

2014年度採択研究代表者

角田 達彦

東京医科歯科大学難治疾患研究所
教授

医学・医療における臨床・全ゲノム・オミックスのビッグデータの解析に基づく
疾患の原因探索・亜病態分類とリスク予測

§ 1. 研究成果の概要

次世代シーケンサーデータから中間サイズの断片の挿入や欠失を予測する新しい方法を開発し、論文出版し、フリーソフトとして公開した (*Scientific Reports* 2018)。またヘテロなオミックスデータを統合しネットワーク解析して分子機構を生物学的に理解と探索するための知識表現として、特に3つの主要なアプローチ、すなわち経路マッピングアプローチ、分子ネットワーク中心アプローチ、および生物学的記述を知識グラフとして表現する方法論について論じた (*Briefings in Bioinformatics* 2018)。応用とし、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)。続き、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)。さらに、独自開発の多次元データ汎用解析ツールのセミパラメトリック階層混合モデル (SP-HMM) の適用研究を推進し、疾患間での多遺伝的構造の定量的評価 (*Front Genet.* 2018)、関連遺伝子多型の発見における必要症例数予測 (*Front Genet.* 2018)、脳卒中や乳がんの進行リスクに関連する一塩基多型の同定 (*Eur J Hum Genet.* 2019) 等の検討を行った。併せ、薬剤効果予測因子の検証に向けて効率的な臨床試験のデザインの検討と、多様な治療効果プロファイルを想定した試験デザインの評価基準を開発した。また臨床現場での治療の礎として、FFPE 臨床シーケンス用の高精度な SNV/indel、fusion、CNA 検出アルゴリズムを開発した (*Genome Medicine* 2018)。国立がん研究センターのがんゲノム医療臨床研究 TOPGEAR と MASTER KEY プロジェクトにて実際に使用、企業技術移転、がんゲノム中核一連携病院体制の先進医療 B の遺伝子検査「個別化医療に向けたマルチプレックス遺伝子パネル検査研究」にて、実際に使用された。また日本初のがんゲノム医療機器として厚生労働省所管 PMDA に公的に承認された。

実データによる疾患ゲノムリスク因子同定では、大規模な国際協力によって様々な集団の喘息のゲノムワイド関連解析の統合メタ解析を行い、9つの新しい喘息関連遺伝子を発見した。そしてeQTLというゲノム-発現相関分子のトランスオミクスを駆使し、それらが免疫学的メカニズムの調節に密接に関わることを解明した (*Nature Genetics* 2018)。また、アルツハイマー患者の脳にA β の蓄積を引き起こす新規の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)。さらに、認知症にはアルツハイマー病とともに、血管性認知症やレビー小体型認知症があるが、世界で初めて認知症前向きコホート1,601人の血清miRNA発現を用いて、それらを区別するバイオマーカーを発見し、リスク予測モデルを構築した (*Communications Biology* 2019)。関節リウマチでは、患者と健常者から得られた末梢血単核細胞からmiRNAを抽出して次世代シーケンサーによりmiRNAの発現量を網羅的に検討し、両者の結果を統合することで、関節リウマチの発症に関わる複数のmiRNAの同定に成功した (*Nucleic Acids Res.* 2018)。また表現型としての関節破壊度に着目して、日本人で最も頻度の高い感受性アレルであるHLA-DRB1*04:05が疾患活動性とは独立して関節破壊と関連することも明らかにした (*Ann Rheum Dis.* 2019)。そして現在、臨床研究と密に連携し、がんの化学療法や免疫療法などの治療法と奏効、予後に関わる因子の実データ解析と予測モデルの構築を行なっている。

【代表的な原著論文】

1. Lysenko A, Sharma A, Boroevich KA, Tsunoda T. An integrative machine learning approach for prediction of toxicity-related drug safety. *Life Science Alliance*, **1**, e201800098 (2018).
2. Yamaguchi-Kabata Y, Morihara T, Ohara T, Ninomiya T, Takahashi A, Akatsu H, Hashizume Y, Hayashi N, Shigemizu D, Boroevich KA, Ikeda M, Kubo M, Takeda M, Tsunoda T. Integrated analysis of human genetic association study and mouse transcriptome suggests LBH and SHF genes as novel susceptible genes for amyloid- β accumulation in Alzheimer's disease. *Human Genetics*, **137**, 521-533 (2018).
3. Shigemizu D, Akiyama S, Asanomi Y, Boroevich KA, Sharma A, Tsunoda T, Matsukuma K, Ichikawa M, Sudo H, Takizawa S, Sakurai T, Ozaki K, Ochiya T, Niida S. Risk prediction models for dementia constructed by supervised principal component analysis using miRNA expression data. *Communications Biology*, **2**, 77 (2019).

§ 2. 研究実施体制

(1)「角田」グループ1

- ① 研究代表者:角田 達彦 (東京医科歯科大学難治疾患研究所 教授)
- ② 研究項目
 - 臨床・ゲノム・オミックス情報を用いたビッグデータ解析基盤の開発とバンク・多層オミックス・コホートサンプルの解析

(2)「角田」グループ2

- ① 研究代表者:角田 達彦 (理化学研究所生命医科学研究センター チームリーダー)
- ② 研究項目
 - 臨床・ゲノム・オミックス情報を用いたビッグデータ解析基盤の開発とバンク・多層オミックス・コホートサンプルの解析

(3)「野間」グループ

- ① 主たる共同研究者:野間 久史 (統計数理研究所 准教授)
- ② 研究項目
 - ゲノムビッグデータの統計解析における疾患関連遺伝子探索のための方法論・ストラテジーの開発と評価
 - 稀な遺伝的変異の疾患発症リスクの評価
 - 個人ゲノムの秘匿性確保

(4)「松井」グループ

- ① 主たる共同研究者: 松井 茂之 (名古屋大学医学系研究科 教授)
- ② 研究項目
 - オミックスデータを用いた関連遺伝子の探索と予測解析
 - 階層混合モデルと経験ベイズ解析の開発と各種疾患への適用
 - 疾患の遺伝的異質性の解析

(5)「猪狩」グループ

- ① 主たる共同研究者:猪狩 勝則 (東京女子医科大学膠原病リウマチ痛風センター 准教授)
- ② 研究項目
 - リウマチ患者の大規模前向き観察研究 IORRA コホートの調査項目とゲノムデータの整備、提供、解析と結果の医学的考察

(6)「三木」グループ

- ① 主たる共同研究者:三木 大樹 (広島大学大学院医歯薬保健学研究科(医) 講師)
- ② 研究項目

- 肝疾患臨床データベースの統一化と診療情報収集登録作業の効率化推進

(7)「加藤」グループ

- ① 主たる共同研究者:加藤 護(国立がん研究センター研究所 部門長)
- ② 研究項目
 - がんオミックスデータを活用したリスク予測マーカーの探索と臨床応用の可能性