

A1-1 データベースの構築とホモロジーモデリングによる機能推定の支援

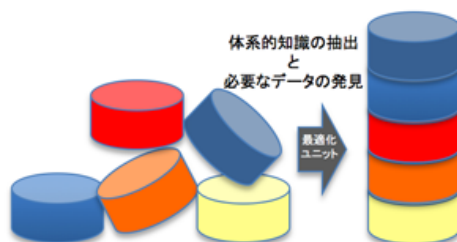
測定データのウェブデータベース構築、タンパク質の立体構造モデリング、タンパク質の機能推定

[1] 支援担当者

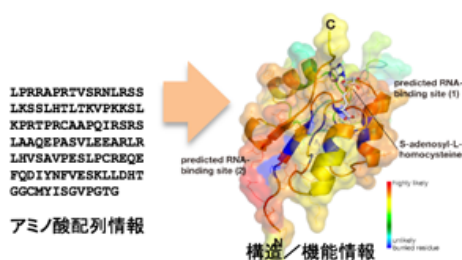
所属		①早稲田大学 先進理工学部
氏名		①鈴木 博文、由良 敬
AMED 事業	ユニット／領域名 課題名	プラットフォーム機能最適化ユニット 創薬等ライフサイエンス研究を促進する研究支援とデータサイエンス
	代表機関 代表者	早稲田大学 由良 敬
支援技術のキーワード		データベース、ホモロジーモデリング、タンパク質機能予測、データ発信用ウェブ開発

[2] 支援技術の概要

- 多くの数値／文字データのデータベース化と、データベースからの知識発見解析



- 研究対象としているタンパク質の立体構造モデリングと機能残基の推定



[3] 支援技術の利用例

- MySQLやPostgreSQL、PHP、Pythonなどを利用したWeb公開型データベースの構築とデータベースにもとづくデータの解析（ケーススタディー）
- 独自のアラインメント技術を用いた遠縁タンパク質をテンプレートとしたホモロジーモデリング
- ホモロジーモデリング構造に基づくタンパク質の機能部位予測。特にRNAとの相互作用残基の推定
- DNAの点変異がもたらすタンパク質構造変化の推定と疾患発症機構の推定など

[4] 支援担当者の研究概要

- ABCトランスポートの多数の立体構造情報と遺伝子のミスセンス情報を組み合わせることで、ミスセンス変異がタンパク質の動的構造の要部分に存在することを見いだした(Sakamoto M., Suzuki H., Yura, K. (2019) *Biophysics and Physicobiology*, **16**, 68-79.)。
- 神経細胞で発現する遺伝子情報を解析し、細胞の形態変化に関与する遺伝子候補を抽出した。その結果に基づく実験により、ほ乳類脳の層構造形成に関与する遺伝子群を特定した(Ohtaka-Maruyama C., ... Yura, K., ... Maeda, N. (2018) *Science*, **360**, 313-317.)。
- ヒトデのアポトーシス関連タンパク質の立体構造をモデリングし、ほ乳類のアポトーシス機構との構造的および機能的差異を推定した (Tamura, R., ... Yura, K., Chiba, K. (2018) *Scientific Reports*, 1611.)。
- Autophagy関連タンパク質を収集したデータベースを公開開始した(Mizutani, H., Sugawara, H., ... Yura, K. (2017) *BMC Structural Biology*, **17**, 4.)。
- 遺伝子発現データを解析し、異なる環境におかれた生物に、どのような変化が起こっているかを推定した(Masakiyo, Y., ... Yura, K., Shimada, S. (2016) *Algal Resources*, **9**, 61-76.)。
- 様々な公共データベースを連結し、新しい知見を発見するためのアプリケーションを開発した。その上でどのような知見が抽出できるかをケーススタディーした (Gojobori, T., ... Yura, K. (2016) *Structural and Functional Genomics*, **17**(4), 69-81.)。
- タンパク質の複合体界面がどのように進化してきたかを複合体構造データに基づき解析し、タンパク質複合体構造予測に適用した (Aoto, S., Yura, K., (2015) *Biophysics and Physicobiology*, **12**, 103-116.)。
- 化学合成された低分子が、核内受容体にどのような構造で相互作用しているかを推定した (Inoue, K., ... Yura, K., ... Tanatani, A. (2015) *European Journal of Medicinal Chemistry*. **102**, 310-319.)。
- 立体構造未知のタンパク質において、ホモロジーモデリング法で構造を推定し、特にループ構造の特徴に基づき、ターゲットとしているタンパク質固有の機能を推定した (Takahashi, K., ... Yura, K., ... Sakuta, M. (2015) *Natural Product Communications*, **10**, 717-719.)。
- ステロイド骨格をもつ低分子がタンパク質とどのように相互作用するかを、その複合体立体構造すべてを分類しパターンを抽出し上で、機能との相関を見いだした (Hori-Tanaka, Y., Yura, K., ... Tanaka, H. (2015) *Steroids*, **96**, 81-88.)。
- 過去に存在した転写因子のアミノ酸配列と立体構造を推定し、転写因子が結合するDNA配列と活性化ドメインの変化を推定した (Kamijyo, A., Yura, K., Ogura, A. (2015) *Gene*, **555**, 73-79.)。

A2-1 電子顕微鏡画像・計算結果のアーカイブ及び構造関連データベースの運用

PDBとBMRB及び独自の二次データベースを活用した構造解析支援、画像データと計算のセキュアなアーカイピング

[1] 支援担当者

所属		①大阪大学 蛋白質研究所
氏名		①栗栖 源嗣、川端 猛、Gert-Jan Bekker、横地 政志
AMED 事業	ユニット／領域名 課題名	プラットフォーム機能最適化ユニット 創薬等ライフサイエンス研究を促進する研究支援とデータサイエンス
	代表機関 代表者	大阪大学 栗栖 源嗣
支援技術のキーワード		EMDB、MD計算、PDB、BMRB、ホモロジーモデリング

[2] 支援技術の概要

Protein Data Bank Japan (PDBj)の運営実績と経験をベースに、電子顕微鏡画像および分子動力学計算やホモロジーモデリングの計算結果をアーカイブするサイトを運営する（<https://empiar.pdbj.org> および<https://bsma.pdbj.org>）。並行して、運営管理をおこなっているProtein Data Bank（PDB）およびBioMagResBank（BMRB）、さらに関連する二次データベースやツール群を活用した構造解析を支援する。支援に供する基盤技術は以下の通りである。

- 論文に紐付いたクライオ電子顕微鏡画像データ（EMPIAR）の公開・登録サイトの運営
- 論文に紐付いた計算結果データ（BSM-Arch）の収集・公開サイトの運営
- ヒト・タンパク質の立体構造モデルの構築
- データベースを活用した自動解析パイプラインによるNMR研究者の構造決定支援
- 複合体立体構造の検索・モデリング（HOMCOSサーバ）
- 立体構造に基づく配列解析
- 電子顕微鏡3次元マップデータへの原子モデルの構造比較と重ね合わせ計算



Electron Microscopy Public Image Archive

[EMPIAR home](#) [Deposition](#) [REST API](#) [FAQ](#) [About EMPIAR](#)

[Feedback](#)

EMPIAR, the Electron Microscopy Public Image Archive, is a public resource for raw, 2D electron microscopy images. Here, you can browse, upload, and download and reprocess the thousands of raw, 2D images used to build a 3D structure. [More ...](#)

Deposit your data in EMPIAR to share it with the structural biology community.

Browse and [download](#) EMPIAR datasets using the table below.

50

entries

Search:

Dataset	Title	Authors	Related EMDB/PDB entries	Size	Resolution
 EMPIAR-10258	LRRCA-DCPIB in MSP1E3D1 nanodiscs (2482 multi-frame micrographs composed of 40 frames each in TIFF format)	Kern DM, Oh S, Hite RK, Brohawn SG [PubMed: 30775971] [DOI: 10.7554/eLife.42636]	EMD-0562 , 6nzw	984.1 GB	3.21
 EMPIAR-10252	Catalytic subunit of protein kinase A bound to ATP, manganese, and IP20 movies obtained using Talos Arctica operating at 200 kV equipped with a K2 (4812 multi-frame micrographs composed of 44 frames each in MRC format)	Herzik Jr MA, Wu M, Lander GC [PubMed: 30833564] [DOI: 10.1038/s41467-019-08991-8]	EMD-0409	5.5 TB	6.0
 EMPIAR-10250	Human methemoglobin movies obtained using Talos Arctica operating at 200 kV	Herzik Jr MA, Wu M, Lander GC [PubMed: 30833564] [DOI: 10.1038/s41467-019-08991-8]	EMD-0407 , EMD-0408 , 6nbc , 6nbd	1.9 TB	2.8-3.2

Quick links

[EMPIAR PDBj](#)
[PDBj](#)
[EMDataBank](#)
[EM Navigator](#)

[EMPIAR: Quick tour](#)

EMPIAR citations

Retrieving high-resolution information from disordered 2D crystals by single-particle cryo-EM.
Righetto RD, Biyani N, Kowal J, Chami M, Stahlberg H. (2019)
[Cryo-EM structure of a light chain-derived amyloid fibril from a patient with systemic AL amyloidosis.](#)
Radamaker L, Lin YH, Annamalai K, Huhn S, Hegenbart U, Schönland SO, Fritz G, Schmidt M, Fändrich M. (2019)
[High-resolution structure determination of sub-100 kDa complexes using conventional cryo-EM.](#)
Herzik MA, Wu M, Lander GC. (2019)
[Cryo-EM structure of phosphodiesterase 6 reveals insights into the allosteric regulation of type I phosphodiesterases.](#)

[3] 支援技術の利用例

データベース運営、生物情報科学、MD計算、生体NMR構造解析の経験と実績から、以下のような技術利用例がある。

- クライオ電子顕微鏡画像データの大量保存
- 分子科学計算データの大量保存
- PDBを利用したヒト・タンパク質の立体構造モデル構築
- 酵素の活性向上のための基質結合立体構造のモデリング(Kawabata et al., *J. Biosci. Bioeng.*, 2017, **124**:28-35)
- 予測立体構造に基づくアミノ酸変異の影響評価(Kimura et al., *Transl Psychiatry*, 2017, **7**:e1214; Ishizuka et al., *Transl Psychiatry*, 2017, **7**:e1184)
- 原子モデル・3次元マップの構造比較・3次元フィッティング
- 分子動力学計算による酵素のダイナミクス解析 (Inaba et al., *J. Thermal Anal. Calorimetry*, 2017, **135**:2655-2663; Numoto et al., *Biochemistry*, 2017, **57**:5289-5300)

[4] 支援担当者の研究概要

結晶構造解析、NMR分光法、クライオ電子顕微鏡構造解析の3手法により構造決定された生体高分子の原子座標は、付随する実験データとともに大阪大学蛋白質研究所が運営するPDBj (Protein Data Bank Japan:大阪大学に措置された共同利用・共同研究拠点経費及びJSTのライフサイエンスデータベース統合化推進プログラムによって支援) が、国際組織wwPDB (worldwide PDB: <http://wwpdb.org/>) のメンバーとしてPDBにアーカイブしており、本ユニットでもPDBjと連携してデータ登録を支援する。原子座標を伴わないクライオ電子顕微鏡法によるポテンシャルマップとNMR分光法によるスペクトルデータについても、大阪大学蛋白質研究所が運営するPDBjがwwPDBの基準に従ってEMDB (Electron Microscopy Data Bank) およびBMRB (BioMagResBank) にアーカイブしており、本ユニットでもPDBjと連携してデータ登録を支援する。

クライオ電子顕微鏡画像データを蓄積する試みはEMPIAR (Electron Microscopy Public Image Archive: <https://www.ebi.ac.uk/pdbe/emdb/empiar/>) としてEBI (欧州バイオインフォマティクス研究所) において開始されたが、国際的に連携した運用はされていなかった。本ユニットでは、EBIと学術交流協定を締結し2018年からEMPIAR-PDBjとしてデータの公開を開始した (<https://empiar.pdbj.org>)。大容量のEMPIARの画像データを多国間で頻繁にデータ交換することは現実的でないため、大阪大学サイバーメディアセンター及び早稲田大学との連携によって、日本で開発されたCold storage技術を活用して国内にバックアップサイトを立ち上げセキュアなデータ保存体制を確立している。

NMRの化学シフト情報が蓄積されているBMRBを活用し、NMRスペクトルからの自動的なピーク・ピッキングとAI技術による自動アサインメント、distance geometry計算まで含めたパイプラインが阪大蛋白研のPDBj-BMRBにおいてすでに構築されており、その仕組みを用いてNMR研究者の構造決定をサポートする。

大阪大学でこれまでに開発・公開しているタンパク質構造データベース関連ツールを活用し、巨大分子のWeb上での可視化技術(Molmil)、電子顕微鏡による構造データのブラウザ(EM navigator)、分子の形についてPDBとEMDBを横断的に検索するサービス(Omokage search)、複合体立体構造の検索・モデリングのサーバ(HOMCOS)、のサービスに加えて、総合的な電子顕微鏡による構造データベースとして本事業への支援とする一方、本事業の成果の一般社会への公開・発信を行う。

A3-1 ゲノム・発現量・タンパク質情報をつなぐ変異の遺伝子機能変化への影響推定

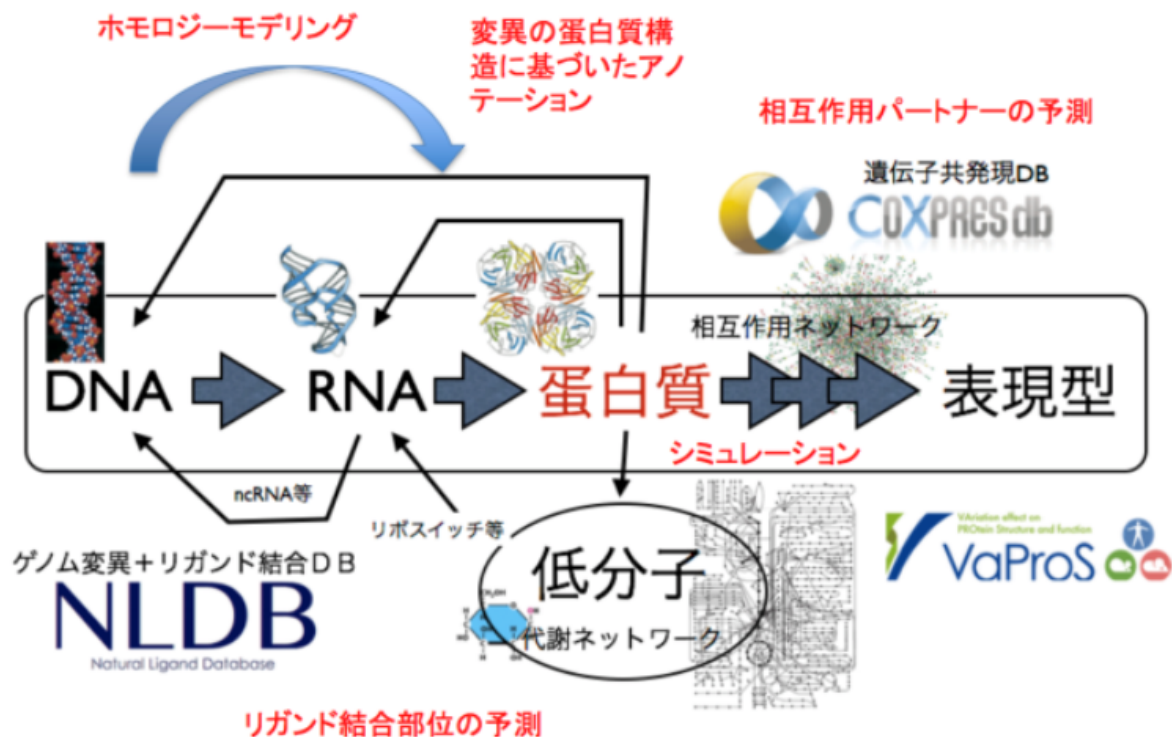
ゲノム情報解析、発現量解析、構造モデリング、タンパク質相互作用解析、MDシミュレーションを中心とした情報解析全般

[1] 支援担当者

所属		①東北大学 大学院情報科学研究科
氏名		①木下 賢吾、大林 武、西 羽美
AMED 事業	ユニット／領域名 課題名	プラットフォーム機能最適化ユニット 創薬等ライフサイエンス研究を促進する研究支援とデータサイエンス
	代表機関 代表者	東北大学 木下 賢吾
支援技術のキーワード		ゲノム変異、遺伝子発現量、タンパク質立体構造、分子動力学シミュレーション

[2] 支援技術の概要

- ゲノム変異をタンパク質立体構造上にマップすることで、変異のタンパク質機能への影響を推定。
- 構造情報がない場合は、ホモロジーモデリングで構造情報を補う。
- 機能部位情報が十分でない場合は、リガンド-タンパク質の結合部位の予測やシミュレーションを行い、機能部位を推定する。
- 遺伝子の共発現情報をもとに、相互作用しているタンパク質のパートナー予測も行う。



[3] 支援技術の利用例

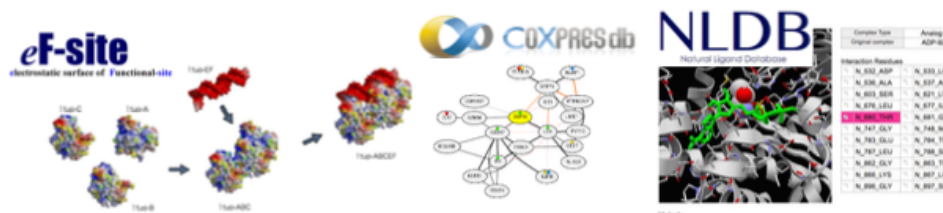
- 変異の構造上へのマッピング
Nishi et al, Protein Sci., 2016
Tadaka et al, NAR, 2018
- ホモロジーモデリング
Tsugita et al, Cell Rep., 2017
- リガンド-タンパク質相互作用面への影響評価
- シミュレーションによる機能変化の評価
Kondo et al, J. Phys. Chem, B, 2018
- 粗視化シミュレーションによる膜結合部位の予測
Fijiwara et al, Sci Rep., 2016
- 遺伝子共発現DB利用支援
Adachi et al, Genome Res, 2015

[4] 支援担当者の研究概要

私たちは一貫して「機能未知の遺伝子の機能を明らかにする」ことを目指した研究を行ってきた。機能と言っても、遺伝子1つで決まる分子機能と相互作用ネットワーク上の位置という文脈で決まる細胞機能があるが、最初は、遺伝子産物であるタンパク質の立体構造から機能を明らかにする研究を行ってきた。その後、遺伝子の共発現とタンパク質複合体の構造予測から、細胞機能を明らかにする研究へと発展させてきた。その過程で「実際に生命科学研究を行っている研究者に役に立つ」ことをコンセプトに、数多くのWebデータベースの開発を行っているが、そのいずれもが年間10万から60万程度のアクセスを獲得し、世界的にもプレゼンスのあるデータベースをととなっている。

分子機能に関しては、立体構造と機能の関係を系統的に解析し、立体構造の対称性と機能の関係、リン酸結合パターンの俯瞰的な描像と網羅的な構造モチーフの同定、分子表面の類似性による機能部位の予測、DNA結合部位の予測など多様な手法を開発、細胞機能に関しては、タンパク質複合体構造の予測、多次元共発現指標など独自の手法による遺伝子共発現データベースの構築など、情報科学的にも独創的な手法で、実用的な手法の開発を行ってきた。

最近では、東北メディカル・メガバンクプロジェクトに参画し、ゲノムの多様性を明らかにするとともに、ゲノムの変異とタンパク質立体構造と血中代謝物の関係の解析を行うなど、変異の意味を見いだすことが困難な低頻度変異に対して、タンパク質の立体構造情報を活用した第一原理的な変異のアノテーション手法の開発を目指している。



A4-1 講習会・セミナーの実施、広報活動

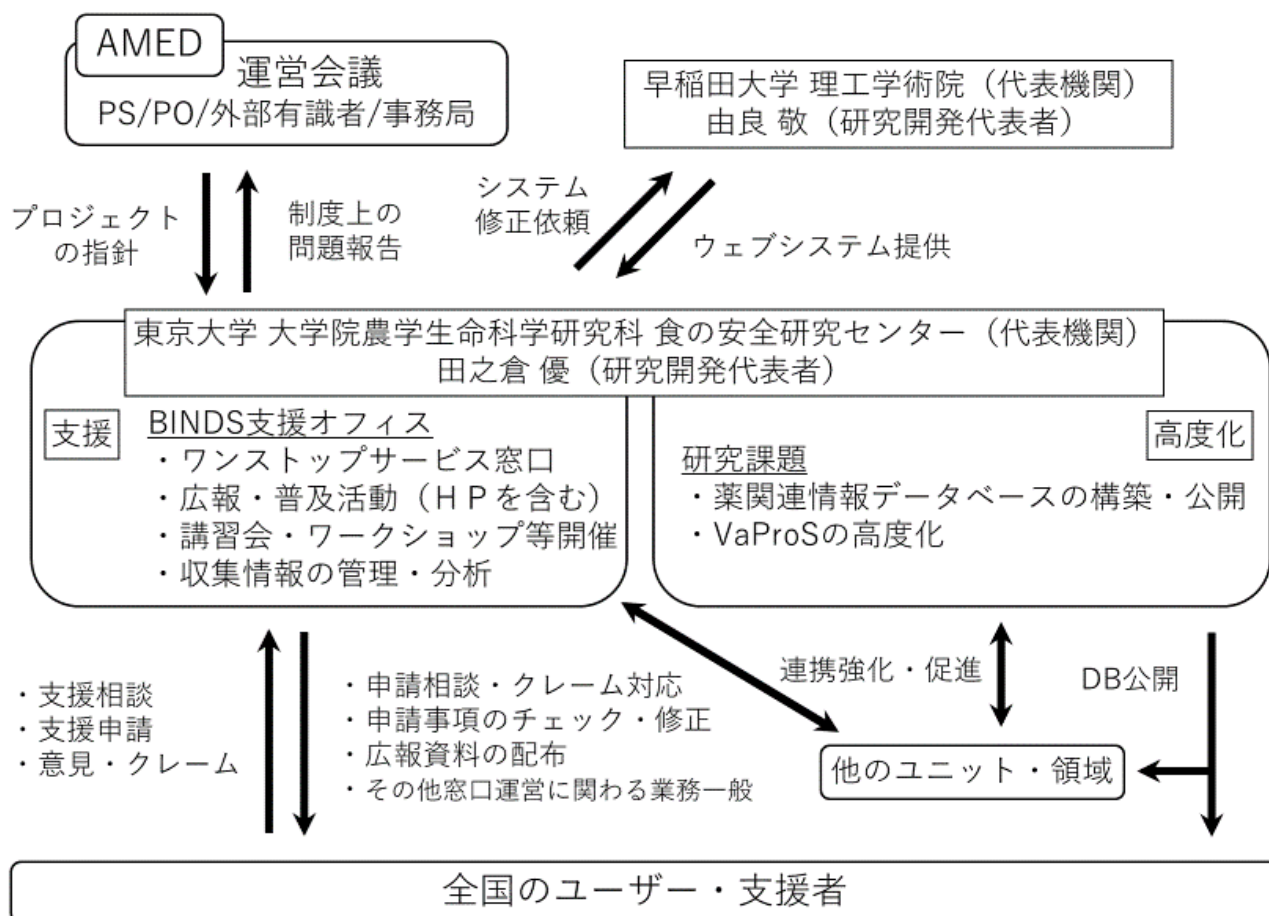
広報活動、啓蒙活動、社会還元、ならびに各ユニット間の連携促進

[1] 支援担当者

所属	①東京大学 大学院農学生命科学研究科食の安全研究センター	
氏名	①田之倉 優	
AMED 事業	ユニット／領域名 課題名	プラットフォーム機能最適化ユニット 創薬等ライフサイエンス研究を促進する研究支援とデータサイエンス
	代表機関 代表者	東京大学 田之倉 優
支援技術のキーワード		ワークショップ、講習会、ユニット間連携、支援申請、広報

[2] 支援技術の概要

BINDSのヘッドクォーター機能を担う運営会議と連携し、本事業に関する実務支援、広報活動等諸事業を推進



(1)プラットフォーム運営の実務支援

創薬やライフサイエンス研究を支援するAMED「創薬等ライフサイエンス研究支援基盤事業」の一環として構築された「創薬等先端技術支援基盤プラットフォーム(BINDS)」の円滑な運営のために、支援オフィスとして、AMEDの運営会議・事務局や各ユニットと連携し、ワンストップコンサルティング・支援申請の窓口や放射光施設ビームタイムの調整システムへの参画等実務支援を行う。またプラットフォーム内の情報交換の促進・連絡調整を行う。

(2)広報活動による社会還元

本事業の主旨、事業内容、成果等を社会に広く周知するために、BINDS事業の紹介・広報資料の発行・配布、BINDSホームページの運用、各種イベント（公開シンポジウム、ワークショップ、利用促進のための講習会・説明会など）の企画・参加・支援を通じて広報活動を展開する。

(3)プラットフォーム内の各ユニット間の連携の強化

さまざまな分野の外部研究者がかかわる「ワンストップサービス」の関連業務、領域を横断する報告会等の各種イベント等を通じて、各ユニット/領域間の連携の強化を図るとともに、創薬プロセスに資する構造生命科学のシーズやニーズの探索を通して新しい研究課題の発掘・開拓を行い、プラットフォームの発展に寄与する。

A4-2 医薬関連情報データベースの構築・公開とVaProSの高度化

医薬関連各種データベースの構築と統合、ならびにその利用に関する支援

[1] 支援担当者

所属		①東京大学 大学院農学生命科学研究科食の安全研究センター
氏名		①田之倉 優
AMED 事業	ユニット／領域名 課題名	プラットフォーム機能最適化ユニット 創薬等ライフサイエンス研究を促進する研究支援とデータサイエンス
	代表機関 代表者	東京大学 田之倉 優
支援技術のキーワード		医薬関連データベース

[2] 支援技術の概要

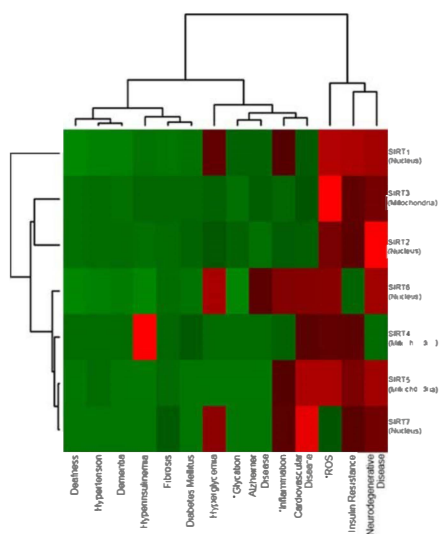
医薬関連情報データベースの構築・公開とVaProSの高度化

[3] 支援技術の利用例

- VaProSによるタンパク質可溶化データベースREFOLDdbの利用 (Mizutani et al. BMC Structural Biology, 17:4 (2017))
- 医食同源老化関連データベースの利用 (Kwon et al. PLOS ONE 12:e0183534 (2017))
- 特定の疾病に対する関連薬抽出ならびに新たな関連薬の推定（ドラッグリポジショニング）
- 各種関連性の推定

[4] 支援担当者の研究概要

これまでに、データサイエンス研究として、タンパク質リフォールディングデータベースではタンパク質のリフォールディングに関する文献情報を網羅し、老化データベースでは、老化関連用語の選定および、全てのPubMedエントリーについて老化関連用語と遺伝名との各アブストラクト内での共起を調べることににより、それらの関連性を明らかにした（図）。さらに、これらの手法を疾病と関連性検索手法に発展させ、疾病と薬の報告のある関連性のみならず、新たな関連性の推定をも行うことができるようにした。この手法の特徴は、生体物質、遺伝子、タンパク質、代謝反応などのデータベースであるKEGGデータベースから検索語を作成し、PubMedアブストラクト中での各検索語の共起関係をKEGGコード間の関連性に交換、これらの関連性をつなぎ合わせることににより、未だ報告のない関連性をも推定できるようにしたことにある。最終的な目的が疾病―薬関連性であったとしても、検索語を全てのKEGGコードから作成することにより、例えば、タンパク質を介した関連性をも推定できるようになっている。検索の際には英文の処理として必要なステマーを医学・生命科学に対応するように独自に開発するなどの工夫も行った。この検索・推定システムについては、PubMed、KEGGの最新版に対応すべくアップデートを常に行っている。そして、その重要な応用例として、COVID-19に対する関連薬（治療薬等）推定データベースを構築した。



老化に関わる各現象/疾病(横項目)と遺伝子(縦項目)の関連性(一部)

A5-1 知識ベース超分子モデリングによる疾患メカニズム解析

超分子モデリングパイプラインで立体構造モデルを構築し、疾患変異マッピングで疾患メカニズムの解析・ドラッグデザインを支援

[1] 支援担当者

所属		①長浜バイオ大学 バイオサイエンス学部
氏名		①土方 敦司、塩生 真史、白井 剛
AMED 事業	ユニット／領域名 課題名	プラットフォーム機能最適化ユニット 創薬等ライフサイエンス研究を促進する研究支援とデータサイエンス
	代表機関 代表者	長浜バイオ大学 白井 剛
支援技術のキーワード		超分子モデリング、疾患分子メカニズム解析、分子間相互作用予測

[2] 支援技術の概要

- 超分子モデリングテンプレート探索システムSIRDによる知識ベースモデリング・ドメイン=ドッキングシミュレーション・粗視化シミュレーション・複合体モデル可視化解析システムを組み合わせたモデリングパイプラインで、多面的・多角的に超分子複合体の立体構造モデル構築を支援する。SIRDはPDBのタンパク質・リガンドの構造と相互作用データの分類データを保持しており、モデルを作成したいタンパク質サブユニットのアミノ酸配列やリガンドの構造をマルチクエリとして検索することで、類似性に基づいた超分子モデリングに利用可能なテンプレート(既知構造情報)の情報が得られる(Fig.1)。
- 疾患関連変異を超分子モデル上にマッピングするツールMutation@A Glanceを高度化し、疾患メカニズムの解析やドラッグデザインを支援する。Mutation@A GlanceはdbSNP、ClinVar、CONSMICなどの疾患関連変異や多型の情報と、それらの既知タンパク質立体構造上へのマッピングの情報を保持しており、新たに発見された変異と既知変異の立体構造上での関係を解析できる(Fig.2)。
- 承認薬(Drug)、ヒトタンパク質(Target)、および疾患医療統計の情報を保持し、Drug-Target-疾患・医療統計の組み合わせを検索することができるデータベースDTX(Drug Target Excavator)を開発した。承認薬(Drug)情報には、薬効(ATCコード)・化合物分類(CASコード)・国立医薬品食品衛生研究所既存化学物質毒性データベース(JECDB)からの化合物毒性情報などが、ヒトタンパク質(Target)情報には、ホモログを含めた立体構造の解明の程度・マップされる疾患関連変異の個数・GWASにより疾患関連性が指摘された回数などの情報が含まれる。疾患医療統計は国際疾患分類(ICD-10)・医薬品医療機器総合機構(PMDA)の副作用情報、厚生労働省NDBオープンデータからのレセプト情報(薬価・医療費)・患者数・死亡原因の情報が含まれる。ターゲットタンパク質のアミノ酸配列やドラッグ候補の構造をクエリとすることで、類似したドラッグターゲット・既存ドラッグ・関連した疾患の関係をグラフとして解析できる(Fig. 3)。

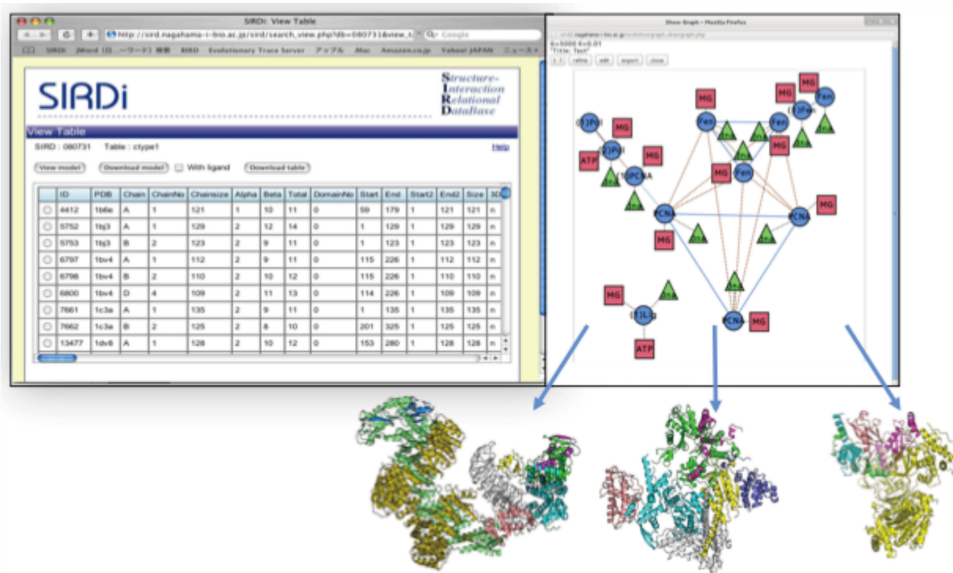


Fig. 1 SIRDシステムによる超分子モデリング

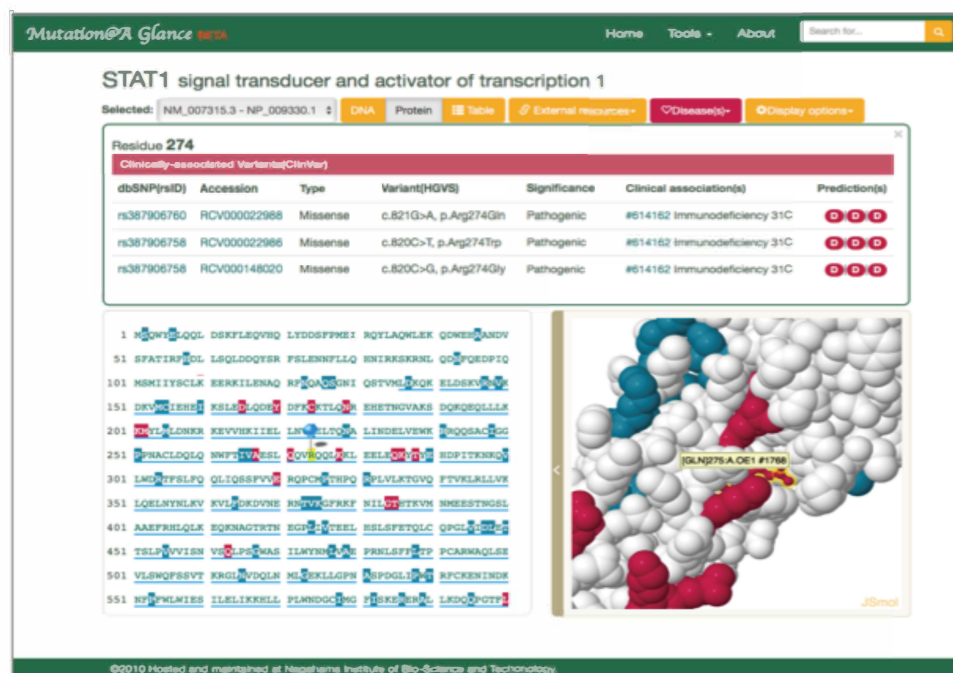


Fig. 2 疾患関連変異マッピング

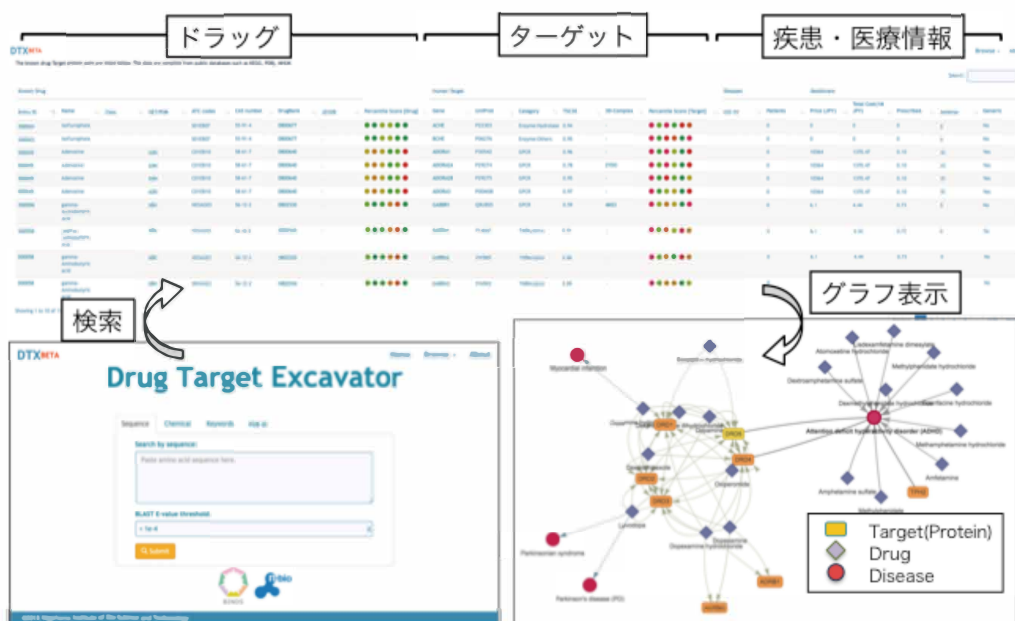


Fig. 3 DTXによるドラッグターゲット-ドラッグ-疾患の解析

[3] 支援技術の利用例

- 超分子モデリングパイプラインを構築し、疾患関連遺伝子産物のモデリングによる疾患メカニズム解析を支援した。これまでにプリオン病(BSE・CJD)の原因とされるPrP^{Sc}の構造 (東北大学、*Structure*, 22, 560-571, 2014)、Aicardi-Goutieres症候群関連タンパク質IFIH1-RNA複合体(京都大学、*Am. J. Hum. Genet.* 95, 121-125, 2014)、カンジダ症の原因遺伝子STAT1 (理研、*J. Biol. Chem.* 292, 6240-6254, 2017)など、50件以上の疾患関連分子モデルを構築して支援研究者に提供した。
- 古細菌から新規に発見された、バックグラウンド配列に依存せず塩基ミスマッチDNAを特異的に切断するエンドヌクレアーゼ EndoMS(NucS)-PCNA-DNA複合体モデルを構築し、ミスマッチ塩基認識メカニズムの解析を支援した。この支援研究はEndoMSを用いたPCRによる簡便な遺伝子診断技術の開発研究に発展した (九州大学・タカラバイオ株式会社、*NAR* 2016:Breakthrough Article, *Structure* 2017, 「耐熱性ミスマッチエンドヌクレアーゼの利用方法」 PCT/JP2015/075603)(Fig. 4)。
- 超分子モデリングパイプラインを用いたモデリングにより、flap endonuclease (FEN)-DNA ligase (Lig)-PCNA(DNA clamp)-DNA複合体など、クライオ電顕解析における超分子モデル構築を支援した(*Sci. Rep.* 8:16209, 2018) (Fig. 5)

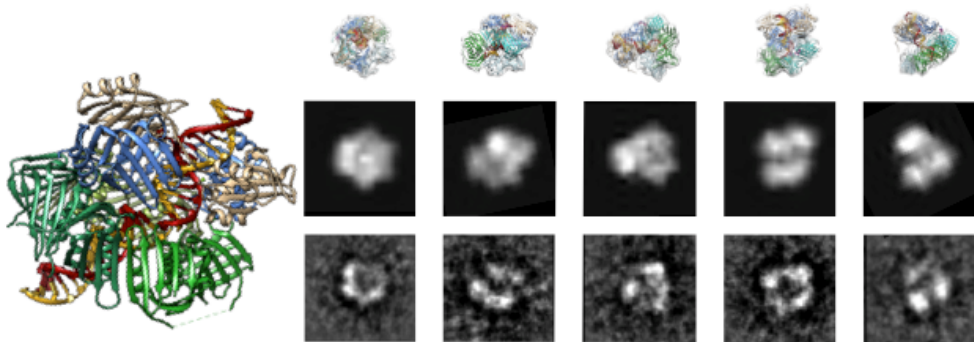


Fig.4 EndoMS(NucS)-PCNA-DNA複合体モデル(左)と電顕像との比較(右)

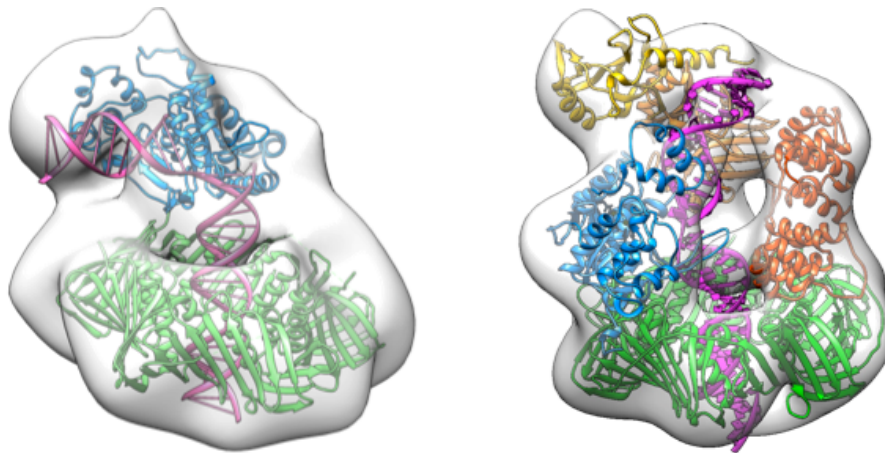


Fig.5 FEN-PCNA-DNA複合体モデル(左)とFEN-Lig-PCNA-DNA複合体モデル(右)の電顕マップへのフィッティング

[4] 支援担当者の研究概要

- 構造データベースと相互作用データベースの統合による知識ベース超分子モデル構築法を開発した。この方法により約4,500のヒトタンパク質超分子モデルを構築し、1,001の新規PPI(Protein-Protein Interaction)インターフェースを予測し、10種のヒト疾患についてPPIの関与を始めて示した (Sci Rep. 5:16341, 2015、IIBMP2015 Excellent Research Award)。Mutation@A Glanceを利用し、これらのモデルの相互作用インターフェースと疾患関連変異の形質発現様式(ドミナントネガティブ、機能獲得、ハプロ不全)の相関を解析したところ、機能獲得変異はドメイン間相互作用、ハプロ不全はタンパク質-DNA相互作用に対するオッズ比が高いことが判明した(Sci Rep. 7:8541, 2017)。また、超分子複合体内の複数のサブユニットにGWAS(全ゲノム相関解析)データをマッピングし、同一GWAS形質が有意に多くマップされる超分子モデル(疾患関連超分子)を1,226個特定した。これらのモデル中には、ある疾患や形質に関連した変異がマップされた複数のサブユニットと相互作用していながら、自身には同じ疾患や形質のアノテーションが付与されていない(未発見の疾患関連遺伝子の可能性がある)サブユニットが174種類存在した。これらのサブユニットの多くは、リン酸化シグナル伝達系の必須遺伝子にコードされるハブタンパク質であることが示された。また、複数のGWAS形質がアサインされるが、関連変異がマップされるサブユニットがGWAS間で重複しないモデルは642個存在した。これらのモデルは異なる疾患の関連性を解析するために利用できると考えられる(Fig. 6)。
- プラットフォーム機能最適化ユニットで開発中のVaProSの高度化を目的として、タンパク質配列やドラッグ構造をクエリとして、クエリ配列の立体構造の新規性やドラッグターゲットとしての新規性などの調査を行うデータベースシステムDrug Target Excavator (DTX)を開発した。新規創薬ターゲットの発見のために、DTXを用いて「疾患」-「疾患関連遺伝子(タンパク質)」-「既知のDrug Target」?「承認薬(Drug)」の間の関連性について調査を行った結果、合理的創薬の典型的なパターンである「疾患関連遺伝子が既知であり、承認薬が存在し、なおかつターゲットは疾患関連遺伝子(タンパク質)そのものか、それと直接相互作用するか、あるいは同一パスウェイ上に存在する」に該当する疾患は、全体の8%であった。最も多いパターンは、ターゲット発見がボトルネックになっていると考えられる「疾患関連遺伝子(タンパク質)が既知であるが、承認薬が存在せずターゲットも未知である」場合であり、全疾患の63%がこれに該当した。興味深い点として、「疾患関連遺伝子(タンパク質)が既知であり、承認薬が存在しターゲットも既知であるが、疾患関連遺伝子とターゲットの分子間相互作用は自明ではない」パターンが3%程度存在し、これらを解析することで、効率的ターゲット探索法へのヒントが得られるのではないかと予想される(Fig.7)。

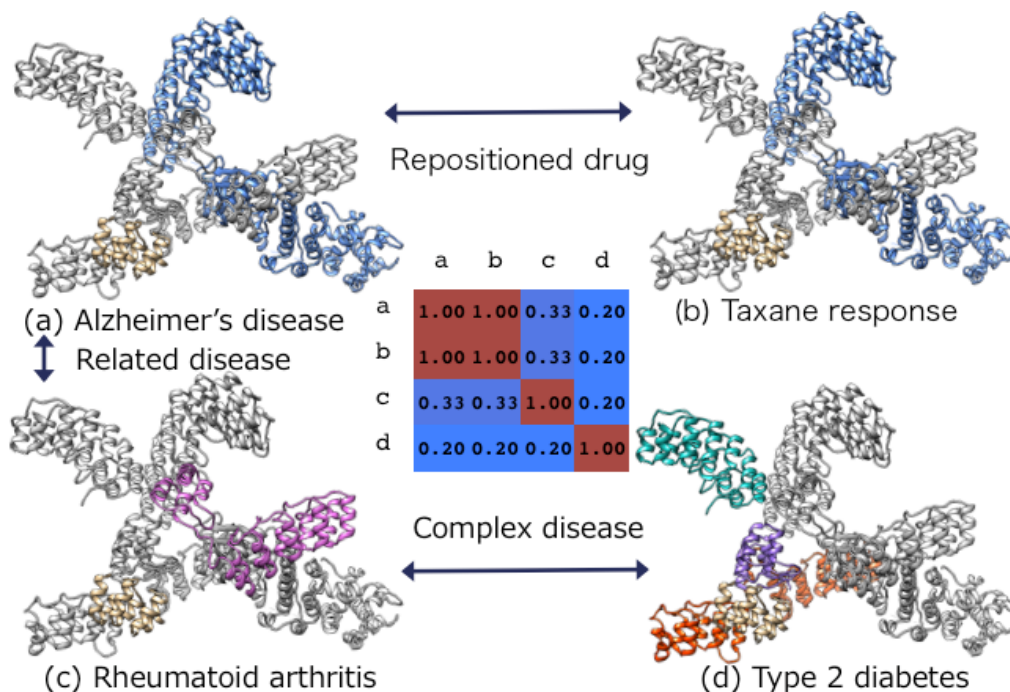


Fig.6 超分子モデルへのGWASマッピング

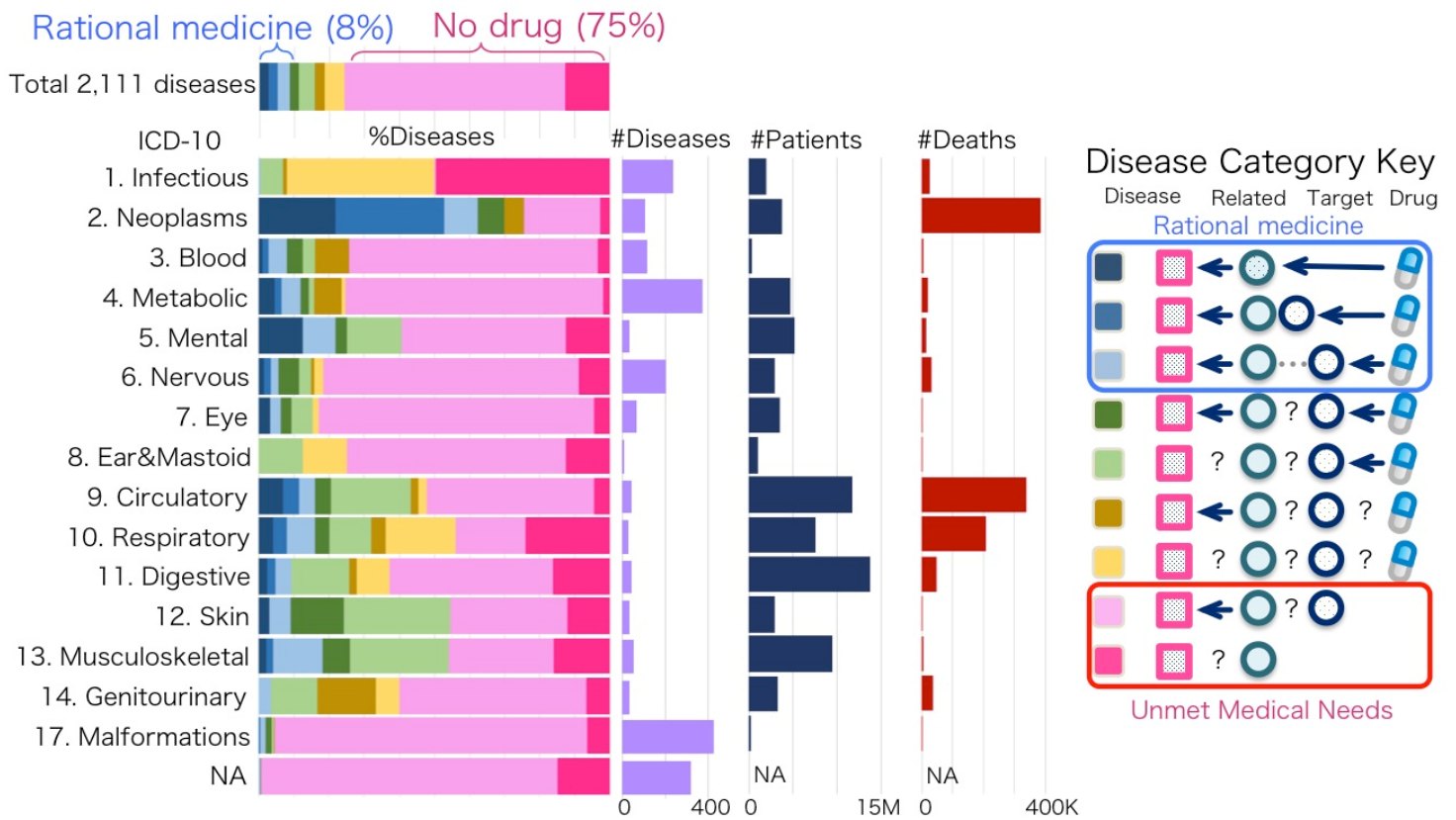


Fig.7 DTXによる関連遺伝子-疾患-ドラッグ解析