

201428016B

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—

(H24-化学-指定-006)

平成 24 年度～26 年度

総合研究報告書

研究代表者 菅野 純

平成 27(2015)年 3 月

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—

(H24-化学-一般-006)

平成 24 年度～26 年度

総合研究報告書

研究代表者 菅野 純

平成 27(2015)年 3 月

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—

(H24-化学-一般-006)

平成 24 年度～26 年度 総合研究報告書

研究代表者 菅野 純

平成 27(2015)年 3 月

目 次

I. 総合研究報告書	
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	
ー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のインフォマティクス技術開発ー	
菅野 純	..... 1
(資料 1) 総合報告資料	..... 27
(資料 2) Percellome Database -PercellomeExplorer 解析結果-	..... 61
II. 研究成果の刊行に関する一覧表	..... 229
III. 研究成果の刊行物・別刷	..... 235

別添 3

# I . 総合研究報告書

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—  
(H24-化学-指定-006)  
総合研究報告書

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<sup>(\*)</sup> 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、新型反復暴露<sup>(\*\*)</sup> が 2 種類の反応を遺伝子発現に及ぼすことが判明した。それは、過渡反応：暴露の都度に概ね 24 時間以内に収まる速い変化、及び、基線反応：暴露を重ねるに連れ、日を追って発現値の基線が徐々に移動する緩徐な反応、である。平成 24 年度は四塩化炭素の単回暴露と比較し 14 日間新型反復暴露で過渡反応の振幅が増減した遺伝子では、それらの基線反応も同様の増減傾向を示すことを見出し、さらに 2 及び 4 日間の新型反復暴露でも過渡反応と基線反応との間に同様の関係がある事を明らかにした。

平成 25 年度はバルプロ酸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た。ただし四塩化炭素では過渡反応が急速に消失する遺伝子が多いのに対し、バルプロ酸ではそのような傾向は弱いという差異が認められた。平成 26 年度はクロフィブレート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基線反応関係を確認し、その

上流に共通の分子機構が働いている強い可能性が示された。

(2) では、平成 24 年度に無処置野生型マウス全胚の胎生 6.25~9.75 日 (12 時点) の網羅的トランスクリプトームデータを対象として、全期間を 1 周期と仮想して周波数解析を実施した結果、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、高調波 (ハーモニクス) 成分の扱いによる効率の良い抽出法についての課題が残った。この代案として、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、発現値の微分解析を実施し、その有用性、及び転写制御の時系列評価では、発現ピークよりも、発現が開始する時点が重要であること等を確認した。平成 25 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計算技術の開発を進め、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が出来る事を示し、その過程において、Shh の標的遺伝子である Gli 遺伝子の発現は、Shh の受容体を介さずに TGF  $\beta$ 2-Smad3 を介して制御されるという、発生初期における非標準的な(non-canonical) 新たな Shh シグナル制御系を見出した可能性を報告した。平成 26 年度は、全ての発現変動起点について候補遺伝子を抽出・分類し、網羅的に発生初期過程に絡むシグナルネットワークを描出する事を検討した。

(3) では、平成 24 年度に医薬基盤研トキシコゲノミクスプロジェクト (TGP) のデータを統合し解析するためのソフトウェア、Percellome データベースを利用した非 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 などの開発を行った。平成 25 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、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た。平成 26 年度は、マウスのみを対象としていた非 Percellome データの絶対量推定を、参照データベースに TGP の一部溶媒群データ (\*\*\*) を追加・拡充することにより、ラットにも拡大したほか、実験間の共通変動遺伝子を高速抽出するプログラム PercellomeExplorer の改良を進め、各化学物質に特異的な生体反応要素のハイスループット抽出解析を実現した。

(4) では、平成 24 年度に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と機能実装を進めた。平成 25 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 (A Geometric Clustering Tool) の Garuda (\*\*\*) 対応を進めた。平成 26 年度は遺伝子制御関係推定システムを Web-service の Inference cloud として提供できるように開発を行った。また、クラスタリングソフトウェア AGCT の Percellome データへの最適化や計算規模及び速度の向上を実施し、高速化を達成した。さらに Percellome データベースの、国際的なソフトウェアの共通基盤である Garuda Platform 上への実装を進め、そのソフトウェアを Garuda platform の一般公開バージョンとして配

布を開始した。

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 4415079 号。

(\*\*) 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回行う。

(\*\*\*) 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

ミクスデータベース (TGDB) と単回暴露時の毒性ネットワーク解析技術を基盤に、これらを維持・拡充しつつ、反復暴露のネットワーク解析、及び、そ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】

#### 研究協力者

Natalia Polouliakh ソニーコンピュータ  
サイエンス研究所

#### B1-1: 試薬及び動物 :

平成 24 年度は、四塩化炭素 (Carbon tetrachloride; 分子量 : 153.8、Cas No. : 56-23-5、純度 99.8%、Wako) について既存データの解析を進めた。また新たに 12 週齢の雄性 C57BL/6J マウス (日本チャールスリバー) に 4 用量 (溶媒 : コーンオイル) の四塩化炭素を、金属ゾンデを用いて強制経口投与を行い、経時的 (投与 2、4、8 及

#### 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



び24時間後)に肝及び肺を採取した。

単回暴露(0日間反復暴露後に単回暴露、以降、[0+1]と表記)時の四塩化炭素の投与量は0、0.7、2、7 mg/kgとした。新型反復暴露である14日間反復暴露(14日間反復暴露後に単回暴露、以降、[14+1]と表記)時の四塩化炭素の投与量は全例に対し5 mg/kg、単回暴露は0、0.7、2、7 mg/kgとした。

平成25年度は、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99.8%、Sigma-Aldrich)について既存データ(単回投与および14日反復投与)の解析を進めた。また新たに12週齢の雄性C57BL/6Jマウス(日本チャールスリバー)に4用量(溶媒: 0.5%メチルセルロース水溶液)のバルプロ酸ナトリウムを、金属ゾンデを用いて強制経口投与を行い、経時的(投与2、4、8及び24時間後)に肝を採取した。

単回暴露([0+1])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日間反復暴露([14+1])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。

平成26年度はクロフィブレート(Clofibrate; 分子量: 242.7、Cas No.: 637-07-0、純度98.0%、Enzo)について既存データの解析を進めた。また新たに12週齢の雄性C57BL/6Jマウス(日本チャールスリバー)に4用量(溶媒: 0.1%DMSO含有0.5%メチルセルロース水溶液)のクロフィブレートを、金属ゾンデを用いて強制経口投与を行い、経時的(投与2、4、8及び

24時間後)に肝を採取した。

単回暴露([0+1])時のクロフィブレートの投与量は0、10、30、100 mg/kgとした。新型反復暴露である14日間反復暴露([14+1])時のクロフィブレートの投与量は全例に対し70 mg/kg、単回暴露は0、10、30、100 mg/kgとした。

B1-2:Total RNA の分離精製:

マウス組織を採取後すみやかにRNA later (Ambion社)に4°Cで一晩浸漬し、RNaseを不活化する。肝は5mm径の生検トレパンにより3ヶ所を各々別チューブに採取した。その後、RNA抽出操作までは-80°Cにて保存した。抽出に当たっては、RNA laterを除いた後、RN easy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用いて破砕液を調製した。得られた破砕液の10 µL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 (Bacillus由来RNA5種類の濃度を変えて混合した溶液)を添加し、TRIZOL (ライフテクノロジーズ・ジャパン社)により水層を得、RN easy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。

B1-3:GeneChip 解析:

全RNA 5 µg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、ビオチン化 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。GeneChip には Mouse Genome 430 2.0 を用いた。ハイブリダイゼーションは 45°C 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。得られた肝サンプルについて、我々が開発した Percellome 手法 (遺伝子発現値の絶対化手法) を適用した網羅的遺伝子発現解析を行った。遺伝子発現データを、我々が開発した「RSort」を用いて、網羅的に解析した。このソフトは、各遺伝子 (probe set: ps) につき、用量、経時変化及び遺伝子の発現コピー数を各軸とした 3 次元グラフにおいて、発現を表す平面の凹凸を評価し、全ての ps を生物学的に有意な順に並び替えるソフトである。また、シグナルネットワークの探索は、Ingenuity Pathways Analysis (IPA) (Ingenuity Systems Inc.) を用いて検討した。

## (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】

### 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ステージのマウス胚 (全胚、ただし卵黄囊膜は除去した) を 1 腹分プールした 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、TIME POINT: 12 点) を作製した。マウス胚は、直接、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・溶解させた。

### B2-2:全胚サンプルの cDNA 合成

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、及び TGP プロジェクトで生成したラットの遺伝子発現データを用いた。データ標準化(絶対量化)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 及び Gene Ontology Consortium、独 BIOBASE 社等が提供している情報を参照した。

#### 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、これらの現存手法より、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。

#### 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う。

#### 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。

### 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 年度は、新型反復暴露実験により、四塩化炭素の肝に対する反復影響を解析した。14 日間の先行反復投与後の単回暴露（[14+1]と表記）では、単回暴露のみの場合（[0+1]と表記）に比べて、反応が減弱ないし消失する遺伝子が多数、反応が増強する遺伝子が少数認められた。また、その際

に溶媒対照群の発現値から読み取れる基線が、反応が減弱した遺伝子では低下し、反応が増強した遺伝子では上昇する傾向が認められた。そこで、最終投与後 2、4、8、24 時間の間にほぼ終結する速い変動を過渡反応（Transient Response）、反復投与により徐々に引き起こされるベースラインの変動を基線反応（Baseline Response）と定義した。これらの反応がいつ完成するかを明らかにするために新たに、1 日間、2 日間、或いは 4 日間の先行反復暴露を実施した翌日に単回暴露を実施した新型反復暴露実験（それぞれ[1+1]、[2+1]、[4+1]と表記）の結果を加え、[0+1]、[1+1]、[2+1]、[4+1]、及び[14+1]の 5 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を解析した。その結果、単回暴露 [0+1]時に発現変動した遺伝子の多くが反復暴露後に、反応減弱ないし消失を示し、基線の低下傾向を伴っていた。また、反応が増強する遺伝子が少数認められ、その際には基線が上昇する傾向が確認された。これらの変化は、多くの遺伝子で[1+1]において完成していた。単回暴露 [0+1]時に発現変動しない遺伝子の多くに、反復暴露による基線の低下を認め、それらは、[1+1]、[2+1]、[4+1]、及び[14+1]の何れにおいても基線の低下のみであり、過渡反応は見られなかった。肺も、ほぼ同様の傾向を示すが、基線が上昇した遺伝子の発現増強が乏しい傾向が認められた。

平成 25 年度は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、[0+1]、[1+1]、[1+2]、[4+1]、及び[14+1]の 5 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、基線反応の関連傾向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。一旦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。また、反復暴露後に異なった化学物質を単回暴露した新型反復毒性実験の結果の解析も進め、ネットワーク交叉に関する情報を得た。

平成 26 年度はクロフィブレートについて同様の新型反復暴露実験を行い、[0+1]、[1+1]、[2+1]、[4+1]、及び[14+1]の 5 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。

過渡反応の有無(最終日の投与の影響の有無)に関わらず、約 4,000 の遺伝子に、有意な基線反応が認められ、うち 1,700 遺伝子に 1 回から 14 回反復に亘り基線の低下が認められた。この遺伝子群は既知遺伝子情報 (IPA) により、四塩化炭素と共通する特定の経路を構成する遺伝子を高率に含んでいることが判明した。また、過渡反応が基線反応の影響を受けるものは比較的少数であり、受けないものが多いことが判明した。

これら 3 化学物質の解析結果を総合し、基線反応と過渡反応の組み合わせによる遺伝子リストを完成させ、その経時的な変動を調節する上下流の遺伝子発現ネットワークを明らかにし、慢性毒性の分子背景の解明

を進めた。その結果、単回暴露時の遺伝子発現ネットワークからの基線反応の出現の推測を行うには、3 化合物からでは件数がまだ不足していることが示唆され、①今後、過渡反応と基線反応の関係を捉えるのに [4+1] の実験を実施すれば十分である可能性が高いこと、②現段階で、基線反応の強い抑制を示す遺伝子群の発現を支配する上流遺伝子群が 3 物質に共通していることが明らかになった事、③その更なる上流の候補遺伝子が単回投与から数個示唆された事、④新型反復暴露実験による過渡反応遺伝子群と基線反応遺伝子群の重なり具合が、化学物質固有であることから、この重なりの内容が反復毒性の予測に重要である可能性が示唆された事、が成果としてあげられる。

## (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、発生過程にお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ことを利用している。

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、従来技術では要素分析が困難であった。

そこで平成 24 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。各時点 4 サンプルのデータを 1 サン

プルずつの 4 周期に展開し、その周期性を持って、それがマスター遺伝子候補であるか否かの程度を推定するパラメータとなり得るかを試みた。周期性を検討する為に、データベース中の全ての遺伝子発現変動を、フーリエ変換を利用する独自の解析ソフト MF Wave analyzer を用いてフーリエ変換した上で、その波長分布についてピアソン相関等の解析を行った。モデルとして、既知の Shh (四肢や、脳・脊髄正中線構造などのデザイン形成に関与) 及び Nkx2-5 遺伝子 (心筋前駆細胞マーカーであり、心臓発生過程に関与) が関与するシグナルネットワークを選択し、本解析手法の性能評価を検討した。その結果、既知関連遺伝子の多くは検出可能であったが、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、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。この代案として、発現の立ち上がりの勾配に着目する方策の可能性を見だし、一階微分及び二階微分を暫定的に適用したところ、既存情報と整合性のある結果を得たことから、その有用性が示唆された。

そこで平成 25 年度は、一階微分及び二階微分することで発現変動基点及び発現ピークの時点を同定する技術を開発し解析を行うこととした。3 次元 Spline 補間とその微分関数を解析に利用する独自の解析ソフト SeekESP\_S 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に着目し、条件設定及び抽出パラメータの最適化を検討後に解析したところ、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、Ccnd1 遺伝子等が含まれ、また Tgf $\beta$  シグナルが、Shh の受容体を介さずに、Shh の標的遺伝子である Gli 遺伝子の発現を制御する可能性を示唆する結果を得、発生初期における非標準的な(non-canonical)新たな Shh シグナル制御系を見いだした可能性を報告した。

平成 26 年度は、モデル遺伝子のシグナルネットワークに着目した平成 25 年度の成果を踏まえ、この発現変動基点及び発現ピークの時点を同定する技術を利用しつつ、全ての発現変動起点について候補遺伝子を抽出・分類し、網羅的に発生初期過程に絡むシグナルネットワークを描出する事を検討した。無処置野生型マウス全胚の胎生 6.25~9.75 日 (12 時点) の網羅的トランスクリプトームデータを対象に、全ての遺伝子発現変動につき、解析ソフト SeekESP\_s.exe を用いて、発現変動起点ごとに自動抽出後 (発現変動起点における発現コピー数が 0.3 以上のものを選択)、目視により生物学的変動と考えられる 2,218 ps を抽出した。この目視の過程で、2,218 ps の内、胚採取開始時点である胎生 6.25 日において、多くの遺伝子 (2,166 ps) の発現が認められる事を見いだした。また、発現のピークが単一の 1 峰性のものと、ふたつである 2 峰性のものに分類できる事が明らかとなった。2 峰性の発現変動を示す遺伝子の場合、はじめの発現変動と次の発現変動とが重なり、境目が不明確となり、次の発現変動起点がどの時点となるのか定

めることはできなかった。そこで本解析では、2 峰性の発現変動を示す遺伝子は選択せず、1 峰性の発現変動を示す遺伝子（60%、1,323 ps）について検討を進めた。具体的には、胎生 6.25 日において発現が認められない遺伝子の内、1 峰性を示す遺伝子は 47 ps、2 峰性を示すものは 5 ps、他方、胎生 6.25 日において発現が認められる遺伝子の内、1 峰性を示すものは 1,244 ps、2 峰性を示すものは 922 ps であった。この内、一例として、胎生 6.25 日での発現の有無に関わらず 1 峰性を示す遺伝子につき、発現変動起点ごとに抽出される遺伝子数を記載すると、胎生 6.50、6.75、7.25、7.50、7.75 及び 8.25 日ではそれぞれ、3、7、239、837、209 及び 28 ps となった。これらの結果により、胎児発生時の遺伝子発現ネットワークの網羅的解析手法とそれに資する網羅的かつ有効なデータの抽出と発現パターンによる分類が完了し、全体像抽出の為の最も重要な段階を完了した。更に、Shh の手法をモデルとした網羅的解析を進めるとともに執筆準備を開始した。

### （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】

平成24年度は、化学物質評価用の遺伝子発現データベースの有用性を高めるために、医薬基盤研トキシコゲノミクスプロジェクト（TGP）のラットデータをマウスの遺伝子発現データからなる Percellome データベースに統合すべく、マウスとラットのデータをシームレスに一括解析する手法（マウス・ラットの相同遺伝子対を一意に示す統合IDとその効率

的な処理アルゴリズム）を生成し、TGP データを Percellome 変換（絶対量化）した上で、Percellome データベースとの統合を進めた。

またデータベース及び解析ツールの公開に向け、Percellome 化せず取得した外部のマイクロアレイデータ（非 Percellome データ）と Percellome 絶対量化データとの比較を可能にする絶対量推定ソフトウェア SnCalc を開発したほか、外部研究システムから直接、Percellome データベースを利用するための WebAPI を作成し、運用を開始した。

平成 25 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 の改良を行い、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、平成 24 年度に作成した Percellome 公開用 Web サーバーの 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）インターフェイスについては、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機能強化を行った。

平成 26 年度は、マウスのみを対象としていた非 Percellome データの絶対量推定をラットにも拡大適用した。具体的には、医薬基盤研トキシコゲノミクスプロジェクト（TGP）のラットデータを利用し（TGP の発足時研究計画を当研究者らが Percellome 法を基軸に設計したため、そのデータは絶対



量化が可能。ただし TGP では Percellome 法適用に必要な品質管理を行っていなかったため、溶媒群データを中心に Percellome 計算ソフトウェア SCal4 による品質チェックを行い、明らかな外部スパイク RNA 添加ミスなどの大きな問題が無い 2072 枚の GeneChip データのみを収集した)、これらを Percellome 法により絶対量化し、SnCalc 参照データベースに追加・拡充することにより、ラットに於いても非 Percellome データの絶対量推定が可能となった(ただし、適用は TGP データと同様のトランスクリプトーム分布を有するデータに限定される)。

また実験間の共通変動遺伝子の抽出など、多数の化学物質暴露データを用いた複雑な集合計算を簡便に行うためのソフトウェア PercellomeExplorer の改良を進め、各化学物質に特異的な生体反応要素のハイスループット解析を実現した。

#### (4) システムトキシコロジー解析基盤の研究開発【北野】

遺伝子制御関係推定アルゴリズムについては、複数の計算手法の組み合わせを利用するに当たって、データに応じた適切な手法の選択が重要である。そこで我々は、データの全体的な特徴で利用するアルゴリズム群を決定し、個別の因果関係に関して各々の手法からの推定信頼度を参照して、最終的な推定ネットワークを構築する方法を開発した。すなわち、データの特徴分析から、そのデータの解析に最適なアルゴリズム群を決定する段階と、そこで選択されたアルゴリズム群の推論結果から、同じデータにおいても推定信頼度が変わる側面を

捉え、最も信頼度の高い相互作用の推定結果を集約する段階の 2 段階を経るものである。また、これらの成果は、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ついては、実例としてプリオン病のデータを利用した遺伝子相互作用ネットワークの分析から、プリオン病と同等の遺伝子発現変化を生じさせる少数のドライバー遺伝子を同定し、その効果をシミュレーションで確認している。ここでは、上で開発された遺伝子制御推定システムの推論結果をもとに、ネットワーク制御推定アルゴリズムを適応してその効果を確認するとともに手法の改良を行った。特に大規模なネットワークを扱うことを考慮して、個別遺伝子の制御可能性の推定に加え、多くの遺伝子を群として扱い、その遺伝子群の関係において制御可能性を推定するという手法を開発した。これにより、従来より大規模なネットワークの制御可能性の解析が可能となった。

さらにプロモーター領域の解析も付加するシステム SHOE を開発し、ゲノムデータの裏付けを基に制御ネットワークの推定を行うシステムも開発した。この過程において、全データからの用量選択、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加え、より正確な状態制御推定を可能とした。

大規模文献情報などの集約手法と伝達手法の研究に関しては、文献情報からパスウェイ構築を効率化するために、パスウェイエディタ CellDesigner を改良し、パスウェイの各

パスをモジュール化して扱えるように、グラフィックデータベースにモデルを格納する Ver. 5 の試作を行った。この結果、大量データを扱う際の実行速度や安定性などが向上した。

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 年度から導入するなど、普及が加速している。また、日本や欧州の製薬会社、化粧品会社、食品会社が強い興味を示しており、本研究のインパクトは大きいと考えられた。

#### 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ベースライン

（基線）が徐々に変動する反応成分）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）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。

平成 25 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お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。またバルプロ酸は基線反応をほとんど誘発せず、従って過渡反応は影響を受けない傾向が強かった。これらの特徴は、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、長期使用で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。

さらに平成 26 年度は、高脂血症治療薬として使用されているクロフィブレートについても同様の実験を行った結果、反復暴露により基線の低下が認められた遺伝子群が属する遺伝子発現経路群は四塩化炭素のそれと酷似していることが確認された。他方、過渡反応については、四塩化炭素が基線反応の影響を強く受けたのに対し、クロフィブレートの過渡反応は基線反応の影響を受けないものの比率が多かった。

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、特徴的な基線反応と過渡反応の両極端のパターンを示す四塩化炭素とバルプロ酸ナトリウム、その中間的な反応を示すクロフィブレートの三者の所見の差異を利用し、基線反応を引き起こす上流機構、転写制御領域の解析、基線反応と過渡反応の連鎖・非連鎖の分子基盤、等を分担研究者の協力のもとに推進する。また、単回暴露実験の結果の中から、反復暴露時の基線反応と過渡反応と連関す

る、即ち、予測に使用可能な、所見の存否を明らかにする。

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、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）を見出し、遺伝子発現変動波の微分関数を利用し、発現変動基点及び発現ピークの時点に着目する解析手法の開発を進め、モデルシグナルネットワークについて検討し、本解析手法により、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。加えて、発現変動起点を基に精査したところ、発生初期における非標準的な(non-canonical)新たな Shh シグナル制御系を見いだした可能性を平成 25 年度に報告した。具体的には、Shh の標的遺伝子である Gli 遺伝子の発現は、Shh の受容体を介さずに TGF $\beta$ 2・Smad3 を介して制御されるという可能性である。

平成 26 年度は、この発現変動基点及び発現ピークの時点を同定する技術を利用しつつ、全ての発現変動起点について候補遺伝子を抽出・分類し、網羅的にシグナルネットワークを推定する事を検討中である。自動抽出後の目視による検索の結果、発現のピークが 1 峰性のものと、2 峰性のものに分類できる事が明らかとなったので、自動

抽出と目視を組み合わせることで 2 回目のピークも含めたより網羅的な検索が実施できたものとする。1 峰性を示す遺伝子につき、発現変動起点ごとに抽出される遺伝子数を検討した結果、発現変動起点が胎生 7.50 日に存在する遺伝子数が最も多い結果となった (Shh 遺伝子も含まれる)。この時期は中胚葉誘導や器官形成が活発な時期であり生物学的に妥当と考えられる。今後、2 峰性を示す遺伝子の二つの発現変動起点の扱い方を確立しつつ、発現の局在を検討する whole mount ISH ならびに *in silico* プロモーター解析を組み合わせることで、発生初期過程に絡むシグナルネットワークの網羅的抽出を計る。

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、新規アルゴリズムの開発や既存ソフトウェア、データベースの改良を行い、3 年間を通じて、性能向上や適用範囲の拡大(汎用化)を実現した。特にマウスデータだけでなくラットデータにも解析対象を広げたこと、非 Percellome データについても絶対量推定がある程度の精度をもって可能になったことは、Percellome 技術の普及、利用拡大を推進し、以てトキシコゲノミクス研究の発展に寄与するものと期待される。

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ており、複数のソフトウェアの連動が実現されつつある。特に Garuda は、今後極めて重要なプロジェクトに成長すると考えられ、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者にとっても、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3年間を通じ、ほぼ計画通りに進捗した。

平成24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、平成25年度にバルプロ酸ナトリウム、平成26年度にクロフィプレートで、同様の解析を実施し、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ピ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る。これらの研究成果は、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の予測評価技術の開発を推進するものである。

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基点及び発現ピークの時点に着目する微分関数を利用した解析手法の開発と解析を進め、すでに 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が出来る事を報告したが、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し、網羅的に発生初期過程に絡むシグナルネットワークを描出し、信頼性と効率の向上と共に、応用範囲の拡充を目指す。

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、研究計画通りソフトウェア、データベースの改良・拡張を進め、網羅的遺伝子発現解析の効率向上を実現している。また平成24年度からインターネット上で一般公開している Percellome WebAPI の提供により、Garuda 経由を中心に Percellome データベースの利用が拡大しつつあり、引き続きサービスの継続や機能向上が求められている。

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研究については、システムバイオロジーを毒性学に応用しようとするものであり、そのために一連の方法論とソフトウェアを開発し、応用することにより、計画通り性能向上や機能拡張が進んだ。現在、この目標に対して、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な適用例を作る事や普及などの展開にも力を入れ、実用レベルでのシステム開発と世界的な普及に注力する計画である。