

# 日本医療研究開発機構 創薬等ライフサイエンス研究支援基盤事業 事後評価報告書



## I 基本情報

補助事業課題名：（日本語）創薬等ライフサイエンス研究支援基盤事業  
（プログラム名）（英語）Platform Project for Supporting Drug Discovery and Life Science Research

実施期間：平成29年4月1日～令和4年3月31日

補助事業担当者 氏名：（日本語）由良 敬  
（英語）Kei Yura

補助事業担当者 所属機関・部署・役職：  
（日本語）学校法人早稲田大学・理工学術院・教授  
（英語）Professor, Faculty of Science and Engineering, Waseda University

## II 補助事業の概要

補助事業の成果およびその意義等

創薬等ライフサイエンス研究を促進する研究支援とデータサイエンス

・「ワンストップサービス」のシステム構築運用と支援依頼の窓口業務：

BINDSは外部研究者の研究を支援することが目的であることより、PS・POおよびAMEDはBINDSがどのような支援をどの程度行ったかを把握する必要がある。事業開始当初において、PS・POおよびAMEDとの話し合いの結果、本事業を俯瞰する立場の者は支援の全体を、また支援実施者は自身の支援課題を管理する仕組みが必要であることが明確になった。そこでプラットフォーム機能最適化ユニット全メンバーによる協議により、支援申請と支援進捗管理を実施する「ワンストップ窓口」を構築することになり、システム開発は早稲田大学が中心となって行い、窓口業務は東京大学が中心となることが決まった。開発の1年目（2017年度）においては、外部研究者が本事業に支援を依頼する仕組みを構築し、支援前のコンサルティング状況や支援の実施状況をPS・POに提供する仕組みを開発できた。第1四半期にテスト版の運用を開始し、第2四半期の8月30日からコンサルティング申請と支援申請を受け付け承認するワンストップ窓口の運用を開始した。ワンストップ窓口の開始前には、東京大学との密な連携のもとで、徹底的にシステムトラブルの可能性を洗い出し、運用開始後にトラブルが発生することがないように努めた。その結果としてワンストップ窓口の運用を停止することなく、事業終了まで継続的に利用することを実現できた。支援研究進捗状況を管理する仕組みは、第3四半期から本格運用を開始した。また運用開始直後から発生した多少のトラブルを回避することを中心にして、ワンストップ窓口利用電子版マニュアルを作成し、全支援実施者に配付するとともに、最新版を常に広報用ウェブサイト（後述）に掲載した。開発の2年目（2018年度）においては、ユニ

ット間連携が明確になるような仕組みを構築した。また広報用ウェブサイト（後述）から申請依頼ができるように、研究支援メニューと連携する工夫をおこなった。2018年8月7日にはPS・PO会議に出席し、ワンストップ窓口の全機能を解説した。またPS・PO及びAMEDが求める情報は、オンタイムで取得できるように、ワンストップ窓口内にデータ表示またはダウンロードボタンを設置した。支援成果の論文情報や被引用件数情報などは、ウェブ上での提供が難しいため、早稲田大学においてデータ収集と調査を行い提出した。2018年度においては、データの更新やユーザーからの問い合わせ対応とシステムトラブルの対応などが279件発生したが、それらすべてを問題なく解決することができた。その結果1633件の支援案件（テスト申請数件を含む）の情報を把握することができた。開発の3年目（2019年度）は、AMEDとの協議のもと、「企業成果占有利用でコンサルティング終了」の処理を追加し、BINDSの企業利用を促進することができた。BINDSでの成果は公開を原則としているため、企業からの施設利用が進んでいなかったことを受け、コンサルティングのみはBINDSで行い、研究そのものはコンサルティングを担当した研究機関との共同研究で実施できるしくみとした。その結果、「企業成果占有利用でコンサルティング終了」となった案件は2022年4月現在で49件となり、BINDSの企業利用を促進することができた。この間に、147件のトラブル等にも対応した。以上の活動により、2019年度末までに2249件の支援案件（テスト申請数件を含む）の情報を把握することができた。開発の4年目（2020年度）には、より正確な情報を収集する要請に応えるために、毎月1日の深夜零時にその時点での全支援案件の状況をファイルに自動保存するしくみを導入し、毎月初めにAMEDに報告することとした。またCOVID-19禍でのインターネット上ハッカー攻撃増加にともない、ワンストップ窓口の攻撃対応力を測定し、問題となるしくみの改修作業も実施した。年度末においてはAMEDからの要請で、中分子創薬ライブラリーを取り扱う部署をワンストップ窓口内に追加した。この間に、48件のトラブル等にも対応した。以上の活動により、2715件の支援案件（テスト申請数件を含む）の情報を安全に把握することができるようになった。最終年度（2021年度）においては、ウェブ窓口の利用頻度が高くなるとともに、PSP0が所有するワンタイムパスワード装置（後述）が劣化してきたことなどを受け、1年間を通して、73件のトラブルに対応したがそれらをすべて解決した。最終的には総計3166件の支援案件（テスト申請数件を含む）の情報を安全に把握することができ、全支援データを5年間にわたり一元的に管理運用することができた。

ワンストップ窓口では、全支援案件をオンタイム・オンラインで把握できるとともに、情報の漏洩が発生しないようにワンタイムパスワード（トークン）による情報セキュリティ向上に努めた。ワンストップ窓口によって、従来の国家プロジェクトにはなかった新しい情報共有基盤を構築することができた。またCOVID-19禍においても、プロジェクトの情報が共有できるしくみを構築できた。

#### ・過去の有用なデータベース等の継承・維持・管理：

創薬標的候補など有用なタンパク質の発見を支援できると考えられるデータベースを、当該事業における有用なデータベースと定義し、当該プロジェクトの初年度第1四半期に、これまでの事業で整備された既存のデータベースやソフトウェア（解析ツール等）の選別を実施した。初年度第2四半期終了時に継承を完了し、それぞれのデータベースを各機関で公開し、これらデータベースの運営を通して、構造生命科学コミュニティの支援を実施した。

初年度における継承データベース選別の結果、タンパク質物性データベースとしてeSOL、RefoldDB、AutophagyDB、SOSUIを、タンパク質構造DBとしてhGtoP、HOMCOS、SIRDを、タンパク質とリガンドの相互作用データベースとしてNLDB、Het-PDB Navi.を、ゲノムとタンパク質の関連データベースとして、VaProS（事業の途中でVaProS-PDISと改名）、COXPRESdb、mutation@AGlance、ThermusDBを継承することとした。早稲田大学では、VaProS-PDIS、eSOL、RefoldDB、AutophagyDB、ThermusDBを継承した。いずれのデータベースにおいて、継承後もコミュニティからの利用が続いており、毎月500～1000件程度のアクセス検索が行われ続けている。つまり、事業当初にコミュニティが必要とするデータベースを適切に選別でき、維持管理することができたと考えられる。

#### ・本事業で得られた成果の収集・蓄積・管理に関する対応：

支援・高度化の実験で得られた測定データに関して、2017年度に2PBの大型ストレージ（コールドストレージ）を早稲田大学に導入し、これらデータの蓄積を開始した。事業構成員はデータの出し入れを自由に行うことができるが、他人のデータを見ることはできない仕組みとした。プロジェクトでの支援によって稼働しているデータベースのバックアップとして機能させたとともに、研究支援によって生まれた巨大データ（電子顕微鏡測定画像データやゲノムデータなど）の保管場所として機能させた。構造解析ユニットの測定データ、バイオロジカルシーズ探索ユニットのゲノム決定データ、およびインシリコユニットの分子動力学計算データが格納完了した。この仕組みにより各事業実施者が大型データの保存場所に困ることがなくなったとともに、現在トップジャーナルで求められている測定データの保存を実現することができるようになった。また、構造解析ユニットから産出された電子顕微鏡解析データは国際データベース（EMPIAR）に格納されることから、EMPIARの維持管理を実施する大阪大学と協力し、EMPIARのデータのバックアップの役割も担うこととした。大阪大学と早稲田大学との間で、高速ネットワークを利用して巨大データを転送することで、遠隔地でのデータ保存も実現し、自然災害などによるデータ紛失のリスク低減にも貢献できるようになった。以上により2022年4月末現在、約600TBのデータが大型ストレージに保存されている。

#### ・本事業の広報活動：

外部研究者に対するBINDSの広報として、<https://www.binds.jp>に広報ウェブを初年度から立ち上げた。広報ウェブサイトでは、BINDSからのニュース発信、BINDSを構成する研究者（支援を実施する研究者）の紹介、研究支援の受け方、どのような支援が可能かを示した支援メニュー、および今までのBINDSからの成果を掲載した。またCOVID-19に対応し、BINDSからの関連成果を独立のページとして発信した。これらの活動により、BINDS事業に支援依頼が行われる環境を整えた。広報ウェブサイトの開発は早稲田大学で実施し、運営はネットワーク型分担機関の東京大学が実施した。ウェブサイトにおいては、支援の内容を説明するページが重要であることから、このページの情報は頻繁に更新し、ユーザーの意見を取り入れるようにした。支援内容を紹介する支援メニューは、2019年以降毎年度初めに更新をすることで、各支援実施者の最新支援技術を掲載してきた。また軽微な内容変更に関しては、変更が発生するたびに対応した。支援メニューは、そのダイジェスト版印刷体を毎年度初めに作成し、東京大学から全国に配付した。2018年度に実施した事業実施者およびコミュニティーからの情報収集により、本事業のコンサルティングおよび支援の手続きが浸透していないことが明らかになった。そこで研究支援を受ける手続きをわかりやすく説明する解説図「スタートアップキット」をAMEDと協力して作成した。スタートアップキットは広報ウェブサイトに掲載するとともに、BINDSを説明する場面（学会のブースなど）で印刷物を配布し広報に利用した。その結果として、研究支援への申込から支援完了までの流れが明確になり、先に示した通り、BINDSとして多くの支援を実施することができた。広報においては、AMEDが学会年会等で実施するブース展示によるBINDS事業の紹介に協力してきた。早稲田大学のメンバーがブースに立ち、AMEDと協力して事業の内容を来訪者に説明した。特に研究の内容の説明に関しては、研究者がその場に居合わせることが重要であり、研究者による説明のおかげで、その場でコンサルティング申請が行われた場合もあった。

#### ・研究支援内容：

バイオインフォマティクスによる研究支援は、その内容が実験研究者にはわかりにくいため、ワンストップ窓口によるコンサルティングが始まる前に学会年会などの場でプレコンサルティングを実施することが必要であった。事業内でプレコンサルティングは数十件おこなった。プレコンサルティングによって、被支援者と支援者が支援の内容を理解する素地ができた段階で、コンサルティングに入るようにした。またプレコンサルティングの段階で、被支援者が想定している問題が解決する場合も非常に多かった（データベース/ツールの紹介など）。上記の活動を通して、ワンストップ窓口を通じた本格的な研究支援を2件実施した。ひとつはタンパク質生産領域との連携支援であり、一般企業からの支援申請であった。もう1件は、京都大

学からの基礎生命科学研究の支援申請であった。シロイヌナズナの形態形成異常変異体が、特定のタンパク質の相互作用の変化によってもたらされることを京都大学のグループが発見できていたが、それらがどのような複合体を形成するかを調べることができていなかった。そこで、バイオインフォマティクス解析により複合体構造を推定する支援を実施した。推定された複合体は種々の実験結果と整合し、実験結果とあわせて論文発表することができた。以上をもって支援終了としたが、関連する研究課題を引き続きプレコンサルティングしているところである。

#### ・新しい情報を取得する技術の開発：

創薬等支援技術基盤プラットフォーム事業(PDIS)情報拠点から継承した VaProS の高度化を担当した。ユニット内協議において、VaProS-BINDS により多くの機械学習の仕組みを導入することとした。近年の機械学習技術の普及により、その重要性が広く認知されるとともに、ユーザーからの要望もあがってくるようになった。そこで VaProS-BINDS において、分散データベースに存在するデータを収集した後に、収集したデータを再解析し重要なキーワードを抽出する機械学習装置を早稲田大学で組み込んだ。その結果、検索を実施したユーザーが想定していなかったキーワードが登場することがあり、ユーザーに新たな視点をもたらす仕組みができあがった。

#### ・構造解析ターゲティングデータベースの開発：

タンパク質の立体構造を基礎とした創薬研究を支援するためには、ヒトのタンパク質において、どこまで立体構造が明らかになっており、どのタンパク質の立体構造を決定することで、創薬研究にインパクトを与えることができるかを見いだすためのデータベースが必要であると考えた。ヒトゲノムのビルド38版には、選択的スプライシング産物も独立のタンパク質と見なした際には、107,844本のタンパク質アミノ酸配列が登録されていた。この全タンパク質を最新のPDBに対して、jackhmmerで検索ができる仕組みを作った。その結果、88,056本のタンパク質(約81%)にはPDBと類似配列部分が存在し、立体構造がわかることを明らかにした。しかし、多くの場合はタンパク質の全長ではなくその一部分のみであることから、アミノ酸残基単位の詳細な情報が必要であることがわかった。このデータ収集の段階で、特に創薬研究に関係するSLCタンパク質とABCトランスポーターで新たな立体構造情報が必要であることがわかった。SLCトランスポーターは、細胞内への選択的な低分子取り込みに関与する膜タンパク質で、このタンパク質に見られる変異が、疾患の原因になっていることが多く知られている。またこのタンパク質が創薬ターゲットになっている場合も知られている。SLCトランスポーターは、共通祖先ではない異なるタンパク質で形成させるファミリーであるため、SLCトランスポーターは互いに異なる立体構造をしている場合があることが想定される。しかしその全容は明らかになっておらず、構造解析を必要とするタンパク質ファミリーのひとつであると言える。これらの情報をデータベースにまとめ、すでに<http://cib.cf.ocha.ac.jp/slc/>で公開しているとともに、論文として発表済みである。ABCトランスポーターは薬剤の輸送などに関与することがわかっており、ABCトランスポーターに見られる変異が薬物動態に関わってくることが知られている。その機能の詳細を知るためには、立体構造情報が必要であるが、5種類ある膜タンパク質型ABCトランスポーターサブファミリーのうち、ABCDは立体構造が明らかになっていないことが判明した。ABCDトランスポーターは脂質の輸送などにも利用されているタンパク質であり、この立体構造を決定することで、全ABCトランスポーターの立体構造が明らかになることがわかった。これらの情報も論文として発表済みである。

#### ・マイクロアレーデータの高度解析：

前事業(PDIS)において、マウス脳で発現する遺伝子のマイクロアレーデータ解析を支援として行っていた。このデータの解析が簡単に実行することができなかったことと、前事業が終了することになったことを受け、本事業では、マイクロアレーデータの解析もデータベース解析の高度化の一環として、かつもとの支援先との共同研究として、引き続き行うこととした。測定の質があまりよくないマイクロアレーデータから、統計的に問題がない遺伝子発現データを抽出する方法を開発し、マウス脳の発現データに適用したところ、従来

言われていなかったプロテアーゼの発現を見いだすことができた。その遺伝子の発現を実験的に確認することができ、本結果をまとめた論文を Science 誌に掲載することができた。

・ VaProS-SARS-CoV-2 の開発 :

2019年12月に現れた SARS-CoV-2 により世界は大混乱し、2021年4月現在には、その変異ウイルスも加わり、多くのヒトの命が失われている。WHO によって COVID-19 と命名された現象は、どのようにして発生しているかを知ることは、治療薬やワクチンの開発の基礎知識となる。そこでプラットフォーム機能最適化ユニットによって開発されている VaProS-BINDS を、SARS-CoV-2 の解析に適用した。VaProS-BINDS は、ヒトゲノムに様々なデータベースの情報を結合したアプリケーションであるので、ヒトゲノムの代わりに SARS-CoV-2 ゲノム塩基配列の情報をアラインメントし、タンパク質の変異部位は見だし、タンパク質立体構造のどの部位に変異が発生しているかを可視化した。それらの変異部位がタンパク質の機能と関連する可能性を示すために、SARS-CoV-2 のタンパク質と相互作用するタンパク質 (SARS-CoV-2 自身のタンパク質とヒトのタンパク質) の情報、および相互作用する低分子の情報を蓄積し、可能な限り可視化した。これらの情報を集積したデータベースを VaProS-SARS-CoV-2 と命名し、<https://www.vapros-sars-cov-2.org/> から公開することで、SARS-CoV-2 の理解に資する情報提供を行った。

**Research Support with Data Science for Promotion of Drug Development Related Life Science**

In the Platform Operation Optimisation Unit, Waseda university supported the management of the whole operation of the platform as well as the life science research in Japan based on data science technology. Waseda University also developed database and bioinformatics technologies necessary for research support. Specific results of Waseda University follow:

- "One-Stop Service" operation: The aim of BINDS project was to support research activities conducted outside the project; hence achievement of the project should be assessed by the degree of supports. The Optimisation Unit discussed with PS/PS and AMED, and decided to build a web database enabling real time check of the research support status by PS/PO, AMED and the project members. Waseda University, hence built "One-Stop Service" and University of Tokyo took the task of running One-Stop Service. One-Stop Service started its operation in 30th August, 2017 and with frequent update. The stable operation has been achieved by the end of 2017. In 2018, data acquisition buttons for PS/PO and AMED were added and a stable operation has been carried out. At the end of 2018 fiscal year, 1633 supports were handled by One-Stop Service. There happened 279 operation problems, but all of them were appropriately managed. In 2019, a new routine for supporting researches in companies was added which resulted in 49 supports by the end of 2021 fiscal year. By the end of 2019 fiscal year, 147 operation problems have happened, but none of them made the operation stumble, and 2249 supports were handled without problems. Under the COVID-19 situation, the traffic of the internet has been being heavy, but all of the "One-Stop Service" have been kept running and no claim from the users was received. In the end of 2021 fiscal year, 3166 supports were handled without serious issues. We have succeeded in managing the national project that disseminate the opportunity of using the state-of-the-art facilities to all the life science researchers in Japan.

- Bioinformatics Database Operation: The Platform Operation Optimisation Unit was supposed to inherit beneficial databases from previous national projects and to manage their operation. We have selected the beneficial databases and carried on the operation of the following databases; eSOL, RefoldDB, AutophagyDB, SOSUI, hGtoP, HOMCOS, SIRD NLDB, Het-PDB Navi., VaProS-PDIS, COXPRESdb mutation@AGlance.

- Data Storage Operation: The storage and beneficial usage of the stored data are one of the fundamental issues in the era of data science. In 2017, Waseda University introduced 2PB cold storage and started storing data produced in the

project. We built a mechanism to safely and swiftly store the electric data over Internet. Image data of cryo electron microscopy, genome sequence data, and molecular dynamics trajectory data were stored in the cold storage. The realisation of the central storage system freed all the project members from preparing and keeping storage for data in their own lab. By the end of 2021 fiscal year, approximately 600TB were in use.

- Platform Public Relations Operation: PR of the project was required to disseminate the contents of the project and let all the life science researchers in Japan use the project. The Platform Operation Optimisation Unit has launched a website (<https://www.binds.jp/>) and achieved the mission. The development of the website was handled by Waseda University and the operation was carried out by University of Tokyo. The most frequently updated contents of the website were support menu, seminar/lecture information, and COVID-19 research information. In addition to the website, the Optimisation Unit made a brochure to promote support activity of the project.

- Research Support Operation: Waseda University has accomplished a couple of research support during the project. A style of bioinformatics support is quite different from the support by wet-lab experiment support, and hence initial discussion was essential before starting the support. We have conducted many initial discussions and some of the issues have been solved during the initial discussion by introducing web database and/or web tools, which means that some of our effort has not been counted as support. Two discussions have expanded to support, one of them was a bioinformatics support for a start-up company and the other was a support of the lab at Kyoto University. This support resulted in publishing two manuscripts on a prestigious plant biology journal.

- Database Building Operation: The Platform Operation Optimisation Unit has inherited VaProS from PIDS, the previous project and enhanced its utility. The inherited database was named VaProS-PDIS and the enhanced one was named VaProS-BINDS. VaProS-PDIS was a database tool that retrieves results of key word search of many different databases. We enhanced the mechanism by increasing the target databases, introducing post processing of the collected data and introducing graphical presentation of gene/protein expression locations in a cell and in an organism. VaProS-BINDS made much use of machine learning technology and post-processed the collected data to gain new insight. As a result, VaProS-BINDS have had about 500 to 1,000 monthly accesses and consolidated the basis of life science research.