

野生型マウス胚における遺伝子発現経時データベースを用いた
T遺伝子に関わる局所シグナルネットワークの描出

- ・ 経時変化を基としたピアソンの相関係数の利用
- ・ ISHによる発現部位の可視化

31遺伝子の中から9遺伝子を選択しISHを検討後、抽出されてきた5遺伝子

- ・ T遺伝子との関連が既知の
Sp5とZic3遺伝子
- ・ 原始線条部位に発現し、T遺伝子との関連が強く示唆される
Dl13、FrzbおよびLefty2遺伝子

*In silico*でのプロモーター解析：

- Zic3以外の4遺伝子の転写開始点上流にT-box結合配列が存在した
- ・ Zic3はT遺伝子による転写制御を受けているとの報告あり

T遺伝子によるDl13、FrzbおよびLefty2遺伝子発現制御は未報告

経時データベース、ISHおよび*In silico*でのプロモーター解析を、
組み合わせる本解析を通し、当該遺伝子ネットワークに関連する既
知だけでなく**制御未知**の遺伝子を効率よく探索できるものとする

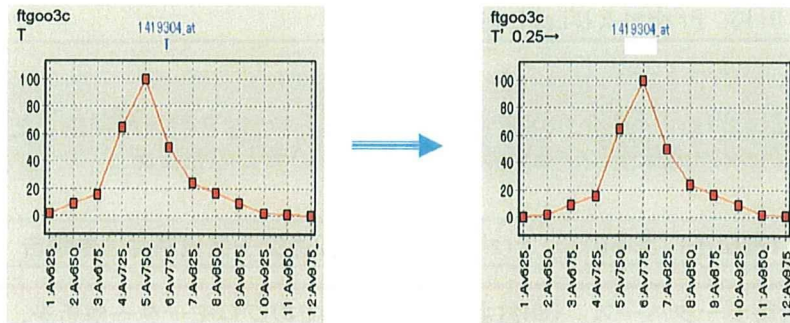
野生型マウス胚における遺伝子発現経時データベースを用いた
T遺伝子に関わる局所シグナルネットワークの描出

さらに下流のレギュラトリーネットワークの探索

T遺伝子の発現経時変化を
0.25日分シフトさせたものを基とした解析

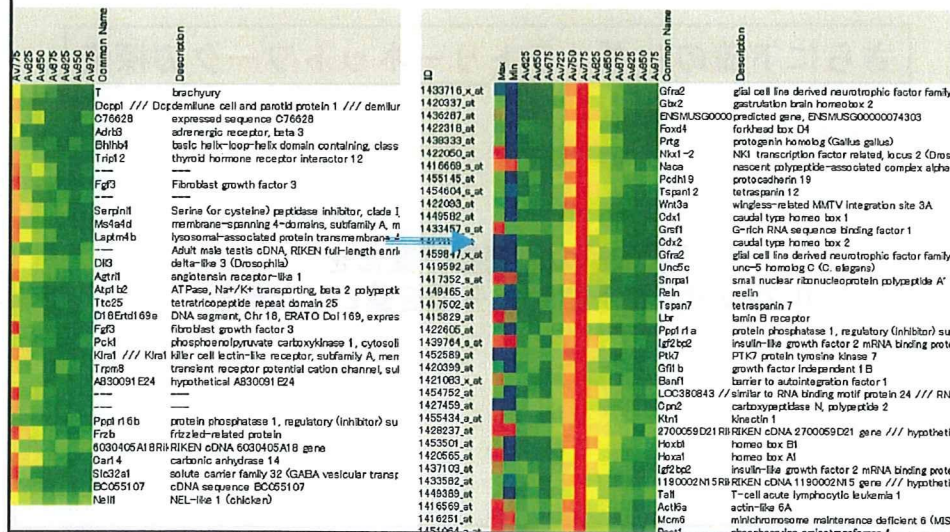
さらに下流のレギュラトリーネットワークの探索

T遺伝子の発現経時変化を
0.25日分シフトさせたものを基とした解析



さらに下流のレギュラトリーネットワークの探索 T遺伝子の発現経時変化を0.25日分シフトさせたものを基とした解析

→ 新たな遺伝子リストが抽出できる



野生型胚およびES・EB分化系における
T (Brachyury) 遺伝子の局所カスケードの描出を行う

I 野生型マウス胚における遺伝子発現経時データベースを用いたT遺伝子に関わる局所シグナルネットワークの描出

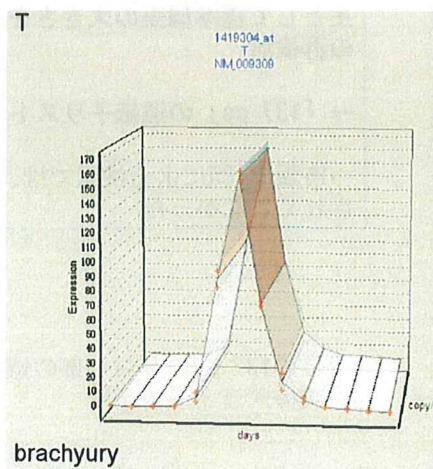
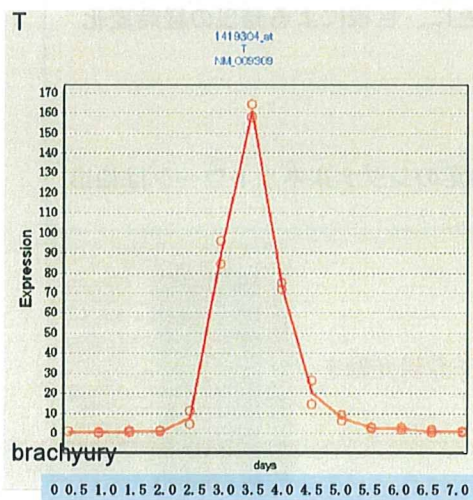
II マウスES・EB分化系の遺伝子発現変動経時データベースを用いたT遺伝子に関わる局所シグナルネットワークの描出

T遺伝子:

脊索を含む中胚葉の形成と原腸陥入に関わる転写因子

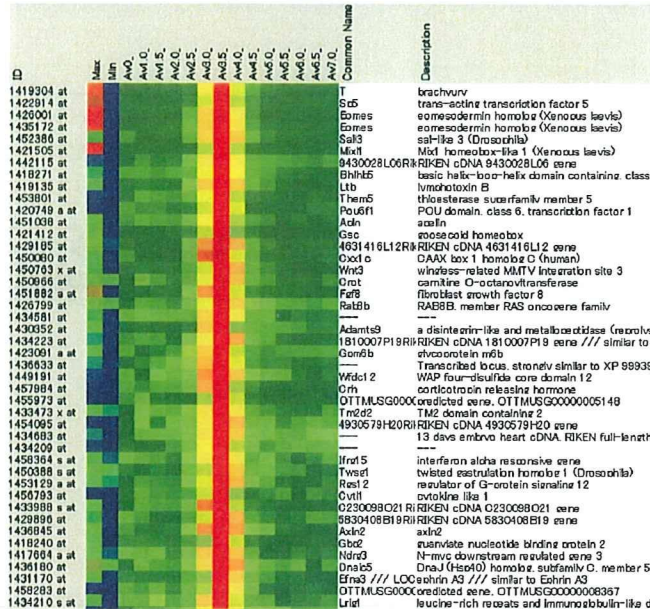
マウスES・EB分化系の遺伝子発現経時データベースにおける
T遺伝子の発現変動

[EB形成0から7日: TIME POINT: 14点]



T遺伝子と似た経時変化を示す遺伝子の抽出 [ES・EB分化系]

標準化(normalize)したT遺伝子の各時点の平均値の経時変化を基にピアソンの相関係数を求め α 値が高い順に並び替えた



T遺伝子と似た経時変化を示す遺伝子の抽出 [ES・EB分化系]

ピアソンの相関係数を求め α 値が高い順に並び替えた

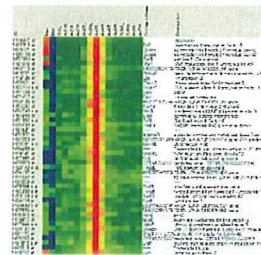
発現コピー数に制限を設けなかった

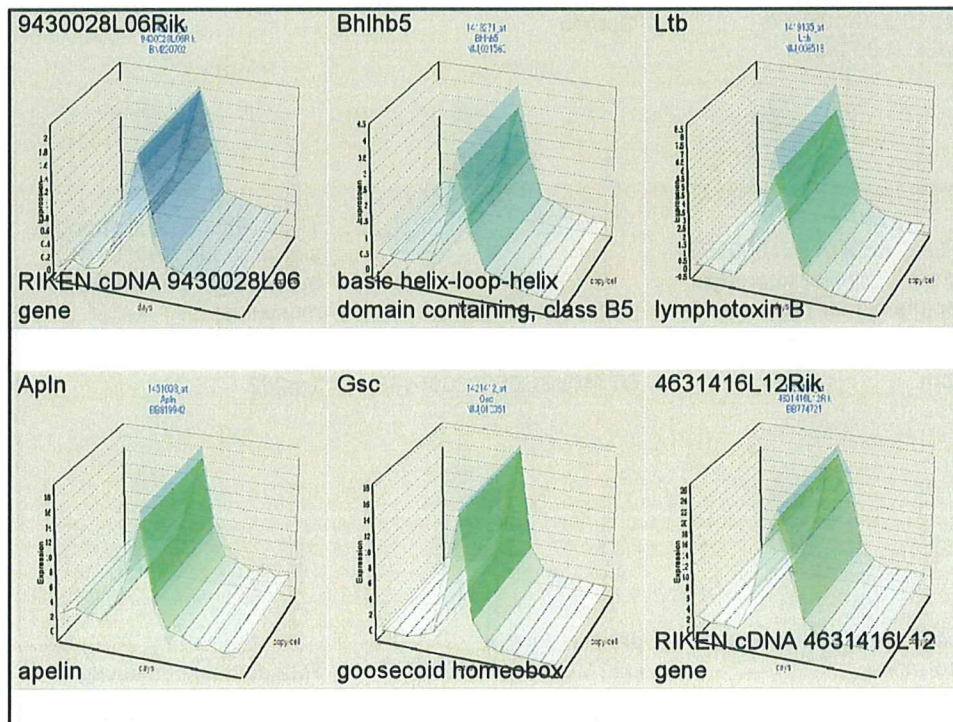
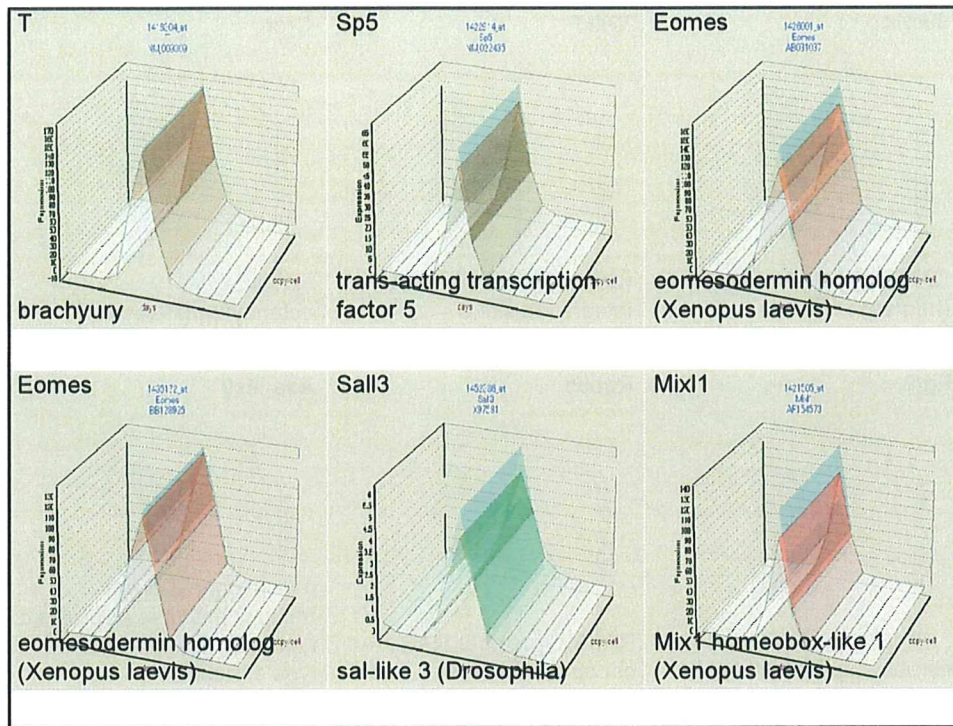
α 値が0.90以上を示す遺伝子リスト「287 ps」について、主として標準偏差の大きさをもとに、目視による発現の経時変化の再検討

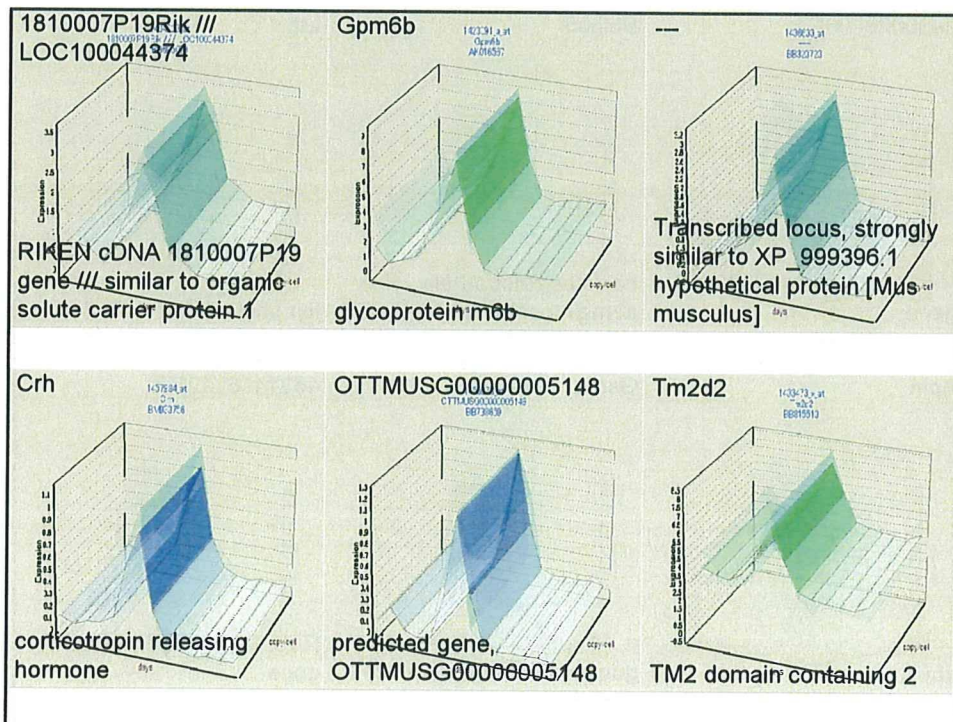
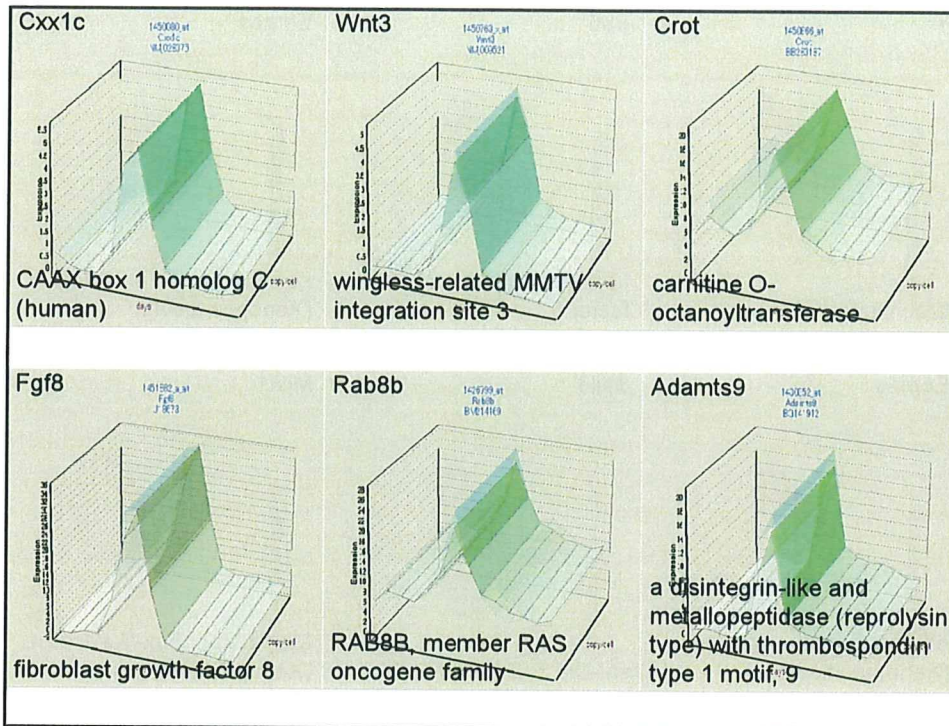
→ 「137 ps」の遺伝子リスト

→市販のIPAIによる検索では、特定のシグナルネットワークは抽出されてこなかった

この137 psのうち一部の遺伝子発現変動を次のスライドで示す





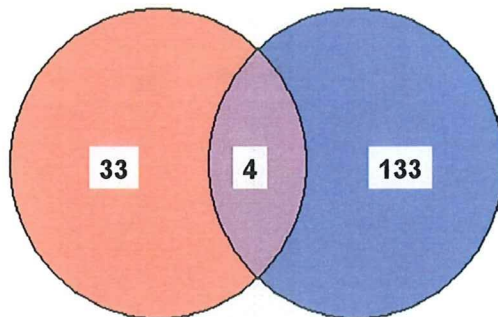


T遺伝子と似た経時変化を示す遺伝子の抽出

- ・ ES・EB分化系の遺伝子発現の経時データベースから得た
137 psの遺伝子リスト
- ・ 野生型胚の遺伝子発現の経時データベースから得た
37 psの遺伝子リスト

比較・検討

赤：野生型胚の経時データベース
青：ES・EB分化系の経時データベース



両者に含まれていた
4遺伝子：

Sp5, Dapk1, Cer1, Dock11

T遺伝子と似た経時変化を示す遺伝子の抽出

- ・ ES・EB分化系における遺伝子発現の経時データベースから得た
137 psの遺伝子リスト
- ・ 野生型胚における遺伝子発現の経時データベースから得た
37 psの遺伝子リスト

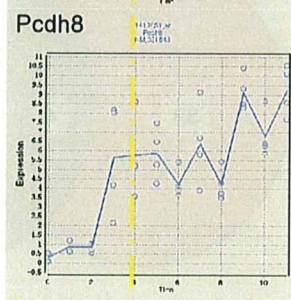
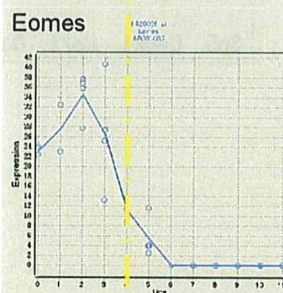
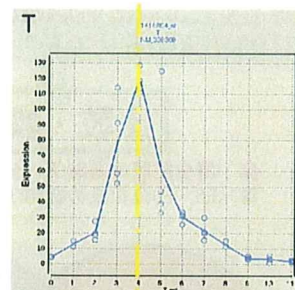
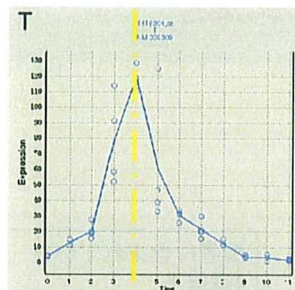
→ Q: 両者が、かなり異なる理由は??

ES・EB分化系の経時データベースから得た137 psの遺伝子リスト



各遺伝子の発現変動を野生型胚の経時データベースを用いて検討

ES・EB分化系の遺伝子発現の経時データベースから得たリストの中の多くの遺伝子の発現変動は、野生型胚では、T遺伝子のものと異なっていた：より早期に発現増加、持続的な発現増加等



T遺伝子と似た経時変化を示す遺伝子の抽出[ES・EB分化系]

・ES・EB分化系は、マウス胚の採取が困難な着床前後の胎生4～6日の代替となりうるため、その遺伝子発現の経時データベースは有用

・ES・EB分化系における分化過程は、野生型胚における発生過程と、必ずしも一致しない可能性が考えられた。

→ ES・EB分化系から得られた結果については、野生型胚において再度検証する必要があるものとする。

目 的

「胎児、ES細胞、概日変動等の自律的なカスケードの描出」

- ・自律的な遺伝子発現が認められる発生過程や概日リズムのシグナルカスケードを描出する
- ・マウス胚およびES・EB培養分化系に化学物質を暴露し、遺伝子発現データを得て、発生過程のシグナルカスケードに対する影響を検証する

平成21年度：

野生型胚およびES・EB分化系の局所カスケードの描出を行う。ES・EB分化系にレチノイン酸受容体 (RAR) 関連物質を暴露し、データを得て、ES分化系における局所カスケードを描出する。

平成22年度：

マウス肝の網羅的な概日カスケードを描出する。引き続きRARカスケードを描出し、発生の自律的なカスケードと比較・検証する。

平成23年度：

腎、肺等の多臓器の概日変動カスケードを描出する。野生型胚におけるRAR関連局所カスケードを描出し、先の結果と比較・検証する。

T遺伝子に関わる局所シグナルネットワークの描出

ま と め[野生型胚]

野生型胚の遺伝子発現経時データベースを利用：

- ・経時変化を基としたピアソン相関係数の利用
- ・ISHによる発現部位の可視化 [31遺伝子の中から9遺伝子を選択]
- ・得られた遺伝子について *In silico*でのプロモーター解析

- ・T遺伝子との関連が既知の
Sp5とZic3遺伝子
- ・原始線条部位に発現し、T遺伝子との関連が強く示唆される
Dl13、FrzbおよびLefty2遺伝子
→Zic3以外の4遺伝子の転写開始点上流にT-box結合配列が存在した
- ・Zic3はT遺伝子による転写制御を受けているとの報告あり

T遺伝子によるDl13、FrzbおよびLefty2遺伝子発現制御は未報告

経時データベース、ISHおよび *In silico*でのプロモーター解析を、組み合わせる本解析を通し、当該遺伝子ネットワークに関連する既知だけでなく **制御未知**の遺伝子を効率よく探索できるものとする

「遺伝子に関わる局所シグナルネットワークの描出

ま と め [ES・EB分化系]

マウスES・EB分化系の遺伝子発現経時データベースを利用：

- ・ 経時変化を基としたピアソンの相関係数の利用
- ・ 野生型胚の遺伝子発現経時データベースでの解析結果との比較



・ ES・EB分化系における分化過程は、野生型胚における発生過程と必ずしも一致しない可能性が考えられた

ES・EB分化系から得られた結果については
野生型胚において再度検証する必要があるものとする

来年度[平成22年度]の予定

・ ES・EB分化系において、アンタゴニスト適用時のRAR分子ネットワークを描出し、非適用時の自律的ネットワークと比較

・ 既に遺伝子発現の経時データベースを豊富に有している成熟期マウス肝における、網羅的な概日ネットワークの描出

厚生労働科学研究費補助金(化学物質リスク研究事業)

「化学物質の有害性評価手法の迅速化、高度化に関する研究
—網羅的定量的大規模トキシコゲノミクスデータベースの維持・拡充と毒性
予測評価システムの実用化の為にインフォマティクス技術開発—
(H21-化学—一般-001)

分担研究報告書

【トキシコゲノミクスのためのインフォマティクス開発】

複数の化学物質のクラスタ交差点からのカスケード描出

研究分担者 相崎 健一

国立医薬品食品衛生研究所・安全性生物試験研究センター・毒性部・室長

研究要旨

既に100余の化学物質による遺伝子発現変動情報を収録し、現在もさらに拡張しつつあるPercellomeデータベースを有効活用して、特に複数の化学物質のクラスタ交差点からのカスケード描出を実現するための基本技術開発を行う。初年度は特徴的な発現変動パターンを自動抽出アルゴリズム(RSort)を活用した総当たりの反応類似度計算を試行し、その有効性を評価した。

A. 研究目的

全遺伝子の発現データから生物学的に有意な情報を効率的に抽出し、これを元に遺伝子発現誘導カスケードを描出する技術を生成し、安全性評価におけるトキシコゲノミクスの有機的活用促進の一助とする。特に、複数の化学物質暴露により共通して誘導・抑制される遺伝子群に着目し、毒性に関わる主たる要素を分離すると共に、それらの関連の解明を目指す。

B. 研究方法

我々は遺伝子発現データの絶対量化技術である Percellome 手法を考案し、これに最適化した”Millefeuille”(ミルフィーユ)データ処

理システムを開発している。

開発・改良したアルゴリズムは、研究者本人がソフトウェアを作成し、実際に本研究計画で得られた生データをサンプルに用いて、ワークステーション上で性能検証を行う。これは本研究におけるプログラム開発はアルゴリズムの改良・微調整を兼ねた try and error 方式で進めなければならない、職業プログラマーに開発委託する際に必要となる仕様の事前決定が出来ないためである。開発言語には、研究的なプログラミングに適した、つまり生成したプログラムの実行速度が早だけでなく、実行プログラムの生成(コンパイル速度)も速い Delphi 言語(=object pascal 言語, 開発環境としては Borland Delphi7 もしくは Delphi2007,

Delphi2010)を使用した。また一般的なグラフ描画には TeeChart コンポーネント (Steema 社) を利用し、100 万件以下の小規模なデータベース操作には組み込み型リレーショナルデータベースコンポーネントである DBISAM (ElevateSoftware 社) を利用して、プログラム開発の効率化を図った。

大規模計算が必要な場合は、専用計算サーバーシステム (MF サーバーシステム) にアプリケーションソフトウェアを移植して、大規模計算用グリッド PC や高速データベースエンジンである Teradata データベース (日本 Teradata 社) を活用して計算処理を行った。

C. 研究結果

既に Percellome データベースに登録されているほぼ全ての遺伝子発現変動データを 3 次元平面波 (Surface グラフ) に変換し、その凹凸の多寡を評価して単純なものから複雑なものへ並び替える RSort アルゴリズム (先行研究で独自に開発) を利用して、凹凸の少ない単純なパターンを示す遺伝子、つまり、より強力もしくはよりユニークな発現制御を受ける遺伝子を化学物質および臓器ごと (肝、肺、腎) に網羅的に自動抽出した。

この結果、209 の Surface データセット (4 万 5 千遺伝子の発現変動を示す三次元曲面波からなるデータセット) から延べにして約 58 万の候補遺伝子情報を得た。

これらの候補遺伝子群は各化学物質が誘導する生体反応の詳細を直接示すものであり、通常は各々について、さらに精密な解析を行う対象であるが、それと平行して、ここでは総合的な解析を行うためのプロファイル情報として利用した。

具体的には、総当たり比較で共通する遺伝

子を抽出し、その数を母数 (化学物質ごとの候補遺伝子総数) で除した値を「反応類似度 (以下、類似度)」として評価して、化学物質ごとに共通遺伝子発現の多い (= 類似の遺伝子発現変動を呈する) 化学物質リストを得ることに成功した。このプロセスは、類似度と共に、類似根拠となる詳細情報 (遺伝子) を同時に生成する点で優れている。

さらに化学物質同士の関連性を概覧するために、化学物質ごとの類似度リストを元にして類似度の正方向列を生成し、可視化手段の 1 つとして主成分解析を行った。

まず臓器を問わず、肝・肺・腎すべての Surface データセットを一括計算したところ、主成分空間では臓器毎のクラスが形成された。次に肝サンプルによる Surface データセットのみで主成分計算したところ、各要素は暴露実験に用いた化学物質の特性を反映した座標にプロットされ、特に Valproic acid や Clofibrate など、単回もしくは新型慢性暴露実験において複数回用いられた化学物質 (を用いたプロジェクトの Surface データセットの要素点) がクラスを形成する様子が観察された。

D. 考察

RSort アルゴリズムの実装アプリケーションソフトウェア RSort.exe は予定通り改良を進めており、本年度は大量のプロジェクトデータの連続自動計算を可能とし、今回実施したような Percellome データベースに登録されているほぼ全てのデータを対象とした一括解析を簡便に実施できるようになった。

しかしながら、現在の RSort.exe は見落としを避ける様に基本設計されているため、計算結果である候補遺伝子リスト中に、相当数の擬陽性データを含んでしまう。

この特性は、今回実施したような自動一括解析において誤差の発生原因となるため、今後も偽陰性データの発生を回避しつつ、擬陽性データの発生を抑制するように、アルゴリズムの微調整を進める。具体的には第二ピークの抽出条件設定機能を追加するなど、複合的な絞り込み機能を拡充し、候補遺伝子の自動抽出精度や網羅性をより一層向上させる計画である。

本年度の研究では Surface データセットを 209 個用いた。これらから得られた候補遺伝子の延べ数は前述の通り58万個であったが、和集合サイズは全遺伝子数相当の4万5千個、遺伝子毎の頻度は最大でも62件、逆に頻度1の遺伝子はわずか80個しかなく、化学物質暴露の際に常に発現誘導されるような特別な遺伝子はないことが分かった。

これは、少数のバイオマーカーだけでは情報不足であって、見落としのない毒性評価を行うためには、本班研究が最終目標とする遺伝子発現誘導カスケードの解明が不可欠であることを強く支持するものである。

また反応類似度の総当たり評価は、既存化学物質の毒性機序解明ばかりでなく新規化学物質の毒性評価にも応用可能であり、部分的であっても既知化学物質と同様の毒性機序を動かす場合、高感度で危険性を検出することが出来る。

将来的には、検出した反応類似性の根拠である共通遺伝子リストをカスケード情報に当てはめて解析・評価することにより、感度が高く、なおかつ確度の高い毒性評価・予測が可能となるものと期待される。

E. 結論

先行研究により世界有数のトキシコゲノミクス

データベースとなった Percellome データベースを用いることで、化学物質暴露時に誘導される遺伝子情報を総合的に捉えることが出来るようになってきた。

これは複数の化学物質暴露毒性で共通する遺伝子や、化学物質ごとに特徴的な暴露影響に対応する遺伝子の同定の足がかりとなる。今後はこれら候補遺伝子リストの解析を進め、毒性の発現誘導カスケード描出の一助とする。

またこれらをプロファイルとして用いることで、新規の化学物質についても、既存化学物質との反応類似性を遺伝子レベルで評価することが出来るようになった。これは長らく待たれていたトキシコゲノミクス研究成果の実用化の端緒となるものである。

この技術の幅広い活用を促進すべく、今後はバイオインフォマティクスを専門としない研究者や専門行政官が、研究・業務に利用することのできる専用解析アプリケーションプログラムの開発も行う。

F. 研究発表

1. 論文発表

Suzuki A, Igarashi K, Aisaki K, Kanno J, Saga Y. NANOS2 interacts with the CCR4-NOT deadenylation complex and leads to suppression of specific RNAs. Proc Natl Acad Sci U S A., 107, 3594-3599, 2010.

Tanemura K, Igarashi K, Matsugami TR, Aisaki K, Kitajima S, Kanno J. Intrauterine environment-genome interaction and children's development (2): Brain structure impairment and behavioral disturbance

induced in male mice offspring by a single intraperitoneal administration of domoic acid (DA) to their dams. J Toxicol Sci., 34 Suppl 2, SP279-286, 2009

2. 学会発表

菅野 純, 相崎健一 Percellome トキシコゲノミクスプロジェクトの進捗-インフォマティクス構築へ- 第36回日本トキシコロジー学会学術年会、2009年7月7日、岩手

種村健太郎, 松上稔子, 五十嵐勝秀, 相崎健一, 北嶋聡, 菅野純 脳発生-発達期の神経シグナルかく乱による遅発性中枢影響解析-幼若期雄マウスへのトリアゾラム投与による学習記憶障害について- 第36回日本トキシコロジー学会学術年会、2009年7月7日、岩手

高木 篤也, 北嶋 聡, 五十嵐 勝秀, 相崎 健一, 菅野 純 Percellome 手法によるマウス ES 細胞分化過程における遺伝子発現の経時データベースの構築と活用 第32回日本分子生物学会年会、2009年12月9日、東京

G. 知的財産所有権の出願・登録情報

1. 特許取得

特許 4415079 「遺伝子の絶対発現量測定方法」

2. 実用新案登録

なし

3. その他

なし

厚生労働科学研究費補助金(化学物質リスク研究事業)

化学物質の有害性評価手法の

迅速化、高度化に関する研究

一網羅的定量的大規模トキシゲノミクスデータベースの維持・拡充と
毒性予測評価システムの実用化のためのインフォマティクス技術開発—
(H21-化学-一般-001)

平成22年3月19日
於. 東京国際フォーラム

分担研究課題

複数の化学物質のクラスタ交差点からの カスケード描出

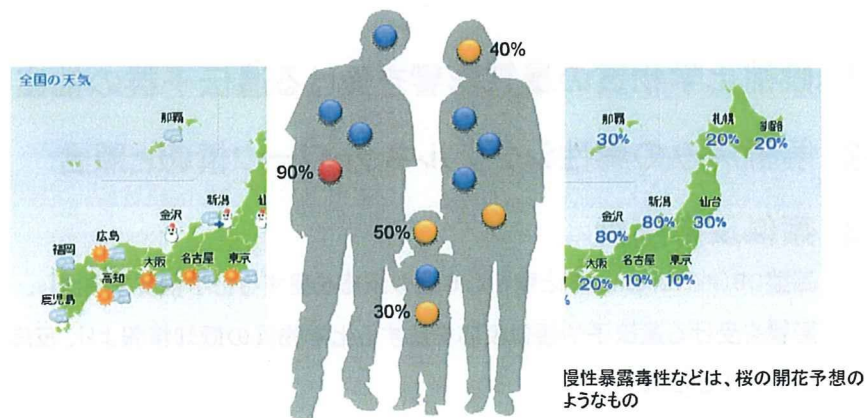
相 崎 健 一

国立医薬品食品衛生研究所
安全性生物試験研究センター
毒性部

1

トキシゲノミクス応用による毒性予測・評価

あるべき出力形式



- 必要な出力情報: 毒性の種類、強度および発生確率と信頼度

2

トキシコゲノミクス応用による毒性予測・評価

(完成型=毒性シグナルネットワークの全容解明後)

- ① 候補化学物質の暴露影響を受ける遺伝子群の抽出
- ② 毒性シグナルネットワーク全情報と照合
- ③ 反応シグナルのin silico合成
- ④ 毒性予測・評価

3

トキシコゲノミクス応用による毒性予測・評価

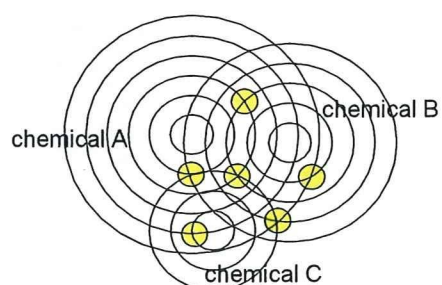
(暫定型=毒性シグナルネットワークの全容解明前)

- ① 候補化学物質の暴露影響を受ける遺伝子群の抽出
 - ② 解析済みの毒性シグナルネットワーク情報と照合
 - ③ 類似反応解析
基盤DB(PercellomeDB)と照合して類似反応を呈する化学物質を抽出し、影響を受ける遺伝子や類似反応を呈する化学物質の既知情報より、反応シグナルを推測
 - ④ 毒性推測
- >>>>> ②, ③から、毒性シグナルネットワークの描出も進める

4

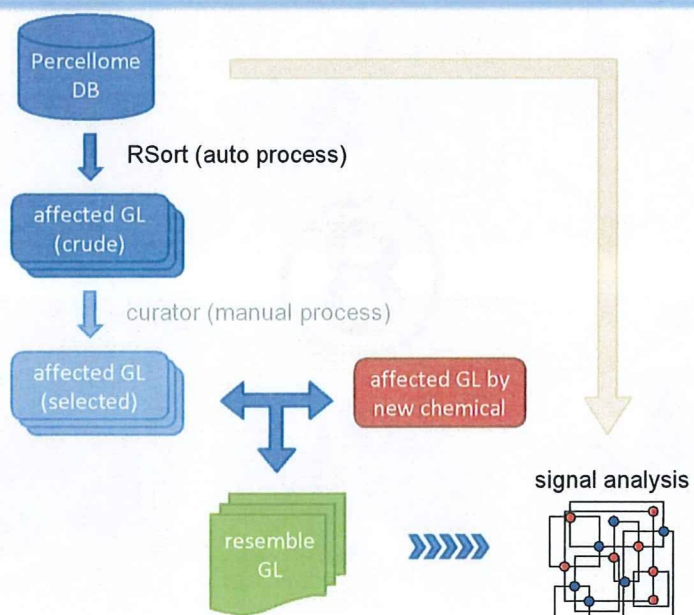
複数の化学物質のクラスタ交差点からの カスケード解析

カスケードの主要素の抽出と、それを元にしたカスケード推測



5

類似反応解析



6

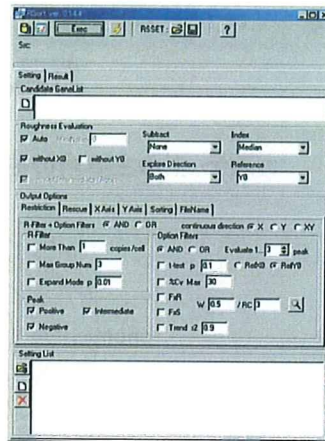
特徴的な遺伝子発現パターンの自動抽出



発現パターンの凹凸の程度・
特徴点の位置(=実験条件)を
評価する。

→ 特定の実験条件において、
強い発現誘導パターンを呈す
る遺伝子を自動抽出する。

H22年度にアルゴリズムを強
化する予定



7

