

令和元年5月29日現在

機関番号：82401

研究種目：挑戦的萌芽研究

研究期間：2016～2018

課題番号：16K14600

研究課題名（和文）表現形質の異常を高精度に検出可能な手法の開発

研究課題名（英文）Development of a method to detect phenotypic abnormalities with high accuracy

研究代表者

田中 信彦（Nobuhiko, Tanaka）

国立研究開発法人理化学研究所・バイオリソース研究センター・開発研究員

研究者番号：10525596

交付決定額（研究期間全体）：（直接経費） 2,900,000円

研究成果の概要（和文）：本研究は、マウスの表現型（ヒトの病気の症状に相当）間の関係性について、これまでで最高品質かつ最大規模のデータを提示した。具体的には、345種類の表現型、60種類の生物学的機能で構成される、3,686種類の表現型間の関係性を提示した。さらに、これらのデータの理解を助けるために、345種類の各表現型について、例えば、「血糖値が異常であるとき、他にどのような表現型の異常が生じるか？」をノード（表現型）とエッジ（関係性）で表示することができるアプリを開発し、インターネット上で公開した。

研究成果の学術的意義や社会的意義

本研究で得られた、実験結果に基づくマウス表現型間の関係性のデータセットは、その正確性と網羅性でこれまでの報告を圧倒するものあり、哺乳動物の表現型間の関係性についての研究者間での共通認識のための参照用データセットとして提示される。これらのデータセットは、ヒトの病気の症状の理解を深めることに役立つとともに、現在進展しているトランスオミックスの研究分野（遺伝子発現から病気の症状が現れるまでの、生体内分子を網羅的に調べる研究分野）でのフェノーム（病気の症状の総体）の参照用データセットとしての使用が期待される。

研究成果の概要（英文）：In this study, we present the largest, highest quality dataset describing relationships among phenotypes in mammals to date, alongside a detailed review of their characteristics. Features, based on 2,050 measured parameters from 113 phenotyping analyses of 3,100 mutant mouse strains, were organized into 532 ontology-annotated phenotypes, and 3,686 significant phenotypic associations were identified among them by association rule mining. This mammalian phenome-wide data resource will be useful for elucidation of mechanisms underlying phenotypes and deepening understanding of both pleiotropic gene expression and human disease pathology.

研究分野：バイオインフォマティクス

キーワード：マウス 表現型 フェノーム オミックス アソシエーション分析

様式 C - 19、F - 19 - 1、Z - 19、CK - 19 (共通)

1. 研究開始当初の背景

集団や個体の表現形質異常の検出が発端となって、様々な遺伝学的解析研究がはじまる。したがって、表現形質が異常か否かの判定は、その原因遺伝子に対する詳細な機能解析研究(2次解析)への動機として、極めて重要な意味を持つ。そこで、研究開始当初では、「表現形質の異常を高精度に検出するためのワークフローの開発」を目的とした。

2. 研究の目的

上述したように、研究開始当初は、「表現形質の異常を高精度に検出するためのワークフローの開発」を目的とした。平成29年度の後半から、よりインパクトの大きい成果が期待される、これまでの研究過程で取得した網羅的な表現型データを使用した「マウス表現型間の関係性の全体像の提示」を目的とした研究へシフトした。

3. 研究の方法

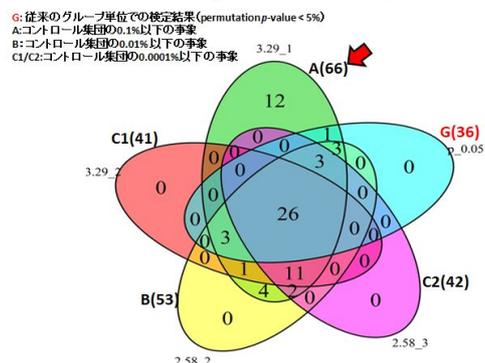
データ解析用のデータには、上述した前者の目的には、理研バイオリソース研究センターのマウス表現型解析開発チームで解析した表現型データを、後者の目的には、国際マウス表現型解析プロジェクト(IMPC)のウェブサイトからダウンロードして所得した表現型データを用いた。本研究での統計解析、結果の可視化、ウェブアプリの作成等のすべての情報学的アプローチは、統計解析ソフト R で行われた。表現型間の関係性の抽出には、アソシエーション分析を適用した。結果の可視化のためのアプリの作成には、RのShinyパッケージを用いた。

4. 研究成果

(1) 表現形質の異常を高精度に検出するためのワークフローの開発

マウスの行動、形態、生理などに関する検査(10種類、合計342の測定項目)の表現型データを用いて、変異遺伝子による異常を高感度で検出可能なワークフローの確立を目指し、統計解析を行った。多様な測定項目の一括評価を、個体単位で可能にするため、変異の程度を表す統計値(改変zスコア)を、[(**変異群 or コントロール群の各個体の値**) - (**内部コントロール群の平均値**)] / (**内部コントロール群の標準偏差**)の計算式で定義し、全個体に対して算出した。このスコアの絶対値が3.29以上、つまり、内部コントロール群の99.9%の範囲外の閾値であれば、変異群のみに見られる顕著な値であったことから、この閾値を超えるスコアが各個体でどれほどあるのかを「異常」の判定材料とした。本法を悪性リンパ腫を呈するp53変異系統に適用した結果、66の測定項目で「異常」と判定され、通常のグループ単位による比較検定での36項目よりも、約1.8倍高い「異常」の検出に成功した。

グループ(G)単位の検定による異常な測定項目(36: $p < 0.05$)との類似点・相違点

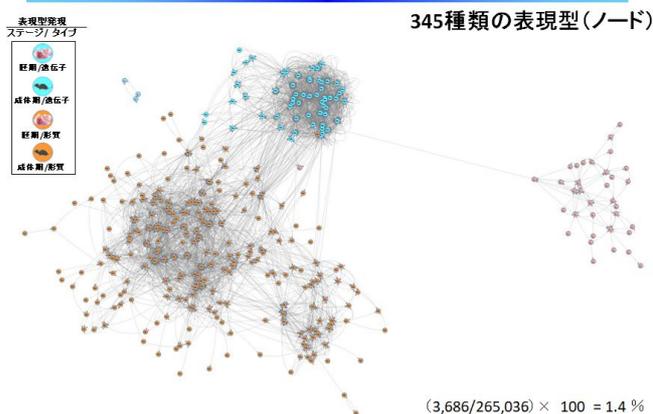


従来のグループ単位での異常の判定(36)よりも、最大で約1.8倍(66)の異常を検出

(2) 表現型間の関係性の提示

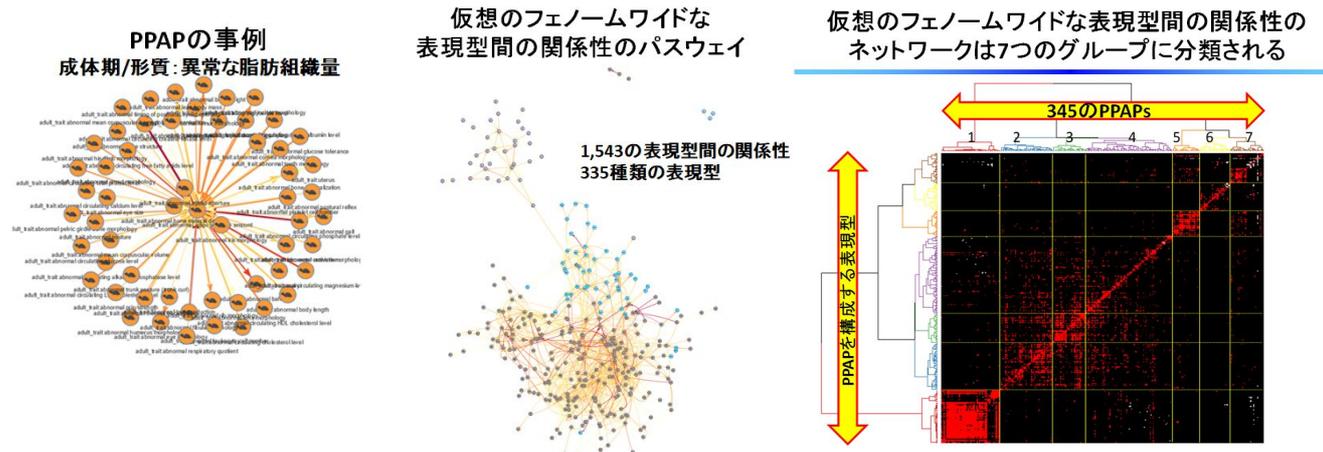
正確な表現型間の関係性を提示するため、遺伝的背景、実験手法、ワークフローやコントロール戦略などが標準化された環境下で、網羅的な表現型解析が系統的に行われ、表現型異常の判定に係る様々なバイアスの影響が低減された、国際マウス表現型解析プロジェクト(IMPC)のデータを用いた。また、表現型間の関係性を調査する際には、「正常」と「異常」の2値データ(「調査なし」のデータを含まない)のみで構成されるデータセットを作成し、これらを解析に用いることで、より正確な関係性が提示できる工夫を行った。結果として、345種類の表現型(60種類の生物学的機能)から構成される3,686種類の表現型間の関係性を明らかにした。例えば、「眼の形成が異常であるならば、歯の形成も異常」といったような関係性が提示された。これらの網羅的なマウス表現型間の関係性データは、単に表現型間の関係性の一般則を提示するだけでなく、ヒト疾患の症状の発現のメカニズムの理解に役立つことが期待される、また、様々な哺乳動物種に対するマルチオミックス研究のデータセットとして、また、ヒト疾患の発症メカニズムの解明のためのデータセットとしての利用も期待される。

3,686種類の有意な表現型間の関係性



(3) 表現型間の関係性の詳細な評価

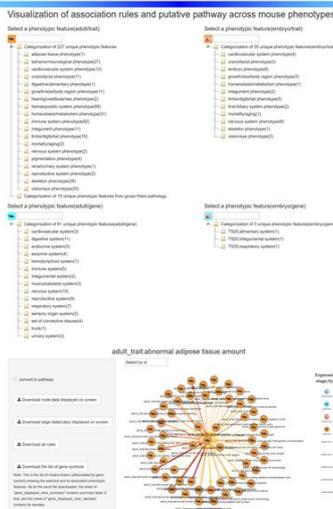
3,686 の表現型間の関係性の理解を容易にするために、このルールを構成する 345 種類の各表現型別に、表現型発現のセット (a set of phenotype-phenotype association pairs: PPAP) を定義した (下図: 左)。この 345 種類の PPAP の調査により、最大の表現型数や生物学的機能数を持つ PPAP (サブネットワーク) を同定した。さらに、表現型発現を連続的に理解するために、各 PPAP をパスウェイ (経路) の様な形状に変換する手法を開発し、変換可能な 283 種類のサブパスウェイを使って、仮想のフェノームワイドな表現型間の関係性のパスウェイを構築した (下図: 中央)。また、この PPAP を構成する表現型の類似性を基にした階層的クラスター分析により、345 種類の表現型間の関係性の全体像を提示し、これが 7 つのグループに分類されることを明らかにした (下図: 右)。



(4) アプリケーションの開発

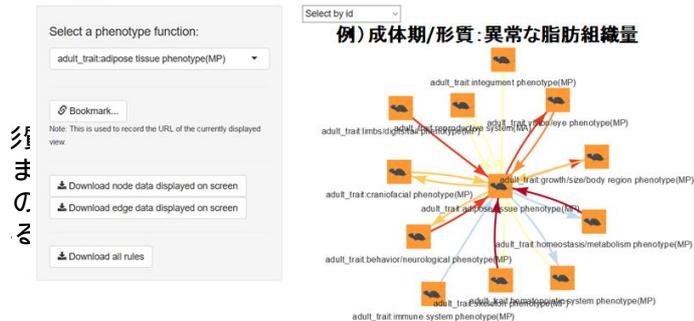
本研究で得られたマウスの表現型間及び生物学的機能間の関係性データセットの利用促進のため、ユーザーが任意に選択した表現型及び生物学的機能に関するものを可視化するアプリケーションを開発した (それぞれ、下図の左と右)。

表現型間の関係性の可視化



生物学的機能間の関係性の可視化

Enrichment analysis of the selected 3,656 association rules based on phenotype functions



5 . 主な発表論文等

[学会発表](計 4 件)

マウスの表現型間の関係性の全体像、[田中 信彦](#), [榎屋 啓志](#), 第 41 回日本分子生物学会年会、口頭発表、2018 年 11 月

マウスの表現型間の関係性の全体像、[田中 信彦](#), [榎屋 啓志](#), 第 41 回日本分子生物学会年会、ポスター発表、2018 年 11 月

マウス表現型間の関係性の全景、[田中 信彦](#), [榎屋 啓志](#), 第 40 回日本分子生物学会年会、ポスター発表、2017 年 11 月

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、[田中 信彦](#), [榎屋 啓志](#), 第 39 回日本分子生物学会年会、ポスター発表、2016 年 11 月

[図書](計 1 件)

Mouse phenome as biological resource, [Nobuhiko Tanaka](#) & [Hiroshi Masuya](#), Impact, Volume 2018, Number 12, pp. 93-95(3)、査読無、2018 年 12 月

〔その他〕

マウスの表現型間の関係性を提示するアプリケーション

https://brc-riken.shinyapps.io/phenotypic_associations_across_the_mouse_phenome/

マウスの生物学的システム（機能）間関係性を提示するアプリケーション

https://brc-riken.shinyapps.io/associations_between_biological_systems/

6．研究組織

(1)研究分担者

なし

(2)研究協力者

なし

科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属されます。