

平成 2 6 年 5 月 3 0 日現在

機関番号： 8 2 4 0 1

研究種目： 基盤研究(B)

研究期間： 2011 ~ 2013

課題番号： 2 3 3 0 0 1 6 1

研究課題名（和文）マウス・ラット表現型の情報統合のための基盤技術の確立

研究課題名（英文）Development of fundamental technology for phenotype data integration in mouse and rat

研究代表者

榎屋 啓志（Mausuya, Hiroshi）

独立行政法人理化学研究所・バイオリソースセンター・ユニットリーダー

研究者番号： 4 0 3 2 1 8 1 4

交付決定額（研究期間全体）：（直接経費） 15,500,000 円、（間接経費） 4,650,000 円

研究成果の概要（和文）：マウス・ラットにおける個体レベルの特性情報を統合整理し、疾患との繋がりを示すような情報統合技術の開発を行った。予め定義された知識リンクを辿り、検査値と疾患を関連づけるアルゴリズムを開発し、これを用いて、1)モデル動物の表現型から関連する疾患の探索、2)疾患から、関連するモデル動物を探索できるデータベースを作成した。生物種、表現型判定基準、疾患など全ての元データは定義変更と追加が可能で、生物横断的な大規模疾患探索データベースへの発展が期待される。

研究成果の概要（英文）：This study aimed to develop a data-integration technology to elucidate hidden relationships between phenotype data of experimental animals and human diseases. We developed an algorithm to relate measurement data to disease via multiple knowledge libraries (ontologies). This algorithm enables to build data base system which can explore 1) related diseases from animal phenotypes, and 2) animal phenotypes from diseases. In this system data of biological species, criteria of phenotype assessment and disease are flexibly defined. It is expected that basic concept of this algorithm will be utilized in the development of the cross-species database which elucidates phenotype-disease relationships from large-scale data set.

研究分野： 総合領域

科研費の分科・細目： 実験動物学、実験動物学

キーワード： データベース 表現型 疾患 マウス ラット バイオリソース 情報統合 オントロジー

1. 研究開始当初の背景

全ゲノム解読、遺伝子改変、大規模解析技術の向上により、疾患モデル研究は研究リソースを再整備する新たな局面を迎えている。欧、米、カナダを中心とした国際マウス大規模表現型解析コンソーシアム (International Mouse Phenotyping Consortium: IMPC) は、全遺伝子の KO マウスの基本表現型を網羅的に解析し、新たな研究プラットフォームを創出しようとしている。

国内では、マウス・ラットの多くのリソースが開発されており、リソース管理基盤が整備されつつある。また、情報技術面では、生物・医学情報を統合的に扱うことを目的として、遺伝子オントロジー (Gene Ontology: GO) に代表される情報統合の技術開発が進行している。

しかしながら、マウス・ラットの特性情報に注目し、精度良く相互運用する情報技術は実用化されておらず、優れた国内技術を用いてこの課題に取り組み、国内をはじめ国際的にも貢献する必要がある。

2. 研究の目的

マウスおよびラットは、遺伝子改変、生殖工学をはじめ様々な技術を用いて実験解析が可能な疾患モデル動物として、疾患研究における極めて重要な役割を担っている。今後、ゲノム情報や大規模解析技術の向上によって機能解析の爆発的進展が見込まれており、これらの情報の提供により疾患や創薬研究に対して大きく貢献することが期待されている。しかし、これらのニーズに十分にこたえるためには、情報面の整備、つまり、多種多様な情報を整理統合し、互いの関連性が明らかになるようなかたちで研究コミュニティへ情報提供することが必須である。

本課題では、マウス・ラットにおける個体レベルの特性情報をデータベース化し、生物種や、実験条件、目的など、様々な違いを明示しながら自動的に連結するアルゴリズムの開発、検証を行った。また、ヒト疾患情報データとの自動的な関連づけを行う実用的な情報提供データベースとしての実装を試みた。

3. 研究の方法

本課題では下記の(1)~(3)を行い、表現型統合のアルゴリズムと、表現型と疾患の関連づけができるデータベースシステムを開発することを計画した。

(1: 1 年目) データベースの基本設計。先行研究として発表されている、オントロジー研究の成果に基づき、モデル動物の表現型や、疾患といった、生物の特性情報をデータとしてモデル化する方法論、生物種の違いと、器官臓器や、遺伝子の相同性による相互関連性を扱う枠組みを設計する。

(2: 2 年目) プロトタイプデータベースによる検証実験。連携研究者の研究機関で公開されているマウスおよびラットの特性データを、上記データベースに格納し、50 種類前後の疾患をモデルケースとして正しく連結されるかの検証実験を行う。

(3: 3 年目) (2) の結果を(1) に反映させ、正しい関連づけができるよう調整を行うとともに、先行研究として発表されている臨床医学オントロジー (<http://lodc.med-ontology.jp>) より、大規模に疾患情報を導入することで、モデル動物の表現型と疾患とを結ぶ関連データベースを開発する。

4. 研究成果

平成 23 年度は、データベースの基本設計および、プロトタイプ作成を行った。

1) データベースの基本設計

先行研究の成果に従い、大阪大学で開発されているオントロジー記述言語および構築ツール「法造 (<http://www.hozo.jp/hozo/>)」を用いて、性質の分類 (表現型、遺伝型を含む)、測定尺度、生物種、個体、生物群、解剖学的部位、定性値の大小関係を決定する文脈などを定義した。

2) プロトタイプ開発

上記の設計に従って格納されたデータは、複雑な相互関係によって記述されているため、シンプルな表形式で閲覧できるデータベース閲覧ソフトウェアのプロトタイプを作成した。テストデータには、NIG Mouse Phenotype Database (<http://molossinus.lab.nig.ac.jp/phenotype/index.html>)、NBRP Rat Database (http://www.anim.med.kyoto-u.ac.jp/nbr/default_jp.aspx) の一部を用いた。表現型データは、測定対象と測定項目のマトリクスで示され、各項目はデータベースの情報に従って分類表示された。また、データベースに登録してある複数組の定性値の判定基準に沿って、定量値を複数の種類の定性値に変換することができた。さらに、測定対象および測定項目に付随するメタデータを表形式で出力することができた。

H24 年度は、前年度に行ったデータベースの基本設計およびプロトタイプを拡張して、ヒト疾患データを導入し、マウスおよびラットの表現型から関連疾患を検索する基本プログラムを作成した。

1) ヒト疾患データの導入

東京大学にて開発された疾患定義と、疾患に含まれる病態 (異常状態) の推移のデータベースである臨床医学オントロジーを本シ

このデータベースでは、全てのデータはフレキシブルに定義可能であり、モデル生物種とその類縁関係を新たに増やしたり、正常/異常の表現型を判定する基準を複数導入したり、新たな疾患定義を増やして多くの情報に関連づけたりといったことが可能である。この技術を発展させることで、モデル生物の表現型データと疾患とをより広く関連づける大規模データベースの開発が可能になると

期待される。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

〔雑誌論文〕(計 8 件)

- 1) 榎屋啓志, 溝口理一郎 “遺伝学オントロジー”, 人工知能学会論文誌, 査読有, 29: 2014 in press
- 2) Makita Y, Kobayashi N, Yoshida Y, Doi K, Mochizuki Y, Nishikata K, Matsushima A, Takahashi S, Ishii M, Takatsuki T, Bhatia R, Khadbaatar Z, Watabe H, Masuya H, Toyoda T. “PosMed: ranking genes and bioresources based on semantic web association study”, Nucl Acids Res, 査読有, 41:W109-W114, 2013, DOI:10.1093/nar/gkt474
- 3) Suzuki T, Furuse T, Yamada I, Motegi H, Kozawa Y, Masuya H, Wakana S. “Pheno-Pub: a total support system for the publication of mouse phenotypic data on the web”, Mamm Genome, 査読有, 24:473-483, 2013, DOI:10.1007/s00335-013-9482-y
- 4) Toki H, Inoue M, Motegi H, Minowa O, Kanda H, Yamamoto N, Ikeda A, Karashima Y, Matsui J, Kaneda H, Miura I, Suzuki T, Wakana S, Masuya H, Gondo Y, Shiroishi T, Akiyama T, Yao R, Noda T. “Novel mouse model for Gardner syndrome generated by a large-scale N-ethyl-N-nitrosourea mutagenesis program”, Cancer Sci, 査読有, 104:937-44, 2013, DOI: 10.1111/cas.12161
- 5) Masuya H. “Roles and applications of biomedical ontologies in Experimental Animal Science”, Exp Anim, 査読無, 61:365-373, 2012, DOI: 10.1538/expanim.61.365
- 6) 豊田哲郎, 榎屋啓志 “フェノーム統合データベース”, 細胞工学, 査読無, 31:240-243, 2012
- 7) Furuse T, Yamada I, Kushida T, Masuya H, Miura I, Kaneda H, Kobayashi K, Wada Y, Yuasa S, Wakana S. “Behavioral and neuromorphological characterization of a novel Tubal mutant mouse”,

Behav Brain Res., 査読有, 227:167-174, 2012, DOI: 10.1016/j.bbr.2011.11.002

- 8) Furuichi T, Masuya H, Murakami T, Nishida K, Nishimura G, Suzuki T, Imaizumi K, Kudo T, Ohkawa K, Wakana S, Ikegawa S. “ENU-induced missense mutation in the C-propeptide coding region of Col2a1 creates a mouse model of platyspondylic lethal skeletal dysplasia, Torrance type”, Mamm Genome, 査読有, 22:318-328, 2012, DOI: 10.1007/s00335-011-9329-3

〔学会発表〕(計 25 件)

- 1) 榎屋啓志 “マウス・ラットの表現型から関連疾患を探すデータベース”, 第 7 回ラットリソースリサーチ研究会, 2014 年 1 月 31 日, 京都
- 2) Masuya H. “Overview of current pipeline in RIKEN BioResource Center”, IMPC Phenotyping Workshop, 2013 年 10 月 28 日, Toronto, Canada
- 3) 榎屋啓志, 古崎晃司, 大江和彦, 溝口理一郎 “コンテキストに依存した定性値を扱う生物表現型統合データベースの試作”, 2013 年度人工知能学会 全国大会 (第 27 回), 2013 年 6 月 6 日, 富山
- 4) 榎屋啓志, 田中信彦, 鈴木智広, 若菜茂晴 “Examinations of Phenotyping Data in Japan Mouse Clinic”, 第 60 回日本実験動物学会総会, 2013 年 5 月 17 日, つくば
- 5) 榎屋啓志, 古崎晃司, 高田豊行, 真下知土, 若菜茂晴, 大江和彦, 溝口理一郎 “マウス・ラット表現型から疾患を推論するソフトウェアの開発”, 第 60 回日本実験動物学会総会, 2013 年 5 月 15 日, つくば
- 6) 田中信彦, 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, 土岐秀明, 井上麻紀, 美野輪治, 野田哲生, 若菜茂晴, 榎屋啓志 “網羅的マウス表現型解析データを用いたデータマイニング:ワークフローの検討”, 第 35 回日本分子生物学会年会, 2012 年 12 月 14 日, 福岡
- 7) 榎屋啓志, 高月照江, 斎藤実香子, 松庫義弘, 蒔田由布子, 吉田有子, 望月芳樹, 小林紀郎, 吉木淳, 中村幸夫, 豊田哲郎, 小幡裕一 “哺乳類バイオリソースの

Linked Open Data データベースのアップデート”, 第 35 回日本分子生物学会年会, 2012 年 12 月 12 日, 福岡

- 8) Masuya H, Takatsuki T, Makita Y, Yoshida Y, Mochizuki Y, Kobayashi N, Yoshiki A, Nakamura Y, Toyoda T, Obata Y. “Development of Linked Open Data for bioresources”, JIST2012: The 2nd Joint International Semantic Technology Conference, 2012 年 12 月 3 日, 奈良
- 9) Masuya H. “A trial of the development of Linked Open Data (LOD) of bioresources”, ANRRC 2012 The 4th ANRRC International Meeting, 2012 年 10 月 19 日, Jeju, Korea
- 10) Masuya H. “Internationally cooperative development of the informatics infrastructure for IMPC”, IMPC International Symposium in Tokyo, 2012 年 9 月 28 日, 東京
- 11) 榎屋啓志, 溝口理一郎 “遺伝子の意味を情報モデルとして記述する”, 日本遺伝学会第 84 回大会, 2012 年 9 月 24 日, 福岡
- 12) Masuya H, Mizoguchi R. “An Ontology of Gene”, ICBO & FOIS 2012, 2012 年 7 月 23 日, Graz, Austria
- 13) 榎屋啓志 “マウス表現型の情報統合へ向けた研究”, 第 26 回モロシヌス研究会, 2012 年 6 月 16 日, 東京
- 14) 榎屋啓志, 溝口理一郎 “遺伝子のオントロジー”, 2012 年度人工知能学会全国大会 (第 26 回), 2012 年 6 月 12 日, 山口
- 15) 榎屋啓志, 田中信彦, 脇和規, 高月照江, 齋藤実香子, 小幡裕一 “理研 BRC IMPC (国際マウス表現型解析コンソーシアム) 参画計画 3 -表現型解析情報の公開と統合-”, 日本実験動物科学・技術九州 2012, 2012 年 5 月 26 日, 別府
- 16) 田中信彦, 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, 土岐秀明, 井上麻紀, 美野輪治, 野田哲生, 若菜茂晴, 榎屋啓志 “マウス表現型の類似度解析ツールの開発: ワークフローの検討”, 日本実験動物科学・技術九州 2012, 2012 年 5 月 26 日, 別府
- 17) 鈴木智広, 古瀬民生, 小澤恵代, 山田郁子, 田中信彦, 榎屋啓志, 若菜茂晴 “日

本マウスクリニック表現型情報公開サイト「Pheno-Pub」の改良” 日本実験動物科学・技術九州 2012, 2012 年 5 月 25 日, 別府

- 18) Masuya H. “MUSDB: LIMS for Japan Mouse Clinic”, IMPC Informatics Workshop, 2012 年 5 月 23 日, Washington DC, USA
- 19) 榎屋啓志 “The Ontological Representation of Genetics”, 5th Interdisciplinary Ontology Conference (InterOntology12), 2012 年 2 月 23 日, 東京
- 20) Masuya H, “Semantic Web 技術を利用した哺乳類バイオリソースのデータベース”, 日本分子生物学会第 34 回年会, 2011 年 12 月 15 日, 横浜
- 21) 榎屋啓志, “遺伝学オントロジーを基盤にしたバイオリソース情報統合”, 日本遺伝学会第 83 回大会, 2011 年 9 月 22 日, 京都
- 22) Masuya H, “An Advanced Strategy for Integration of Biological Measurement Data”, International Conference on Biomedical Ontology, 2011 年 7 月 29 日, Buffalo, USA
- 23) 榎屋啓志, “マウス・ラットのフェノーム情報統合へ向けて”, 2011 年 7 月 8 日, 新潟
- 24) Masuya H, “Recent progress of IT in JMC and AMPC, 1st Asian Mouse Phenotyping Consortium Council Meeting”, 1st Asian Mouse Phenotyping Consortium Council Meeting (招待講演), 2011 年 6 月 13 日, Seoul, Korea
- 25) 榎屋啓志, “マウスでの表現型情報の国際的統合の動向について”, 第 58 回日本実験動物学会総会 (招待講演), 2011 年 5 月 26 日, 東京

〔図書〕(計 2 件)

- 1) 榎屋啓志, “表現型データの国際標準化とデータの利用 (書名: 論文ができてしまう! 疾患モデルマウス表現型解析指南)” 中山書店, 2012, 348-359
- 2) 榎屋啓志, “第三節 バイオインフォマティクス (書名: 生物機能モデルと新しいリソース・リサーチツール, series モデル動物利用マニュアル)”エル・アイ・シ

〔その他〕

ホームページ等

PhenosSITE

(http://ja.brc.riken.jp/lab/bpmp/index_j.html)

PATO2YAMATO オントロジー

(http://ja.brc.riken.jp/lab/bpmp/ontology/ontology_pato2yato_j.html)

6. 研究組織

(1) 研究代表者

榎屋 啓志 (MASUYA, Hiroshi)

独立行政法人理化学研究所・バイオリソースセンター・ユニットリーダー

研究者番号：4 0 3 2 1 8 1 4

(2) 研究分担者

溝口 理一郎 (MIZOGUCHI, Riichiro)

北陸先端科学技術大学院大学・サービスサイエンス研究センター・教授

研究者番号：2 0 1 1 6 1 0 6

古崎 晃司 (KOZAKI, Kouji)

大阪大学・産業科学研究所・准教授

研究者番号：0 0 3 6 2 6 2 4

(平成 25 年度より研究分担者)

(3) 連携研究者

若菜 茂晴 (WAKANA, Shigeharu)

独立行政法人理化学研究所・バイオリソースセンター・チームリーダー

研究者番号：9 0 1 9 2 4 3 4

高田 豊行 (TAKADA, Toyoyuki)

国立遺伝学研究所・大学共同利用機関等の部局等・助教

研究者番号：2 0 3 5 6 2 5 7

真下 知士 (MASHIMO, Tomoji)

京都大学・医学(系)研究科(研究院)・准教授

研究者番号：8 0 3 9 7 5 5 4