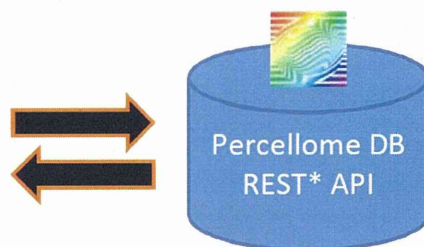
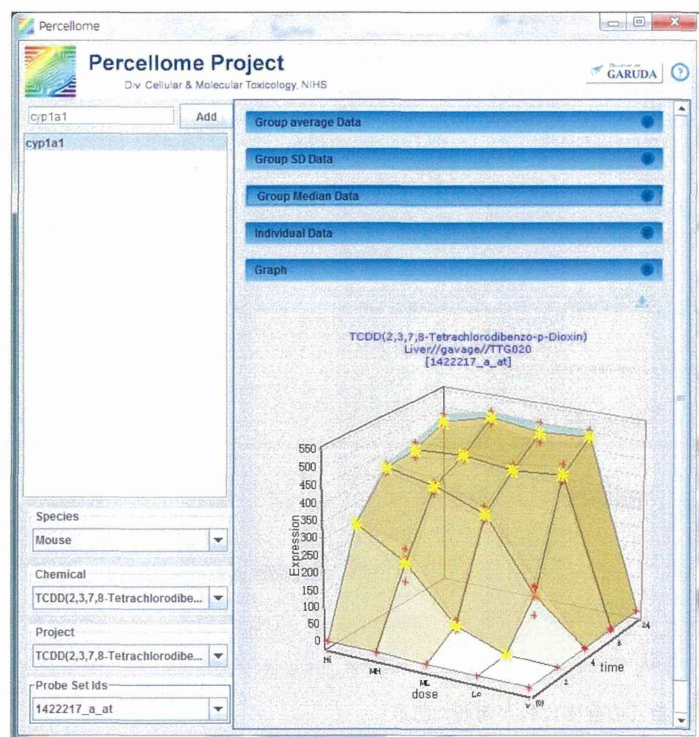


# Percellome Gadget



\*REST  
(Representational State Transfer)  
ハイパーメディアシステムのためのソフトウェア  
アーキテクチャのスタイルのひとつ

API  
(Application Programming  
Interface)  
アプリケーション(ユーザのプログラム)がOS等  
の公開機能を使うためのインターフェース。

19

## Percellomeデータベース 一般公開手続

### • 職務著作物として登録 → 国の所有とする

厚生労働省試験研究機関職務発明等規定 第14条に基づく届出。

第5章 著作物の取扱い  
(届出)

第14条 作成した著作物または作成中の著作物が次の各号の一に該当するに至った 職員(以下、「著作物を作成した職員」という。)は、速やかにその著作物の内容を詳記した著作物届(様式第4号)を、所属部長を経由して所長に届け出なければならない。

- 一 著作物を公表もしくは職員以外に移転する必要が生じた場合
- 二 著作物に関連して本規程に定める他の届出を行う場合

(帰属)

第16条 前条において、プログラムの著作物について、職務著作物であると認定したときの著作者は、国とする。

1 前条において、データベースの著作物について、職務著作物であると認定し、かつ、国は自己の著作の名義の下に公表するものの著作者は、国とする。

毒性部は実務担当。

20

# Percellomeデータベース 一般公開ライセンス

- データ: Creative Commonsライセンス

表示・継承 4.0国際 (CC BY-SA 4.0)

あなたの従うべき条件は以下の通りです。

**表示** — あなたは [適切なクレジット](#)を表示し、ライセンスへのリンクを提供し、[署名があったらその署名](#)を消さなければなりません。あなたはこれを合理的などのような方法で行っても構いませんが、許諾者があなたやあなたの利用行為を支持していると示唆するような方法は避けます。

**継承** — もしあなたがこの資料をリミックスしたり、改変したり、加工した場合に、あなたはあなたの貢献部分を元の作品と[同ライセンス](#)の下に頒布しなければなりません。

**追加的な制約は課せません** — あなたは、このライセンスが他の者に許諾することを法的に制限するよういかなる法的規定も**技術的手段**も適用してはなりません。

**ご注意**

あなたは、資料の中でパブリックドメインに属している部分に関して、あるいはあなたの利用が著作権法上の**権利制限規定**にもとづく場合には、ライセンスの規定に従う必要はありません。

保証は提供されていません。ライセンスはあなたの利用に必要な全ての許諾を与えないかも知れません。例えば、[パブリシティ権](#)、[肖像権](#)、[人権](#)などの他の諸権利はあなたがどのように資料を利用するかを制限することがあります。

- ソフトウェア: Apacheライセンス ver.2

21

# Percellomeデータベース ミラーサイト設置

- データベースの公開維持
- トラフィック増大対策
- 設置予定
  - 理研(イノベーションハブ\*: 小安先生、古関先生)  
\* 公的DBの集約ハブ化を目指す
  - システムバイオロジー研究機構(北野先生)

22

## 今後の展開

---

- Percellomeデータベース一般公開サイトの維持
- GARUDA Gadgetの拡充  
MFtools、特にPercellomeExplorerの移植
- PercellomeWeb REST API拡充
- GeneChip CELデータの公的DB登録

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  
— 新型反復暴露実験と単回暴露実験の網羅的定量的遺伝子発現情報の対比による毒性予測の精緻化と実用版毒性予測評価システムの構築 —

分担研究課題：「システムトキシコロジー解析基盤の研究開発」

研究要旨：

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

A. 研究目的

本研究は、これまでの蓄積を基盤に、システム毒性学のさらなる発展と実用レベルに達する成熟化を目指し、その中で本プロジェクトの成果を国際的に大きく普及させることを目指していく。特に、本研究フェーズでは、技術実証のフェーズから、より現実的な応用を意識し、広範な化学物質の低濃度長期反復暴露時の生体毒性の計算機予測を可能とし、OECD AOP (Adverse Outcome Pathway) との連動性を高めた解析技術を開発する。

ここまでの研究では、解析の基本技術の開発を中心に行ってきたが、生体内分子相互作用ネットワークは、大規模データからの推定を行う方式であった。しかし、広範な毒性予測を行うには、事前に高精度な計算モデルを構築し、そのモデルと大規模データから推定された分子間相互作用の変化を比較参照しながらより精度の高い毒性予測を行う必要がある。さらに、マウスでの実験結果と人間での違いや、マウスを利用した毒性試験ができない場合に、どのように人間での毒性リスクを予測するかの技術を

開発することがより重要性を増している。

そこで本研究計画では、主要な毒性に係るアウトカムに関連するモデルを構築し、その挙動を予測するための技術と情報基盤群の開発を行うこととする。これは、これまでのプロジェクトが、最初に特定の化学物質を取り上げ、その影響を順次解析するという流れで研究を進めたアプローチとは、逆のアプローチであり、毒性アウトカム(Adverse Outcome)という生理学的な結果から、それを引き起こす原因にさかのぼるという方法論である(図1)。すでに、今までの研究で、化学物質からその影響を予測する方法は、一定の開発が進められており、その方法のさらなる洗練化は継続する必要があるものの、この段階で、逆側からのアプローチを新たに導入することで、全体の解析系が確実に連動し、実用的な結果をもたらすことをより確実にすると考える。また、複数の化学物質の複合的効果を考える際に、Forward Effect Analysis のみでは、検討すべきターゲットの範囲が広範になると同時に、実際に Adverse Outcome として関連するターゲットはそのごく一部であ



るという極めて非効率的な解析を行わざるを得ない。これが、今まで複合的な暴露に対して計算論的アプローチが十分な有効性を持ちにくかった理由の一つである。しかし、実際には知りたい **Adverse Outcome** は明確に定義することができ、そのような結果をもたらすであろう遺伝子や分子の変動は、一定の組み合わせに限定することが可能である。その組み合わせが、ある程度広範囲にわたる場合においても、その範囲が定義されることは非常に重要である。これは、例えば、代謝工学で使われる **Metabolic Flux Analysis(FBA)**などの制約駆動型システム状態分析手法を、逆方向に利用するというイメージである。この考えをもとに、**AOP** の中間段階の分子の変動や遺伝子変動の注意すべきパターンを類型化し、そこから、複数の候補化合物の組み合わせを使った **Forward Effect Analysis** を実行することで、より効率的かつ精度の高い複合暴露の計算予測ができると期待される。

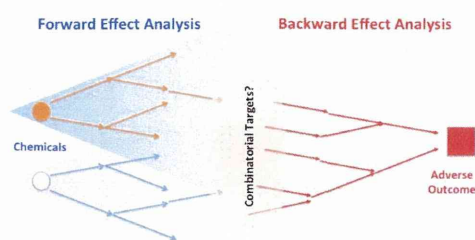


図 1 : Forward Effect Analysis と Backward Effect Analysis の融合

精密な計算機モデルとしては、(1) ミトコンドリアモデル、(2) 肝毒性モデル、(3) 皮膚毒性モデルを構築する。初期モデルは、文献情報やデータベースなどを利用して構築し、次の段階で、**Percellome** も含めた実験データを利用して精度向上を図る。このモデルは、分子間相

互作用と遺伝子制御ネットワークを包含する大規模なモデルとなる。すでに我々は、**EGFR Pathway**, **TLR Pathway**, **mTOR Pathway**, **Parkinson's Disease Pathway** など高精度モデルの構築の経験があり、この経験とそこで開発した一連の技術を毒性モデルの開発に応用する。

### (C) AOP 連携

**Adverse Outcome Pathway(AOP)** は、**Pathway** の名称を使っているものの、分子間相互作用ネットワークのレベルでの **Pathway** ではなく、より抽象的な、生理学的変化がどのように引き起こされるかを概念化したネットワーク化知識体系である。そこで、一連の毒性に関連する **AOP** を **Garuda Platform** 上におけるツールである **PhysioDesigner** で定義し、分子間相互作用ネットワークモデル構築ツールである **CellDesigner** によるモデルと連動させる。ここで重要なのが、分子レベルや遺伝子レベルでのどのような変動が、**AOP** レベルでの変動に関連するののかの同定と検証である。

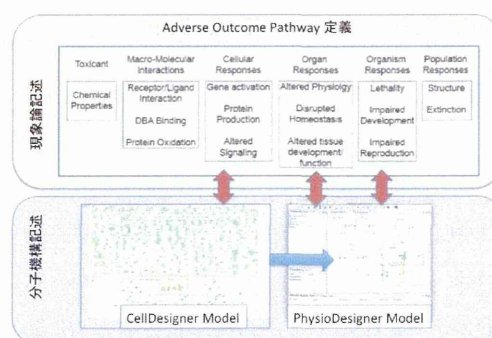


図 2: 現象論的記述が主体の AOP 定義と分子機構記述モデルとの連携

**AOP** レベルの記述が、生理学的なレベルであるので、最終的に分子レベルや遺伝子レベルのモデルでの予測を **AOP** レベルに連関させる必要がある。逆に、これが実現できれば、マウス実験からのヒトでの影響推

定やマウスを使わない場合でのヒトでの影響推定を可能とする技術に一步踏み出すことができる。

#### (D) Garuda Systems Toxicology Platform の開発

これらの研究と今までの研究成果を統合的に運用し、さらに実用的な毒性予測研究に資することができるプラットフォームを開発する。基盤は、これまでの研究で開発をしている Garuda Platform として、そこに毒性研究に特化したツール群、データベースなどを装備し、標準運用手順(Standard Operating Procedure)を定義する。これによって、標準的な毒性の計算機予測の手順が定義される。これらの機能を実現したシステムは、Garuda Systems Toxicology Terminal として実用化される。(図3)



図 3 : Garuda Systems Toxicology Terminal

#### B. 研究方法

本課題で、成し遂げようとしている目標は広範であり、多元的アプローチとそれらを統合することが必要となる。そのため、(1) 大規模データ解析技術の開発と (2) ゲノム解析とその関連データベースの整備を行った。

##### (1) 大規模データ解析技術の開発

近年、大規模かつ多次元的な生物学データが蓄積されつつある。特に、化合物が毒性を引き起こすメカニズムは非常に複雑であり、Percellome 等の化合物の毒性に関する発現データベースは、この複雑性を内包している。したがって、Percellome データベースから、有用な情報を抽出するためには、この様なデータの複雑性に対処できる解析法が必要不可欠である。しかしながら、この様な複雑性の高いデータから有用な情報を引き出すことは、従来の研究で用いられてきた統計解析法などでは不可能である。

このようなデータの複雑性に対処するために、Machine learning の手法が活用されつつある。しかしながら、一つの Machine learning の手法のみを使用するだけでは、複雑性の高いデータの全貌を捉えることはできず、適切な解析は難しい。

この様な問題に対しては、ensemble learning と呼ばれる、多数の machine learning アルゴリズムを統合する手法が有用である。実際に、多数の machine learning の手法を統合した IBM Watson は、Jeopardy クイズショーにおいて、人間のチャンピオンに勝利するという成果を上げている。

そこで、今年度、我々は、大規模な毒性データの複雑性に対応できる解析手法の確立を目的として、多数の machine learning の手法の実装を行い、さらに、これらの手法を統合して解析を行う ensemble learning system の開発を行った。

##### (2) ゲノム解析とその関連データベースの整備

この領域では、従来から開発を行っていた ACGT (図4) と SHOE (図5) の強化と Garuda Platform 上への実装を行った。



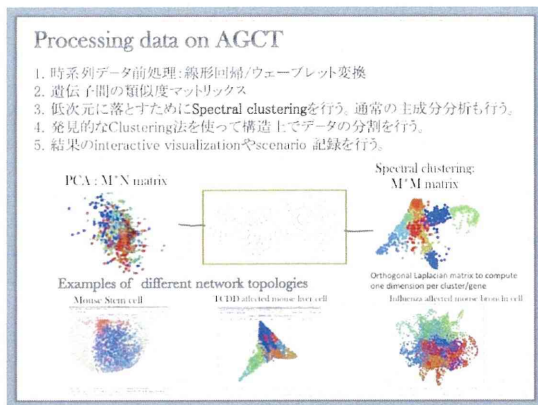


図 4: ACGT

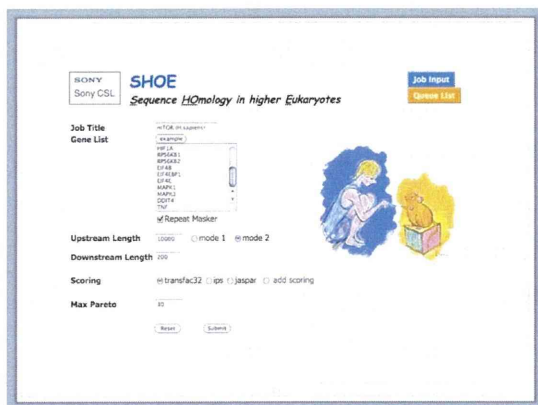


図 5: SHOE

これらソフトウェアの Garuda Platform への対応により、より広範な毒性解析パイプラインの構築が可能となる。

### C. 研究結果

#### (1) 大規模データ解析技術の開発

今回、多数の machine learning の手法を統合して解析を行う、ensemble learning system の開発を行った。今回実装を行った machine learning の手法は、(1) クラスの予測を行う分類器、(2) 数値の予測を行うアルゴリズム、の二種類に分類される。

(1) の分類器は、例えば、図 6 に示すように、遺伝子の各種 cell line における発現データを入力データとして、それぞれの遺伝子が“薬剤ターゲット候補”か“薬剤ターゲット候補で無い”などの予測を行う

ものである。一方、(2) の数値の予測を行うアルゴリズムは、例としては、図 7 のように、各患者の健康診断データ（年齢や肥満度など）から、一年後の血糖値の数値の予測を行うものである。

今回、我々は、全部で 58 種（表 1）の分類器の手法を実装した。また、これらの 58 種のアルゴリズムを並列に実行し、結果を統合して予測を行うシステムの開発を行った（図 8）。図 2 の例は、3 つのアルゴリズムを並列に実行している場合である。この例では、3 つのアルゴリズムは、それぞれ、遺伝子 A が、“薬剤候補となる”か“薬剤候補とならない”かの予測を行っている。その後、この 3 つのアルゴリズムの多数決を取って、最終的な予測を行う。

今回実装したシステムを、大規模な薬剤投与下における発現データ（cmap, <https://www.broadinstitute.org/cmap/>）に対して使用し、薬剤候補遺伝子を正確に予測できることを確認した。このシステムは、遺伝子や化合物が毒性を出すかどうかの予測や、パスウェイが毒性に関わっているかどうかについての予測に対し応用可能であり、これらに対しても精度の高い予測が出来ると期待される。

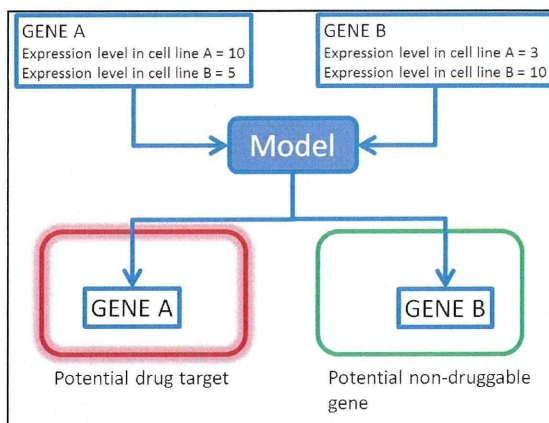


図 6. 分類器の例





の強化と Garuda への準拠を行った。  
ACGT は、転写解析ツールであり、転写産物の動態から遺伝子間の相互作用を情報幾何学的に推定し、可視化する。このツールの Garuda 準拠が完了した。(図 1 0)

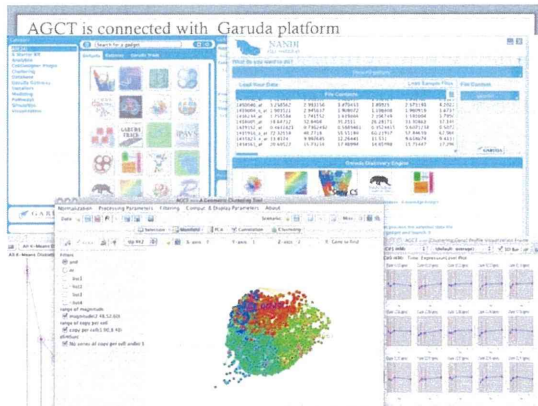


図 1 0 : Garuda に準拠した ACGT

これにより、ACGT で解析した遺伝子相互作用のうち重要な部分を、既存の大規模知識ベースや Pathway map 上に投射することで生体内での影響の推定が容易になった。(図 1 1)

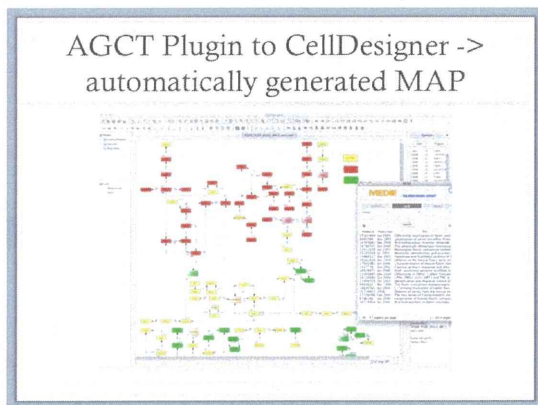


図 1 1 : ACGT-CellDesigner 連携

この機能を利用して、TCDD と TCDF の毒性面での同一性・差異を Pathway 上にマップして解析することができた。(図 1 2)

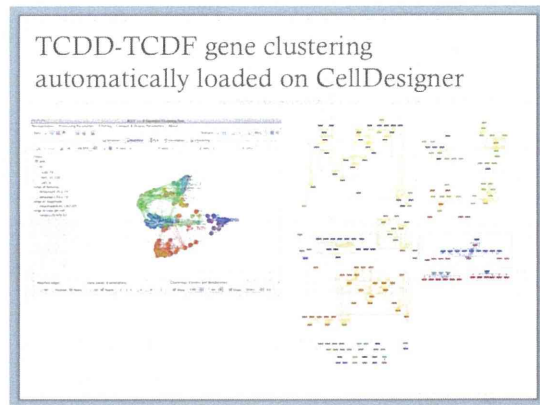


図 1 2 : 連携機能を使った TCDD-TCDF の毒性解析例

さらにホモロジー解析を行う SHOE の機能強化を行った。(図 1 3)



図 1 3 : SHOE の機能強化: ホモロジー解析機能の強化

これは解析精度の向上やデータベースの整備とともに遺伝子とそれに対する転写因子のゲノム上での位置を可視化するなどの機能を強化した。(図 1 4)



図 1 4 : 転写因子のゲノム上の位置の表示

さらに SHOE も Garuda に準拠することで、他ツールとの連携を強化した。(図 1 5)

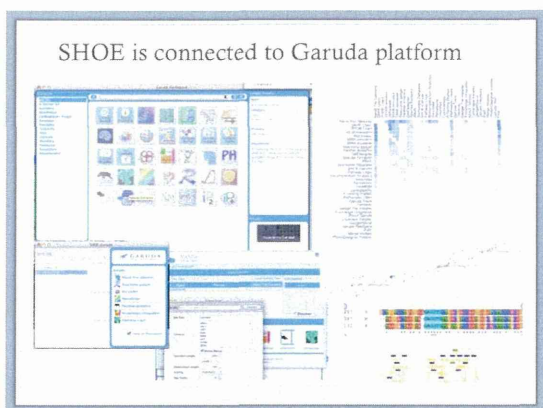


図 1 5 : Garuda に準拠した SHOE

#### D. 考察

各々の機能の強化、Garuda 準拠による連動性の強化は、毒性パイプラインの構築において目覚ましい効果があることが確認された。

#### E. 結論

本年度の成果は、概ね順調であり、これを基盤に最終的なプロジェクトの目標を達成する道筋がより明確になった。

#### F. 健康危険情報 なし

#### G. 研究発表

##### 1. 論文発表

- ① Hieu T Nim, Milena B Furtado, Mauro W Costa, Nadia A Rosenthal, Hiroaki Kitano and Sarah E Boyd. VISIONET: intuitive visualisation of overlapping transcription factor networks, with applications in cardiogenic gene discovery. *BMC Bioinformatics*. 16, 141, doi:10.1186/s12859-015-0578-0, May 1, 2015.
- ② 北野宏明. システム・トキシコロジーの展開. *QIGEN eyes*. 12, 7-9, May 14, 2015.
- ③ 北野宏明. システム・トキシコロジーの展開 (第 2 回). *QIGEN eyes*. 13, 7-9, Aug. 27, 2015.
- ④ Tiago J. S. Lopes, Jason E. Shoemaker, Yukiko Matsuoka, Yoshihiro Kawaoka, Hiroaki Kitano. Identifying problematic drugs based on the characteristics of their targets. *frontiers in Pharmacology*. 6, 186, doi: 10.3389/fphar.2015.00186, Sep. 1, 2015.
- ⑤ Matsuoka, Y.; Fujita, K.; Ghosh, S.; Kitano, H. Weaving Knowledge into Biological Pathways in a Collaborative Manner. *Computational Systems Toxicology* (eds. Julia Hoeng and Manuel C. Peitsch, Humana Press, Springer), 181-208, Sep. 14, 2015.
- ⑥ Kitano, H. Accelerating systems biology research and its real world deployment. *npj Systems Biology and Applications*, 1, doi:10.1038/npjbsa.2015.9, Sep. 28, 2015.
- ⑦ Takahiro Amemiya, Masashi Honma, Yoshiaki Kariya, Samik Ghosh, Hiroaki Kitano, Yoshihisa Kurachi, Ken-ichi Fujita, Yasutsuna Sasaki, Yukio Homma, Darrel R Abernethy, Haruki Kume & Hiroshi Suzuki. Elucidation of the molecular mechanisms underlying adverse reactions