

厚生労働科学研究費補助金研究報告書表紙

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）

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

令和 2 年度 総括・分担研究報告書

研究代表者 夏目 やよい

令和 4 年（2022）年 5 月

厚生労働科学研究費補助金研究報告書目次

目 次

I. 総括研究報告	
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 夏目 やよい	3
II. 分担研究報告	
1. 「新薬創出を加速する症例データベースの構築」に関する研究 小倉 高志	12
2. 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に関する研究 高村 大也	14
3. 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 西村 紳一郎	17
4. 肺がん統合データベース構築及びAI技術を用いたオミックス解析 浜本 隆二	20
5. 肺がん患者のシングルセル・マルチオミックス解析 西岡 安彦	26
6. 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に関する研究 奥野 恭史	30
7. 層別化AIによる解析および解釈の精度向上のための研究 熊ノ郷 淳	38
III. 研究成果の刊行に関する一覧表	40

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）

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 2 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）を収集した。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。これらのタンパク質のバイオインフォマティクス解析により、IPF の創薬標的候補を提示した。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のプロトタイプ構築や格納するデータの匿名加工条件設定、運用体制を策定するための先行事例調査やニーズ調査を実施した。

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. 既知 (背景) 情報の収集と知識ベース化: 公共・市販のデータベースを TargetMine システムに追加統合した。追加したデータベースは下記の通りである。

- ・ MGeND: 疾患名・年齢・性別などの臨床情報と遺伝子変異データとを統合的に扱うデータベース
- ・ PoSSuM: タンパク質上の低分子結合サイトを推定するためのデータベース

論文からの自動知識抽出を可能にするモデル構築のため、臨床系文献のアノテーション基準となるガイドラインの作成・修正を行った。対象となる文献 (167 件) は当所専門家により選抜され、アノテーション作成は産総研人工知能研究センターのアノテーター 3 名が独立して業務に当たることにより頑健なアノテーションを実施し、随時アノテーション中のデータを解析して一つのエンティティに複数の異なる ID がアサインされていないか確認・修正をおこなった。このプロセスは産総研人工知能研究センターのプロトコルに則って

実施された。アノテーションガイドライン作成は当所専門家及び産総研人工知能研究センターの担当研究者・アノテーターと議論の上、本事業の対象疾患である呼吸器疾患の専門用語を適切にアノテーションできるように実践に即して策定した。モデル構築は産総研人工知能研究センターへの委託により実施した。

2. 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、下記のオ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

昨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血清と、血清にひもつく診療情報（診察記録、CT 画像およびその読影所見、血液検査値、患者基本情報および初診時間診票）を大阪大学病院医療情報部より収集した。また、検査の結果、器質的な呼吸器疾患を有しないと診断された方々を健常者とし、同様に血清および診療情報を収集した。なお、血清はバイオバンク番号、診療情報は、いずれも研究用の匿名化 ID が付与された形で入手した。今年度は、上記診療情報の中で、病院側の COVID-19 対応のため、入手が遅れていた CT 画像の残り欠損データを同様に匿名化 ID が付与された形で入手した。また免疫染色を行うために、肺がんを合併して手術に至った IPF 患者肺における線維化部位および正常部位のホルマリン固定標本も入手した。呼吸機能検査値およびフローボリュームカーブについては、匿名化作業を行うため、医療情報部から呼吸内科へ転送することとした。

i. 診療情報の収集と構造化

診療情報のうち、診察記録は、必要な情報 102 項目をあらかじめリストして作成したテンプレートを用いて医師が入力もしくは自由記載事項テキストデータよりマニュアルでキュレーションを行って構造化データとした。読影所見は、マニュアルあるいは自然言語処理手法を用いて、重要表現にタグ付けを行い、部位と病変のペアとそれらのペアが認められる（positive）、認められない（negative）および疑われる（suspect）の 3 つの場合分けからなる構造化データを得た。血液検査値は、173 項目の重要項目を選択してキュレーションを行い構造化した。初診時間診票、基本情報は、診察記録のテンプレート項目に情報をキュレーションして追加した。また、構造化にあたり、欠損値補填の手法について、欠損の意味を確認するとともに、主として健常者基準値を用いて補填を行った。なお、呼吸機能データについては、匿名化処理の手法構築を行った。

ii. プロテオーム解析

血清 200 μ 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。

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＋オミックスデータまたは肺組織や血液）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℃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℃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×g で 10 分遠心した。上清をチューブに 250μL ずつ分注して速やかに-80℃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×g で 10 分遠心した。上清をチューブに 250μL ずつ分注して速やかに-80℃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×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℃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ックスデータ収集

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、シーケン斯拉イブラリとした。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 700µl を添加して RNA 抽出後、14µl 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 5µl 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 (DIA)

プロテオーム解析には血清を用いた。エクソソーム抽出条件検討の結果、大阪大学コホートの血清中エクソソームプロテオーム解析と同じく MagCapture isolation kit (Fujifilm Wako) を用いることに決定した。

3. 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：本事業において理研 AIP との共同研究により開発された患者層別化 AI を用いて、これまでに収集した大阪大学コホートの臨床情報（602 件の血清中エクソソーム プロテオームデータ及びそれと紐づく診療情報）のデータ解析を実施した。診療情報の欠損値は、健常者の代表値を用いて補填してから解析に用いた。

患者層別化 AI を用いた臨床情報の解析

患者層別化 AI を用いた解析では、紐付けるデータ 2 種類を行列の形で入力する必要がある。そこで上記臨床情報の解析時には①胸部 CT 画像の読影所見（one-hot ベクトル化による構造化済みデータ）とプロテオームデータ、②基本患者情報などからなる診療情報とプロテオームデータ、③血液検査値とプロテオームデータ、の 3 通りの組み合わせで入力し、それぞれにおいて出力された結果の中から特に IPF の特徴（①「網状影」「牽引性気管支拡張」「蜂巣肺」、②「SpO2」「労作時呼吸器困難」など、③「KL-6」「肺サーファクタントプロテイン D」「アンギオ変換酵素」など）との連関が示唆されるタンパク質を探索した。

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。

IPF 患者肺組織・病変部位におけるタンパク質の免疫染色

大阪大学呼吸器内科より入手したホルマリン固定済みの肺組織から FFPE (Formalin fixed paraffin embedded ; ホルマリン固定パラフィン包埋) ブロックを作製し、滑走式マイクロームを用いて、厚さ 4 μ m に薄切して未染色標本を作製した。未染色標本は HE (ヘマトキシリンエオジン) 染色、MT (マッソントリクローム) 染色を実施し、炎症および線維化を確認した。さらに、未染色標本を用いて、解析によって見出されたタンパク質の特異抗体を用いる IHC (Immunohistochemistry;免疫組織化学)法による染色を実施した。検出試薬は OptiView DAB ユニバーサルキットを使用し、また、抗体ごとに陰性コントロールとして抗体希釈液を用いた染色を同時に実施し、各抗体に対する IHC 染色の評価を行った。なお、染色標本は明視野スライドスキャナーを用いて、対物レンズ 20 倍の設定で撮影し、画像ファイルを SVS 形式で保存した。この標本作製、染色作業はジェネティックラボ株式会社（北海道）に委託して実施した。

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- β 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、 α -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 1~7.5 $\times 10^4$ 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- β (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- β 添加 48 時間後に、SuperPrep® II Cell Lysis & RT Kit for qPCR (東洋紡株式会社) を用いて細胞溶解および 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トフィック株式会社) 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 例および健常者 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。

4. 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）を構築するため、OPF システム基盤プロトタイプ of 構築・診療情報の個人特定リスク評価及び二次匿名加工条件検討・想定ユーザーへのヒアリングや先行事例調査を行なった。

OPF システム基盤プロトタイプ of 設計と構築

特定非営利活動法人 システム・バイオロジー研究機構（SBI）との連携により、OPF システム基盤のプロトタイプを構築した。SBI が開発したプラットフォームである Garuda (<http://www.garuda-alliance.org/>) で用いられている技術（SBI が特許保有）により、OS やプログラム言語に依存せずデータベースや AI を相互連結させることができるように設計・構築した。OPF へのアクセスはアカウント発行により制御し、OPF 上で閲覧・解析可能なデータはカテゴリーごと（「Medical Record」「miRNA」「mRNA」「Genome」「Epigenome」「Protein」）、プロジェクトごと（例えばコホートによる分類）に選択できるように設計し、AI は「ガジェット」として逐次追加可能な仕様とした。また、本システム基盤上にガジェット化した患者層別化 AI を搭載し、簡便な操作により実行可能となっていることを確認した。

診療情報の個人特定リスク評価及び二次匿名加工条件検討

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、また匿名加工作業の実施が求められる。そのため、当該年度は IPF の診療情報の各項目について匿名加工作業の観点からの類型化を行い、匿名加工を必要とする項目の絞り込みを行なった。さらに、絞り込まれた項目に対して個人特定リスクの評価を行なった。この結果をもとに、匿名加工作業を実施する必要があると判断された各項目について、その個人特定リスクに応じた具体的な匿名加工手法を検討・設定した。

想定ユーザーへのヒアリングや先行事例調査

本事業の成果公開にあたり、既に OPF やそれに類似したシステムを運用している団体・企業の外部公開化の手法を把握・参考にする（先行事例調査）とともに、本事業 OPF の想定ユーザーとなる産業界などからのニーズを把握（想定ユーザーへのヒアリング）しておく必要がある。そのため、OPF の適切な運営体制の策定、ユーザーニーズの高いシステムの構築、持続運営可能なモデルの実現などに向けて情報収集を行なった。先行事例調査では、外部公開しているプラットフォームのシステム構築に関する全体像や構築方法、構築時の課題、法規制への対応状況、収益構造等を調査対象とした。想定ユーザーへのヒアリングでは、製薬・食品・ヘルスケア・診断分野・メディカル IT 分野の企業などを対象とし、各企業で求められているデータベースの内容、データベース利用に際して拠出可能な費用、データベースを利活用する上での社内・団体内での情報管理体制、データベースを利活用する上での課題点・問題点等を調査対象とした。さらに、医療用医薬品市場の市場規模および患者動向に関する情報に加えて、OPF 構築に必要な医療 IT 技術（インシリコ創薬・創薬支援システム、医療ビッグデータ分析システム、医療ビッグデータ向けクレンジングサービスなど）の市場規模についても併せて調査を行なった。

（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・遺伝子解析研究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. 既知（背景）情報の収集と知識ベース化：

計画通りにデータベースの統合を完了した。

論文から分子・病態に関する知識抽出を行う AI 開発のために必要な言語リソースの構築において、産総研人工知能研究センターと共同でアノテーションガイドラインの作成を行い、肺がん・臨床系論文 167 件のアノテーションを完了した。当該年度までに構築した言語リソースは研究分担者（高村大也）が担当する論文からの知識自動抽出技術開発において学習データとして使用された。言語リソース構築・知識自動抽出技術に関する研究成果全般の詳細については、研究分担者（高村大也）による報告書を参照されたい。

2. 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および大阪大学医学部附属病院医療情報部より、IPFを含む間質性肺炎およ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2年度はCT画像156名およびこれまでに入手した情報の中の欠損分を入手した。なお、呼吸機能データについては、匿名化処理の手法構築のみを行った。

これまでに入手済みの血清および採血日にひもついた診療情報を構造化して、患者層別化AIでの解析を試行した（本研究結果は次項「3. 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」を参照）。

神奈川循環器呼吸器病センターより、下記の臨床情報を取得した（2021年3月31日現在）。

- ・ 研究同意取得 令和2年度：376名（同意撤回4例）、累計：874名（同意撤回5例）
- ・ 血液 令和2年度：429例（累計：855例）
- ・ 肺組織（クライオバイオプシー・外科的肺生検・外科手術検体）令和2年度：111例
- ・ 臨床情報 累計：715例

肺組織126症例（176検体）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）

PBMC197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301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）

また、COVID-19がSARS-CoV-2感染による間質性肺炎であると考えられる所見を得たことから、COVID-19患者のPBMC30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）

神奈川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。

3. 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：

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を複数個見出すことに成功した。これらのタンパク質は、IPFや線維化に関する既知情報と照合して矛盾がなく、かつこれまでIPFとの関与が報告されていないパスウェイがIPF発症に寄与している可能性を示唆する結果が得られた。また、これらのタンパク質は血清中エクソソームだけでなくIPF患者の肺組織・病変部位においても量的変動が認められた。

4. オープンプラットフォームの構築：

オープンプラットフォーム（OPF）システムの設計と構築

OPFシステム基盤プロトタイプを医薬基盤・健康・栄養研究所サーバー上に構築し、搭載された患者層別化AIガジェットの動作確認を完了した。すなわち、テストデータを患者層別化AIガジェットで読み込み、基盤上でプログラムを実行して出力を画面上に表示させることができることを確認した。また、格納する臨床情報へのアクセス画面を設計・確認した。

所内で別途行っているデータ連携基盤とシステム等を共通とすべきかどうかを検討したが、コンピュータシステムの統合は行わず、運用体制だけを共通の方式で行うこととした。上記のごとくプロトタイプでシステムの動作が確認できたので、プロトタイプと同様に医薬基盤・健康・栄養研究所サーバー上にon premiseで本番環境を構築した。

また、事業終了後も OPF に AI やデータベースを追加することができるように、これらのガジェット化を行うための環境も用意した。

診療情報の個人特定リスク評価及び二次匿名加工条件検討

IPF の診療情報（大阪大学コホート、神奈川県立循環器呼吸器病センターコホート）の全項目について類型化、匿名加工の必要性の有無に関する評価、個人特定リスク評価、匿名加工が必要であると判断された項目に対する匿名加工条件設定を完了した。ただし、オミックスデータ、画像、医療テキストについては当該年度は対象外であり、次年度の実施を予定している。

想定ユーザーへのヒアリングや先行事例調査

先行事例調査については国内外の組織または機関 6 件を対象に調査を完了した。また、想定ユーザーへのヒアリングについては 25 社を対象に調査を完了した。本調査結果をもとに、情報セキュリティと利便性のバランスやエコシステム構築の重要性を熟慮しながら OPF 国内運用の基本構想策定を開始した。

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 プロトタイプ構築と運用準備開始である。

①では、これまでに収集した臨床情報、開発した患者層別化 AI、結果解釈を支援するツールの拡充の全てを結集し、IPF の特徴（診療情報の項目）と関連のあるタンパク質及びその結果から導き出される IPF 創薬標的候補を見出すことに成功した。この結果は、IPF のように分子レベルで疾患発症・重症化メカニズムの理解が進んでいない難病であっても臨床情報を収集し適切な AI による解析を行うことで創薬標的探索が可能であることを示唆する重要な成果である。本解析によって見出されたタンパク質・IPF 創薬標的候補について、IPF との関連や創薬標的としての有望性をより詳細に検討するほか、その結果を AI 開発にフィードバックすることで更なる精度向上を目指す。

また、②では世界でも稀な IPF マルチオミッ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、本データから創薬標的探索や発症メカニズム推定を行う AI 開発を並行して遂行していく。

③ではシステム基盤のプロトタイプ構築が完了したことから、今後は本基盤に搭載するデータや AI の加工（データの匿名加工や AI のガジェット化）を本格化し、令和 3 年度末国内運用開始を目指す。

E. 結論

コロナ禍により IPF 肺組織収集症例数を下方修正したものの、当該年度研究計画の達成状況は概ね良好である。AI による臨床情報の解析により、創薬標的候補を提示するに至った。本事業の成果であるデータベース、AI を広く共有するためのプラットフォームのプロトタイプが完成し、国内運用開始に向けた情報収集も予定通り完了した。

G. 研究発表

1. 論文発表

- 1) Esaki T., Horinouchi T., Natsume-Kitatani Y., Nojima Y., Sakane I., Matsui H., Estimation of relationships between gene expression and chemical substructures: Adapting canonical correlation analysis for small sample data by gathered features using consensus clustering, *Chem-Bio Informatics Journal* 20:58-61, 2020
- 2) Allendes Osorio R., Nyström-Persson J., Nojima Y., Kosugi Y., Mizuguchi K., Natsume-Kitatani Y., Panomicon: A web-based environment for interactive, visual analysis of multi-omics data, *Heliyon* 6(8): e04618, 2020
- 3) Nojima Y., Takeda Y., Maeda Y., Bamba T., Fukusaki E., Itoh M., Mizuguchi K., Kumanogoh A., Metabolomic analysis of fibrotic mice combined with public RNA-Seq human lung data reveal potential diagnostic biomarker candidates for lung fibrosis, *FEBS Open Bio* 11(10):2427-2436, 2020
- 4) Teranishi Y., Kuwahara H., Ueda M., Takemura T., Kusumoto M., Nakamura K., Sakai J., Kimura T., Furutani Y., Kawashima M., Imokawa G., Nogami-Itoh M., Sphingomyelin Deacylase, the Enzyme Involved in the Pathogenesis of Atopic Dermatitis, Is Identical to the β -Subunit of Acid Ceramidase, *International Journal of Molecular Sciences* 22(21):8789, 2020

- 5) Tripathi L., Itoh M., Takeda Y., Tsujino K., Kondo Y., Kumanogoh A., Mizuguchi K., Integrative Analysis Reveals Common and Unique Roles of Tetraspanins in Fibrosis and Emphysema, *Frontiers in Genetics* 11:585998, 2020

2. 学会発表

- 1) Koba T., Takeda Y., Narumi R., Nojima Y., Futami Y. N., Edahiro R., Itoh M. N., Shirai Y., Adachi J., Tomonaga T., Kumanogoh A., Targeted Proteomics of Serum Extracellular Vesicles to Identify Novel Biomarkers For COPD, ISEV2020, Web 開催, 2020.5.22
- 2) 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 67 回日本実験動物学会総会, 講演要旨集のみ発行, 2020.5.25
- 3) 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 令和元年度成果報告会, Web 開催, 2020.7.20
- 4) Koba T., Takeda Y., Narumi R., Nojima Y., Futami Y. N., Edahiro R., Itoh M. N., Shirai Y., Adachi J., Tomonaga T., Kumanogoh A., Targeted Proteomics of Serum Extracellular Vesicles To Identify Novel Biomarkers For COPD, ATS2020, Web 開催, 2020.8.5
- 5) 伊藤眞里, 機械学習を活用した創薬研究のテーマ創出, 技術情報協会セミナー, Web 開催, 2020.8.25
- 6) 夏目やよい, 臨床情報からのデータ駆動的な患者層別化に向けて, 第 9 回生命医薬情報学連合大会 IIBMP2020, Web 開催, 2020.9.1
- 7) 滝本宜之, 武田吉人, 木庭太郎, 玄山宗到, 広瀬雅樹, 新井徹, 伊藤眞里, 朝長毅, 熊ノ郷淳, 井上義一, エクソソームのプロテオミクスによるリンパ脈管筋腫症の新規血液バイオマーカーの同定, 第 60 回日本呼吸器学会学術講演会, Web 開催, 2020.9.20
- 8) 白井雄也, 武田吉人, 中山真菜, 木庭太郎, 二見悠, 足立淳, 朝長毅, 伊藤眞里, 水口賢司, 熊ノ郷淳, 血清エクソソームの最新定量プロテオミクスを用いた線維化マーカーの探索, 第 60 回日本呼吸器学会学術講演会, Web 開催, 2020.9.20
- 9) 安部祐子, 武田吉人, 吉村華子, 木庭太郎, 管泰彦, 福島清春, 白山敬之, 岩堀幸太, 平田陽彦, 足立淳, 伊藤眞里, 熊ノ郷淳, エクソソームの最新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探索, 第 7 回日本細胞外小胞学会, Web 開催, 2020.10.26
- 10) 木庭太郎, 武田吉人, 鳴海良平, 野島陽水, 伊藤眞里, 中山真菜, 吉村華子, 白水崇, 足立淳, 朝長毅, 熊ノ郷淳, 血清エクソソームのターゲットプロテオミクスによる新規 COPD バイオマーカー探索, 第 7 回日本細胞外小胞学会, Web 開催, 2020.10.26
- 11) 伊藤眞里, AI を用いる創薬研究テーマ探索:臨床情報収集と知識処理の自動化への挑戦, CBI 学会 2020 年大会, Web 開催, 2020.10.28
- 12) 夏目やよい, 官民研究開発投資拡大プログラム PRISM 創薬 AI 社会実装への展望, IQVIA ジャパンメディアセミナー, Web 開催, 2020.12.8

F. 知的財産権の出願・登録状況 (予定を含む。)

1. 特許取得

複数の項目を関係付けるための方法、システム、およびプログラム、夏目やよい、上田修功、2021.1.29、特願 2021-13264、国立研究開発法人医薬基盤・健康・栄養研究所、理化学研究所

2. 実用新案登録

なし

3. その他

なし

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))
分担研究報告書

課題名 : 「新薬創出を加速する症例データベースの構築」に関する研究
研究分担者名 : 小倉 高志 神奈川県立循環器呼吸器病センター 所長

研究要旨

指定難病の一つである特発性肺線維症を主体とした間質性肺炎患者に関する診療情報と患者サンプル（血液、肺組織）の収集を行う。さらに収集した患者サンプルのオミックス解析を実施する。それにより得られた臨床情報（診療情報および患者サンプルから得たオミックスデータ）を格納した疾患統合データベースを構築する。さらに得られた臨床情報を医薬基盤研にて解析することにより、創薬ターゲット分子の発見・同定を行う。

A. 研究目的

特発性肺線維症 (IPF) 及び間質性肺炎（疑いを含む）患者に関する診療情報および患者サンプルの収集を行い、臨床情報（診療情報および患者サンプルから得たオミックスデータ）を統合し格納した疾患統合データベースを完成させる。

B. 研究方法

IPF を含む間質性肺炎（疑いを含む）と診断した患者から臨床情報、各種検査結果の抽出、及び研究用に血液 10mL を採取する。また、研究同意を得られた患者からの気管支鏡検査、外科的肺生検からの肺組織採取を行う。臨床情報は、初診時間診票の記載データと電子カルテ日常診療医師入力記録から抽出を行う。各種検査結果は、胸部 CT、X 線画像、血液検査データ（血算、肝機能、腎機能、炎症マーカーなど）、呼吸機能検査および生理機能検査データ、画像読影所見（CT 画像、X 線画像に付随の読影所見）、病理所見から得られた結果を抽出する。さらに、外科生検とクライオ生検による得られた肺組織から DNA、RNA の同時抽出を行い、オミックス解析を行う。

(倫理面への配慮)

課題名「新薬創出を加速する人工知能の開発」について、当院の倫理委員会の承認を得て研究開始。同意説明文書を用いて医師が患者へ説明を行い、同意を取得した患者に研究協力をして頂いている。また、同意撤回の機会を保障している。

C. 研究結果

今年度実績：

同意取得 376 例（うち同意撤回 4 例）

組織採取：血液 429 例、クライオ生検 74 例、VATS 32 例、手術肺 1 例

コロナ禍の影響により組織採取例が少なくなる傾向だが、順調に症例数が増加している。

研究開始後の全実績：

同意取得 874 例（うち同意撤回 5 例）

組織採取：血液 855 例、クライオ生検 160 例、VATS 54 例、手術肺 1 例

これまでに収集した症例の MDD 診断結果によると、IPF が全体の 9%、UCIP が 22%、cHP が 12%、NSIP が 9% であった。IPF が少ないことが課題であるが、分類不能型 (UCIP) の中に IPF が含まれていることと未診断の症例が多数 (37%) あることから、今後 IPF の比率は増えてくると想定している。

収集・分析を行う臨床情報の質的向上のために、外部医師による定期的なコンサルを受けた。また、検査技師を増員し、新型コロナウイルス感染症により遅延していた、対象患者の採取血液・画像検査及び病理の解析データの取得・解析の遅延を補った。

D. 考察

上でも述べたが、MDD 診断結果によると、IPF が全体の 9%、82 例であった。その他、分類不能型間質性肺炎(UCIP)が 22%、慢性過敏性肺臓炎(cHP)が 12%、非特異性間質性肺炎(NSIP)が 9%であったが、IPF の比率が少ないことが課題であった。典型的な IPF は組織採取の対象とならないことから、血液についても解析の対象とすることとし、医薬基盤研と協力して今年度は 196 例の末梢血単核球からの DNA 抽出と全ゲノム解析、及び miRNA の解析を行った。

まだ、未診断が 319 例ほどあり、IPF の症例は増えると予想されるが、今後の症例収集の方向性としてはより多くの情報が得られる組織採取可能な症例の収集に努めるとともに、IPF の比率を高めるべく努力を続ける必要がある。

E. 結論

コロナ禍により症例の収集が遅延したものの、これまでに 874 例の同意取得まで完了しており、来年度には目標としている 1000 例の症例収集が完了する見込みである。医薬基盤研との協力によりこれらの症例の臨床方法を格納した新薬創出に資する世界に類をみない疾患統合データベースの構築を引き続き推進する予定である。

G. 研究発表

1. 論文発表

現時点ではなし

2. 学会発表

現時点ではなし

F. 知的財産権の出願・登録状況（予定を含む。）

1. 特許取得

なし

2. 実用新案登録

なし

3. その他

なし

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))
分担研究報告書

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に関する研究
研究分担者名 : 高村 大也 国立研究開発法人産業技術総合研究所
人工知能研究センター・研究チーム長

研究要旨

基礎生物学系文献からの分子、病態の知識抽出については、言語リソースの構築を行い、エンティティ、関係及びイベント抽出モデルの開発および学習と評価を行った。また、論文から情報を抽出するキュレーション技術の開発を行った。

さらに、基礎生物学情報と抽出情報との融合を目指し、当研究所で開発された PoSSuM 等の蛋白質に関するデータベースを利用し、検索・推論基盤の構築に向けた準備を実施した。

A. 研究目的

本研究プロジェクトは、論文などの技術文献から、生命系のエンティティ(蛋白質、疾患など)やイベント、およびそれらの間の関係を、自動的に抽出する技術を開発することを目的とする。さらに、それを用いて実際に知識抽出を行い、抽出した知識を、エンティティやイベントをノードとし、関係を持つノード同士を結合した、巨大な高次ネットワーク構造(疾患ネットワーク)で表す。また、既存の蛋白質データベースの情報などを合わせて、蛋白質の配列比較や化合物の構造比較に関するバイオインフォマティクス/ケモインフォマティクスの技術と自然言語処理の技術を統合して、蛋白質などのエンティティ間の類似度も測定可能にする。疾患ネットワークと、エンティティ間の類似度を利用することで、柔軟かつ広範囲にエンティティ間の相互作用を推論する技術を開発する。

より具体的には、次の 3 つの課題を解く。

- 1 基礎生物学系文献からの分子、病態の知識抽出
- 2 基礎生物学情報と抽出情報との融合
- 3 知識の表現方法および柔軟な推論技術

B. 研究方法**① 基礎生物学系文献に関する言語リソース構築 :**

論文などの技術文献から、生命系のエンティティ(蛋白質、疾患など)やイベント、およびそれらの間の関係を、自動的に抽出する技術を開発する。そのために、アノテーションガイドラインの作成や実際のアノテーションを含むコーパス開発を行う。

エンティティのアノテーション ; 特発性肺線維症(IPF)を中心とする肺疾患に関する文献要旨データ 120 件に対して、Disorder (疾患名)、GGPs (遺伝子・遺伝子産物)、Cell (細胞)、Anatomical entity (組織・器官など解剖学上のエンティティ)、Pharmacological substance (医薬品) など主要なエンティティのアノテーションとノーマリゼーションを済ませているが、更に追加の 30 件のデータに対して、エンティティとノーマリゼーションを行う。

関係性、イベントのアノテーション ; 医薬基盤研の肺疾患の専門家と議論・検討しながら、アノテーションのガイドラインとスキーム作成を進める。作成されたガイドライン、スキームに従って、エンティティ、ノーマリゼーションのアノテーションが完了した文献データに対して、アノテーションできる関係性やイベントのアノテーションを進めていく。

② 基礎生物学系文献からの分子レベルの知識抽出：

上の基礎生物学系文献に関する言語リソース構築で開発したアノテーションコーパスを基に、機械学習手法(特にニューラルネットワークに基づく手法)を開発し、実際に学習させることでエンティティ抽出技術の開発を行う。また、アノテーションコーパスにより、精度評価を行う。同様に、エンティティリンキング(ノーマライゼーション)の技術、またイベント抽出技術を開発する。

③ 知識の表現方法及び柔軟な推論技術の開発：

抽出されたエンティティや関係で構築される疾患ネットワークにおいて、そこに含まれる知識の表現技術や、ネットワーク上の柔軟な推論技術を開発する。ネットワーク上の推論技術として、知識グラフの埋め込み表現などを通じたイベント列を抽出する方法を開発する。特に、不確かな情報も有効に利用できるようなロバストなネットワーク上の推論手法の開発を進める。

④ 基礎生物学情報と抽出情報との融合：

新規標的候補等の同定に向けた文献情報と生物学情報との融合を目指し、当研究所で開発された PoSSuM 等の蛋白質に関するデータベースを利用し、知識抽出および推論技術により構築された疾患ネットワークの検証および拡充(新規情報の追加等)を可能とする検索・推論基盤の構築に資する。近年、蛋白質立体構造決定技術の向上により利用可能な構造情報が急増している。これらの全ての情報を利用するため新規サーバーを導入し、昨年度より PoSSuM データベースの更新作業を行っている。本年度は PoSSuMds データベースの更新にも着手する。これと同時に、検索・推論基盤構築に向けたケーススタディとして、肺線維症などの発症や治療に関わる免疫システムの論文調査を行い、抽出データを基に検索・推論基盤の構築を目指す。

(倫理面への配慮)

人及び動物を研究対象としていない研究であるため、倫理面の問題はないと判断した。

C. 研究結果

基礎生物学系文献に関する言語リソース構築のため、アノテーションスキームの設計、ノーマライゼーションのためのデータベースの調査、アノテーションガイドラインの作成、および実際のマニュアルアノテーションを行った。

また、上で開発された言語リソースを用い、エンティティ抽出モデル、エンティティリンキングモデル(ノーマライゼーションモデル)、関係抽出モデル、イベント抽出モデルの開発および学習と評価を行った。また、上記モデルで抽出した情報から、疾患ネットワークの構築し視覚化するデモシステムを構築した。また、論文から情報を抽出するキュレーション技術の開発を行った。

さらに、当研究所で開発された蛋白質の基質結合(候補)部位に関するデータベース PoSSuM を例として、検索・推論基盤の構築に向けた準備(PoSSuM の更新)を実施した。PDB データベースに 2020 年 9 月時点で登録されていた約 16 万件の蛋白質の立体構造およびアノテーションデータに対する、基質結合(候補)部位の類似性検索等の計算およびウェブサイトの更新作業を完了した。また、本プロジェクトで対象疾患の発症や治療に関わると予測された蛋白質の学術論文調査について、調査結果の再抽出および抽出結果の再フォーマットを実施した。

D. 考察

言語リソース構築については、アノテーションの質を測るためにアノテータ間一致度の算出が必要である。また、情報抽出については、実用性を高めるための技術開発が次の段階になる。また、PoSSuM と TargetMine の接続について今後さらなる検討を要すると考えられる。

E. 結論

コーパス構築、情報抽出技術の開発、検索・推論基盤の構築のそれぞれについて、予定していた開発目標を達成した。

G. 研究発表

1. 論文発表

なし

2. 学会発表

- 1) Mohammad Golam Sohrab (産総研), Anh-Khoa Duong Nguyen (産総研), Makoto Miwa (産総研, 豊田工業大学), and Hiroya Takamura (産総研), “Neural Exhaustive Approach for Entity and Relation Recognition Over Wet Lab Protocols”, The 6th workshop on Noisy User-generated Text (W-UT), 2020.
- 2) Mohammad Golam Sohrab (産総研), Anh-Khoa Duong Nguyen (産総研), Makoto Miwa (産総研, 豊田工業大学), Goran Topic (産総研), Masami Ikeda (産総研), and Hiroya Takamura (産総研), “BENNERD: A Neural Named Entity Linking System for COVID-19”, The 28th International Conference on Empirical Methods in Natural Language Processing (EMNLP), System Demonstration. 2020.

F. 知的財産権の出願・登録状況（予定を含む。）

1. 特許取得
なし
2. 実用新案登録
なし
3. その他
なし

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））
分担研究報告書

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

研究分担者名 : 西村 紳一郎 北海道大学・教授

研究要旨

本研究チームは「血中エクソソームの由来臓器を糖鎖解析により推定する基盤技術」という研究テーマを分担している。令和2年度は計画していた正常 ddY マウス主要臓器（脳・肝臓・肺・骨格筋）と大腿骨、腎臓、膵臓、脾臓、心臓、胃、腸（大腸および小腸）、皮膚、甲状腺、卵巣、子宮、睾丸および血清（エクソソームを含む）の定量的糖鎖プロファイリングを完了した。その結果、特定の組織・臓器や細胞から放出される定量的紐つきのエクソソーム糖鎖プロファイルデータベース構築の重要性が改めて明確に示された。世界初の「マウス組織・臓器糖鎖アトラス」の完成と発信に向けて論文を作成中である。

A. 研究目的

エクソソームが積載する様々な生体分子はいずれもそのエクソソームを放出した細胞内で合成されているため、エクソソームはそれらの細胞が分布する組織や臓器での異常を示唆する有効なバイオマーカーとなる可能性が高い。しかし、血液中のエクソソームは明らかに複雑なナノ微粒子の混合物であり、それらの由来組織・臓器を特定することはもとより推定すること自体困難である。我々は由来する組織や臓器と定量的に紐づいたエクソソームの情報を抽出できれば特定の組織・臓器での疾患特異的な変化を検知できると考えた。

本研究チームでは「血中エクソソームの由来臓器・組織を糖鎖解析により推定する基盤技術」を確立し、由来臓器・組織と定量的に紐づいたエクソソームの糖鎖構造情報を用いる新たな AI 支援型診断技術への応用を目的とした。

B. 研究方法

図1の基本プロトコルに基づいて主要臓器と組織の糖鎖構造解析を行ってそれらの詳細な構造パターンを定量的にプロファイリングした。正常 ddY マウスを解剖して摘出した主要臓器（脳・肝臓・肺・骨格筋）と大腿骨、腎臓、膵臓、脾臓、心臓、胃、腸（大腸および小腸）、皮膚、甲状腺、卵巣、子宮、睾丸および血清（エクソソームを含む）の定量的糖鎖プロファイルの比較を行った。血清以外の臓器についてはマウス5検体の糖鎖発現量の平均と標準誤差を求め、血清や他の臓器との比較を行った。

マウス主要臓器・組織の糖タンパク質の糖鎖プロファイルを獲得するための標準プロトコルを確立

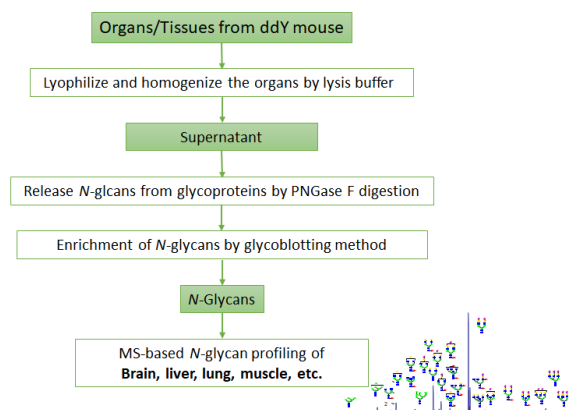


図1 マウス組織・臓器糖鎖解析基本プロトコル

（倫理面への配慮）

なお動物実験に際しては、動物愛護管理法および関係基準等を遵守し、動物福祉の基本概念である3R: Replacement, Reduction, Refinementの原則および動物実験倫理に基づいて適正に実施した。

C. 研究結果

その結果、各臓器・組織間で糖鎖プロファイルにおける大きな差異が認められた。一例として図 2 には主要 4 臓器（脳・肝臓・肺・骨格筋）と 8 臓器（脳・肝臓・肺・筋肉、骨、腎臓、膵臓、脾臓）の糖鎖プロファイルと血清の糖鎖プロファイルと比較した結果を示した。また、血清糖鎖プロファイルは糖鎖ごとに比較した場合に特定の臓器・組織の発現プロファイルとの相関がみられるものも確認できるが、基本的に血清糖鎖プロファイルとの差異が大きく組織・臓器のみで発現しているユニークな構造の糖鎖も存在することが初めて明らかとなった。例えば非常に興味あることに、脳と肺においては、他の臓器あるいは培養細胞などには見られない特徴的な糖鎖が数種類存在していた。

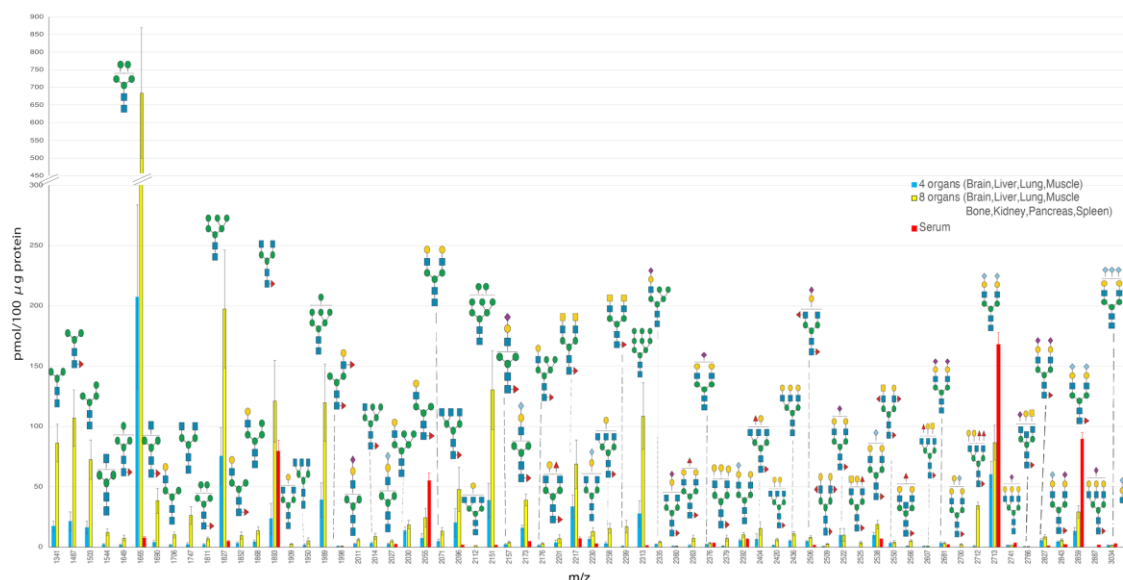


図 2 ddY マウス脳・肝臓・肺・筋肉（青）、脳・肝臓・肺・筋肉、骨、腎臓、膵臓、脾臓（黄）、およびエクソソームを含む血清（赤）の全糖鎖の比較（n=5）。各糖鎖構造の発現量の平均と標準偏差を示している。

D. 考察

今回の結果から糖鎖（構造）プロファイルはそれぞれの臓器（細胞）および組織間で大きく異なっていることが改めて証明できた。一方、血液から抽出・精製されるエクソソームの糖鎖プロファイルは由来する臓器・組織の糖鎖プロファイルの総和として観測されることを意味している。その際に各臓器・組織に由来するエクソソームの重みづけは個々の臓器や組織の定量的な糖鎖プロファイルとの相関（紐づけ）から予測可能であると考えられる。マウス血液中のエクソソーム糖鎖プロファイルと今回取得した個々の臓器・組織糖鎖プロファイルの比較からエクソソームの由来臓器・組織の割合を推定することができる。言い換えると、何らかの疾患により血液中のエクソソーム糖鎖プロファイルに変化が見られた場合に、それがどの臓器・組織の異常・変調によるのかAI などにより容易に検出できることを示唆している（論文投稿準備中）。

E. 結論

特定の組織・臓器や細胞から放出されるエクソソームの血中濃度は主要な臓器や組織の糖鎖プロファイルデータベースとの相関を解析することによって推定できる。すなわち、由来する組織や臓器と定量的に紐づいたエクソソームの情報を抽出できれば特定の組織・臓器での疾患特異的な変化を検出できるという仮説が支持された。正常マウスの臓器・組織糖鎖アトラスを健常ヒト臓器・組織糖鎖アトラスに拡張すれば患者血清のエクソソームの糖鎖プロファイルからの新しい診断技術の研究開発に展開できる可能性が高い。

G. 研究発表

1. 論文発表

- 1) Takahashi H., Kamiyama T., Hirane N., Kobayashi N., Aiyama T., Nagatsu A., Shimada S., Orimo T., Kakisaka T., Fukai M., Yokoo H., Kamachi H., Nishimura S-I., Taketomi A., “Analysis of the correlation between alterations in *N*-glycans and invasiveness in liver cancer cell lines” *Oncol. Rep.* 44, 2757-2769 (2020)
- 2) Pablo A. Guillen-Poza, Elena Matilde Sánchez-Fernández, Gerard Artigas, José Manuel García Fernández, Hinou H., Carmen Ortiz Mellet, Nishimura S-I., Fayna Garcia-Martin, “Amplified detection of breast cancer autoantibodies using MUC1-based Tn antigen mimics” *J. Med. Chem.* 63, 8524-8533 (2020)
- 3) Yokoi Y., Nishimura S-I., “Effect of Site - specific *O*-Glycosylation on the Structural Behavior of NOTCH1 Receptor Extracellular EGF - like Domains 11 and 10” *Chem. Eur. J.* 26, 12363-12372 (2020)
- 4) Wakui H., Tanaka Y., Ose T., Matsumoto I., Kato K., Yao M., Tachibana T., Sato M., Naruchi K., García-Martín F., Hinou H., Nishimura S-I. “A straightforward approach to antibodies recognising cancer specific glycopeptidic neoepitopes” *Chem. Sci* 11, 4999-5006 (2020)

2. 学会発表

なし

F. 知的財産権の出願・登録状況（予定を含む。）

4. 特許取得

なし

5. 実用新案登録

なし

6. その他

なし

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))
分担研究報告書

課題名 : 肺がん統合データベース構築及び AI 技術を用いたオミックス解析
研究分担者名 : 浜本 隆二 国立がん研究センター研究所 医療 AI 研究開発分野・分野長

研究要旨

次世代シーケンサーや遺伝子編集技術に象徴されるバイオテクノロジーの加速的な進化の一方で、医薬品開発を巡る近年の研究開発の生産性は急激に低下し、「イノベーションの欠乏」が叫ばれて久しい。10 年前後の期間にわたって開発される新薬にかかる費用はますます増大し、一つの薬を上市するために必要な開発コストは 1000 億円以上に上る。また、候補化合物から医薬品として世に出されるものは 2 万から 3 万に一つであり、その成功確率は極めて低い。特に、新薬候補物質の効果を始めて患者で示す”Phase II 試験“での開発中止 (Phase II attrition) が顕著であり、開発した薬剤の多くがマウスでは効くものの、人間ではその有効性が証明されないという、マウス実験を基盤とする現状の医薬品開発パイプラインの構造的な課題がある。本研究課題申請者が対象疾患とする肺がんにおいても、患者の遺伝子を調べることで最適な治療薬を選択する「がんゲノム医療」が本邦でも開始されたものの、遺伝子の異常が検出されながらも適合する薬剤が存在しないケースが多数出現することが判明しており、社会問題となっている。以上の事から、新薬開発プロセスにおける構造的な課題を克服し、疾患治療で現状打破を得るためのイノベーションが現在期待されている。そのための鍵となるものが、ヒトの薬剤投与時や疾患罹患時における網羅的分子プロファイルに現れる生体変化の統合的理解と、実際の医療から得られるリアルワールド・データの解析に基づいた薬剤のヒトにおける有効性や毒性予測、更にこうした解析を可能にする最新の計算機アルゴリズムであると考えられる。

近年、「50 年来の技術的ブレークスルー」とも言われる深層学習と最新の機械学習手法の組み合わせが、生命情報処理に対するアルゴリズム革命をもたらすものと期待されている。特に、生物の脳神経系にヒントを得た情報処理メカニズムである人工ニューラルネットワークを多層化した深層学習は、社会や産業の形を変え得る画期的な情報技術として大きな注目を集めている。深層学習を既存の機械学習手法と比して特徴付けるものとしては、多種のデータを入力として取り扱えるマルチモーダル学習、複数の異なるタスクをモデルの一部として共有できるマルチタスク学習、少数の教師ありデータから汎化性能を得ることのできる半教師あり学習や教師無し学習、階層的な特徴を自動的に獲得できる表現学習などが挙げられる。近年のコンピュータ処理能力の飛躍的向上と加速的に蓄積されるビッグデータを基盤として、深層学習は画像や音声といった比較的高度な知的タスクで、既に人間を上回る性能を示すに至り、生命情報処理ならびに創薬の分野においてもその応用に関する世界的な研究開発競争が激化している。そこで、本邦における分子標的治療薬の創生及び生命情報処理に関わる第一線の研究者が一つのチームとして協力し合うことで、がんの生体時空間にわたるシステムの統合理解に基づいた創薬候補因子の探索と、がん治療におけるあるべき不均一性を包含した、臨床のリアルワールド・データに基づいたヒトにおける有効性や毒性の予測を可能にすると捉え、医薬品開発における「イノベーションの欠乏」を解決するための新しい技術開発に取り組んだ。その結果 AI を活用した複数の解析プラットフォームの開発に成功した。また AI 解析において最重要な事項の一つとして、質の高いデータベースの構築が挙げられるが、我々は臨床データを効率的に収集するプラットフォームの構築に成功した。その結果 AI 解析を志向した世界最大規模の肺がん統合データベースを、本研究課題で構築することに成功した。本データベースは質の高いデータが保存されていることから、本邦の財産と考えられる貴重な研究成果であると判断している。

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。

今年度は、i) データ収集：世界最大規模の肺がん統合データベースの拡充、ii) 肺がんデータプラットフォームの構築、iii) 解析アルゴリズムの開発を目標とする。

B. 研究方法

- ① 肺がんオミックス統合データベースの拡充：特にエピゲノム解析 (ChIP-seq 解析、DNA メチル化解析) を優先して解析を進める。
- ② 肺がんデータプラットフォームの構築：AI 解析を志向した効率的なデータ及び臨床情報収集システムを構築する。
- ③ 解析アルゴリズムの開発：AI を用いてマルチオミックスデータ解析システムを開発する。

(倫理面への配慮)

本研究課題は、国立がん研究センター・研究倫理審査委員会で承認をうけ (研究開発課題：2019-108、“新薬創出を加速する人工知能の開発”)、実施している。

C. 研究結果

1.

概要：AI の学術研究を行う上で最重要事項の一つとして、質の高いデータベースを構築することが挙げられる。

我々は本研究課題で肺がんのオミックスデータベース構築に取り組み、全エクソン解析及び診療情報においては、1500 例以上という米国の TCGA を超える、世界最大規模の肺がん統合データベースの構築に成功した。本データベースは本邦の財産ともいえる、貴重な研究成果であると判断している。また、新型コロナウイルス感染症の蔓延に伴い研究計画の一部変更が余儀なくされ、2021 年4月 1 日から9月 30 日まで研究費を繰越したが、その間に、当初令和2年度に計画していた全ゲノム解析及び ChIP-seq 解析を計画通り全て終了した。

2.

概要：AI 研究により優れた研究成果を出していく為には、質の高い構造化データを準備する事は非常に重要である。しかし、多くの医療データは不均質で構造化されておらず、価値創造の機会を逸しているのが現状である。我々は AI 技術と最先端の ICT 技術を組み合わせる事で、効率的にデータを構造化するプラットフォームの構築に成功した。本プラットフォームはメディカル AI 研究を発展させる上で、重要なイノベーションであると判断している。

3.

概要:がんは複雑な疾患である為、その本態解明を行い創薬に発展させる為には、臨床検体から得られた膨大なオミックスデータを、詳細な診療情報と共に効率的に解析する技術を開発することが必須であると考えている。そこで、我々は機械学習・深層学習技術を活用して、がんに関する医療ビッグデータを解析する手法の開発に取り組んだ。その結果肺がんに関する、3 種類の新規プラットフォームの開発に成功し、その成果を国際誌に論文発表した。

4.

概要:肺がん統合データベース構築においては、ゲノム解析のみならず、エピゲノム解析も施行し、特にヒストン修飾解析に関しては、FFPE サンプルからも効率的かつ精度の高い ChIP-seq 解析を行う事が可能な、新しい ChIP-seq 解析手法 (RCRA ChIP-seq 法)を開発した。FFPE サンプルを用いた ChIP-seq 解析は世界中で高い需要があり、本手法は科学技術イノベーションに大きく貢献する成果であると判断している。また、世界最大規模の肺がん統合データベースも、本邦にとって貴重な財産で、科学技術イノベーションに大きく寄与する可能性があるかと判断している。

5.

概要:上述のようにメディカル AI 研究を行う上で、質の高い構造化された医療データを蓄積していくプラットフォームを構築することは重要である。我々は AI を用いて、自動的に放射線画像にアノテーション付けを行うプラットフォームを開発し特許を取得後、既に大手医療機器メーカーに導出している。我々が開発した技術が社会実装される事により、AI 技術を活用した放射線画像解析に関して、研究及び臨床双方に大きく貢献すると考えられ、科学技術イノベーションに大きく寄与する成果であると判断している。

6.

概要:我々はがんに関する大規模オミックスデータを、効率的かつ高精度に解析する AI 技術を用いたプラットフォームを複数種類開発した。医療データは、サンプル数に比してパラメータ数が多いという新 NP 問題と呼ばれる課題など、AI 技術を有効に活用する上で解析しないといけないいくつかの問題がある。実際ゲノム・エピゲノム解析においては、医用画像解析と比して未だ AI 技術の導入は進んでいないという現実もある。一方、今回我々が開発した手法は、大規模ながんオミックスデータを診療情報と合わせて効率的かつ精度高く解析することが可能であり、科学技術イノベーションに大きく寄与する成果であると判断している。

E. 研究発表

1. 論文発表

- 1) Ken Asada, Kazuma Kobayashi, Samuel Joutard, Masashi Tubaki, Satoshi Takahashi, Ken Takasawa, Masaaki Komatsu, Syuzo Kaneko, Jun Sese, Ryuji Hamamoto: Uncovering Prognosis-Related Genes and Pathways by Multi-Omics Analysis in Lung Cancer. *Biomolecules*, 10, 524 (2020)
- 2) Theodore Vougiouklakis, Riyue Bao, Sohyoung Kim, Songjoon Baek, Makda Zewde, Benjamin Bernard, Kyunghye Burkitt, Nupur Nigam, Evgeny Izumchenko, Naoshi Dohmae, Ryuji Hamamoto, Yusuke Nakamura: WHSC1 monomethylates histone H1 and induces stem-cell like features in squamous cell carcinoma of the head and neck. *Neoplasia*, 22, 283-293 (2020)
- 3) Ken Asada, Amina Bolatkan, Ken Takasawa, Masaaki Komatsu, Syuzo Kaneko, Ryuji Hamamoto: Critical Roles of N6-Methyladenosine (m6A) in Cancer and Virus Infection. *Biomolecules*, 10, 524 (2020)
- 4) Shunichi Jinnai, Naoya Yamazaki, Yuichiro Hirano, Yohei Sugawara, Yuichiro Ohe, Ryuji Hamamoto: The Development of a Skin Cancer Classification System for Pigmented Skin Lesions Using Deep Learning. *Biomolecules*, 10, 1123 (2020)
- 5) Kazuma Kobayashi, Amina Bolatkan, Shuichiro Shiina, Ryuji Hamamoto: Fully-Connected Neural Networks with Reduced Parameterization for Predicting Histological Types of Lung Cancer from Somatic Mutations. *Biomolecules*, 10, 1249 (2020)
- 6) Satoshi Takahashi, Ken Asada, Ken Takasawa, Ryo Shimoyama, Akira Sakai, Amina Bolatkan, Norio Shinkai, Kazuma Kobayashi, Masaaki Komatsu, Syuzo Kaneko, Jun Sese, Ryuji Hamamoto: Predicting Deep Learning Based Multi-Omics Parallel Integration Survival Subtypes in Lung Cancer Using Reverse Phase Protein Array Data. *Biomolecules*, 10, 1460 (2020)

- 7) Machiko Kojima, Kenbun Sone, Katsutoshi Oda, Ryuji Hamamoto, Syuzo Kaneko, Shinya Oki, Asako Kukita, Akira Kawata, Harunori Honjoh, Yoshiko Kawata, Tomoko Kashiya, Masakazu Sato, Ayumi Taguchi, Yuichiro Miyamoto, Michihiro Tanikawa, Tetsushi Tsuruga, Kazunori Nagasaka, Osamu Wada-Hiraike, Yutaka Osuga, Tomoyuki Fujii: The histone methyltransferase SMYD2 is a novel therapeutic target for the induction of apoptosis in ovarian clear cell carcinoma cells. *Oncology Letters*, 20, 153 (2020)
- 8) Ai Dozen, Masaaki Komatsu, Akira Sakai, Reina Komatsu, Kanto Shozu, Hidenori Machino, Suguru Yasutomi, Tatsuya Arakaki, Ken Asada, Syuzo Kaneko, Ryu Matsuoka, Daisuke Aoki, Akihiko Sekizawa, Ryuji Hamamoto: Image Segmentation of the Ventricular Septum in Fetal Cardiac Ultrasound Videos Based on Deep Learning Using Time-Series Information. *Biomolecules*, 10, 15426 (2020)
- 9) Vasiliki Papadaki, Ken Asada, Julie K Watson, Toshiya Tamura, Alex Leung, Jack Hopkins, Margaret Dellett, Noriaki Sasai, Hongorzul Davaapil, Serena Nik-Zainal, Rebecca Longbottom, Makoto Nakakido, Ryo Torii, Abhi Veerakumarasivam, Syuzo Kaneko, Mandeep S Sagoo, Gillian Murphy, Akihisa Mitani, Kohei Tsumoto, John D Kelly, Ryuji Hamamoto, Shin-Ichi Ohnuma: Two Secreted Proteoglycans, Activators of Urothelial Cell-Cell Adhesion, Negatively Contribute to Bladder Cancer Initiation and Progression. *Cancers*, 12, 3362 (2020)
- 10) Ryuji Hamamoto, Kruthi Suvarna, Masayoshi Yamada, Kazuma Kobayashi, Norio Shinkai, Mototaka Miyake, Masamichi Takahashi, Shunichi Jinnai, Ryo Shimoyama, Akira Sakai, Ken Takasawa, Amina Bolatkan, Kanto Shozu, Ai Dozen, Hidenori Machino, Satoshi Takahashi, Ken Asada, Masaaki Komatsu, Jun Sese, Syuzo Kaneko: Application of Artificial Intelligence Technology in Oncology: Towards the Establishment of Precision Medicine. *Cancers*, 12, 3532 (2020)
- 11) Miku Wada, Asako Kukita, Kenbun Sone, Ryuji Hamamoto, Syuzo Kaneko, Masaaki Komatsu, Yu Takahashi, Futaba Inoue, Machiko Kojima, Harunori Honjoh, Ayumi Taguchi, Tomoko Kashiya, Yuichiro Miyamoto, Michihiro Tanikawa, Tetsushi Tsuruga, Mayuyo Mori-Uchino, Osamu Wada-Hiraike, Yutaka Osuga, Tomoyuki Fujii: Epigenetic Modifier SETD8 as a Therapeutic Target for High-Grade Serous Ovarian Cancer. *Biomolecules*, 10, 1686 (2020)
- 12) Kanto Shozu, Masaaki Komatsu, Akira Sakai, Reina Komatsu, Ai Dozen, Hidenori Machino, Suguru Yasutomi, Tatsuya Arakaki, Ken Asada, Syuzo Kaneko, Ryu Matsuoka, Akitoshi Nakashima, Akihiko Sekizawa, Ryuji Hamamoto: Model-Agnostic Method for Thoracic Wall Segmentation in Fetal Ultrasound Videos. *Biomolecules*, 10, 1691 (2020)
- 13) Shun-Ichiro Kageyama, Junyan Du, Syuzo Kaneko, Ryuji Hamamoto, Shigeo Yamaguchi, Riu Yamashita, Masayuki Okumura, Atsushi Motegi, Hidehiro Hojo, Masaki Nakamura, Katsuya Tsuchihara, Tetsuo Akimoto: Identification of the mutation signature of the cancer genome caused by irradiation. *Radiotherapy and Oncology*, 155, 10-16 (2021)
- 14) Kazuma Kobayashi, Naoya Murakami, Kana Takahashi, Koji Inaba, Hiroshi Igaki, Ryuji Hamamoto, Jun Itami: Tensor regression-based model to investigate heterogeneous spatial radiosensitivity after I-125 seed implantation for prostate cancer. *In vivo*, 35, 489-497 (2021)
- 15) Masaaki Komatsu, Akira Sakai, Reina Komatsu, Ryu Matsuoka, Suguru Yasutomi, Kanto Shozu, Ai Dozen, Hidenori Machino, Hirokazu Hidaka, Tatsuya Arakaki, Ken Asada, Syuzo Kaneko, Akihiko Sekizawa, Ryuji Hamamoto: Detection of Cardiac Structural Abnormalities in Fetal Ultrasound Videos Using Deep Learning. *Applied Sciences*, 11, 371 (2021)
- 16) Tatsuya Ozawa, Syuzo Kaneko, Frank Szulzewsky, Zhiwei Qiao, Mutsumi Takadera, Yoshitaka Narita, Tadashi Kondo, Eric C Holland, Ryuji Hamamoto, Koichi Ichimurai: C11orf95-RELA fusion drives aberrant gene expression through the unique epigenetic regulation for ependymoma formation. *Acta Neuropathol Commun*, 9, 36 (2021)
- 17) Daisuke Takayanagi, Sou Hirose, Ikumi Kuno, Yuka Asami, Naoya Murakami, Maiko Matsuda, Yoko Shimada, Kuniko Sunami, Masaaki Komatsu, Ryuji Hamamoto, Mayumi Kobayashi Kato, Koji Matsumoto, Takashi Kohno, Tomoyasu Kato, Kouya Shiraishi, Hiroshi Yoshida: Comparative Analysis of Genetic Alterations, HPV-Status, and PD-L1 Expression in Neuroendocrine Carcinomas of the Cervix. *Cancers (Basel)*, 13, 1215 (2021)

- 18) Satoshi Takahashi, Masamichi Takahashi, Manabu Kinoshita, Mototaka Miyake, Risa Kawaguchi, Naoki Shinojima, Akitake Mukasa, Kuniaki Saito, Motoo Nagane, Ryohei Otani, Fumi Higuchi, Shota Tanaka, Nobuhiro Hata, Kaoru Tamura, Kensuke Tateishi, Ryo Nishikawa, Hideyuki Arita, Masahiro Nonaka, Takehiro Uda, Junya Fukai, Yoshiko Okita, Naohiro Tsuyuguch, Yonehiro Kanemura, Kazuma Kobayashi, Jun Sese, Koichi Ichimura, Yoshitaka Narita, Ryuji Hamamoto: Fine-Tuning Approach for Segmentation of Gliomas in Brain Magnetic Resonance Images with a Machine Learning Method to Normalize Image Differences among Facilities. *Cancers (Basel)*, 13, 1415 (2021)
- 19) Takahashi S, Takahashi M, Tanaka S, Takayanagi S, Takami H, Yamazawa E, Nambu S, Miyake M, Satomi K, Ichimura K, Narita Y, Hamamoto R. A New Era of Neuro-Oncology Research Pioneered by Multi-Omics Analysis and Machine Learning. *Biomolecules*, 11, 565 (2021)
- 20) Kaneko S, Mitsuyama T, Shiraishi K, Ikawa N, Shozu K, Dozen A, Machino H, Asada K, Komatsu M, Kukita A, Sone K, Yoshida H, Motoi N, Hayami S, Yoneoka Y, Kato T, Kohno T, Natsume T, Keudell GV, Saloura V, Yamaue H, Hamamoto R. Genome-Wide Chromatin Analysis of FFPE Tissues Using a Dual-Arm Robot with Clinical Potential. *Cancers (Basel)*, 13, 2126 (2021)
- 21) Asada K, Kaneko S, Takasawa K, Machino H, Takahashi S, Shinkai N, Shimoyama R, Komatsu M, Hamamoto R. Integrated Analysis of Whole Genome and Epigenome Data Using Machine Learning Technology: Toward the Establishment of Precision Oncology. *Front Oncol*, 12;11:666397 (2021)
- 22) Kobayashi K, Miyake M, Takahashi M, Hamamoto R. Observing Deep Radiomics for the Classification of Glioma Grades. *Sci Rep*, 11, 10942 (2021)
- 23) Lee K, Kim K, Ryu TY, Jung CR, Lee MS, Lim JH, Park K, Kim DS, Son MY, Hamamoto R, Cho HS. Dysregulation of EHMT1 affects lung cancer proliferation by controlling CDKN1A expression. *Mol Oncol*, 15, 2989-3002 (2021)
- 24) Komatsu M, Sakai A, Dozen A, Shozu A, Yasutomi S, Machino H, Asada K, Kaneko S, Hamamoto R. Towards Clinical Application of Artificial Intelligence in Ultrasound Imaging. *Biomedicines*, 9, 720 (2021)
- 25) Yamada M, Saito Y, Yamada S, Kondo H, Hamamoto R. Detection of flat colorectal neoplasia by artificial intelligence: A systematic review. *Best Pract Res Clin Gastroenterol*, 52-53, 101745 (2021)
- 26) Kawaguchi RK, Takahashi M, Miyake M, Kinoshita M, Takahashi S, Ichimura K, Hamamoto R, Narita Y, Sese J. Assessing Versatile Machine Learning Models for Glioma Radiogenomic Studies across Hospitals. *Cancers (Basel)*, 13, 3611 (2021)
- 27) Kobayashi K, Hataya R, Kurose Y, Miyake M, Takahashi M, Nakagawa A, Harada T, Hamamoto R. Decomposing normal and abnormal features of medical images for content-based image retrieval of glioma imaging. *Med Image Anal*, 74, 102227 (2021)
- 28) Kaneko S, Takasawa K, Asada K, Shinkai N, Bolatkan A, Yamada M, Takahashi S, Machino H, Kobayashi K, Komatsu M, Hamamoto R. Epigenetic Mechanisms Underlying COVID-19 Pathogenesis. *Biomedicines*, 9, 1142 (2021)
- 29) Asada K, Komatsu M, Shimoyama R, Takasawa K, Shinkai N, Sakai A, Bolatkan A, Yamada M, Takahashi S, Machino H, Kobayashi K, Kaneko S, Hamamoto R. Application of Artificial Intelligence in COVID-19 Diagnosis and Therapeutics. *J Pers Med*, 11, 886 (2021)
- 30) Kuno I, Takayanagi D, Asami Y, Murakami N, Matsuda M, Shimada Y, Hirose S, Kato MK, Komatsu M, Hamamoto R, Okuma K, Kohno T, Itami J, Yoshida H, Shiraishi K, Kato T. TP53 mutants and non-HPV16/18 genotypes are poor prognostic factors for concurrent chemoradiotherapy in locally advanced cervical cancer. *Sci Rep*, 11, 19261 (2021)

2. 学会発表

- 1) 浜本 隆二、PRISM 肺がんプロジェクトの進捗と今後の戦略、第 9 回生命医薬情報学連合大会 (IIBMP2020) 企画セッション【PRISM 創薬 AI】: オンライン、2020 年 9 月
- 2) 浜本 隆二、臨床応用を志向した医療 AI 研究の現状と課題、第 9 回生命医薬情報学連合大会 (IIBMP2020) 企画セッション【医療 AI】: オンライン、2020 年 9 月
- 3) 浜本 隆二、人工知能 (AI) 技術の医療応用への取組-内閣府/PRISM プロジェクトを中心として-、第 37 回日本呼吸器外科学会学術集会・特別企画 1: オンライン、2020 年 9 月

- 4) 浜本 隆二、Medical AI Research for Clinical Application、第 79 回日本癌学会学術総会・モーニングレクチャー：オンライン、2020 年 10 月
- 5) 浜本 隆二、ビッグデータ時代のがん研究：AI の利活用とその課題、第 79 回日本癌学会学術総会・SSP プログラム：オンライン、2020 年 10 月
- 6) 浜本 隆二、Precision Medicine 実現に向けたビッグデータ解析時代のがん研究、第 24 回日本がん分子標的治療学会学術集会：オンライン、2020 年 10 月
- 7) 浜本 隆二、臨床応用を志向した AI 研究：Precision Medicine 時代のビッグデータ解析、第 28 回日本乳癌学会学術総会・特別企画：オンライン、2020 年 10 月
- 8) 浜本 隆二、Precision Medicine と AI、第 58 回日本癌治療学会学術集会・会長企画シンポジウム：オンライン、2020 年 10 月
- 9) 浜本 隆二、AI and Healthcare in Japan: Overview、2nd Annual Conference of Korea Society of Artificial Intelligence in Medicine：オンライン、2020 年 10 月
- 10) 浜本 隆二、Medical AI Research in the Era of Precision Medicine、Seoul Bio-Economy Forum 2020：オンライン、2020 年 11 月
- 11) 浜本 隆二、臨床応用を志向した医療 AI 研究：その可能性と課題、日本オミックス医学会シンポジウム：オンライン、2020 年 11 月
- 12) 浜本 隆二、Medical AI Research for Clinical Applications、2nd French-German-Japanese Symposium：オンライン、2020 年 11 月
- 13) 浜本 隆二、AI の臨床診断への応用：その可能性と課題、第 5 回 Liquid Biopsy 研究会：オンライン、2021 年 1 月
- 14) 浜本 隆二、AI 技術を用いた医療機器の臨床応用、第 20 回日本再生医療学会総会：オンライン、2021 年 3 月

F. 知的財産権の出願・登録状況（予定を含む。）

1. 特許取得
「スケッチによる類似画像検索アルゴリズム」（予定）
2. 実用新案登録
該当無し
3. その他
該当無し

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))
分担研究報告書

課題名 : 肺がん患者のシングルセル・マルチオミックス解析

研究分担者名 : 西岡 安彦 徳島大学大学院医歯薬学研究部 呼吸器・膠原病内科学分野

研究要旨

がん細胞とその周囲の間質細胞にはサイトカインや増殖因子のパラクライン作用を介した相互作用が存在し、がんの治療抵抗性やがん転移の頻度等にも関与している。そのため肺がん組織の間質細胞のシングルセル解析データを基に間質細胞のプロファイル解析や遺伝子発現解析を行うことで肺がん治療開発のための有用なデータとなり得る可能性がある。本研究では間質性肺炎合併/非合併肺がん患者の肺がん手術検体から採取した間質細胞に対してscRNA-seqおよびscATAC-seqによりシングルセル解析のデータを回収すると同時に、肺組織のmRNA発現解析や患者の血清サンプルのエクソソーム解析を行うことで、間質性肺炎合併肺がん組織に特異的な間質細胞プロファイルやがんの進行に関わる生体分子を見出し、創薬ターゲット候補を特定するAI開発および、その元となるデータベース作製につなげることを目的とした。外科手術を受けた12症例（間質性肺炎合併例4例、非合併例8例）の肺がん組織を用いてscRNA-seqを行い、正常肺組織、肺がん組織および間質性肺炎組織を比較することで、それぞれに特徴的な線維芽細胞のクラスターパターンを同定した。肺がんおよび間質性肺炎に特徴的な線維芽細胞クラスターの遺伝子発現解析から間質細胞を標的とした創薬ターゲット分子の同定を進めている。

A. 研究目的

創薬ターゲット候補を特定するAI開発につなげることを目的として、間質性肺炎合併/非合併肺がん患者の手術検体より間質細胞のシングルセル解析を行い、これら疾患の臨床的・生物学的データの収集を行う。

シングルセル解析においては過去に特発性肺線維症の肺の解析 (Am J Respir Crit Care Med. 2019;199:1517-36) や肺がん以外のがん間質細胞の解析 (Cancer Discov. 2019;9:1102-23) について報告があり、それぞれで新規細胞クラスターや特異的な遺伝子発現を指摘しており、その疾患の成り立ちや微小環境の解明につながっている。一方で肺がんにおいてはシングルセル解析の報告はまだなされていない。がん細胞とその周囲の間質細胞にはサイトカインや増殖因子のパラクライン作用を介した相互作用が存在し、がんの治療抵抗性やがん転移の頻度等にも関与している。本研究により肺がん組織の間質細胞のシングルセル解析データを基に間質細胞のプロファイル解析や遺伝子発現解析を行うことで肺がん治療開発のための有用なデータとなり得る可能性がある。そのため、本研究では間質性肺炎合併/非合併肺がん患者の肺がん手術検体より間質細胞を収集し、がん組織と非がん組織で比較するためのシングルセル解析のデータを回収して、肺がん組織に特異的な間質細胞プロファイルやがんの進行に関わる生体分子を見出し、創薬ターゲットの探索に発展させることを目標とする。本研究では多様な間質細胞の中でも線維芽細胞に着目し検討を進める。また、同一患者の血清サンプルのエクソソームなどの解析や組織からのRNA抽出も行い、多角的にデータベースを構築していく。

B. 研究方法

当研究では間質性肺炎合併および非合併肺がん患者の肺がん切除検体を用いてシングルセル解析を主に実施し、癌周囲の間質細胞の遺伝子発現を網羅的に検討して間質性肺炎合併肺がんに特徴的な間質細胞プロファイルを明らかにするとともに、癌の進展に影響を与える可能性があり治療標的となり得る生体分子の同定を目指したデータベース構築を行った。具体的には、①肺がん切除検体を用いたシングルセル解析 (scRNA-Seq; single-cell RNA-sequencing, scATAC-Seq; single-cell assay for transposase accessible chromatin with high-throughput-sequencing)、②肺がん切除検体の組織の mRNA 発現解析、③同一症例の血液サンプル解析、④臨床情報収集を下記のように実施した。

- ① シングルセル解析は間質性肺炎合併/非合併肺がん切除検体の肺がん部位、健常肺部位、間質性肺炎部位より、それぞれの間質細胞を Epcam⁺CD45⁻CD31⁻CD235⁻細胞として cell sorter を用いて採取し、10x genomics 社の Chromium Controller を用いて cDNA ライブラリーを作成する。調整した cDNA ライブラ

リーは次世代シーケンサーでシーケンスした後、Cell ranger を用いて Fastq ファイルや遺伝子発現マトリックスを作成し、Seurat もしくは Loupe Browser といった pipeline を用いて細胞クラスタリングや遺伝子発現状況について解析した。

- ② 組織の mRNA 発現解析は、シングルセル解析と同様に肺切除検体の肺がん部位、健常肺部位、間質性肺炎部位より mRNA を抽出し、遺伝子発現を網羅的に解析した。
- ③ 血液サンプルはシングルセル解析の対象となった肺がん患者より採血を行い、エクソソーム解析用に基盤・健康・栄養研究所に送付した。
- ④ 臨床情報収集はシングルセル解析を行った症例について患者背景（年齢、性別、既往歴、常用薬など）、身体所見、検査データ（血液、尿、呼吸機能、生理機能検査など）、胸部画像所見、病理組織所見を収集した。

（倫理面への配慮）

本検討では前例において全例インフォームドコンセントを行い、研究内容を理解しご同意いただいた被験者のみが対象となる。日常診療以上の侵襲を被験者に与えないが、プライバシーの保護に留意し試料および情報は対応表を作成した上で匿名化し、鍵のかかる保管場所に保管される。対応表は徳島大学のみで保管され、データの供与先には提供されない。本検討は人を対象とする医学系研究に関する倫理指針を遵守し徳島大学病院医学系研究倫理審査委員会の承認を得て実施した。

C. 研究結果

12 症例（間質性肺炎合併例 4 例、非合併例 8 例）にて scRNA-seq 解析を実施し、正常肺部分 3 件、間質性肺炎部分 4 件、間質性肺炎合併肺がん部分 4 件、間質性肺炎非合併肺がん部分 5 件の sc RNA-seq のデータが得られ 29472 個の間質細胞のデータが得られた（図 1）。

間質細胞は図 2 の如く A~H の 8 つのクラスターに分けられた（図 2A）。これらの内、クラスター A とクラスター B は正常肺組織で見られ、他のクラスターは肺線維症の部位や肺がんなど病的肺組織に見られた（図 2B）。病的肺組織で見られたクラスター C~F では間質性肺炎で報告されている *Cthrc1* 高発現の Collagen1 産生能の高い間質細胞と類似していた（Nat Commun 2020;11:1920）。間質性肺炎合併肺がんにて特異的な間質細胞クラスターは指摘できず、その間質細胞の分画は線維症と肺がんの間であり、正常組織で見られる間質細胞も多く含むことが特徴であった。一方で同一の間質性肺炎合併肺がん症例で線維症部位と肺がん部位で比較するとクラスター E とクラスター F の割合ががん部位で増加しており、これらのクラスターが特に肺がん組織に特徴的な間質細胞クラスターと考えられた（図 2C）。クラスター E と F では高発現の遺伝子プロファイルが類似しており、これらを統合して高発現遺伝子の GO 解析を行ったところ ECM organization、Vasculature development、Cell adhesion に関わる遺伝子発現レベルが高くがん組織における微小環境調整に関わっている可能性が示唆された（図 2D）。

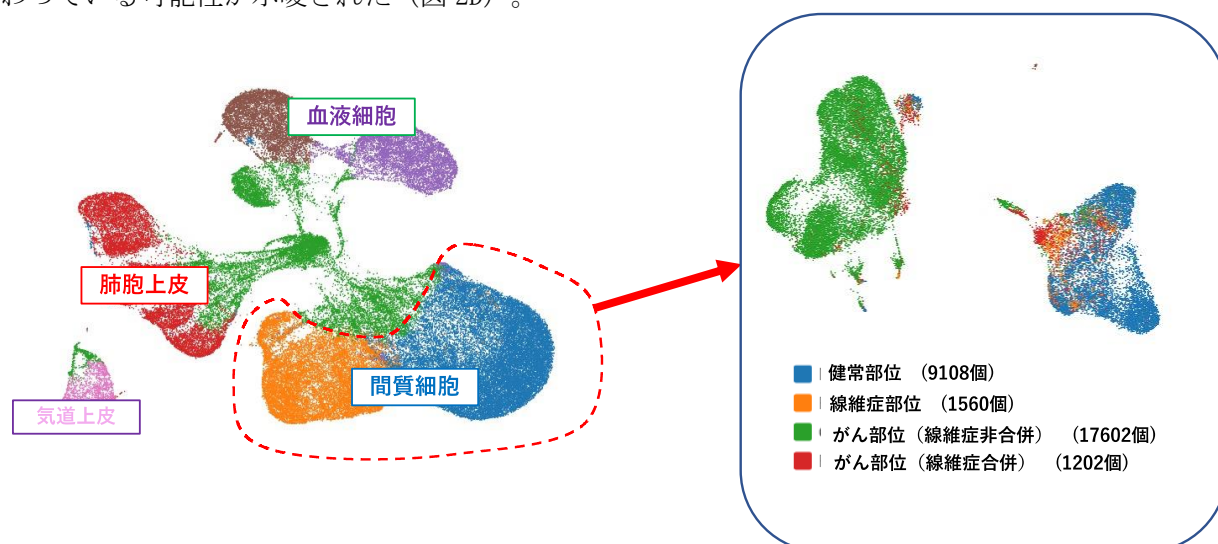


図 1. 肺がん手術検体の間質細胞の sc RNA-seq (UMAP)

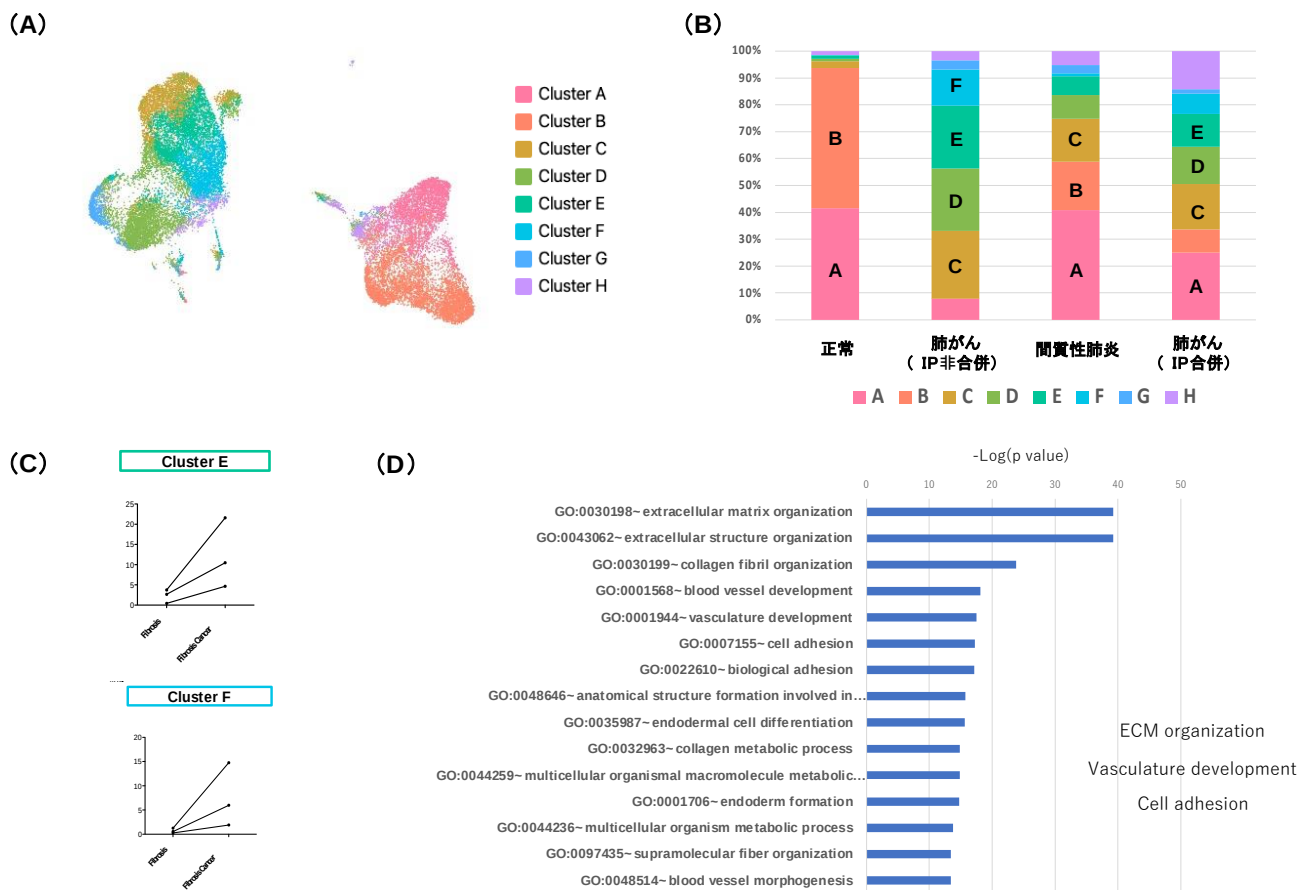


図2. 間質細胞のクラスタリング

(A) 肺がん手術検体より得られた sc-RNA seq データから抽出された間質細胞のクラスタリング (UMAP)。(B) 各採取組織における間質細胞のクラスターの割合。(C) 同一症例における間質性肺炎合併肺がんの線維症部位とがん部位での Cluster E、F の割合の比較。(Fibrosis : 線維症部位、Fibrosis Cancer : 間質性肺炎合併肺がんの肺がん部位) (D) Cluster E、F の GO 解析データ。

これら scRNA-seq の解析データより間質性肺炎合併肺がんの間質細胞では①特異的なクラスターは見られず、②その分画は線維症と肺がんの中間であり、肺がん組織であるものの正常組織もしくは線維症肺でみられる線維芽細胞ががん生存・増殖の微小環境調整に作用している可能性が考えられた。また、肺線維症で指摘された *Cthrc1* 陽性線維芽細胞は肺がんでも見られ疾患肺に普遍的に見られる線維芽細胞クラスターであること、中でも今回指摘されたクラスター E、F はがんに比較的特徴的に見られる細胞集団でありその ECM 産生や血管増生などの機能から治療標的となりうることを示唆された。

これら収集された scRNA-seq のデータ及び 8 症例（間質性肺炎非合併例 5 例（正常部位 5 件、肺がん部位 5 件）、間質性肺炎合併例 3 例（正常部位が 1 件、間質性肺炎部位が 4 件、肺がん部位が 3 件））での組織 bulk の mRNA 発現解析より創薬ターゲット候補を特定するデータベース作製につなげる。

E. 研究発表

1. 論文発表

- 1) Mitsuhashi A, Kondoh K, Horikawa K, Koyama K, Nguyen N, Afroj T, Yoneda H, Otsuka K, Ogino H, Nokihara H, Shinohara T, Nishioka Y. PD-1/PD-L1 blockade mediates anti-angiogenic effects by tumor-derived CXCL10/11 as a potential predictive biomarker. *Cancer Sci* 12(12):4853-4866, 2021.
- 2) Sato S, Chong SG, Upagupta C, Yanagihara T, Saito T, Shimbori C, Belleys PS, Nishioka Y, Kolb M. Fibrotic extracellular matrix induces release of extracellular vesicles with profibrotic miRNA from fibrocytes. *Thorax* 76(9):895-906, 2021.
- 3) Afroj T, Mitsuhashi A, Ogino H, Saijo A, Otsuka K, Yoneda H, Tobiume M, Nguyen TN, Goto H, Koyama K, Sugimoto M, Kondoh O, Nokihara H, Nishioka Y. Blockade of PD-1/PD-L1 pathway enhances the antigen-presenting capacity of fibrocytes. *J Immunol* 206(6):1204-1214, 2021.

4)Kawano H, Koyama K, Nishimura H, Toyoda Y, Kagawa K, Sato S, Naito N, Goto H, Inagaki Y, Nishioka Y. Development of improved method to identify and analyze lung fibrocytes with flow cytometry in a reporter mouse strain. Immun Inflamm Dis 9(1):120-127, 2021.

2. 学会発表
該当なし

F. 知的財産権の出願・登録状況（予定を含む。）

1. 特許取得
該当なし

2. 実用新案登録
該当なし

3. その他
該当なし

**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）
分担研究報告書**

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に関する研究

研究分担者名 : 国立大学法人 京都大学 奥野恭史

研究要旨

創薬標的を探索する AI 開発

薬剤反応性の識別、バイオマーカー推定、分子メカニズムの解明と創薬標的探索を行う AI の開発

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・利用して創薬標的を探索する 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) を含む間質性肺炎およ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標的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、創薬標的候補の実験的検証に有益となる基盤構築を包括的に遂行する。当該年度は、i) データ収集: IPF を含む間質性肺炎で 250 検体及び肺がんデータ収集の完結、ii) データ解析: これまでに開発した患者層別化 AI の改良および創薬標的探索に有用な AI の多角的開発、iii) 結果解釈・仮説創出: データウェアハウス TargetMine への公共・市販データベース追加と基礎科学系文献からの自動知識抽出に必要な言語リソースおよび AI 手法構築の継続、iv) エクソソーム糖鎖データベース構築に向けた基盤技術の確立、v) 肺がんを併発した IPF における一細胞解析を介した対象疾患の分子基盤分析を目標とする。

A. 研究目的

患者個別の疾患タイプ、薬剤反応性等について、識別・分類、分子メカニズムの解明、バイオマーカー推定と創薬標的分子探索を可能にする AI 技術を開発する。特に、生命科学や医療における様々な情報を創薬に応用する上で問題となる以下の課題の克服を目指す。

「克服すべき技術的な課題」

- ・全ゲノム配列 (超次元データ)、マルチオミックスのデータ処理
- ・薬剤反応性などの患者層別化・個別化
- ・多種多様情報 (実験データ、臨床データ、文献データ等) の統合
- ・分子・パスウェイ・細胞・臓器・動物・ヒトのマルチスケールモデル
- ・AI モデルへの時間的・空間的・定量的概念の導入

B. 研究方法

① DNN や機械学習による薬剤反応性の識別とバイオマーカー推定

限られたデータを有効活用し、効率的に薬剤反応性の予測やバイオマーカーの推定を行うためには、ゲノムやオミックスデータ、薬剤 (化合物)、臨床情報の情報を組み合わせた超高次元のデータを扱う必要がある。

そのため、適切に feature engineering、次元圧縮、特徴選択等の技術を用いて、本質的なデータを抽出することが重要である。また、これらのデータを組み合わせて利用する方法として、オミクスデータ、化合物、臨床情報などの複数種類のデータを入力するマルチモーダルアプローチと複数の出力薬剤反応性を同時に扱うマルチタスクアプローチが知られている。今年度では、前年度で開発した実装を元に、実際に公共で利用可能なデータを用いてモデル構築を行い、これら二つのアプローチを薬剤反応性識別やバイオマーカーの推定に利用する場合にどのような長所や短所があるかといったことを明らかにする。

本年度では、公開ベンチマークデータを用いて、上記の既存手法に関する調査および実装を行い、これらのアプローチの優位性について評価を行った。

② GCN によるネットワーク構造を考慮した薬剤反応性の識別、バイオマーカー推定、創薬ターゲット分子探索

限られたコストで効率的に実験を行い、得られた実測データを効果的に活用する方法の一つは、文献や既存の公共データベースの知識を利用することである。文献や既存の公共データベースにある網羅的な情報（文献空間の情報）を実験によって得られた実測空間の情報を統合した機械学習の方法を開発することで、少ない実測データからの効率的な予測を実現する。これらの異なる性質を持つ情報を柔軟に表現する方法として、本研究課題ではグラフ表現（ネットワーク）を利用する。

今年度ではプロトタイプモデルを元に、より実装の強化を行い、翌年度以降の実データを用いた予測モデルの構築に向けたベースとなる手法開発を行った。

③ ベイジアンネットワークによる分子メカニズムの解明、バイオマーカー推定、創薬ターゲット分子探索

ベイジアンネットワークを用いた研究のうち、今年度は薬剤感受性に基づく患者層別化が可能な分子メカニズム解明・解析手法の確立を目指す（全体計画 3-1）。前年度においてトランスクリプトーム・データを用いたサンプル単位でのベイジアンネットワーク推定法は確立している。これを応用し、新規に取得するオミクス・データと薬剤感受性試験データを用いて、患者ごとの薬剤への反応の違いを、ベイジアンネットワークを用いてモデル化する。得られたネットワークモデルを用いてゲノムデータ及び薬剤感受性との相関を解析する。以上により、各薬剤において感受性の違いを説明するトランスクリプトーム・ネットワークを抽出する方法を確立し、薬剤感受性に違いが発生するメカニズムの、トランスクリプトーム・ネットワークレベルでの解明及び薬剤感受性に基づく患者の層別化を目指す。

（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本学において倫理審査、承認を得た後、ヒトゲノム・遺伝子解析研究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 DNN や機械学習による薬剤反応性の識別とバイオマーカー推定

Omics ベースのアプローチ (Ding, Michael Q., et al. "Precision oncology beyond targeted therapy: combining omics data with machine learning matches the majority of cancer cells to effective therapeutics." *Molecular Cancer Research* 16.2 (2018): 269-278.) と薬剤ベースのアプローチ (Chang, Yoosup, et al. "Cancer drug response profile scan (CDRscan): a deep learning model that predicts drug effectiveness from cancer genomic signature." *Scientific reports* 8.1 (2018): 1-11.) の二つのアプローチに関する二つの既存手法の評価を行った。

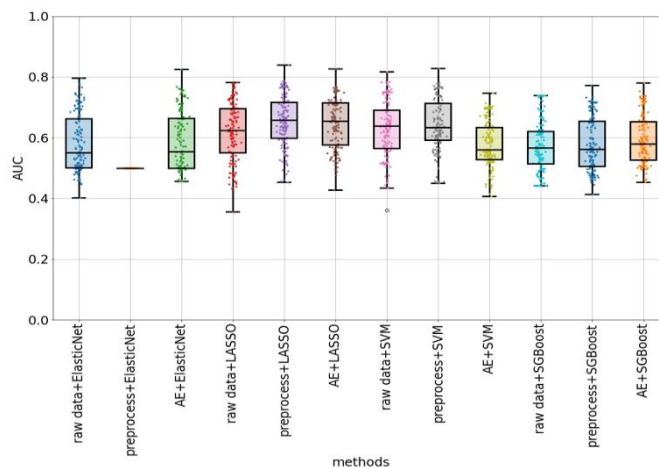


図 1：Omics ベースのアプローチ：従来法における薬剤活性予測の正答率比較

Omics ベースのアプローチでは、手法が有効な薬剤とそうでない薬剤が存在していることがわかり、薬剤によって手法を使い分ける必要性などが示唆された（図 1）。また、薬剤ベースのアプローチでは、従来モデルに加えいくつかの新規モデルを用いて評価を行った（図 2）。データとしては、現状公開データ G D S C・C C L Eを用いたものである。評価結果としてはデータのバランスを調整するなどのデータに合わせた処理がいくつか効果的であることが判明した。

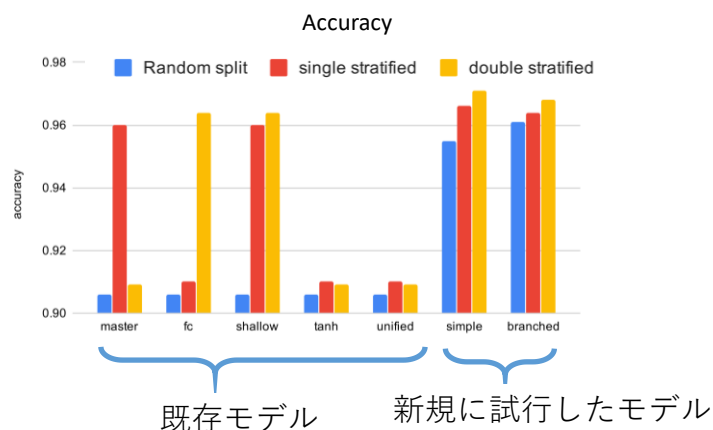


図 2：薬剤ベースのアプローチ：ゲノム変異と薬剤(抗がん剤)の情報を用いたマルチモーダル薬剤応答予測モデルの比較

② GCN によるネットワーク構造を考慮した薬剤反応性の識別、バイオマーカー推定、創薬ターゲット分子探索

遺伝子・薬剤・疾患の文献等から構築したグラフから新たな関係を予測する手法と予測結果を説明する部分を可視化する手法を開発し、GCN を用いた手法が他の手法よりも精度よく予測を行うことができることを示した（図 3-1）。また、この時の予測結果を説明する部分を可視化する手法を開発した（図 3-2）。

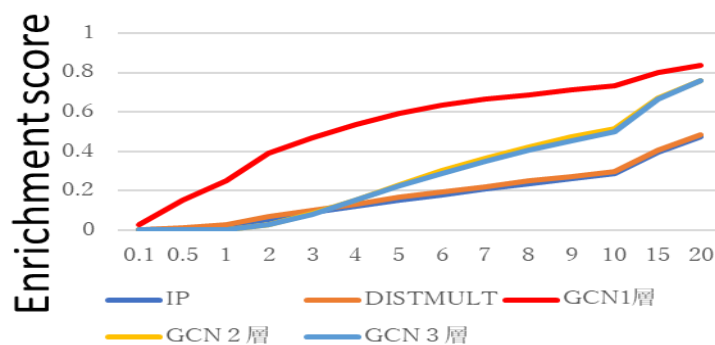


図 3-1：遺伝子・薬剤・疾患グラフの補間問題における手法間比較

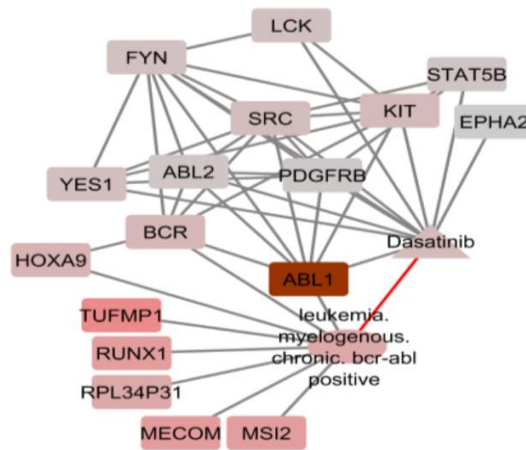


図 3-2：遺伝子・薬剤・疾患グラフの予測結果の可視化

また、このグラフ補間問題をさらに、予測問題へと適用する手法開発を進めており、文献情報と実験情報を合わせたモデルの開発が完了しており、詳細な評価および次年度に向けた実データへの適用可能性について検討を進めている。

③ ベイジアンネットワークによる分子メカニズムの解明、バイオマーカー推定、創薬ターゲット分子探索 (1) がん患者層別化手法の確立

昨年度開発した遺伝子ネットワークからの重要サブネットワーク抽出技術を改良し、分子メカニズム解明に有効ながんサブタイプ分類法を開発した。TCGA 胃がんデータセット (The Cancer Genome Atlas Research Network, *Nature*, 2014) では、そのオミクスデータによる 4 つの新規サブタイプが提案されているが、患者の予後との関連は見出されていなかった。ベイジアンネットワークを用いた重要サブネットワーク抽出技術では、患者サンプル毎に、ネットワークの枝 (エッジ) 毎の重要度を定量化することが可能であるが、この定量化方法によるデータ行列 (= 患者毎の遺伝子ネットワークのプロファイル) をクラスタリングすることにより、生存時間に統計的な差がある 3 つのサブタイプを同定することができた (図 4)。同定されたサブタイプ毎に重要なサブネットワークは異なっていることもわかり、これらの患者グループ毎に分子メカニズムが異なることが示唆された。本研究は BioRxiv に投稿済みで (Nakazawa et al. *BioRxiv*, 2021.03.24.436731)、学術雑誌への投稿も目前である。

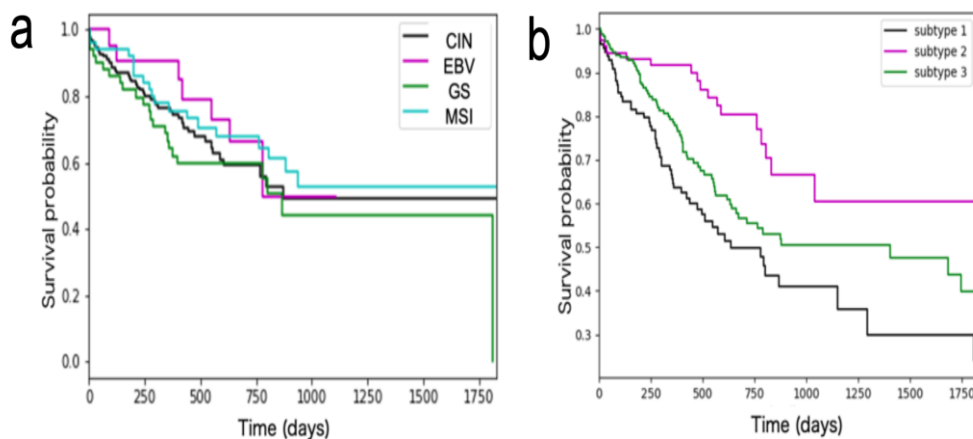


図 4. (a) 既存 4 サブタイプの生存時間解析。(b) 新規開発手法により同定した 3 サブタイプの生存時間解析 (任意の 2 グループ間の p -value < 0.05)。

(2) 遺伝子ネットワーク・サブネットワーク抽出法の検証

次に、開発した手法のさらなる検証を目的として、最近公開された SARS-CoV-2/COVID-19 データを用いて遺伝子ネットワーク解析を行い、創薬ターゲット探索への適用可能性を検証した。結果は論文にまとめ arXiv に登録後、すでに学術雑誌に投稿済みである (Tanaka et al., 2020, arXiv:2008.09261)。抽出した遺伝子

ネットワークには既存薬ターゲット遺伝子が多く含まれていることが確認でき、既存知識も矛盾のない結果が得られた（図5）。このサブネットワークに含まれる遺伝子は基本的に全て新規の創薬ターゲットになりうると言えるが、特にモジュールとして得られた境界の遺伝子や、子供遺伝子の多いハブ遺伝子、既存標的遺伝子の近傍の遺伝子は特に有力な新規候補遺伝子と言える。

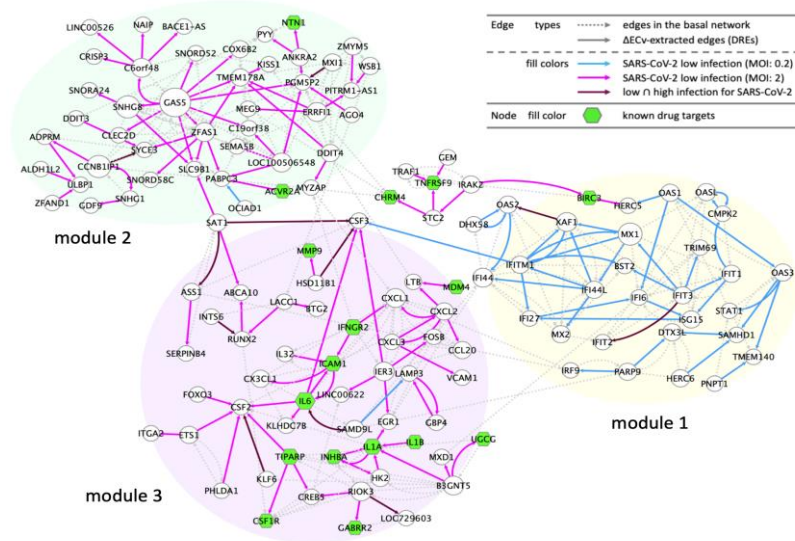


図5. SARS-CoV-2 データの遺伝子ネットワーク解析結果。ウィルス量の違いによるサンプル間の比較によりサブネットワークを抽出した。3つのモジュールが観測され、これらの境界に存在する遺伝子が、新規創薬ターゲットとしては有力である。緑色のノードが既存薬の標的遺伝子である。

(3) 薬剤感受性に基づくネットワーク抽出技術の開発

10患者からの複数薬剤毎の体積変化（腫瘍増殖率）とトランスクリプトームデータ32サンプルを入手することができたので、その解析を行なった。目的は遺伝子ネットワーク解析により薬剤感受性を予測・説明できる遺伝子間の関係性（＝遺伝子ネットワーク中のエッジ）の抽出が可能かどうか検証することである。これまで開発した部分ネットワーク抽出法は、2群間、あるいは2サンプル間での推定されたシステム上での振舞いの違いにより、部分ネットワークを抽出する方法だったが、今回新たな方法として、薬剤ごとの腫瘍増殖率と強く相関するエッジを抽出する方法を考案し、試行した。この方法は、コントロールサンプルなどの比較対象や薬剤感受性・耐性のラベル付けが不要であるため、応用範囲が広いと思われる。偽相関により単一の非連結エッジが多数抽出される可能性も考えていたが、薬剤に関連する連結部分ネットワークがモジュールとして抽出可能であることが示された（図6）。次に抽出した遺伝子ネットワークモジュールによって、腫瘍増殖率が予測できるかどうかの検証を行なった。予測がある程度可能であれば、抽出した遺伝子ネットワークモジュールに薬剤応答性を説明できる情報が十分あることが言える。図7は抽出した連結成分ごとに、leave-one-out 交差検証法による予測能力の結果である。腫瘍増殖率のデータをクラスタリングすることにより、サンプルを薬剤応答群と非応答群に二分し、予測問題の正解ラベルとした。上段は連結成分の枝のECvを特徴量として用いた場合の予測率、下段は連結成分を構成する遺伝子の発現量を特徴量として用いた場合の予測率を示している。サンプル数が10と非常に少ないが、連結成分によっては十分な予測精度を示しており、抽出したサブネットワークのサンプルごとに使われ方の違いにより薬剤の効果を予測できるという結果になった。以上より、抽出したサブネットワークが薬剤反応機序を説明できている可能性が確かめられた。今後はより大規模なデータセットでの検証を引き続き進める予定である。

以上のほか、理化学研究所のスーパーコンピュータ「富岳」を一般利用開始前に利用できる早期利用課題に採択されたため、遺伝子ネットワーク推定ソフトウェアの「富岳」への対応を行い、大規模ベイジアンネットワーク推定アルゴリズムであるNNSRアルゴリズムおよびサンプル毎の枝貢献量計算の富岳対応を完了している。

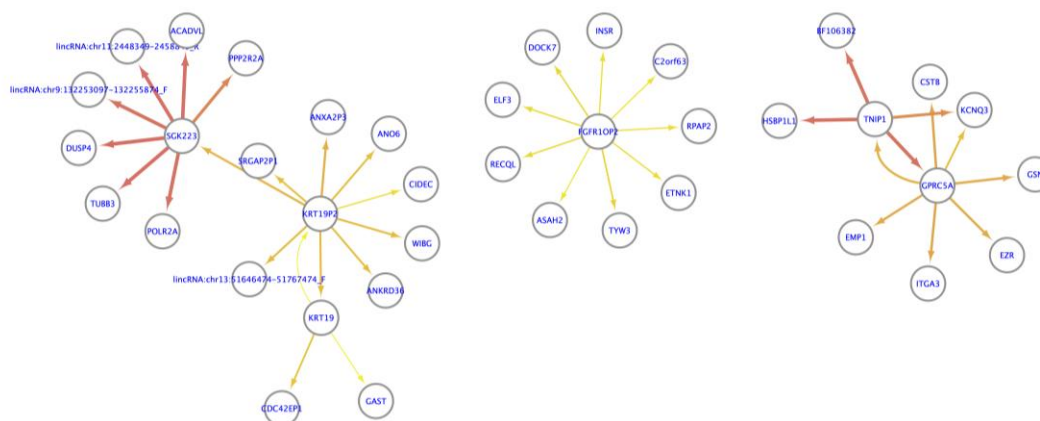


図6 セツキシマブ投与に対する腫瘍体積変化率と相関する遺伝子ネットワークエッジ抽出結果例 (preliminary results)。エッジ単体ではなくハブ的な遺伝子を中心とした部分ネットワークがモジュールとして抽出できていることがわかる。

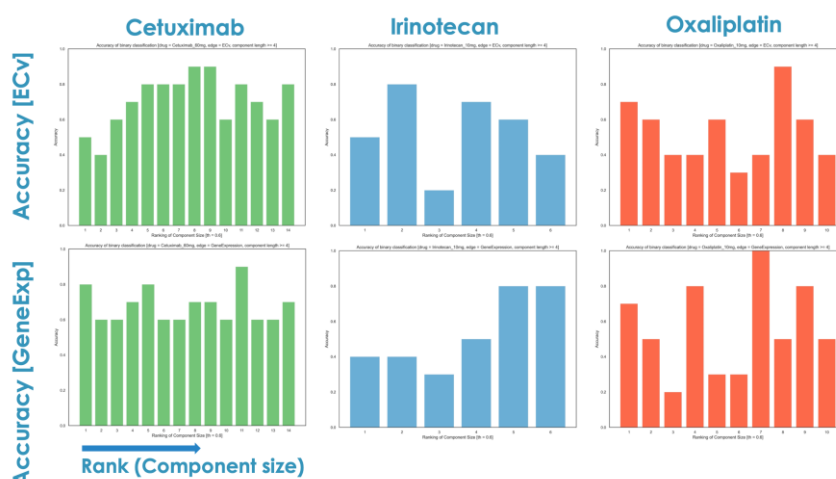


図7. 抽出したサブネットワークの連結成分ごとの腫瘍増殖率予測結果。上段は枝貢献量による予測。下段は連結成分に含まれる遺伝子遺伝子発現量による予測結果。サブネットワークにより8~9割程度の予測能力があることがわかった。

D. 考察

① DNN や機械学習による薬剤反応性の識別とバイオマーカー推定

Omics ベースのアプローチおよび薬剤ベースのアプローチの結果より、Omics ベースのアプローチでは、手法が有効な薬剤とそうでない薬剤が存在していることがわかり、薬剤による影響を強く受けてしまうことが分かった。また、薬剤ベースのアプローチでは、データの偏りや前処理にモデルの性能が影響を受けるため、これらを考慮した手法の性能が高いことが分かったため、より精度の高いモデルを構築するためには、データのの前処理等を含めたシステム全体の設計を検討する必要があることが分かった。

② GCN によるネットワーク構造を考慮した薬剤反応性の識別、バイオマーカー推定、創薬ターゲット分子探索

ネットワーク予測に関しては、遺伝子・薬剤・疾患グラフの補間問題におけるベンチマークでの本アプローチの有効性を示した。今後、ベンチマークデータではなくよりリアルなデータを用いての評価を行い、有用性を評価する必要がある。

③ ベイジアンネットワークによる分子メカニズムの解明、バイオマーカー推定、創薬ターゲット分子探索
今年度は、昨年度開発したサンプルごとに遺伝子ネットワークを推定可能な方法を応用し、患者サブタイプ層別化、薬剤耐性層別化、薬剤標的遺伝子探索への応用の方法論の開発と検証を行なった。開発した方法論により、これまでできなかった遺伝子ネットワークを用いたサブタイプや薬剤耐性などが可能であることが確認でき、また抽出したサブネットワークは既存薬剤標的遺伝子を多く含むことがわかった。今後は計算機による結果を実データで検証する必要がある、それを通して手法の更なるブラッシュアップを図る必要がある。

E. 結論

① DNN や機械学習による薬剤反応性の識別とバイオマーカー推定

薬剤ベースのアプローチをベースにしつつ、適宜 Omics 情報を入れていくアプローチが有用なのではないかと推察され、今後、この薬剤ベースのモデルに②で検討中の文献情報やより精緻な特徴量を用いてさらに精度の高いモデルを目指す。

② GCN によるネットワーク構造を考慮した薬剤反応性の識別、バイオマーカー推定、創薬ターゲット分子探索

ネットワーク構造を考慮する手法である GCN の有効性を確かめることができたため、これを用いて、①とも合わせて、薬剤予測問題へと適用する手法開発を進めて、文献情報と実験情報を合わせたモデルの構築を行う。さらに今後、この手法を強化し、バイオマーカー推定、創薬ターゲット分子探索へも利用していく方針である。また、実データへの適用に向けて、ツールとしてのさらなる整備を進めている。

③ ベイジアンネットワークによる分子メカニズムの解明、バイオマーカー推定、創薬ターゲット分子探索

トランスクリプトームデータを用いて、メカニズム解明、バイオマーカー推定、創薬ターゲット探索がこれまで以上に可能になることがわかった。今後は、プロテオームデータやゲノムデータなどの統合解析が鍵であり、計算による仮説の実験による検証を行いつつ、データ統合のための方法論の開発を進めたい。

G. 研究発表

1. 論文発表

- 1) Matsumoto S, Ishida S, Araki M, Kato T, Terayama K, Okuno Y " Extraction of protein dynamics information from cryo-EM maps using deep learning " Nat Mach Intell 3, 153-160 (2021).
- 2) Ryosuke Shibukawa; Shoichi Ishida; Kazuki Yoshizoe; Kunihiro Wasa; Kiyosei Takasu; Yasushi Okuno; Kei Terayama; Koji Tsuda, "CompRet: a comprehensive recommendation framework for chemical synthesis planning with algorithmic enumeration," J Cheminform 12, 52 (2020)
- 3) Masuda, N., Murakami, K., Kita, Y., Hamada, A., Kamada, M., Teramoto, Y., Matsumoto, K., Sano, T., Saito, R., Okuno, Y., Ogawa, O., Kobayashi, T. Trp53 mutation in Krt5-expressing basal cells facilitates the development of basal squamous-like invasive bladder cancer in the chemical carcinogenesis of mouse bladder, Am. J. Pathol., 190(8), 1752-1762 (2020)
- 4) Saito, Y., Koya, J., Araki, M., Kogure, Y., Shingaki, S., Tabata, M., McClure, M., Yoshifuji, K., Matsumoto, S., Isaka, Y., Tanaka, H., Kanai, T., Miyano, S., Shiraishi, Y., Okuno, Y., Kataoka, K. Landscape and function of multiple mutations within individual oncogenes, Nature, 582, 95-99 (2020)
- 5) Ono, F., Chiba, S., Isaka, Y., Matsumoto, S., Ma, B., Katayama, R., Araki, M., Okuno, Y. Improvement in predicting drug sensitivity changes associated with protein mutations using a molecular dynamics based alchemical mutation method, Sci. Rep., 2020, 10(1):2161.
- 6) Tanaka, Y., Tamada, Y., Ikeguchi, M., Yamashita, F., Okuno, Y. System-based differential gene network analysis for characterizing a sample-specific subnetwork, Biomolecules, 2020, 10(2):306.
- 7) Hatae, R., Chamoto, K., Young, H., Kim, S., Sonomura, K., Taneishi, K., Kawaguchi, S., Yoshida, H., Ozasa, H., Sakamori, Y., Akrami, M., Fagarasan, S., Masuda, I., Okuno, Y., Matsuda, F., Hirai, T., Honjo, T. Combination of host immune metabolic biomarkers for the PD-1 blockade cancer immunotherapy, JCI Insight, 2020, 5(2), 133501.
- 8) Iwata, H., Kojima, R., Okuno, Y. An in Silico Approach for Integrating Phenotypic and Target-

2. 学会発表

- 1) 日本学術会議 未来からの問い—日本学術会議 100 年を構想する (2020 発行)
第 4 章 医療の未来社会、4-3 医療におけるビッグデータ・AI、(2) ビッグデータ・AI が拓く医療・創薬の未来
- 2) 「スーパーコンピュータ「富岳」による COVID-19 治療薬探索」 “Drug Screening for COVID-19 using Supercomputer “Fugaku” 第 94 回日本薬理学会年会 特別講演 AI 創薬の現状と未来 Today and future of AI-based drug development シンポジウム (2021. 03. 08)
- 3) 「AI・シミュレーションによる新型コロナウイルス治療法開発への挑戦」 特別講演 AI ネットワーク社会フォーラム ～AI-Ready 社会の実現に向けて～ (2021. 03. 01)
- 4) 「データ駆動型 医療・医薬品・ヘルスケア産業」 IQVIA セミナー (2021. 02. 15)
- 5) 「スーパーコンピュータ「富岳」・AI による新型コロナウイルス治療法開発への挑戦」 第 38 回コロイド界面技術シンポジウム 「みんなを元気にするすごい技術 アフターコロナの研究開発 ～動向/指針/変化する研究」 (2021. 02. 04)
- 6) 「AI が拓く創薬イノベーション」 第 16 回がんトランスレーショナルリサーチ (TR) ワークショップ AI が創る医薬品開発のカッティング・エッジ (2021. 01. 19)
- 7) 「AI・シミュレーションが拓く創薬・医療の未来」 けいはんな R&D イノベーションフォーラム (2020. 12. 15.)
- 8) 「「富岳」がもたらす創薬・医療へのインパクト」 第 13 回 スーパーコンピューティング技術 産業応用シンポジウム (2020. 12. 10.)
- 9) 「スーパーコンピュータ・AI が拓く創薬の未来」 Life Science Startup Ecosystem (2020. 12. 10.)
- 10) 「スーパーコンピュータ・AI が拓く創薬の未来」 第 3 回産学官連携情報交流セミナー (2020. 12. 03.)
- 11) 20 年度精神・神経疾患研究開発費班会議 (2020. 12. 01.)
- 12) 「AI・シミュレーションが拓く創薬・医療の未来」 CHUGAI DIGITAL DAY～ヘルスケア×デジタルの 2030 未来予想図 (2020. 11. 27)
- 13) 「スーパーコンピュータ「富岳」・AI による新型コロナウイルス治療法開発への挑戦」 総務省 情報通信政策研究所主催「AI ネットワーク社会推進会議」 (2020. 11. 12.)
- 14) 創薬における AI の現状と未来 奥野恭史 第 21 回日本毒性病理学会・教育セミナー 2020. 11. 7
- 15) スパコン「富岳」・AI による新型コロナウイルス治療法開発への挑戦 奥野恭史 数理 AI データサイエンス市民講座 (2020. 10. 23)
- 16) スーパーコンピュータ・AI が拓くがん分子標的治療戦略 奥野恭史 第 24 回日本がん分子標的治療学会学術集会 (2020. 10. 8)
- 17) スーパーコンピュータ「富岳」vs 新型コロナウイルス 奥野恭史 STS フォーラム公開シンポジウム (2020. 10. 3)
- 18) スーパーコンピュータ・AI が拓く創薬・医療の未来 奥野恭史 大阪府工業協会・役員会講演 (2020. 9. 28)
- 19) AI・シミュレーションによる薬効・毒性予測 奥野恭史 第 47 回日本毒性学会学術年会シンポジウム (2020. 6. 29)

F. 知的財産権の出願・登録状況 (予定を含む。)

1. 特許取得

なし

2. 実用新案登録

なし

3. その他

- 1) Life Intelligence Consortium (LINC) 代表

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))
分担研究報告書

課題名 : 層別化 AI による解析および解釈の精度向上のための研究
研究分担者名 : 熊ノ郷 淳 大阪大学大学院医学系研究科 呼吸器・免疫内科

研究要旨

PRISM 課題である「新薬創出を加速する人工知能の開発」において、大阪大学間質性肺炎コホート臨床データ（診療データおよびエクソソームプロテオームデータ）を収集、構造化し、データ駆動的に IPF 患者の特徴を示すクラスターが有する特徴的な分子群の抽出を行った。抽出された分子群の生物学的意味の知識整理やさらなるパスウェイ解析を行い、線維化と関連するパスウェイやそのパスウェイにおいてハブとなる分子の特定に至った。見出したパスウェイや特定したハブ分子については、有効な治療法やバイオマーカーの早期開発が望まれている致死疾患である IPF の新規創薬ターゲットとなる可能性がある。そこで、本研究においては、見出したパスウェイや特定したハブ分子が、IPF 患者において発現しているか、あるいは機能しているかを明らかにする。

A. 研究目的

見出したパスウェイや特定したハブ分子が、IPF 患者において発現しているか、あるいは機能しているかを明らかにする。すなわち、①線維化肺組織における当該分子の発現、リン酸化パスウェイ経路の存在の証明、および②特定したパスウェイが活性化した場合に産生される代謝物の測定を行う。そのことにより、当該分子の臨床的意義を明らかにするとともに、PRISM プロジェクトにおいて開発した新規創薬ターゲットおよび層別化バイオマーカー探索のための戦略の妥当性を科学的に証明する。

B. 研究方法

対象となる患者群は、文書同意を取得した下記の患者

①については、肺がん患者（線維症合併および非合併）

②については、IPF 確定診断患者および器質的呼吸器疾患を有さない受診者

① 手術肺の収集：手術時の摘出肺より線維化部分および正常部分を採取し、ただちに凍結する。

② 血液および尿の採取：受信時に採取する。可能な限り時間や食事後の時間、投薬前後の条件をそろえる。

(倫理面への配慮)

大阪大学医学部付属病院における倫理規定に基づき、研究計画を行った。

C. 研究結果

今年度は、検体採取により解析依頼中であり、解析結果をもとに今後の計画を立案予定である。

D. 考察

解析結果をもとに今後の計画を立案予定である。

E. 結論

見出したパスウェイや特定したハブ分子が、IPF 患者において発現を検討中である。

G. 研究発表

1. 論文発表

なし

2. 学会発表
なし

F. 知的財産権の出願・登録状況（予定を含む。）

1. 特許取得
なし

2. 実用新案登録
なし

3. その他
なし

研究成果の刊行に関する一覧表

雑誌

発表者氏名	論文タイトル名	発表誌名	巻号	ページ	出版年
Esaki T., Horinouchi T., Natsume-Kitatani Y., Nojima Y., Sakane I., Matsui H.	Estimation of relationships between gene expression and chemical substructures: Adapting canonical correlation analysis for small sample data by gathered features using consensus clustering	Chem-Bio Informatics Journal	20	58-61	2020
Allendes Osorio R., Nyström-Persson J., Nojima Y., Kosugi Y., Mizuguchi K., Natsume-Kitatani Y.	Panomicon: A web-based environment for interactive, visual analysis of multi-omics data	Heliyon	6(8)	e04618	2020
Nojima Y., Takeda Y., Maeda Y., Bamba T., Fukusaki E., Itoh M., Mizuguchi K., Kumanogoh A.	Metabolomic analysis of fibrotic mice combined with public RNA-Seq human lung data reveal potential diagnostic biomarker candidates for lung fibrosis	FEBS Open Bio	11(10)	2427-2436	2020
Teranishi Y., Kuwahara H., Ueda M., Takemura T., Kusumoto M., Nakamura K., Sakai J., Kimura T., Furutani Y., Kawashima M., Imokawa G., Nogami-Itoh M.	Sphingomyelin Deacylase, the Enzyme Involved in the Pathogenesis of Atopic Dermatitis, Is Identical to the β -Subunit of Acid Ceramidase	International Journal of Molecular Sciences	22(21)	8789	2020
Tripathi L., Itoh M., Takeda Y., Tsujino K., Kondo Y., Kumanogoh A., Mizuguchi K.	Integrative Analysis Reveals Common and Unique Roles of Tetraspanins in Fibrosis and Emphysema	Frontiers in Genetics	11	585998	2020
Takahashi H., Kamiyama T., Hirane N., Kobayashi N., Aiyama T., Nagatsu A., Shimada S., Orimo T., Kakisaka T., Fukai M., Yokoo H., Kamachi H., Nishimura S-I., Taketomi A.	Analysis of the correlation between alterations in N-glycans and invasiveness in liver cancer cell lines	Oncol. Rep.	44	2757-2769	2020
Pablo A. Guillen-Poza, Elena Matilde Sánchez-Fernández, Gerard Artigas, José Manuel García Fernández, Hinou H., Carmen Ortiz Mellet, Nishimura S-I., Fayna Garcia-Martin	Amplified detection of breast cancer autoantibodies using MUC1-based Tn antigen mimics	J. Med. Chem.	63	8524-8533	2020

発表者氏名	論文タイトル名	発表誌名	巻号	ページ	出版年
Yokoi Y., Nishimura S-I.	Effect of Site - specific O-Glycosylation on the Structural Behavior of NOTCH1 Receptor Extracellular EGF - like Domains 11 and 10	Chem. Eur. J.	26	12363-12372	2020
Wakui H., Tanaka Y., Ose T., Matsumoto I., Kato K., Yao M., Tachibana T., Sato M., Naruchi K., García-Martín F., Hinou H., Nishimura S-I.	A straightforward approach to antibodies recognising cancer specific glycopeptidic neopeptides	Chem. Sci	11	4999-5006	2020
Ken Asada, Kazuma Kobayashi, Samuel Joutard, Masashi Tubaki, Satoshi Takahashi, Ken Takasawa, Masaaki Komatsu, Syuzo Kaneko, Jun Sese, Ryuji Hamamoto	Uncovering Prognosis-Related Genes and Pathways by Multi-Omics Analysis in Lung Cancer.	Biomolecules	10	524	2020
Theodore Vougiouklakis, Riyue Bao, Sohyoung Kim, Songjoon Baek, Makda Zewde, Benjamin Bernard, Kyunghye Burkitt, Nupur Nigam, Evgeny Izumchenko, Naoshi Dohmae, Ryuji Hamamoto, Yusuke Nakamura	WHSC1 monomethylates histone H1 and induces stem-cell like features in squamous cell carcinoma of the head and neck.	Neoplasia	22	283-293	2020
Ken Asada, Amina Bolatkan, Ken Takasawa, Masaaki Komatsu, Syuzo Kaneko, Ryuji Hamamoto	Critical Roles of N6-Methyladenosine (m6A) in Cancer and Virus Infection.	Biomolecules	10	524	2020
Shunichi Jinnai, Naoya Yamazaki, Yuichiro Hirano, Yohei Sugawara, Yuichiro Ohe, Ryuji Hamamoto	The Development of a Skin Cancer Classification System for Pigmented Skin Lesions Using Deep Learning.	Biomolecules	10	1123	2020
Kazuma Kobayashi, Amina Bolatkan, Shuichiro Shiina, Ryuji Hamamoto	Fully-Connected Neural Networks with Reduced Parameterization for Predicting Histological Types of Lung Cancer from Somatic Mutations.	Biomolecules	10	1249	2020
Satoshi Takahashi, Ken Asada, Ken Takasawa, Ryo Shimoyama, Akira Sakai, Amina Bolatkan, Norio Shinkai, Kazuma Kobayashi, Masaaki Komatsu, Syuzo Kaneko, Jun Sese, Ryuji Hamamoto	Predicting Deep Learning Based Multi-Omics Parallel Integration Survival Subtypes in Lung Cancer Using Reverse Phase Protein Array Data.	Biomolecules	10	1460	2020

発表者氏名	論文タイトル名	発表誌名	巻号	ページ	出版年
Machiko Kojima, Kenbun Sone, Katsutoshi Oda, Ryuji Hamamoto, Syuzo Kaneko, Shinya Oki, Asako Kukita, Akira Kawata, Harunori Honjoh, Yoshiko Kawata, Tomoko Kashiya, Masakazu Sato, Ayumi Taguchi, Yuichiro Miyamoto, Michihiro Tanikawa, Tetsushi Tsuruga, Kazunori Nagasaka, Osamu Wada-Hiraike, Yutaka Osuga, Tomoyuki Fujii	The histone methyltransferase SMYD2 is a novel therapeutic target for the induction of apoptosis in ovarian clear cell carcinoma cells.	Oncology Letters	20	153	2020
Ai Dozen, Masaaki Komatsu, Akira Sakai, Reina Komatsu, Kanto Shozu, Hidenori Machino, Suguru Yasutomi, Tatsuya Arakaki, Ken Asada, Syuzo Kaneko, Ryu Matsuoka, Daisuke Aoki, Akihiko Sekizawa, Ryuji Hamamoto	Image Segmentation of the Ventricular Septum in Fetal Cardiac Ultrasound Videos Based on Deep Learning Using Time-Series Information.	Biomolecules	10	15426	2020
Vasiliki Papadaki, Ken Asada, Julie K Watson, Toshiya Tamura, Alex Leung, Jack Hopkins, Margaret Dellett, Noriaki Sasai, Hongorzul Davaapil, Serena Nik-Zainal, Rebecca Longbottom, Makoto Nakakido, Ryo Torii, Abhi Veerakumarasivam, Syuzo Kaneko, Mandeep S Sagoo, Gillian Murphy, Akihisa Mitani, Kohei Tsumoto, John D Kelly, Ryuji Hamamoto, Shin-Ichi Ohnuma	Two Secreted Proteoglycans, Activators of Urothelial Cell-Cell Adhesion, Negatively Contribute to Bladder Cancer Initiation and Progression.	Cancers	12	3362	2020
Ryuji Hamamoto, Kruthi Suvarna, Masayoshi Yamada, Kazuma Kobayashi, Norio Shinkai, Mototaka Miyake, Masamichi Takahashi, Shunichi Jinnai, Ryo Shimoyama, Akira Sakai, Ken Takasawa, Amina Bolatkan, Kanto Shozu, Ai Dozen, Hidenori Machino, Satoshi Takahashi, Ken Asada, Masaaki Komatsu, Jun Sese, Syuzo Kaneko	Application of Artificial Intelligence Technology in Oncology: Towards the Establishment of Precision Medicine.	Cancers	12	3532	2020
Miku Wada, Asako Kukita, Kenbun Sone, Ryuji Hamamoto, Syuzo Kaneko, Masaaki Komatsu, Yu Takahashi, Futaba Inoue, Machiko Kojima, Harunori Honjoh, Ayumi Taguchi, Tomoko Kashiya, Yuichiro Miyamoto, Michihiro Tanikawa, Tetsushi Tsuruga, Mayuyo Mori-Uchino, Osamu Wada-Hiraike, Yutaka Osuga, Tomoyuki Fujii	Epigenetic Modifier SETD8 as a Therapeutic Target for High-Grade Serous Ovarian Cancer.	Biomolecules	10	1686	2020

発表者氏名	論文タイトル名	発表誌名	巻号	ページ	出版年
Kanto Shozu, Masaaki Komatsu, Akira Sakai, Reina Komatsu, Ai Dozen, Hidenori Machino, Suguru Yasutomi, Tatsuya Arakaki, Ken Asada, Syuzo Kaneko, Ryu Matsuoka, Akitoshi Nakashima, Akihiko Sekizawa, Ryuji Hamamoto	Model-Agnostic Method for Thoracic Wall Segmentation in Fetal Ultrasound Videos.	Biomolecules	10	1691	2020
Shun-Ichiro Kageyama, Junyan Du, Syuzo Kaneko, Ryuji Hamamoto, Shigeo Yamaguchi, Riu Yamashita, Masayuki Okumura, Atsushi Motegi, Hidehiro Hojo, Masaki Nakamura, Katsuya Tsuchihara, Tetsuo Akimoto	Identification of the mutation signature of the cancer genome caused by irradiation.	Radiotherapy and Oncology	155	10-16	2021
Kazuma Kobayashi, Naoya Murakami, Kana Takahashi, Koji Inaba, Hiroshi Igaki, Ryuji Hamamoto, Jun Itami	Tensor regression-based model to investigate heterogeneous spatial radiosensitivity after I-125 seed implantation for prostate cancer.	In vivo	35	489-497	2021
Masaaki Komatsu, Akira Sakai, Reina Komatsu, Ryu Matsuoka, Suguru Yasutomi, Kanto Shozu, Ai Dozen, Hidenori Machino, Hirokazu Hidaka, Tatsuya Arakaki, Ken Asada, Syuzo Kaneko, Akihiko Sekizawa, Ryuji Hamamoto	Detection of Cardiac Structural Abnormalities in Fetal Ultrasound Videos Using Deep Learning.	Applied Sciences	11	371	2021
Tatsuya Ozawa, Syuzo Kaneko, Frank Szulzewsky, Zhiwei Qiao, Mutsumi Takadera, Yoshitaka Narita, Tadashi Kondo, Eric C Holland, Ryuji Hamamoto, Koichi Ichimurai	C11orf95-RELA fusion drives aberrant gene expression through the unique epigenetic regulation for ependymoma formation.	Acta Neuropathol Commun	9	36	2021
Daisuke Takayanagi, Sou Hirose, Ikumi Kuno, Yuka Asami, Naoya Murakami, Maiko Matsuda, Yoko Shimada, Kuniko Sunami, Masaaki Komatsu, Ryuji Hamamoto, Mayumi Kobayashi Kato, Koji Matsumoto, Takashi Kohno, Tomoyasu Kato, Kouya Shiraishi, Hiroshi Yoshida	Comparative Analysis of Genetic Alterations, HPV-Status, and PD-L1 Expression in Neuroendocrine Carcinomas of the Cervix.	Cancers (Basel)	13	1215	2021

発表者氏名	論文タイトル名	発表誌名	巻号	ページ	出版年
Satoshi Takahashi, Masamichi Takahashi, Manabu Kinoshita, Mototaka Miyake, Risa Kawaguchi, Naoki Shinojima, Akitake Mukasa, Kuniaki Saito, Motoo Nagane, Ryohei Otani, Fumi Higuchi, Shota Tanaka, Nobuhiro Hata, Kaoru Tamura, Kensuke Tateishi, Ryo Nishikawa, Hideyuki Arita, Masahiro Nonaka, Takehiro Uda, Junya Fukai, Yoshiko Okita, Naohiro Tsuyuguch, Yonehiro Kanemura, Kazuma Kobayashi, Jun Sese, Koichi Ichimura, Yoshitaka Narita, Ryuji Hamamoto	Fine-Tuning Approach for Segmentation of Gliomas in Brain Magnetic Resonance Images with a Machine Learning Method to Normalize Image Differences among Facilities.	Cancers (Basel)	13	1415	2021
Takahashi S, Takahashi M, Tanaka S, Takayanagi S, Takami H, Yamazawa E, Nambu S, Miyake M, Satomi K, Ichimura K, Narita Y, Hamamoto R.	A New Era of Neuro-Oncology Research Pioneered by Multi-Omics Analysis and Machine Learning.	Biomolecules	11	565	2021
Kaneko S, Mitsuyama T, Shiraishi K, Ikawa N, Shozu K, Dozen A, Machino H, Asada K, Komatsu M, Kukita A, Sone K, Yoshida H, Motoi N, Hayami S, Yoneoka Y, Kato T, Kohn T, Natsume T, Keudell GV, Saloura V, Yamaue H, Hamamoto R.	Genome-Wide Chromatin Analysis of FFPE Tissues Using a Dual-Arm Robot with Clinical Potential.	Cancers (Basel)	13	2126	2021
Asada K, Kaneko S, Takasawa K, Machino H, Takahashi S, Shinkai N, Shimoyama R, Komatstu M, Hamamoto R.	Integrated Analysis of Whole Genome and Epigenome Data Using Machine Learning Technology: Toward the Establishment of Precision Oncology.	Front Oncol	12 (11)	666397	2021
Kobayashi K, Miyake M, Takahashi M, Hamamoto R.	Observing Deep Radiomics for the Classification of Glioma Grades.	Sci Rep	11	10942	2021
Lee K, Kim K, Ryu TY, Jung CR, Lee MS, Lim JH, Park K, Kim DS, Son MY, Hamamoto R, Cho HS.	Dysregulation of EHMT1 affects lung cancer proliferation by controlling CDKN1A expression.	Mol Oncol	15	2989-3002	2021
Komatsu M, Sakai A, Dozen A, Shozu A, Yasutomi S, Machino H, Asada K, Kaneko S, Hamamoto R.	Towards Clinical Application of Artificial Intelligence in Ultrasound Imaging.	Biomedicines	9	720	2021
Yamada M, Saito Y, Yamada S, Kondo H, Hamamoto R.	Detection of flat colorectal neoplasia by artificial intelligence: A systematic review.	Best Pract Res Clin Gastroenterol	52-53	11745	2021
Kawaguchi RK, Takahashi M, Miyake M, Kinoshita M, Takahashi S, Ichimura K, Hamamoto R, Narita Y, Sese J.	Assessing Versatile Machine Learning Models for Glioma Radiogenomic Studies across Hospitals.	Cancers (Basel)	13	3611	2021
Kobayashi K, Hataya R, Kurose Y, Miyake M, Takahashi M, Nakagawa A, Harada T, Hamamoto R.	Decomposing normal and abnormal features of medical images for content-based image retrieval of glioma imaging.	Med Image Anal	74	102227	2021

発表者氏名	論文タイトル名	発表誌名	巻号	ページ	出版年
Kaneko S, Takasawa K, Asada K, Shinkai N, Bolatkan A, Yamada M, Takahashi S, Machino H, Kobayashi K, Komatsu M, Hamamoto R.	Epigenetic Mechanisms Underlying COVID-19 Pathogenesis.	Biomedicines	9	1142	2021
Asada K, Komatsu M, Shimoyama R, Takasawa K, Shinkai N, Sakai A, Bolatkan A, Yamada M, Takahashi S, Machino H, Kobayashi K, Kaneko S, Hamamoto R.	Application of Artificial Intelligence in COVID-19 Diagnosis and Therapeutics.	J Pers Med	11	886	2021
Kuno I, Takayanagi D, Asami Y, Murakami N, Matsuda M, Shimada Y, Hirose S, Kato MK, Komatsu M, Hamamoto R, Okuma K, Kohno T, Itami J, Yoshida H, Shiraishi K, Kato T.	TP53 mutants and non-HPV16/18 genotypes are poor prognostic factors for concurrent chemoradiotherapy in locally advanced cervical cancer.	Sci Rep	11	19261	2021
Mitsuhashi A, Kondoh K, Horikawa K, Koyama K, Nguyen N, Afroj T, Yoneda H, Otsuka K, Ogino H, Nokihara H, Shinohara T, Nishioka Y.	PD-1/PD-L1 blockade mediates anti-angiogenic effects by tumor-derived CXCL10/11 as a potential predictive biomarker.	Cancer Sci	12(12)	4853-4866	2021
Sato S, Chong SG, Upagupta C, Yanagihara T, Saito T, Shimbori C, Belle PS, Nishioka Y, Kolb M.	Fibrotic extracellular matrix induces release of extracellular vesicles with profibrotic miRNA from fibrocytes.	Thorax	76(9)	895-906	2021
Afroj T, Mitsuhashi A, Ogino H, Saijo A, Otsuka K, Yoneda H, Tobiume M, Nguyen TN, Goto H, Koyama K, Sugimoto M, Kondoh O, Nokihara H, Nishioka Y.	Blockade of PD-1/PD-L1 pathway enhances the antigen-presenting capacity of fibrocytes.	J Immunol	206(6)	1204-1214	2021
Kawano H, Koyama K, Nishimura H, Toyoda Y, Kagawa K, Sato S, Naito N, Goto H, Inagaki Y, Nishioka Y.	Development of improved method to identify and analyze lung fibrocytes with flow cytometry in a reporter mouse strain.	Immun Inflamm Dis	9(1)	120-127	2021
Matsumoto S, Ishida S, Araki M, Kato T, Terayama K, Okuno Y	Extraction of protein dynamics information from cryo-EM maps using deep learning	Nat Mach Intell	3	153-160	2021
Ryosuke Shibukawa; Shoichi Ishida; Kazuki Yoshizoe; Kunihiro Wasa; Kiyosei Takasu; Yasushi Okuno; Kei Terayama; Koji Tsuda	CompRet: a comprehensive recommendation framework for chemical synthesis planning with algorithmic enumeration	J Cheminform	12	52	2020
Masuda, N., Murakami, K., Kita, Y., Hamada, A., Kamada, M., Teramoto, Y., Matsumoto, K., Sano, T., Saito, R., Okuno, Y., Ogawa, O., Kobayashi, T.	Trp53 mutation in Krt5-expressing basal cells facilitates the development of basal squamous-like invasive bladder cancer in the chemical carcinogenesis of mouse bladder	Am. J. Pathol.	190(8)	1752-1762	2020
Saito, Y., Koya, J., Araki, M., Kogure, Y.,	Landscape and function of multiple	Nature	582	95-99	2020

発表者氏名	論文タイトル名	発表誌名	巻号	ページ	出版年
Shingaki, S., Tabata, M., McClure, M., Yoshifuji, K., Matsumoto, S., Isaka, Y., Tanaka, H., Kanai, T., Miyano, S., Shiraishi, Y., Okuno, Y., Kataoka, K.	mutations within individual oncogenes				
Ono, F., Chiba, S., Isaka, Y., Matsumoto, S., Ma, B., Katayama, R., Araki, M., Okuno, Y.	Improvement in predicting drug sensitivity changes associated with protein mutations using a molecular dynamics based alchemical mutation method	Sci. Rep.	10(1)	2161	2020
Tanaka, Y., Tamada, Y., Ikeguchi, M., Yamashita, F., Okuno, Y.	System-based differential gene network analysis for characterizing a sample-specific subnetwork	Biomolecules	10(2)	306	2020
Hatae, R., Chamoto, K., Young, Hak, Kim., Sonomura, K., Taneishi, K., Kawaguchi, S., Yoshida, H., Ozasa, H., Sakamori, Y., Akrami, M., Fagarasan, S., Masuda, I., Okuno, Y., Matsuda, F., Hirai, T., Honjo, T.	Combination of host immune metabolic biomarkers for the PD-1 blockade cancer immunotherapy	JCI Insight	5(2)	133501	2020
Iwata, H., Kojima, R., Okuno, Y.	An in Silico Approach for Integrating Phenotypic and Target-Based Approaches in Drug Discovery	Mol. Inform.	39(1-2)	1900096	2020

厚生労働大臣
(国立医薬品食品衛生研究所長) 殿
(国立保健医療科学院長)

機関名 国立研究開発法人
医薬基盤・健康・栄養研究所

所属研究機関長 職 名 理事長

氏 名 米田 悦

次の職員の令和2年度厚生労働科学研究費の調査研究における、倫理審査状況及び利益相反の管理状況については以下のとおりです。

1. 研究事業名 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）

2. 研究課題名 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

3. 研究者名 (所属部局・職名) 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・

サブプロジェクトリーダー

(氏名・フリガナ) 夏目 やよい ・ナツメ ヤヨイ

4. 倫理審査の状況

	該当性の有無		左記で該当がある場合のみ記入 (※1)		
	有	無	審査済み	審査した機関	未審査 (※2)
ヒトゲノム・遺伝子解析研究に関する倫理指針	☒	☐	☒	医薬基盤・健康・栄養研究所	☐
遺伝子治療等臨床研究に関する指針	☐	☒	☐		☐
人を対象とする医学系研究に関する倫理指針 (※3)	☒	☐	☒	医薬基盤・健康・栄養研究所	☐
厚生労働省の所管する実施機関における動物実験等の実施に関する基本指針	☐	☒	☐		☐
その他、該当する倫理指針があれば記入すること (指針の名称:)	☐	☒	☐		☐

(※1) 当該研究者が当該研究を実施するに当たり遵守すべき倫理指針に関する倫理委員会の審査が済んでいる場合は、「審査済み」にチェックし一部若しくは全部の審査が完了していない場合は、「未審査」にチェックすること。

その他 (特記事項)

(※2) 未審査に場合は、その理由を記載すること。

(※3) 廃止前の「疫学研究に関する倫理指針」や「臨床研究に関する倫理指針」に準拠する場合は、当該項目に記入すること。

5. 厚生労働分野の研究活動における不正行為への対応について

研究倫理教育の受講状況	受講 ☒ 未受講 ☐
-------------	---

6. 利益相反の管理

当研究機関におけるCOIの管理に関する規定の策定	有 ☒ 無 ☐ (無の場合はその理由:)
当研究機関におけるCOI委員会設置の有無	有 ☒ 無 ☐ (無の場合は委託先機関:)
当研究に係るCOIについての報告・審査の有無	有 ☒ 無 ☐ (無の場合はその理由:)
当研究に係るCOIについての指導・管理の有無	有 ☐ 無 ☒ (有の場合はその内容:)

(留意事項) ・該当する□にチェックを入れること。
・分担研究者の所属する機関の長も作成すること。

令和3年4月27日

厚生労働大臣
(国立医薬品食品衛生研究所長) 殿
(国立保健医療科学院長)

機関名 地方独立行政法人神奈川県立循環器呼吸器医療センター
所属研究機関長 職 名 所長 田尻 道彦
氏 名

次の職員の令和2年度厚生労働科学研究費の調査研究における、倫理審査状況及び利益相反等の管理については以下のとおりです。

1. 研究事業名 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）
2. 研究課題名 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(19AC5001)
3. 研究者名 (所属部局・職名) 副院長兼呼吸器内科部長
(氏名・フリガナ) 小倉 高志・オグラ タカシ

4. 倫理審査の状況

	該当性の有無		左記で該当がある場合のみ記入 (※1)		
	有	無	審査済み	審査した機関	未審査 (※2)
ヒトゲノム・遺伝子解析研究に関する倫理指針	☒	☐	☒	神奈川県立循環器呼吸器病センター研究倫理審査委員会	☐
遺伝子治療等臨床研究に関する指針	☐	☒	☐		☐
人を対象とする医学系研究に関する倫理指針 (※3)	☒	☐	☒	神奈川県立循環器呼吸器病センター研究倫理審査委員会	☐
厚生労働省の所管する実施機関における動物実験等の実施に関する基本指針	☐	☒	☐		☐
その他、該当する倫理指針があれば記入すること (指針の名称:)	☐	☒	☐		☐

(※1) 当該研究者が当該研究を実施するに当たり遵守すべき倫理指針に関する倫理委員会の審査が済んでいる場合は、「審査済み」にチェックし一部若しくは全部の審査が完了していない場合は、「未審査」にチェックすること。

その他 (特記事項)

(※2) 未審査に場合は、その理由を記載すること。

(※3) 廃止前の「疫学研究に関する倫理指針」や「臨床研究に関する倫理指針」に準拠する場合は、当該項目に記入すること。

5. 厚生労働分野の研究活動における不正行為への対応について

研究倫理教育の受講状況	受講 ☒ 未受講 ☐
-------------	---

6. 利益相反の管理

当研究機関におけるCOIの管理に関する規定の策定	有 ☒ 無 ☐ (無の場合はその理由:)
当研究機関におけるCOI委員会設置の有無	有 ☒ 無 ☐ (無の場合は委託先機関:)
当研究に係るCOIについての報告・審査の有無	有 ☒ 無 ☐ (無の場合はその理由:)
当研究に係るCOIについての指導・管理の有無	有 ☐ 無 ☒ (有の場合はその内容:)

(留意事項) ・該当する□にチェックを入れること。
・分担研究者の所属する機関の長も作成すること。

令和3年 4月 7日

厚生労働大臣
(国立医薬品食品衛生研究所長) 殿
(国立保健医療科学院長)

機関名 国立研究開発法人産業技術総合研究所

所属研究機関長 職 名 理事長

氏 名 石村 和彦

次の職員の令和2年度厚生労働科学研究費の調査研究における、倫理審査状況及び利益相反等の管理については以下のとおりです。

1. 研究事業名 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）
2. 研究課題名 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
3. 研究者名 （所属部局・職名）人工知能研究センター・知識情報研究チーム 研究チーム長
（氏名・フリガナ）高村 大也・タカムラ ヒロヤ

4. 倫理審査の状況

	該当性の有無		左記で該当がある場合のみ記入（※1）		
	有	無	審査済み	審査した機関	未審査（※2）
ヒトゲノム・遺伝子解析研究に関する倫理指針	☐	☒	☐		☐
遺伝子治療等臨床研究に関する指針	☐	☒	☐		☐
人を対象とする医学系研究に関する倫理指針（※3）	☐	☒	☐		☐
厚生労働省の所管する実施機関における動物実験等の実施に関する基本指針	☐	☒	☐		☐
その他、該当する倫理指針があれば記入すること （指針の名称： ）	☐	☒	☐		☐

（※1）当該研究者が当該研究を実施するに当たり遵守すべき倫理指針に関する倫理委員会の審査が済んでいる場合は、「審査済み」にチェックし一部若しくは全部の審査が完了していない場合は、「未審査」にチェックすること。

その他（特記事項）

（※2）未審査の場合は、その理由を記載すること。
（※3）廃止前の「疫学研究に関する倫理指針」や「臨床研究に関する倫理指針」に準拠する場合は、当該項目に記入すること。

5. 厚生労働分野の研究活動における不正行為への対応について

研究倫理教育の受講状況	受講 ☒ 未受講 ☐
-------------	---

6. 利益相反の管理

当研究機関におけるCOIの管理に関する規定の策定	有 ☒ 無 ☐ （無の場合はその理由： ）
当研究機関におけるCOI委員会設置の有無	有 ☒ 無 ☐ （無の場合は委託先機関： ）
当研究に係るCOIについての報告・審査の有無	有 ☒ 無 ☐ （無の場合はその理由： ）
当研究に係るCOIについての指導・管理の有無	有 ☐ 無 ☒ （有の場合はその内容： ）

（留意事項） ・該当する□にチェックを入れること。
・分担研究者の所属する機関の長も作成すること。

令和3年4月8日

厚生労働大臣 殿

機関名 北海道大学

所属研究機関長 職 名 総長

氏 名 寶 金 清 博

次の職員の令和2年度厚生労働科学研究費の調査研究における、倫理審査状況及び利益相反等の管理については以下のとおりです。

1. 研究事業名 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）
2. 研究課題名 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

3. 研究者名 （所属部局・職名） 大学院先端生命科学研究院・教授
（氏名・フリガナ） 西村 紳一郎・ニシムラ シンイチロウ

4. 倫理審査の状況

	該当性の有無		左記で該当がある場合のみ記入（※1）		
	有	無	審査済み	審査した機関	未審査（※2）
ヒトゲノム・遺伝子解析研究に関する倫理指針	☐	☒	☐		☐
遺伝子治療等臨床研究に関する指針	☐	☒	☐		☐
人を対象とする医学系研究に関する倫理指針（※3）	☐	☒	☐		☐
厚生労働省の所管する実施機関における動物実験等の実施に関する基本指針	☐	☒	☐		☐
その他、該当する倫理指針があれば記入すること （指針の名称： ）	☐	☒	☐		☐

（※1）当該研究者が当該研究を実施するに当たり遵守すべき倫理指針に関する倫理委員会の審査が済んでいる場合は、「審査済み」にチェックし一部若しくは全部の審査が完了していない場合は、「未審査」にチェックすること。

その他（特記事項）

（※2）未審査の場合は、その理由を記載すること。

（※3）廃止前の「疫学研究に関する倫理指針」や「臨床研究に関する倫理指針」に準拠する場合は、当該項目に記入すること。

5. 厚生労働分野の研究活動における不正行為への対応について

研究倫理教育の受講状況	受講 ☒ 未受講 ☐
-------------	---

6. 利益相反の管理

当研究機関におけるCOIの管理に関する規定の策定	有 ☒ 無 ☐ （無の場合はその理由： ）
当研究機関におけるCOI委員会設置の有無	有 ☒ 無 ☐ （無の場合は委託先機関： ）
当研究に係るCOIについての報告・審査の有無	有 ☒ 無 ☐ （無の場合はその理由： ）
当研究に係るCOIについての指導・管理の有無	有 ☐ 無 ☒ （有の場合はその内容： ）

（留意事項） ・該当する□にチェックを入れること。
・分担研究者の所属する機関の長も作成すること。

令和3年4月1日

厚生労働大臣
(国立医薬品食品衛生研究所長) 殿
(国立保健医療科学院長)

機関名 国立研究開発法人国立

所属研究機関長 職 名 理事長

氏 名 中釜 斉

次の職員の令和2年度厚生労働科学研究費の調査研究における、倫理審査状況及び利益相反等の管理については以下のとおりです。

- 研究事業名 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）
- 研究課題名 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
- 研究者名 (所属部局・職名) 国立がん研究センター研究所医療 AI 研究開発分野・分野長
(氏名・フリガナ) 浜本 隆二・ハマモト リュウジ
- 倫理審査の状況

	該当性の有無		左記で該当がある場合のみ記入 (※1)		
	有	無	審査済み	審査した機関	未審査 (※2)
ヒトゲノム・遺伝子解析研究に関する倫理指針	☒	☐	☐	国立がん研究センター	☐
遺伝子治療等臨床研究に関する指針	☐	☒	☐		☐
人を対象とする医学系研究に関する倫理指針 (※3)	☒	☐	☐	国立がん研究センター	☐
厚生労働省の所管する実施機関における動物実験等の実施に関する基本指針	☐	☒	☐		☐
その他、該当する倫理指針があれば記入すること (指針の名称:)	☐	☒	☐		☐

(※1) 当該研究者が当該研究を実施するに当たり遵守すべき倫理指針に関する倫理委員会の審査が済んでいる場合は、「審査済み」にチェックし一部若しくは全部の審査が完了していない場合は、「未審査」にチェックすること。

その他 (特記事項)

(※2) 未審査に場合は、その理由を記載すること。

(※3) 廃止前の「疫学研究に関する倫理指針」や「臨床研究に関する倫理指針」に準拠する場合は、当該項目に記入すること。

5. 厚生労働分野の研究活動における不正行為への対応について

研究倫理教育の受講状況	受講 ☒ 未受講 ☐
-------------	---

6. 利益相反の管理

当研究機関におけるCOIの管理に関する規定の策定	有 ☒ 無 ☐ (無の場合はその理由:)
当研究機関におけるCOI委員会設置の有無	有 ☒ 無 ☐ (無の場合は委託先機関:)
当研究に係るCOIについての報告・審査の有無	有 ☒ 無 ☐ (無の場合はその理由:)
当研究に係るCOIについての指導・管理の有無	有 ☐ 無 ☒ (有の場合はその内容:)

(留意事項) ・該当する□にチェックを入れること。
・分担研究者の所属する機関の長も作成すること。

令和3年4月1日

厚生労働大臣 殿

機関名 徳島大学

所属研究機関長 職 名 学長

氏 名 野 地 澄 晴

次の職員の令和2年度厚生労働科学研究費の調査研究における、倫理審査状況及び利益相反等の管理については以下のとおりです。

- 研究事業名 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）
- 研究課題名 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
- 研究者名 （所属部局・職名） 大学院医歯薬学研究部・教授
（氏名・フリガナ） 西岡 安彦・ニシオカ ヤスヒコ

4. 倫理審査の状況

	該当性の有無		左記で該当がある場合のみ記入（※1）		
	有	無	審査済み	審査した機関	未審査（※2）
ヒトゲノム・遺伝子解析研究に関する倫理指針	☐	☒	☐		☐
遺伝子治療等臨床研究に関する指針	☐	☒	☐		☐
人を対象とする医学系研究に関する倫理指針（※3）	☒	☐	☒	徳島大学	☐
厚生労働省の所管する実施機関における動物実験等の実施に関する基本指針	☐	☒	☐		☐
その他、該当する倫理指針があれば記入すること （指針の名称：実施機関における動物実験等の実施に関する基本指針）	☒	☐	☒	徳島大学	☐

（※1）当該研究者が当該研究を実施するに当たり遵守すべき倫理指針に関する倫理委員会の審査が済んでいる場合は、「審査済み」にチェックし一部若しくは全部の審査が完了していない場合は、「未審査」にチェックすること。

その他（特記事項）

（※2）未審査の場合は、その理由を記載すること。

（※3）廃止前の「疫学研究に関する倫理指針」や「臨床研究に関する倫理指針」に準拠する場合は、当該項目に記入すること。

5. 厚生労働分野の研究活動における不正行為への対応について

研究倫理教育の受講状況	受講 ☒ 未受講 ☐
-------------	---

6. 利益相反の管理

当研究機関におけるCOIの管理に関する規定の策定	有 ☒ 無 ☐ （無の場合はその理由： ）
当研究機関におけるCOI委員会設置の有無	有 ☒ 無 ☐ （無の場合は委託先機関： ）
当研究に係るCOIについての報告・審査の有無	有 ☒ 無 ☐ （無の場合はその理由： ）
当研究に係るCOIについての指導・管理の有無	有 ☐ 無 ☒ （有の場合はその内容： ）

（留意事項） ・該当する□にチェックを入れること。

・分担研究者の所属する機関の長も作成すること。

2021年 4月 1日

厚生労働大臣
(国立医薬品食品衛生研究所長) 殿
(国立保健医療科学院長)

機関名 国立大学法

所属研究機関長 職 名 医学研究科

氏 名 岩井 一

次の職員の令和2年度厚生労働科学研究費の調査研究における、倫理審査状況及び利益相反の管理状況については以下のとおりです。

- 研究事業名 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）
- 研究課題名 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
- 研究者名 （所属部局・職名） 医学研究科 人間健康科学系専攻 教授
(氏名・フリガナ) 奥野 恭史 ・ オクノ ヤスシ

4. 倫理審査の状況

	該当性の有無		左記で該当がある場合のみ記入 (※1)		
	有	無	審査済み	審査した機関	未審査 (※2)
ヒトゲノム・遺伝子解析研究に関する倫理指針	☒	☐	☒	医薬基盤・健康・栄養研究所	☐
遺伝子治療等臨床研究に関する指針	☐	☒	☐		☐
人を対象とする医学系研究に関する倫理指針 (※3)	☒	☐	☒	医薬基盤・健康・栄養研究所	☐
厚生労働省の所管する実施機関における動物実験等の実施に関する基本指針	☐	☒	☐		☐
その他、該当する倫理指針があれば記入すること (指針の名称:)	☐	☒	☐		☐

(※1) 当該研究者が当該研究を実施するに当たり遵守すべき倫理指針に関する倫理委員会の審査が済んでいる場合は、「審査済み」にチェックし一部若しくは全部の審査が完了していない場合は、「未審査」にチェックすること。

その他（特記事項）

(※2) 未審査の場合は、その理由を記載すること。

(※3) 廃止前の「疫学研究に関する倫理指針」や「臨床研究に関する倫理指針」に準拠する場合は、当該項目に記入すること。

5. 厚生労働分野の研究活動における不正行為への対応について

研究倫理教育の受講状況	受講 ☒ 未受講 ☐
-------------	---

6. 利益相反の管理

当研究機関におけるCOIの管理に関する規定の策定	有 ☒ 無 ☐ (無の場合はその理由:)
当研究機関におけるCOI委員会設置の有無	有 ☒ 無 ☐ (無の場合は委託先機関:)
当研究に係るCOIについての報告・審査の有無	有 ☒ 無 ☐ (無の場合はその理由:)
当研究に係るCOIについての指導・管理の有無	有 ☐ 無 ☒ (有の場合はその内容:)

(留意事項) ・該当する□にチェックを入れること。
・分担研究者の所属する機関の長も作成すること。

令和3年4月1日

厚生労働大臣 殿

機関名 国立大学法

所属研究機関長 職 名 大学院医

氏 名 熊ノ郷

次の職員の令和2年度厚生労働科学研究費の調査研究における、倫理審査状況及び利益相反等の管理については以下のとおりです。

1. 研究事業名 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）

2. 研究課題名 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

3. 研究者名 （所属部局・職名） 医学系研究科・教授

（氏名・フリガナ） 熊ノ郷 淳・クマノゴウ アツシ

4. 倫理審査の状況

	該当性の有無 有 無	左記で該当がある場合のみ記入（※1）		
		審査済み	審査した機関	未審査（※2）
ヒトゲノム・遺伝子解析研究に関する倫理指針	☐ 有 ☒ 無	☐		☐
遺伝子治療等臨床研究に関する指針	☐ 有 ☒ 無	☐		☐
人を対象とする医学系研究に関する倫理指針（※3）	☒ 有 ☐ 無	☒	大阪大学医学部附属病院	☐
厚生労働省の所管する実施機関における動物実験等の実施に関する基本指針	☐ 有 ☒ 無	☐		☐
その他、該当する倫理指針があれば記入すること （指針の名称： ）	☐ 有 ☒ 無	☐		☐

（※1）当該研究者が当該研究を実施するに当たり遵守すべき倫理指針に関する倫理委員会の審査が済んでいる場合は、「審査済み」にチェックし一部若しくは全部の審査が完了していない場合は、「未審査」にチェックすること。

その他（特記事項）

（※2）未審査に場合は、その理由を記載すること。

（※3）廃止前の「疫学研究に関する倫理指針」や「臨床研究に関する倫理指針」に準拠する場合は、当該項目に記入すること。

5. 厚生労働分野の研究活動における不正行為への対応について

研究倫理教育の受講状況	受講 ☒ 未受講 ☐
-------------	---

6. 利益相反の管理

当研究機関におけるCOIの管理に関する規定の策定	有 ☒ 無 ☐ （無の場合はその理由： ）
当研究機関におけるCOI委員会設置の有無	有 ☒ 無 ☐ （無の場合は委託先機関： ）
当研究に係るCOIについての報告・審査の有無	有 ☒ 無 ☐ （無の場合はその理由： ）
当研究に係るCOIについての指導・管理の有無	有 ☐ 無 ☒ （有の場合はその内容： ）

（留意事項） ・該当する□にチェックを入れること。
・分担研究者の所属する機関の長も作成すること。