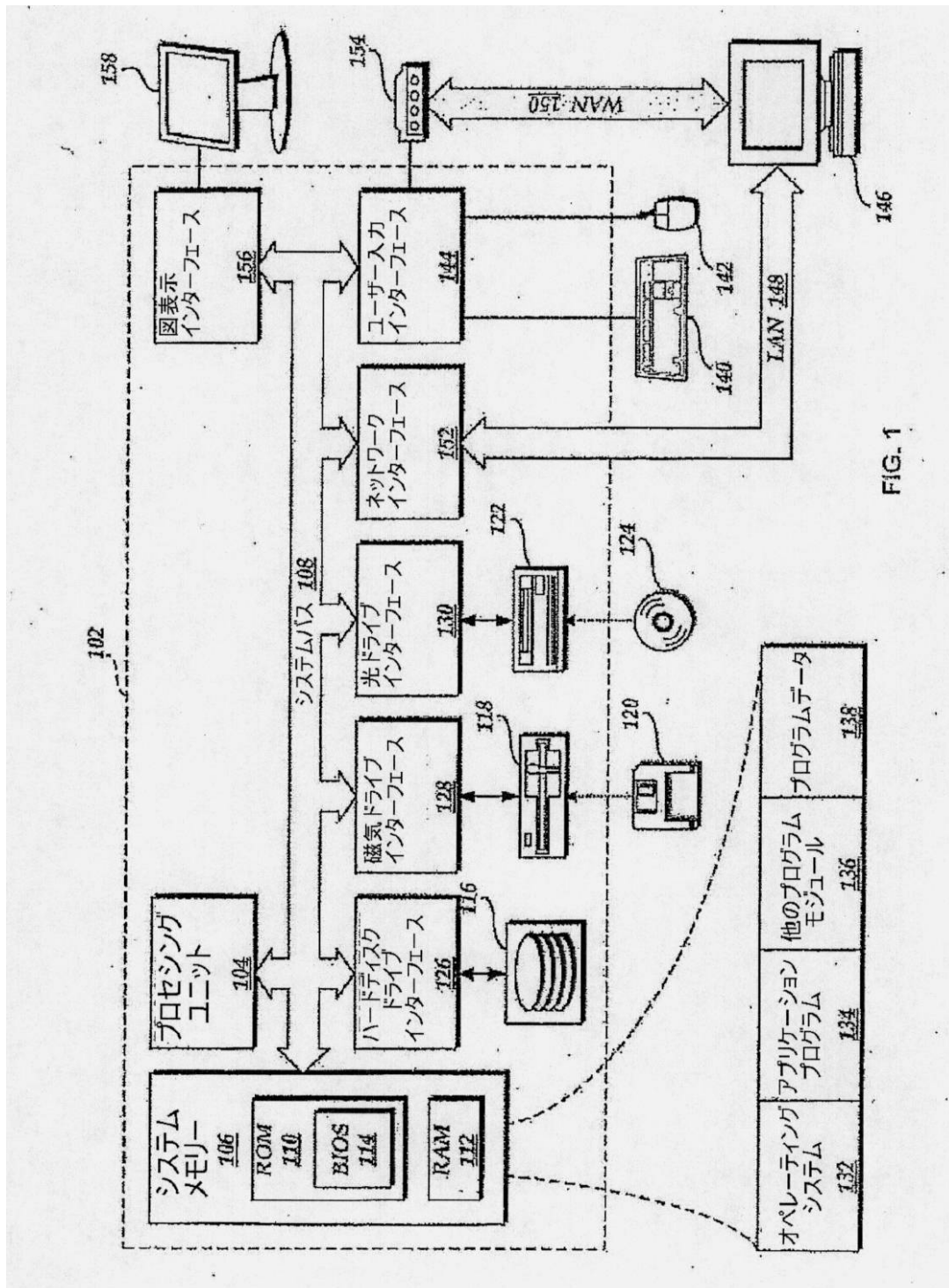


FIG. 1

【図 2】

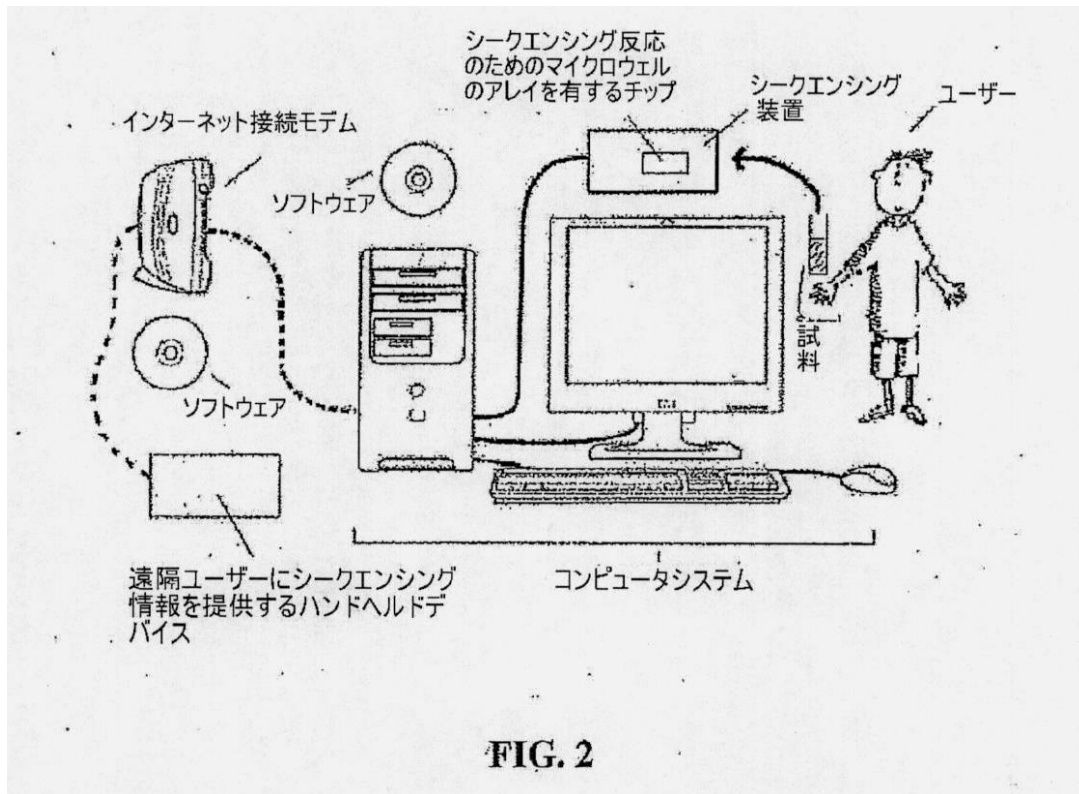
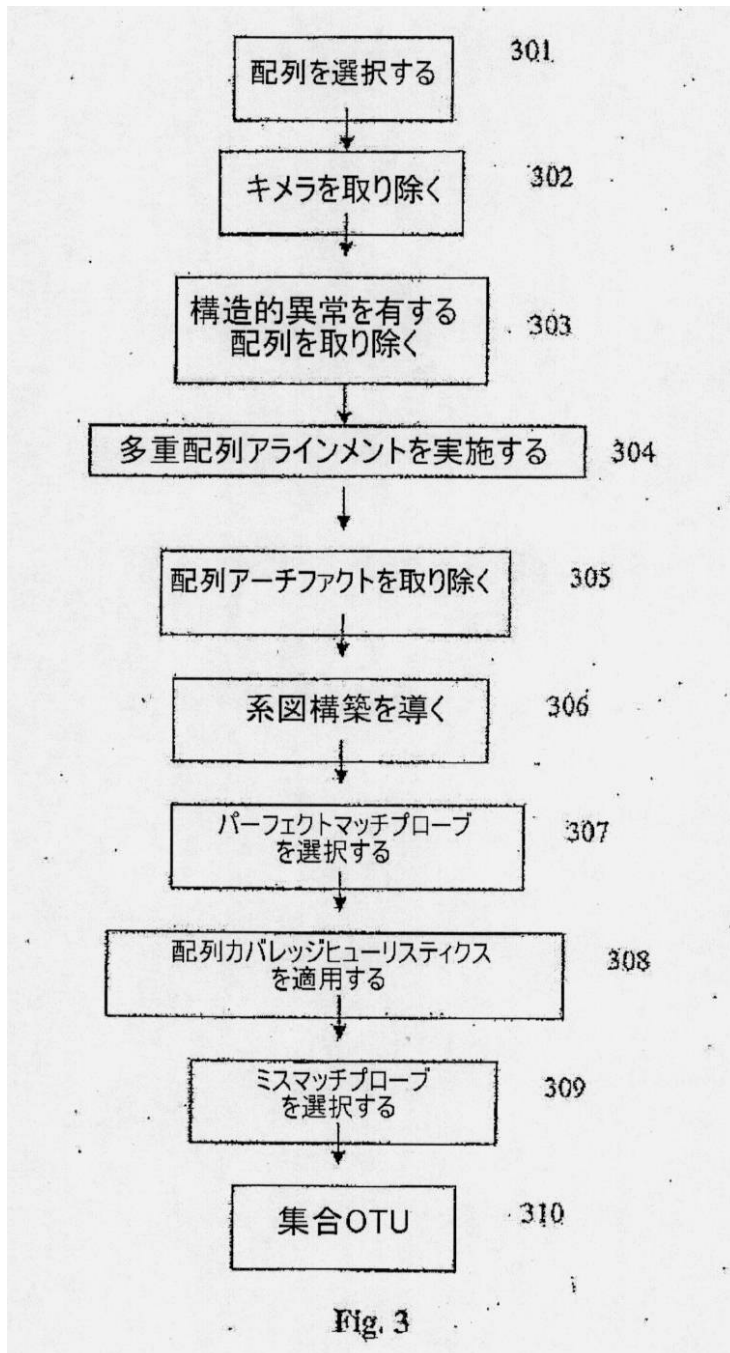
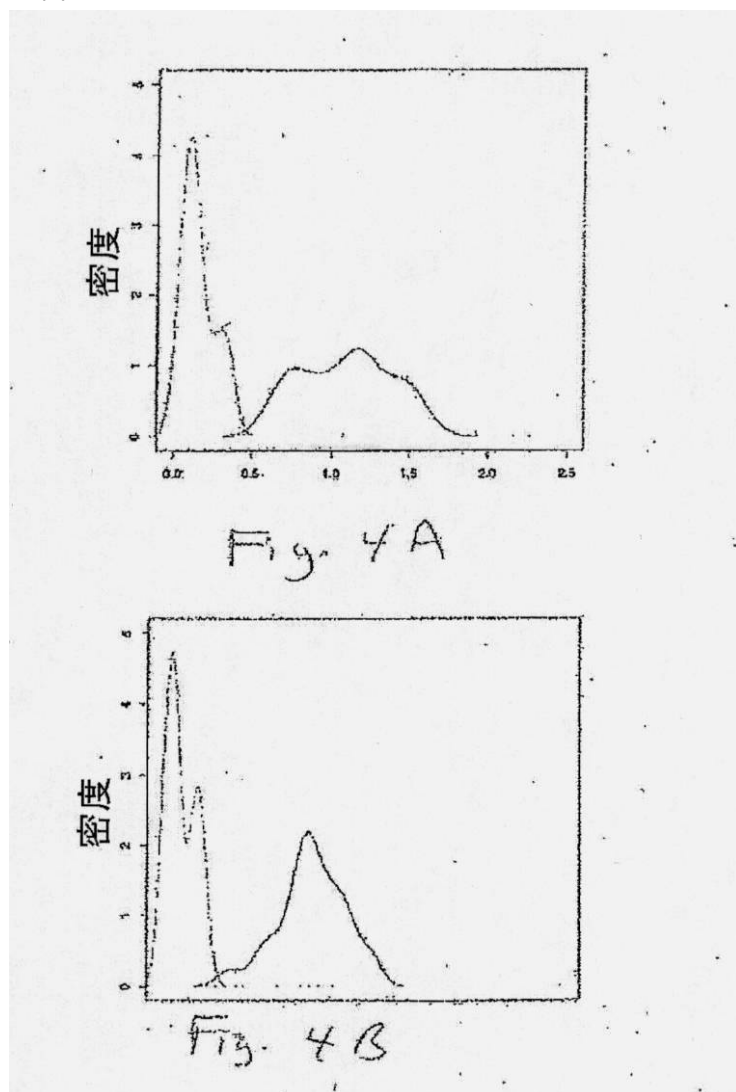


FIG. 2

【図 3】



【図 4】



【図 5】

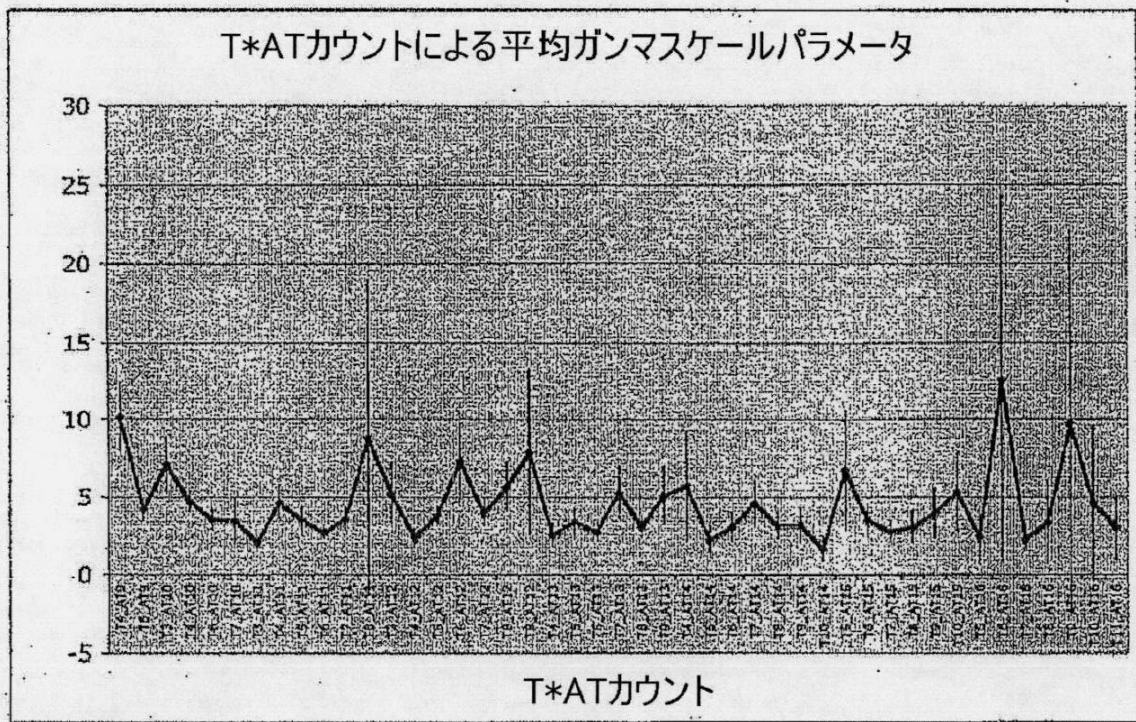
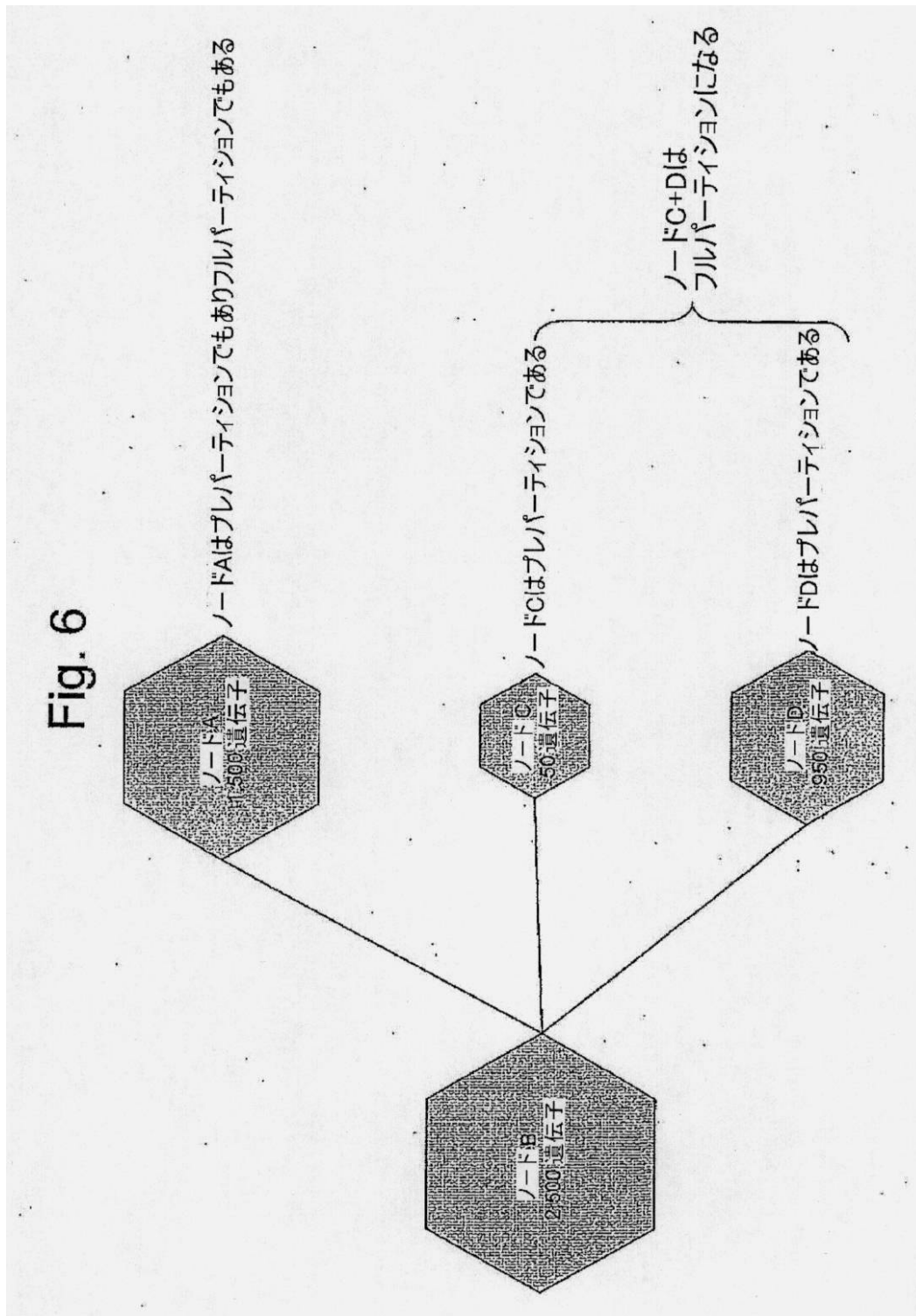
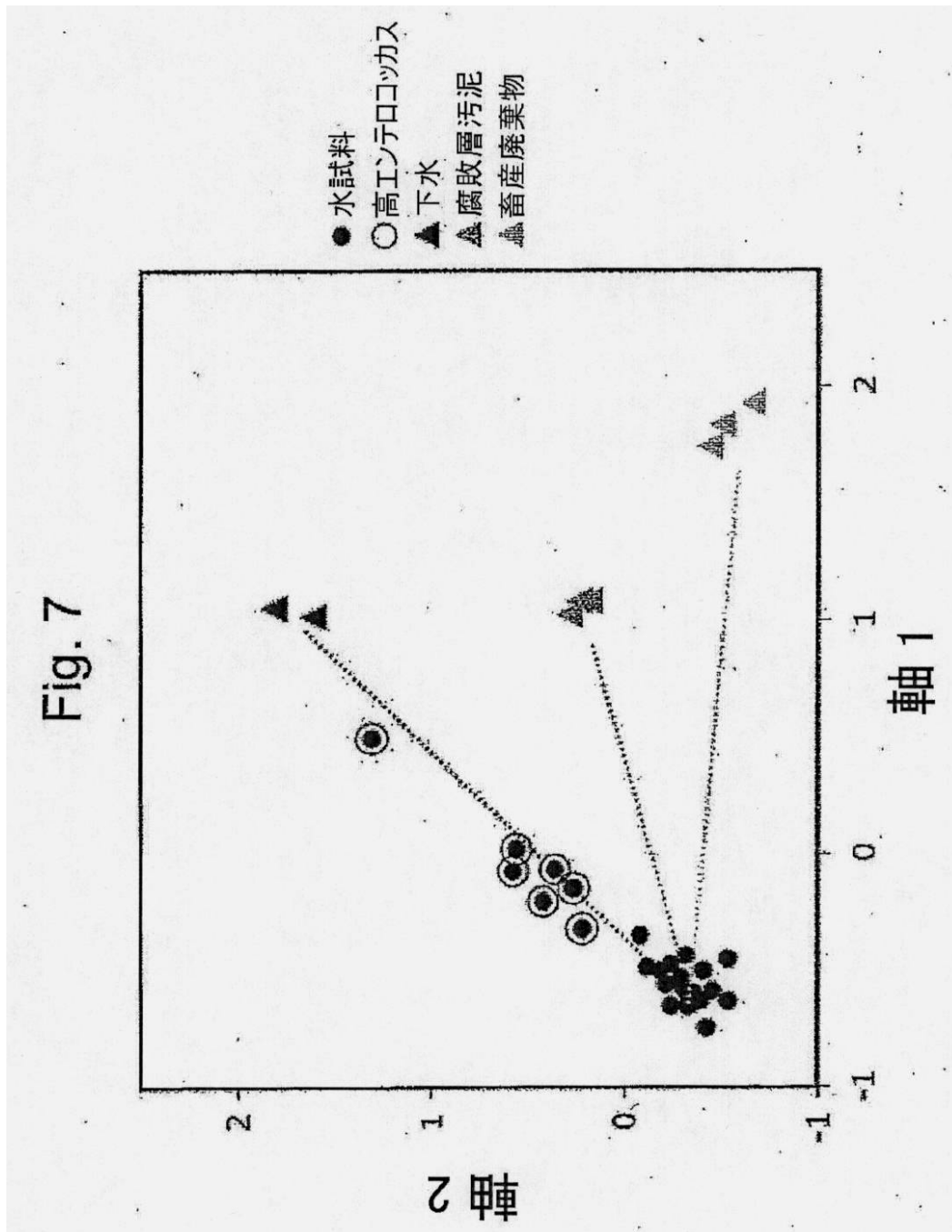


Fig. 5

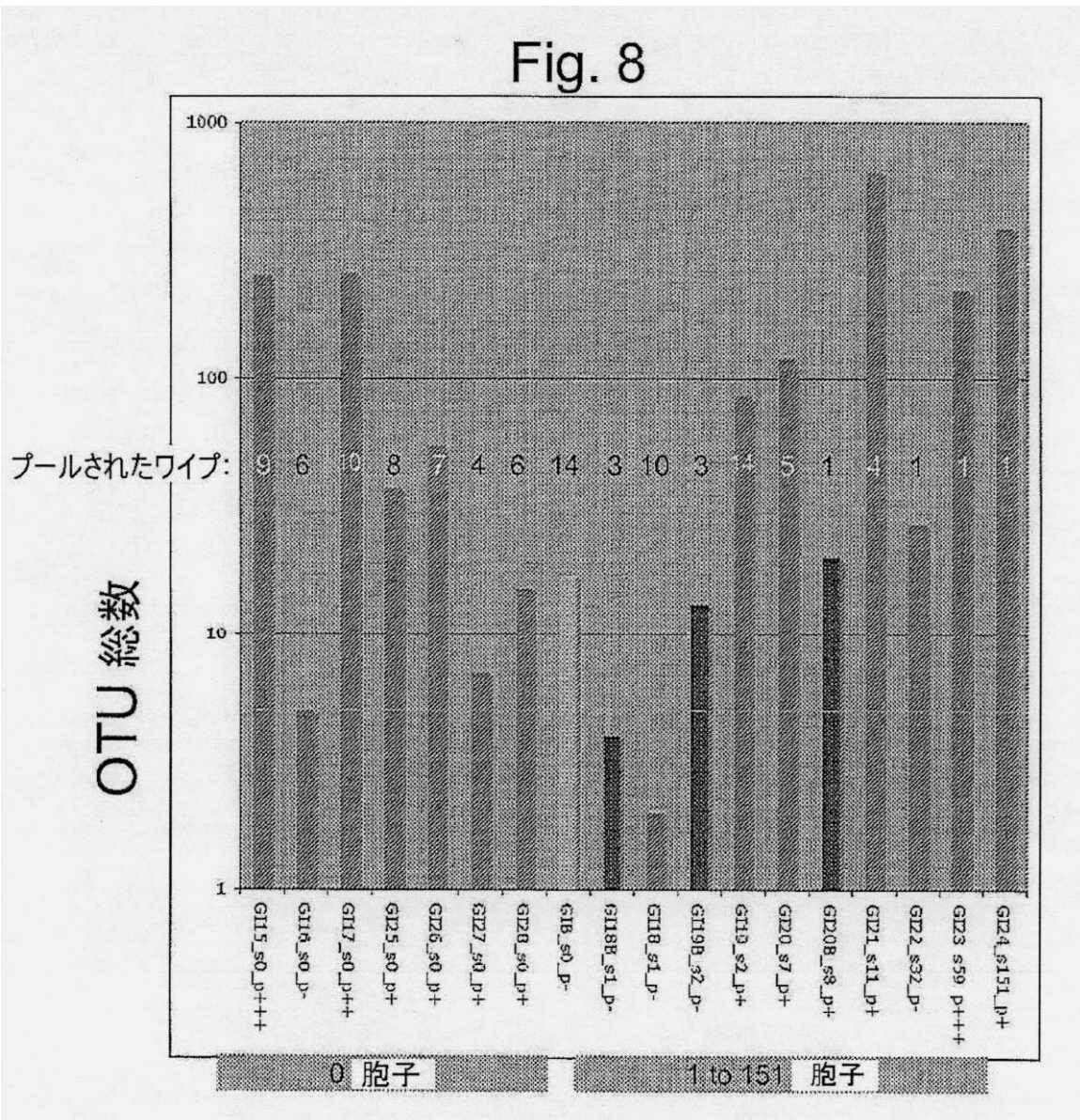
【図 6】



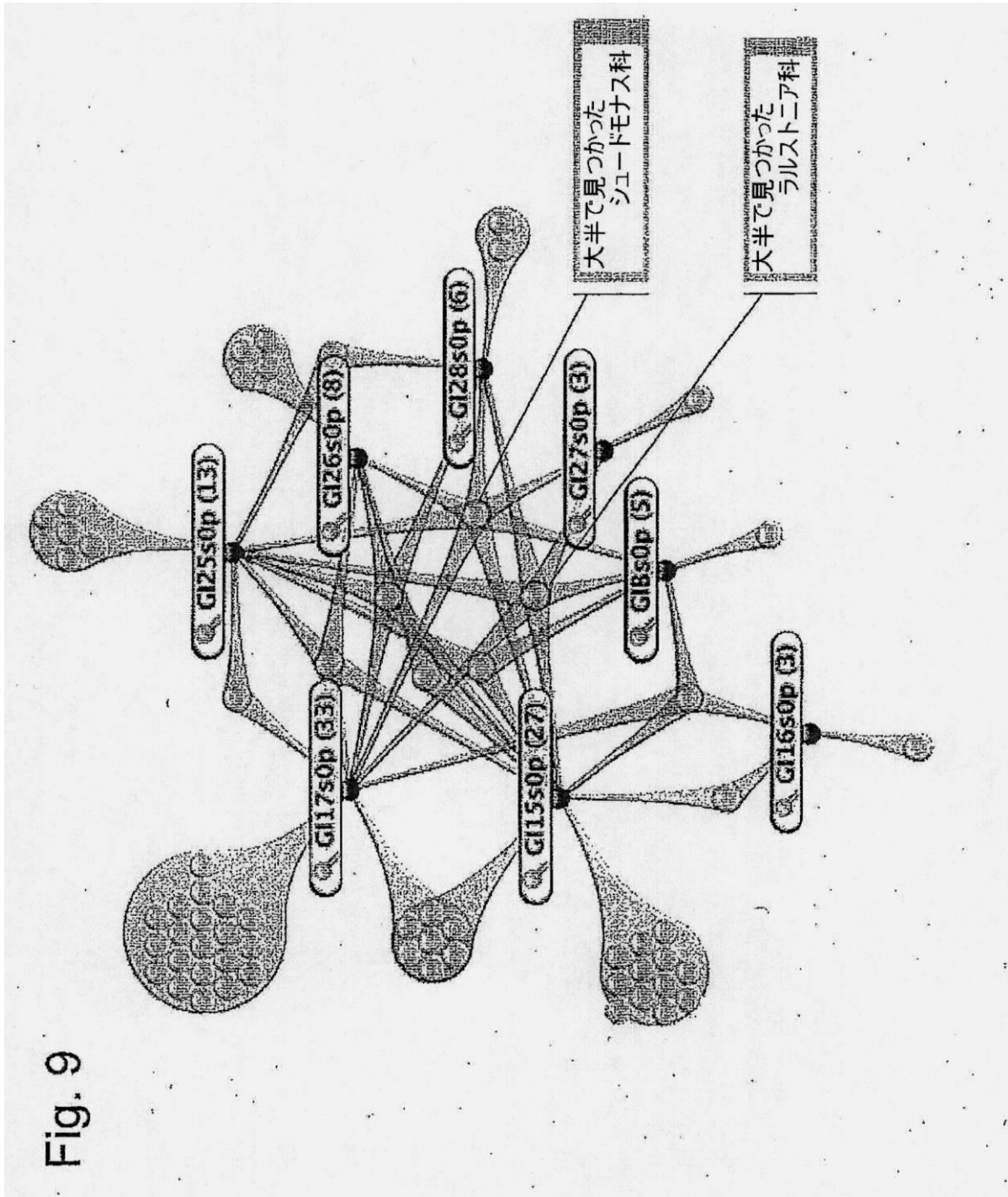
【図 7】



【図 8】

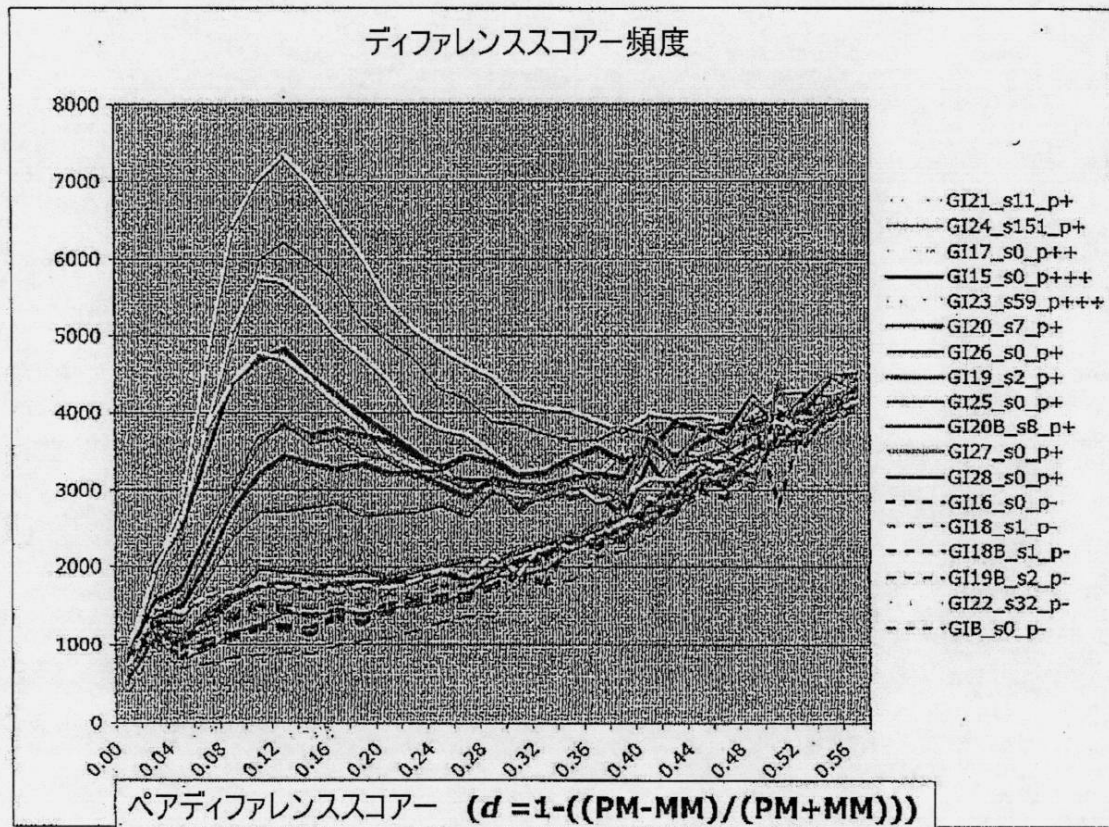


【図 9】



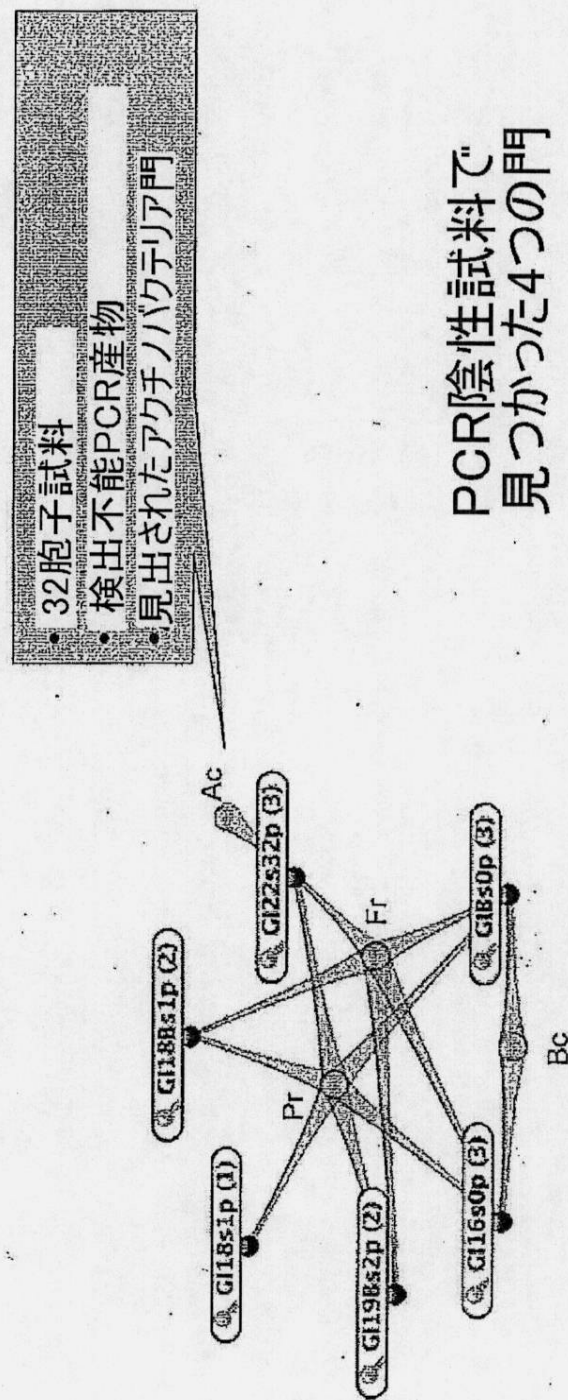
【図10A】

Fig10A



【図10B】

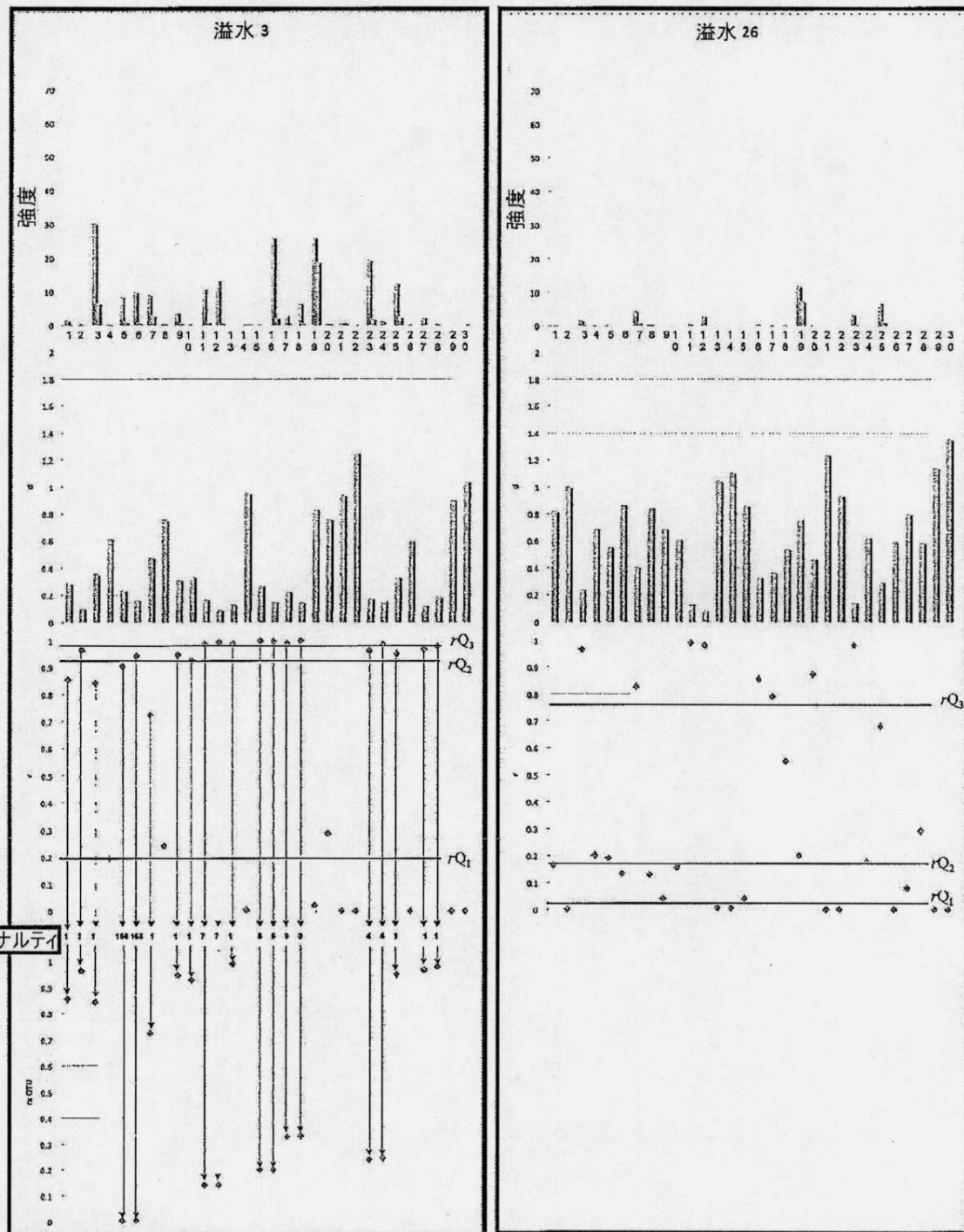
Fig. 10B



Pr: プロテオバクテリア門 Fr: ファーミキューテス門 Bc: バクテロイデス門 AO: ABY1_OD1
Ac: アクチノバクテリア門 Gyl: シアノバクテリア門 Pl: プランクトミケス門

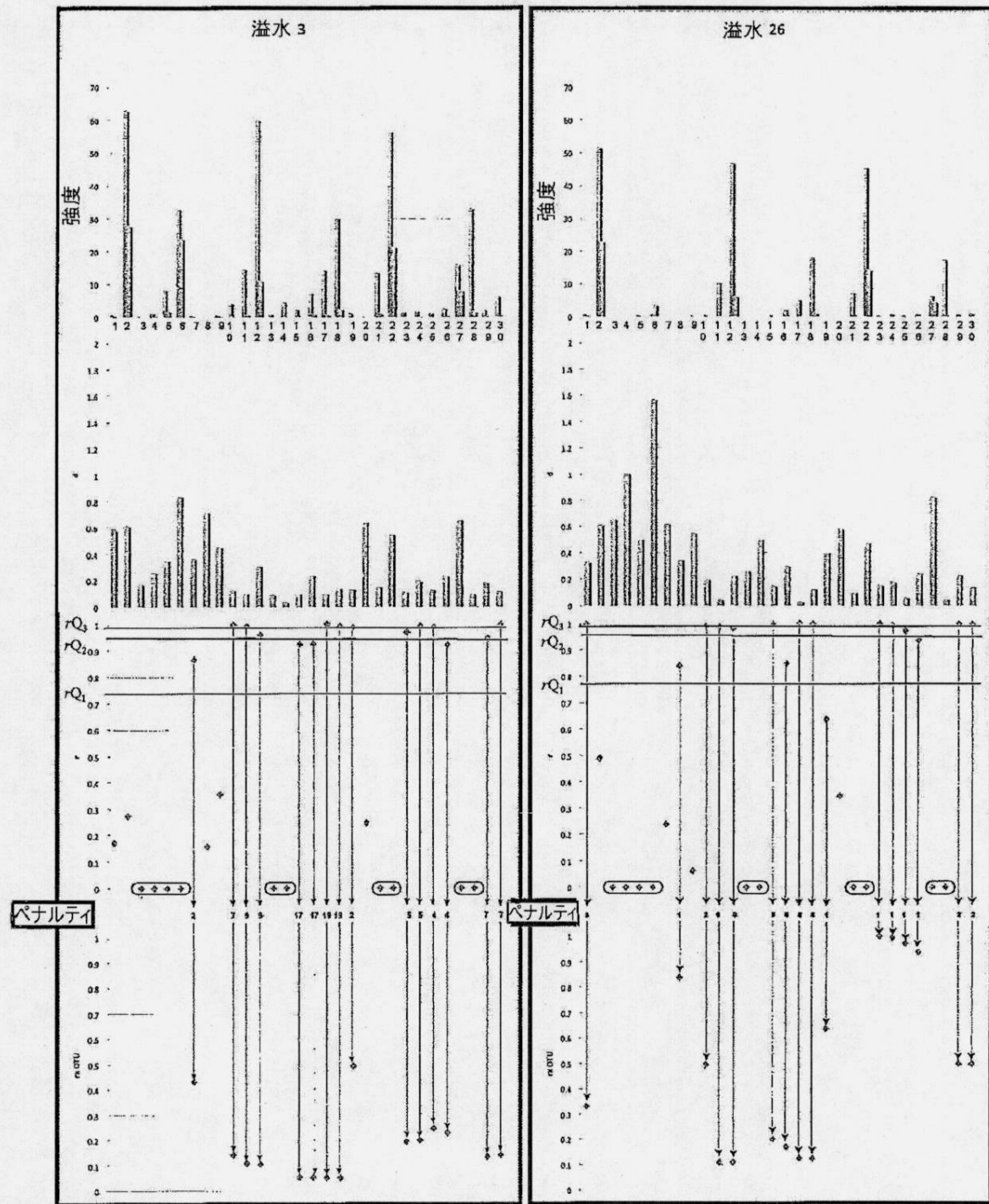
【図 11】

Figure 11 糞便細菌OTU36742



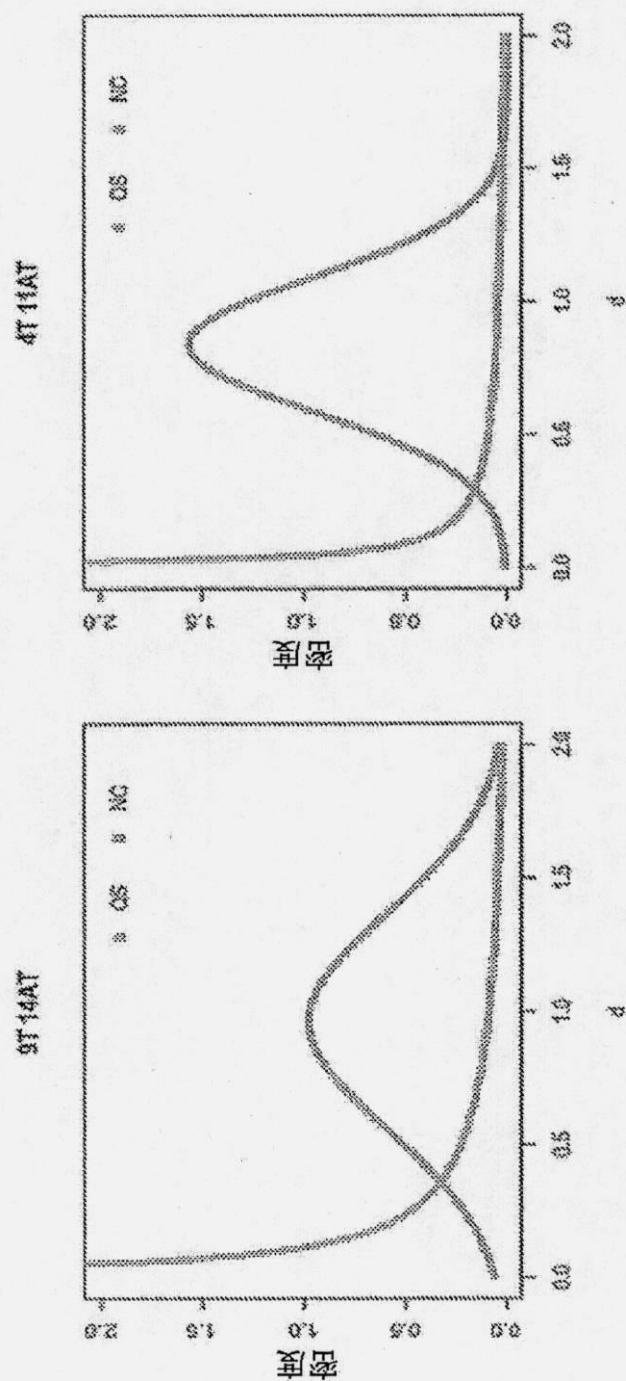
【図 12】

Figure 12 ルミノコッカスOTU38712



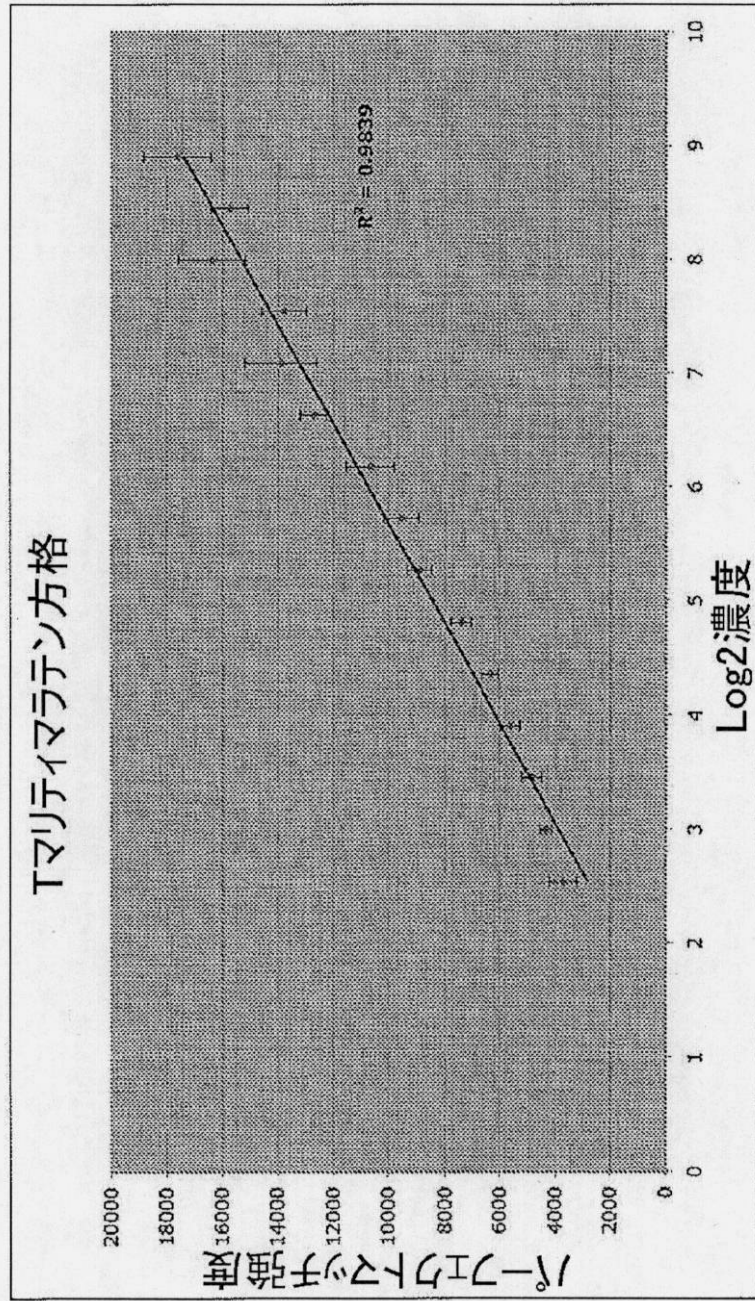
【図 13】

Fig. 13



【図 14】

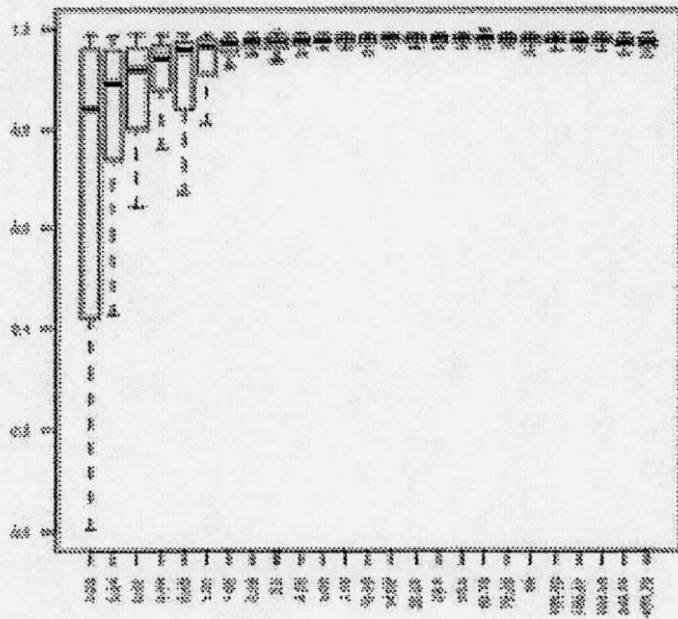
Fig. 14



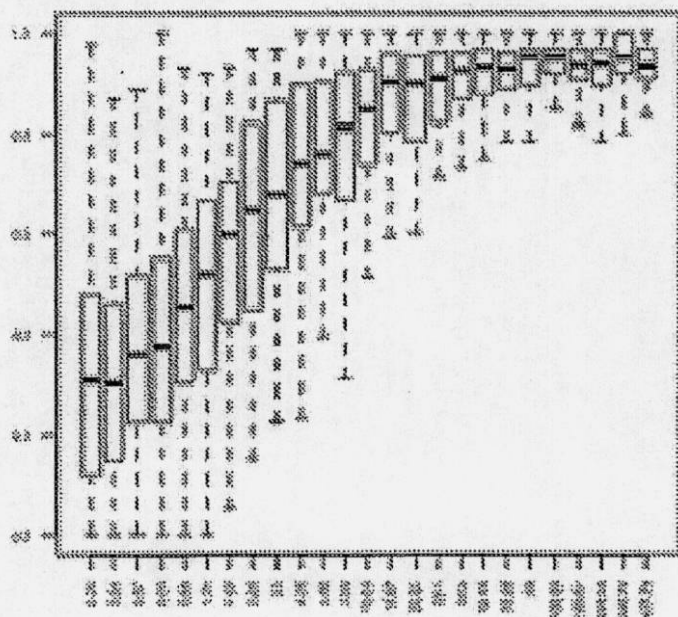
【図 15】

Fig. 15

q2_g2qn.pdf 濃度による

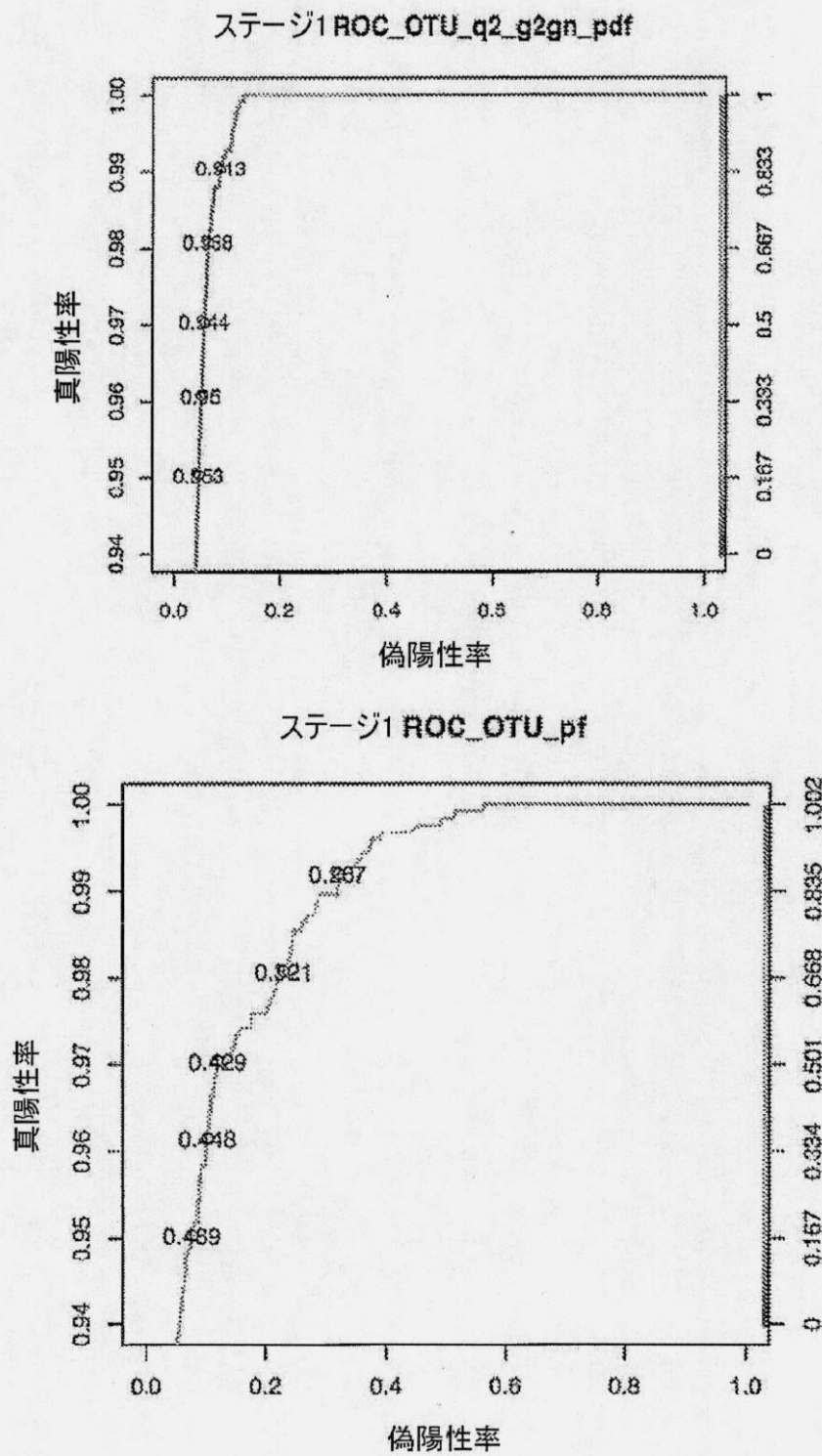


p1 濃度による



【図 16】

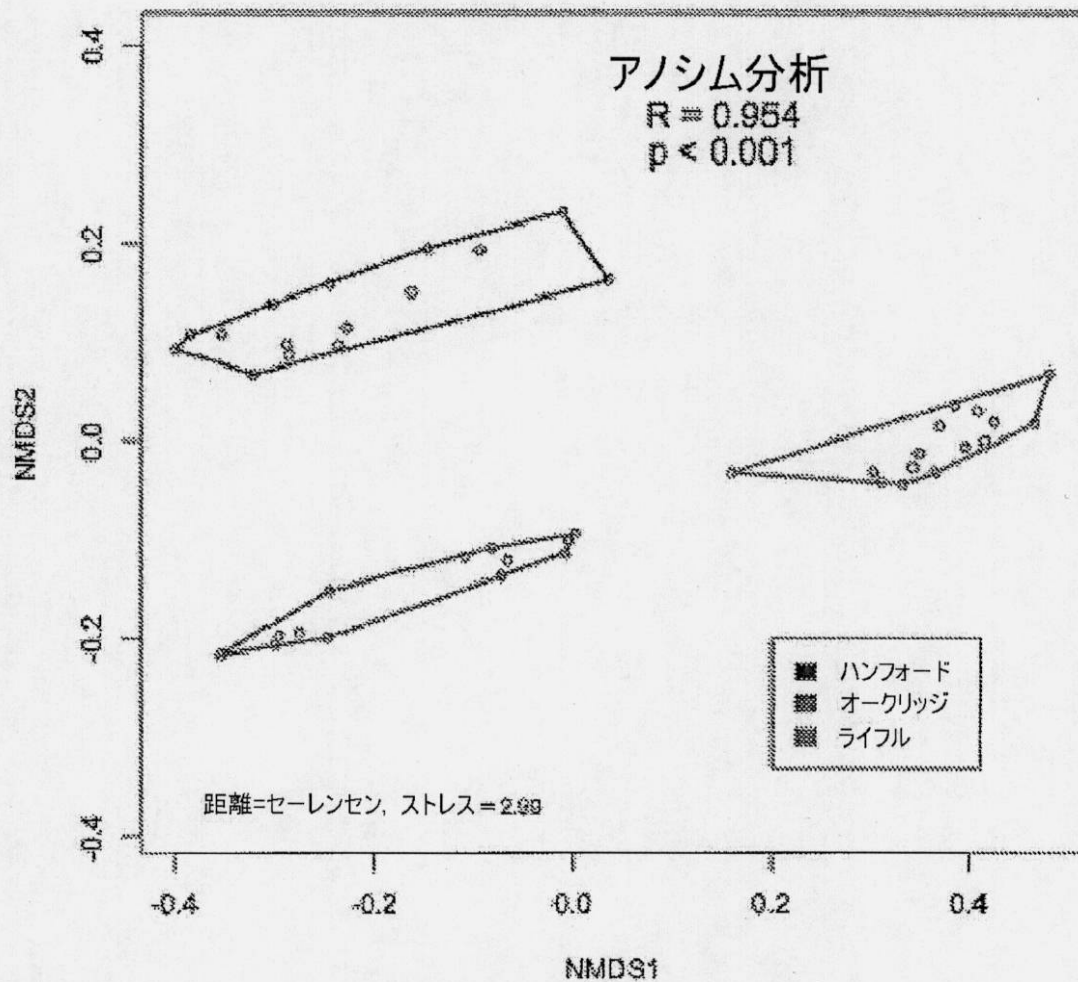
Fig. 16



【図 17】

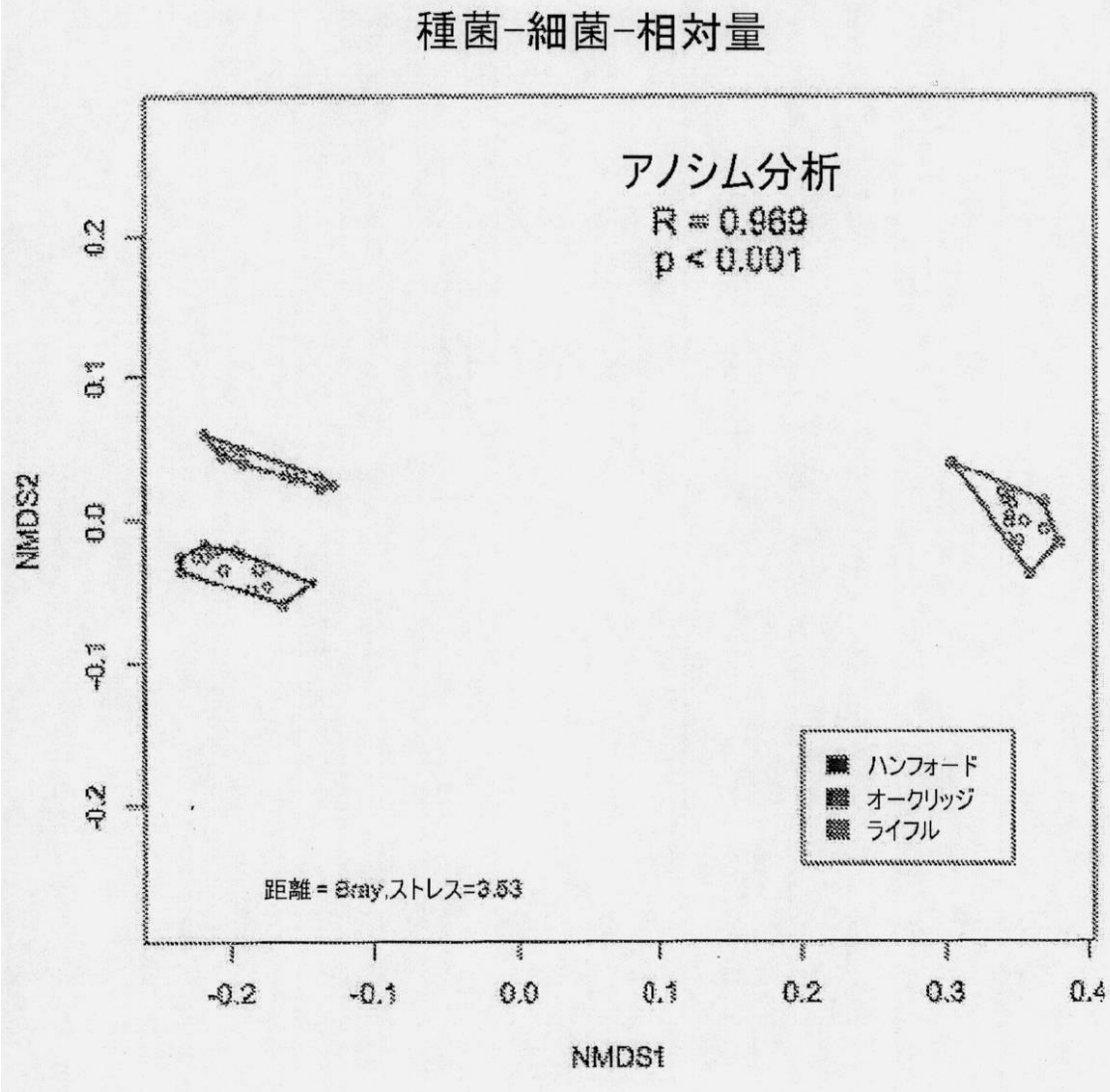
Fig. 17

種菌-細菌-存在/非存在

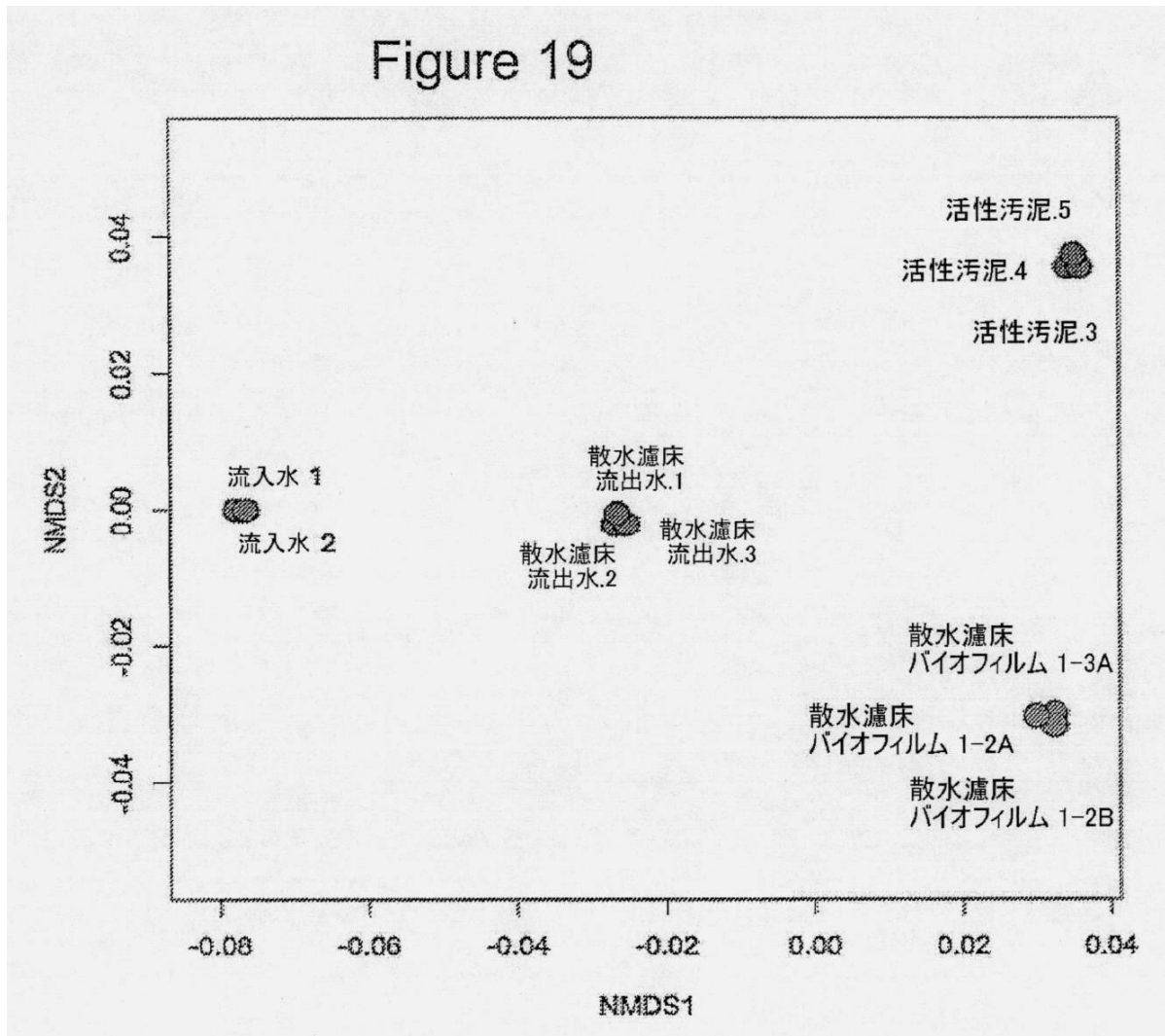


【図18】

Fig. 18

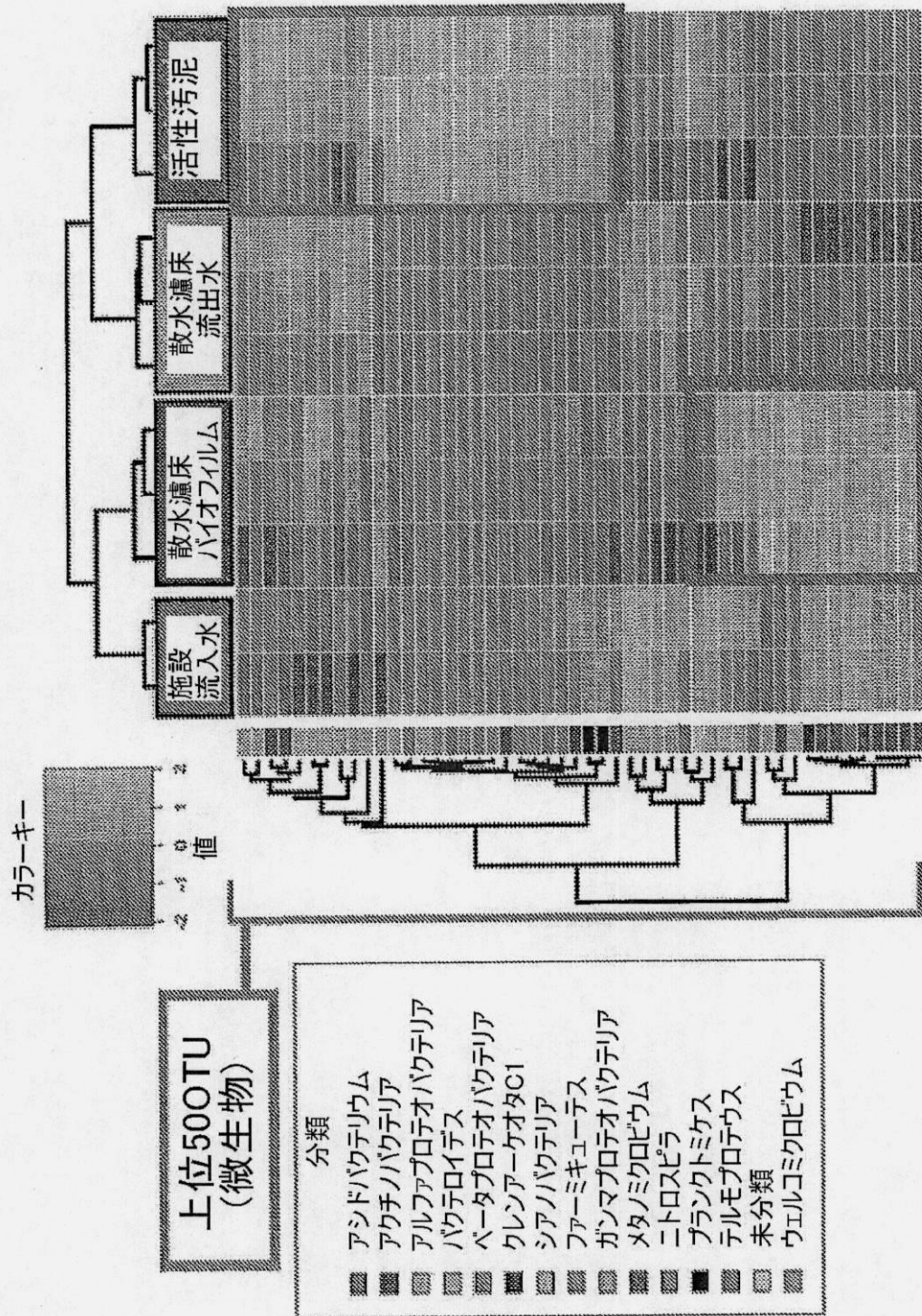


【図 19】



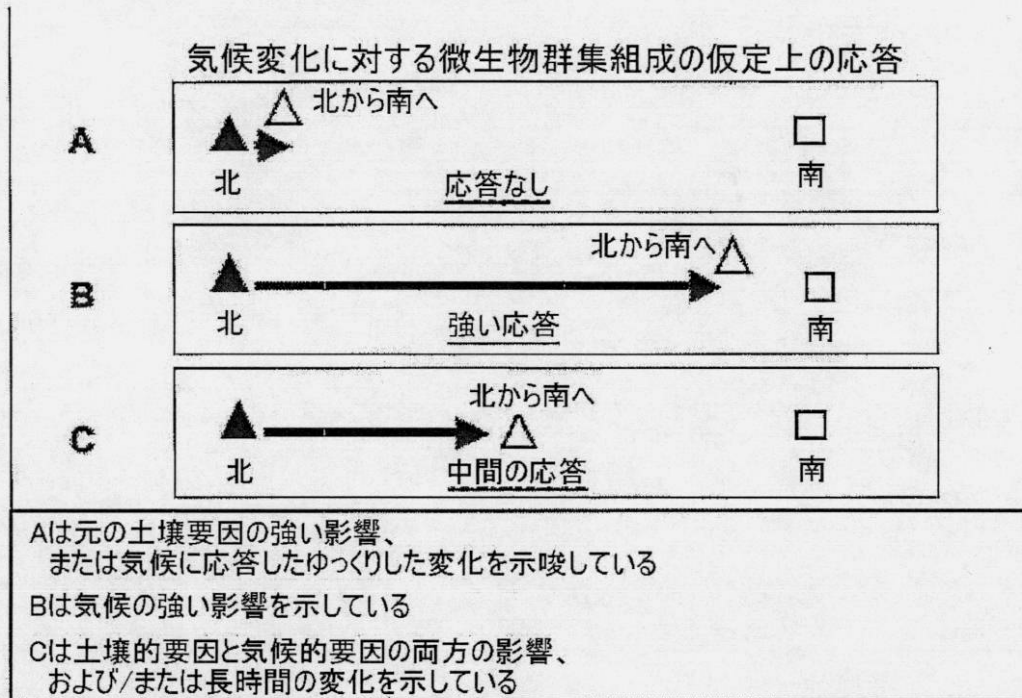
【図20】

Figure 20



【図 21】

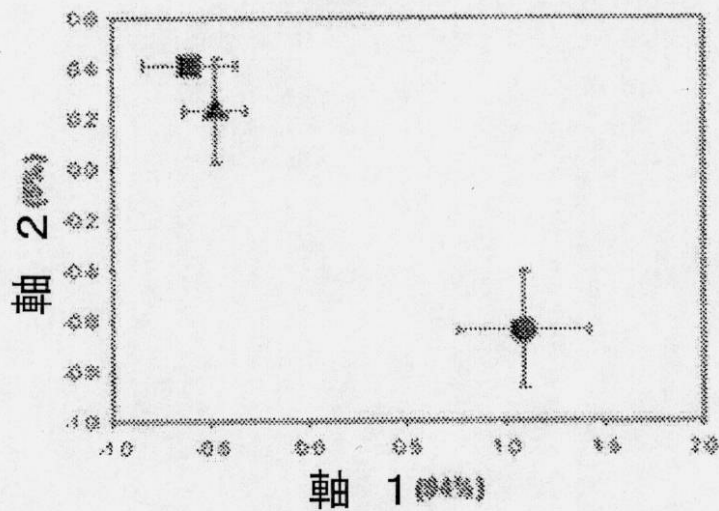
Figure 21



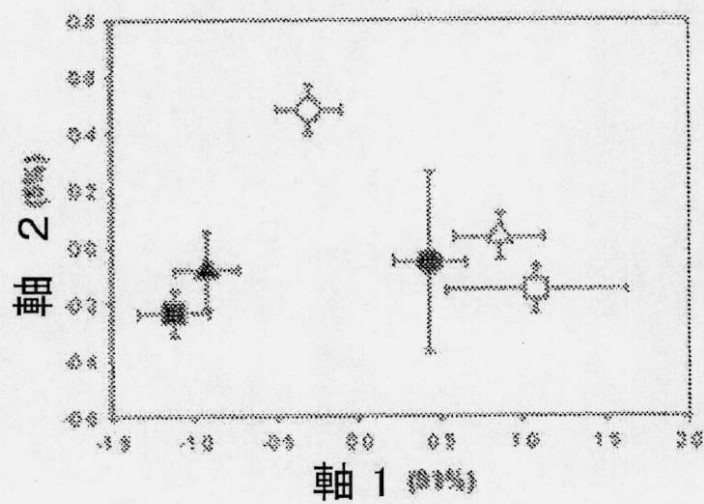
【図 22】

Figure 22

PhyloChip細菌OTU新鮮な試料のNMS、
2005年8月



PhyloChip細菌OTU新鮮な試料および対照のNMS、
1年



レジェンド

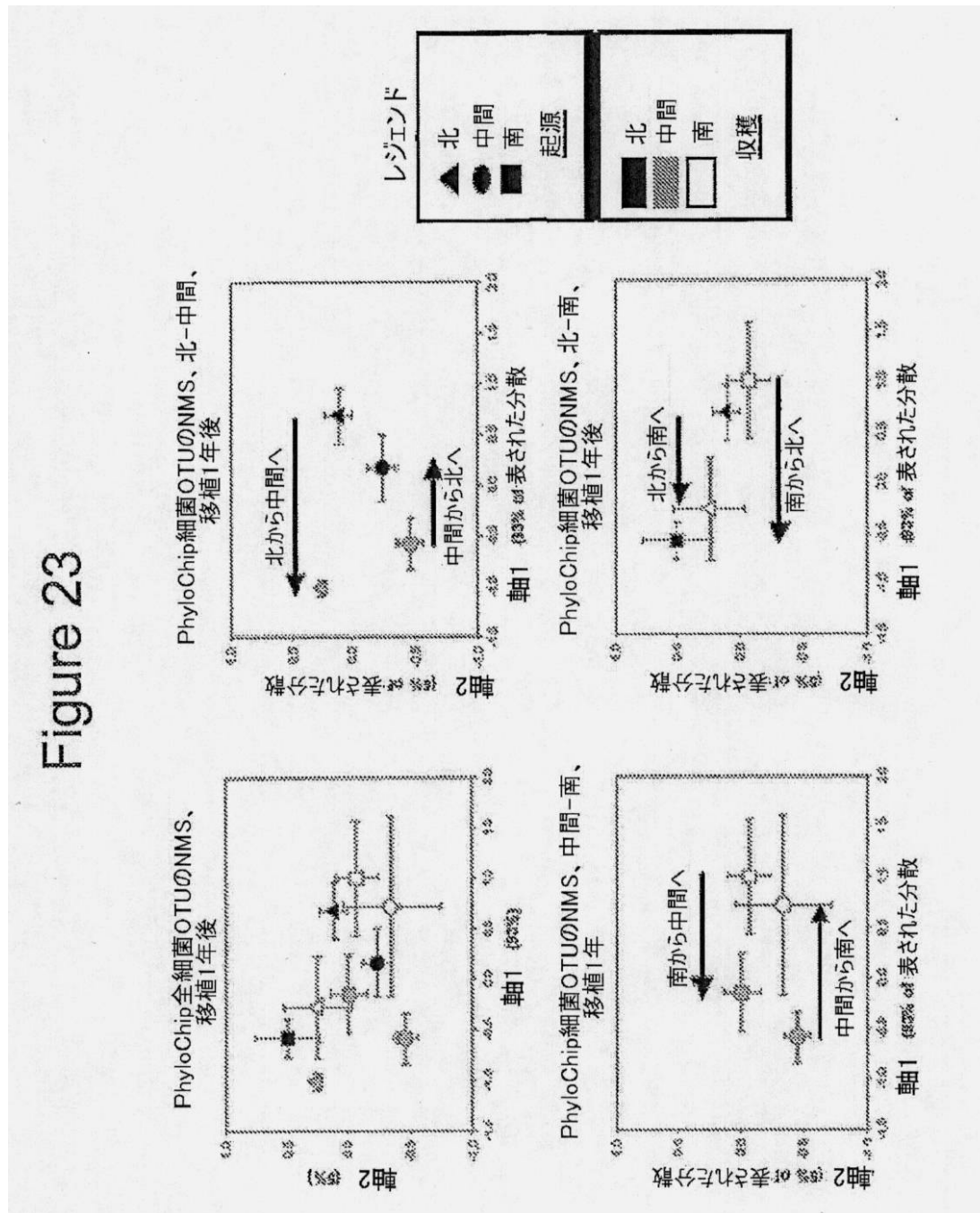
▲ 北
● 中間
■ 南

地点

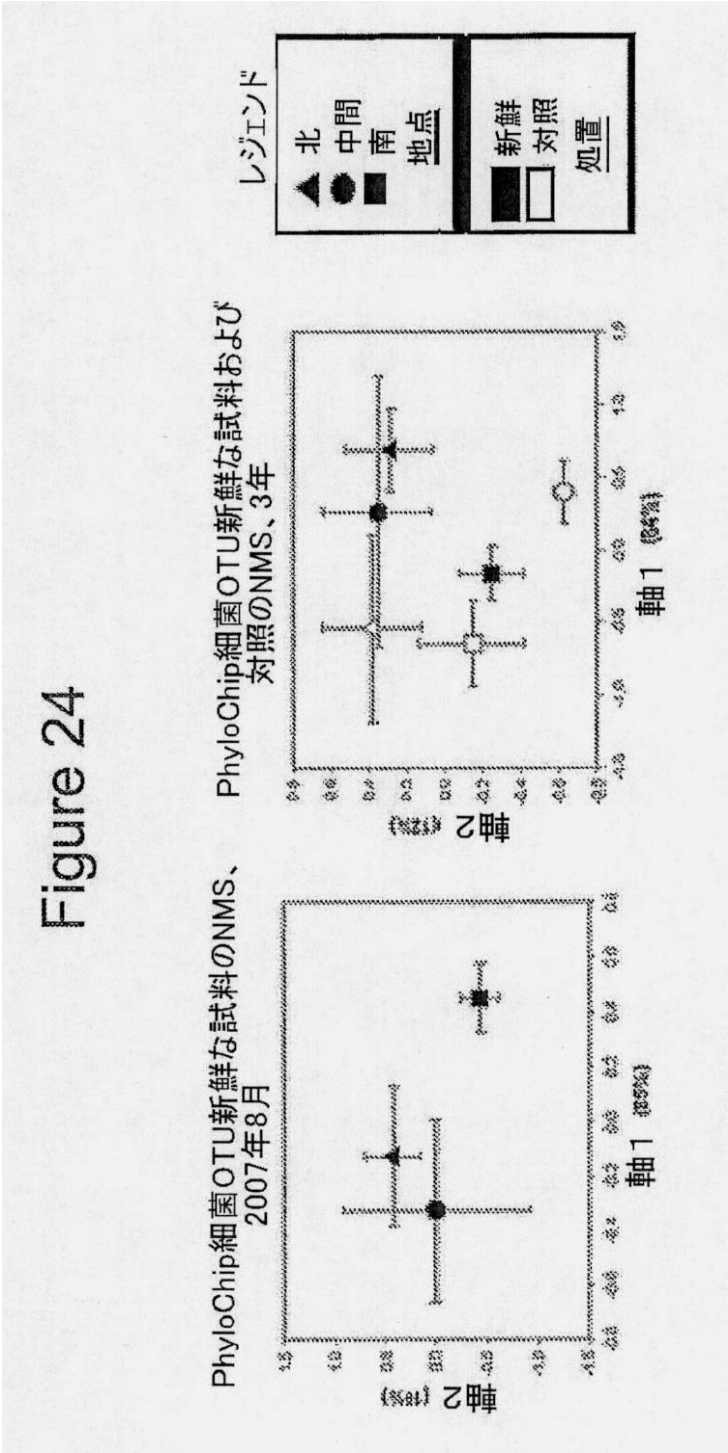
■ 新鮮
□ 対照

処置

【図 23】

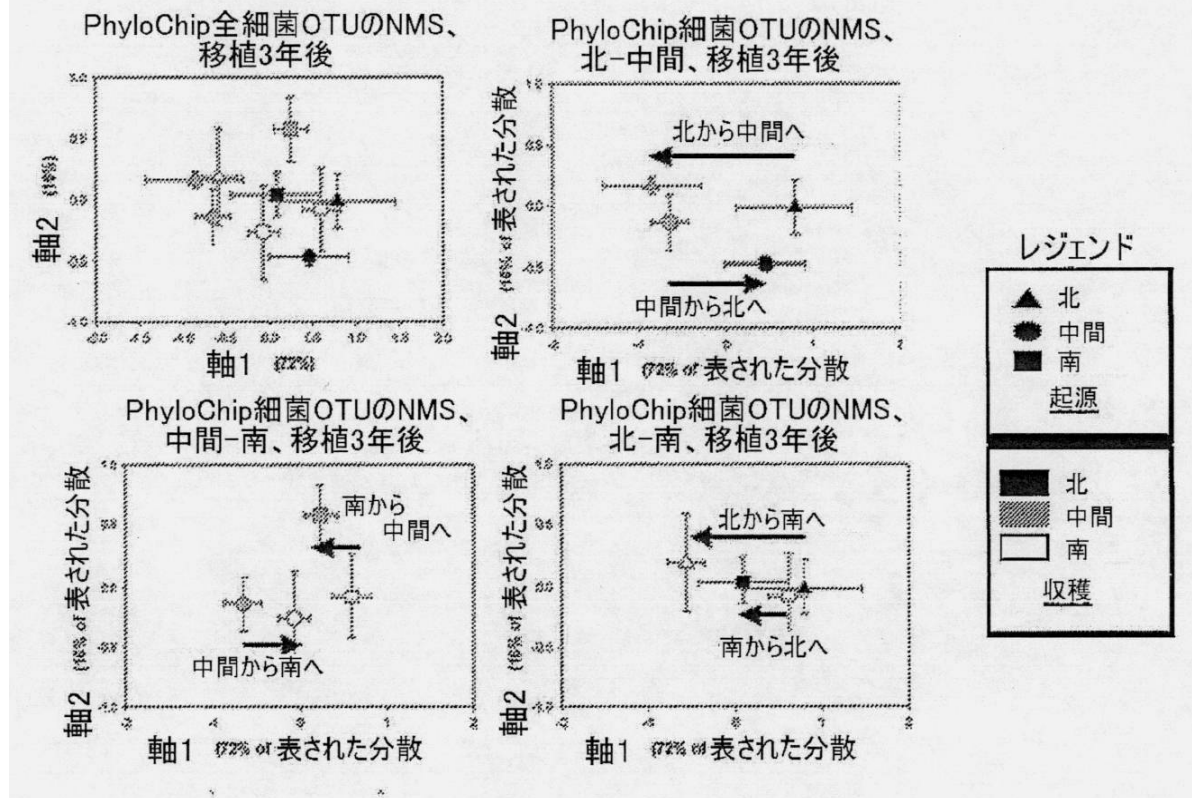


【 図 2 4 】

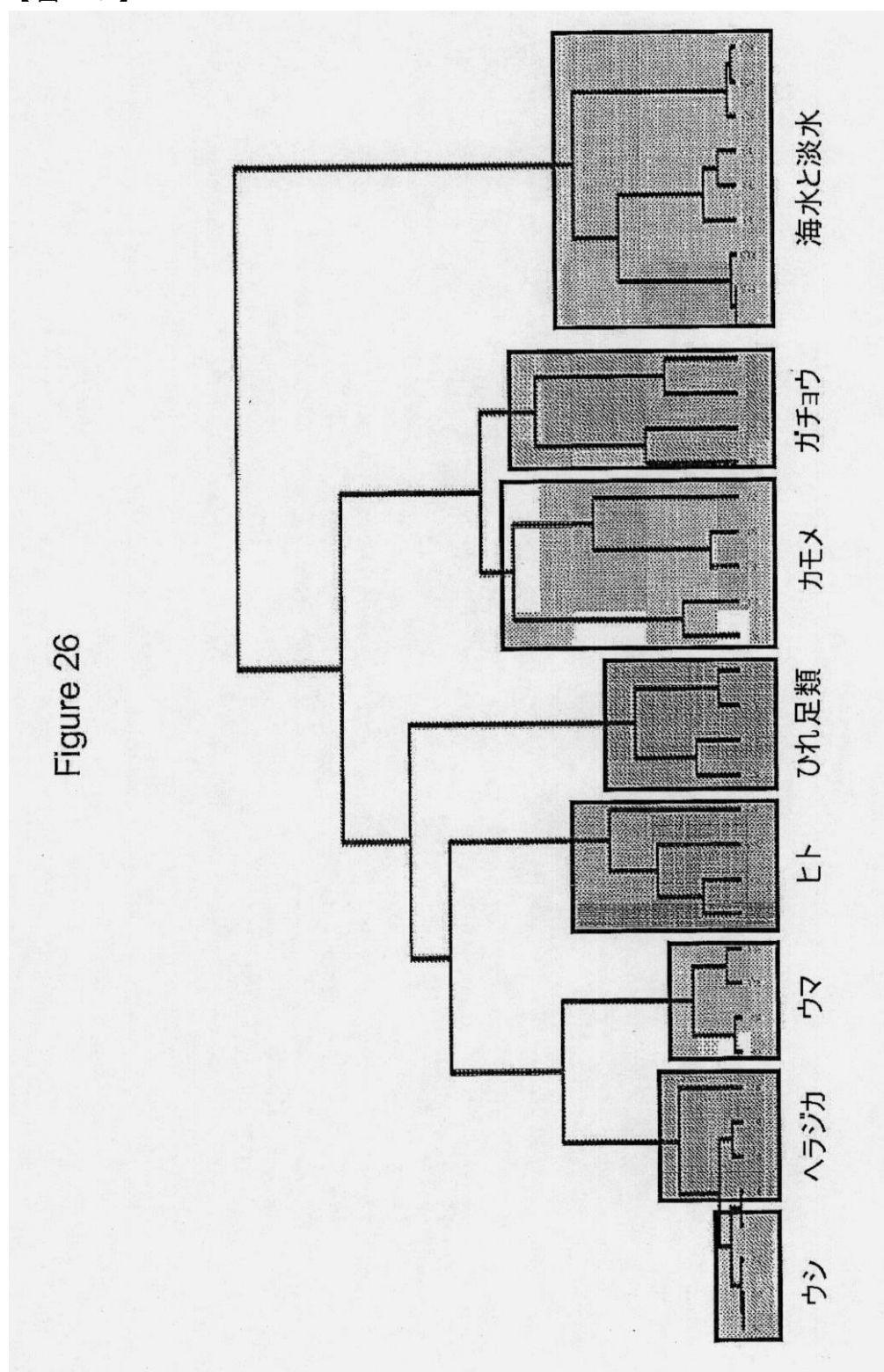


【図 25】

Figure 25

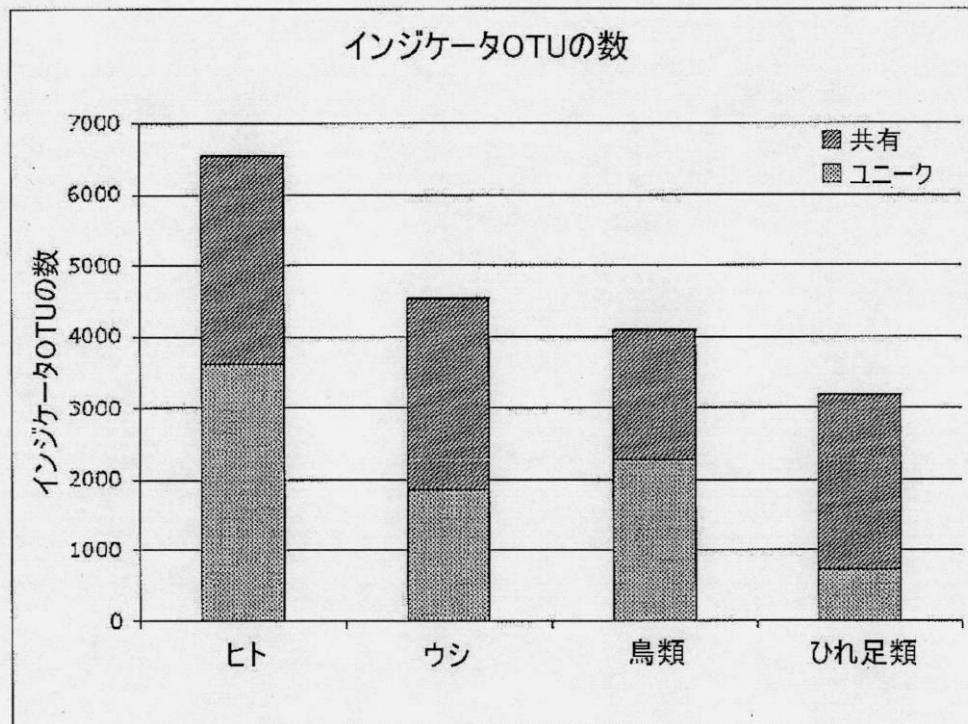


【 图 2 6 】



【図 27】

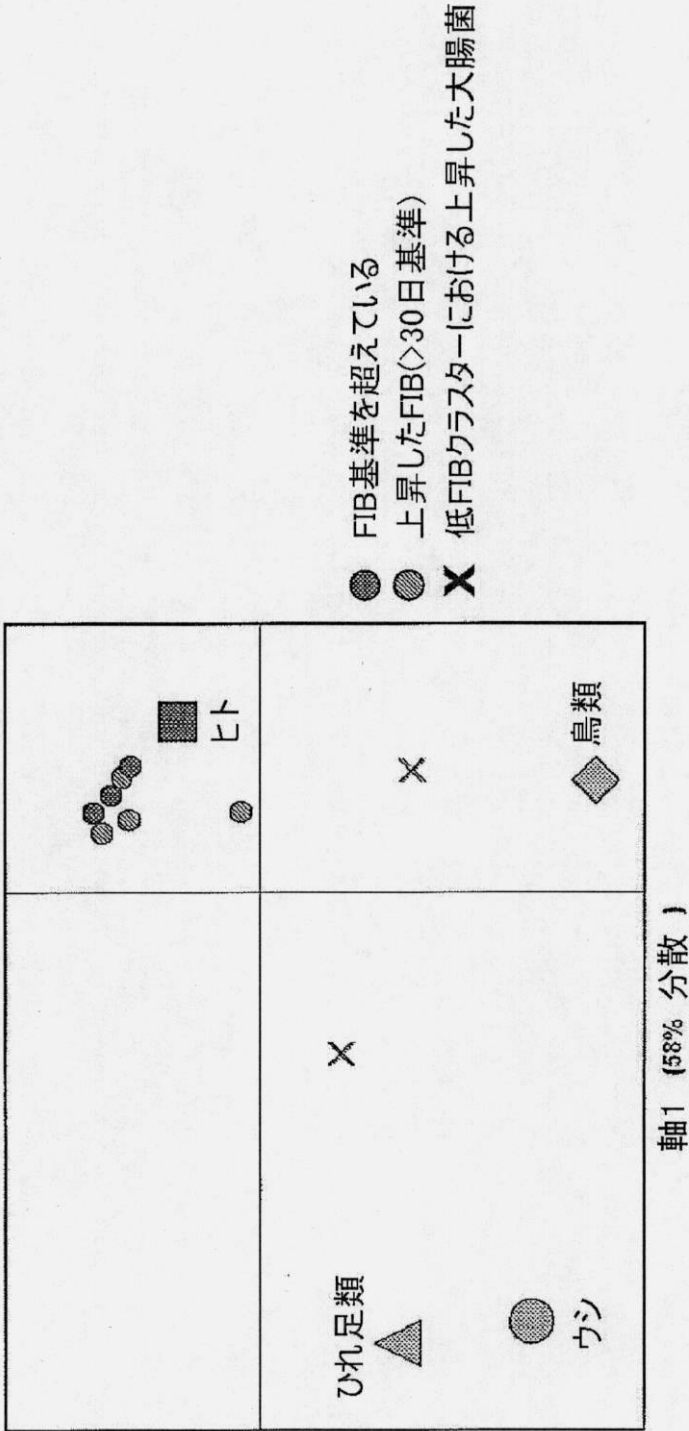
Figure 27



【 図 2 8 】

Figure 28

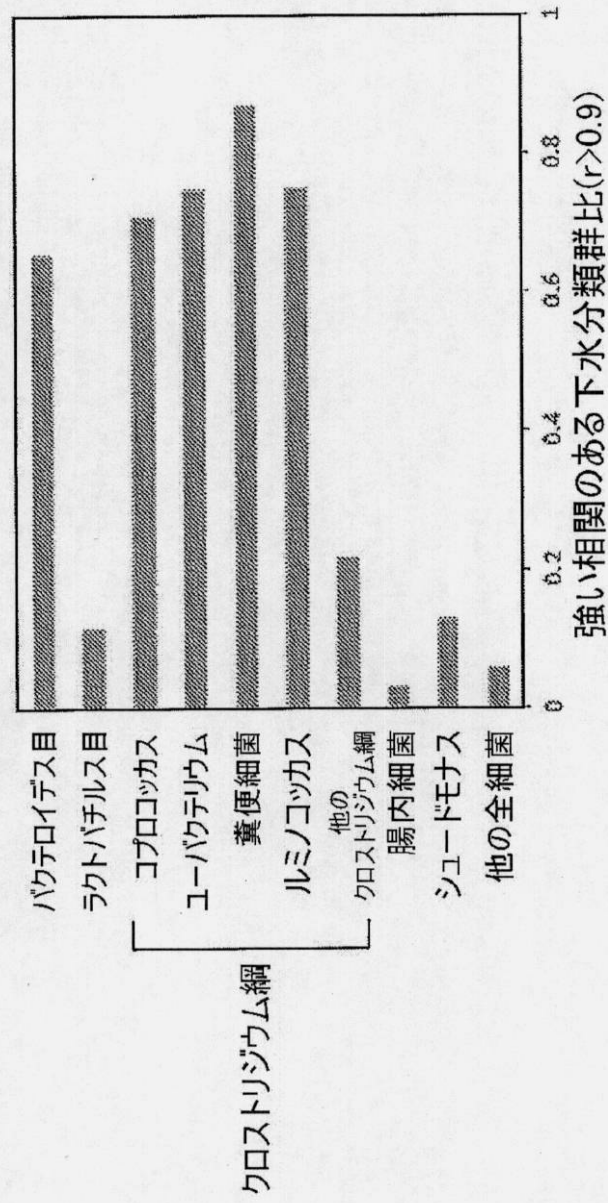
インジケータ群集と各試料由来の上昇したOTUを比較する



【図 29】

Figure 29

湾水中の糞便インジケータと強く相関している下水細菌



【図 3 1】

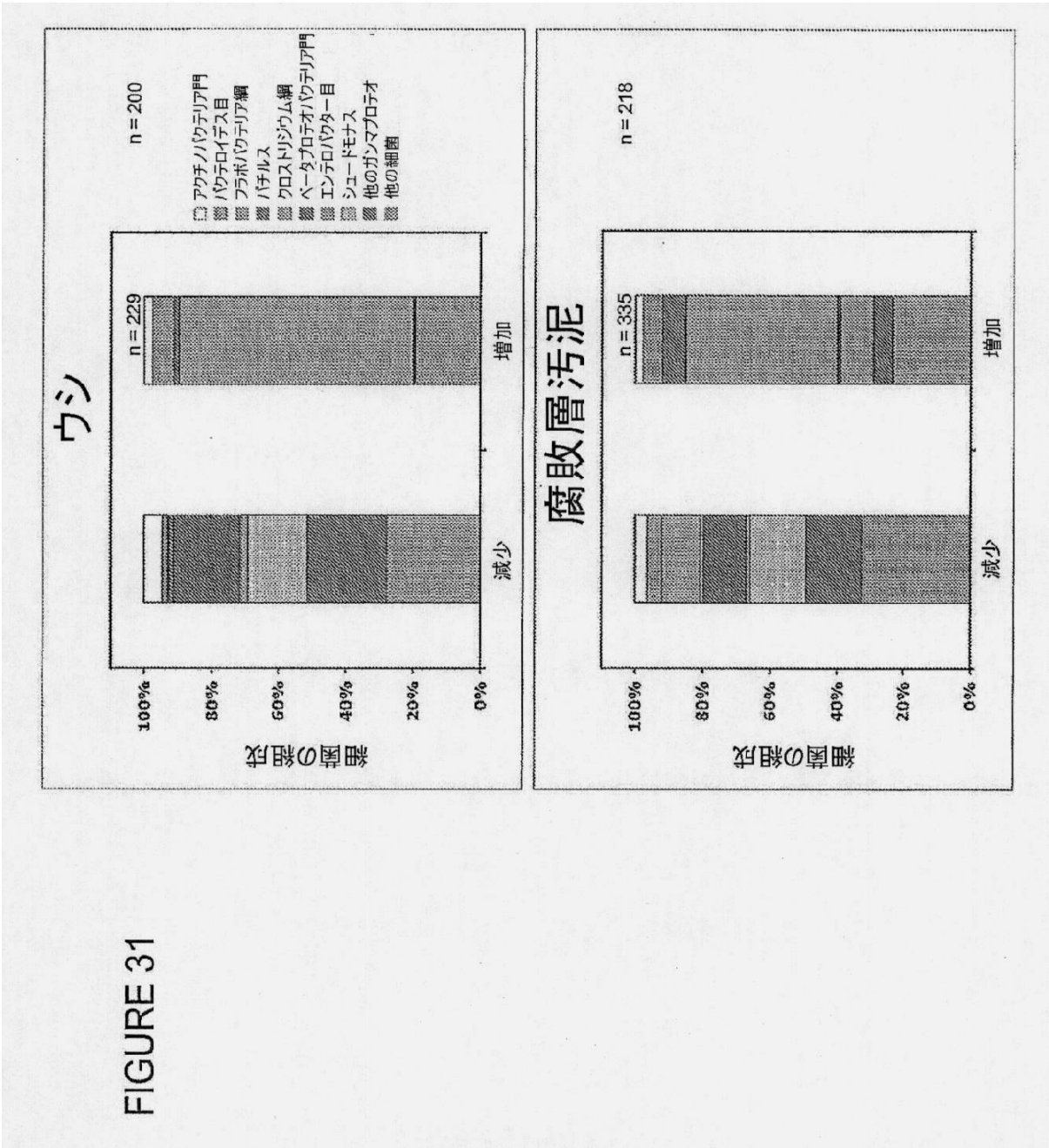
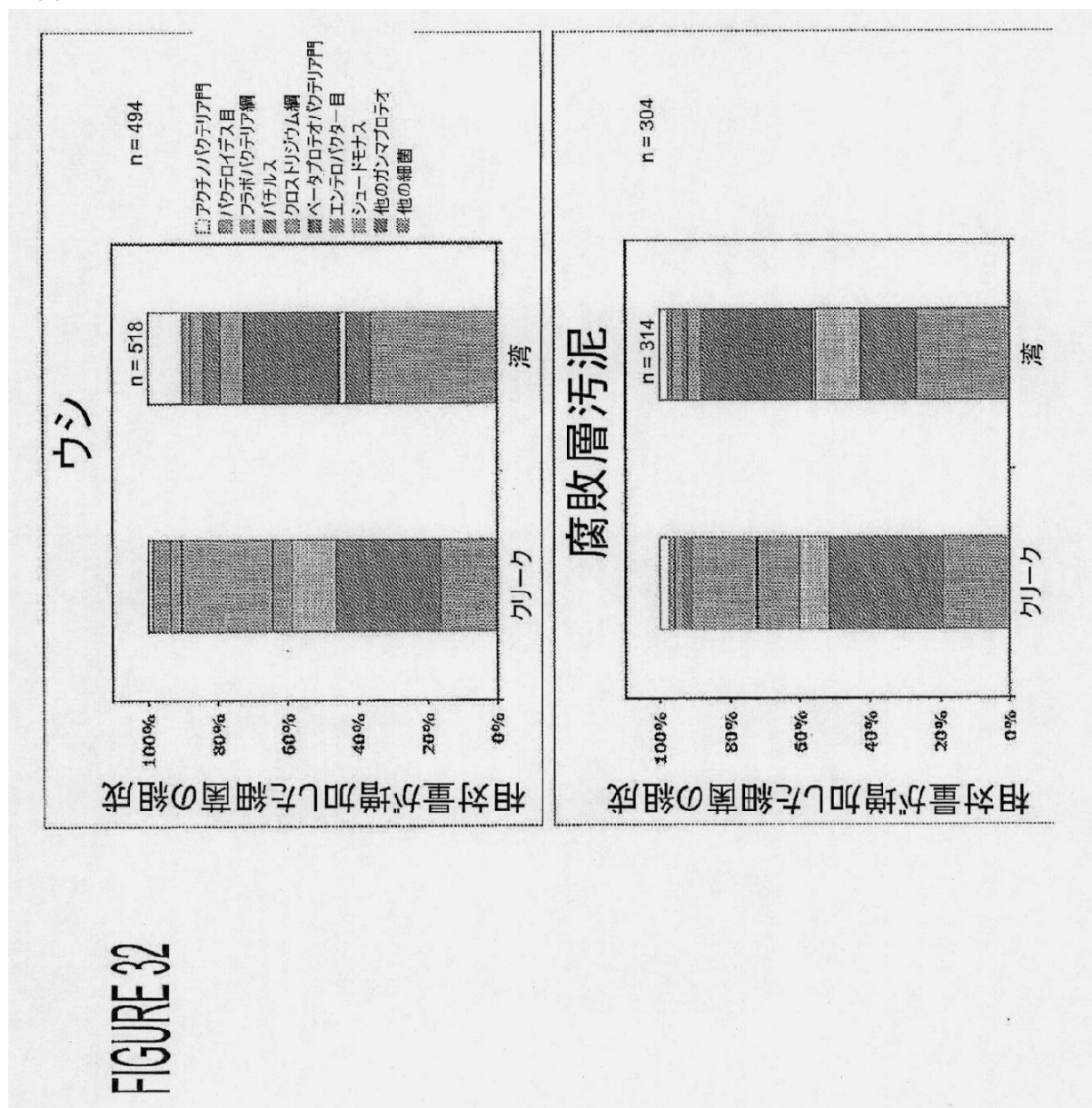
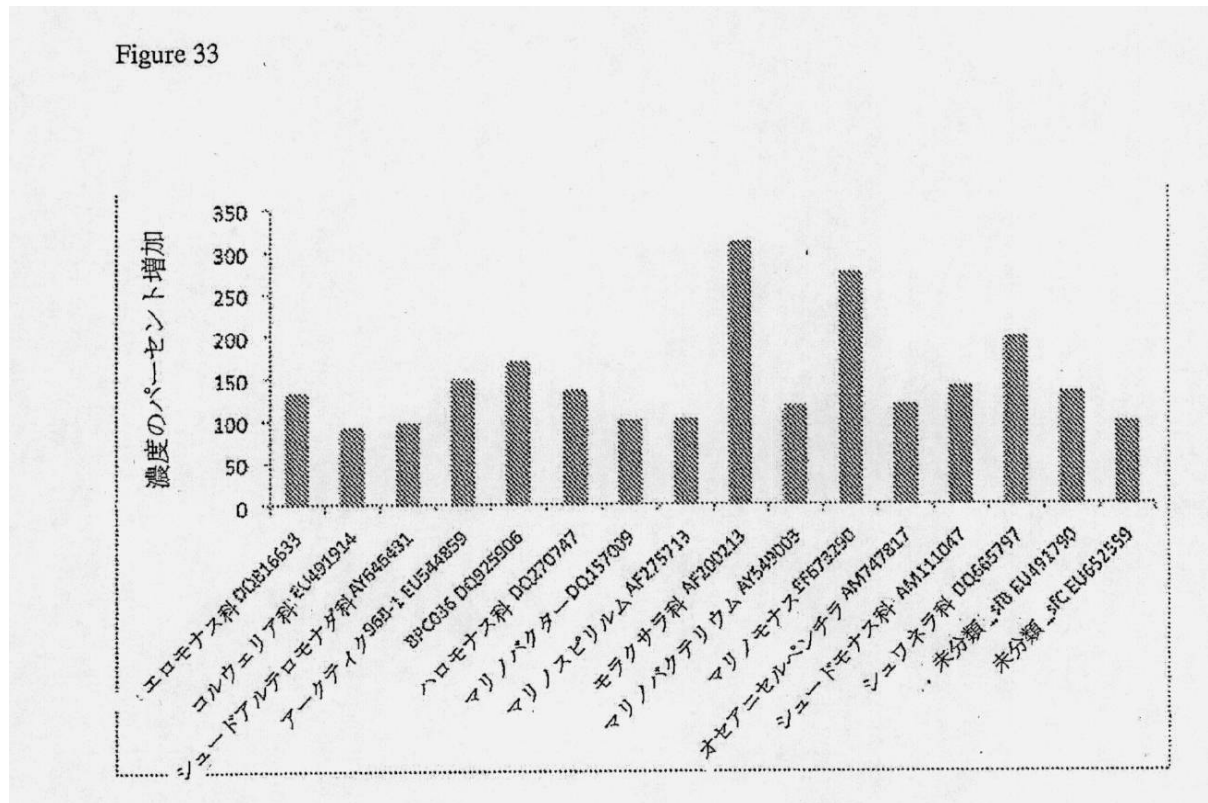


FIGURE 32



【図 33】



【配列表】

0005795579000001.app

フロントページの続き

(51)Int.Cl.		F I		
G 0 1 N 37/00 (2006.01)		G 0 1 N 33/53		M
G 0 1 N 33/569 (2006.01)		G 0 1 N 37/00	1 0 2	
G 0 6 F 19/20 (2011.01)		G 0 1 N 33/569		A
		G 0 1 N 33/569		B
		G 0 6 F 19/20		

(31)優先権主張番号 61/220,937

(32)優先日 平成21年6月26日(2009.6.26)

(33)優先権主張国 米国(US)

(31)優先権主張番号 61/347,817

(32)優先日 平成22年5月24日(2010.5.24)

(33)優先権主張国 米国(US)

(72)発明者 ガーリー・エル・アンデルセン

アメリカ合衆国・カリフォルニア・94708・パークリー・ケーラー・アヴェニュー・780

(72)発明者 トッド・ゼッド・デサンティス

アメリカ合衆国・カリフォルニア・94551・リヴァモア・セイレム・コート・626

(72)発明者 エオイン・プロディー

アメリカ合衆国・カリフォルニア・94114・サン・フランシスコ・トゥエンティフィフス・ストリート・3857

審査官 一宮 里枝

(56)参考文献 特表2005-512527(JP,A)

特表2007-506429(JP,A)

特開2004-357550(JP,A)

特開2007-060953(JP,A)

国際公開第2008/130394(WO,A1)

国際公開第2007/018563(WO,A1)

(58)調査した分野(Int.Cl.,DB名)

C12Q 1/00-3/00

C12N 15/00-15/90

JSTPlus/JMEDPlus/JST7580(JDreamIII)

CAplus/MEDLINE/EMBASE/BIOSIS(STN)