

厚生労働科学研究費補助金（化学物質リスク研究事業）
総括研究報告書

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 -
（H24-化学-指定-006）

研究代表者 菅野 純

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・部長

研究要旨

本研究は、先行実施された Percellome[®]トキシコゲノミクス研究を基盤に、分子メカニズムに依拠した網羅的毒性評価手法を構築し、毒性予測と評価の一層の迅速化、高精度化を進めることを目的とする。この為に、分子生物学・分子毒性学の専門家とバイオインフォマティクスの専門家の緊密な共同研究体制の下、以下の4研究を実施した。

- (1) 反復暴露実験の分子機序解析による、既存の単回暴露実験データベースからの反復毒性予測の性能評価
- (2) 胎児発生過程におけるマスター遺伝子を基軸とした遺伝子発現ネットワークの網羅的解析
- (3) Percellome 3次元データ等の為に専用解析ソフトウェアの開発研究
- (4) システムトキシコロジー解析基盤の研究開発

(1) では、今年度（平成25年度）は、a) 初年度（平成24年度）実施した四塩化炭素による新型反復暴露実験^(**)のデータ解析を網羅的に進め、遺伝子発現調節に係る情報を抽出する手法を検討した。b) バルプロ酸ナトリウムにて同設計の新型反復暴露実験セットを行い、四塩化炭素と類似した過渡反応 - 基線反応関係を確認しつつ、バルプロ酸ナトリウム固有の所見を得た。バルプロ酸ナトリウムの単回暴露時と比較し、14回反復(1日1回投与)の新型反復暴露時に、過渡反応成分(暴露の都度の変化を示す成分 (Transient Response ; T-Res)) が増加あるいは減少した遺伝子の抽出を行ったところ、肝で単回暴露時に発現した遺伝子の多くが反復暴露時にも発現すること、ただし、基線反応成分 (Baseline Response ; B-Res : 回を重ねるに連れて発現値の基線を徐々に移動させる成分と定義) の変動を示す遺伝子は多く、その増減に同調して、過渡反応が影響を受けることが確認された。すなわち、反復投与の回数と、過渡反応と基線反応の関係を検討したところ、四塩化炭素と同様に多くの遺伝子については、2回反復暴露の時点で基線変動が完成すること

が明らかとなった(例外あり)。そこで投与期間をさらに短くし、1回の先行暴露を行う実験を終了し、現在マイクロアレイデータの測定を実施中である。また、TGP ラットデータ(**)を Percellome 変換し、類似の解析を加えマウスデータと比較した。反復投与により基線反応を示す遺伝子が、どのような転写制御を受けるかを検討する為に、当該遺伝子の転写開始点上流において共通する転写因子や結合配列を *in silico* プロモーター解析により検討し、一定の傾向の存在が示唆された。

(2)では、無処置野生型マウス全胚の胎生 6.25~9.75 日(12 測定)の網羅的トランスクリプトームデータを対象に、初年度、全期間を1周期と仮想して周波数解析を実施したが、経時的に離れた関連遺伝子のうち発現パターンが異なるものについての、高調波(ハーモニクス)成分の扱いに課題が残った。この問題は発現の立ち上がりの勾配に着目すれば解決する可能性を見だし、今年度は、一階微分及び二階微分により発現変動起点及び発現ピークの時点を同定する技術を開発し、Shh 遺伝子シグナルネットワークをモデルに、発生過程のマスター遺伝子である可能性の高い遺伝子を網羅的に抽出する技術開発を進めた。

(3)では、今年度は各班員のデータ解析の効率と精度の向上を優先して平成 26 年度予定の計画を実施した。3次元グラフ表示データからの候補遺伝子抽出プログラム RSort の候補遺伝子の自動抽出精度を向上させると共に、線グラフ表示データから候補遺伝子を抽出するアルゴリズムを開発したほか、初年度の成果(非 Percellome データの絶対量推定ソフトウェア、Percellome データベース公開用 WebAPI など)の拡張・改良を進めた。

(4)では、遺伝子発現ネットワーク推定アルゴリズム及び遺伝子クラスター解析手法 AGCT (A Geometric Clustering Tool) アルゴリズムの性能向上と機能実装を進めた。今年度は、遺伝子制御関係推定では、実際の網羅的トランスクリプトームデータを用いて最適化を進めた。遺伝子クラスター解析手法に関しては、開発中の AGCT の Garuda^(****)対応を順調に進めている。さらに Percellome DB 対応の Garuda ガジェットプログラムを作成・改良するなど、その国際的普及活動を進めた。Garuda Platform は、国際的に非常に高い評価を得ており、普及は予想以上に順調に進んでいる。また、遺伝子上流領域の配列検索手法の導入を加えた。

尚、動物実験の計画及び実施に際しては、科学的及び動物愛護的配慮を十分行い、国立医薬品食品衛生研究所の「動物実験の適正な実施に関する規程」に従い実施した。

(*) mRNA 発現値を細胞 1 個当たりのコピー数として絶対定量する方法。特許取得済み。

(**)全動物に同量の検体を反復投与し、遺伝子発現測定直前の投与時に、溶媒群、

低用量群、中用量群、高用量群に分けて最終投与を一に行う。

(***) TGP はその発足時に当研究者らが Percellome 法を基軸に設計したものであり、その生データから絶対量化や後述の 3 次元波面表示・解析が可能である。

(****) 各種の生物学的研究ソフトウェアの Web 公開型統合プラットフォーム。

<http://www.garuda-alliance.org/>

研究分担者

北野 宏明 特定非営利活動法人
システム・バイオロジー研究
機構 会長

北嶋 聡 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 第 5 室長

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 第 1 室長

A . 研究目的

本研究は、化学物質による生体影響の分子メカニズムに依拠した毒性評価手法の迅速化、高度化、及びその実用の為のインフォマティクス開発を目的とする。即ち、先行研究にて構築済みの延べ 5.8 億遺伝子情報からなる高精度トキシコゲノミクスデータベース (TGDB) に蓄積された単回暴露時データからの毒性ネットワークを解析する技術を基盤に、これらを維持・拡充しつつ、反復暴露時のネットワーク解析、及び、その単回データからの予測評価技術を開発する。ここにインフォマティクス専門家によるシステムトキシコロジーの概念を導入し、網羅的毒性予測評価システムの機能強化と精度向上を図る。

B . 研究方法

(1) 反復暴露影響の分子機序解析による、

既存の単回暴露実験データベースからの反復毒性予測の性能評価【菅野】

B1-1: 試薬及び動物 :

バルプロ酸ナトリウム (Valproic acid sodium salt; 分子量 : 166.19、Cas No.: 1069-66-5 、 純度 98 % 以上、Sigma-Aldrich) について既存データ (単回投与および 14 回反復投与 (1 日 1 回投与)) の解析を進めた。

単回暴露 (0 日間反復暴露後に単回暴露、以降、[0+1] と表記) 時のバルプロ酸ナトリウムの投与量は 0、50、150、500 mg/kg とした。新型反復暴露である 14 回反復暴露 (14 日間反復暴露後に単回暴露、以降、[14+1] と表記) 時のバルプロ酸ナトリウムの 14 回反復投与の用量は全例に対し 100 mg/kg、最終の単回暴露の用量は 0、50、150、500 mg/kg とした。

B1-2: 肝 cRNA ターゲット液の作製 :

マウス組織を採取後すみやかに RNA later (Ambion 社) に 4 で一晩浸漬し、RNase を不活化する。肝は 5mm 径の皮膚生検用トレパンにより 3 ヶ所を打ち抜き、各々別チューブに採取した。その後、RNA 抽出操作までは -80 にて保存した。抽出に当たっては、RNA later を除いた後、RNeasy キット (キアゲン社) に添付される RLT buffer を添加し、ジルコニアビー

ズを加え、シェーカーにて破砕液を調製した。その破砕液の 10 μ L を取り、DNA 定量蛍光試薬 Picogreen を用いて DNA 含量を測定した。DNA 含量に応じ、臓器毎にあらかじめ設定した割合で Spike cocktail (細菌由来の mRNA 5 種類の濃度を変えて混合した溶液) を添加し、TRIZOL により水層を得、RNeasy キットを用いて全 RNA を抽出した。100ng を電気泳動し RNA の純度及び分解の有無を検討した。

全 RNA 5 μ g を取り、アフィメトリクス社のプロトコールに従い、T7 プロモーターが付加したオリゴ dT プライマーを用いて逆転写し cDNA を合成し、得た cDNA をもとに第二鎖を合成し、二本鎖 DNA とした。次に T7 RNA ポリメラーゼ (ENZO 社キット) を用い、ピオチン化 UTP, CTP を共存させつつ cRNA を合成した。cRNA はアフィメトリクス社キットにて精製後、300-500bp となるよう断片化し、GeneChip ターゲット液とした。

B1-3:GeneChip 解析:

GeneChip には Mouse Genome 430 2.0 (マウス) を用いた。ハイブリダイゼーションは 45 にて 18 時間行い、バッファーによる洗浄後、phycoerythrin (PE) ラベルストレプトアビジンにて染色し、専用スキャナーでスキャンしてデータを得た。肝サンプルから得られたデータに、我々が開発した Percellome 手法(遺伝子発現値の絶対化手法)を適用し、網羅的遺伝子発現データを得て、我々が開発した「RSort」ソフトウェアを用いて、自動的なノイズデータの除去を行った。このソフトは、全遺伝子 (probe set: ps) の用量、時間及び遺伝子

の発現コピー数を各軸とした 3 次元グラフの凹凸を評価し、ノイズデータを除去し、生物学的に有意な反応を示したものを順番にリストアップするソフトである。また、既知のシグナルネットワークの探索には、Ingenuity Pathways Analysis (IPA) (Ingenuity Systems Inc.) を用いた。

(2) 胎児発生過程におけるマスター遺伝子を基軸とした遺伝子発現ネットワークの網羅的解析【北嶋】

B2-1:無処置野生型マウス胚・全胚の RNA 採取

C57BL/6CrSlc マウス (日本エスエルシー) を実験に用いた (プラグが確認された日の 15 時を胎生 0.5 日とした)。経時的にサンプリングした各発生ステージのマウス全胚 (ただし卵黄嚢膜は除去した) を 1 腹分ブールし、1% の 2-メルカプトエタノール含有 RLT バッファー (QIAGEN 社) に変性・溶解させ、得た RNA サンプルを用い、マイクロアレイ (Affymetrix GeneChip Mouse Genome 430 2.0) を用いて、網羅的遺伝子発現変動解析を検討することにより、遺伝子発現経時データベース (胎生 6.25 ~ 9.75 日、観測時点: 12 点) を作製した。

B2-2:全胚サンプルの cRNA ターゲット液の作製

全胚サンプルについては、2 段階増幅により cDNA を得た。すなわち、全 RNA の 100ng をとり、T7 プロモーターの付加したオリゴ dT プライマーを用い逆転写し cDNA を得、2 本鎖とした後、T7 RNA ポ

リメラーゼ (Ambion 社) を用いて cRNA を合成 (この段階ではビオチン化塩基は用いない) した (増幅 1 回目)。その cRNA を鋳型に random primer を用いて逆転写して cDNA を得て 2 本鎖とし、T7 RNA ポリメラーゼ (ENZO 社) を用い、ビオチン化 CTP、UTP 共存下で cRNA を合成、断片化した後、GeneChip へのハイブリダイゼーションに供した。

B2-3: Whole mount ISH

当該プローブセット (ps) について GeneChip にて使用している塩基配列を調べ、その配列を基に、ジゴキシゲニンでラベルした dNTP を用いて RNA プローブを作製した。作製した RNA プローブを用いて、固定後プロテナーゼ K 処理したマウス胚とハイブリダイゼーションを行い、検出は抗ジゴキシゲニン抗体、発色は BM purple で行った。

(3) Percellome 3 次元データ等の為の専用解析ソフトウェアの開発研究【相崎】

B3-1: 参照データ

a) 遺伝子発現データ: Percellome プロジェクトで生成したマウスの遺伝子発現データを用いた。データ標準化には Percellome 法を利用した。

b) 遺伝子のアノテーション情報: 米 Affymetrix 社および Gene Ontology Consortium、独 BIOBASE 等が提供している情報を参照した。

c) 原著論文 (J Toxicol Sci. 38(4):643-54) 等のヒューマンキュレーションを経た候補遺伝子リストを参照した。

B3-2: ソフトウェア作成に用いた開発言語及びコンポーネント

開発効率と生成する実行バイナリの実行速度を重視して、Win32/64 開発は RAD (Rapid Application Development) 対応の Delphi ver.7 もしくは XE3 (Object Pascal 言語、USA, Embarcadero Technologies, Inc.) を用いた。Web アプリケーション開発には Delphi もしくは PHP、JavaScript を用いた。1 億件以下の小～中規模のデータベース処理に際して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の DBISAM (USA, Elevate Software, Inc.) を、一般的なグラフ描画には TeeChart (Spain, Steema Software SL) を利用した。また WebAPI の構築には DataSnap (USA, Embarcadero Technologies, Inc.) を用いた。

B3-3: データ解析計算

主たる計算は独自に開発したプログラムにより実施した。検証は必要に応じて Excel (USA Microsoft Corporation) や R 言語 (オープンソース R Development Core Team) で実施し、浮動小数点誤差以上の乖離がないことを確認した。

大規模計算が必要な場合は、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) を装備する研究計算用サーバーシステム (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実施した。

(4) システムトキシコロジー解析基盤の研究開発【北野】

本研究は、システムバイオロジーをトキシ

コロジーに応用する研究であり、従来の手法ではとらえきれなかったネットワークを基盤とした毒性の理解を可能とする。このため、大規模データを利用する手法と大規模文献情報などをいかに集約するかに関する研究、そしてそれらを統合的なプラットフォームに実装し広く利用できるようにする研究によって構成される。この研究は、以下の項目に区分できる：

- (4-1) 大規模データからの遺伝子制御関係推定アルゴリズムの開発と改良
- (4-2) 大規模データからの状態制御遺伝子群推定アルゴリズムの開発と改良
- (4-3) 大規模文献情報などの集約手法と伝達手法の研究
- (4-4) Garuda Platform とその関連ソフトウェアの研究開発

B4-1: 大規模データからの遺伝子制御関係推定アルゴリズムの開発と改良

Percellome や多くの研究で利用されている遺伝子発現データを利用し、遺伝子間の発現変動の相互関係と因果関係ネットワークの構築を可能とする手法の研究を行う。現在、遺伝子発現データからのネットワーク抽出の研究は多く行われているが、その精度は高くない。例えば、広く使われているベイズ統計を利用した方法は、比較的少数の遺伝子に対して大量のデータが存在するときには適切な方法とは言えない。このような状況では、相互情報量 (Mutual Information) を利用する手法の方が適しているが、その計算の具体的な手法により、推定結果の精度などは大きく違ってくる。我々は、既に、複数の計算手法の組み合わせを利用して、これらの現手法より、精

度の高い手法を開発した。本研究では、この手法をさらに改良し、より精度の高いものに改良する。同時に、発現の位相差、時間差、ゲノム情報などを統合し、相関関係ではなく因果関係に関するネットワークを高精度で構築可能な手法の開発を行う。

B4-2: 大規模データからの状態制御遺伝子群推定アルゴリズムの開発と改良

システム毒性学で重要なことは、ネットワークとして記述された生体システムが、毒性物質の影響でどのように「制御」されてしまうかを理解すること (フォワードアナリシス) であり、また、ネットワークを特定の状態に遷移させた原因は、どの遺伝子群への影響によるものであるかを特定できるということ (バックワードアナリシス) である。フォワードアナリシスは、研究項目 4-1 と 4-3 で実現する。この研究項目では、ネットワークを制御することができる遺伝子群の同定を大規模データから行う手法の研究を行う。

B4-3: 大規模文献情報などの集約手法と伝達手法の研究

フォワードアナリシスでは、大規模データのみならず、大量の文献やデータベースなどを総動員して精度の高い分子相互作用ネットワークと遺伝子制御ネットワークを構築する必要がある。すでに我々は、これらに関係するソフトウェアである CellDesigner や Payao などを構築して、このプロセスの効率化と高精度モデルの構築ノウハウを蓄えてきた。この蓄積から、次のステップとして大量に存在する文献などの情報を、知的に集約し、モデル構築を行

う研究手法などを提示し、可能な範囲で自動的にモデルを構築する情報を抽出する手法の開発が必要である。この研究項目では、このプロセスに関する研究とそのコンセプトを実証するソフトウェアの試作を行っている。

B4-4: Garuda Platform に準拠する毒性解析ソフトウェアの開発

本研究で開発される手法を実装するソフトウェアや Percellome などは、国内外の幅広い研究者に利用されることで、トキシコロジーに関連する研究の発展とその成果の利用を促進する必要がある。我々は、国際的なソフトウェアの共通基盤として Garuda Platform を開発している。ここでは、この Garuda Platform をトキシコロジーの分野へと応用することを可能とする一連の開発と普及・啓蒙を行う。

従来の手法では捕え切れなかったネットワークを基盤とした毒性の理解を可能とするために、生体活動をシステムとして把握するシステムバイオロジーをトキシコロジーに応用した「システムトキシコロジー」解析基盤を開発する。その成果をライフサイエンスソフトウェアの国際的共通基盤 (Garuda Platform) を介して国内外の研究者に提供し、国際的な研究活動を促進する。

H25 年度は、大規模データからの状態制御遺伝子群推定アルゴリズムの開発として、状態制御遺伝子群 (ネットワークを制御することができる遺伝子群) を大量データから同定する手法を開発する。

H26 年度において、大規模データからの

遺伝子制御関係抽出及び状態制御遺伝子群推定により、Percellome トキシコゲノミクスデータベース等の大規模データベースを対象としたネットワーク情報の抽出ならびこれらデータの応用の手法を検討する。

倫理面への配慮

動物実験の計画及び実施に際しては、科学的及び動物愛護的配慮を十分行い、所属の研究機関が定める動物実験に関する指針のある場合は、その指針を遵守している。

(国立医薬品食品衛生研究所は国立医薬品食品衛生研究所・動物実験委員会の制定になる国立医薬品食品衛生研究所・動物実験の適正な実施に関する規程 (平成 19 年 4 月版) 及び国立医薬品食品衛生研究所 遺伝子組換え実験安全管理規則の承認を受けて行う。)

C . 研究結果

当初計画に沿って研究を行い、下記の成果を得た。

(1) 反復暴露影響の分子機序解析による、既存の単回暴露実験データベースからの反復毒性予測の性能評価【菅野】

平成 24 年度の四塩化炭素における新型反復暴露実験に引き続き、バルプロ酸ナトリウム (VPA) について同様の新型反復暴露実験を行い ([1+1]、[4+1]は完了。[2+1]は1月末に実施済み、データ採集中)、[0+1]、[1+1]、[4+1]、及び[14+1]の4通りの反復投与期間の長短による過渡反応と基線反応のそれぞれの推移と両反応の関連性の解析を肝について進めた。VPA は、反復暴露後

も過渡反応が完全消失しない遺伝子の数が多い特徴を有していた。その上で、基線反応は基本的に四塩化炭素と同様であり、反復投与により基線が低下した際には、過渡反応のピークが低下する傾向が認められた。反復暴露回数が増すことで発現が抑制されるものの、[14+1]において再び基線が上昇し、それに伴い過渡反応が回復する遺伝子が検出されたほか、基線反応、過渡反応ともに反復暴露に影響されない遺伝子も多数検出された。現在、[2+1]の新型暴露実験を1月下旬に実施した。そのマイクロアレイデータを待って、[0+1]、[1+1]、[2+1]、[4+1]、[14+1]から得られている諸傾向を確認し、分子機序の解析を進める予定である。また、ラット TGP データを Percellome 変換し、類似の解析を加えマウスデータと比較した。

反復投与により基線反応を示す遺伝子が、どのような転写制御を受けるかを検討する為に、当該遺伝子の転写開始点上流において共通する転写因子や結合配列を *in silico* プロモーター解析により検討し、一定の傾向が存在することが示唆された。基線反応低下には、Xbp1 の関与が示唆される経路と、Eif2 の関与が示唆される経路とが区別される可能性がある。前者は単回暴露時に遺伝子発現が誘発され、反復暴露によりそれが減弱する遺伝子群、後者は、単回暴露による遺伝子発現誘導が見られず、反復暴露により基線のみが低下する遺伝子群を有意に多く含むことが判明した。

(2) 胎児発生過程におけるマスター遺伝子を基軸とした遺伝子発現ネットワークの網羅的解析【北嶋】

本研究は、発生過程に於いて重要なマスタ

ー遺伝子は一般的に、この期間中に一過性、即ち単相性の発現を示し、一群の下流遺伝子に連鎖的な発現を誘発し、特定の形質の形態形成を制御するという特徴を利用している。

胎児発生過程における遺伝子発現については、自律的な遺伝子発現が多く認められる為、遺伝子発現シグナルの流れを解析するのに適しているが、活発な細胞増殖や多様な細胞分化が同時並行で起こるため、入手可能な従来技術では詳細な要素分析が困難であった。

そこで初年度は、経時データベースを有する胎生 6.25~9.75 日の遺伝子発現変動を 1 周期と仮想して、周波数解析を実施したが、その結果、既知関連遺伝子の多くは検出可能であったものの、経時的に離れた関連遺伝子のうち発現パターンが異なるものについての、高調波（ハーモニクス）成分の扱いによる効率の良い抽出についての課題が残った。この解決として、発現の立ち上がりの勾配に着目する方法、すなわち一階微分及び二階微分を適用し、既存情報と整合性のある結果を得ることを確認した。

今年度は、3 次元 Spline 補間関数を一階微分及び二階微分することで発現変動起点及び発現ピークの時点を同定する独自の解析ソフト SeekESP_s.exe を用いてデータベース中の全ての遺伝子発現変動につき解析を行った。モデルとして初年度に検討した Shh 遺伝子シグナルネットワークを用いた。抽出パラメータの最適化を実施し、Shh シグナル関連候補遺伝子として 8,306 ps が自動抽出され、この内、目視により生物学的変動と考えられるものとして 648 ps が抽出された。この中には、Shh シグナルと関

連することが報告されている Irs1、Foxf1、Foxc2、Ccmd1 遺伝子等が含まれていた。今後、発現の局在が Shh と十分に近いことを whole mount ISH 情報により確認しつつ、シグナルの流れと、ネットワーク構成要素の遺伝子の発現起点を精査したところ、Shh シグナルを障害する可能性のある Shh 外のネットワークとして、Smad3 及び TGF β 2 の関係するものが見出された。今後、*in silico* プロモーター解析を組み合わせ、分析精度の向上を計る。

(3) Percellome 3次元データ等の為の専用解析ソフトウェアの開発研究【相崎】

今年度は、各班員のデータ解析の効率と精度の向上を優先してH26年度予定であった計画を実施し、候補遺伝子抽出プログラム RSort の改良と、前年度に作成した WebAIP の機能拡張を行った。具体的には以下の通り。

RSort は、2 条件（主に用量、時間）の組み合わせにおける遺伝子発現量を 3 次元平面で表現した遺伝子発現変動パターンの凹凸を評価し、全遺伝子に対応する総数 4~5 万個の発現データから、ノイズや解析不能パターンを除去して、候補遺伝子リスト（通常、総数の 1/10 程度）を短時間（1~2 分）で自動抽出し、データ解析の作業効率を大幅に高めている。ただし、毒性評価上の必要から偽陰性を極力排除するよう最適化されていることから、従来版（ver.0.17）では偽陽性が多発し、自動抽出結果の 2/3~3/4 を占めていたため、一層の改良が求められていた。

そこで偽陽性データの混入を減らすべく、

既存の roughness フィルタと独立した追加フィルタの作成を検討し、3 次元の遺伝子発現変動パターンを構成する各 2 次元切片を対象として、増減パターン組み合わせ数や変曲点数、連続性指標、グラフ密度などを検討して、最適条件を求めた。

その結果、偽陰性データの発生数を数個~200 個程度に抑えつつ、従来版 RSort の抽出結果から数百~2000 個程度の偽陽性データを削減することに成功した。

また、RSort は、時間と用量と発現量の 3 次元表示データを効率的に処理すべく開発されてきたものなので、経時変化データ等の線グラフ表示データの処理については精度が低下する傾向があった。そこで、分担研究（2）が対象とする個体発生の経時変動データをサンプルとし、効率的な候補遺伝子抽出プログラム SeekESP_S.exe を作成した。

経時変化を追跡する実験から得られた線グラフデータに於いて、データ点が十分に多い場合は、各遺伝子発現変動のピーク（山頂や谷底の位置）を自動検出して前後関係を中心に解析するよりも、遺伝子発現の起点（発現変動起点）を自動検出し解析した方がシグナルネットワーク描出の制度が向上することが生物学的に想定される。

そこで、各群の平均値をなめらかに繋ぐ 3 次スプライン曲線を求め、その一階微分曲線と二階微分曲線、三階微分曲線を算出して、各遺伝子の発現変動起点を推定するアルゴリズムを作成した。

実証のためのプログラムを作成し、性能試験を行ったところ、分担研究（2）の成果で詳細を記載したとおり、既知のシグナルネットワークに対応する候補遺伝子リスト

を適切に抽出する事に成功した。

初年度に Percellome 公開用 Web サーバーに実装した REST (現時点で最も普及しているソフトウェアアーキテクチャの1つ) WebAPI については、当該システムを利用したオンライン解析能力を向上させるべく、機能を拡張し、RSort で抽出した候補遺伝子リストをインターネット経由で提供するなどの機能強化を行った。

(4) システムトキシコロジー解析基盤の研究開発【北野】

遺伝子制御関係推定アルゴリズムでは、複数のアルゴリズムの併用を行う方式を開発し、現状での最も優れた方式と同等の結果を安定的に維持することに成功している。さらに、より一層の大規模化に備え、この性能を維持しつつ、情報エントロピーを利用しながら計算量を減少させる手法の開発に成功した。また、これらを Inference cloud という Web-Service の形で実装し、オンライン上で利用可能な形態にした。

状態制御遺伝子群推定アルゴリズムに関しては、以前から開発している A Gene geometric Clustering Tool (AGCT) で Percellome データの特徴を考慮に入れた解析を容易に実行できるようにするための開発を行った。具体的には、全データからのドーズ選択(細胞選択)、時計遺伝子効果削除、スペクトラルクラスタリングや主成分分析の表示の上で6つのアルゴリズムによるクラスタリング等の機能を追加した。これらの機能では Percellome データのニーズに即し、事前データ削除を行わずに全データを計算した後、必要な部分だけをユーザ

が取り出せるようになっている。更に大規模データをより短時間で計算できるようにした。

制御遺伝子群の同定についての予備検討では、幾つかの事例において、実験的に確認されている遺伝子群と合致することなどを確認した。

さらに、Garuda Platform の開発者向けパッケージ 6.2 版の配布開始(2013年12月)にあわせ、ネットワークの推定や AGCT など一連のアプリケーションを Garuda に準拠し、毒性学に利用できるように対応したほか、オルソログ制御領域分析ソフトウェア SHOE (Sequence Homology in Higher Eukaryotes) に関してはオンラインで Garuda 経由でサービス提供できるように準備を進めた。また Percellome DB の Garuda 上への実装に協力し、Percellome DB の国際的普及への準備を行った。

Garuda Platform は、国際的に非常に高い評価を得ており、普及は予想以上に順調に進んでいる。実際に、武田薬品工業株式会社が、2014年度からの導入を決めるなど、普及が加速している。

D. 考察

反復暴露影響の分子機序解析による、既存の単回暴露実験データベースからの反復毒性予測の性能評価においては、肝及び肺における四塩化炭素の新型反復暴露実験により、単回暴露時に発現変動した遺伝子のほぼ全てについて、基線反応成分(暴露回数を重ねるに連れて発現値の基線(ベースライン)が徐々に変動する反応成分(Baseline Response; B-Res))は、過渡反応成分(単回暴露時の2, 4, 8, 24時間のうちに発現

が変動する速い変化の成分 (Transient Response ; T-Res))が増加する場合は増加、減弱する場合は減少することを見いだした。増加する事例があることから、反復投与による代謝誘導による化学物質の分解促進では説明できない事象であると考えられた。むしろ、この過渡反応と基線反応の連関性に関する知見は、生物学的・毒性学的に新規性が高くエピジェネティクスに関わる分子機序の関与が示唆されることから、これを明らかにすることは、反復毒性の分子毒性学的理解の促進、及び、単回暴露実験データベースからの反復毒性予測法を開発するにあたり重要と考えられる。

今年度は抗けいれん薬として使用されているバルプロ酸ナトリウムについて同様の実験を行ったところ、四塩化炭素で見られた所見と基本的に同傾向であるものの、発現抑制を受けない遺伝子の数が多いという特徴が見られた。この特徴は、治療薬として反復投与されている化学物質の特性として、長期使用によっても薬効が維持されるために必要なものであるとの見方が可能性であるものとして注目される。また、TGPのラットにより解析したデータから、種間に共通な分子機構が示唆された点も外挿性の点から注目される。

今後、慢性毒性の理解のための過渡反応と基線反応の分子機序解明に、四塩化炭素とバルプロ酸ナトリウムの差異を利用した転写制御領域の解析等を分担研究者の協力のもとに推進する。

胎児発生過程におけるマスター遺伝子を基軸とした遺伝子発現ネットワークの網羅的解析においては、初年度は、周波数解析

(高速フーリエ変換)の性能と限界、微分関数の利用の有用性、及び胎児発生過程における発現解析の要点(例えば転写制御の時系列評価では、発現ピークを与える時点よりも、発現が開始する時点に注目しなければならないこと、など)を明らかにした。今年度は、初年度の結果を受け、微分関数を利用し、発現変動起点及び発現ピークの時点を解析する手法の開発を進め、モデルシグナルネットワークにつき検討し、網羅的・効率的に発生過程のマスター遺伝子である可能性の高い遺伝子の抽出が出来る事を示唆した。このノウハウは、胎児に限らず、一般的な毒性標的臓器の経時的な遺伝子発現データの変動解析にも適用可能であり有用であると考えられる。

Percellome 3次元データ等の為の専用解析ソフトウェアの開発研究においては、本年度、自動的にノイズデータを除外し、研究者が操作する候補遺伝子数を減らすことで解析効率の向上に大きな役割を果たしている RSort の性能を一層向上させた。今までは、生物学的な意味のある遺伝子発現データを除外してしまわないように、若干の余裕をもってノイズデータを除去していたことから、研究者が操作する段階で、若干のノイズデータが残存していた。今年度は、この残存ノイズデータを更に高精度に除去する改良を加えた。この改良は、個々のデータの網羅的遺伝子解析作業の効率を高めるのみならず、複数の候補遺伝子リストの和集合計算を行う必要性が生じる大規模解析、例えば複数の化学物質毎に特異的に発現する遺伝子群の比較解析時に、その効果が明瞭に現れる重要な技術である。

今まで、時間と用量と発現量の 3 次元表示データを中心に開発してきた解析ソフトウェアに加え、経時変化データ等の線グラフ表示データに特化した候補遺伝子抽出アルゴリズムによるソフトウェアを用意することが出来た。これにより 3 次元表示データに匹敵する解析精度や作業効率の向上が見込まれ、この一部の技術は 3 次元表示データ解析にも還元可能であり、相互に統合的解析が可能となる。

システムトキシコロジー解析基盤の研究開発では、毒性学にシステムバイオロジーを適用する一連の手法が具体的な形となってきたおり、複数のソフトウェアの連動が実現されつつある。

Garuda は、今後極めて重要なプロジェクトに成長すると思われる。ソフトウェアを開発する側の研究者のみならず、ユーザとなる製薬企業や生命科学の基礎研究に携わるものにとっても、重要な Platform であるとの認識が広まっている。この Garuda Platform にシステムトキシコロジーの研究を可能とするソフトウェア群を実現する事は、今後の研究の展開にとって非常に重要であると考えられる。

E . 結論

本研究は、ほぼ計画通りに進捗した。

初年度に実施した、四塩化炭素による新型反復暴露解析では、単回暴露時の過渡反応成分と反復暴露時の基線反応成分の基本的な関連性を見だし、今年度、同様の解析をバルブプロ酸ナトリウムで実施し、大筋での同様の所見、及び、化学物質特有の所見を得た。この過渡反応と基線反応に関する

知見は生物学的・毒性学的に新規性が高くエビジェネティクス等の機序の関与が示唆されることから、これを明らかにすることは、反復毒性の分子毒性学的理解の促進、及び、単回暴露実験データベースからの反復毒性予測法を開発するにあたり重要と考えられる。

来年度は、クロフィブレートについて同プロトコールの実験を実施し、単回暴露（急性）の毒性ネットワーク解析結果から反復暴露による生体影響を予測評価する技術の開発を推進する。

胎児発生過程におけるマスター遺伝子を基軸とした遺伝子発現ネットワークの網羅的解析については発現変動起点及び発現ピークの時点に着目する微分解析手法の開発と解析を進め、網羅的・効率的に発生過程のマスター遺伝子である可能性の高い遺伝子を抽出する事が出来た。今後は、実際に網羅的の獲得と、信頼性と効率の向上を目指す。

Percellome 3 次元データ等の為の専用解析ソフトウェアの開発研究では、初年度成果の異種動物由来のデータ統合技術や絶対量化されていない非 Percellome データの絶対量推定技術、及び今年度成果の RSort 技術の改良により、網羅的遺伝子発現解析の効率向上を押し進めている。

また初年度から公開している Percellome データベースの WebAPI の改良により、外部のオンライン解析ソフトウェアからも、RSort による候補遺伝子情報にアクセス出来るようになり、解析効率の向上が期待される。

システムトキシコロジー解析基盤の研究開発も順調に推移した。

遺伝子制御関係推定アルゴリズム及び状態制御遺伝子群推定アルゴリズムの性能向上や機能拡張が進んでおり、システムバイオロジーを毒性学に応用しようとする目標に対して、順調に展開している。すでに幾つかのソフトウェアが、Garuda Platform上に実装され連動するなどの成果をあげている。今後、これらのソフトウェアを利用した具体的適用例を作る事や普及などの展開にも力を入れたい。

F . 健康危険情報

特になし

G. 研究発表

(1) 論文発表

Kanno J, Aisaki K, Igarashi K, Kitajima S, Matsuda N, Morita K, Tsuji M, Moriyama N, Furukawa Y, Otsuka M, Tachihara E, Nakatsu N, Kodama Y. (2013) Oral administration of pentachlorophenol induces interferon signaling mRNAs in C57BL/6 male mouse liver. *J Toxicol Sci.* 38(4):643-54.

Si Y, Inoue K, Igarashi K, Kanno J, Imai Y. (2013) Autoimmune regulator, Aire, is a novel regulator of chondrocyte differentiation. *Biochem Biophys Res Commun.* 437(4):579-84.

Fujimoto N, Takagi A, Kanno J. (2013) Neonatal exposure to 2,3,7,8-tetrachlorodibenzo-p-dioxin increases the mRNA expression of prostatic proteins in C57BL mice. *J*

Toxicol Sci. 38(2):279-83.

Hase T, Ghosh S, Yamanaka R, Kitano H. (2013) Harnessing Diversity towards the Reconstructing of Large Scale Gene Regulatory Networks. *PLOS Computational Biology.* 9(11): e1003361, doi:10.1371/journal.pcbi.1003361.

Matsuoka Y, Matsumae H, Katoh M, Eisfeld AJ, Neumann G, Hase T, Ghosh S, Shoemaker JE, Tiago Lopes, Watanabe T, Watanabe S, Fukuyama S, Kitano H, Kawaoka Y. (2013) A comprehensive map of the influenza A virus replication cycle. *BMC Systems Biology.* 7, 97, doi:10.1186/1752-0509-7-97.

Yamashita F, Sasa Y, Yoshida S, Hisaka A, Asai Y, Kitano H, Hashida M, Suzuki H. (2013) Modeling of Rifampicin-Induced CYP3A4 Activation Dynamics for the Prediction of Clinical Drug-Drug Interactions from In Vitro Data. *PLOS ONE.* 8, 9, e70330.

Fujita K, Ostaszewski M, Matsuoka Y, Ghosh S, Glaab E, Trefois C, Crespo I, Perumal TM, Jurkowski W, Antony PMA, Diederich N, Buttini M, Kodama A, Satagopam VP, Eifes S, Del Sol A, Schneider R, Kitano H, Balling R. (2014) Integrating Pathways of Parkinson's Disease in a Molecular

Interaction Map. *Molecular Neurobiology*. 49:88-102. DOI 10.1007/s12035-013-8489-4.

Naito T, Yatsunami A, Kaji N, Ando T, Sato, K, Moriya H, Kitano H, Yasui T, Tokeshi M, Baba Y. (2013) Parallel Real-Time PCR on a Chip for Genetic Tug-of-War (gTOW) Method. *Analytical Sciences*. 29(3): 367-71.

Makanae K, Kintaka R, Makino T, Kitano H, Moriya H. (2013) Identification of dosage-sensitive genes in *Saccharomyces cerevisiae* using the genetic tug-of-war method. *Genome Research*. 23, 300-311.

Schaefer MH, Lopes T, Mah N, Shoemaker JE, Matsuoka Y, Fontaine JF, Louis-Jeune C, Eisfeld AJ, Neumann G, Perez-Iratxeta C, Kawaoka Y, Kitano H, Andrade-Navarro MA. (2013) Adding Protein Context to the Human Protein-Protein Interaction Network to Reveal Meaningful Interactions. *PLOS Computational Biology*. 9, 1.

Polouliakh N (2013) Reprogramming resistant genes: in-depth comparison of gene expressions among iPS, ES and somatic cells. *Frontier of Physiology* doi:10.3389/fphys.2013.00007.

(2) 学会発表

Jun Kanno, Progress in Japanese Percellome Project and incorporation of TGP data, 11th International Conference of Environment Mutagens (11th ICEM), (2013.11.4), Foz do Iguassu, Brazil, invited

Jun Kanno, Percellome Toxicogenomics, A Quantitative and Comprehensive Approach for Basic and Applied Toxicology. ICT2013 The XIII International Congress of Toxicology, (2013.7.2), Seoul, Korea, distinguished lecture

菅野 純、"Percellome Project ケミカルバイオロジ - の視点からのトキシコゲノミクス—Percellome Project の進捗とその応用性—"、第 40 回日本毒性学会学術年会、2013 年 6 月 18 日、千葉、シンポジウム

菅野 純、網羅的絶対量化遺伝子発現解析による外来物質生体影響の動的ネットワークマーカー描出：Percellome Project、第 102 回日本病理学会総会、2013 年 6 月 7 日、札幌、シンポジウム

Jun Kanno, Percellome toxicogenomics for more comprehensive and quantitative toxicology-extending the analysis among different organs and different species, the workshop on Moving Forward in Human Cancer Risk Assessment in the Genomics Era 2.0, (2013.5.16), Paris, France, invited

Jun Kanno, Percellome Toxicogenomics application to Sick House Syndrome-level inhalation toxicity , the Meeting of the Extended Advisory Group on Molecular Screening, (2013.5.15), Paris, France, invited

北嶋 聡,小川幸男, 大西 誠 ,相磯成敏,相崎健一,五十嵐勝秀, 高橋祐次, 菅野 純
シックハウス症候群レベルの極低濃度暴露の際の海馬における Percellome 法による吸入トキシコゲノミクス
第 40 回日本毒性学会学術年会
(2013.6.18)

Kitajima S, Aisaki K, Igarashi K, Kanno J. Application of Percellome Toxicogenomics approach to food safety: A flavor, estragole appears to be a PPAR-alpha agonist.
The XIII International Congress of Toxicology 2013 (ICT 2013)] (2013.7.3.), Seoul, Korea

Tanemura K, Igarashi K, Furukawa Y, Otsuka M, Aisaki K, Kitajima S, Sato E, Kanno J. Delayed Effects on CNS Induced by Disturbance of Neural Activity during Development – Behavioral Impairment in Male Adult Mice Induced by Postnatal Oral Intake of Acephate.
The XIII International Congress of Toxicology 2013 (ICT 2013)] (2013.7.3.), Seoul, Korea

Kitano, H. Systems drug discovery and Neuro-degenerative diseases. CHDI's 8th Annual Huntington's Disease therapeutics Conference: A Forum for Drug Discovery & Development, Molino Stucky Hilton Hotel, Venice, Italy, Apr. 9, 2013. (invited)

Kitano, H. Systems Biomedicine for Drug Discovery and Personalized Healthcare. Seminar at Institute for Infocomm Research(I2R), Institute for Infocomm Research(I2R), Singapore, Apr. 12, 2013. (invited)

Kitano, H. Systems Biology: past, present, future. Talk at AIBN at the University of Queensland, Australian Institute for Bioengineering and Nanotechnology, University of Queensland, Brisbane, Australia, May 9, 2013. (invited)

Kitano, H. Systems biology in the context of systems and precision engineering. Talk at Australian Institute of Marine Science(AIMS), AIMS, Townsville, Australia, May 10. 2013. (invited)

北野宏明. システムトキシコロジープラットフォーム. 第 40 回日本毒性学会学術年会 シンポジウム 7「毒性オミクス」, 幕張メッセ, 千葉, June18, 2013. (invited)
Kitano, H. Systems biology and

systems biomedicine: integrative systems sciences and biomedical sciences. 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Osaka International Convention Center, Osaka, July 7, 2013. (invited keynote)

Kitano, H. Systems Biology Platform for Drug Discovery. Talk at Imperial College, London, UK, July 30, 2013. (invited)

Kitano, H. Systems Biology and Software Platform. Nature Publishing Group Asian Academic Publishing team meeting, Sheraton Miyako Hotel, Sep. 6, 2013. (invited)

Kitano, H. Garuda Platform: An integrated inter-operability for biomedical software and data resources. COMBINE 2013, Institut Curie, Paris, France, Sep. 16, 2013. (invited)

北野宏明. システムバイオロジーを活用したバイオマーカー評価. 日経バイオテク/日経バイオテク ONLINE プロフェッショナルセミナー「実例から学ぶクリニカルバイオマーカー時代の到来」, コクヨホール, Oct. 2, 2013. (invited)

Kitano, H. Systems Drug Discovery and Systems Biomedicine. Systems Biology Seminar co-hosted by Garvan Institute and UNSW, Garvan Institute, Sydney, Oct. 11, 2013. (invited)

北野宏明. システムバイオロジーを活用

したバイオマーカー評価. 講演@東北メディカルメガバンク機構. Nov. 1, 2013. (invited)

北野宏明. Systems drug discovery and its software platform システム創薬とソフトウェアプラットフォーム. セミナー@武田薬品工業株式会社 湘南研究所, Dec. 3, 2013. (invited)

北野宏明. 多階層生体機能学: モデル創薬から臨床推論サポートまで. 金沢医科大学 第 15 回 KMU 研究推進セミナー, 金沢医科大学病院, Dec.4, 2013. (invited)

北野宏明. Current Status and Perspectives on Systems Drug Discovery. システム創薬の現状と今後の展望. アスピオファーマ社内セミナー, アスピオファーマ株式会社, Dec. 24, 2013. (invited)

H . 知的所有権の取得状況

1 . 特許取得

特許第 5177712 号、2013 年 1 月 18 日登録、特許権者：国立医薬品食品衛生研究所、NTT データ、発明者：菅野純、相崎健一ら、「競合的ハイブリダイゼーションにおける遺伝子データの補正方法及び補正装置」

2 . 実用新案登録

なし

3 . その他

なし