

## 科学研究費助成事業 研究成果報告書

平成 28 年 6 月 22 日現在

機関番号：82708

研究種目：基盤研究(B) (一般)

研究期間：2013～2015

課題番号：25292130

研究課題名(和文) ユーザフレンドリーな新規海洋生物多様性測定システムの開発

研究課題名(英文) Development of a new user-friendly software to analyze marine biodiversity

研究代表者

長井 敏 (Nagai, Satoshi)

国立研究開発法人水産総合研究センター・中央水産研究所・グループ長

研究者番号：80371962

交付決定額(研究期間全体)：(直接経費) 14,500,000円

研究成果の概要(和文)：本研究では、専門知識が無くても扱える、操作者の技量差が結果に反映しない、海洋試料への適用に向けて最適化されたデジタルDNAチップアナライザ(DDCA)を用いた海洋生物多様性測定システムの構築を主目的とした。内容としては、海洋の真核生物を広く検出できるユニバーサルプライマーの開発・選定、標的遺伝子群のシーケンスデータ収集と種名情報付加DNAプローブセットの作成、DDCAで検出された生物種の出現組成を表やグラフでユーザーフレンドリーに表示する機能をプログラムに付加し、汎用性の向上を目指すことで、高精度の海洋生物モニタリングが可能な新規システムの開発を行い、目標を達成した。

研究成果の概要(英文)：The Digital DNA chip analyzer (DDCA) is a computer program that detects target genes or species, utilizing digital (in silico) hybridization process between target probe sets and massively parallel sequencing (MPS)-data. The aim of this study was to develop a conventional, highly sensitive and objective system to analyze marine biodiversity using DDCA with MPS-data. To achieve this goal, we studied the following topics; 1) development of new universal primer pairs in 18S-rRNA to analyze MPS data in DDCA, 2) development of DNA probes in the 18S-rRNA gene and 16S-rRNA genes, i.e. downloads of the target sequences from public genetic databases, selections and editing of the target regions, 3) development of output formats for visualization of the analyzed data in GUI such as tables by excel or graphics drawn in R package.

研究分野：海洋分子生物学

キーワード：デジタルDNAチップアナライザ DDCA 海洋生物 生物多様性 次世代シーケンサー プランクトン 海洋細菌 メタゲノム解析

1. 研究開始当初の背景

或る海洋空間中の微生物多様性(何種類の微生物が存在しているか)を網羅的に把握することは、その環境の物質循環・生物群集構造・赤潮発生確率等の「生物学的ポテンシャル」を理解・予測する上できわめて重要である。各自治体の水産海洋研究機関では1970年代から精力的なプランクトンモニタリングが実施され、約半世紀に亘る動植物プランクトンの出現データが蓄積されてきた。しかし、そのほとんどは光学顕微鏡観察により得られた情報であり、記録に残っているのは「同定可能な優占種」に限定される。実際、わずか1mLの海水であっても、そこに存在する生物を光学顕微鏡で網羅的に検出・同定することは不可能である。観察結果には、作業者の習熟度が大きく影響を及ぼす上、天然海水中には形態に基づく同定が困難なプランクトン種や微小すぎて光学顕微鏡観察では識別できない微生物も多く含まれることがその原因である。

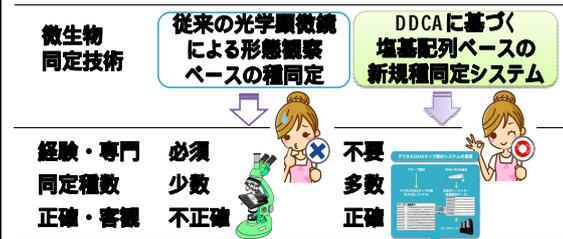
一方、近年の次世代型シーケンサー(NGS)の飛躍的普及により、様々な環境下の生物相を「形態」ではなく「遺伝子断片の塩基配列」で網羅的に測定・評価するメタゲノム解析技術が拡がりつつある。こうした背景の下、沿岸生物モニタリングに、メタゲノム解析技術の早期導入を望む声は非常に多い。そこで本研究では、「取扱が容易な(=専門知識が無くても扱える)」、「エラーの少ない(=操作者の技量差が結果に反映しない)」、「海洋試料への適用に向けて最適化された」デジタルDNAチップアナライザ(DDCA)を用いた海洋生物多様性測定システムの構築(図1)を主目的とする。このシステムの構築により、体長1μm以下の細菌から数ミリまでの動物プランクトンを含む数千種の原核・真核生物群について簡易かつ高精度なメタゲノム解析が可能となる。本研究の実施により、同システムの構築に不可欠なソフトウェアおよび標準手法が構築されれば、海域の生物多様性の効率的な把握にきわめて強力なツールがもたらされるものと期待される。

2. 研究の目的

或る海洋空間中にどれだけの種類の生物が存在しているか(生物学的多様性)を把握することは、その環境の物質循環・生物群集構造等の「生物学的ポテンシャル」を理解する上できわめて重要である。従来の光学顕微鏡観察では、その記録が「同定可能な優占種」に限定される上、観察者の習熟度等により結果が異なるという問題がある。一方、近年の次世代型シーケンサー(NGS)の飛躍的普及により、様々な環境下の微生物相を「形態」ではなく「遺伝子配列」により網羅的に測定するメタゲノム解析技術が拡がりつつある。本研究では、この技術を応用することで、海洋由来試料に適用可能なデジタルDNAチップアナライザ(DDCA)を利用した海洋生物多様

性測定システムの構築を目指す。これにより、体長<1μmの細菌から数mmまでの動物プランクトンを含む数千種以上の原核・真核生物群について、従来法に比べ飛躍的に簡易かつ高精度なモニタリングが可能となるものと期待される(図1)。

図1. 沿岸生物モニタリングの現状と本研究への期待



3. 研究の方法

(1) 海洋生物を網羅検出するためのユニバーサルプライマーの追加開発

海洋細菌の16S-rDNA遺伝子のユニバーサルプライマーは既に報告があり、これを参照したい。18S-rDNAおよび28S-rDNA遺伝子については、既にユニバーサルプライマーを選定済みであるが、海草・海藻類ではリブロービスリン酸カルボキシルーゼ遺伝子(rbcL)、海産無脊椎動物ではミトコンドリア-チトクロームCオキシダーゼ1遺伝子(COX1)等の方がrDNA遺伝子よりも進化速度が速く種同定精度が高いため、関連の文献調査を行いながら、rbcLやCOX1の領域に新たなユニバーサルプライマーを開発する。開発したプライマーセットの性能評価のため、NGS(Roche 454 GS FLX Titanium system)によるメタゲノム解析を行い、カバーできる範囲や精度検証を実施する。

(2) 種名情報を付加したDNAプローブセットの作成

細菌16S-rDNA、真核生物用18S-rDNA、28S-rDNA、rbcLおよびCOX1等のユニバーサルプライマーの決定後、これらの領域を含む配列をGenbank、DDBJ、EMBL等の核酸配列データベースから取得する。なるだけ種同定の漏れが生じないように、網羅的に配列を取得する一方で、重複配列の除去や、sp.などの曖昧な配列やシノニムの配列を精査・厳選して配列データを収集する。DDCAを用いたNGSメタゲノム解析において、>100万個のプローブセットを作成する。ユニバーサルプライマーとしては、細菌検出・種同定用の16S-rDNA<sup>11)</sup>、ほぼ全ての真核生物(菌類、動植物プランクトン、原生動物、大型藻類)検出・種同定用の18S-rDNA、渦鞭毛藻・珪藻等の植物プランクトン検出・種同定用の28S-rDNA<sup>12)</sup>、大型藻類検出・種同定用の葉

緑体 rbcL 遺伝子、動物プランクトン検出・種同定用ミトコンドリア COXI 遺伝子の合計 5 個の領域を標的としたセットを構築する。

### (3) NGS メタゲノムデータを用いた新規システムの精度検証

実際に海水やプランクトンネットサンプルから NGS メタゲノムデータを取得し、Mothur 等で推奨されているメタゲノム解析プラットフォーム<sup>9)</sup>による種同定結果と本提案の新規システムによる種同定結果を比較し、特に DNA プローブ配列の登録に偏りはないか、検出出来ていない生物群はないか、その精度検証を行い、プローブ登録に漏れのある場合、これを実施する。

### (4) プローブセットの自己成長機能システムの開発

Genbank、DDBJ、EMBL 等の核酸配列データベースでは、毎日、登録データの更新が行われている。作成するプローブセットを常に最新の状態にするため、定期的に新規登録配列を取得し、プローブセットを自動更新する為のパイプラインを設計する。パイプライン設計においては、2) の課題にて実施するプローブ作成フローをもとに、定期的な新規登録配列の取得、重複登録の精査、新規プローブの登録、プローブセットの更新を自動で実施するパイプラインを構築する。

### (5) 結果出力用の集計表 + グラフ機能を付加したプログラム開発

DDCA から出力された解析結果について、出現種を取りまとめた表とグラフ作成、サンプル間や海域間の主成分分析や種毎の出現頻度を示すヒートマップ等を出力する機能を持ったプログラムを開発する。DDCA + 自己成長機能 + グラフ機能を完備したソフトウェアの作成は外注する。

## 4. 研究成果

**(1) ユニバーサルプライマーの開発：** これまで開発した 18S-rRNA 遺伝子を標的にしたプライマーに加え、節足動物だけを標的にした 28S-rRNA 遺伝子を増幅するためのプライマー、葉緑体の rbcL 遺伝子についても、より広範囲の藻類に対応したプライマーの開発を行った。

**(2) 標的遺伝子群のシーケンスデータの収集 + 種名情報を付加した DNA プローブセットの作成：** 公的核酸配列データベースから真核生物の 18S-rRNA 遺伝子、細菌の 16S-rRNA 遺伝子として登録されている配列を全て抽出した。各生物群あるいは種毎に登録されている配列数に大きな差があるので、

このバイアスを除去するため、100、98、95、90% でクラスタリングを行い、代表配列を抜き出すこと、必要な領域だけを切り出すこと、一連の工程をパイプライン化することで、DDCA で用いるプローブを、自動的に作成する技術確立に成功した。

**(3) NGS メタゲノムデータを用いたシステムの精度検証：** これまで取得した次世代シーケンスデータを用いて、申請者らが使用するメタゲノム解析手法と DDCA の結果を比較したところ、DDCA で検出された OTU (Operational Taxonomic Units) 数は約 2/3 であり、予定通り、公的データベースに未登録種については、検出されなかった。

八代海 16 サンプル、鹿児島湾 10 サンプル (500mL の海水をろ過紙、全プランクトンを捕集し、DNA 抽出) から 18S-rRNA 遺伝子のユニバーサルプライマーを用いた Roche 454 titanium FLX system により NGS データを得た。DDCA による検出感度について他の手法である核酸クロマトチップや multiplex-PCR と比較したところ、有意な際は見られず、他の手法と遜色のない高感度検出が可能であることを確認した。渦鞭毛藻 *Alexandrium* 属の場合、DNA 量で 0.1-10pg、リボソーム RNA 遺伝子のコピー数で 5-500 コピーあれば検出可能であった。

公的核酸データベースからダウンロードした配列をそのまま DNA プローブとして DDCA に登録し、NGS データに対して DDCA 上でハイブリダイゼーションすると、NGS 数 (5 万-20 万配列) とプローブ数 (18S-rRNA 遺伝子で 3 万個) があまりに多くなり、通常のパソコンでは計算が終了しないことが判明した。このため、NGS 数とプローブ数の関係について検討した結果、ワークステーションクラスのパソコンであれば、NGS 数 (5 万-20 万配列) に対して 1,000-2,000 プローブであれば、1 日程度でハイブリが終了することが判明した。このため、使用する研究者が、自分の研究対象生物を中心とした 100-500 種程度の DNA プローブセットを作成することで、DDCA を用いた効果的な解析ができること期待される。DNA プローブ数の使用制限については、今後、パソコンのスペックの上昇とともに解決されると思われる。

**(4) プローブセットの自己成長機能システムの開発：** 公的核酸配列データベースから必要な配列の抽出方法、クラスタリング、プローブとして用いるのに必要な部分の切り出し等の作業の自動化するプラットフォームを完成させ、必要な時に、標的の遺伝子領域のプローブ配列の自動的な更新を可能とするプログラムを完成させることに成功した。

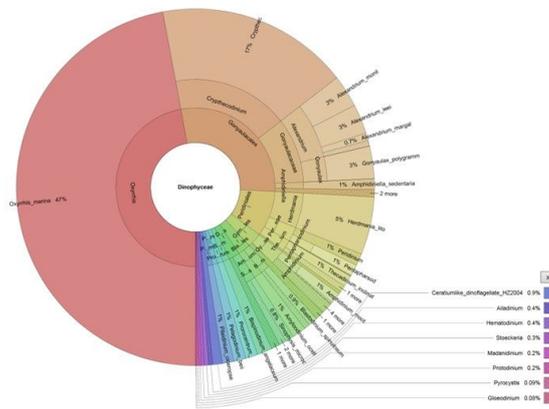


図 2 . 組成円グラフの出力結果

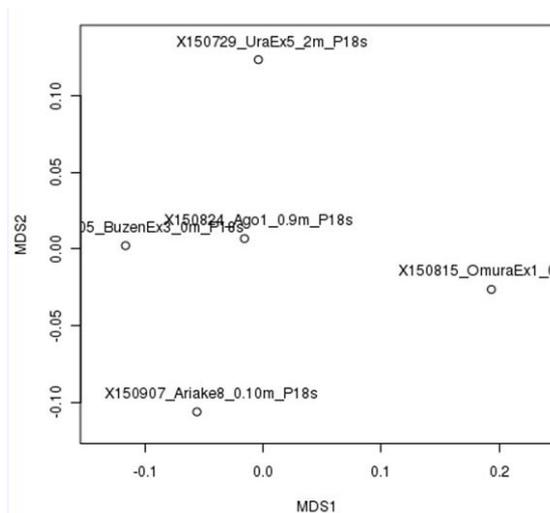


図 3 . 類似度指数を用いた NMDS プロットの出力結果

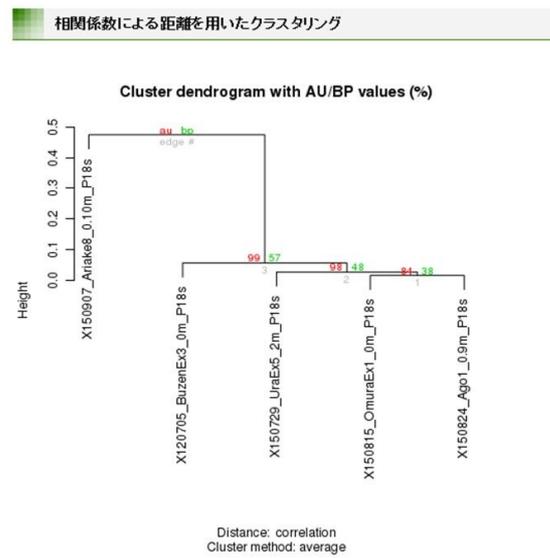


図 4 . 類似度指数によるクラスタリングデンドログラムの出力結果

**(5) 集計表 + グラフ機能を付加したプログラム開発:** DDCA から出力された解析結果について、出現種の情報を取りまとめた集計表としての出力、スーパーグループ、門、綱、

科、属レベルでの組成円グラフ作成 (図 2)、サンプル間や海域間の差を検討するための類似度指数を用いた NMDS (Non-metric multidimensional scaling) 解析の出力 (図 3)、同じく類似度指数を用いたクラスタリングデンドログラムの出力 (図 4)、ヒートマップ・クラスタリングを行うためのグラフィカルインターフェースの作成を行うなど、ソフトウェアを作成することができた。

## 5 . 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文](計 2 件)

Nagai S, Urushizaki S, Yasuike M, Nakamura Y, Fujiwara A, Takano Y, Tanabe AS, Hongo Y, Katakura S. Plankton metagenomics in Okhotsk Sea. The proceedings of the 30<sup>th</sup> international symposium on Okhotsk Sea & Sea Ice, 31-34 (2015). February 17, 2015 Mombetsu City Hall (Hokkaido, Mombetsu city). 査読無

Nagai S, Miyamoto S, Ino K, Tajimi S, Nishi H, Tomono J. Easy detection of multiple *Alexandrium* species using DNA chromatography chip. Harmful Algae 51: 97-106 (2016). 査読有

[学会発表](計 4 件)

陳 虹諺・長井 敏・漆崎慎吾・山本圭吾・及川 寛・安池元重・中村洋路・田邊晶史・本郷悠貴・藤原篤志・岸野洋久、植物プランクトンブルームとその系統的集積性、平成 27 年度日本水産学会春季大会、平成 27 年 4 月 2 日、東京海洋大学 (東京都港区)  
陳 虹諺・長井 敏・漆崎慎吾・山本圭吾・及川 寛・安池元重・中村洋路・田邊晶史・本郷悠貴・藤原篤志・岸野洋久、植物プランクトンブルームとその系統的集積性、2015 年度計量生物学会年会、平成 27 年 3 月 29 日、京都大学 (京都市)

Nagai S, Urushizaki S, Yasuike M, Nakamura Y, Fujiwara A, Takano Y, Tanabe AS, Hongo Y, Katakura S. Plankton metagenomics in Okhotsk Sea. The the 30<sup>th</sup> international symposium on Okhotsk Sea & Sea Ice. February 17, 2015 Mombetsu City Hall (Hokkaido, Mombetsu city)

Nagai S, Miyamoto S, Takahashi H, Tomono J. Easy detection of multiple *Alexandrium* species using DNA chromatography chip. The 16<sup>th</sup> International Conference on Harmful Algae. October 29, 2014. Wellington, New Zealand.

[その他]

ホームページ <http://www.fra.affrc.go.jp/>

## 6 . 研究組織

(1)研究代表者

長井 敏 (NAGAI, Satoshi)  
国立研究開発法人 水産総合研究センター  
中央水産研究所 水産遺伝子解析センター  
メタゲノム研究グループ・グループ長  
研究者番号：80371962

(2)研究分担者

安池 元重 (YASUIKE Motoshige)  
国立研究開発法人 水産総合研究センター  
中央水産研究所 水産遺伝子解析センター  
構造研究グループ・研究員  
研究者番号：20604820

(3)研究分担者

中村 洋路 (NAKAMURA Yoji)  
国立研究開発法人 水産総合研究センター  
中央水産研究所 水産遺伝子解析センター  
構造研究グループ・研究員  
研究者番号：90463182

(4)連携研究者

藤原 篤志 (FUJIWARA Atushi)  
国立研究開発法人 水産総合研究センター  
中央水産研究所 水産遺伝子解析センター  
構造研究グループ・グループ長  
研究者番号：30443352

(5)連携研究者

里見 正隆 (SATOMI Masataka)

国立研究開発法人 水産総合研究センター  
中央水産研究所 水産物応用開発研究センター  
衛生管理グループ・主任研究員  
研究者番号：00344325