

平成21年3月31日現在

研究種目：基盤研究（C）

研究期間：2007～2008

課題番号：19500254

研究課題名（和文） 非計量多次元尺度構成法による生体生命情報解析

研究課題名（英文） Bioinformatics using non-metric multidimensional scaling method

研究代表者

田口 善弘（TAGUCHI YOSHIHIRO）

中央大学・理工学部・教授

研究者番号：30206932

研究成果の概要：

近年、遺伝子に関する網羅的な実験結果が広く公開されるようになってきた。本研究では数千個の遺伝子に対して行われた遺伝子発現、蛋白質発現、塩基配列、アミノ酸配列などのデータを非計量多次元尺度構成法を用いて解析して、遺伝子の機能などを解明する研究を行った。この結果、この様な様々なデータに対して、非計量多次元尺度構成法を用いた研究が非常に有効であることを確認した。

交付額

（金額単位：円）

	直接経費	間接経費	合計
2007年度	1,500,000	450,000	1,950,000
2008年度	500,000	150,000	650,000
年度			
年度			
年度			
総計	2,000,000	600,000	2,600,000

研究分野：総合領域

科研費の分科・細目：情報学・生体生命情報学

キーワード：非計量多次元尺度構成法、ゲノム科学、マイクロアレイ、Barcode of Life、DNA copy number 異状、プロテオミクス、シアノバクテリア、概日周期、

## 1. 研究開始当初の背景

ゲノム科学の進展に伴い、非常に多くの網羅的なデータが公開されるようになってきた。この様な状況で、数千個の遺伝子に対する大規模データを解析する手法の開発が求められてきた。

## 2. 研究の目的

非計量多次元尺度構成法は、本来、社会科学で広く用いられてきた解析手法であるが、これをゲノムサイエンスに用いることにより、unsupervised(先見知なき)な解析手法と

してこれを用いることを目的とした

## 3. 研究の方法

インターネット上で公開されている様々なゲノム科学における網羅的な実験データを収集し、非計量多次元尺度構成法を用いて、遺伝子の機能などを予測したりできるかどうかを試し、どの様な場合に、非計量多次元尺度構成法を用いることができるのかを解明した。

## 4. 研究成果

以下のような諸データに対して、非計量多次

元尺度構成法を適用して、良い結果を得た。Barcode of Life と高次分類群の関係、がん細胞のマイクロアレイによる判別、プロテオミクスとマイクロアレイのメタ解析、DNA copy number 異状のマイクロアレイデータに折る解析、シアノバクテリアの概日周期など

#### (1)Barcode of Life と高次分類群の関係、

Barcode of Life は種分類を形態ではなく、DNA の塩基配列で行おうという画期的なプロジェクトであり、動物についてはDNAのどの部分をBarcode of Life とみなせば、種同定がよくなるかという研究が行われている。今回、われわれは、これを高次分類にも用いることを計画し、塩基配列間の順位を距離として用いることで高次分類群の判別効率が著しく上昇することを見出した。また、順位を距離とした場合の可視化には非計量多次元尺度構成法が適当であるということを見出した。

#### (2)がん細胞のマイクロアレイによる判別、

良性と悪性（転移）の腫瘍の判別は難しく、これをマイクロアレイの結果を持ちいて行おうという研究がなされている。我々は、マイクロアレイによる遺伝子発現プロファイルを非計量多次元尺度構成法で解析してから判別分析を行うことで、従来の遺伝子選択を伴う方法と同等以上の成果を上げることができると示した。

#### (3)プロテオミクスとマイクロアレイのメタ解析、

プロテオミクスとマイクロアレイは、共にコーディング RNA の発現レベルを解析する実験手段であるが、転写と翻訳レベルでの観測であるため、必ずしも一致しない。我々は非計量多次元尺度構成法を用いてからメタ解析を行うことで両者の共通部分を抽出することに成功した。

#### (4)DNA copy number 異状のマイクロアレイデータに折る解析、

DNA copy number 異状は、重要な原因である。本研究ではマイクロアレイのデータから、コピー数異状を起こしている遺伝子を同定することに成功した。Copy number 異状を起こしている遺伝子は、染色体ごとに自動的に非計量多次元尺度構成法による埋め込み空間内でクラスタ化した。

#### (5)シアノバクテリアの概日周期

シアノバクテリアは、概日周期をもつ、

唯一の原核生物である。ルシフェラーゼを用いた遺伝子発現変異体を網羅的に作成した発光実験の結果を非計量多次元尺度構成法で解析し、振動が日周期と夜間の抑制によることを見出した。

#### 5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計10件)

①Fumiaki Kataoka, Y-h. Taguchi, Analysis of DNA copy number alterations with non-metric multidimensional scaling method, IPSJ SIG Technical Report, 2008-BIO-12, pp.55-62 (2008) 査読無

②Y-h. Taguchi, M. Michael Gromiha, Gene Ontology Term Prediction Based upon Amino Acid Occurrence, 2008 International Joint Conference on Neural Networks, pp.616-621 (2008) 査読有

③Masaru Konno, Y-h. Taguchi, Cancer discrimination based upon non-metric multidimensional scaling method, IPSJ SIG Technical Report, 2008-BIO-13, pp.13-16 (2008) 査読無

④Hisamitsu Akiba, Y-h. Taguchi, Markov random field models for protein function prediction, IPSJ SIG Technical Report, 2008-BIO-14, pp.51-54 (2008) 査読無

⑤ Hisamitsu Akiba, Y-h. Taguchi, Nonmetric Distances for Barcode of Life, IPSJ Transactions on Bioinformatics, 1, pp. 35-41, (2008) 査読有

⑥Y-h. Taguchi, M. Michael Gromiha, A Novel Discrimination of GO term Annotated Proteins based on Amino Acid Occurrence and Composition, Pattern Recognition in Bioinformatics, Thirs IAPR Internatiocnal Workshop, PRIB2008, Supplementary Proceedings, (2008) 査読有

⑦ Hisamitsu Akiba, Y-h. Taguchi, Nonmetric distances for Barcode of Life, IPSJ SIG Technical Report, 2007-BIO-9, pp.31-38 (2007) 査読無

⑧ Y-h. Taguchi, A robust measure of correlation between two genes on a microarray using non-metric multidimensional scaling method, IPSJ SIG Technical Report, 2007-BIO-10, pp.1-8

(2007) 査読無

⑨ Y-h. Taguchi, M. Michael Gromiha, Protein Fold Recognition based upon the Amino Acid Occurrence, Pattern Recognition in Bioinformatics, Second IAPR International Workshop, PRIB2007, Proceedings, Springer, 4774, pp.120-131 (2007) 査読有

⑩ Y-h. Taguchi, MM Gromiha, Application of amino acid occurrence for discriminating different folding types of globular protein, BMC Bioinformatics, 8, 404 (2007) 査読有

[学会発表] (計 20 件)

① Masaru Konno, Y-h. Taguchi, Cancer discrimination based upon non-metric multidimensional scaling method, 16th Annual International Conference Intelligent Systems for Molecular Biology, 2008/7/19-23, Tront

② Fumiaki Kataoka, Y-h. Taguchi, Analysis of DNA copy number alterations with non-metric multidimensional scaling method, 16th Annual International Conference Intelligent Systems for Molecular Biology, 2008/7/19-23, Tront.

③ Hisamitsu Akiba, Y-h. Taguchi, Markov random field models for protein function prediction, 16th Annual International Conference Intelligent Systems for Molecular Biology, 2008/7/19-23, Tront

④ Y-h. Taguchi, ESTIMATION OF CELL CYCLE REGULATED GENES USING ANCHOR GENES AND NON-METRIC MULTIDIMENSIONAL SCALING, 16th Annual International Conference Intelligent Systems for Molecular Biology, 2008/7/19-23, Tront.

⑤ 田口善弘, 非計量多次元尺度構成法の生体生命情報学への応用, バイオ アカデミック フォーラム, 2008/7/3, 東京

⑥ Y-h. Taguchi, GO Term Prediction for UniProt Using Amino Acid Occurrence (invited), Pepcon2008: Peptide & Protein Technology: From Concept to Market, 2008/4/22, 深セン

⑦ 廣瀬綾子, 田口善弘, nMDS を使ったマイクロアレイデータとプロテオミクスデータの比較, 第 18 回システムバイオロジー研究

会, 2008/3/14, 東京

⑧ Y-h. Taguchi, Estimation of cell cycle regulated genes using anchor genes and non-metric multidimensional scaling, The Sixth Asia Pacific Bioinformatics Conference, 2008/1/14-17, 京都

⑨ Fumiaki Kataoka, Y-h. Taguchi, Analysis of DNA copy number alterations with non-metric multidimensional scaling method, The Sixth Asia Pacific Bioinformatics Conference, 2008/1/14-17, 京都

⑩ 田口善弘, 非計量多次元尺度構成法の生体生命情報解析への応用, 第 36 回分子生物情報研究会 (SIG-MBI), 2008/1/11, 東京

⑪ Fumiaki Kataoka, Y-h. Taguchi, Analysis of DNA copy number alterations with non-metric multidimensional scaling method, The 2007 Annual Conference of Japanese Society for Bioinformatics, 2007/12/17, 東京

⑫ Y-h. Taguchi, Estimation of cell cycle regulated genes using anchor genes and non-metric multidimensional scaling, The 2007 Annual Conference of Japanese Society for Bioinformatics, 2007/12/17, 東京

⑬ 田口善弘, 非計量多次元尺度構成法とその生体生命情報解析への応用, 統計数理研究所研究会 非線形科学と統計科学の対話, 2007/11/27, 東京

⑭ Hisamitsu Akiba, Y-h. Taguchi, Nonmetric Distance for Barcode of Life, Second IAPR International Workshop, PRIB2007, 2007/10/1-2, Singapore.

⑮ Kazuhisa Makabe, Y-h. Taguchi, Gene Arrangement for Cell Division Cycle Microarray Experiments without Sinusoidal Fittings, Second IAPR International Workshop, PRIB2007, 2007/10/1-2, Singapore.

⑯ 田口善弘, 非計量多次元尺度構成法を用いた生体生命情報解析, バイオ アカデミック フォーラム, 2007/6/21, 東京

⑰ Kazuhisa Makabe, Y-h. Taguchi, Gene arrangement for cell division cycle microarray experiments without sinusoidal

fittings, Eleventh Annual International Conference on Research in Computational Molecurea Biology, 2007/4/21-25, サンフランシスコ

⑱ Tokitaka Oyama, Hiroshi Ito, Takao Kondo, Y-h. Taguchi, Gene expression profile analysis for circadian promoter activities of cyanobacterial bioluminescent reporter strains using non-metric multidimensional scaling, Eleventh Annual International Conference on Research in Computational Molecurea Biology, 2007/4/21-25, サンフランシスコ

⑲ Hisamitsu Akiba, Y-h. Taguchi, Nonmetric distances for Barcode of Life, Eleventh Annual International Conference on Research in Computational Molecurea Biology, 2007/4/21-25, サンフランシスコ

⑳ Y-h. Taguchi, M. Michael Gromiha, Recognition of protein folds using amino acid occurrence, Eleventh Annual International Conference on Research in Computational Molecurea Biology, 2007/4/21-25, サンフランシスコ

[図書] (計 3 件)

① T. Oyama, T. Kondo, The Kai Oscillator In "Bacterial circadian clock and rhythms book". Edited by Johnson CH, Mackey SR, Ditty JL. Springer Publishing New York (2009 in press) 査読無

② 田口善弘、日本評論社、高校で教わりたかった物理、(2009)、170ページ

③ 田口善弘、岩波書店、文庫で読む科学、(2007) pp.93-104

[その他] (計 3 件)

① 田口善弘、非計量多次元尺度構成法とその生体生命情報解析への応用、物性研究, 91, pp. 137-143 (2008)

② 伊庭幸人、末谷大道、青柳富誌生、田口善

弘、小松崎民樹、研究会報告「非線型科学と統計科学の対話」、91, pp.113-174, (2008)

③ 田口善弘、フロ桶はいついっばいになる？、数学セミナー、9月号、pp.44-47, (2007)

## 6. 研究組織

### (1) 研究代表者

田口 善弘 (TAGUCHI YOSHIHIRO)

中央大学・理工学部・教授

研究者番号：30206932

### (2) 研究分担者

小山 時隆 (OYAMA TOKITAKA)

京都大学・大学院理学研究科・准教授

研究者番号：30324396

### (3) 連携研究者