

システム分子医学と複雑系疾病理論

—がん転移のアトラクター理論を例として—

東北大学 東北メディカル・メガバンク機構

東京医科歯科大学 名誉教授 (臨床腫瘍学)

田中 博



ゲノム・オミックス医療

3つの概念/パラダイム

根拠となる疾患形成機序の概念

第1 ゲノム医療

- 生得的ゲノム情報の変異・多型性

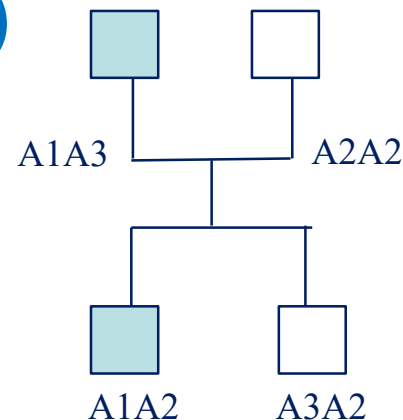
第2 オミックス医療

- 後天的体細胞の網羅的分子プロファイル

第3 システム（分子）医学

- 細胞分子ネットワークの歪み/構造変化

網羅的分子医学 第1パラダイム ゲノム医療 (1990~現在)



- 「生得的な(germ-line)」ゲノム変異や

多型性に基づいた個別化医療

- 生得的ゲノム (constitutional genome)は全細胞で生涯を通じて同一 99.9%は共通

- 疾患原因遺伝子 (Disease causative gene)

- 家系調査/Linkage解析、Positional Cloning

- 1980年代から: ハンチントン病 (CAGリピート, HTT遺伝子, 99%浸透率), その他にデシャンヌ型筋ジストロフィー, 嚢胞性線維症など
- 当時400程度のDNAマーカ → ヒトゲノム解読計画へ

- 疾患感受性遺伝子 (Disease susceptibility gene)

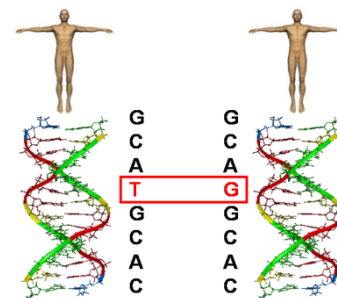
- 多型性: 一塩基多型(SNP), 1億個位,

- そのほかにマイクロサテライト、CNVなど

- 全ゲノム関連解析 (GWAS) 疾患相対リスク

- HAPMAPプロジェクト、1000ゲノムプロジェクトなど

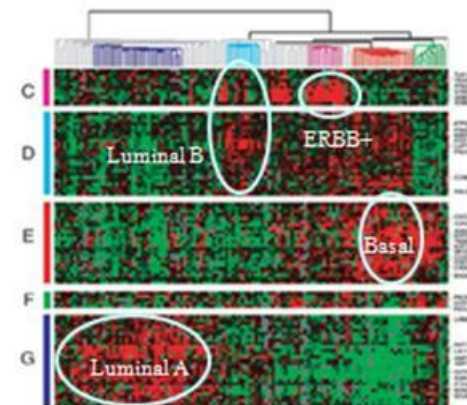
- 薬剤代謝酵素の多型性 → 有効性・副作用の評価



1塩基の変異が疾患や薬剤への応答性に関与する場合がある

網羅的分子医学 第2パラダイム オミックス医療 (2000~)

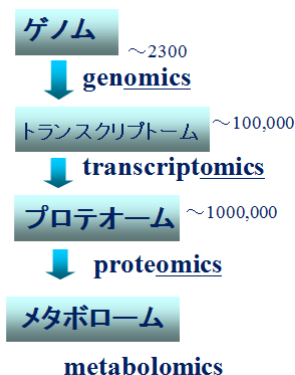
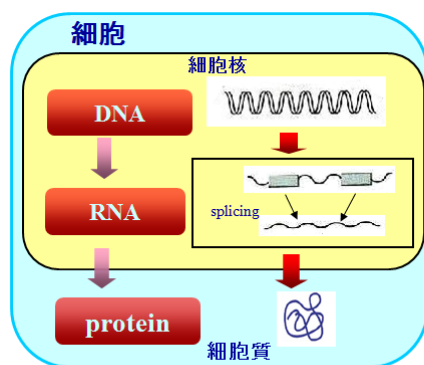
- 「**後天的・疾病状態依存的に変化する体細胞変異**」
 - 後天的体細胞ゲノム変化・オミックス変化に基づく医療
 - 遺伝子発現プロファイル・プロテオーム・メタボロームなど
 - 網羅的分子表現型 "molecular phenome" — 分子的病態像
 - 疾患の進行度の評価**
 - 疾患組織や病態進行の段階に依存
 - 臨床症状や病理変化より早く変化→予測医療・先制医療



- 疾患オミックスプロファイルによる疾患の内在的分類**
 - 網羅的分子病態像によるサブタイプ分類
 - 遺伝子発現プロファイルによる分類
 - 臨床・病的な分類よりもintrinsicな分類
 - 乳がんの臨床を根底的に変革

| intrinsic分類 | ER | PgR | HER2 | 予後 |
|---------------|---------------------|-----|------|--------|
| Luminal A | ER(+) and/or PgR(+) | (-) | (-) | 予後良好 |
| Luminal B | ER(+) and/or PgR(+) | (-) | (+)* | A型より不良 |
| HER2 enriched | (-) | (-) | (+) | 予後不良 |
| Basal-like | (-) | (-) | (-) | 予後不良 |

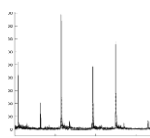
乳がんのオミックスプロファイルと亜型分類と予後・治療



Next generation sequencer



DNA microarray



Mass spectrometry



網羅的分子医学 第3パラダイム システム分子医学 (2010~)

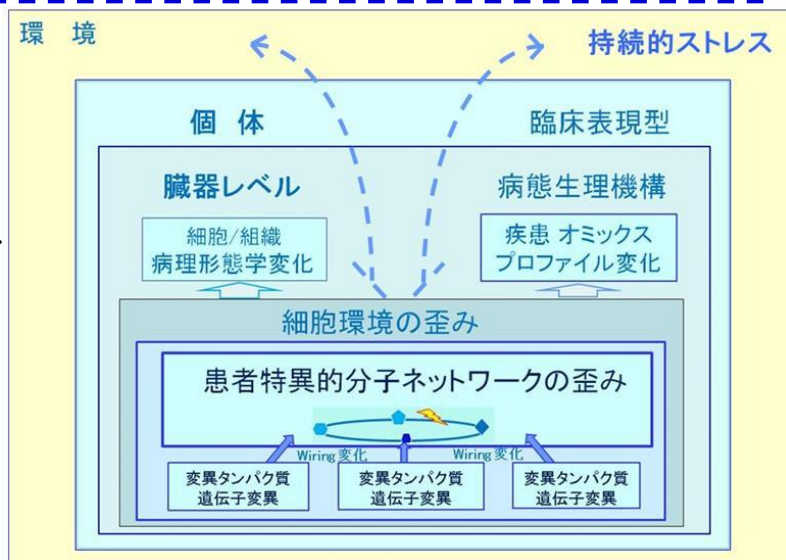
システム科学の疾患への応用
「疾患をシステムとして理解する」

単因子性の疾患を除いて、大半の疾患は1個や2個の遺伝子の変異ではなく多数の遺伝子の変異やタンパク質の機能異常による
分子パスウェイ/ネットワークの調節機能不全や歪み
distortion (dysregulation) of molecular network

個別化医療・予測医療・先制医療・創薬のための
疾患のシステムの理解に基づく医療

疾患オミックス
網羅的分子表現型
(molecular phenome)
成り立たせる基底としての
「細胞分子ネットワークの構造変化」

病態生理層と病理形態層による
臓器・個体レベルの階層システム化



システム分子医学を支える 2つの疾患システム科学

第1の疾患システム科学 疾患システム・バイオロジー

- システム生物学の疾患領域への応用
- 細胞病態過程のシステム・モデリング
 - 計算機シミュレーション・実験値との突合
 - 実験データのシステム・モデル分析
 - 欠点：モデル化可能な既知の生物過程に限る
 - 特徴：システム好像は不変、変量だけが時間的に変化
- 疾患オミックス情報からシステム同定
 - 決定的な同定法を探求中
 - 多層疾患オミックス情報の利用

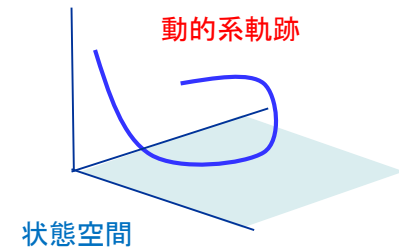


培風館 2012

システム分子医学を支える 2つの疾患システム科学

第2の疾患システム科学

複雑系疾病論（病態力学系）



- 複雑系・動的（非線形）システム論
- 非線形動的システム概念からの疾病理解
 - 生命システム構造（枠）の動的変化（「遷移」）
 - 生命システム構造の「転移」としての疾患発症
 - 発生や生命進化との疾病との関係
- 具体的疾患過程への応用は探索中
 - 数理モデルとしての限界の超克
 - 具体的な疾患過程への適用—現象的なシステム量へ

I 疾患システムバイオロジー

疾患システムバイオロジーの現状

例 がん疾患

- **Cancer Systems Biology**

- 「疾患システムバイオロジー」の最も活発な分野

- 疾患システムバイオロジー (disease systems biology)

- システム生物学の疾患への応用

- 「システム(分子)医学」の基礎を支える学問分野

- システム生物学:

- 「生命をシステムとして理解する」生命情報学の第3世代

- 第1世代: 単分子の核酸・タンパク質の配列・進化解析

- 第2世代: 網羅的分子 (genome/omics) の配列・進化解析

- **学術的領域としてのすでに確立**

- NIH, 2004年にグラント開始

- 00'の中頃からシステム生物学の生命科学での興隆に影響を受けて疾患分野での応用を目指す。

- 全米で12の研究施設 (Center for Cancer Systems Biology)

- NCIのICBP(integrative cancer biology program)より助成

- Dana-Faber(Harvard), Stanford, Columbia, MIT, MSKCC等

- 国際学会・シンポジウム: CSBの名前で多数開催

- Cancer Systems Biology Consortiumが2015年からNIHの援助 (37.8M\$)のもとに活動開始



Center for Cancer Systems Biology

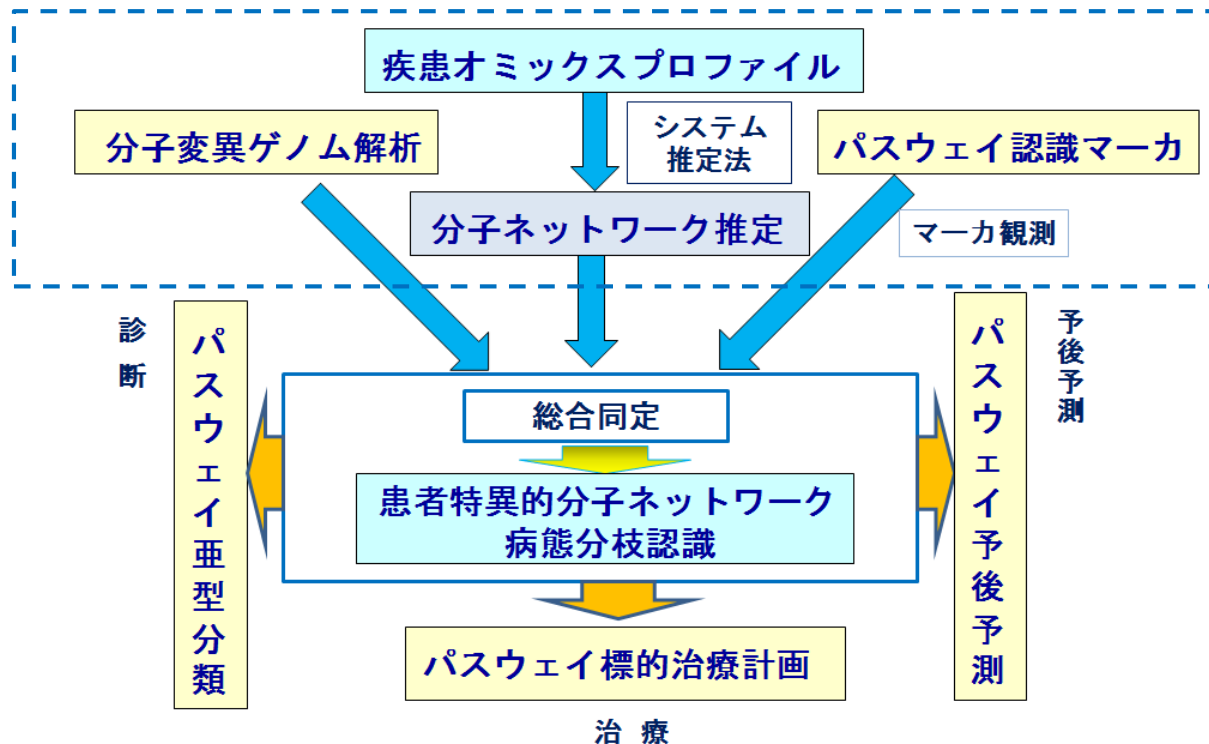
がん医学・医療の世代論

- **第1世代：がん関連遺伝子の発見(1976年から80年代,90年代前半)**
 - Bishopらの細胞性がん遺伝子srcの発見(1976), がん遺伝子の概念提唱
 - 1982年初めてがん遺伝子rasがヒトのがん細胞から発見。
 - がん抑制遺伝子、網膜芽腫タンパク質Rbのクローニング(1986)。
 - 「少数のがん遺伝子・がん抑制遺伝子の変異としてがんを理解」
- **第2世代：がんの基幹パスウェイの解明 (90年代から00年代前半)**
 - がん関連遺伝子が続々と発見。がん関連遺伝子が含まれるパスウェイの解明とパスウェイへ遺伝子変異の影響へ関心移った。
 - 1980年代の後半から1990年代にかけてシグナル伝達系や遺伝子制御系の研究が爆発的に発展
 - 増殖因子受容体/Ras/MAPKカスケードよりなる増殖系の全容が解明され(1992)
 - がん遺伝子EGFRやRasの変異によるがん化の機構も明らかになる。
 - 細胞周期系やアポトーシス系も、ほぼ全貌解明
 - 「がん関連遺伝子の変異によつ主要パスウェイの調節不全としてがんを理解」
- **第3世代：がんの網羅的分子情報解析・分子標的薬の開発 (00年代から現在)**
 - 遺伝子発現プロファイルやプロテオーム情報によるがんの早期発見や予後・再発・抗がん剤感受性の予測など (液性がん、乳がんなど) が進む
 - がんの網羅的分子研究が大きく進展。
 - 「がんゲノムアトラス (TCGA)」「国際がんゲノムコンソーシアム (ICCG)」などの計画も世界各地で推進され、
 - がん腫ごとの変異、ドライバー変異の実態が明確に
 - パスウェイの基軸となるがん関連分子を標的にする「分子標的薬」が登場する。
 - **がんの臨床ゲノム医療の開始** (シーケンス→ドライバー変異→分子標的薬)

疾患システムバイオロジーの基本戦略

マルチオミックスからの分子ネットワーク病態認識へ

1. 疾患オミックスプロファイル (disease omics profile) から
 - 患者特異的分子ネットワーク (個別化医療) の調節不全分枝 同定
 - Dysregulated pathway/subnetwork (“disease-perturbed network” Hood) の同定
2. マルチオミックスによる臨床的実践の戦略
 - 遺伝子発現プロファイル 推定法による分子ネットワークの同定 (80%)
 - 次世代シーケンシング 転写因子や信号パスウェイ分子の変異
 - プロテオーム (パスウェイバイオマーカ) リン酸化情報



疾患システムバイオロジーの臨床実践の フレームワーク

目標：患者特異的な細胞分子ネットワークの歪構造変化を同定し
これを基礎に診断・治療・予後を組み立てる

基本原則

臨床実践レベルで実現可能な分子ネットワーク歪構造同定

基本的枠組み

1. 遺伝子発現プロファイルを、**主要疾患オミックスプロファイル**として分子ネットワーク推定アルゴリズムを適応してNWの大枠を決定
2. Clinical sequencingにより遺伝子・ゲノム変異を得てNWを詳細指定
3. リン酸化タンパク質抗体などのパスウェイバイオマーカより活性化分枝特定

問題点：単一の疾患オミックスプロファイルから患者特異的な分子ネットワークは
同定不可能

方向：疾患オミックスプロファイルのクラスタリングから分子ネットワーク歪構造の
亜型を決定 患者のオミックスプロファイルから類似パターンを検索

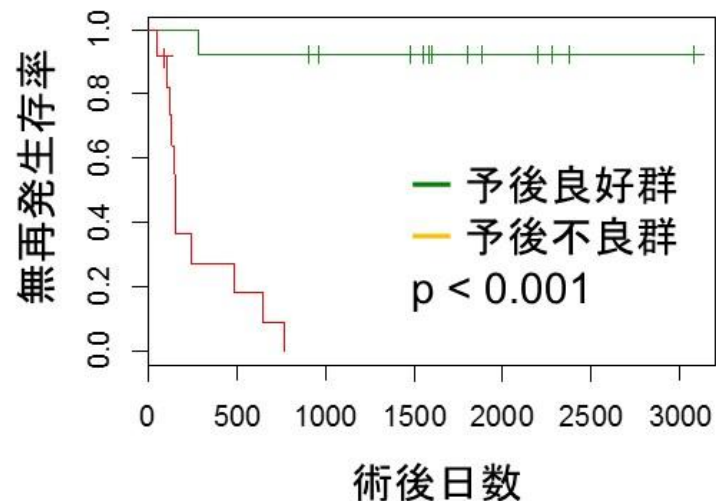
分子ネットワーク歪構造同定

疾患オミックスプロファイルの乱れの
親近度クラスター分離

分子ネットワーク差異的部分の推定

疾患ネットワーク同定方法論の試み

- 東京医科歯科大学附属病院 肝胆膵・総合外科にて肝切除を受けた40症例
 - 予後良好群20例 vs. 予後不良群20例
 - 臨床病理学的因子としては門脈侵襲(vp: $p < 0.001$)などに強く相関
- 遺伝子発現プロファイル的差異
 - DNAマイクロアレイによる発現解析
 - SNPアレイによるコピー数解析
 - エクソーム解析 (変異遺伝子)
 - プロテオーム解析 (リン酸化)

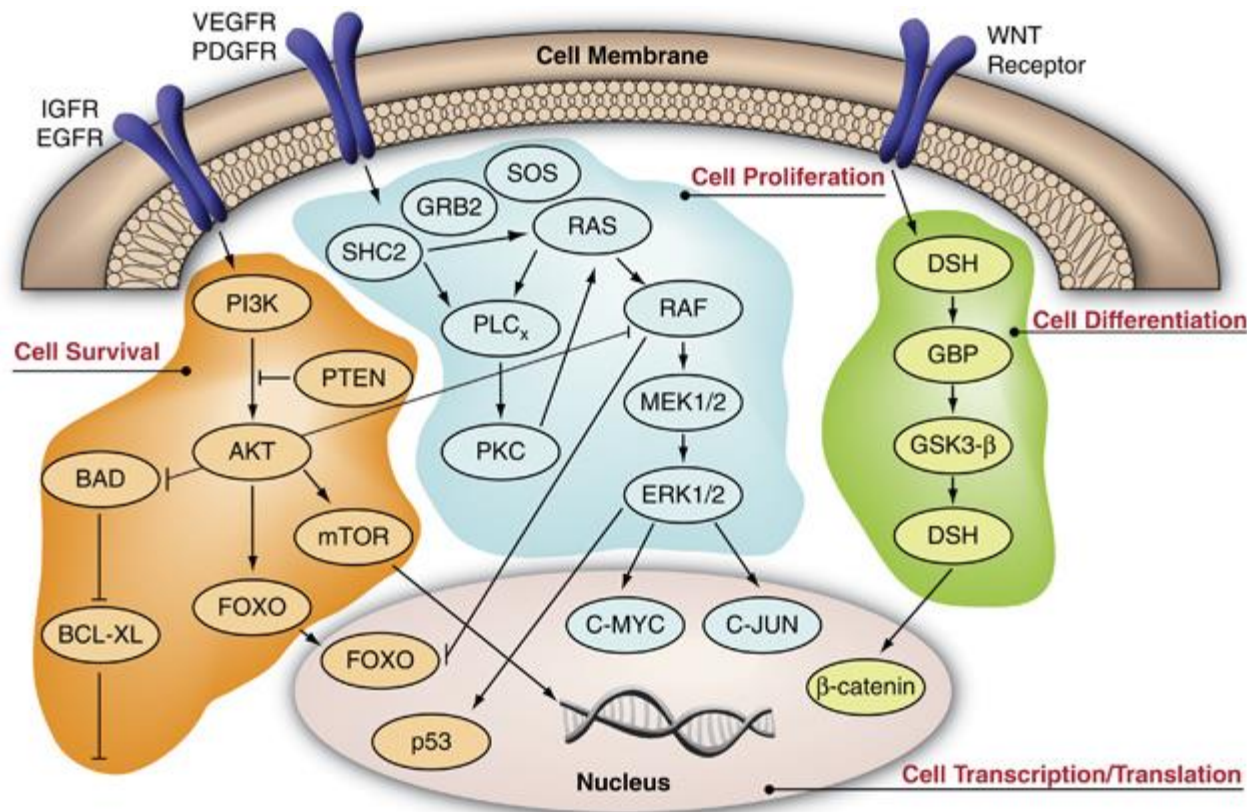


細胞分子ネットワーク同定方法論

- 展開する細胞分子ネットワーク（候補NW）を選択
- 遺伝子発現プロファイルのクラスター（亜集合）に遺伝子ネットワーク同定アルゴリズム（例えばGenie3）を適用、NW同定
 - 臨床統合オミックスDB（例えばiCOD）が存在すれば、患者の疾患オミックスと類似な症例（類似性検索）を検索して、一定の患者亜集合を構成する。
- タンパク質間相互作用NWなどを用いてFalse Positive 枝を刈り込む
- 亜集合間(例えば健常 - 疾患群) で、Differential 法でNWの歪構造サブネットワークを抽出

分子ネットワークモデル

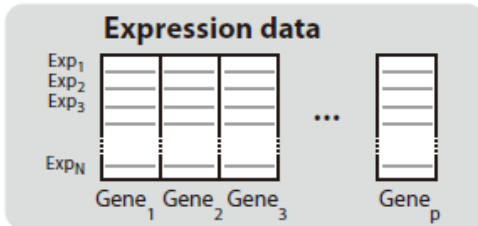
- Whittaker S, Marais R, Zhu AX. The role of signaling pathways in the development and treatment of hepatocellular carcinoma. *Oncogene*. 2010;29(36):4989-5005.
- 1遺伝子に複数のプローブセットが該当する場合には、四分位偏差(IQR)が最大のものを選択



GENIE3

GEne Network Inference with Ensemble of trees

①遺伝子 p 個の発現データ



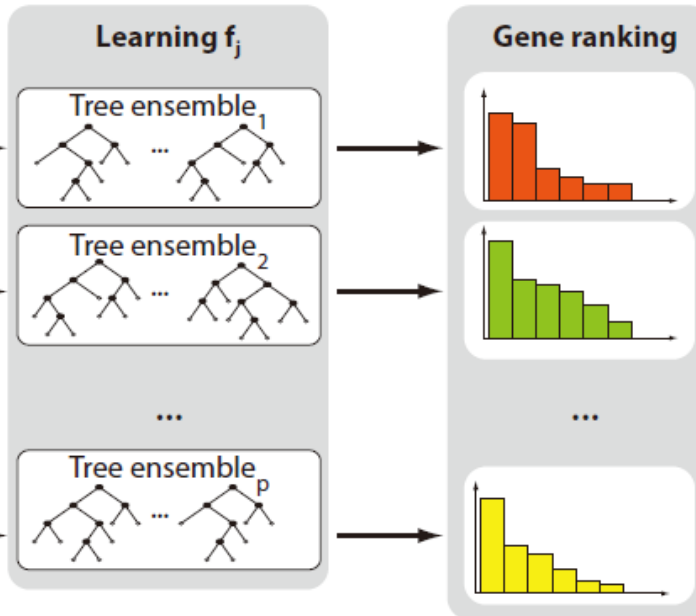
LS¹ 部分問題 ...
目的関数はターゲット遺伝子の発現量

他の幾つかの遺伝子の発現量で目的関数を説明する

■ Output gene □ Input gene

②ネットワーク推定の問題を p 個の部分問題として扱う。

③それぞれの部分問題をrandom forestで解く。
Random forestは多数の回帰木を作成して問題を解く方法。

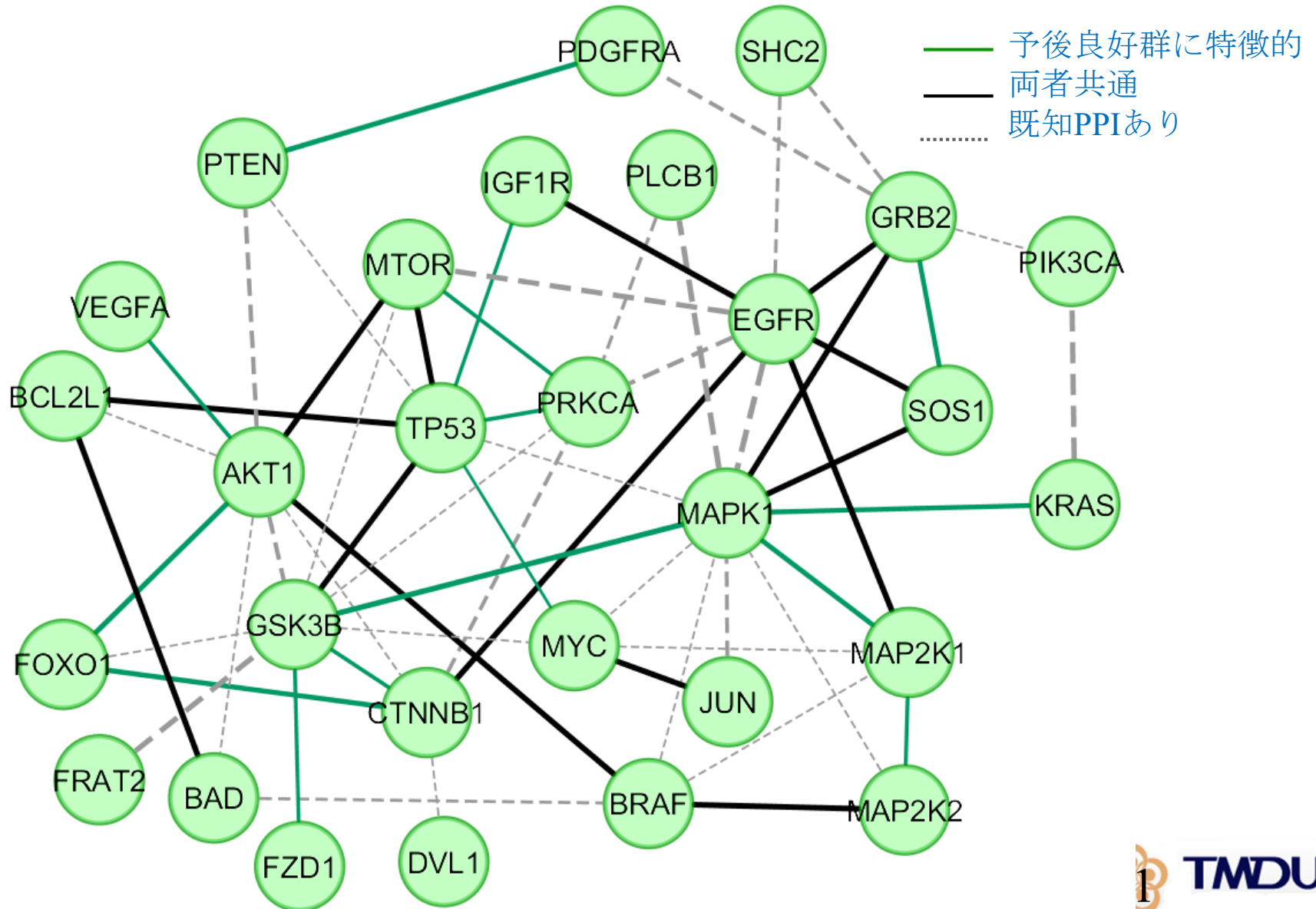


④ p 個のリンクのスコアリストを統合して、全遺伝子のリンク全てに対するスコアリストを作成する。

部分問題の結果はそれぞれの遺伝子のスコアとして表される。高いスコアの遺伝子がターゲット遺伝子を制御する遺伝子とみなされる。そのスコアが、ターゲット遺伝子とその遺伝子の間のリンクのスコアとなる。

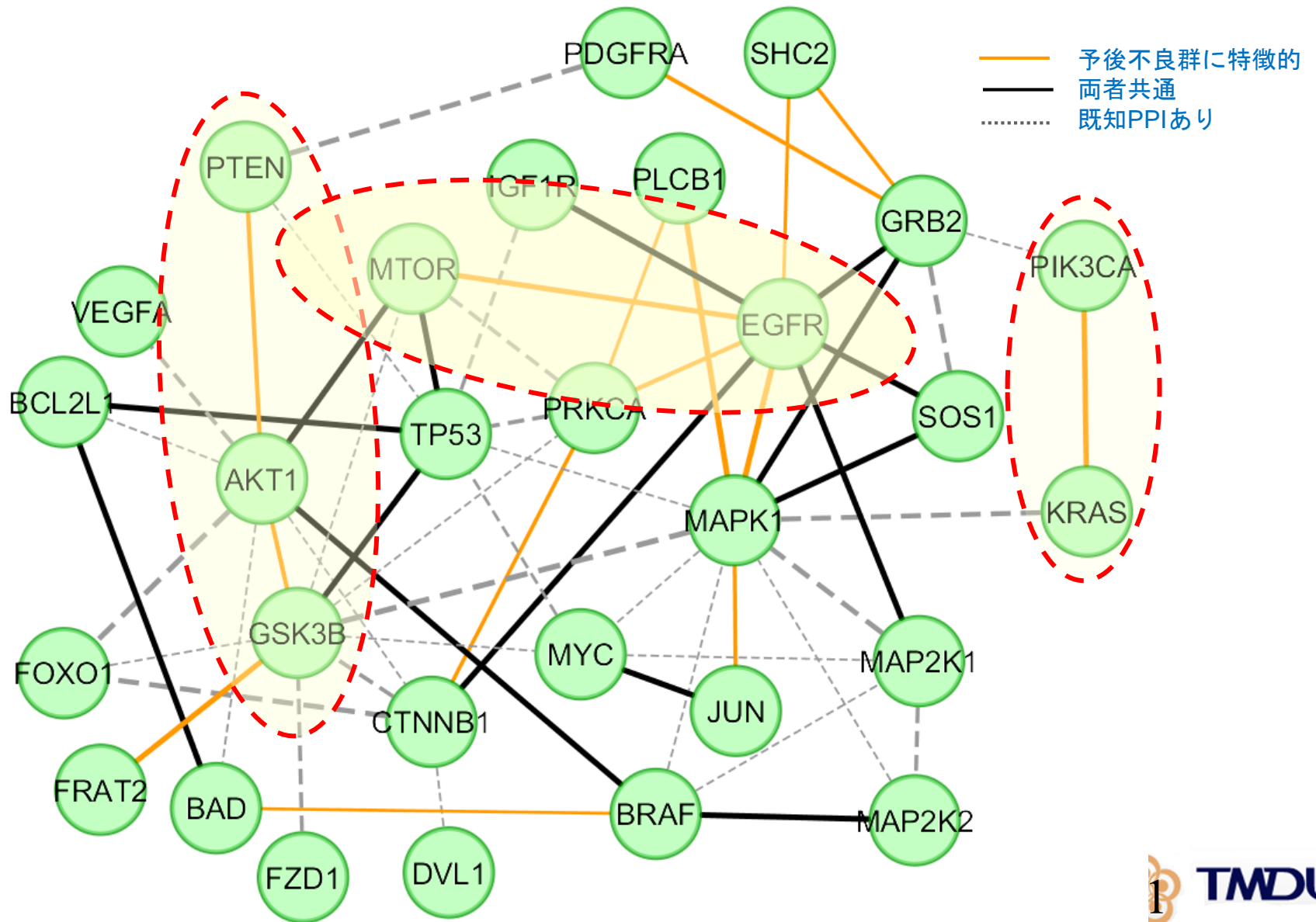
Huynh-Thu VA, Irrthum A, Wehenkel L, Geurts P. Inferring regulatory networks from expression data using tree-based methods. PLoS ONE (2010) 28:e12776.

細胞増殖分子ネットワーク推定 予後良好群ネットワーク



細胞増殖分子ネットワーク推定

differential branches 増殖系の過剰活性

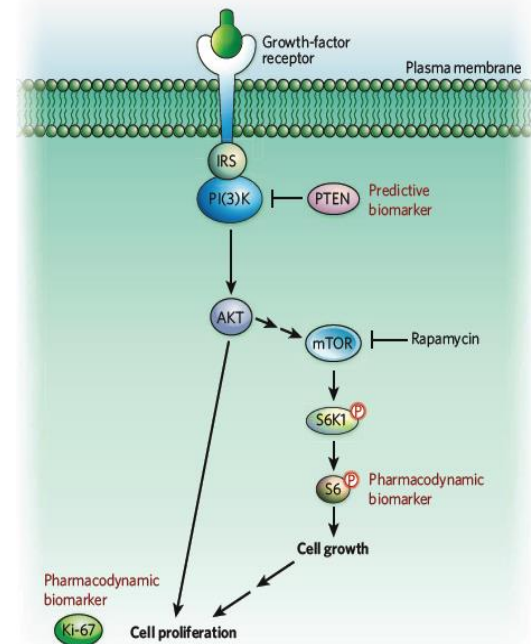
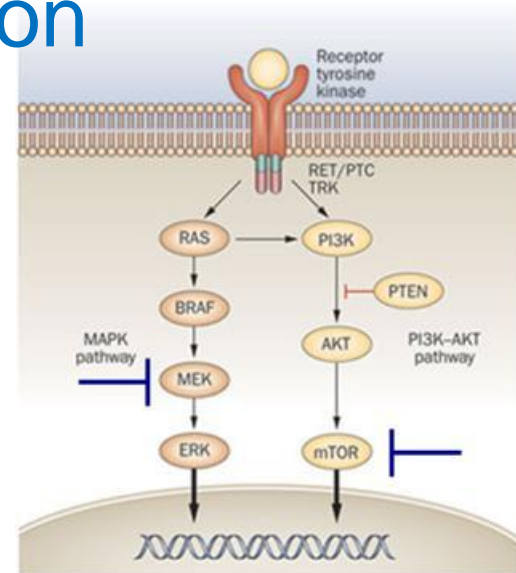


疾患のシステム理解 がんとPathway Addiction

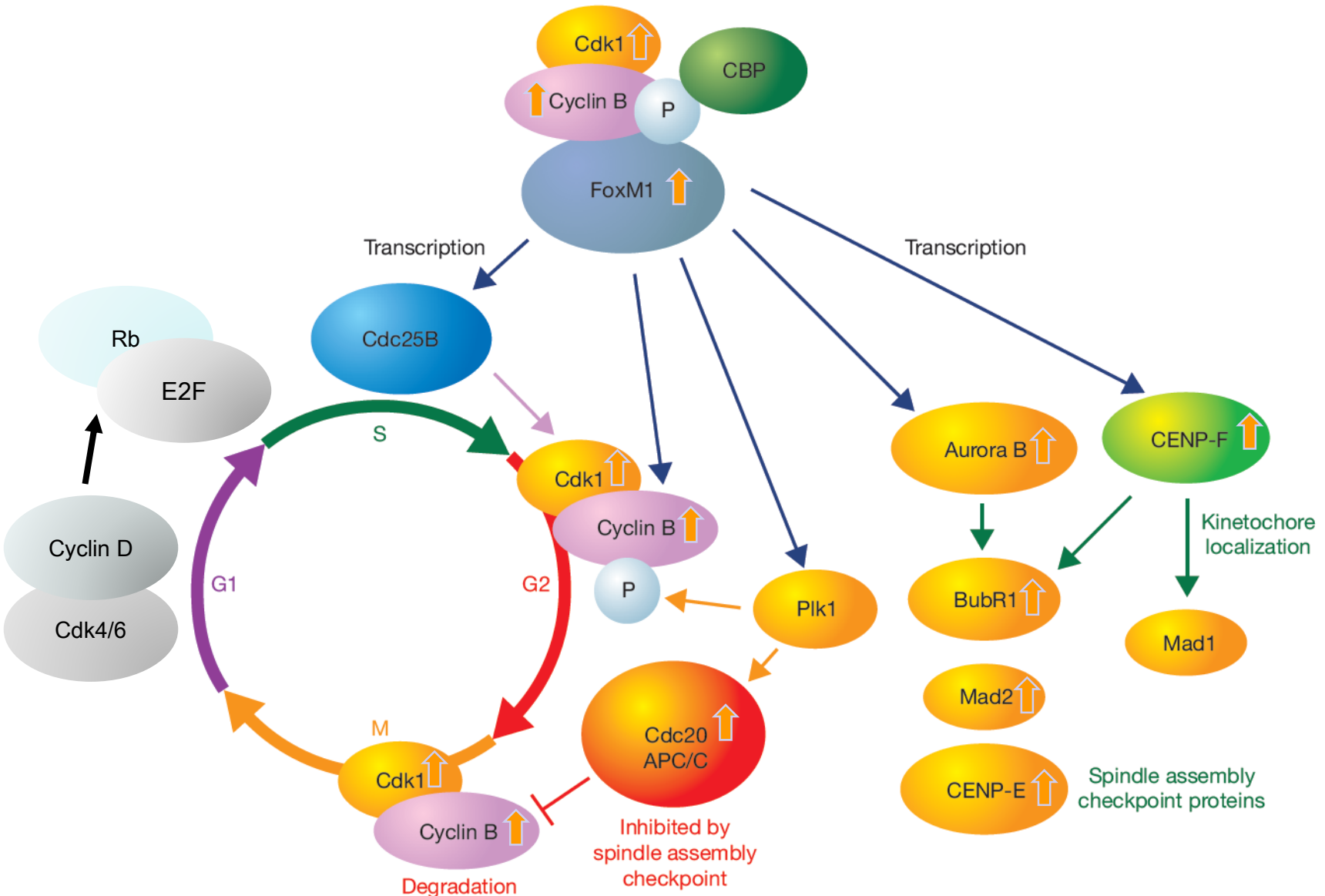
■ 多重シグナル伝達経路の存在

EGFRパスウェイの下流

- Ras-Raf-MEK-ERK
 - PI3K-Akt-mTOR-S6
- がん抑制遺伝子 PTEN 機能喪失
 - PI3K/Akt パスウェイは通常 PTENで抑制
 - がんではPTEN変異のため PI3K/Akt 系活性
 - 全ゲノム配列解析やExomeシーケンシング
- PI3K/Akt/mTORにPathway Addiction
 - MEK 阻害剤がしばしば無効
 - パスウェイ過剰活性化は非線形システム現象
- “パスウェイ活性化バイオマーカー”を
 - リン酸化した S6の抗体マーカー
 - 患者特異的パスウェイ異常の同定
- mTOR inhibitor を使用する
 - PI3K/Akt/mTOR 過剰活性化がんパスウェイの抑制



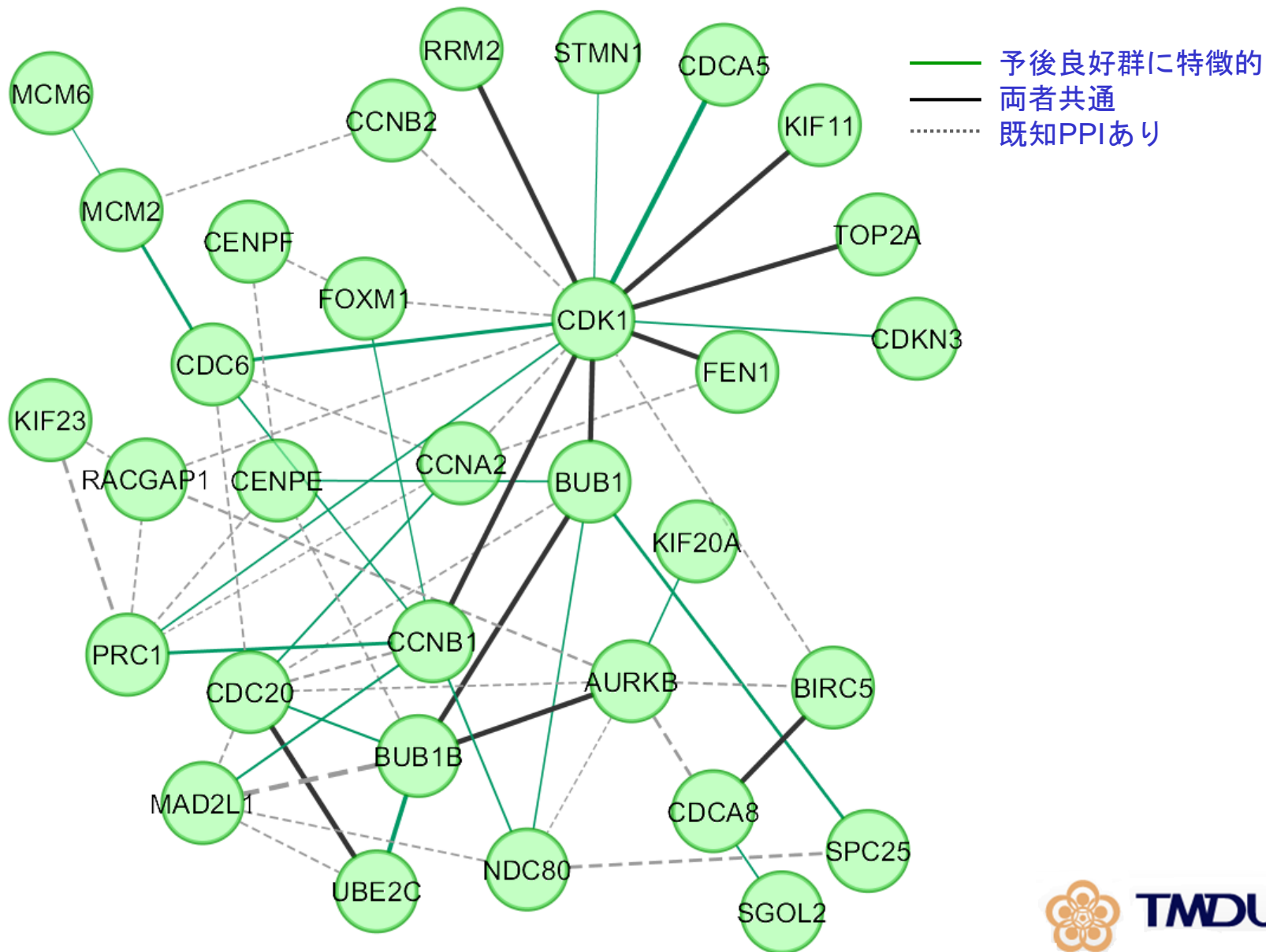
S期～M期に関わる遺伝子群の発現変動



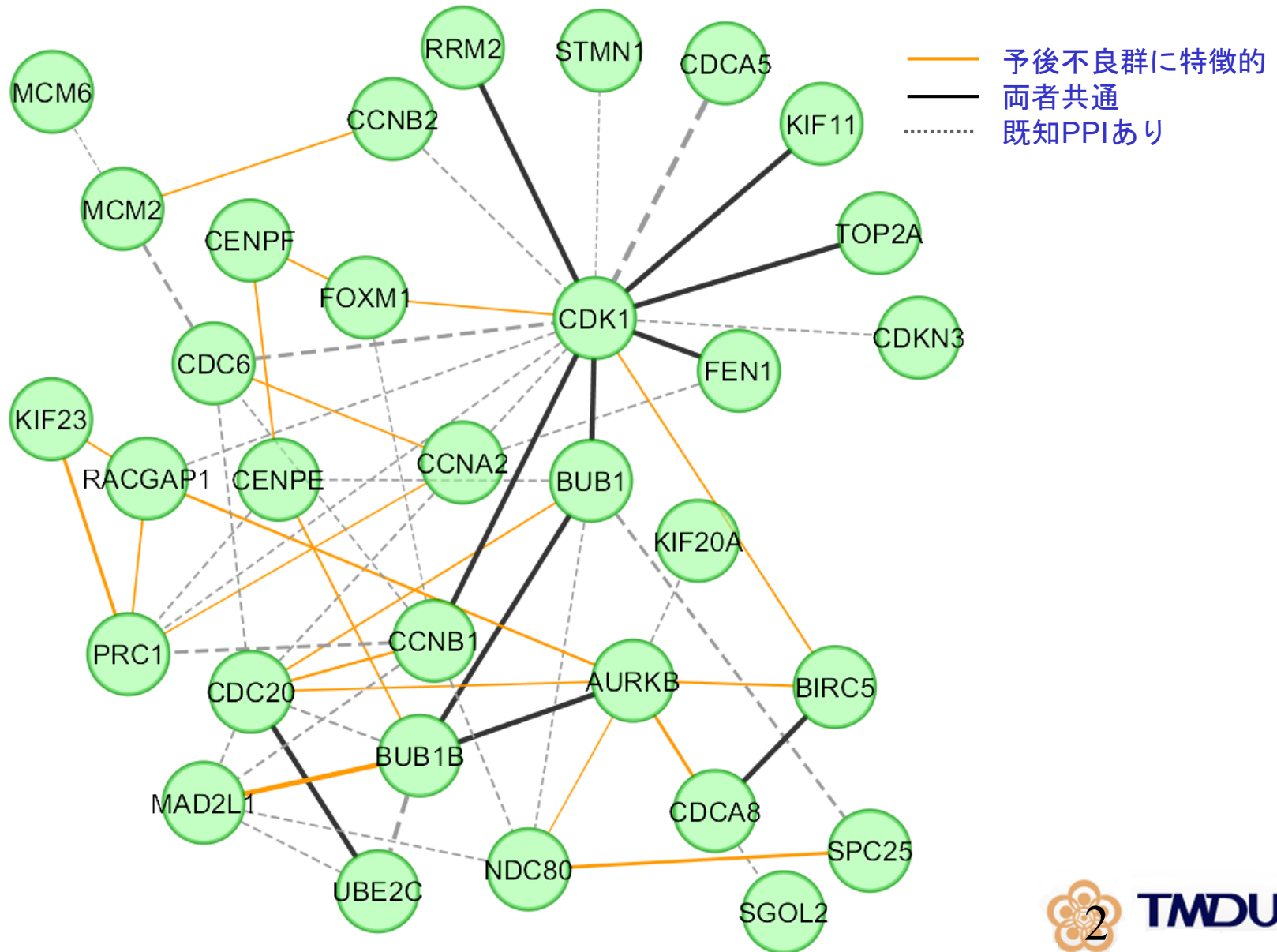
Costa RH. Nat Cell Biol. (2005)

細胞周期 分子ネットワーク推定

予後良好群ネットワーク

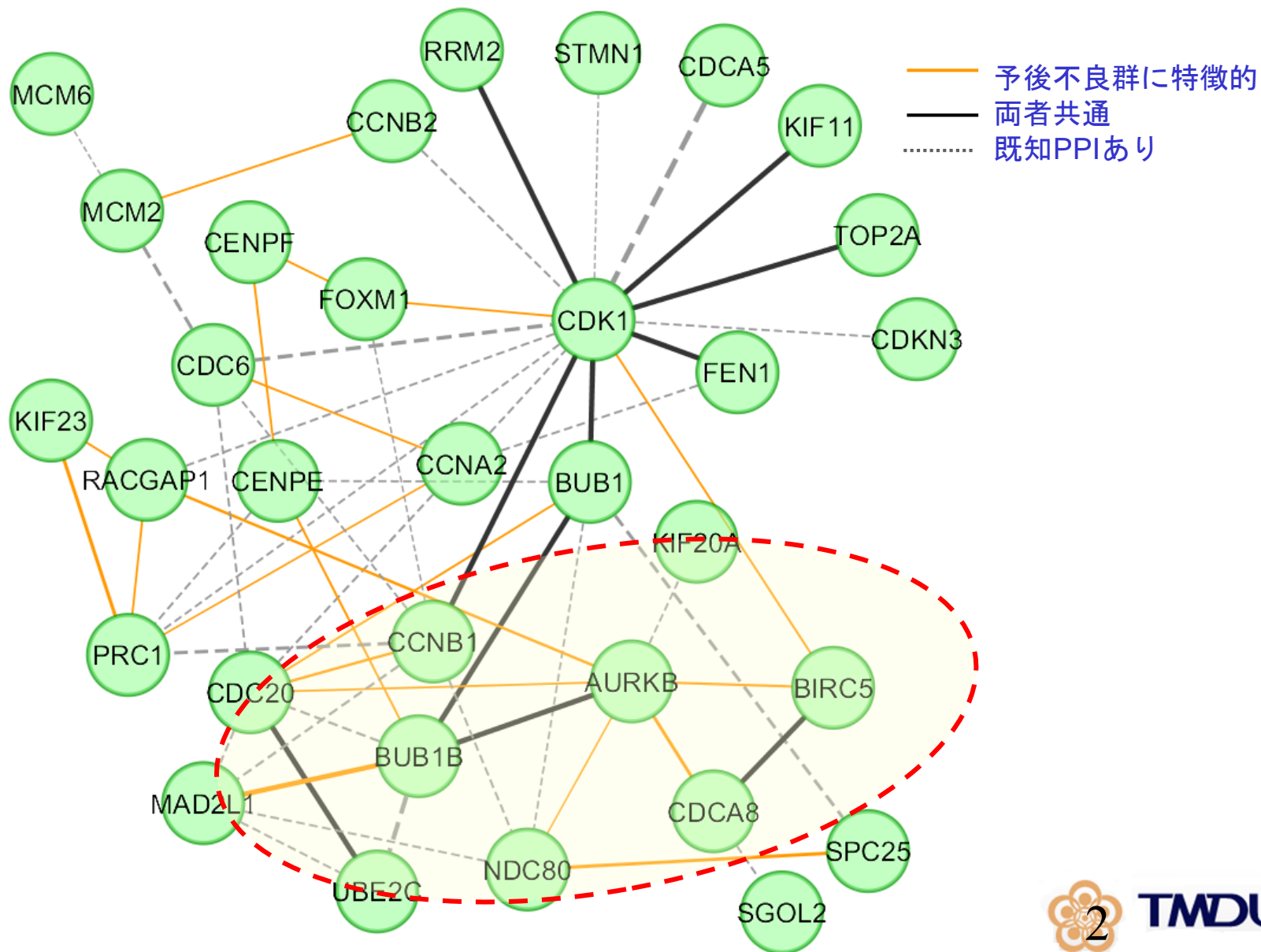


細胞周期 分子ネットワーク推定 予後不良群ネットワーク



細胞周期 分子ネットワーク推定

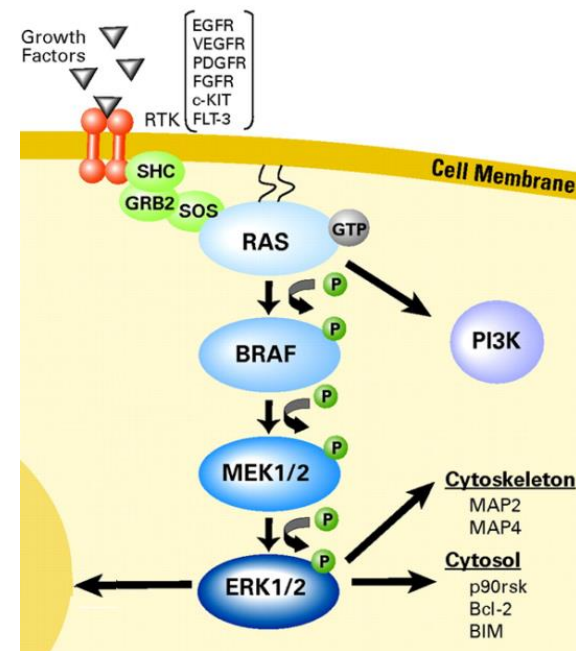
Differential branches 結果 AurkBを起点としたM期NW



システム分子医学と創薬

oncogene addictからpathway addictへ

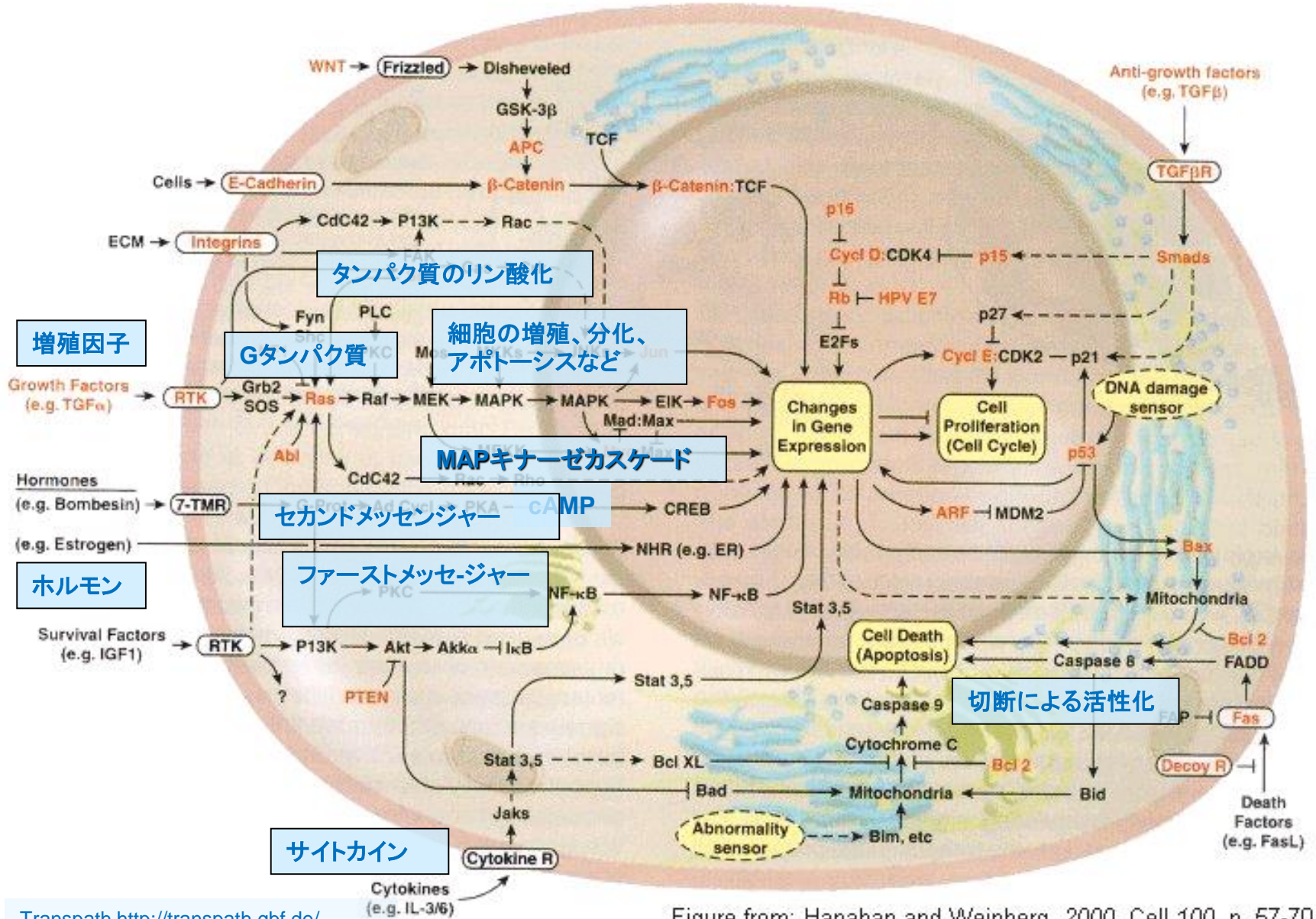
- 分子標的薬と oncogene addiction
 - 分子標的薬の有効性理論根拠
 - 乳がんHer2, 肺がんEGFR
- パスウェイ過剰依存
 - Pathway addiction
 - 標的分子というより、並行パスウェイ間の競合による選別
 - 非線形システム現象（差異拡大型正帰還）
 - EGFRとRASの両方にaddictの場合
 - EGFR阻害剤は腫瘍抑制効果が低い



II 複雑系疾病論

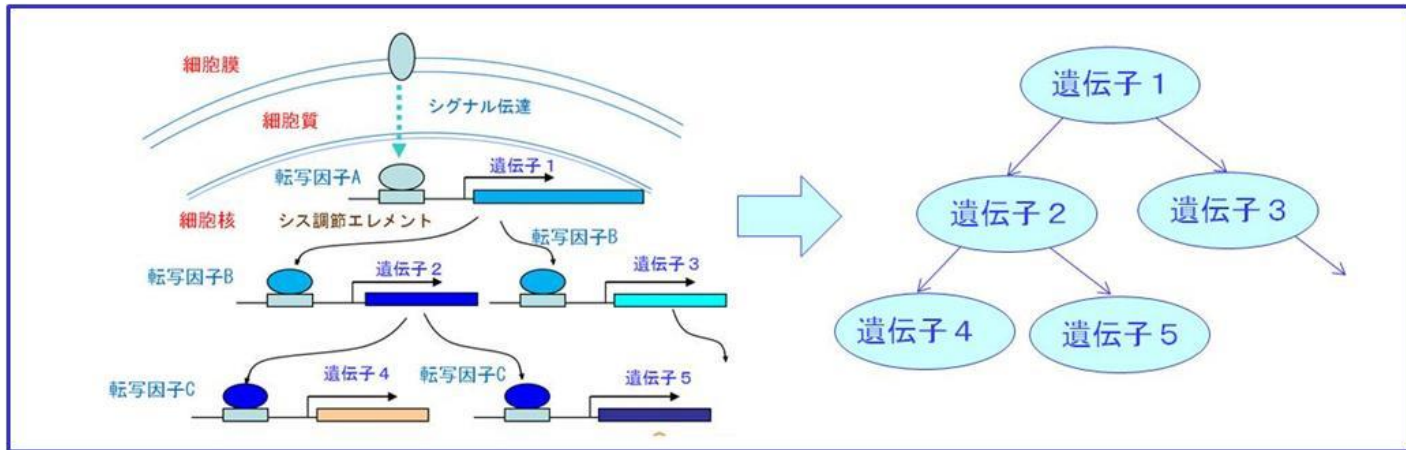
(病態力学系アプローチ)

1. 細胞内シグナル伝達系



2. 遺伝子発現調節ネットワーク gene regulatory network (GRN)

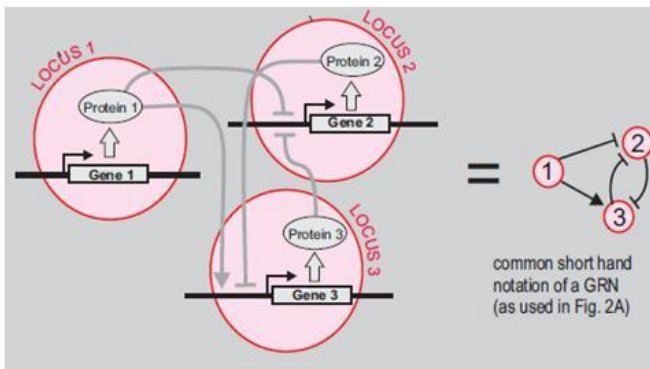
遺伝子発現調節ネットワーク（GRN）は生命系の機能的構造を形成。シグナル伝達系は情報伝達



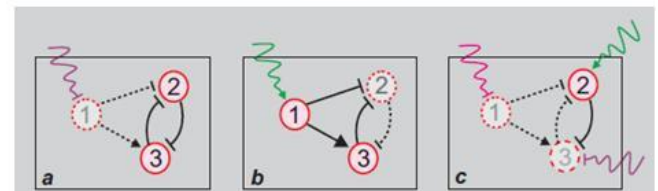
ゲノム
遺伝子座



遺伝子
発現調節
ネットワーク
構造



遺伝子発現調節
ネットワークの
様々な発現パターン



病態進行の機序：分子ネットワーク構造変化

Prion accumulation network

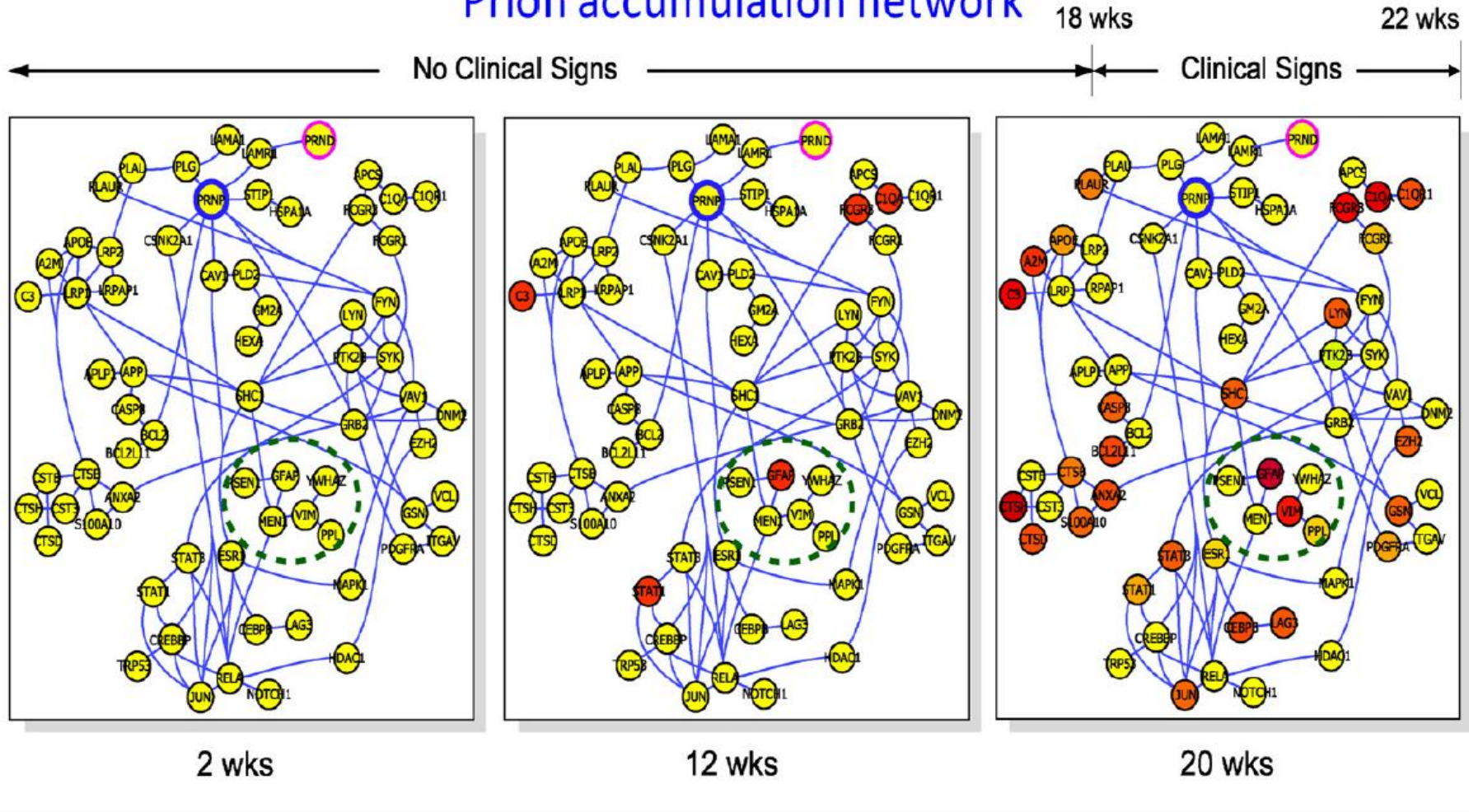


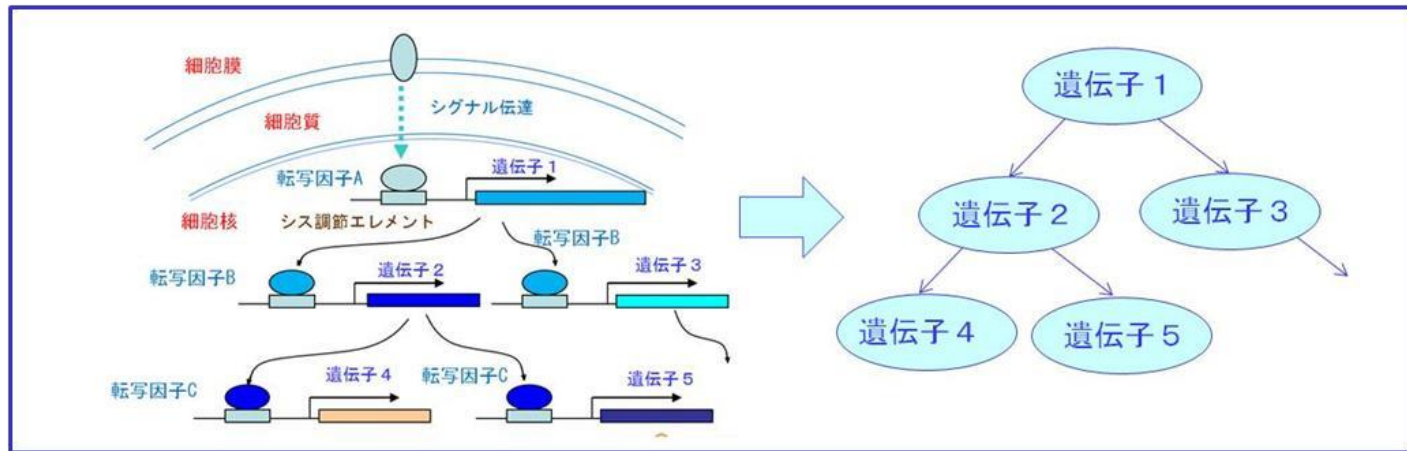
FIGURE 4

A schematic of the prion accumulation and replication network in the prion-induced mouse neurodegenerative disease. The red indicates transcript levels that have been increased in the brains from prion-infected animals as compared with normal control brains. The yellow indicated transcripts that are the same in control and diseased animals. The three panels represent the network at 2, 12 and 20 weeks in animals that live about 22 weeks with this disease. The disease-perturbed networks appear about eight weeks before the clinical signs appear in these animals.

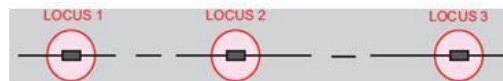
がんと再生医療の
細胞分子ネットワーク空間分析
定量的 Waddington エピゲノム地形分析

遺伝子発現調節ネットワーク (再掲) gene regulatory network (GRN)

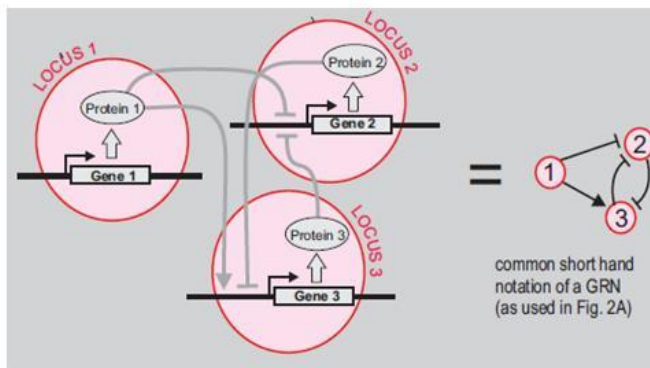
遺伝子発現調節ネットワーク (GRN) は生命系の機能的構造を形成。シグナル伝達系は情報伝達



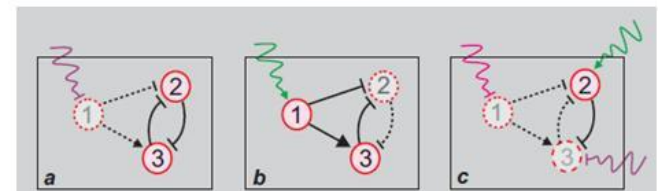
ゲノム
遺伝子座



遺伝子
発現調節
ネットワーク
構造

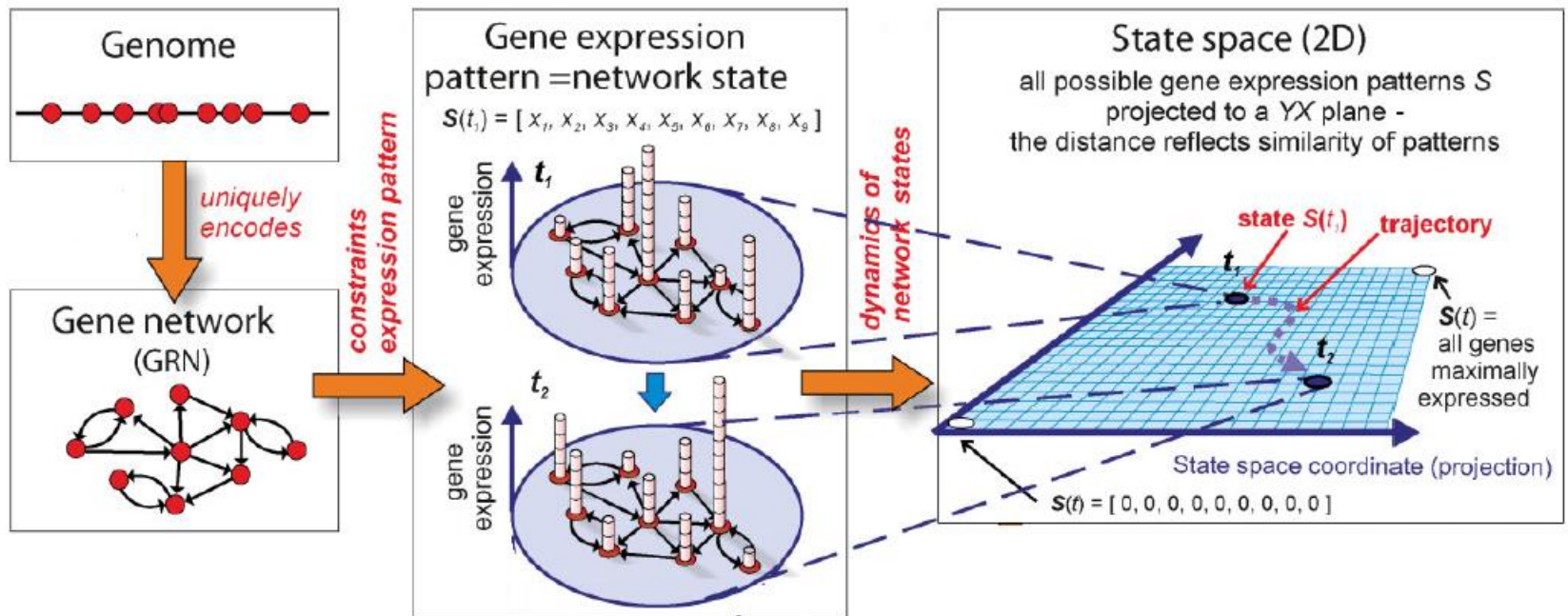


遺伝子発現調節
ネットワークの
様々な発現パターン



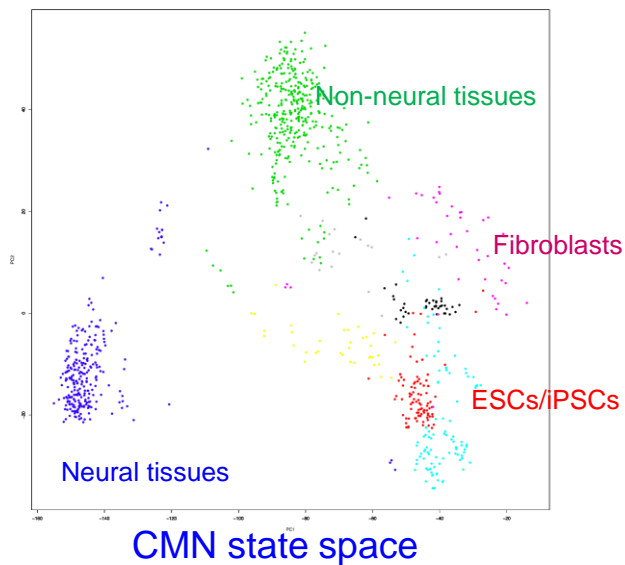
細胞分子ネットワークの状態空間

- 細胞分子ネットワーク (GRN) の状態空間を考える
- 細胞分子ネットワークの結線構造はゲノムにコードされて生得的/固定的である
- GRNの中でどの遺伝子が発現しているかは、細胞型および状態によって異なる

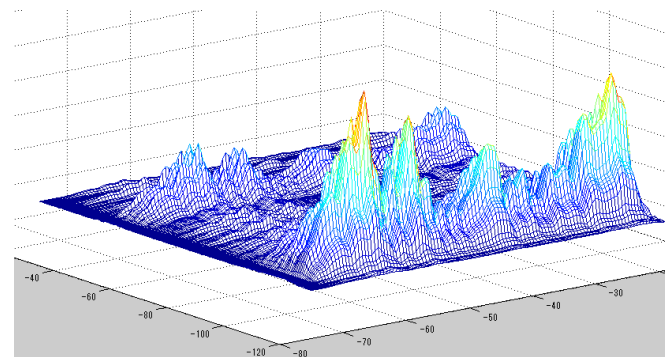


細胞型とGRN空間での分布

- GRN空間内での220種類の細胞型 (cell type)の発現
- 一定の細胞分子ネットワークの発現パターン (安定な発現プロファイル) に対応する点の周りに集積
- GRN空間に頻度分布 (確率構造)



CMN state space



各細胞型: 共通に高頻度で観測される
遺伝子発現プロファイルとして集積

定量的Waddington Epigenetic Landscape (qWEL 理論)

- 細胞分子ネットワーク空間
 - 遺伝子のネットワークの活性化状態を表す
 - 遺伝子発現プロファイル (=GRN) の空間
- GRN空間における確率分布 $\Psi(r)$
 - GRN空間における遺伝子発現プロファイルの頻度分布 $\Psi(r)$
- 確率分布を準ポテンシャル表示に変換

定量的Waddingtonエピゲノム地形

- quantified Waddington Epigenetic Landscape
- ここでは“qWEL (理論)”とよぶ
- ボルツマンの関係式 $f(x) = \exp(-E/kT)$

$$U(r) = -\log \Psi(r)$$

Waddington のエピゲノム地形 (Epigenetic Landscape)

- 発生学者Waddingtonが細胞運命の譬喩として提唱(1942,1957)

- 細胞運命 (cell fate) : 細胞の発生過程

多能性幹細胞(ES/iPSC)から

前駆細胞 (progenitor cell) を経由して

成熟分化細胞型 (matured cell type)

球が稜線間の谷間を経由して転がる

- 細胞は、分水嶺 (Dividing ridge) の麓で

どの谷間を転げ落ちるか

選択する

- Dividing Ridge (分水嶺) :

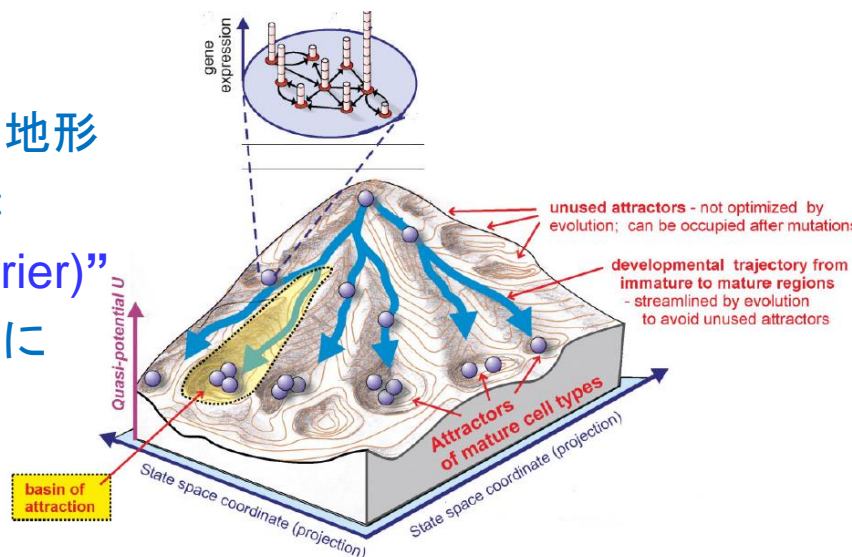
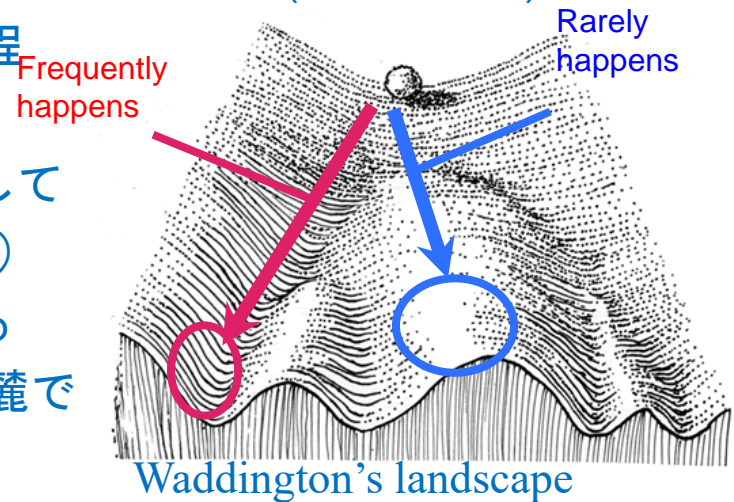
細胞分子ネットワークの構造が作る地形

エピゲネティックな遺伝子調節機構

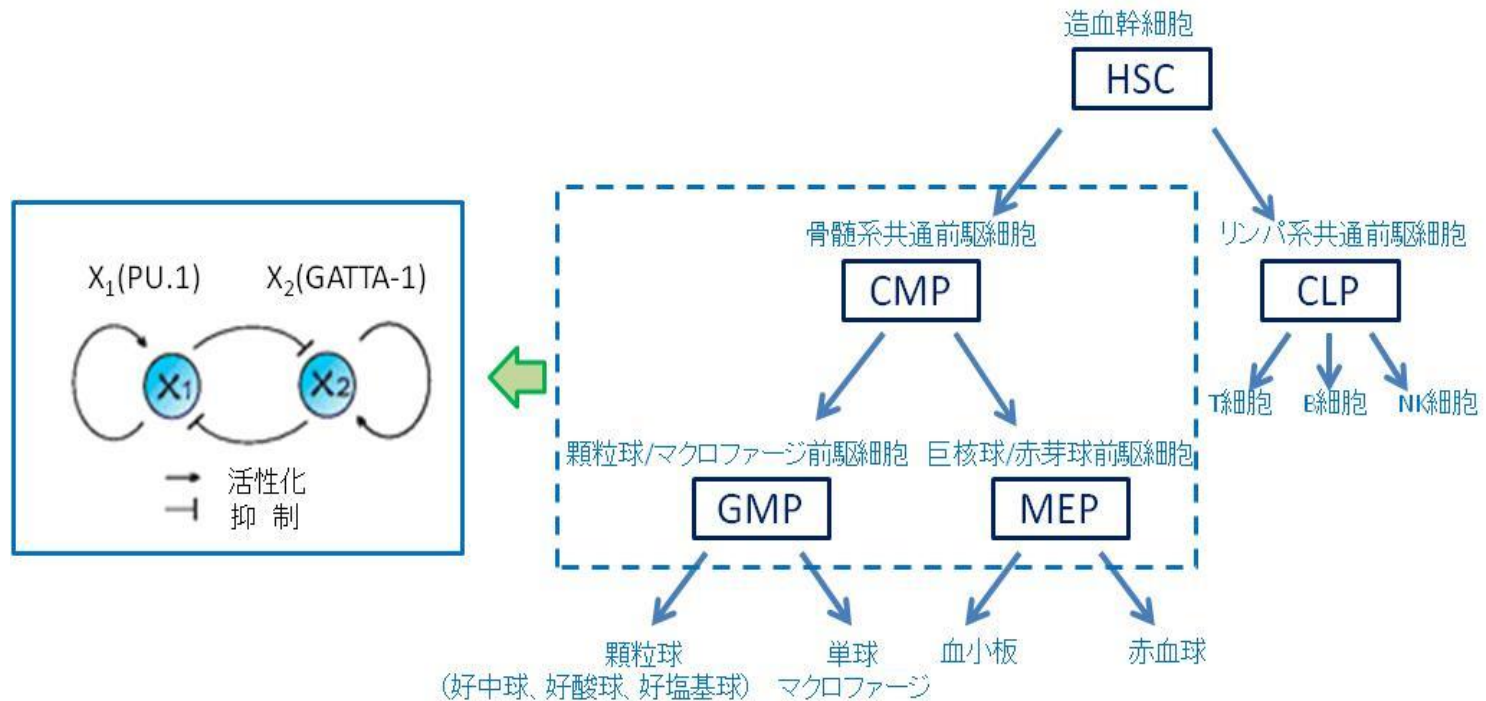
→”エピゲノム障壁 (Epigenetic Barrier)”

- 成熟細胞型はWaddington landscape に

Basinを作る



数理解析によるqWEL地形の導出 Wangのモデル



造血幹細胞からの血球成分への分化と骨髓系共通前駆細胞CMPの発生運命を決定する転写因子PU.1とGATA-1の相互抑制動態モデル

定量的Waddington Epigenetic Landscape (qWEL 理論)

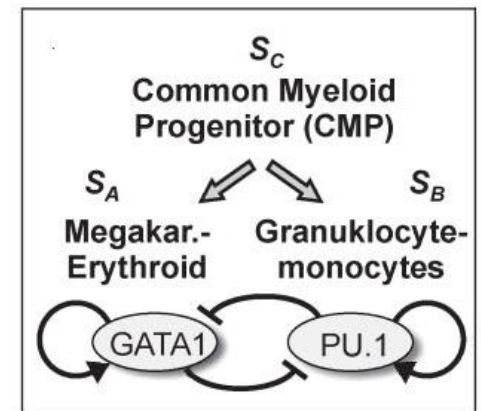
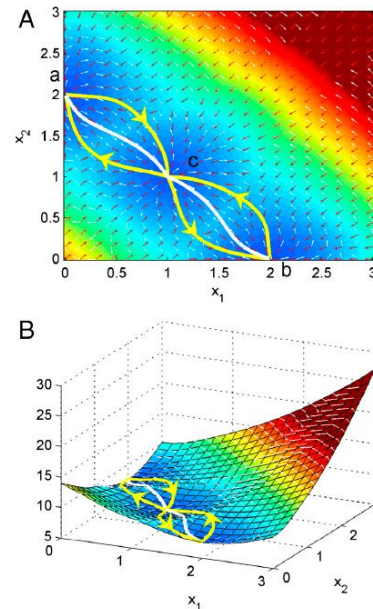
- 多能性幹細胞（前駆細胞）の運命選択
 - 骨髓系前駆細胞 Myeloid Progenitor Cell (CMP)
 - 赤芽球系前駆細胞 Megakaryocyte-Erythroid Progenitor Cell (GATA1)
 - 顆粒球マクロファージ前駆細胞 Granulocyte-Macrophage Progenitor Cell (PU.1)
- 方程式

$$\frac{dx_1}{dt} = \frac{a_1 x_1^n}{S^n + x_1^n} + \frac{b_1 S^n}{S^n + x_2^n} - k_1 x_1 = F_1(x_1, x_2)$$

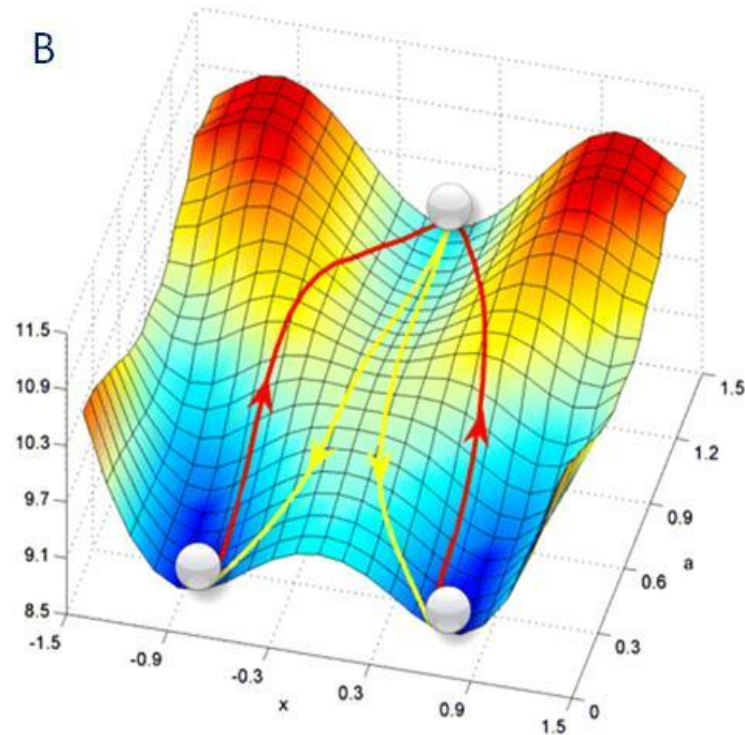
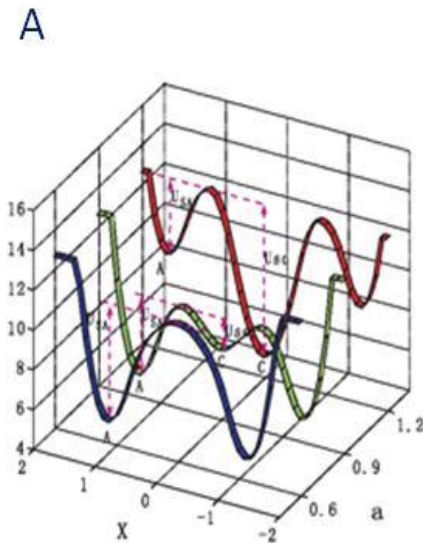
$$\frac{dx_2}{dt} = \frac{a_2 x_2^n}{S^n + x_2^n} + \frac{b_2 S^n}{S^n + x_1^n} - k_2 x_2 = F_2(x_1, x_2)$$

$$dx/dt = \mathbf{F}(\mathbf{x}) = [F_1(x_1, x_2), F_2(x_1, x_2)]$$

$$\frac{da}{dt} = -\lambda a$$



Wangの2遺伝子モデルとWaddington地形

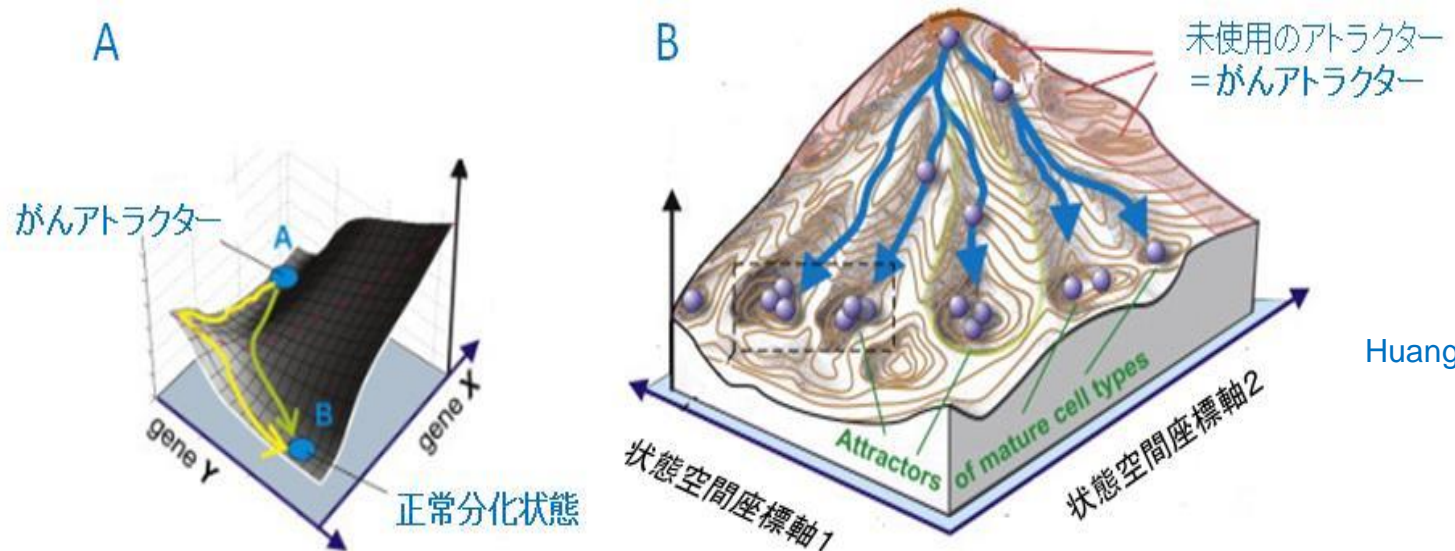


発生進行パラメータaとWaddingtonエピジェネティック地形

図A パラメータaによる準ポテンシャルの横断面の変化 (aが大きくなると中央のCMPのアトラクターが不安定になり、両脇のMEP、GMPのアトラクターが深くなる。図B 発生に伴うaの変化を取り入れたアトラクター地形の変化と細胞状態の分化過程。オリジナルなWaddingtonエピジェネティック地形と同等の地形が構成できた。

がんアトラクターと病態力学系治療方針

- Kaufman : がんアトラクター理論
 - がんアトラクターとは正常の分化パス（青色のパス）から離れたところのアトラクターで、通常は未使用のアトラクターである
 - 遺伝子発現調節ネットワークの可能解として存在
- がん治療への方針
 - 状態空間でのがんアトラクターの位置（A）と正常分化状態（B）
 - A : 現在のがん病態の状態空間での位置
 - B : 正常分化状態
- qWEL地形を参照した治療方針
 - A→BのパスをX方向に継いでY方向に遷移させるよりもX,Y同時に動かす方がバリアを登らなくてよい。
 - 複数抗がん剤による最適治療制御の可能性



Huangから著者改変

我々の方法

経験的ポテンシャル法 (empirical qWEL)

- GEOなどの遺伝子発現プロファイルDBの遺伝子発現プロファイルのデータから
GRN空間に経験的頻度分布を描く
- 頻度分布を正規化
 - リサンプロリングによって緻密にする
 - 細胞型のアプリアリな事前確率を乗して調整する
- 主成分分析を行い2成分系の座標を計算
 - 確率分布をスプライン関数などで補完する
- 確率密度を量子化してボルツマン関係式より
- $U = -\ln \Psi$ より準ポテンシャルを計算

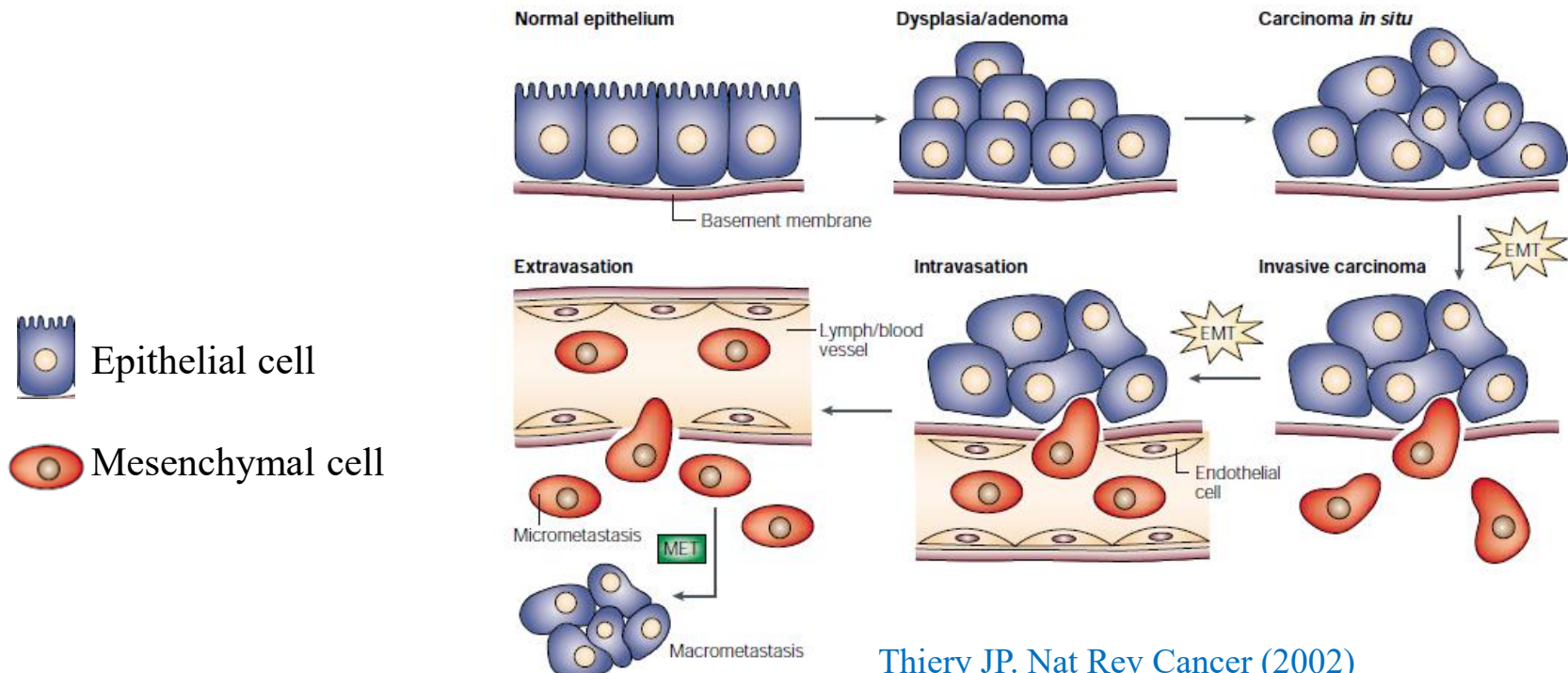
がんの転移過程

上皮間葉転換(EMT)の病態力学

Epithelial-Mesenchymal Transition

上皮間葉転換(EMT)

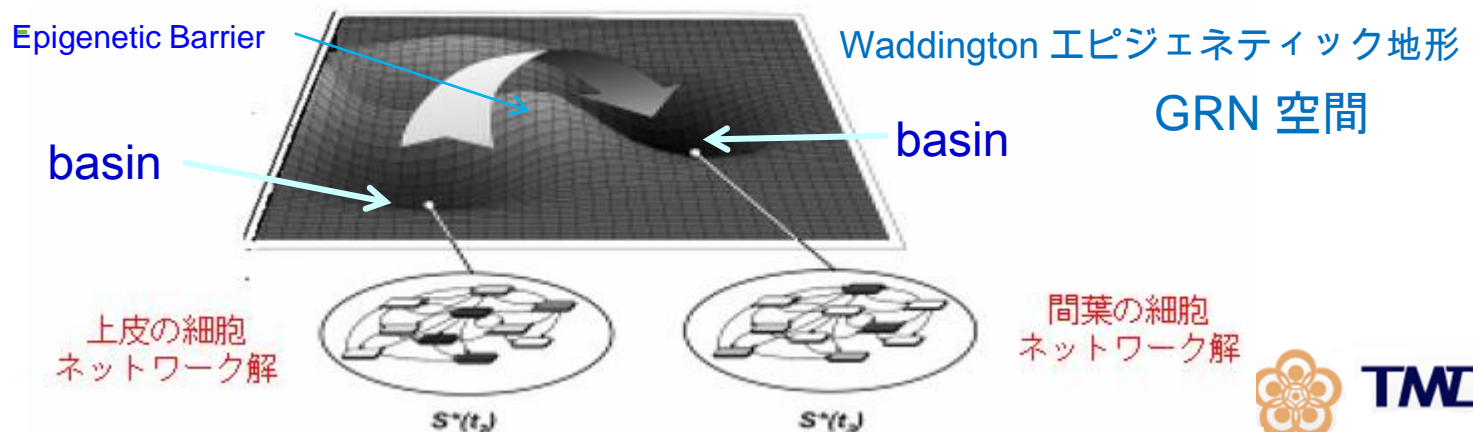
- Epithelial-mesenchymal transition (EMT)
 - non-motile, polarized epithelial cells, embedded via cell-cell junctions, dissolve their cell-cell junctions and convert into individual, non-polarized, motile and invasive mesenchymal cells
- EMT は生物に取っても重要な発生現象
 - Gastrulation (原腸陥入)、神経堤形成など
- EMT は、細胞分子ネットワーク発現の大域的变化



がんの浸潤転移

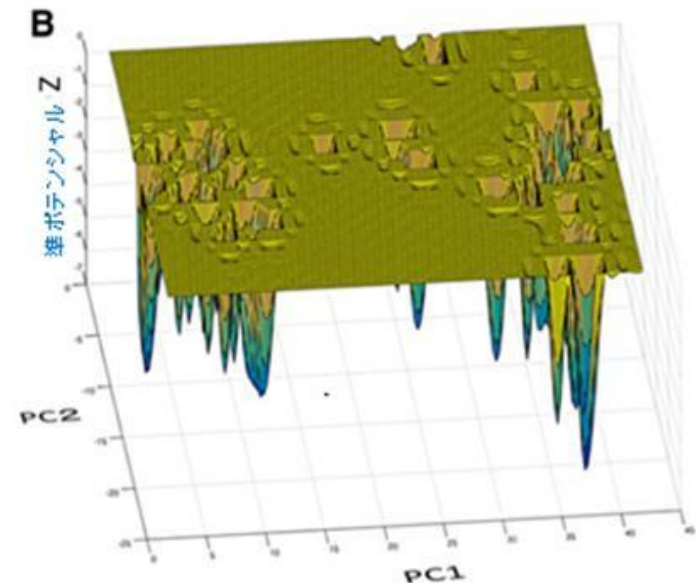
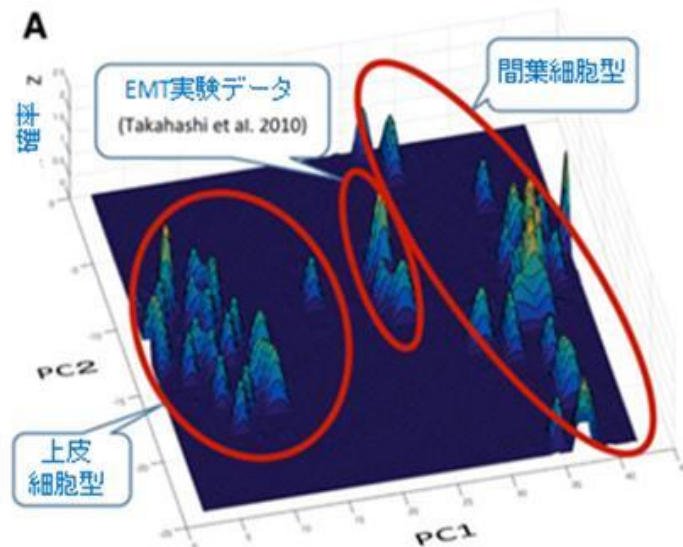
細胞分子ネットワーク空間の大域変異

- 細胞分子ネットワーク状態 (GRN状態空間)
 - 細胞分子ネットワークはいくつかの安定解を持つ
 - 各安定解は細胞種類に相当する(約220種類の安定解)、細胞ネットワークが作るCMN空間のエピゲノム地形(準ポテンシャル)「アトラクター状態」これらは多能性幹細胞から分化解である
- EMT (上皮・間葉転移) Epithelial-mesenchymal Transition
 - 細胞分子ネットワークの上皮安定解が間葉安定解へと移行
 - 全ネットワーク状態の転移であるから「相転移」といえる
- がんの転移とは相転移
 - 安定解を不安定化し(ゲノム不安定性)、障壁を越えるためにはゲノムの不安定性が必要である。



GRN部分空間でのEM状態の表示

- **統計的頻度分布の構築**
 - GEO, ArrayExpressより上皮細胞（E）および間葉細胞（M）の遺伝子発現プロフィール・データを収集
 - EMT過程の時系列データとして利用する高橋らの実験的EMTデータを追加
 - Takahashi E, Nagano O, et al. (2010); J Biol Chem. 285(6):4060-73.
- **EMT部分空間の構築**
 - そのままでは4万次元GRN空間、EとMの識別性の高い遺伝子で部分空間を抽出
 - 61遺伝子よりなるGRN部分空間を得る
- **可視化3次元表示のための主成分分析**
 - 61遺伝子の主成分分析を行い、第1-2主成分で2次元座標上に、遺伝子発現データを表す3次元表示系を取得
 - 頻度分布および、そのボルツマン変換により準ポテンシャル表示



第 1, 2 主成分の意味

- **PC1: E状態とM状態を一般的に分離する遺伝子**
 - FN1 (fibronectin), COL1A1, COL13A1(Collagen), NEGR1, TNC(Cell adhesion) , ARHGAP24(cell motility)など
- **PC2: がんアトラクターにおけるE状態とM状態を分離する遺伝子**
 - HMMR, PBK, CKMT1A/1B, and BRI3BP (がん関連遺伝子)
 - 細胞周期関連遺伝子cyclin B1, MAD2L1

(a) The top 20 PC1 loadings

| Gene symbol | Gene name | PC1 loading |
|-----------------|--|-------------|
| ELTD1 | EGF, latrophilin and seven transmembrane domain containing 1 | 0.261823 |
| <u>FN1</u> | <u>Fibronectin</u> | 0.238429 |
| CNN1 | Calponin 1, basic, smooth muscle | 0.206869 |
| <u>COL1A1</u> | <u>Collagen, type I, alpha 1</u> | 0.203478 |
| <u>NEGR1</u> | <u>Neuronal growth regulator 1</u> | 0.190520 |
| FBN1 | Fibrillin 1 | 0.187782 |
| ARMCX1 | Armadillo repeat containing, X-linked 1 | 0.182721 |
| WNT5B | Wingless-type MMTV integration site family, member 5B | 0.182222 |
| XK | X-linked Kx blood group | 0.172636 |
| BAALC | Brain and acute leukemia, cytoplasmic | 0.163815 |
| GAP43 | Growth associated protein 43 | 0.157189 |
| ENPP1 | Ectonucleotide pyrophosphatase/ phosphodiesterase 1 | 0.156683 |
| <u>COL13A1</u> | <u>Collagen, type XIII, alpha 1</u> | 0.156403 |
| DSC2 | Desmocollin 2 | 0.154287 |
| PLSCR4 | Scramblase 4 | 0.148105 |
| <u>TNC</u> | <u>Tenascin C</u> | 0.143167 |
| RCAN2 | Regulator of calcineurin 2 | 0.140531 |
| 1560208_at | | 0.135836 |
| <u>ARHGAP24</u> | <u>Rho GTPase activating protein 24</u> | 0.135123 |

(b) The top 20 PC2 loadings

| Gene symbol | Gene name | PC2 loading |
|-------------|---|-------------|
| DSC2 | Desmocollin 2 | 0.266742 |
| ENPP1 | Ectonucleotide pyrophosphatase/ phosphodiesterase 1 | 0.264823 |
| ASPM | Asp (abnormal spindle) homolog, microcephaly associated | 0.258160 |
| HMMR | <u>Hyaluronan-mediated motility receptor (RHAMM)</u> | 0.246326 |
| PBK | <u>PDZ binding kinase</u> | 0.238582 |
| CKMT1A/1B | <u>Creatine kinase, mitochondrial 1A/1B</u> | 0.221564 |
| AP1S3 | Adaptor-related protein complex 1, sigma 3 subunit | 0.217856 |
| MIR222HG | mir host gene | 0.216734 |
| CNN1 | Calponin 1, basic, smooth muscle | 0.192942 |
| MAD2L1 | <u>MAD2 mitotic arrest deficient-like 1</u> | 0.189140 |
| CCNB1 | <u>Cyclin B1</u> | 0.173976 |
| CLIP3 | CAP-GLY domain containing linker protein 3 | 0.171328 |
| KCND3 | Potassium channel, voltage gated Shal related subfamily D, member 3 | 0.169711 |
| ELTD1 | EGF, latrophilin and seven transmembrane domain containing 1 | 0.159304 |
| HPS3 | Hermansky-Pudlak syndrome 3 | 0.144361 |
| GAP43 | Growth associated protein 43 | 0.133483 |
| PVRL3 | Poliovirus receptor-related 3 | 0.132407 |
| BRI3BP | <u>BRI3 binding protein</u> | 0.132175 |
| XK | X-linked Kx blood group | 0.128981 |

網膜色素上皮細胞(培養細胞) におけるEMT過程と 遺伝子発現ネットワーク(Cell line of EMT)

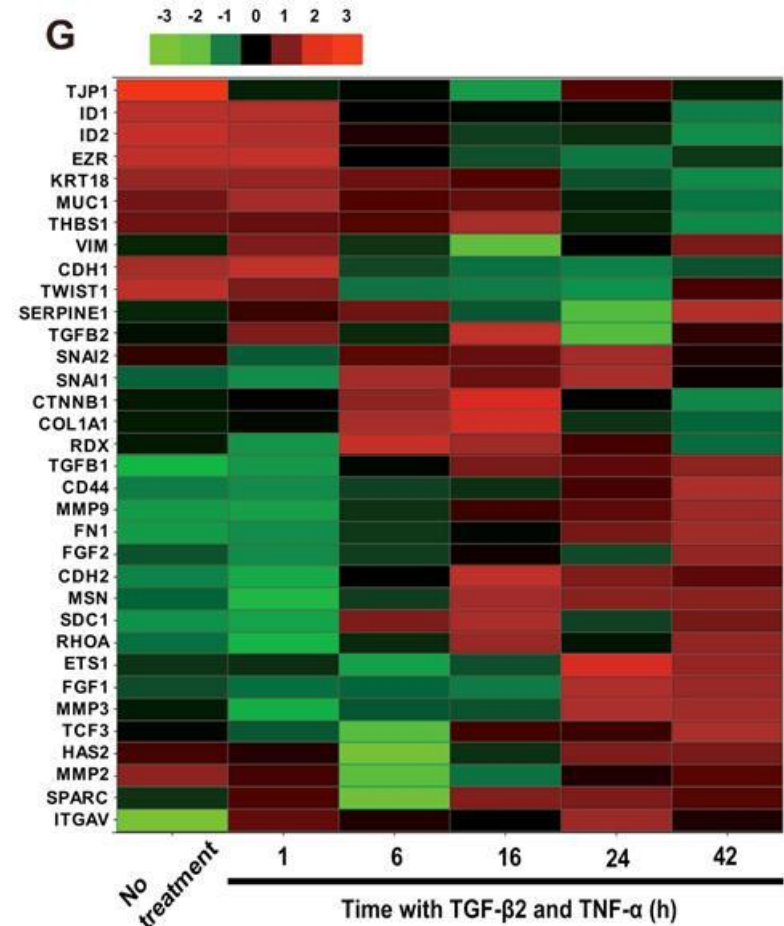
高橋らの網膜色素上皮細胞 (retinal pigment epithelial cell) の実験データを使用。
TGF- β , TNF- α を添加すると自動的にEMTを
起こす。我々研究室で再実験。

観測時点

0h, 1h, 6h, 16h, 24h, 42h, 60h
2,218 probe sets (1,797 genes)

Takahashi E, Nagano O, et al. (2010)
Tumor necrosis factor-alpha regulates
transforming growth factor-beta-dependent
epithelial-mesenchymal transition by
promoting hyaluronan-CD44-moesin
interaction.

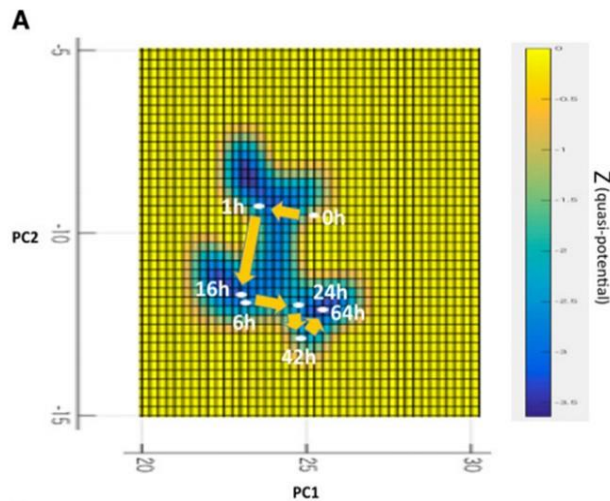
J Biol Chem. 285(6):4060-73.



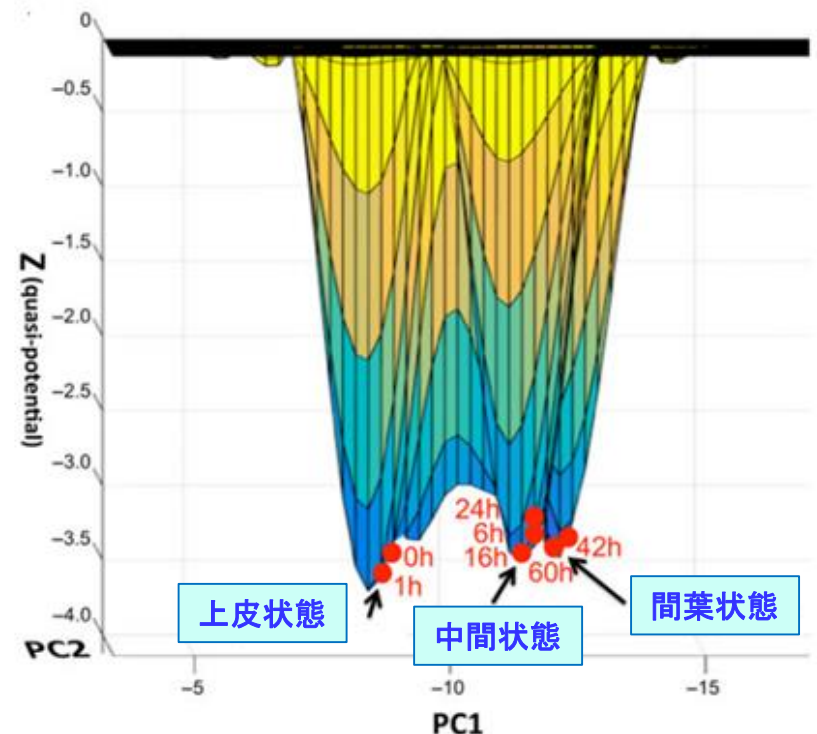
定量Waddington地形でのアトラクター遷移

- 3次元Waddington地形での表示
 - がん転移過程のアトラクター遷移
 - 準ポテンシャルの3次元表示
- 中間状態への停留を明示
 - 「がん転移3状態論」 確証

準ポテンシャル上の遷移の3次元表示
(側面ビュー：横からの眺め)



準ポテンシャルの深さの情報を表示
(俯瞰ビュー：上からの眺め)



材料と方法

- Gene expression data of a human retinal pigment epithelium cell line (ヒト網膜色素上皮細胞 ARPE-19) in response to treatment of TGF, which induces EMT; total RNA was extracted and gene expression was measured at each time point (0, 1, 6, 16, 24, 42, 60h) using GeneChip (Affy HGU133 Plus 2.0).

Takahashi E, Nagano O, Ishimoto T, Yae T, et al. (2010) Tumor necrosis factor-alpha regulates transforming growth factor-beta-dependent epithelial-mesenchymal transition by promoting hyaluronan-CD44-moesin interaction. J Biol Chem. 285(6):4060-73.

- ARACNe algorithm to infer gene regulatory network based on time-series gene expression data with 1000 bootstraps.

2つの遺伝子発現 (X、Y) の相互情報量 $I(X,Y)$ を計算し、そのリンクのp値を計算する。間接的な関連を除去するためにデータ処理不等式を使用する。1000回のブートストラップにより90%以上を採用

Margolin AA, Nemenman I, Basso K, Wiggins C, et al. (2006) ARACNE: an algorithm for the reconstruction of gene regulatory networks in a mammalian cellular context. BMC Bioinformatics 7 (Suppl 1): S7.

Results: EMT遺伝子関連ネットワークを推定 するための遺伝子の選択

Identification of target genes to infer gene-gene influence network:

total **5183** probe sets (2988 genes)

1) Identification of (DEM) differentially expressed genes SAM;
comparison b/w successive two time points; $p < 0.001$; $FC > 1.5$

3471 probe sets (1766 genes)

2) Identification of upward/downward trend genes (accumulated chi-squared test; $p < 1.0 \times 10^{-10}$)

1689 probe sets (1203 genes)

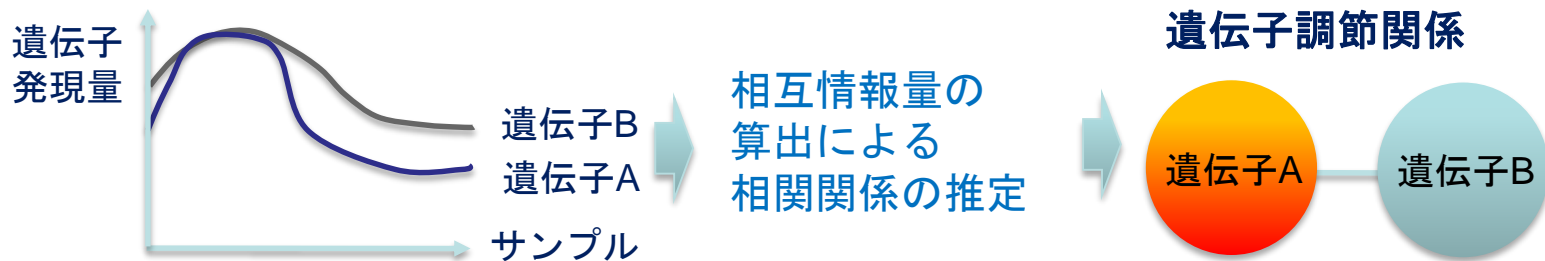
3) Collection of known epithelial/mesenchymal marker genes and EMT related genes

34 genes

ARACNeによる遺伝子調節ネットワークの推定

- ARACNeは遺伝子調節ネットワークの推定アルゴリズムで、各遺伝子の発現量を確率変数にし、2遺伝子の同時確率を個別確率の積で除算して相互情報量を算出し、2遺伝子の相関関係を推定

ARACNeによる遺伝子調節ネットワークの推定

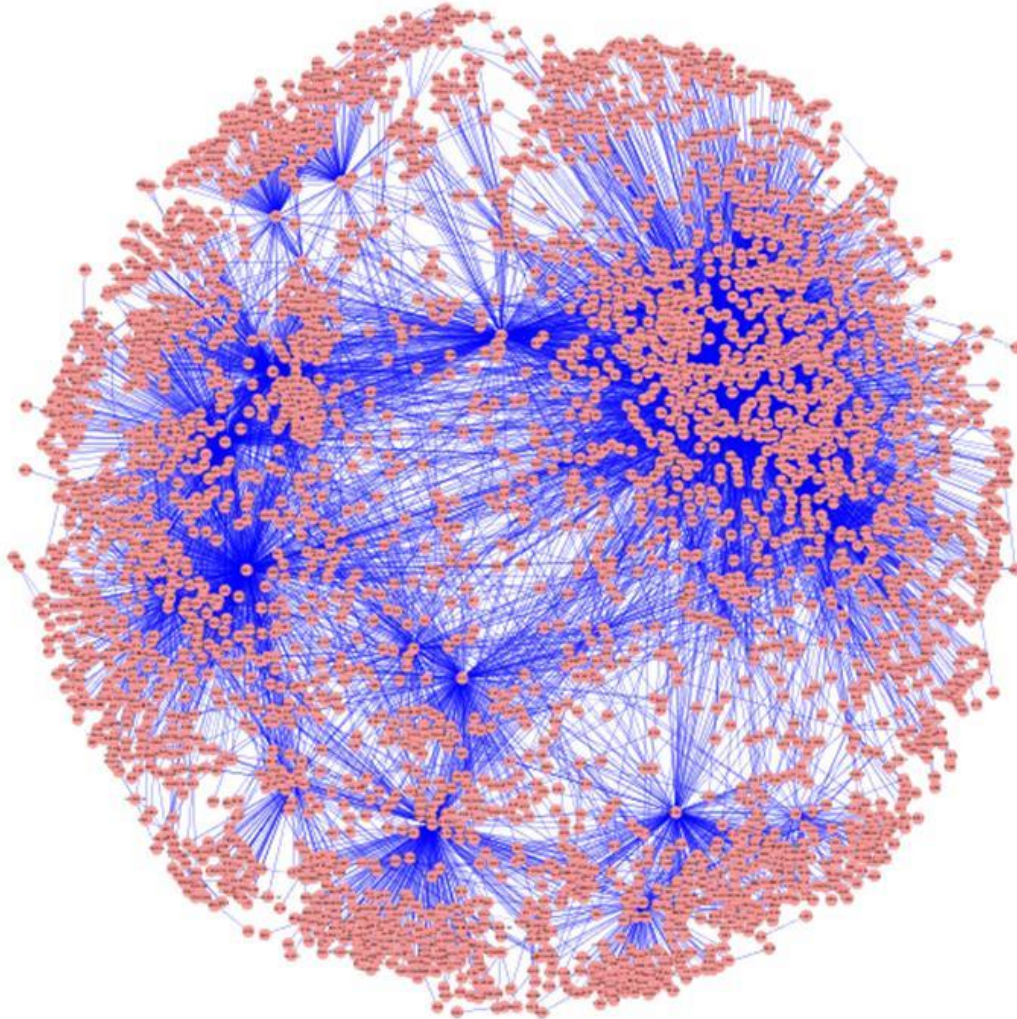


- P値とブートストラップ検定により推定した遺伝子調節ネットワークの信頼性を評価
- 不等式により間接的な相互作用も評価:

$$I(g_1, g_3) \leq \min [I(g_1, g_2); I(g_2, g_3)]$$

Cancer EMT 遺伝子関連ネットワーク

ARACNe による cancer EMT gene-gene influence network (17,368 regulations) ($p < 1.0 \times 10^{-10}$; bootstrap >90%) の推定.



ARACNe; $p < 1.0 \times 10^{-10}$;
bootstrap >90%

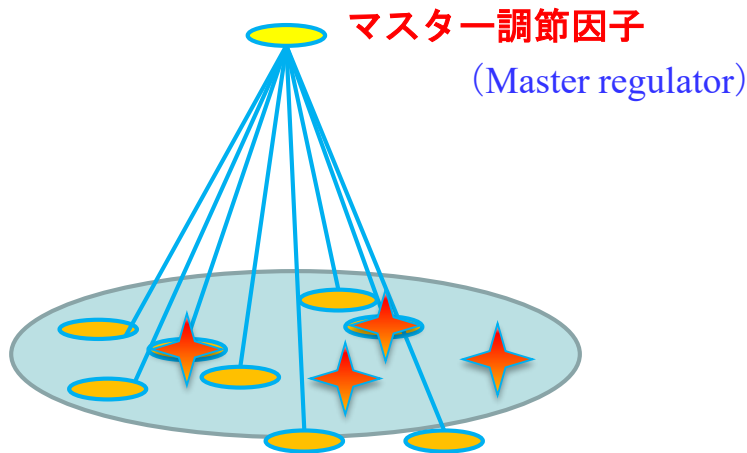
がんのEMTのマスター調節因子分析(11因子)

マスター調節因子

差別的発現遺伝子を平均的期待数より有意に多く制御している転写因子 (Fisher exact test $p < 0.05$).

がん EMT GRN

ARACNe algorithmによって推測される



- マスター調節因子候補
- 差別的発現遺伝子 (DEGs) (+ マスター遺伝子)
- 転写因子の制御される遺伝子

EMT 関連遺伝子のマスター調節因子

| Transcription factors | # of DEG (SAM; $p < 0.0001$, $FC > 1.5$) | P-value (Fisher exact test) |
|-----------------------|--|-----------------------------|
| TCF3 | 86/135 | 2.20×10^{-16} |
| ZEB1 | 46/592 | 8.58×10^{-16} |
| SMAD2 | 36/317 | 0.000209 |
| TWIST1 | 16/154 | 0.005634 |
| TP63 | 6/7 | 0.01302 |
| FOSL2 | 5/6 | 0.02516 |
| PPARA | 5/7 | 0.03732 |
| ARNTL2 | 15/38 | 0.03843 |
| MXD1 | 4/5 | 0.04918 |
| MITF | 4/5 | 0.04918 |
| NR2F2 | 4/5 | 0.04918 |

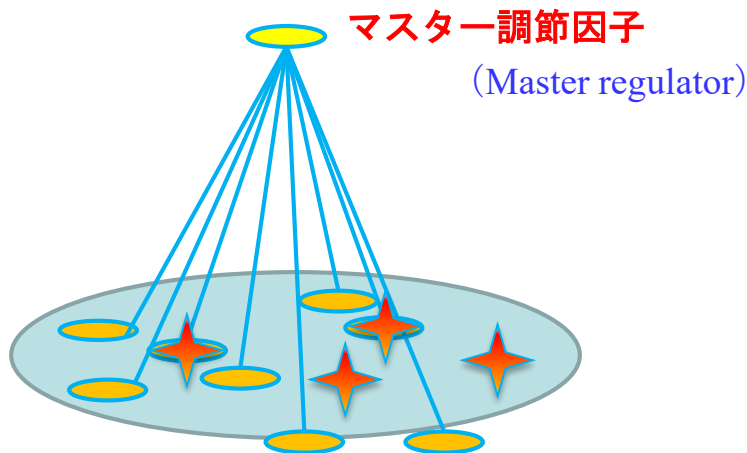
がんのEMTのマスター調節因子分析(11因子)




マスター調節因子

差別的発現遺伝子を平均的期待数より有意に多く制御している転写因子 (Fisher exact test $p < 0.05$).

がん EMT GRN

ARACNe algorithmによって推測される



-  マスター調節因子候補
-  差別的発現遺伝子 (DEGs) (+マーカー遺伝子)
-  転写因子の制御される遺伝子

| Transcription factors | # of DEG (SAM; $p < 0.0001$, $FC > 1.5$) | P-value (Fisher exact test) |
|-----------------------|--|-----------------------------|
| TCF3 | 86/135 | 2.20×10^{-16} |
| ZEB1 | 46/592 | 8.58×10^{-16} |
| SMAD2 | 36/317 | 0.000209 |
| TWIST1 | 16/154 | 0.005634 |
| TP63 | 6/7 | 0.01302 |
| FOSL2 | 5/6 | 0.02516 |
| PPARA | 5/7 | 0.03732 |
| ARNTL2 | 15/38 | 0.03843 |
| MXD1 | 4/5 | 0.04918 |
| MITF | 4/5 | 0.04918 |
| NR2F2 | 4/5 | 0.04918 |

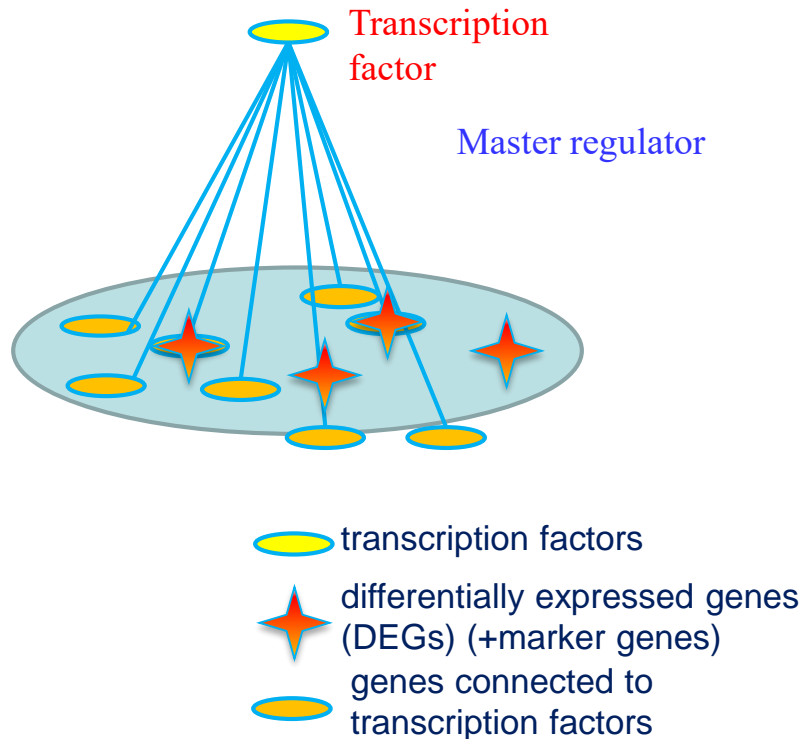
EMT 関連遺伝子のマスター調節因子

A : 11遺伝子 (5%有意) B : 4遺伝子(1%有意)

がんのEMTにおけるマスター調節因子分析（4因子）

Inference of master regulators (MRs) which regulate more differentially expressed genes than expected (Fisher exact test).

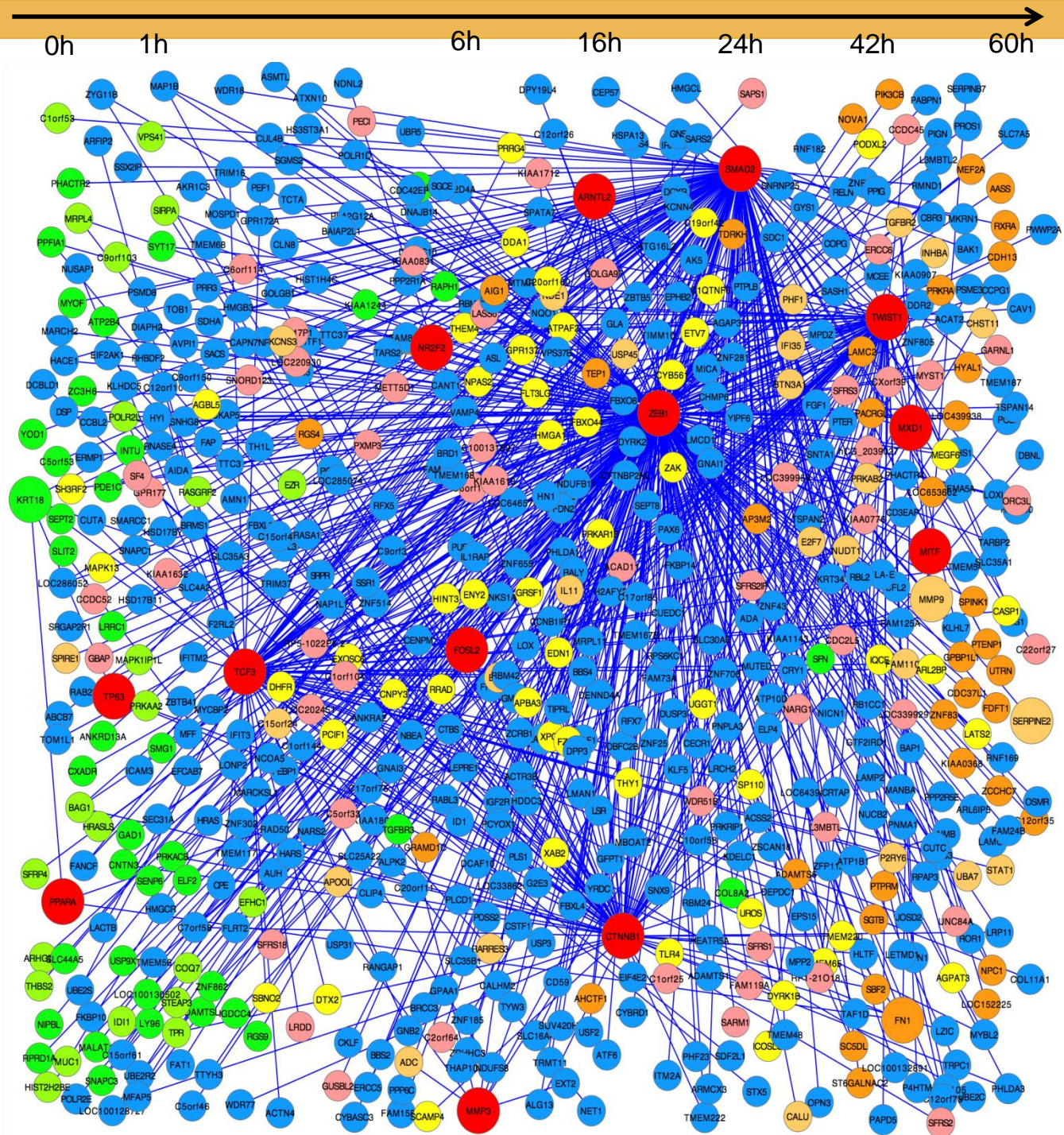
Cancer EMT gene GIN
inferred by ARACNe algorithm



| transcription factors | # of DEG (SAM; p<0.0001, FC>1.5) | P-value (Fisher exact test) |
|-----------------------|----------------------------------|-----------------------------|
| TCF3 | 86/135 | 2.20×10^{-16} |
| ZEB1 | 46/592 | 8.58×10^{-16} |
| SMAD2 | 36/317 | 0.000209 |
| TWIST1 | 16/154 | 0.005634 |
| TP63 | 6/7 | 0.01302 |
| FOSL2 | 5/6 | 0.02516 |
| PPARA | 5/7 | 0.03732 |
| ARNTL2 | 15/38 | 0.03843 |
| MXD1 | 4/5 | 0.04918 |
| MITF | 4/5 | 0.04918 |
| NR2F2 | 4/5 | 0.04918 |

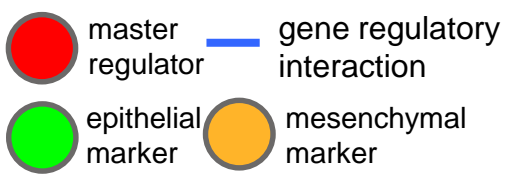
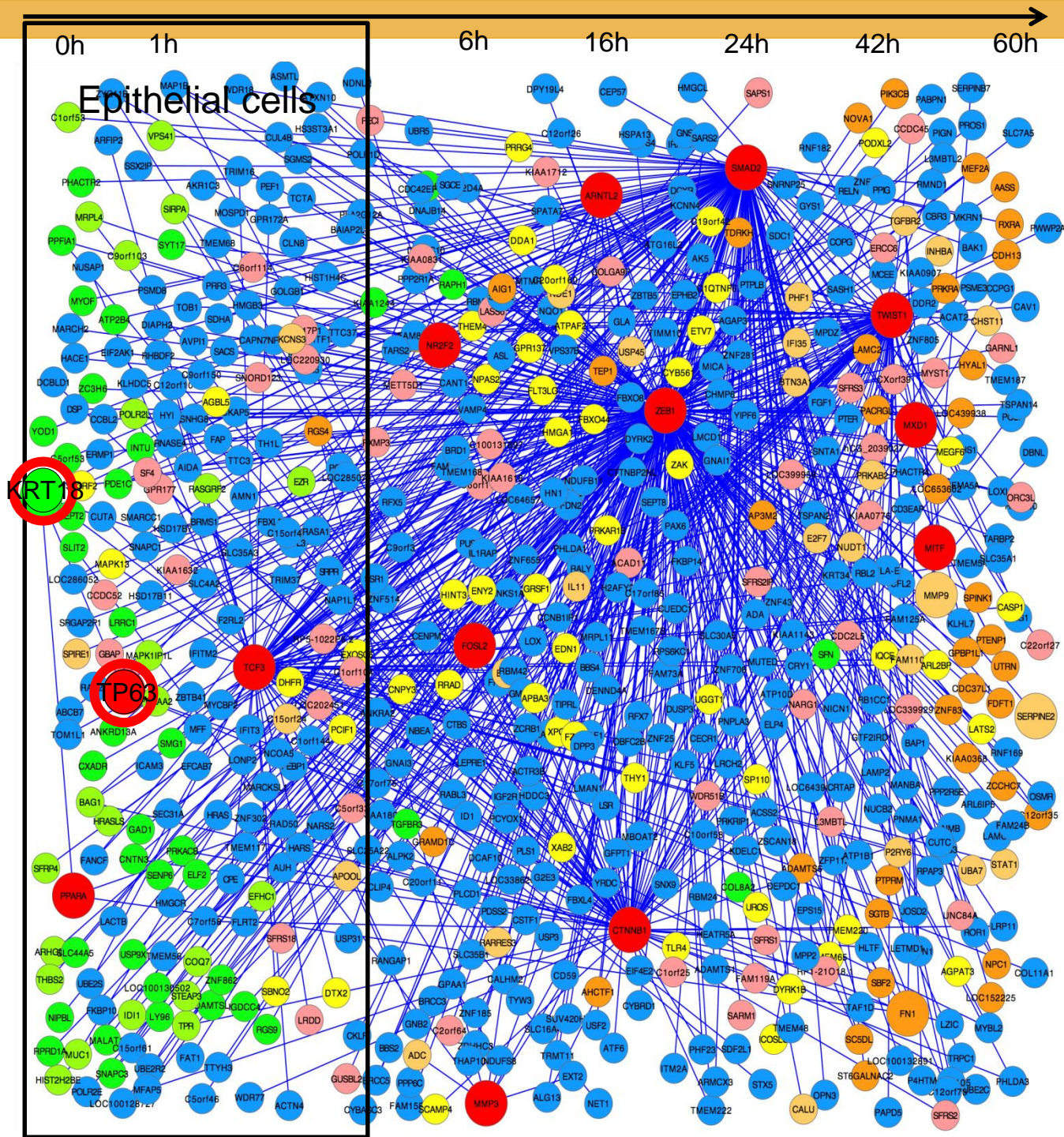
がん EMT network

Gene regulatory networks directly regulated by 11 master regulators.



がん EMT network 初期状態

KRT18 and TP63 expressed in the epithelial cells as reported before. A master regulator TP63 was predicted to regulate KRT18.

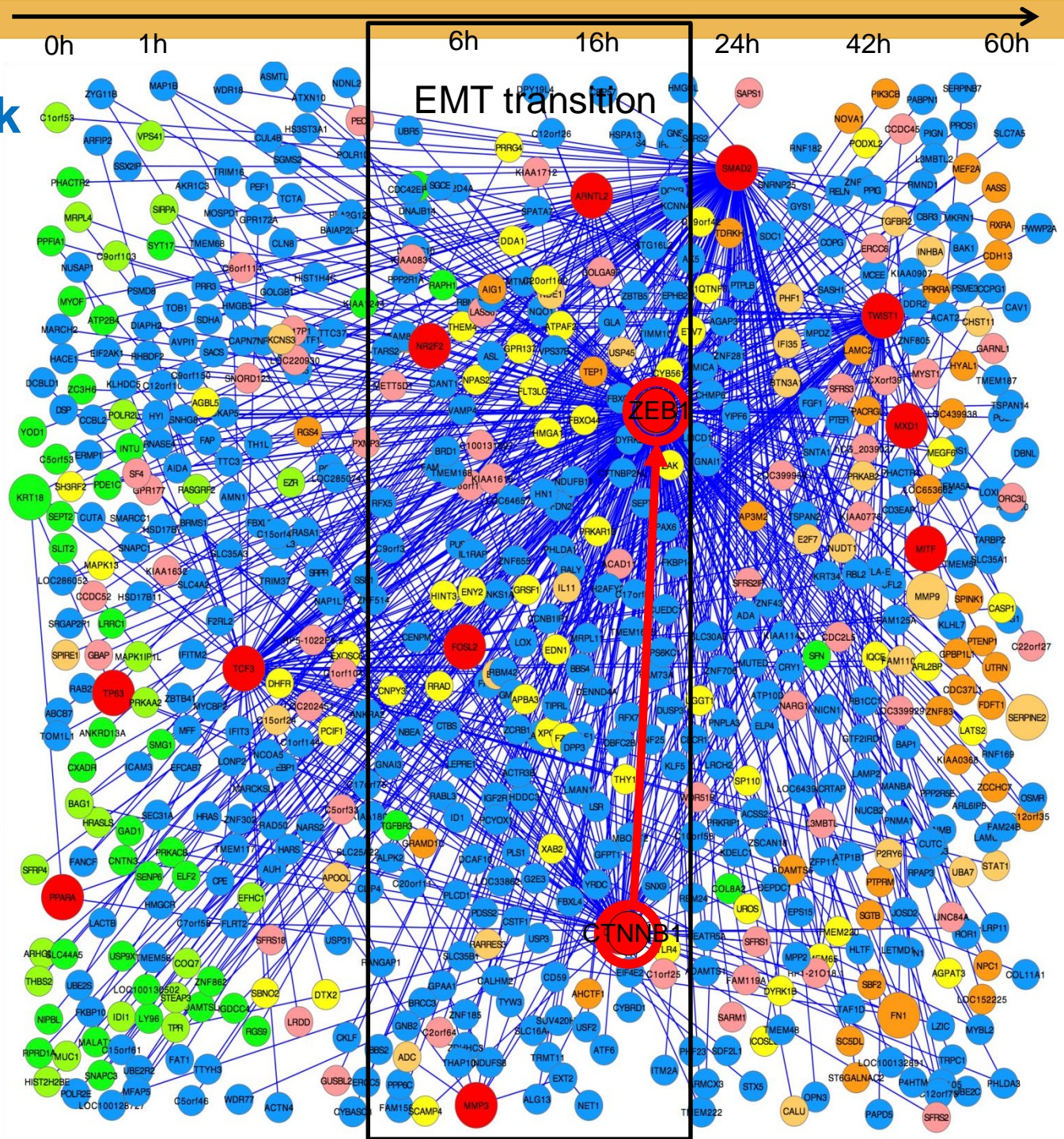


がん EMT network 中期状態

A master regulator CTNNB1 was predicted to induce the expression of ZEB1 as reported before. The CTNNB1 was known to be key factor inducing the EMT process.

ZEB1 was also a master regulator which expressed from the beginning of EMT process, and was known to down-regulate CDH1 gene expressions.

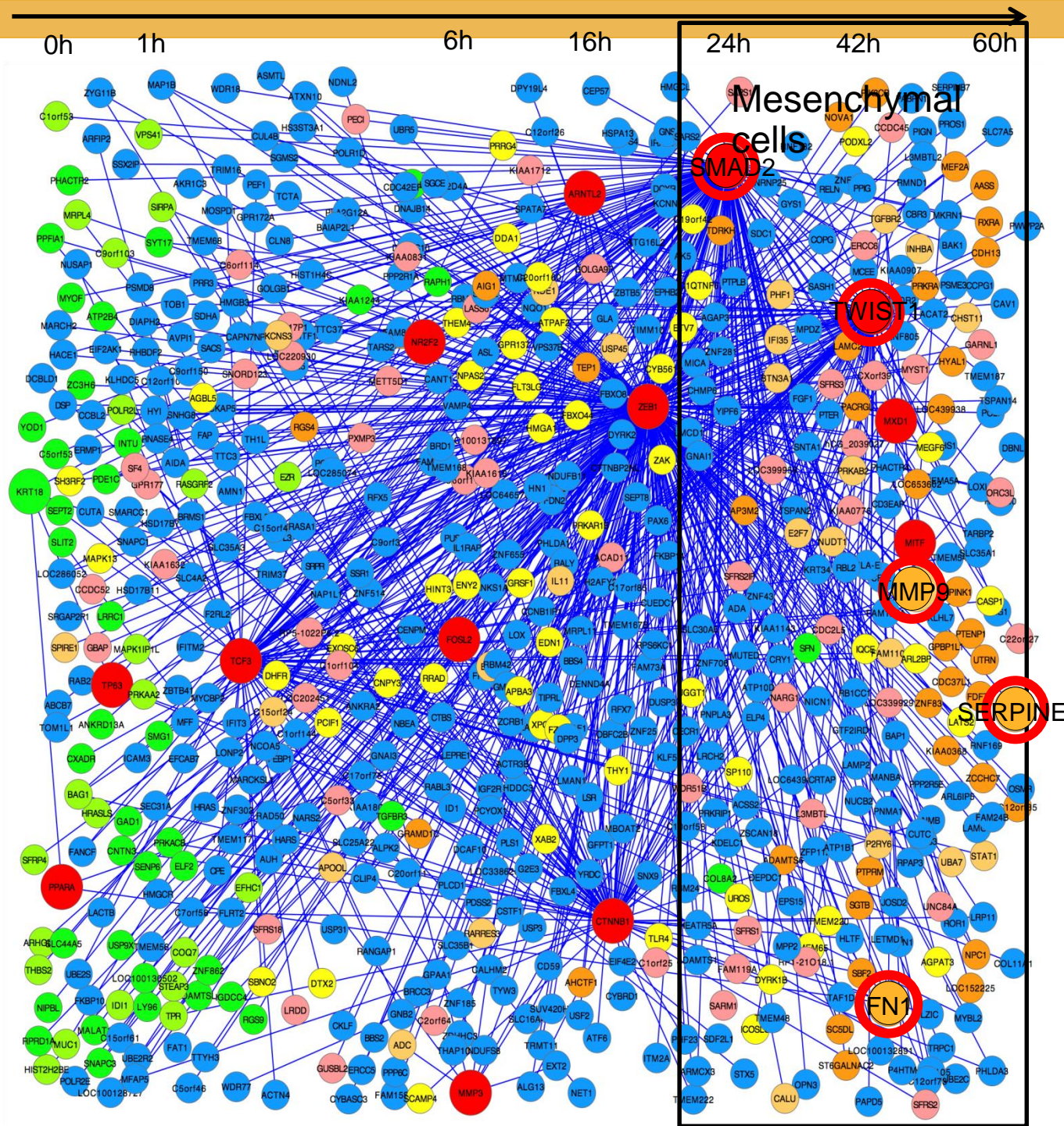
- master regulator
- epithelial marker
- gene regulatory interaction
- mesenchymal marker



がん EMT network 後期状態

SMAD2 was a master regulator, and was known to induce the mesenchymal marker genes; e.g., MMP9 and FN1, and SERPINE2 genes.

TWIST1 was also a master regulator, and in fact, was reported to be essential in induction of the EMT process.



- master regulator
- epithelial marker
- mesenchymal marker
- gene regulatory interaction

タンパク質相互作用ネットワーク と疾患・創薬

Evaluation of drug targetness from protein-
protein interaction network

分子ネットワークの構造と疾病・創薬

Principle of Systems Molecular Medicine
*Disease is caused by
Distortion of the cellular molecular network*

*Drugs which cure diseases would be
more efficiently discovered by
Systems approach*

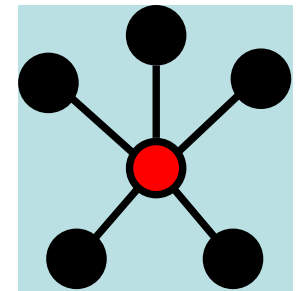
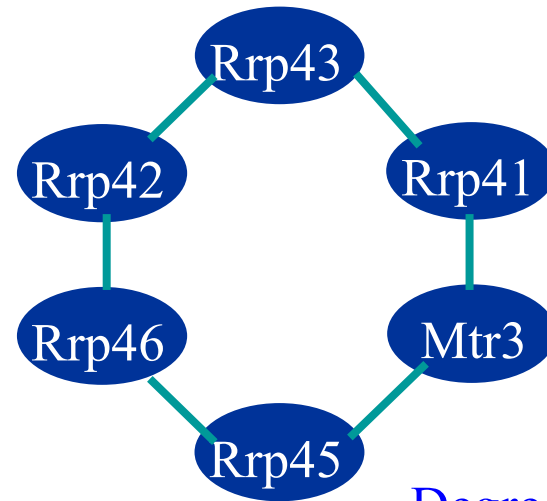
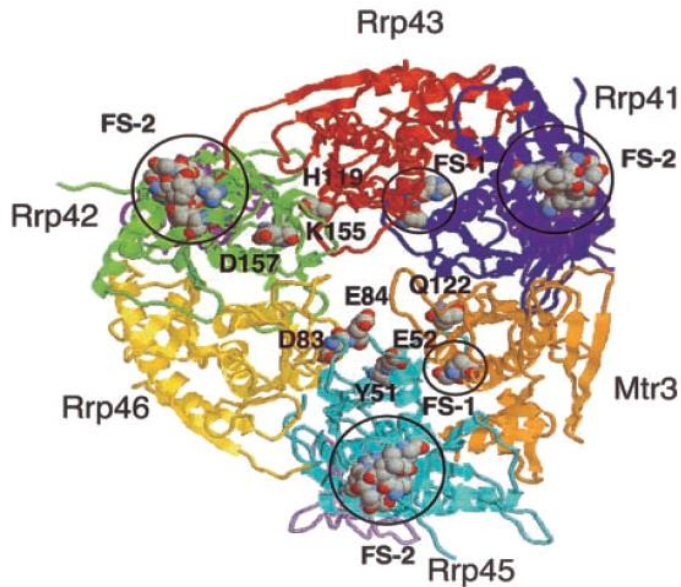
systems-oriented drug design

Taking protein-protein interaction network (PIN) as
typical example of cellular network
Systems properties of drug are investigated

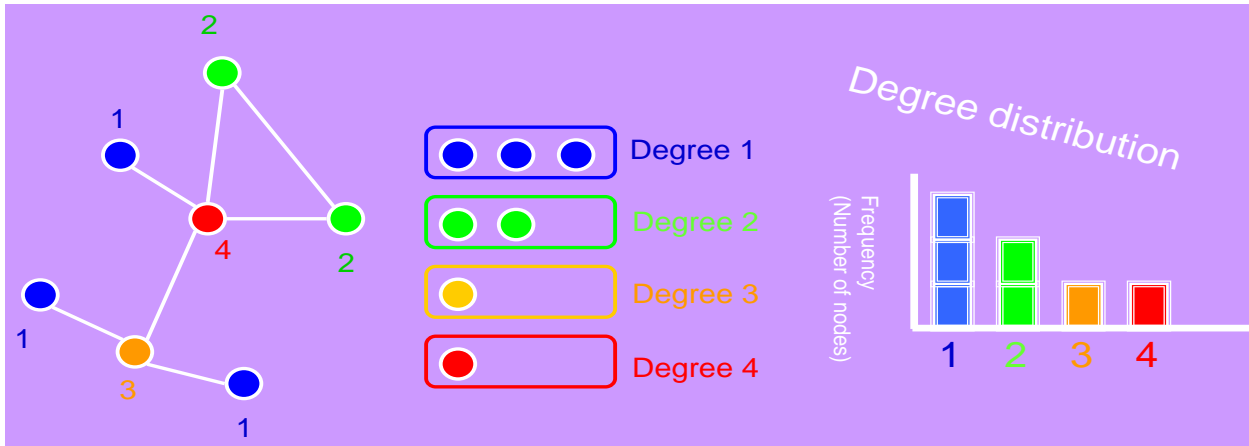
Hase, T., Tanaka, H., Suzuki, Y., Nakagawa, S., Kitano H (2009)
Structures of protein protein interaction network and their
implications on drug design. *PLoS Comput Biol* 5(10):e1000550.

タンパク質間相互作用ネットワーク

In protein-protein interaction networks (PINs), a node and a link represent a protein and an interaction between two proteins, respectively. The number of links connected to a node is called a degree.



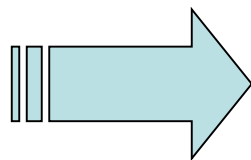
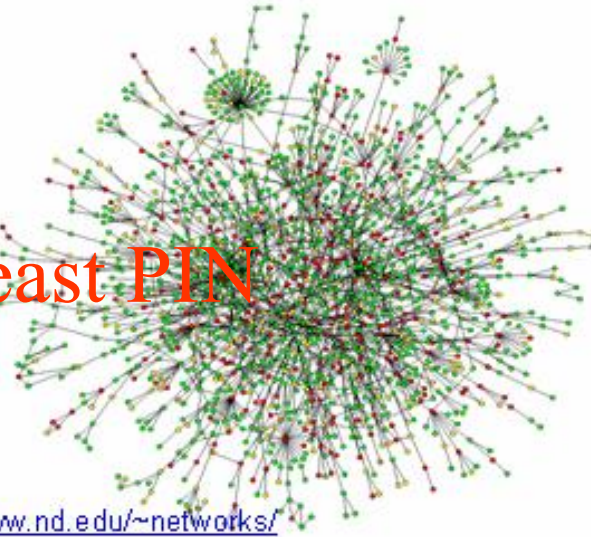
Degree of red node is 5.



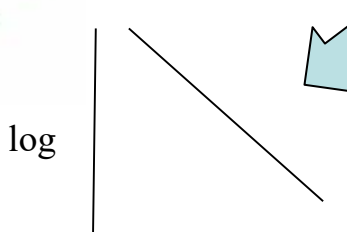
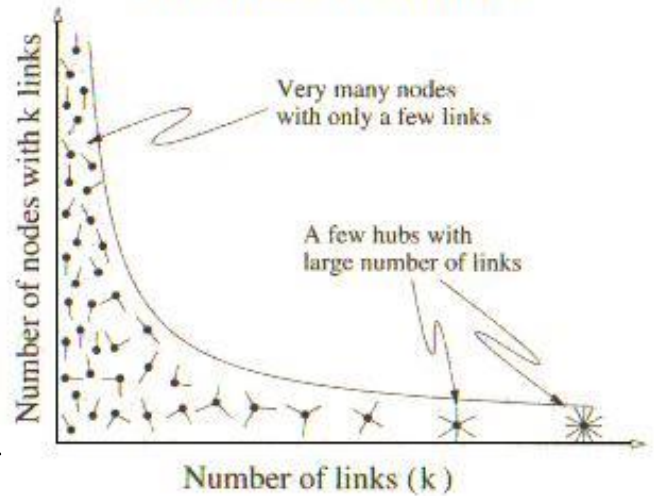
In PINs, there are a large number of low-degree nodes (branches) and a small number of high-degree nodes (hubs). Hub-branch (scale free) structure

Yeast PIN

<http://www.nd.edu/~networks/>



Power Law Distribution



Scale free distribution



HIPPIE » Human Integrated Protein-Protein Interaction rEference

[PROTEIN QUERY](#)[NETWORK QUERY](#)[SCREEN ANNOTATION](#)[DOWNLOAD](#)[INFORMATION & CONTACT](#)[CBDM GROUP](#)

HIPPIE Howto

Confidence scoring of experimentally measured interactions

A core component of HIPPIE is the confidence scoring of interactions based on the amount and reliability of evidence supporting each interaction. This score is calculated as a weighted sum of the number of studies in which an interaction was detected, the number and quality of experimental techniques used to measure an interaction and the number of non-human organisms in which an interaction was reproduced.

The parameters of this scoring scheme were jointly optimized by a group of experts and a computer algorithm: we first assigned quality scores to each experimental technique that measures protein-protein interactions. This experimental quality score is supposed to reflect the reliability and the error rate of the techniques. The list of experimental quality scores can be found in the [download](#) section. We are aware that the estimation of experimental quality is led by the individual perception and experience of our expert team to a high degree and can only reflect our subjective beliefs. If you feel that we misjudged a technique fundamentally, we would like to ask you to give us feedback. Also, there is the possibility for users of HIPPIE to modify the scores and to apply our rescoring tool to get an individually scored version of HIPPIE. Please, find a stand-alone version of the tool and instructions on how to run it in the [download](#) section.

Query interactions

The [query mode](#) allows for querying HIPPIE with a [UniProt](#) identifier (id or accession), [gene symbol](#) or [Entrez](#) gene id. All literature interactions in HIPPIE that match the query dprotein are displayed.

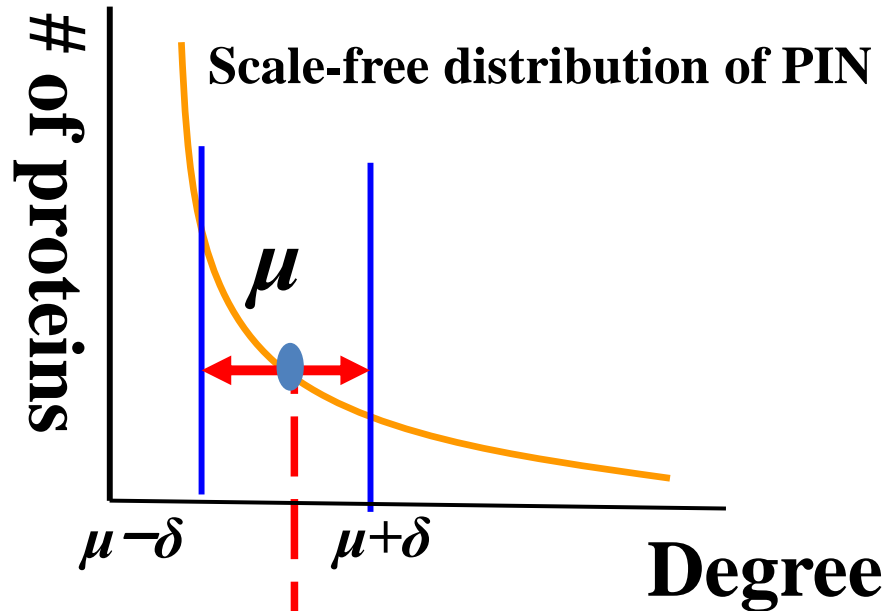
On the result page UniProt id, Entrez gene id and gene symbol of the interactors of the query protein are indicated. The confidence score of the interaction is given in the last column. Clicking on it opens an evidence page where the source databases of the interaction are listed, the publications which mention the interaction, the experimental systems used to detect the interaction and in which species the interaction is conserved.

We use a high confidence protein-interaction dataset that several integrate representative interaction databases (e.g., BioGrid, IntAct, MINT, etc).

The high-quality protein-interaction network is composed of 28,889 interactions among 7,890 proteins.

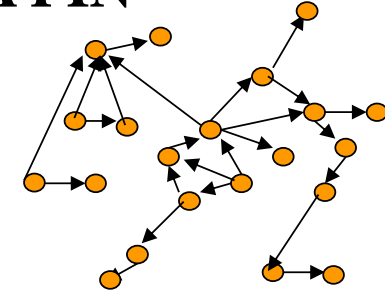


Moving stratification by connection degree (MSD)



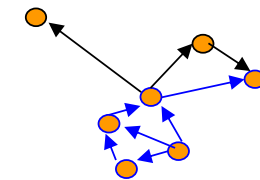
MSD was used to extract sub-networks consisting of nodes with degrees from $\mu - \delta$ to $\mu + \delta$.

A PIN

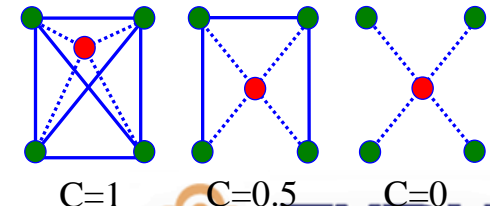


MSD
From 3 to 5

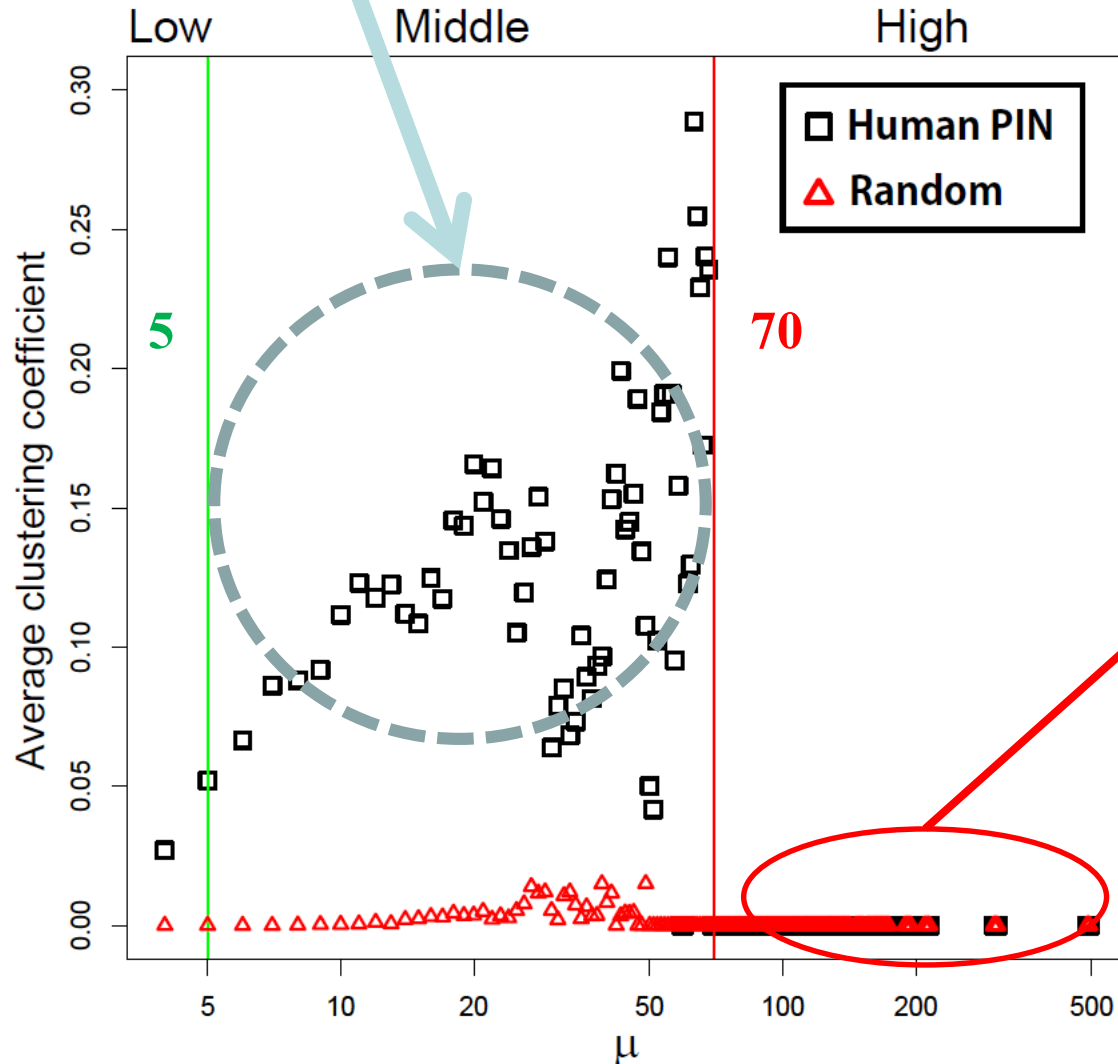
A subnetwork in the PIN



Cluster coefficient

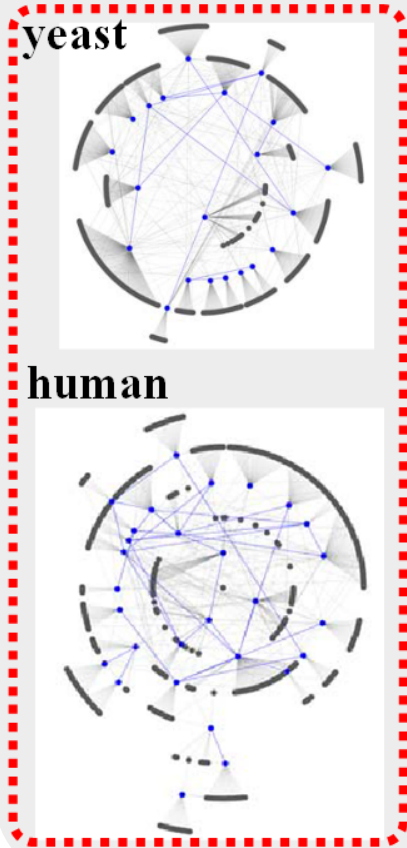


There are extensive interconnections among middle degree that form core backbone of the network.



There are a small number of connections among high-degree nodes. A high-degree nodes connects to a large number of low-degree nodes to form a hub-centric functional modules.

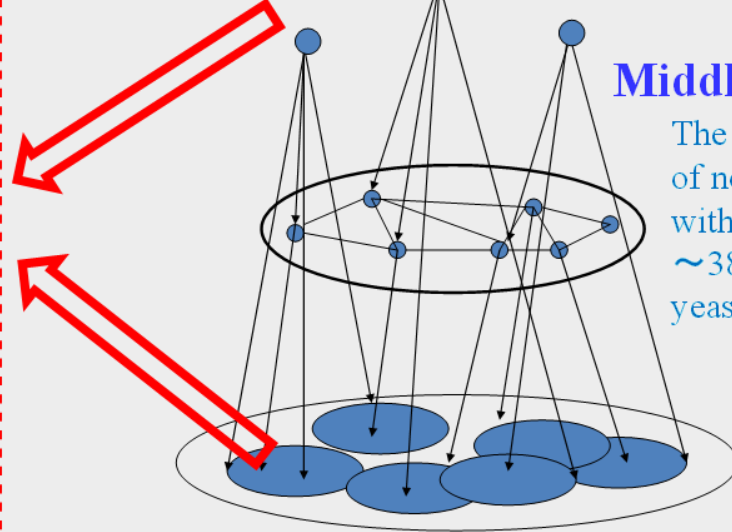
Three-Layer Structure in the yeast and human PINs



High-degree layer

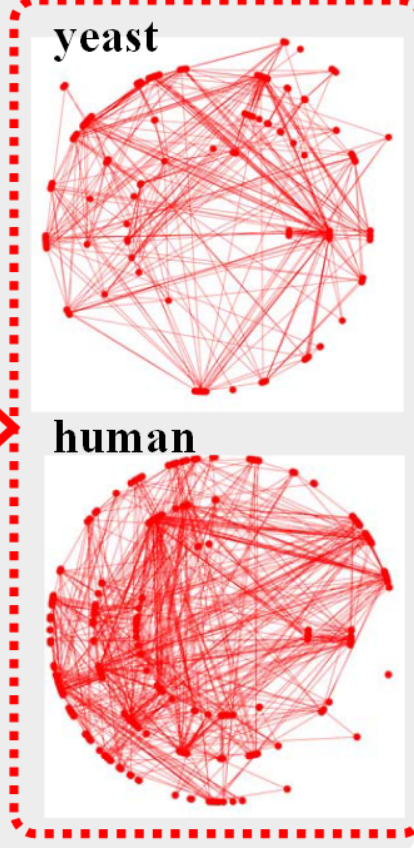
The layer composed of nodes with degree greater than 31 and 39 for human and yeast, respectively.

Same as former report*



Middle-degree layer

The layer composed of nodes with degree with 6 ~ 30 and 6 ~ 38 for human and yeast, respectively



Low-degree layer

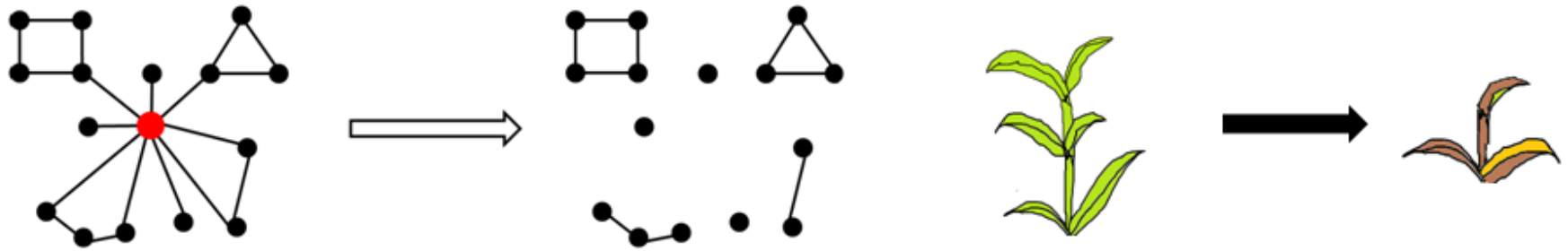
The layer composed of nodes with degree lower than 5.

The layer composed of nodes with degree lower than 2
Low-degree layer

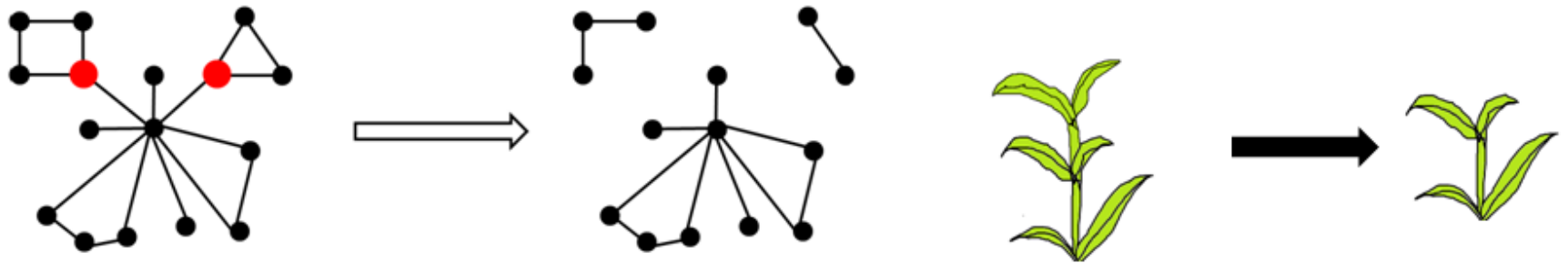


*Hase, T., Tanaka, H. et.al (2009) Structures of protein protein interaction network and their implications on drug design. PLoS Compt Biol. 5(10)

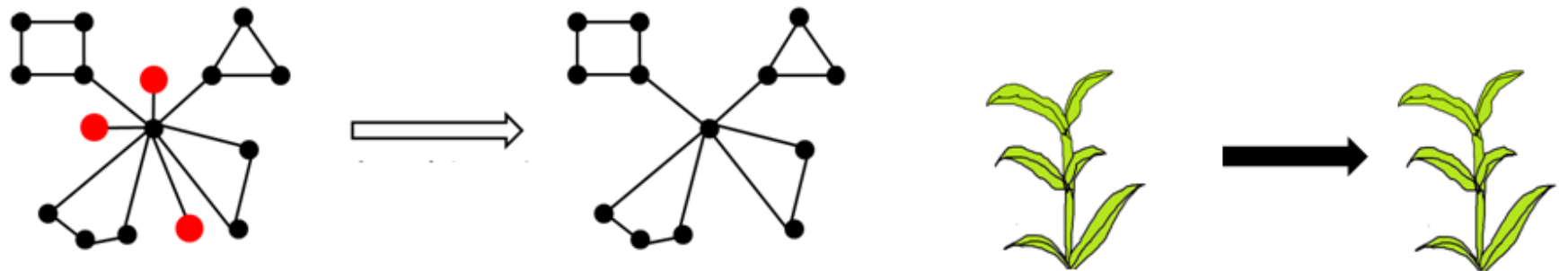
High-degree genes tend to be essential.



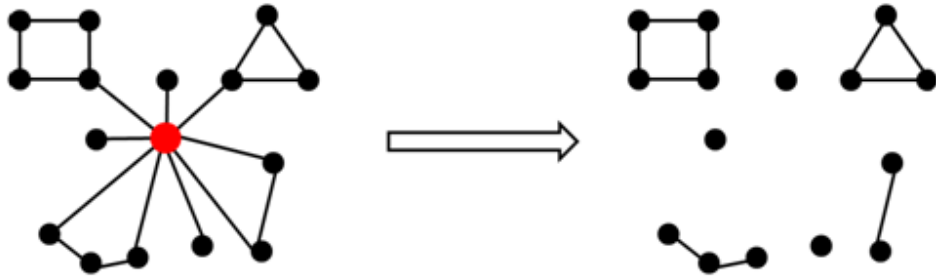
Middle-degree genes may affect phenotype.



Genes with less severe impacts are tend to be lower degree regions.

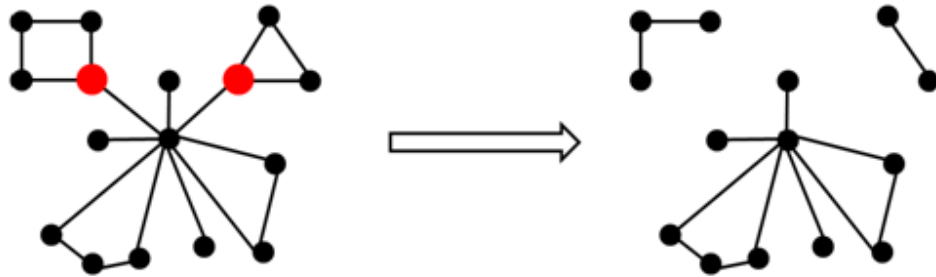


High-degree genes tend to be essential.



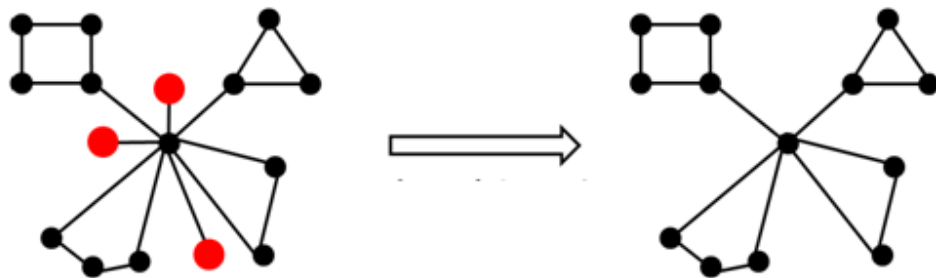
Disease genes?

Middle-degree genes may affect phenotype.



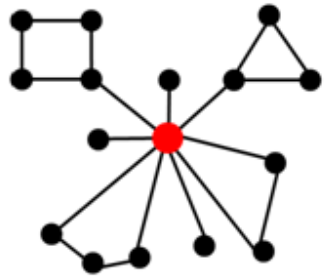
Disease genes?

Genes with less severe impacts are tend to be lower degree regions.



Disease genes?

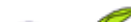
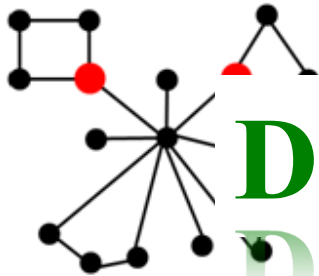
High-degree genes tend to be essential.



Not Disease genes!



Middle-degree genes affect phenotype.



Disease genes!

Goh et al. 2007 PNAS
Friedman et al 2008 PNAS



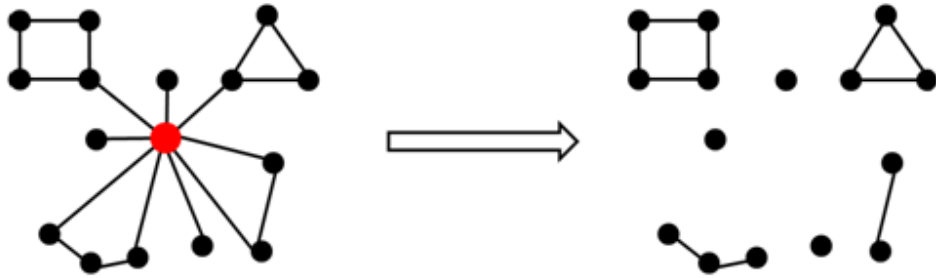
Genes with less severe impacts are tend to be lower degree regions.



Disease genes!



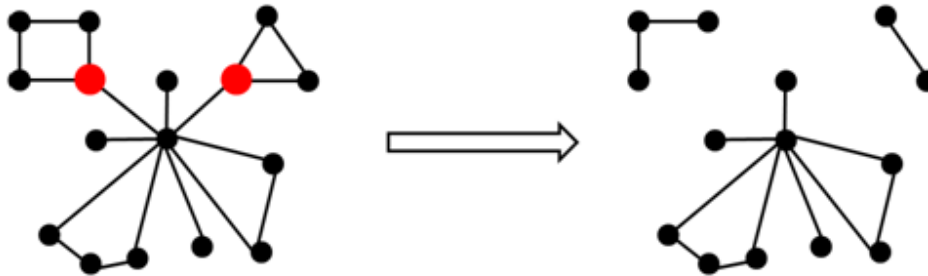
High-degree genes tend to be essential.



Drug targets?



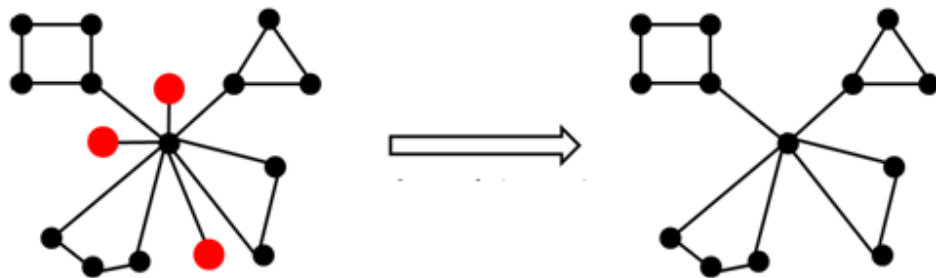
Middle-degree genes may affect phenotype.



Drug targets?



Genes with less severe impacts are tend to be lower degree regions.



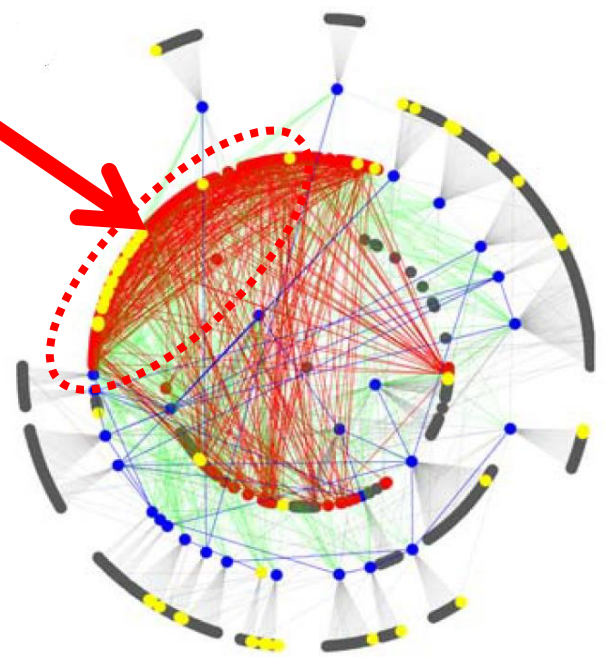
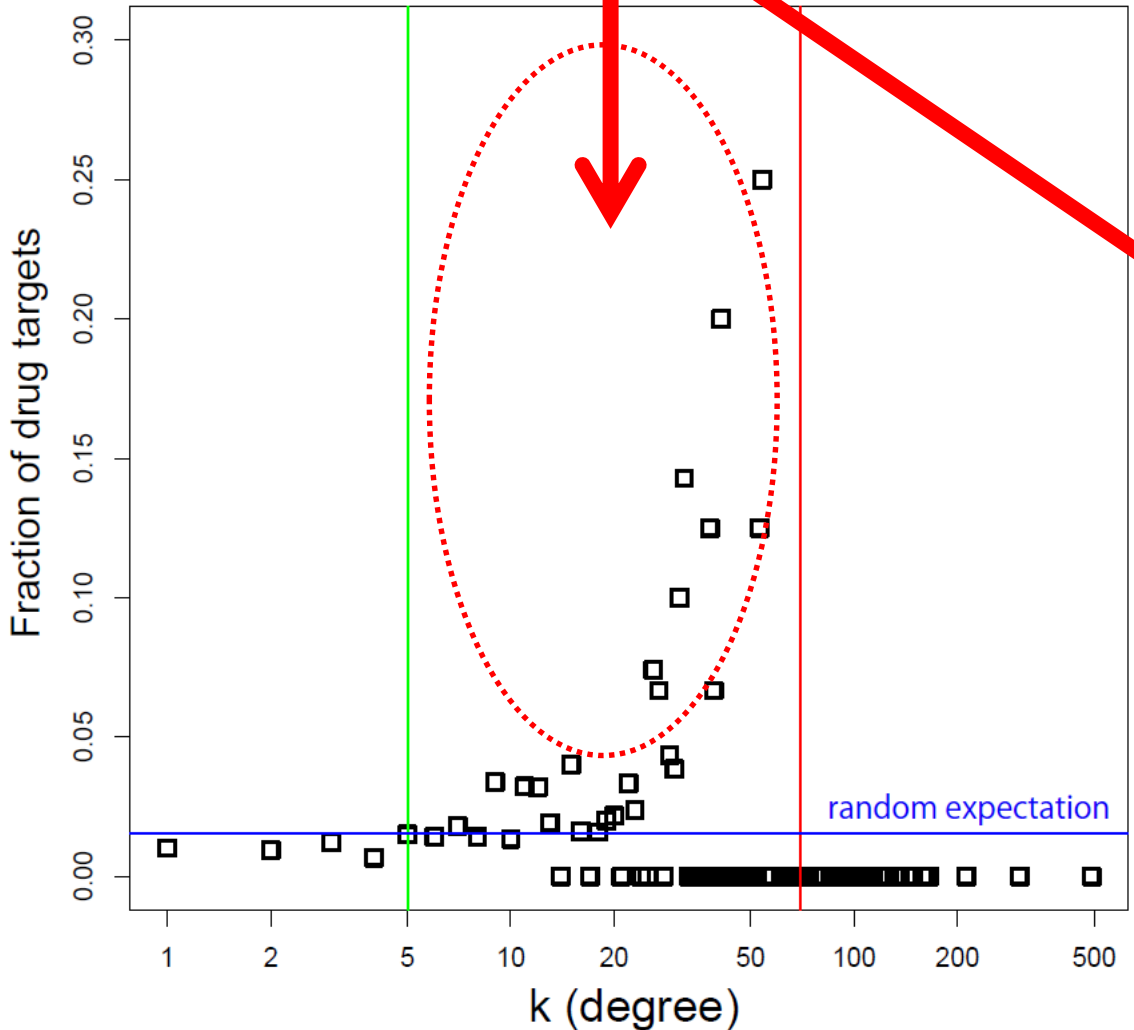
Drug targets?

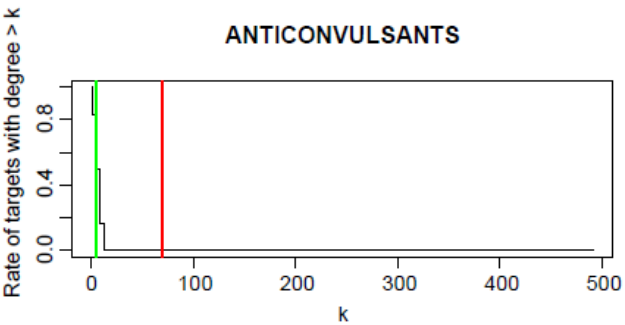
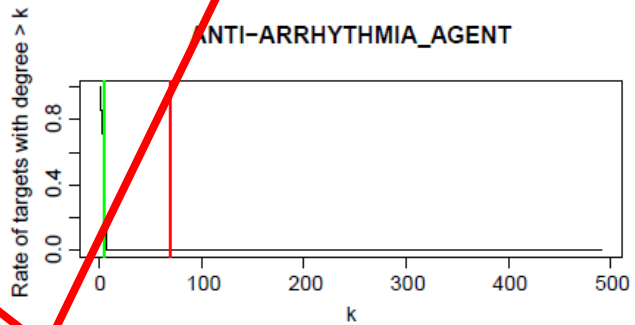
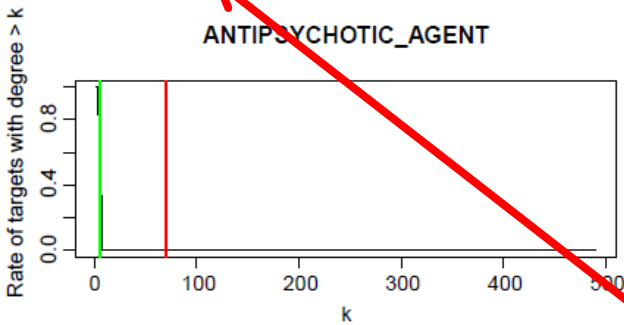
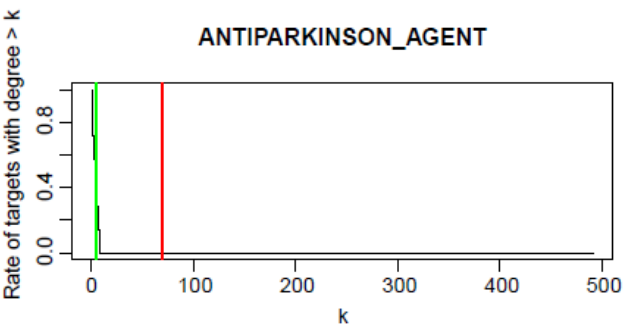
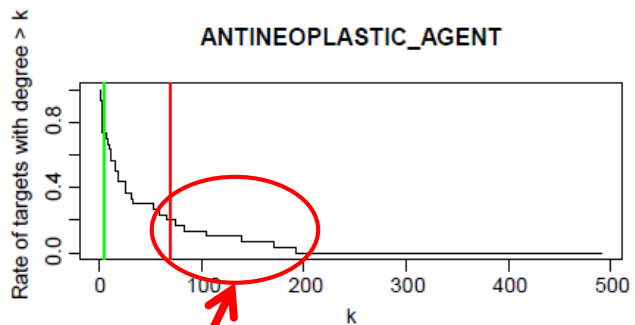
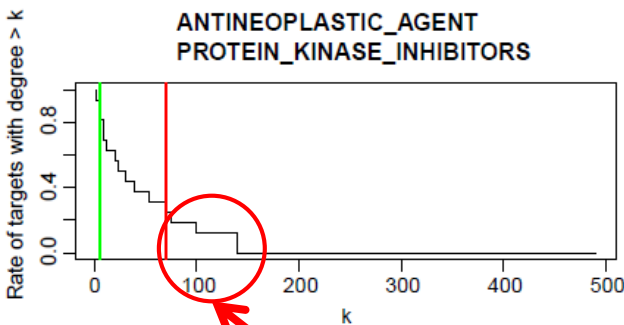


中間度数タンパク質

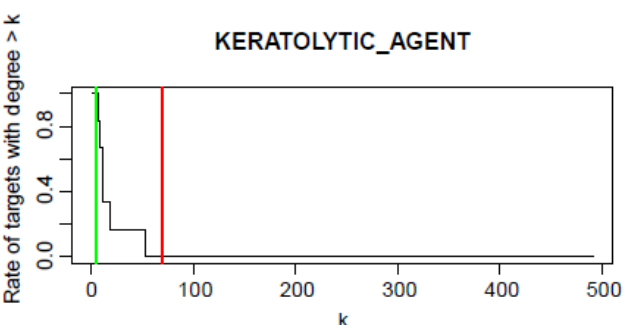
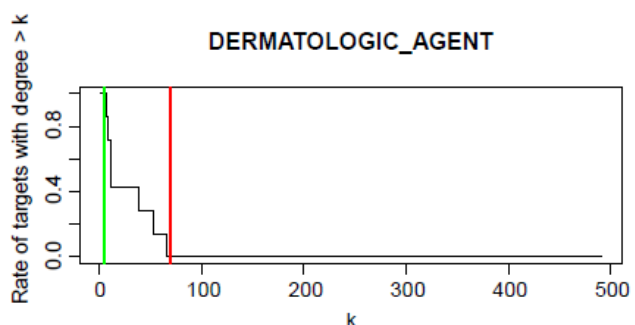
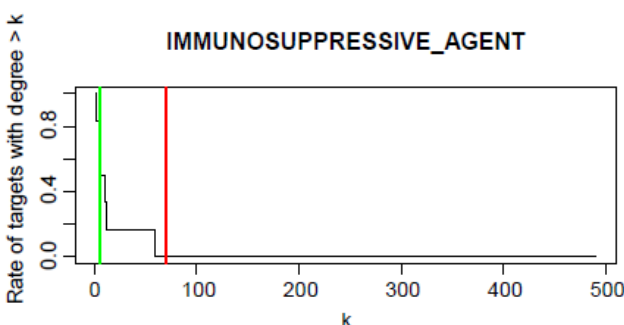
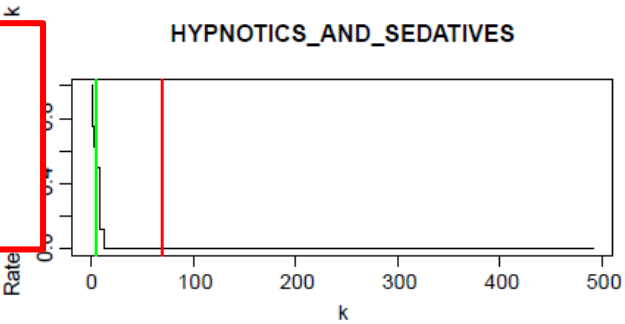
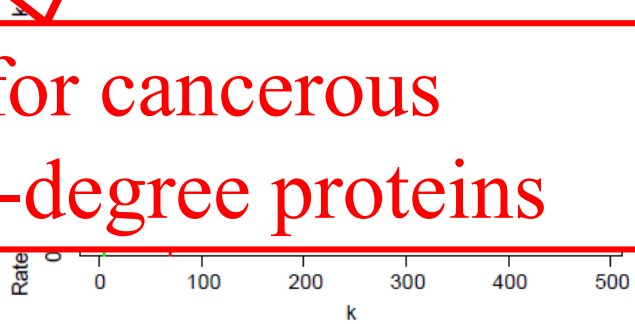
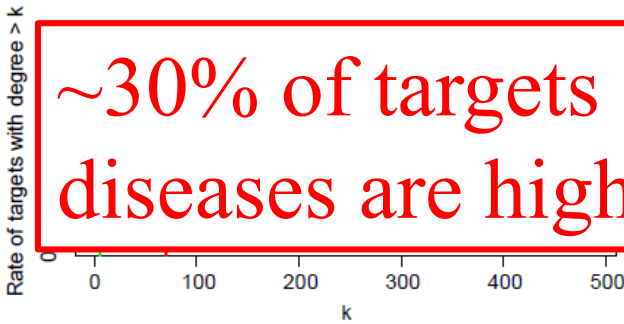
the optimal target for therapeutic drugs.

Therefore, many targets for drugs that are currently on the market are concentrated on backbone of the human PIN.

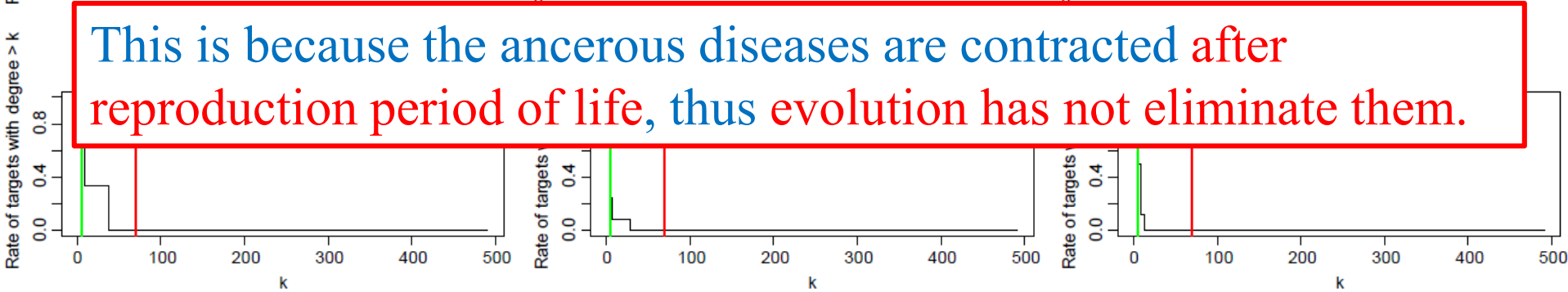
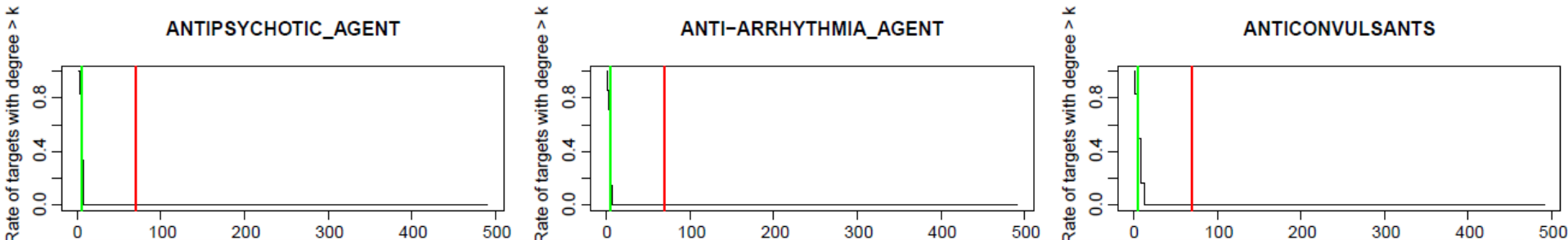
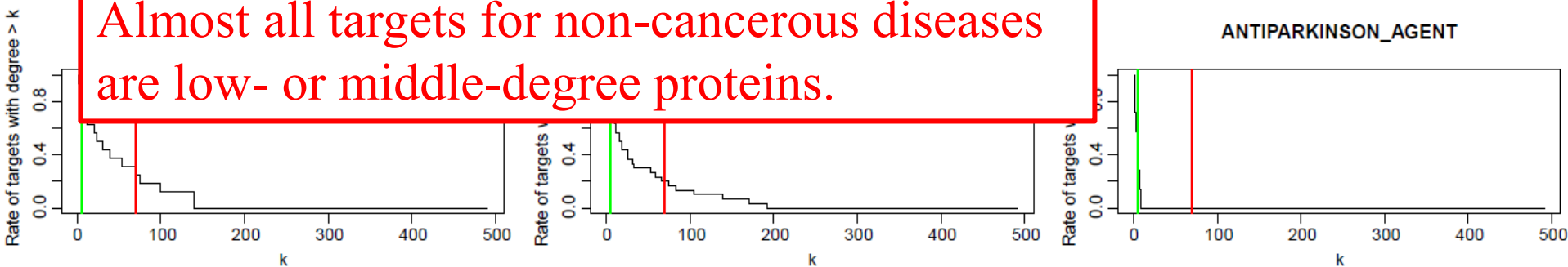




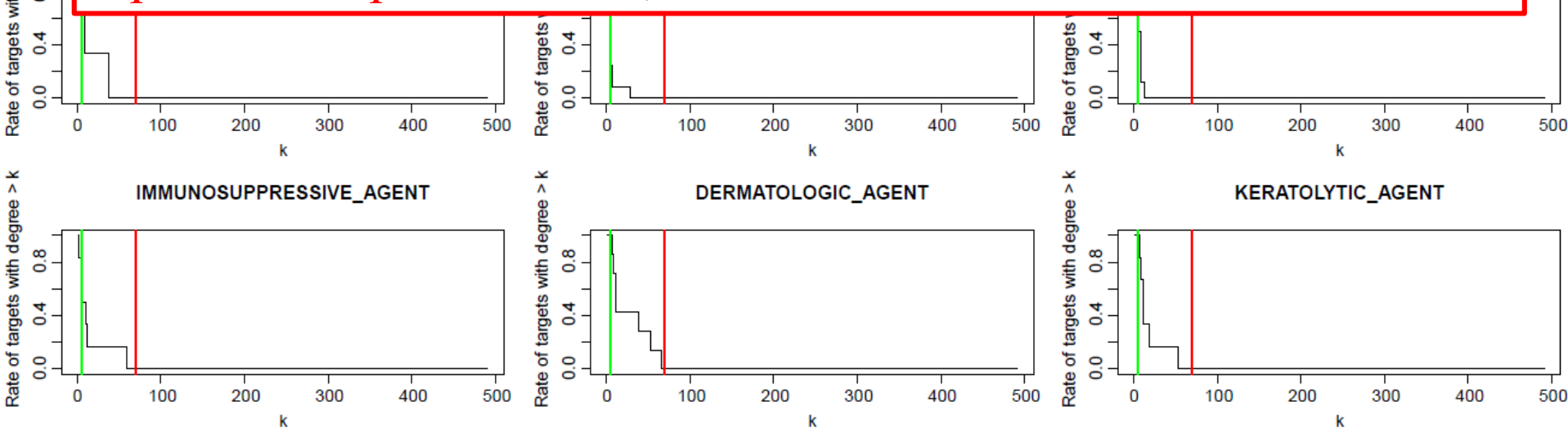
~30% of targets for cancerous diseases are high-degree proteins



Almost all targets for non-cancerous diseases are low- or middle-degree proteins.



This is because the cancerous diseases are contracted after reproduction period of life, thus evolution has not eliminate them.



まとめ

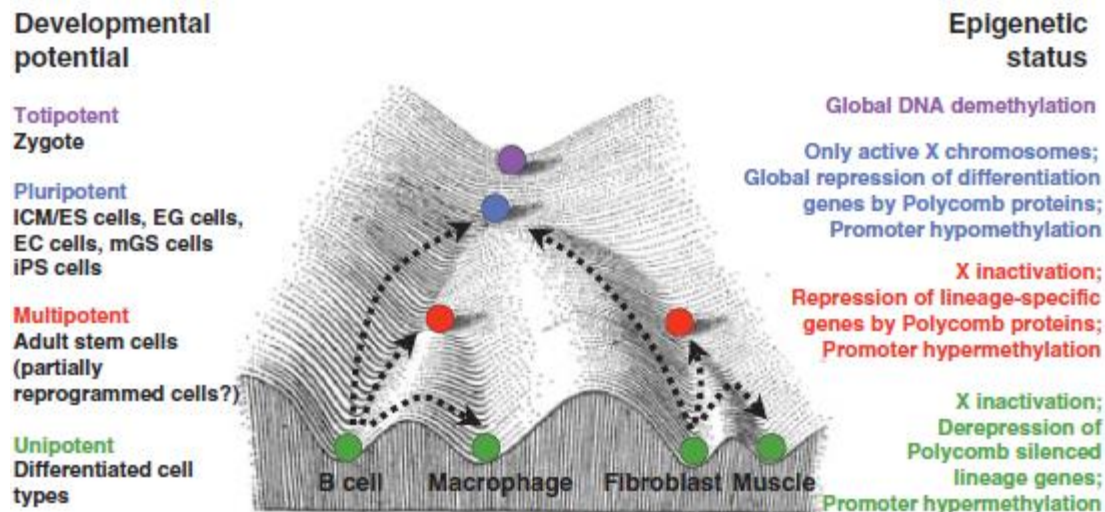
- ゲノム/オミックス医療は、疾患を細胞分子ネットワークの歪みと捉える「システム分子医学」の枠組みによって十全に把握され、個別化医療、先制医療を可能にする基盤となる。
- 細胞分子ネットワークの状態空間と定量的 Waddington エピゲノム地形（細胞状態確率ポテンシャル）は、発生・分化・疾患進行を統一的に捉える基礎理論となる。
- タンパク質間相互作用ネットワークは3層構造をなし、疾患・薬剤標的となるタンパク質は中間層に存在する。がんは高度数層が標的であるがこれは生殖期以降に罹患するので淘汰されないためである。

Acknowledge

- Hideyuki Saya, *Keio University*, EMT
- Soichi Ogishima, *TMDU*, EMT network identification
- Takeshi Hase, *TMDU*, PIN and Drug
- Kaoru Mogushi, *TMDU*, gene expression analysis
- Shigeki Arii, *TMDU*, hepatocellular carcinoma
- Shinji Tanaka, *TMDU*, hepatocellular carcinoma

ダイレクト・リプログラミング

- 皮膚などの採取しやすい細胞から、神経細胞や心筋細胞に分化する幹細胞をiPS細胞を介さず、必要な幹細胞を直接（ダイレクトに）作成する。
- 繊維芽細胞→MyoD→骨格筋細胞
繊維芽細胞→iHEP細胞→肝細胞
心筋繊維芽細胞→iCM細胞→心筋細胞など
- iPS細胞に戻すよりも成功率高く時間もかからない(0.01-0.1%に対して2-4%程度)

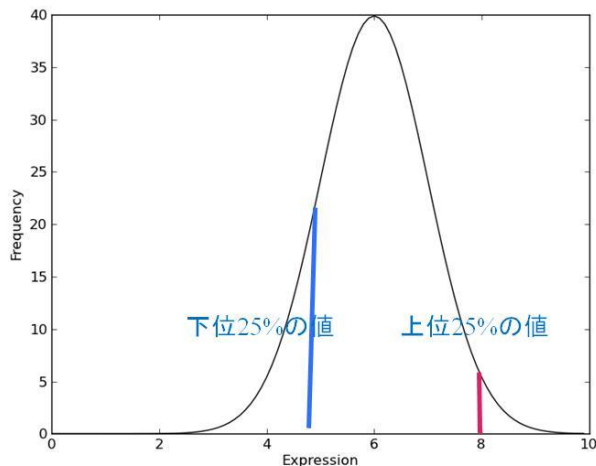


遺伝子発現プロファイル

Affymetrix GeneChip Human Genome U133 Plus 2.0 Array

- 成熟組織の網羅的データセット (疾患状態含む)
 - GSE7307
- fibroblast/keratinocytesからiPSへのリプログラミングデータ
 - GSE12583, 18226, 23583, 9832, 16694, 9865, 37258
- ES, iPS細胞のデータ
 - GSE18618
- ES, iPSからの分化誘導データ
 - GSE9169(ニューロン), 14897(肝細胞), 9440, 29115(肝)
- 4, 5, 6, 7, 8, 9週目の胚
 - GSE15744
- 1, 2, 4, 8分割受精卵、桑実胚、胚盤胞(ヒトのみ)
 - GSE18290

合計 956サンプル



$$\begin{aligned} IQR &= \log_2 \frac{\text{上位 25 \% の値}}{\text{下位 25 \% の値}} \\ &= \log_2 \frac{8}{4} \\ &= 1 \end{aligned}$$

今回は $IQR > 4$ の
プルーブ・遺伝子を抽出。

プルー部数・遺伝子数:
839プルーブ 603遺伝子

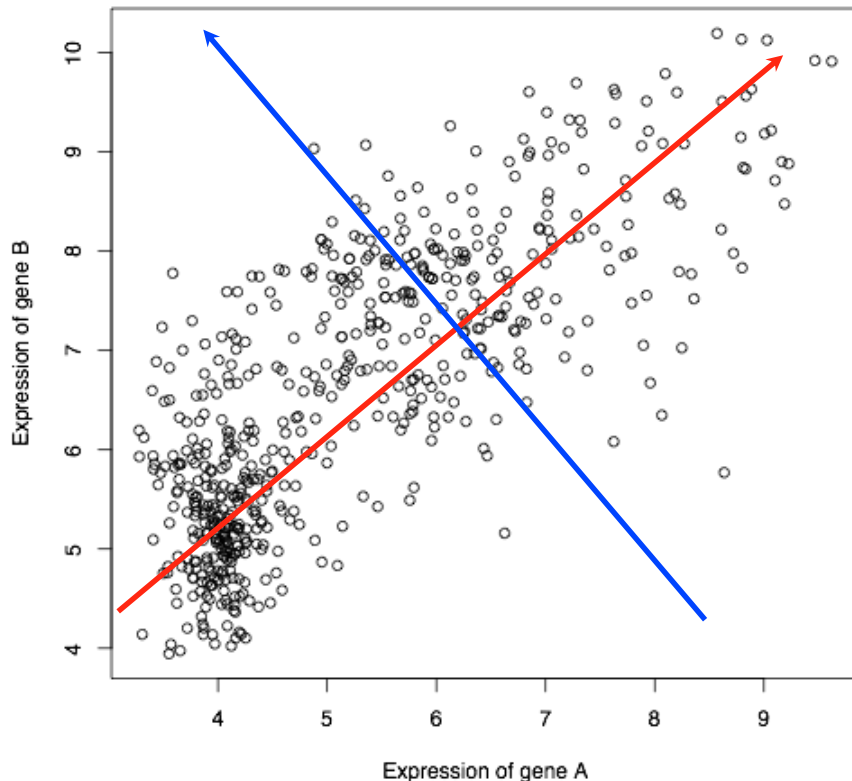
GO Biological Processes

| Rank | GO Biological Processes | Count | % | P-value | Benjamini |
|------|--------------------------------------|-------|------|-----------------------|-----------------------|
| 1 | Anatomical structure development | 179 | 30.4 | 4.6×10^{-25} | 1.3×10^{-21} |
| 2 | Nervous system development | 107 | 18.2 | 1.7×10^{-24} | 2.4×10^{-21} |
| 3 | System development | 166 | 28.2 | 3.7×10^{-23} | 3.4×10^{-20} |
| 4 | Developmental process | 195 | 33.2 | 1.4×10^{-20} | 9.9×10^{-18} |
| 5 | Multicellular organismal development | 177 | 30.1 | 4.0×10^{-18} | 2.2×10^{-15} |
| 6 | Multicellular organismal process | 228 | 38.8 | 3.0×10^{-16} | 1.5×10^{-13} |
| 7 | Neurogenesis | 63 | 10.7 | 2.1×10^{-15} | 8.3×10^{-13} |
| 8 | Cell adhesion | 63 | 10.7 | 2.1×10^{-12} | 7.1×10^{-10} |
| 9 | Biological adhesion | 63 | 10.7 | 2.2×10^{-12} | 6.8×10^{-10} |
| 10 | Generation of neurons | 55 | 9.4 | 2.2×10^{-12} | 6.1×10^{-10} |

 603遺伝子は分化・発達の指標

主成分分析 (PCA)による次元圧縮

遺伝子A, Bの発現量の分布が次のようなとき



もし、gene A, Bが神経分化に関与する遺伝子なら、**第一主成分**は神経分化の程度を表す指標となる。

ポテンシャルの2次元表示で特徴を概ね表現できる。
(データの低次元化)

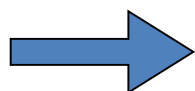
軸の解釈 GO Biological Processes

| Rank | GO Biological Processes | Count | % | P-value | Benjamini |
|------|-------------------------------|-------|------|----------------------|----------------------|
| 1 | Synaptic transmission | 6 | 21.4 | 7.9×10^{-5} | 3.4×10^{-2} |
| 2 | Transmission of nerve impulse | 6 | 21.4 | 1.7×10^{-4} | 3.6×10^{-2} |
| 3 | System process | 10 | 35.7 | 2.3×10^{-4} | 3.4×10^{-2} |
| 4 | Nervous system development | 8 | 28.6 | 9.6×10^{-4} | 1.0×10^{-1} |
| 5 | Cell communication | 7 | 25.0 | 1.1×10^{-3} | 9.0×10^{-2} |



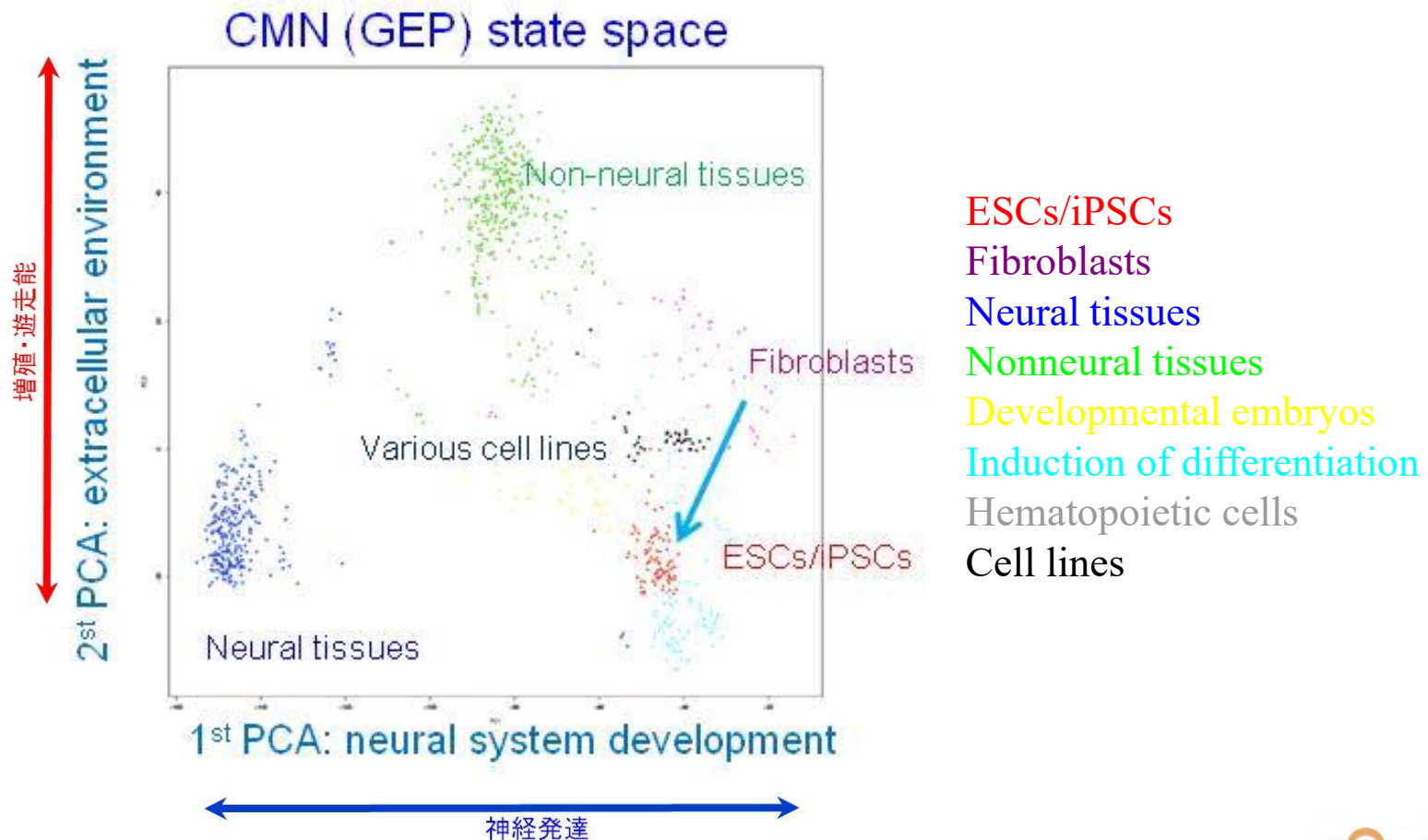
第1主成分は神経発達の指標

| Rank | GO Biological Processes | Count | % | P-value | Benjamini |
|------|---|-------|------|----------------------|----------------------|
| 1 | Extracellular matrix organization | 6 | 20.7 | 7.5×10^{-7} | 2.4×10^{-4} |
| 2 | Extracellular structure organization | 6 | 20.7 | 6.9×10^{-6} | 1.1×10^{-3} |
| 3 | Positive regulation of response to stimulus | 5 | 17.2 | 6.2×10^{-4} | 6.4×10^{-2} |
| 4 | Collagen fibril organization | 3 | 10.3 | 1.1×10^{-3} | 8.3×10^{-2} |
| 5 | Immune effector process | 4 | 13.8 | 1.5×10^{-3} | 8.9×10^{-2} |

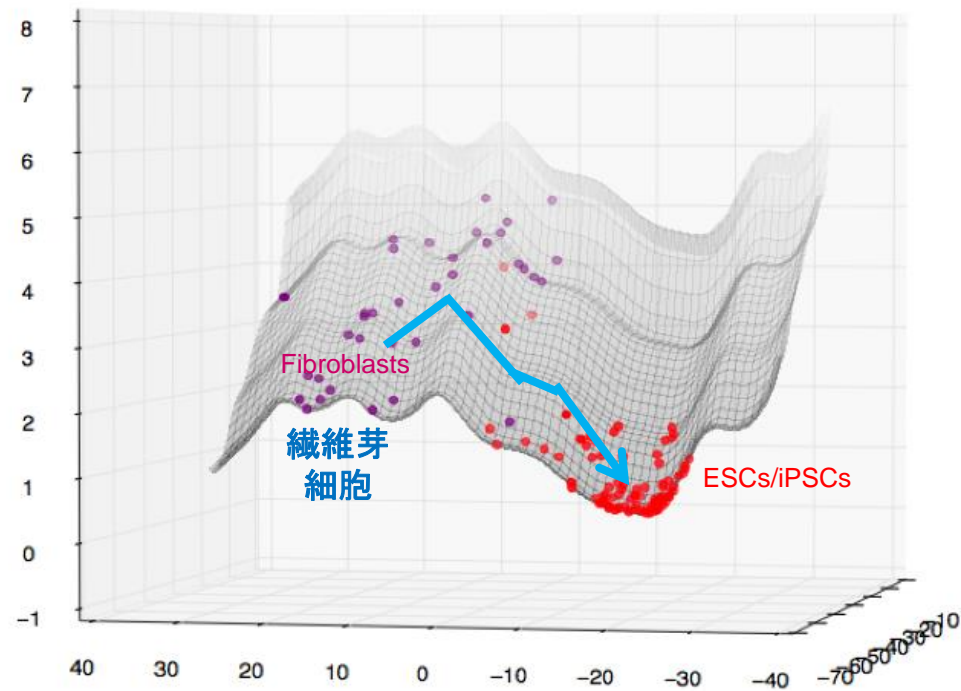
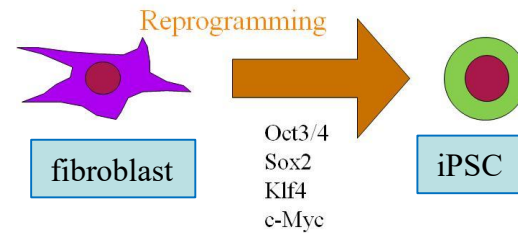
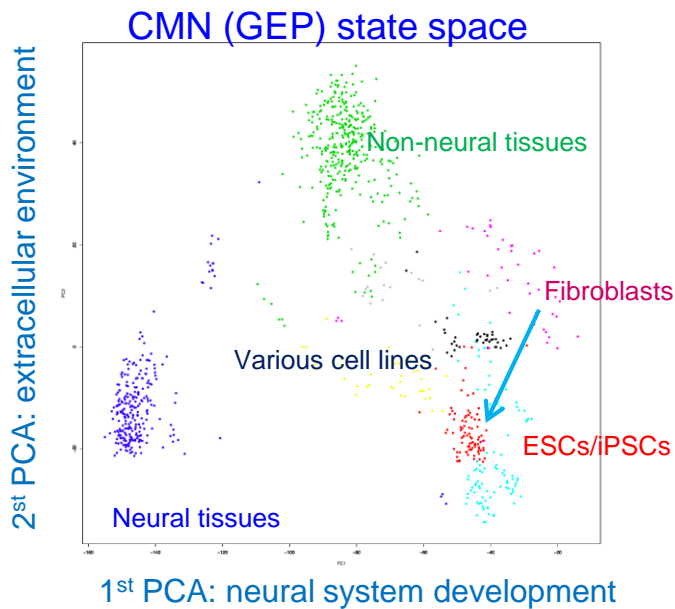


第2主成分は増殖・遊走能の指標

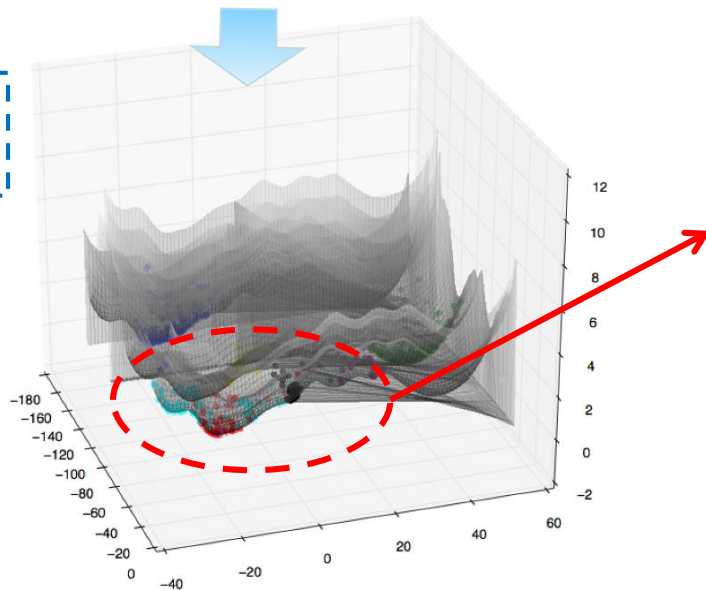
第1・2主成分によって構築された 細胞状態平面



iPSC reprogramming プロセス



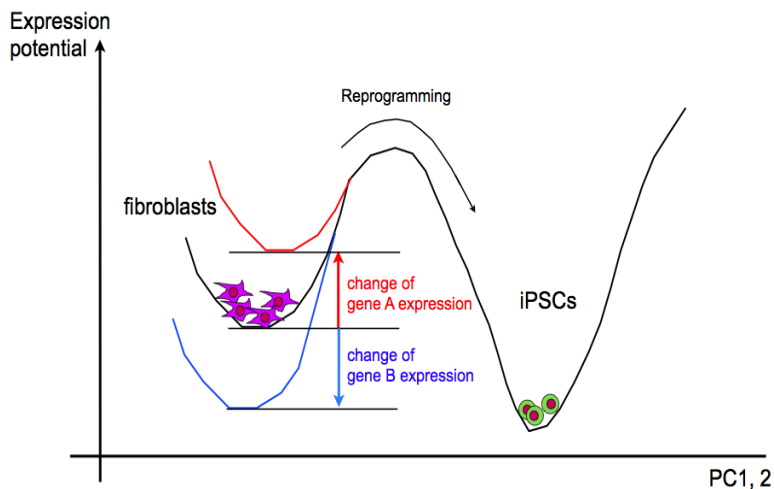
qWEL picture



Oxford press-JBiC prize 2012

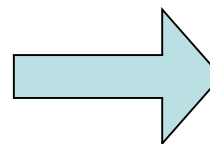
準ポテンシャルへの寄与度

1. 繊維芽細胞とiPSCの各遺伝子の発現の平均値を計算
2. 繊維芽細胞の発現値を iPSC の発現値に置き換えてポテンシャルのBasinの深さが変化するか調べる



| | |
|-----|-----------|
| 1 | CFH |
| 2 | DCN |
| 3 | MIR143HG |
| 4 | LMOD1 |
| 5 | NMU |
| 6 | SCG3 |
| 7 | C10orf116 |
| 8 | ASPN |
| 9 | PRELP |
| 10 | IGJ |
| ... | |
| 30 | SOX11 |
| ... | |
| 50 | SOX2 |

Transcription factors only



| | | |
|---|----|--------|
| 1 | 23 | HoxA10 |
| 2 | 30 | Sox11 |
| 3 | 34 | Zic2 |
| 4 | 50 | Sox2 |
| 5 | 85 | Emx2 |
| 6 | 91 | FoxM1 |

inc: iPScell increase
dec: iPScell decrease

考察と結論

- Sox11はES細胞におけるH3K9/14AcとH3K4me3の活性化に関係しており、ES細胞やiPS細胞でのH3K9AcやH3K4me3のレベルは線維芽細胞に比べて高いという報告がある。
 - iPS細胞でのSox11の発現量は上昇していたことから、Sox11はそれらの修飾レベルを高めることでリプログラミングに貢献している可能性がある。
- HoxA10はhomeobox遺伝子の一つであり、発達や形態形成に関与している。
 - HoxA10の発現量はiPS細胞では減少していたことから、HoxA10の発現をRNA干渉で抑制し発達を妨げることで、iPS細胞の樹立効率が向上する可能性がある。
- Emx2はSox2エンハンサーを抑制調節していることが知られている。
 - Emx2の発現量はiPS細胞で低くなっているため、Emx2 siRNAでその発現を抑制することでリプログラミング効率が向上する可能性がある。
- 発現ポテンシャル値に対する遺伝子の寄与度を計算することで、目的細胞への誘導において重要な遺伝子を推定できる
 - 線維芽細胞からiPS細胞への誘導においては、Sox2やSox11が転写因子の上位にランキングされた

ご清聴ありがとうございました

