# 大規模遺伝子ネットワーク推定ソフトウェア SiGN

## システムのデジタル化による細胞の薬剤応答ダイナミズムの追跡

データ解析融合研究開発チーム
宮野　悟
（東大医科研グループ）

# システムを見るって、どういうこと？

- システムのダイナミズムを追跡するには
- システムを比較するには
- システムの予測シミュレーションモデルを創るには
- システムの機能モジュールを描出するには
- システムの個性・多様性をデジタル化するには
- これらを自在につないで作戦をつくる

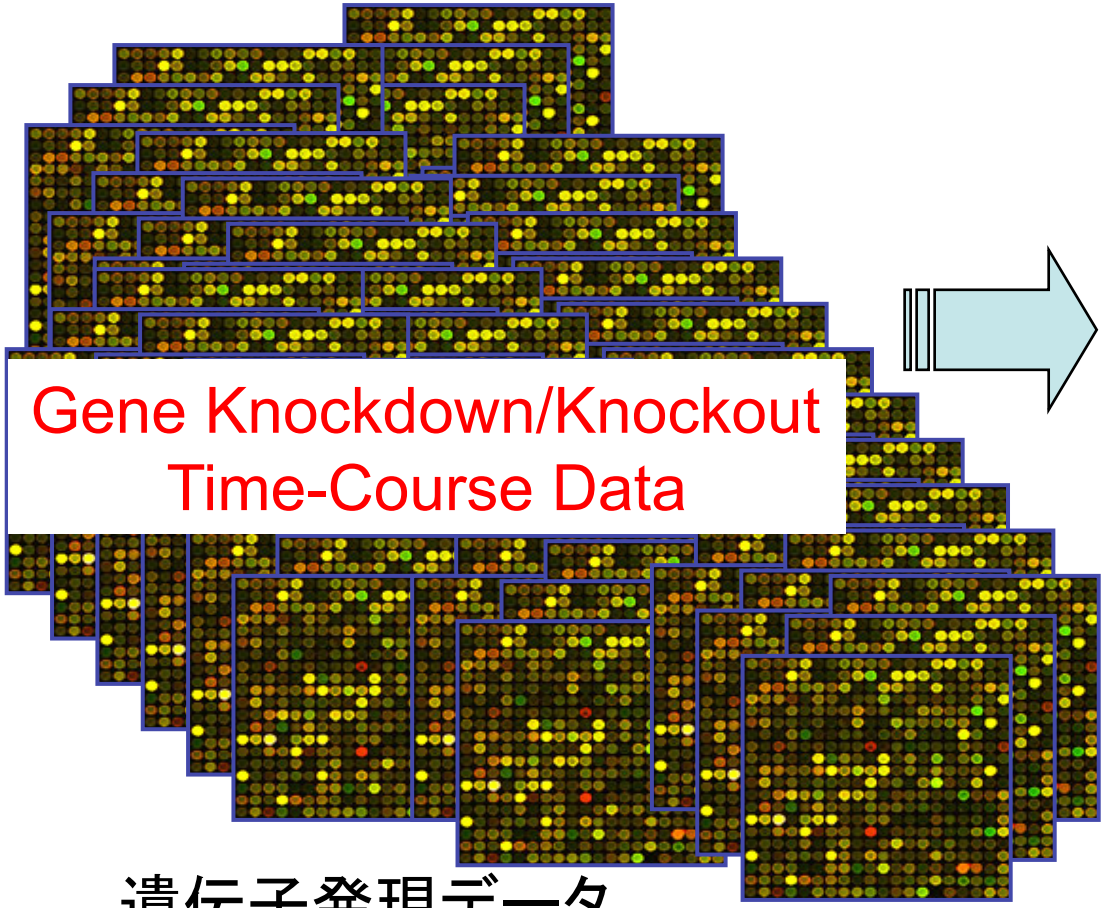
# SiGNに入っているソフトウェア

- SiGN-BN（ベイジアンネットワークモデル）
  - 動的遺伝子ネットワークアルゴリズム(Dynamic Bayesian)
  - 大規模全ゲノム遺伝ネットワーク推定アルゴリズム(NNSR)
  - 中規模高効率並列ブートストラップアルゴリズム(HC+Bootstrap)
  - 小規模並列版静的・動的全体最適化アルゴリズム(Para-OS)
- SiGN-SSM（状態空間モデル）
  - 遺伝子ネットワーク推定アルゴリズム
  - モジュールネットワーク推定アルゴリズム
  - 遺伝子発現予測アルゴリズム
- SiGN-L1（L1正則化ネットワーク）
  - NetComparator による複数ネットワーク同時推定・比較アルゴリズム
  - NetworkProfiler
  - SGGM による静的無向グラフ推定アルゴリズム
  - SVAR による時系列有向グラフ推定アルゴリズム

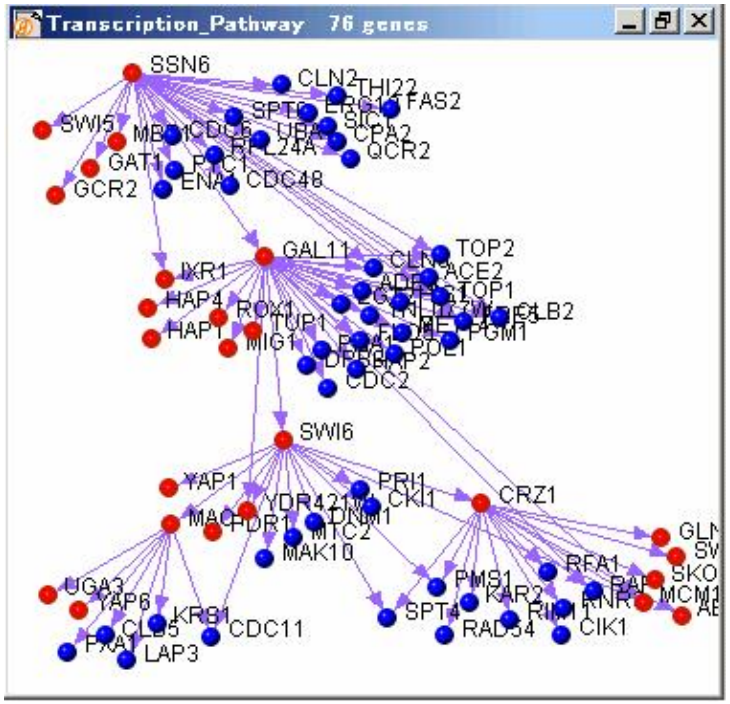
# SiGN-BN（ﾍﾞｲｼﾞｱﾝﾈｯﾄﾜｰｸﾓﾃﾞﾙ）

**Bayesian Network and Nonparametric Regression**

**ベイジアンネットワークにノンパラメトリック回帰を組み合わせた方法**



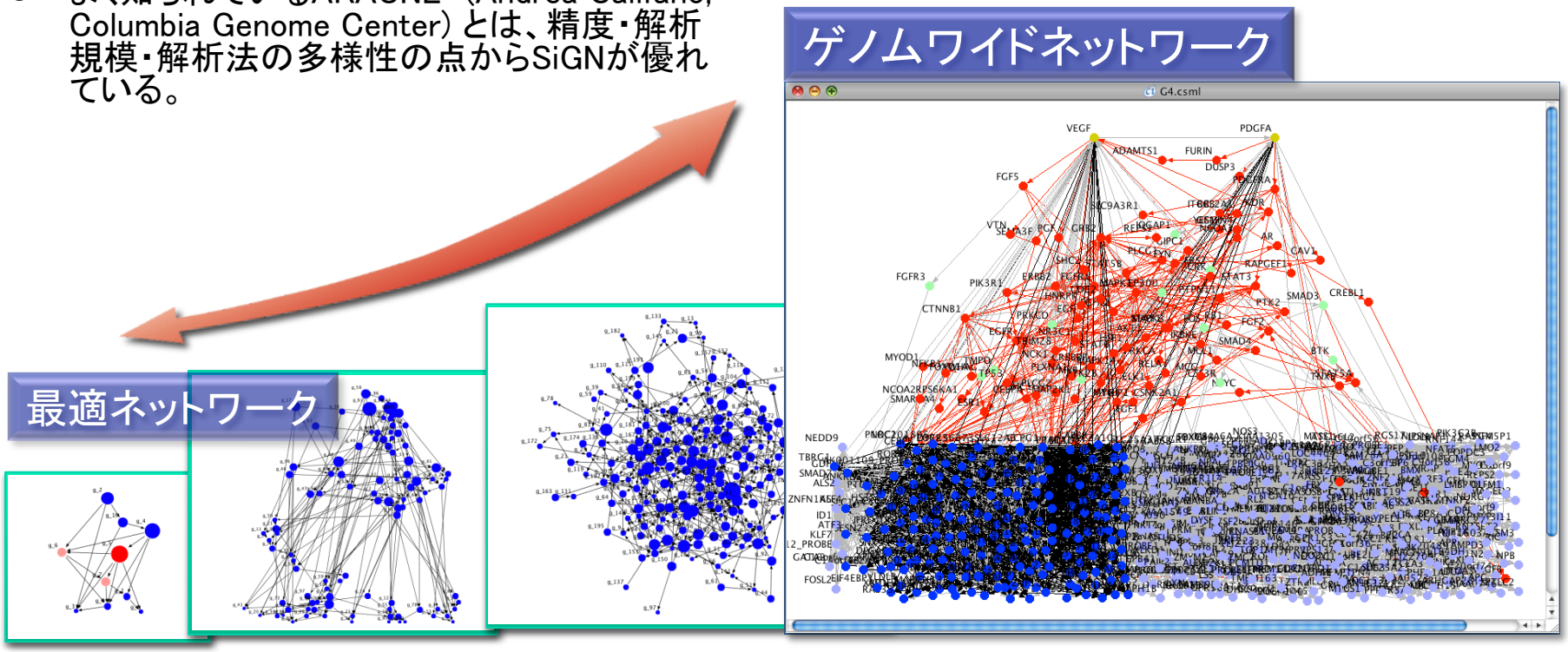

遺伝子発現データ




遺伝子ネットワーク

# 最適な数十遺伝子規模のネットワークからゲノムワイド（2万遺伝子）ネットワークまで、動的・静的ネットワークの推定が可能なソフトウェアパッケージ

- 次世代スパコンに代表される超並列計算機を用いる遺伝子ネットワーク推定プログラムパッケージは現時点ではSiGNを除いて世界に無い。
- よく知られているARACNE （Andrea Califano, Columbia Genome Center）とは、精度・解析規模・解析法の多様性の点からSiGNが優れている。



# SiGN-BNでできること

- 薬に応答して動的に変化する遺伝子ネットワークの描出（メラノーマとタキソール）
- 遺伝子ネットワークによる既存薬を用いたドラッグターゲット探索（HUVECとFenofibrateを用いたPOF）

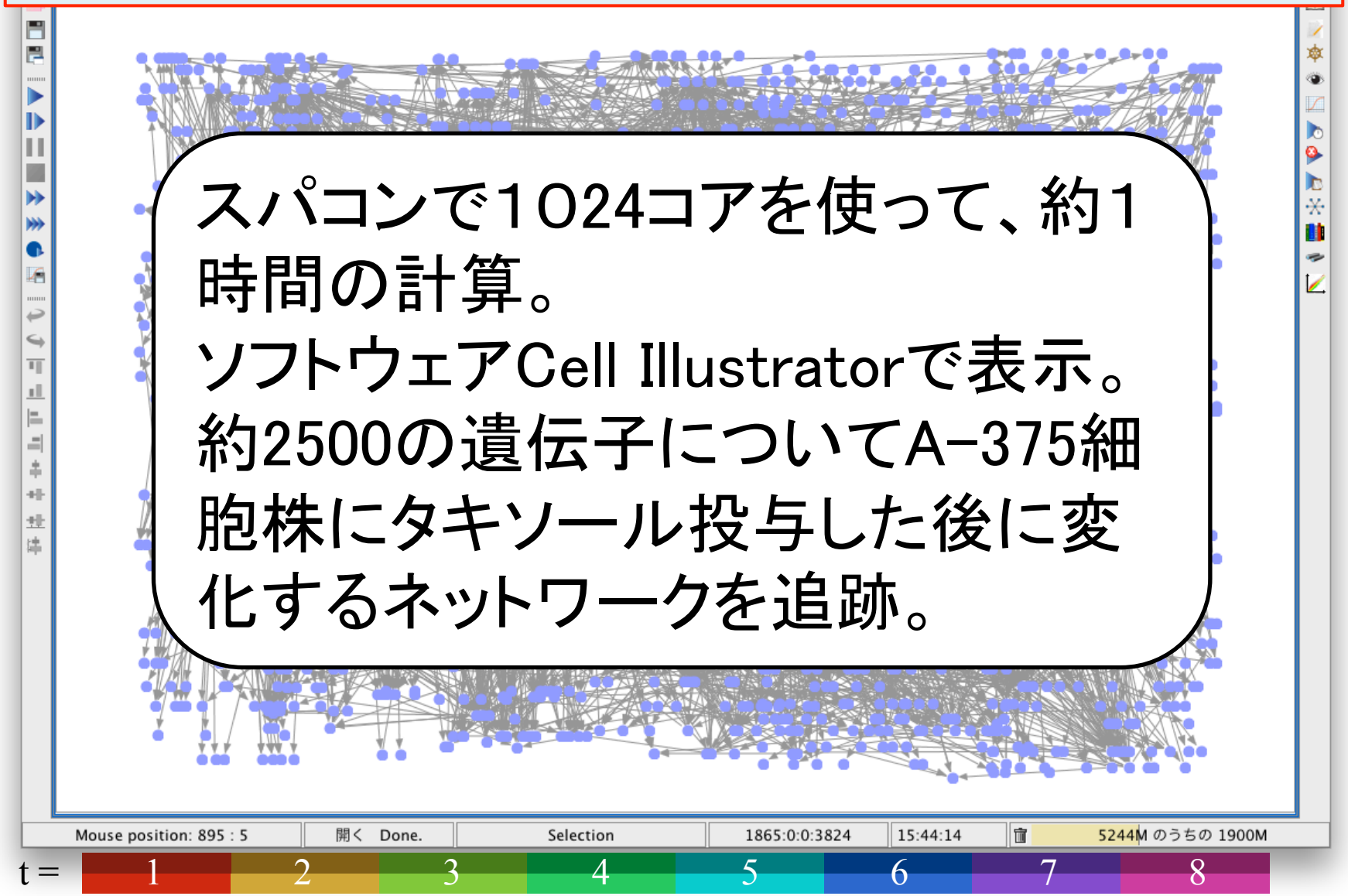
# 薬に応答して動的に変化する遺伝子ネットワークの描出
## （メラノーマ細胞の遺伝子ネットワークとタキソール）
**具体的にどんなことが今の段階でできているのか－ダイナミックベイジアン＋非線形回帰**

- 細胞：メラノーマ細胞株 A-375
- 薬：タキソール(Paclitaxel、抗がん剤）
- Microarray：Illumina® HumanHT-12 v3
- データ：14 時点、3回計測、計84アレイデータ
  - 0h, 15min, 30min, 45min, 1h, 1.5h, 2h, 3h, 4h, 6h, 8h, 12h, 18h, 24h
  - タキソール投与と非投与の２種類のデータ
- 目的：タキソールを投与したメラノーマ細胞の遺伝子ネットワークの動的変化を見ることで、新たなターゲット遺伝子を探索

SiGN-BN

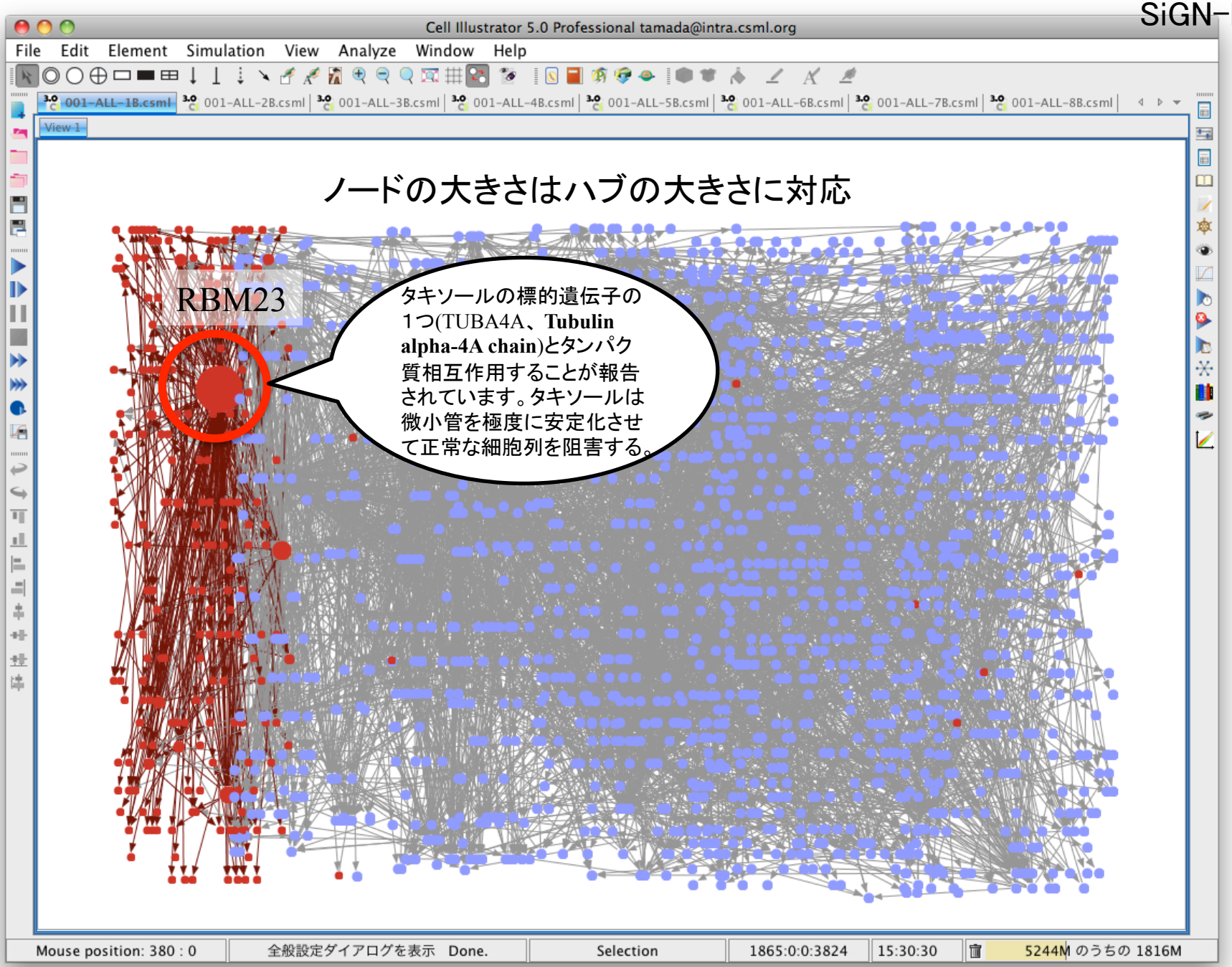
# メラノーマ細胞の遺伝子ネットワークとタキソール

## ダイナミックベイジアン＋非線形回帰によるネットワーク推定の一例



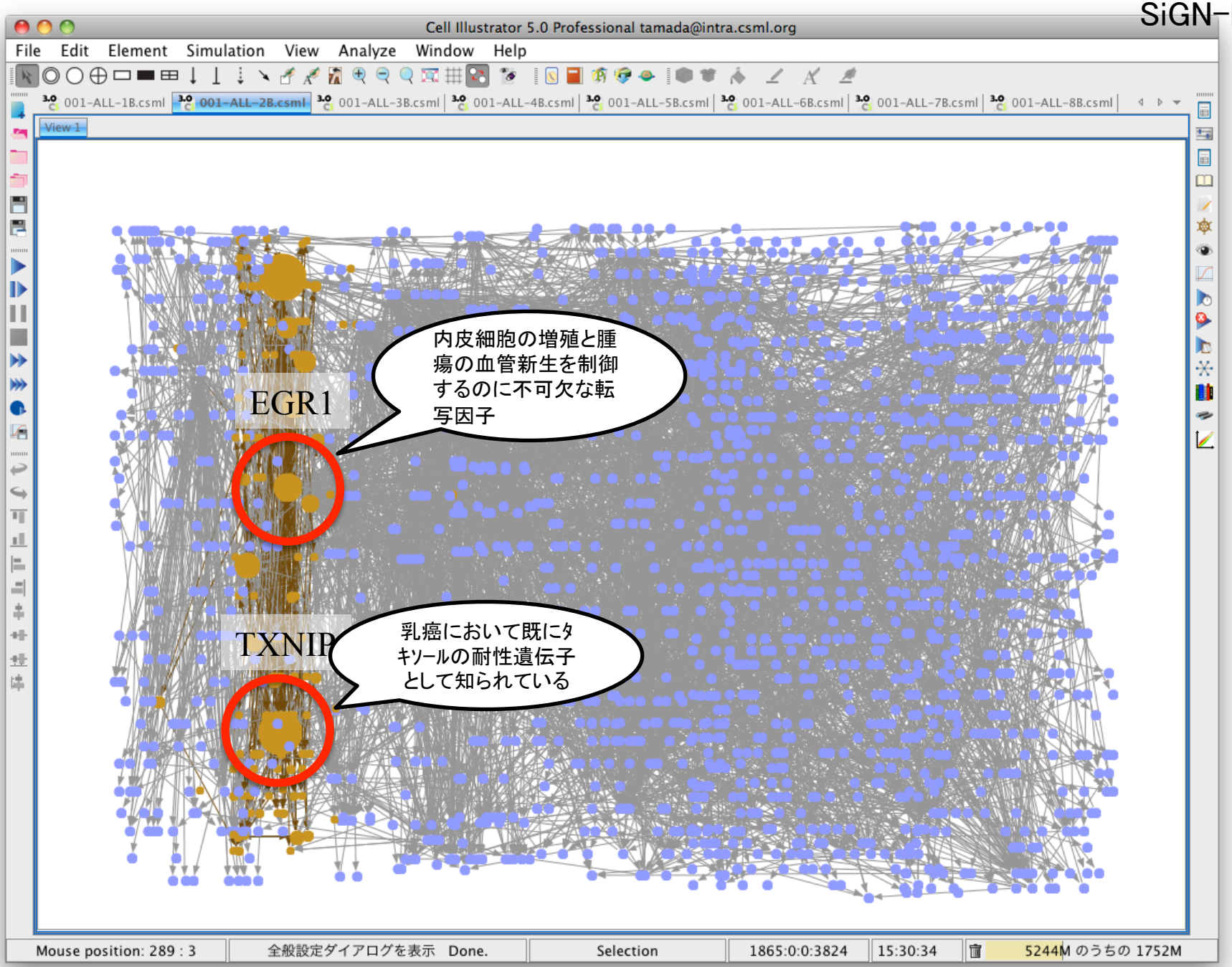

スパコンで1024コアを使って、約1時間の計算。
ソフトウェアCell Illustratorで表示。
約2500の遺伝子についてA-375細胞株にタキソール投与した後に変化するネットワークを追跡。

SiGN-BN



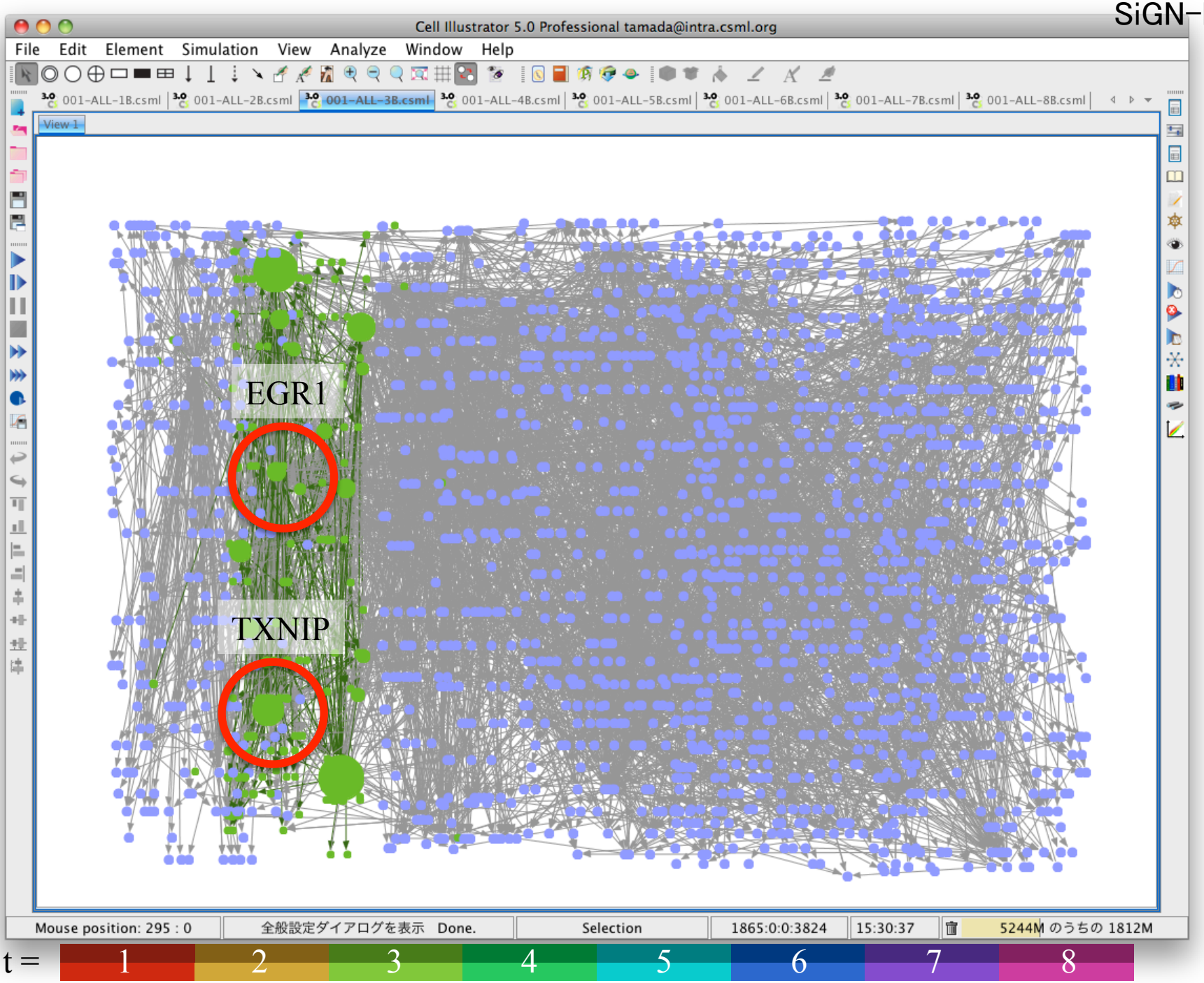

t = 1 2 3 4 5 6 7 8

SiGN-BN



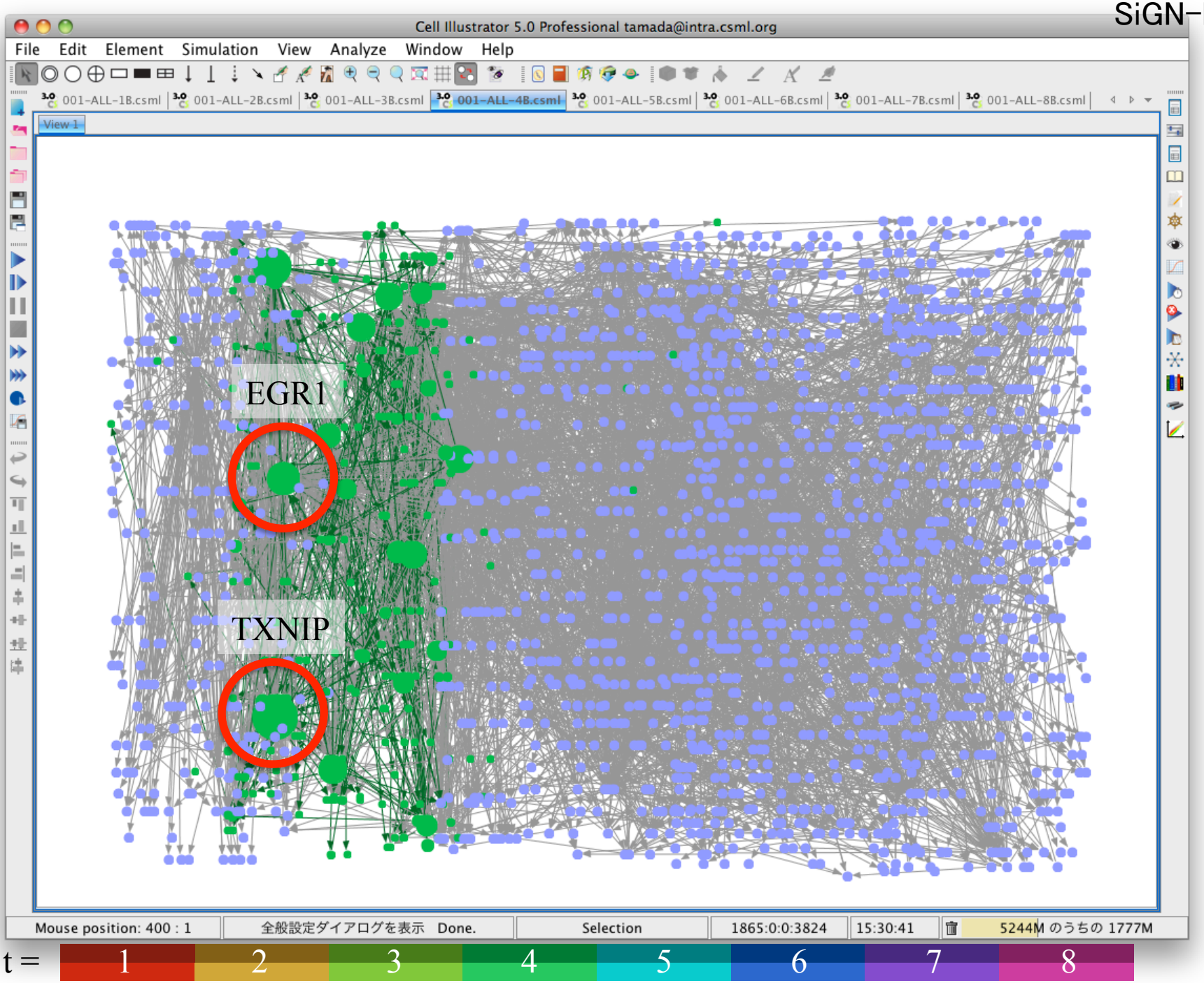

t = 1 2 3 4 5 6 7 8

SiGN-BN



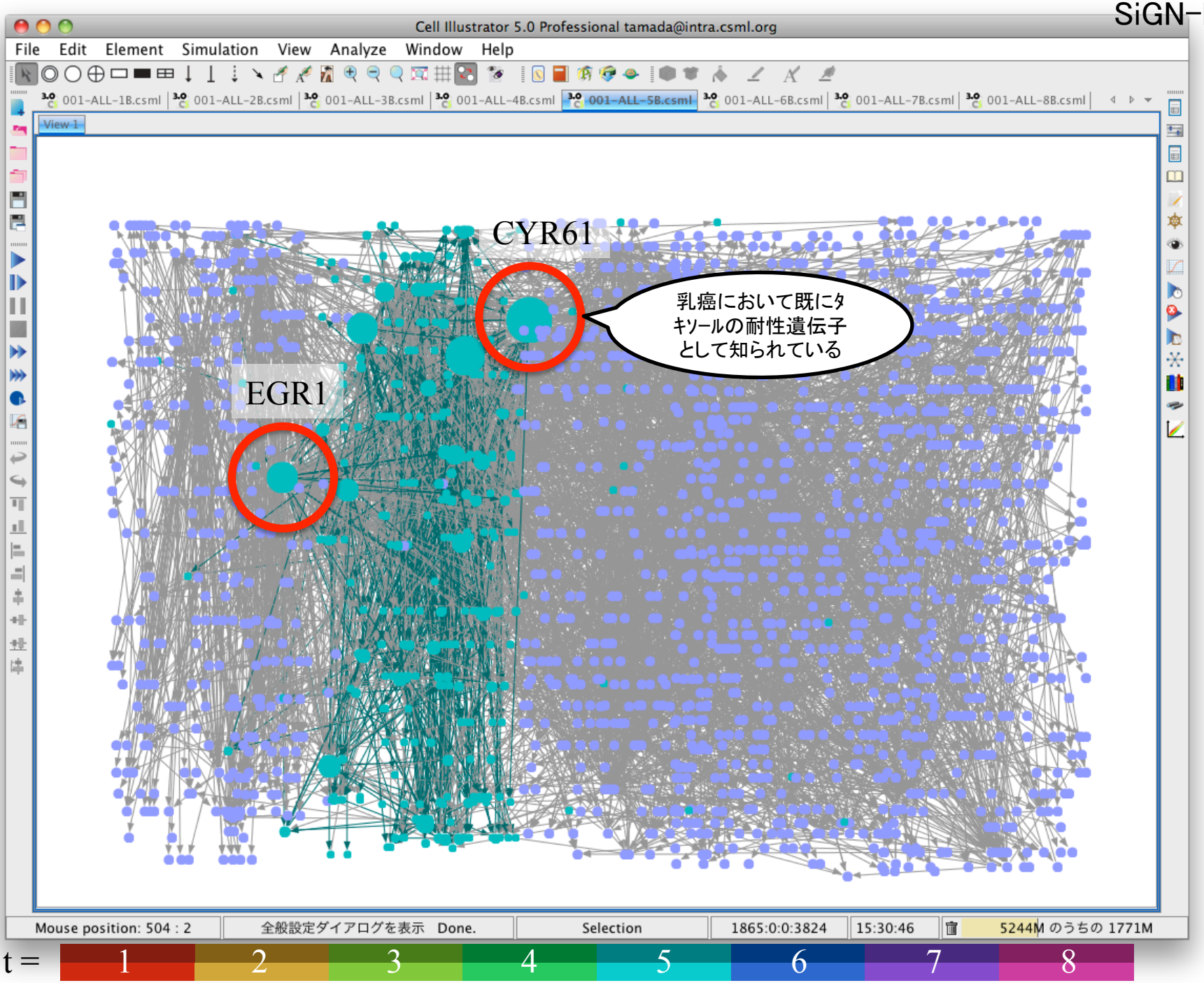

t = 1 2 3 4 5 6 7 8

SiGN-BN



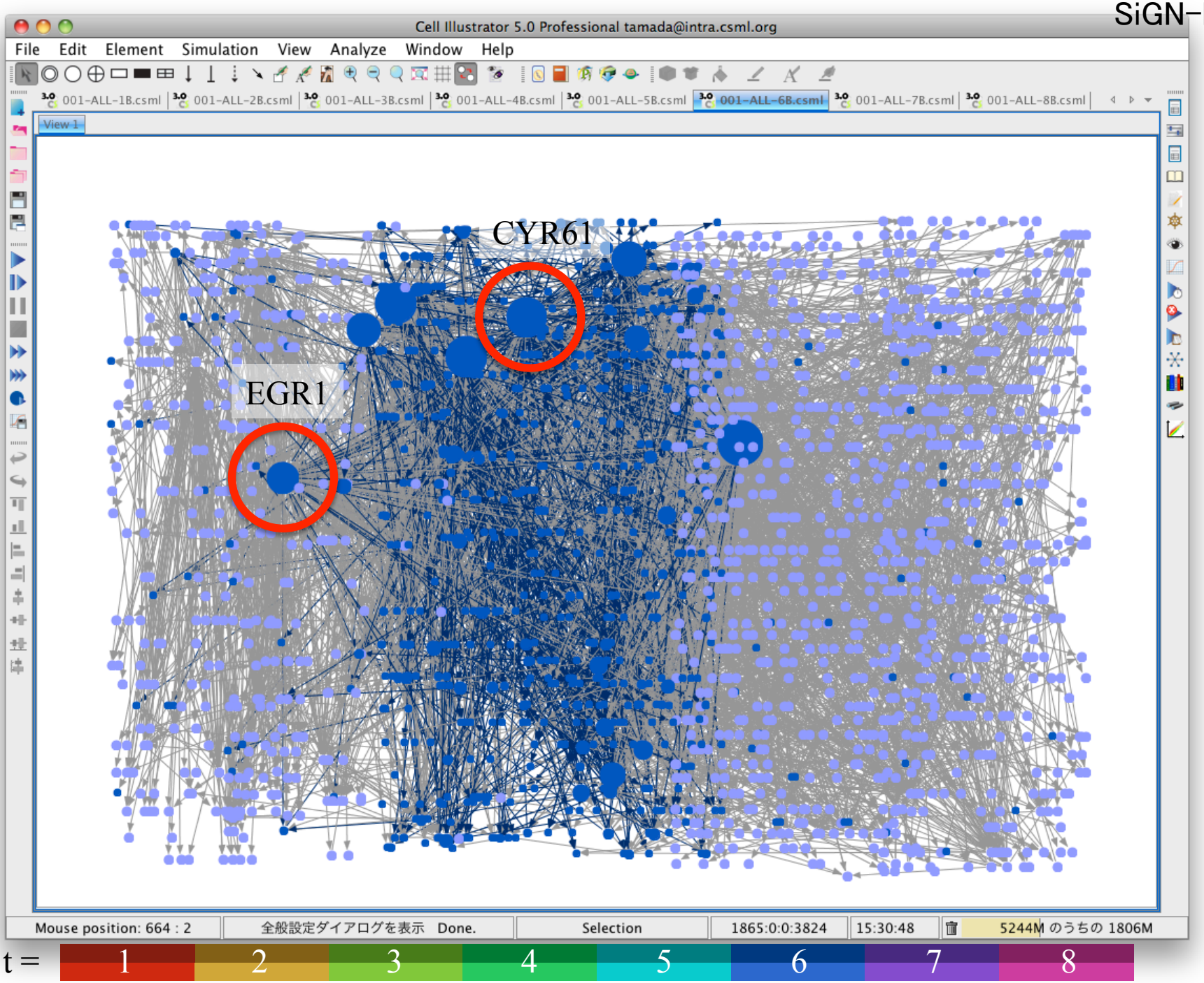

t = 1 2 3 4 5 6 7 8

SiGN-BN



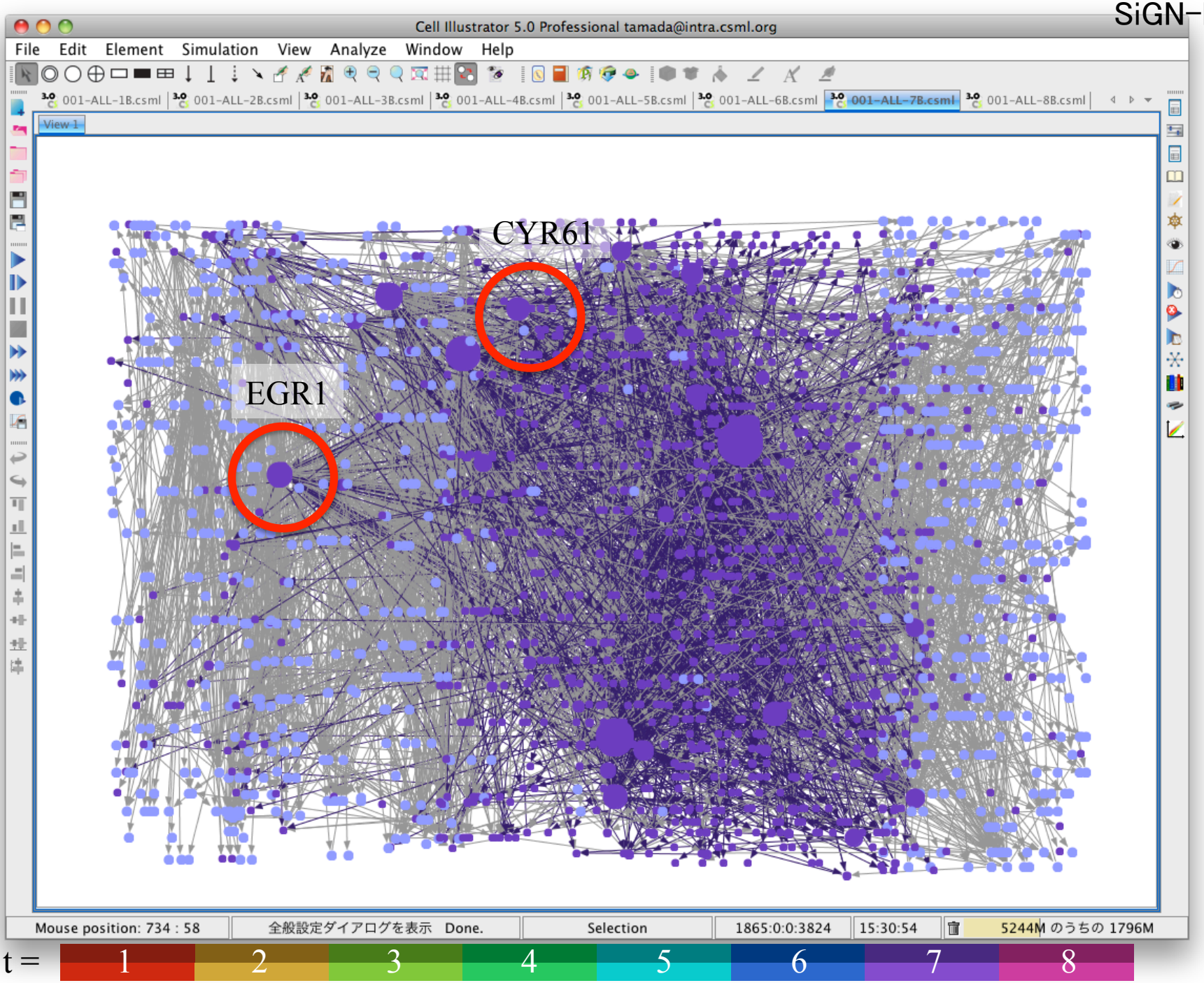

t = 1 2 3 4 5 6 7 8

SiGN-BN



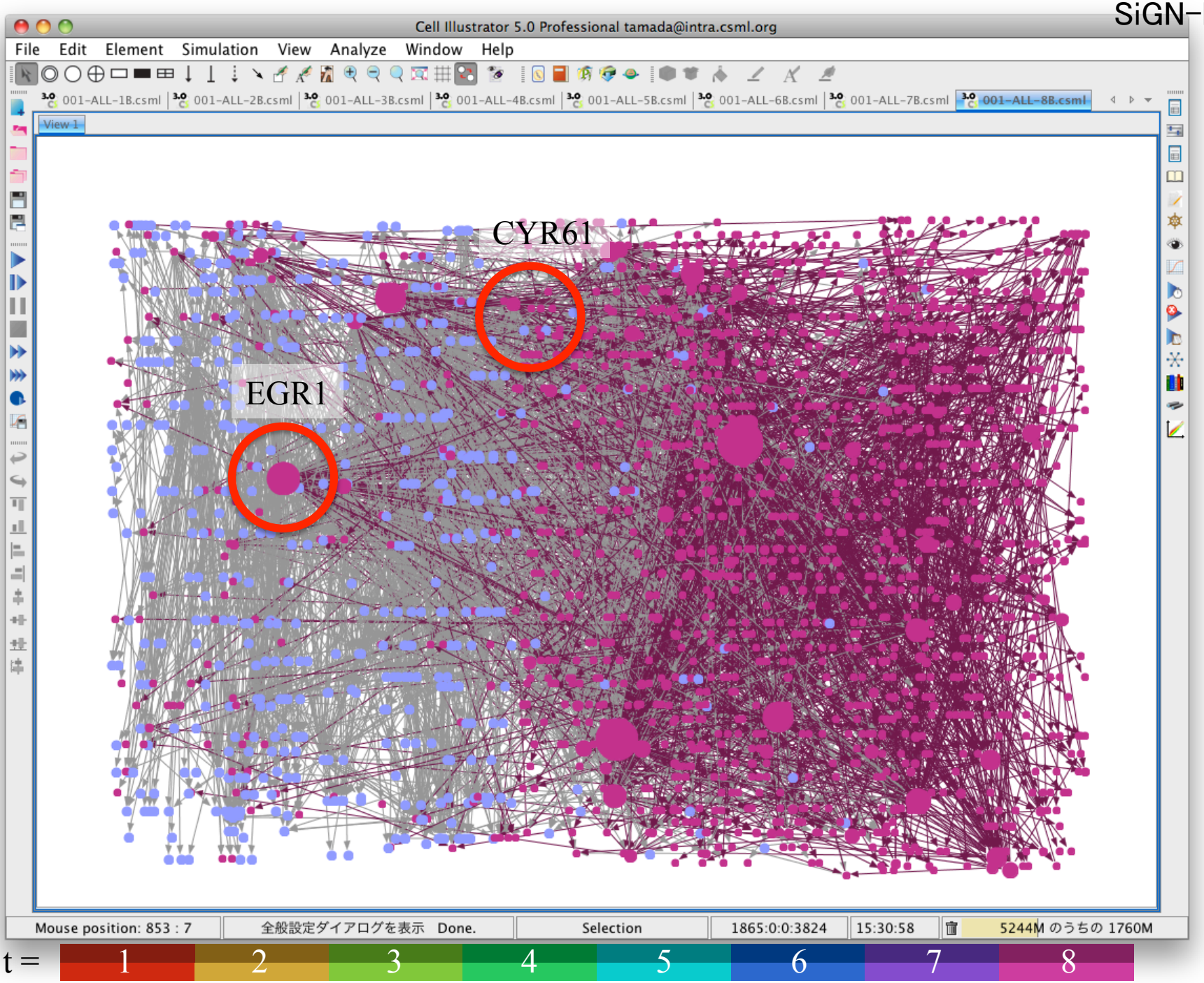

t = 1 2 3 4 5 6 7 8

SiGN-BN



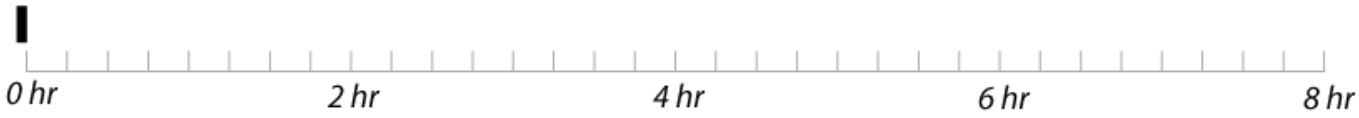

t = 1 2 3 4 5 6 7 8

SiGN-BN



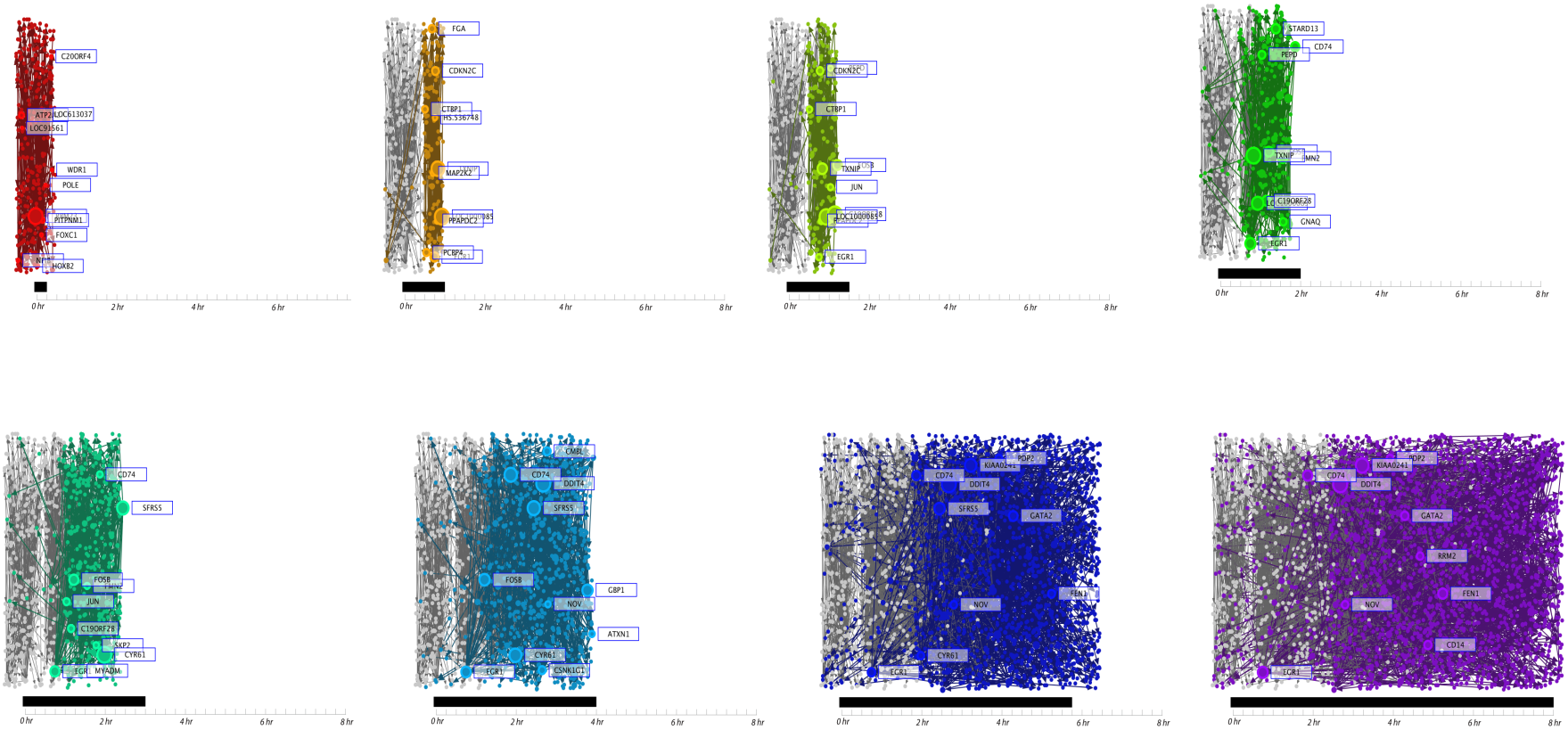

t = 1 2 3 4 5 6 7 8

SiGN-BN



SiGN-BN

# がんシステムのダイナミズムを追跡　SiGN
## Dynamic Bayesian Network

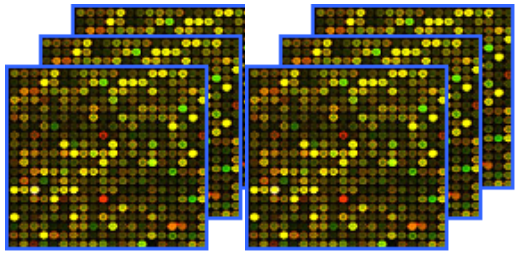


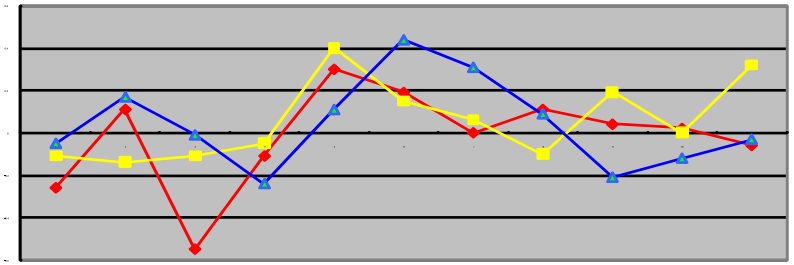
# 遺伝子ﾈｯﾄﾜｰｸによる既存薬を用いたﾄﾞﾗｯｸﾞﾀｰｹﾞｯﾄ探索（HUVEC遺伝子ネットワークとFenofibrate）

**具体的にどんなことが今の段階でできているのか－静的ﾍﾞｲｼﾞｱﾝﾈｯﾄﾜｰｸ＋非線形回帰**

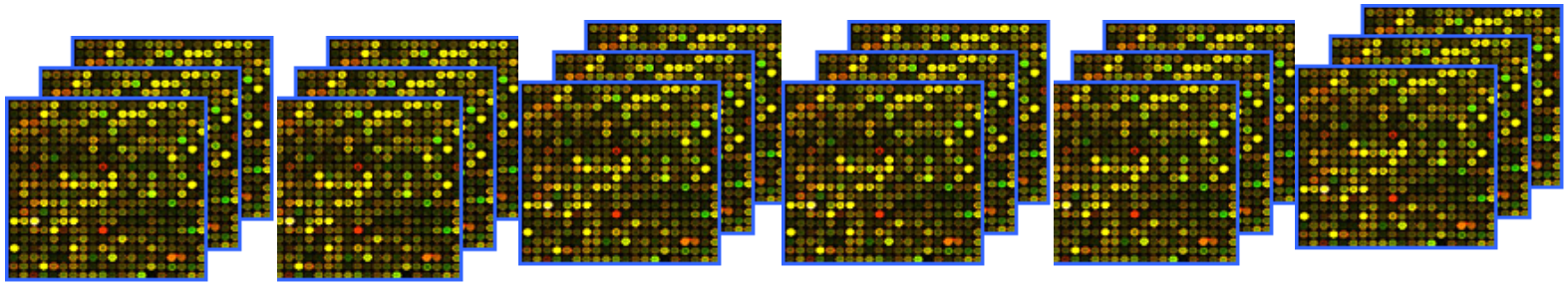
- 細胞：HUVECサンプル（Cambridge大学Rosie Hosptial）
- 薬：Fenofibrate(脂質代謝異常症薬)（PPARαのagonist）
- Microarray：Code Link 20K
- データ
  - siRNAによる遺伝子ノックダウン遺伝子発現データ（270遺伝子ノックダウン）
  - Fenofibrate (25μM)をHUVECに投与して、6時点（0, 2, 4, 6, 8,18 h)、2回計測
- 目的：Fenofibrateに応答して動く遺伝子群の遺伝子制御ネットワーク構造の描出と、PPARαの下流ネットワークの探索

# 2種類の遺伝子発現データ





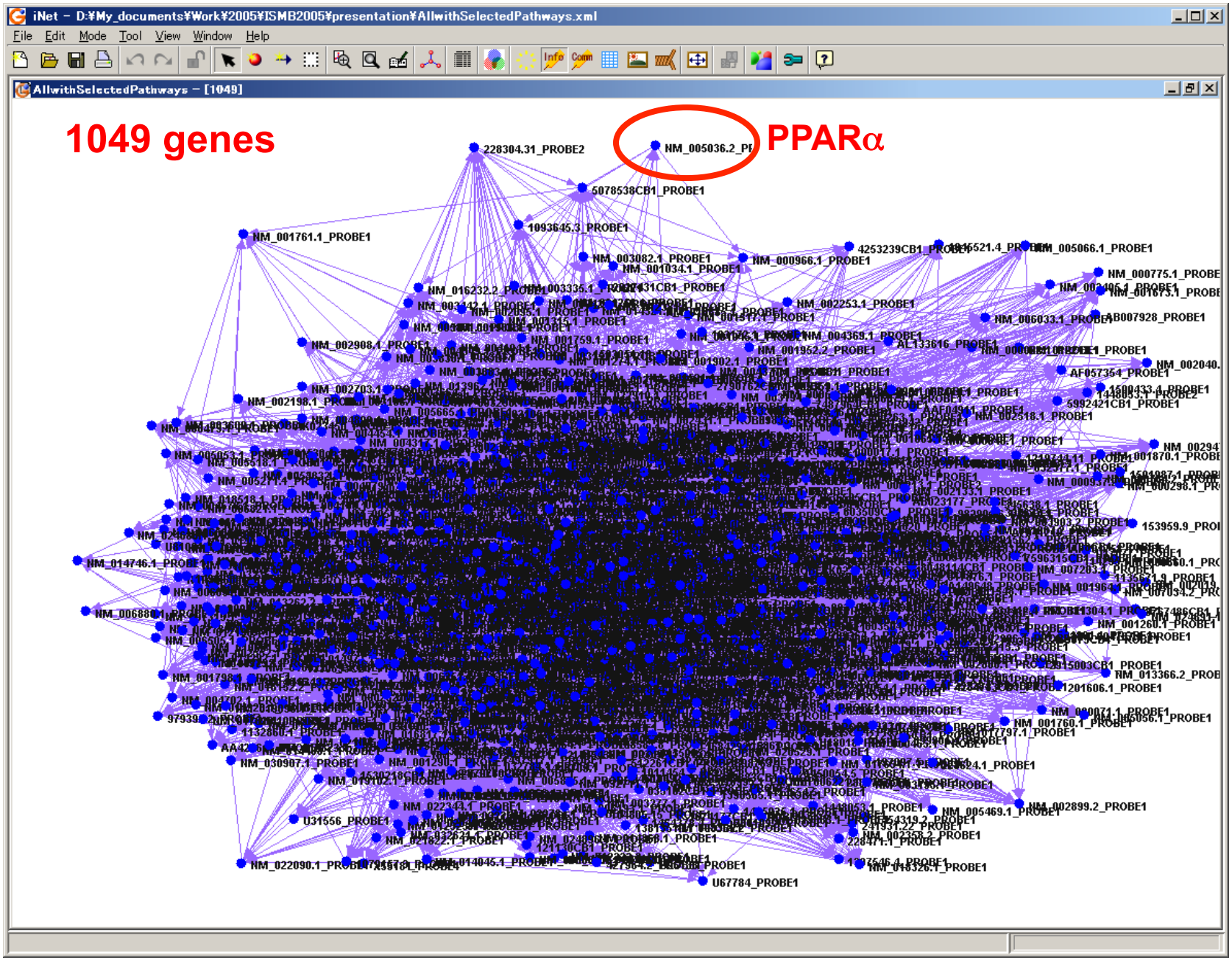
1. Fenofibrateを投与して時系列遺伝子発現を計測したマイクロアレイデータ



2. 270種の遺伝子をノックダウンしたマイクロアレイデータ

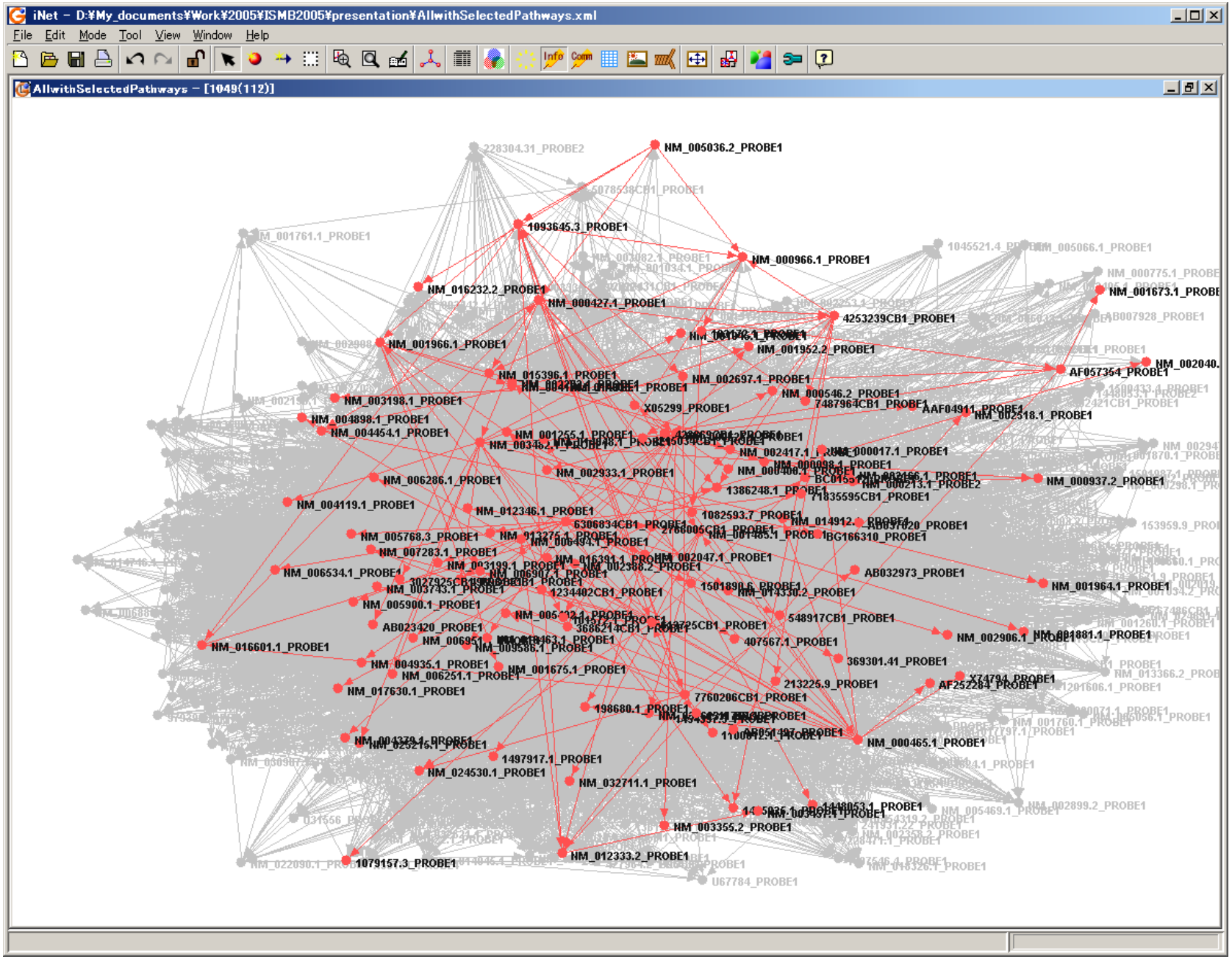
SiGN-BN

# Fenofibrateに反応した遺伝子のネットワーク

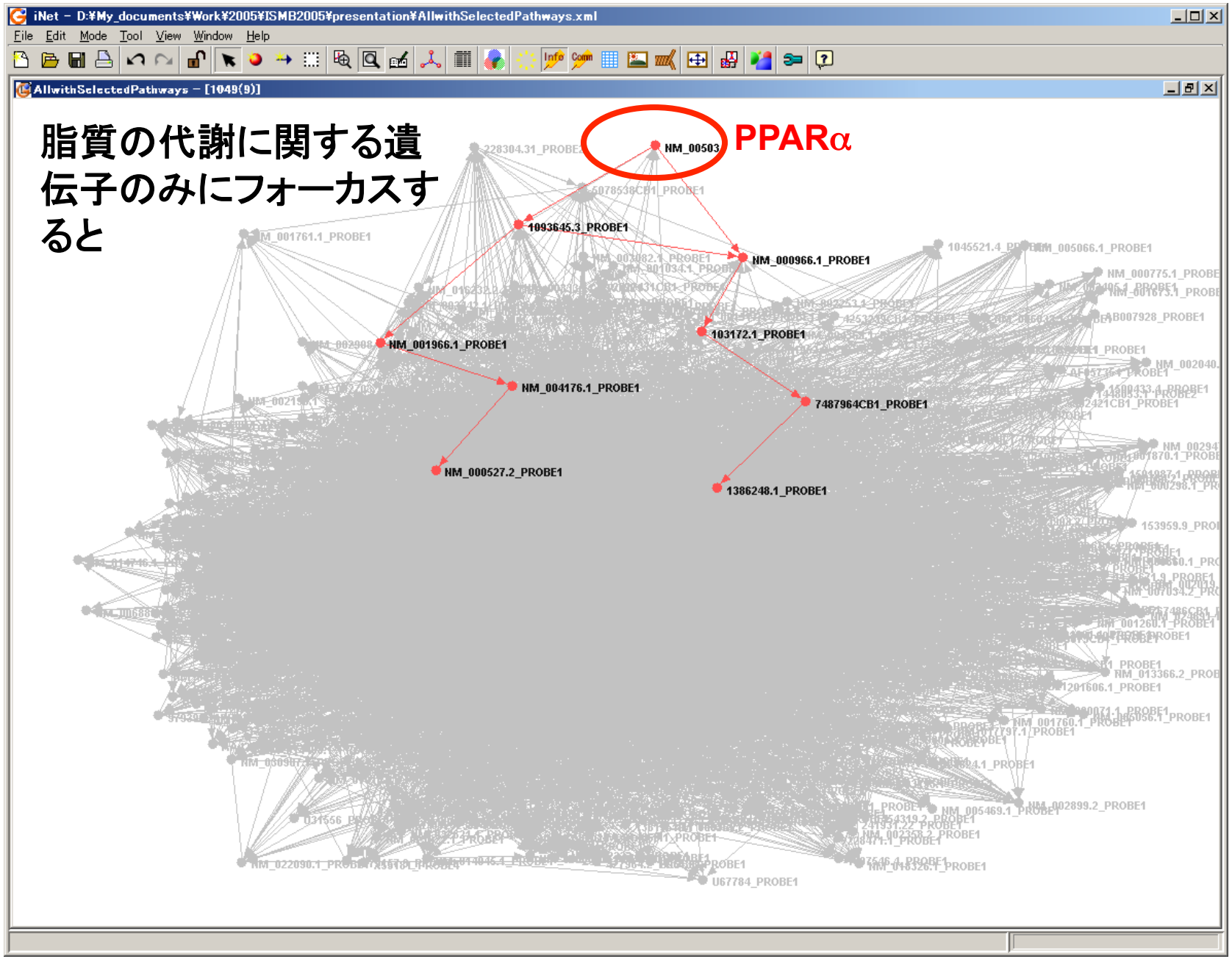


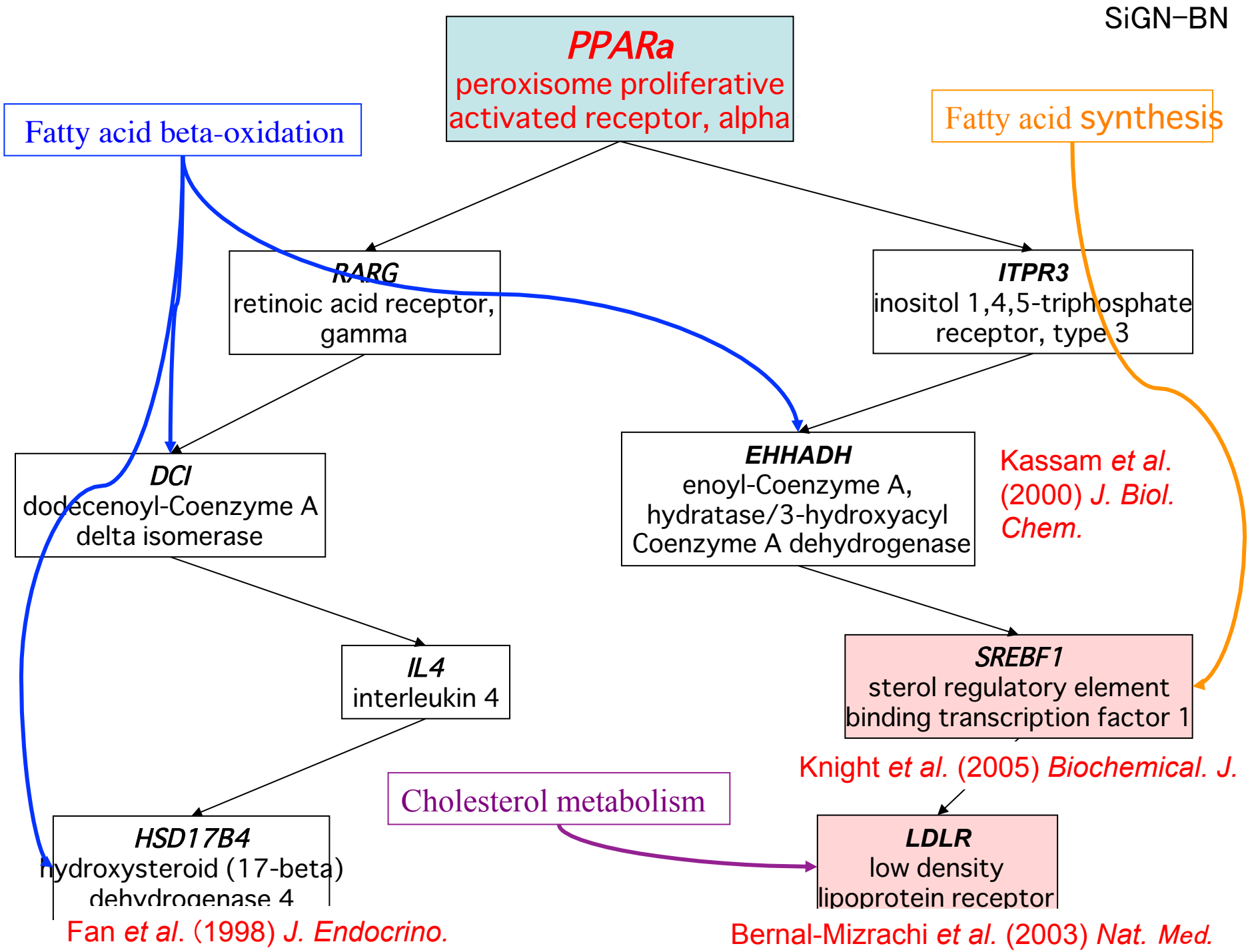
SiGN-BN

# PPARαの影響下にある遺伝子群



# 評価


iNet - D:¥My_documents¥Work¥2005¥ISMB2005¥presentation¥AllwithSelectedPathways.xml
File Edit Mode Tool View Window Help
AllwithSelectedPathways - [1049(9)]
脂質の代謝に関する遺伝子のみにフォーカスすると
PPARα
NM_00503
1093645.3_PROBE1
NM_000966.1_PROBE1
NM_001966.1_PROBE1
103172.1_PROBE1
NM_004176.1_PROBE1
7487964CB1_PROBE1
NM_000527.2_PROBE1
1386248.1_PROBE1


SiGN-BN
PPARa
peroxisome proliferative activated receptor, alpha
Fatty acid beta-oxidation
Fatty acid synthesis
RARG
retinoic acid receptor, gamma
ITPR3
inositol 1,4,5-triphosphate receptor, type 3
DCI
dodecenoyl-Coenzyme A delta isomerase
EHHADH
enoyl-Coenzyme A, hydratase/3-hydroxyacyl Coenzyme A dehydrogenase
Kassam et al. (2000) J. Biol. Chem.
IL4
interleukin 4
SREBF1
sterol regulatory element binding transcription factor 1
Knight et al. (2005) Biochemical. J.
Cholesterol metabolism
HSD17B4
hydroxysteroid (17-beta) dehydrogenase 4
LDLR
low density lipoprotein receptor
Fan et al. (1998) J. Endocrino.
Bernal-Mizrachi et al. (2003) Nat. Med.

SiGN-BN

# 遺伝子ネットワークを利用した創薬ターゲット遺伝子のイン・シリコ探索

| Children# | GeneTitle | GeneName | Druggable | Description |
|---|---|---|---|---|
| 30 | Gene1 | Gene1 | | |
| 22 | Gene2 | Gene2 | | |
| 20 | Gene3 | Gene3 | | |
| 16 | LIPG | lipase, endothelial | Druggable | NatGenet 21:424-8 |
| 15 | Gene4 | Gene4 | | |
| 13 | Gene5 | Gene5 | | |
| 13 | HMGCR | 3-hydroxy-3-methylglutaryl-Coenzyme A reductase | Druggable | Many Company's Targets |
| 11 | Gene6 | Gene6 | | |
| 10 | Gene7 | Gene7 | Druggable | |
| 9 | Gene8 | Gene8 | | |
| 9 | Gene9 | Gene9 | | |
| 9 | LSS | lanosterol synthase (2,3-oxidosqualene-lanosterol cyclase) | Druggable | Roche's Target |
| 8 | AKR1C3 | aldo-keto reductase family 1, member C3 | Druggable | CancerRes 63:505-512 |
| 7 | PTGS2(COX2) | prostaglandin-endoperoxide synthase 2 | Druggable | Many Company's Targets |
| 7 | Gene10 | Gene10 | | |
| 7 | PPARA | peroxisome proliferative activated receptor, alpha | Druggable | Fournier Pharma's Target |

# HMG-CoA Reductase Inhibitors

- Major drug target for cholesterol lowering and hyperlipidemia
  - Fluvastatin@Novartis
  - Pravastatin@Sankyo
  - Simvastatin@Merck
  - Lovastatin @Merck
  - Atorvastatin@Pfizer

1: Am Fam Physician. 1993 May 15;47(7):1623-7. Related Articles, Links

**HMG CoA reductase inhibitors for treatment of hyperlipidemia.**

Kantner TR.

Texas Tech University Health Sciences Center, Lubbock.

Lowering serum cholesterol levels in patients with hyperlipidemia has been shown to decrease deaths due to coronary heart disease. While diet is the mainstay of treatment for hyperlipidemia, drug therapy is necessary to achieve treatment goals in certain patients. The HMG CoA reductase inhibitors are the most effective cholesterol-lowering agents currently available. These agents are well tolerated and have few serious side effects. Their major disadvantage is their high cost.

# COX-2/PTGS2 Inhibitors

***COX-2 inhibitors cause heart disease and cerebral stroke!!***

The NEW ENGLAND JOURNAL of MEDICINE

AJP - Regulatory, Integr
and Comparative Physio
HOME HELP FEEDBACK SUBSCR
Institution: KYUSHU UNIV 00A
*Am J Physiol Regul Integr Comp*
0363-6119/05 $8.00

EDITORIAL FOCUS

**Vioxx, Celebrex, Bextra....Do we have a new target for anti-inflammatory and antipyretic therapy?**

**Andrej A. Romanovsky**

*Systemic Inflammation Laboratory, Trauma Research, St. Joseph's Hospital, Phoenix, Ari*

TOGETHER WITH THE ARTICLES by Ivanov et al. (11) in the previous issue and by Thompson et al.
*American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, the
opens the Call for Papers on Physiology and Pharmacology of Temperature Regulation. Th
indispensable role for the prostaglandin (PG) $E_2$-synthesizing enzyme, microsomal PGE sy
models of experimental fever: interleukin (IL)-1$\beta$ induced and turpentine induced. These fi
the context of the major breakthroughs and setbacks that have marked the development of a
inflammatory pharmacotherapy in the recent past.

HELP

Next ▸

THIS ARTICLE
- Return to Search Result
- PDF
- PDA Full Text

COMMENTARY
- Letters

b, and valdecoxib. The Adenomatous Polyp Prevention on Vioxx (APPROVe)
h a history of colorectal adenomas, was stopped early because rofecoxib doubled the
r events (relative risk, 1.92; 95 percent confidence interval, 1.19 to 3.11). These
eased risk of myocardial infarction previously seen in the Vioxx Gastrointestinal

THE LANCET
Volume 359, Issue 9315 , 20 April 2002, Page 1410
doi:10.1016/S0140-6736(02)08393-9 Cite or Link Using DOI
Copyright © 2002 Published by Elsevier Ltd.

News

**Why do COX-2 inhibitors increase risk of cardiovascular events?**

**Helen Frankish**

Available online 3 May 2002.

US researchers have this week shed light on a mechanism that could explain why some new generation non-steroidal anti-inflammatory drugs (NSAIDs), which selectively inhibit cyclo-oxygenase 2 (COX-2) but not COX-1, might increase the rate of cardiovascular events compared with traditional NSAIDs, which inhibit both COX-1 and COX-2.

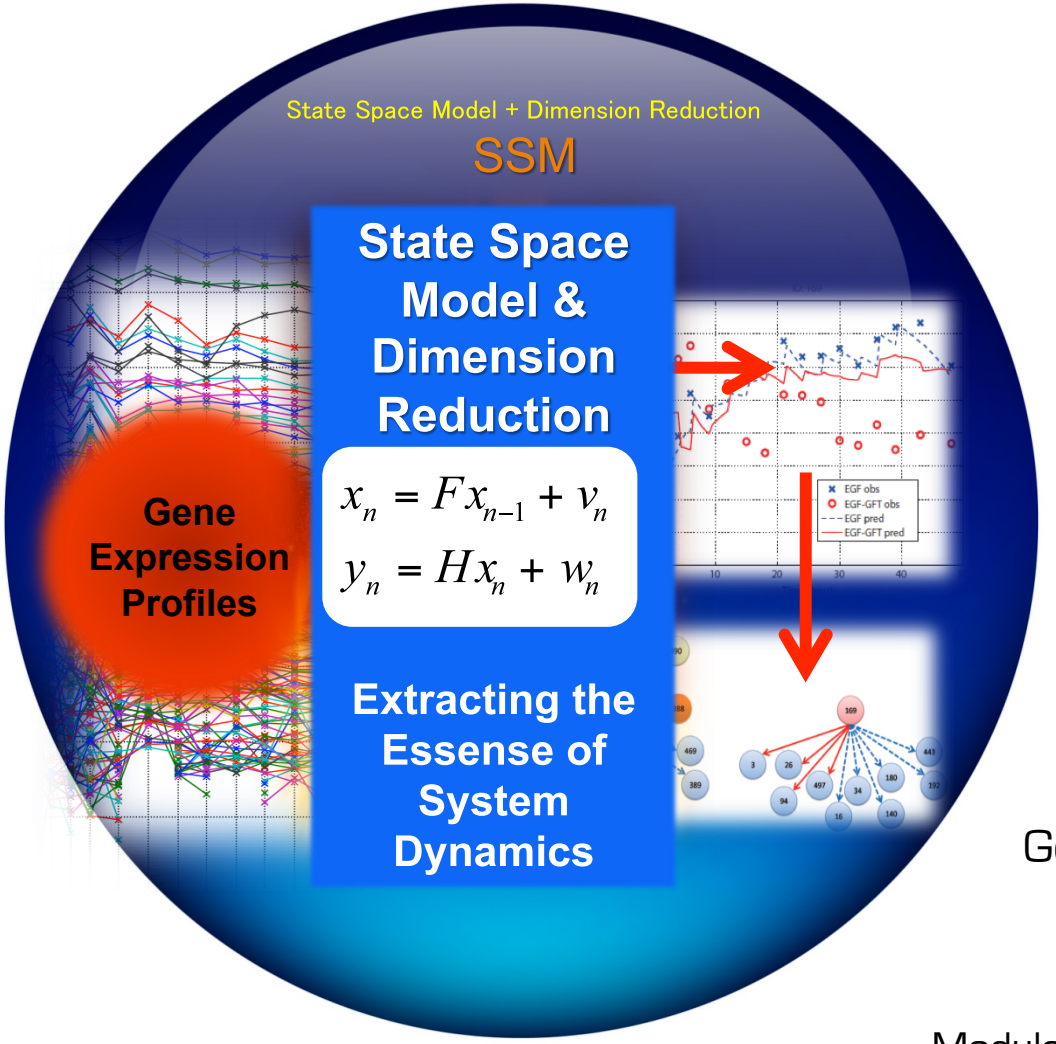
"Attention to this issue was drawn by the outcome of a comparison of rofecoxib, a COX-2 inhibitor, and naproxen, a COX-1 and COX-2 inhibitor, in the VIGOR trial", explains lead researcher Garret FitzGerald (University of Pennsylvania School of Medicine, Philadelphia, PA, USA). The VIGOR investigators reported that rofecoxib treatment increased the incidence of myocardial infarction five-fold compared with patients who received naproxen (*N Engl J Med* 2000; 343: 1520–28).

**Various COX-2 inhibitors**

- **Vioxx@Merck**
- **Celebrex@Pfizer**
- **Bextra@Pfizer**

# SiGN-SSM（状態空間モデル）

## 状態空間モデルと次元圧縮による時系列データからの予測モデルの構築




State Space Model:

$$x_n = Fx_{n-1} + v_n \in R^k \quad \text{System Model}$$

$$y_n = Hx_n + w_n \in R^p \quad \text{Observation Model}$$

High-Dimensional Short Time Series Data:

$$Y = \{y_1, \cdots, y_{N_{\text{obs}}}\} \qquad N_{\text{obs}} << p \approx 10^3$$

System Estimation with Dimension Reduction:

$$\dim(x_n) = k < \dim(y_n) = p$$

$$L(\theta) = \int p(Y, X \mid \theta) dX$$

$$H^T R^{-1} H = \Lambda = \text{diag}(\lambda_1, \cdots, \lambda_k)$$

Gene Expression Prediction:

$$y_{n|n-1} = H \int x_n p(x_n \mid y_1, \cdots, y_{n-1}, \theta) dx_n$$

Module-Based Gene Network Estimation:

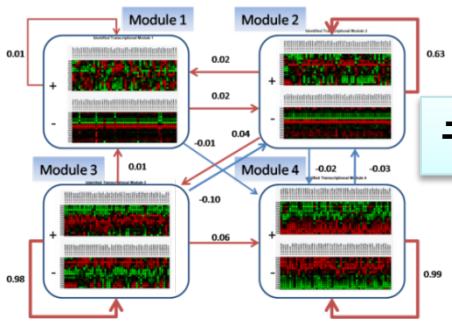
$$R^{-1/2}(y_n - w_n) = \Psi R^{-1/2}(y_{n-1} - w_{n-1}) + R^{-1/2} H v_n$$

# 状態空間モデルによる遺伝子制御ネットワークの推定

SSM

$$x_n = F x_{n-1} + v_n,\ n \in \mathcal{N},$$
$$y_n = H x_n + w_n,\ n \in \mathcal{N}_{\mathrm{obs}},$$

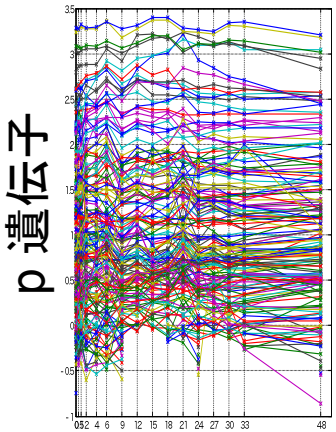
$$\mathrm{SSM}(\theta) \quad \theta = \{H, F, R, x_0\}$$




モジュール抽出

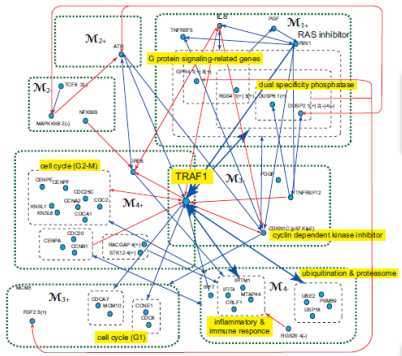
時系列遺伝子発現データ

高次元(p遺伝子)
少数時点データ



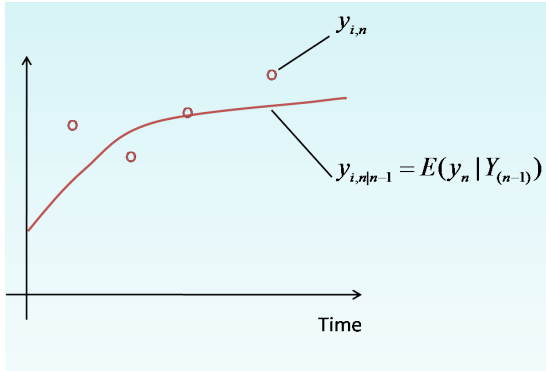

$N_{obs} << p$ $\quad$ $N_{obs}$ 時点数

$$Y_{\mathcal{N}_{\mathrm{obs}}} = \{y_n\},\ n \in \mathcal{N}_{\mathrm{obs}}$$




遺伝子ネットワーク

予測モデル



# SiGN-SSMでできること

- 薬に応答して動く遺伝子ネットワーク予測モデルの構築とシミュレーションによる薬の効果ポイントの推定（肺上皮細胞(SAEC)とイレッサ）

⇒　薬の投与で変化するシステムの違いを同定

# 薬の投与でシステムのどこが変わったのかをシミュレーションで抽出（SAECとイレッサとEGF）

**具体的にどんなことが今の段階でできているのか－状態空間モデルによるシミュレーション**

- 細胞：SAEC
- 薬：イレッサ
- 刺激：EGF
- Microarray：Agilent
- データ：19時点（48時間）
  - SAECをEGFで刺激した時系列データ（コントロールデータ）
  - SAECにイレッサを投与したあとEGFで刺激した時系列データ（ケースデータ）
- 目的：イレッサ投与と非投与、それぞれのもとで、EGFで刺激した時の時系列遺伝子発現データから大規模遺伝子ネットワークの予測モデルを構築し、シミュレーションによりシステムの違いを抽出

SiGN-SSM

**SAEC #52** mRNA

(hours)

- 24 | - 2 | 0 | 0.5 | 1 | 2 | 4 | 6 | 9 | 12 | 15 | 18 | 21 | 24 | 27 | 30 | 33 | 36 | 39 | 43 | 48

Starvation Start

Gefitinib (0.5 μM) Treatment

EGF (100 ng/ml) Stimulation

RNA Sampling

( - ) : 0 hr ~ 48 hr

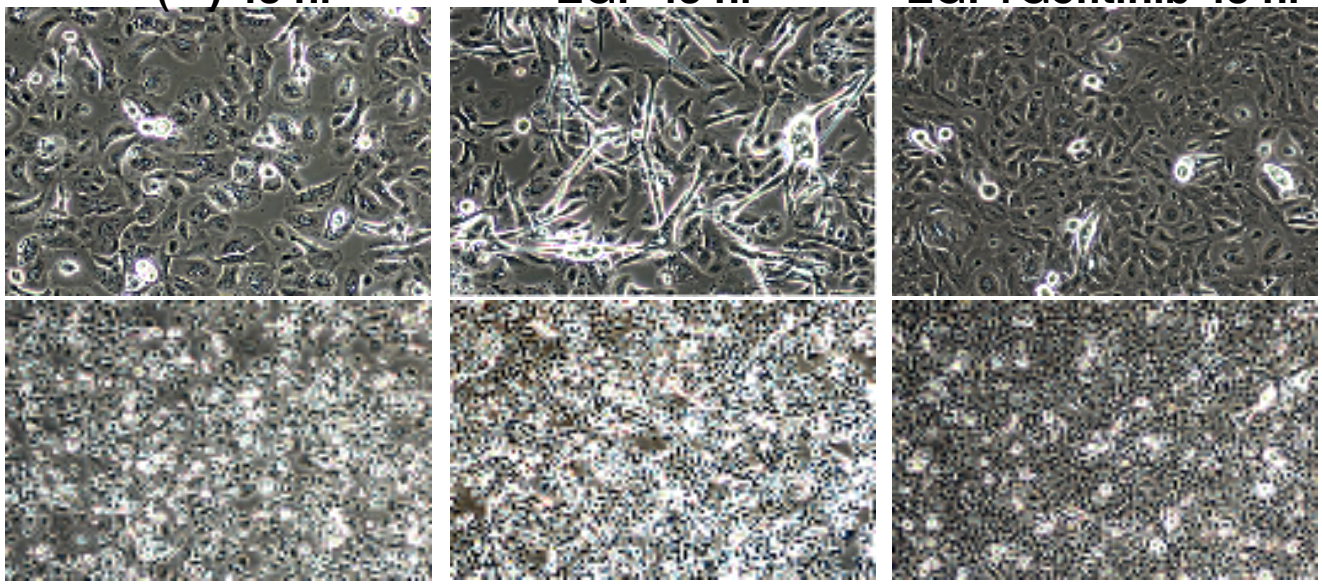
EGF : 0 hr ~ 48 hr

EGF+Gefitinib : 0 hr ~ 48 hr

Gefitinib : 0 hr ~ 9 hr

## Cell Morphology

( - ) 48 hr | EGF 48 hr | EGF+Gefitinib 48 hr



# Strategy to Predict Differentially Regulated Genes

1. Train SSM by CTRL time-course data and Estimate an CTRL-SSM System

$x_n = Fx_{n-1} + v_n$

$y_n = Hx_n + w_n$

$Y_{N_{obs}}^{CTRL} = [y_1^{CTRL}, \cdots, y_{N_{obs}}^{CTRL}]$

$\hat{\theta}^{CTRL} = [\hat{H}^{CTRL}, \hat{F}^{CTRL}, \hat{R}^{CTRL}]$

2. Predict CASE time-course data by the CTRL-SSM System

$\hat{\theta}^{CTRL}$

$Y_{n-1}^{CASE} = [y_1^{CASE}, \cdots, y_{n-1}^{CASE}]$

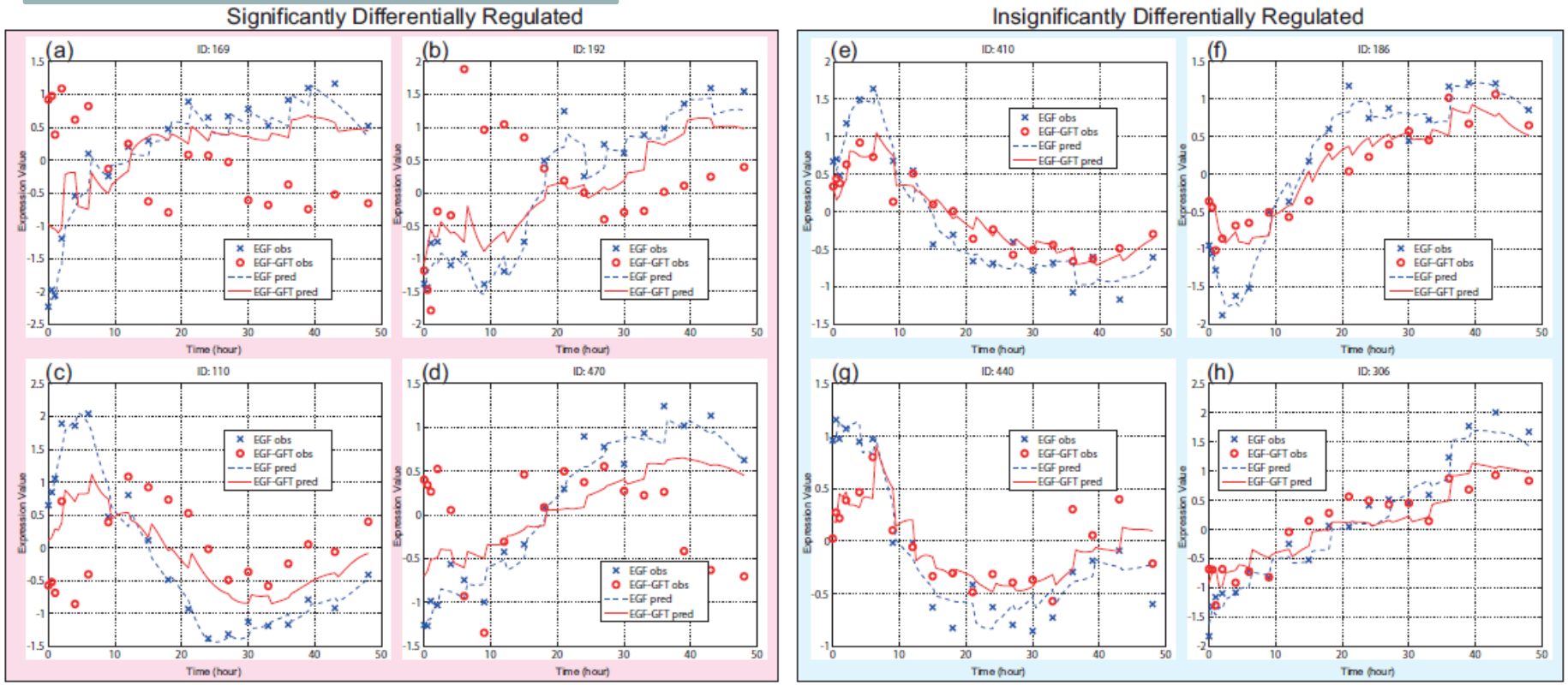
Predicted Gene Expression at *n*

3. Compare Predicted and Observed Time Courses

$n = 1, \cdots, N_{obs}$

Candidates of Differentially Regulated Genes:
**Unpredictable genes in the Case data by the Control's SSM**.

X : EGF obs (Control)
O : EGF+GFT obs (Case)
--- : EGF pred by SSM(EGF)
__ : EGF+GFT pred by SSM(EGF)

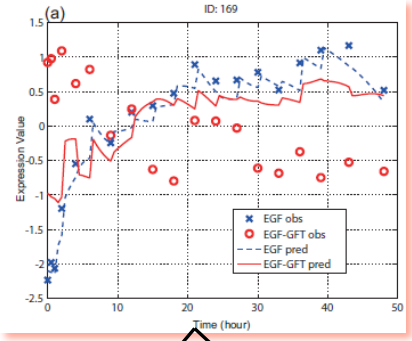



Prediction of
EGF-GFT by SSM(EGF) (red solid line)
deviate from the observations (circles)

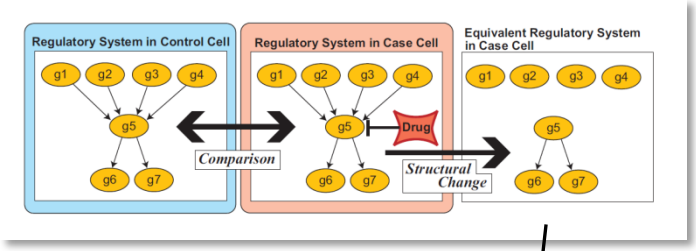
Prediction of
EGF-GFT by SSM(EGF) (red solid line)
trace the observations (circles)

# 状態空間モデルで推定されたネットワーク

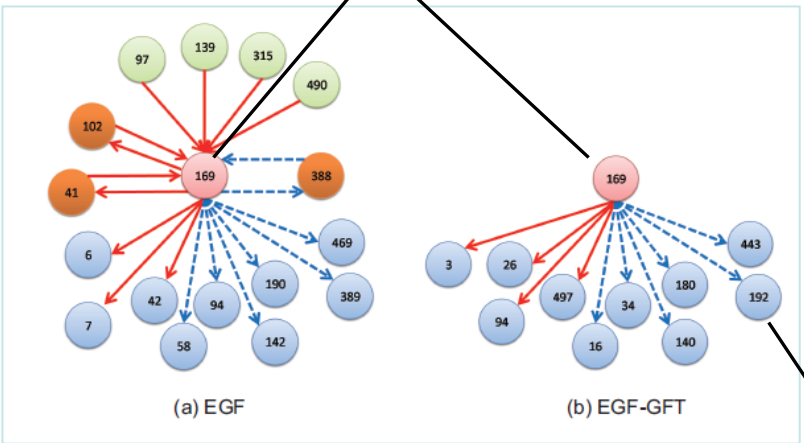
Diff gene
(gene 162)




$$R^{-1/2}(y_n - w_n) = \Psi R^{-1/2}(y_{n-1} - w_{n-1}) + R^{-1/2} H v_n$$



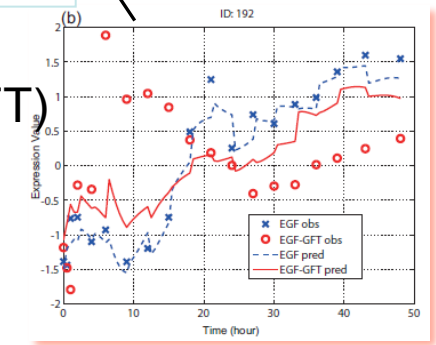

Predicted disappearance of parent nodes by drug dosing




Control
(EGF)

Case
(EGF+GFT)

Diff gene
(gene 192)



# SiGN-L1（L1正則化ネットワーク）

- スパース学習法による遺伝子ネットワーク推定
  - L1-regularization for parameter estimation
    - Graphical Gaussian models
    - Vector autoregressive models
- アルゴリズム
  - Weighted lasso (~ 100,000 molecules)
  - Recursive elastic net (~100,000 molecules)
  - Relevance-weighted recursive elastic net (~ 1,000)

# SiGN-L1でできること

- NetComparator

遺伝子ネットワーク推定から遺伝子ネットワーク比較へ

  - 異なる成長因子の刺激によって誘導される乳癌細胞株MCF-7の遺伝子ネットワーク比較）

- NetworkProfiler

構造方程式モデルで見る細胞の個性・多様性

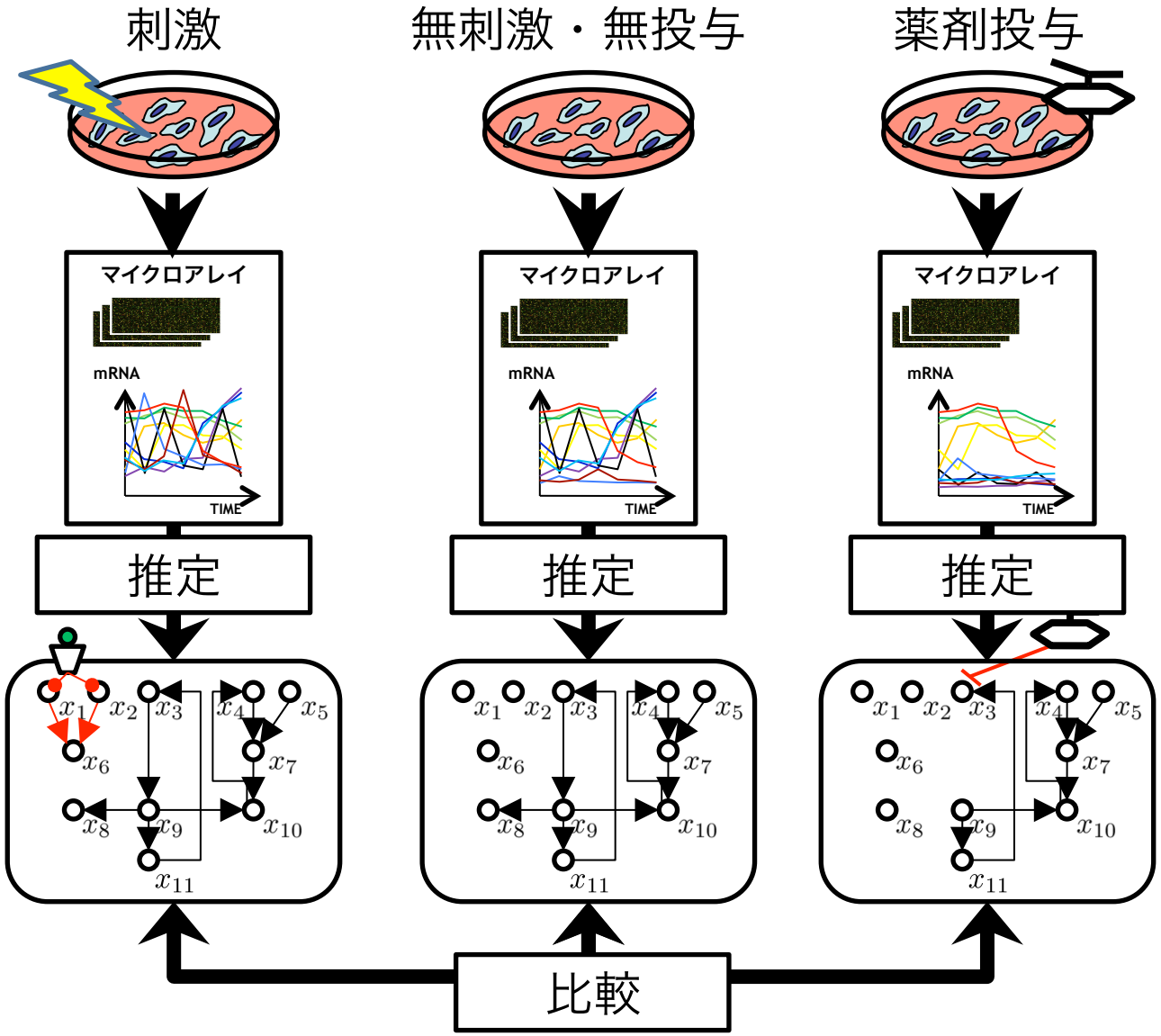
  - 大規模臨床サンプルから炙り出す個人個人のがんの遺伝子ネットワーク
  - EMTで変化する遺伝子ネットワークを描出

# NetComparator

SiGN-L1

遺伝子ネットワーク推定から遺伝子ネットワーク比較へ

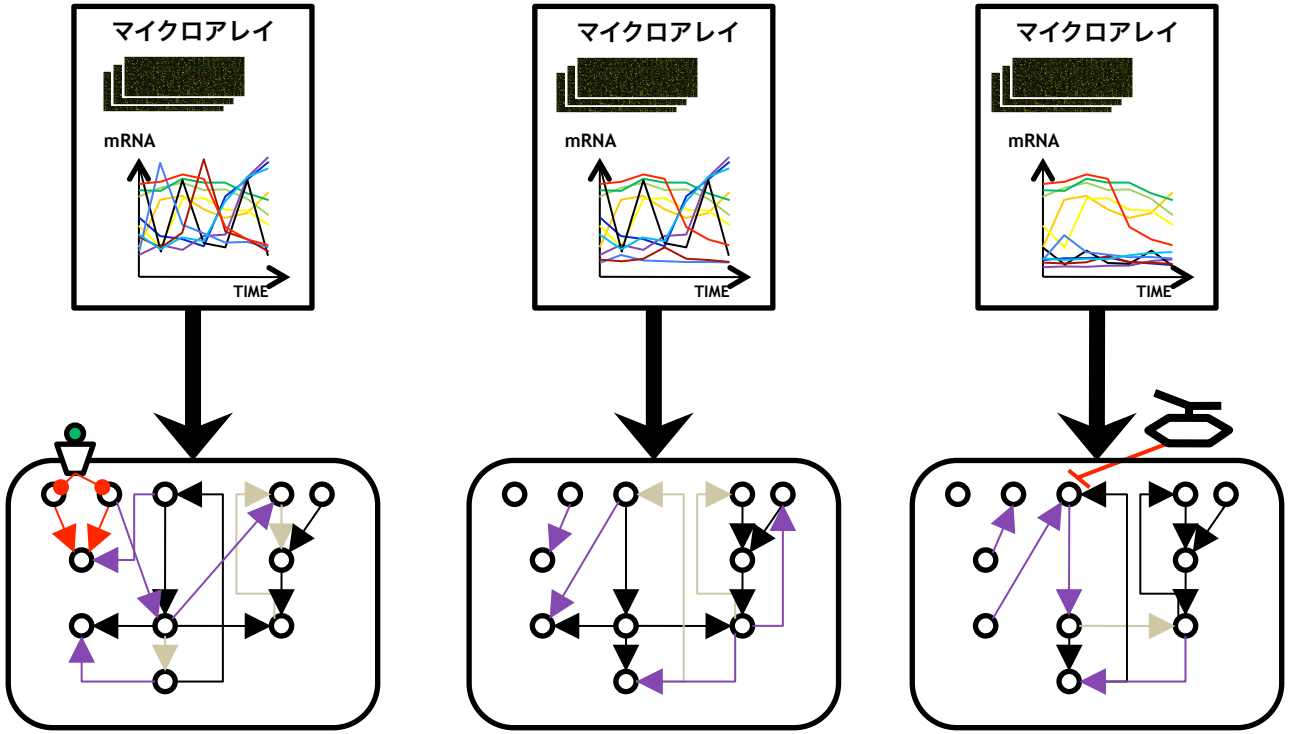
データから細胞内のシステムをリバースエンジニアリングする




**細胞内システムの違い**がどこにあるかを見つけたい

# 問題点

各実験条件で観測される**時点数（標本数）**<<　**変数（遺伝子数）**
（多くの場合、１０以下）

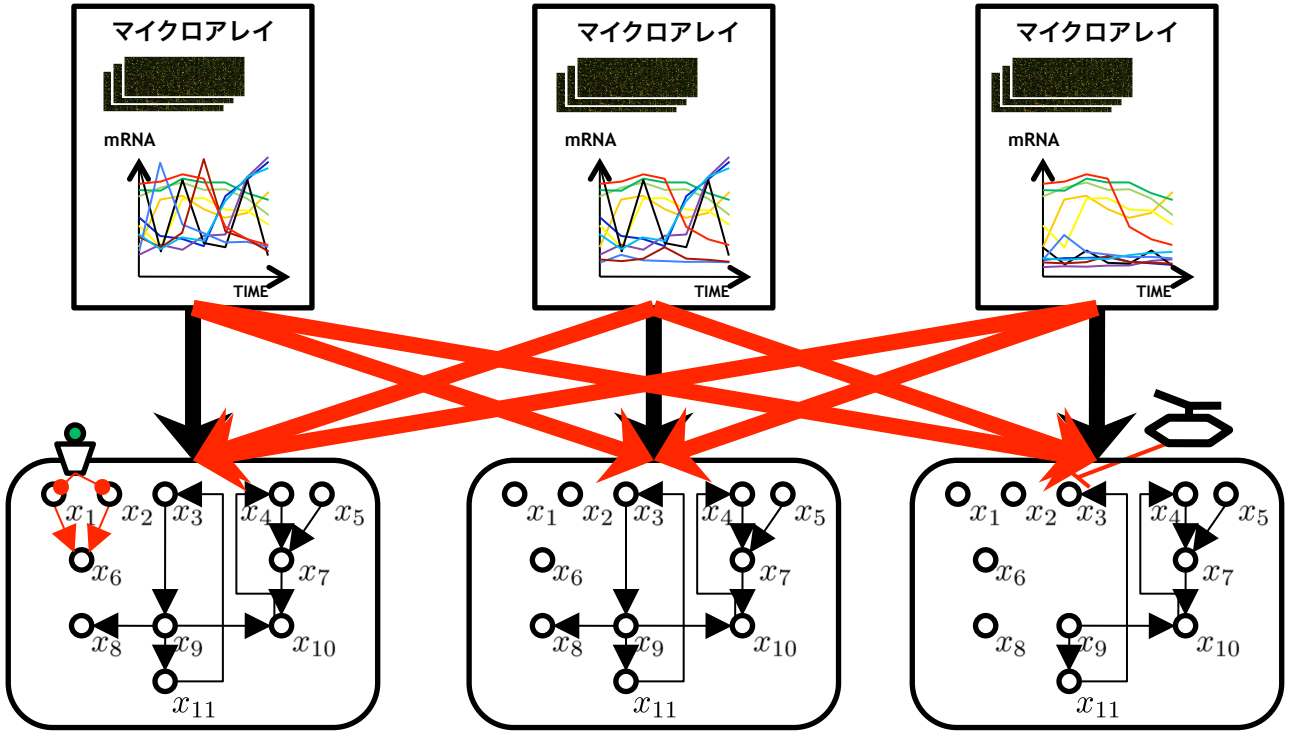



**偽陽性**（false positive）と**偽陰性**（false negative）の辺が増える

**真のシステムの違い**を見つけることが困難になる

# 解決策

仮定：システム（ネットワーク）の大部分は実験条件によって変わらない




**他の実験条件からの標本を統合**することで、
ネットワークの推定精度を改善する
Regulaized Weighted Recursive Elastic Net
を考案

# Recursive Elastic Net

Shimamura *et al.* (2009)

目的: いまある推定値をひな型に、よりスパースな推定量を求める

**Elastic net**

$$\hat{\mathbf{b}} = \arg\min_{\mathbf{b}} \left\{ \frac{1}{2}\|\mathbf{y} - \mathbf{X}\mathbf{b}\|_2^2 + \boxed{\lambda_1 \sum_{j=1}^{p} |b_j|} + \frac{\lambda_2}{2} \sum_{j=1}^{p} b_j^2 \right\}$$

$$\boxed{\lambda_1 |\beta_1| + \lambda_1 |\beta_2| + \ldots + \lambda_1 |\beta_m|}$$

**Recursive elastic net**

$$\hat{\mathbf{b}} = \arg\min_{\mathbf{b}} \left\{ \frac{1}{2}\|\mathbf{y} - \mathbf{X}\mathbf{b}\|_2^2 + \lambda_1 \sum_{j=1}^{p} w_j |b_j| + \frac{\lambda_2}{2} \sum_{j=1}^{p} b_j^2 \right\}$$

**各変数にどれだけスパース制約をかけるかを決める重み**

- 大きい $w_j$ は $b_j$ を 0 にする
- 小さい $w_j$ は $b_j$ を非0 にする

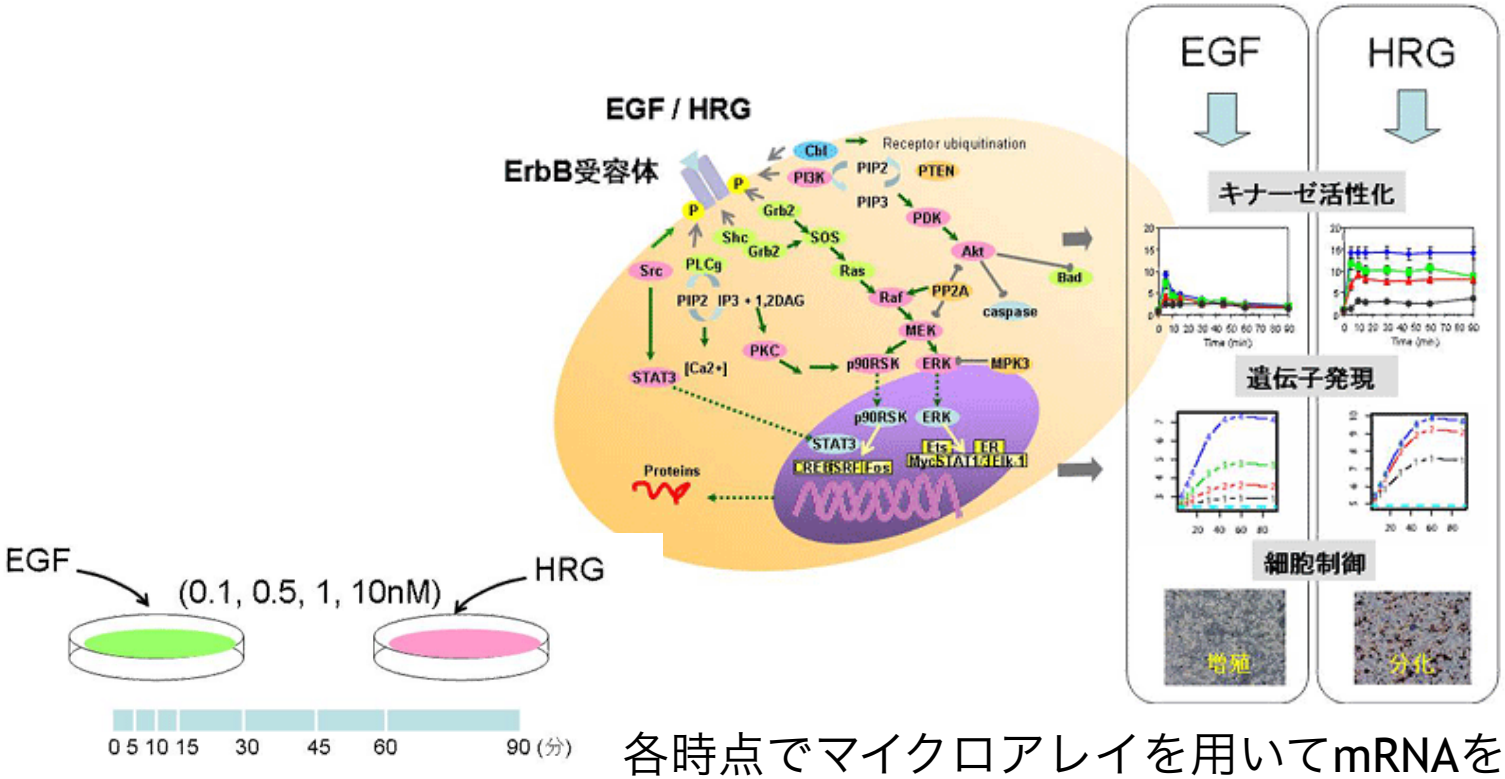
$w_k$をどうやって推定するか？

$$w_j^{(l+1)} = \frac{1}{|\hat{b}_j^{(l)}| + \delta}, \quad j = 1, \ldots, p.$$

モデル評価規準が小さくなる限り、**b** と **w** の更新を繰り返す

# 異なる刺激下による乳癌細胞株MCF-7のマイクロアレイデータ解析

二つの異なる成長因子（EGFとHRG）による細胞内の情報伝達



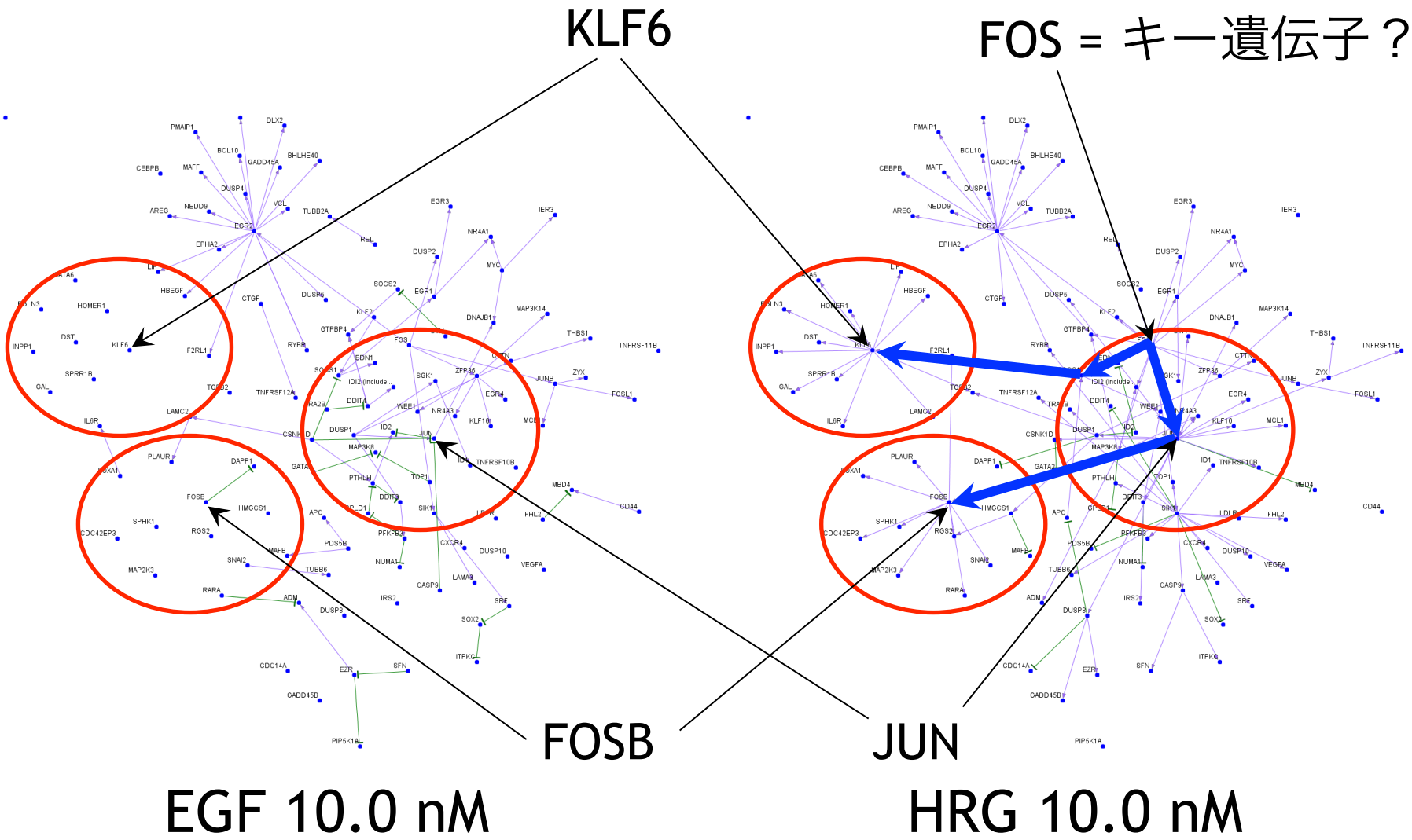

各時点でマイクロアレイを用いてmRNAを計測

データ:　**8実験条件**下において**8時点**で計測した時系列遺伝子発現データ
目的: EGF、HRG各々の成長因子によって誘導される８実験条件下での遺伝子ネットワークを推定し、**実験条件間のシステムの違い**を見つけたい

SiGN-L1

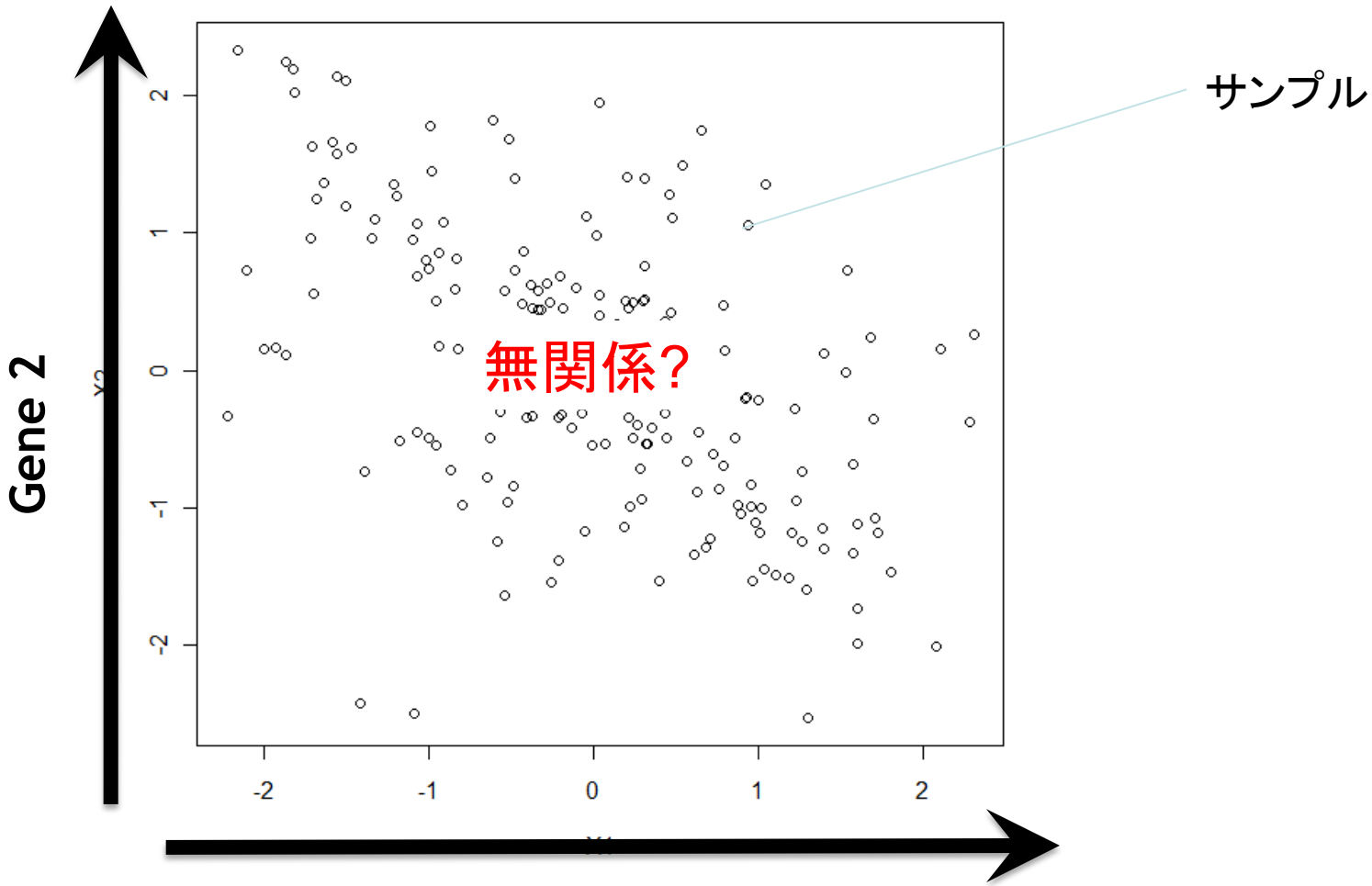
# 実験条件間で異なる遺伝子間相互作用の比較

EGF刺激とHRG刺激によって誘導されるネットワークの違いを比較

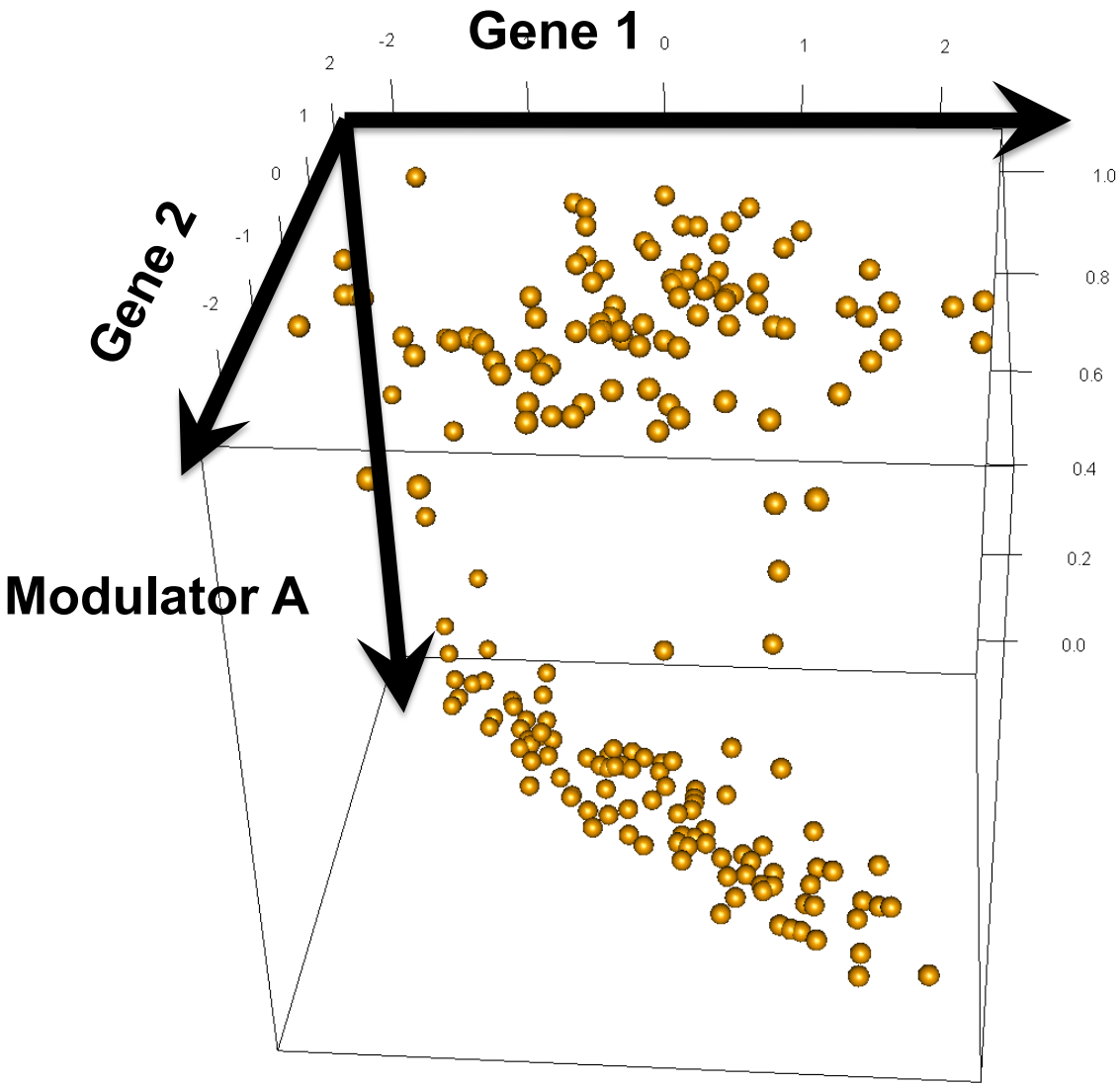


# NetworkProfiler

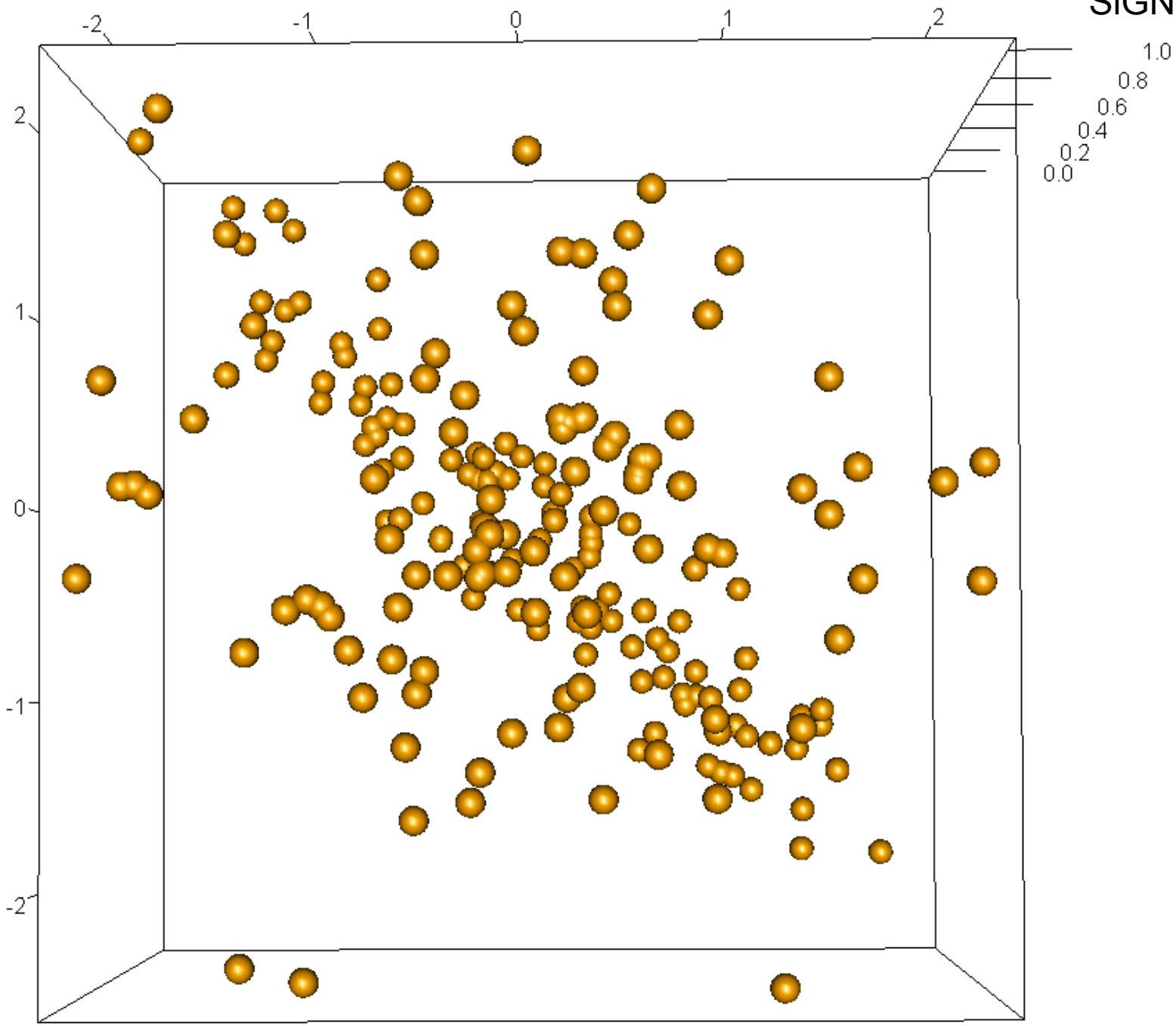
## 構造方程式モデルで見る細胞の個性・多様性



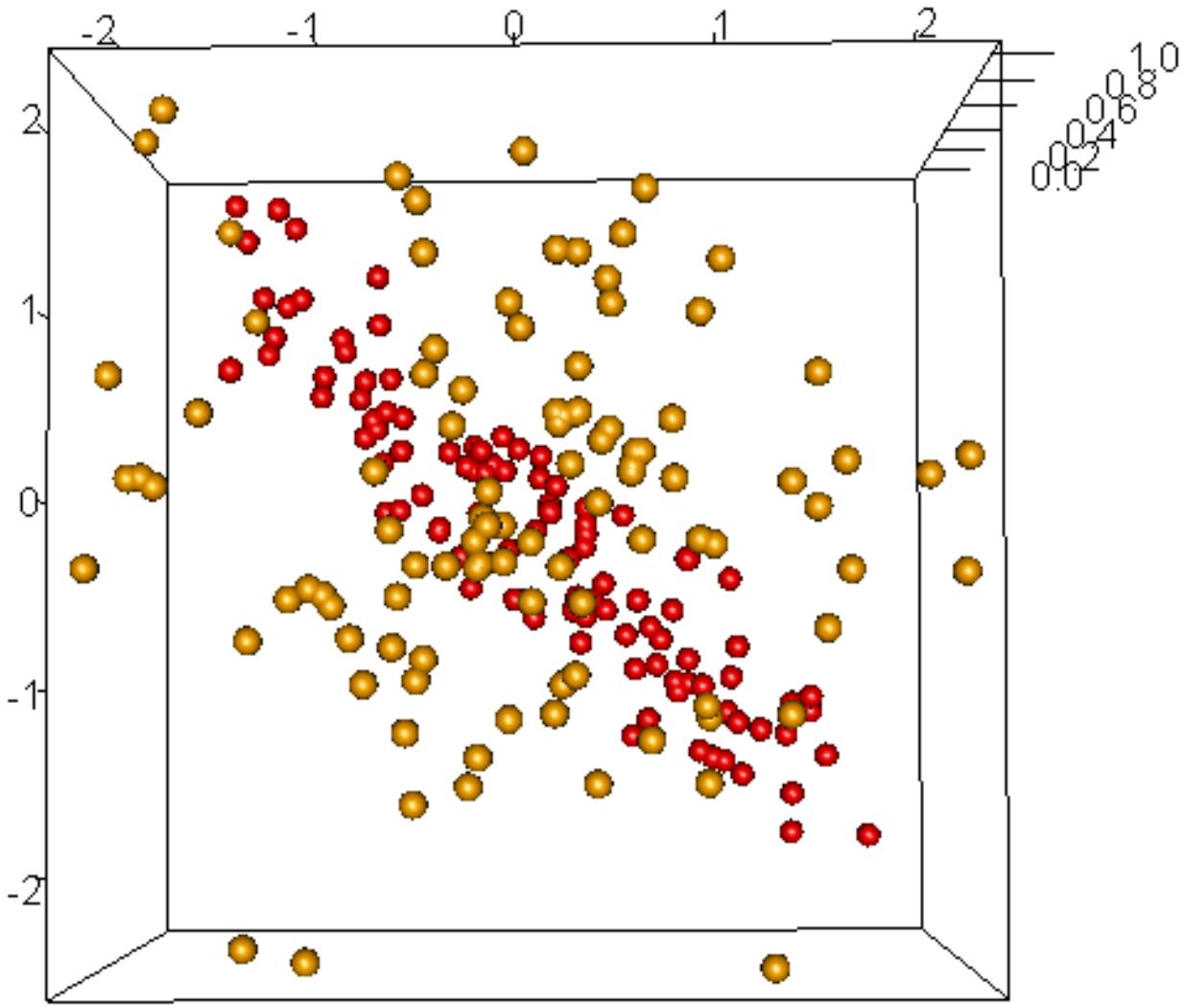
# Change can be found if we look with a modulator



SiGN-L1



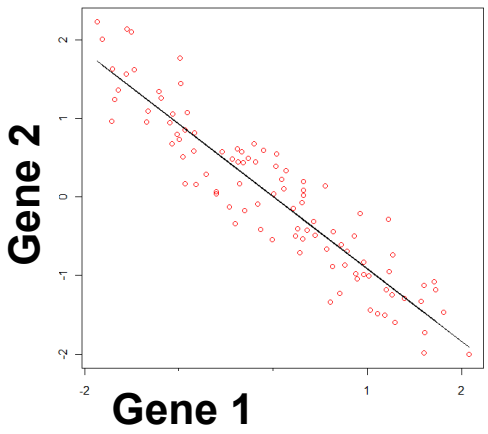
SiGN-L1

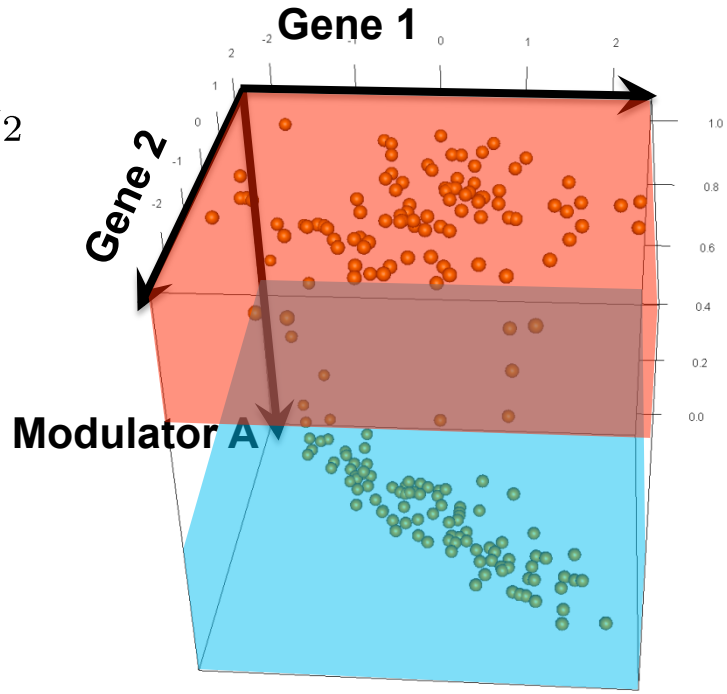


# Finding Gene Relation by Data Stratification

$$X_2 = -0.92 \times X_1 + \epsilon_2$$

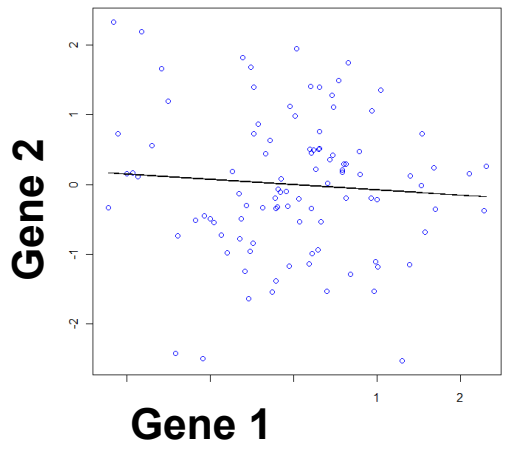
Modulator A is low



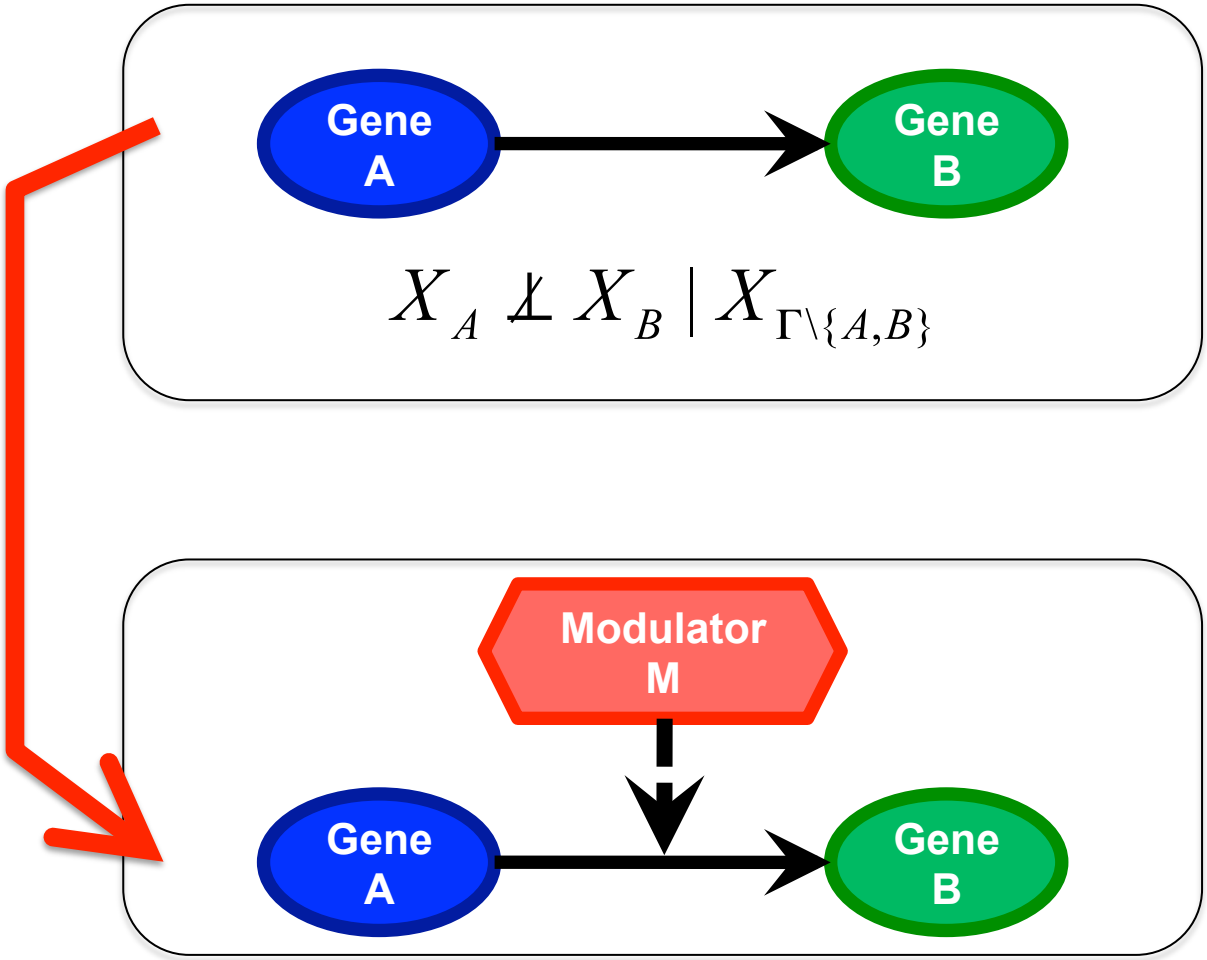

$$X_2 = -0.07 \times X_1 + \epsilon_2$$

Modulator A is high



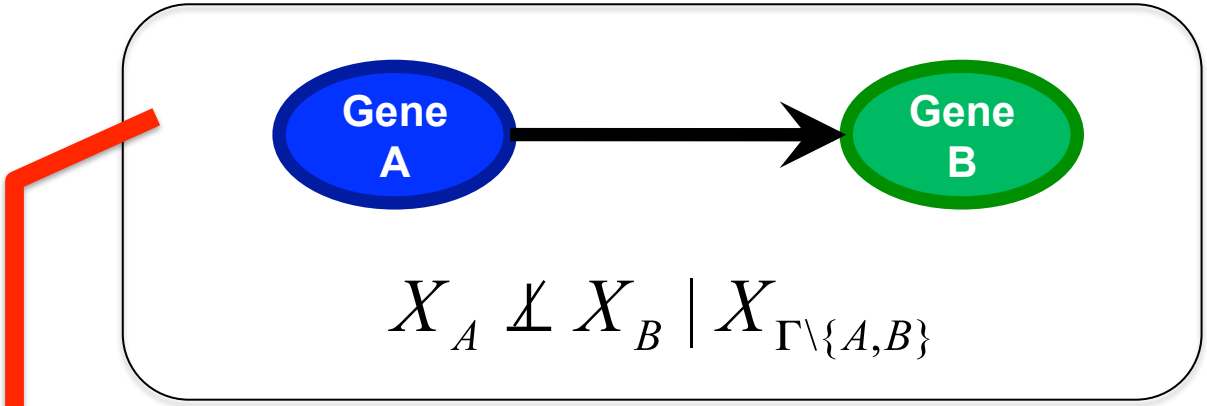

**Stratification of Data**

SiGN-L1

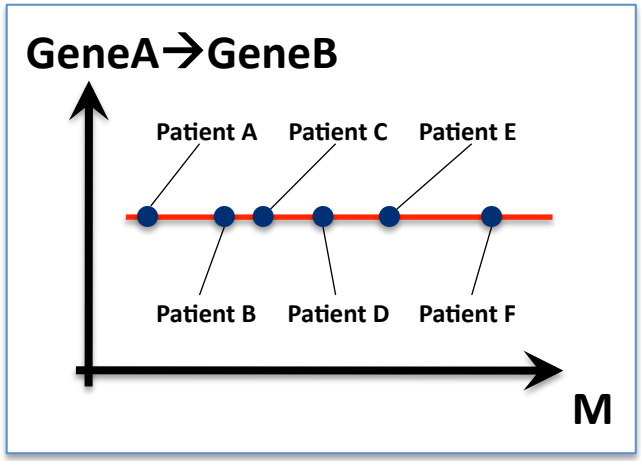

Gene
A
Gene
B
$X_A \not\perp X_B \mid X_{\Gamma\setminus\{A,B\}}$
Modulator
M
Gene
A
Gene
B

# Concept of NetworkProfiler

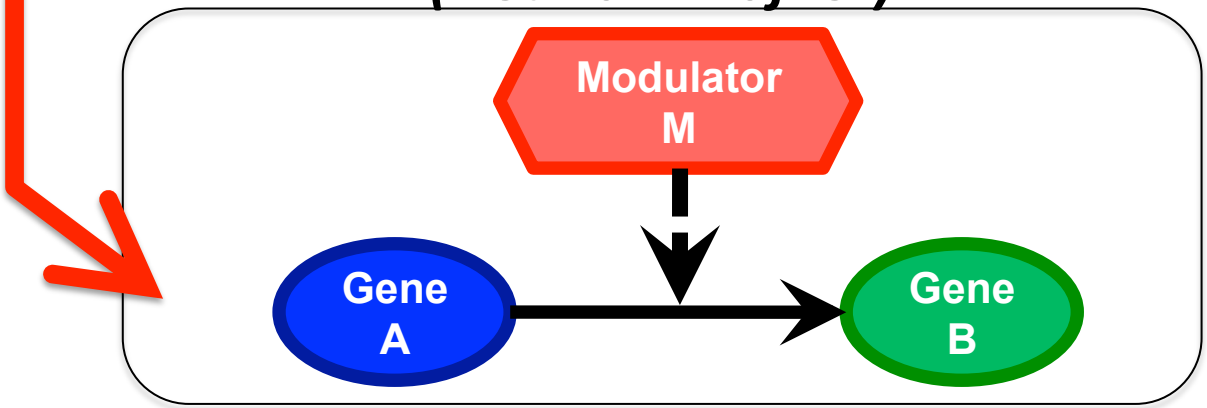
***Usual graphical model***

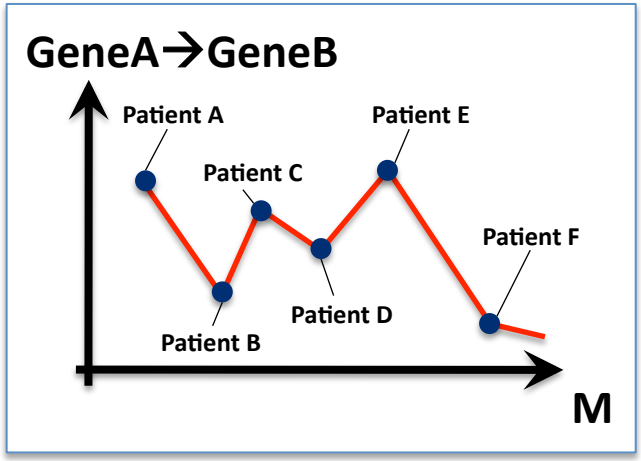



$$X_A \not\perp X_B \mid X_{\Gamma\setminus\{A,B\}}$$




***Modulator-dependent graphical model (NetworkProfiler)***



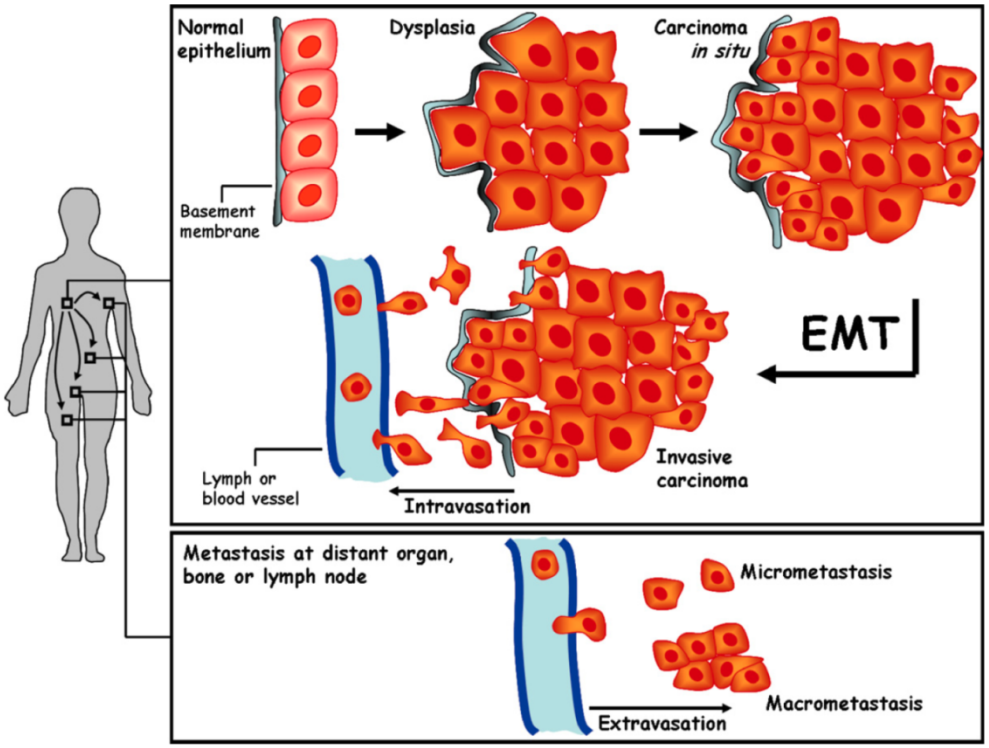




**Modulator**: **Cofactor** modulating relationship between genes A and B

# Epithelial-Mesenchymal Transition (EMT)

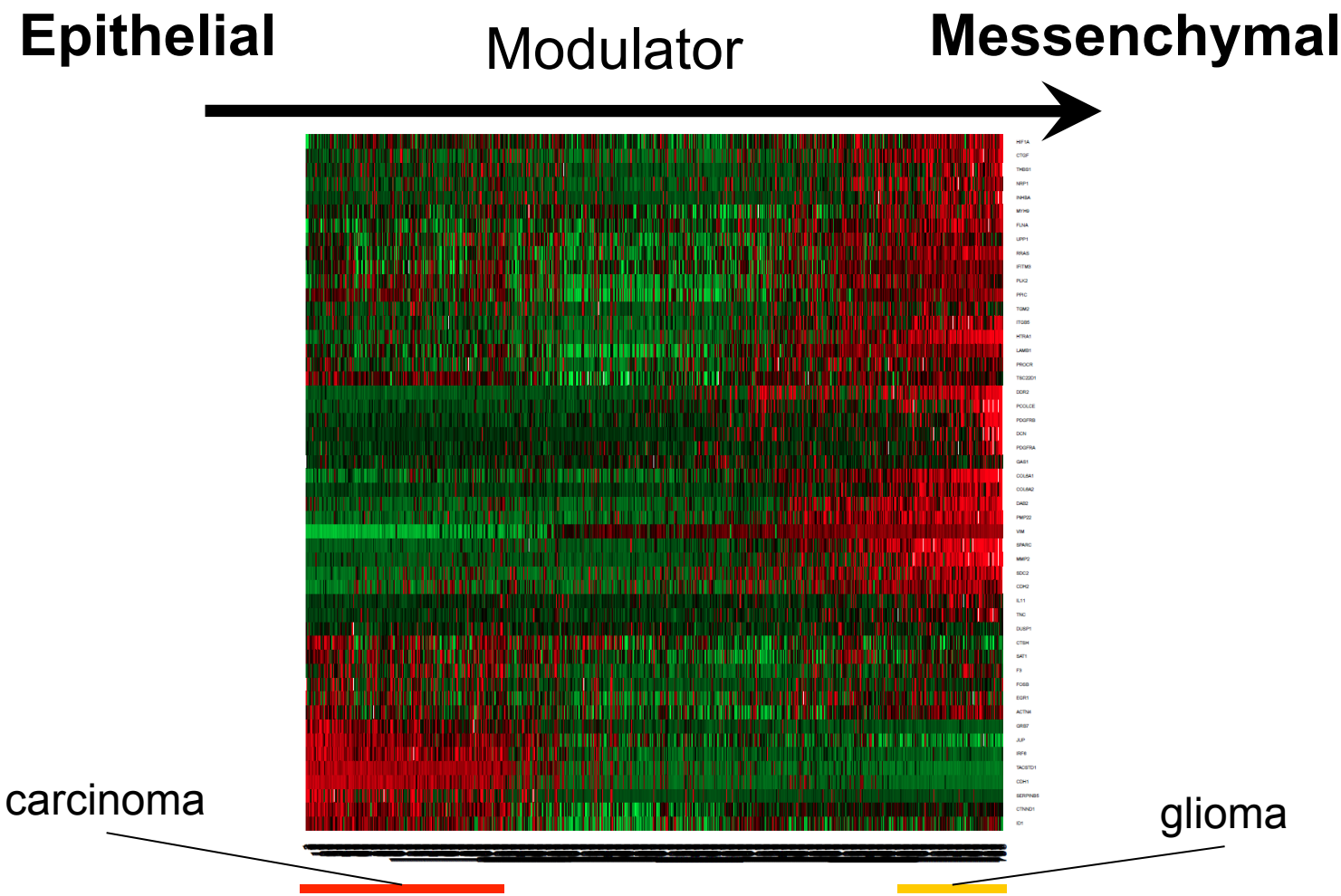
- 上皮系細胞が間葉系細胞に形態変化する現象
- がんの悪性度、がんの浸潤性に深く関連
- ゲフィチニブなどの抗がん剤を効かなくする原因の一つ

Chua et al. (2008)



# Modulator for EMT

We selected coherent 50 genes from 121 EMT signature genes to define the modulator for EMT (EEM, Niida et al., 2009)



# 上皮間葉移行に関わるメカニズムの探索

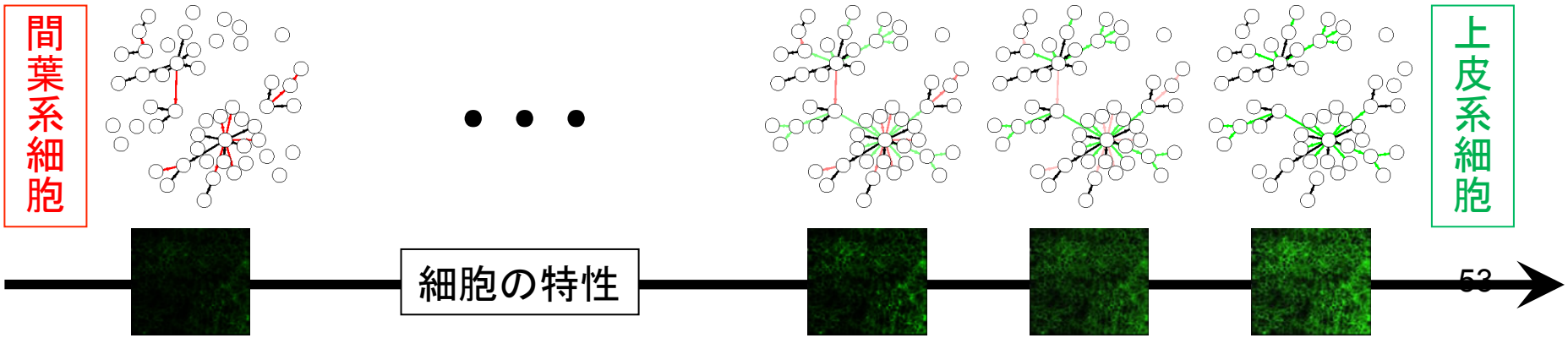
入力

- 762 のがん細胞（肺がんは254）の22858転写物のプロファイル
  （22277 mRNA+581 miRNA）（Sanger Centerのデータ）
- モジュレーター：EMTの尺度を定義（がんの転移やｹﾞﾌｨﾁﾆﾌﾞの効きにくさに関係）

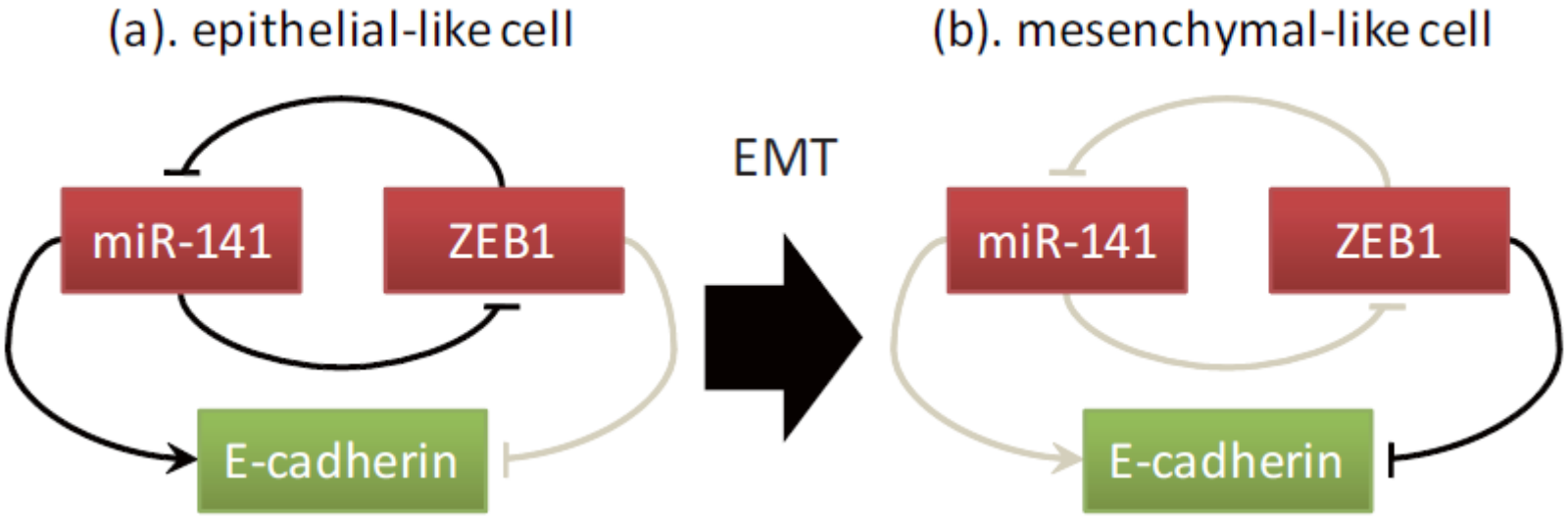
**ヒトゲノム解析センターのスパコン6,000コアのうち**
**1,024コア(12.3TFLOPS) で3カ月計算**

出力

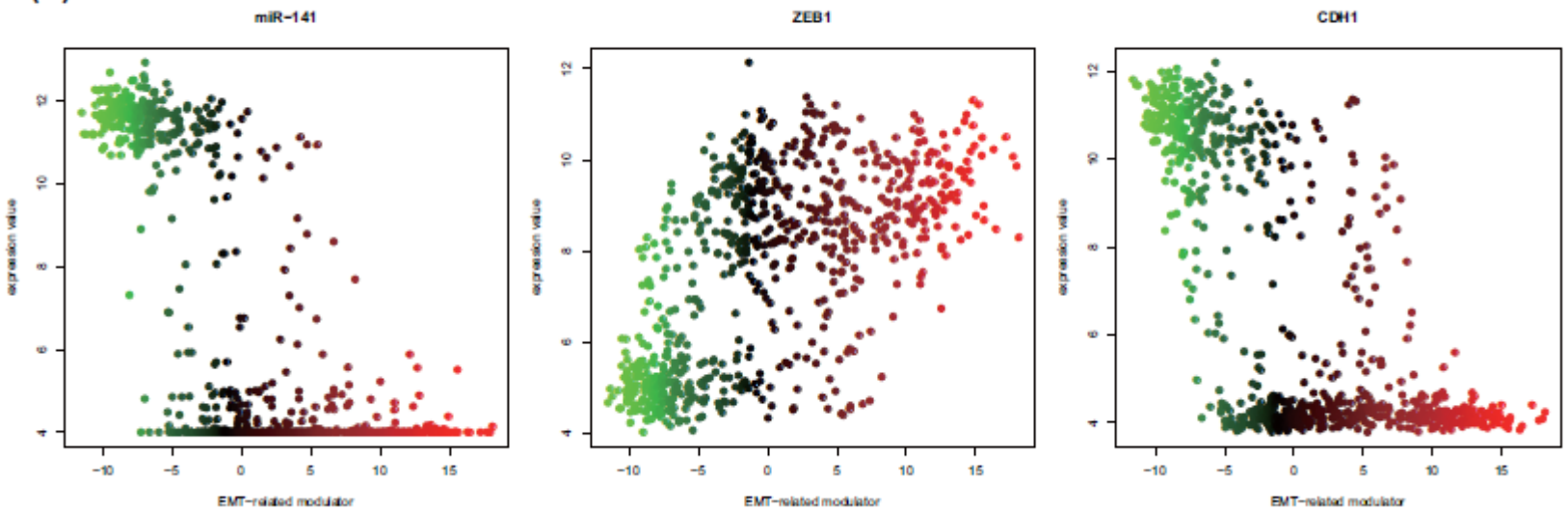
上皮間葉移行の変化に関わる 22,858 ノードの遺伝子ネットワークが 762 個



# Upstream Regulatory Changes of E-cadherin




(c).

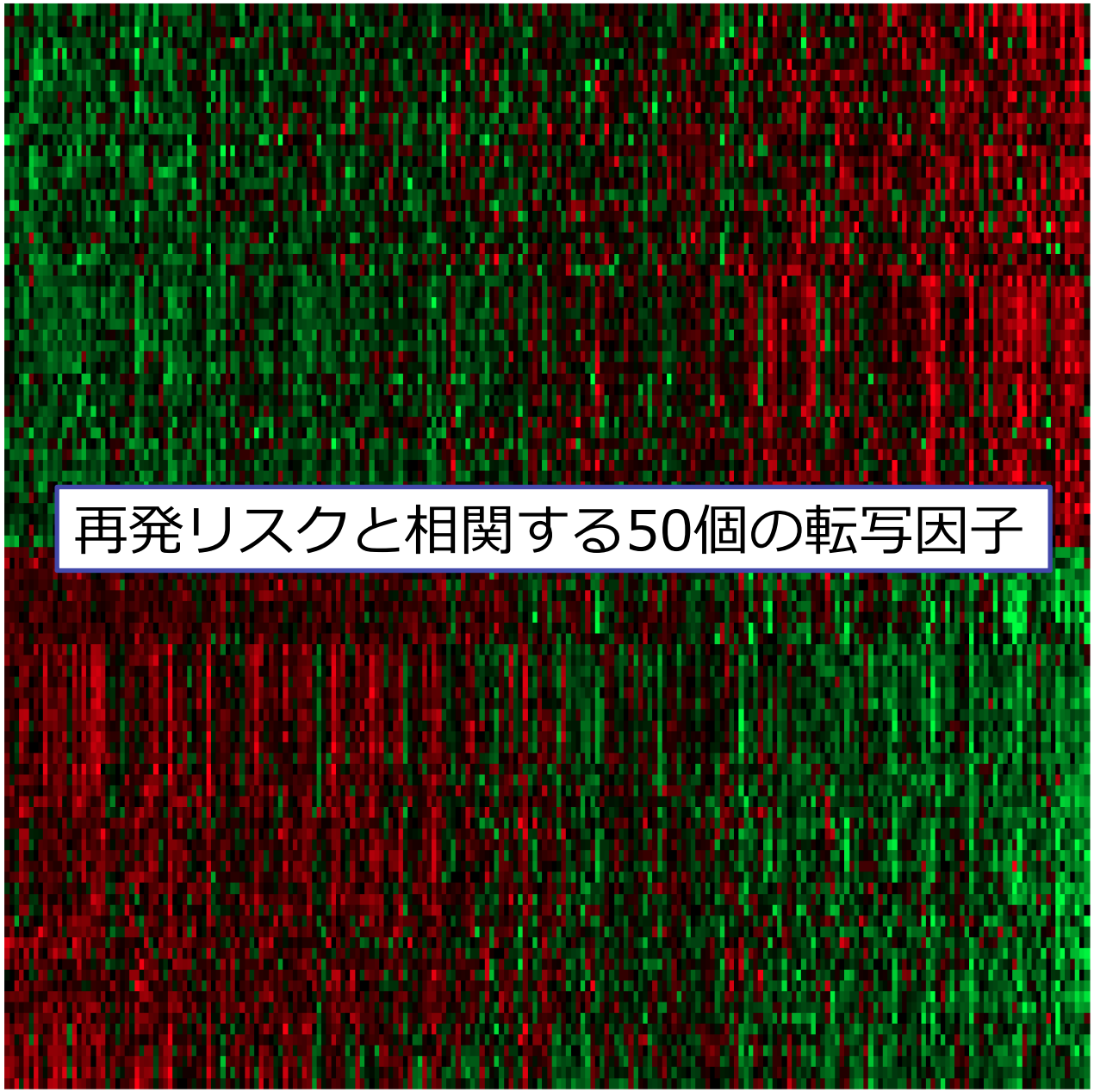


# Examples of Modulator

- Tumor progression (Stage I, Stage II, …)
- Drug sensitivity (IC50, GI50, …)
- Disease-free survival
- Molecular characteristics (Metastasis, EMT…)
- Pathway activity
- …

SiGN-L1

低 ← 再発リスク → 高



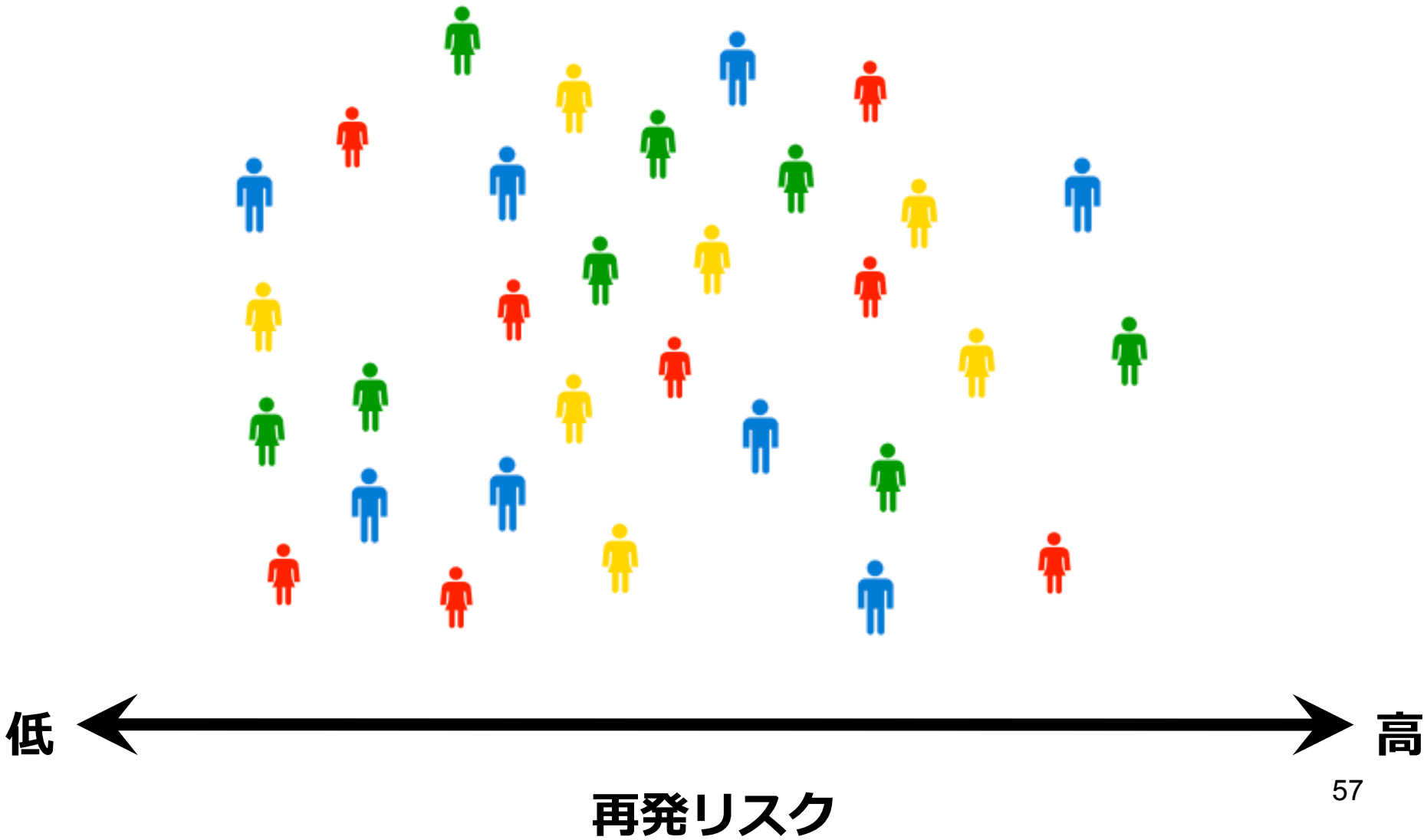
再発リスクと相関する50個の転写因子

ANKIB1
BUD31
ZNF282
ZKSCAN5
MTERF
ZNF92
GTF3A
ARNTL2
ETV6
CARM1
SHOX2
WT1
HOXA1
SAP30
RUNX2
TWIST1
NKX3-2
SOX11
HOXB7
ZMIZ2
PPARD
CBX2
VDR
HIF1A
ATF5
THOC4
UHRF1
FOXM1
EZH2
BRCA1
MYBL2
CCNE1
TRIP13
E2F7
E2F2
TCF19
ZNF367
SUV39H2
EED
TFAP2A
CLPB
TIAL1
HMGA1
PTTG1
UBE2K
ZNF207
GPN1
SSRP1
TARDBP
HDAC1
MECOM
TOB1
NFIX
IRX1
HLF
PRDM16
IRX2
NR3C2
THRB
SOX7
EPAS1
LDB2
TAL1
FHL5
DACH1
TCF21
TBX5
ERG
TBX2
TBX4
SMAD6
NOSTRIN
ZBTB16
PIAS1
MYST4
GATA6
PGR
NFIA
MYOCD
SMAD9
SMARCA2
KAT2B
CBFA2T3
STAT5B
MXD4
MECP2
CITED2
BTG2
JMY
SHPRH
ARID4A
KLF15
AFF3
RORA
KLF9
KLF2
FOS
FOSB
CSRNP1
JUND

EMTマーカー

浸潤マーカー

# 数百のがんサンプル



# 再発リスク

構造方程式モデルで各個人のシステムを解析

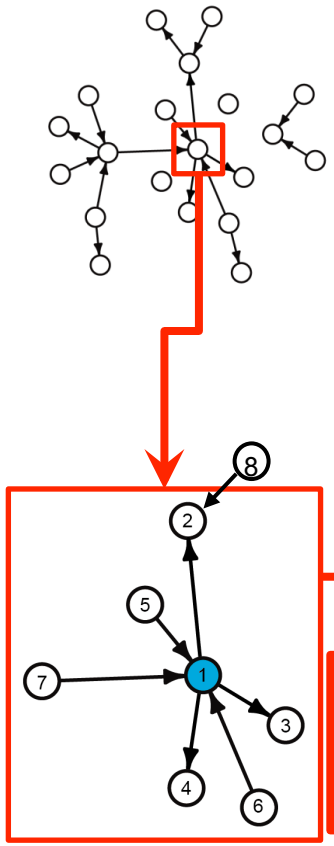
$p$:遺伝子数, $n$：マイクロアレイ数
$x_{ij}$：$i$ 番マイクロアレイにおける遺伝子$j$ の発現量
$G_i$：$i$ 番サンプルの構造方程式モデルによる遺伝子ネットワーク

$$x_{ij} = \sum_{k=1}^{p} \beta_{jk} x_{ik} + \varepsilon_{ik}, \quad (k \neq j)$$

パラメータ$\beta_{jk}$ はElastic net型 重み付きロス関数を最小化することで推定。$m_1,\ldots,m_n$ は尺度値。

$$S(\beta_{j1},\ldots,\beta_{jp} \mid i) = \sum_{\alpha=1}^{n} K_h(m_\alpha - m_i)\left\{x_{\alpha j} - \sum_{\substack{k=1 \\ k\neq j}}^{p} \beta_{jk} x_{\alpha k}\right\}^2 + \lambda \sum_{k\neq j} w_{jk}\left|\beta_{jk}\right| + \frac{\nu}{2}\sum_{k\neq j}\beta_{jk}^2$$

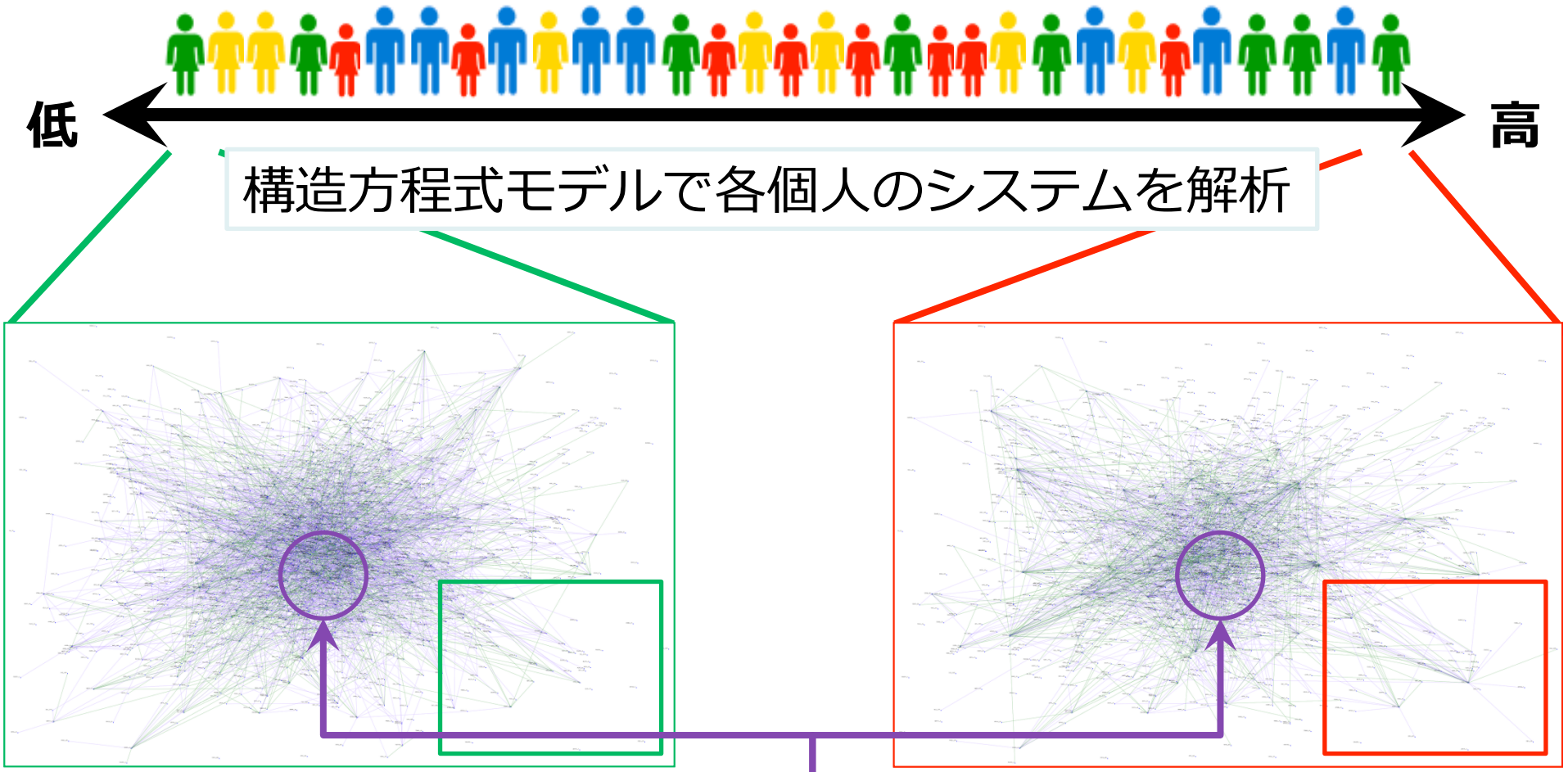
Kernel function $K_h(m_\alpha - m_i)$ は$i$ 番サンプルの近傍を定義し、これはbandwidth parameter $h$ を選ぶことで最適化される。




構造方程式モデル

$$\begin{aligned} x_1 &= 4.6 \times x_5 + 4.2 \times x_6 + 0.7 \times x_7 + \epsilon_1 \\ x_2 &= -2.7 \times x_1 + 0.95 \times x_8 + \epsilon_2 \\ x_3 &= -1.4 \times x_1 + \epsilon_3 \\ x_4 &= -0.5 \times x_1 + \epsilon_4 \end{aligned}$$

# 再発リスク



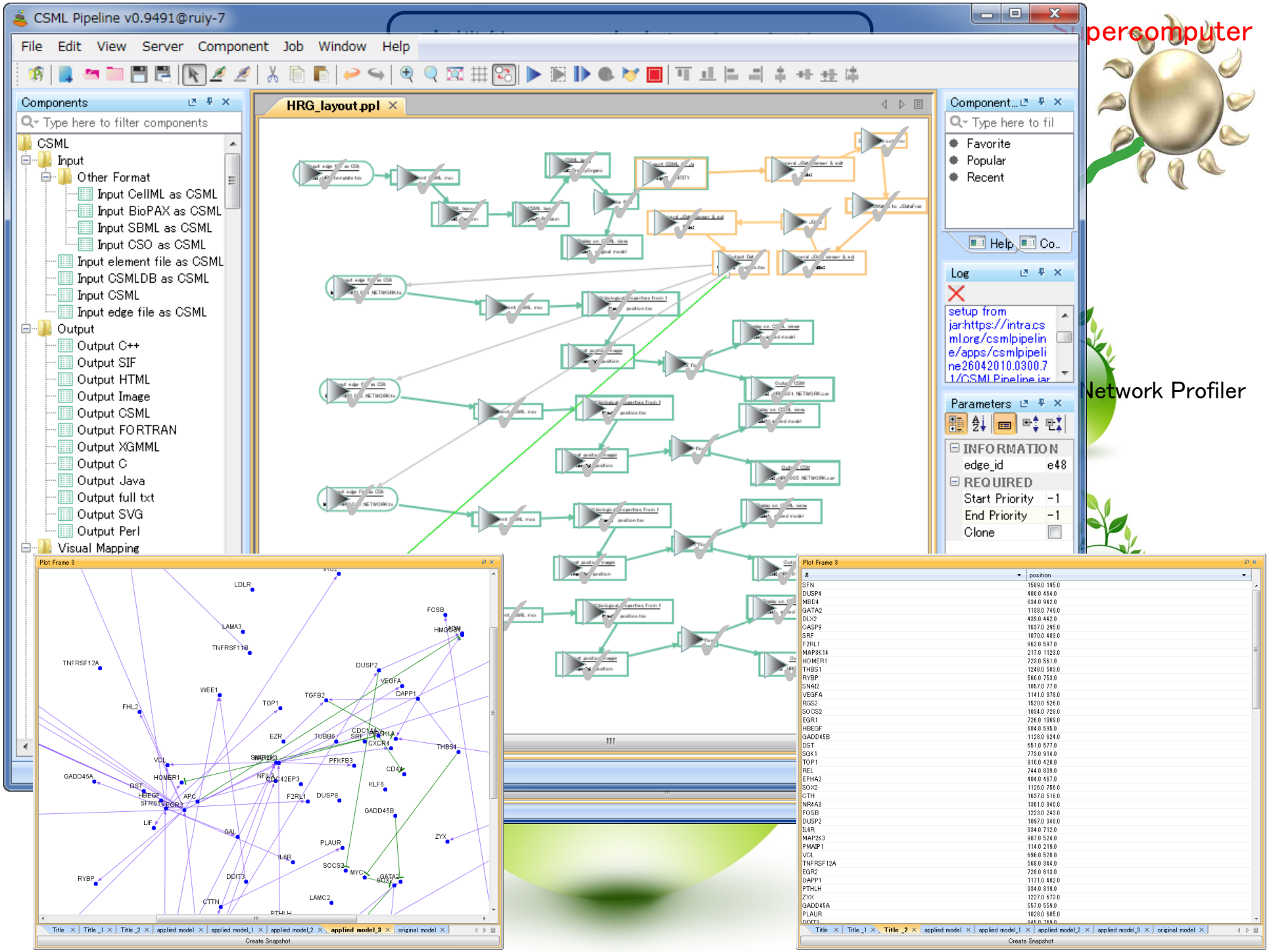

**CTGF (Connective Tissue Growth Factor)**
増殖・分化を制御するTGF-βの下流因子で、TGF-βの間質繊維化促進作用を仲介
増殖促進、遊走、細胞外基質算出、血管新生作用を呈する

# データ解析とシミュレーションパイプラインをONE STOPで

Systems Biology integrative Pipeline

SBiP

(Java software)


CSML Pipeline v0.9491@ruiy-7
File Edit View Server Component Job Window Help
Components
Type here to filter components
CSML
Input
Other Format
Input CellML as CSML
Input BioPAX as CSML
Input SBML as CSML
Input CSO as CSML
Input element file as CSML
Input CSMLDB as CSML
Input CSML
Input edge file as CSML
Output
Output C++
Output SIF
Output HTML
Output Image
Output CSML
Output FORTRAN
Output XGMML
Output C
Output Java
Output full txt
Output SVG
Output Perl
Visual Mapping
HRG_layout.ppl
Component...
Type here to fil
Favorite
Popular
Recent
Help
Log
setup from jar:https://intra.csml.org/csmlpipeline/apps/csmlpipeline26042010.0300.71/CSMLPipeline.jar
Parameters
INFORMATION
edge_id e48
REQUIRED
Start Priority -1
End Priority -1
Clone
Plot Frame 3
Create Snapshot
percomputer
Network Profiler