

科学研究費助成事業 研究成果報告書

令和 4 年 5 月 25 日現在

機関番号：12601

研究種目：新学術領域研究(研究領域提案型)

研究期間：2017～2021

課題番号：17H06307

研究課題名(和文)次世代ヒト全ゲノム・オミクスの解析方法論の開発と応用

研究課題名(英文)Development and application of analysis methodology for next generation human omics

研究代表者

角田 達彦(Tsunoda, Tatsuhiko)

東京大学・大学院理学系研究科(理学部)・教授

研究者番号：10273468

交付決定額(研究期間全体)：(直接経費) 78,400,000円

研究成果の概要(和文)：代謝アダプテーションで疾患へ至る全体像と挙動をとらえるには、マルチオミクスデータを、階層をまたいで統合する技術が必要である。そこで本研究では、トランスオミクス統計解析手法を知識型、因果・階層型、統合型に基づき整備し、オミクスデータを解析することで、トランスオミクスを読み解き、疾患や生体における代謝状態の変化を体系的網羅的に明らかにした。その結果、深層学習に基づく新たな方法論を導き、また臨床検体とモデル動物を用いて、糖尿病、アルツハイマー、がん免疫・微小環境のトランスオミクス解析を行い、ヒトとマウスにまたがる新規遺伝子やパスウェイ、相互作用、新たな亜分類やその背景となる機序を発見した。

研究成果の学術的意義や社会的意義

本研究により、オミクスの多階層データを繋ぐ方法論の基礎を築いた。知識型、因果・階層型、統合型という柱を基本とし、中でも特に派生した深層学習によるオミクス解析の方法論の進展は、世界でも唯一無二である。また組織検体からゲノムや遺伝子発現などのプロファイルからがん微小環境をトランスオミクス解析し、免疫編集を解明した。本研究を基盤とし、個体レベルのネットワーク再構築により、糖代謝疾患などの多因子疾患やがん転移などの環境適応の全体像の解明、薬剤の動的な選択などのプレシジョン医療の基盤の形成、複数の創薬標的分子のネットワークによる同定、そして新規治療法開発に繋がるのが期待できる。

研究成果の概要(英文)：To capture the whole picture and behavior leading to diseases by metabolic adaptation, it is necessary to integrate multi-omics data across the omics layers. In this study, we arranged trans-omics statistical analysis methods based on knowledge-based, causal-hierarchical, and integrated methods, and analyzed omics data to decipher trans-omics and systematically and comprehensively clarify changes in metabolic states in diseases and living organisms. As a result, we derived a new original methodology based on deep learning and conducted trans-omics analysis of diabetes, Alzheimer's disease, cancer immunity and tumor-microenvironment using clinical samples and model animals to discover novel genes, pathways, interactions, new sub-classes and their underlying mechanisms across humans and mice.

研究分野：オミクス医科学

キーワード：トランスオミクス 代謝アダプテーション オミクス解析

1. 研究開始当初の背景

生命は環境に応じて動的に代謝を調整することにより恒常性を維持している。代謝疾患、がん、炎症性疾患や薬剤耐性などで見られる特有の代謝状態は、それぞれの環境変化に対して、生体が代謝を調整して適応した結果（代謝アダプテーション）である。代謝アダプテーションは、状況に応じてトランスオミクスネットワークを動的に切り替えることでなされる。しかし、生体の基底状態から疾患へ至る全体像と挙動はとらえられていなかった。その原因は、従来の各種オミクス計測・解析は、それぞれの階層単独で開発されており、階層を繋ぐことが考えられていないからである。そこで、マルチオミクスデータを、階層をまたいで統合する技術が必要である。

ところで研究代表者らは、世界初の GWAS (*Nature Genetics* 2002)、国際 HapMap 計画への参画と全世界の GWAS の加速に貢献 (*Nature* 2007)、次世代シーケンサーを用いた世界初の日本人全ゲノムの解析 (*Nature Genetics* 2010) に携わってきた。そして国際がんゲノムコンソーシアム、特に肝がん全ゲノム解析を導き (*Nature* 2010、*Nature Genetics* 2012、*Nature Communications* 2015、*Nature* 2020)、日本の大きな貢献をしたものとして、肝がん 300 症例の全ゲノム解析の成果とし、肝がんの予後に対する患者層別化に成功した (*Nature Genetics* 2016)。特にゲノム変異にタンパク質間相互作用を加えて情報量を増し、クラスタリングを行ったところ、再発が非常に起こりにくい新たなクラスタを発見した。さらにメチル化とトランスクリプトームのデータも取得し、融合遺伝子の同定結果を加え、全オミクスにまたがるデータとして同時に扱い、クラスタリングを行った。するとゲノムのみの場合よりも、再発、全生存期間ともに、クラスタ分類とのより有意な相関が認められた。おそらくは、がん細胞の状態をより色濃く反映しているものと推察した。しかし、その解釈と普遍性にあたり、オミクス層間をより有機的に分子的に繋げる方法の必要性を覚え、多階層オミクス間を繋げる統計的方法の整備と解析という着想に至った。

2. 研究の目的

トランスオミクス統計解析手法を整備し、オミクスデータを解析することで、トランスオミクスを読み解き、疾患や生体における代謝状態の変化を体系的網羅的に明らかにすることを目標とした。ヒトのデータを中心とした解析、そしてマウスなどの知見をヒトへマッピング・変換し、疾患特異的機序を探ることが主軸である。そこで以下の 3 つを柱とした。

- (1) 知識型、因果・階層型、統合型の 3 種類のトランスオミクス統計解析方法の整備と実証。
- (2) 実データの統計的解析による疾患の多因子バイオマーカーや多因子標的分子の候補の推定。
- (3) 実データによるがん、炎症疾患、糖代謝、薬剤応答のネットワーク解明と医学応用基盤構築。これらにより、本領域にトランスオミクス解析アルゴリズムを提供し、技術面と、解析・解読・解釈での統計解析面で貢献することを目的とした。

3. 研究の方法

- (1) トランスオミクス解析の方法論とし、知識型、因果・階層型、統合型を整備した。
 - ① 知識型：ゲノムやオミクス解析の課題の一つは、偶然により原因候補が多数挙がることである。真の原因に迫るため、トランスクリプトームや染色体構造などのアノテーションデータにより、ゲノムから代謝にいたる制約を見出すことで疾患原因を絞る体系的手法を整備した。またそのための公共のアノテーション情報や分子データを外部から収集した。
 - ② 因果・階層型：オミクスの違いが代謝に及ぼす影響を調べるため、ゲノムとトランスクリプトーム間から始まる層間の分子の依存関係を様々な組織・臓器で求め、ネットワークにマップすることで代謝や疾患に関わるゲノム多様性を同定した。既存の代謝パスウェイデータベースに加え、タンパク質間相互作用をもとにしたネットワークによる解析手法を整備した。
 - ③ 統合型：疾患のヘテロ性と原因の相補性を解くため、高次元空間からの次元圧縮、特徴抽出、特徴選択を行う数学的手法と、多階層オミクスデータを同時に扱う枠組みについて整備した。特に、深層学習により特徴抽出を行い、疾患の原因もしくは関連する因子とパスウェイ・ネットワークを同定する手法を独自・新規に開発した。
- (2) 上記の方法論を用い、シグニチャーとフェノタイプとを相関解析すること、および疾患トランスオミクスネットワークを統計的に解析することにより、疾患の多因子バイオマーカーや多因子標的分子の候補を推定した。そして、パスウェイにマッピングし統合することによって、疾患や生体におけるパスウェイや状態の変化を明らかにした。
- (3) 実データに対し、統計的な解析を主とするデータ駆動型の方法を用いて解析し、疾患や免疫系の挙動、代謝アダプテーションのメカニズムの解明に貢献した。以下、主なものを挙げる。

① 解析対象の一つとして、代謝疾患の代謝アダプテーション（黒田班）の解析として、経口糖負荷前後の野生型および代謝疾患のモデルマウスからバイオマーカーを抽出し、ネットワーク解析を行い、多因子バイオマーカーモデルを構築した。このために、数十万人のヒト GWAS 症例とマウスのオミクスデータを用いた。同様の枠組みをアルツハイマー病の解析にも適用すべく、ヒト GWAS とモデルマウス実験データを整備した。

② 次に、がんの全体像の解明には、従来のがん細胞集団の解析に加え、それらの周囲の環境も重要である。本トランスオミクス計画では、そのような環境のオミクスも同時にとらえ解析できることから、がん免疫・微小環境の臨床検体のがんのデータを、臨床医との調整、フィージビリティスタディと検証実験を行った上で取得した。そしてがん免疫・微小環境の臨床検体のオミクス状態の変化を網羅的に明らかにし、ネットワーク解析を行い、多因子バイオマーカーモデルを構築した。

4. 研究成果

(1) 解析の方法論に関し、知識型解析手法では、トランスクリプトーム、染色体構造などのアノテーションデータにより、ゲノムから代謝にいたる制約を見出すことで因子候補の制約・絞り込みを行う体系的手法を整備した。そのために、組織・臓器特異的発現 (GTE_x, GEO)、ゲノム構造、SNP/SNV/変異 (WGS, WES, COSMIC, dbSNP) などの公共 DB を用いた。

(2) 因果・階層型手法では、計測データを用い、ゲノムとトランスクリプトーム間から始まる層間の分子の依存関係を同定する方法を基礎にした。様々な細胞種の eQTL, GTE_x (全身の各組織部位) による組織特異的なゲノム配列と遺伝子発現の階層構造を用いてトランスオミクス解析を行う方法、発現から免疫細胞プロファイルを推定する方法を整備し、また、ゲノム配列データからタンパク質・アミノ酸配列構造・修飾を高精度に予測する手法を開発した (*Scientific Reports* 2018, *BMC Genomics* 2019)。翻訳後修飾が生む多様なプロテオームは、様々な病因となる。近年多くの翻訳後修飾が報告される中でも、リシンホスホグリセリル化が最も注目されるが、それを検出する実験的方法は高価で時間がかかり、非効率的である。そのためホスホグリセリル化リジンを正確に予測する計算方法が提案されているが、性能は依然として低かった。そこで、それを予測する新しい2つの方法を提案し、データとソフトウェアを公開した。一つ目は、多層パーセプトロン分類子とアミノ酸の構造情報 (アクセス可能な表面積、バックボーンのねじれ角、局所構造の立体配座) を利用する新しい手法で (*Scientific Reports* 2018, 図1)、二つ目は、アミノ酸に関する構造情報と進化的情報を利用する手法である (*BMC Genomics* 2019)。三つ目は、位置特異的スコア行列を用いる手法である (*BMC Molecular and Cell Biology* 2019)。これらに加え、アミノ酸残基行列やアミノ酸の構造的な属性によりリシンのホスホグリセリル酸化翻訳後修飾を予測する手法 (*Genes* 2020a, *Genes* 2020b)、HMM によってシングルストランドとダブルストランドの DNA にタンパク質が結合するかを予測する手法 (*Analytical Biochemistry* 2020)、また深層学習によってタンパク質とペプチドが結合する部位を予測する手法 (*Journal of Theoretical Biology* 2020) を開発した。いずれの手法も、以前の予測手法と比べ有意に優れていた。

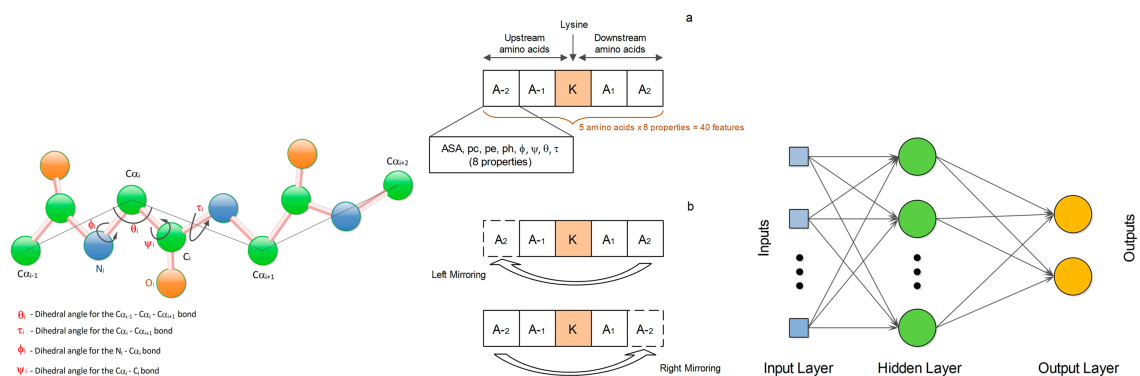


図1. ホスホグリセリル化リジンを予測する、物理的構造と神経回路を用いた新しい手法。

(3) 統合型解析手法では、高次元空間からの次元圧縮、特徴抽出、特徴選択を行う数学的手法と、多階層オミクスデータを同時に扱う枠組みを構築し、疾患のヘテロ性と原因の相補性を解く手法の一つを提案した。タンパク質間相互作用、パスウェイ、Gene Ontology ネットワークによって分子間相互作用を統合して解析する方法を実現した。ビッグデータを解析すると、疾患の亜分類に関わる分子シグネチャが同定でき、発症を機構的に解釈できるようになるが、ヘテロなトランスオミクスデータの統合解析が必要である。そのための知識表現として、経路マッピング、分子ネットワーク、知識グラフの3つのアプローチについて論じ、それぞれの有用性、応用方法、課題を、データを用いて論じた (*Briefings in Bioinformatics* 2018, 図2)。

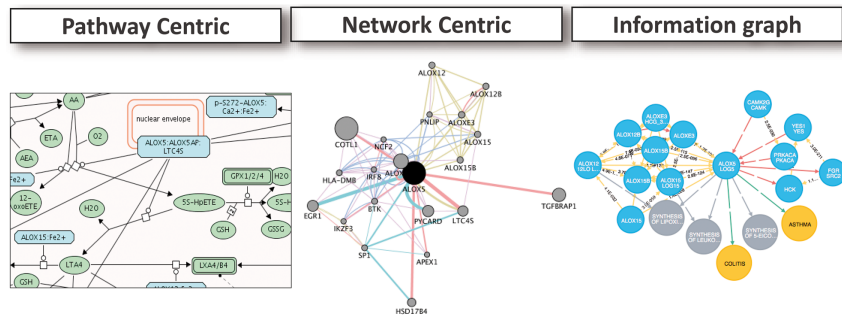


図 2. トランスオミクスデータの統合解析の3つのアプローチ.

(4) そして全く新たな試みとして、オミクスデータを深層学習で解析・学習するためのオミクス等の非画像データを画像データに変換して用いる方法論と深層学習を活用したオミクス解析を開発した (*Scientific Reports* 2019、図 3 左)。特に高次元空間からの次元圧縮、特徴抽出、特徴選択を行う方法を深層学習で確立し、がん解析に適用し、特徴づけを行なった (*Briefings in Bioinformatics* 2021、図 3 右)。これらの方法論により、深層学習で疾患に特有な特徴の抽出と科学的な発見が可能であることを示した。

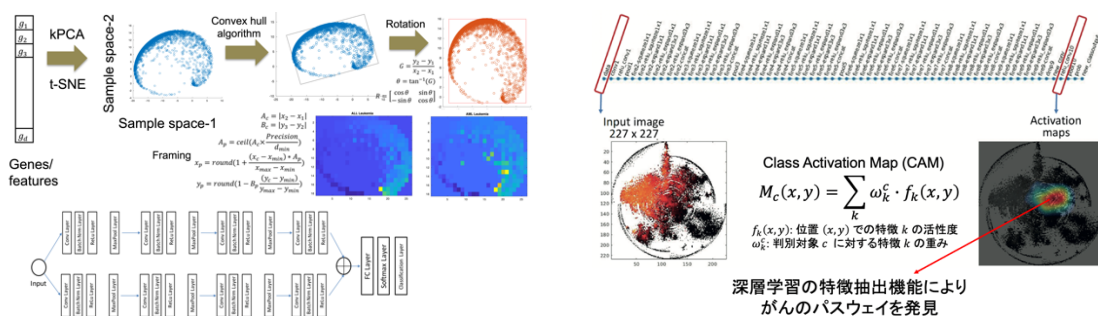


図 3. 深層学習でオミクスデータを扱い疾患の特徴の抽出・科学的発見を導く独自の的方法論.

(5) また miRNA 発現と機械学習による認知症のレビー小体型の診断手法と実データによる解析・予測モデル構築 (*BMC Medical Genomics* 2019)、そしてエクソームシーケンスデータによる腫瘍転移時の細胞数の定量推定を行う手法を開発した (*International Journal of Cancer* 2020)。

(6) 開発した手法を用いた臨床検体と実験モデルを用いたデータの実際の解析として、従来のゲノムワイド関連解析 (GWAS) を発展させ、遺伝子発現などのオミクスデータとの統合、大規模国際化、分子データベース統合による機能メカニズムの解明という新しい方法論を展開した。その一環としてまず、アルツハイマー病の新規の原因遺伝子を同定するため、ヒト GWAS とマウス脳のトランスクリプトーム解析を組み合わせる独自手法を提案し、実データを用いた解析を行った結果、新たな2つの原因遺伝子を発見した (*Human Genetics* 2018)。マウスでは、アルツハイマー病モデルマウスで各遺伝子の発現量と $A\beta$ 蓄積量との相関を調べ、ヒトでは日本人のアルツハイマー病の GWAS データを活用した。マウスとヒトとの同祖遺伝子ペアで結合し、統合した。その結果、統計的に有意である複数の遺伝子を検出した。これらの遺伝子に対し剖検脳を用いてアルツハイマー病患者群と対照群との間での遺伝子発現量の差をみた結果、2つの遺伝子で、対照群に比べ患者群で有意に発現量が低くなることが分かった。またこのプロジェクトではそれらの結果を GTE_x によるトランスオミクス解析により強固な知見にした。さらにこのプロジェクトでは国際コンソーシアム IGAP のデータも活用してマウスデータと統合し、これらの遺伝子に加え新たなアルツハイマー病関連遺伝子候補も挙げる事ができた。このように、ヒトとモデル動物からの異なるオミクスデータを統合解析および階層をまたがる解析をすると、ヒトの GWAS だけでは検出されなかった新しい疾患遺伝子を発見できる可能性がある。この方法論は一般性が高く、黒田班とのヒト GWAS とマウス代謝による糖尿病の共同研究の基礎になった。

(7) 実データを用いたもう一つの方向性とし、薬剤安全性予測モデルを開発した。近年、医薬品開発はコストの上昇と新薬承認の減少による収益の減少が顕著である。臨床試験中での開発中止や、市場発売後の販売中止の大きな原因は薬剤毒性である。これを予測できる計算方法があれば、膨大な労力と時間の無駄を減らせる。そこで、薬物標的と非標的の性質、Gene Ontology で計算した機能的影響スコア、分子相互作用ネットワークを使って薬物毒性を予測する独自の機械学習方法を提案した (*Life Science Alliance* 2018、図 4)。毒性のある薬物を安全な薬物と区別でき、また異なる薬剤を組み合わせるとき最も低毒性になるように設計を行うことも可能になる。

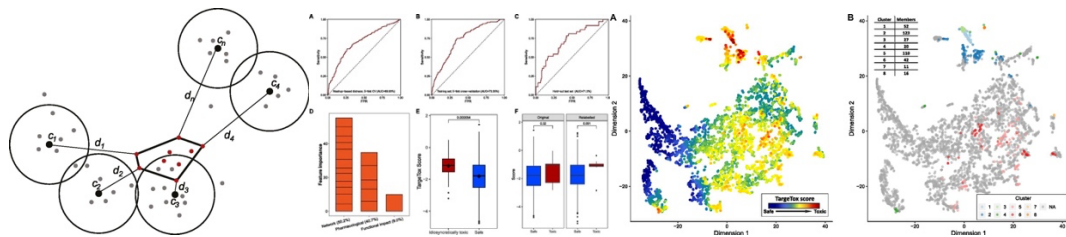


図 4. 薬物標的の性質、機能的影響、分子ネットワーク、機械学習による薬物毒性の予測手法。

(8) 糖尿病に関する実データ解析として、ヒトのゲノムワイド関連解析とマウスを用いた代謝疾患解析を行い、実データ解析手法の確立を進めた。その結果、糖尿病の代謝オミクスでは黒田班と共同でヒト GWAS とマウスのトランスオミクスの情報乗り入れによる新規発見に至り（医学のあゆみ 2021）、論文化中である。またこれらを総合し、ゲノムワイド関連解析の発展形としてのトランスオミクス解析に関する総説を出版した。研究の一つとして、I型もしくはインスリン抵抗性糖尿病の診断方法の解析に貢献できた（*Diabetes Research and Clinical Practice* 2020）。

(9) がん免疫に関し、臨床検体の WES と RNA-Seq データから、がん浸潤免疫細胞の特性を推測する、がん免疫解析パイプラインの計画と開発を行った。それを進行性大腸がんの 89 検体のオミクスデータに適用しトランスオミクス解析を行なったところ、免疫回避と非常に悪い全生存率を特徴とする独特のがんのサブタイプを新たに発見した（図 5; *iScience* 2022）。この分類の患者の予後が極端に悪い原因を免疫編集の観点から解析すると、このサブタイプでは、ネオ抗原のクローン性高発現が少なく、染色体不安定性が高いため、免疫チェックポイント分子を介した適応免疫抵抗性やネオアンチゲン提示異常が生じる。また、免疫再編成によるネオ抗原の枯渇とクローン性ネオアンチゲンの高負荷は、良好な全生存率と相関していることが観察された。これらから、腫瘍の微小環境の状態とネオ抗原の組成が、進行性大腸がんの治療計画決定に関連する可能性のある、有望な新しい予後バイオマーカーであることを示すことができた。さらにこれらの方法論を胃がんや腎がんなどの他の各がん種の免疫学的解析に応用し、新規発見を報告した。

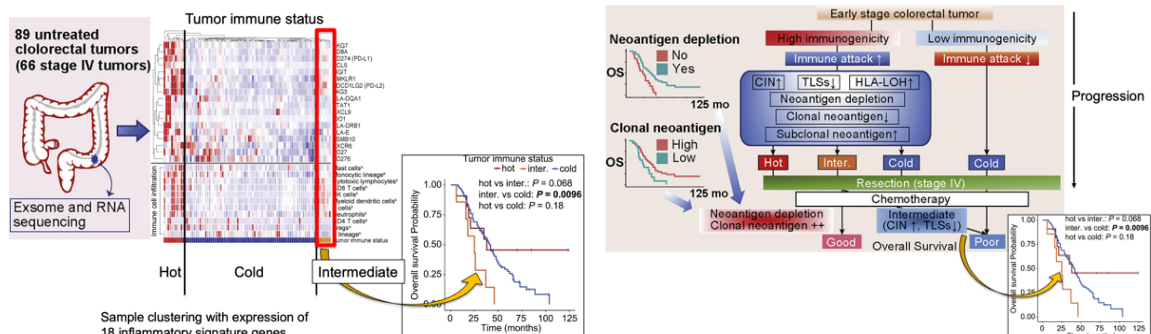


図 5. 進行性大腸がんの免疫学的解析による新たな分類・機序の発見と予後との関連。

(10) 国内外の位置づけとインパクトとして、このプロジェクトのトランスオミクス解析は国内外で類を見ない。その中で、知識型、因果・階層型、統合型の考え方で進めることができた。これらにはアノテーション、多くの検体の遺伝子型と表現型の対のデータ、そしてネットワーク解析や深層学習が基盤となるが、特に深層学習による解析方法論の進展は世界でも唯一無二である。またこれらを基盤とし実データで得られた成果はいずれも科学的インパクトが大きい。とくに、研究開始当初予期していなかったこととして、深層学習の発展とそのオミクスへの適用可能性、また免疫や線維芽細胞、血管などを含むがん微小環境の解析の進展などがあげられる。新たなメトリック生成によるオミクスデータの深層学習への適用と科学的発見への応用は本研究代表者が考え出したもので世界的に反響が大きい。また上記大腸がんの解析などのがん微小環境の解析では、患者由来の組織検体のオミクスに免疫細胞が浸潤していることを利用する解析であり、ゲノムや遺伝子発現など多くの分子が複雑に絡み合う現象だが、今回は免疫編集という観点からまとめることができ、トランスオミクス解析がちょうど時期を得たものとなった。

(11) 今後の展望として、トランスオミクスのより深い階層・より複雑な構造のデータの解析や、それらから得られる科学的発見についてさらなる発展が期待できる。すなわち、1細胞データや空間オミクスデータ、またそれらの動態などの新たな次元を持つデータが多く検体で蓄積されることが期待され、それらに対する本研究の適用方法の発展と、新たな細胞、細胞間・分子間相互作用、ネットワークが見出される。そのような一連の研究の結果、個体レベルのネットワーク再構築により、糖代謝疾患などの多因子疾患やがんの環境適応の全体像の解明や、薬剤の動的な選択などのプレジジョン医療の基盤の形成が期待される。また、複数の創薬標的分子をネットワークで同定できれば、新規治療法開発につながると思われる。

5. 主な発表論文等

〔雑誌論文〕 計21件（うち査読付論文 20件 / うち国際共著 15件 / うちオープンアクセス 16件）

1. 著者名 DU Ran, Xie Shiqi, Fang Yuxin, Igarashi-Yokoi Tae, Moriyama Muka, Ogata Satoko, Tsunoda Tatsuhiko, Kamatani Takashi, Yamamoto Shinji, Cheng Ching-Yu, Saw Seang-Mei, Ting Daniel, Wong Tien Y., Ohno-Matsui Kyoko	4. 巻 -
2. 論文標題 Deep Learning Approach for Automated Detection of Myopic Maculopathy and Pathologic Myopia in Fundus Images	5. 発行年 2021年
3. 雑誌名 Ophthalmology Retina	6. 最初と最後の頁 -
掲載論文のDOI (デジタルオブジェクト識別子) 10.1016/j.oret.2021.02.006	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 該当する
1. 著者名 Chandra Abel Avitesh, Sharma Alok, Dehzangi Abdollah, Tsunoda Tatushiko	4. 巻 11
2. 論文標題 RAM-PGK: Prediction of Lysine Phosphoglycerylation Based on Residue Adjacency Matrix	5. 発行年 2020年
3. 雑誌名 Genes	6. 最初と最後の頁 1524 ~ 1524
掲載論文のDOI (デジタルオブジェクト識別子) 10.3390/genes11121524	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 該当する
1. 著者名 Singh Vineet, Sharma Alok, Dehzangi Abdollah, Tsunoda Tatushiko	4. 巻 11
2. 論文標題 PupStruct: Prediction of Pupylylated Lysine Residues Using Structural Properties of Amino Acids	5. 発行年 2020年
3. 雑誌名 Genes	6. 最初と最後の頁 1431 ~ 1431
掲載論文のDOI (デジタルオブジェクト識別子) 10.3390/genes11121431	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 該当する
1. 著者名 Hosoe Jun et al.	4. 巻 169
2. 論文標題 Clinical usefulness of multigene screening with phenotype-driven bioinformatics analysis for the diagnosis of patients with monogenic diabetes or severe insulin resistance	5. 発行年 2020年
3. 雑誌名 Diabetes Research and Clinical Practice	6. 最初と最後の頁 108461 ~ 108461
掲載論文のDOI (デジタルオブジェクト識別子) 10.1016/j.diabres.2020.108461	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Sharma Ronesh, Kumar Shiu, Tsunoda Tatsuhiko, Kumarevel Thirumananseri, Sharma Alok	4. 巻 612
2. 論文標題 Single-stranded and double-stranded DNA-binding protein prediction using HMM profiles	5. 発行年 2021年
3. 雑誌名 Analytical Biochemistry	6. 最初と最後の頁 113954 ~ 113954
掲載論文のDOI (デジタルオブジェクト識別子) 10.1016/j.ab.2020.113954	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 該当する

1. 著者名 Wardah Wafaa, Dehzangi Abdollah, Taherzadeh Ghazaleh, Rashid Mahmood A., Khan M.G.M., Tsunoda Tatsuhiko, Sharma Alok	4. 巻 496
2. 論文標題 Predicting protein-peptide binding sites with a deep convolutional neural network	5. 発行年 2020年
3. 雑誌名 Journal of Theoretical Biology	6. 最初と最後の頁 110278 ~ 110278
掲載論文のDOI (デジタルオブジェクト識別子) 10.1016/j.jtbi.2020.110278	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 該当する

1. 著者名 Mason Mike J. et al.	4. 巻 34
2. 論文標題 Multiple Myeloma DREAM Challenge reveals epigenetic regulator PHF19 as marker of aggressive disease	5. 発行年 2020年
3. 雑誌名 Leukemia	6. 最初と最後の頁 1866 ~ 1874
掲載論文のDOI (デジタルオブジェクト識別子) 10.1038/s41375-020-0742-z	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 該当する

1. 著者名 Nishino Jo, Watanabe Shuichi, Miya Fuyuki, Kamatani Takashi, Sugawara Toshitaka, Boroevich Keith A., Tsunoda Tatsuhiko	4. 巻 146
2. 論文標題 Quantification of multicellular colonization in tumor metastasis using exome sequencing data	5. 発行年 2020年
3. 雑誌名 International Journal of Cancer	6. 最初と最後の頁 2488 ~ 2497
掲載論文のDOI (デジタルオブジェクト識別子) 10.1002/ijc.32910	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Chandra Abel, Sharma Alok, Dehzangi Abdollah, Shigemizu Daichi, Tsunoda Tatsuhiko	4. 巻 20
2. 論文標題 Bigram-PGK: phosphoglycerylation prediction using the technique of bigram probabilities of position specific scoring matrix	5. 発行年 2019年
3. 雑誌名 BMC Molecular and Cell Biology	6. 最初と最後の頁 57
掲載論文のDOI (デジタルオブジェクト識別子) 10.1186/s12860-019-0240-1	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 該当する

1. 著者名 Shigemizu Daichi, Akiyama Shintaro, Asanomi Yuya, Boroevich Keith A., Sharma Alok, Tsunoda Tatsuhiko, Sakurai Takashi, Ozaki Kouichi, Ochiya Takahiro, Niida Shumpei	4. 巻 12
2. 論文標題 A comparison of machine learning classifiers for dementia with Lewy bodies using miRNA expression data	5. 発行年 2019年
3. 雑誌名 BMC Medical Genomics	6. 最初と最後の頁 150
掲載論文のDOI (デジタルオブジェクト識別子) 10.1186/s12920-019-0607-3	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 該当する

1. 著者名 Sharma Alok, Vans Edwin, Shigemizu Daichi, Boroevich Keith A., Tsunoda Tatsuhiko	4. 巻 9
2. 論文標題 DeepInsight: A methodology to transform a non-image data to an image for convolution neural network architecture	5. 発行年 2019年
3. 雑誌名 Scientific Reports	6. 最初と最後の頁 11399
掲載論文のDOI (デジタルオブジェクト識別子) 10.1038/s41598-019-47765-6	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 該当する

1. 著者名 Chandra Abel Avitesh, Sharma Alok, Dehzangi Abdollah, Tsunoda Tatsuhiko	4. 巻 19(Supp19)
2. 論文標題 EvoStruct-Phogly: incorporating structural properties and evolutionary information from profile bigrams for the phosphoglycerylation prediction	5. 発行年 2019年
3. 雑誌名 BMC Genomics	6. 最初と最後の頁 984
掲載論文のDOI (デジタルオブジェクト識別子) 10.1186/s12864-018-5383-5	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 該当する

1. 著者名 Chandra Abel, Sharma Alok, Dehzangi Abdollah, Ranganathan Shoba, Jokhan Anjeela, Chou Kuo-Chen, Tsunoda Tatsuhiko	4. 巻 8
2. 論文標題 PhoglyStruct: Prediction of phosphoglycerylated lysine residues using structural properties of amino acids	5. 発行年 2018年
3. 雑誌名 Scientific Reports	6. 最初と最後の頁 17923
掲載論文のDOI (デジタルオブジェクト識別子) 10.1038/s41598-018-36203-8	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 該当する

1. 著者名 Lysenko Artem, Sharma Alok, Boroevich Keith A, Tsunoda Tatsuhiko	4. 巻 1
2. 論文標題 An integrative machine learning approach for prediction of toxicity-related drug safety	5. 発行年 2018年
3. 雑誌名 Life Science Alliance	6. 最初と最後の頁 e201800098
掲載論文のDOI (デジタルオブジェクト識別子) 10.26508/lisa.201800098	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 該当する

1. 著者名 Yamaguchi-Kabata Yumi, Morihara Takashi, Ohara Tomoyuki, Ninomiya Toshiharu, Takahashi Atsushi, Akatsu Hiroyasu, Hashizume Yoshio, Hayashi Noriyuki, Shigemizu Daichi, Boroevich Keith A., Ikeda Manabu, Kubo Michiaki, Takeda Masatoshi, Tsunoda Tatsuhiko	4. 巻 137
2. 論文標題 Integrated analysis of human genetic association study and mouse transcriptome suggests LBH and SHF genes as novel susceptible genes for amyloid- accumulation in Alzheimer 's disease	5. 発行年 2018年
3. 雑誌名 Human Genetics	6. 最初と最後の頁 521 ~ 533
掲載論文のDOI (デジタルオブジェクト識別子) 10.1007/s00439-018-1906-z	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Mansoor Saqi, Artem Lysenko, Yi-Ke Guo, Tatsuhiko Tsunoda, Charles Auffray	4. 巻 -
2. 論文標題 Navigating the disease landscape: knowledge representations for contextualizing molecular signatures.	5. 発行年 2018年
3. 雑誌名 Briefings In Bioinformatics	6. 最初と最後の頁 -
掲載論文のDOI (デジタルオブジェクト識別子) 10.1093/bib/bby025	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 該当する

1. 著者名 Sugawara Toshitaka, Miya Fuyuki, Ishikawa Toshiaki, Lysenko Artem, Nishino Jo, Kamatani Takashi, Takemoto Akira, Boroevich Keith A., Kakimi Kazuhiro, Kinugasa Yusuke, Tanabe Minoru, Tsunoda Tatsuhiko	4. 巻 25
2. 論文標題 Immune subtypes and neoantigen-related immune evasion in advanced colorectal cancer	5. 発行年 2022年
3. 雑誌名 iScience	6. 最初と最後の頁 103740 ~ 103740
掲載論文のDOI (デジタルオブジェクト識別子) 10.1016/j.isci.2022.103740	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -
1. 著者名 Sharma Alok, Lysenko Artem, Boroevich Keith A, Vans Edwin, Tsunoda Tatsuhiko	4. 巻 22
2. 論文標題 DeepFeature: feature selection in nonimage data using convolutional neural network	5. 発行年 2021年
3. 雑誌名 Briefings in Bioinformatics	6. 最初と最後の頁 -
掲載論文のDOI (デジタルオブジェクト識別子) 10.1093/bib/bbab297	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 該当する
1. 著者名 Kumar Shiu, Sharma Ronesh, Tsunoda Tatsuhiko, Kumarevel Thirumananseri, Sharma Alok	4. 巻 22
2. 論文標題 Forecasting the spread of COVID-19 using LSTM network	5. 発行年 2021年
3. 雑誌名 BMC Bioinformatics	6. 最初と最後の頁 -
掲載論文のDOI (デジタルオブジェクト識別子) 10.1186/s12859-021-04224-2	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 該当する
1. 著者名 TSUNODA Tatsuhiko	4. 巻 60
2. 論文標題 DeepInsight: Methodology to Handle Non-image Data, such as Genomics Data, with Deep Learning	5. 発行年 2020年
3. 雑誌名 Seibutsu Butsuri	6. 最初と最後の頁 149 ~ 152
掲載論文のDOI (デジタルオブジェクト識別子) 10.2142/biophys.60.149	査読の有無 無
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Matsuo Hitoshi、Kamatani Takashi、Hamba Yu、Boroevich Keith A.、Tsunoda Tatsuhiko	4. 巻 23
2. 論文標題 Association between high immune activity and worse prognosis in uveal melanoma and low-grade glioma in TCGA transcriptomic data	5. 発行年 2022年
3. 雑誌名 BMC Genomics	6. 最初と最後の頁 -
掲載論文のDOI (デジタルオブジェクト識別子) 10.1186/s12864-022-08586-6	査読の有無 有
オープンアクセス オープンアクセスとしている(また、その予定である)	国際共著 -

[学会発表] 計32件(うち招待講演 30件/うち国際学会 17件)

1. 発表者名 角田 達彦
2. 発表標題 新たな知によるゲノム医科学
3. 学会等名 東京大学ゲノム医科学研究機構2020年度シンポジウム(招待講演)
4. 発表年 2021年

1. 発表者名 Tatsuhiko Tsunoda
2. 発表標題 Exploring etiologies, sub-classification, and risk prediction of diseases based on big-data analysis of clinical and whole omics data in medicine
3. 学会等名 ERCIM-JST Joint Symposium on Big Data and Artificial Intelligence. 2021(招待講演)(国際学会)
4. 発表年 2021年

1. 発表者名 Tatsuhiko Tsunoda
2. 発表標題 Exploring etiologies, sub-classification, and risk prediction of diseases based on big-data analysis of clinical and whole omics data in medicine
3. 学会等名 CREST International Symposium on Big Data Application 2021(招待講演)(国際学会)
4. 発表年 2021年

1. 発表者名 角田 達彦
2. 発表標題 深層学習が拓く新たなゲノム医科学
3. 学会等名 第16回日本がん分子標的治療学会TRワークショップ(招待講演)
4. 発表年 2021年

1. 発表者名 角田 達彦
2. 発表標題 全ゲノム解析による糖尿病などの疾患の発症や治療・予後予測に向けて
3. 学会等名 AI医療活用研究会(招待講演)
4. 発表年 2019年

1. 発表者名 Tatsuhiko Tsunoda
2. 発表標題 Data-driven Medical Sciences with Omic Analysis
3. 学会等名 Moonshot International Symposium(招待講演)(国際学会)
4. 発表年 2019年

1. 発表者名 Tatsuhiko Tsunoda
2. 発表標題 Personalized Cancer Medicine with Heterogeneity and Immunological Analysis
3. 学会等名 CREST International Symposium on Big Data Application(招待講演)(国際学会)
4. 発表年 2020年

1. 発表者名 Tatsuhiko Tsunoda
2. 発表標題 Omic Big Data Analysis Drives Precision Medicine.
3. 学会等名 The 8th Annual Translational Bioinformatics Conference/2018 Annual Conference of Korean Society for Bioinformatics (招待講演) (国際学会)
4. 発表年 2018年

1. 発表者名 Tatsuhiko Tsunoda
2. 発表標題 Omic analysis with AI drives precision medicine.
3. 学会等名 The 5th RIKEN-KI/SciLifeLab Joint Symposium (招待講演) (国際学会)
4. 発表年 2018年

1. 発表者名 Tatsuhiko Tsunoda
2. 発表標題 Trans-omic and Trans-species Analysis Drives Precision Medicine.
3. 学会等名 The 41st Annual Meeting of the Molecular Biology Society of Japan (招待講演) (国際学会)
4. 発表年 2018年

1. 発表者名 Tatsuhiko Tsunoda
2. 発表標題 Medical Big Data Analysis for Precision Medicine.
3. 学会等名 CREST Joint International Symposium for Big Data (招待講演) (国際学会)
4. 発表年 2019年

1. 発表者名 Tatsuhiko Tsunoda
2. 発表標題 Cancer heterogeneity and immunology for precision medicine.
3. 学会等名 CREST International Symposium on Big Data Application (招待講演) (国際学会)
4. 発表年 2019年

1. 発表者名 Tatsuhiko Tsunoda
2. 発表標題 Trans-omic Analysis Strategy for Precision Medicine.
3. 学会等名 The 2nd International Symposium for Trans-Omics (招待講演) (国際学会)
4. 発表年 2018年

1. 発表者名 Tatsuhiko Tsunoda
2. 発表標題 Public Big Data Accelerate Medical Science Research.
3. 学会等名 Symposium "Public database and application for medicine in Japan and Asia", The 63rd Annual Meeting of the Japan Society of Human Genetics (招待講演) (国際学会)
4. 発表年 2018年

1. 発表者名 Tatsuhiko Tsunoda
2. 発表標題 Exploring etiologies, sub-classification, and risk prediction of diseases based on big-data analysis of clinical and whole omics data in medicine.
3. 学会等名 CREST International Symposium on Big Data Application (招待講演) (国際学会)
4. 発表年 2018年

1. 発表者名 角田 達彦
2. 発表標題 ゲノム医科学の進展とプレジジョン医療の推進.
3. 学会等名 東北メディカルメガバンク第16回ゲノム・オミックス連携推進セミナー（招待講演）
4. 発表年 2018年

1. 発表者名 角田 達彦
2. 発表標題 臨床研究におけるビッグデータの活用.
3. 学会等名 臨床研究統計コース, 統計数理研究所（招待講演）
4. 発表年 2018年

1. 発表者名 Tatsuhiko Tsunoda
2. 発表標題 Exploring etiologies, sub-classification, and risk prediction of diseases based on big-data analysis of clinical and whole omics data in medicine
3. 学会等名 CREST Big Data Fields Joint Meeting（招待講演）（国際学会）
4. 発表年 2017年

1. 発表者名 Tatsuhiko Tsunoda
2. 発表標題 Trans-omic analysis drives precision medicine
3. 学会等名 The 1st International Symposium for Trans-Omics（招待講演）（国際学会）
4. 発表年 2017年

1. 発表者名 Tatsuhiko Tsunoda
2. 発表標題 Multi-omic analysis for precision cancer medicine
3. 学会等名 DNA sequencing technologies and their application in practice WS (招待講演) (国際学会)
4. 発表年 2017年

1. 発表者名 Tatsuhiko Tsunoda
2. 発表標題 Omic analysis drives precision medicine
3. 学会等名 International Conference for Precision Cancer Medicine (招待講演) (国際学会)
4. 発表年 2017年

1. 発表者名 Tatsuhiko Tsunoda
2. 発表標題 Multiomics and clinical analysis of cancer
3. 学会等名 CREST International Symposium on Big Data Application (招待講演) (国際学会)
4. 発表年 2017年

1. 発表者名 角田 達彦
2. 発表標題 医学・医療ビッグデータによるマルチオミックス解析
3. 学会等名 東京医科歯科大学 早稲田大学連携ワークショップ (招待講演)
4. 発表年 2017年

1. 発表者名 角田 達彦
2. 発表標題 オミックス解析によるゲノム医療の新たな展開
3. 学会等名 第 8 回生物統計ネットワークシンポジウム (招待講演)
4. 発表年 2017年

1. 発表者名 角田 達彦
2. 発表標題 ゲノム医科学を推進するデータ駆動型研究
3. 学会等名 AMED「脳とこころ」疾患メカプロジェクト進捗報告会 (招待講演)
4. 発表年 2021年

1. 発表者名 角田 達彦
2. 発表標題 ゲノム医科学を推進する深層学習と数理科学
3. 学会等名 第25回日本がん分子標的治療学会学術集会・シンポジウム2「AI」(招待講演)
4. 発表年 2021年

1. 発表者名 鎌谷 高志
2. 発表標題 DNA, RNA解析を用いたEGFR-TKIの効果予測モデルの構築
3. 学会等名 第60回日本呼吸器学会学術講演会
4. 発表年 2020年

1. 発表者名 鎌谷 高志
2. 発表標題 粘液型脂肪肉腫におけるがん微小環境の免疫学的解析
3. 学会等名 第79回日本癌学会学術総会
4. 発表年 2020年

1. 発表者名 鎌谷 高志
2. 発表標題 生物情報科学の視点から行うがん研究 ~DNA, RNA, 病理画像を用いた免疫環境解析や多領域解析を中心に~
3. 学会等名 第19回関東骨軟部腫瘍の基礎を語る会（招待講演）
4. 発表年 2021年

1. 発表者名 鎌谷 高志
2. 発表標題 生命情報科学の観点から、癌を研究する ~ multi-omics, deep learningを用いて ~
3. 学会等名 第293回がん研有明病院学術研究講演会（招待講演）
4. 発表年 2019年

1. 発表者名 鎌谷 高志
2. 発表標題 医学、医療におけるAIのインパクト
3. 学会等名 第3回 統計・機械学習若手シンポジウム（招待講演）
4. 発表年 2018年

1. 発表者名 鎌谷 高志
2. 発表標題 基礎と臨床に活かすAIの使い方
3. 学会等名 第1回 Meeting of Allergology Research in Tokyo (招待講演)
4. 発表年 2018年

〔図書〕 計3件

1. 著者名 Tsunoda T, Tanaka T, Nakamura Y	4. 発行年 2019年
2. 出版社 Springer Singapore	5. 総ページ数 218
3. 書名 Genome-Wide Association Studies	

1. 著者名 Lopez Y, Kamola PJ, Sharma R, Shigemizu D, Tsunoda T, Sharma A	4. 発行年 2018年
2. 出版社 Elsevier	5. 総ページ数 22
3. 書名 Encyclopedia of Bioinformatics and Computational Biology (Computational Pipelines and Workflows in Bioinformatics)	

1. 著者名 柚木克之、角田達彦、黒田真也 (岡田随象 編)	4. 発行年 2018年
2. 出版社 株式会社 メディカル ドゥ	5. 総ページ数 10
3. 書名 遺伝統計学と疾患ゲノムデータ解析 (2章-10 「GWASをトランスオミクスで読み解く」)	

〔産業財産権〕

〔その他〕

代謝統合オミクス
<http://transomics.umin.jp>
 東京大学 角田研究室
<http://mesm.bs.s.u-tokyo.ac.jp>

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究分担者	重水 大智 (Shigemizu Daichi) (70617464)	国立研究開発法人国立長寿医療研究センター・メディカルゲノムセンター・副部長 (83903)	

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究協力者	宮 冬樹 (Miya Fuyuki)		

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関			
フィジー	University of the South Pacific			
オーストラリア	Griffith University			
英国	Imperial College London			
フランス	EISBM			
米国	Morgan State University			