

# オミックス創薬

—特に疾患ネットワークを用いたDR—

東京医科歯科大学 名誉教授

東北大学 東北メディカル・メガバンク

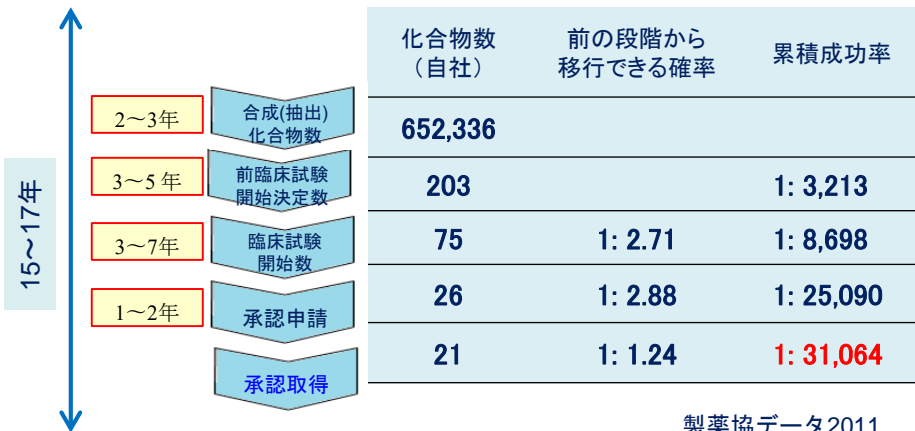
機構長特別補佐

田中 博



# 創薬を巡る状況

- 医薬品の開発費の増大
  - 1 医薬品を上市するのに約700~1000億円
- 開発成功率の減少
  - 2万~3万分の1の成功率
  - とくに非臨床試験から臨床試験への間隙
  - phase II attrition (第2相損耗)
- 臨床的予測性
  - できるだけ早い段階でのヒトでの有効性・毒性の予測
- 臨床予測性の向上
  - 罹患者のiPS細胞を使う
  - ヒトの網羅的分子情報を使う



# 網羅的分子情報に基づいた 計算論的創薬の基礎戦略

# 計算論的創薬

## computational drug discovery

### これまでの計算論的創薬

- 分子(構造)中心 (molecular-structure oriented)

### 分子構造解析・分子設計 (*in silico* drug design)

- Structure-based rational drug design
- 標的分子の分子構造解析
  - 薬剤(リガンド)との結合構造の分子構造解析
- リガンドの分子設計—分子力学・量子化学
- リード化合物の構造最適化
- 定量的構造活性相関(QSAR)の利用

### 新しい計算論的創薬のアプローチ (*mol. profile* drug design)

- オミックス創薬・システム創薬
- 網羅的分子プロファイル・分子ネットワーク変化中心
- 薬剤—標的分子の結合を取り巻く **genom-wide**な分子環境
- 標的分子より「疾患システム」という対象の把握
- 化合物—疾患の反応関係の「ビッグデータ」利用

# ネットワーク計算論的創薬の前提 「疾患」の網羅的分子把握の世代論

- 疾患の網羅的分子把握のパラダイム変化
  - ゲノム・オミックス医学の世代論

## 第1世代 (生得的) ゲノム医学

- 生殖系列細胞のゲノムの変異・多様性
- 個別化医療(personalized medicine)、テイラーメイド投薬
- 疾患原因遺伝子、感受性遺伝子による遺伝子型相対リスク

## 第2世代 オミックス医学

- 疾患罹患体細胞の網羅的分子情報
- 予測的・先制医療(predictive and preemptive medicine)
- 遺伝子発現プロファイルやプロテオームによる疾患予測・早期発見

## 第3世代 システム分子医学

- 疾患の分子パスウェイの知識増大
- 疾患：細胞分子ネットワークの歪み⇒調節不全分枝の同定
- 疾患の分子ネットワーク病態学に基づいた個別化予測医療
- 分子ネットワークに基づいた創薬戦略

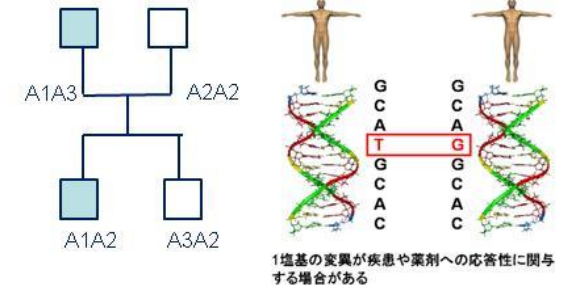
## (第3.5世代 環境統合システム分子医学)

- 環境・生活習慣との相互作用の場（エピゲノム機構や腸内細菌のメタゲノム機構）を取り入れた超生物的ネットワークによる疾患理解

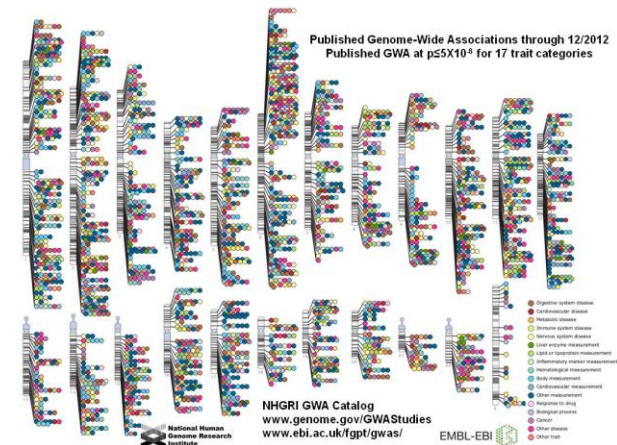


# 網羅的分子機序 第1世代 ゲノム医療 (1990~)

- 「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の疾患理解
  - 生得的ゲノムは全細胞/生涯を通じて同一
- 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 (99%浸透率) デシヤンヌ型筋ジストロフィー、嚢胞性線維症
    - 当時400程度のDNAマーカー ヒトゲノム解読計画へ
  - rare disease: 相対リスク大、遺伝難病で治療手段無
- 疾患感受性遺伝子(Disease susceptibility gene)
  - 多型性: 一塩基多型(SNP), 1000万,
    - 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) 患者対照分析
    - HAPMAPプロジェクト、1000ゲノムプロジェクト
  - common variant 仮説、相対リスク小: 1.1-1.5
    - 環境因子: 例, 喫煙肺がん相対リスク4.8
  - 発症予測に効果低い: multiple rare variants 仮説
  - 疾患の発症メカニズムなどに寄与

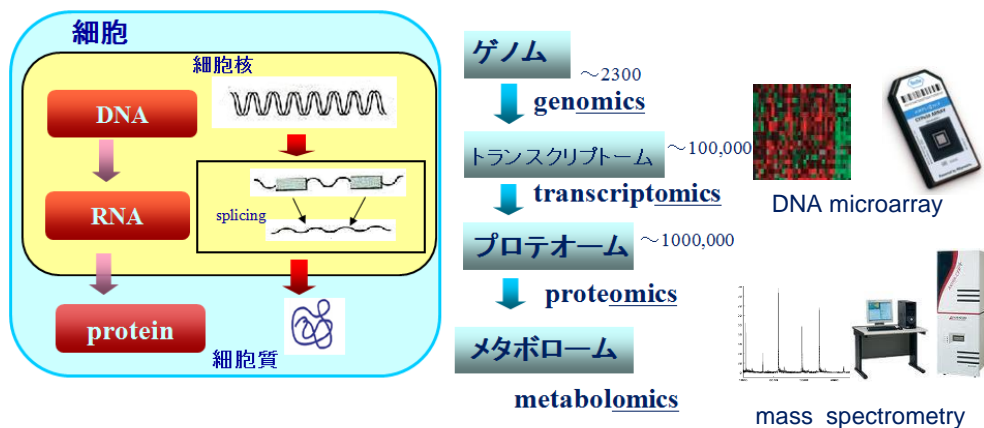
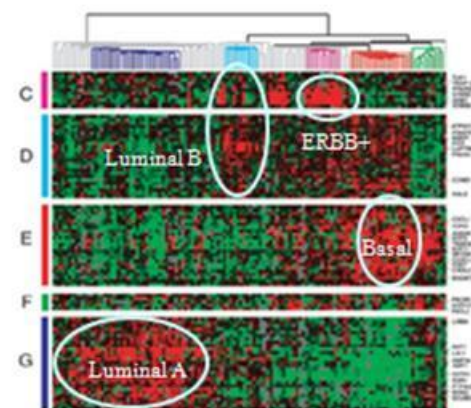


疾患	原因遺伝子・感受性遺伝子	座位
ハンチントン病	HTT(huntingtin) 遺伝子	4q16.3
筋ジストロフィー	DMD(dystrophin)遺伝子	Xp21.2
嚢胞性線維症	CFTR遺伝子	7q31.2
家族性大腸ポリポシス	APC遺伝子	5q21-q22
色素性乾皮症	XPA遺伝子(グループ)	9q22.3
炎症性大腸炎	CARD1 遺伝子など	16q21
統合失調症	COMT 遺伝子など	22q11.21-23
2型糖尿病	CAPN1 の遺伝子/ADIPOQ遺伝子	2q37.3/3q27
晩発性アルツハイマー	APOE遺伝子など	19q13.2
高血圧	AGT 遺伝子/ACE遺伝子など	1q42-43/17q23
乳がん	BRCA1/2 遺伝子など	17q21/13q12
慢性骨髄性白血病	BCR-ABL融合遺伝子	22q11
腎臓がん	VHL遺伝子	3p25.3



# 網羅的分子機序 第2世代 オミックス医療 (2000~)

- 「後天的・疾病依存的に変化する」  
(体細胞) ゲノム・オミックスに基づく医療
  - 疾患 Omics Profile: 後天的ゲノム変異 (がん)、遺伝子発現プロファイル  
プロテオーム・メタボローム。網羅的分子表現型 "molecular phenome"
  - 病態進行に依存: 臨床症状や病理変化より早く変化
  - 疾患の予後予測・早期診断・先制医療
- 疾患のオミックスプロファイル (網羅的分子病態像)  
サブタイプ分類 “疾患の患者個性”
  - 臨床・病的な分類に対してintrinsicな分類 (例,乳がん)



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

乳がんの亜型分類と予後・治療

# 網羅的分子機序 第3世代

## システム分子医学

個別化医療・予測先制医療のための  
疾患のシステムの理解に基づく医療

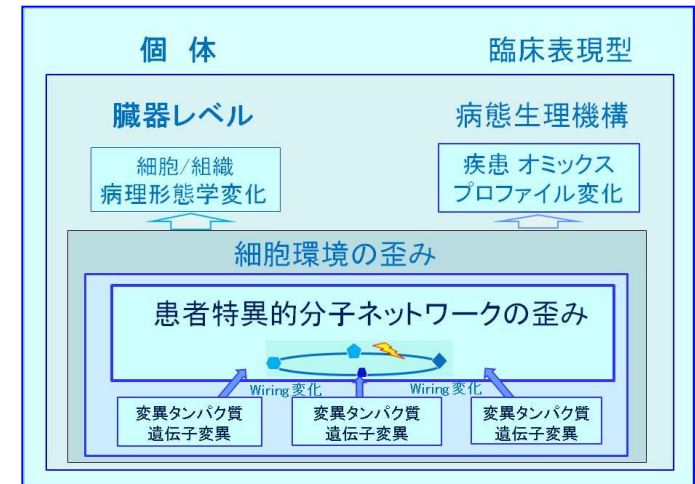
「疾患をシステムとして理解する」

システムバイオロジーの疾患への応用

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不全や歪み **distortion (dysregulation) of molecular network**

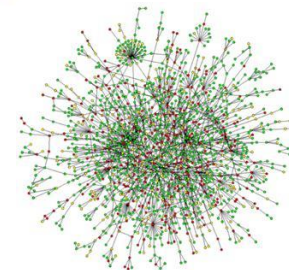
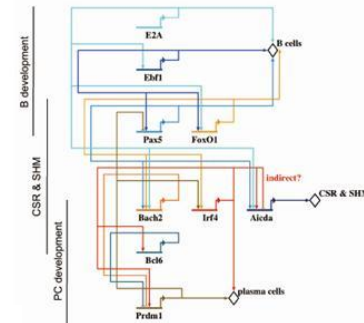
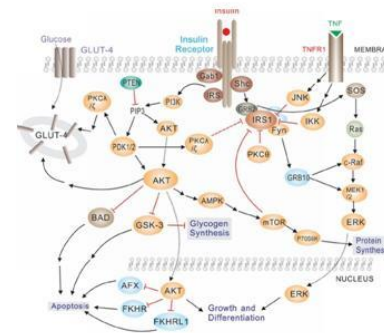
オミックスプロファイルをそのままではなく細胞分子ネットワークに投影して、疾患による歪み（過剰発現、機能不全パスウェイ分枝）を同定

Tanaka H.: *Meth. Info. Med.* 49:173-185, 2010



シグナル伝達ネットワーク

遺伝子発現調節ネットワーク



タンパク質相互作用ネットワーク



# オミックス創薬の原理

# オミックス創薬の基礎

## もっと基本的な方針

薬剤および疾患による網羅的分子表現型のGenome-Wideな変化

### ● 薬剤特異的遺伝子発現 (Drug-induced SDE)

#### — CMAP : Connectivity Map (現在はLINCSなど)

- 薬剤投与による遺伝子発現プロファイルの変化
- 米国 Broad Institute, 1309化合物, MCF7, PC5など5 がんセルライン, 7000 遺伝子発現プロファイル
- Signature (遺伝子発現刻印: 差別的発現遺伝子の代表的集合)  
Signature of Differential gene Expression
- DB利用: SDEをquery, 順位尺度で類似性の高い順に化合物を提示
- 最近LINCS 100万サンプルへ拡張

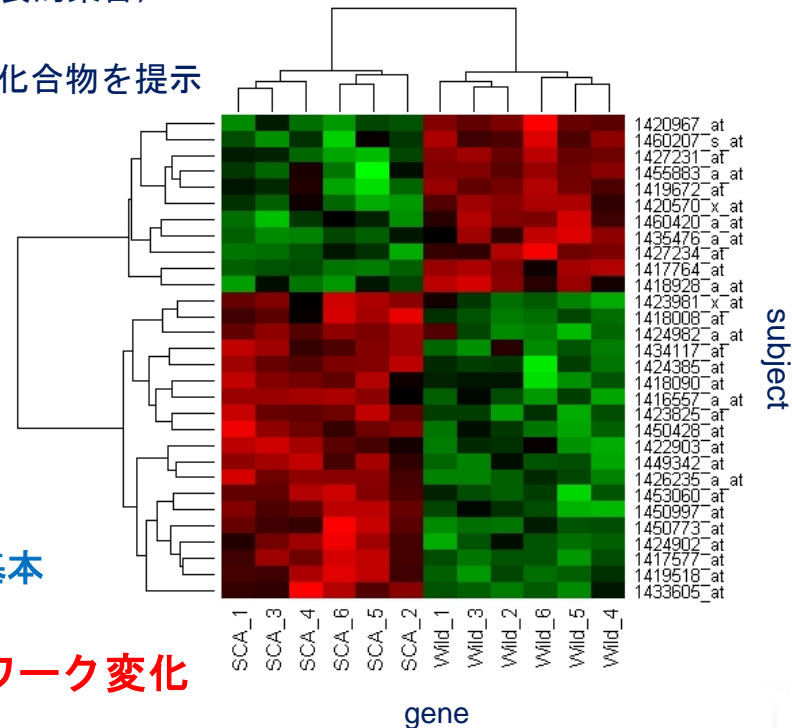
### ● 疾病特異的遺伝子発現

#### (Disease-associated SDE)

- GEO (gene expression omnibus),
  - 疾病罹患時の遺伝子発現プロファイルの変化
  - 米国NCBI作成・運用 2万5千実験, 70万プロファイル
  - ArrayExpress EBI作成DB、サンプル数同程度

本来は、分子ネットワークの疾病/薬剤特異的変化が基本  
(システム医学・ネットワーク医学)。

遺伝子発現プロファイル変化  $\approx$  分子ネットワーク変化

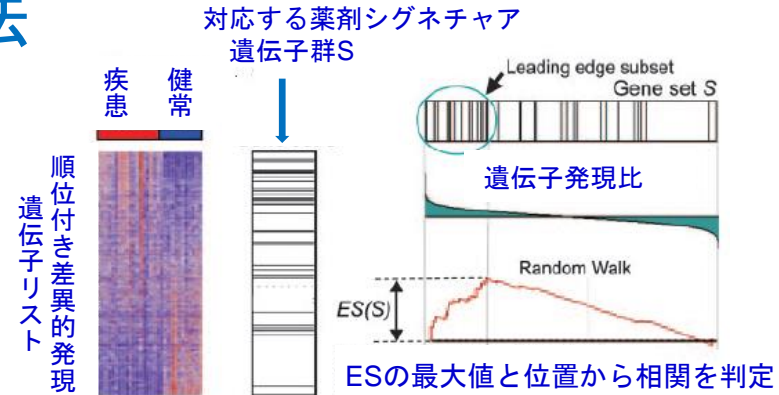


# 遺伝子発現プロファイルによる有効性予測

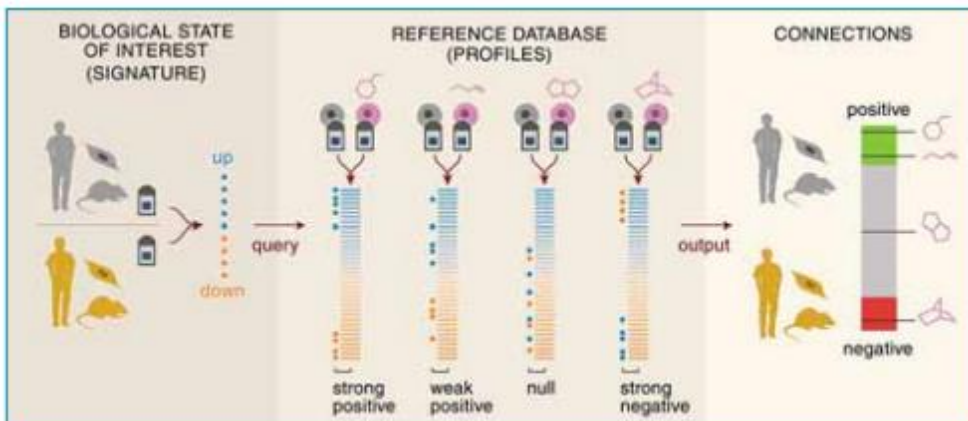
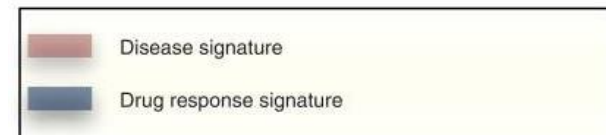
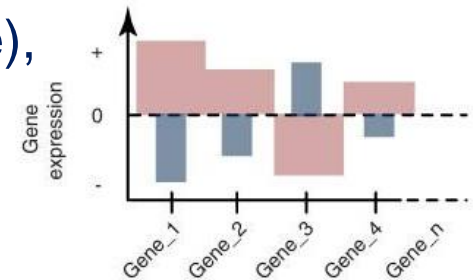
## 遺伝子発現シグネチャ逆位法 (signature reversion)

- 薬剤特異的遺伝子発現シグネチャ
- 疾患特異的遺伝子発現シグネチャ
- 有効性予測**：両者が負に相関する
- Non-parametric な相関尺度で評価
  - Gene Set Enrichment Analysis (GSEA) :
    - 対照と比較して順位づけられた遺伝子リストの上位に密集しているかの尺度
- 例：炎症性腸疾患IBDに 抗痙攣剤(topiramate), 骨格筋委縮にウルソール酸

GSEA

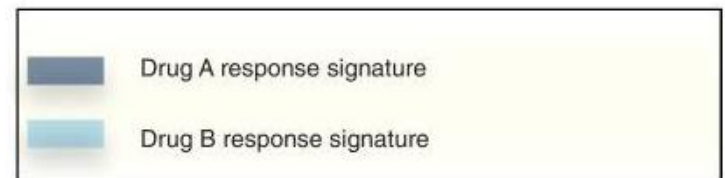
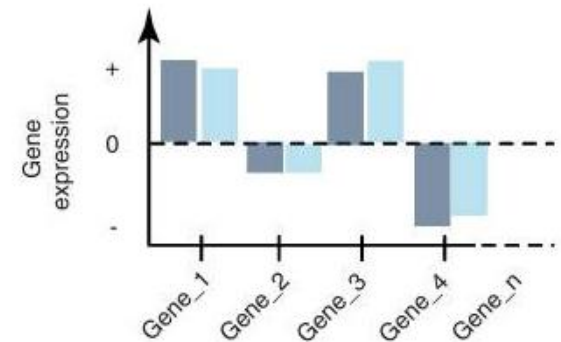


発現比ランクの高い順から遺伝子を調べ 遺伝子リストS中に該当する遺伝子が存在したらESを加算、無ければ減算



# 遺伝子発現プロファイルによる毒性予測

- 連座法 guilt-by-association :
- 薬剤－疾患間 副作用予測
  - 薬剤特異的シグネチャと
  - 疾患特異的シグネチャが
  - ノンパラメトリック相関 (ES) 正
  - 毒性・副作用の予測
- 薬剤－薬剤間
  - 薬剤ネットワークからのDR
  - Connectivity map から薬剤特異的遺伝子発現の薬剤間の類似性をノンパラメトリック親近性尺度 (GSEA)で評価
  - この類似性のもとに薬剤ネットワーク構築
  - 近隣解析によりDR
  - 例：抗マラリア剤をクローン病に適応



# 計算論的DRの現状

# Drug Repositioning

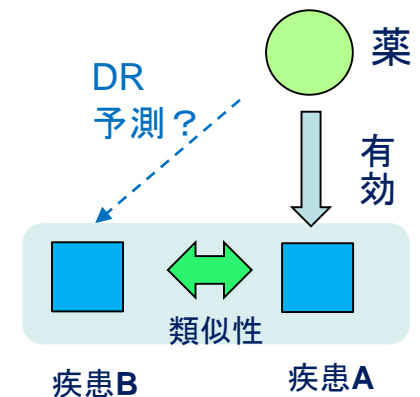
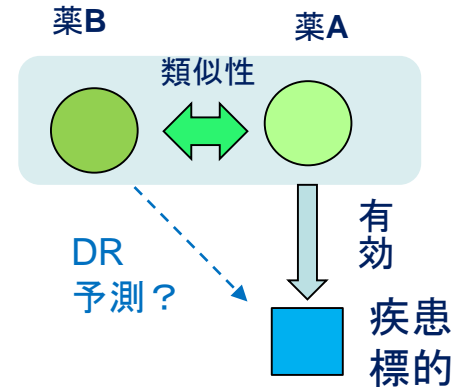
ヒトでの安全性と体内動態が十分に分かっている  
既承認薬の標的分子や作用パスウェイなどを、体系的・論理的・網羅的に解析することにより新しい  
薬理効果を発見し、その薬を別の疾患治療薬として  
開発する創薬戦略

## 利 点

- (1) 既承認薬なので、ヒトでの安全性や体内動態などが既知で臨床試験で予想外の副作用や体内動態の問題により開発が失敗するリスクが少なく**開発の成功確率が高い**
- (2) 既にあるデータや技術（動物での安全性データや製剤のGMP製造技術など）を再利用することで、**開発にかかる時間とコストを大幅に削減できる**

# 合理的DRへのアプローチ

- 医薬品中心 Drug-based (drug-centric)
  - 医薬品の構造・特徴の類似性に基づいて別の医薬品の適応を予測
    - ① 化合物の化学的構造・特徴の類似性
    - ② 薬物投与時の遺伝子発現プロファイル
- 疾患中心 Disease-based(disease-centric)
  - 疾患の発症機序の類似性に同一の医薬品が別の疾患の適応を予測
    - ① 疾患原因/感受性遺伝子の共有
    - ② 疾病遺伝子発現プロファイル
    - ③ 疾患を起こす分子ネットワークの類似性
- 両者の融合的アプローチ



# 網羅的分子情報を用いた計算論的DR

## (1) 疾患ネットワーク準拠法

- 疾患のゲノム・オミックス機序の類似性をネットワークで表す

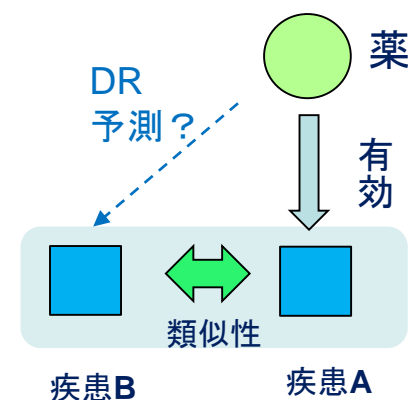
### DR候補予測

- 既存の<疾患-薬剤>を決め  
近傍の疾患をDR候補とする

## (2) 疾患-薬剤相関準拠法

疾患-薬剤profile相関より  
有効性・毒性の確認

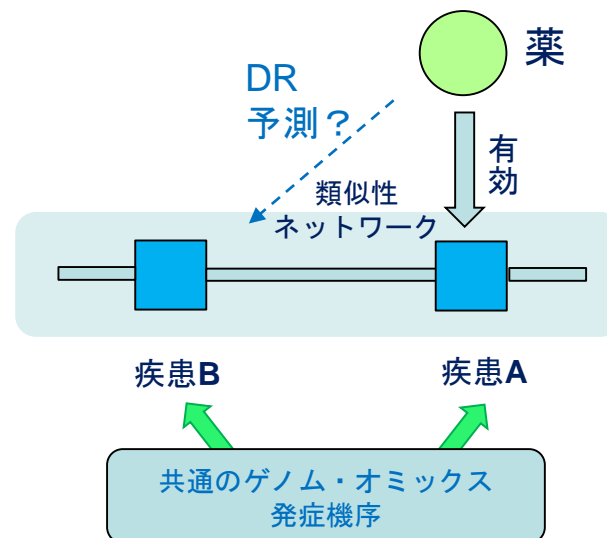
- シグネチャア逆位法、連座法で有効性、毒性を予測





# 疾患ネットワークに基づいたDR

- 従来の疾患体系 nosology
  - Linne以降300年に亘って表現型による疾病分類
  - 臓器別・病理形態学別の疾患分類学
- ゲノム・オミックスレベルでの発症機構による疾患分類
  - 発症機構類似性を基準に疾患ネットワーク
  - ゲノム・オミックス医学の疾病概念が基礎



## 疾患の成立機序における主要機序

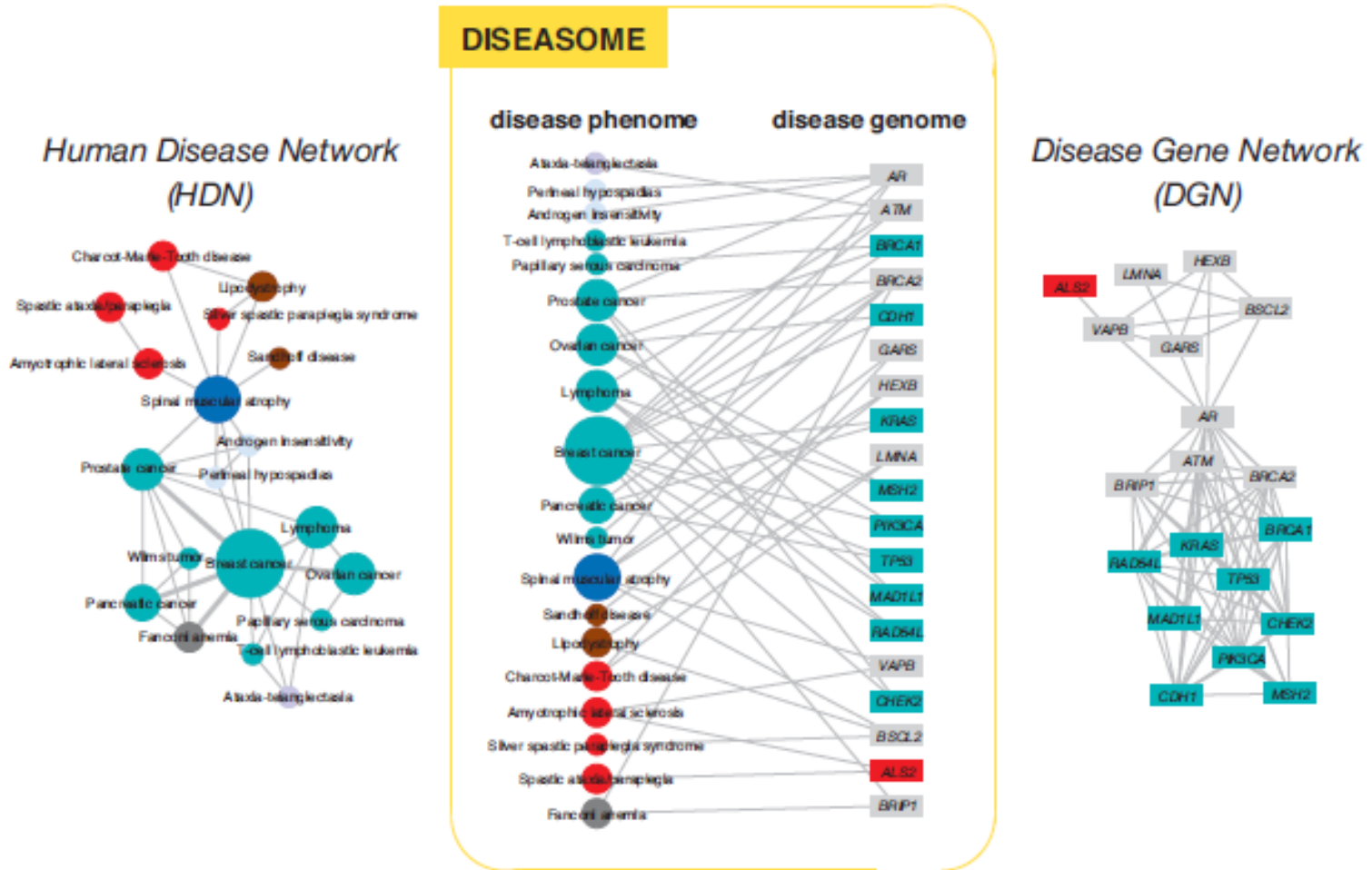
- 疾患関連遺伝子型（第一世代型）
  - 原因遺伝子、疾患感受性遺伝子の変異・多型が主要発症機序
- 疾患オミックス型（第2世代型）
  - 疾患オミックスプロファイルの変容が主要発症機序
  - Transdisease omics
- 疾患分子ネットワーク型（第3世代型）
  - 分子ネットワークの歪みが主要発症機序
  - がんなどで遺伝子型（肺腺がん等）でない通常のがん

# 第1世代型

## Diseasomeと疾患遺伝子

- OMIMから 1,284 疾患と 1,777 疾患遺伝子を抽出
- ヒト疾患ネットワーク (HDN)
  - 867疾患は他疾患へリンクを持つ 細胞型や器官に非依存
  - 516疾患が巨大クラスターを形成
    - 大腸がん、乳がんがハブ形成
    - がんはP53 やPTENなどにより最結合疾患 がんなどは後天的変異
  - 疾患を網羅的に見る見方：臓器や病理形態学に非依存
  - リンネ (12疾患群分類) 以来300年続いた分類学を越える
- 疾患遺伝子ネットワーク (DGN)
  - 1377遺伝子は他遺伝子へ結合
  - 903遺伝子が巨大クラスター
    - P53がハブ
- ランダム化した疾患/遺伝子ネットワークに比べ
  - 巨大クラスターのサイズが有意に小さい
- 疾患遺伝子は機能的なモジュール構造
  - 同じモジュールに属する遺伝子は相互作用し
  - 同一の組織で共発現し、同じGOを持つ

# 疾患ネットワーク Diseasome (Goh, Barabasi et al.)



1つ以上の疾患関連遺伝子を共有する疾患

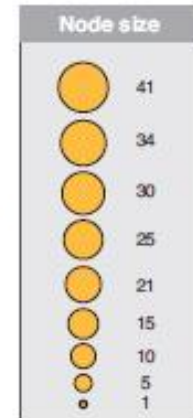
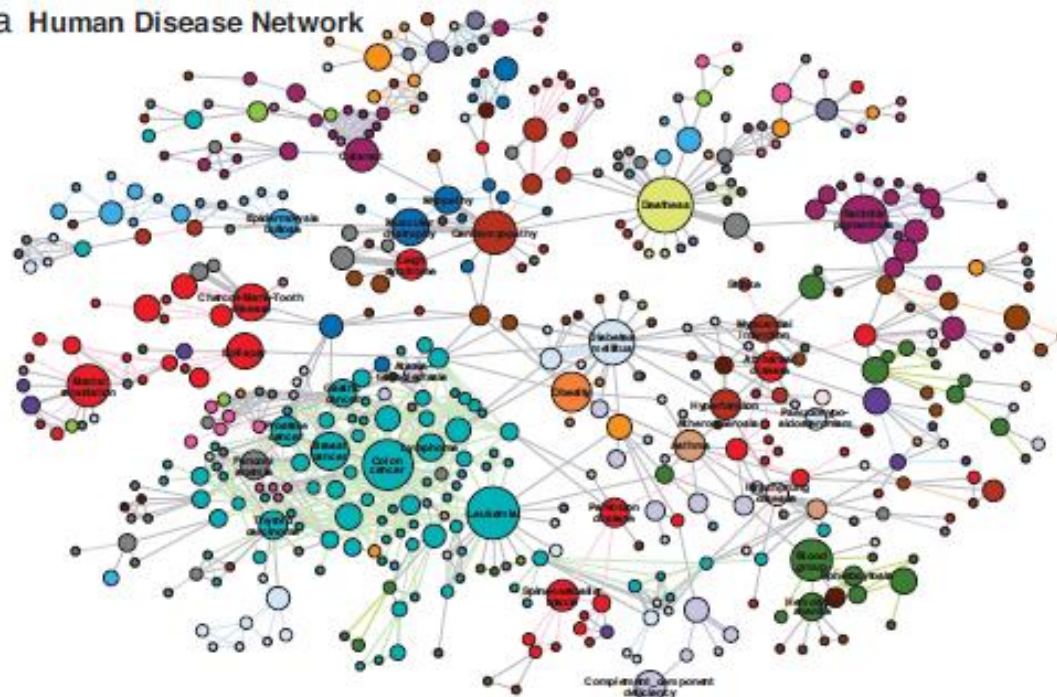
1つ以上の疾患を共有する疾患関連遺伝子

Kwang-Il Goh\*, Michael E. Cusick, David Valle, Barton Childs, Marc Vidal, and Albert-Laszlo Barabasi The human disease network PNAS2007

# 疾患ネットワーク (HDN)

**Nodeの直径**  
 疾患に関与している原因遺伝子の数に比例  
**リンクの太さ**  
 疾患間で共有している原因遺伝子の数

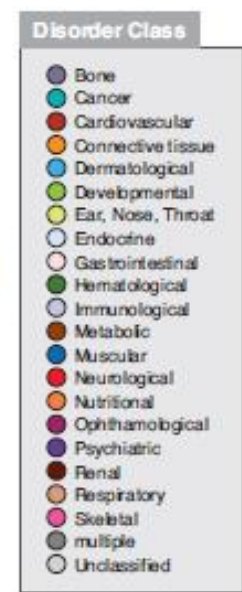
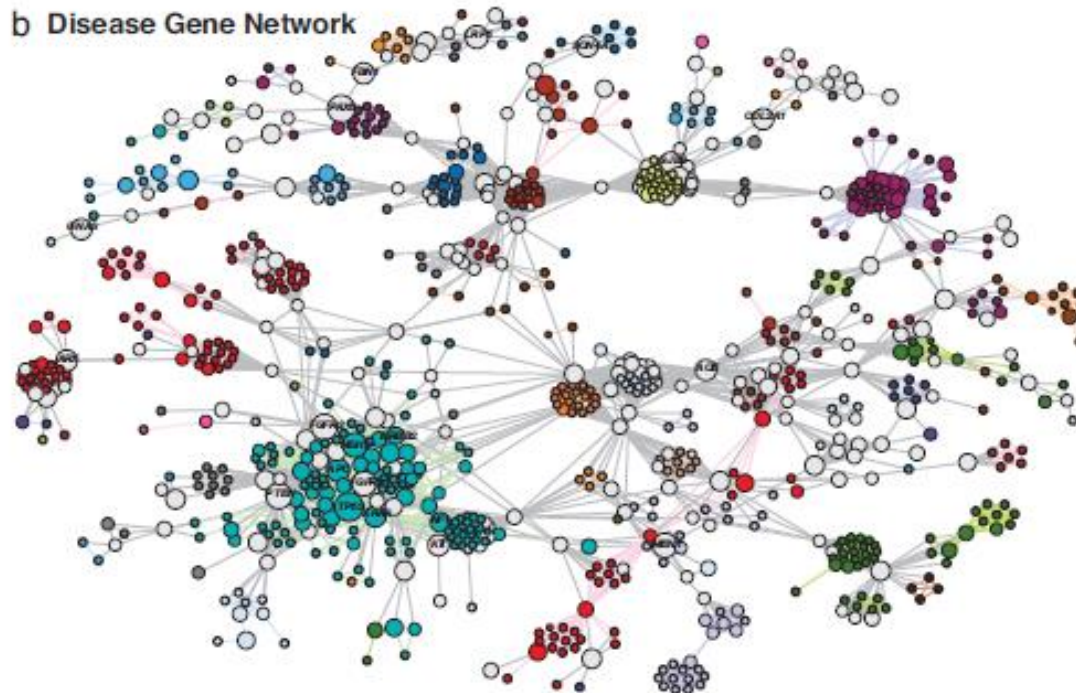
a Human Disease Network

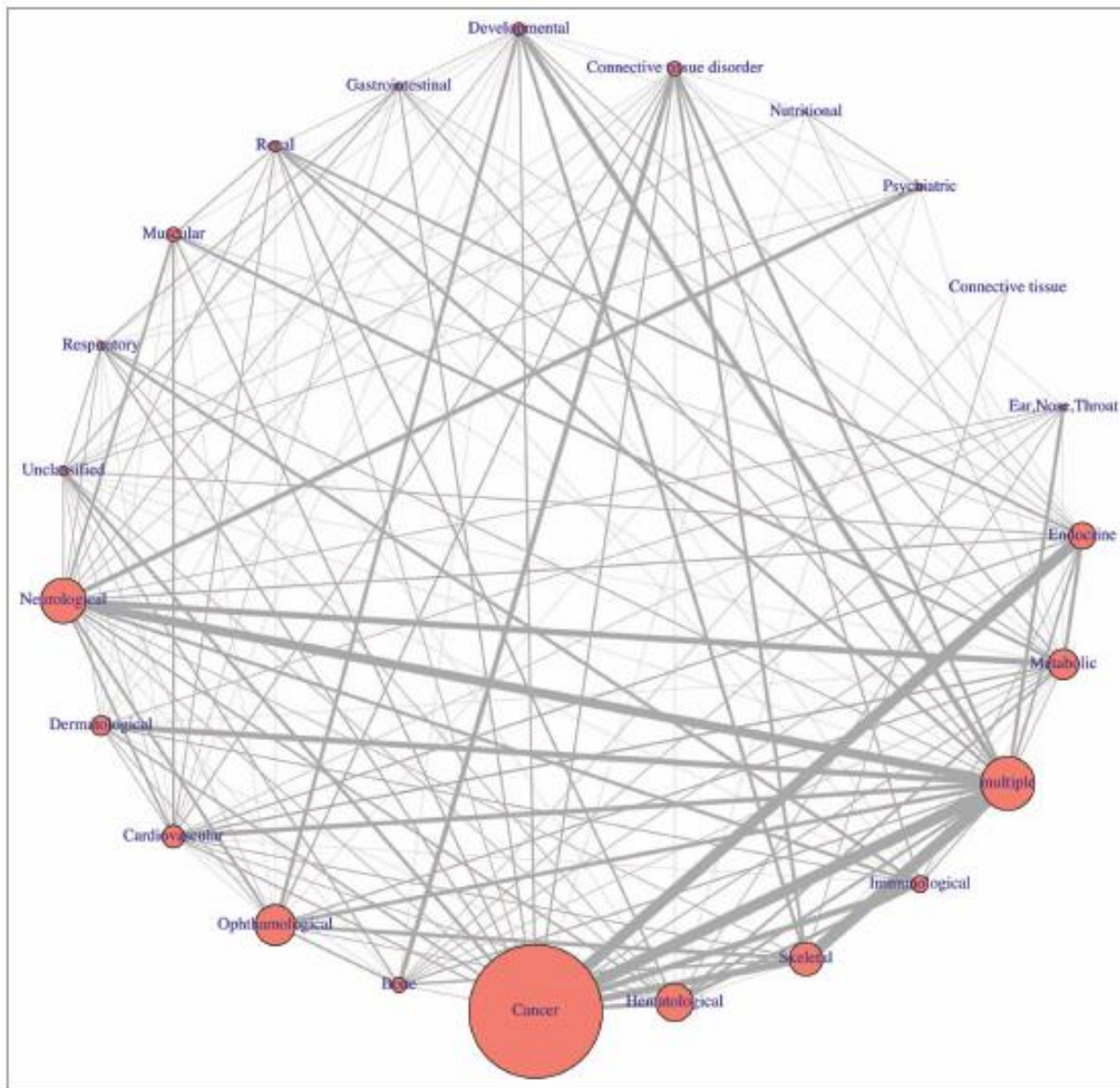


# 疾患遺伝子ネットワーク (DGN)

**Nodeの径**  
 その遺伝子を原因にしている疾患の数に比例  
 2つ以上の疾患に関与していると明灰色の遺伝子ノード

b Disease Gene Network





23グループのOMIM疾患のネットワーク  
 ノードの径：共通リンクの合計、リンクの太さ：共通遺伝子数

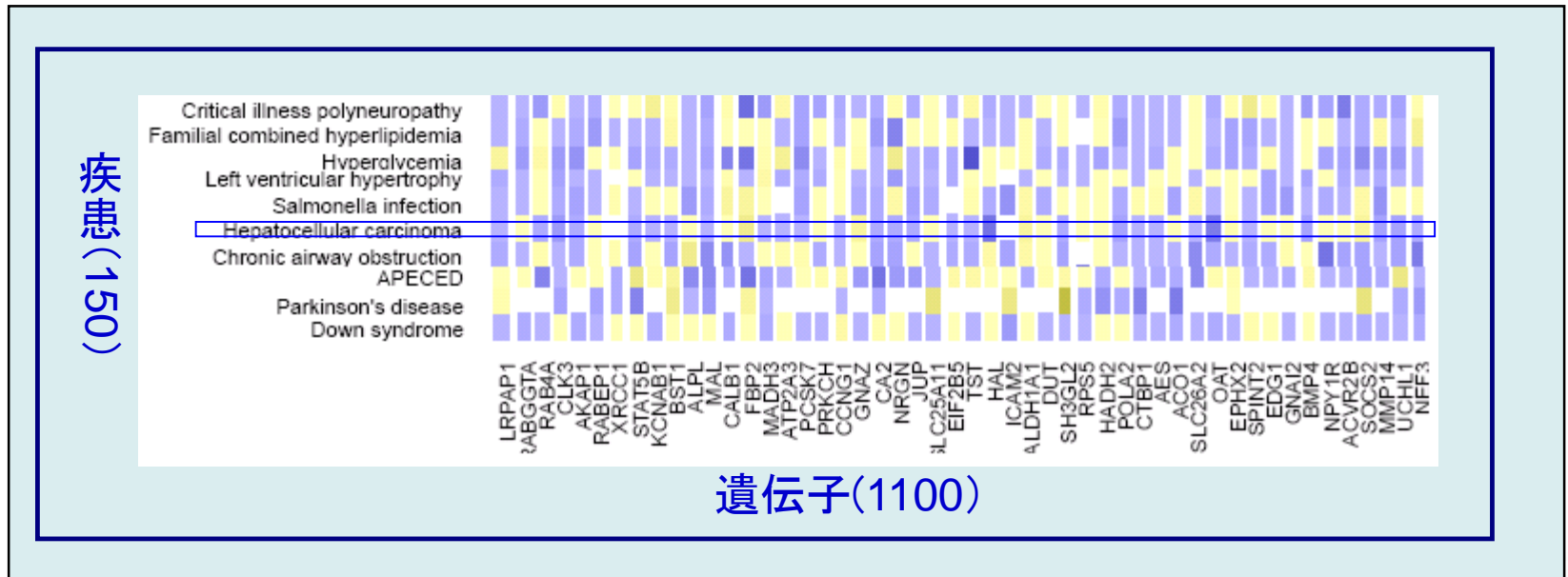
# Diseasomeを巡る状況

- Mendel疾患、複雑疾患、環境疾患へと発展
- **他のネットワークと融合**
  - タンパク質相互ネットワーク、代謝ネットワーク
    - PPIの近傍(Vanunu),代謝網での酵素の基質の共有
  - GWAS (WTCCC、NIH-GAD)のSNPの共有
    - すべてがつながり偽陽性のネットワークで有効性低い
  - miRNA, 環境因子 (annotation MEDLINE)
    - 電子カルテから時系列病歴収集
    - 進化的直系的表現型性 (他の動物も利用)
  - パスウェイ準拠型の疾患ネットも
- **表現型疾患ネットワークも存在する**
  - Phenotype : MeSHの頻度をベクトルとする(van Driel)
- **Diseasomeは、臨床表現型ネットワークと分子ネットワークを繋げる機構**
  - 遺伝子を通して疾患間を移動できる
  - Systems pathobiology, nosology, personalized medicine

# 第2世代型

## GENOMED (A. Butte et al)

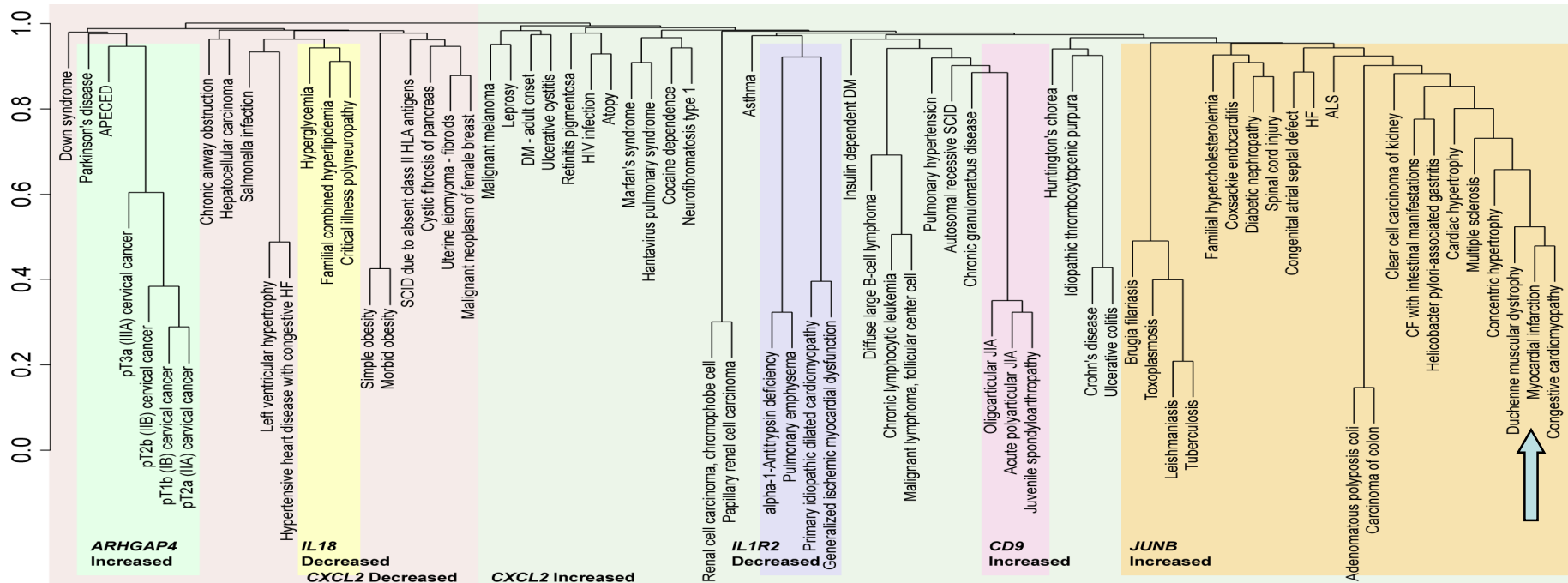
- 遺伝子発現DBのGEO (Gene Expression Omnibus) 利用
  - 約20万のサンプル
- 疾患名は注釈文より用語集UMLSを用いて抽出
- 疾患ごとに多数の遺伝子発現パターンを平均化





# Gene-Expression Nosology of Medicine

- 疾患を平均遺伝子発現パターンよりクラスター分類
  - 臓器別疾患分類では予想できない疾患間の親近性
  - 分類項目はサイトカインの遺伝子発現と相関
  - 疾患の再体系化に基づいた医薬の repositioning
- さらに656種類の臨床検査を結合した分析
- 心筋梗塞・デュシャンヌ型筋ジストロフィーに近い



# Transcriptomeの変化をPPIに投影した疾患ネットワーク (Butte)

- ネットワークモジュール

遺伝子発現プロファイルではなく4620に分解したPPIネットワークの機能moduleでの疾病での平均発現変化をもとに疾患ネットワーク構築

- 基本方法

- GEOから信頼性などより54の疾患を選択
- 各疾患について各moduleに含まれる遺伝子群の **疾患時と健常時の発現差のt統計量の平均**
- MRS: Molecular Response Score 各疾患に各モジュールで定義 (ベクトル量)
- **疾患間の相関**は、両疾患の健常時発現を制約とした
- **MRSの偏相関係数**

- 疾患ネットワークの性質

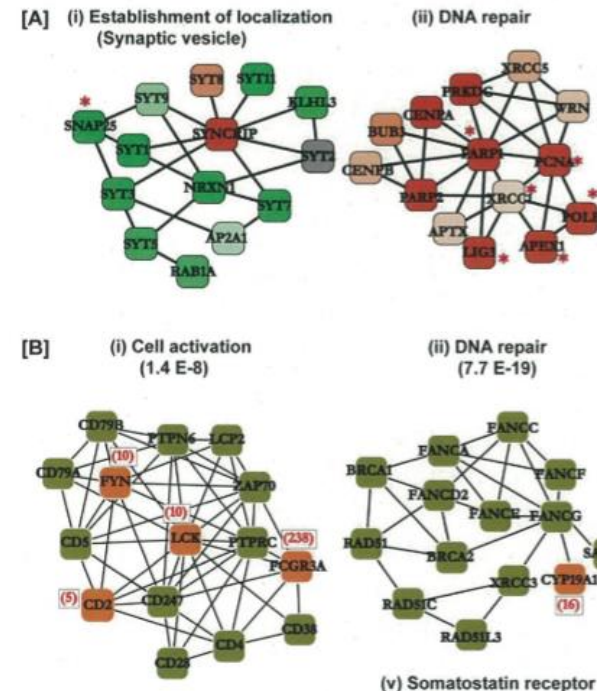
- **138の有意な相関**: ランダム化ネットに対し有意
- $p < 0.01, FDR = 0.1$
- **疾患類似性**: 肺がん群: 修復パスウェイ, 精神疾患: SNP-25

- 138の疾患相関のうち

- 17は少なくとも1つの共通薬 (14は共通の薬剤に有意)
- Fluorarcil (日光性角化) ⇒ 大腸がん、ほかDoxorubicin

- 疾患の大半を占める59モジュール

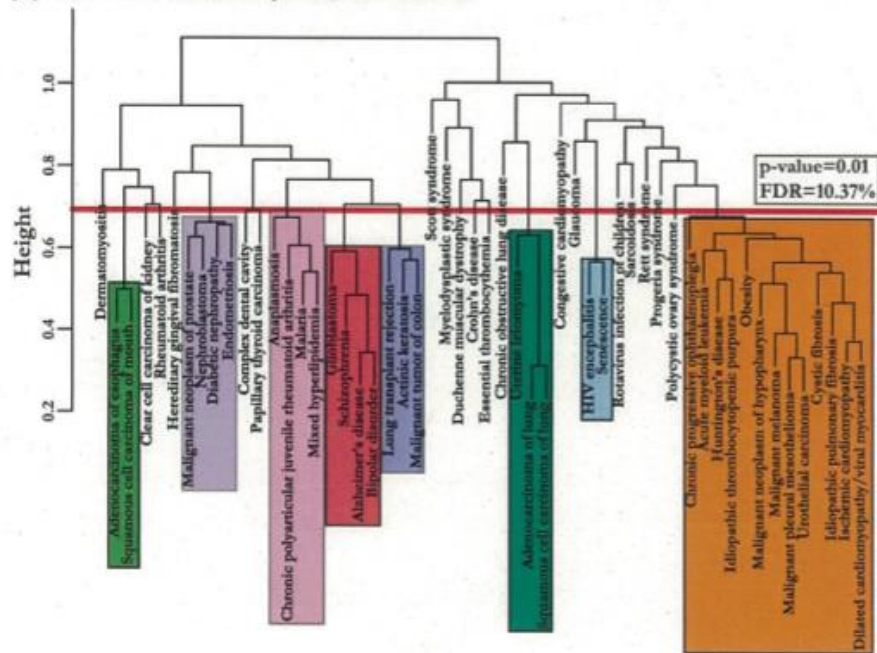
- 共通疾患シグネチャ



# Transcriptomeの変化をPPIに投影した 疾患ネットワーク (Butte)

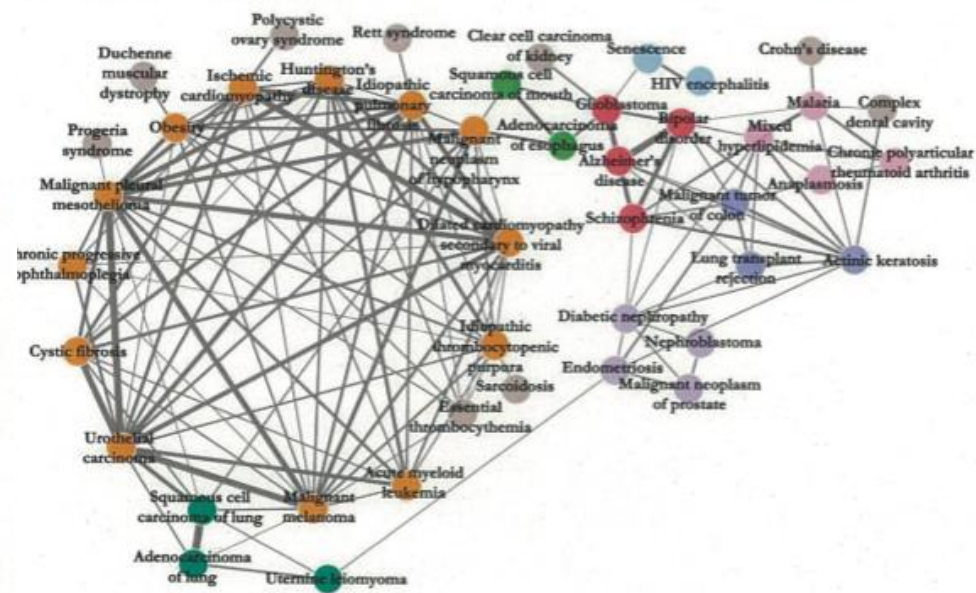
- アルツハイマー症、統合失調症、双極性障害がグループ化
- 子宮筋腫と肺がん、マラリアとクローン病
- 17のがんが1つの群ではない。がんの異質性
- 疾患ネットワーク間の遺伝子共有は高くない (遺伝子外効果)

[A] Hierarchical relationships between diseases



階層的クラスタリング

[B] All significant disease correlations



疾患ネットワーク

# Transcriptional Profiling による疾患ネットワーク (Hu, Agarwal)

遺伝子発現プロファイルの類似性を、相関係数、ESによるリンク

疾患-疾患、薬剤-薬剤、疾患-薬剤のネットワークを発現プロファイルより構成

疾患 (disease-disease) 645 組  
 疾患-薬 (disease-drug) 5008 組  
 薬 - 薬 (drug-drug) 164,374 組

## 結果

①疾患関連の60%はMeSH、  
 その他はTranscriptomeによる発見

## ②主な発見

### <疾患 - 疾患>

HSP (Hereditary Spastic Paraplegia

(遺伝性痙攣性対麻痺)

⇒bipolar双極性うつ病

Solar keratosis 日光性角化症

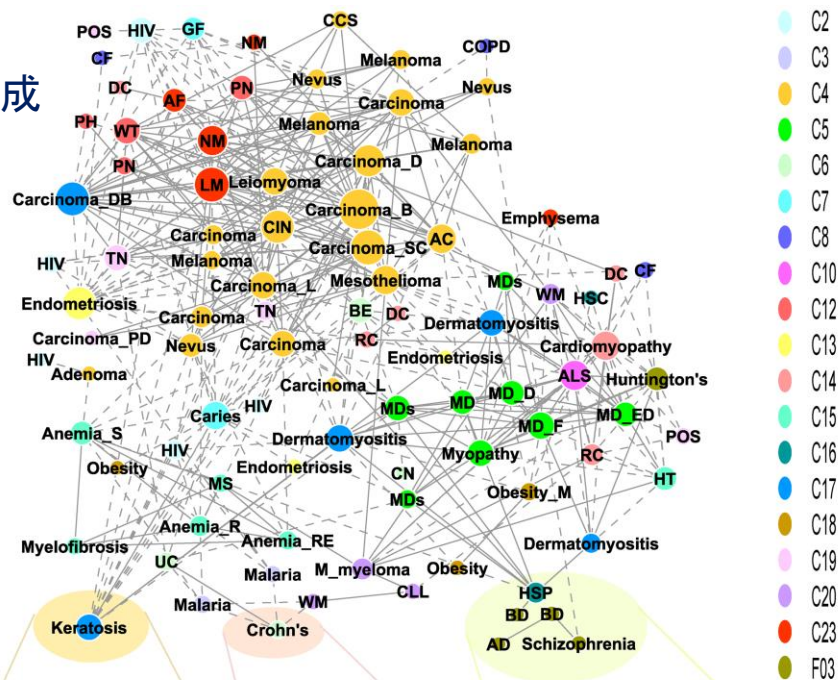
⇒ cancer(squamous)

### <疾患 - 薬>

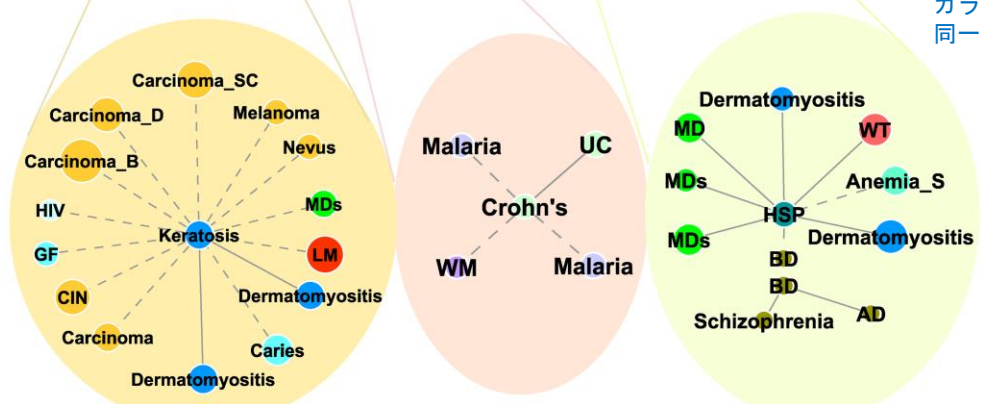
マラリア治療薬

⇒ Crohn's disease

ハンチントン病に種々の薬剤



カラーはMeSH  
 同一カテゴリー



# Transcriptional Profiling による Drug-Diseaseネットワーク

Disease-drug network

橙色 病気

緑色 医薬品

Tamoxifen (breast cancer)

(負の値をもっている)

⇒アトピー,

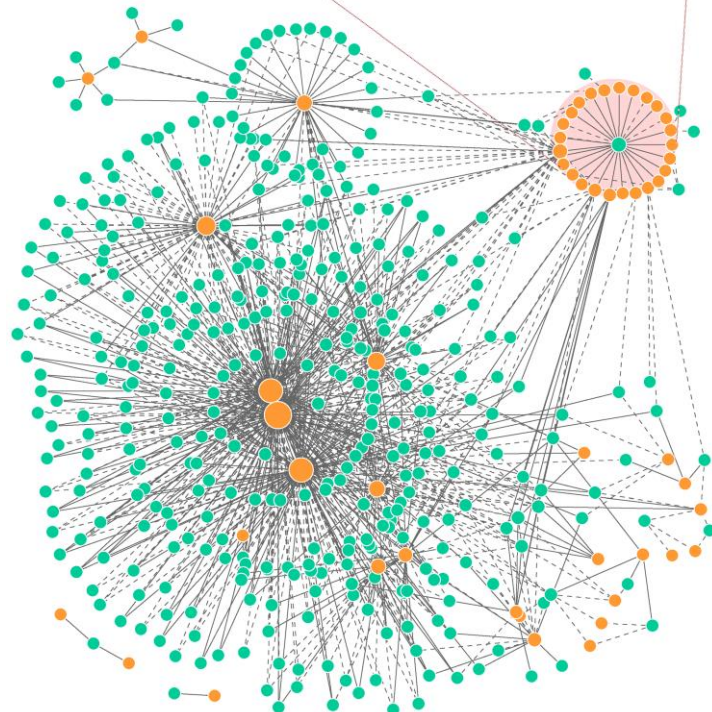
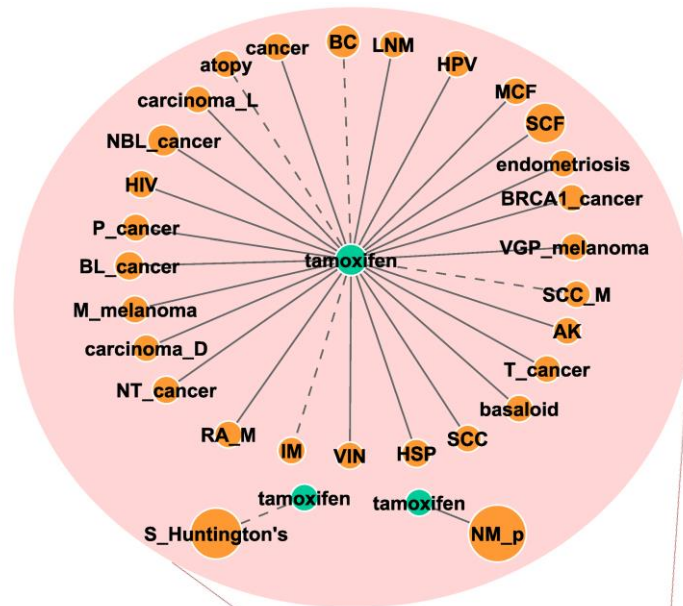
⇒マスト細胞分泌抑制、

アレルギー抑制

(正の値をもっている)

副作用の予測

⇒ 発癌性



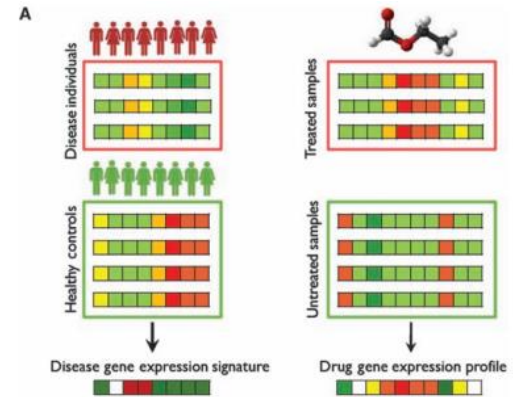
# Drug-Disease interactionを 基礎とした計算DR（Sirota, Butte2011）

## アプローチ

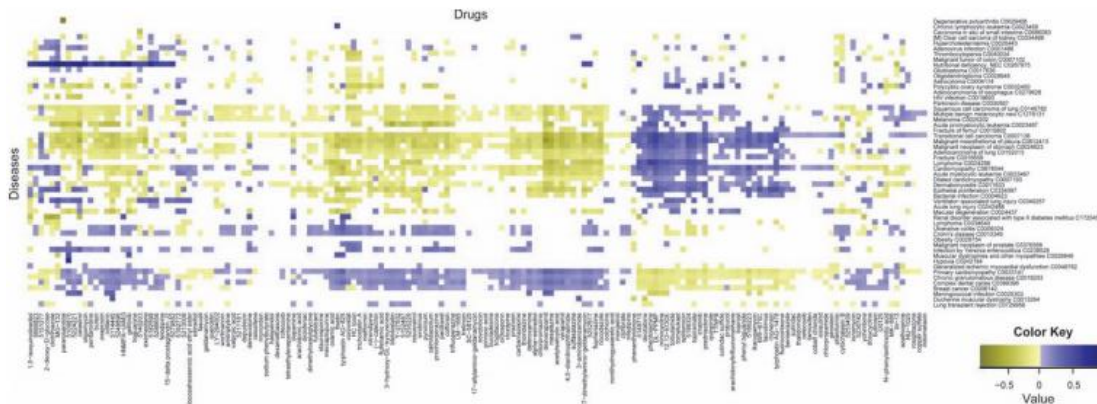
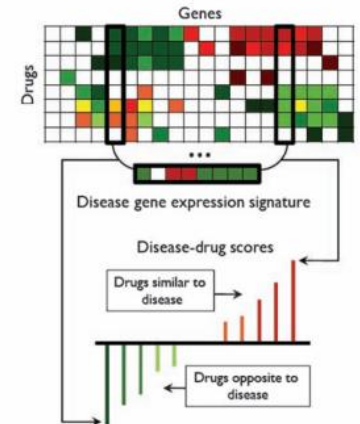
- NCBI・GEOを用いる
- 100疾患のシグネチャを取得
- 164の薬剤・化合物
- c-Mapより得た参照遺伝子発現プロファイル
- 疾患の発現をqueryにして疾患-薬剤間で類似性スコアを計算
- 16000組の疾患-薬剤間の2664組が有意、半数以上が治療的関連性あり
- 164疾患の内、53疾患が有意に薬剤と関連

Table 1. Drugs and diseases with the most indications.

Drugs with most indications		Diseases with most indications	
Vorinostat	21	Transitional cell carcinoma	95
Gefitinib	18	Melanoma	79
HC toxin	18	Cardiomyopathy	73
Colforsin	17	Adenocarcinoma of lung	72
17-Dimethylamino-geldanamycin	16	Multiple benign melanocytic nevi	68
Trichostatin A	16	Squamous cell carcinoma of lung	67
3-Hydroxy- $\alpha$ -kynurenine	15	Malignant neoplasm of stomach	66
5114445	15	Dermatomyositis	63
Dexverapamil	15	Malignant mesothelioma of pleura	53
Prochlorperazine	15	Primary cardiomyopathy	48

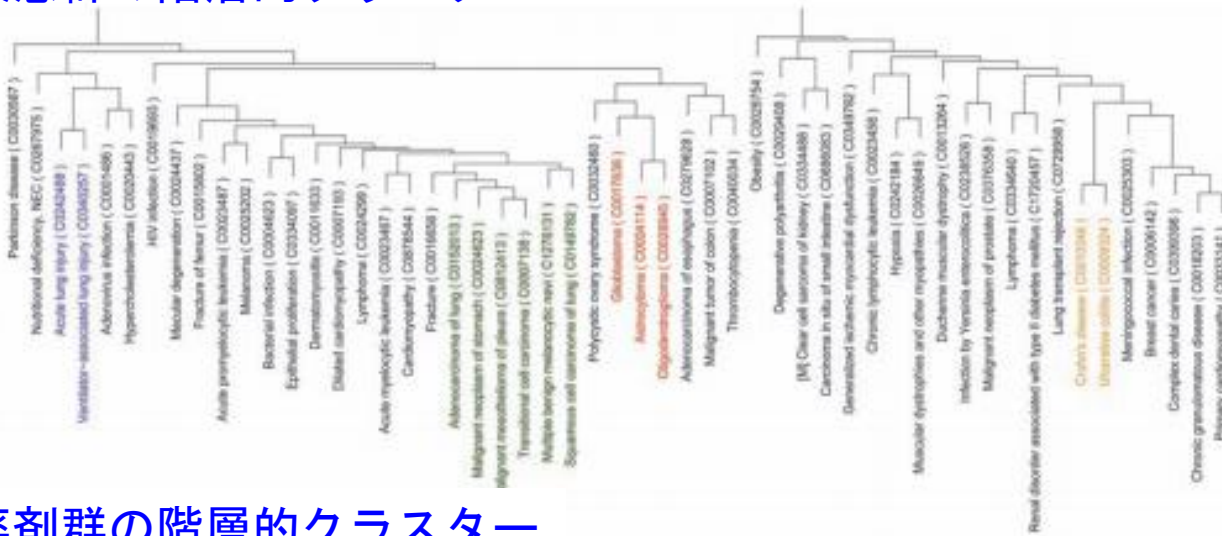


B Reference database of drug gene expression



# Drug-Disease interactionを 基礎とした計算DR (Sirota, Butte2011)

## 疾患群の階層的クラスター



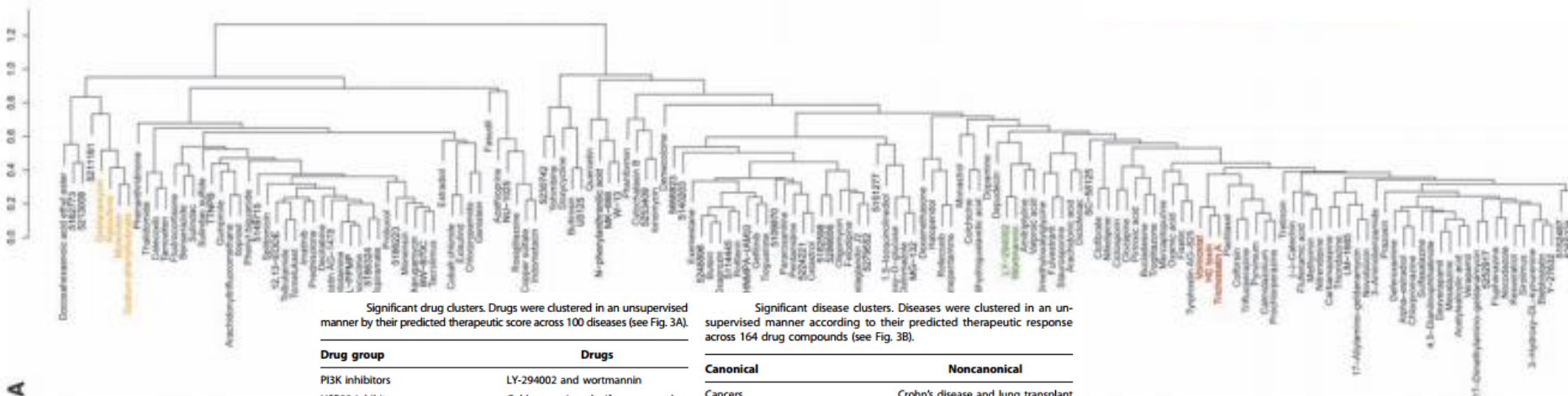
疾患のクラスター解析

- ・がんの大クラスター
- ・IBD、潰瘍性大腸炎、クローン病のクラスター

薬剤のクラスター解析

- ・HDAC阻害剤  
Hctoxin, trichostatin
- ・PI3K, 抗炎症剤クラスター
- ・HSP90関連薬剤

## 薬剤群の階層的クラスター



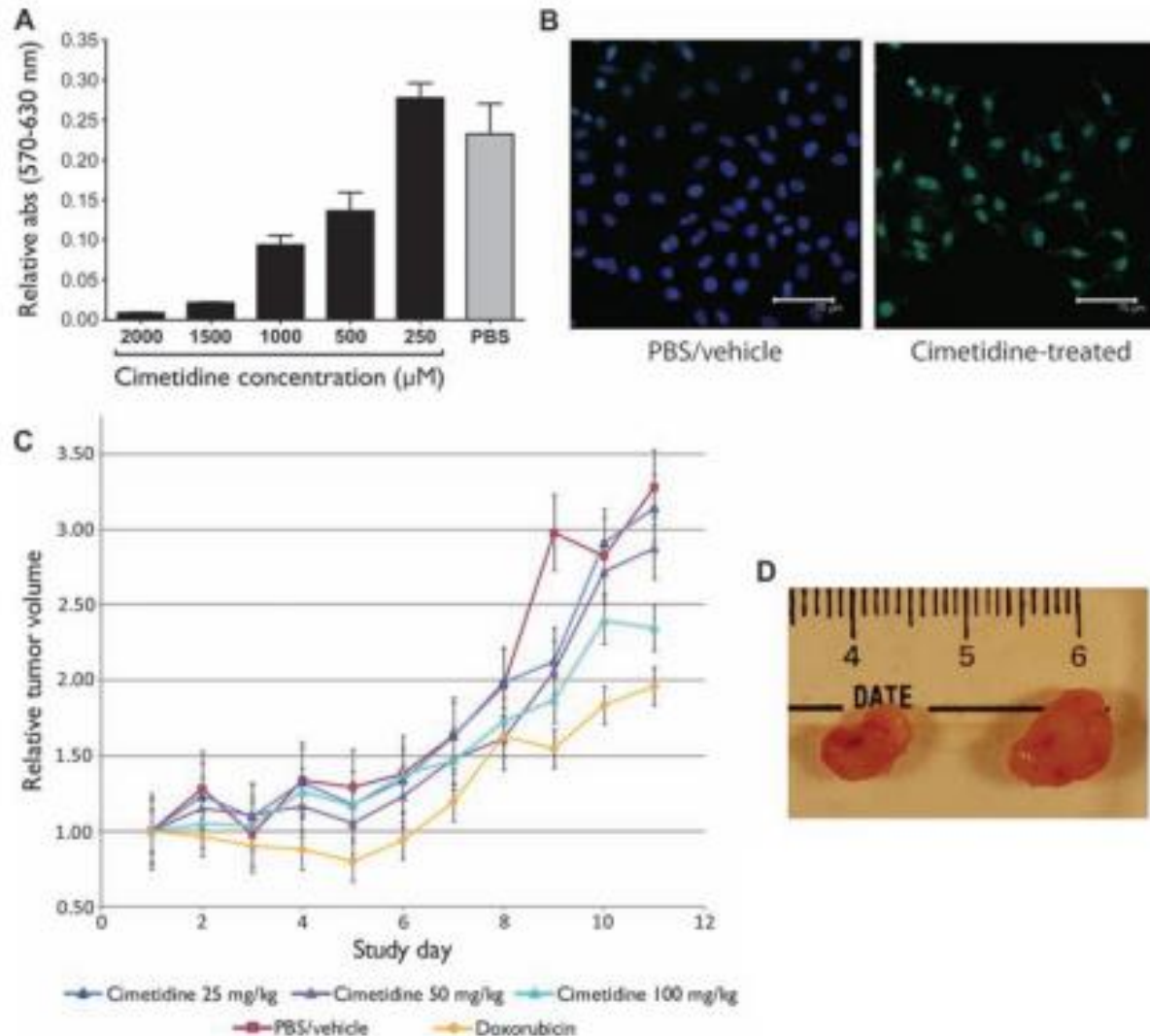
Significant drug clusters. Drugs were clustered in an unsupervised manner by their predicted therapeutic score across 100 diseases (see Fig. 3A).

Drug group	Drugs
PI3K inhibitors	LY-294002 and wortmannin
HSP90 inhibitors	Geldanamycin, raloxifene, monorden, and sodium phenylbutyrate
HDAC inhibitors	Vorinostat, HC toxin, and trichostatin A
Salicylate anti-inflammatory agents	Sulfasalazine, mesalazine, and acetylsalicylic acid

Significant disease clusters. Diseases were clustered in an unsupervised manner according to their predicted therapeutic response across 164 drug compounds (see Fig. 3B).

Canonical	Noncanonical
Cancers	Crohn's disease and lung transplant
Ulcerative colitis and Crohn's disease	Polycystic ovary and glioblastoma
	Cardiomyopathy and cancer

# 動物実験での実証





# 最近の話題 LINCS

- LINCS (library of Integrated network-based cellular signatures)
  - GE-HTS(gene expression high throughput screening)の1つ
  - 摂動（化合物添加）を与えて、調節系を介して細胞表現型を観察する
  - 遺伝子発現変化⇒差別的発現 **signature**
  - cMAP (2006, Lamb)に比べてスケール拡大
    - cMAPは、4つの細胞系列～ 1300化合物 FDA認可薬剤  
Micro array (mRNA) Affymetrix U 113で遺伝子発現測定
- NIHから助成, 百万の遺伝子発現プロファイルを L1000 技術で測る
  - Broad Institute cMAPと同じメンバーが考案
  - 1000遺伝子の発現しか測定しない ゲノムワイドな遺伝子発現プロファイル（～全遺伝子 22000 genesの発現）をGEOから作ったモデルで推定する
  - 相互依存性高い⇒1000遺伝子にすべて情報が含まれている
- L1000技術
  - 細胞溶液からリガンド媒介増幅によってmRNA増幅
  - 遺伝子特異的なProbeはcDNA (mRNA) にtaqリガーゼでアニールする
  - ProbeはPCRで増幅され、ルミネックスビーズと遺伝子特異的部分で対形成する
  - 対形成した差異染色ビーズはレーザを用いて検出され定量化される
  - ビーズの上の対形成したprobeの密度を測る80の恒常的発現校正遺伝子
- 22412 摂動遺伝子発現
  - 56 細胞コンテキスト（ヒト初代培養細胞、がん培養細胞）について
  - 16425 化合物、薬剤
  - 5806 遺伝子ノックアウト(RNAi, miRNA)、過剰発現
  - 総計で100万ぐらい遺伝子発現プロファイルがある
- Genometry が**L1000**™ Expression Profiling技術でヤンセンと契約
  - 25万種類の化合物

# LINCSの問合せ画面

## --- LINCS Canvas Browser ---

### Gene Lists

#### Up List

- EEF1A2
- UBE2S
- FAM64A
- FGFR1
- PAXIP1
- SPARC
- SNRPA1
- ADAMTS1
- EIF4EBP1
- PFKP
- BTG2
- CDK16
- ERRFI1
- ARPC4
- IFI30

clear

#### Down List

clear

Up  Down

Search Example Enrich

Aggravate  Reverse

### Top 50 Consensus Experiments (Down/reverse)

Overlap	Info (Perturbation, Dose, Time, Cell, Batch)
0.5000	<a href="#">Tyrphostin AG 1478.56.78 μm 24 h A375 CPC006</a>
0.5000	<a href="#">PD0332991.2 μm 24 h MDAMB231.LJP001</a>
0.5000	<a href="#">PD0332991.10 μm 24 h MDAMB231.LJP001</a>
0.5000	<a href="#">PD0332991.10 μm 24 h MCF10A.LJP001</a>
0.5000	<a href="#">Aminopurvalanol A.10 μm 24 h PC9 CPC002</a>
0.5000	<a href="#">3,5-dichloro-2-hydroxyphenyl(phenyl)benzenesulfonamide</a>
0.4800	<a href="#">PD0332991.2 μm 24 h BT20</a>
0.4800	<a href="#">PD0332991.10 μm 24 h BT20</a>
0.4800	<a href="#">MLN2238.10 μm 24 h BT20</a>
0.4800	<a href="#">2-(6,6-dimethoxy-3-oxo-1,2,3,4-tetrahydrophthalazine-5-yl)carbamoyl)phenylpropan-1-amine</a>

Showing 1 to 10 of 47 entries

### Average Change - Time Point - Drugs - Dose

IL1  
100 ng/μL, 6 h  
in BT20

contrast:

Avg. Z-score:

Select a cell line: BT20

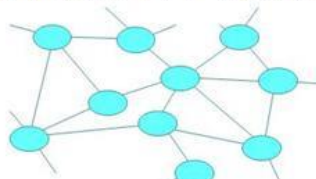
Select a batch: LJP004

Multiple Selections:

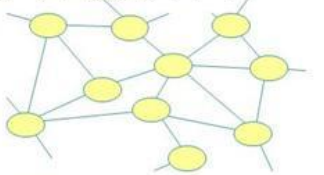
# 疾患ネットワークとDrug Projection Map

## DR informaticsの構築

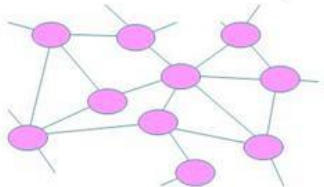
第1世代疾患ネット (原因遺伝子親近性)



第2世代疾患ネット (OmicsProfile親近性)

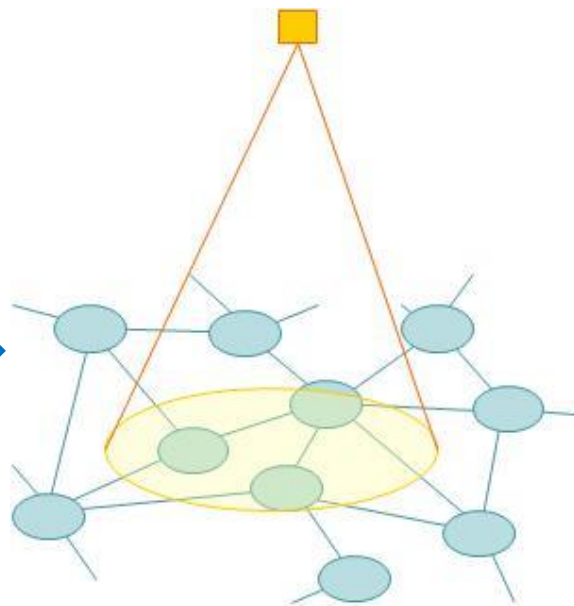


第3世代疾患ネット (Pathway親近性)



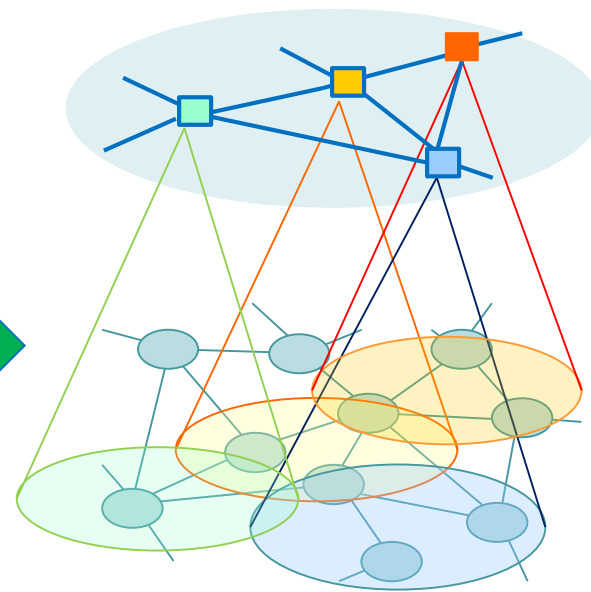
統合化

薬剤



疾患ネットワーク

薬剤ネットワーク

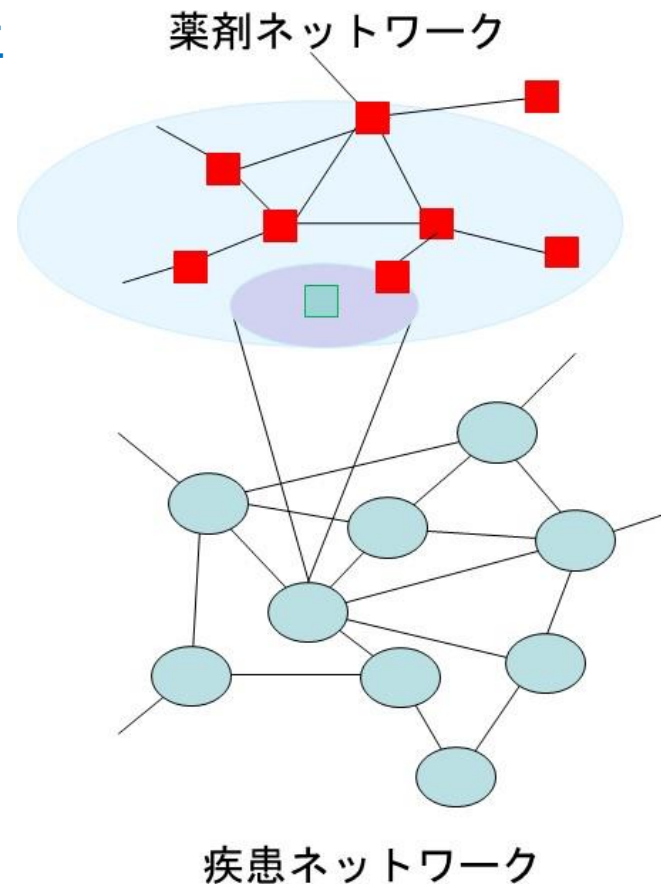
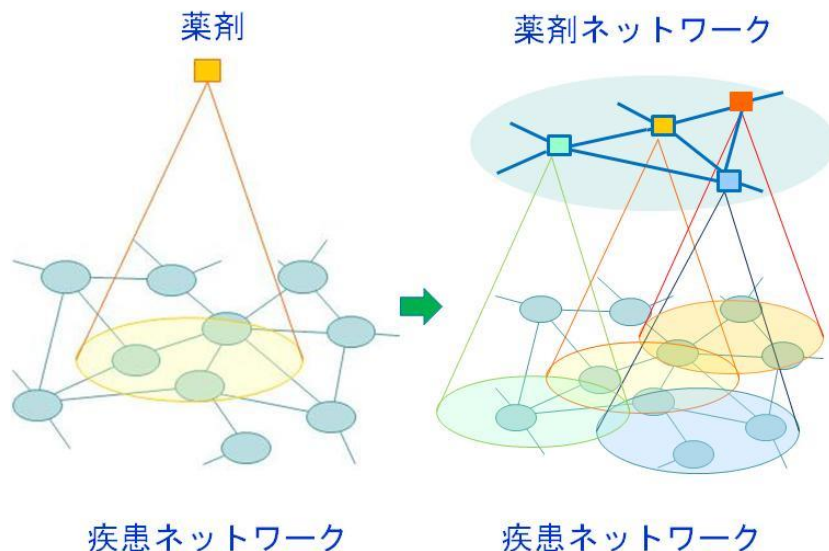


疾患ネットワーク

# DRの方法論から創薬方法論へ

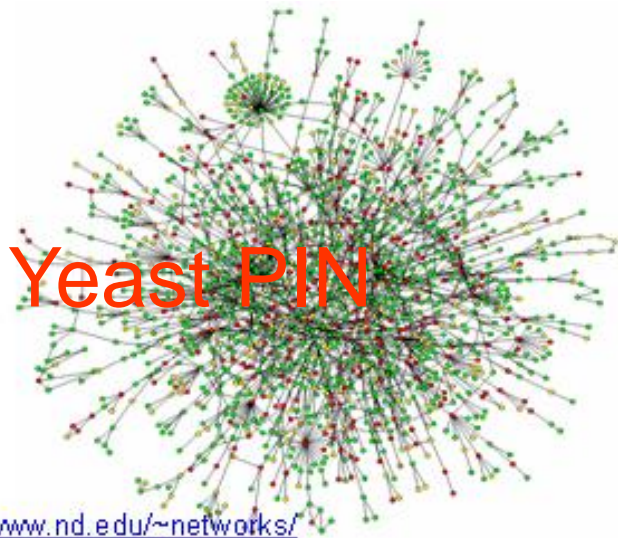
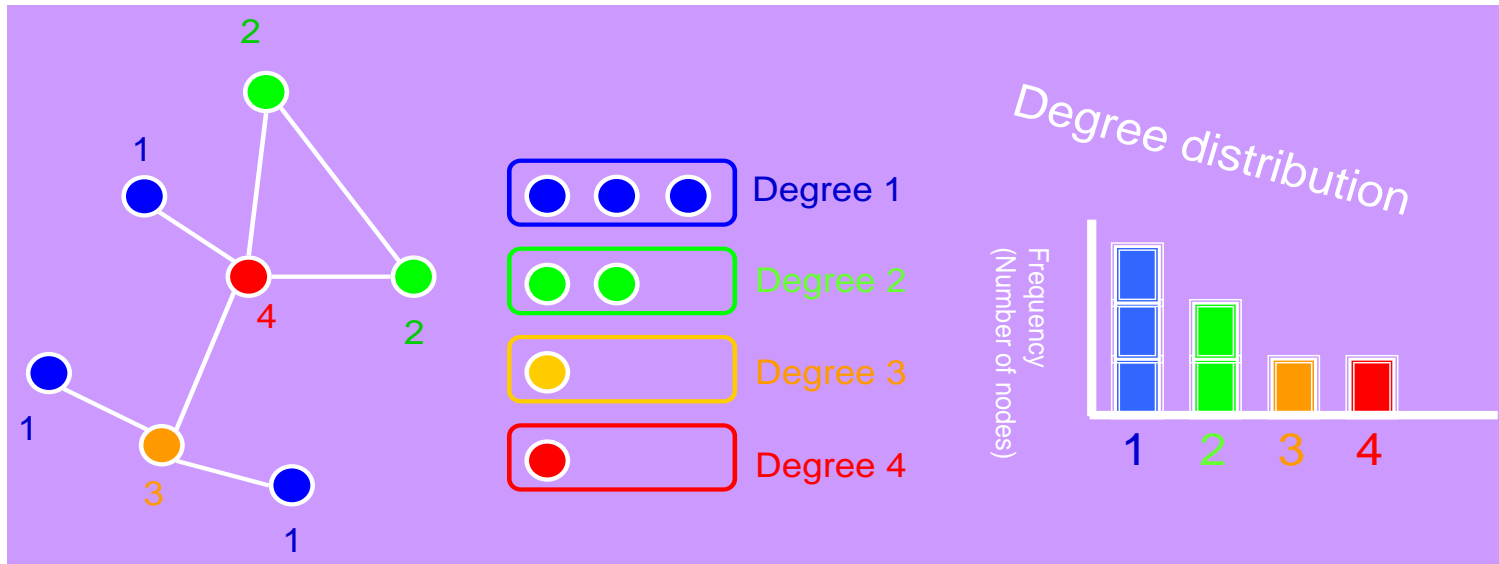
- **疾患ネットワークの十全な形成**
  - 多層的な疾患ネットワークとその統合
  - 医薬品の有効性・毒性近傍Projection  
⇒ DRにおけるfeasibilityは証明
- **創薬への発展**
  - 薬剤・化合物空間のネットワークは既に確立
  - cMapでは不十分・LINCS(2014)
  - 疾患ネットワークがの確立が重要
  - 疾患から逆投影。創薬の可能性探索
- **疾患トポロジーと薬剤トポロジーの双対性**
  - Dual Topology-based Drug Discovery

疾病から薬剤ネットワークへの逆投影  
Double Topology 双投影 創薬方法論

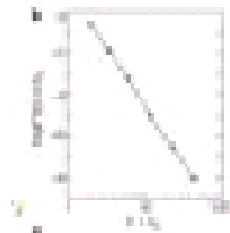
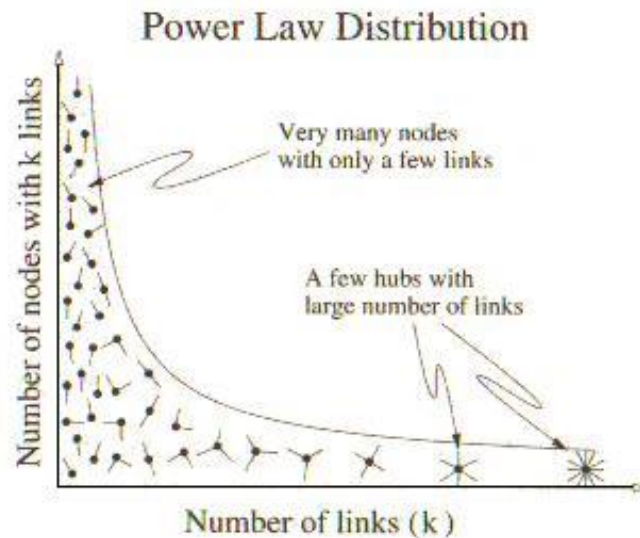


# タンパク質相互作用ネットワーク と創薬

タンパク質相互作用ネットワーク(PIN)では数少ない相互作用が集中したタンパク質(hub)と相互作用が1や2の多数の末端タンパク質(branch)が存在する



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

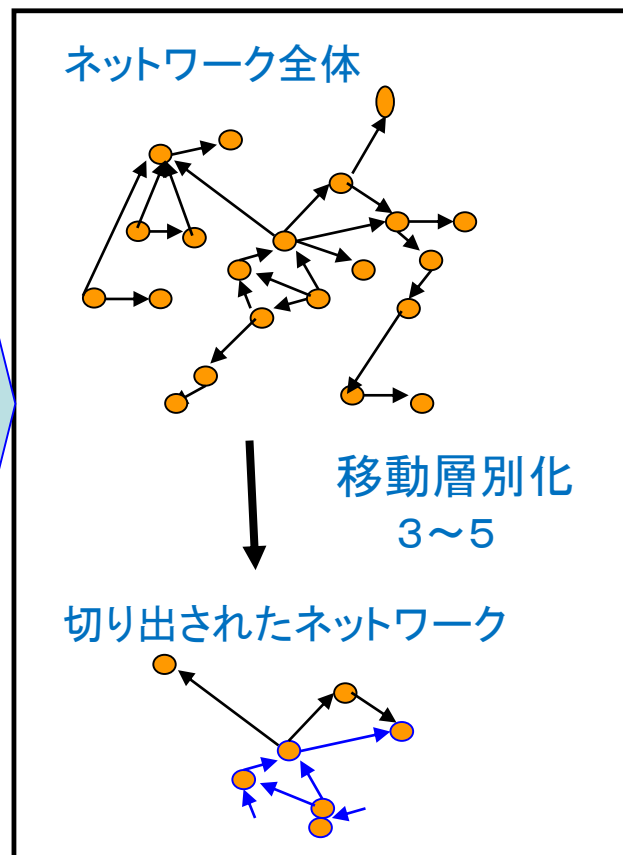
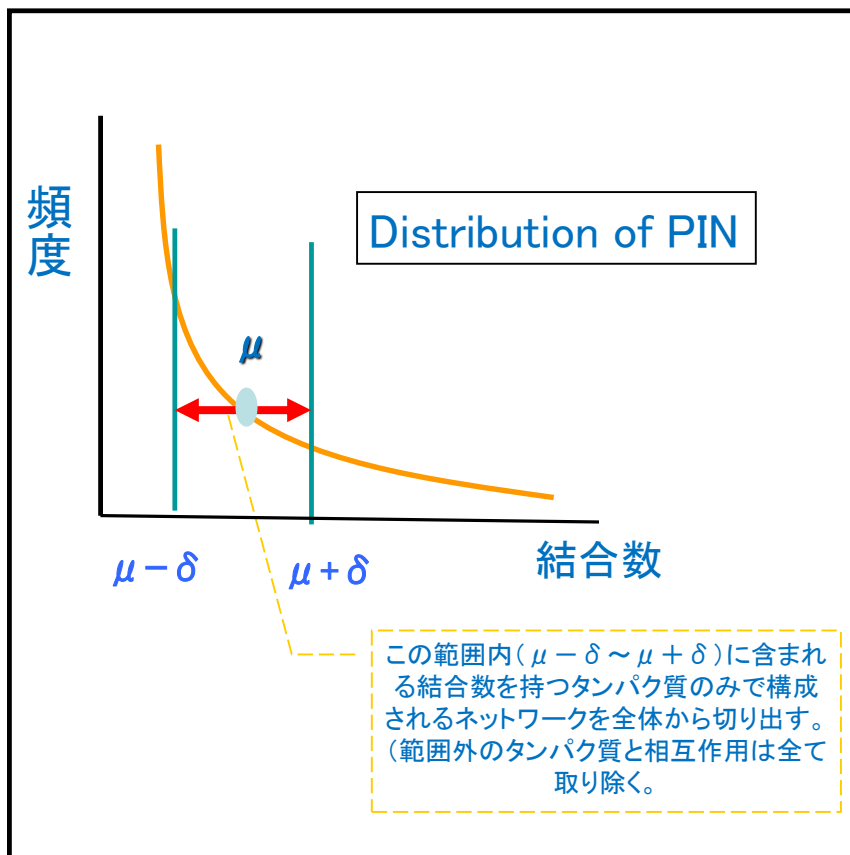


両Logで直線



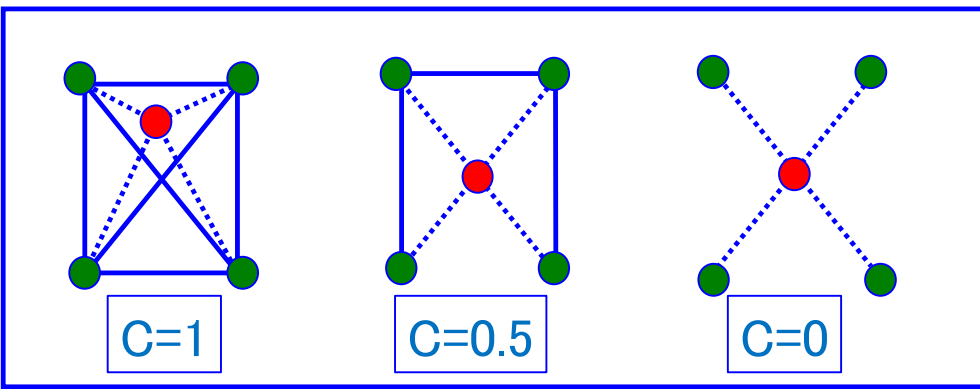
<http://www.macs.hw.ac.uk/~pdw/topology/ScaleFree.html>

# 結合次数ごとの層別化ネットワーク構造の解析



## クラスター係数

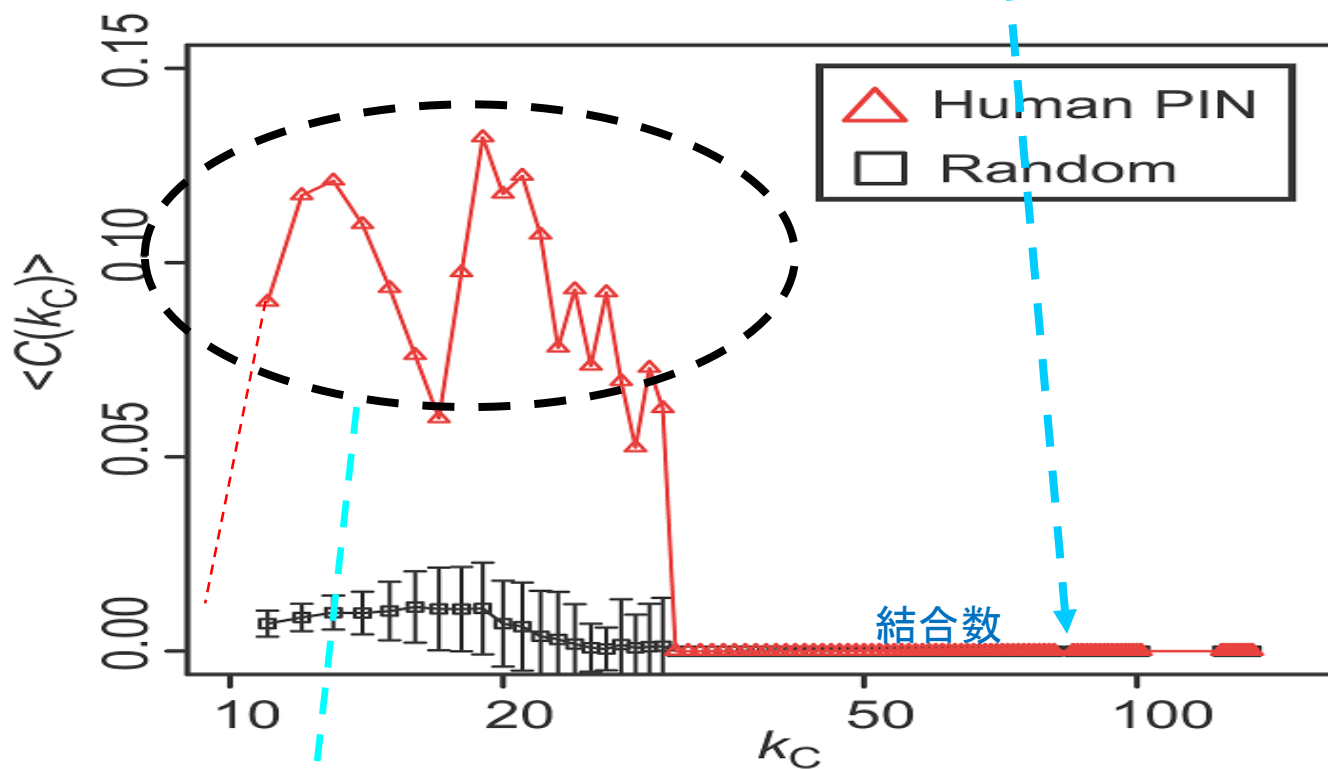
Hase, T., Tanaka, H et.al (2009) Structures of protein protein interaction network and their implications on drug



design. *PLoS Comput Biol.* 5(10):

# 移動層別化によるクラスター係数

高いConnection Degreeを持つノード(スーパーハブ)は  
お互いに密に結合しない

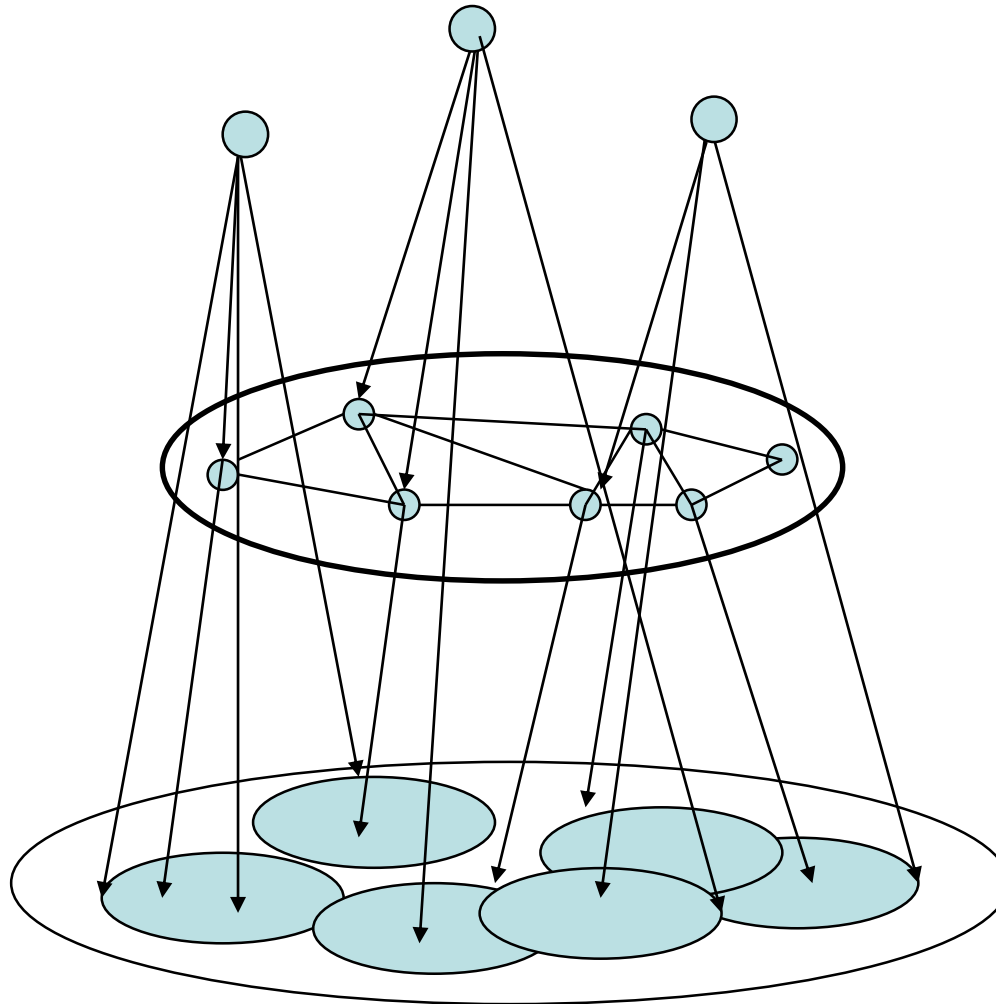


中程度の結合数(7~42)を持つタンパク質は、  
多数の密なモジュールを構成している。



タンパク質相互作用から見られる

# 生命情報ネットワークの構造



高層

高デGREEハブ

度数

> 31 ヒト

> 39 酵母

中間層

中程度デGREE

度数

6 ~ 30 ヒト

6 ~ 38 酵母

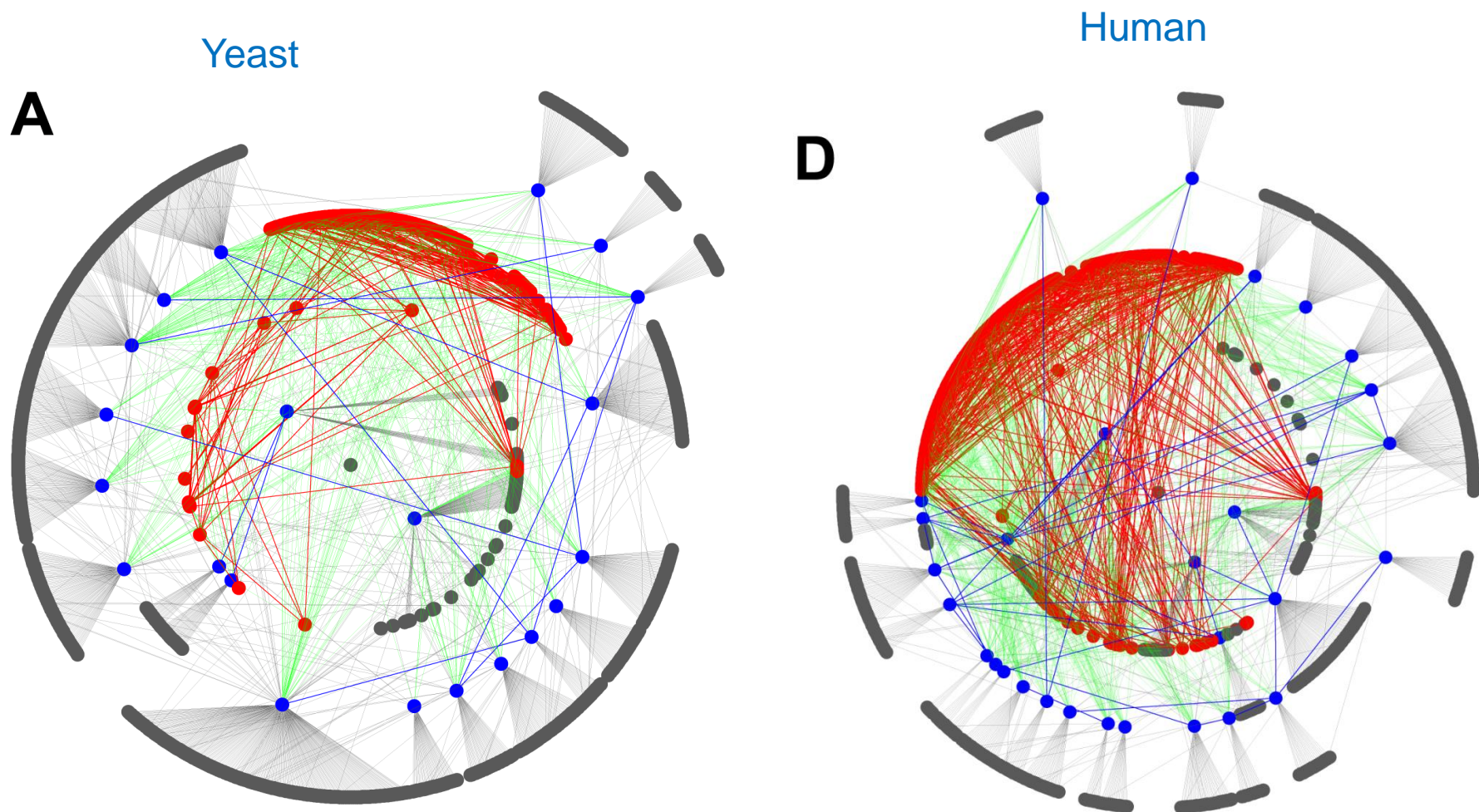
低層

低デGREE

ブランチ

度数 < 6

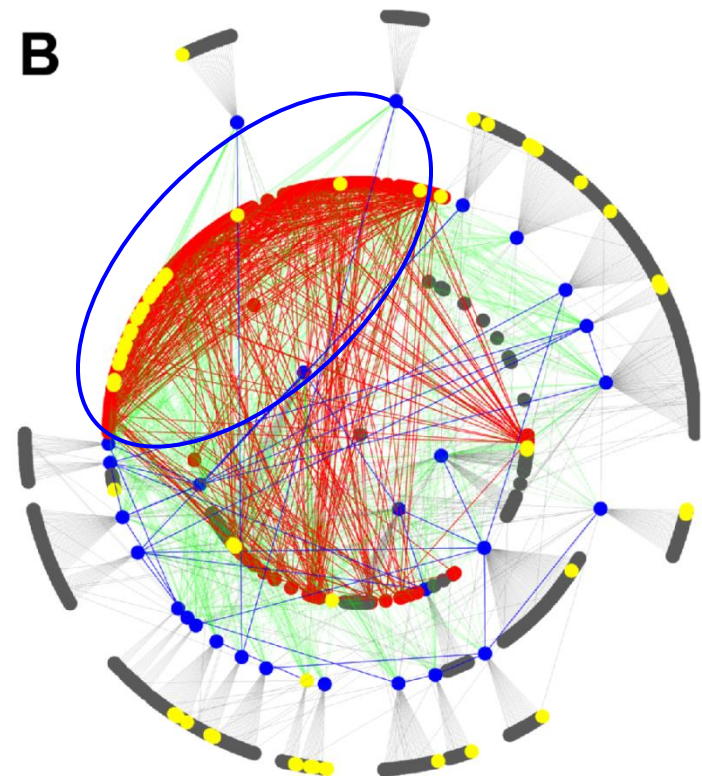
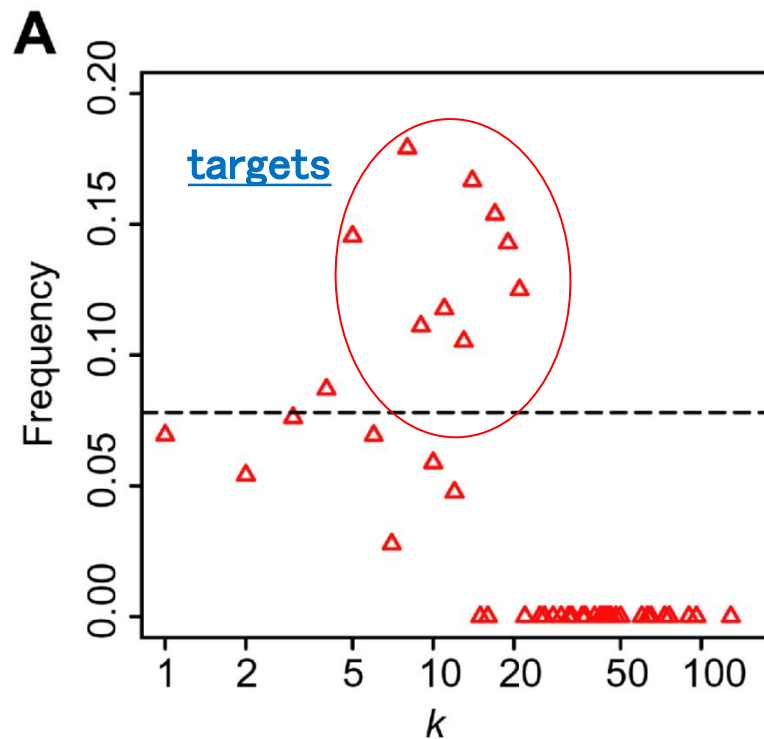
# PINのCloud Topology (3環トポロジー)



**Middle-degree ノードは PIN backbone を形成する。**

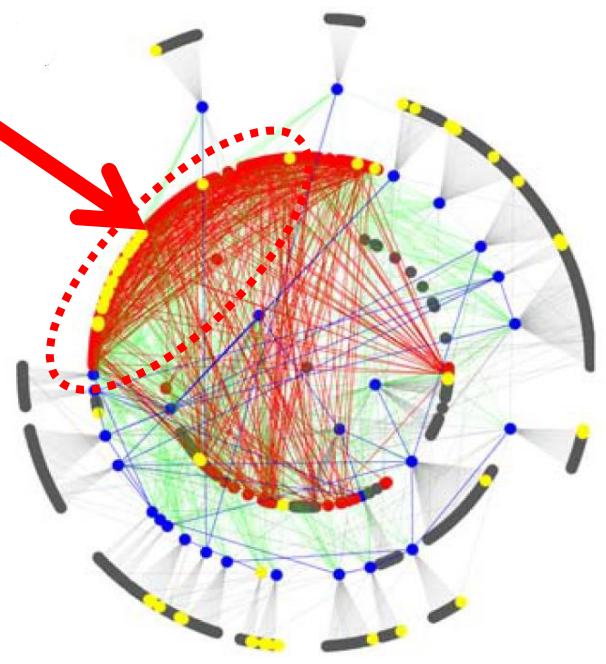
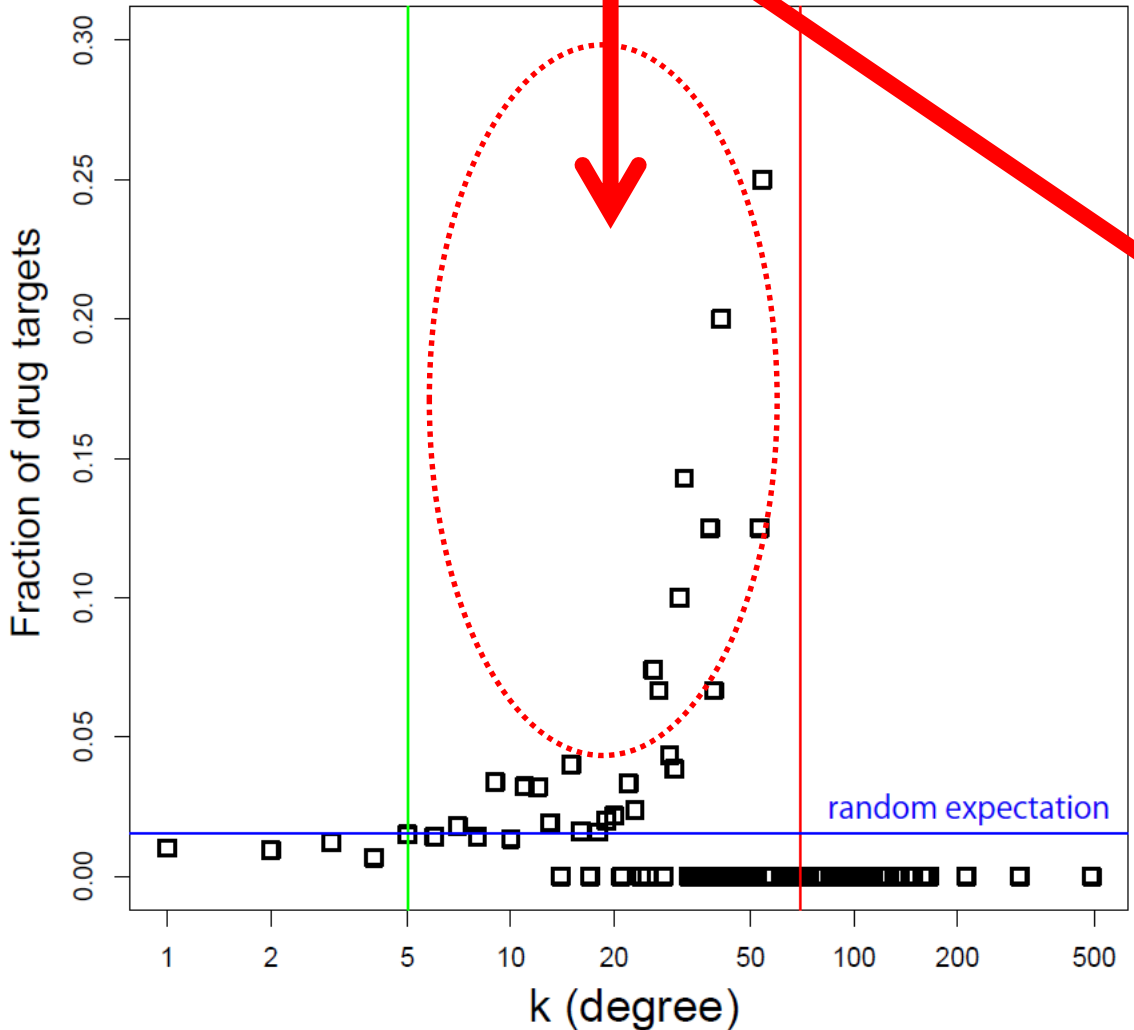
灰色, 赤, 青は、それぞれ低層、中層、高層のdegreeのノードをそれぞれ表す。

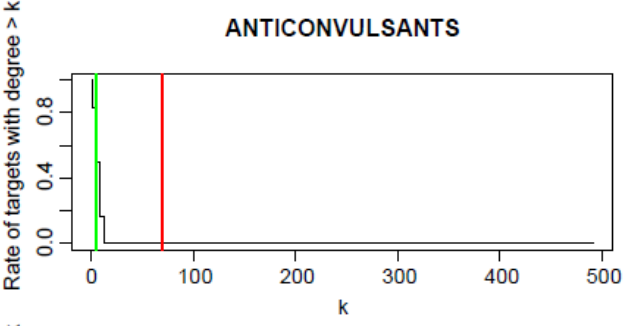
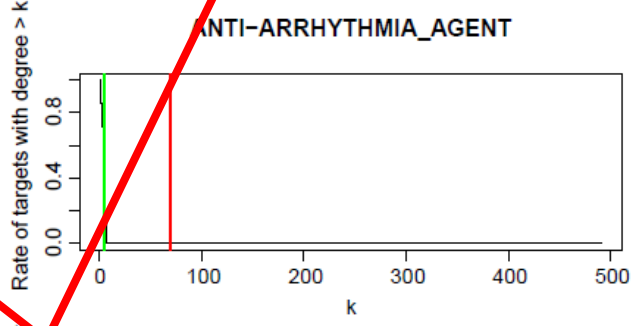
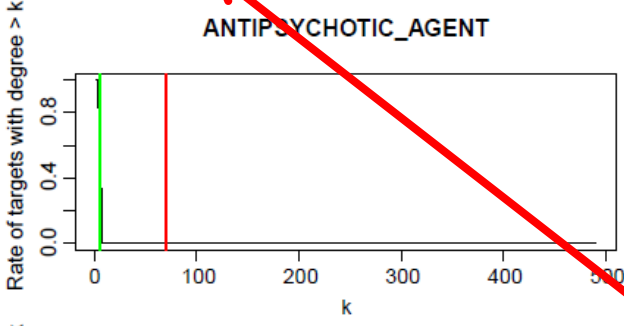
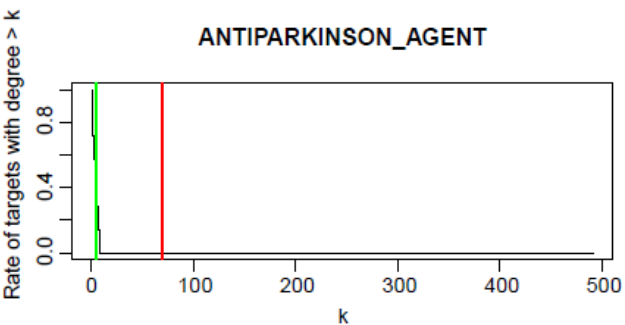
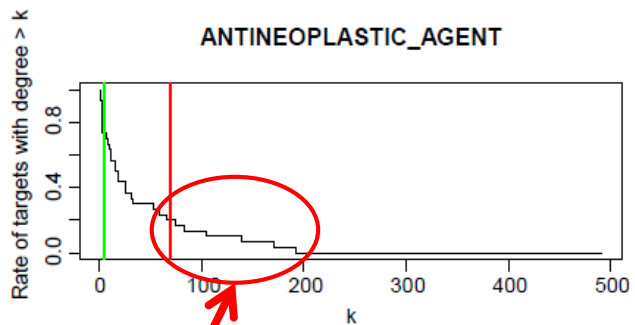
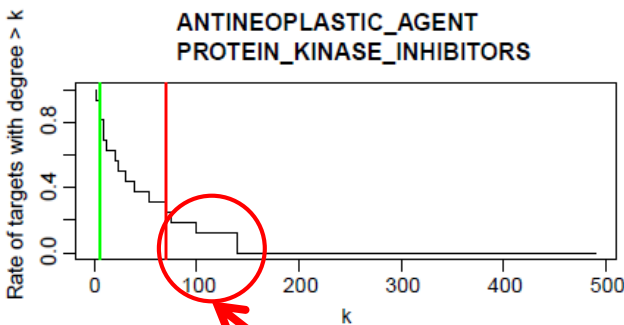
# 薬剤標的分子と結合度数



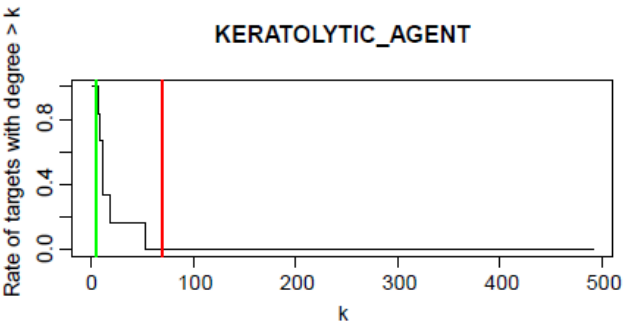
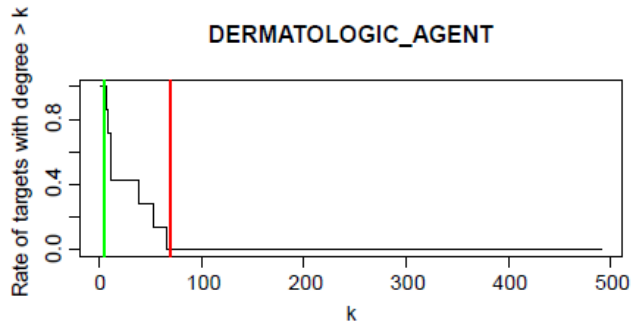
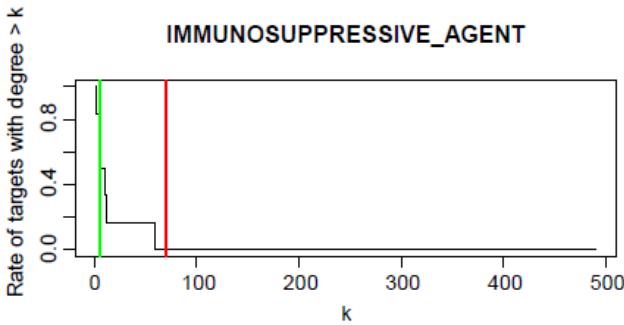
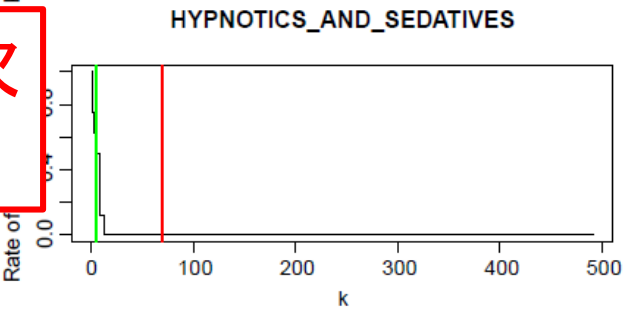
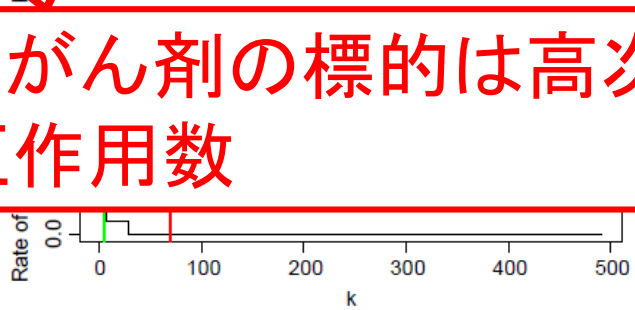
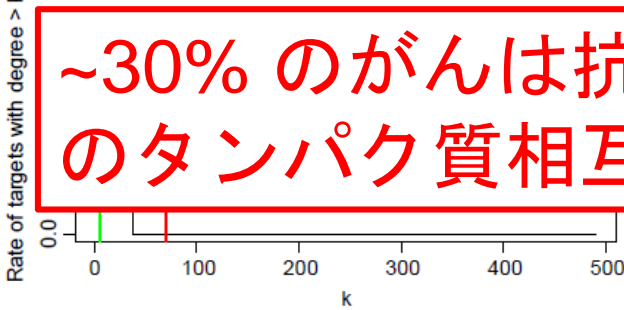
中層レベルのノードは治療薬として最適な標的である。それゆえ、多くの市場にある薬剤標的は、ヒトのバックボーンタンパク質に集中している

**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.

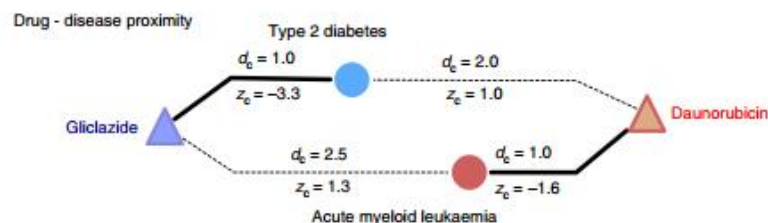
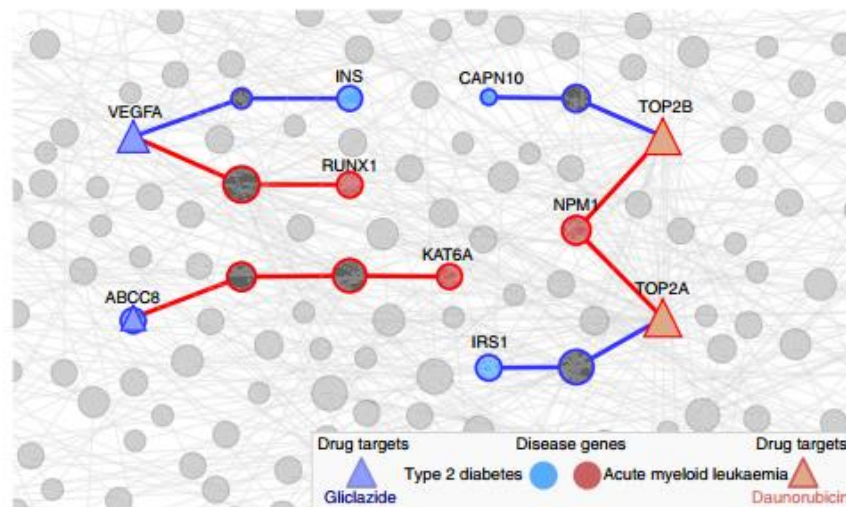
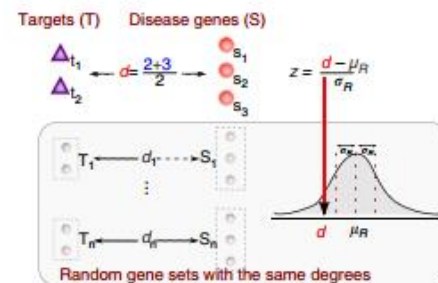
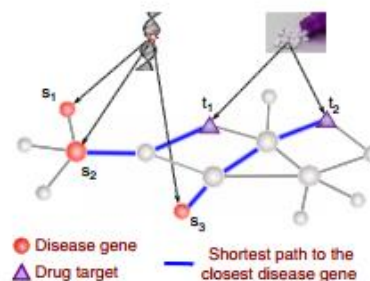
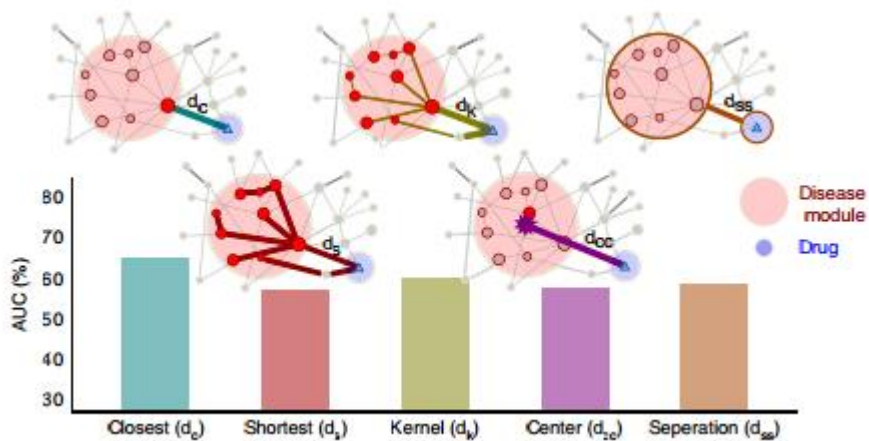




**~30% のがんは抗がん剤の標的は高次のタンパク質相互作用数**



# タンパク質相互作用ネットワークでの 近隣性によるDR



(Guney2016, Nat. Com)

# AI医療・AI創薬

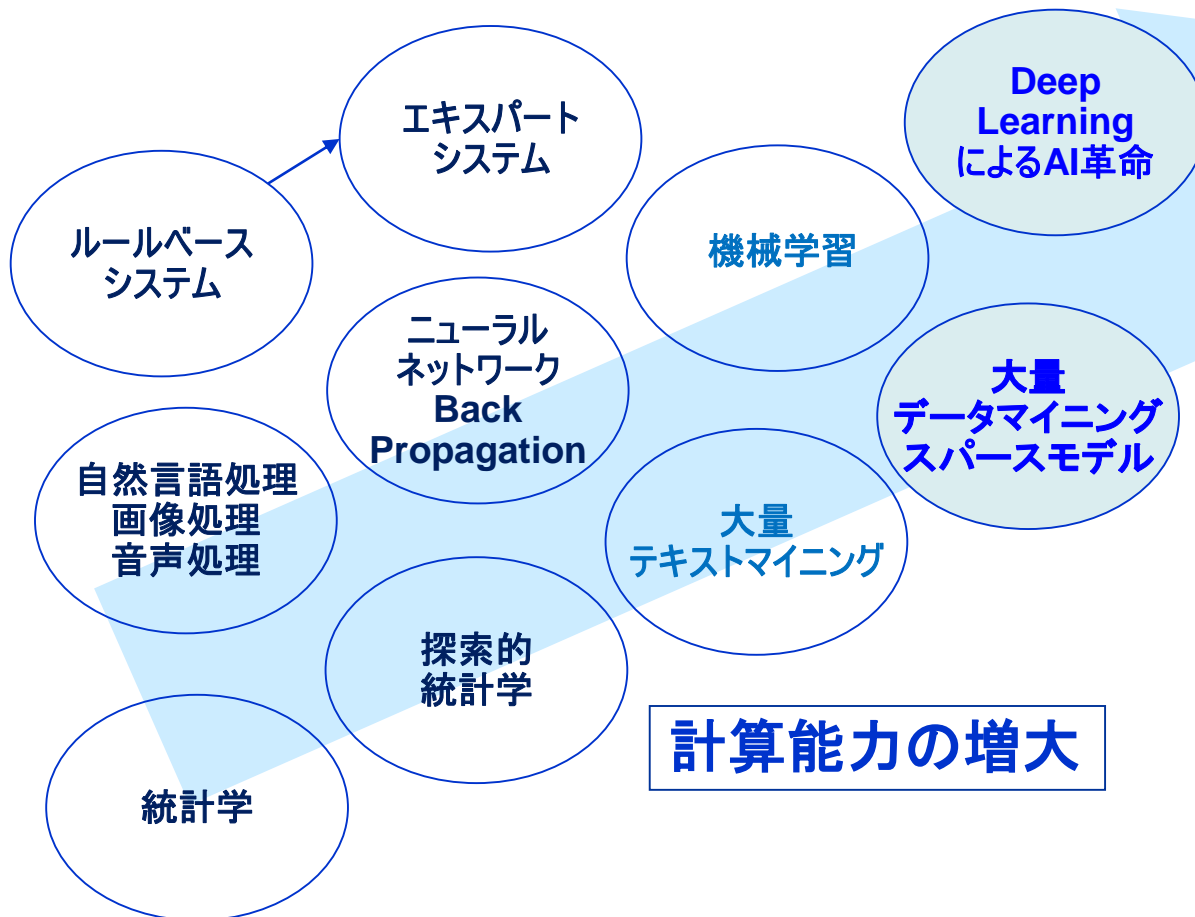
## —人工知能への期待—

# 人工知能への期待

人工知能 (AI) の分野

データの増大

ビッグデータ  
人工知能による  
知的処理



計算能力の増大



# 医療分野の人工知能の歴史

記号（シンボル）的知識処理

ニューロネットワーク処理

1970

問題解決の一般探索手法 **GPS**  
解決木の高速探索（ゲーム）

ニューロネットワーク  
3層の学習機械 **Perceptron**  
入力層、隠れ層、出力層

1980

推論システム（if-thenルールシステム）  
知識の表現と利用（専門家システム）  
医療診断システム（Mycin, Internist-I）  
大ブーム 医療から産業応用の期待波及

多層型ニューロネット  
後方伝播 **Back Propagation**  
結合係数修正アルゴリズム

1990

期待消滅！

知識発見 機械学習  
Machine Learning, KDD  
診断知識のDBからの学習

しばらく停滞！

2000

知識準拠診療支援（DSS）  
医療ターミノロジー  
医療オントロジー

ニューロネットワーク型  
多層型ニューロネット  
**深層学習 Deep Learning**  
結合係数修正アルゴリズム  
画像処理から創薬まで



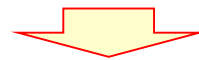
# 「ビッグデータ」のData 原理

問題点 属性値数(p) ≫ サンプル数(n)

p: 数億になる場合あり    n: 多くても数万、通常数千



これら膨大な属性変数がすべて独立ならばビッグデータの構造解析は不可能。単変量解析の羅列 (GWASのManhattan Plot) しか可能でない



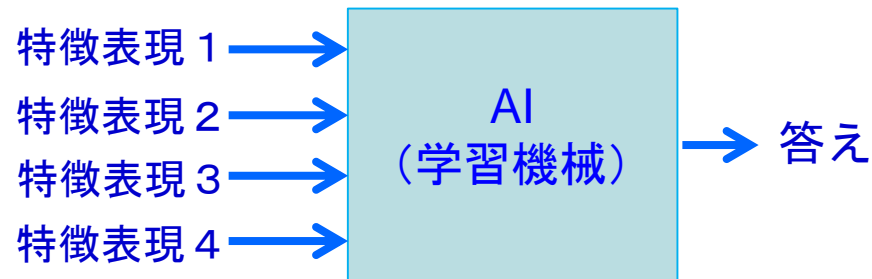
**ビッグデータ・スパース仮説**

ビッグデータは、多数であるが属性値数より少ない独立成分が基底となって、相互にModificationして構成されている。  
(独立成分の推定は、サンプル数とともに増加する)

データ次元縮約の原理 (**principle of compositionality**)

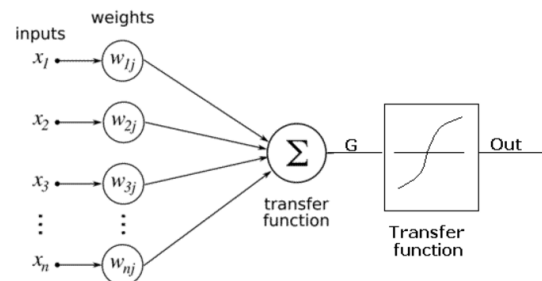
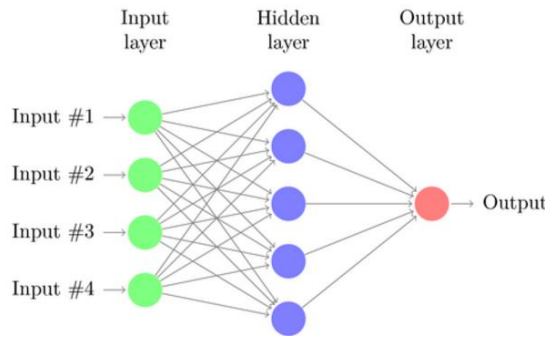
# Deep Learning による 人工知能革命

- 機械学習のこれまでの限界
  - 分類・判別する学習機械（システム）
  - 対象の特徴表現ベクトルを与えて分類
  - 与え方に関して細かな技法にとらわれる
- 「教師あり学習」
  - 分類対象の特徴と正解を与え学習機械（AI）を構築
  - 対象の表現(画像等)と概念を結合できない

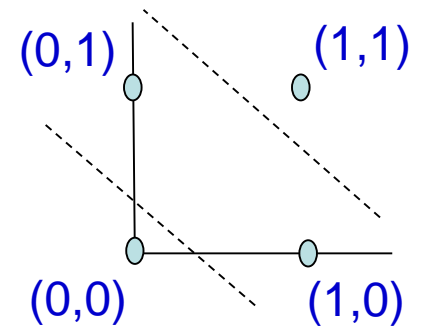


# 従来のニューロネットワーク

## 古典的Neural Network・パーセプトロン(1970年代)

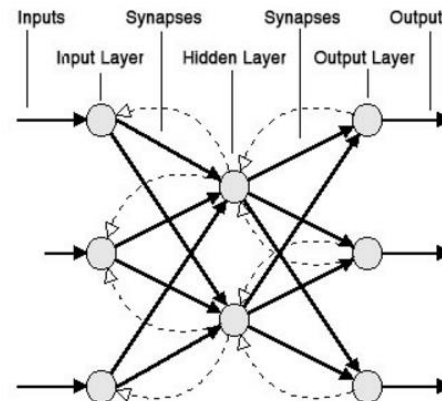
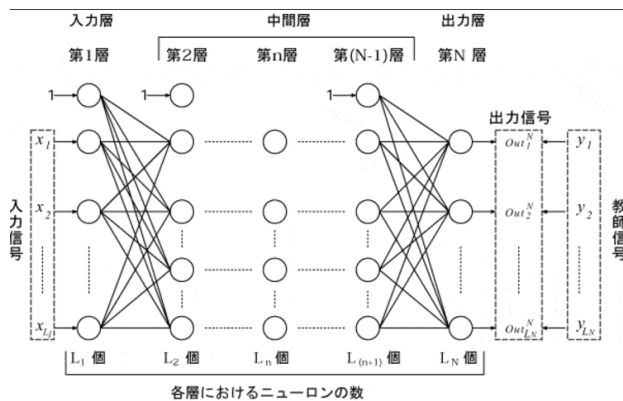


XOR



## 多層Neural NetworkとBack projection (1980年代)

線形分離できない

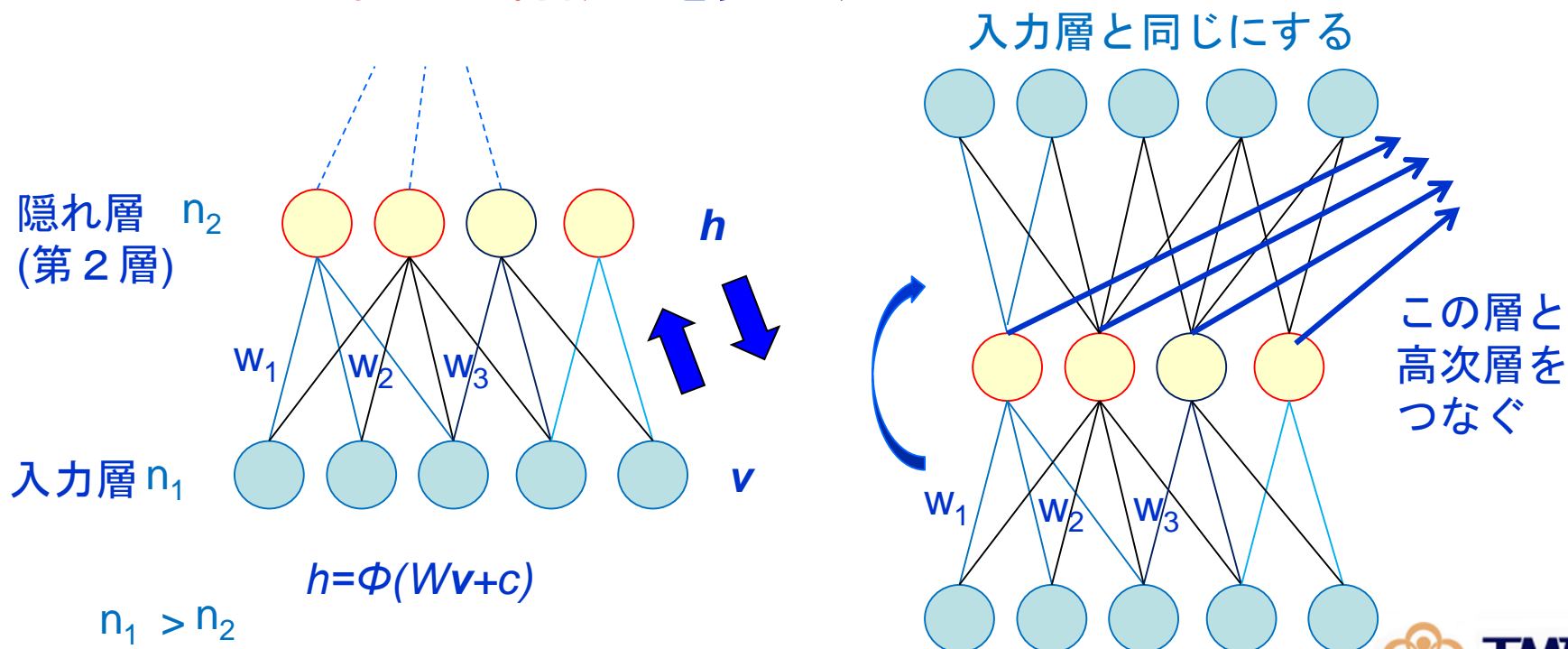


Back Propagation (1986 Rumelhart)  
 望ましい出力との誤差を教師信号として与える事により、次第に結合係数を変化させ、最終的に正しい出力が得られるようにする。結合係数を変える事を学習と呼ぶ。この学習方法には、最急降下法(勾配法)が使われる。出力層へ寄与の高いノードの重みの変更。

多層にわたる逆伝搬で修正感度減衰

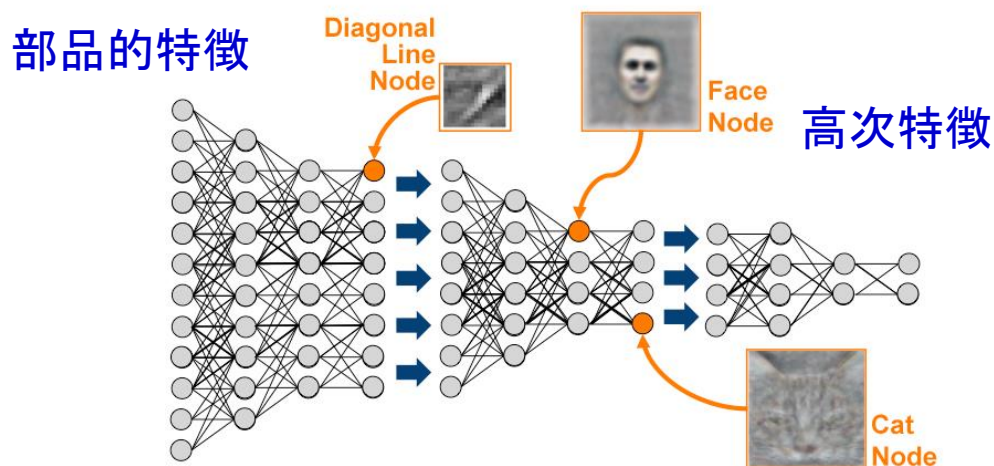
# DLの革命点 Autoencode 1

- 対象に固有な**内在的特徴**を学ぶ自己符号化の原理
- 格段ごとに入力を少ない中間層を介して復元できるかを行なう
- 次元を圧縮されて可及的に復元する
  - できるだけ復元に**効果的な**特徴量を探索する
  - 内在的な特徴量**を見出す



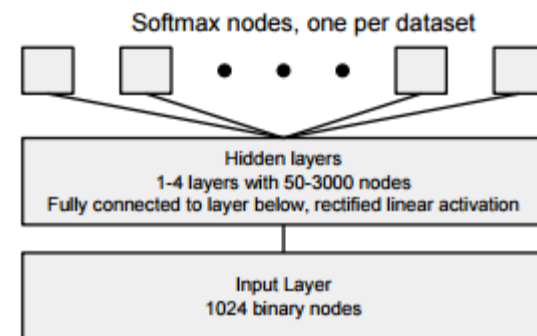
# DLの革命点 Autoencode 2

- 各層ごとに自己符号化を行うので**何層でもネットを組める**→Deep Learning
- 第一層で学習した特徴量を使って、つぎの階層を作るので**高次の特徴量**が作られる
- **特徴的表現**と**概念**を結びつけるため「**教師あり学習**」が最後に必要である
- **自動特徴抽出**によってこれまでの学習手法の限界を克服した（構造的理解）



# Deep learning : 創薬からの注目

- 創薬を巡る状況
  - 平均14年、約2000億円 (\$1.7 B) の費用
  - 市場化された新薬の減少
  - 創薬に費やす期間・コストを低減したい
- Kaggle (データサイエンス競技会)にMerck社が出題  
Molecular Activity Challenge (2012).
  - 15データセットから異なった分子の生物学的活動を予測するモデルの開発コンテスト
  - 勝利したモデルは深層学習 deep learning を用いたモデル
- Google in collaboration with Stanford (2015)
  - Stanford 大学の Pande 研究室と共同研究  
バーチャルドラッグスクリーニングに対する  
deep learningによるツール開発  
"Massively Multitask Networks for Drug  
Discovery"



# Artificial Intelligenceと創薬

- 標的分子選択と妥当性検証
  - 適切な分子標的の選択
- Virtual screening と選択 ←
  - 適切な化合物に対するクラス判定
  - 研究例：ChEMBLに対するdeep learning
    - 13 M 化合物特徴量 (ECFP12), 1.3M 化合物, 5k 薬剤標的
    - Ligand-based 標的予測, 7種の予測法とAUC比較
    - Deep learning: SVM, k-nearest nb, logistic回帰より優位
    - DLで構造活性相関を学習する
    - 特徴量の抽出、薬理機序への理解
  - リード最適化
- システム薬理学
  - ネットワーク病態学よりの創薬戦略
  - 他のシステムへの影響(毒性, 副作用)

## Pharmacophoreの抽出

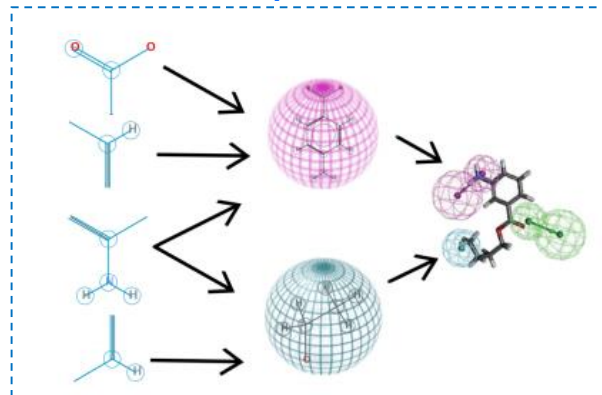


Figure . Hierarchical nature of fingerprint features: by combining the ECFP features we can build reactive centers. By pooling specific reactive centers together we obtain a pharmacophore that encodes a specific pharmacological effect.



# そのほかのAI創薬の話題

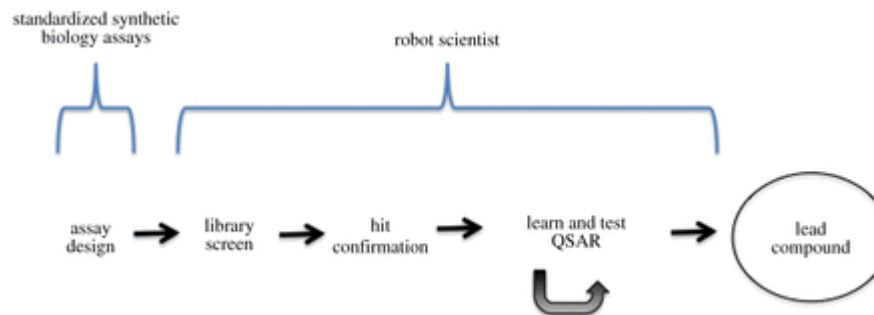
- Berg社のAI創薬
  - AIを方法として膵臓がんの抗がん剤を開発中
  - 膵臓がんと非患者の14兆のゲノム・オミックス情報を比較。
  - 調節不全パスウェイのシステム推定
  - システム薬理学的AIによる創薬
- マンチェスター大学（Cambridgeとも共同）

## Artificially-intelligent Robot Scientist for new drugs

- ライブラリースクリーニング, ヒット化合物の確証, リード化合物などの自動化
- 構造活性相関（Quantitative Structure Activity Relationship (QSAR)）を反復学習する
- 熱帯病、寄生体のDHFR（ジヒドロ葉酸還元酵素：薬剤耐性）を標的にして学習、細胞を合成生物学操作  
血管新生阻害因子（抗がん剤）をDR候補を探索
- 最上位にコンセプト木（“root: assay triple screen”など）



Robot scientist Eve at work



ご清聴有難う御座います

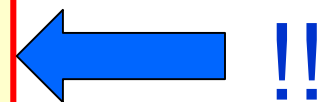


# 參考資料

# 医療ビッグデータ時代の到来

- (1) 次世代シーケンサなどによる「ゲノム/オミックス医療」による網羅的分子情報蓄積
- (2) モバイルヘルス(mHealth) によるWearable センサ情報の継続的蓄積 (unobstructed monitoring)
- (3) Biobankによるゲノム・コホート情報

大量データの急激な  
コストレス化かつ高精度化



ゲノム : 13年→1日(1/5000) 3500億→10万円(1/350万)

個別化医療・予測医療  
健康・医療の適確性の飛躍的な増大



# 医療の「ビッグデータ革命」

## ～何が新しいのか～

### 1) 臨床診療情報

- 従来型の医療情報
  - 臨床検査、医用画像、処方、レセプトなど

### 2) 社会医学情報

- 従来型の社会医学情報
  - 疫学情報・集団単位での疾患罹患情報

### 3) 新しい種類の医療ビッグデータ

- 網羅的分子情報・個別化医療
  - ゲノム・オミックス医療
  - システム分子医学・Precision Medicine
- 生涯型モバイル健康管理 (mHealth)
  - ウェアラブル・生体センシング

旧来のタイプの  
医療データの  
大容量化

新しいタイプの  
医療ビッグデータ

# 医療の「ビッグデータ革命」

～ゲノム・オミックスデータの基軸的な特徴～

＜目的もデータ特性も従来型と違う＞

従来の医療情報の「ビッグデータ」

**Big “Small Data”** ( $n \gg p$ )

医療情報・疫学調査では属性数：10項目程度

— 目的：Population MedicineのBig Data

⇒個別を集めて「集合的法則」を見る

網羅的分子情報などのビッグデータ

**Small “Big Data”** ( $p \gg n$ )

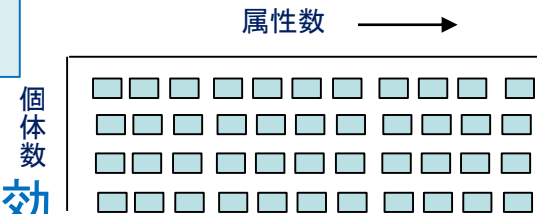
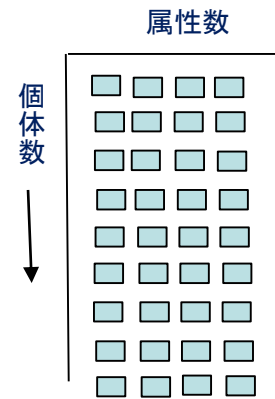
1個体に関するデータ属性種類数が膨大

属性に比べて個体数 少数:従来の統計学が無効

「新NP問題」：多変量解析:GWASで単変量解析の羅列

— 目的：例えば医療の場合Personalized Medicine

⇒大量データを集めて「個別化パターン」の多様性を抽出



新しいデータ科学の必要性

# 医療の「ビッグデータ」革命は どんな既存のパラダイムに挑戦しているか

- Population medicineのパラダイム転換
  - <One size fits for all>のPopulation医療はもはや成り立たない
  - 個別化医療 “Personalized (Precision) medicine”
    - 個別化医療を実現するために<個別化・層別化パターン>を網羅的に調べる：どこまでの粒度で個別化・層別化すればよいか
- Clinical research（臨床研究）のパラダイム転換
  - 臨床研究を科学にする従来の範型RCTは、個別化概念に破綻した
  - <statistical evidence based>呪縛からの解放
  - 「標本」統計・「推測」統計学に限定されない臨床研究
  - Real World Data: ビッグデータ知識生成（BD2K）
- 創薬の戦略パラダイムの転換
  - ビッグデータ創薬の可能性
  - 網羅的分子データからの計算機創薬・システム創薬
  - Transdisease Omics, Drug networkのDual Network Topologyによる創薬

# 次世代シーケンサのインパクト

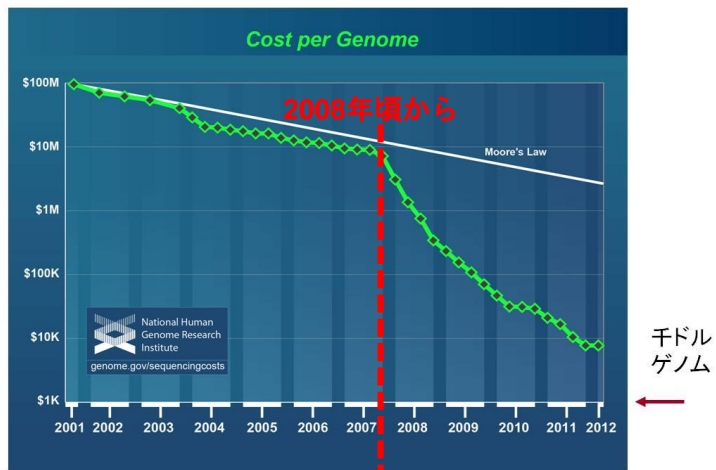
## 次世代シーケンサを始めとするhigh-throughput分子情報収集の急激な発展

急速な高速化と廉価化 ヒトゲノム解読計画13年,3500億円⇒1日,10万円

2005~ NGS 454 (LS,Roche)  
2007/8~454, Solexa (Illumina),  
SOLiD (LT,TF)  
**シーケンス革命**



	HiSeq2500		Ion Proton
本体価格	約1億円		約3500万円
モード / チップ	ハイアウトプット	ラピッドラン	Ion Proton I
解析時間	11日	27時間	2時間
リード長 (bp)	2 x 100	2 x 150	200
データ産出量 (Gb)	約600	約120	10
試薬コスト (ヒト1人全ゲノム)	数十万円		不可 エクソームのみ



HiSeq X システム 10台構成 (経費1/5)



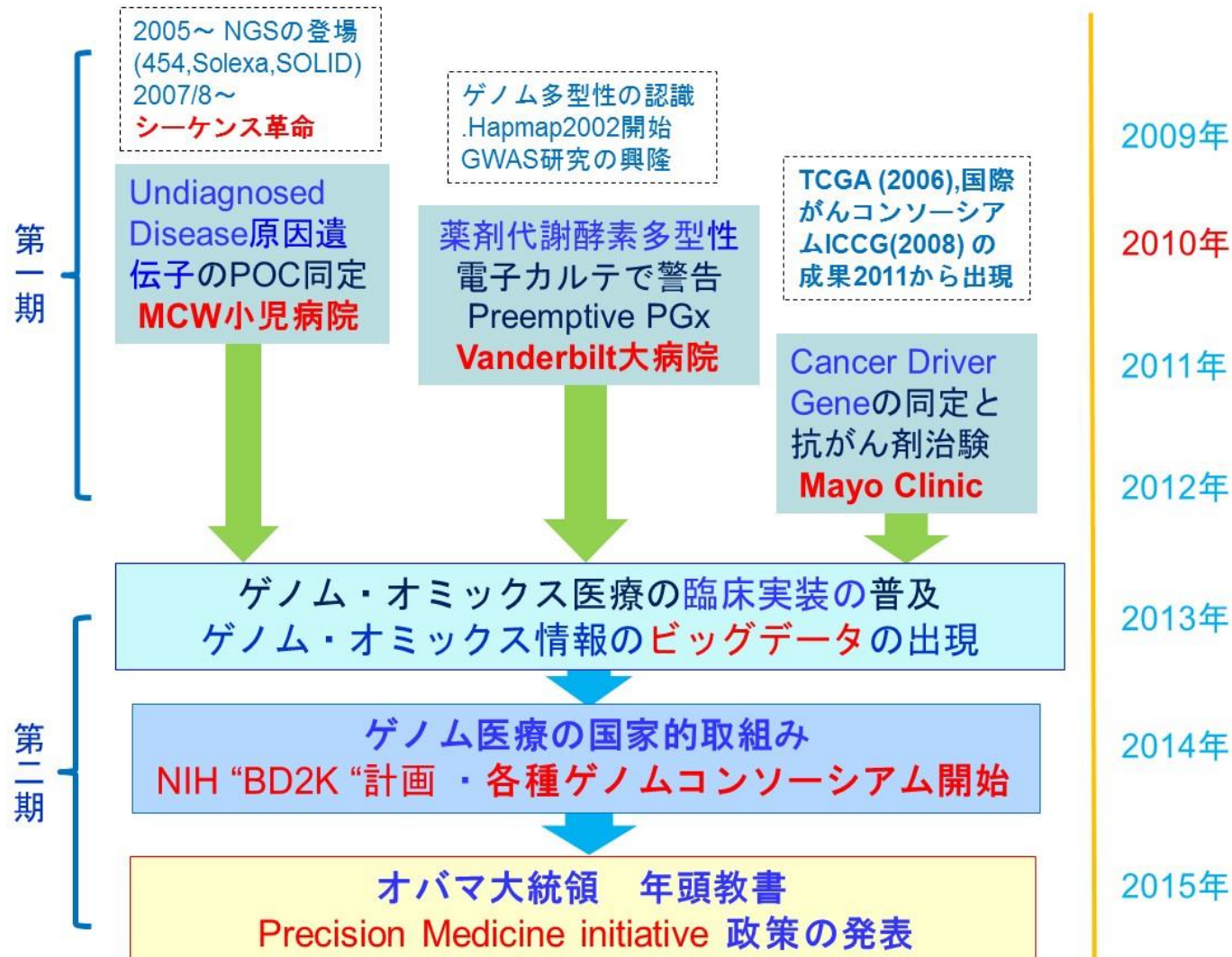
DNA Sequencing Cost: the National Human Genome Research Institute

**シーケンス革命 2007/8**

ゲノム(配列決定)機器の進歩は、計算機のムーアの法則を越えている!



# ゲノム・オミックス医療の流れ



# 医療ビッグデータ時代の到来（米国）

ゲノム医療の実践

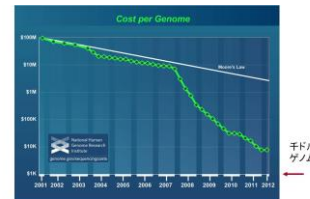
## 第1段階 ゲノム医療の発展

次世代シーケンシングの臨床普及 (2010~)

全ゲノム (X30 : 100Gb)・エキソーム解析 (X100 : 6Gb)

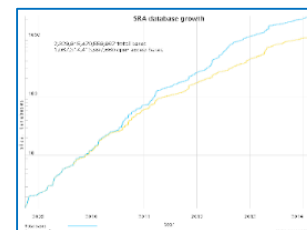
米国では数十の著名病院で実施

ゲノム・オミックス情報の蓄積



DNA Sequencing Cost: the National Human Genome Research Institute

2000兆塩基 (2 Pb)  
が登録(NCBI:SRA)



## 第2段階 医療ビッグデータ時代

医療情報との統合

電子カルテからの  
臨床フェノタイプ

医療ビッグデータ

学習アルゴリズム

ゲノム医療知識

人工知能AI



MayoClinicでは  
10万人患者WGS

医療ビッグデータ

# Precision Medicine

個人の遺伝素因・環境素因に合わせた (tailored) 医療  
One size fits for all のPopulation 医療とは異なる

趣旨：基本は、個別化医療 Personalized Medicine の概念と変わらないが、目指していたのは診断/治療の個人化ではなく層別化であることを明確化

概念の拡張：Personalized Medicineが標榜された時から10数年経っている

## 医療ビッグデータ時代の到来による個別化医療の拡張

- (1) 遺伝素因 X 環境(生活習慣)要因のスキーマ重視  
SNPや変異 (Genome)だけでなく環境・生活習慣要因(Exposome) の重視、疾患発症は2つの要因の相互作用を明快に強調。電子カルテの臨床表現型 (Clinical Phenome)も疾患発症後には不可欠。3つの成因の重視
- (2) 日常生理モニタリング情報の包摂  
モバイルヘルス(mHealth)・wearable sensorによる大量継続情報収集の重視
- (3) ゲノムコホート・Biobankの重視  
Precision Medicineを実現する基礎として、ゲノムコホート/Biobankが必要であることを認識。Real world dataの重視

# Biobank/ゲノムコホートへの期待

- 疾患型バイオバンク/ゲノムコホート
  - 個別化医療の層別化パターンの網羅的摘出
  - 病院ゲノム・オミックス医療DBを補う
- Population型（健常者）コホート
  - (1) 前向きコホートの長所により発症要因同定  
疾患発症相対リスク 「個別化予防」  
＝遺伝子要因 × 環境生活習慣要因  
上記の相互作用を評価 (exposome, expotype)
  - (2) 「健康から疾患発症に至る過程」を多数収集  
「先制医療受攻状態」 (vulnerable period) 同定  
⇒ 先制医療薬の開発, QOL・医療経済的にも良策
  - (3) 慢性疾患患者のコホート  
⇒ GWASが可能、重症化・合併症のリスク因子

# Biobankとゲノムコホート

## • バイオバンクの目的・機能の変化

- 従来は再生医療のための生体標本や臨床研究の資料保存、近年はゲノム医療の基盤としての役割
- **ゲノム/オミックス個別化医療、創薬の情報基盤**
  - **疾患型BioBank**：全国的・全世界規模で疾患罹患患者の網羅的分子情報（ゲノムなど）とそれに対応する臨床表現型（臨床検査、医用画像、処方歴、手術歴、病態経過、転帰など）の収集。**疾患ゲノムコホート**
- **個別化予防の情報基盤**
  - **Population型BioBank**：「健常者」前向きコホート。調査開始時の網羅的分子情報（ゲノム）と臨床環境情報（exposome）を集めて、生涯を追跡するゲノム・コホート

## • 欧米のBiobank

- **英国 UK biobank**
  - 50万人の健常者。40～69歳（2006-2010, 62Mポンド), 2011-16, 25Mポンド
  - 健診データ（血液・尿・唾液サンプル、生活情報）を集め、健康医療状況を追跡する。
- **英国 Genomics England,**
  - 2013開始、2017年までに 10万人のゲノム 配列収集。
  - 最初の対象は稀少疾患（患者・家族）、がん患者、最初はEnglandのみ
- **欧州 BBMRI (Biobank/Biomole. Res. Infra.)**
  - 250以上の欧州各国のBioBankを統合
- **オランダ Lifeline**
  - 165000人北部オランダ 2006年開始 30年間の追跡、3世代コホート（世界初）
- **Precision Medicine Initiative Genome Cohort**
  - 100万人のゲノムを集める

# 「学習する医療システム」 Learning Health System

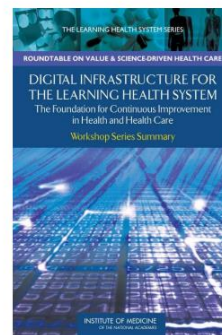
新しい生物医学知識が臨床実践に給されるまで17年  
臨床データを用いて医療を実施しながら医療を改善

- IOM “Clinical Data as a Basic Staple of Health Learning”
- 医療システムのデジタル化（IT化）は必然の傾向である
- 「ルーチンの医療活動から集められたデータ（形式的臨床研究と違って）がLHSを支える鍵である」
- データを共有することによって学習して医療システムを改善
- RCTは「黄金基準」であるが、通常の医療システムの外で実施されている。医療が実際対象とする患者集団を代表しているのか。
- RCTは時間が掛かり費用もかかる
- 有効な知識の蓄積の速度が加速する

IOM(Institute of Medicine)のレポート  
2007年にEBM/RCT（無作為試験）に  
変わるパラダイムとして提案

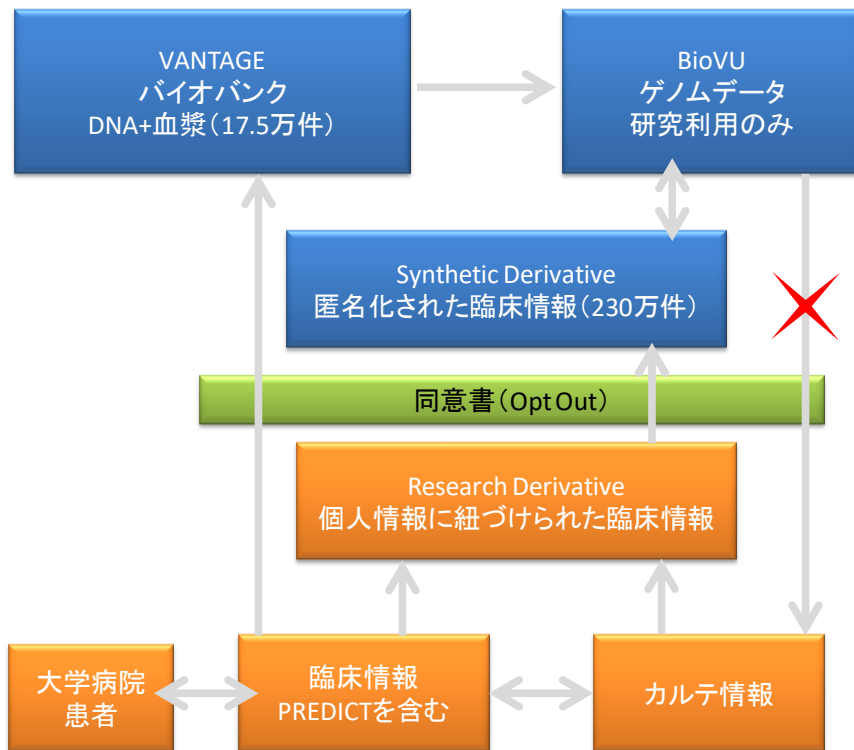
*Digital Infrastructure for the Learning Health System: The Foundation for Continuous Improvement in Health and Health Care*

*Best Care at Lower Cost: The Path to Continuously Learning Health Care in America*



# LHSの代表例 BioVU

## ゲノム情報と電子カルテ情報を用いた Vanderbilt大学病院の医療情報システム



### 電子カルテ

**Synthetic Derivative** : 電子カルテから匿名化臨床表現型のデータベース 230万件。Opt out 形式

### バイオバンクと遺伝子解析

**BioVU** : Synthetic Derivativeと連結可能な Genome DNA情報

**VANTAGE Core** : 検体17.5万件、血液検からDNA抽出・ゲノム解析、バイオバンク運営

**PREDICT** : 臨床レベルの遺伝子解析情報により、薬物副作用防止などを実現するシステムを自らの医療システムにより知識抽出して実現する

クロビドグレル（抗血栓剤）の遺伝子多型に関してABCB1, CYP2C19、さらにPON1の多型が知られていたが、ヒトを対象とした臨床実験の報告はなかった。SDから循環器疾患で clopidogrelの投与歴の対象者（ケース群）およびコントロール群を選出。BioVUから遺伝型を決定する。この条件に合致するケース群は255件。解析の結果、CYP2C19\*2とABCB1の関与は有意。PON1は非有意が判明した。

# 個別化（層別化）医療の概念の普及とRCTの限界

- 個別化・層別化の概念の浸透
- RCTの治験集団とReal World Dataの乖離
  - 全ての個別化パターンを包摂した治験集団は現実には不可能
  - 現在の治験集団
    - 大半のRCTは医療現実の外の「人工的な環境」
    - 高齢者・妊婦はいない、欧米では黒人とくに青年は含まれない
- 将来へ向けたプラットフォームの確立
  - 母集団に近いReal World 医療データが収集可能
    - ⇒データの大規模化の「相転移」
  - Real World Data時代の臨床研究のプラットフォームを形成。
    - ⇒製造後第3相試験でReal World Dataを使うか
    - RCTとRWDの融合としての
    - registry-based clinical randomized trial
  - 我が国の戦略 段階的移行





# Biobank準拠の創薬・治験

- 疾患レジストリー/疾患型バイオバンク  
準拠型ランダムイズ治験
  - スウェーデンのTASTE(ST-Elevation MI in Scand.)
  - Registry-based randomized clinical trial
  - 疾患レジストリーの登録患者から治験に適した治験対象者を選び
  - 選んだ集団で治験薬・対照薬をランダムイズして割付ける(50\$)
  - 治験のエンドポイントは疾患レジストリーの追跡で観測される
  - 観測研究であるPopulation 型コホートでは困難か

# 次の段階へ：現在のゲノム医療の致命的限界

成功した臨床実装（米国では2010年から今日までこれだけ）

1. 希少先天遺伝疾患の原因遺伝子を病院の現場でシーケンサにより同定
2. がんのドライバー遺伝子変異を同定、適切な分子標的薬を処方
3. 患者の薬剤の代謝酵素の多型性を先制的に同定し、副作用を防ぐ

しかし

多因子疾患の機序/発症予測などには**全く無力**である

- 「単一遺伝的原因」帰着アプローチの限界
- 「行方不明の遺伝力」の主要な原因  
複数の疾患関連遺伝子間の相互作用:  $G \times G$   
環境と遺伝子の相互作用が:  $G \times E$

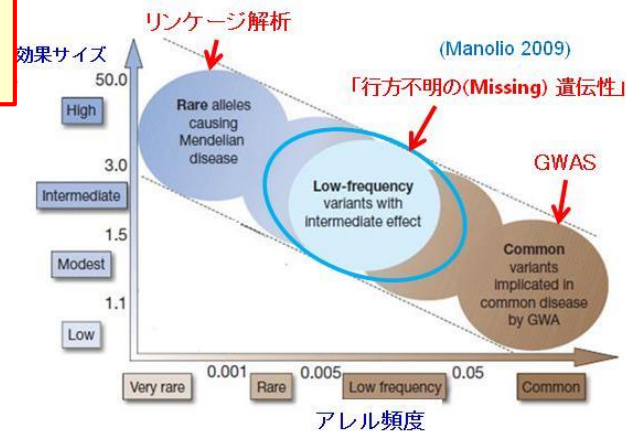
SNPの相対リスク  
低い(1.1~1.3)理由  
 $G \times E$  組合せ特異的効果  
を環境要因の平均



多因子疾患は個人の<遺伝的体質と環境要因>の  
<相互作用の結果。シーケンスだけでは解明不能

疾患発症の遺伝要因と環境要因の相互作用は  
加算的 ( $G \oplus E$ ) でもなく乗算的 ( $G \otimes E$ ) でもない  
< ( $G, E$ ) 組合せ特異的な効果 > である

例 大腸がんの遺伝要因と環境（生活習慣）要因



# 多因子疾患機序解明の地平

＜遺伝子要因と環境との相互作用の基底＞はどんな機序で行われているか

## エピゲノム

### 環境によるエピゲネティック修飾

オランダ  
飢饉 (1944)



DOHaD(Developed Origin of Health and Diseases)学説

オランダ飢饉のとき、母親の胎内にいた人々  
出生30年後、肥満、糖尿病、心疾患、高罹患率

過度な低栄養：肝臓のPPARα/γ（儉約遺伝子）メチル化低下・遺伝子発現がオン  
エピジェネティック変化は可変：短期的変化、長期的「記憶」次の世代も

環境因子



Epigenome変化



遺伝子発現調節

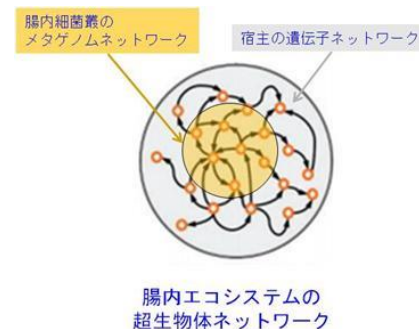
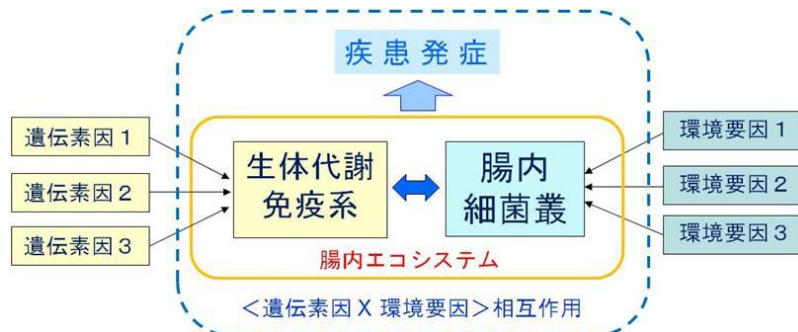


疾病発症

## メタゲノム

### Microbiomeにおける生体細菌叢相互作用

- ・ 食事などの栄養物質による環境要因は、**腸内細菌叢の代謝物**を介して、宿主の生体機構に相互作用
- ・ 心筋梗塞や糖尿病、**腸内細菌が産出する代謝物**（短鎖脂肪酸やTMAOなど）が**生体シグナル物質**や**生体活性物質**となって**受容体や転写因子の活性化**して生体側の**遺伝子ネットワーク**に働きかける。
- ・ 腸内細菌叢と生体の＜**超生物系**（supra-organization）＞において＜**環境要因x遺伝素因**＞の相互作用



メタゲノム