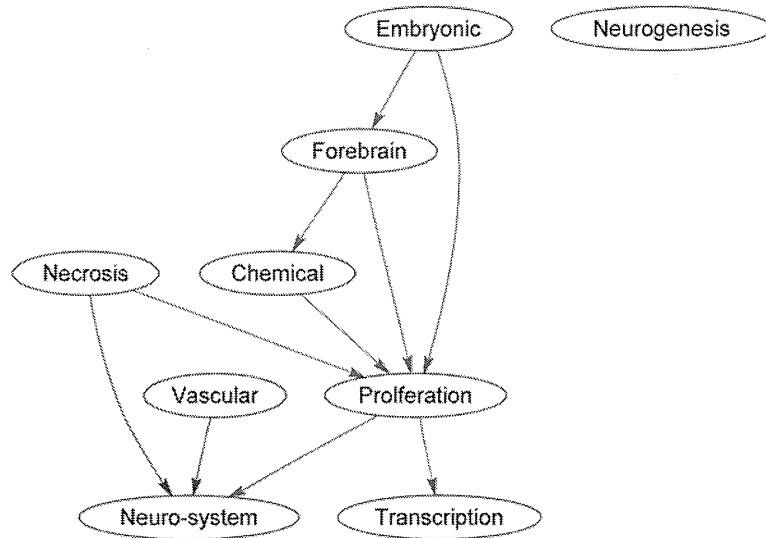


EB_TMD



NS_TMD

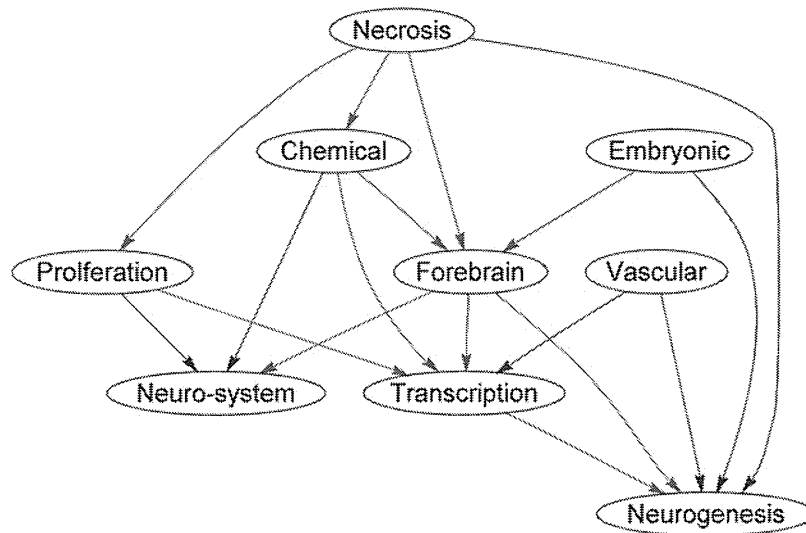


図12. ベイジアンネットワーク解析結果

ES 細胞試験における遺伝子ネットワークを入力に用いた 未知毒性化合物のサポートベクターマシン毒性解析

研究分担者

藤 淵 航

独立行政法人産業技術総合研究所 研究チーム長

研究要旨

限られた数の既知の毒性化合物を用いた ES 細胞試験から、未知の化合物の生体への影響を高性能に予測する新しい一連の手法を開発した。特に、経験的なマーカー遺伝子を全く用いず、10 種の毒性既知の化合物のマイクロアレイの結果から多変量解析の手法を用いて 10 のマーカー候補遺伝子を選択した。その 10 遺伝子で 20 種の毒性化合物の予測を行った結果、神経毒、発癌毒、変異原性のどのカテゴリーにおいても 90%以上の高性能な予測率が得られた。さらに、ベイジアンネットワークを推定し、入力ベクトルに加えることで予測率が向上する結果が得られた。本手法の成功は、ES 細胞を用いた発達毒性試験の精度向上に確率推論（ベイジアンネットワーク解析）の利用が有効であることを示している。

A. 研究目的

本研究では生体への影響が類型化された毒性既知の化学物質をヒト ES 細胞に曝露し、これにより攪乱される遺伝子発現量を測定した時系列データを作成する。このデータを、数理科学手法を用いて機械学習し、未知の化合物に対する毒性評価を行うためのインフォマティクス技術を開発する。

B. 研究方法

1) マイクロアレイからのマーカー遺伝子の探索

10 種の化合物の ES 細胞への曝露実験によるマイクロアレイデータ（発達神経毒性 5 種：Valproic acid, Cyclopamine, Phenytoin, Methylmercury, Acrylamide、変異原性 5 種：Benzo [a] anthracene, 3-Methylcholanthrene, Benzo [a] pyrene, Diethylnitrosamine (DEN, 55-18-5), Diethylstilbestrol）から、多変量統計解析の手法の 1 つである正準相関解析（または対応分析と呼ばれる）を用いて、毒性カテゴリー名は情報として加えず（教師なし）に自動的に相関最大化に基づいて、各化合物（場合によっては化合物群）と相関のある遺伝子を抽出した。正準相関解析とは、カーネ

ルを使用した正則化非線形正準相関の式で書けば、

図 1 の様に表される。今回、カーネル部分は単純な線形のまま、通常の前則化を用いない正準相関解析を行った。その場合には単純な固有値問題で解くことが可能である。すなわち、元の行列の共分散行列を S_{aa} , S_{ab} , S_{bb} とすると、

$$(S_{ab}S_{bb}^{-1}S_{ba} - \lambda^2 S_{aa})a=0$$

または、

$$(S_{ba}S_{aa}^{-1}S_{ab} - \lambda^2 S_{bb})b=0$$

の固有値 λ を求めることで線形変換係数 a , b が求められる。

2) 遺伝子ネットワークの予測率への影響解析

確率的な遺伝子ネットワークモデルとして有名なベイジアンネットワークを推定し、このネットワークの遺伝子結合パラメーターである β 値のセットを、遺伝子発現量に加えて入力ベクトルにしたサポートベクターマシンを構築した。

今回、使用したベイジアンネットワークは、国立環境研究所で開発された TAOgen をベースとして、従来の問題

点である、不安定解の解消のために、レプリカ交換法を用いた RX-TAOgen を開発した(図2)。さらに、今回は 10 遺伝子と少数のため使用する必要がなかったが、もう一つの問題点である大規模遺伝子ネットワークの計算にも対応するため、産総研で開発済みの高速遺伝子モジュール探索ソフトウェア SAMURAI(System for Assembling Modules by Ultra Rapid Algorithm on Itemsets)と連結させることで、初期値を従来よりも最適解に近いところにとること事後確率の収束を速くすることに成功した(CBRC2010 ポスター発表)。

ベイジアンネットワークの推定だけでは、本当に遺伝子ネットワークが情報を加えているのか不明である。予測率の変動が有為であるかどうかを判定するには、i) ランダムなネットワークを構成した場合の予測率の平均と分散を用いて統計評価する、ii) 異なる手法による遺伝子ネットワークでも予測率が増減するか、を試す必要がある。i) に基づく予備実験をやったところ、ランダムなネットワークを加えても予測率の上昇は 1%未満であった。また、手法的に特徴ベクトルの選択には t 検定を用いているため、情報の高くない β 値は使用されない。このため、予測率がそれほど悪くはなる事はないと考えられる。さらに、ii) の検証のため、ここでは、偏相関係数を求める GGM(グラフィカルガウシアンモデル)と、時間差を入れて相関係数を計算する CrossR(交差相関係数)を用いた(図3)。両ネットワークとも遺伝子間(偏)相関係数の絶対値での閾値 0.8 と比較して高ければ 1、そうでなければ、0 のネットワークを求めた。求めた値は、ベイジアンネットワークと同様、元の遺伝子発現量に加えて入力ベクトルにしたサポートベクターマシンを構築した。

3) 探索したマーカー遺伝子での 20 化合物毒性予測

上述の正準相関解析を元に得られた 10 マーカー遺伝子を用いた 20 化合物(発達神経毒性 9 種: 上記 5 種に加え 4-hydroxy PCB107, 2,5-hexanedione, Warfarin, Thalidomide、変異原性 5 種: 上記、発癌性 6 種: TCDD, Lithocholic acid (434-13-9), Thioacetamide (62-55-5), Butylated hydroxyanisole (25013-16-5), Methapyrilene hydrochloride (135-23-9), Phenobarbital) による大規模

RT-PCR 実験を行った(東京大学、国立環境研究所)。RT-PCR 実験では、各化合物で濃度を基準値の 1/1、1/2、1/4、1/8、1/16 倍の 5 段階、時間を暴露後 24、48、72、96 時間の 4 段階、これを 2 リピートした計 40 点を計測した。計測した値は同時点の b-Actin の値で除算して補正した。これを産業技術総合研究所で上記 10 遺伝子による遺伝子ネットワークを推定し、SVM による 20 化合物の毒性予測を行った。10 遺伝子のネットワークを推定するのに 40 データ点を利用しており、データ点が予測する遺伝子数の 4 倍になっている。

ベイジアンネットワークの予測には細かなパラメーター設定が必要であり、これを簡便に行えるように、図4のような WWW ページを作成し、容易にクラスターマシンで並列処理を行う CGI スクリプトをサーバー側で走らせて研究の効率化を図った。

サポートベクターマシンは、重み線形判別関数を使用して、判別面とその直近のベクトル(サポートベクトル)とのマージンを最大化する手法を用いた判別解析法である。このマージン最大化を行う技術としてラグランジュの未定乗数法を用いるが、この解には元のベクトル間で総当たりの内積を計算することになる。この内積項を取る前に様々なマッピングを行わせることが可能で、マッピング後の内積関数をカーネルと呼んでいる。このカーネルには、非線形マッピング関数を含むことが可能であり、線形のサポートベクターマシンの内部に非線形マッピングをすることで非常に強力な判別能力を有することとなる(図5)。

(倫理面への配慮)

産総研ではヒト由来試料倫理委員会を通してその倫理面の審査を行っている。共同研究者から得られた遺伝子発現データは全てこの倫理委員会で承認されたものであり、倫理面での問題はない。

C. 研究結果

1) マイクロアレイからのマーカー遺伝子の探索

10 種のマイクロアレイデータ(発達神経毒性 5 種、変異原性 5 種)から、多変量統計解析の手法の 1 つである正準相関解析(または対応分析と呼ばれる)を用いて、毒性

カテゴリ名は情報として加えず（教師なし）に自動的に
相関最大化に基づいて、各化合物（場合によっては化合物
群）と相関のある遺伝子を抽出した。これにより、固有
値空間に毒性化合物と遺伝子の関係が視覚化できた
（図6）。また、各固有値において寄与率の高い遺伝子を
1～数個ずつ計 10 遺伝子を毒性マーカー候補として選
択した。

今回のデータは、数万遺伝子×10化合物でのマイクロ
アレイデータ行列1つのみが入力であり、この行と列の相
関を最大とする様な線形変換係数を求めた。マイクロア
レイデータは、統計ソフトRのBioConductorを用いて、qu
antile normalizationによる正規化を行った。これを、
同じくRのcorresp関数を用いて対応分析を行った。その
結果、ES固有値成分の分散への寄与率が、

1	15.1835327691295
2	11.2167458799198
3	10.6471761915481
4	10.2045078232397
5	9.28845969866881
6	8.97545500256852
7	8.90753832370446
8	8.80673835314728
9	8.57616096854483
10	8.19368498952898
11	4.17145385144158e-29

となった。この寄与率を考慮に入れながら、各成分で線
形変換係数の絶対値の高いものから、1つか2つずつマ
ーカー候補遺伝子を計 10 個選択した。今回、選択した
遺伝子は、Nanog、SOX2、ZNF208、ADRM1、TRIB1、
CRY1、DMFT1、hVHL、SMAD7、SMAD6 である。また、
内部標準として b-Actin を使用した。

2) 遺伝子ネットワークの予測率への影響解析

国立環境研究所と共同開発してきたレプリカ交換法ベ
イジアンネットワーク推定プログラム RX-TAOfgen 以外に
GGMとCrossRを用いた予測を行った。予測はまず、H22
年度の15化合物に対して行い、結果をRX-TAOfgenと比
較し、日本バイオインフォマティクス学会で口頭発表に採

択され報告を行った（論文3）。その時に用いた遺伝子は、
Oct3/4、Nanog、Nodal、Tuj1、GATA2、Nestin、Lmx1A、
MAP2、Pax6 であり、ES細胞からEB形成後、神経幹細胞
を誘導して神経に至るまでに重要な遺伝子を経験的に
選択している。その結果、ベイジアンネットワークと
CrossR ネットワークを加えた場合に、予測が有意に向上
したが、GGM を使用した場合には予測が向上せず、むしろ
低下することも見られた。今回は 3 種の遺伝子ネットワ
ークのうち、CrossR を用いたものが最も予測率が向上し
た（表1）。

この時のベイジアンネットワークによる遺伝子ネットワ
ークを図7に示した。目視による遺伝子ネットワークのパター
ンの違いを探したが、簡単にはわかるものではなかった。
そこで、このネットワークのエッジで推定された β 値 $9 \times$
 $9=81$ 個を用いて、ネットワーク間での相関係数を計算し
た。その結果、図8に示すように、同じ毒性化合物グル
ープ内では異なるグループより、相関がより高くなる傾向が
見られた。

3) 探索したマーカー遺伝子での 20 化合物毒性予測

最終年度に正準相関解析に基づいて選択した 10 遺
伝子での大規模 RT-PCR による 20 化合物（各 2 リピート計
40 サンプルデータ）の SVM 予測を、遺伝子ネットワークを
使用した方法を用いて行った。（表2）。その結果、予備研
究段階で前回は最も良い予測率の向上があった CrossR
ネットワークだが、今回は予測が向上せずむしろ低下
することが見られた。このため、CrossR および前回あまり
良い予測率が得られなかった GGM は今回の予測から省
略した。一方、ベイジアンネットワークを用いた場合には、
前回と同様に予測率が向上することが観察された。向上
した予測率は3種のどのカテゴリー（発達神経毒性、発癌
毒性、変異原性）においても 90-100%と非常に高い予測
率を示した。この結果から、ベイジアンネットワークは安定
して予測率を向上させる情報を含んでいるのではないかと
いうことが示唆された。

D. 考察

単純に SVM だけを用いるよりも遺伝子ネットワークを用

いた方が、予測率が向上する結果が得られたのは学術的に意義が高い。予測率が向上する理由として、ネットワークパターン（ β 値の分布）に同一毒性グループ内で相関があり、これが情報を付加している可能性が高いことが示された。また、試した 3 種の遺伝子ネットワークのうち、ベイジアンネットワークによる予測率の向上は安定していた。また、H22 年度の結果では、CrossR(交差相関法)による予測率の上昇が観察された。もしかすると、経験的選択遺伝子は数学的に相関のあまりない遺伝子を選んでいったため、単純な相関係数に基づく GGM では相関が検出できずに予測が却って悪くなったのかも知れないと高察される。

また、今回、10 化合物のマイクロアレイだけで知識を入れない遺伝子の選択法を行った。実験も、前回の ESC→EB→NSC→NC という発生をとらず、ESC そのままで判定ができるかという系に変更した。それにも関わらず 90%以上と非常に高性能な予測率が得られたことは、ES 細胞を用いるだけで少数の化合物からしか遺伝子発現の挙動が得られていなくても、未知の化合物の影響を予測することが可能だと示唆を与えている。この結果は統合的な ES 細胞系による化合物毒性予測のパイプラインを組み上げることができ、本研究の目的を十分に達したと考えられる。

今回の研究で、遺伝子ネットワークを推定するのに、遺伝子の数の約4倍以上という制約を設けた。本研究は予算が限られており、また上記の制約があるため、使用できる毒性化合物の種類が当初の計画書よりも大幅に縮小せざるを得なかった。今回は精度の高い実験値を得ることを目標としたため、大規模な RT-PCR 実験を行った。このデータ数が、コントロールや内部標準も合わせると H22 年度は 7,200、H23 年度は 9,768 点を必要とし、多大な実験者の労力がかかったため、この測定過程をできるだけ自動化する必要があると考えられる。これらの膨大な点を用いて研究を遂行したが、概ね優れた結果が得られた。

E. 結論

当初は ES 細胞から神経細胞を分化させた系での毒性

予測を行う計画であったが、研究期間3年間の間に実状に合わないことが判明し、また、機械学習には学習データ数がかかなり必要だが、予算の限界があることから、ES 細胞そのままを用いた最低限の化合物数での予測システムを構築した。前人未到の研究であり、結果が出るまでは、これでうまくいくのか全くの未知数であったが、基礎研究としてこれだけの結果が出たことは稀なことであった。3年間の本研究を通して、遺伝子の数が数 10 となってもデータが揃えば、遺伝子ネットワークを推定し、化合物の毒性予測が可能なシステムの原型が完成したと言える。

F. 研究発表

1. 論文発表

1. Jean-Francois Pessiot, Hyeryung Kim, Wataru Fujibuchi, Pairwise Ranking Component Analysis, submitted to *Knowledge and Information Systems* (under revision review).
2. CELLPEDIA: a repository for human cell information for cell studies and differentiation analyses, Hatano, A., Chiba, H., Moesa, H.A., Taniguchi, T., Nagaie, S., Yamanegi, K., Takai, T., Tanaka, H., *Fujibuchi, W., *Database*, 2011:bar046.
3. Prediction of Chemical Toxicity by Network-based SVM on ES-cell Validation System, *Fujibuchi, W., Aburatani, S., Yamane, J., Imanishi, S., Akanuma, H., Sone, H., Ohsako, S., *The Proceedings of the 2011 Joint Conference of CBI-Society and JSBi*, Nov. (2011).
4. Jean-Francois Pessiot, Hyeryung Kim, Wataru Fujibuchi, Learning similarity functions for multi-platform microarray data, *SIG-BIO*, 2011.
5. Review: Alternative pre-mRNA splicing in digestive tract malignancy. Miura K, Fujibuchi W, Sasaki I. *Cancer Sci.* 102(2):309-16 (2011).
6. In search of true reads: A classification approach to next generation sequencing data, E. Wijaya, J.F. Pessiot, M.C. Frith, W. Fujibuchi, K. Asai & P.

- Horton, *Intl. Conference on Bioinformatics and Biomedicine Workshops*, 37:561-6 (2011).
7. Jean-Francois Pessiot, Hirokazu Chiba, Hiroto Hyakkoku, Takeaki Taniguchi, Wataru Fujibuchi PeakRegressor identifies composite sequence motifs responsible for STAT1 binding sites and their potential rSNPs, *PLoS ONE*, 5(8):1-8, 2010.
 8. Edward Wijaya, Jean-Francois Pessiot, Martin Frith, Wataru Fujibuchi, Kiyoshi Asai and Paul Horton, In Search of True Reads: A Classification Approach to Next Generation Sequencing Data Selection, in *Proc. 2010 IEEE International Conference on Bioinformatics and Biomedicine (BIBM) Next Generation Sequencing Workshop*, 561-566, 2010.
 9. Miura K., Kinouchi M., Ishida K., Fujibuchi W., Naitoh T., Ogawa H., Ando T., Yazaki N., Watanabe K., Haneda S., Shibata C., and Sasaki I. Cell Death and Cancer (Edited by Prof. Samali A) 5-FU Metabolism in Cancer and Orally-Administrable 5-FU Drugs, *Cancers Special Issue*, 2, 1717-1730, (2010).
 10. 千葉啓和、藤渕航「細胞情報解析に役立つツール—幹細胞研究の進展とその創薬応用に向けて」、*実験医学* Vol.28(19)(12月号)2010.
 11. Kato, T. and Fujibuchi, W. Kernel Classification Methods for Cancer Microarray Data, *Medical Biostatistics for Complex Diseases*, pp. 279-303, Wiley-Blackwell, Germany, 2010.
 12. 藤渕航. シミュレーテッドアニーリングによる多重プライマー配列デザイン法, シングルセル解析の最前線, pp. 265-73, シーエムシー出版, 2010.
 13. Sone H, Okura M, Zaha H, Fujibuchi W, Taniguchi T, Akanuma H, Nagano R, Ohsako S, and Yonemoto J. pCEC: a toxicogenomics database with a toxicoinformatics system for risk evaluation and toxicity prediction of environmental chemicals. *J Toxcol Sci*, 35, 115-23, 2010
 14. 幡野晶子、Harry Amri Moesa、千葉啓和、谷口丈晃、永家聖、山根木康嗣、藤渕航. 細胞分化解析を目指した網羅的ヒト細胞データベース「CELLPEDIA」, 情報処理学会研究報告 2010-BIO-20, No.12, 2010.
 15. 金蕙鈴、加藤毅、茂櫛薫、田中博、藤渕航. 正則化正準相関解析を用いた抗がん剤の影響による共通パスウェイ解析, 情報処理学会研究報告 2010-BIO-20, No.9, 2010.
 16. Jean-François Pessiot, Hirokazu Chiba, Hiroto Hyakkoku, Takeaki Taniguchi, Wataru Fujibuchi PeakRegressor identifies composite sequence motifs responsible for STAT1 binding sites and their potential rSNPs, *Proceedings of Critical Assessment of Massive Data Analysis 2009*, pp. 4-11, 2009.
 17. Wataru Fujibuchi, Hirokazu Chiba, Hideo Akiyama, and Hitoshi Shiku Designing Pyro-Primer Sequences Using a Simulated Annealing Algorithm, to Critically Target mRNAs in Quantitative Cell Analysis, *Proceedings of the 6'th International Forum on Post-genome Technologies*, pp. 253-7, 2009.
 18. Hideaki Karasawa, Koh Miura, Wataru Fujibuchi, Kazuyuki Ishida, Naoyuki Kaneko, Makoto Kinouchi, Mitsunori Okabe, Toshinori Ando, Yukio Murata, Hiroyuki Sasaki, Kazuhiro Takami, Akihiro Yamamura, Chikashi Shibata and Iwao Sasaki Down-regulation of cIAP2 enhances 5-FU sensitivity through the apoptotic pathway in human colon cancer cells, *Cancer Sci.*, 100, 903-13, 2009.
 19. 百石弘澄、杉原稔、諏訪牧子、加藤毅、山名早人、藤渕航. 2-way prediction 法による GPCR リカントの結合予測, 情報処理学会研究報告 2009-BIO-18, No.2, 2009.
 20. Wataru Fujibuchi, Hyeryung Kim, Yoshifumi Okada, Takeaki Taniguchi, Hideko Sone, High-performance gene expression module analysis tool and its application to chemical toxicity data, *Methods in Mol. Biol. 577: Reverse Chemical Genetics*, pp. 55-65, Humana Press, U.S.A., 2009.

2. 学会発表

1. Towards iPS cell drug discovery—Prediction of chemical toxicity by gene network learning based on ES cell validation system, *Systems Biology: Global Regulation of Gene Expression*, CSHL, New York, 2012.
2. 平成 23 年度 第 11 回 産総研・産技研 LS-BT 合同研究発表会、「幹細胞インフォマティクスによる化合物毒性予測」、産業技術総合研究所(2012)
3. 幹細胞インフォマティクスによる化合物毒性予測、平成 23 年度 第 11 回 産総研・産技研 LS-BT 合同研究発表会ポスターセッション、産業技術総合研究所(2012)
4. Wataru Fujibuchi, Towards iPS cell drug discovery: prediction of chemical toxicity based on gene network learning, *CBRC Workshop 2011*, Feb. 2012, Tokyo.
5. Prediction of Chemical Toxicity by Gene Network Learning on ES-cell Validation System, Wataru Fujibuchi, Sachiyo Aburatani, Junko Yamane, Satoshi Imanishi, Hiromi Akanuma, Reiko Nagano, Hideko Sone, Seiichiroh Ohsako, poster presentation at *CBRC Workshop 2011*, Feb. 2012, Tokyo.
6. CELLPEDIA: a repository for human cell information for cell studies and differentiation analyses, Akiko Hatano, Hirokazu Chiba, Harry Amri Moesa, Takeaki Taniguchi, Satoshi Nagaie, Koji Yamanegi, Takako Takai-Igarashi, Hiroshi Tanaka, Wataru Fujibuchi, poster presentation at *CBRC Workshop 2011*, Feb. 2012, Tokyo.
7. Reverse Engineering of Genetic Regulatory Networks, Jean-francois Pessiot, Wataru Fujibuchi, poster presentation at *CBRC Workshop 2011*, Feb. 2012, Tokyo.
8. Prediction of Chemical Toxicity by Network-based SVM on ES-cell Validation System, *Fujibuchi, W., Aburatani, S., Yamane, J., Imanishi, S., Akanuma, H., Sone, H., Ohsako, S., *poster presentation at the 2011 Joint Conference of CBI-Society and JSBi*, Nov. 2011, Kobe.
9. Wataru Fujibuchi, Takeaki Taniguchi, Hideko Sone, Seiichiro Ohsako, Analysis of toxic effects on ES cell differentiation by integrated gene networks, poster presentation at *CBRC*, (2011) July, Tokyo.
10. Jean-Francois Pessiot, Hirokazu Chiba, Hiroto Hyakkoku, Takeaki Taniguchi, Wataru Fujibuchi, PeakRegressor identifies composite sequence motifs responsible for STAT1 binding sites and their potential rSNPs, oral presentation at *CBRC*, (2010) July, Tokyo.
11. Jean-Francois Pessiot, Hyeryung Kim, Wataru Fujibuchi, Cross-platform analysis of microarray data, poster presentation at *CBRC*, (2010) July.
12. Jean-Francois Pessiot, Hyeryung Kim, Wataru Fujibuchi, Learning similarity functions for gene expression data, poster presentation at *LS-BT*, (2011) February, Tsukuba.
13. Jean-Francois Pessiot, Hyeryung Kim, Wataru Fujibuchi, Learning similarity functions for multi-platform microarray data, oral presentation at *SIG-BIO*, (2011) March, Kyoto.
14. Wataru Fujibuchi, Takeaki Taniguchi, Hideko Sone, Seiichiro Ohsako, Analysis of toxic effects on ES cell differentiation by integrated gene networks, poster presentation at *CBRC*, (2010) July, Tokyo.
15. Hirokazu Chiba, Taku Tanaka, Masato Miyake, Wataru Fujibuchi, A Method for Optimizing Gene Combination to Induce Adipocyte Differentiation from Mesenchymal Stem Cells, *Cold Spring Harbor Laboratory Meeting—Systems Biology*, NY, U.S.A., Mar. 2010.
16. Taku Tanaka, Hirokazu Chiba, Wataru Fujibuchi, Masato Miyake, Adipogenesis Induced by Short Interference RNAs Predicted by Informatics, *Cold Spring Harbor Laboratory Meeting—Systems Biology*, NY, U.S.A., Mar. 2010.

17. 藤渕航. 生命情報工学の医学への応用. 第9回重粒子医科学センターシンポジウム「先端科学と社会の接点」、静岡(2009)12月
18. Wataru Fujibuchi, Hirokazu Chiba, Hideo Akiyama, and Hitoshi Shiku Designing Pyro-Primer Sequences Using a Simulated Annealing Algorithm, to Critically Target mRNAs in Quantitative Cell Analysis, 6th International Forum on Post-genome Technologies, Beijing, China, Sep. 2009.
19. 藤渕航. 産総研の細胞形態・遺伝子発現情報 統合データマイニングシステムの紹介. 国際バイオEXPO、東京(2009)7月
20. 千葉啓和、幡野晶子、永家聖、山根木康嗣、Larisa Kiseleva、谷口丈晃、Paul Horton、藤渕航. CELLPEDIA/CellMontage 細胞統合解析データベース. 「データベースが拓くこれからのライフサイエンス」、東京大学(2009)6月

G. 知的財産権の出願・登録状況

1. 特許取得

藤渕 航、千葉 啓和, プライマーセット探索装置、プライマーセット探索方法およびプログラム, 特願2009-212703、2009/09/15.

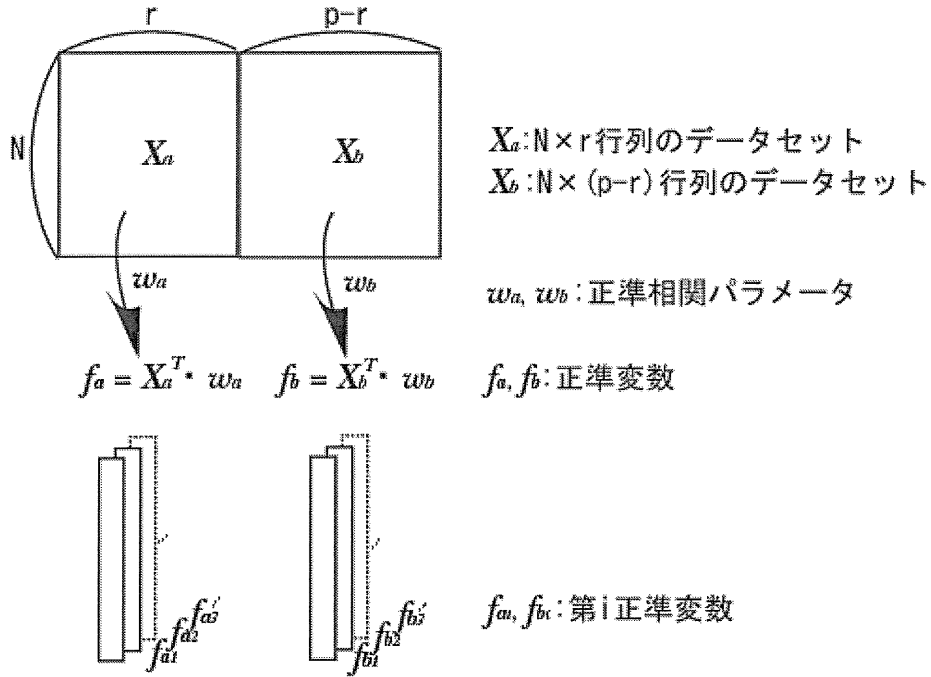
2. 実用新案登録

無し

3. その他

改良版 RX-TAOgen による遺伝子ネットワーク推定ページ:

<http://med.cbrc.jp/rxtaogen/> (要パスワード)



$$\mathbf{k}_a(\mathbf{x}_a, \mathbf{x}'_a) = \langle \phi_a(\mathbf{x}_a), \phi_a(\mathbf{x}'_a) \rangle, \quad \mathbf{k}_b(\mathbf{x}_b, \mathbf{x}'_b) = \langle \phi_b(\mathbf{x}_b), \phi_b(\mathbf{x}'_b) \rangle$$

$$\mathbf{f}_a \equiv \phi_a(\mathbf{X}_a)^T \phi_a(\mathbf{X}_a) \boldsymbol{\alpha}_a = \mathbf{K}_a \boldsymbol{\alpha}_a, \quad \mathbf{f}_b \equiv \phi_b(\mathbf{X}_b)^T \phi_b(\mathbf{X}_b) \boldsymbol{\alpha}_b = \mathbf{K}_b \boldsymbol{\alpha}_b$$

$$\operatorname{argmax} \rho(\mathbf{f}_a, \mathbf{f}_b) = \frac{\boldsymbol{\alpha}_a^T \tilde{\mathbf{K}}_a \tilde{\mathbf{K}}_b \boldsymbol{\alpha}_b}{\sqrt{\boldsymbol{\alpha}_a^T \tilde{\mathbf{K}}_a^2 \boldsymbol{\alpha}_a + \gamma_a \boldsymbol{\alpha}_a^T \boldsymbol{\alpha}_a} \sqrt{\boldsymbol{\alpha}_b^T \tilde{\mathbf{K}}_b^2 \boldsymbol{\alpha}_b + \gamma_b \boldsymbol{\alpha}_b^T \boldsymbol{\alpha}_b}}$$

$$\text{subject to } \boldsymbol{\alpha}_a^T \tilde{\mathbf{K}}_a^2 \boldsymbol{\alpha}_a + \gamma_a \boldsymbol{\alpha}_a^T \boldsymbol{\alpha}_a = \boldsymbol{\alpha}_b^T \tilde{\mathbf{K}}_b^2 \boldsymbol{\alpha}_b + \gamma_b \boldsymbol{\alpha}_b^T \boldsymbol{\alpha}_b = 1$$

図 1 : 正則化を入れたカーネル正準相関解析

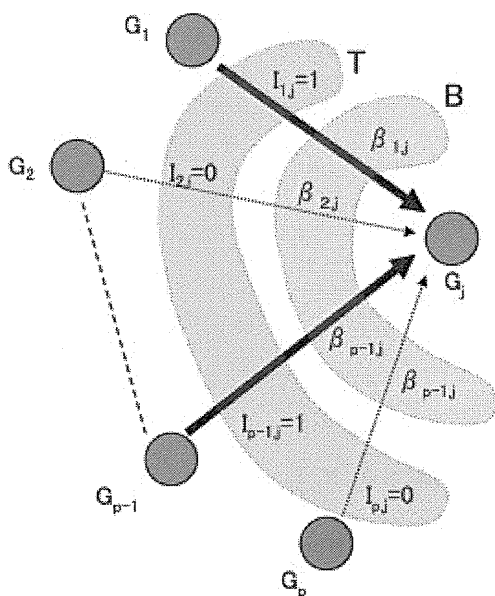
レプリカ交換法による 高性能ベイジアンネットワーク推定法

基本ネットワークモデル

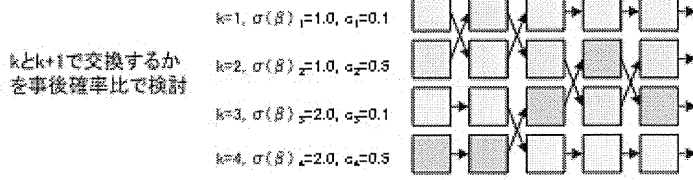
$$E[\log(G_j)] = \sum_{i=1, i \neq j}^p I_{ij} \beta_{ij} \log(g_i)$$

事後確率を最大化することでモデルを推定

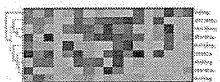
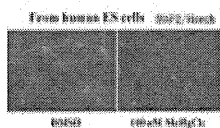
$$f(T, B, S | g) \propto f(g | T, B, S) * h(T) * h(B | T) * h(S)$$



産総研で高性能並列化：レプリカ交換法の導入
(Swendsen and Wang 1986, Earl and Deem 2005)



kとk+1で交換するか
を事後確率比で検討



安定した
ネットワーク
推定が可能
になった。

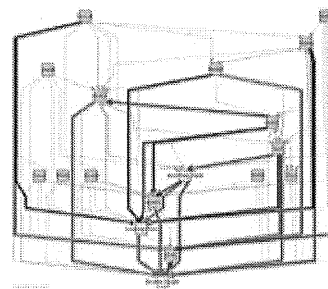


図2：ベイジアンネットワークモデルとレプリカ交換法による高性能化

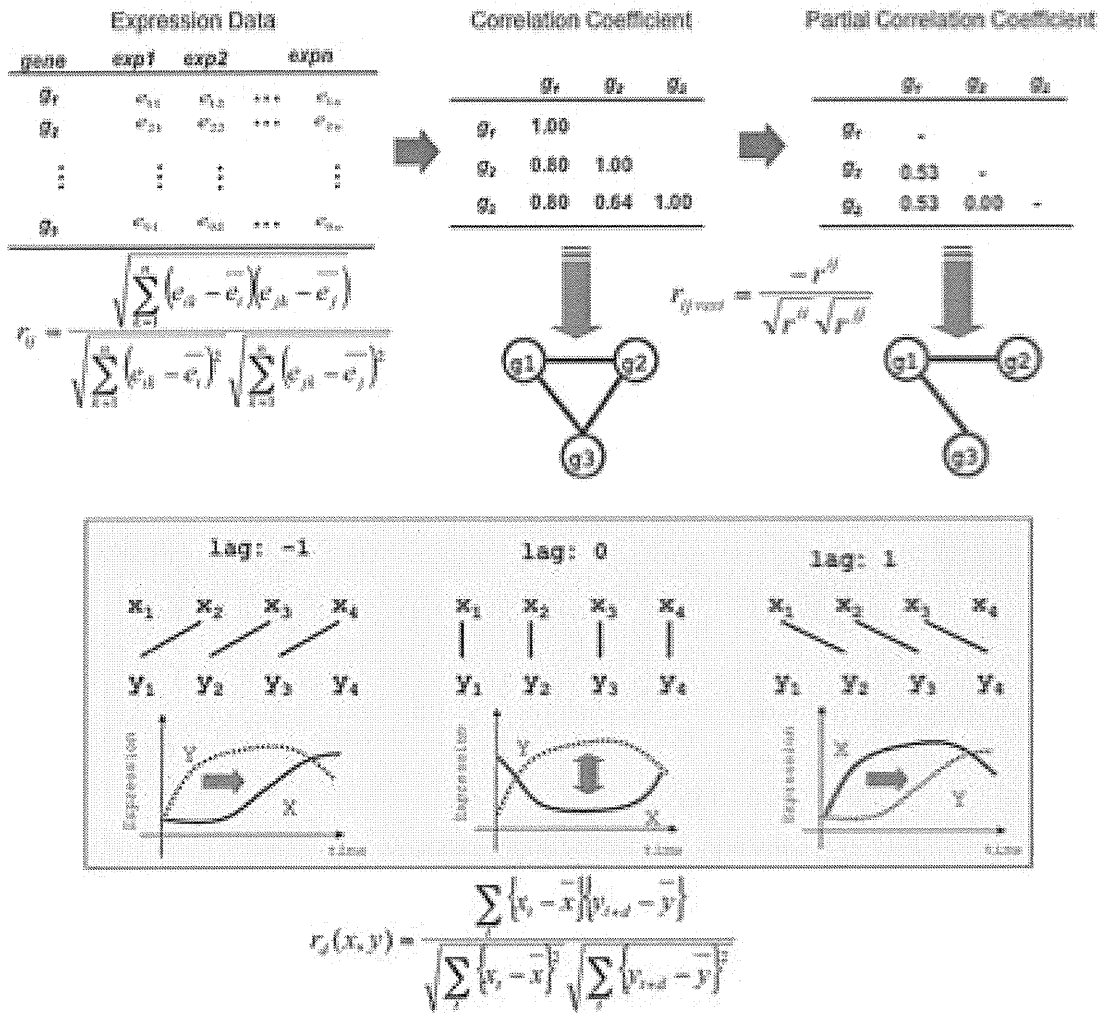


図3：グラフィカルガウシアンモデル（上）と交差相関係数（下）

The screenshot shows the RX-Taogen web interface. At the top, there is a browser window with the URL `http://med.cbrc.jp/rxtaogen/cgi-bin/top.cgi?jobId=`. The page title is "RX-Taogen". Below the title, there are navigation links for "Home" and "Submitted jobs".

The main configuration area includes several input fields and buttons:

- Expression profile**: `sample` (with a help icon), a file path `/Users/wataru/Desktop/厚...`, and a "選択..." button. Default: 18.tsv
- Beta**: `sample` (with a help icon) and a "選択..." button.
- T**: `sample` (with a help icon) and a "選択..." button.

Below these are host and replica settings:

	gb01	gb02	gb03	gb04	gb05	gb06	gb07	gb08
Host / replica(#)	4	4	0	0	0	0	0	0
degree1	1	1	1	1	1	1	1	1
degree2	0.0001	0.0002	0.0004	0.0008	0.002	0.005	0.01	0.02

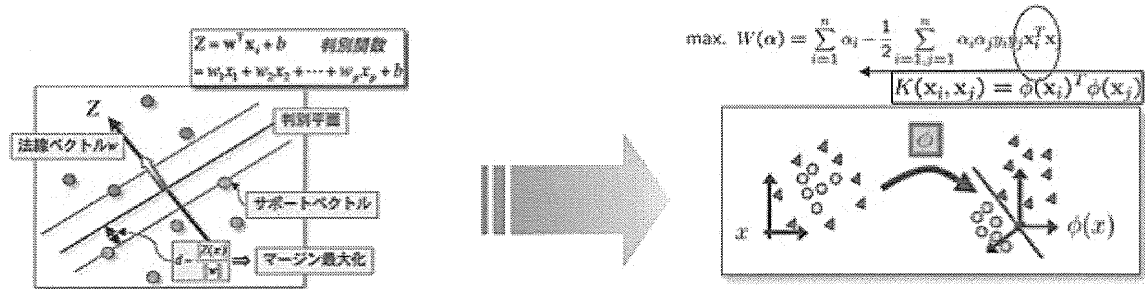
Other parameters include:

- sampling**: `10` ≥ 1
- replica_exchange**: off on
- exchange**: `2000` ≥ 1
- r**: `0.5` (0.0 - 1.0)
- measure**: log posterior propability log likelihood
- burn_in**: `0` ≥ 0
- count_in**: `10000` ≥ 1
- verbose**: .out+.log .out+.log+.param+.measure
- verbose_measure**: `500` ≥ 1
- verbose_parameter**: `500` ≥ 1
- email**: `w.fujibuchi@aist.go.jp`

A "submit" button is located at the bottom left of the configuration area. A status bar at the very bottom of the browser window shows the text "完了".

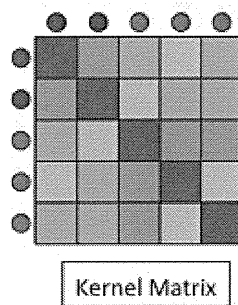
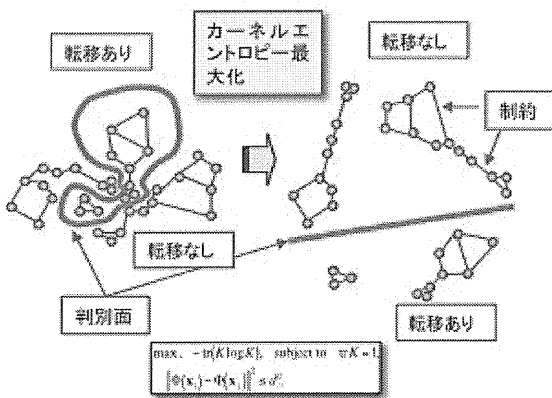
図 4 : RX-TA0gen の WWW 入力ページ。簡便にパラメーターを設定でき、8 コア × 8 台のクラスターマシンの実行が容易にできる。

SVM and Maximal Entropy Kernel



SVM (Linear discrimination)

Kernel can allow non-linear discrimination



Used Kernels:

- Linear
- RBF
- Polynomial
- Maximal Entropy
- EKM(Schölkopf 2002)
- Saigo(Saigo 2004)

Maximal Entropy Kernel (Tsuda and Noble 2004, Fujibuchi and Kato 2007)

図5：カーネルサポートベクターマシン法

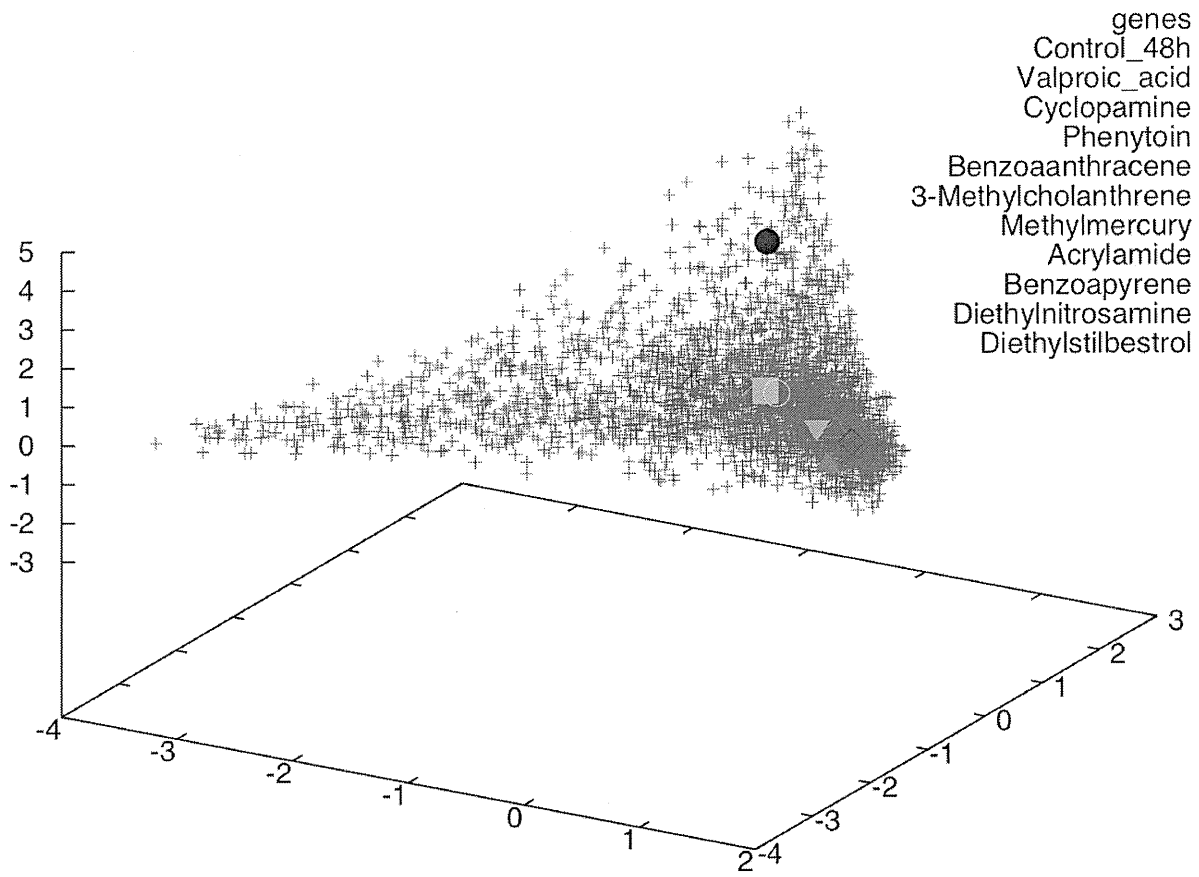


図6: 正準相関解析による10種(+コントロール)の化合物マイクロアレイデータと特徴遺伝子(赤い点)の第3固有値空間による可視化

Inferred Bayesian Networks

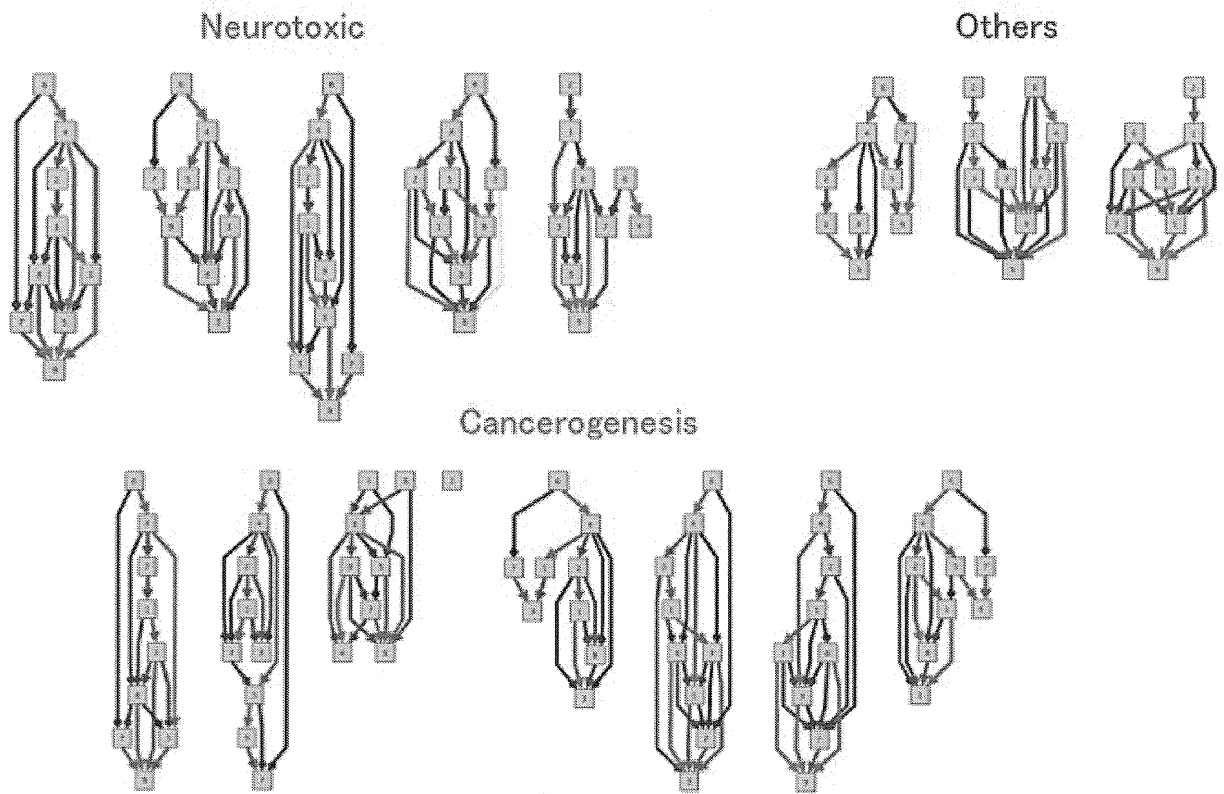


図7:9遺伝子による15化合物を誘導神経細胞に暴露した時の遺伝子ネットワーク

厚生労働科学研究費補助金（化学物質リスク研究事業）
研究分担報告書

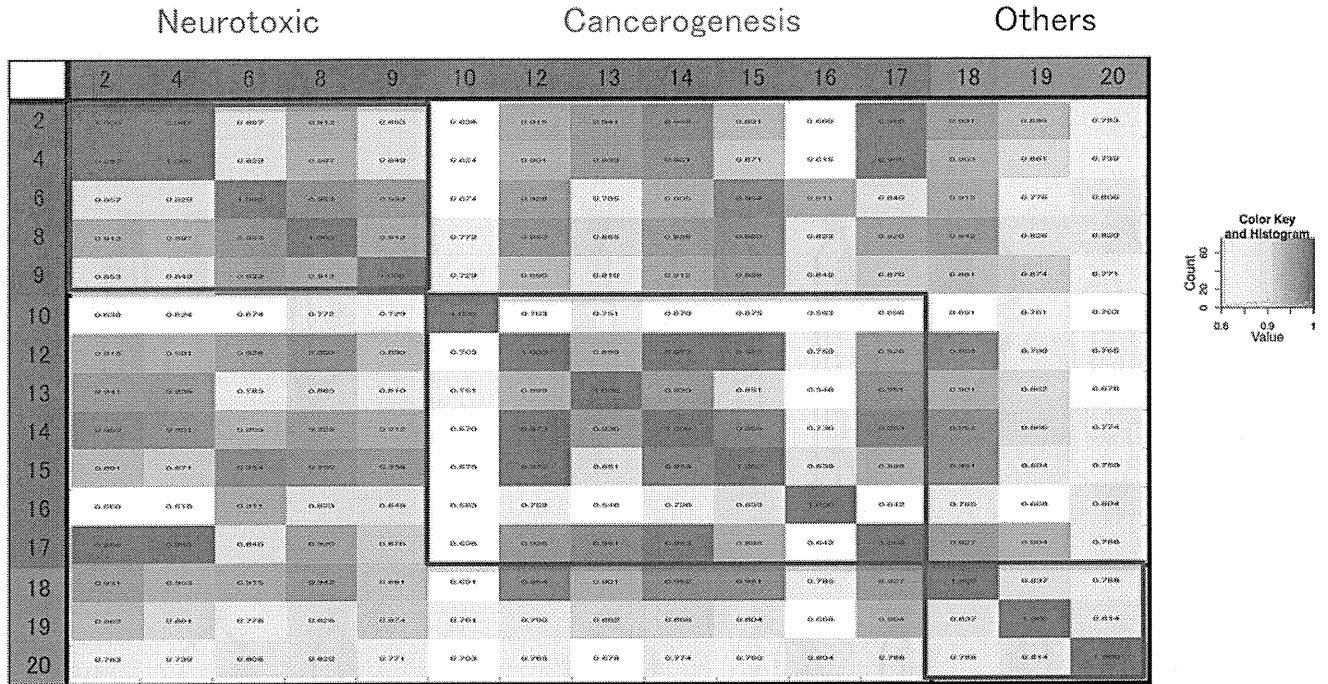


図8: 遺伝子ネットワークのβ値を用いたピアソン相関係数

Toxic Category	#Data	Random	SVM	BN+	GGM+	CrossR+
神経毒性判定	10/30	66.6%	90.0%	93.3%	86.7%	96.7%
発癌毒性判定	14/30	53.3%	96.6%	100.0%	96.6%	100.0%

表1:3種の遺伝子ネットワーク予測法(RX-TAOgen, GGM, CrossR)での15毒性化合物におけるSVM予測率向上の比較

Toxic Category	#Data	Random	SVM	BN+	GGM+	CrossR+
神経毒性判定	18/40	55.0%	87.5%	90.0%	-	-
発癌毒性判定	12/40	70.0%	100.0%	100.0%	-	-
変異原性判定	10/40	75.0%	87.5%	95.0%	-	-

表2:3種の遺伝子ネットワーク予測法(RX-TAOgen, GGM, CrossR)での20毒性化合物におけるSVM予測率向上の比較

ES 細胞試験における化学物質影響の種間・細胞間比較に関する研究

研究分担者

大迫誠一郎

東京大学 疾患生命工学センター 准教授

研究要旨

ヒト胚性幹細胞試験（EST）を利用した毒性影響評価のための、被験モデル化合物としてメチル水銀を用いてその有効性を検討した。ヒト ES 細胞（hESC）ならびにマウス ES 細胞（mESC）から各々の神経系分化培養系を確立し、未分化神経細胞のステージでメチル水銀を各種濃度で曝露した。

MEF 上で大量培養したマウスおよびヒト ES 細胞（B6G2, KhES3）をマイクロデバイス上に播種し、均一な胚葉体を形成させ、オルニチン・ラミニンコート上に再び播種し、神経誘導培地にて神経分化を誘導させた。MeHg（1, 10, 100, 1000 nM）を神経誘導培地に添加し、細胞死を経時的に観察、また最終分化に相当すると思われる時点（マウス：Day23；ヒト：Day50）で、MAP2 陽性細胞の定量的形態観察、各種分化マーカーの発現レベル測定を行った。

MTT アッセイによる細胞死では、マウスで 1 nM からヒトでは 100 nM から有意な低下を示し、マウス細胞の感受性のほうは高いという結果を得た。しかし、マルチチャンネル全自動細胞画像解析装置による MAP2 陽性細胞の形態解析では、ヒトにおいて、神経突起伸長がより低濃度から抑制がかかることがわかった。分化マーカーに対する影響もマウスとヒトでは MeHg に対する感受性が大きく異なることもわかり、細胞死と分化後の細胞形態表現型の影響において、両種間でメチル水銀の濃度に対する感受性が異なることがわかった。この違いを評価するため、ベイジアンネット解析を実施したところ、より低用量から分化マーカーへの影響の見られた hESC 由来細胞のネットワーク解析でメチル水銀ノード（Node）が最上層に来ることがわかった。

今回実施したマウスならびにヒト ES 細胞からの神経分化培養系による比較検討実験は、環境化学物質のヒトへの感受性を検討する上において有効な手段となると思われる。

A. 研究目的

受精卵から胚、胎児を経て成熟個体に至るまで各細胞の分化過程をヒト胚性幹細胞（hESC）の分

化培養系を利用して再現することは可能であり、発生影響試験の理想的モデルと言える。本研究では化学物質の安全性評価で重要な問題であるヒ

トへの生体影響を予測することを目的に、ヒト胚性幹細胞試験（hEST）を利用する。このサブテーマでは古くから発達神経毒性の知られているメチル水銀を用いて、ヒト ES 細胞の神経系細胞分化培養において毒性影響を観察し、形態情報と遺伝子発現情報を取得して確率推論アルゴリズム（ベイジアンネット解析）で評価を行った。比較検討対象にマウス ESC を用い、同一濃度での曝露実験を実施した。

B. 研究方法

hESCは京大再生医科学研究所幹細胞センターの末盛博文博士から提供されたKhES-3 株（XY genotype）を使用した。ES細胞の維持分化にはマウス胚性線維芽細胞（MEF）をフィーダー細胞として継代を行った。胚様体（EB）の形成にはマイクロスフィアアレイ（MSA; 1020 holes, ϕ 300 $\mu\text{m}/\text{hole}$ ）を用いた。MEFを酵素的に浮遊させ、ES細胞集塊のみをディッシュ上に残した上で、Rock阻害剤Y-27632 を添加したEB培地に懸濁し、 2.0×10^4 個/mlのhES細胞をMSAに播種した（Day1）。Day3, Day5 で同じ培地で交換し培養を継続、EBの成長を待った。Day7 でY-27632 非添加のEB培地に交換し、Day9 にNIM培地に交換した。Day11 で形成されたEBを 20 個/wellの密度でオルチニン・ラミニン（O/L）コート 24-well-plateにNPM培地を用いて播種した。2 日おきにNPM培地を交換し、Day18 でNDM培地に交換してその後Day50 に到達するまでNDM培地で培養し続けた。一方、mES細胞にはGreen mouse FM131 mouseから樹立したB6G-2（XY genotype）を使用した。ES細胞の維持培養にはLIFを用い、MEFを支持細胞として用いた。最終継代から 48 時間目に酵素的にmES細胞のみを回収し、EB培地に懸濁してピペッティングによりシングルセルとし、同じくMSAに 1.0×10^5 個/mlの密度で播種しEB形成を行った（Day 1）。

Day3 で 10 nMのRAを含むEB培地（EB+RA）に交換した。1 日おきに、EB+RA で培地交換しDay7 には再びEB培地に交換したDay9 で形成されたEBを 80 個/wellの密度でO/Lコートの 24-well-plateにEB培地を用いて播種した。Day10 に相当する翌日、NIM培地に交換してその後Day23 に到達するまで 2 日おきにNIM培地を交換しながら培養し続けた。

メチル水銀の曝露はヒト培養系では NDM 培地培養期間中の Day27（O/L day 16）から最終培養到達時点の Day50 まで、2 日おきに培地を交換しつつ曝露し続けた。マウス培養系では NIM 培地培養期間中の Day12（O/L day 3）から最終培養到達時点の Day23 まで、2 日おきに培地を交換しつつ曝露し続けた。曝露スタート時点から MTT アッセイにより経時的に細胞活性を測定した。また、培養最終段階で RNA を採取し、*Nanog*, *Pou5f1*, *Nodal*, *Nes*, *Pax6*, *Emx2*, *Mtap2*, *En1*, *Otx1*, *Hoxb1*, *Hoxb4*, *Olig1*, *Olig2*, *Gapdh* の mRNA を定量 RT-PCR で測定した。また MAP2 抗体による免疫染色と Hoechst 33342 による核染色を施し、IN Cell Analyzer 1000（GE Healthcare UK Ltd., Buckinghamshire, UK）で MAP2 陽性細胞のからの神経突起長と分岐点数を測定した。

（倫理面への配慮）

東京大学ライフサイエンス委員会倫理審査専門委員会にて 2009 年 12 月機関承認を得、それにより、2010 年 1 月文部科学省より使用許可を得た。

C. 研究結果

メチル水銀は 10 nM からマウス胚様体（EB）の生存率を低下させる作用を示したが（図 1）、生存した神経細胞の dendrite 伸長には影響がなかった（図 2 および図 3）。一方ヒト EB では、10 nM メチル水銀は生存率には影響がないものの（図 1）、分化した神経細胞の dendrite 伸長に

は 1nM から統計学的に有意な抑制効果を示した（図2 および図3）。ヒトおよびマウス間で十分に比較可能なオーソログ遺伝子 13 種を定量 PCR で測定したデータセットと形態情報（IN Cell Analyzer 1000 解析）による確率推論モデルにおいても、マウスではメチル水銀に起因するネットワーク構造が見られないのに対して、ヒトではメチル水銀を最上位（親 Node）にもつネットワークを描くことがわかった（図4）。

D. 考察

メチル水銀曝露は低濃度（10nM）でマウスの EB 生存率低下を引き起こしたが、ヒトではこの濃度では EB 生存率に影響はなかった。しかしながら、1nM から分化神経細胞のデンドライト伸長抑制効果が観察できたことは、この物質のヒト神経細胞への特有の効果だと考えられる。マウスではこれまで、実験動物による胎生期メチル水銀曝露で神経系影響モデルを作ることが困難とされてきたが、これは神経細胞機能異常に先んじて細胞死が起きていた可能性がある。またヒトでは細胞死よりも神経系機能異常が生じるため胎生期影響が顕著に出るのではないかと考察できる。

E. 結論

上記の結果より、本研究課題で開発した神経細胞分化系は、同一化合物であっても、発生という観点からヒトにおける特徴的影響を観察できるポテンシャルをもつことが実証された。また、評価法として導入するベイズ推定モデルも遺伝子発現変動とその結果である細胞表現型に予測性を持つことが提示できた。これら研究成果はヒトの感受性を評価するためにヒト ES とマウス ES 細胞を利用した本システムの有効性を示している。

F. 健康危険情報

なし。

G. 研究発表

1. 論文発表

1. He X, Imanishi S, Sone H, Nagano R, Qin XY, Yoshinaga J, Akanuma H, Yamane J, Fujibuchi W, and Ohsako S. Effects of Methylmercury Exposure on Neuronal Differentiation on Mouse and Human Embryonic Stem Cells. *Toxicol Lett* (2012) (in press)
2. Yoshioka W, Aida-Yasuoka K, Fujisawa N, Kawaguchi T, Ohsako S, Hara S, Uematsu S, Akira S, and Tohyama C. Critical role of mPGES-1 in the pathogenesis of hydronephrosis caused by lactational exposure of mice to dioxin. *Toxicol Sci* (2012) (in press).
3. Nagano R, Akanuma H, Qin XY, Imanishi S, Toyoshiba H, Yoshinaga J, Ohsako S, and Sone H. Multi-Parametric Profiling Network Based on Gene Expression and Phenotype Data: A Novel Approach to Developmental Neurotoxicity Testing. *Int J Mol Sci* 13(1), 187-207, (2012).
4. Ohsako S. Perinatal exposure to environmental chemicals induces epigenomic changes in offspring. *Genes Environ* 33, 43-49, (2011).
5. Sone H, Okura M, Zaha H, Fujibuchi W, Taniguchi T, Akanuma H, Nagano R, Ohsako S, and Yonemoto J. pCEC: a toxicogenomics database with a toxicoinformatics system for risk evaluation and toxicity prediction of environmental chemicals. *J Toxicol Sci* 35, 115-123, (2010).
6. Alam MS, Ohsako S, Matsuwaki T, Zhu XB, Tsunekawa N, Kanai Y, Sone H, Tohyama C, and Kurohmaru M. Induction of spermatogenic cell apoptosis in prepubertal rat testes irrespective of