

平成 22年 5月 20日現在

研究種目：若手研究（B）  
 研究期間：2007～2009  
 課題番号：19710166  
 研究課題名（和文） CAGE 絶対値遺伝子発現プロファイルによるヒトとマウスの組織特異性の網羅的解析  
 研究課題名（英文） Tissue Specific Analysis of Human and Mouse using CAGE absolute gene expression profiles  
 研究代表者  
 竹中 要一（Takenaka Yoichi）  
 大阪大学・大学院情報科学研究科・准教授  
 研究者番号：00324830

## 研究成果の概要（和文）：

タンパク質の設計図である遺伝子は、染色体という形で全ての細胞に存在します。その設計図である遺伝子は、心臓と目といった組織ごとに異なります。これが遺伝子の量に基づくその組織の特徴です。本研究では組織ごとに遺伝子の利用された回数を測定可能な技術であるCAGEを用い、ヒトとマウスにおける組織特性、および組織特異性を調べる方法について研究しました。その結果、情報量という評価尺度を用いることが有効なことを明らかにしました。

## 研究成果の概要（英文）：

Genes are known as the blueprints of proteins. They exist in all the cell of an individual. The amount of genes the cell used in is different among different tissues. This is the tissue-specificity in the view point of gene expression. The number of expressions or mRNAs can be measured by the CAGE technology. I studied the way how to compare the tissue-specificity among tissues and found information-content is a good measure.

## 交付決定額

（金額単位：円）

	直接経費	間接経費	合計
2007年度	1,000,000	0	1,000,000
2008年度	500,000	150,000	650,000
2009年度	700,000	210,000	910,000
年度			
年度			
総計	2,200,000	360,000	2,560,000

研究分野：複合新領域

科研費の分科・細目：ゲノム科学・ゲノム情報科学

キーワード：CAGE, バイオインフォマティクス, 発現プロファイル, 組織特異性

## 1. 研究開始当初の背景

ゲノム計画も一段落つき、ゲノム配列自身から新たな知見を得る静的なゲノム解析から、遺伝子やタンパク質の発現量に基づく動的なゲノム解析へと進展していた。これは

DNA マイクロアレイ, Gene チップをはじめとする遺伝子発現量計測技術の進歩により、遺伝子発現量の網羅的な観測が安価に可能となったためである。この観測結果を用いて応募者はコンピュータによる各種解析手法

の提案をおこなってきた。DNA マイクロアレイや Gene チップは非常に強力な技法として、多数の実験条件間における遺伝子発現量の違いの網羅的計測に用いられていた。しかしながら応募者は、ある一つの実験条件における2つの遺伝子の発現量比較が行えないこと、すなわち遺伝子 A と遺伝子 B のいずれが多く発現しているのかを知ることができないことを不満に感じていた。2 遺伝子間の発現量比較が可能になれば、新たな知見が多く得られると考えられたからである。

この不満は、応募者が参加していたマウスの完全長 cDNA を大量に取得するプロジェクト(通称 FANTOM プロジェクト)で確立し観測された CAGE(Cap Analysis Gene Expression)と呼ばれる遺伝子発現計測手法で解消される運びとなった。CAGE とはある実験条件において発現している mRNA の 5' 端より切り出した約 20 塩基(CAGE タグと呼ぶ)を網羅的にシーケンスする方法であり、遺伝子に対応する CAGE タグの個数を数えることによって発現量を計測することが可能となる。そのため、同一実験条件下において2 遺伝子の発現量の違いを測定することが可能となる。以降ではこの性質を有するデータを絶対値発現プロファイルと呼ぶ。当時、ヒトとマウスに対する CAGE 絶対値発現プロファイルが公開されていた。そこで本研究では、絶対値発現プロファイルの特長を活かしてヒトゲノムとマウスゲノムの動的な解析を行うことを目的として設定した。

## 2. 研究の目的

本研究では、絶対値発現プロファイルの特長を活かしてヒトゲノムとマウスゲノムの動的な解析を行うことを目的とする。絶対値発現プロファイルの特長、言い換えると絶対値発現プロファイルでしか行えないことは以下の2点である。

- A. 1 実験条件における 2 遺伝子の発現量比較が可能であり、発現量の加算が可能である。
- B. 転写開始点ごとの発現量比較が可能である。

特長 A は遺伝子の発現量が CAGE タグの個数で表現される事に起因する特長であるが、これは複数遺伝子の発現量を加算することが可能であることも同時に意味している。この特長を利用することで、「ある実験条件(ex. 成体の肝臓)におけるタンパク質の分子機能(ex. 酵素や転写調節)の発現量が全発現量に占める割合」のような動的なゲノム解析を行うことができる。現在、遺伝子のアノテーションとして GO(Gene Ontology)を利用することが多いが、大まかな機能しか判らない遺伝子と詳細な機能が判明している遺伝子とが混在している状態である。これは GO の階層において上位概念、下位概念の関係(例

えば「酵素」と「酸化還元酵素」の関係)にある。下位概念は上位概念に包含されるため、遺伝子機能を GO - Slim のような一定の上位概念に束ねての静的なゲノム解析が多く行われている。遺伝子発現量の加算が可能になることにより、DNA マイクロアレイなどでは不可能であった上位概念に束ねての動的なゲノム解析が可能になる。

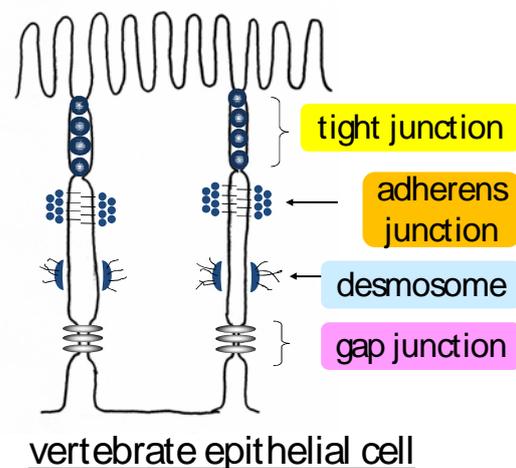
次に特長 B について述べる。CAGE タグにおける 20 塩基長の配列は遺伝子の転写開始点を表しており、その転写開始点ごとに発現量、すなわち CAGE タグの個数を数えることができる。そのため転写開始点を複数個もつ遺伝子の転写開始点ごとの発現量を測定することが可能である。複数の転写開始点はアルタナティブプロモーターに起因すると考えられているため、CAGE によってアルタナティブプロモーターに関する多くの知見が得られることが期待されている。

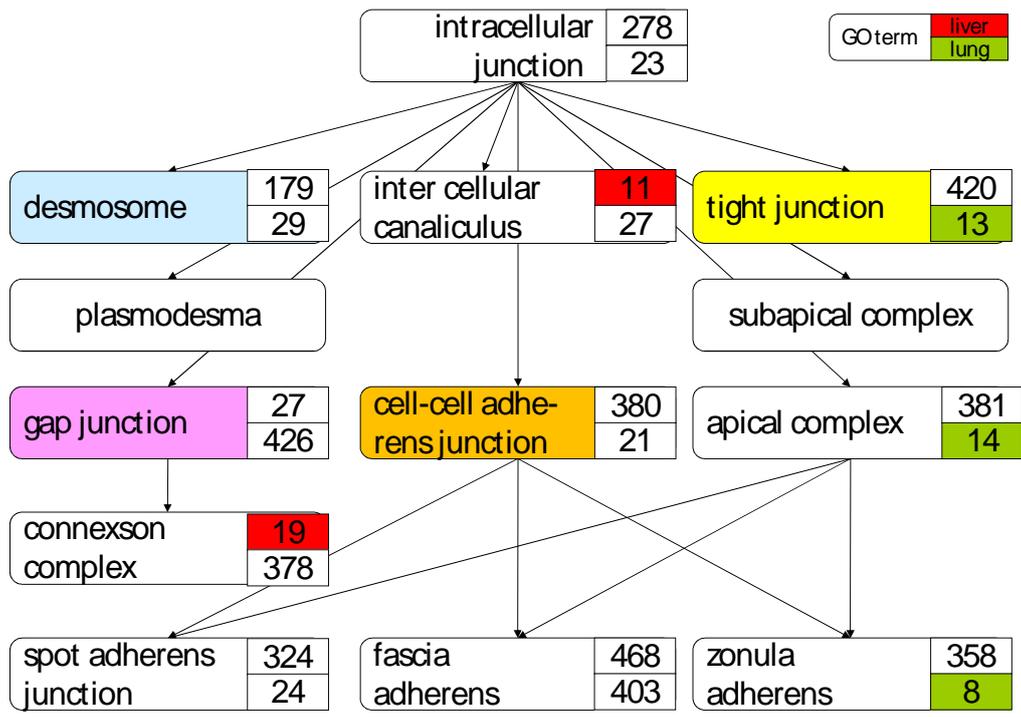
## 3. 研究の方法

遺伝子の役割、機能について記述するための語彙体系として GO(Gene Ontology)がある。GO は遺伝子の機能を自然言語でなく計算機で容易に扱うことができる単一根非循環有効グラフで表している。本研究では、GO のノードごとに各組織の発現量を収集する。収集にあたっては、加算可能である CAGE の性質を利用し、下位概念の GO-term の遺伝子発現量を上位概念の GO-term に加算した。組織ごとに集計した各 GO-term の発現量に対して、各組織における情報量を計算する。計算されたある組織における GO-term の情報量を、比較した組織の情報量から引くことにより、対象 GO-term の組織特異性の尺度とした。この比較をマウスの各組織間で比較した(学会発表[3,4])。またある組織におけるヒトとマウスの違いを計測した(雑誌論文[3]等)。

## 4. 研究成果

- (1) マウスにおける細胞の小構造(GO における Cellular Component)より、組織特性が顕





著である例を記載する。肝臓と肺における細胞間結合の組織特異性

前頁の図に上皮細胞における細胞間結合の模式図を記す。図では、Tight Junction, Adherens Junction, Desomose, Gap Junctionの位置を指している。これらの場所および、関連する GO-term について,肝臓と肺における特異性を示した図を本頁上部に記す。図のノードは GO-term を、ノード右上の数字が肝臓の GO-term の特異度を、ノード右下の数字が肺の特異度を記している。この数字は、各組織における全 GO-term の特異度を順位化したものである。値として自然数をとり、数値(順位)が小さいほど特異的である。

本結果より、肺において tight junction をターゲットとするタンパク質が特異的に働いていることがわかる。これは、肺がその上皮組織において外界と直接接しており、毒性物質の侵入を阻止するための機能を有していることを密接に関係することが示唆される。

(1) ヒトとマウス間で組織比較を行った際にヒト特徴的、マウス特徴的だとみられる遺伝子のリストを記載する。

ヒト肝臓特異的 GO-term

GO-term
Calcium channel inhibitor activity
Plasmin inhibitor activity
Calcium oxalate binding
IgA binding
Lysine carboxypeptidase activity
Trypsin inhibitor activity
Cymotrypsin inhibitor activity
(S)-limonene 7-monooxygenase activity

Water binding
Sterol 12-alpha-hydroxylase activity

マウス肝臓特異的 GO-term

GO-term
Urocanate hydratase activity
Retinal dehydrogenase activity
Endogenous peptide antigen binding
Ornithine carbamoyltransferase activity
Beta-ureidopropionase activity
Butyrate-CoA ligase activity
Acyl-CoA ligase activity
Alcohol sulfontransferase activity
Urate Oxidase activity
Alkane 1-monooxygenase activity

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7件)

[1] Harukazu Suzuki, Alistair R R Forrest, Erik van Nimwegen, Carsten O Daub, Piotr J Balwierz, Katharine M Irvine, Timo Lassmann, Timothy Ravasi, Yuki Hasegawa, Michiel J L de Hoon, Shintaro Katayama, Kate Schroder, Piero Carninci, Yasuhiro Tomaru, Mutsumi Kanamori-Katayama, Atsutaka Kubosaki, Altuna Akalin, Yoshinari Ando, Erik Arner, Maki Asada, Hiroshi Asahara, Timothy Bailey, Vladimir B Bajic, Denis Bauer, Anthony G Beckhouse,

Nicolas Bertin, Johan Bjoerkegren, Frank Brombacher, Erika Bulger, Alistair M Chalk, Joe Chiba, Nicole Cloonan, Adam Dawe, Josee Dostie, Paer G Engstroem, Magbubah Essack, Geoffrey J Faulkner, J Lynn Fink, David Fredman, Ko Fujimori, Masaaki Furuno, Takashi Gojobori, Julian Gough, Sean M Grimmond, Mika Gustafsson, Megumi Hashimoto, Takehiro Hashimoto, Mariko Hatakeyama, Susanne Heinze, Winston Hide, Oliver Hofmann, Michael Hoernquist, Lukasz Huminieck, Kazuho Ikeo, Naoko Imamoto, Satoshi Inoue, Yusuke Inoue, Ryoko Ishihara, Takao Iwayanagi, Anders Jacobsen, Mandeep Kaur, Hideya Kawaji, Markus C Kerr<sup>1</sup>, Ryuichiro Kimura, Syuhei Kimura, Yasumasa Kimura, Hiroaki Kitano, Hisashi Koga, Toshio Kojima, Shinji Kondo, Takeshi Konno, Anders Krogh, Adele Kruger, Ajit Kumar, Boris Lenhard, Andreas Lennartsson, Morten Lindow, Marina Lizio, Cameron MacPherson, Norihiro Maeda, Christopher A Maher, Monique Maqungo, Jessica Mar, Nicholas A Matigian, Hideo Matsuda, John S Mattick, Stuart Meier, Sei Miyamoto, Etsuko Miyamoto-Sato, Kazuhiko Nakabayashi, Yutaka Nakachi, Mika Nakano, Sanne Nygaard, Toshitsugu Okayama, Yasushi Okazaki, Haruka Okuda-Yabukami, Valerio Orlando, Jun Otomo, Mikhail Pachkov, Nikolai Petrovsky, Charles Plessy, John Quackenbush, Aleksandar Radovanovic, Michael Rehli, Rintaro Saito, Albin Sandelin, Sebastian Schmeier, Christian Schoenbach, Ariel S Schwartz, Colin A Semple, Miho Sera, Jessica Severin, Katsuhiko Shirahige, Cas Simons, George St Laurent, Masanori Suzuki, Takahiro Suzuki, Matthew J Sweet, Ryan J Taft, Shizu Takeda, Yoichi Takenaka, Kai Tan, Martin S Taylor, Rohan D Teasdale, Jesper Tegne'r, Sarah Teichmann, Eivind Valen, Claes Wahlestedt, Kazunori Waki, Andrew Waterhouse, Christine A Wells, Ole Winther, Linda Wu, Kazumi Yamaguchi, Hiroshi Yanagawa, Jun Yasuda, Mihaela Zavolan, David A Hume, Takahiro Arakawa, Shiro Fukuda, Kengo Imamura, Chikatoshi Kai, Ai Kaiho, Tsugumi Kawashima, Chika Kawazu, Yayoi Kitazume, Miki Kojima, Hisashi Miura, Kayoko Murakami, Mitsuyoshi Murata, Noriko Ninomiya, Hiromi Nishiyori, Shohei Noma, Chihiro Ogawa, Takuma Sano, Christophe Simon, Michihira Tagami, Yukari Takahashi, Jun Kawai, Yoshihide Hayashizaki:  
The transcriptional network that controls growth arrest and differentiation in a

human myeloid leukemia cell line, Nature Genetics, Vol.41, No.5, pp.553-562 (May, 2009) ,査読有.

[2] Yoshiyuki Kido, Shigeto Seno, Susumu Date, Yoichi Takenaka, Hideo Matsuda: A Distributed-Processing System for Accelerating Biological Research using Data-Staging, IPSJ Transactions on Bioinformatics, Vol.49, No.SIG5(TBI04), pp.58-64 (March, 2008) ,査読有.

[3] Sami Maekawa, Atsuko Matsumoto, Yoichi Takenaka, Hideo Matsuda: Tissue-specific functions based on information content of gene ontology using cap analysis gene expression, Medical and Biological Engineering and Computing, Vol.45, No.11, pp.1029-1036 (November, 2007),査読有.

[学会発表](計 25 件)

[1] Shigeto Seno, Yoichi Takenaka, Hideo Matsuda, A Method for Analyzing Gene Expression Profiles based on the Underlying Structures, Proceedings of 17th Annual International Conference on Intelligent Systems for Molecular Biology and 8th European Conference on Computational Biology (ISMB/ECCB2009), M21, Stockholm, Sweden (June 29, 2009).

[2] Mitsuru Jikeya, Shigeto Seno, Yoichi Takenaka, Hideo Matsuda: A Method for Analysis of Tissue-Specific Transcriptional Control Using Distributions of Transcription Starting Sites, Proceedings of the 2008 Annual Conference of the Japanese Society for Bioinformatics (JSBi2008), P044, Senri-Chuo (December 15, 2008).

[3] Yoichi Takenaka, Hideo Matsuda, Tissue specificity measured by differences in information content between target tissue and whole tissue, Proceedings of 16th International Conference on Intelligent Systems for Molecular Biology (ISMB2008), A21, Toronto (July 21, 2008).

[4] Yoichi Takenaka, Akiko Matsumoto, Hideo Matsuda, GO based Tissue Specific Functions of Mouse using Countable Gene Expression Profiles, Genome Informatics, Vol.19, pp.154-165, Singapore (December 4, 2007).

6 . 研究組織

(1)研究代表者

竹中 要一 (Takenaka Yoichi )

大阪大学・大学院情報科学研究科・准教授

研究者番号：00324830