

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）

「生命と環境の

フェノーム統合データベース」

研究開発期間：平成26年4月～平成29年3月

## 研究開発終了報告書

研究代表者：榎屋啓志

(理化学研究所 バイオリソースセンター  
マウス表現型知識化研究開発ユニット  
ユニットリーダー)

## §1. 研究開発実施の概要

本課題では、表現型情報を研究分野横断的に利活用し、新たなイノベーションの情報活用基盤とすることを目標として、【提供データ拡充】、【データ収集技術】、【データ活用技術】の3つのテーマについて研究開発を実施した。

【提供データ拡充】提供するフェノーム情報の質と量の拡大として、各コミュニティの表現型データベースから、約 4,000 系統、表現型にして合計約 134 万データポイントを取集した。これらのデータは、変異アレルや系統/株と関連させ、NBDC 統合化推進プログラムの他の研究開発課題と連携できるような RDF (5,400 万トリプル)とした上で、標準化・統合化・体系化して公開した。RDF データ公開の基盤には、理研メタデータベース (<http://metadb.riken.jp/metadb/front>) を用いたことで、RDF ファイル公開、SPARQL による任意の検索と他の RDF データベースとの連携 (SPARQL エンドポイント)、Web によるデータ閲覧が可能である。当初の計画になかった国際コンソーシアムのマウス網羅的表現型データの提供を受けたことで同一の枠組みを用いて利用できる「表現型ビッグデータ」が形成されつつある。これらのデータは、ポータルサイト J-Phenome (<http://jphenome.info>) よりリンクされている。

【データ収集技術】一般研究者から直接表現型データ登録を受け付けるシステムとして、今後爆発的に増加すると見込まれる「ゲノム編集技術」による多様な生物の変異体の表現型を収集するシステム J-Phenome Genome Editing Organism Registry を開発した。予め用意しておいた語彙 (オントロジー) を画面遷移に沿って選んで行くことで、対象のゲノム改変生物、およびその表現型について、本プロジェクトで設計したデータスキーマに準拠し、かつ適切なオントロジーでアノテーションされた RDF が生成される。このアプリケーションを利用して、今後ゲノム編集コンソーシアムと連携してデータ収集を行っていく。

【データ活用技術】モデル生物表現型と疾患との関連づける可視化ツールとして、器官組織、形質、遺伝子の相同性を通じて、ヒト疾患 (複合表現型) と、モデル生物の表現型 (単純表現型の組み合わせ) の生物学的対応が評価できるような関連化および可視化を行うアプリケーション J-Phenome Disease Model Finder (<http://diseasemodel.riken.jp/>) を開発した。本アプリケーションでは、UberPheno オントロジーのリンクを通じて、疾患語彙と関連する表現型を示すモデル生物 (マウス、ラット、メダカ等) を検索することができる。また、下記の機能を備える。1) ヒト疾患名を記入して、疾患モデル動物の候補を検索する機能。2) 遺伝子オルソログデータを元にヒト疾患研究のモデル動物の候補を検索する機能、3) 部位名から、関係するヒト疾患研究のモデル動物の候補を検索する機能、4) 1)~3) の検索結果に至るプロセスをノードとエッジのグラフ構造で可視化する機能を持っている。

さらにコミュニティ連携を行った。AMED が推進する未診断疾患イニシアチブ (IRUD) のデータベース (慶應大学)、及びシステムバイオロジー研究所 (SBI: <http://sbi.jp/indexJ.html>) とのデータ連携の合意、国際的には、マウス網羅的表現型解析を推進する国際マウス表現型解析コンソーシアム (IMPC: <http://www.mousephenotype.org>)、生物種横断的な表現型情報統合データベース Monarch Initiative (<https://monarchinitiative.org>) との連携を開始した。

## §2. 研究開発実施体制

### 1. 研究グループ

#### (1) 研究代表者グループ

##### 人員構成

種別	氏名	所属機関	役職	研究開発項目	参加時期
○	梶屋 啓志	理化学研究所バイオリソースセンター	ユニットリーダー	総括	H26.4～ H29.3
	田中 信彦	同上	開発研究員	システム開発	H26.4～ H29.3
*	高月 照江	同上	テクニカルスタッフ	提供するフェノーム情報の質と量の拡大	H26.4～ H29.3
*	斎藤 実香子	同上	パートタイマー	アノテーション作業、データ管理および変換	H26.4～ H28.3
*	高山 英紀	同上	パートタイマー	アノテーション作業、データ管理および変換	H26.4～ H27.3
*	大島 和也	同上	パートタイマー	アノテーション作業、データ管理および変換	H27.4～ H28.7
*	谷川 紀子	同上	パートタイマー	アノテーション作業、データ管理および変換	H28.4～ H29.3
*	佐藤 道比古	同上	パートタイマー	アノテーション作業、データ管理および変換	H28.3～ H28.12
*	大城 望	同上	派遣	サーバー管理作業	H26.4～ H26.10 H27.4～ H27.11 H28.4～ H28.10
*	矢田 由加里	同上	派遣	アノテーション作業	H26.6～ H27.3

##### 担当項目

本課題の全項目を本グループで行った。

### 2. 有識者会議等

該当なし

### §3. 研究開発の目的、実施内容及び成果

#### 1. 研究開発の背景

##### 本研究開発を実施するに至った経緯

表現型は、生命の理解に極めて重要な要素である。開発当初、ゲノムやパスウェイといった分子レベルの生物種横断的な統合データベースがライフイノベーションを牽引していた一方で、表現型データベースは、特に国内では各生物種で独自に公開されているにすぎず、標準化や横断的な統合が行われていないために、各生物種の研究コミュニティ内で限定的にしか用いられていなかった。本プロジェクトのねらいは、RDF 技術を用いることで、表現型とその関連データの研究分野の垣根を越えた情報共有を実現させることである。これが実現されれば、異なる生物種の表現型情報を組み合わせた情報解析を従来に比べて格段に行いやすくなる。このような統合と活用の実例を積み重ねることにより、生命科学におけるデータ主導のイノベーションを推進することができると考えた。また、多様なデータを統一規格で利用することによる成果の拡大、分野間での技術共有によるシナジー効果、一元的管理によるメンテナンスの低減などが期待できた。

##### 類似または関連するデータベース

生物毎に中心的なデータベースセンター、例えば、マウスでは Mouse Genome Informatics (MGI: <http://www.informatics.jax.org>: ジャクソン研究所)、ラットでは Rat Genome Database (RGD: <http://rgd.mcw.edu>: ウィスコンシン大学)、ゼブラフィッシュでは Zebrafish Model Organism Database (ZFIN: <http://zfin.org>: オレゴン大学)等が、現在に至るまで文献や各国のプロジェクトからのデータ収集を行っている。本研究の開始時点で、これらのデータベースは、各生物における標準オントロジー、とそれらの間のリンクを整備しつつあるが、その動きは混乱を最小化するために緩やかであった。

我が国の各プロジェクトとこれらの国際センターとの連携も十分には行われていなかった。上記のように国際データベースはフェノタイプ情報のフォーマットを統一化しつつあるが、その動きは国内ではほとんど注目されていない。国際データベースへのデータ寄託には独自データベースからのデータ変換が必要だが、労力の問題から消極的な国内データベースも多く、メダカのように国際センターが存在しないものもあった。

ヒトに関しては、国内でもコホート解析等のフェノタイプの遺伝解析プロジェクトが進展していたが、モデル生物とのデータ連結等は進められていない状態であった。国外では、ヒト遺伝学情報の収集と、他種の生物および細胞系を用いた解析モデル系は疾患と創薬研究の両輪と考えられており、例えば、マウスの表現型解析プロジェクトである International Mouse Phenotyping Consortium (IMPC: <http://www.mousephenotype.org>: 国際コンソーシアム及び、NIH ファンディング)と 1000 人ゲノムプロジェクト (<http://www.internationalgenome.org>: The International Genome Sample Resource コンソーシアム)との情報連携が模索されている状態であった。

##### 研究開発開始以降、背景状況の変化

本研究開発を開始して以降、特にヒト疾患と関連させた表現型データの利用の発展があった。標準病名を共有する努力が従来から行われてきたが、これに加えて Human Phenotype Ontology (HPO)が注目されるようになった。医学分野において、表現型は病名よりもさらに細かい「病態」や「検査結果」に対応する。病名は医師の診断により決定するが、HPO を用いることで、根拠となった病態や検査結果の情報共有が可能である。例えば、遺伝的な希少疾患には、「未診断疾患」と呼ばれるものがあり、症状が明確であっても病名がつかず、治療法開発が遅れるなどの問題があった。HPOを用いた情報共有はそこに光をさすものであり、国際希少疾患コンソーシアム (International Rare Diseases Research Consortium: IRDiRC) や国際未診断疾患ネットワーク (Undiagnosed Disease Network International) UDNI では、HPO を用いた症状の共有を進めている。国内では、国立研究開発法人日本医療研究開発機構 (AMED) では、未診断疾患イニシアチブ (Initiative on Rare and Undiagnosed Diseases: IRUD) を通じて、HPO を用いた

全国レベルでの希少疾患関連医療機関の情報ネットワークの作成を開始した。

また、生物種横断的な表現型情報統合では、米英豪の研究チームによる大規模表現型統合プロジェクト Monarch Initiative (URL: <https://monarchinitiative.org>) が開始されて、MGI, RGD, ZFIN を始めとするモデル生物データベースから集めた表現型データを統合し、疾患との関連性から検索できるようなデータベースを公開するとともに、IRDIRCとも連携している。マウス分野では、実験解析レベルで標準化した表現型解析を行う International Mouse Phenotyping Consortium (IMPC: URL)が、Phase 2 に移行し、Phase 1 における大規模データを公開した。前述の Monarch Initiative は、IMPC とも連携している。

以上のような背景状況の変化に応じて、本課題は、国際的、及び国内の表現型データ連携の枠組みとの連携を行うような戦略の変更が必要となった。本研究グループは元々IMPCのメンバーであったが、これに加えて IRUD データベース及び Monarch Initiative との連携の合意、Monarch Initiative で開発された表現型統合データライブラリの利用、IMPC 大規模表現型データの発信において本課題が RDF 版データの作成／公開の役割を担う等の実施内容変更を行った。

また、RDF 技術は生命科学分野において着実に拡がりを見せており、2013 年に Uniprot から始まった EBI RDF platform では、EBI の 7 つのデータセットが公開されている。特に 2016 年に Ensembl のデータが公開されたことで、本課題の表現型データと、公共ゲノムデータの連携が容易に行えるようになった。

## 2. 研究開発対象のデータベース・ツール

### (1) データベース

#### 主要なもの

正式名称	略称	概要
J-Phenome	なし	遺伝子の多様性の結果として現れる生命の表現型情報を、モデル動物(マウス、ラット、ゼブラフィッシュ、メダカ)、ゲノム編集研究など、幅広い研究コミュニティから収集し、研究分野の垣根を超えて標準化・統合化・体系化して公開するデータベースへのリンクのポータルサイト。

#### 上記以外のもの

正式名称	略称	概要
BioResource Schema	BRS	本課題で扱う、表現型とゲノムのバリエーションのデータの関係性を記述するための RDF スキーマ。(J-Phenome の一部としてリンク)
NBRP メダカ表現型メタデータ	なし	NBRP メダカ系統の持つ遺伝子の変異や表現型を公共データやオントロジーにリンクして整理。(J-Phenome の一部としてリンク)
IMPC RDF データ	なし	国際マウス表現型解析コンソーシアム(IMPC)のマウス網羅的表現型データ (J-Phenome の一部としてリンク)
NBRP ラット表現型メタデータ	なし	NBRP ラット系統の持つ遺伝子の変異や表現型を公共データやオントロジーにリンクして整理。(J-Phenome の一部としてリンク)
遺伝研 マウス表現型データベースメタデータ	なし	国立遺伝学研究所、マウス表現型データベースのデータを公共データやオントロジーにリンクして整理。(J-Phenome の一部としてリンク)
ACGG-DB: ノック	なし	Asian Community of Glycoscience and

アウトマウスを用いた機能糖鎖科学データベースのメタデータ		Glycotechnology-Database (ACGG-DB)の一部である「KO Mouse DB」(ノックアウトマウスを用いた機能糖鎖科学データベース)の RDF メタデータ。統合化推進プログラムの成松チームと連携して作成。(J-Phenome の一部としてリンク)
遺伝研 ゼブラフィッシュ表現型データ	なし	国立遺伝学研究所、ゼブラフィッシュトラップ遺伝子データベース zTrap のを公共データやオントロジーにリンクして整理。(J-Phenome の一部としてリンク)
Japanese Cre Resouce Expression Database	JCRED	コンディショナルノックアウトマウスにおける遺伝子不活性化部位及びステージを決定するための、Cre-driver マウスの情報共有のためのデータベース (J-Phenome の一部としてリンク)
Rare disease vocabulary in Japanese)	RDVJ	厚生労働省指定難病と難病オントロジーORDO との関連性データからなるオントロジー (J-Phenome の一部としてリンク)
BRC マウスリソース表現型メタデータ	なし	理研バイオリソースセンターから提供されるマウス系統の持つ遺伝子の変異や表現型が公共データやオントロジーにリンクして整理されている。(J-Phenome の一部としてリンク)
BRC 細胞リソース特性メタデータ	なし	理研バイオリソースセンターから提供される培養細胞株特性データ。(J-Phenome の一部としてリンク)
JCM 微生物リソース特性メタデータ	なし	理研バイオリソースセンターから提供される微生物株について、株としての情報を特性や表現型情報を含めて提供しています。(J-Phenome の一部としてリンク)
シロイヌナズナフェノームデータベース	なし	文献キュレーションによって収集したシロイヌナズナフェノーム情報および理研内で開発されたシロイヌナズナのバイリソース(変異体)のフェノームを統合。(J-Phenome の一部としてリンク)

## (2) ツール等

正式名称	略称	概要
J-Phenome Disease Model Finder	なし	異常の起こる器官組織、形質、遺伝子の相同性を通じて、ヒト疾患(複合表現型)と、モデル生物の表現型の生物学的対応を可視化するアプリケーション。
J-Phenome Genome Editing Organism Registry		ゲノム編集生物を対象として一般研究者から直接表現型データ登録を受け付けるシステム。
Stanza for Paging Table		ページング付きテーブル型汎用スタンザ。公開エンドポイントの RDF データを、任意のテーブル型で可視化する汎用ツール。

※データベース、ツールの詳細は別紙参照。

### 3. 達成目標及び実施計画

#### (1) 当初の実実施計画・達成目標

本課題では、表現型情報を研究分野横断的に利活用し、新たなイノベーションの情報活用基盤とすることを目標として、【提供データ拡充】、【データ収集技術】、【データ活用技術】の3つのテーマについて研究開発を計画した。

##### 1)【提供データ拡充】提供するフェノーム情報の質と量の拡大

第1期における、マウス系統、細胞株、微生物株の表現型情報の更新に加え、表現がtを、遺伝子発現などの「分子表現型」、シンプルな計測結果等の「単純表現型」、複数のフェノタイプが組み合わさった「複合表現型」の階層に分類し各コミュニティのデータベースから収集する。これらの収集に際して、各データベースおよび、NBRP データベースと、各プロジェクトへ働きかけてデータ生産レベルでの議論を行うなどの協力関係を確立する。また、NBDC の各グループとも連携し、各グループの情報と RDF レベルで相互にリンクされるようにする。

収集対象データは以下である。

分子フェノタイプ:

マウス Cre ドライバー系統マーカー遺伝子の発現情報:約 100 系統

ゼブラフィッシュのジレントラップ系統マーカー遺伝子の発現情報:約 500 系統

単純フェノタイプ:

マウス網羅的表現型解析データ:約 100 系統

ラット網羅的表現型解析データ:約 1,000 系統

メダカ表現型:約 500 系統

複合フェノタイプ

疾患オントロジー:約 5,000 疾患(リンク作成用の参考データとして整備)

国際的な標準病名 ICD10:約 12,000 疾患(リンク作成用の参考データとして整備)

(合計 2,200 系統 17,000 疾患)

平成 26 年度は、各データベースとの連携の大枠を確立し、データインポートのワークフローの確立と、RDF 化に必要な各オントロジーの選定や、基盤として必要なマッピングデータの作成を行う。

平成 27 年度には、データの RDF 化と公開を行う。データは整備できたものから順次公開し、平成 27 年度中に、上記全ての RDF データを、NBDC あるいは、理研の RDF レポジトリからの公開を目指す。

平成 28 年度には、各データのゲノム情報や、NBDC 内各プロジェクトとの連結を行い、そのデータを NBDC あるいは、理研の RDF レポジトリから公開する。

##### 2)【データ収集技術】生物学者向けの表現型データ入力システム開発

上記のデータベースからの直接のデータ収集とは別に、一般研究者から直接表現型データ登録を受け付けるシステムを開発する。多様な生物の遺伝子改変内容と表現型を、オントロジー語彙を選択することで簡単に登録できるようにする。今後爆発的に増加すると見込まれる「ゲノム編集技術」による多様な生物の変異体の表現型を主な収集対象とする。ゲノム編集コンソーシアムにおいて、データ登録によるコミュニティ活性化等のインセンティブが働くように、他サイトから引用できるデータ閲覧インターフェースを設計する。

平成 26 年度は、ゲノム編集コンソーシアムとの協議を行い、ユーザーが選択肢と使用するオントロジーの収集や統合化の要件の策定など、仕様調査と設計を行う。平成 27 年度は、プロトタイプの作成およびゲノム編集コンソーシアムに限定公開し、要望の受付を行う。平成 28 年度は、さらなるアプリケーションの改変と公開、データ収集を行う。

##### 3)【データ活用技術】モデル生物表現型と疾患との関連づけと可視化

これまで利用者側で結び付けにくかった異種データの間新たな結合として、異常の起こる器官組織や、形質、遺伝子の相同性を通じて、ヒト疾患(複合表現型)と、モデル生物の表現型(単純表現型の組み合わせ)の生物学的対応が、より高い信頼性をもって評価できるような関連化および可視化方法を開発する。さらに、この推論アルゴリズムを用いて、疾患研究者が、大量のデータに基づいて、実験解析に最適なモデル系を探ることができるようなアプリケーション開発を行う。

平成 26 年度は、アルゴリズムの基本設計とオントロジーの収集と検証を行う。平成 27 年度は、各種データの連結の検証作業と、表現型センターやゲノム編集コンソーシアムからも利用できる Web アプリケーションの設計を開始する。平成 28 年度は、Web アプリ開発とその公開を行い、データの収集を開始する。

## (2) 期間中に追加・削除・変更した実施計画・達成目標

本計画における上記 3 つのテーマ及びその他に分けて記述する。

### 1)【提供データ拡充】提供するフェノーム情報の質と量の拡大 ・追加

#### ● Rare-disease vocabulary in Japanese (RDVJ)

厚生労働省指定難病名は、ある程度のカテゴリを示すものの、必ずしも公共の疾患オントロジーと対応していない。そこで、IRDIRC のデータベース Orphanet (<http://www.orpha.net>)で開発された希少疾患オントロジーORDO との関連性のオントロジー作成を追加した。

#### ● IMPC RDF データ

計画では、理研 BRC の保有するマウス網羅的表現型解析データ:約 100 系統の収集を予定していたが、理研 BRC が参画する国際マウス表現型解析コンソーシアム (International Mouse Phenotyping Consortium: IMPC) の国際連携を通じて解析された各遺伝子の機能をノックアウトしたマウスの 120 万件に及ぶ表現型データを RDF 化して公開することとした。

IMPC は、マウス全遺伝子の機能欠失の表現型を世界共通の基準で解析し、そのデータとマウスを世界の研究者に提供することを目的として、2011 年に開始した国際共同開発プロジェクト。米国国立衛生研究所、英国医学研究評議会、欧州委員会など、世界の有力機関・組織で構成されている。現在、日本の理研バイオリソースセンターを含め、13 か国 18 研究施設が参画している。

本データは、約 3,000 種類の変異(ノックアウト)系統の血液検査、血圧、行動、形態など多項目にわたる表現型解析結果が含まれている。本データの提供を受けたことで、公開データは、表現型データとして 120 万を超え、「表現型ビッグデータ」と呼べる規模となっている。

### ・削除

#### ● 理研 BRC の保有するマウス網羅的表現型解析データ:約 100 系統

上記 IMPC データに本データが含まれるため、削除した

#### ● 国際的な標準病名 ICD10:約 12,000 疾患(リンク作成用の参考データとして整備)

疾患との関連性データに Monarch initiative データベース由来の Disease Ontology 中心のデータを使用することになったため、当初の予定である ICD10 は、DO からの公共データを用いた間接的なリンクも利用可能であることから直接は使用しないことにした。

### 2)【データ収集技術】生物学者向けの表現型データ入力システム開発 特に変更なし



### 3)【データ活用技術】モデル生物表現型と疾患との関連づけと可視化

#### ・追加

- ページング付きテーブル型汎用スタンザ(Stanza for Paging Table)の開発  
J-Phenome Disease Model Finderに加え、ユーザーの要望に応じたデータの可視化が必要となり、SPARQL 検索式を与えることで、任意のテーブル形式でデータを可視化することができるツールを開発することとした。

### 4)その他

#### ・追加

- データベース連携の交渉  
RDF でデータを整備したことで、他の RDF 基盤データベースとの連携が比較的容易になった。データの利用率を向上させるため、本課題で整備したデータを外部新規ユーザー実際に利用しているシステムに組み込んで連携するための交渉を行うこととした。

## 4. 実施内容

### (1) 実施内容

#### 【提供データ拡充】提供するフェノーム情報の質と量の拡大(高月、斎藤、高山、大島、谷川、大城、矢田、柗屋)

コミュニティの表現型データベースから、約 4,000 系統(目標 2,200)、表現型にして合計約 130 万の表現型データを収集し、RDF 化および公開を行った。詳細を表 1 に示す。RDF データ公開の基盤には、理研メタデータベース(<http://metadb.riken.jp/metadb/front>)を用いた。本データベースを用いることで、RDF ファイル公開、SPARQL エンドポイントの公開、Web によるデータ閲覧が可能となった(図 1)。

表 1: RDF として公開したデータセットのリスト及びその概要

カテゴリ	データセット名	データ提供者	URL	表現型データポイント数	総トリプル数	担当者
スキーマ	バイオリソース記述 RDF スキーマ	(本プロジェクト)	<a href="http://metadb.riken.jp/metadb/db/bioresource_schema">http://metadb.riken.jp/metadb/db/bioresource_schema</a>	—	531	柗屋
単純表現型	NBRP メダカ表現型メタデータ	基生研	<a href="http://metadb.riken.jp/metadb/db/NBRP_medaka">http://metadb.riken.jp/metadb/db/NBRP_medaka</a>	581	12,268	斎藤、矢田
単純表現型	IMPC RDF データ	IMPC	<a href="http://metadb.riken.jp/metadb/db/IMPC_RDF">http://metadb.riken.jp/metadb/db/IMPC_RDF</a>	1,225,386	53,907,630	柗屋、高山
単純表現型	NBRP ラット表現型メタデータ	京大	<a href="http://metadb.riken.jp/db/NBRP_rat">http://metadb.riken.jp/db/NBRP_rat</a>	70,927	655,177	高月
単純表現型	遺伝研 マウス表現型メタデータ	遺伝研	<a href="http://metadb.riken.jp/db/Nig_consomic_mouse">http://metadb.riken.jp/db/Nig_consomic_mouse</a>	2,296	26,341	高月
単純表現型	ACGG-DB: ノックアウトマウスを用いた機能糖鎖科学データベースのメタデータ	統合化推進プログラム成松チーム	<a href="http://metadb.riken.jp/metadb/db/Glycomics_mouse">http://metadb.riken.jp/metadb/db/Glycomics_mouse</a>	91	1,462	高月
分子表現型	zTrap 遺伝子発現メタデータ	遺伝研	<a href="http://metadb.riken.jp/db/zTrap">http://metadb.riken.jp/db/zTrap</a>	1,383	39,717	斎藤、柗屋
分子表現型	マウス Cre ドライバー系統マーカー遺伝子の発現情報	熊本大、理研 BRC	<a href="http://jered.info">http://jered.info</a>	3,279	43,680	高月、柗屋

複合表現型	Rare-disease vocabulary in Japanese	(本プロジェクト)	<a href="http://metadb.riken.jp/metadb/ontology/RDVJ">http://metadb.riken.jp/metadb/ontology/RDVJ</a>	684(語彙)	2,376	斎藤、矢田
小計				1,303,943	54,689,182	-
第1期データ	BRC マウスリソース表現型メタデータ	理研 BRC	<a href="http://metadb.riken.jp/metadb/db/rikenbrc_mouse">http://metadb.riken.jp/metadb/db/rikenbrc_mouse</a>	13,152	511,120	高月、谷川
第1期データ	BRC 細胞リソースメタデータ	理研 BRC	<a href="http://metadb.riken.jp/metadb/db/rikenbrc_cell">http://metadb.riken.jp/metadb/db/rikenbrc_cell</a>	6,200	206,155	高月、谷川
第1期データ	JCM 微生物リソースメタデータ	理研 BRC	<a href="http://metadb.riken.jp/metadb/db/rikenbrc_jcm_microbe">http://metadb.riken.jp/metadb/db/rikenbrc_jcm_microbe</a>	0	688,356	高月、谷川
第1期データ	シロイヌナズナフェノームデータベース	理研 ACCC	<a href="http://metadb.riken.jp/metadb/db/SciNetS_ria143i">http://metadb.riken.jp/metadb/db/SciNetS_ria143i</a>	14,629	143,130	梶屋
合計				1,337,924	1,337,924	-

全ての表現型データは、バイオリソース記述 RDF スキーマ (BioResource Schema: BRS) に沿って記述されており、例えば、『すべての表現型データを列挙』など、データセット横断的に検索が可能である。さらには、SPARQL エンドポイントの機能である Federated query (連邦検索) を利用して、例えば『EBI RDF Platform の Reactome データセットから 1 つのパスウェイを指定し、そのパスウェイに含まれるマウス遺伝子のノックアウトについて、IMPC データセットに含まれる表現型を列挙する』など、データベース (データレポジトリ) 横断的な検索も可能である。

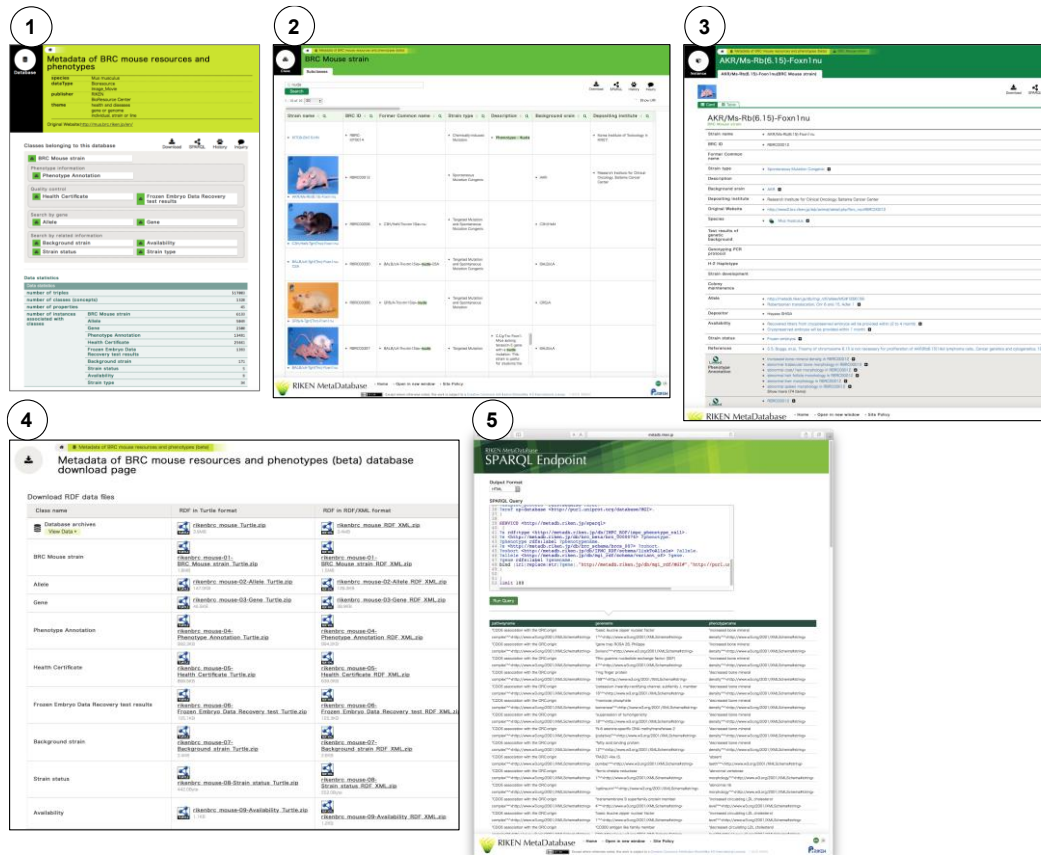


図 1：理研メタデータベースを利用したデータの表示の例

BRC マウスリソース表現型メタデータの例を示す。

- ① トップページ。データセットに含まれるクラス、及びデータセットの統計を示す。1 つのクラスを選択することで②の表画面に遷移する。
- ② 表画面。個々のデータをテーブル型に表示。1 つの行を選択することで、③のカード型画面に遷移する。
- ③ カード型画面。データセットを超えたデータのつながりを閲覧可能なように工夫されている。
- ④ ダウンロード画面。
- ⑤ SPARQL エンドポイント画面。検索式による任意の検索が可能。

各データセットは、ポータルである J-Phenome (<http://jphenome.info>) よりリンクされており、J-phenome データベースの一部として位置付けられている(図 2)。また、オリジナルの提供者からリンクして個別に扱えるように、個別の URL を持っている。

表 1 のすべてのデータは、理研メタデータベースからの公開に加え、順次 NBDC RDF ポータルからの公開を進めるために、ライフサイエンス統合データベースセンター(DBCLS)によるレビューを受けている状態である(2017年1月現在)

### 【データ収集技術】生物学者向けの表現型データ入力システム開発(梶屋、田中、高月)

J-Phenome Genome Editing Organism Registry の開発を完了した(図 2)。予め用意しておいた語彙(オントロジー)を画面遷移に沿って選んで行くことで、対象のゲノム改変生物、およびその表現型について、本プロジェクトで設計したデータスキーマに準拠し、かつ適切なオントロジーでアノテーションされた RDF が生成される。大量の種類別の語句の選択を実用的に行うインターフェースを開発した。本課題終了後も本ツールを用いたデータ収集を行っていく予定。

### 【データ活用技術】モデル生物表現型と疾患との関連づけと可視化(梶屋、田中、高月)

本アプリケーション(J-Phenome Disease Model Finder: <http://diseasemodel.riken.jp/>)では、UberPheno オントロジーのリンク(図 3)を通じて、疾患語彙と関連する表現型を示すモデル生物(マウス、ラット、メダカ等)を検索することができる。下記の機能を備える。

1. ヒト疾患名を記入して、疾患モデル動物の候補を検索する機能  
ヒト疾患語彙を記入(サジェスト補助あり)あるいはツリーから選択すると、関連する表現型を示すモデル生物を横断的にリスト表示する(図 4)。
2. 遺伝子オルソログデータを元にヒト疾患研究のモデル動物の候補を検索する機能  
遺伝子名を記入(サジェスト補助あり)すると、その遺伝子のオルソログに変異を持つモデル生物を横断的にリスト表示する。
3. 部位名から、関係するヒト疾患研究のモデル動物の候補を検索する機能  
生物種を指定して部位名をオントロジーから選択すると、その相同器官に異常を示すモデル生物を検索して表示する。

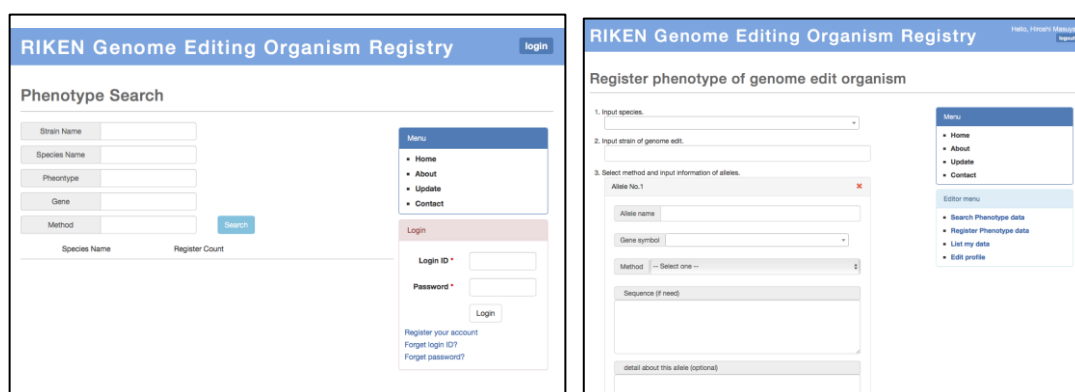


図 2 生物学者向けの表現型データ入力システムのトップおよび登録画面

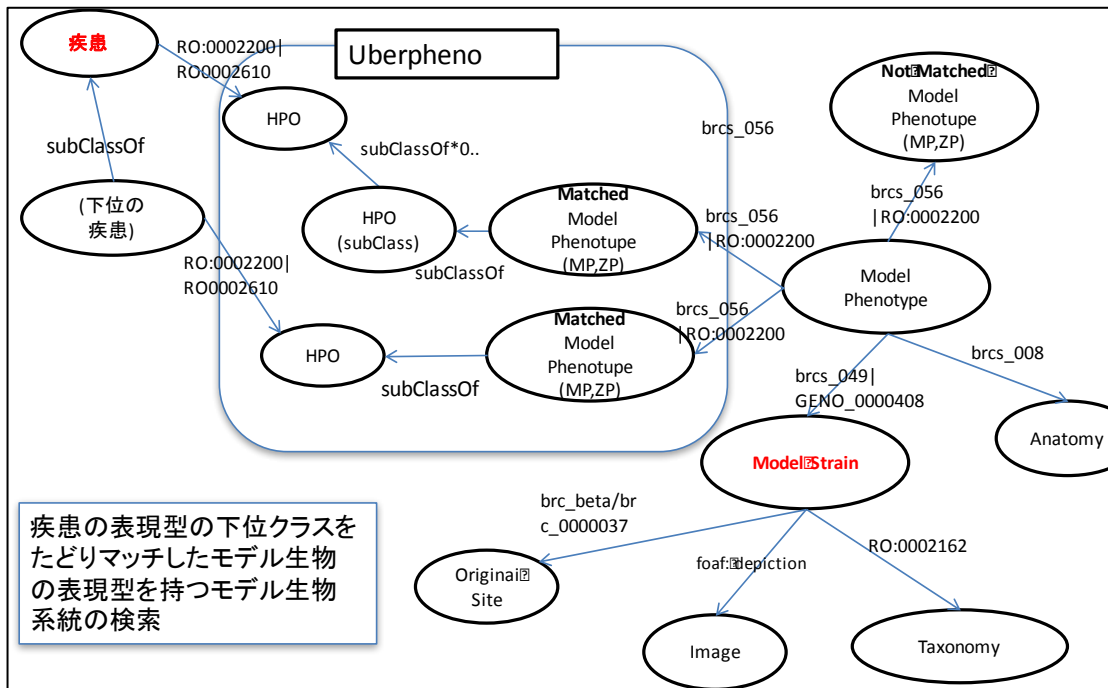


図3 UberPheno を介したヒト疾患とモデル生物表現型のリンクの概要

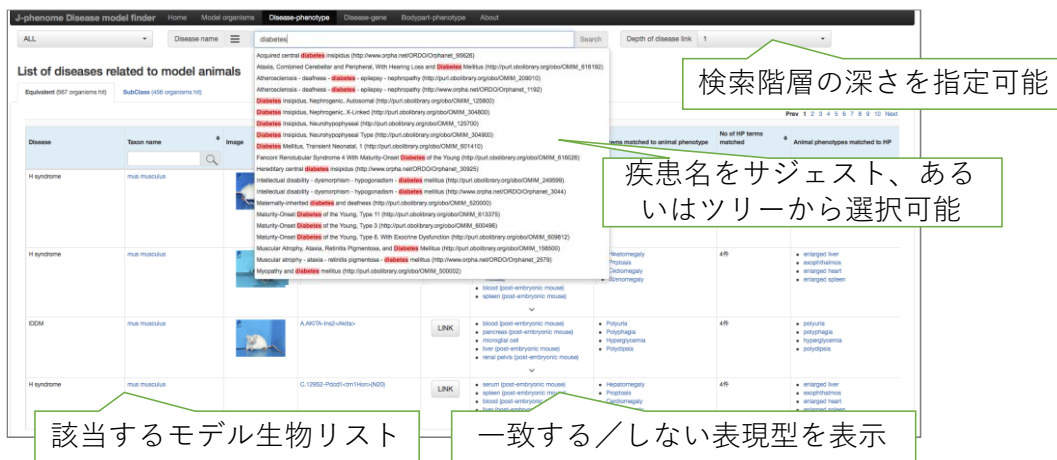


図4 疾患モデル動物候補検索画面の概要

また、ページング付きテーブル型汎用スタンザ (Stanza for Paging Table) の開発を行った。DBCLS の提案する Togo Stanza の仕組みを利用することで、任意のウェブページ内にプログラミングなしに容易に実装でき、本課題以外からも利用可能な汎用ツールとして開発した (3 月公開予定)。

**【データベース連携の交渉】(柵屋)**

- 本課題で整備した国内発の表現型データが世界で用いられるように、上述の Monarch Initiative に対してデータ供給を行うことで合意した。
- 創薬をターゲットにシステムバイオロジー基盤開発を行っているシステムズバイオロジー

- ー研究所(SBI)に対して、本課題のデータを供給することで合意した。
- AMED が組織する未診断疾患イニシアチブ (Initiative on Rare and Undiagnosed Diseases: IRUD) の情報システムに疾患モデル動物の表現型-遺伝子型の関連性データを供給することで合意した。

## (2) データベースの利便性に関する利用者ニーズと具体的な対応(榎屋、田中、高月)

分子生物学会の NBDC データベース展示ブース出展を通じて、ユーザーからの意見収集を行った。これに準じて、下記の対応を行った。

- 1) J-Phenome Disease Model Finder の機能修正  
モデル動物を表現型から検索できるシンプルな機能が必要との要望を受け、画面構成を変更した。
- 2) Stanza for Paging Table の開発  
ユーザーより、単なる RDF のビューワである理研メタデータベース表示への不満や、J-Phenome Disease Model Finder で対応できないデータの表示に関する要望が寄せられたため、任意のデータ表示に素早く対応できるツールとして、ページング付きテーブル型汎用スタンザ (Stanza for Paging Table) を開発した。

## (3) 持続的なデータベース運用体制の構築に向けた取り組み

下記を行った。

- 1) RDF をデータ形式として採用することで、現在データレポジトリとして利用している理研メタデータベース以外にも、NBDC RDF Portal を含め、公共 RDF レポジトリであれば容易にデータの引き継ぎが可能と考えている。
- 2) 第 1 期の細胞株表現型データについて、海外の提供機関とともに国際国際コンソーシアムとして、世界標準の候補となる RDF スキーマ: Cell Line Ontology (CLO) を提案し、論文発表を行った(原著論文 1)。
- 3) DBCLS 及び研究開発課題「ゲノム・メタゲノム情報統合による微生物DBの超高度化推進(研究代表者:黒川頭)」と連携して、RDF によるリソース情報統合の有用性や、後述のオントロジーの共通利用、国際標準化に向け、アジア(アジアリソースセンターネットワーク:ANRRC)、世界(微生物世界データセンター:WDCM)を通じて、働きかけを行った。
- 4) IMPC データの RDF 化について、IMPC のデータ公開を行う NIH プロジェクト Mouse Phenotyping Informatics Infrastructure (MPI2) との協議(§ 10-1 参照)を行い、データ更新ワークフローを構築した。
- 5) 表現型とつながる次世代シーケンサー(NGS)解析のメタデータ記述について、DDBJ、理研情報基盤センターと連携し、スキーマ設計を行った。
- 6) 国内での RDF データ統合を推進させるため、榎屋は、理研内有識者会議「バイオインフォマティクス委員会、データベース作業部会」部会長として、世界のデータベース動向と、RDF 利用を推進する提言を行った。(別紙)理研内データ利用基盤として、RDF を採用する「理研メタデータベース」の開発について、理研情報基盤センターと協力した。

## (4) 統合化推進プログラムの他のチームや DBCLS との連携

下記を行った。

- 1) 定期的な SPARQLthon への参加  
参加する統合化プログラムの各研究開発課題の推進メンバーとの間で、データ連携の可能性、ツールの再利用、共用に関する議論、打ち合わせを行った。
- 2) ACGG-DB: ノックアウトマウスを用いた機能糖鎖科学データベースのメタデータ「糖鎖統合データベースおよび国際糖鎖構造リポジトリの開発」(研究代表者:成松久)

と連携し、当該プロジェクトにおける表現型データを「ノックアウトマウスを用いた機能糖鎖科学データベースのメタデータ」として、本課題の RDF スキーマに沿って作成した。

- 3) 「JCM 微生物リソース特性メタデータ」に関して、DBCLS 及び「ゲノム・メタゲノム情報統合による微生物DBの超高度化推進(研究代表者:黒川頭)」の提案する RDF スキーマ、Microbial Culture Collection Vocabulary (MCCV)及び、オントロジーである Growth Media Ontology (GMO)、Metagenome/Microbe Environmental Ontology (MEO)に進じてデータを再構築した。本プロジェクトで作成したデータが、そのまま流通し活用されるネットワークができつつある。さらに上記の国際連携に向けた ANRRC, WCM への働きかけを行った。
- 4) SPARQLthon に参加する統合化推進プログラムの研究チーム、DBCLS と連携し、共通利用性の高い RDF スキーマ策定を行った。測定値(数値、単位)、バイオサンプル、文献等。
- 5) 表現型データは画像を参照することが多いため、画像メタデータの形式について、研究開発課題「生命動態情報と細胞・発生画像情報の統合データベース(研究代表者:大浪修一)」と画像メタデータの RDF スキーマ作成の協議を行った。
- 6) 研究開発課題「植物ゲノム情報活用のための統合研究基盤の構築(研究代表者:田畑哲之)」の RDF データ構築のアドバイスをを行った(§ 10-1: 田畑グループ合同ミーティング及び SPARQLthon)また、梶屋は同課題のオントロジー整備ワーキンググループ員を務めた。
- 7) RDF データ利用者が、インターネット上のデータの所在、リンクの有無などを考慮して SPARQL クエリを作成するためのツール「SPARQL Builder」開発において、主にデータ利用例、応用システム考案の面で連携を行った。
- 8) Stanza for Paging Table 開発において、DBCLS の提唱する「Togo Stanza」フレームワークを採用し、開発のアドバイスを受けた。また、SPARQLthon を通じて本ツールの利用の呼びかけを行った。

#### (5) データ産出を行う研究組織や研究室、プロジェクトとの連携

下記を行った。

- 1) ゼブラフィッシュ(遺伝研)、マウス(遺伝研)、ラット(京大 NBRP)、ゲノム編集コンソーシアム(広島大学)とは研究打ち合わせを行い、データ提供方法、公開方法等について議論を行った。
- 2) NBRP メダカの表現型データアノテーションを、基生研のテクニカルスタッフと連携して行った。

#### (6) 人材の育成

§ 10-1 に示す勉強会を行った。

#### (7) その他

特になし

## §4. 主要なデータベースの利活用状況

### 1. アクセス数

#### (1) 実績

2016年11月までに公開し、アクセス数が得られているデータベースの統計を表1に示す。

名称	種別	平成 25(2013) 年度	平成 26(2014) 年度	平成 27(2015) 年度	平成 28(2016) 年度
J-Phenome	訪問者数	(公開前)		1,133 <sup>*3</sup>	1,658
	訪問数			1,367 <sup>*3</sup>	2,095
	ページ数			2,177 <sup>*3</sup>	4,861
NBRP メダカ表現型 メタデータ	訪問者数	(公開前)		351 <sup>*4</sup>	366
	訪問数			1,181 <sup>*4</sup>	1,595
	ページ数			15,195 <sup>*4</sup>	14,219
IMPC RDF データ	訪問者数	(公開前))		128 <sup>*5</sup>	460
	訪問数			306 <sup>*5</sup>	1,595
	ページ数			4,393 <sup>*5</sup>	70,601
NBRP ラット表現型メタ データ	訪問者数	(公開前)			86 <sup>*6</sup>
	訪問数				147 <sup>*6</sup>
	ページ数				8,324 <sup>*6</sup>
遺伝研 マウス表現型デ ータベースメタデータ	訪問者数	(公開前)			91 <sup>*6</sup>
	訪問数				147 <sup>*6</sup>
	ページ数				6,998 <sup>*6</sup>
ACGG-DB: ノックア ウトマウスを用いた 機能糖鎖科学デー タベースのメタデータ	訪問者数	(公開前)		100 <sup>*5</sup>	251
	訪問数			171 <sup>*5</sup>	599
	ページ数			3,003 <sup>*5</sup>	10,131
BRC マウスリソー ス表現型メタデータ	訪問者数	8,037 <sup>*1</sup>	128	2,243 <sup>*5</sup>	4,228
	訪問数	11,328 <sup>*1</sup>	321	7,867 <sup>*5</sup>	6,943
	ページ数	103,321 <sup>*1</sup>	811	131,890 <sup>*5</sup>	162,202
BRC 細胞リソース特 性メタデータ	訪問者数	23,649 <sup>*1</sup>	183	1,533 <sup>*5</sup>	1,787
	訪問数	36,023 <sup>*1</sup>	454	7,821 <sup>*5</sup>	3,708
	ページ数	290,314 <sup>*1</sup>	1,271	166,040 <sup>*5</sup>	162,202
JCM 微生物リソース 特性メタデータ	訪問者数	(公開前)		N.D. <sup>*2</sup>	296
	訪問数			N.D. <sup>*2</sup>	1,146
	ページ数			N.D. <sup>*2</sup>	12,029
シロイヌナズナフェ ノームデータベース	訪問者数	(公開前)		N.D. <sup>*2</sup>	366
	訪問数			N.D. <sup>*2</sup>	1,297
	ページ数			N.D. <sup>*2</sup>	14,219

表 1-1 研究開発対象の主要なデータベースの利用状況(年度別)

脚注: 1: 前システムによる前課題による報告。2: 前システム閉鎖にともなう問題によりアクセス数データ得られず。3: 2015.5月公開、4: 2015年7月公開、5: 2015年10月公開、6: 2016年7月より公開。

名称	種別	平成 25(2013) 年度	平成 26(2014) 年度	平成 27(2015) 年度	平成 28(2016) 年度
J-Phenome	訪問者数	(公開前)		103	184
	訪問数	(公開前)		124	233
	ページ数	(公開前)		198	540
NBRP メダカ表現型 メタデータ	訪問者数	(公開前)		39	41
	訪問数	(公開前)		131	177
	ページ数	(公開前)		1,688	1,580
IMPC RDF データ	訪問者数	(公開前)		21	51
	訪問数	(公開前)		51	177
	ページ数	(公開前)		732	7,845
NBRP ラット表現型メ タデータ	訪問者数	(公開前)			86
	訪問数	(公開前)			147
	ページ数	(公開前)			8,324
遺伝研 マウス表現型デ ータベースメタデータ	訪問者数	(公開前)			91
	訪問数	(公開前)			147
	ページ数	(公開前)			6,998
ACGG-DB: ノックア ウトマウスを用いた 機能糖鎖科学デー タベースのメタデータ	訪問者数	(公開前)		17	28
	訪問数	(公開前)		29	67
	ページ数	(公開前)		501	1,126
BRC マウスリソー ス表現型メタデータ	訪問者数	670	11	470	4,228
	訪問数	944	27	771	6,943
	ページ数	8,610	68	18,022	162,202
BRC 細胞リソース特 性メタデータ	訪問者数	1,971	15	199	1,787
	訪問数	3,002	38	412	3,708
	ページ数	24,193	106	18,022	162,202
JCM 微生物リソース 特性メタデータ	訪問者数	(公開前)		N.D.	83
	訪問数	(公開前)		N.D.	375
	ページ数	(公開前)		N.D.	2,262
シロイヌナズナフェ ノームデータベース	訪問者数	(公開前)		N.D.	77
	訪問数	(公開前)		N.D.	572
	ページ数	(公開前)		N.D.	3,930

表 2-2 研究開発対象の主要なデータベースの利用状況(月間平均)

## (2) 分析

- 本課題で開発しているデータベース群は、J-Phenome ポータルからリンクされていると同時に、個々に Google 等をたどってアクセスが可能のため、2016 年 12 月までに公開されたデータベースのリストを示した。
- 前課題における 2013 年のアクセス数は、PosMed, BioLOD 等の本課題で対象としていないツールのアクセスを含んでおり、他の年度と比較評価することが難しい。2014 年は、これらのツールがコスト面から運用困難になり、運用停止となったこと、暫定のサーバーによる不十分な運用形態であったこと等からアクセス数が激減していた。また、本課題で対象となっているほとんどのデータが作成中であった。
- RDF データの公開基盤として 2015 年 5 月に運用を開始した「理研メタデータベース」を選定し、順次公開を行った。2014 年に激減したアクセス数は、全てのデータベースにおいて、月



ごとに徐々に回復していった。分子生物学会出展でのアピールも影響していると思われる。2016年度以降は月毎のアクセス数は全体的に安定傾向にある。

- アクセス数は分野による利用ユーザーの数を反映していると考えられる。培養細胞、マウスは多く、メダカは少ない。データ自体は複数のオントロジーを介して繋がっているため、分野横断的に検索を可能にするツール等によって「相互乗り入れ」をより推進することでメダカ等、利用者が少ない分野のデータに関しても、増加が期待できる。
- 今後、J-Phenome Disease Model Finder 等の統合的な可視化を実現するツールを通して、データの利用者をさらに増やしていきたいと考えている。
- 本課題のデータ利用は、GUIからの閲覧のみでなく、SPARQL等を用いた検索による利用が大きいと考えられる。SPARQL 検索は、外部(及び課題内部)におけるデータ利用ツールの開発の際にも大量に行えるため、評価方法が難しいが、今後は SPARQL からの利用が増加するような方策を考えていきたい。

## 2. その他

特になし

### §5. 研究開発期間中に得られた科学・技術や産業に対する波及効果

研究開発の背景に記した通り、近年、ヒト疾患研究分野、とりわけ単一遺伝子の寄与が大きい希少疾患、難病、未診断疾患の分野において、オントロジーを用いた表現型(病態)情報の共有が注目されるようになってきている。本課題は疾患モデル動物となりうる脊椎動物の表現型と遺伝子型の関連情報の統合を行っているが、この情報はヒト希少疾患の原因遺伝子特定に大きく役立つ情報である。本課題が IMPC の大規模マウス表現型データを扱うこととなったことを受けて、AMED が組織する未診断疾患イニシアチブ (Initiative on Rare and Undiagnosed Diseases: IRUD) の情報センター(慶應大学)との連携を開始することとなり、このことを端緒として、H27年度よりAMED 臨床ゲノム情報統合データベース整備事業「真に個別患者の診療に役立つ領域横断的に高い拡張性を有する変異・多型情報データベースの創成」に班員として参加することとなった。主に、本課題で収集したモデル動物の遺伝子型-表現型関連性データの医療分野の提示方法に関して検討を行っていく予定である。

### §6. 今後の展開

- データベース連携の交渉では、Monarch Initiative, SBI, IRUD に対する合意形成を行った。今後、Monarch Initiative, SBI に対しては、実際にデータ連携のための組み込み作業を、IRUD に対しては本データベースを検索するツールの提供等を行っていく予定である。これらの連携を通して、医療研究分野へのモデル生物表現型の利用がより高めることができると考えている。SBI との連携は創薬方面での本データ利用の可能性を示しており、データ提供を進めていくことが必要である。
- 本課題で開発した J-Phenome Genome Editing Organism Registry を用いて、ゲノム編集コンソーシアムからのデータ収集を行っていくことができると期待される。単に各研究者の自由意思に任せるだけでなく、ゲノム編集コンソーシアムの活動アピール等のモチベーションを利用していく。
- 我が国として独自に整備すべき表現型統合データベースとして、かつ、世界の表現型データベースと連携しながら、多くの国内ユーザーが J-Phenome に対して希望している網羅的検索を実現するために、モデル生物の A) 生物種横断及び B) 種内の 2 つの方向で網羅性を向上させる。
- DBCLS では、100 万件の希少疾患症例報告をテキストマイニングし、症例毎に Disease Ontology (DO), Orphanet Rare Disease Ontology (ORDO)等の疾患語彙を Human

Phenotype Ontology(HPO)による表現型との対応データを保持している。このデータの解析により、各疾患における病態(表現型)の重み付けが行える。この結果をRDF化して、重み付けが行えるようなデータライブラリとして公開するとともに、本課題の J-Phenome Disease Model Finder にも適用し、より精度の高い疾患モデル動物検索を実現する。

- RDF で記述された他のオープンなデータセットと連結した検索を駆使し、表現型や疾患に絞り込みをかけた場合に、それに伴って生じる関連データの統計的な偏り(濃縮:エンリッチメント)を検出して示すアプリケーション、また、未だ特定されていないような遺伝子機能の予測、特定の研究分野にしか用いられなかった実験動物のほか分野における利用を提案する推論方法の開発を行う。

## §7. 自己評価

**【提供データ拡充】** 本課題以前は、フェノーム情報を生物種横断的に統合するというアイデアは、特に国内では意識されていなかった。国外ではオントロジーに基づく基盤技術が確立されつつあり、本課題の実施中にも進展があった。本プロジェクトで掲げる「研究分野の垣根を越えた情報共有」は、これらの RDF 関連技術を利用することでより大きく推進することができた。

データ量からみると、当初計画になかった IMPC データ統合を行ったことで、目標を達成し、計画以上の成果を出すことができた。表現型の総データポイント数は 130 万を超え、同一の枠組みを用いて利用できる「表現型ビッグデータ」が形成されつつある。

基盤技術としての RDF、共通 RDF スキーマの導入は、本成果の大きなアドバンテージとなった。研究分野の垣根を越えた情報共有は、データのシンプルなビューワー、及び SPARQL 検索及び連邦検索 (federated query) 等により目標のレベルは実現できた。また、世界的及び、統合化推進プログラム及び、SPARQLthon 等を通じた国内での RDF 技術の拡がりを背景に様々な発展があった。本課題が扱う表現型と遺伝子型、これらを持つバイオソースの関係性は、他のデータベースとの接続点が予想以上に多く、実施内容(4)に述べたように、個々のデータの URI 共用、スキーマの共同開発等、多岐にわたる連携を行った。これらの連携は、今後のデータ統合の推進力になっていくと期待される。

今後はこれらのデータの利用率をさらに向上させることが課題である。

**【データ収集技術】** 生物学者向けの表現型データ入力システム開発システムの開発完了に留まり、ゲノム編集コンソーシアムからのデータ収集のトライアルを十分行うことができなかった。本課題終了後もこのツールを用いたデータ収集を行っていく予定である。

**【データ活用技術】** モデル生物表現型と疾患との関連づけと可視化アプリケーションを開発した。HPO オントロジー、及び Monarch Initiative で開発された表現型統合データライブラリを利用することで、より効率的な開発が行えたが、一方で、データのリンクが深い場合の計算量が問題となり、その解決のためにほぼ計画通りの開発量となった。モデル生物と疾患とを結びつけ、分野の違いによる部位の呼び名の違い等を吸収した検索等が可能となった。本技術は、分野横断性を実現するツールの方向性の一つを示したと考えている。

**【データベース連携の交渉】** RDF によるデータ整備は、データベース連携の可能性を大きく広げた。本課題で整備したデータを広く利用させるための方策として、外部のデータベースに本課題のデータを組み込むというアイデアに着目し、希少疾患研究、国際的なフェノーム統合、システムズバイオロジーという本課題とは目的が(外見上)異なる 3 つのデータベースとの連携を交渉した。実施期間中には、実際のデータ連携には至らなかったが、「実際にユーザーがいる場所にデータを持っていく」という意味では、データ利用拡大の 1 つの方策を示すことができたと考えている。

## §8. 外部発表等

### 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)	9 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. Sarntivijai S, Lin Y, Xiang Z, Meehan TF, Diehl AD, Vempati UD, Schürer SC, Pang C, Malone J, Parkinson H, Liu Y, Takatsuki T, Saijo K, Masuya H, Nakamura Y, Brush MH, Haendel MA, Zheng J, Stoeckert CJ, Peters B, Mungall CJ, Carey TE, States DJ, Athey BD, He Y. “CLO: The cell line ontology”, *J Biomed Semantics*. vol. 5 No. 37. 2014 (DOI: 10.1186/2041-1480-5-37)
2. 梶屋啓志, 溝口理一郎 “遺伝学オントロジー”, *人工知能学会論文誌* vol. 29, No. 3, pp. 311-327, 2014 (DOI: <http://doi.org/10.1527/tjsai.29.311>)
3. Karp N A, Meehan T F, Morgan H, Mason J C, Blake A, Kurbatova N, Smedley D, Jacobsen J, Mott R F, Iyer V, Matthews P, Melvin D G, Wells S, Flenniken A M, Masuya H, Wakana S, White J K, Lloyd K C K, Reynolds C L, Paylor R, West D B, Svenson K L, Chesler E J, de Angelis M H, Tocchini-Valentini G P, Sorg T, Herculat Y, Parkinson H, Mallon A M, Brown S D M, “Applying the ARRIVE Guidelines to an In Vivo Database” *PLOS Biology*, 13, e1002151, 2015 (DOI:10.1371/journal.pbio.1002151)
4. Maier H, Schütt C, Steinkamp R, Hurt A, Schneltzer E, Gormanns P, Lengger C, Griffiths M, Melvin D, Agrawal N, Alcantara R, Evans A, Gannon D, Holroyd S, Kipp C, Raj NP, Richardson D, LeBlanc S, Vasseur L, Masuya H, Kobayashi K, Suzuki T, Tanaka N, Wakana S, Walling A, Clary D, Gallegos J, Fuchs H, de Angelis MH, Gailus-Durner V, Principles and application of LIMS in mouse clinics, *Mammalian Genome*, 26, 467-481, 2015 (DOI 10.1007/s00335-015-9586-7)
5. Kimura M, Ichimura S, Sasaki K, Masuya H, Suzuki T, Wakana S, Ikegawa S, Furuichi T, Endoplasmic reticulum stress-mediated apoptosis contributes to a skeletal dysplasia resembling platyspondylic lethal skeletal dysplasia, Torrance type, in a novel Col2a1 mutant mouse line, *Biochemical and Biophysical Research Communications*, 468, 86-91, 2015 (doi:10.1016/j.bbrc.2015.10.160)
6. Satoshi Kume, Hiroshi Masuya, Yosky Kataoka and Norio Kobayashi, Development of an Ontology for an Integrated Image Analysis Platform to enable Global Sharing of Microscopy Imaging Data, *Proceedings of the 15th International Semantic Web Conference*, P23, 2016 (<http://ceur-ws.org/Vol-1690/paper93.pdf>)
7. Kai Lenz, Hiroshi Masuya and Norio Kobayashi, RIKEN MetaDatabase: a database publication platform for RIKENs life-science researchers that promotes research collaborations over different research area, *Proceedings of the 15th International Semantic Web Conference*, P24, 2016 (<http://ceur-ws.org/Vol-1690/paper34.pdf>)
8. Terue Takatsuki, Mikako Saito, Sadahiro Kumagai, Eiki Takayama, Kazuya Ohshima, Nozomu Ohshiro, Kai Lenz, Nobuhiko Tanaka, Norio Kobayashi and Hiroshi Masuya, A RDF based Portal of Biological Phenotype Data produced in

Japan, Proceedings of the 15th International Semantic Web Conference, P25, 2016 (<http://ceur-ws.org/Vol-1690/paper10.pdf>)

9. Kobayashi N, Lenz K and Masuya H, RIKEN MetaDatabase: a database platform as a microcosm of linked open data cloud in the life sciences, Joint International Semantic Technology Conference JIST 2016: Semantic Technology pp 99-115, id:21, 2016 (DOI: 10.1007/978-3-319-50112-3\_8)
10. Atsuko Yamaguchi, Kouji Kozaki, Kai Lenz, Yasunori Yamamoto, Hiroshi Masuya and Norio Kobayashi, Data Acquisition by Traversing Class-Class Relationships over the Linked Open Data, Joint International Semantic Technology Conference JIST 2016: Semantic Technology pp 136-151, 2016, (DOI: 10.1007/978-3-319-50112-3\_11)
11. Kobayashi N, Kume S, Lenz K and Masuya H, RIKEN MetaDatabase: a database platform as a microcosm of linked open data cloud in the life sciences, International Journal on Semantic Web and Information Systems (IJSWIS), accepted (JIST2016 ベストペーパー (改変バージョン) として推薦論文)

## 2. その他の著作物（総説、書籍など）

1. 榎屋啓志、表現型オントロジー～バイオリソースのさらなる活用のために～、バイオリソースニューズレターBioResource now!, Vol.12, No.8, 2016  
(URL:  
[http://shigen.nig.ac.jp/shigen/news/n\\_letter/2016/newsletter\\_v12\\_n8.html#main](http://shigen.nig.ac.jp/shigen/news/n_letter/2016/newsletter_v12_n8.html#main))
2. 榎屋啓志・若菜茂晴、マウスの解析に役立つデータベース・ウェブツール、実験医学別冊マウス表現型解析スタンダード、36-46、2016、羊土社

## 3. 国際学会発表及び主要な国内学会発表

### (1) 概要

種別	国内外	件数
招待講演	国内	4 件
	国際	6 件
口頭発表	国内	7 件
	国際	4 件
ポスター発表	国内	20 件
	国際	11 件

### (2) 招待講演

#### 〈国内〉

1. 榎屋啓志. “モデル生物をより使いやすくするための表現型情報整備”、第 11 回生命資源研究・支援センターシンポジウム、熊本、2015 年 3 月 13 日
2. 榎屋啓志. “表現型情報で生物をつなぐフェノーム統合データベース”、トーゴの日シンポジウム 2014、東京、2014 年 10 月 5 日
3. 榎屋啓志. “生物表現型の RDF データ作成とそのメリット”、第 5 回生命医薬情報学連合大会、東京、2016 年 9 月 29 日
4. 榎屋啓志. “生物の表現型が研究コミュニティをつなぐ”、トーゴの日シンポジウム 2016、東京、2016 年 10 月 6 日

#### 〈国際〉

1. Masuya H. “Bioinformatics –Mouse Phenotyping Databases”, AMMRA-AMPC Student Workshop, Canberra Australia, October 23, 2014

2. Masuya H. “Development of the mouse phenotype database”, The 11th KOGO Winter Symposium 2015, Hongcheon Korea, February 5, 2015
3. Masuya H, Development of an integrated phenotype database of experimental animals, Korean Association for Laboratory Animal Science (KALAS) 2015 International Symposium KALAS-RIKEN BRC Joint Symposium, Incheon, Korea, August 12, 2015
4. Masuya H, Data integration of phenotype data of disease model animals using Resource, 2016 AMMRA & AMPC meeting in Hakone, Hakone, May 20, 2016
5. Masuya H, A portal of biological phenotype data of experimental animals produced in Japan, The 8th ANRRC International Meeting, Kyoto, September 22, 2016
6. Masuya H, RDF-Based Data sharing of phenotype data of experimental animals produced from Japan, The Fourth International Conference on Rare and Undiagnosed Diseases, Tokyo, Japan, November 17, 2016

### (3) 口頭講演

#### 〈国内〉

1. 榎屋啓志、高月照江、齋藤実香子、高山英紀、吉田有子、蒔田由布子、望月芳樹、土井考爾、小林紀郎、豊田哲郎. “セマンティックWeb技術を用いた、生物表現型統合データベース”、JSAI2014 人口知能学会全国大会(第28回)、松山、2014年5月12日
2. 榎屋啓志、古崎晃司、高山英紀、高田豊行、真下知士、若菜茂晴、大江和彦、溝口理一郎. “マウス・ラットの表現型と関連するヒト疾患を調べられるデータベース”、日本実験動物科学技術 さつぽろ 2014、札幌、2014年5月17日
3. 榎屋啓志、池内達郎、布山喜章. “遺伝学用語集の改訂における課題と問題”、日本遺伝学会第86回大会、長浜、2014年9月20日
4. 榎屋啓志、高山英紀、古崎晃司、今井健、大江和彦、溝口理一郎、生物表現型情報と、疾患情報をつなげるデータベース、2015年度 人工知能学会全国大会、函館、2015年5月30日
5. 榎屋啓志、高月照江、齋藤実香子、高山英紀、大島和也、田中信彦、戀津魁、小林紀郎、実験動物の利用拡大に向けたデータベース構築、第63回日本実験動物学会総会、川崎、2016年5月18日
6. 榎屋啓志、高月照江、齋藤実香子、高山英紀、大島和也、田中信彦、戀津魁、小林紀郎、RDF技術を用いた疾患モデル動物の表現型データの統合と利用拡大、2016年度 人工知能学会全国大会、北九州、2016年、6月7日
7. 榎屋啓志、高月照江、齋藤実香子、高山英紀、大島和也、田中信彦、表現型データ共有のためのデータベースポータル: J-phenome、日本遺伝学会 第88回大会、三島、2016年9月9日

#### 〈国際〉

1. Masuya H, Takatsuki T, Saito M, Mori H, Kawashima S, Data coordination of bio-resources related information using resource description framework technology, The 7th ANRRC International Meeting, Incheon, Korea, September 18, 2015
2. Masuya H, J-phenome: a portal of Japanese phenotype data, BioHackathon 2016 Symposium, Tsuruoka, June 12, 2016
3. Kobayashi N, Lenz K and Masuya H. “RIKEN MetaDatabase: a database platform as a microcosm of linked open data cloud in the life sciences”, Joint International Semantic Technology Conference 2016 (JIST 2016), Singapore, November 3, 2016
4. Yamaguchi A, Kozaki K, Lenz K, Yamamoto Y, Masuya H and Kobayashi N.

“Semantic Data Acquisition by Traversing Class-Class Relationships over the Linked Open Data”, Joint International Semantic Technology Conference 2016 (JIST 2016), Singapore, November 4, 2016

#### (4) ポスター発表

(国内)

1. 田中信彦、土岐秀明、茂木浩未、鈴木智広、金田秀貴、三浦郁生、山田郁子、古瀬民生、小林喜美男、井上麻紀、美野輪治、野田哲生、若菜茂晴、榎屋啓志. “新規の疾患モデルマウスや生命現象の発見を目的としたデータマイニング:ワークフローの検討”、日本実験動物科学技術 さっぽろ 2014、札幌、2014年5月15日
2. 田中信彦. “網羅的マウス表現型データの解析ワークフローの検討”、第28回モロシヌス研究会、三島、2014年6月27日
3. 高月照江、矢田有加里、齋藤実香子、榎屋啓志. “表現型情報でつなぐフェノーム統合データベース”、トーゴの日シンポジウム 2014、東京、2014年10月5日
4. 田中信彦、土岐秀明、茂木浩未、鈴木智広、金田秀貴、三浦郁生、山田郁子、古瀬民生、小林喜美男、井上麻紀、美野輪治、若菜茂晴、榎屋啓志. “網羅的マウス表現型データの解析ワークフローの検討”、第37回日本分子生物学会年会、横浜、2014年11月26日
5. 高月照江、矢田有加里、齋藤実香子、高田豊行、真下知士、成瀬清、若菜茂晴、田中信彦、榎屋啓志、生物種を超えた表現型情報統合のためのデータ作成、第62回日本実験動物学会総会、京都、2015年5月28日
6. 廣田和之、小澤恵代、榎屋啓志、田村勝、若菜茂晴、日本マウスクリニックにおける骨形態評価の国際標準化への取り組み、第62回日本実験動物学会総会、京都テルサ、2015年5月28日
7. 三浦郁生、篠木晶子、臼田大輝、鈴木智広、金田秀貴、古瀬民生、小林喜美男、山田郁子、田村勝、榎屋啓志、若菜茂晴、C57BL/6 亜系統間を検出可能な SNPs マーカーセットの構築、第62回日本実験動物学会総会、京都テルサ、2015年5月29日
8. 榎屋啓志、セマンティック Web 技術を用いたバイオリソースデータの標準化と整備、第1回理化学研究所・産業技術総合研究所 共同シンポジウム、産業技術総合研究所臨海副都心センター、2015年6月29日
9. 榎屋啓志、高月照江、齋藤実香子、高山英紀、大島和也、大城望、田中信彦、モデル生物の表現型統合データベースの作成に向けて、第29回モロシヌス研究会、かんばの宿 有馬、2015年7月3日
10. 高月照江、齋藤実香子、大島和也、高山英紀、金子裕代、成瀬清、若菜茂晴、田中信彦、榎屋啓志、表現型情報統合のためのデータ作成、トーゴの日シンポジウム 2015、東京大学弥生講堂、2015年10月6日
11. 高月照江、齋藤実香子、大島和也、高山英紀、金子裕代、成瀬清、若菜茂晴、田中信彦、榎屋啓志、表現型情報統合データベースのためのデータ作成、BMB2015、神戸ポートアイランド、2015年12月3日
12. 田中信彦、土岐秀明、鈴木智広、金田秀貴、三浦郁生、山田郁子、古瀬民生、小林喜美男、井上麻紀、美野輪治、若菜茂晴、榎屋啓志、表現型の異常を高感度で検出する統計解析ワークフローの開発、BMB2015、神戸ポートアイランド、2015年12月3日
13. 榎屋啓志、小林紀郎、生物表現型ビッグデータの形成に向けて:セマンティック Web 技術を用いた表現型データの統合、第2回理化学研究所・産業技術総合研究所 共同シンポジウム、産業技術総合研究所つくばセンター、2016年2月2日
14. 高月照江、齋藤実香子、高田豊行、庫本高志、成瀬清、若菜茂晴、田中信彦、榎屋啓志、実験動物の利用拡大に向けたデータベース構築、第63回日本実験動物学会総会、川崎、2016年5月18日

15. 熊谷禎洋、高月照江、斎藤実香子、高山英紀、榎屋啓志、RDF を利用して実験動物の表現型と疾患との関連性を示すシステム、トーゴの日シンポジウム 2016、東京、2016 年 10 月 5 日
16. 高月照江、斎藤実香子、高山英紀、大島和也、田中信彦、熊谷禎洋、榎屋啓志、生物種横断的な表現型データポータルサイト J-phenome について、トーゴの日シンポジウム 2016、東京、2016 年 10 月 6 日
17. 山口敦子、山本泰智、古崎晃司、戀津魁、榎屋啓志、小林紀郎、信頼性が確保された SPARQL エンドポイントでのクエリ生成支援、トーゴの日シンポジウム 2016、東京、2016 年 10 月 5 日
18. 戀津魁、榎屋啓志、小林紀郎、理研メタデータベースの運用とデータ統合の実際、トーゴの日シンポジウム 2016、東京、2016 年 10 月 6 日
19. 田中信彦、榎屋啓志、網羅的マウス表現型データのアソシエーション分析による表現形質間の関係性の提示と予測、第 39 回日本分子生物学会年会、横浜、2016 年 12 月 2 日
20. 高月照江、斎藤実香子、高山英紀、大島和也、田中信彦、熊谷禎洋、榎屋啓志、生物種横断的な表現型データポータルサイト J-phenome について、第 39 回日本分子生物学会年会、横浜、2016 年 12 月 2 日

〈国際〉

1. Takatsuki, T. Yada, Y. Saito, M. Masuya, H. , Toward development of integrated phenotype database across biological species, Towards precision medicine, Phenotyping human diseases in mice. The 16th Frank & Bobbie Fenner Conference, Canberra, 20-21 October 2014
2. Suzuki T, Kaneda H, Kobayashi K, Miura I, Furuse T, Yamada I, Minowa O, Toki H, Tanaka N, Tamura M, Masuya H, Wakana S, RIKEN BRC Japan Mouse Clinic: Comprehensive mouse phenotyping for research supports and for IMPC, Phenotyping human diseases in mice. The 16th Frank & Bobbie Fenner Conference, Canberra, 20-21 October 2014
3. Masuya H, Takatsuki T, Saito M, Lents K, Takayama E, Ohshima K, Tanaka N, Naruse K, Takada T, Kuramoto T, Wakana S, Kobayashi N, Development of Semantic Web/RDF based integrated database of experimental animals, 29th International Mammalian Genome Conference, Yokohama, Japan, November 9, 2015
4. Tanaka N, Masuya H, Exploring of novel mouse models for human disease with comprehensive mouse phenotyping data, 29th International Mammalian Genome Conference, Yokohama, Japan, November 9, 2015
5. Masuya H, Takatsuki T, Saito M, Takayama E, Ohshima K, Ohshiro N, Lenz K, Tanaka N, Mori H, Kawashima S, Kobayashi N, RDF-based data sharing of bio-resource related information, SWAT4LS International Conference, Cambridge, England, December 8, 2015
6. Kobayashi N, Lenz K, Masuya H, RIKEN Meta Database: a life-science metadata database based on the Semantic Web, SWAT4LS International Conference, Cambridge, England, December 8, 2015
7. Kobayashi N, Yamaguchi Y, Masuya H, RIKEN MetaDatabase: a life-science database integration and publication platform based on the Semantic Web, 4th INCF Japan Node International Workshop and 14th INCF Nodes Workshop, Wako, May 28, 2016
8. Takatsuki T, Saito M, Kumagai S, Takayama E, Ohshima K, Ohshiro N, Lenz K, Tanaka N, Kobayashi N, Masuya H, A RDF-Based portal of biological phenotype data, The 15th International Semantic Web Conference (ISWC2016),



Kobe, October 19, 2016

9. Yamaguchi A, Kozaki K, Lenz K, Yamamoto Y, Masuya H, Kobayashi N, Data acquisition by traversing Class-Class relationships over the linked open data, The 15th International Semantic Web Conference (ISWC2016), Kobe, October 19, 2016
10. Kume S, Masuya H, Kataoka Y, Kobayashi N, Development of an ontology for an integrated image analysis platform to enable global sharing of microscopy imaging data, The 15th International Semantic Web Conference (ISWC2016), Kobe, October 19, 2016
11. Lenz K, Masuya H, Kobayashi N, RIKEN MetaDatabase: a database publication platform for RIKENs life-science researchers that promotes research collaborations over different research area, The 15th International Semantic Web Conference (ISWC2016), Kobe, October 19, 2016

#### 4. 知財出願

なし

#### 5. 受賞・報道等

##### (1) 受賞

1. Best Paper Award of Joint International Semantic Technology Conference 2016 (JIST 2016), Norio Kobayashi, Kai Lentz and Hiroshi Masuya, 2016年11月4日
2. 網羅的マウス表現型データのアソシエーション分析による表現形質間の関係性の提示と予測、優秀ポスター賞、第39回日本分子生物学会年会、横浜、2016年12月2日

##### (2) メディア報道

1. 理研、JST 合同プレスリリース、マウスの大規模解析データを世界へ – 国際標準規格の技術を活用した生命科学の新たなビッグデータ、2016年9月20日 (URL: [http://www.riken.jp/pr/press/2016/20160920\\_2/](http://www.riken.jp/pr/press/2016/20160920_2/))
2. 日経産業新聞、遺伝子改変マウス実験 データベース統合 理研など、2016年9月21日
3. 化学工業日報、マウスの「表現型」データ 国際規格に変換、発信、2016年9月27日

##### (3) その他

該当なし

## §9. 研究開発期間中の活動

### 1. 進捗ミーティング

年月日	名称	場所	参加人数	目的・概要
2014年 5月9日	田畑グループ合同ミーティング(非公開)	かずさDNA研究所	7人	動物、植物でのRDF構成の調整のためのミーティング。田畑グループと共催
2014年 5月20日	チーム内ミーティング(非公開)	理研筑波研究所 研究棟121室	5人	研究室内の進捗ミーティング: 疾患データ ICD10 の状況確認
2014年 6月17日	チーム内ミーティング(非公開)	理研筑波研究所 研究棟121室	6人	研究室内の進捗ミーティング: RDF 概要説明。
2014年 7月8日	チーム内ミーティング(非公開)	理研筑波研究所 研究棟121室	7人	研究室内の進捗ミーティング: RDF データ不具合の検証
2014年 7月11日	IMPC-MPI2 マウスデータ統合グループとの打ち合わせ(非公開)	European Bioinformatics Institute (EBI)	10人	EBI グループとの共催。マウス表現型データのRDF 国際共有のための打ち合わせ
2014年 7月24日	チーム内ミーティング(非公開)	理研情報基盤センター	3人	理研内 RDF データ基盤を利用した、表現型データベース開発について打ち合わせ
2014年 8月19日	チーム内ミーティング(非公開)	理研筑波研究所 研究棟121室	7人	研究室内の進捗ミーティング: RDF 変換ツールの動作確認
2014年 9月12日	研究打ち合わせ(非公開)	理研情報基盤センター	3人	理研内 RDF データ基盤との擦り合わせ
2014年 9月30日	チーム内ミーティング(非公開)	理研筑波研究所 研究棟121室	6人	研究室内の進捗ミーティング: ラットデータの確認
2014年 11月20日	研究打ち合わせ(非公開)	理研情報基盤センター	3人	理研内 RDF データ基盤との擦り合わせ
2014年 12月9日	チーム内ミーティング(非公開)	理研筑波研究所 研究棟121室	7人	研究室内の進捗ミーティング: 微生物 RDF データスキーマ確認
2014年 12月18日	ゼブラフィッシュデータ研究打ち合わせ(非公開)	国立遺伝学研究所	3人	ゼブラフィッシュデータのコミュニティへの貢献について議論
2014年 12月19日	マウス毛色データ収集に関するミーティング(非公開)	国立遺伝学研究所	10人	遺伝研に保存されている各国のマウス毛色表現型のデータ化に関する議論
2015年 1月13日	研究打ち合わせ(非公開)	理研情報基盤センター	3人	理研内 RDF データ基盤との擦り合わせ
2015年 1月19日	表現型データベース勉強会(非公開)	理研筑波研究所 研究棟121室	5人	ゼブラフィッシュデータベース ZFIN の全体像を確認
2015年 2月4日	韓国マウス表現型解析グループとの打ち合わせ	Seoul National University	5人	ソウル大学グループと共催: 韓国におけるマウス表現型解析解析の状況確認と、データ化についての議論
2014年 5月9日	田畑グループ合同ミーティング(非公開)	かずさDNA研究所	7人	動物、植物でのRDF構成の調整のためのミーティング。田

				畑グループと共催
2014年 5月20日	チーム内ミーティング (非公開)	理研筑波研究所 研究棟 121 室	5 人	研究室内の進捗ミーティング: 疾患データ ICD10 の状況 確認
2014年 6月17日	チーム内ミーティング (非公開)	理研筑波研究所 研究棟 121 室	6 人	研究室内の進捗ミーティング: RDF 概要説明。
2014年 7月8日	チーム内ミーティング (非公開)	理研筑波研究所 研究棟 121 室	7 人	研究室内の進捗ミーティング: RDF データ不具合の検証
2014年 7月11日	IMPC-MPI2 マウスデ ータ統合グループとの 打ち合わせ(非公開)	European Bioinformatics Institute (EBI)	10 人	EBI グループとの共催。マウ ス表現型データの RDF 国際 共有のための打ち合わせ
2014年 7月24日	チーム内ミーティング (非公開)	理研情報基盤セ ンター	3 人	理研内 RDF データ基盤を利用 した、表現型データベース 開発について打ち合わせ
2014年 8月19日	チーム内ミーティング (非公開)	理研筑波研究所 研究棟 121 室	7 人	研究室内の進捗ミーティング: RDF 変換ツールの動作確認
2014年 9月12日	研究打ち合わせ(非公 開)	理研情報基盤セ ンター	3 人	理研内 RDF データ基盤との 擦り合わせ
2014年 9月30日	チーム内ミーティング (非公開)	理研筑波研究所 研究棟 121 室	6 人	研究室内の進捗ミーティング: ラットデータの確認
2014年 11月20日	研究打ち合わせ(非公 開)	理研情報基盤セ ンター	3 人	理研内 RDF データ基盤との 擦り合わせ
2014年 12月9日	チーム内ミーティング (非公開)	理研筑波研究所 研究棟 121 室	7 人	研究室内の進捗ミーティング: 微生物 RDF データスキーマ 確認
2014年 12月18日	ゼブラフィッシュデータ 研究打ち合わせ(非公 開)	国立遺伝学研究 所	3 人	ゼブラフィッシュデータのコミ ュニティへの貢献について議 論
2014年 12月19日	マウス毛色データ収集 に関するミーティング (非公開)	国立遺伝学研究 所	10 人	遺伝研に保存されている各国 のマウス毛色表現型のデータ 化に関する議論
2015年 1月13日	研究打ち合わせ(非公 開)	理研情報基盤セ ンター	3 人	理研内 RDF データ基盤との 擦り合わせ
2015年 1月19日	表現型データベース勉 強会(非公開)	理研筑波研究所 研究棟 121 室	5 人	ゼブラフィッシュデータベース ZFIN の全体像を確認
2015年 2月4日	韓国マウス表現型解析 グループとの打ち合 わせ	Seoul National University	5 人	ソウル大学グループと共催: 韓国におけるマウス表現型解 析解析の状況確認と、データ 化についての議論
2015年 3月2日	表現型データベース勉 強会(非公開)	理研筑波研究所 研究棟 121 室	5 人	国際マウス表現型解析コンソ ーシアムデータベースの全体 像を確認
2015年 3月16日	表現型データベース勉 強会(非公開)	理研筑波研究所 研究棟 121 室	5 人	RDF スキーマの詳細につい て確認
2015年 3月17日	チーム内ミーティング (非公開)	理研筑波研究所 研究棟 121 室	7 人	研究室内の進捗ミーティング: メダカデータの表現型アノテ ーション状況確認

2015年 4月17日	チーム内ミーティング (非公開)	理研筑波研究所 研究棟 121 室	6 人	研究進捗報告のためのミーテ ィング
2015年 4月8日	研究打ち合わせ(非公 開)	理研情報基盤セ ンター	3 人	理研内 RDF データ基盤との 擦り合わせ
2015年 4月20日	表現型データベース勉 強会(非公開)	理研筑波研究所 研究棟 121 室	5 人	Rat Genome Database (RGD)
2015年 4月21日	研究打ち合わせ(非公 開)	理研情報基盤セ ンター	3 人	理研内 RDF データ基盤との 擦り合わせ
2015年 4月27日	チーム内ミーティング (非公開)	理研筑波研究所 研究棟 121 室	6 人	研究進捗報告のためのミーテ ィング
2015年 5月11日	研究打ち合わせ(非公 開)	Jackson 研究 所(米国)	5 人	MGI データの RDF 化に関す る打ち合わせ
2015年 5月18日	表現型データベース勉 強会(非公開)	理研筑波研究所 研究棟 121 室	5 人	IMPC データの統計解析方 法について
2015年 5月19日	研究打ち合わせ(非公 開)	広島大学	2 人	ゲノム編集データの RDF 化 に関する打ち合わせ
2015年 5月20日	チーム内ミーティング (非公開)	理研筑波研究所 研究棟 121 室	5 人	研究進捗報告のためのミーテ ィング
2015年 6月3日	チーム内ミーティング (非公開)	理研筑波研究所 研究棟 121 室	7 人	研究進捗報告のためのミーテ ィング
2015年 6月8日	表現型データベース勉 強会(非公開)	理研筑波研究所 研究棟 121 室	5 人	Web Ontology Language (OWL)
2015年 6月30日	表現型データベース勉 強会(非公開)	理研筑波研究所 研究棟 121 室	5 人	RDF 作成ツールの利用
2015年 7月1日	チーム内ミーティング (非公開)	理研筑波研究所 研究棟 121 室	6 人	研究進捗報告のためのミーテ ィング
2015年 7月13日	表現型データベース勉 強会(非公開)	理研筑波研究所 研究棟 121 室	5 人	統計解析環境 R
2015年 7月17日	研究打ち合わせ(非公 開)	理研バイオリソ スセンター	4 人	中国科学院との微生物 RDF に関する打ち合わせ
2015年 7月21日	チーム内ミーティング (非公開)	理研筑波研究所 研究棟 121 室	6 人	研究進捗報告のためのミーテ ィング
2015年 7月22日	研究打ち合わせ(非公 開)	東京工業大学	4 人	中国科学院、東工大、 DBCLS との微生物 RDF に 関する打ち合わせ
2015年 8月5日	チーム内ミーティング (非公開)	理研筑波研究所 研究棟 121 室	6 人	研究進捗報告のためのミーテ ィング
2015年 8月11日	研究打ち合わせ(非公 開)	Seoul National University (韓 国)	4 人	Korean Mouse Phenotyping Consortium との連携に関する打ち合わせ
2015年 9月7日	表現型データベース勉 強会(非公開)	理研筑波研究所 研究棟 121 室	5 人	SQL の初歩
2015年 9月9日	チーム内ミーティング (非公開)	理研筑波研究所 研究棟 121 室	7 人	研究進捗報告のためのミーテ ィング
2015年 9月11日	ANRRC の IT に関す る打ち合わせ	Incheon (韓 国)	6 人	アジアにおけるバイオリソ ス情報の標準化に関する意見 交換

2015年 9月24日	研究打ち合わせ(非公開)	東北大学医学部	5人	マウス表現型及び、疾患語彙の共通利用に関する意見交換
2015年 9月28日	表現型データベース勉強会(非公開)	理研筑波研究所 研究棟 121 室	5人	JCM のデータの詳細解説
2015年 9月1日	研究打ち合わせ(非公開)	理研情報基盤センター	3人	理研内 RDF データ基盤との打ち合わせ
2015年 10月7日	チーム内ミーティング(非公開)	理研筑波研究所 研究棟 121 室	6人	研究進捗報告のためのミーティング
2015年 10月16日	研究打ち合わせ(非公開)	理研横浜研究所	3人	理研内 RDF データ基盤を用いた情報統合に関する打ち合わせ
2015年 10月19日	表現型データベース勉強会(非公開)	理研筑波研究所 研究棟 121 室	5人	IMPC ウェブサイトのデータの詳細解説
2015年 10月22日	理研データベース基盤に関する意見交換	理研情報基盤センター	6人	理研内データベース基盤における RDF の利点に関する意見交換
2015年 11月4日	チーム内ミーティング(非公開)	理研筑波研究所 研究棟 121 室	6人	研究進捗報告のためのミーティング
2015年 11月6日	研究打ち合わせ(非公開)	慶応大学	3人	AMED-IRUD との連携に関する意見交換
2015年 12月11日	IMPC データ連携打ち合わせ	European Bioinformatics Institute (EBI)	5人	IMPC との RDF 連携に関する打ち合わせ
2016年 1月8日	SBI とのデータ連携打ち合わせ	システム・バイオロジー研究機構(東京)	4人	RDF を用いた、データ連携及び利用に関する意見交換、打ち合わせ
2016年 1月13日	チーム内ミーティング(非公開)	理研筑波研究所 研究棟 121 室	6人	研究進捗報告のためのミーティング
2015年 1月25日	表現型データベース勉強会(非公開)	理研筑波研究所 研究棟 121 室	4人	非生物学者のための遺伝学解説
2016年 2月9日	チーム内ミーティング(非公開)	理研筑波研究所 研究棟 121 室	6人	研究進捗報告のためのミーティング
2015年 2月22日	表現型データベース勉強会(非公開)	理研筑波研究所 研究棟 121 室	4人	MySQL の初歩
2015年 3月1日	細胞データベース打ち合わせ	慶応大学	8人	培養細胞のデータ共有に関する意見交換
2015年 3月7日	表現型データベース勉強会(非公開)	理研筑波研究所 研究棟 121 室	5人	統計解析の初歩
2016年 3月22日	チーム内ミーティング(非公開)	理研筑波研究所 研究棟 121 室	5人	研究進捗報告のためのミーティング
2016年 3月28日	表現型データベース勉強会(非公開)	理研筑波研究所 研究棟 121 室	5人	HPO レビュー
2016年 4月5日	チーム内ミーティング(非公開)	理研筑波研究所 研究棟 121 室	8人	研究進捗報告のためのミーティング
2016年 4月11日	表現型データベース勉強会(非公開)	理研筑波研究所 研究棟 121 室	5人	FM アドバンスの機能について

2016年 4月12日	チーム内ミーティング (非公開)	理研筑波研究所 研究棟 121 室	8 人	研究進捗報告のためのミーテ ィング
2016年 5月17日	チーム内ミーティング (非公開)	理研筑波研究所 研究棟 121 室	7 人	研究進捗報告のためのミーテ ィング
2016年 5月24日	チーム内ミーティング (非公開)	理研筑波研究所 研究棟 121 室	7 人	研究進捗報告のためのミーテ ィング
2016年 5月31日	チーム内ミーティング (非公開)	理研筑波研究所 研究棟 121 室	7 人	研究進捗報告のためのミーテ ィング
2016年 6月13日	Monarh Initiative と のデータ連携打ち合わせ(非公開)	ホテルたちばな や(BH2016 会 場)	4 人	Monarch Intiative と J-Phenome のデータ連携に 関する電話会議
2016年 6月21日	チーム内ミーティング (非公開)	理研筑波研究所 研究棟 121 室	7 人	研究進捗報告のためのミーテ ィング
2016年 7月5日	チーム内ミーティング (非公開)	理研筑波研究所 研究棟 121 室	7 人	研究進捗報告のためのミーテ ィング
2016年 7月11日	表現型データベース勉 強会(非公開)	理研筑波研究所 研究棟 121 室	5 人	MySQL について
2016年 7月12日	チーム内ミーティング (非公開)	理研筑波研究所 研究棟 121 室	8 人	研究進捗報告のためのミーテ ィング
2016年 7月19日	チーム内ミーティング (非公開)	理研筑波研究所 研究棟 121 室	7 人	研究進捗報告のためのミーテ ィング
2016年 7月23日	Neuroinformatics International 日本ノ ードとの打ち合わせ(非 公開)	湘南国際村 (NIJC ハッカソ ン会場)	3 人	画像メタデータに関する打ち 合わせ
2016年 7月27日	チーム内ミーティング (非公開)	理研筑波研究所 研究棟 121 室	7 人	研究進捗報告のためのミーテ ィング
2016年 8月2日	チーム内ミーティング (非公開)	理研筑波研究所 研究棟 121 室	7 人	研究進捗報告のためのミーテ ィング
2016年 8月8日	チーム内ミーティング (非公開)	理研筑波研究所 研究棟 121 室	7 人	研究進捗報告のためのミーテ ィング
2016年 9月13日	チーム内ミーティング (非公開)	理研筑波研究所 研究棟 121 室	7 人	研究進捗報告のためのミーテ ィング
2016年 9月27日	チーム内ミーティング (非公開)	理研筑波研究所 研究棟 121 室	7 人	研究進捗報告のためのミーテ ィング
2016年 10月3日	表現型データベース勉 強会(非公開)	理研筑波研究所 研究棟 121 室	5 人	MySQL について
2016年 10月4日	チーム内ミーティング (非公開)	理研筑波研究所 研究棟 121 室	7 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)	理研筑波研究所 研究棟 121 室	7 人	研究進捗報告のためのミーテ ィング
2016年 11月8日	チーム内ミーティング (非公開)	理研筑波研究所 研究棟 121 室	7 人	研究進捗報告のためのミーテ ィング
2016年 11月15日	チーム内ミーティング (非公開)	理研筑波研究所 研究棟 121 室	7 人	研究進捗報告のためのミーテ ィング
2016年 11月29日	チーム内ミーティング (非公開)	理研筑波研究所 研究棟 121 室	7 人	研究進捗報告のためのミーテ ィング

2016年 12月6日	チーム内ミーティング (非公開)	理研筑波研究所 研究棟 121 室	6 人	研究進捗報告のためのミーテ ィング
2016年 12月13日	チーム内ミーティング (非公開)	理研筑波研究所 研究棟 121 室	7 人	研究進捗報告のためのミーテ ィング
2016年 12月20日	チーム内ミーティング (非公開)	理研筑波研究所 研究棟 121 室	7 人	研究進捗報告のためのミーテ ィング
2016年 12月27日	チーム内ミーティング (非公開)	理研筑波研究所 研究棟 121 室	7 人	研究進捗報告のためのミーテ ィング

## 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
2014年 4月18~19 日	理化学研究所 バイオリ ソースセンター 一般公 開 ミニ講座「生き物の 研究でかつやくするコン ピューター」	理研筑波研究 所 バイオリソ ース棟小会議 室	30 人	一般の方向けに、生物学研 究におけるデータベースの重 要性を講義。
2014年 8月29日	The 3rd RIKEN BRC/NANJING UNIVERSITY MARC International Summer Intensive Course of the Mouse	理研筑波研究 所 森脇和郎 ホール	20 人	マウス研究における表現型統 合データベースの重要性と、 主なデータベースの利用方法 について講義
2015年 7月29日	The 4th RIKEN BRC / NANJING UNIVERSITY MARC International Short Summer Course of the Mouse	南京大 学、 Model Animal Resource Center	100 人	マウス研究における表現型統 合データベースの重要性と、 主なデータベースの利用方法 について講義
2016年 4月22~23 日	理化学研究所 バイオリ ソースセンター 一般公 開 ミニ講座「生き物の 研究でかつやくするコン ピューター」	理研筑波研究 所 バイオリソ ース棟小会議 室	30 人	一般の方向けに、生物学研 究におけるデータベースの重 要性を講義。
2016年 7月27日	The 5 <sup>th</sup> RIKEN BRC-Nanjing University MARC Mouse Resource Workshop Joint with RIKEN Symposium	理研筑波研究 所 森脇和郎 ホール	20 人	マウス研究における表現型統 合データベースの重要性の 講義と、J-phenome データベ ースの利用方法実習

以上

別紙 研究開発対象のデータベース等

No.	正式名称	別称	概要	URL	公開日	状態	分類	生命科学系データベースアーカイブ	NBDCヒトデータベース	NBDC RDFポータル	関連文献 (論文リストに記載があれば、その番号でも可)
1	BRC マウスリソース表現型メタデータ		生物遺伝材料としてのマウス系統を収録したデータベースです。系統の持つ遺伝子の変異や表現型が公共データやオントロジーにリンクされています。ここに登録されているマウス系統は、理研バイオリソースセンターより提供されています。	http://metadb.riken.jp/db/rikenbrc_mouse	2015年6月1日	継続・発展	データベース等	公開済	対象外	調整中	Yoshiki A, Ike F, Mekada K, Kitaura Y, Nakata H, Hiraiwa N, Mochida K, Ijuin M, Kadota M, Murakami A, Ogura A, Abe K, Moriwaki K, Obata Y. The mouse resources at the RIKEN BioResource center. Exp Anim. 2009 Apr;58(2):85-96.
2	BRC細胞リソース特性メタデータ		生物遺伝材料としての培養細胞株を収録しています。ヒトやマウス等の哺乳類をはじめ様々な生物種にわたる、多種の細胞株を公開しています。ここに登録されている細胞株は、理研バイオリソースセンターより提供されています。	http://metadb.riken.jp/db/rikenbrc_cell	2015年6月1日	継続・発展	データベース等	調整中	対象外	調整中	Nakamura Y. Exp Anim. 2010;59(1):1-7. Review. Bio-resource of human and animal-derived cell materials.
3	IMPC RDFデータ		国際マウス表現型解析コンソーシアム(IMPC)が収集するノックアウトマウスの網羅的表現型データのRDFメタデータです。ES細胞から作製したノックアウトマウスの表現型解析に関連する解析手法、表現型アノテーション、解析対象遺伝子型などのデータがRDF化されています。	http://metadb.riken.jp/metadb/db/IMPC_RDF	2015年10月14日	新規	データベース等	調整中	対象外	調整中	Dickinson ME, Flenniken AM, Ji X, Teboul L, et al High-throughput discovery of novel developmental phenotypes. Nature. 2016 Sep 22;537(7621):508-514. doi: 10.1038/nature19356. Kobayashi N, Lenz K and Masuya H, RIKEN MetaDatabase: a database platform as a microcosm of linked open data cloud in the life sciences, Joint International Semantic Technology Conference JIST 2016: Semantic Technology pp 99-115, id:21, 2016 (DOI: 10.1007/978-3-319-50112-3_8)
4	ACGG-DB: ノックアウトマウスを用いた機能糖鎖科学データベースのメタデータ		Asian Community of Glycoscience and Glycotechnology-Database (ACGG-DB)の一部である「KO Mouse DB」(http://integbio.jp/dbcatalog/record/nbdc01172)のRDFメタデータです。「機能糖鎖科学研究のためのKOマウス系統」と「表現型アノテーション」について、RDFおよびスプレッドシート形式のデータがダウンロード可能です。	http://metadb.riken.jp/metadb/db/Glycomics_mouse	2015年10月1日	新規	データベース等	調整中	対象外	調整中	
5	JCM微生物リソースメタデータ		理研バイオリソースセンター 微生物材料開発室(JCM)が保有する微生物株について、株の特性や表現型情報と培養状況を収録しています。ここに登録されている微生物株は、理研バイオリソースセンターより提供されています。	http://metadb.riken.jp/db/rikenbrc_jcm_microbe	2015年6月1日	継続・発展	データベース等	調整中	対象外	公開済	
6	J-phenome	A portal of phenotype data analyzed in Japan	遺伝子の多様性の結果として現れる生物の表現型情報を集約したポータルサイトです。バイオサイエンスデータベースセンター(NBDC)における統合化推進プログラムの採択課題「生命と環境のフェノーム統合データベース」の成果です。哺乳類、シロイヌナズナ、微生物など、幅広い研究コミュニティから収集した情報を、研究分野の垣根を超えて標準化・統合化・体系化し、オープンに公開しています。	http://jphenome.info	2015年5月15日	新規	データベース等	公開済	対象外	対象外	Terue Takatsuki, Mikako Saito, Sadahiro Kumagai, Eiki Takayama, Kazuya Ohshima, Nozomu Ohshiro, Kai Lenz, Nobuhiko Tanaka, Norio Kobayashi and Hiroshi Masuya, A RDF based Portal of Biological Phenotype Data produced in Japan, Proceedings of the 15th International Semantic Web Conference, P25, 2016 (http://ceur-ws.org/Vol-1690/paper10.pdf)
7	シロイヌナズナフェノームデータベース		文献キュレーションによって収集したシロイヌナズナフェノーム情報および理研内で開発されたシロイヌナズナのバイリソース(変異体)のフェノームを統合しました。	http://jphenome.info/?page_id=95	2015年6月1日	継続・発展	データベース等	調整中	対象外	調整中	
8	NBRPメダカ表現型メタデータ		メダカ系統のリソースデータベースであるNBRP Medaka (http://integbio.jp/dbcatalog/record/nbdc00147)の表現型データのメタデータです。RDFおよび、スプレッドシート形式のデータがダウンロードページから利用できます。	http://metadb.riken.jp/metadb/db/NBRP_medaka	2015年7月16日	継続・発展	データベース等	調整中	対象外	調整中	Sasado T1, Tanaka M, Kobayashi K, Sato T, Sakaizumi M, Naruse K. The National BioResource Project Medaka (NBRP Medaka): an integrated bioresource for biological and biomedical sciences. Exp Anim. 2010;59(1):13-23. Review.
9	NBRPラット表現型メタデータ		ライフサイエンス分野で広く用いられている哺乳類動物モデルである、ラット系統の持つ遺伝子の変異や表現型が公共データやオントロジーにリンクされて整理されています。ここに登録されているマウス系統は、NBRPプロジェクト(京都大学)より提供されています。	http://metadbdev.riken.jp/db/NBRP_rat	2016年7月	新規	データベース等	調整中	対象外	調整中	Serikawa T, Mashimo T, Takizawa A, Okajima R, Maedomari N, Kumafuji K, Tagami F, Neoda Y, Otsuki M, Nakanishi S, Yamasaki K, Voigt B, Kuramoto T. "National BioResource Project-Rat and related activities." Exp Anim. 2009 Jul;58(4):333-41. (PMID: 19654430)
10	遺伝研 マウス表現型メタデータ		実験用マウスの遺伝的背景に、野生由来系統MSMの染色体を導入した「MSM-B6コンソミックマウス系統」の持つ遺伝子の変異や表現型が公共データやオントロジーにリンクされて整理されています。ここに登録されているマウス系統は、国立遺伝学研究所より提供されています。	http://metadbdev.riken.jp/db/Nig_consomic_mouse	2016年7月	新規	データベース等	調整中	対象外	調整中	Takada T, Mita A, Maeno A, Sakai T, Shitara H, Kikkawa Y, Moriwaki K, Yonekawa H, Shiroishi T. "Mouse inter-subspecific consomic strains for genetic dissection of quantitative complex traits." Genome Res. 2008 Mar;18(3):500-8. (doi: 10.1101/gr.7175308.)
11	zTrap遺伝子発現メタデータ		脊椎動物の発生・器官形成・行動等の複雑な生命現象を遺伝学的に研究するためのモデル動物、小型熱帯魚ゼブラフィッシュ(Danio rerio)におけるマーカー遺伝子発現公共データやオントロジーにリンクされて整理されています。ここに登録されている系統は、国立遺伝学研究所より提供されています。	http://metadbdev.riken.jp/db/zTrap	2016年12月	新規	データベース等	調整中	対象外	調整中	Kawakami K, Abe G, Asada T, Asakawa K, Fukuda R, Ito A, Lal P, Mouri N, Muto A, Suster ML, Takakubo H, Urasaki A, Wada H, Yoshida M. "zTrap: zebrafish gene trap and enhancer trap database." BMC Dev Biol. 2010 10:105. (doi: 10.1186/1471-213X-10-105.)
12	Japan Cre Resource and Expression Database	JCRED	Creマウス系統は、部位や時期特異的に遺伝子の機能を不活性化化するコンディショナルKO作成に必要なマウス系統です。Cre導入遺伝子の発現情報が、公共データやオントロジーによって整理されています。ここに登録されているマウス系統は、熊本大学、および理研バイオリソースセンターより提供されています。	http://jcred.info	2016年12月	新規	データベース等	調整中	対象外	調整中	
13	Rare disease vocabulary in Japanese	RDVJ	厚生労働省指定難病と難病オントロジーORDOとの関連性データ:約400疾患。厚生労働省の指定する130難病に含まれる400疾患について、国際的に用いられている難病オントロジーORDOとの関連性を整理した。	http://metadbdev.riken.jp/ontology/RDVJ	2017年1月	新規	データベース等	調整中	調整中	調整中	
14	J-Phenome Disease Model Finder		異常の起こる器官組織、形質、遺伝子の相同性を通じて、ヒト疾患(複合表現型)と、モデル生物の表現型の生物学的対応を可視化するアプリケーション。	http://diseasemodel.riken.jp/	2016年12月	新規	ツール等	調整中	対象外	対象外	



No.	正式名称	別称	概要	URL	公開日	状態	分類	生命科学系データベースアーカイブ	NBDCヒトデータベース	NBDC RDFポータル	関連文献 (論文リストに記載があれば、その番号でも可)
15	J-Phenome Genome Editing Organism Registry		ゲノム編集生物を対象として一般研究者から直接表現型データ登録を受け付けるシステム。	未定	2017年3月	新規	ツール等	調整中	対象外	対象外	
16	Stanza for Paging Table		ページング付きテーブル型汎用スタンザ。公開エンドポイントのRDFデータを、任意のテーブル型で可視化する汎用ツール。	未定	2017年3月	新規	ツール等	調整中	対象外	対象外	