

研究報告書

「エコファーマによる高速かつ省エネ創薬を実現する情報技術の構築」

研究タイプ: 通常型

研究期間: 2015年10月 ~ 2019年3月

研究者: 山西 芳裕

1. 研究のねらい

近年、新薬の開発が低迷している。新薬を一個開発するのに数千億円の研究開発費と10年以上の歳月を要すると云われており、新薬開発の成功率は非常に低い。有効な治療薬が無い疾患に今苦しんでいる患者に対して、迅速な救済措置が必要である。一方で、我が国は少子高齢化の到来による医療費の増大にも直面している。つまり、疾患治療のための必要な新薬は開発しつつ、一人当たりの医療費を抑制すべきという難しい社会的課題に迫られている。そのためには、我が国の創薬体制を根本的に改革し、新しい21世紀型医療へのパラダイムシフトが不可欠である。

本研究では、新薬候補を人類の資産である既承認薬から見出す創薬戦略「エコファーマ」(文脈によりドラッグリポジショニングやスマート創薬などとも呼ばれる)を提唱し、それを促進する情報基盤技術を構築する。既承認薬はヒトでの安全性や体内動態、製造方法が確認されているため、エコファーマによる創薬は失敗するリスクが低いという特長がある。行程を大幅にスキップできるため開発期間やコストを劇的に削減でき、高速かつ省エネ創薬を実現できる。例えば、世界的に大流行して多くの死者を出したエボラ出血熱の治療薬として、インフルエンザ治療薬のアビガンの有効性が報告されたことは記憶に新しい。

本研究期間内に、近年の生命医科学で大量に生み出されてきた既承認薬、遺伝子、タンパク質、化合物、疾患に関する医薬ビッグデータを融合解析し、様々な疾患に対して効能を持つ既承認薬の候補を予測する機械学習アルゴリズムを開発する。これによって、疾患に苦しむ患者に、必要な薬を、早く、安く、届ける創薬サービスを大幅に効率化する。更に、既承認薬の薬理作用を模倣するような天然化合物を、日常生活において身近な漢方薬や食品の中から計算機的に探索し、それらの科学的根拠を示すことによって予防医療への道を切り開く。医療に伴う患者の経済的・精神的・肉体的負担を減らし、誰もが健康長寿を全うできる社会の実現を目指す。患者に優しく(安全)、エコノミカル(安く効率的)で、エコロジカル(省エネ)なエコファーマ創薬サービスと、病気になる前に予防する健康維持型医療サービスを促進することによって、資源の浪費を回避し、人件費の削減、新薬開発期間の短縮などの効果により国家の医療費削減にも貢献することが狙いである。

2. 研究成果

(1) 概要

本研究では、医薬ビッグデータ解析を通じて創薬の高速化・効率化や医療費の抑制という社会的課題を解決するために「エコファーマによる高速かつ省エネ創薬を実現する情報技術の構築」を提案した。新薬候補を人類の資産である既承認薬から見出す創薬戦略「エコファーマ」を促進するための情報基盤技術を構築した。具体的には、医薬ビッグデータ解析から既

承認薬の標的タンパク質を推定し、疾患に効能のある既承認薬の予測を行う機械学習手法を開発した。

既承認薬への応用では、薬が標的タンパク質を制御して疾患への効能を発揮する点に注目し、薬の効能を、薬、タンパク質、疾患の3者間の関連ネットワークとして情報科学的に捉えて、薬、タンパク質、疾患の三種類のノードから構成される三部グラフのエッジの有無を予測する問題として定式化した。国際疾病分類で定義された疾患に対して、日本や欧米の全ての既承認薬の中から治療薬候補を大規模予測し、その予測結果の妥当性を実験で検証した。

更に、本研究では、標的タンパク質の制御だけでなく、パスウェイの制御による創薬(パスウェイ創薬)という新しい創薬コンセプトを提唱した。一つのタンパク質を制御する薬を探索する従来のアプローチではなく、タンパク質相互作用ネットワークの機能モジュールであるパスウェイを制御する薬を探索するという新たなアプローチを提案した。大規模な薬物応答遺伝子発現データと生体分子相互作用ネットワークの融合解析により、薬物に応答して制御されるパスウェイを統計的に推定した。がんへの応用では、抗がん剤の多くに見られる特徴的なパスウェイを見出し、薬物の潜在的な抗がん作用を予測する統計解析手法を開発した。

更に、日常生活において身近な漢方薬、生薬、食品から健康効果を見出し、それらの科学的根拠を示す情報技術を開発した。漢方薬への応用では、漢方薬、生薬、成分化合物、標的タンパク質の階層関係を整備し、漢方薬の成分化合物が標的とするタンパク質群のパスウェイ解析により、漢方薬のメカニズムを推測した。標的タンパク質や標的パスウェイの情報を基に、漢方薬の新しい適応可能疾患を予測し、解析結果を視覚化した KampoDB というデータベースを構築して、一般公開した。

これらの研究成果は、欧文学術雑誌で7本の論文を出版し、国内学会で 49 件、国際学会で8件の口頭発表を行った。

(2) 詳細

さきがけの研究提案書の「4.研究内容」で記述していた研究項目に対応する形で、研究成果の達成状況を以下に記す。

研究テーマ A1. 生命医薬関連ビッグデータの収集

既承認薬、遺伝子、タンパク質、疾患に関する様々なデータは、インターネットや文献、データベースから適時収集して、機械学習で情報解析可能な形に整備した。【達成済】

研究テーマ A2. 既承認薬の標的タンパク質を予測する手法の実装

薬の情報(化学部分構造、副作用報告、薬物応答遺伝子発現変化など)とタンパク質の情報(アミノ酸配列、機能モチーフなど)の連結やテンソル積によって、全ての薬・タンパク質ペアのフィンガープリント(超高次元のバイナリ特徴ベクトル)を構築した。高次元データに有効なスパース分類器(L1 正則化サポートベクターマシンやロジスティック回帰など)を拡張し、薬がそのタンパク質を標的とするかを予測する手法を開発することができた。【達成済】

実際に、ベンチマークデータの約 537 万個の薬物-タンパク質ペア(84,195,800 次元のフィンガープリント)の相互作用予測を行うため、簡潔データ構造 LOUDS を用いてフィンガープリントをコンパクトに表現し、分類モデルを学習する方法を提案した。その成果として、メモリ

を 57GB から 16GB に圧縮した状態でモデル学習が可能となり、高い予測精度を達成した (AUC: 0.98, AUPR: 0.63) (Tabei et al, BMC Systems Biology, in press)。【達成済】

研究テーマ A3. 疾患の治療標的タンパク質を予測する手法の実装

病態に特徴的な分子的特徴は疾患間で共通する場合もある。ここでは、約 80 種類の疾患 (副腎白質ジストロフィー、白血病、アルツハイマー病、喘息、アトピー性皮膚炎、乳がん、嚢胞性線維症、デング熱、炎症性腸疾患など) に対して、患者の遺伝子発現プロファイルから疾患特異的な遺伝子発現プロファイルを構築し、疾患間の相関解析を行なった。マルチタスク学習の枠組みで、疾患の類似性を用いて治療標的タンパク質の共通性を予測する新しい機械学習法を開発した。現在論文準備中である。【論文執筆中】

研究テーマ A4. 疾患に効能のある既承認薬を予測する手法の実装

薬・タンパク質関連ペアと、疾患・タンパク質関連ペアを比較し、疾患の候補薬を予測する手順を確立した。標的タンパク質が阻害されるか活性化されるか区別して実行することで予測精度が改善することを確認し、論文で発表した (Scientific Reports, 8:156, 2018)。【達成済】

更に、標的タンパク質の制御だけでなく、パスウェイの制御による創薬 (パスウェイ創薬) という世界初のコンセプトを提唱した。従来の創薬では、1つのタンパク質を制御する薬を探索するアプローチが一般的だが、タンパク質がなす相互作用の影響を考慮できないという問題があった。そこで、タンパク質相互作用ネットワークの機能モジュールであるパスウェイを制御する薬を探索するという新たなアプローチを提案した。まず、大規模な薬物応答遺伝子発現データと生体分子相互作用ネットワークの融合解析により、抗がん剤の多くに見られる特徴的なパスウェイ (細胞周期パスウェイを抑制し、p53シグナリングパスウェイ、細胞死パスウェイなど) を見出した。次に、薬物の潜在的な抗がん作用を自動的に予測するための統計解析手法を開発した。この研究成果は創薬科学のトップジャーナルで論文を発表することができた (Journal of Medicinal Chemistry, 61, 9583–9595, 2018)。【達成済】

研究テーマ B1. あらゆる疾患に対して治療薬候補を大規模予測

本研究で開発した手法を、日本や欧米で承認されている全ての既承認薬 (約 8000 個) に適用し、様々な疾患 (約 1500 種類) に対する潜在的な効能を大規模予測した結果、約 900 種類の疾患に対して予測結果を得ることができた。【達成済】

研究テーマ B2. 臨床データの収集

米国 FDA (食品医薬局) の市販後調査報告システムで、患者に対する薬理効果や副作用が記録されている報告書をテキストマイニングすることによって、既承認薬の副作用情報を収集した。これを機械学習アルゴリズムの入力の一部として用いることによって、既承認薬の標的タンパク質予測に対する有用性を示し、提案するアルゴリズムとともにその結果を論文として発表することができた (Tabei et al, BMC Systems Biology, in press)。【達成済】

研究テーマ B3. 予測結果の妥当性を検証

本研究の提案手法で予測された疾患・候補薬ペアに関して、近年の臨床報告書や最新の臨床系学術論文で、スコアの高い予測結果の妥当性を検証した。候補薬によって症状の改善が報告されているかなどの考察を行い、予測された多くの薬物・疾患ペアに対する妥当性を検証することができた。【達成済】

更に、本研究で開発したパスウェイ創薬手法の予測結果を実験的に検証した。特に、抗がん作用が新たに予測された薬物について、細胞死誘導能などの観点から実験的に検証した。例えば、統合失調症治療薬であるペンフルリドールなどいくつかの既存薬が、期待する抗がん作用をもつことが観察され、情報科学的な予測の妥当性を示すことができた (Journal of Medicinal Chemistry, 61, 9583–9595, 2018)。【達成済】

研究テーマ C1.天然化合物データの収集

疾患の治療効果や予防効果を持つ天然化合物(漢方薬、生薬、食品などの成分化合物)を探索するため、天然化合物データは、KNApSAcK データベースなどの化合物の情報を収集して解析することができた。漢方薬や生薬のデータは、富山大学和漢医薬学総合研究所で長年蓄積されているデータを共同研究の形で入手して、漢方薬、生薬、成分化合物、標的タンパク質の階層関係を整備し、情報解析することができた。【達成済】

研究テーマ C2.疾患予防効果を持つ天然化合物の大規模予測

漢方薬の各成分化合物が相互作用すると推定される全てのタンパク質をパスウェイ (KEGG Pathway など) にマッピングし、漢方薬が影響を及ぼすパスウェイを推測した。また、数百万個の化合物・タンパク質間相互作用情報を用いた教師付き学習と化学構造類似性検索を組み合わせることで、漢方薬の成分化合物の標的タンパク質候補を探索した。標的タンパク質の情報を基に、漢方薬の新しい適応可能疾患を予測し、これらの解析結果を視覚化するシステムを構築した。例えば、腹痛の漢方薬「大建中湯」の抗炎症作用や炎症関連大腸発癌に対する効能と作用機序を予測した。予測システムの第一弾を論文発表することができた (Sawada et al, Scientific Reports, 8:11216, 2018)。【達成済】

研究テーマ C3.解析結果のデータベース化

西洋医薬品に対する解析結果は、標的タンパク質などの科学的根拠を付加して系統的に整備し、データベース化した。知財の点で問題がある結果(公開すると特許が取れなくなるもの)を除いてデータベース化したものを、論文発表する予定である。天然化合物を成分とする漢方薬や生薬の解析結果は、KampoDB (<http://wakanmoview.inm.u-toyama.ac.jp/kampo/>) というデータベースを構築して、誰でも利用できるように一般公開し、予測システムとともに論文発表することができた (Sawada et al, Scientific Reports, 8:11216, 2018)。【達成済】

3. 今後の展開

今後は、解析対象とするデータの種別を拡大し、多様な疾患のヒトのマルチオミクスデータや分子ネットワークデータや、既承認薬以外の化合物データ(開発中止化合物、合成化合物、生理活性化合物など)を扱えるような手法の開発に、研究を展開していく予定である。

さがけ研究のアドオン研究として、内閣府 PRISM プロジェクト創薬分野「新薬創出を加速する人工知能の開発」に採択された。内閣府 PRISM プロジェクトでは、肺がんおよび特発性肺線維症を対象疾患とし、医薬基盤・健康・栄養研究所、理化学研究所、国立がん研究センター、産業技術総合研究所など多くの研究機関が連携して行う。その中で、研究代表者(山西)チームの研究課題名は「創薬標的分子の確からしさを検証するツール物質の探索」であり、特

発性肺線維症および肺がんの創薬標的候補の「確からしさ」の検証実験に利用可能なツール化合物を探索していく予定である。

このアドオン研究は、疾患の新しい創薬ターゲットの発見に貢献するものであり、疾患の分子機序の理解や適応可能疾患の探索空間の拡大に繋がる。更に、解析対象を既承認薬だけでなく任意の化合物(開発中止化合物、合成化合物、生理活性化合物など)にも広げるので、化合物の探索空間の拡大や疾患類似性の考慮による予測精度の向上が期待できる。本アドオン研究は、これまでの申請者のさきがけ研究を大きく発展させ、創薬現場への社会実装に繋がるものである。

4. 自己評価

研究目的の達成状況:

さきがけの研究提案書の「4.研究内容」で記述していた研究項目は、予定していた内容は全て達成することができた。しかしながら、さきがけの研究期間内に論文出版までできなかった項目もあるので、現在できるだけ早く出版できるように論文執筆中である。

研究の進め方(研究実施体制及び研究費執行状況):

研究実施体制は、研究補助者を数人雇用してデータ収集・整備を進めながら、本研究を遂行した。初年度や大学異動時期など、研究補助者の公募などで雇用が途切れた時期もあるが、概ね同じような研究実施体制を保つことができた。研究費は主に、医薬関連ビッグデータの収集と整備、開発手法の計算リソースの構築や確保に、概ね予定通り執行できた。

研究成果の科学技術及び社会・経済への波及効果(今後の見込みを含む):

さきがけの研究成果の情報技術を社会・経済に波及させるため、現在、多くの日本企業(製薬会社、化粧品会社、食品会社など)と共同研究を遂行している。さきがけで開発した情報技術は、実際に企業内の開発現場において医薬品開発や健康食品開発に利用されているので、将来的な製品化の可能性は非常に高く、日本経済への貢献が期待できる。また、医療現場に波及させるため、現在、多くの大学病院の研究者と共同研究を遂行している。予測された薬の人に対する効果を確認する臨床研究を進めて、最近増えつつある高額な薬(抗体医薬品など)の効果を、低価格かつ低副作用の既存薬で代用できることを示すことができれば、国家の医療費削減にも貢献することが期待できる。

その他領域独自の評価項目に基づいて、研究者自身の評価:

さきがけの研究成果を、欧文学術雑誌で7本の論文を出版し、国内学会で49件、国際学会で8件の口頭発表を行い、学術研究の成果を出すことができた。特に創薬科学のトップジャーナルである Journal of Medicinal Chemistry で論文を出版できたことは大きな成果であり、実験系・臨床系の創薬系研究者のコミュニティでも注目を浴びた。本さきがけ領域で目標とする社会実装は、多くの企業との共同研究(創薬現場での開発手法の利用)や大学病院との共同研究(医療現場での開発手法の利用)によって、達成できた。

今後の見込み:

今後は、医薬品開発工程の研究を進めて、最終的な医薬品の製品化や、保険適用のための治験まで進めることを目標としたい。

5. 主な研究成果リスト

(1) 論文(原著論文)発表

- | |
|--|
| 1. Iwata, M., Hirose, L., Kohara, H., Liao, J., Sawada, R., Akiyoshi, S., Tani, K., and <u>Yamanishi, Y.</u> , "Pathway-based drug repositioning for cancers: computational prediction and experimental validation", <i>Journal of Medicinal Chemistry</i> , 61, 9583–9595, 2018
doi:10.1021/acs.jmedchem.8b01044 (Oct. 2018) |
| 2. Sawada, R., Iwata, M., Umezaki, M., Usui, Y., Kobayashi, T., Kubono, T., Hayashi, S., Kadowaki, M., and <u>Yamanishi, Y.</u> , "KampoDB, database of predicted targets and functional annotations of natural medicines", <i>Scientific Reports</i> , 8:11216, 2018,
doi:10.1038/s41598-018-29516-1 (Jul. 2018) |
| 3. Sawada, R., Iwata, M., Tabei, Y., Yamato, H., and <u>Yamanishi, Y.</u> , "Predicting inhibitory and activatory drug targets by chemically and genetically perturbed transcriptome signatures", <i>Scientific Reports</i> , 8:156, 2018 doi:10.1038/s41598-017-18315-9 (Jan. 2018) |

(2) 特許出願

研究期間累積件数:0件

(3) その他の成果(主要な学会発表、受賞、著作物、プレスリリース等)

主要な学会発表:

1. Yamanishi, Y., "Data-driven drug discovery and repositioning by machine learning methods", ACS Skolnik Symposium, "De novo design – Automating drug discovery" session, Boston, USA, Aug.21–Aug.21, 2018. 【招待講演】(Aug. 2018)
2. 山西芳裕, "AI創薬:薬効や副作用を予測するデータ駆動型アプローチ", 構造活性フォーラム 2018「創薬におけるビッグデータの活用とAI戦略」(京都), 6/15, 2018.【招待講演】(2018年6月)

受賞:

1. 情報計算化学生物学会 2017年大会・ポスター賞(2017年10月)
2. 第6回生命医薬情報学連合大会・最優秀口頭発表賞(2017年9月)

著作物(本など):

1. Yamanishi, Y., "Linear and Kernel Model Construction Methods for Predicting Drug-Target Interactions in a Chemogenomic Framework", *Computational Chemogenomics, Methods in Molecular Biology Series* (Brown, J.B., eds.), Springer, 355–368, 2018. doi:10.1007/978-1-4939-8639-2_12 (Oct. 2018)
2. 山西芳裕, "ファーマコインフォマティクス", *ゲノム創薬科学*(田沼靖一編集), 159–182, 裳華房, 2017. (2017年10月)

プレスリリース:

1. 「医薬ビッグデータから抗がん作用薬を自動的に予測する情報技術を開発」
九州工業大学 (2018年11月)

以上