

**厚生労働科学研究費補助金**  
**(政策科学総合研究事業 (臨床研究等ICT基盤構築・人工知能実装研究事業))**  
**総括研究報告書**

課題名 : 新薬創出を加速する症例データベースの構築・拡充/創薬ターゲット推定アルゴリズムの開発  
研究代表者名 : 夏目やよい

国立研究開発法人医薬基盤・健康・栄養研究所 AI健康・医薬研究センター  
バイオインフォマティクスプロジェクト プロジェクトリーダー

**研究要旨**

本事業では、「創薬ターゲットの枯渇問題」を克服すべく、動物からではなくヒトの情報から創薬ターゲット分子を探索する AI の開発実装を目的とする。

令和3年度は、肺がん手術検体及びバイオプシー検体のオミックス解析を行うとともに、神奈川県立循環器呼吸器疾患センターで IPF を含む間質性肺炎患者の臨床情報（オミックスデータ及びそれと紐づけられた診療情報）を引き続き収集した。これまでに本事業で開発された患者層別化 AI を用い、大阪大学で収集された臨床情報の解析を行うことで IPF の特徴に該当する診療情報の項目に紐づけられる血中タンパク質をデータ駆動的に複数個見出したが、これらのタンパク質の上流因子を阻害することにより EMT（上皮間葉転換）が抑制されることを確認した。さらに、本事業の成果であるデータベースや AI を広く共有するための基盤であるオープンプラットフォームに新たに 9 つの AI ガジェットと大阪大学コホートデータ（匿名加工済み診療情報、プロテオームデータ）を搭載した。

**A. 研究目的**

医薬品開発において、近年国内外を問わず創薬ターゲットの枯渇が問題となっている。現在残されているのは高難易度の創薬ターゲットのみであるがために、新薬の研究開発には多額の費用が必要となっており、これが高薬価、ひいては医療費の高騰の要因となっている。

更に、臨床試験段階で期待していた薬効が得られず開発が中断する例が増えていることも問題点として挙げられる。特に医薬品開発の 70~80%が Phase2 で中止となっており、この約 60%が、薬効が得られなかったことが原因との報告がある。つまり、「動物では効くが、ヒトでは効かなかった」という事案が多発している。これは現在の創薬研究開発スキームの限界であると考えられる。

このような現状を打開する解決策として、人工知能 (AI; Artificial intelligence) が注目されている。AI のパフォーマンスと可能性に創薬・医療・ヘルスケア分野が大きな期待を寄せており、今後国際競争が激化することが必至である。

これらの現状を背景に、本事業では、「創薬ターゲットの枯渇問題」を克服すべく、動物からではなくヒトの情報から創薬ターゲット分子を探索する AI の開発実装を目的とする。つまり、臨床情報 (=電子カルテを始めとする診療情報+オミックスデータ) を収集・利用して創薬ターゲットを探索する AI 手法の開発をおこなう。本事業では、対象疾患として難病指定の IPF (特発性肺線維症) を含む間質性肺炎及び部位別がん死亡者数 1 位である肺がんを選択し、これらの臨床情報収集とそれを支援する基盤構築、異種かつ大量のデータを統合して創薬ターゲット候補となる生体分子群を自動的に抽出する AI 手法の開発を行う。また、本事業で作成される IPF/肺がんの疾患統合データベース、機能分子を特定するための AI 及び知識ベース等を多くの研究者等に利用してもらうための環境 (オープンプラットフォーム) の構築を目指す。

**B. 研究方法**

1. 臨床情報の収集と機械可読表現の開発 : 肺がん手術検体及びバイオプシー検体のオミックス解析を行うとともに、大阪大学医学部呼吸器・免疫内科バイオバンク、大阪大学医学部附属病院医療情報部大阪大学病院バイオバンク及び神奈川県立循環器呼吸器疾患センターで IPF を含む間質性肺炎患者の臨床情報を収集した。

**肺がん手術検体・バイオプシー検体のマルチオミックスデータ収集**

国立がん研究センターにおいて収集された肺がん手術検体及びバイオプシー検体 (検体採取・処理・保存は国立がん研究センターにおけるプロトコルに則って行われた) を用いて、下記のおミックスデータを収集した。オミックスデータ取得はタカラバイオ株式会社への外部委託により実施した。

- ・ 全ゲノム解析
- ・ 全エクソーム解析
- ・ DNA メチル化解析
- ・ ヒストン修飾 (ChIP-seq) 解析
- ・ トランスクリプトーム (RNA-seq)

#### 大阪大学コホートにおける臨床情報収集

令和元年度までに大阪大学医学部呼吸器・免疫内科のバイオバンクより、同意取得済みの IPF を含む間質性肺炎患者及び器質的な呼吸器疾患を有しないと診断された方々の血清と、血清にひもつく診療情報（診察記録、CT 画像及びその読影所見、血液検査値、患者基本情報及び初診時間診票）を大阪大学病院医療情報部より収集している。当該年度は、COVID-19 の影響により入手が遅れていた呼吸機能検査値及びフローボリュームカーブと MDD 結果を同様に匿名化 ID が付与された形で入手した。

#### 神奈川県立循環器呼吸器病センターコホートにおける臨床情報収集

神奈川県立循環器呼吸器疾患センターにおいても、下記の臨床情報（診療情報＋オミックスデータ又は肺組織や血液）を収集した。診療情報及び血液は同意を得られた対象疾患患者全員より取得し、採血のタイミングから最も近い診療記録と紐づけた。肺組織は、診断の上でクライオバイオプシーが必要と判断された場合及び外科手術検体が得られた場合においてのみ収集した。クライオバイオプシーを収集する場合、1 患者より数カ所からクライオバイオプシーの採取を行い、縦に半割して一方を病理診断、もう一方をオミックスデータ取得に用いた。採取した肺組織及び血液は、下記のプロトコルで処理・保存し、1 ヶ月に 1 回当所に輸送（三井倉庫ホールディングス株式会社に委託）し、オミックスデータ取得はタカラバイオ株式会社に委託した。

##### 収集した診療情報

- ・ 診察記録 (97 項目)
- ・ 所見
  - 超音波 (5 項目)
  - 気管支鏡 (19 項目)
  - 外科生検 (16 項目)
  - CT 画像検査 (19 項目)
- ・ 検査
  - 血液＋尿 (103 項目)
  - 生理機能 (35 項目)
  - CT 画像
  - 肺組織
- ・ クライオバイオプシー：採取後チューブに回収し、速やかに $-80^{\circ}\text{C}$ で保存した。
- ・ 外科手術検体 (VATS)：採取後チューブに回収し、速やかに $-80^{\circ}\text{C}$ で保存した。
- ・ 上記と紐づけられた病理所見
- ・ 上記と紐づけられた病理画像からの特徴量
- ・ 血液
  - 血清 (プロテオーム解析用)：採血後、室温で静かに 5 回転倒混和し、室温で 30 分間静置した後にスイング式ローターを用いて、室温、 $1,300\times g$  で 10 分遠心した。上清をチューブに  $250\mu\text{L}$  ずつ分注して速やかに $-80^{\circ}\text{C}$ で保存した。
  - 血漿 (miRNA トランスクリプトーム解析用)：採血後、室温で静かに 10 回転倒混和し、スイング式ローターを用いて、室温、 $1,300\times g$  で 10 分遠心した。上清をチューブに  $250\mu\text{L}$  ずつ分注して速やかに $-80^{\circ}\text{C}$ で保存した。
  - 末梢血単核細胞 (PBMC、全ゲノム解析用)：採血後、室温で静かに 5 回転倒混和し、スイング式ローターを用いて室温、 $1,600\times g$  で 20 分遠心した。まとめて採血する場合、最初の検体から最後の検体の時間が 1 時間以内になるようにした。遠心後、ゲルバリアの上にある細胞層を乱さないようにして血漿層を約半分吸引し、パスツールピペットを用いて、細胞層全量をチューブに回収して速やかに $-80^{\circ}\text{C}$ で保存した。

## 肺組織からの核酸 (DNA・RNA) 同時抽出

肺組織サンプル (56 症例、80 検体) について、DNA・RNA 同時抽出キット (NucleoSpin RNA、NucleoSpin RNA/DNA Buffer Set) を用いて核酸抽出を行った。抽出された核酸はバイオアナライザーを用いて RNA 分画における DNA 混入の有無を確認し、RNA 分画は NucleoSpin RNA/DNA Buffer Set のマニュアル通りに DNase 処理を行なった。

## 肺組織からのマルチオミクスデータ収集

### i. 全ゲノム解析 (depth 30)

#### (1) TruSeq DNA ライブラリの作成

アコースティックソルビライザー Covaris (コバリス社) を用いて物理的に数百 bp に断片化した DNA の両末端を平滑化とリン酸化処理をした後、3'-dA 突出末端処理を行い、Index 付きアダプターを連結した。その後、アガロースゲル電気泳動によりアダプターを連結した DNA をサイズ選別し、それをを鋳型とし、PCR による増幅を行い、シーケンスライブラリとした。ライブラリの品質は Agilent2100 BioAnalyzer を用いて測定した。

#### (2) シーケンス解析

Illumina 社の NovaSeq6000 を用いて 150base 長ペアエンドシーケンスを行い、シーケンサー付属ソフト (NovaSeq Control Software v1.6.、Real Time Analysis (RTA) v3.4.、bcl2fastq2 v2.20) を用いて塩基配列 (リード配列) を得た。タグ配列に基づき塩基配列 (リード配列) を検体ごとに分類して fastq ファイルを取得した。

### ii. DNA メチル化解析

Illumina 社の Infinium MethylationEPIC BeadChip を用いて、Infinium HD Methylation Protocol Guide, Manual Protocol (15019519 v01) に従い、ゲノム DNA を用いてメチル化解析を実施した。

#### (1) Bisulfite 処理

ゲノム DNA (250ng) を EZ DNA Methylation Kit (ZYMO RESEARCH 社) を用いて Bisulfite 変換し、精製、回収を行った。

#### (2) 全ゲノム増幅、断片化、及び精製

Bisulfite 処理した DNA はアルカリ処理後に酵素による全ゲノム増幅を行なった。その後、酵素による断片化、イソプロパノール沈殿による精製の後、バッファーに再懸濁した。これを熱変性させ、Infinium MethylationEPIC BeadChip にアプライして、48℃で約 23 時間ハイブリダイゼーションを行なった。

#### (3) 一塩基伸長反応とスキャニング

ハイブリダイゼーション後、BeadChip をバッファーで洗浄し、一塩基伸長反応によってプローブ末端に一塩基の標識ヌクレオチドを導入、ハイブリダイズした DNA を変性・除去し、取り込ませた標識ヌクレオチドに対する蛍光色素標識抗体染色を行い、蛍光イメージを取得した。

#### (4) データ解析

取得した蛍光イメージデータを GenomeStudio/Methylation Module を用いて、Background Subtraction 及び Normalization を実施して解析を行なった。

### iii. トランスクリプトーム (RNA-seq)

#### (1) TruSeq Stranded mRNA ライブラリの作製

検体より PolyA+RNA を単離し、断片化により得られる RNA を鋳型とした逆転写反応により一本鎖 cDNA を合成した。これを鋳型として、dUTP を取り込ませた二本鎖 cDNA を合成した。得られた二本鎖 cDNA の両末端を平滑化・リン酸化処理した後、3'-dA 突出処理を行い、Index 付きアダプターを連結した。アダプターを連結した二本鎖 cDNA を鋳型とし、dUTP を持つ鎖を選択的に増幅しないポリメラーゼにより PCR 増幅を行い、シーケンスライブラリとした。ライブラリの品質は Agilent2100 BioAnalyzer を用いて測定した。

#### (2) シーケンス解析

Illumina 社の NovaSeq6000 を用いて 150base 長ペアエンドシーケンスを行い、シーケンサー付属ソフト

(NovaSeq Control Software v1.6.、Real Time Analysis (RTA) v3.4.、bcl2fastq2 v2.20) を用いて塩基配列 (リード配列) を得た。タグ配列に基づき塩基配列 (リード配列) を検体ごとに分類して fastq ファイルを取得した。

#### 血液からのマルチオミックスデータ収集

##### i. 全ゲノム解析 (depth 30)

肺組織を取得していない症例においては、PBMC (末梢血単核細胞) を用いた全ゲノム解析を実施した。プロトコルは、「肺組織からのマルチオミックスデータ収集 1. 全ゲノム解析」と同一である。

##### ii. トランスクリプトーム (miRNA-seq)

miRNA トランスクリプトーム解析には血漿を用いた。RNA 抽出は QIAGEN miRNeasy micro kit を用いた。各サンプルに QIAzol 700 $\mu$ l を添加して RNA 抽出後、14 $\mu$ l の nuclease free water で溶出した。ライブラリ調製キットは QIAseq miRNA Library Kit を使用し、input 量は 5 $\mu$ l とした。こうして調製したライブラリの quality check はバイオアナライザーと qPCR (2nM サンプルを使用) にておこなった。miRNA-seq は Nextseq 75bp シングルエンドで行い、1 サンプルあたり 2000 万リード以上取得することとした。RNA 抽出から miRNA-seq までの一連の作業は株式会社 DNA チップ研究所に委託した。

##### iii. プロテオーム解析 (DIA)

血清 200  $\mu$ l から MagCapture isolation kit (Fujifilm Wako) を用いて細胞外小胞を精製した。細胞外小胞中のタンパク質をトリス (2-カルボキシエチル) ホスフィンによる還元・ヨードアセトアミドによるアルキル化を行い、トリプシン消化、脱塩を行った。前処理したサンプルを Data independent acquisition (DIA) 法を用いた LC-MS/MS 解析を実施した。データ解析は、DIA 解析ソフト Spectranout を用いて実施し、欠損値には run-wise imputation を行った。Quality control として、15 検体毎に市販血清 1 検体を加えて、サンプル調製からデータ解析までの品質保証を行った。また質量分析の Quality control として HeLa 細胞消化物の DIA 解析を実施した。

2. 統合データ解析の手法開発とそれを用いた新規創薬ターゲット候補の同定：神奈川県立循環器呼吸器病センターにおいて収集を進めているマルチオミックスデータの解析に対応するため、本事業において理研 AIP との共同研究により開発された患者層別化 AI の改良を実施した。一部作業については政策基礎研究所への委託により実施した。

#### 3. オープンプラットフォームの構築：

本事業で作成される IPF/肺がんの疾患統合データベース、機能分子を特定するための AI 及び知識ベース等を多くの研究者等に利用してもらうための環境 (オープンプラットフォーム、以後 OPF) の構築と運営開始準備を進めている。OPF を介した公開を予定している診療情報の個人特定リスクを評価し、匿名加工情報を作成した。更に、本事業において開発された AI のうち 9 つについてガジェット化 (OPF 上で実行できるようにデプロイ) した。

#### 診療情報の個人特定リスク評価及び匿名加工情報の作成

OPF の構築及び運用にあたり、患者の診療情報など要配慮個人情報を取り扱う上でのリスク評価並びに匿名加工の必要性の評価、また匿名加工情報の作成が求められる。そのため、当該年度は令和 2 年度に実施した診療情報の個人特定リスク評価結果と匿名加工条件を基に匿名加工情報の作成を行なった。

#### AI ガジェットの作成

OPF システム基盤の開発を行なった特定非営利活動法人 システム・バイオロジー研究機構 (SBI) への委託により、本事業成果の AI から下記の 9 つを選択して AI ガジェットの作成を行なった。

JaMIE: A Japanese Medical Information Extraction System (研究分担者 黒橋禎夫)

kGCN: A graph-based deep learning framework for life science (研究分担者 奥野恭史)

Modified Diet Networks (研究分担者 浜本隆二)

Molenc: A molecular encoder using counted-unfolded fingerprints (研究分担者 山西芳裕)

Multiomics Analyzer (研究分担者 浜本隆二)

SFMEDM: space-efficient feature maps for string alignment kernels (研究分担者 田部井靖生)

Vanishing Rabking Kernels: A single-parameter ranking method with an applicability domain

(研究分担者 山西芳裕)

NamedEntityRecognizer (研究分担者 高村大也)

EntityLinker (研究分担者 高村大也)

4. 層別化 AI による解析及び解釈の精度向上: これまでに見出された IPF の特徴と紐づけられるタンパク質について、ネットワーク解析による関連パスウェイや上流因子の探索、EMT への影響、メタボローム解析による前述パスウェイの活性化の確認を実施した。

バイオインフォマティクス解析による出力結果解釈

上記解析によって見出されたタンパク質に関して、TargetMine や IPA (Ingenuity Pathway Analysis, QIAGEN 社) を用いた既存知識との照合による分析を行い、関与が疑われるパスウェイの推定や肺線維化に寄与する因子の絞り込みを行なった。

EMT 試験系の樹立

線維化の発症と進行には肺上皮細胞が上皮間葉転換 (Epithelial-mesenchymal transition ; EMT) により筋線維芽細胞に分化することが重要であることが広く示唆されている。そのため、解析により見出した分子の EMT への関与を評価する目的で、これらの分子の発現が報告されており、また TGF- $\beta$  添加により EMT が誘導されると報告されているヒト気道上皮細胞 BEAS-2B を用いて EMT 評価試験系の樹立に着手した。

EMT は、上皮性マーカーである E-cadherin の発現抑制と、間葉系マーカーである Fibronectin、 $\alpha$ -SMA 及び Snail の発現増強を指標として評価することとした。ATCC より購入した BEAS-2B 細胞を  $1\sim 7.5\times 10^4$  cells/mL (培地 ; Bronchial epithelial cell basal medium、Lonza 株式会社) で 96 well プレートへ播種し、24 時間後に TGF- $\beta$  (Proteintech) を終濃度 2.5、5、10、15 及び 20 ng/mL となるように添加した。TGF- $\beta$  添加 48 時間後に、SuperPrep® II Cell Lysis & RT Kit for qPCR (東洋紡株式会社) を用いて細胞溶解及び RT を行い、上述の EMT マーカー遺伝子群の発現量を qPCR により測定した。qPCR には、THUNDERBIRD Probe qPCR Mix (東洋紡株式会社) を用い、各マーカープローブには TaqMan Probe (サーモフィッシュサイエンティフィック株式会社) を用いた。本試験系の樹立は株式会社ケーエーシー (京都) に委託して実施した。

血清中代謝物の網羅的測定 (メタボローム)

大阪大学より入手した血清のうち、IPF 確定診断 10 例及び健常者 10 例について、キャピラリー電気泳動・フーリエ変換型質量分析計 (CE-FTMS) を用いて、糖リン酸、有機酸、アミノ酸、アミン等のイオン性代謝物質を網羅的に測定した。また液体クロマトグラフィー・飛行時間型質量分析計 (LC-TOFMS) を用いて遊離脂肪酸、アシルカルニチン、胆汁酸、ポリフェノール類等の中性代謝物を網羅的に測定した。なお、分析対象物質の物性に応じて、ポジティブモードとネガティブモードの 2 つのメソッドにて実施した。測定データは、検出ピークを標準データベース (約 1300 物質) と照合することにより帰属するとともに、糖代謝、アミノ酸代謝に関わる 110 物質に関しては濃度定量を行った。本測定は、ヒューマン・メタボローム・テクノロジー株式会社 (山形) に委託して実施した。

(倫理面への配慮)

本研究は、医薬基盤・健康・栄養研究所ならびに分担研究機関において倫理審査、承認を得た後、ヒトゲノム・遺伝子解析に関する倫理指針及び、人を対象とする医学系研究に関する倫理指針に従って遂行した。

## C. 研究結果

### 1. 臨床情報の収集と機械可読表現の開発 :

肺がん手術検体及びバイオプシー検体のオミックス解析の研究結果全般の詳細については、研究分担者 (浜本隆二) による報告書を参照されたい。

大阪大学医学部呼吸器・免疫内科のバイオバンク及び大阪大学医学部附属病院医療情報部より、IPF を含む間質性肺炎及び器質的な呼吸器疾患がないと診断された健常者について、令和 3 年度は呼吸機能検査値及びフローボリュームカーブと MDD 結果を同様に匿名化 ID が付与された形で入手した。MDD 結果を用いた解析は現在進行中である。

肺組織 129 症例を対象に下記のマルチオミックスデータを取得した。

- ・ 全ゲノム解析 (depth 30)
- ・ DNA メチル化解析
- ・ トランスクリプトーム (RNA-seq)

PBMC 526 症例を対象に下記のマルチオミックスデータを取得した。

- ・ 全ゲノム解析 (depth 30)

血漿 449 症例を対象に下記のマルチオミックスデータを取得した。

- ・ トランスクリプトーム (miRNA-seq)

血清 909 症例を対象に下記のマルチオミックスデータを取得した。

- ・ プロテオーム (DIA)

神奈川県循環器呼吸器病センターの臨床情報収集全般の詳細については、研究分担者（小倉高志）による報告書を参照されたい。

## 2. 統合データ解析の手法開発とそれを用いた新規創薬ターゲット候補の同定：

本事業において開発した患者層別化 AI (subset binding) において、メモリ負荷が大きいことが課題の一つであった。そのため、主にメモリ負荷の低減と解析対象データの探索範囲の絞り込みにより、これまでよりも高次元のデータに対して適用することが可能となった。

## 3. オープンプラットフォームの構築：

診療情報の個人特定リスク評価及び匿名加工情報の作成

当該年度は、大阪大学コホート診療情報の匿名加工情報の作成を完了した。また、オミックスデータや CT 画像、医療テキストについて個人特定リスクの評価と（必要に応じて）匿名加工情報作成の条件検討を完了した。

### AI ガジェットの実装

大阪大学コホートデータ（匿名加工済み診療情報、プロテオームデータ）の OPF 搭載を完了した。9 つの AI についてガジェット実装及び OPF 搭載を完了した。

### システムの維持・管理体制の整備

バックアップシステム及びデータベースの高速化システムを導入するとともに、オープンプラットフォームの基幹技術の導入元である企業からのサポート体制を整備した。

## 4. 層別化 AI による解析及び解釈の精度向上：

令和 2 年度に実施した大阪大学コホートの臨床情報を患者層別化 AI の入力データとして用いた解析の結果、患者基本情報を含む診療情報・胸部 CT 画像の読影所見・血液検査で IPF 患者に認められる特徴と紐づけられるタンパク質（以降、IPF 関連タンパク質と表記）を複数個見出すことに成功した。ネットワーク解析により、IPF 関連タンパク質において共通してつながるパスウェイの存在を明らかにし、IPF 患者の血液を用いたメタボローム解析によりこのパスウェイの活性化が起きているか確認を進めている。また、IPF 関連タンパク質の上流因子を見出し、これを阻害する薬剤を炙り出した。昨年度において IPF 関連タンパク質の多くが肺線維化部位において発現が亢進していることを確認していることから、この上流因子を阻害することにより IPF の治療あるいは進行を食い止めることが可能となることが示唆される。そのため、この上流因子の阻害が肺線維化において重要な役割を担うとされている EMT（上皮間葉転換）に対してどのように働くかを確認した。ヒト肺胞基底上皮腺癌由来 A549 細胞及びヒト気道上皮細胞 BEADS2B を用いた in vitro EMT 試験系において、前述の薬剤が濃度依存的に EMT を抑制することを見出した。

## D. 考察

当該年度の成果は、分担研究者との共同研究に関連するものを除くと大きく分けて 3 つである。すなわち、①大阪大学コホートの臨床情報解析による IPF 創薬標的候補の提示、②神奈川県立循環器呼吸器病センターコホートの臨床情報収集継続による疾患 DB 拡充、③OPF 構築と運用準備である。

①では、AIによってデータ駆動的に提示された創薬標的の阻害が実際に EMT を抑制することを確認しており、臨床情報を用いた解析ワークフローの有効性を更に強固に示すことができたと言える。また、②では世界でも稀な IPF マルチオミクスデータベースの構築に向けて着実に症例数を追加しており、次年度には本データの解析による創薬標的探索や発症メカニズム推定が本格化する見込みである。

③ではシステム基盤に大阪大学コホートデータと合計 10 つの AI ガジェットを搭載し、運用開始に向けた手続きが済み次第国内運用を開始する見込みである。

## E. 結論

OPF 国内運用開始時期が数ヶ月後ろ倒しになっているものの、当該年度研究計画の達成状況は概ね良好である。AIによって提示された創薬標的候補が有望であることを支持する結果を得ることができた。

## F. 研究発表

### 1. 論文発表

- 1) Yayoi Natsume-Kitatani, Mari N Itoh, Yoshito Takeda et al. Data-driven patient stratification and drug target discovery by using medical information and serum proteome data of idiopathic pulmonary fibrosis patients, 29 March 2022, PREPRINT (Version 2) available at Research Square [https://doi.org/10.21203/rs.3.rs-405195/v3]
- 2) Hosomi, K., Saito, M., Park, J., Murakami, H., Shibata, N., Ando, M., Nagatake, T., Konishi, K., Ohno, H., Tanisawa, K., Mohsen, A., Chen, Y.A., Kawashima, H., Natsume-Kitatani, Y., Oka, Y., Shimizu, H., Furuta, M., Tojima, Y., Sawane, K., Saika, A., Yonejima, Y., Takeyama, H., Matsutani, A., Mizuguchi, K., Miyachi, M., & Kunisawa, J. (2022). Unique metabolic profiles of *Blautia wexlerae* achieve beneficial effects for the control of obesity and type 2 diabetes. (preprint)
- 3) Hioki K., Hayashi T., Natsume-Kitatani Y., Kobiyama K., Temizoz B., Negishi H., Kawakami H., Fuchino H., Kuroda E., Coban C., Kawahara N, m Ishii H. J., Machine-learning-assisted screening of herbal medicine extracts as vaccine adjuvants, *Frontiers in Immunology*, 13 (2023)
- 4) 佐藤壮, 疋田正喜, 本田晴香, 杉山大介, 中山哲夫, 村上正晃, 内田萌菜, 北條慎太郎, 田中くみ子, 青枝大貴, 熊谷雄太郎, 竹内直志, 大野博司, 河本宏, 永野誠治, 案浦健, 荒木球沙, 久枝一, 高田光輔, 渡辺登喜子, 日比谷健司, 藤田次郎, 森田公一, 柴田岳彦, 村上耕介, 宮澤正顕, 四柳宏, 小原恭子, 小原道法, 渡邊真弥, 氣賀恒太郎, 崔 龍洙, 浜口毅, 山田正仁, 反町典子, 小原乃也, 廣田圭司, 岩淵禎弘, 橋本真一, 伊藤眞里, 武田吉人, 大木伸司, 山村隆, 藤井貴之, 西英一郎, 森本真司, 奥村龍, 竹田潔, 大島茂, 薬研究者・アカデミア研究者が知っておくべき 最新の免疫学とその応用技術, 技術情報協会, ISBN 978-4-86104-856-2, 2021. 8. 30
- 5) 夏目やよい, 新薬創出を加速する人工知能の開発 ~データ駆動型創薬ターゲット探索プラットフォームの構築~, 医学情報誌「あいみつく」第 42 巻 3 号, 2021. 9. 3
- 6) 鎮西清行, 大竹正規, 今関剛, 加藤健太郎, 中川肇, 山本真由美, 牧野奈緒, 貝田章太郎, 上島努, 鎌田英世, 矢島弘士, 阪本剛, 三澤将史, 杉山崇, 森公彦, 牛久保智宏, 長谷公隆, 医療機器・ヘルスケアにおける ICT 技術開発と規制対応, 株式会社情報機構, ISBN 978-4-86502-218-6, 2021. 9. 27
- 7) 夏目やよい, 機械学習によって加速される次世代アジュバント開発, 週刊医学のあゆみ, 第 279 巻 10 号 PP961-964, 2021. 12. 4
- 8) 藤本明洋, 宇津野泰弘, 松本直通, 新井田要, 剛澄仁, 浦大樹, 青木弘太郎, 横井毅, 梅原崇史, 新城恵子, 近藤豊, 大庭成喜, 間野達雄, 岩田淳, 山下暁朗, 今井大達, 田口善弘, 廣明秀一, 近藤格, 松尾光一, 熊代宗弘, 山本英樹, 林宣宏, 郷康広, 栗本一基, 鈴木勉, 高野淳, 伊賀淳一, 一石英一郎, 大澤毅, 松崎芙美子, 久保田浩行, 茅野光範, 西海信, 三輪洋人, 村田唯, 岩本和也, 柴田龍弘, 中村慎吾, 橋本真一, 岩淵禎弘, 岩田岳, 小巻翔平, 清水厚志, 藤田直也, 中井謙太, 伊藤眞里, 笹栗弘貴, 船山学, 松村剛, 成瀬紘也, 辻省次, 田中章景, 土井宏, 佐々木征行, 松林泰毅, 三條伸夫, 木村大樹, 尾崎紀夫, 本成登貴和, 千葉奈津子, 額賀重成, 河野隆志, 徳永英樹, 川内大輔, 白石椋, 波田野琢, 王子悠, 服部信孝, 村山圭, 野々山恵章, 堀之内智子, 野津寛大, 飯島一誠, 鶴木元香, 木村彰方, 宮田崇, 川口頌平, 萩原大輔, 有馬寛, 疾患原因遺伝子・タンパク質の解析技術と創薬/診断技術への応用, 技術情報協会, ISBN 978-4-86104-877-7, 2022. 3. 31

## 2. 学会発表

- 1) 伊藤眞里, 夏目やよい, 黒田正孝, 松村泰志, 武田吉人, 熊ノ郷淳, 水口賢司, 「新薬創出を加速する人工知能(AI)の開発」:大阪大学 IPF コホートデータ収集と知識処理, 第 61 回日本呼吸器学会学術講演会, Web 開催, 2021. 4. 24
- 2) 安部祐子, 武田吉人, 木庭太郎, 福島清春, 白山敬之, 平田陽彦, 足立淳, 伊藤眞里, 横井崇, 南俊行, 栗林康造, 木島貴志, 熊ノ郷淳, エクソソームの次世代プロテオミクスによる悪性胸膜中皮腫の新規バイオマーカー開発, 第 61 回日本呼吸器学会学術講演会, Web 開催, 2021. 4. 24
- 3) 木庭太郎, 平田陽彦, 武田吉人, 伊藤眞里, 吉村華子, 熊ノ郷淳, エクソソームのプロテオミクスによる肺胞径と相関する肺気腫バイオマーカー :Fibulin-3 第 61 回日本呼吸器学会学術講演会, Web 開催, 2021. 4. 24
- 4) 榎本貴俊, 武田吉人, 網屋沙織, 足立雄一, 新津敬之, 原伶奈, 野田成美, 白井雄也, 白山敬之, 三宅浩太郎, 平田陽彦, 足立淳, 伊藤眞里, 夏目やよい, 熊ノ郷淳, 次世代プロテオミクスによる進行性線維化を伴う間質性肺疾患の新規バイオマーカー探索(PRISM), 第 61 回日本呼吸器学会学術講演会, Web 開催, 2021. 4. 24
- 5) 白井雄也, 武田吉人, 足立雄一, 榎本貴俊, 網屋沙織, 野田成美, 福島清春, 白山敬之, 三宅浩太郎, 平田陽彦, 足立淳, 夏目やよい, 伊藤眞里, 熊ノ郷淳, 「PRISM」データから見えてきた新たな線維化バイオマーカー, 第 61 回日本呼吸器学会学術講演会, Web 開催, 2021. 4. 24
- 6) 野田成美, 武田吉人, 網屋沙織, 足立雄一, 榎本貴俊, 原伶奈, 新津敬之, 白井雄也, 白山敬之, 三宅浩太郎, 平田陽彦, 足立淳, 夏目やよい, 伊藤眞里, 熊ノ郷淳, 次世代プロテオミクスによる上葉優位型肺線維症の新規バイオマーカー探索(PRISM), 第 61 回日本呼吸器学会学術講演会, Web 開催, 2021. 4. 24
- 7) 網屋沙織, 武田吉人, 足立雄一, 榎本貴俊, 新津敬之, 野田成美, 原伶奈, 白井雄也, 白山敬之, 三宅浩太郎, 平田陽彦, 足立淳, 伊藤眞里, 夏目やよい, 熊ノ郷淳, エクソソームの次世代プロテオミクスによるサルコイドーシスの新規バイオマーカー探索(PRISM), 第 61 回日本呼吸器学会学術講演会, Web 開催, 2021. 4. 24
- 8) 野田成美, 武田吉人, 吉村華子, 安部祐子, 管泰彦, 福島清春, 白山敬之, 三宅浩太郎, 平田陽彦, 足立淳, 伊藤眞里, 熊ノ郷淳, エクソソームのプロテオミクスによる小細胞肺癌の新規バイオマーカー探索研究, 第 61 回日本呼吸器学会学術講演会, Web 開催, 2021. 4. 24
- 9) 夏目やよい, 伊藤眞里, 松村泰志, 武田吉人, 熊ノ郷淳, 水口賢司, 野田成美, 上田修功, 特発性肺線維症に対する創薬標的探索を目指した人工知能の開発, 第 61 回日本呼吸器学会学術講演会, Web 開催, 2021. 4. 24
- 10) 足立雄一, 武田吉人, 榎本貴俊, 網屋沙織, 野田成美, 原伶奈, 新津敬之, 白井雄也, 福島清春, 白山敬之, 三宅浩太郎, 平田陽彦, 夏目やよい, 足立淳, 伊藤眞里, 熊ノ郷淳, 次世代プロテオミクスとパイオインフォマティクスが紐解 IPF と NSIP の病態解明, 第 61 回日本呼吸器学会学術講演会, Web 開催, 2021. 4. 24
- 11) 伊藤眞里, 薬効薬理研究における動物実験の現状と課題, 第 68 回日本実験動物学会総会, Web 開催, 2021. 5. 21
- 12) 伊藤眞里, 新薬創出を加速する人工知能の開発, 第 3 回日本メディカル AI 学会, Web 開催, 2021. 6. 12
- 13) 夏目やよい, 相崎健一, 北嶋聡, Samik GHOSH, 北野宏明, 水口賢司, 菅野純, 毒性オミクスにおけるエピジェネティクス情報を加えた人工知能解析 PPAR $\alpha$  リガンドの比較毒性オミクス, 第 48 回日本毒性学会学術年会, Web 開催, 2021. 7. 7
- 14) 伊藤眞里, 人工知能技術を用いた特発性肺線維症の創薬標的の探索, 第 15 回 日常診療に役立つ呼吸器セミナー, Web 開催, 2021. 7. 17
- 15) 夏目やよい, 全体概要・特発性肺線維症, 新薬創出を加速する人工知能の開発 令和二年度成果報告会, Web 開催, 2020. 7. 20
- 16) 夏目やよい, 演題なし デモンストレーション, 研究者のための+ $\alpha$ シリーズ Vol.8 How to Give an Impressive Pitch Presentation in Global Situations ~研究者の国際 R&D 戦略としてのセールストーク~, Web 開催, 2021. 8. 13
- 17) 伊藤眞里, エクソソーム内プロテオミクスから難治性呼吸器疾患の病態解明へ, キアゲン IPA ユーザーグループミーティング, Web 開催, 2021. 9. 8
- 18) 夏目やよい, トランスクリプトームデータからの毒性発現メカニズム推定と安全性評価, 幹細胞を用いた化学物質リスク情報共有化コンソーシアム第 3 回 scChemRISC 研究会, Web 開催, 2021. 10. 12



- 19) 伊藤真里, 「新薬創出を加速する人工知能の開発」 特発性肺線維症患者エクソソーム内プロテオミクスと診療情報の活用, 第8回日本細胞外小胞学会学術集会, Web開催, 2021. 10. 19
- 20) 夏目やよい, 新薬創出を加速する人工知能の開発 ～臨床情報を活用した創薬標的探索～, CBI学会2021年大会, Web開催, 2021. 10. 28
- 21) 夏目やよい, 新薬創出を加速する人工知能の開発, 関西共創の場第3回若手人材育成セミナー, Web開催, 2021. 11. 27
- 22) 夏目やよい, 「新薬創出を加速する人工知能の開発」事業の紹介、京都大学大学院医学研究科「医学領域」産学連携推進機構(KUMBL)主催シーズ・ニーズマッチング, 2022. 1. 27
- 23) 夏目やよい, データ駆動型安全性評価の諸問題, 幹細胞を用いた化学物質リスク情報共有化コンソーシアム2022年度年会, 京都, 2022. 3. 1

G. 知的財産権の出願・登録状況(予定を含む。)

1. 特許取得

国際出願番号: PCT/JP2022/003368

発明の名称: 複数の項目を関係付けるための方法、システム、及びプログラム

発明者: 夏目やよい、上田修功

出願人: 国立研究開発法人医薬基盤・健康・栄養研究所

出願日: 2022年1月28日

出願番号: 特願2022-50865

発明の名称: 特発性肺線維症の治療または予防剤

発明者: 夏目やよい、伊藤真里、黒田正孝、水口賢司、足立淳、朝長毅、熊ノ郷淳、武田吉人、上田修功

出願人: 国立研究開発法人医薬基盤・健康・栄養研究所

国立大学法人大阪大学

国立研究開発法人理化学研究所

出願日: 2022年3月25日

2. 実用新案登録

なし

3. その他

なし