

Luxembourg, Sep. 27, 2011. (invited)

Kitano, H. Garuda Goals. Garuda Six Workshop, Pavillon Skip, Luxembourg, Sep. 28, 2011.

Kitano H. Novel scientific approaches in chronic diseases: Systems biology. India-Japan Symposium on Global Challenges in Health and Environment, Indian Embassy Auditorium, Tokyo, Oct. 7, 2011. (invited)

Kitano, H. Systems Drug Design. Invited presentation at Monash University, Monash University, Melbourne, Australia, Oct. 19, 2011. (invited)

Kitano, H. Systems biology and its medical and industrial implications. Systems Biology Forum, Monash University, Australia, Oct. 21, 2011. (invited)

北野宏明. システムバイオロジーと創薬パイプライン. 神奈川科学技術アカデミー (KAST) 教育講座 システムバイオロジー・合成バイオロジーの最新動向コース、Nov. 7、2011. (invited)

Kitano, H. Unmet medical needs in the aging society. Economic Opportunities from the Ageing Society: Policies and Challenges, Asian Development Bank Institute, Tokyo, Nov. 18, 2011. (invited)

Kitano, H. HD-Physiology Project and the Garuda Alliance Software and Resource Platform. The 4th Global COE International Symposium on Physiome and Systems Biology for Integrated Life Sciences and Predictive Medicine, Senri-Hankyu Hotel, Osaka, Nov. 21, 2011. (invited)

北野宏明. システム創薬の展開. 第3回三井情報バイオサイエンスセミナー、三井情報株式会社、東京、Nov. 28、2011. (invited)

Kitano, H. Introduction to Systems biology and to multi-level modeling with CellDesigner. Workshop “Can Systems Biology Aid Personalized Medication?” Linköping University, Sweden, Dec. 5, 2011. (invited)

Kitano, H. Overview of HD-physiology network. Workshop “Can Systems Biology Aid Personalized Medication?” Linköping University, Sweden, Dec. 6, 2011. (invited)

Natalia Polouliakh, Jun Kanno, Yukiko Matsuoka, Ken-Ichi Aisaki, Richard Nock, Frank Nielsen, Keigo Oka, Satoshi Kitajima and Hiroaki Kitano, Discovery of Gene Network Regulated by the Toxicity Equivalent Factor of 2,3,7,8-Tetrachlorodibenzo-p-Dioxin (TCDD) and 2,3,7,8-Tetrachlorodibenzofuran (TCDF) chemicals. the 11th International Conference on System Biology (2010.10.11) (Edinburgh, UK) Poster

北野 宏明、システムバイオロジーによる毒性オミクスへのアプローチ、第37回日本トキシコロジー学会 学術年会 (2010.6.16) (沖縄) 口演

Polouliakh N, Matsuoka Y, Ghosh S, Nock R, Nielsen F, Kitajima S, Takagi A, Aisaki K.I, Kanno J, Kitano H, Signaling network in Mouse Embryonic Stem Cells The 9th International Conference on Systems Biology (Aug. 22-28) , 2008, Sweden, poster

G. 知的所有権の取得状況

1. 特許取得

- (1) 特許第 3995099 号、2007 年 8 月 10 日登録、「高次元データを塊に分割する装置」
- (2) 特許第 4415079 号、2009 年 12 月 4 日登録、「遺伝子の絶対発現量測定方法」
- (3) 特許出願(特願 2010-294175) 2010 年 12 月 28 日 (審査請求 2011 年 1 月 7 日) 「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

(資料 1) 総合報告資料

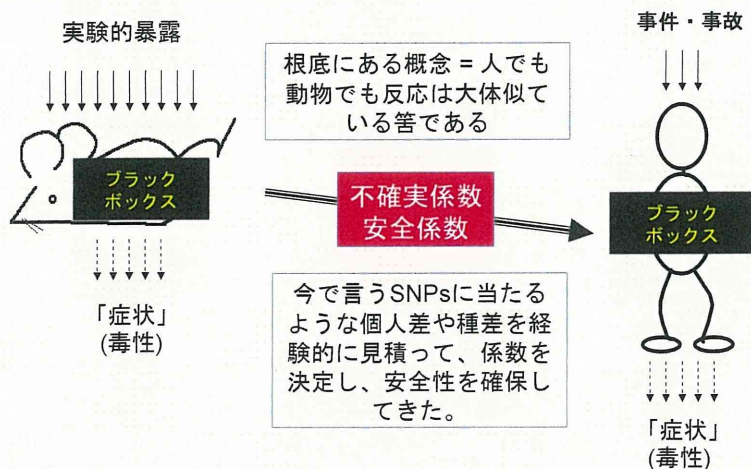
化学物質の有害性評価手法の迅速化、高度化に関する研究

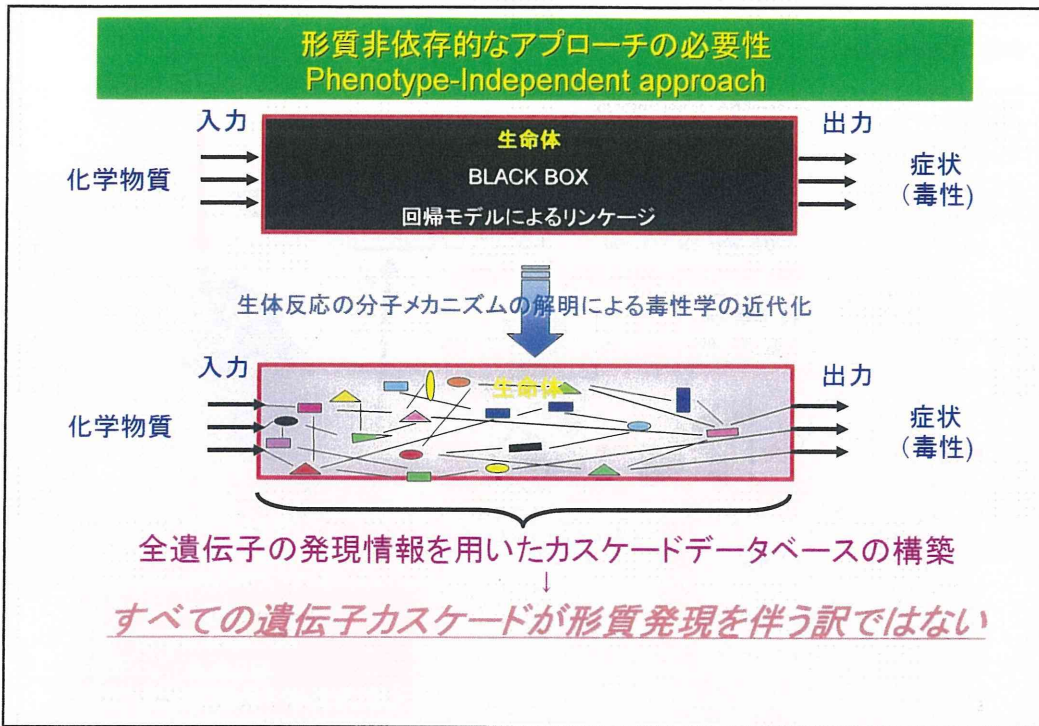
—網羅的定量的大規模トキシゲノミクスデータベースの維持・拡充と毒性予測評価システムの実用化のためのインフォマティクス技術開発—
(H21-化学-一般-001)

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
菅野 純

半数致死量(LD₅₀)や安全係数(不確実係数)からの毒性学の近代化

魔法の数字「安全係数」をいつまでも使い続けるのか？
サリドマイド問題のような事例の回避





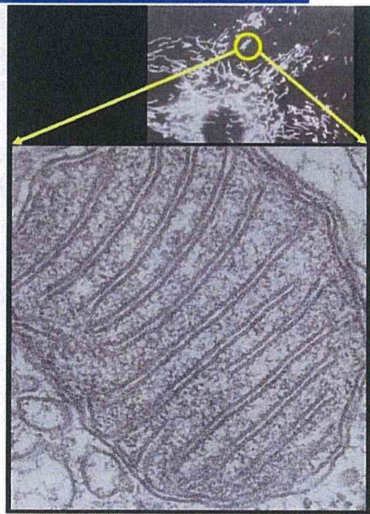
トキシコゲノミクスは現行毒性学にとって、光学顕微鏡しかない時代に現れた電子顕微鏡のようなもの。
 その実用化には、皆が納得する「教科書」が必要。

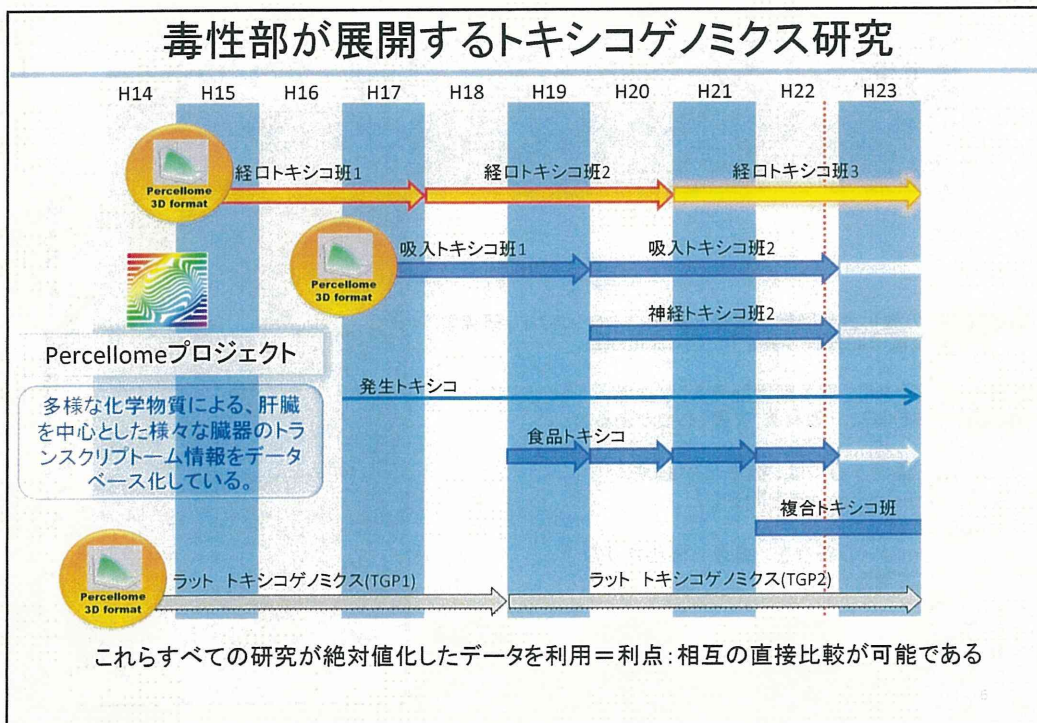
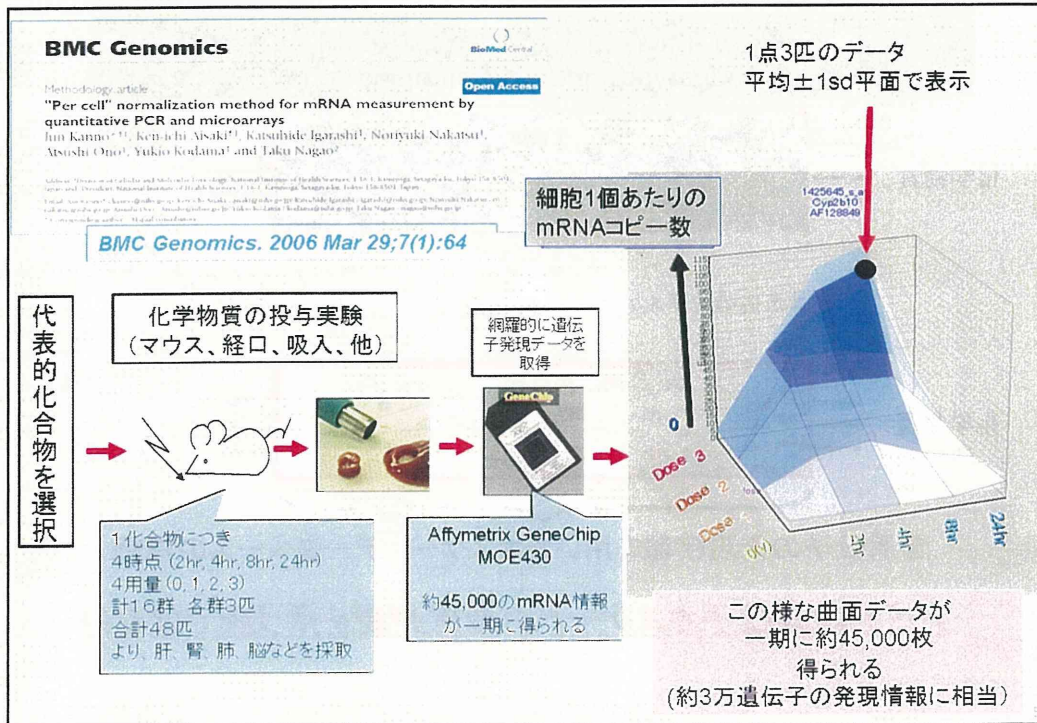
電子顕微鏡の教科書や図譜が出来るまでに、10~20年。研究者が多数の電子顕微鏡写真をアーカイブ化した。

トキシコゲノミクスは、電子顕微鏡のような立場。多数のデータを基に、「教科書」を書く必要がある。

データを蓄積するには、標準化が必須。
 →Percellome法を開発した。

↓
 数万種類のmRNAの各々を、細胞一個当たりのコピー数として測定。

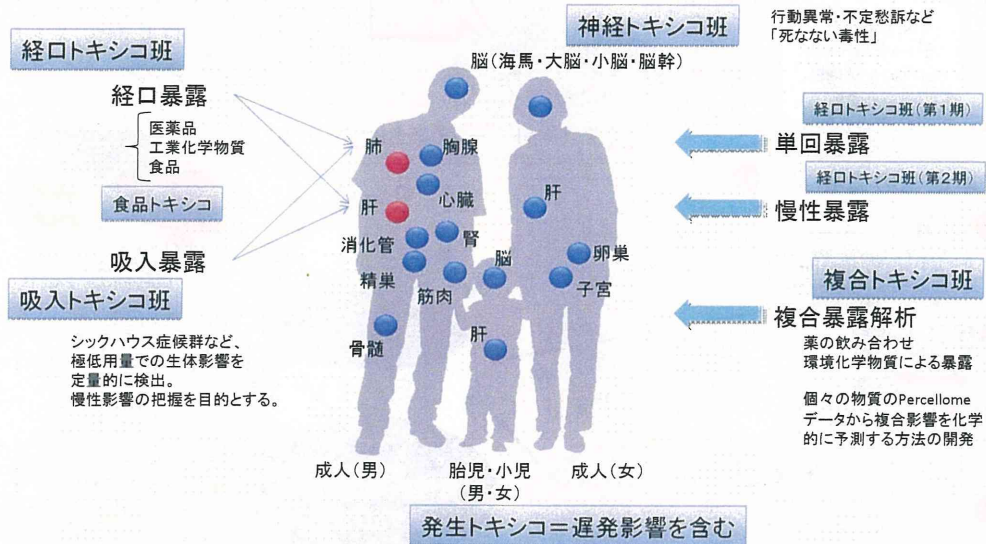






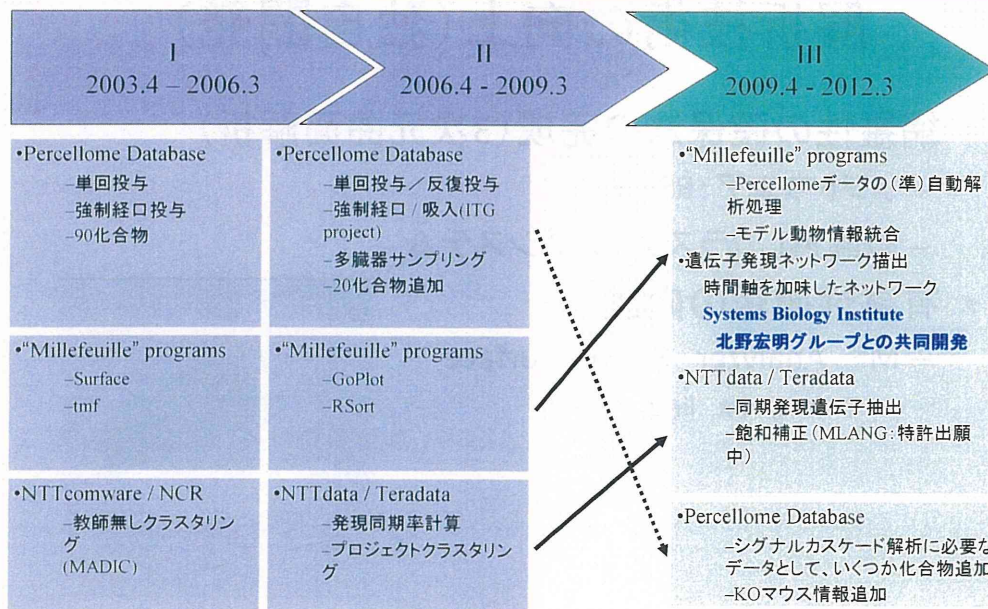
Percellome Toxicogenomics データベース

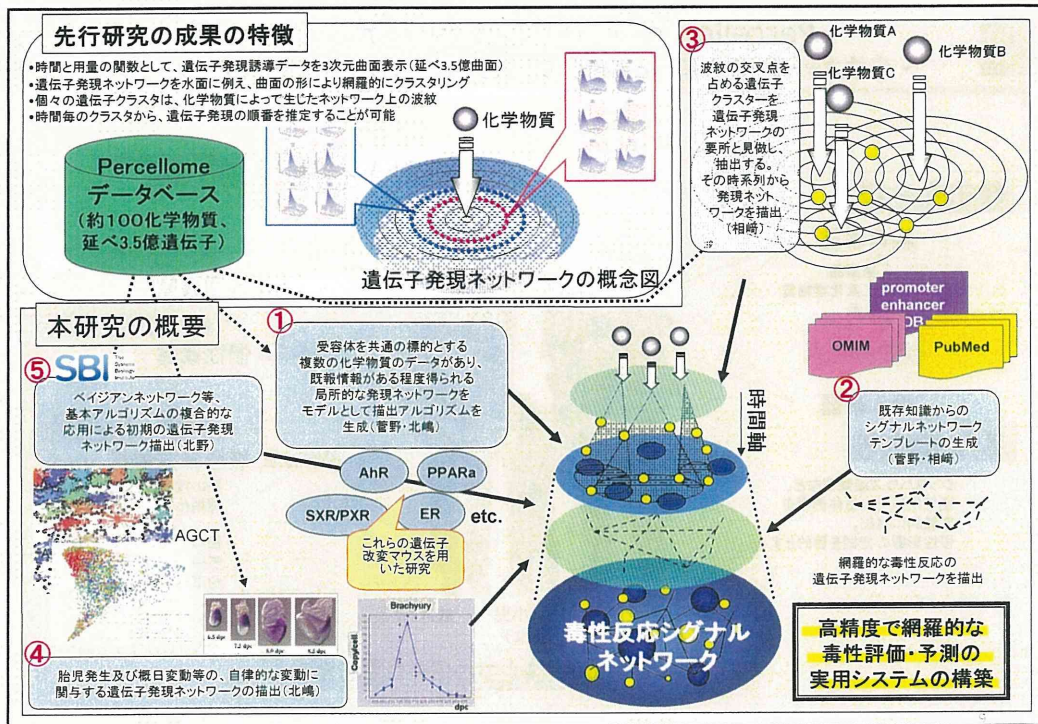
～ 毒性ネットワークの解明と高精度の毒性評価・予測のための基盤データベース ～



Percellome Toxicogenomics Project (第3期)

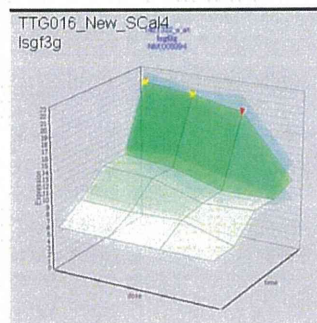
厚生労働科研費





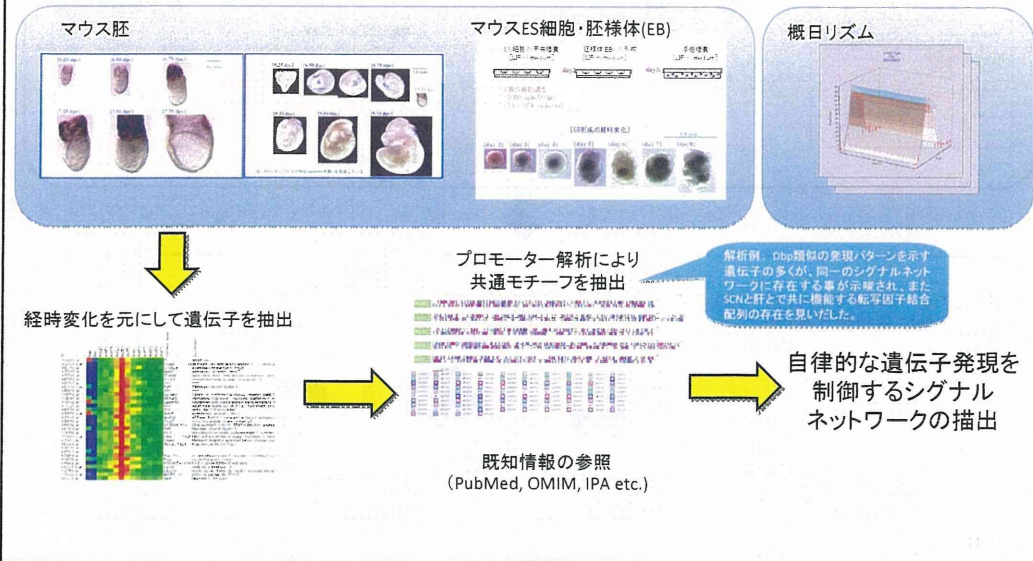
解析技術の向上(独自開発)

- 網羅性の確保法の完成(3次元曲面解析)
 - ソフトウェア: RSort
 - 教師無しクラスタリングシステム
- 有意性確認の向上
 - MF Analyzer / MF Surface
 - T-Test, Fx line



胎児、ES細胞、概日変動等の自律的なシグナルネットワークの描出

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等のシグナルネットワークの描出アルゴリズムを開発し、その成果をより複雑な毒性シグナルネットワークの描出に応用する。



複数の化学物質のクラスタ交叉点からのカスケード描出のための基盤ソフトウェア開発～PercellomeExplorer

本プログラムで、既知のネットワーク情報やGO情報との照合処理も可能となった

時系列を考慮した誘導遺伝子比較
(例)食品添加物Sodium DehydroacetateとAcephateの高い類似性を発見

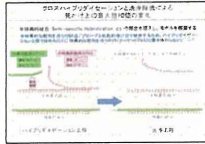
誘導遺伝子情報による化学物質の主成分解析図

>>>>> 全データベースを対象とした時間経過と用量を加味した解析が容易に実施できるようになった。また解析経過及び結果の可視化により、データ検証の精度とスピードが格段に向上した。

in silico処理によるマイクロアレイデータの高精度化

クロスハイブリダイゼーション補正～MLANG

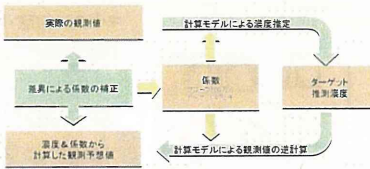
独自モデルの考案



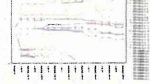
Langmuir吸着等温式の多溶質への拡張

$$N_{ij} = \frac{C_{ij}}{\alpha_{ij}} \cdot K \cdot \left(N - \sum_{j=1}^m N_{ij} \right) + \varepsilon_{ij}$$

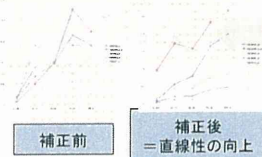
実データによる学習過程を含む係数推定ロジック



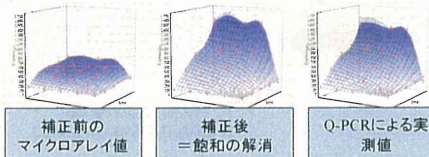
学習効果の推移



飽和したデータ等を補正し、データの定量性(直線性)を高めた

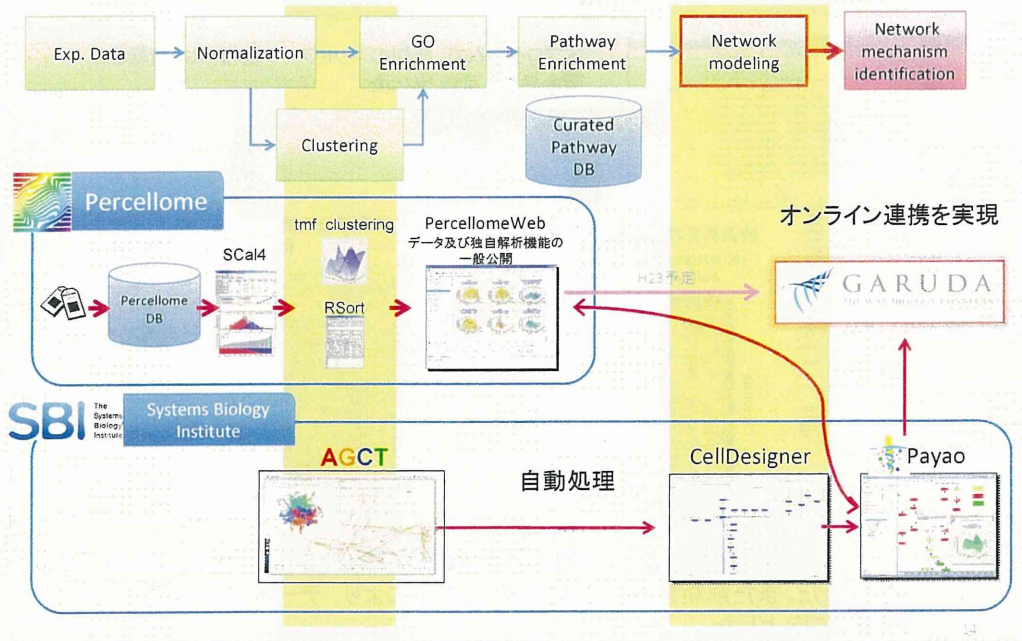


補正処理によって、マイクロアレイデータの飽和が補正され、定量RTPCR と同等の結果を得た

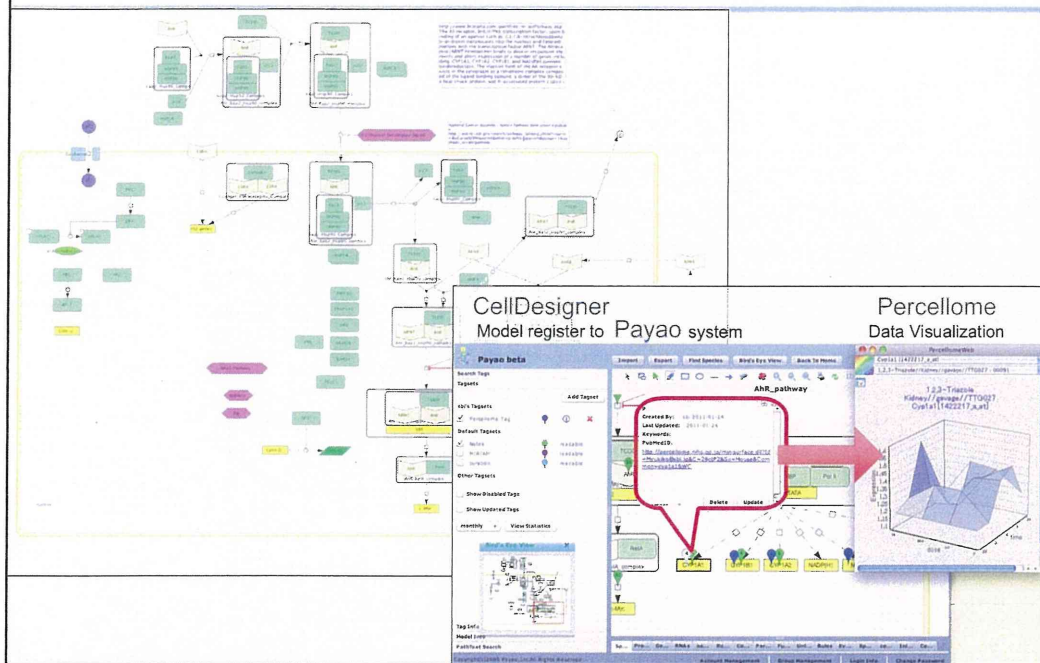


特許出願済み

遺伝子発現ネットワーク描出と国際公開(SBIとの共同開発)



(例) AhR pathway built by CellDesigner and Payao



研究成果

- 遺伝子KOマウス(AhR遺伝子及びER α 遺伝子)を用いた稀少データの取得、およびそれを用いた遺伝子発現ネットワーク解析(反応の欠落部分の利用による高解像度解析)
- 胎児・ES細胞における発生・発達や、概日変動をモデルとした、自律的な遺伝子発現データの網羅的なネットワーク解析技術の生成
- 既存情報(文献やデータベース)の参照システムの作成
- 複数の化学物質の遺伝子発現クラスター解析を効率化するソフトウェア群の開発
- 多次元のデータ構造からの発見的手法を用いたクラスタリングソフトウェアAGCT (A Gene Clustering Tool)など、ネットワーク自動描出に向けたソフトウェア群の開発
- オンラインデータベースPercellomeWebシステムの公開・運用、GARUDA対応準備
- 委託研究
 - 独自クラスタリングを利用した化学物質分類および同期発現遺伝子の抽出技術の高精度化
 - マイクロアレイデータのクロスハイブリダイゼーションによる誤差の補正技術の作成
 - 次世代シーケンサへのPercellome法適用及び、RNAseq技術の性能評価



In silico 安全性評価システムの実用化へ

(資料 2)

委託研究報告書(STEP8～10)

委託研究報告書 (STEP8)

同期発現インフォマティクス研究

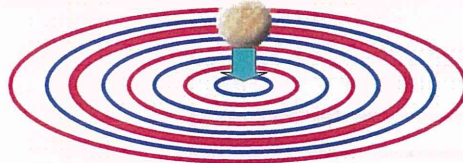
1. 目的

平成17年度において、クラスタリング（化合物影響により似たような反応を示す遺伝子群=クラスタを抽出する技術）を用いて、暴露する化合物によって異なる反応パターンを示しているも、高率に同じクラスタを形成している遺伝子群(同期発現遺伝子群)を抽出する研究を行った。

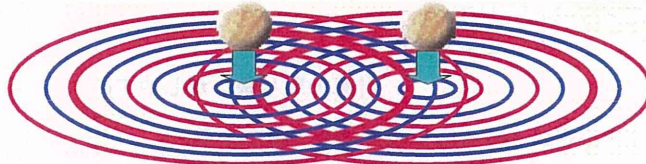
平成20年度において、半特異的結合を考慮したマイクロアレイデータの補正方法（MLANG）の基本部分を開発した。この手法を用いることにより、発現値の推定精度が向上したと期待される。
この高精度の推定発現値を用いて、同期発現判定がどのように変更されるか調査した。

2. 1. 同期率計算の基本概念

池に石を投げ入れると同心円状に波紋が広がる様に、化合物を投与することにより、RNA転写への影響が遺伝子カスケード上に広がっていく。



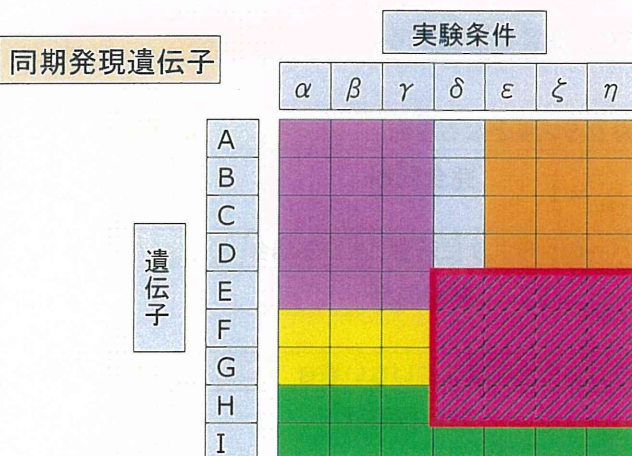
化合物の影響により、遺伝子の発現量が時間とともに変化していく。変化のパターンが似ている遺伝子をクラスタリング手法によってまとめる。これは、同心円状に存在する遺伝子をまとめることを意味する。同心円のどちら側に存在するのかわからない。



複数の化合物に対する反応から、常に同期して発現するような遺伝子を見つけ出すことにより、特定の遺伝子グループの同定が可能であると考えられる。

Copyright(C)2009-2010 NTT DATA Corporation

2. 1. 同期率計算の基本概念



単一の実験条件で、遺伝子軸方向の指標としてクラスタリングを実施できる。単一の遺伝子で実験条件軸方向でまとまりを示す指標を作成する。

Copyright(C)2009-2010 NTT DATA Corporation

2. 1. 同期率計算の基本概念

各遺伝子の周囲の状況を考慮し、実験条件をまたがっても、周囲の状況の変化が少ない場合に、大きな値を示す数値を作成する。

	α	β	γ	δ	ϵ	ζ	η
A							
B							
C							
D							

(1) 値域：0以上1以下
同期しているとみなせる場合に1

計算の取り扱いがしやすい

(2) 性質 実験条件を増やした場合に低下する

$$0 \leq \text{Sync}_{\alpha\beta\gamma\delta}(A) \leq \text{Sync}_{\alpha\beta\gamma}(A) \leq 1$$

2. 2. 持っているべき性質①

重複を含まないクラスタ結果で、性質を定義する。

(1) 全実験条件でUNIQUEなクラスタとなる場合

$$\text{Sync}_{\alpha \dots \omega} = 1 \quad \text{完全同期}$$

(2) 全実験条件で同一メンバーのクラスタに所属している場合

$$\text{Sync}_{\alpha \dots \omega} = 1 \quad \text{完全同期}$$

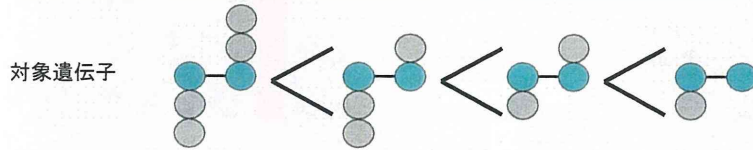
(3) ある実験条件でどのクラスタにも所属していない場合

$$\text{Sync}_{\alpha \dots \omega} = 0 \quad \text{完全非同期}$$

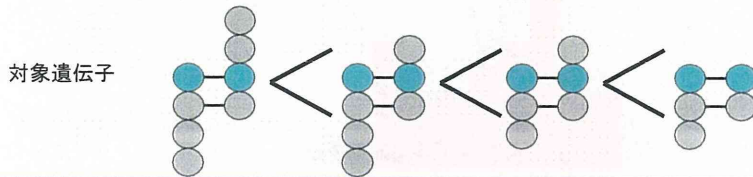
2. 2. 持っているべき性質②

2 実験条件間、重複を含まないクラスタの場合

(1) 対象遺伝子だけ重複



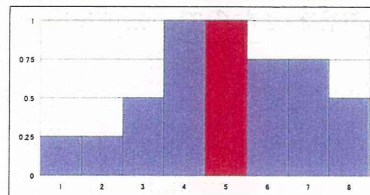
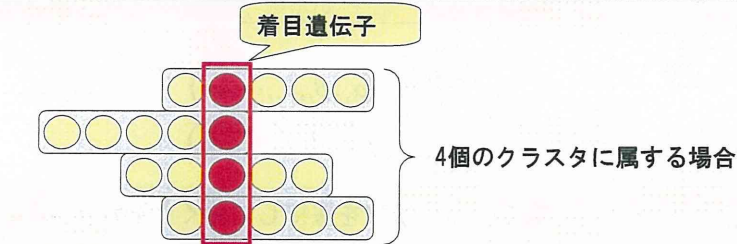
(2) 対象遺伝子以外でも重複



Copyright(C)2009-2010 NTT DATA Corporation

2. 3. 同期率の構成 重複クラスタの対処①

着目する遺伝子に関わるクラスタ全部を対象にして分析を行う。

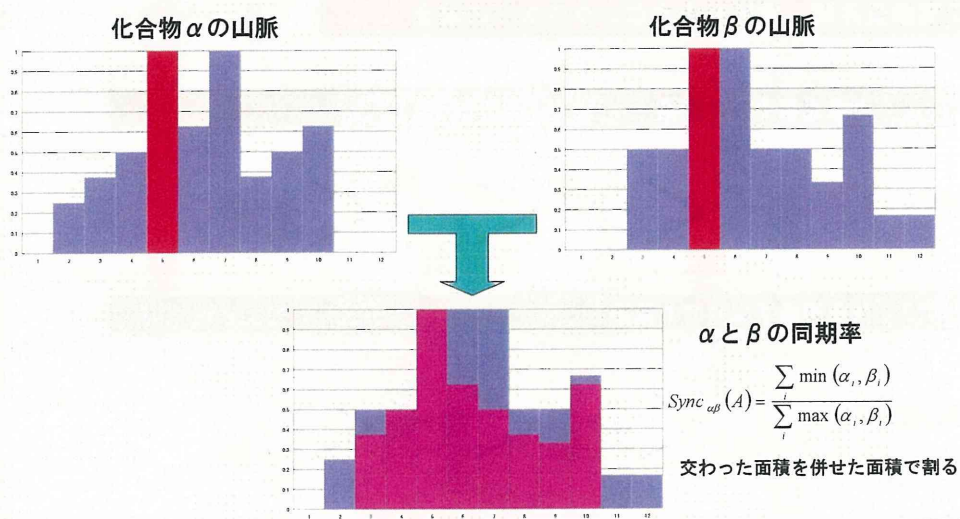


属する個数の割合から、山脈のような形状が描ける

この山脈の形状が似ていれば、近くの遺伝子の存在状況が似ているといえるのではないかな？

Copyright(C)2009-2010 NTT DATA Corporation

2. 3. 同期率の構成 重複クラスタの対処②



Copyright(C)2009-2010 NTT DATA Corporation

2. 3. 同期率の構成 3以上の実験条件の対処

計算方法を拡張する

$$Sync_{\alpha\beta\gamma\cdots\omega}(A) = \frac{\sum_i \min(\alpha_i, \beta_i, \gamma_i, \cdots, \omega_i)}{\sum_i \max(\alpha_i, \beta_i, \gamma_i, \cdots, \omega_i)}$$

計算を行うためには、元の山脈情報を保持しておく必要がある

遺伝子2個の組み合わせなので、数が多い

Copyright(C)2009-2010 NTT DATA Corporation

3. 1. 同期率計算の改良

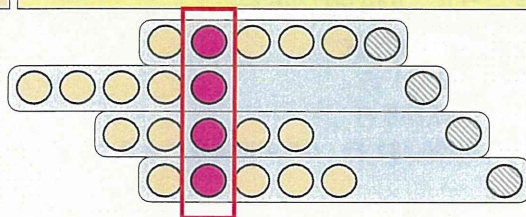
- 初期の同期率計算では次のような問題を抱えていた
 - UNIQUEとなった遺伝子に対する同期率が他の遺伝子と比較して飛びぬけて大きな値を示す。
 - クラスタリングの際に、クラスタリング対象から外れていると同期率が計算できず、同期率の比較ができない。

3. 1. 同期率計算の改良

基本の計算方法では、ユニーク遺伝子、ISOLATE遺伝子、および、クラスタリングを実施しなかった遺伝子が突出した結果となる。

対応策

各クラスターにダミーの遺伝子がついたものとみなす。何個のダミー遺伝子を与えるかをパラメータとして指定する。

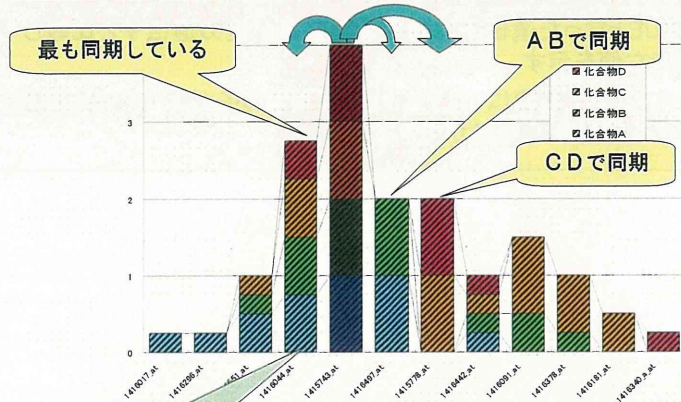


クラスタリング対象外とした化合物が存在した場合

同期率を「0」と定義する

3. 2. 同期率計算の改良

同期率そのものは、化合物を貫いて、同じ動きをしている遺伝子が存在するかの指標だが、どの遺伝子と同じ動きをしているか不明であった。



対象プローブ

本手法で得られた数値は可換ではない

同期している対象化合物の調査は計算量の問題で困難

Copyright(C)2009-2010 NTT DATA Corporation

4. クラスタリング計算途中での停止

- 同期率計算の途中で、2TBytesを使いきり、容量不足のため計算が停止した。
- データ量
 - クラスタリング結果テーブル 249,907,396行
 - TTG044-Lのクラスタリング結果行数
 - ・ Sfc.no 1: 13473 probes, 56,151 行
 - ・ Sfc.no 2: 21866 probes, 147,843 行
 - ・ Sfc.no 3: 31747 probes, 3,253,781 行 (MLANG補正によるクラスタリング)
 - 特殊な結果
 - ・ RIGOROUSプローブが1315個からなるクラスターが存在する

Copyright(C)2009-2010 NTT DATA Corporation

4. クラスタリング計算途中での停止 TTG044-Lのクラスタ状況



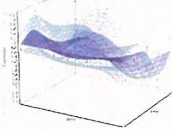
Snf1lk
1419766_at



最も多くのクラスタ(366個)に属するプローブ(1419766_at)に対する各クラスタのRIGOROUSプローブ(近い10個)を確認する。

ここに挙げたRIGOROUSプローブは、別々のクラスタに所属するが、ほぼ同じ形となっている。何らかのバイアスがかかり、大きなクラスタが出来上がり、その周辺にも多くのクラスタが出来上がったものと思われる。

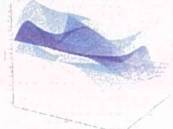
Rnmt
1436328_at



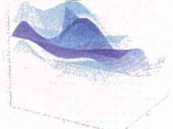
Paqr5
1456654_at



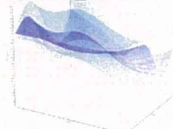
Rbm9
1418247_s_at



Krtap16-8
1425237_at



Unc119
1418123_at



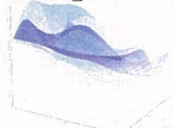
2310057J18Rik ///
2310065H11Rik
1419643_s_at

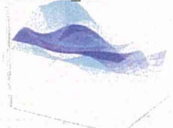


Cdca7l
1451306_at

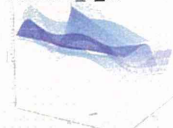


Pde11a
1432999_at





Mmab
1452468_s_at



Copyright(C)2009-2010 NTT DATA Corporation

5. 1. 同期率計算 対象化合物



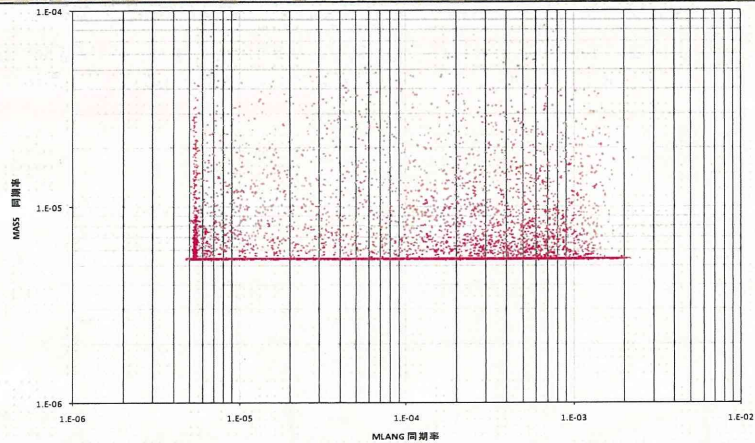
次の化合物に対して、同期率計算を実施する

化合物	溶媒	餌種類	実験日
TTG020-L TCDD_TEST20040323	corn oil	STD	2003/10/16
TTG026-L TCDF	corn oil	STD	2003/11/13
TTG047-L Bisphenol A	MC + 0.1%DMSO	PLD	2004/6/17
TTG048-L Genistein	MC + 0.1%DMSO	PLD	2004/7/1

Copyright(C)2009-2010 NTT DATA Corporation

5. 2. 同期率計算 結果比較（同期率値）

MAS5とMLANGを用いて推定された発現値を用いて、同期率を計算した。各プローブセットが両者でどのような同期率となったかを示す。

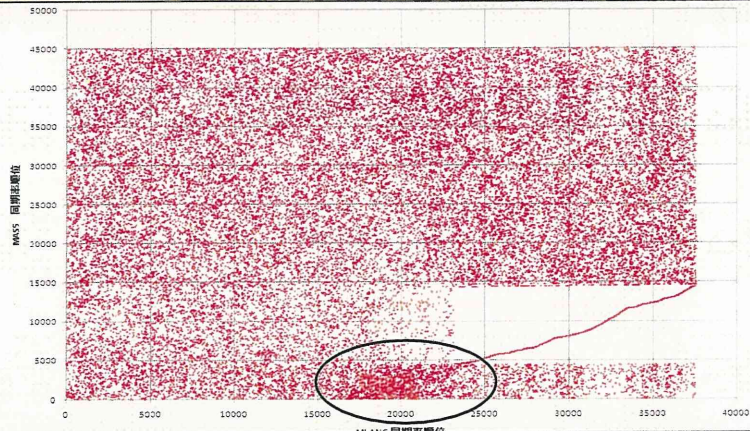


MAS5とMLANGで相関がないように思われる。

Copyright(C)2009-2010 NTT DATA Corporation

5. 3. 同期率計算結果比較（結果順位）

各プローブセットが両者でどのような同期率順位となったかを示す。



中央下のほうに若干濃い領域が存在する。ここに意味が強いものが集まっていると思われる。

Copyright(C)2009-2010 NTT DATA Corporation