

ライフサイエンスデータベース統合推進事業（統合化推進プログラム）

平成 28 年度 研究開発実施報告書

研究代表者 大浪 修一

理化学研究所 生命システム研究センター 発生動態研究チーム チームリーダー

「生命動態情報と細胞・発生画像情報の統合データベース」

§1. 研究実施体制

(1) 研究代表者グループ

- ① 研究代表者： 大浪 修一（理化学研究所生命システム研究センター チームリーダー）
- ② 研究項目：
 - ・生命動態システム科学の最新データの継続的な統合
 - ・RDF/オントロジーを利用したデータベース連携の実現
 - ・データベースツールおよび解析ソフトウェアの充実
 - ・生命動態の定量データを記述するフォーマットの大規模データへの対応と国際連携
 - ・細胞生物学および発生生物学の画像データベースの統合
 - ・画像処理ソフトなどに対応したアプリケーション・プラグインの開発

§2. 研究実施内容

(1) 通期及び本年度の研究開発計画と達成目標

我々は平成 24 年度採択の統合化推進プログラムの研究開発課題として、生命動態システム科学のデータベースの統合を行ってきた。本研究開発課題は、これまで進めてきた我が国の生命動態システム科学のデータベースの統合化の発展的な継続と、我が国の生命動態システム科学のデータベース統合の中長期的な継続を可能にする体制と仕組みの構築を目標とする。加えて、我が国の細胞生物学および発生生物学の画像データベースと生命動態システム科学のデータベースの統合を進める。第2年次である平成 28 年度は以下の研究を計画した。

(1a) 生命動態システム科学の最新データの継続的な統合

前年度に引き続き、国内外の生命動態システム科学のデータベースの統合を進める。生命動態システム科学の 3 推進策(理研 QBiC、CREST 生命動態(1 期)、生命動態推進拠点)が生産し、平成 28 年 3 月までに論文発表され公開が可能になった全ての時空間定量データを、平成 28 年 10 月のメジャーアップデート時に SSBD から公開する。CREST 生命動態の 2 期と 3 期の研究室について、当研究課題の説明と生産予定のデータのヒアリングを実施する。

(1b) RDF/オントロジーを利用したデータベース連携の実現

定量・画像データのメタ情報を効率的に記述できるオントロジーを新たに整備する。SSBD に登録された定量・画像データに対するメタ情報を RDF/オントロジーで記述し、RDF トリプルストアを公開する。

(1c) データベースツールおよび解析ソフトウェアの充実

当初計画を変更し、(2d)データベースのオープンソース化 を平成 28 年度の研究内容に追加したため、本項目の開発は実施しない。

(2a) データベース統合化の中長期的な体制の構築

プロジェクト・グループの委員交代のルール等を整備し、生命動態システム科学の定量データおよび細胞・発生生物学の画像データの統合を推進する持続可能な体制を構築する。

(2b) 生命動態の定量データを記述するフォーマットの大規模データへの対応と国際連携

前年度に引き続き BDML の HDF5 バイナリフォーマットへの対応を行い、完了する。Omics BDML の開発を推進し、上記 BDML の変更箇所を引き継ぐかたちでアップデートを完了する。

(2c) データベース登録作業等の簡素化・効率化

データベースに格納した情報から所定のフォーマットの定量データを自動生成するシステムを構築し、BDML のバージョンアップ時の作業を簡素化する。

(2d) データベースのオープンソース化(新規追加項目)

データベースの持続可能性を模索する中で、データベースのオープンソース化を決定した。これにより、国内外の研究室・研究所が独自に定量データのデータベースを構築し、データの公開・共有を行うことが可能になる。これによりゲノムデータベースにおける DDBJ, GenBank, EMBL のような関係を構築することにより、データの分散共有を推進することが可能になり、ストレージ・データ処理・ダウンロード負荷を分散させることが可能になる。平成 28 年度中に、オープンソースの共有を開始する。

(3a) 細胞生物学および発生生物学の画像データベースの統合

生命動態システム科学の 3 推進策の研究室の代表的な動画像データの統合を完了する。加え

て、細胞生物学会、発生生物学会と、動画像の公開について議論する。さらに、理研 CDB で産出される顕微鏡画像の収集を開始する。

(3b) 画像処理ソフトなどに対応したアプリケーション・プラグインの開発

国内で使用されている主要な画像処理ソフトウェアに対して、SSBD に登録された画像の利用を促進するアプリケーション・プラグインの開発を開始する。OMERO との連携をさらに深める

(2) 本年度の研究開発進捗状況

(1a) 生命動態システム科学の最新データの継続的な統合

時空間定量データの公開:理研 QBiC と CREST 生命動態 1 期、生命動態拠点の定量データの産出の状況を確認し、東京大学飯野研究室(CREST)が計測した線虫の神経細胞の挙動に対する 14 BDML ファイルや理研 QBiC 高橋研究室がシミュレーションで予測した細胞生粘菌内やラットの神経培養細胞内(PC12)における一分子の挙動に対する 4 BDML ファイル等、139 BDML ファイルを新しく公開した。

CREST 生命動態の2期3期へのヒアリング:CREST 生命動態 2 期の研究室について、当研究課題の説明と生産予定のデータのヒアリングを実施した。

レポジトリ利用への対応:論文公開時における、定量データおよび画像データのレポジトリとしての SSBD の利用への対応を開始した。

(1b) RDF/オントロジーを利用したデータベース連携の実現

OME-XML に準拠する RDF/オントロジーの設計とサンプルデータの作成を開始した。

(2a) データベース統合化の中長期的な体制の構築

体制の構築:日本細胞生物学会および日本発生生物学会との協力体制を構築した。また、ABiS との具体的な連携内容について協議を開始した。

データ収集体制の構築:データベースに登録するデータの急増に備え、データ収集の作業を担当する人材を確保した。

システムの更新:データベースを運用しているサーバーの OS を長期サポートがある OS に変更した。さらに、定量データを管理する SSBD プラットフォームと画像データを管理する OMERO プラットフォームをそれぞれ Docker 環境に移行することで、サーバーに障害が起きた場合への柔軟な対応を可能にし、中長期的にみて円滑な稼働が可能な体制を構築した。

(2b) 生命動態の定量データを記述するフォーマットの大規模データへの対応と国際連携

大規模データへの対応:BDML フォーマットの HDF5 バイナリフォーマットへの拡張版である BD5 フォーマットの開発を継続して行った。また BDML フォーマットから BD5 フォーマットに変換するツール(BDML2BD5)を開発し、SSBD データベースで公開している全ての定量データの BD5 への変換を完了した。BD5 開発に伴うメタ情報記述の変更箇所を Omics BDML に適用した。

国際連携:EuBI が主宰する GBI Project の第一回の Exchange of Experience workshop に画像データのマネージメント担当の日本代表として参加し、EuBI との連携を構築した。EuBI のリーダーの 1 人である Jan Ellenberg 博士と H29 年 2 月に打ち合わせを行い、EuBI との連携を強化した。当該分野に関連する主要な国際学会で発表を行い、データベースおよびフォーマットの宣伝と国際連携を模索した。

(2c) データベース登録作業等の簡素化・効率化

システムの更新:BDML の HDF5 バイナリフォーマット拡張に伴い、SSBD データベースシステムも HDF5 フォーマットを基盤としたシステムに変更することにした。このシステム変更により、データベース登録作業等を飛躍的に簡素化・効率化できることが期待される。このため、データベースに格納した情報から所定のフォーマットの定量データを自動生成するシステムの開発は H29 年度に変更し、HDF5 フォーマットを基盤としたデータベースシステムおよび REST API のプロトタイプを開発した。

(2d) データベースのオープンソース化

定量データを管理する SSBD データベースシステムのオープンソース化を行い、GitHub にて公開した (<https://github.com/openssbd/OpenSSBD>)。また、画像を管理する OMERO システムの Docker ファイルを GitHub に日本語での解説付きで公開した (<https://github.com/openssbd/omero-web>)。

(3a) 細胞生物学および発生生物学の画像データベースの統合

画像データの公開:理研 QBiC と CREST 生命動態 1 期、生命動態拠点の画像データの産出の状況を確認し、線虫の神経細胞を計測した共焦点顕微鏡画像、スピニングディスク共焦点顕微鏡画像、落射蛍光顕微鏡画像 (14 dataset)や、マウス由来 iPS 細胞などを計測した位相差顕微鏡画像 (4 dataset)等、6 種類、計 143 dataset の画像データを公開した。

理研 CDB へのヒアリング:理研 CDB の全研究室について、当研究課題の説明と生産予定のデータのヒアリングを実地した。

学会等との連携:日本細胞生物学会、日本発生生物学会との連携体制を構築した。また、ABiS との具体的な連携内容について協議を開始した。

システムの更新:これまで公開したすべての画像に生物種名などのメタ情報を追加し、追加したメタ情報を用いて画像を検索できるシステムを構築し公開した。

(3b) 画像処理ソフトなどに対応したアプリケーション・プラグインの開発

OMERO と SSBD のシステム連携の改善:OMERO Users Meeting 2016 に参加し、最新の OMERO を基盤とする効率的な画像管理法とアプリケーション開発法に関する情報を入手した。OMERO を SSBD データベースに効果的に組み込むために、定期的に要望を送るなど密に連絡を取り合っている。

OMERO と ImageJ 間の連携に関する情報や OME-XML の有効活用法の情報の取得:ImageJ との連携については予定を変更し、H29 年度に実施することとした。OME-XML の情報をより有効に活用できるようにするために、OME-XML に準拠した RDF/オントロジーの開発を開始した ((1b)参照)。

§3. 成果発表等

(3-1) 原著論文発表

- ① 発行済論文数 (国内 (和文) 0 件、国際 (欧文) 3 件)
- ② 未発行論文数 (国内 (和文) 0 件、国際 (欧文) 0 件)
- ③ 論文詳細情報

1. Yukako Tohsato, Kenneth H. L. Ho, Koji Kyoda, and Shuichi Onami, “SSBD: a

- database of quantitative data of spatiotemporal dynamics of biological phenomena”, *Bioinformatics*, vol. 32, No. 22, pp.3471-3479, 2016 (DOI: 10.1093/bioinformatics/btw417).
2. Xian-Hua Han, Yukako Tohsato, Koji Kyoda, Shuichi Onami, Ikuko Nishikawa, and Yen-Wei Chen “Nuclear detection in 4D microscope images of a developing embryo using an enhanced probability map of top-ranked intensity-ordered descriptors”, *IPSJ Transactions on Computer Vision and Applications*, vol. 8, 8, 2016 (DOI: 10.1186/s41074-016-0010-3).
 3. Sihai Yang, Xianhua Han, Yukako Tohsato, Koji Kyoda, Shuichi Onami, Ikuko Nishikawa, and Yenwei Chen, “Phenotype analysis method for identification of gene functions involved in asymmetric division of *Caenorhabditis elegans*”, *Journal of Computational Biology*, vol. 24, No. 5, pp.436-446, 2017 (DOI: 10.1089/cmb.2016.0210).

(3-2) データベースおよびウェブツール等の構築と公開

別紙1参照。

(3-4) その他の著作物(総説、書籍など)

なし。

(3-5) 国際学会および国内学会発表

① 招待講演 (国内 4件、国際 3件)

〈国内〉

1. 大浪修一、線虫 *C. elegans* の受精カルシウム波と受精、第109回日本繁殖生物学会大会、相模原、2016年9月11日-14日
2. 遠里由佳子、京田耕司、ホー・ケネス、大浪修一、BDML/SSBD: 生命動態情報と画像情報の統合データベースの開発、第5回生命医薬情報学連合大会「バイオイメージインフォマティクス: 大規模生命画像データの情報解析に基づく生物学」BoF セッション、東京、2016年09月29日-10月1日
3. 大浪修一、データ駆動型解析による生命科学の革新と可視化の貢献、公開シンポジウム「人工知能と可視化」、日本学術会議、東京、2016年11月30日
4. 大浪修一、バイオイメージ・インフォマティクスが可能にするデータ駆動バイオロジー、BKC バイオインフォマティクス研究会、滋賀県草津市、2017年2月1日

〈国際〉

1. Shuichi Onami, Causality network of biological data, The 9th IEEE Pacific Visualization Symposium (PacificVis 2016), Taipei, Taiwan, Apr 19-22, 2016
2. Shuichi Onami, Data-driven analysis of the mechanism of animal development. CREST Symposium on Big Data Application, Tokyo, Japan, Aug 5, 2016.

3. Shuichi Onami, Data-driven modeling of *C. elegans* embryogenesis. CREST International Symposium on Big Data Application, Tokyo, Japan, Jan 11-12, 2017.

② 口頭講演 (国内 5件、国際 5件)

〈国内〉

1. 京田耕司、遠里由佳子、ホー ケネス、大浪修一、SSBD: 生命動態情報と細胞・発生画像情報の統合データベース、第68回日本細胞生物学会大会(日本ケミカルバイオロジー学会第11回年会合同大会)、京都、2016年6月15-17日
2. 勝間秀人、高山順、遠里由佳子、京田耕司、大浪修一、生命科学研究所の特化型シミュレーターの開発および特化型シミュレーター群の構築、第68回日本細胞生物学会大会(日本ケミカルバイオロジー学会第11回年会合同大会)、京都、2016年6月15-17日
3. 遠里由佳子、岡田初美、高山順、京田耕司、大浪修一、公共データベースの顕微鏡画像を用いた核分裂動態の定量計測とデータ駆動型解析、バイオイメージ・インフォマティクスワークショップ2016、吹田、2016年6月22-23日
4. 京田耕司、遠里由佳子、ホー ケネス、大浪修一、SSBD: 生命動態情報と細胞・発生画像情報の統合データベース、バイオイメージ・インフォマティクスワークショップ2016、吹田、2016年6月22-23日
5. 高山順、大浪修一、線虫 *C. elegans* において精子 TRP-3 チャンネルが受精卵のカルシウム波を引き起こす、日本発生生物学会 秋季シンポジウム 2016、三島、2016年10月19-21日

〈国際〉

1. Ho K.H.L., Tohsato Y., Kyoda K., Onami S.: SSBD: a database for reusing and sharing of quantitative data and microscopy images of biological dynamics. 7th Asia-Pacific *C. elegans* Meeting, Beijing, China, Jun 25-29, 2016
2. Tohsato, Y., Okada, H., Takayama, J., Kyoda, K., Onami, S.: Developing quantitative resource for computational analysis from images of *C. elegans* embryogenesis in a public database Phenobank. The Allied Genetics Conference (TAGC) 2016, Orland, USA, Jul 13-17, 2016.
3. Takayama, J., Onami, S.: The sperm TRP family channel TRP-3 induces a calcium wave in the fertilized oocyte of *C. elegans*. The Allied Genetics Conference 2016, Orlando, USA, Jul 13-17, 2016
4. Kyoda, K., Tohsato, Y., Ho, K.H.L., Onami, S.: SSBD: a database for sharing quantitative data and microscopy images of biological dynamics. Bioimage Informatics Conference 2016, Singapore, Singapore, Oct 10-12, 2016
5. Takayama, J., Okada, H., Onami, S.: The sperm-derived TRP family channel TRP-3 induces a calcium rise in the fertilized oocyte in *C. elegans*. ASCB Annual Meeting 2016, San Francisco, USA, Dec 3-7, 2016.

③ ポスター発表 (国内 10件、国際 8件)

〈国内〉

1. 遠里由佳子、岡田初美、高山順、京田耕司、大浪修一、公共データベースの顕微鏡画像を用いた核分裂動態の定量計測とデータ駆動型解析、バイオイメージ・インフォマティクスワークショップ 2016、吹田、2016年6月22-23日。
2. 京田耕司、遠里由佳子、ホー ケネス、大浪修一、SSBD: 生命動態情報と細胞・発生画像情報の統合データベース、バイオイメージ・インフォマティクス ワークショップ 2016、吹田、2016年6月22-23日
3. 遠里由佳子、ホー・ケネス、京田耕司、大浪修一、SSBD: 細胞・発生画像情報と定量情報の統合データベース、MIRU2016 第19回画像の認識・理解シンポジウム、静岡、2016年8月1-4日
4. ホー・ケネス、遠里由佳子、京田耕司、大浪修一、SSBD: 生命動態情報と細胞・発生画像情報を活用するソフトウェアの開発とその進展、トーゴの日シンポジウム 2016、東京、2016年10月5日-6日
5. 遠里由佳子、ホー・ケネス、京田耕司、大浪修一、SSBD: 生命動態情報と細胞・発生画像情報の統合データベース開発の進展、トーゴの日シンポジウム 2016、東京、2016年10月5日-6日
6. 京田耕司、遠里由佳子、ホー・ケネス、大浪修一、BDML: 生命現象の時空間動態に対する定量データを記述する言語、トーゴの日シンポジウム 2016、東京、2016年10月5-6日
7. 高山順、大浪修一、線虫 *C. elegans* において精子 TRP-3 チャンネルが受精卵のカルシウム波を引き起こす、日本発生生物学会 秋季シンポジウム 2016、三島、2016年10月19-21日
8. 高山順、大浪修一、*C. elegans*の受精における二相性カルシウム波の分子基盤、第54回日本生物物理学会年会、つくば、2016年11月25-27日
9. 勝間秀人、高山順、遠里由佳子、京田耕司、大浪修一、生物科学研究用の特化型シミュレータ群の具現化、第54回日本生物物理学会年会、筑波、2016年11月25-27日
10. 高山順、大浪修一、線虫の精子 TRP-3 チャンネルは「受精のタイムキーパー」である、第39回日本分子生物学会年会、横浜、2016年11月30日-12月2日

〈国際〉

1. Tohsato Y., Ho K., Kyoda K., Onami S.: SSBD: a database of quantitative data and microscopy images of biological dynamics. The 11th Annual OME Users Meeting, Dundee, Scotland UK, May 31- June 1, 2016
2. Tohsato, Y., Okada, H., Takayama, J., Kyoda, K., Onami, S.: Developing quantitative resource for computational analysis from images of *C. elegans* embryogenesis in public databases. Intelligent Systems for Molecular Biology (ISMB) 2016, Orland, Florida, USA, Jul 8-12, 2016
3. Tohsato, Y., Ho, K., Kyoda, K., Onami, S.: SSBD: an open database of quantitative data of biological dynamics. Intelligent Systems for Molecular Biology (ISMB) 2016, Orland,

Florida, USA, Jul 10, 2016

4. Tohsato, Y., Ho, K., Kyoda, K., Onami, S.: SSBD: an open database of quantitative data and microscopy images of biological dynamics. The Allied Genetics Conference (TAGC) 2016, Orland, Florida, USA, Jul 13-17, 2016
5. Kyoda K., Okada H., Onami S.: Digital resources for high-throughput analysis of 3D spatial and temporal cell division dynamics in early embryos. The Allied Genetics Conference 2016, Orlando, Jul 13-17, 2016.
6. Kyoda K., Tohsato, Y., Ho, K., Onami S.: SSBD: a database for sharing and reusing of quantitative data and microscopy images of biological dynamics. The 17th International Conference on Systems Biology (ICSB2016), Barcelona, Sep 16-20, 2016.
7. Onami, S: Analysis of quantitative data of nuclear division dynamics from single gene knockdown embryos for all essential embryonic genes in *C. elegans*. CSHL Meeting on Biological Data Science, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, USA, Oct 26-29, 2016
8. Takayama, J., Okada, H., Onami, S.: The sperm-derived TRP family channel TRP-3 induces a calcium rise in the fertilized oocyte in *C. elegans*. ASCB Annual Meeting 2016, San Francisco, USA, Dec 3-7, 2016.

(3-6) 知的財産権の出願

なし。

(3-7) 受賞・報道等

① 受賞

なし。

② 新聞報道

1. 神戸新聞、理研、米科学誌に発表 卵子活性化 仕組み解明 「受精のメカニズム、理解進む」、2016年4月8日 朝刊

③ その他の成果発表

1. SSBD: 生命動態情報と細胞・発生画像情報の統合データベース、使ってみようバイオデータベース-つながるデータ、広がる世界 (BioDB)、第39回日本分子生物学会年会、横浜、2016年11月30日-12月2日
2. データ駆動型サイエンスで生命を理解する、理研ニュース 2016年12月号、pp.6-9
3. データ駆動型サイエンスで生命を理解する、60秒でわかる？最先端の科学 2016年12月5日、<http://www.riken.jp/pr/videos/60sec/20161205/>

§4. 研究開発期間中に主催した活動(主催したワークショップ等)

年月日	名称	場所	参加人数	概要
2016年4月12日	チーム内ミーティング (非公開)	理化学研究所	4人	研究進捗報告のための ミーティング
2016年5月23日	チーム内ミーティング (非公開)	理化学研究所	4人	研究進捗報告のための ミーティング
2016年6月29日	チーム内ミーティング (非公開)	理化学研究所	4人	研究進捗報告のための ミーティング
2016年8月24日	チーム内ミーティング (非公開)	理化学研究所	4人	研究進捗報告のための ミーティング
2016年11月1日	チーム内ミーティング (非公開)	理化学研究所	4人	研究進捗報告のための ミーティング
2016年12月19日	チーム内ミーティング (非公開)	理化学研究所	4人	研究進捗報告のための ミーティング
2017年1月23日	チーム内ミーティング (非公開)	理化学研究所	4人	研究進捗報告のための ミーティング
2017年3月15日	チーム内ミーティング (非公開)	理化学研究所	4人	研究進捗報告のための ミーティング

以上

別紙1 既公開のデータベース・ウェブツール等

No.	研究開発課題名	正式名称	別称	概要	URL	公開日	状態	分類	関連論文 (本紙論文リストに記載があれば、その番号でも可)
1	生命動態情報と細胞・発生画像情報の統合データベース	SSBD	Systems Science of Biological Dynamics	SSBDは生命動態の定量データとその取得に利用した動画を再利用が容易なかたちで包括的に管理することを目的に構築されたデータベースです。単一分子、細胞、細胞核などのスケールにおける様々な手法により取得された様々な生物種のデータを収集しています。手法としては実験解析およびコンピュータシミュレーションともに対象としています。定量データについては、新規に開発されたBiological Dynamics Markup Language (BDML)というデータフォーマットに従って表現されています。また動画については、OMEROという世界標準になりつつある生命科学分野の動画の管理システムを利用して管理されています。本データベースに収録されているデータセットはダウンロード可能であり、各々の利用許諾にもとづいて利用することができます。	http://ssbd.qbic.riken.jp/	2013/9/2	維持・発展	データベース等	Tohsato, Y., Ho, K. H. L. Kyoda, K., Onami, S.: SSBD: a database of quantitative data of spatiotemporal dynamics of biological phenomena. <i>Bioinformatics</i> 33, 3471-3479 (2016). Kyoda, K., Tohsato, Y., Ho, K. H. L., Onami, S.: Biological Dynamics Markup Language (BDML): an open format for representing quantitative biological dynamics data. <i>Bioinformatics</i> 31, 1044-1052 (2015).
2	生命動態情報と細胞・発生画像情報の統合データベース	WDDD	Worm Developmental Dynamics Database	線虫の初期胚において、野生型と各胚発生必須遺伝子の機能をRNAiで阻害した場合の細胞分裂動態データを、4次元微分干渉顕微鏡(DIC)とコンピュータ画像処理を用いて収集しました。これは動物の発生過程を理解する上で重要な情報となります。	http://so.qbic.riken.jp/wddd/cdd/index.html	2012/5/18	維持・発展	データベース等	Kyoda, K., Adachi, E., Masuda, E., Nagai, Y., Suzuki, Y., Oguro, T., Urai, M., Arai, R., Furukawa, M., Shimada, K., Kuramochi, J., Nagai, E., Onami, S.: WDDD: worm developmental dynamics database. <i>Nucleic Acids Res.</i> 41, D732-737 (2013).