

厚生労働科学研究費補助金（化学物質リスク研究事業）
分担研究報告書

「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ためのインフォマティクス技術開発 - 」
（H24-化学-指定-006）

分担研究課題：「Perce llome 3次元データ等のための専用解析ソフトウェアの開発研究」

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・第一室長

研究要旨

先行研究に於いて遺伝子発現データの絶対量化技術である Perce llome 手法（特許 4415079 号）を実用化し、既に約 140 の化学物質による遺伝子発現変動情報を収録して現在もさらに拡張しつつある網羅的遺伝子発現情報データベース（Perce llome データベース）を構築している。この高精度大規模データベースの有効活用には、時間×暴露用量×遺伝子発現量からなる遺伝子発現変動情報を生物学者によるデータ把握が容易な 3次元波動面（Surface グラフ）に変換して解析する直感的で効率のよい手法（"Millefeuille"（ミルフィーユ、MF）データ処理）を採用している。本分担研究では、Perce llome 手法や Surface グラフの特徴を活かし、遺伝子発現反応の網羅的解析を効率良く行うためのプログラムの開発・改良を行い、トキシコゲノミクス技術を利用した毒性予測・評価システムの実用化を推進する。

平成 24 年度は、Perce llome 法の適用を以て、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、マウスデータからなる Perce llome データベースに結合し、統合トキシコゲノミクスデータベースの構築を進めた。また非 Perce llome データの絶対量推定ソフトウェア SnCalc.exe の開発に着手した。さらに Perce llome データベースシステムの WebAPI を開発し、ライフサイエンス研究用ソフトウェアの国際共通プラットフォーム GARUDA プロジェクト等、外部の学術・行政システムからのオンラインアクセスを可能にする本格的な一般公開の準備を整えた。

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 に新規フィルタールーチンを追加して自動抽出精度を向上させたほか、経時変化データ等の線グラフ表示デ

ータを対象とした発現変動起点の自動抽出アルゴリズムの新規開発や、適用対象がマウスに限定されていた SnCalc.exe のラットデータへの適用の可能性を検討し、また Percellome データベース一般公開システムの WebAPI の機能拡張を実行した。

本年度（平成 26 年度）は、SnCalc.exe による非 Percellome データの絶対量推定を、参照データベースにラット単回投与データに適用拡大すべく、医薬基盤研トキシコゲノミクスプロジェクト（TGP）の単回投与実験の溶媒群データ^(*)を SnCalc リファレンスデータベースに追加・拡充し、絶対量推定性能を確認したほか、遺伝子発現変動評価の標準値とすべく、SnCalc リファレンスデータベースから溶媒別の標準溶媒コントロールデータを生成した。また実験間の共通変動遺伝子を高速抽出するプログラム PercellomeExplorer を改良を進め、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。さらに Percellome 専用解析ソフトウェア群の基本ツールとなっている MFSurface.exe のユーザーインターフェイス、特に解析結果保存機能の拡充を行うなど、本研究班で開発してきた諸ソフトウェアが、より一層ユーザー本位の解析ツールとなるよう改良作業を行った。

(*)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。

A. 研究目的

本研究班が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実用化の推進の為にインフォマティクス開発を目指す中において、本分担研究は、その中核を成す約 140 の化学物質による遺伝子発現変動情報を絶対量化して収録している高精度大規模遺伝子発現情報データベース（Percellome データベース）を活用し、遺伝子発現反応の網羅的解析を効率良く行うためのプログラムの開発・改良を実施する。

B. 研究方法

B-1: 参照データ

遺伝子発現データとしては、Percellome プロジェクトで生成したマウスの遺伝子発現データ、および 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

タ標準化には Percellome 法を利用した。

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。

B-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

コンパイルを繰り返す実験的な開発における作業効率と、生成するソフトウェア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、XE5 (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。

B-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 サーバシステム; NTT コムウェア・NTT データ) にアプリケーションソフトウェアを移植して実施した。

C. 研究結果

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 化せずに取得した外部のマイクロアレイデータ (以下、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。加えて、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計画当初は平成 26 年度予定であった研究を実施した。即ち、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、偽陽性データを従来版 RSort よりも数百 ~ 2000 個程度の削減することに成功した。また、適用対象がマウスに限定されていた SnCalc.exe のラットデータへの適用の可能性を検討したのに加え、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、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。

平成 26 年度は、非 Percellome データの絶対量推定プログラム SnCalc をラットにも拡大適用した。その為に、以下の手順を踏んだ。TGP は、その発足に際し当研究者らが Percellome 法を基軸に設計したものであり、そのラットデータには絶対量化に必要な情報が含まれているが、運営上、Percellome 法の実施に必要な品質管理を行っていなかったため、本研究において溶媒群データを中心に Percellome 計算ソフト

ウエア SCal4 による品質チェックを追加的に行い、データ品質に大きな問題がない 2072 枚分のラット GeneChip データを収集した。これを利用し、SnCalc が必要とする参照データベースを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を可能とした(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)。

また拡充された SnCalc リファレンスデータベースを有効活用し、遺伝子発現変動評価の標準値とすべく、溶媒別の標準溶媒コントロールデータを生成した。これは本研究班本テーマの新型反復暴露試験の解析において、基線反応の変動の有無を評価する際の標準値とした。

さらに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抽出解析を実現したほか、Percellome 専用解析ソフトウェア群の基本ツールとなっている MFSurface.exe のユーザーインターフェイス、特に解析結果保存機能の拡充を行った(図)。これらにより、本研究班で開発してきた諸ソフトウェアが、より一層ユーザー本位の解析ツールとなった。

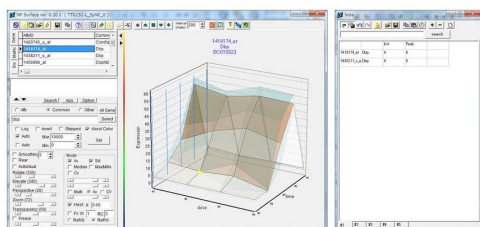


図. 解析結果保存機能を拡充した MFSurface.exe

D. 考察

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、3年間を通じて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、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。

本年度は、今までの研究の総まとめとして、解析対象となるデータを拡大し、評価基準となるような標準データをデータベースから生成し、開発した解析用ソフトウェアのユーザーインターフェイスの改良を行って、解析効率・精度の向上に努めた。

今後は、動作環境の拡大も視野に、GARUDA(北野宏明博士らが進め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>) を介したオンライン化や PercellomeDB の WebAPI の強化を進める。

E. 結論

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、平成 24 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 RSort 技術、PercellomeExplorer 技術の改良により、網

羅的遺伝子発現解析の効率向上を推進している。

また平成 24 年度から継続的に公開している Percellome データベース一般公開システムの WebAPI の改良・強化により、GARUDA 等、外部のオンライン解析ソフトウェアからも、本研究班の成果にアクセス出来るようになり、Percellome データベースの一層の普及と、毒性や創薬といった研究分野の解析能力・効率の向上が期待される。

F . 研究発表

1 . 論文発表

Janesick A, Nguyen TT, Aisaki K, Igarashi K, Kitajima S, Chandraratna RA, Kanno J, Blumberg B. Active repression by RAR signaling is required for vertebrate axial elongation., *Development.* (2014);141(11):2260-70.

Tanaka M, Yamazaki Y, Kanno Y, Igarashi K, Aisaki K, Kanno J, Nakamura T. Ewing's sarcoma precursors are highly enriched in embryonic osteochondrogenic progenitors. *J Clin Invest.* (2014);124(7):3061-74.

2 . 学会発表

Jun Kanno, Ken-ichi Aisaki, Satoshi Kitajima, Percellome Toxicogenomics, 50th Congress of the European Societies of Toxicology (EUROTOX2014)(2014.9.9) Edinburgh, UK, poster

Jun Kanno, Ken-ichi Aisaki, Satoshi Kitajima, Percellome toxicogenomics

project as the 3R-toxicology and the foundation of in vitro- and in silico-toxicology, the 9th World Congress on Alternatives and Animal Use in the Life Sciences (WC9) (2014.8.27), Prague, Czech Republic, Oral

相崎健一、北嶋 聡、菅野 純、遺伝子発現から見た毒性学 Percellome トキシコゲノミクスの進捗、第 36 回日本中毒学会総会・学術集会(2014.7.25) 東京、シンポジウム

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吸入暴露時の海馬 Percellome トキシコゲノミクス - 化学構造が異なる 3 物質の比較 -、第 41 回日本毒性学会学術年会 (2014.7.3) 神戸、口演

菅野 純、相崎健一、北嶋 聡、Percellome Project の進捗 新型反復暴露による慢性毒性の予測に向けての分子背景の解析、第 41 回日本毒性学会学術年会 (2014.7.2) 神戸、シンポジウム

G . 知的財産所有権の出願・登録状況 (予定も含む)

1 . 特許取得

なし

2 . 実用新案登録

なし

3 . その他

なし