

2022年度 研究開発実施報告

□概要

研究開発課題名	バイオイメージングデータのグローバルなデータ共有システムの構築
開発対象データベースの名称(URL)	SSBD データベース (https://ssbd.riken.jp)
研究代表者氏名	大浪 修一
所属・役職	理化学研究所・生命機能科学研究センター・チームリーダー (2023年3月時点)



□目次

§1. 研究実施体制	3
§2. 研究開発対象とするデータベース・ツール等	4
(1) データベース一覧	4
(2) ツール等一覧	4
§3. 実施内容	5
(1) 本年度の研究開発計画と達成目標	5
(2) 進捗状況	8
§4. 成果発表等	12
(1) 原著論文発表	12
① 論文数概要	12
② 論文詳細情報	12
(2) その他の著作物(総説、書籍など)	12
(3) 国際学会および国内学会発表	12
① 概要	12
② 招待講演	13
③ 口頭講演	13
④ ポスター発表	14
(4) 知的財産権の出願 (国内の出願件数のみ公開)	15
① 出願件数	15
(5) 受賞・報道等	15
① 受賞	15
② メディア報道	15
③ その他の成果発表	15
§5. 研究開発期間中に主催した活動(ワークショップ等)	15
(1) 進捗ミーティング	15
(2) 主催したワークショップ、シンポジウム、アウトリーチ活動等	16

§1. 研究実施体制

グループ名	研究代表者または主たる共同研究者氏名	所属機関・役職名	研究題目
大浪グループ	大浪 修一	理化学研究所・チームリーダー	バイオイメージングデータのグローバルなデータ共有システムの構築

§2. 研究開発対象とするデータベース・ツール等

(1) データベース一覧

【主なデータベース】

No.	名称	別称・略称	URL
1	SSBD:database		https://ssbd.riken.jp/

【その他のデータベース】

No.	名称	別称・略称	URL
1	SSBD:repository		https://ssbd.riken.jp/repository/

(2) ツール等一覧

No.	名称	別称・略称	URL
1	BD-Zarr		https://github.com/openssbd/bdz
2	BDML/BD5		https://github.com/openssbd/BDML-BD5

§3. 実施内容

(1) 本年度の研究開発計画と達成目標

1a) SSBD:repository のデータ登録システムのスケーラビリティの拡大とユーザーエクスペリエンスの向上

SSBD:repository は最小限に絞った必須メタデータとともにデータを公開するサービスである。最小限であっても多岐にわたるメタデータを、実験を実施するデータの産出者や論文の著者から適切に獲得するために、これまで、広く一般に利用されている表計算フォーマット(Microsoft Excel 形式)のテンプレートを用いてきた。この仕組みには、誤りを入力時に発見することが難しい、何らかの手段でのファイルのやり取りが必要になる、ファイルからデータベースへのデータ登録作業が必要になるなどの問題がある。そこで、適切なウェブフォームを用いることで、ファイルのやり取りをなくし、誤りによる作業の手戻りやデータ登録の作業を軽減する。人手を介さないことからスケーラビリティが拡大する。さらに、頻出入力例や利用者ごとの入力履歴から、入力内容を自動的に補完する機能を追加し、ユーザーエクスペリエンスを向上させる。

第 1 年度は、1) SSBD:repository のデータ登録システムに利用者および管理スタッフが利用可能なウェブフォームによるデータ入力システムを試作する。データ入力システムを運用し、利用上の問題点などを明らかにする。2) 試作したデータ入力システムに入力の不足、余分、誤りを発見する仕組みを実装し、使用者に対して問題点を指摘する仕組みを構築する。3) 試作したデータ入力システムに頻出入力例から入力を補完する仕組みを実装し、使用者を補助する仕組みを構築する。

1b) SSBD:repository の利用者の拡大

新学術領域「先端バイオイメージング支援プラットフォーム」などのイメージングを支援するプロジェクト、新学術領域「シンギュラリティ生物学」で開発されている AMATERAS など最先端のイメージング装置を開発し利活用を推進している施設、JST-CREST、AMED-CREST などと連携し、そこで得られた画像データを SSBD:repository に登録してもらうことにより、利用者の拡大をはかる。また、AMED と連携し、AMED が支援する研究から得られたデータのうち、イメージング関連データを SSBD:repository で引き受けることにより、データを効率的に格納・共有するスキームを確立する。

SSBD:repository はオリジナルデータを迅速に公開するサービスであることから、顕微鏡画像や生命動態定量データをオリジナルファイルのまま公開しダウンロードさせているが、顕微鏡画像管理システム OMERO の画像自動登録機能による画像プレビューや、一定の基準の生命動態定量データについて簡単な可視化を行うことで、サービスの利便性を向上させ利用者の拡大を図る。

また、SSBD:repository ではデータセットが公開されるごとに SNS (twitter) でアナウンスを行っており、論文やプレプリントの著者によりリツイートされることで論文とデータセットだけでなく SSBD:repository の宣伝にもなっている。これを Image.sc などの専門的な SNS にも展開することで利用者の拡大を図る。

第 1 年度は、1) 新学術領域「先端バイオイメージング支援プラットフォーム」により支援されているプロジェクトの画像、JST-CREST「多細胞」領域で取得された画像、新学術領域「シンギュラリティ生物学」で開発されている AMATERAS による画像などを SSBD:repository に登録を推奨する取り組みを実施する。2) SSBD:repository に顕

顕微鏡画像管理システム OMERO を導入し、画像の自動登録機能による画像のプレビューを実現する。

1c) SSBD:database のデータ登録システムのスケーラビリティの拡大とユーザーエクスペリエンスの向上

SSBD:database は十分なメタデータとともにデータを公開するサービスである。メタデータは内容が多岐にわたり、また数も多いため、キュレーターによるデータの入力・データ産出者や論文の著者による内容の確認と修正・データベースへのデータの登録のために、これまで、広く一般に利用されている表計算フォーマット (Microsoft Excel 形式) のテンプレートを用いてきた。この仕組みには、オントロジー等の正確な表現が必要なデータの誤りを見逃す、データの確認が煩雑かつデータベース登録後のメタデータの扱いや表示に結びつかない、何らかの手段でファイルのやり取りが必要になる、ファイルからデータベースへの登録作業が必要になるなどの問題がある。そこで、適切なウェブフォームと確認用のウェブサイトを用いることで、ファイルのやり取りをなくし、メタデータの誤った登録を防止する。キュレーターは多数のメタデータを入力することから、画面遷移を最小とし、同じデータの繰り返し入力を削減するユーザインタフェースとする。ファイルのやり取りを削減することと合わせて、作業のスケーラビリティを拡大させる。さらに、認証付きのメタデータ確認用ウェブサイトを用意し、データ産出者や論文の著者にデータベース登録後のイメージを伝えるとともに、連絡先などの非公開情報を安全に確認する手段とすることでユーザーエクスペリエンスを向上させる。

第 1 年度は、1) SSBD:database のデータ登録システムに管理スタッフが利用可能なウェブフォームによるデータ入力システムを試作する。データ入力システムを運用し、利用上の問題などを明らかにする。2) 試作したデータ入力システムに入力の不足、余分、誤りを発見する仕組みを実装し、使用者に対して問題点を指摘する仕組みを構築する。3) 入力済みのメタデータの内容をウェブ上で確認できる仕組みを試作し、利用上の問題などを明らかにする。

2a) リポジトリのための最小限のメタデータの標準化や API の共同開発等を含む、SSBD:repository と BioImage Archive とのデータの共有の実現

BioImage Archive は、Global BioImaging が提唱するバイオイメーキングデータのグローバルなデータ共有システムにおけるデータリポジトリの機能を提供するために、2019 年に EMBL-EBI に設立されたバイオイメーキングデータのデータリポジトリである。SSBD:repository と BioImage Archive の連携については、BioImage Archive の開発チームの代表である EMBL-EBI の Matthew Hartley 博士が 2020 年初頭に EMBL-EBI に着任して以降、Harley 博士と大浪の間で継続的に議論が行われている。両データリポジトリの間でデータを共有するためには、各リポジトリにおいてデータに付与するメタデータの標準化が必須である。BioImage Archive と SSBD データベースは 2019 年よりメタデータの標準化を目指した共同研究を開始しており、2019 年からの議論をまとめた、現時点での推奨されるメタデータに関する論文を 2021 年に発表した (Sarkans et al, Nature Methods 18, 1418-1422, 2021)。本研究開発では、これまでの議論を土台にして両データリポジトリ間のメタデータ標準化の議論を加速し、両データリポジトリが採用する標準メタデータを確立する。さらに、メタデータのデータリポジトリへの登録を支援するユーザー支援システムを共同開発し、メタデータを利用して両データリポジトリより共有されているバイオイメー

ジグデータの横断検索が可能な、統合データ検索システムを構築する。

第 1 年度は、1) SSB:repository のメタデータを、Sarkans et al., 2021 で推奨された REMBI 規格に準拠させる改良を行う。2) SSB:repository のメタデータと BioImage Archive のメタデータの項目や内容の比較を行い、相互に検索が可能な標準メタデータについて議論を行う。

2b) グローバルなデータ共有のためのファイルフォーマット、メタデータ、API の標準化等を含む、グローバルなデータ共有システムの構築

SSBD で使用するデータ共有のためのファイルフォーマットは、バイオイメージング分野の国際的なコミュニティで合意したフォーマットに従うものとする。画像データのフォーマットは、欧州の OME チームを中心に開発が進められている OME-NGFF を採用することとし、共通のメタデータを持つ TIFF, HDF5, Zarr を基盤としたフォーマットで共有する方針である。定量データのフォーマットについては、これまでに BDML, BD5 といったフォーマットを開発してきたが、これらを開発したノウハウを活かして、画像データの次世代フォーマットである OME-NGFF と親和性の高いフォーマットを、国際的なコミュニティの同意を得ながら開発する予定である。これらのフォーマットにより格納されたデータに対する API も、フォーマットと同様に、国際的なコミュニティと共同で開発する予定である。共通のファイルフォーマットをグローバルに採用することにより、API の開発やツール群の開発の効率化が期待される。さらに、SSBD では、国際的なメタデータ標準化を目指したコミュニティ、QUAREP LiMi により登録が推奨されているメタデータを採用する。

第 1 年度は、1) 定量データの次世代フォーマットの初版を開発し、国際的なコミュニティへの提案を行う。2) コミュニティからのフィードバックなど、国際的なコミュニティと連携して、フォーマットの改良版を開発する。

3a) バイオイメージングデータに関する新たな情報解析手法の開発のためのデータ基盤の整備

深層学習を含む情報処理解析技術の開発を促進するために、多次元データへの並列アクセスを可能にする次世代のデータ形式で、画像データおよび定量データを提供する。具体的にはクラウドからネイティブにアクセス可能で、柔軟で包括的なメタデータ構造を持つ Zarr を基盤としたデータ形式で、データを公開することにより、解析ツールからのデータアクセスを高速かつ効率的に行えるようにする。TIFF や HDF5 など確立されたデータ形式で格納されたデータファイルとの相互変換ツールを準備することにより、迅速かつ持続可能なデータ提供をはかる。画像データの次世代データ形式である OME-NGFF においては、既に bioformats2raw, raw2ometiff が準備されており、これらのツールを利用する。定量データについても同様の方針で、変換ツールを準備する予定である。

第 1 年度は、SSBD:database で共有されている画像データの OME-NGFF への変換を開始する。OME-NGFF においては、特にメタデータ記述に関するアップデートが急ピッチで進められているため、それらを注視しつつ、データの変換およびメタデータの更新を行う。

3b) バイオイメージングデータに関する主要な画像情報処理ソフトウェアの収集と実行環境の提供

SSBD では、これまでに画像データと定量データを収集・共有を行なってきた。NBDC 主催の分子生物学会でのブースやバイオイメージングコミュニティにおけるユーザーからの要望の多くは、画像データから定量データを計測する要望であった。

この要望に応えるために、本開発期間には、新たに論文で利用された画像解析プログラムを Notebook 形式で収集・共有する。画像データの種類や目的に応じて、利用される画像処理ソフトウェアは異なる場合があるため、バイオイメージング分野における主要なソフトウェア (例えば、Fiji, CellPose, ilastik など) を利用したプログラムを揃える予定である。これにより、画像データ解析結果を再現したり、改変・改良を行ったりすることにより、新たな画像データへの適用が可能になる。

各種プログラムがどのような画像データを利用し、どのような解析結果が得られるのかを容易に理解できるようにするため、Notebook 形式の実行環境の整備も行う。これは、実行環境を整えることが困難な実験生物学者の要望にも応えることにもつながる。最初に、Google Colab や Binder といった無料で利用できる外部の実行環境を使ったサービスを提供する。これらの環境は、比較的軽い処理に適しており、サービス利用者、提供者双方に負担なく、気軽にプログラムを試すことが可能になる。しかしながら、論文で使用される画像処理には比較的重い処理も含まれることが予想される。このようなケースにも対応するために、独自にサーバを準備して BinderHub などを用いることにより、画像処理等を行うことができる実行環境を提供する。

また、SSBD で共有されている画像データは、深層学習用のトレーニングデータとしての利用が期待される。しかしながら、SSBD で共有されている画像データは最先端の技術で撮影されたものが多く、深層学習に利用するにはデータ数が充分でない場合も予想される。そこで、既存のデータ拡張の手法を用いることにより、データ数を増やすことでトレーニングデータとしての活用を促進する。

第 1 年度は、論文で開発または使用された画像処理ソフトウェア、ワークフローについての情報を収集する。使用されている汎用ソフトウェアや独自に開発されたソフトウェアについての情報を集約・整理し、将来、提供を予定している実行環境の要件の洗い出しを行う。

(2) 進捗状況

1a) SSBD:repository のデータ登録システムのスケーラビリティの拡大とユーザーエクスペリエンスの向上

SSBD:repository へのデータ登録システムに、利用者および管理スタッフが利用可能なウェブフォームによるデータ入力システムを試作した。SSBD:repository は最小限に絞った必須メタデータとともにデータを公開するサービスであり、必須メタデータには、データセットの情報、データ登録者の情報、データに関するプレプリントや論文の情報、データの生産や解析に関わった者の情報などが含まれる。必須メタデータのウェブフォームへの入力不足があった場合の指摘、入力されているが想定される入力フォーマットと異なる場合の余分や誤りに関して、自動的に指摘する仕組みを導入した。さらに、入力内容が適切であるかについて、表計算フォーマットに記述された必須メタデータの誤りを発見するスクリプトを作成し、誤りを指摘する仕組みを導入した。特に、データに関する

プレプリントや論文の情報の入力においては、1c) SSBD:database のデータ登録システムにおいて先行して開発を進めた、登録候補となる論文情報を収集するデータベースシステムと連携し、プレプリントの公開サービスである bioRxiv などの文献情報の入力と管理を省略する仕組みとした。

1b) SSBD:repository の利用者の拡大

JST-CREST「多細胞」領域で取得された画像、新学術領域「シンギュラリティ生物学」で開発されている AMATERAS による画像などの SSBD:repository への登録を推奨する取り組みを実施した。この結果、2022 年 4 月以降に公開された 15 件のうち、5 件が JST-CREST「多細胞」に関連する著者を含む論文のデータ、6 件が新学術領域「シンギュラリティ生物学」に関連する著者を含む論文のデータとなった。また、先端バイオイメージング支援プラットフォーム ABiS による画像解析等の支援を行うためのデータプラットフォームを、理化学研究所データ科学基盤上に構築し、支援者および被支援者に提供する仕組みを導入した。これは ABiS 総括支援・画像解析支援を担当する基礎生物学研究所上野特任教授、加藤輝助教らと共同で構築し運用を開始したものであり、同総括支援における画像データベースの構築支援の一環でもあることから、SSBD:repository の紹介を行っている。さらに、SSBD:repository において顕微鏡画像管理システム OMERO の運用を開始し、画像のプレビューを実現した。特に、SSBD と連携して新たに構築された Azebex データベースでは、SSBD:repository に格納されたオリジナル顕微鏡画像を基盤として用い、顕微鏡画像の拡大・縮小・回転といった画像のプレビュー機能を OMERO から利用することで、独自のメタデータを用いた画像データベースの構築を可能とした。

1c) SSBD:database のデータ登録システムのスケーラビリティの拡大とユーザエクスペリエンスの向上

SSBD:database へのデータ登録システムに、管理スタッフが利用可能なウェブフォームによるデータ入力システムを試作した。SSBD:database は豊富なメタデータとともにデータを公開するサービスであり、メタデータには、データセットの情報、データ登録者の情報、データに関する論文の情報、データの生産や解析に関わった者の情報に加えて、データセットの生命科学的な性質を体系化したオントロジーや、データセットに含まれる空間情報・時間情報の単位などが含まれる。メタデータのウェブフォームへの入力不足があった場合の指摘、入力されているが想定される入力フォーマットと異なる場合の余分や誤りに関して、自動的に指摘する仕組みを導入した。さらに、入力内容が適切であるかについて、表計算フォーマットに記述されたメタデータの誤りを発見するスクリプトを作成し、誤りを指摘する仕組みを導入した。特に、データに関する論文については、キュレーターによるデータの発見や登録のスケーラビリティを向上させるために、登録候補となる論文情報を収集するデータベースシステムを SSBD の一部として新たに構築した。生命科学・医科学の論文データベースである NCBI PubMed および PubMed Central データベースを、SSBD に関連するプロジェクトの助成番号、SSBD に関連するプロジェクトの研究者名と所属組織で検索し、整理のためのタグ情報を付与して蓄積する。蓄積された論文情報は SSBD から直接参照可能であり、キュレーターによる収集作業を省力化するとともに、従来の数十倍の論文を整理して扱うことができるようになった。

2a) SSBD:repository のデータ登録システムのスケーラビリティの拡大とユーザーエクスペリエンスの向上

SSBD:repository のメタデータを、Sarkans et al., 2021 で推奨された REMBI 規格に準拠させる改良を開始した。具体的な手順としては、REMBI で提案されたメタデータの枠組みと、SSBD:repository における必須メタデータおよび SSBD:database におけるメタデータを比較し、不足しているメタデータについて導入が可能かを検討した。その結果、観察対象の遺伝子情報、処置の情報、顕微鏡で撮影した際の対象の情報などが不足していたことから、これらに関するメタデータを SSBD:repository および SSBD:database のメタデータへの導入候補とした。さらに、欧州 BIA (BioImage Archive) との連携に先立ち、欧州 IDR (Image Data Resource) を運営する OME (Open Microscopy Environment) チームとのグローバルなデータ共有に関する共同研究を前倒しで開始することとなった。オンラインでの議論などを経て、2023 年 2 月 26 日から 2023 年 3 月 5 日にかけて SSBD チームが英国・ダンディー市にて共同研究の打ち合わせを行い、上記の遺伝子情報を軸としたデータ検索や関連データの提案などの連携を行い、グローバルなデータ検索システムを立ち上げることに合意した。これにともない、SSBD:database のメタデータについても、SSBD:repository にあわせて REMBI 規格を含む形に拡張することとした。

2b) グローバルなデータ共有のためのファイルフォーマット、メタデータ、API の標準化等を含む、グローバルなデータ共有システムの構築

定量データを記述するための次世代フォーマットとして、Zarr (<https://zarr.readthedocs.io/en/stable/>) を基盤としたデータフォーマットである BD-Zarr を開発した (<https://github.com/openssbd/bdz>)。Zarr は、クラウド上に巨大な多次元配列データを格納するために開発されたストレージフォーマットである。BD-Zarr は、主に画像解析により検出された分子や核、細胞といった物体の代表的な座標データ、それらの時間的繋がりを示したトラッキングデータ、さらには検出した物体に付随する特徴に関するデータを記述するためのフォーマットである。これらのデータは、解析する際に単独もしくは組み合わせて利用するため、それぞれ有機的に連携しつつ個別にデータを格納することが望ましい。そこで、我々は、複数の行列データをそれぞれ連携して格納することが可能な AnnData (<https://anndata.readthedocs.io/en/latest/>) を基盤として、座標データとトラッキングデータ、特徴データを格納した。上半期に、BD-Zarr の初版を完成し、バイオイメージングコミュニティのフォーラムである image.sc (<https://forum.image.sc/>) で提案をおこなった。コミュニティからのいくつかのフィードバックを受けて、下半期に、1) OME-Zarr の表形式データを記述するための仕様への対応、2) トラッキングデータの格納方法の改良をおこなった。OME-Zarr の表形式データを記述するための新たな仕様が、BD-Zarr の初版の提案後に発表されたため、その仕様に準拠するよう BD-Zarr の改良をおこなった。また、BD-Zarr の初版では、トラッキングデータを密な行列で記述していたが、コミュニティからのフィードバックを受けて、疎な行列で記述することに変更した。さらに、Python を基盤としたバイオイメージングデータの可視化・解析ツールである napari (<https://napari.org/stable/>) で、BD-Zarr で記述されたデータを可視化するツールの開発をおこなった。

3a) バイオイメージングデータに関する新たな情報解析手法の開発のためのデータ基盤の整備

深層学習を含む情報処理解析技術の開発および検証を促進するために、SSBD:database で共有されている

画像データを、バイオイメージングデータの次世代フォーマットである OME-Zarr の最新版に変換をおこない、いくつかの代表的なデータを公開した (<https://ssbd.riken.jp/ssbd-ome-ngff-samples>)。OME-Zarr への変換には、`bioformats2raw` (<https://github.com/glencoesoftware/bioformats2raw>) を利用し、ファイルからメタデータを適切に取得できないデータに対しては、手作業によりメタデータの修正を行った。OME-Zarr Validator (<https://github.com/ome/ome-ngff-validator>) を利用して、作成したデータ検証をおこなった。これらのデータは、Vizarr (<https://github.com/hms-dbmi/vizarr>) や AVIVATOR (<https://avivator.gehlenborglab.org/>) などのウェブアプリケーションや `napari` といったツールで可視化することが可能である。これらのデータリソースの共有により、国際コミュニティへの貢献をおこなった (Moore et al. 2023, accepted)。

上記のデータを公開するために、理化学研究所の共同利用計算システムである HOKUSAI SailingShip (SS) を利用した。HOKUSAI SS は、Amazon S3 互換のオブジェクトストレージを利用可能であるため、Zarr フォーマットのデータを外部共有するのに適したシステムである。また、これらのデータの公開・非公開を含めたデータの管理をおこなうために、国立情報学研究所が開発している Gakunin RDM システムを利用した。2024 年度に予定していた 4a) SSB D データベースの運用の理化学研究所共同利用計算システムへの移行、2025 年度に予定していた 4b) SSB D データベースと NII 研究データ基盤システムとのシステムレベルでの連携の確立を、一部前倒して実施した。

3b) バイオイメージングデータに関する主要な画像情報処理ソフトウェアの収集と実行環境の提供

論文で開発または使用された画像処理ソフトウェア、ワークフローについての情報を収集した。国内外で利用されている画像処理ソフトウェアとして、`Fiji` (<https://fiji.sc/>)、`CellPose` (<https://www.cellpose.org>)、`CellProfiler` (<https://cellprofiler.org/>) などが利用されていた。また、独自に開発された画像処理ワークフロー等については、`GitHub` (<https://github.com/>) などのリポジトリでコードが公開されていた。これらのソフトウェアとワークフローは、Python 上で動作させることが可能であり、`Google Colab` (<https://colab.research.google.com/>) や `Binder` (<https://mybinder.org/>) など無償で利用できる外部サービスで実行することが可能である。さらに、`Jupyter Notebook` の実行環境を独自に整備、提供することにより、より柔軟なワークフローの実行が可能になることが期待される。このようなワークフローの実行環境の整備に向けて、それを可能にする `BinderHub` 等のツール群の調査を開始した。

§4. 成果発表等

(1) 原著論文発表

① 論文数概要

種別	国内外	件数
発行済論文	国内(和文)	1件
	国際(欧文)	5件
未発行論文 (accepted, in press 等)	国内(和文)	0件
	国際(欧文)	1件

② 論文詳細情報

1. Ito Eisuke, Ueda Takuya, Takano Ryo, Tohsato Yukako, Kyoda Koji, Onami Shuichi, Nishikawa Ikuko. Phenotype anomaly detection for biological dynamics data using a deep generative model. Proceedings of the 31st International Conference on Artificial Neural Networks (ICANN 2022). Lecture Notes in Computer Science (LNCS) 2022 Sept. 1 3530:432-444. doi: 10.1007/978-3-031-15931-2_36.
2. Soya Shinkai, Hiroya Itoga, Koji Kyoda, Shuichi Onami. PHI-C2: interpreting Hi-C data as the dynamic 3D genome state. BIOINFORMATICS 2022. 38(21):4984-4986, doi: 10.1093/bioinformatics/btac613. Web of Science ID : WOS:000853888100001.
3. Yusuke Azuma, Hatsumi Okada, Shuichi Onami. Systematic analysis of cell morphodynamics in *C. elegans* early embryogenesis. Frontiers in bioinformatics 2023 Mar. 3:1082531-1082531. doi: 10.3389/fbinf.2023.1082531. PMID : 37026092.
4. Towako Hiraki-Kajiyama, Nobuhiko Miyasaka, Reiko Ando, Noriko Wakisaka, Hiroya Itoga, Shuichi Onami, Yoshihiro Yoshihara. An atlas and database of neuropeptide gene expression in the adult zebrafish forebrain. bioRxiv 2023. doi: 10.1101/2023.03.29.534505.
5. Yuki Yamagata, Tsubasa Fukuyama, Shuichi Onami, Hiroshi Masuya. Ontology for Cellular Senescence Mechanisms. bioRxiv 2023 Mar. doi: 10.1101/2023.03.09.531883.
6. Yuki Yamagata, Tatsuya Kushida, Shuichi Onami, Hiroshi Masuya. Homeostasis Imbalance Process Ontology: a Study on COVID-19 Infectious Processes. BMC Bioinformatics 2023 (in print).

(2) その他の著作物(総説、書籍など)

1. 京田耕司, 大浪修一「バイオイメージングデータベースの現状と展望」実験医学別冊 最強のステップ UP シリーズ 空間オミクス解析スタートアップ実践ガイド 2022年12月 pp.154-159 ISBN 978-4-7581-2261-0

(3) 国際学会および国内学会発表

① 概要

種別	国内外	2023	件数
招待講演	国内		5件
	国際		2件
口頭発表	国内		4件
	国際		3件
ポスター発表	国内		11件

種別	国内外	2023	件数
	国際		4件

② 招待講演

〈国内〉

1. 大浪修一: バイオイメージングデータの解析、管理、公開に関する情報科学技術の現状と将来展望. 第45回日本分子生物学会年会, 千葉, 2022年11月30日-12月2日.
2. 大浪修一: SSBD データベースによるグローバルなバイオイメージングデータの共有. 第45回日本分子生物学会年会, 千葉, 2022年11月30日-12月2日.
3. 大浪修一: データベース SSBD、データ共有プラットフォーム OMERO が開くオープンサイエンス. 日本学術会議 公開シンポジウム「異なるモダリティを統合するバイオ計測の最前線と展望」, 東京, 2022年11月15日.
4. 大浪修一: バイオイメージングデータのグローバルなデータ共有システムの構築. トーゴーの日シンポジウム 2022, オンライン, 2022年10月5日.
5. 大浪修一: 先端バイオイメージングデータのデータ駆動解析とオープンサイエンス. 第31回日本バイオイメージング学会学術集会, 吹田, 2022年9月3-5日.

〈国際〉

1. Onami, S.: Image data-driven modeling of developmental systems and open bioimaging DX. IPR x RIKEN (BDR) Symposium 2023 Dive into Data of Life, Toyonaka, Japan, Feb 21-22, 2023.
2. Onami, S.: General overview of Image Data Management WG. Exchange of Experience VII, Montevideo, Uruguay, Sep 14-16, 2022.

③ 口頭講演

〈国内〉

1. 京田耕司, 糸賀裕弥, 王放放, 山縣友紀, 遠里由佳子, 大浪修一: SSBD:repository/database: バイオイメージングデータのグローバルなデータ共有, 第64回日本植物生理学会年会, 仙台, 2023年3月15日.
2. 大浪修一: 理研オープンライフサイエンスプラットフォームの紹介および活動報告. 理研シンポジウム: 理研オープンライフサイエンスプラットフォーム 研究課題成果報告会, 和光, 2023年3月14日.
3. 大浪修一: シンギュラリティ生物学のためのデータサイエンスプラットフォーム: 現状と今後の展望. 新学術領域「シンギュラリティ生物学」成果報告シンポジウム, 東京, 2023年3月10日.
4. 東裕介, 大浪修一: データ駆動型および仮説検証型アプローチによる細胞形態変化の定量解析. 第45回日本分子生物学会年会, 千葉, 2022年11月30日-12月2日.

〈国際〉

1. Ito, E., Ueda, T., Takano, R., Tohsato, Y., Kyoda, K., Onami, S., Nishikawa, I.: Phenotype anomaly detection for biological dynamics data using a deep generative model. The 31st International Conference on Artificial Neural Networks (ICANN2022). Bristol, UK, Sep 6-9, 2022.
2. Azuma, Y., Onami, S.: An image processing and deep learning-based pipeline for quantitative analysis of cellular (morpho)dynamics. 55th Annual Meeting of the Japanese Society for Developmental Biologists. Kanazawa, Japan, May 31-Jun 3, 2022.
3. Kyoda, K., Ho, H.L.K., Itoga, H., Tohsato, Y., Onami, S.: BD5: an open data format for representing quantitative biological dynamics data. 2022 European HDF5 Users Group (HUG). Online, Saint Paul-lez-Durance, France, May 31-Jun 2, 2022.

④ ポスター発表

〈国内〉

1. 糸賀裕弥, 王放放, 山縣友紀, 京田耕司, 遠里由佳子, 大浪修一: SSBD:database / SSBD:repository - global sharing of bioimaging data. 日本生理学会第 100 回記念大会. 京都, 2023 年 3 月 15 日.
2. 東裕介, 大浪修一: 画像処理による細胞動態の定量解析, 新学術領域「シンギュラリティ生物学」成果報告シンポジウム. 東京, 2023 年 3 月 10 日.
3. 糸賀裕弥, 王放放, 山縣友紀, 京田耕司, Miranda Miranda Miguel, 山本春菜, 遠里由佳子, 大浪修一: SSBD:database, SSBD:repository, Data platform and Data portal for Singularity Biology - シンギュラリティ生物学のための SSBD、データプラットフォーム、データポータル取り組み, 新学術領域「シンギュラリティ生物学」成果報告シンポジウム. 東京, 2023 年 3 月 10 日.
4. 京田耕司, 糸賀裕弥, 大浪修一: バイオイメージングデータの次世代フォーマット開発と解析ワークフロー基盤の開発構想. 新学術領域「シンギュラリティ生物学」成果報告シンポジウム, 東京, 2023 年 3 月 10 日.
5. 糸賀裕弥, 王放放, 山縣友紀, 京田耕司, 遠里由佳子, 大浪修一: SSBD:database/SSBD:repository バイオイメージングデータのグローバルなデータ共有, 第 14 回光塾, 大阪, 2022 年 1 月 18 日-19 日.
6. 京田耕司, 糸賀裕弥, 王放放, 山縣友紀, 遠里由佳子, 大浪修一: SSBD:database/repository バイオイメージングデータのグローバルなデータ共有. 第 45 回日本分子生物学会年会, 千葉, 2022 年 11 月 30 日-12 月 2 日.
7. 山縣友紀, 大浪修一, 梶屋啓志: オントロジーに基づく細胞老化プロセスの知識記述モデリング, 第 45 回日本分子生物学会年会, 千葉, 2022 年 11 月 30 日-12 月 2 日.
8. 東裕介, 大浪修一: 細胞のセグメンテーションデータ作成と公開. 第 25 回情報論的学習理論ワークショップ, つくば, 2022 年 11 月 20-23 日.
9. 糸賀裕弥, 王放放, 山縣友紀, 京田耕司, 遠里由佳子, 大浪修一: SSBD:database/repository バイオイメージングデータのグローバルなデータ共有, トーゴの日シンポジウム 2022, オンライン, 2022 年 10 月 5 日.
10. 京田耕司, ホーケネス, 糸賀裕弥, 大浪修一: BD-zarr: 生命現象の時空間動態データを記述するための次世代フォーマットの開発. トーゴの日シンポジウム 2022, オンライン, 2022 年 10 月 5 日.
11. 糸賀裕弥, 王放放, 山縣友紀, 京田耕司, 遠里由佳子, 大浪修一: SSBD:database/repository バイオイメージングデータのグローバルなデータ共有, 第 31 回バイオイメージング学会学術集会, 吹田, 2022 年 9 月 3-5 日.

〈国際〉

1. Yamagata, Y., Onami, S., and Masuya H.: Ontology Development for Knowledge Organization of Cellular Senescence Processes. RIKEN BDR Symposium 2023, Kobe, Mar 7-9, 2023
2. Itoga, H., Wang, F., Yamagata, Y., Kyoda, K., Tohsato, Y., Onami, S.: SSBD:database and SSBD:repository - Global sharing of bioimaging data -, RIKEN BDR Symposium 2023, Kobe, Mar 7-9, 2023
3. Yamagata Y., Onami S., Masuya H.: Knowledge Systematization for Cellular Senescence Processes by Homeostasis Imbalance Process Ontology. The 2022 International Conference of Biomedical Ontology (ICBO 2022), Ann Arbor, MI, USA and Online, September 25-28, 2022.

4. Kyoda, K., Ho, H.L.K., Itoga, H., Onami, S.: BD-zarr: a zarr-based format for representing quantitative data of biosystems dynamics. I2K 2022 (From Images to Knowledge – 2022). Online, May 6-10, 2022.

(4) 知的財産権の出願（国内の出願件数のみ公開）

① 出願件数

種別	件数
特許出願(国内)	0件

(5) 受賞・報道等

① 受賞

該当なし

② メディア報道

該当なし

③ その他の成果発表

該当なし

§5. 研究開発期間中に主催した活動(ワークショップ等)

(1) 進捗ミーティング

年月日	名称	場所	参加人数	目的・概要
2022年4月5日	所内ミーティング(非公開)	オンライン	5人	理研 CBS データリポジトリについて情報収集
2022年4月13日	チーム内ミーティング(非公開)	オンライン	6人	研究進捗報告のためのミーティング
2022年4月27日	チーム内ミーティング(非公開)	オンライン	3人	研究進捗報告のためのミーティング
2022年5月12日	OLSP Workshop 2022	オンライン	20人	理研オープンライフサイエンスプラットフォームプロジェクト主催イベントにて国際連携や理研が開発した主要 DB の連携等について意見交換
2022年5月25日	チーム内ミーティング(非公開)	オンライン	6人	同上
2022年6月23日	チーム内ミーティング(非公開)	オンライン	6人	同上
2022年7月25-26日	Euro-BioImaging Scientific Advisory Board	Truku,Finlands	20人	バイオイメージングの国際連携について意見交換、情報収集
2022年8月5日	チーム内ミーティング(非公開)	オンライン	9人	同上

年月日	名称	場所	参加人数	目的・概要
2022年8月9日	David Osumi-Suherland (EMBL-EBI)との意見交換会(非公開)	理研神戸地区内居室	7人	研究に関する意見交換、情報収集
2022年9月1日	チーム内ミーティング(非公開)	オンライン	5人	データ集積に関するミーティング
2022年9月6日	チーム内ミーティング(非公開)	オンライン	7人	研究進捗報告のためのミーティング
2022年10月18日	チーム内ミーティング(非公開)	オンライン	4人	同上
2022年10月19日	チーム内ミーティング(非公開)	オンライン	8人	同上
2022年11月16日	チーム内ミーティング(非公開)	オンライン	4人	同上
2022年11月22日	チーム内ミーティング(非公開)	オンライン	8人	同上
2022年12月21日	チーム内ミーティング(非公開)	オンライン	8人	同上
2023年2月22日	チーム内ミーティング(非公開)	オンライン	8人	同上
2023年2月28日ー3月4日	OME-SSBD ミーティング	University of Dundee	14人	共同研究のための情報交換、打合せ
2023年3月2日	Matthew Hartley(EMBL-EBI)とのミーティング	London, UK	2人	Euro-BioImaging/ Bioimaging UK Meeting においてバイオイメージングの国際連兼について情報収集、意見交換

(2) 主催したワークショップ、シンポジウム、アウトリーチ活動等

年月日	名称	場所	参加人数	目的・概要
2022年8月9日	BDR Seminar(所内セミナー)	理化学研究所神戸地区セミナールーム・オンライン	30人	David Osumi-Suherland 先生による講演

以上

別紙1 既公開のデータベース・ウェブツール等

No.	正式名称	別称・略称	概要	URL	公開日	状態	分類	関連論文
1	SSBD:database		生物ダイナミクスのための付加価値データベース。1分子、細胞、組織、個体などの生物学的対象の定量データや顕微鏡画像を解析するための豊富なオープンリソースと解析用ソフトウェアツールを提供する。実験と計算シミュレーションの両方から得られたデータが含まれる。	https://ssbd.riken.jp/database/		維持・発展	データベース等	https://doi.org/10.1093/bioinformatics/btw417
2	SSBD:repository		バイオイメージングと生物学的定量データセットを保存・公開するオープンデータアーカイブ。他の研究者がこれらのデータセットにアクセスし、ダウンロードすることができる。	https://ssbd.riken.jp/repository/		維持・発展	データベース等	https://doi.org/10.1016/j.celrep.2020.107814
3	Biological Dynamics Markup Language/BD5	BDML/BD5	BDMLのスキーマと仕様、BD5の仕様について	https://ssbd.qbic.riken.jp/bdml/		維持・発展	ツール等	https://doi.org/10.1093/bioinformatics/btu767 https://doi.org/10.1371/journal.pone.0237468
4	BDML4DViewer		BDML形式で記述された時空間定量データをImageJ上で可視化するためのプラグイン	https://ssbd.qbic.riken.jp/BDML4DViewer/index.html		維持・発展	ツール等	https://doi.org/10.1093/bioinformatics/btu767
5	BD5		BD5形式の時空間定量データの入出力および簡単な解析を行うサンプルコード	https://github.com/openssbd/BD5_samples		維持・発展	ツール等	https://doi.org/10.1371/journal.pone.0237468
6	phenochar		BDML形式で記述された時空間定量データから表現型特徴をオフラインで解析するためのツール	https://ssbd.qbic.riken.jp/phenochar/index.html		維持・発展	ツール等	https://doi.org/10.1093/bioinformatics/btu767
7	BDML2BD5		BDMLからBD5に変換するコンバーター	https://github.com/openssbd/BDML2BD5		維持・発展	ツール等	https://doi.org/10.1093/bioinformatics/btu767 https://doi.org/10.1371/journal.pone.0237468
8	BD5LINT		BD5形式で記述されたデータファイルの検証を行うツール	https://github.com/openssbd/bd5lint		維持・発展	ツール等	https://doi.org/10.1371/journal.pone.0237468
9	SSBD OME-NGFF Samples		SSBDデータベースで共有している一部データに対するome-zarrファイルセットを公開している	https://ssbd.riken.jp/ssbd-ome-ngff-samples/		新規	データベース等	https://doi.org/10.1101/2023.02.17.528834
10	BD-Zarr		BD-Zarrの仕様およびI/Oおよび可視化のサンプルコード	https://github.com/openssbd/bdz		新規	ツール等	