

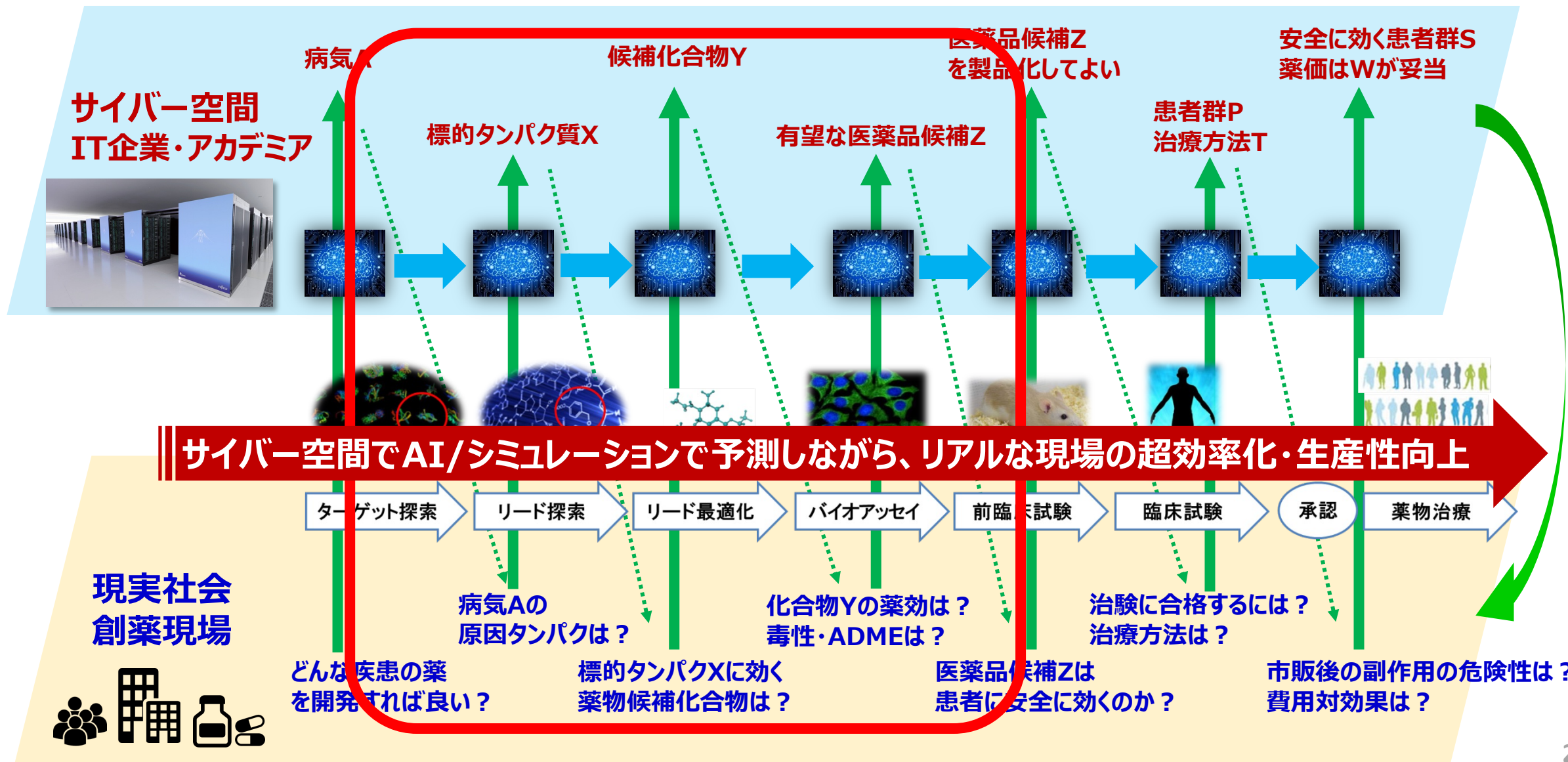
創薬DXプラットフォームの構築と応用

研究課題1

理化学研究所計算科学研究センター
京都大学大学院医学研究科ビッグデータ医科学

櫻木 実, 奥野 恭史

徹底的なDX化による創薬の超効率化・生産性向上を実現する



創薬DXプラットフォーム

標的分子探索・同定

リード探索・最適化

前臨床試験

臨床試験

入力：
患者臨床情報
疾患オミクス情報

標的探索から前臨床までの一貫した予測による高速化・成功率向上

- ・リード探索段階からのADMET・物性予測による薬剤設計の超効率化
- ・患者臨床情報・オミクス情報に基づく治験成功率の向上

出力：
創薬標的分子
薬剤候補構造
予測理由
その他の有用情報

最終波及効果

創薬のスピードアップ・
超効率化、日本の創薬
力向上、医療費削減、
新薬を待つ患者への
迅速な医薬品の提供

多様なモダリティの高精度薬剤設計

低分子・中分子・高分子薬剤の分子生成・モダリティ変換・最適化

患者臨床情報・オミクス情報を入力として、創薬標的分子とその薬剤候補構造を予測理由とともに提案する
高い精度・信頼性・汎用性・利便性をもつ「創薬DXプラットフォーム」の構築と社会実装

最新のAI技術の改良・導入

- ・遺伝子・タンパク・化合物の基盤・言語モデル・生成AI
- ・AlphaFold, 拡散モデルによる構造生成
- ・次世代スパコン・量子コンピューティング技術

世界トップレベルの独自AI・シミュレーション技術

- ・臨床データからの独自の患者層別化・病態マーカー特定技術
- ・世界最高精度の標的探索AI・低分子創薬AI
- ・独自の中分子・核酸創薬DBとAI
- ・AIとの融合による高速・高精度分子シミュレーション技術
- ・独自3D構造生成AIによるAlphaFoldの弱みを克服

- LINC・DAIIAでのこれまでの創薬AI開発実績に基づく、効果的・効率的なプラットフォーム開発
- 十数年に渡りAI・シミュレーション創薬を牽引してきたベテランと新進気鋭の若手研究者による開発体制

産学・アカデミアとの連携

- ・アカデミア創薬・製薬企業テーマへの応用・精度向上・ユーザビリティ向上
- ・国内を代表する医療機関・基礎研究機関との実験・創薬応用での連携

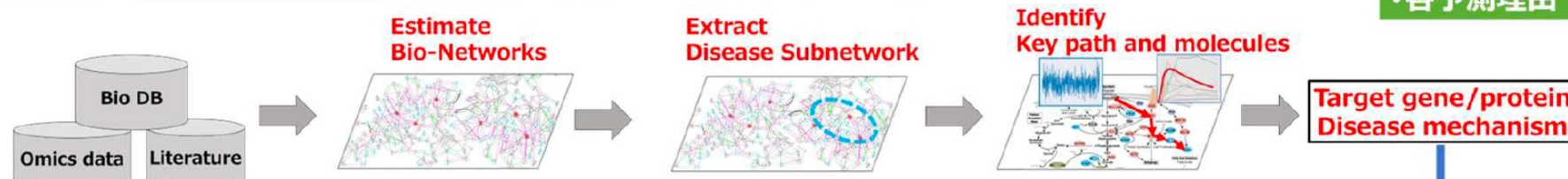
メガファーマに対抗するデータ戦略

- ・製薬企業のデータ・経験知
- ・製薬企業との連合学習体制
- ・大規模言語モデル(LLM)による文献からの知識抽出
- ・富岳を用いた大規模シミュレーションによるデータ生成
- ・AI駆動型による超効率実験体制の構築

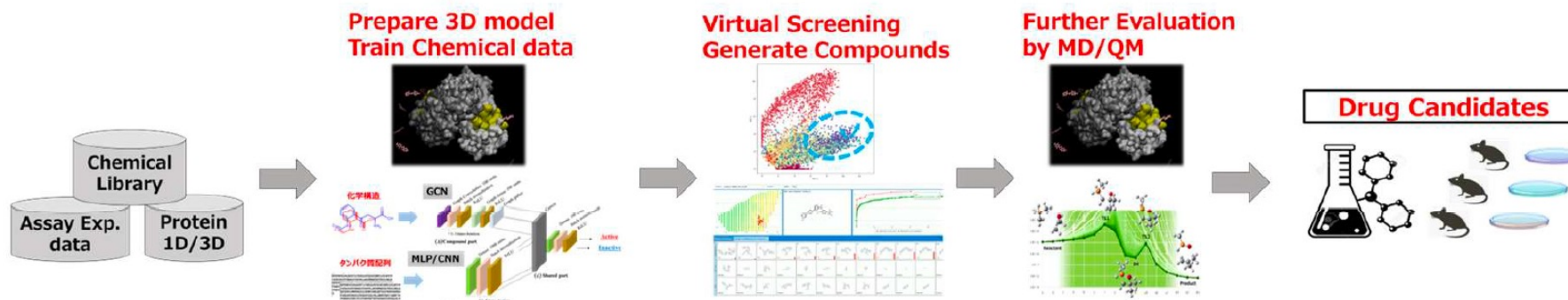
研究代表者：奥野恭史

DAIIA-Xにおける創薬DXプラットフォームの全体像

創薬標的分子探索フロー



薬剤デザインフロー 低分子・中分子（ペプチド系、非ペプチド系、核酸）・高分子（抗体）



創薬標的分子探索フロー

入力：
患者情報
臨床データ
オミクスデータ

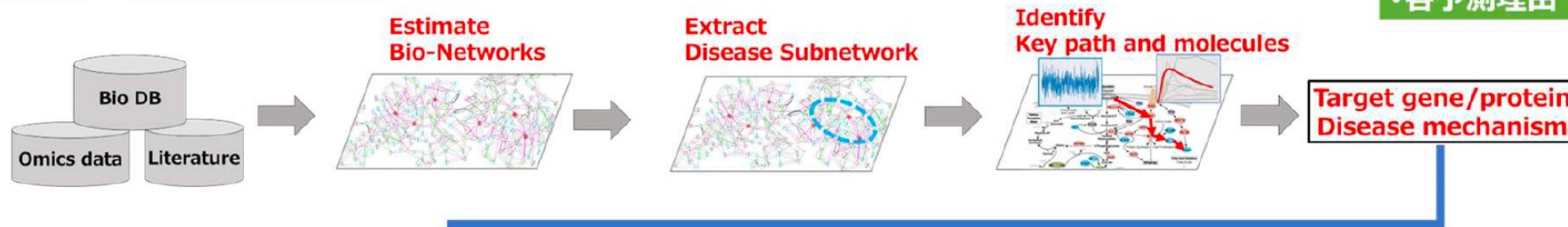
患者層別化
病態マーカー推定
バイオバンクからの
サンプル選択

オミクスプロファイル
分子ネットワーク
文献・既知DB・知識グラフ
からの標的探索

因果シミュレーション・
AI駆動型スクリーニ
ング実験による
標的同定

標的からの
最適モダリティ推定

出力：
標的分子名
モダリティ候補
病態マーカー
各予測理由



薬剤デザインフロー 低分子・中分子（ペプチ ド系、非ペプチド系、核 酸）・高分子（抗体）

入力：
標的分子名

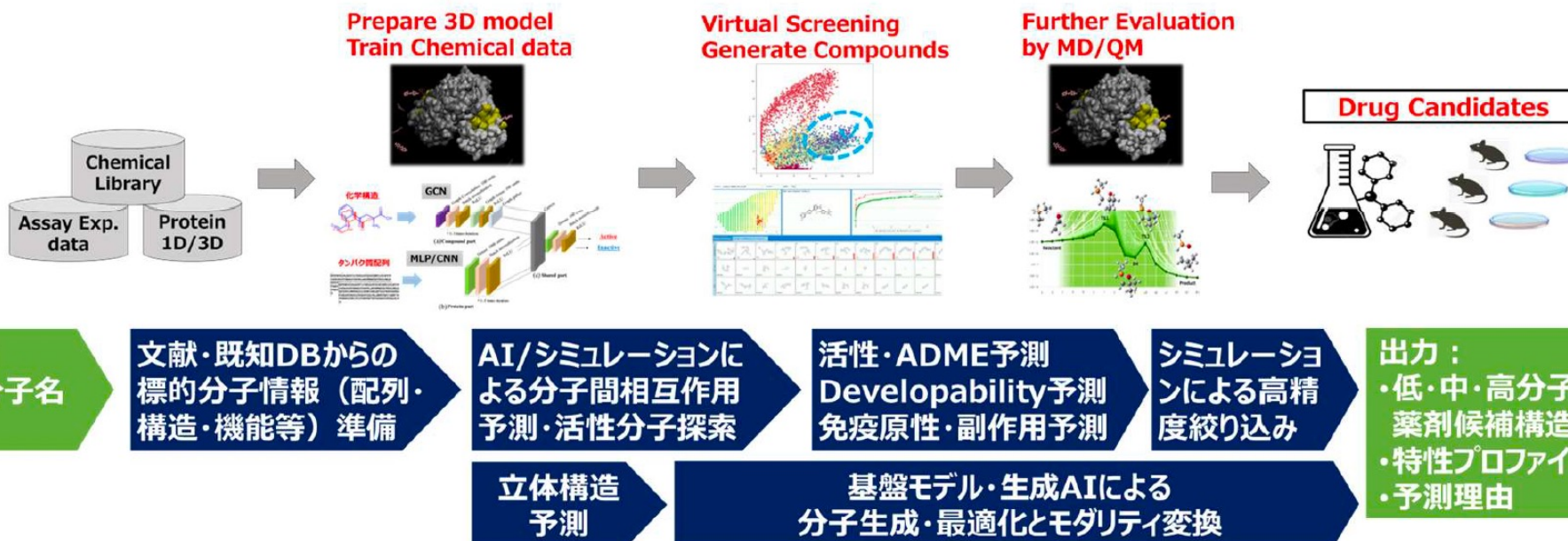
文献・既知DBからの
標的分子情報（配列・
構造・機能等）準備

AI/シミュレーションに
よる分子間相互作用
予測・活性分子探索

活性・ADME予測
Developability予測
免疫原性・副作用予測

シミュレーシ
ョンによる高精
度絞り込み

出力：
低・中・高分子
薬剤候補構造
特性プロファイル
予測理由



患者臨床データ・バイオバンク

ターゲット探索

ターゲット選択

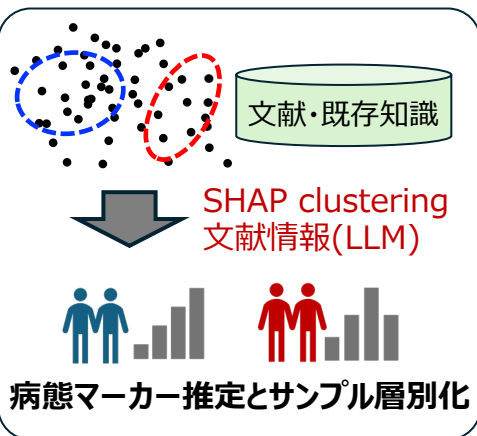
⑪ 患者層別化・病態マーカー抽出法

基盤分子ネットワーク

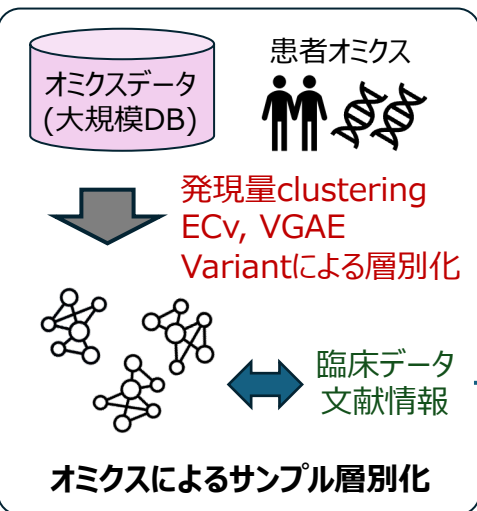
Case/Ctrl差

⑯ 計算フロー構築と実験検証

対象患者群
病態マーカー



オミクスデータ取得



オミクスデータ

DB・文献・KG
(既知情報)

RNAseq
基盤モデル

DNA
基盤モデル

疾患特異的
ネットワーク

標的候補
分子リスト

因果経路最適
化・AI駆動実験

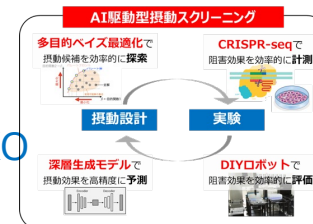
創薬標的
候補分子

候補
モダリティ

⑫ 統合モデルによる標的探索法

- ・従来DEG, scRNA解析
- ・疾患・摂動マルチモーダルAI
- ・ベイジアンネット/TP/Subnet法
- ・SLGCN (GE+KG)
- ・基盤モデル・グラフ学習/virtual KO
- ・RNA動態深層学習モデル

⑭ 因果シミュレーションと
AI駆動型実験システムによる選択



⑮ 最適モダリティ推定法

- ・DB, LLMによる推定

疾患関連遺伝子・変異

疾患関連タンパク構造部位

- ・変異病原性予測 (PathoGen, AF等)

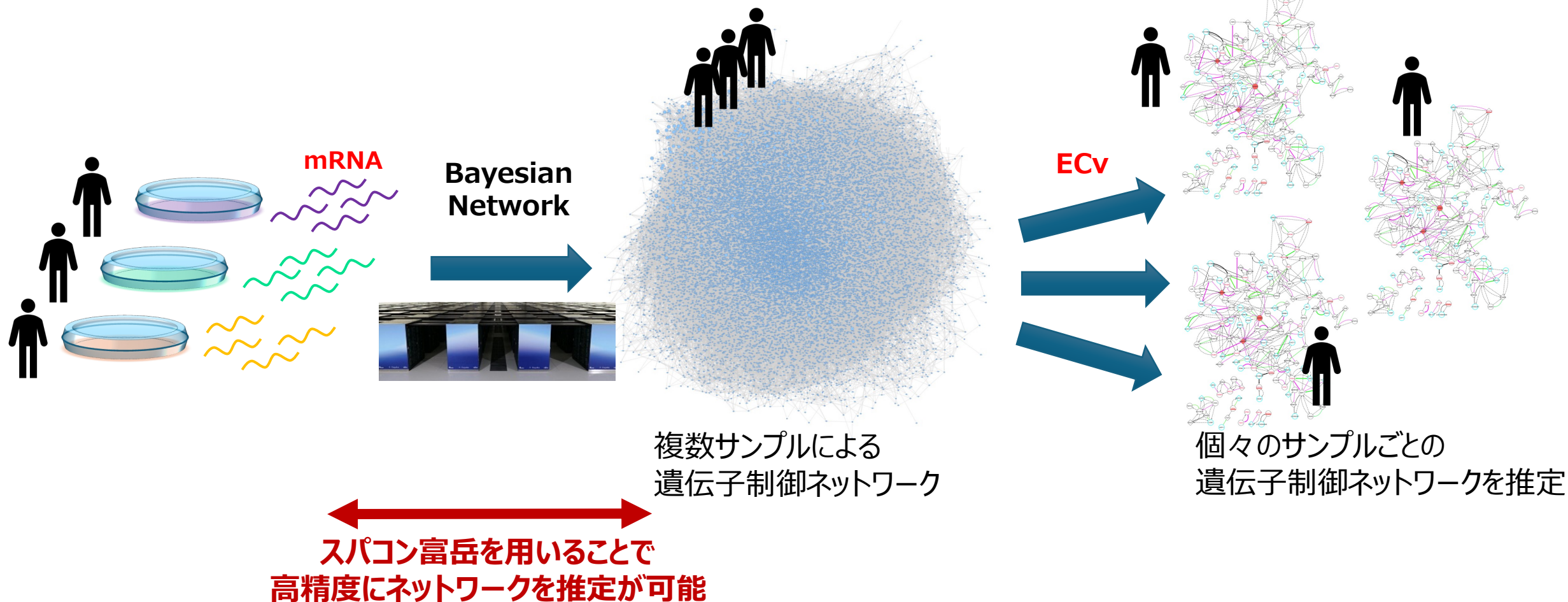
⑬ 病態バリエーション/遺伝子予測法

- ・変異タンパク影響度解析 (AF, MD等)

オミクスデータからのネットワークレベルでの創薬標的遺伝子の特定

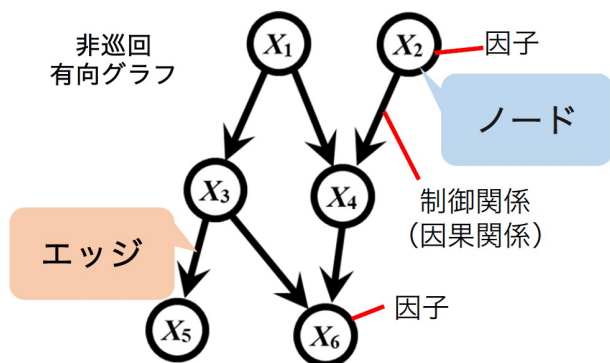
従来法は、複数サンプルを総合した
遺伝子ネットワークを推定するのみ

個々のサンプル（患者）ごとの
ネットワークを抽出し、正常由来ネット
ワークと比較することで、ネットワークレ
ベルでの創薬標的遺伝子の特定が可能



ベイジアンネットワーク

グラフ統計モデルで、因子（＝確率変数）間の
因果関係（＝条件付き独立性）をネットワークで表現したもの



多変量同時確率の分解

6変数の分解例

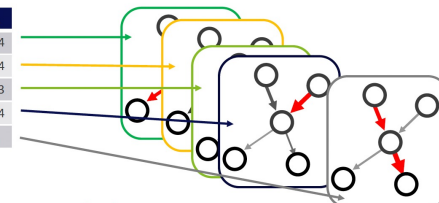
$$\Pr(X_1, X_2, X_3, X_4, X_5, X_6) \\ = \Pr(X_1) \Pr(X_2) \Pr(X_3|X_1) \Pr(X_4|X_1, X_2)$$

$$\Pr(X_5|X_3) \Pr(X_6|X_3, X_4) \\ = \prod_{j=1}^p \Pr(X_j|Pa(X_j))$$

dECv (delta-Edge Contribution value)

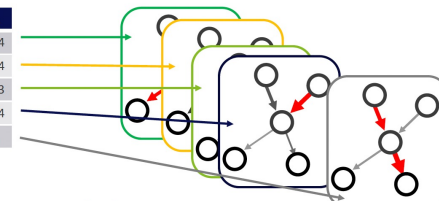
正常サンプルのECv

サンプル	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆
A	156	52	18	141	91	34
B	172	81	24	132	88	54
C	148	49	16	115	81	23
D	152	51	18	121	76	44
...	...	...	...	...	...	...



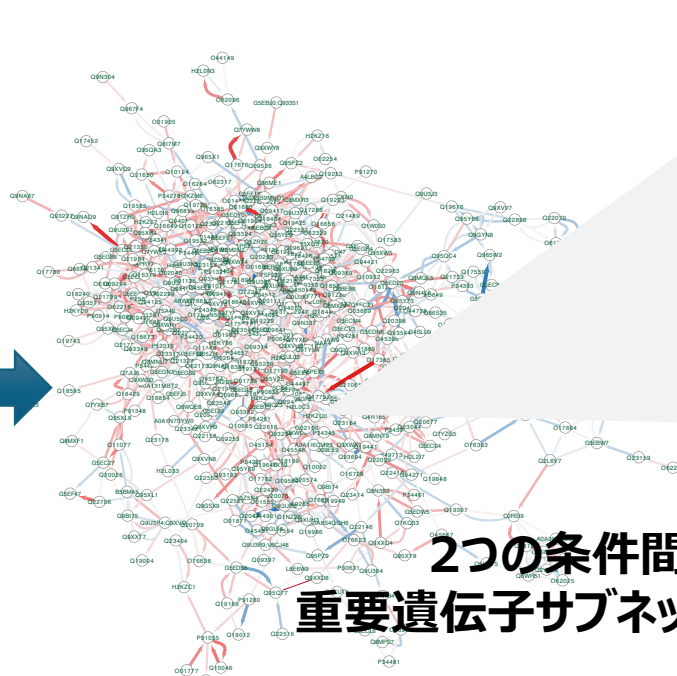
疾患サンプルのECv

サンプル	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆
A	156	52	18	141	91	34
B	172	81	24	132	88	54
C	148	49	16	115	81	23
D	152	51	18	121	76	44
...	...	...	...	...	...	...

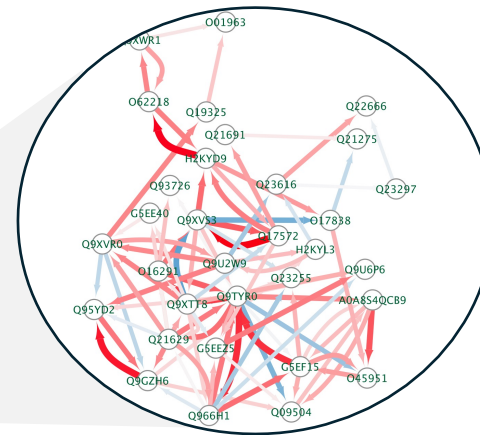


独自開発技術

複数サンプル・全タンパクによる
basal ネットワーク推定



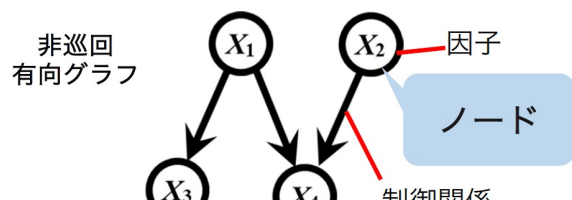
2つの条件間における
重要遺伝子サブネットワークを抽出



病態関連遺伝子
下流Pathway解析

ベイジアンネットワーク

グラフ統計モデルで、因子（＝確率変数）間の
因果関係（＝条件付き独立性）をネットワークで表現したもの



多変量同時確率の分解

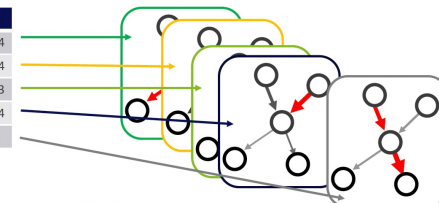
6変数の分解例

$$\Pr(X_1, X_2, X_3, X_4, X_5, X_6) \\ = \Pr(X_1) \Pr(X_2) \Pr(X_3|X_1) \Pr(X_4|X_1, X_2)$$

dECv (delta-Edge Contribution value)

正常サンプルのECv

サンプル	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆
A	156	52	18	141	91	34
B	172	81	24	132	88	54
C	148	49	16	115	81	23
D	152	51	18	121	76	44
...	...	...	...	...	...	...



独自開発技術

スパコン富岳を用いて大規模な遺伝子ネットワークを個人レベルで推定できる
従来の遺伝子発現量の解析では捉えられない、疾患関連分子ネットワークの抽出が可能
DAIIA-X（創薬DXプラットフォーム）において実装済み・アプリケーションとして実運用中

→ 未知の病態関連分子・サブネットワークの同定、新規創薬ターゲット探索に有用

複数サンプル・全タンパクによる
basal ネットワーク推定



2つの条件間における
重要遺伝子サブネットワークを抽出

病態関連遺伝子
下流Pathway解析

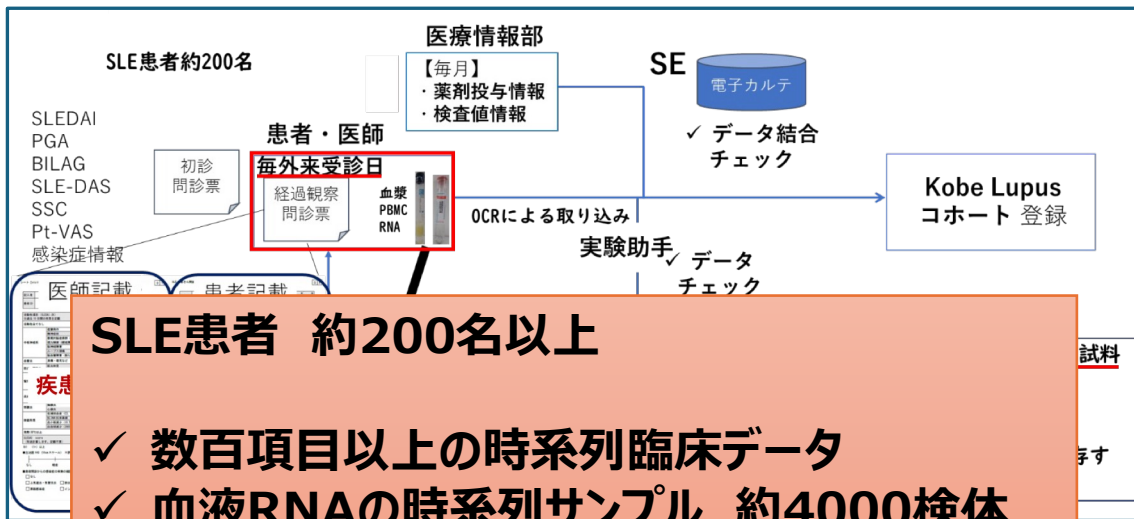
SLE（全身性エリテマトーデス）

若年女性が罹患する自己免疫疾患。皮膚、関節、腎臓、中枢神経、血球、心肺、消化管あらゆる臓器が障害をおこすが、**個々の患者で罹患臓器や重症度が極めて多彩**

ステロイド、免疫抑制薬、生物学的製剤により生命予後は改善したが、**依然課題の多い疾患**

世界で“唯一無二のSLEデータベース”

神戸市立医療センター中央市民病院と、神戸医療都市産業推進機構医療イノベーション推進センター(TRI)が中心となって構築



SLE患者 約200名以上

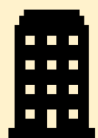
- ✓ 数百項目以上の時系列臨床データ
- ✓ 血液RNAの時系列サンプル 約4000検体
- ✓ DNAサンプル 約800検体

1. ステロイド、免疫抑制薬、生物学的製剤を用いても**寛解導入が困難な難治性患者**が存在する
2. 客観的には寛解状態にあるにもかかわらず、**強い倦怠感などのため社会復帰が困難**な病態が存在する
3. 多数存在する免疫抑制薬、生物学的製剤の中から、各患者に**最適な治療薬を事前に予測することが困難**である

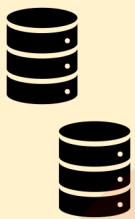


DAIIA-Xで開発している創薬DXプラットフォームを適用・応用することで、治療抵抗性SLEの病態解明・創薬標的探索を実証。さらに、これをモデルケースとした臨床データベースからのAI創薬システムを構築する

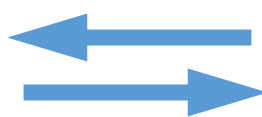
神戸医療都市産業推進機構 医療イノベーション推進センター(TRI)



- DMPに基づくデータベースの管理・運用 (QC/標準化/クリーニングなど)
- 解析用データセットの整備・基本解析



臨床情報・試料情報の抽出
院内で仮名化(対応表管理)



神戸市立医療センター中央市民病院



- 臨床データベース構築、患者試料収集
- 専門医アウトカム定義や、標的候補の臨床的妥当性検証、動物実験の検証



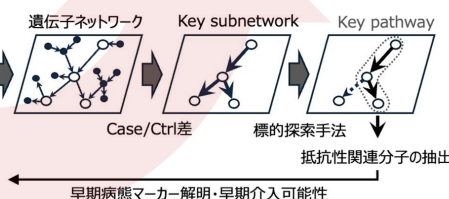
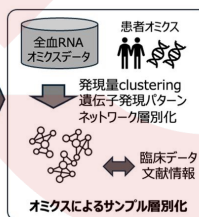
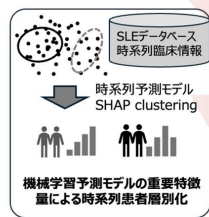
※個人識別情報・原試料は全て院内で管理

解析結果
再現性可能な
手順共有

治療抵抗性SLEの分子メカニズム解明 創薬標的探索・早期マーカー推定

臨床的妥当性
検証フィードバック
in vivo/in vitro実験

仮名化済み解析用データ
(時系列臨床/オミクス /PRO)



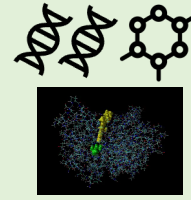
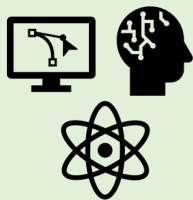
標的候補+層別化/早期予測
バイオマーカー候補
ツール化合物探索(迅速PoC)
新規化合物設計(必要に応じ)

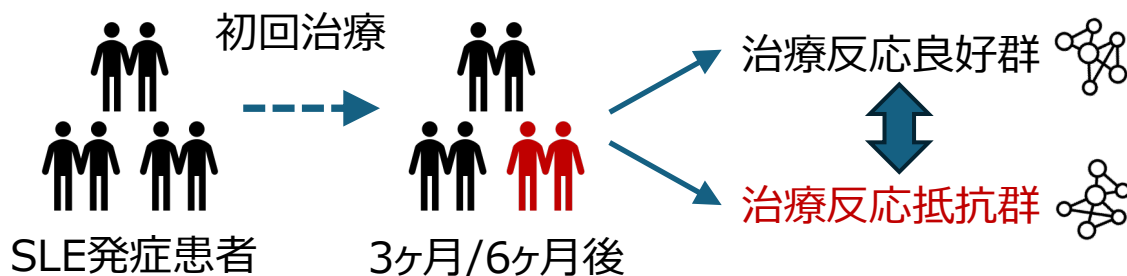
京都大学 (医学研究科ビッグデータ医科学)



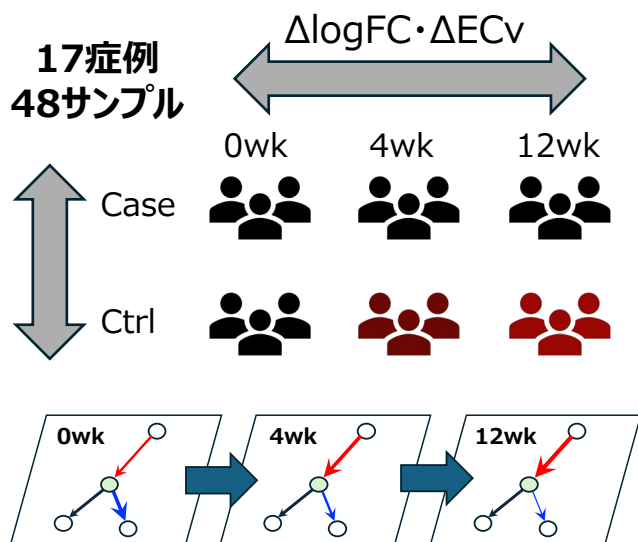
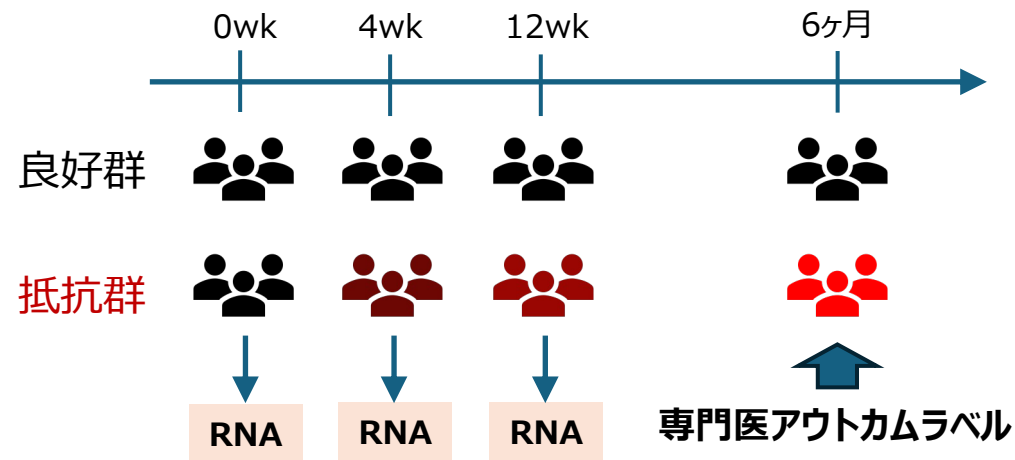
- 構築済み・運用中の創薬DXプラットフォーム (DXPF) 既存計算基盤を活用した標的探索解析・マーカー抽出
- 創薬DXPFの実応用検証

理化学研究所計算科学研究センター (R-CCS)

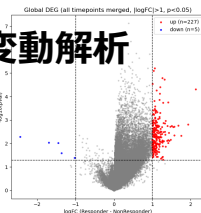




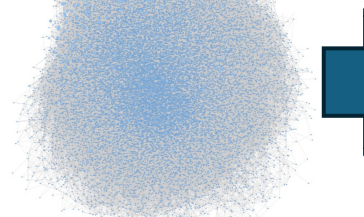
治療抵抗性はいつ頃・どんな分子メカニズムの時間変化で生じるのか？
抵抗性の病態完成よりも早期のオミクス変化から創薬標的探索を実施
FS検証では、時系列RNAseqとベイジアンネットワークで解析



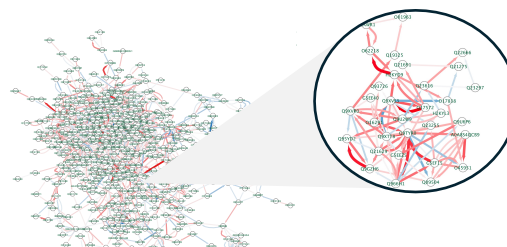
発現変動解析



ベイジアンネットワーク解析



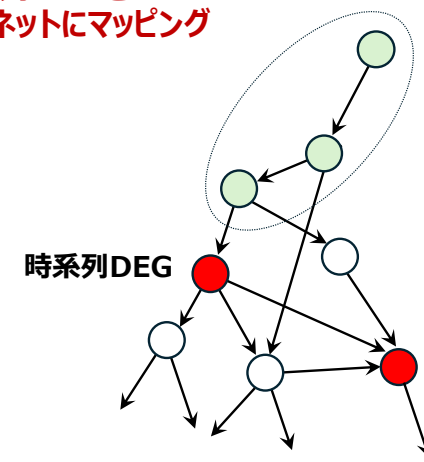
発現量の時間変化の群間差が
大きい遺伝子（時系列DEG）



ネットワーク時間変化の群間差が
大きい（時系列dECvサブネット）

時系列DEGをdECv
サブネットにマッピング

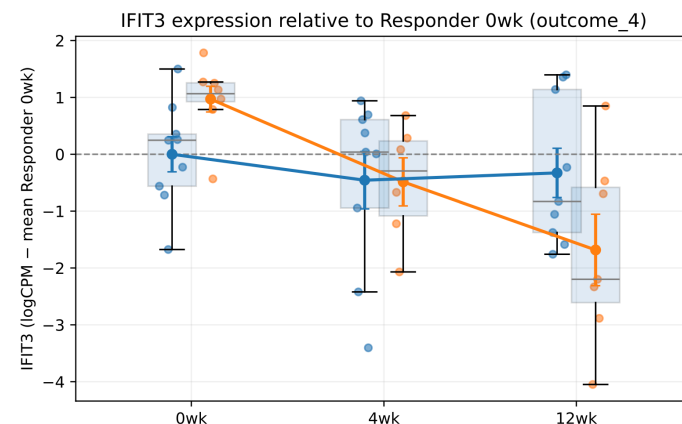
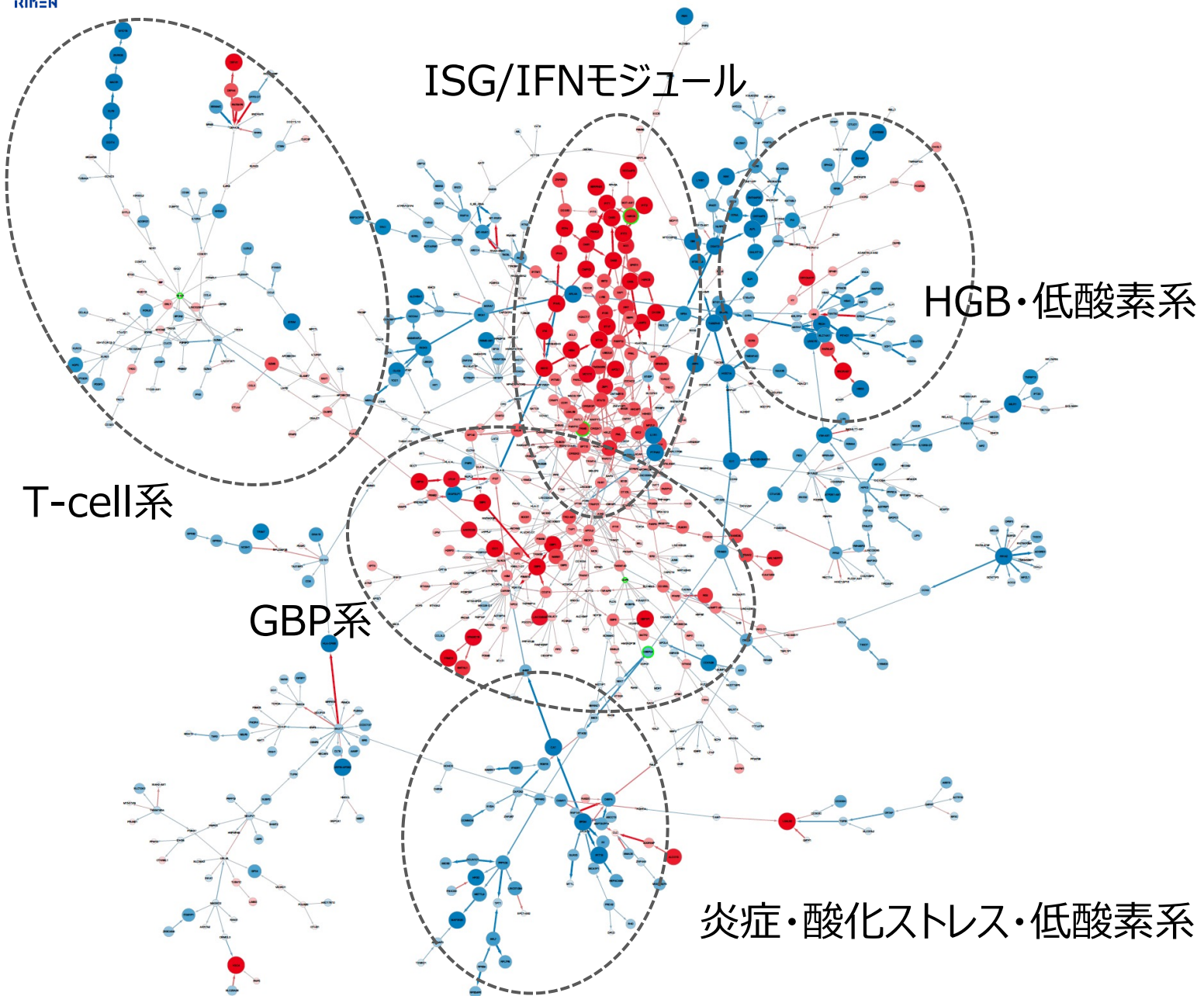
病態関連遺伝子



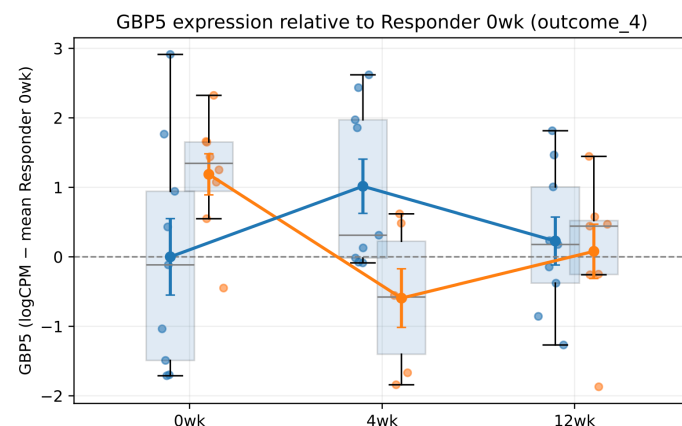
- ① 遺伝子制御関係の時系列変化がCase/Ctrlで大きく異なるもの
- ② Case/Ctrlの差分が時間経過とともに大きくなっていくもの

**治療抵抗性の時系列的な病態
形成に重要な遺伝子群の抽出**

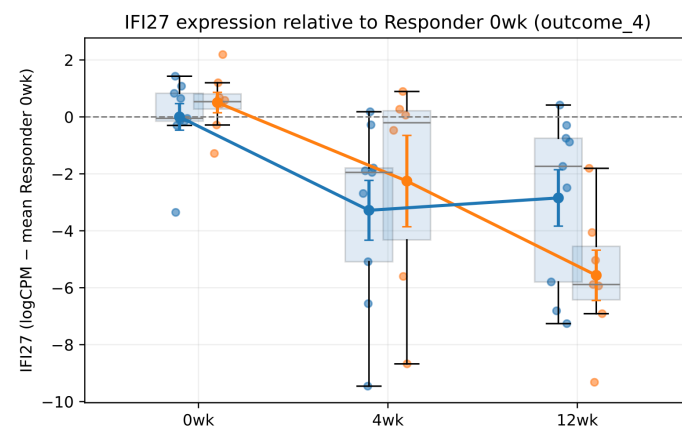
治療抵抗性SLEの病態形成に関わるサブネットワークを抽出



IFIT3



GBP5

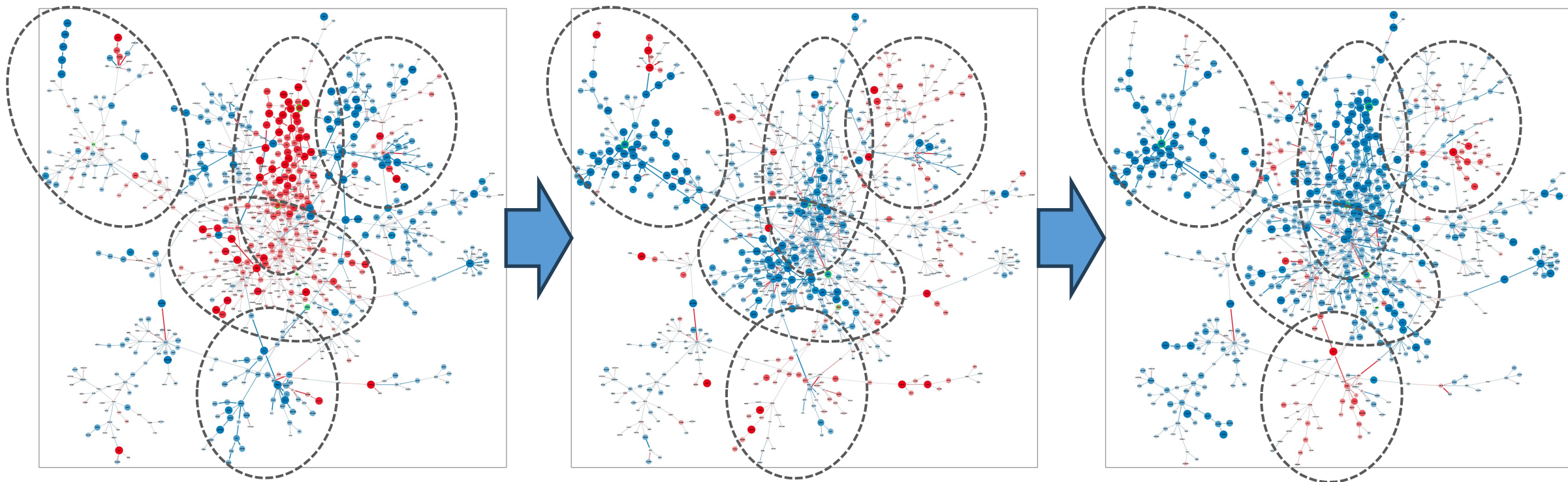


IFI27

0wk

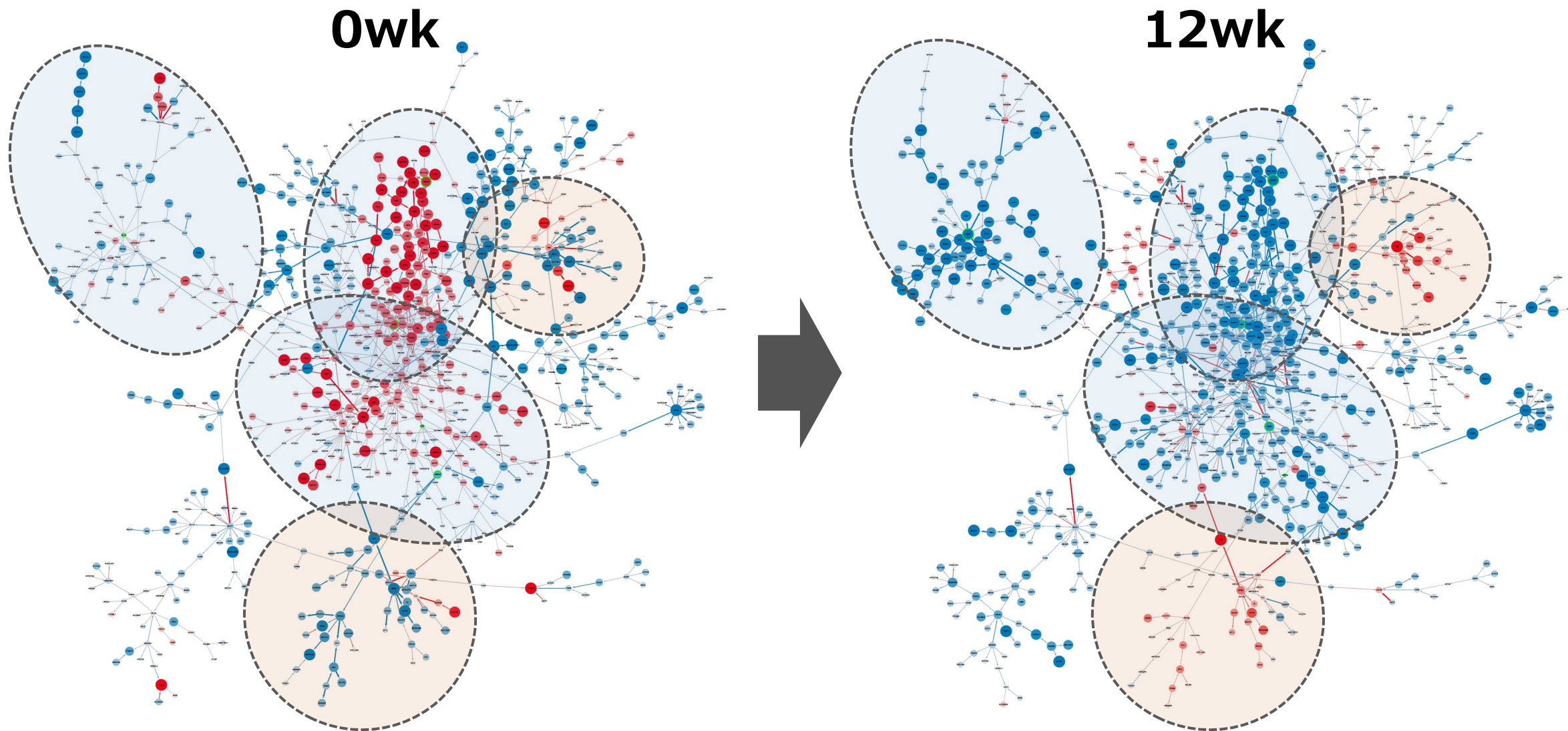
4wk

12wk



**治療抵抗性SLEの病態形成において、時系列的に変化する
病態関連分子モジュールを捉えることに成功**

治療抵抗性SLEの病態形成に関わるサブネットワークを抽出



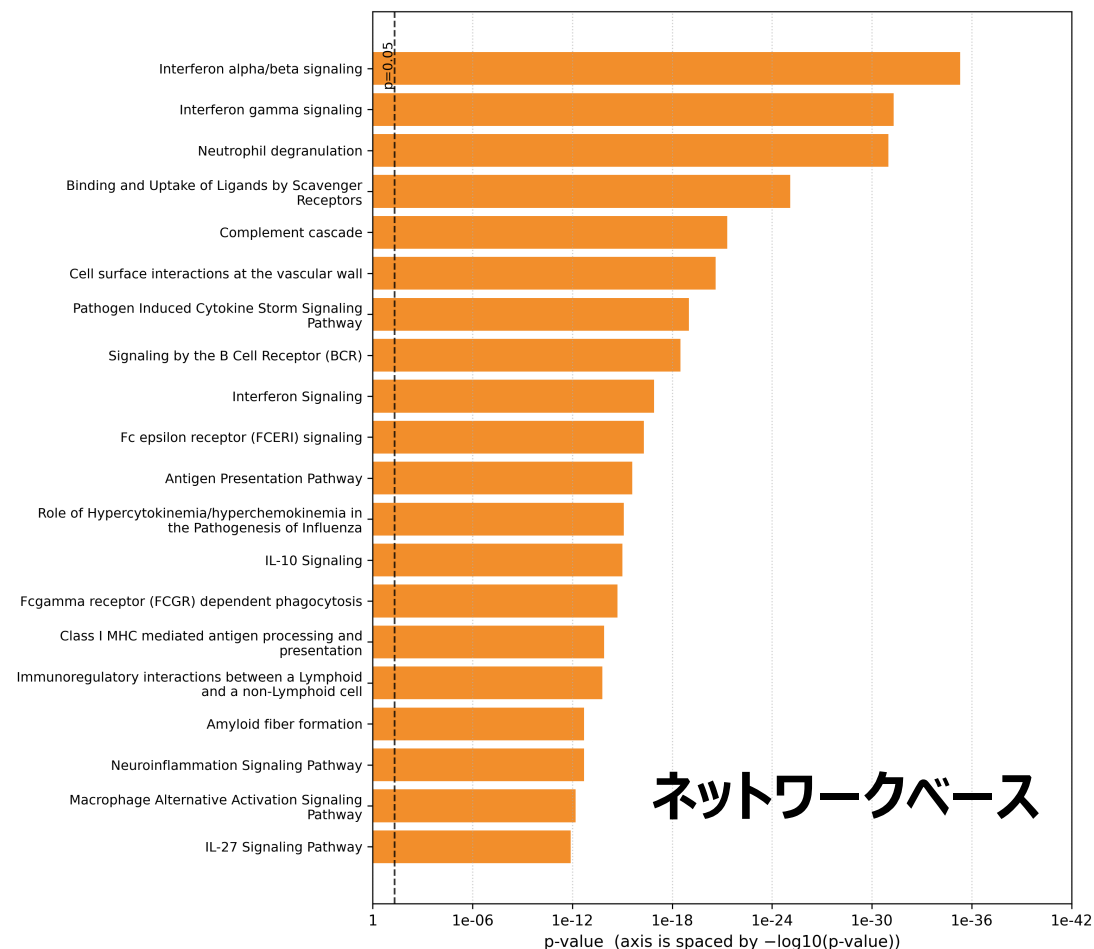
一部の遺伝子サブネットワークが時間経過で強化／減弱していることを同定

時系列DEG + 時系列サブネットのIPA解析上位20位（パスウェイ解析）

IPA Canonical Pathways - Top 20 (DEG)



IPA Canonical Pathways - Top 20 (dECv)



- IFN/ISG・抗ウイルス + 炎症性サイトカイン応答
- エピゲノム/クロマチン・転写（rRNA含む）制御 + DNA修復/組換え
- 造血/血球系（顆粒球・巨核球/血小板・赤血球） + 細胞老化

- I型/II型IFN・抗ウイルス/サイトカインストーム系
- 自然免疫のエフェクター（好中球・補体・貪食・マクロファージ） + 血管壁相互作用
- 獲得免疫・抗原提示/Fc受容体・免疫調節～炎症波及

- FS検証により時系列RNAネットワーク解析で特定された分子ネットワークの**詳細解析と個別ネットワーク解析**を実施中
- 同時系列サンプルを用いて追加で**シングルセル解析**を実施中
- 並行して、数百項目の時系列臨床データ・検査データを用いて、治療抵抗性に関する、**患者時系列層別化解析**を実施中
- DAIIA-Xプロジェクトとも連携し、実臨床データからの創薬DXPF応用検証から、「AI創薬システム構築」のためのシステム構築も並行して整備

**DAIIA-X(創薬DXPF) × 臨床オミクス時系列DB × ネットワーク解析
創薬DXプラットフォームを用いた新規治療標的探索と創薬研究を加速**