

平成28年度医療研究開発推進事業費補助金
(創薬等ライフサイエンス研究支援基盤事業) 補助事業成果報告書

I. 基本情報

事業名：創薬等ライフサイエンス研究支援基盤事業（創薬等支援技術基盤プラットフォーム事業）
Platform Project for Supporting Drug Discovery and Life Science Research
(Platform for Drug discovery, Informatics, and Structural life science)

補助事業課題名：（日本語）構造生命科学データクラウドの構築運用と高度化
（英語）Development, Promotion and Advancement of Data-Cloud for
Structural Life Science

補助事業担当者 （日本語）情報・システム研究機構 国立遺伝学研究所 特任教授 由良敬
所属 役職 氏名：（英語）Kei Yura, Project Professor, National Institute of Genetics,
Research Organization of Information and Systems

実施期間：平成 28 年 4 月 1 日 ～ 平成 29 年 3 月 31 日

II. 成果の概要（総括研究報告）

補助事業期間最終年度にあたりますので、初年度から最終年度までの全補助事業期間における活動総括概要を作成してください。

（1）情報拠点代表機関および分担機関の協力のもと、タンパク 3000 プロジェクトとターゲットタンパク研究プログラム情報プラットフォーム由来のデータベースとツールを継承し、運用更新することができた。25 件のデータベースおよびツールを遺伝学研究所から <http://p4d-info.nig.ac.jp/useful.html> のもとで公開した。これらのデータベースおよびツールには、プロジェクト期間を通して、総計 800 万件以上の利用があった。主なデータベースとツールは、以下の通りである。FUJI:アミノ酸配列から二次構造、三次構造および天然変性領域を予測するツール; SNPs:SNPs を 1D/3D にマップするツール; HOMCOS:相同複合体の検索・モデリングサーバ; S-VAR:*de novo* のミスセンス変異がタンパク質の機能に及ぼす影響を予測するツール; GTOP:プロテオームデータベース; hGTOP:ヒトに特化したプロテオームデータベース; TP Atlas:ターゲットタンパク研究における 35 課題が扱ったタンパク質の鳥瞰図; CASA:日米欧のアミノ酸配列を網羅

的にアノテーション付けしたデータベース ; PCI:タンパク質化合物相互作用データベース ; Autophagy DB:オートファジーに関するタンパク質を網羅するデータベース ; Trypanosomes DB:トリパノソーマのピリミジン代謝系酵素を中心とする酵素データベース ; eSOL:大腸菌全タンパク質の可溶率と合成量のデータベース ; AT Atlas:ターゲットタンパク研究プログラムの技術開発研究課題である「生産」「解析」「制御」の成果概要集 ; ThermusTB:「丸ごと一匹プロジェクト」から継承した高度好熱菌のゲノムデータベース ; Refold Database:タンパク質の巻き戻し技術を論文から収集したデータベース。

(2) 情報拠点代表機関および分担機関の協力により、構造生命科学データクラウド VaProS の構築に成功し、<http://p4d-info.nig.ac.jp/vapros/>より公開することができた。VaProS は、17 個の異なるデータベースの同時検索を実施し、検索結果を関連も含めて可視化することを可能とした。各分担機関が開発した要素技術を代表機関がとりまとめ、ユーザがひとつのインターフェースで検索できるようにした。さらに、VaProS の普及のために、プラットフォーム構成員が参加した VaProS の利用大会を 2 回開催した。また VaProS 講習会を 8 回開催し、総計 139 名の参加者を得た。VaProS はプロジェクト期間中、1280 万件以上の利用があり、VaProS を謝辞する学術論文も出版された。

(3) 情報拠点代表機関および分担機関の協力により、構造生命科学を必要とするニーズを発掘し支援することができた。外部委員により構成される課題選定協議会の審査を通過した 32 件の支援と、研究者のコンサルティング 75 件を実施した。またプラットフォーム内の下支えとして、プラットフォームの成果を発信するためのウェブサイト、および支援申請を容易に実施できるウェブサイトを分担機関と協力して構築するとともに、プロジェクト全体の支援メニュー冊子を作成した。

(1) Representative organization of the Information Core, with the help of subordinate organizations, kept, managed and further developed the databases and tools developed in the previous national project such as P3K and TPRP. In total 25 databases and tools are available at <http://p4d-info.nig.ac.jp/useful.html>. More than eight million public access were recorded during the project term. The following is the main databases and tools. FUJI: A tool for predicting secondary and tertiary structures and disordered regions from amino acid sequence; SNPs: A tool for mapping SNPs in genomes to 1D/3D of proteins; HOMCOS: A tool for searching for and modelling a protein interactions in 3D; S-VAR: A tool for predicting the effect of de novo missense mutation on protein; GTOP: A database for whole proteome; hGTOP: A database for whole human proteome; TP Atlas: A bird's-eye-view database for proteins studied in the 35 teams at Targeted Proteins Research Program; CASA: A database for complete annotation of worldwide sequence databasess; PCI: A database for protein-ligand interactions; Autophagy DB: A database for a comprehensive set of proteins related to autophagy; Trypanosomes DB: A database for Trypanosome enzymes related to pyrimidine pathway; eSOL: A database for the solubility and the amount of the whole proteins in *Escherichia coli*; AT Atlas: A database for the results of Production Group, Analysis Group and Regulation Group in Targeted Proteins Research Program; ThermusTB: A database for Thermus thermophilus genome derived from Whole-Organism Project; Refold Database: A database for protein refold methodology extracted from literature.

(2) Representative organization of the Information Core, with the help of subordinate

organizations, successfully developed VaProS, structural life science data cloud and made the application available at <http://p4d-info.nig.ac.jp/vapros/>. VaProS realized simultaneous search of 17 different databases and enabled visualization of the whole search results with the relationship amongst each search result. Each subordinate organization developed a part of the application including a database and the representative organization combined them under a single interface that enabled the complete search with a single input from the user. For promotion of VaProS, VaProS Meeting was held twice and VaProS Practice was held eight times. 139 people attended VaProS Practice. VaProS was used more than 12.8 million times during the project and there appeared scientific manuscript acknowledging VaProS during the project.

(3) Representative organization of the Information Core, with the help of subordinate organizations, uncovered needs for structural life science and supported the researchers in need. The core supported 32 research teams, which were approved by the project selection committee and consulted 75 research teams to solve the scientific issues. The core developed and managed the website for promoting the project and its achievement, and developed a web-interface for research support application and edited a booklet for cataloging the available technology for support in the project behind the scenes.

III. 成果の外部への発表

(1) 学会誌・雑誌等における論文一覧（国内誌 0 件、国際誌 2 件）

1. Mizutani, H., Sugawara, H., Buckle, A.M., Sangawa, T., Miyazono, K., Ohtsuka, J., Nagata, K., Shijima, T., Nosaki, S., Xu, Y., Wang, D., Hu, X., Tanokura, M., Yura, K. REFOLDdb: a new and sustainable gateway to experimental protocols for protein refolding. BMC Structural Biology. 2017, 17, 4 (8 pages).
2. Gojobori, T., Ikeo, K., Katayama, Y., Kawabata, T., Kinjo, AR., Kinoshita, K., Kwon, Y., Migita, O., Mizutani, H., Muraoka, M., Nagata, K., Omori, S., Sugawara, H., Yamada, D., Yura, K. VaProS: A Database-Integration Approach for Protein/Genome Information Retrieval. Journal of Structural and Functional Genomics. 2016, 17, 69-81.

(2) 学会・シンポジウム等における口頭・ポスター発表

1. Information Platform from the Targeted Proteins Research Program (TPRP) to Structural Life Science (SLS), ポスター、Hideaki Sugawara, Masafumi Muraoka, Hisashi Mizutani, Kazuho Ikeo, Takao Owayanagi, Takashi Gojobori, Gopi Kuppuraj, Kei Yura, Yoichi Murakami, Kengo Kinoshita, Jun Ohtsuka, Koji Nagata, Junko Sato, Akira Kinjo, Haruki Nakamura, ICSG2013, 2013/7/29, 国内.

2. 構造生命科学データクラウド VaProS による生命科学ビックデータへの挑戦、口頭、由良敬、BMB2015, 2015/12/01, 国内.
3. 配列から機能まで、統合解析のための VapRoS、口頭、池尾一穂、日本農芸化学会 2016 年度大会、2016/03/28, 国内.
4. 構造生命科学データクラウド VaProS による生命科学ビックデータへの挑戦、口頭、由良敬、日本農芸化学会 2016 年度大会、2016/03/28, 国内.
5. A data cloud approach for integration of protein and genome information, ポスター、Takashi Gojobori, Kazuho Ikeo, Yukie Katayama, Takeshi Kawabata, Akira R. Kiniyo, Kengo Kinoshita, Yeondae Kwon, Ohsuke Migita, Hisashi Mizutani, Masafumi Muraoka, Koji Nagata, Satoshi Omori, Hideaki Sugawara, Daichi Yamada, Kei Yura, CBI 学会 2016 年大会、2016/10/25, 国内.
6. 配列から機能まで、統合解析のための VaProS、口頭、池尾一穂、第 39 回日本分子生物学会年会、2016/12/01, 国内.

(3) 「国民との科学・技術対話社会」に対する取り組み

1. 「ゲノムを読む」「生命を形作るタンパク質」、菅原秀明、静岡県富士高等学校フィールドワークス、2013/11/07, 国内.
2. データクラウドを用いた構造生命科学：その開発と利用、由良敬、第 2 回創薬等支援技術基盤プラットフォーム公開シンポジウム、2014/08/27, 国内.
3. データクラウドを用いた構造生命科学：その開発と利用、由良敬、PDIS 最先端セミナー：創薬につなぐ日本の創薬基盤技術、2015/02/04, 国内.
4. データクラウド VaProS:異なる階層のデータをつなぐことで見えてくる新たなシーズ、由良敬、第 3 回創薬等支援技術基盤プラットフォーム公開シンポジウム、2015/10/08, 国内.
5. 情報拠点の成果：構造生命科学データクラウドの構築、中村春木、第 4 回創薬等支援技術基盤プラットフォーム公開シンポジウム、2016/12/07, 国内.

(4) 特許出願

該当なし