

科学研究費助成事業 研究成果報告書

平成 28 年 6 月 23 日現在

機関番号：14301

研究種目：若手研究(B)

研究期間：2013～2015

課題番号：25850185

研究課題名(和文) 公共マイクロアレイデータを用いた家畜QTL候補遺伝子の探索システム

研究課題名(英文) A system for identifying QTL candidate genes in livestock using published microarray data

研究代表者

松田 洋和 (Matsuda, Hirokazu)

京都大学・(連合)農学研究科(研究院)・助教

研究者番号：30597810

交付決定額(研究期間全体)：(直接経費) 3,100,000円

研究成果の概要(和文)：モデル生物においては、高品質な遺伝子アノテーション情報が利用でき、量的形質に關与する候補遺伝子の抽出が容易に実施される。しかしながら、家畜種においてはアノテーション情報は充実していない。近年では、マイクロアレイや次世代シーケンサー技術により網羅的な遺伝子発現データが得られ、これらのデータが複数の生物種に關して公共のデータベースに蓄積されている。本研究では、公共マイクロアレイデータを用いた家畜におけるQTL候補遺伝子の探索システムの開発を行った。ここでは、とくにマウスとウシの場合に焦点を絞って、本システムの表現型の予測および信頼性の高い遺伝子ネットワークの推定における有用性について検討を加えた。

研究成果の概要(英文)：In model organisms, high quality of gene annotation is available and it is possible to quickly identify candidate genes for the quantitative trait of interest. But its use is restricted in livestock due to poor annotation. Recently, microarray and next-generation sequencing technologies have generated high-throughput gene expression data that are collected in public database for multiple species. This study develops a system for identifying QTL candidate genes in livestock using published microarray data. Here, focusing on mouse and cattle, we investigated the usefulness of current system for predicting phenotypes and estimating a reliable gene network.

研究分野：動物遺伝育種

キーワード：QTL候補遺伝子 公共データベース 網羅的遺伝子発現データ

1. 研究開始当初の背景

今日では家畜を含む様々な生物種において、高密度 SNP を用いたゲノムワイド関連解析が実施され、経済形質に関する QTL が染色体上にマッピングされつつある。しかしながら、通常利用される高密度 SNP はあくまでマーカーであり、ここから原因遺伝子を同定するためには、検証に値する候補遺伝子の選定が必要とされる。実際に形質に影響を与える遺伝子が特定されれば、その遺伝子型情報は家系を超えて普遍的に利用することができ、より正確な育種価および表現型値の予測や、DNA 育種技術を用いた効率的な選抜が可能となる。

QTL 領域における候補遺伝子探索のプロセスについて、マウスなどのモデル動物では、遺伝子に関する豊富な情報が蓄積されデータベース化されており、これらの既知情報を用いて目的の形質に関連する可能性が高い候補が選ばれる。さらに、近交系や遺伝子改変動物を用いて、形質と遺伝子との間の関連についてより直接的な情報を得ることもできる。一方、家畜においては、モデル動物の遺伝子との相同性に基づく比較ゲノムによる情報を得ることができるものの、種の違いなどから高い信頼性は期待できない。近交系や遺伝子改変動物を利用することが困難な家畜において、原因遺伝子を特定するためには、様々な証拠に基づくより特異度の高い候補遺伝子探索法が必要とされる。

一方で、近年、マイクロアレイなどによる網羅的な遺伝子発現情報を用いた候補遺伝子の絞り込みも実施されており、形質に関連する組織や発生段階での特異的な発現の情報も利用されている。この点について、公共のデータベースに登録されている生のアレイデータは家畜のものも含めて年々増加しており、これらのデータを有効に利用することができれば、家畜における候補の特異度を高めることができると期待される。

2. 研究の目的

本研究では、このような背景を踏まえ、モデル生物であるマウスに加えて、ウシを対象として、量的形質の育種価および表現型の予測や、量的形質に関与する遺伝子遺伝子間のネットワークの推定に対して有用な QTL 候補遺伝子探索システムの開発を試みた。

3. 研究の方法

(1) Gene Expression Omnibus (<http://www.ncbi.nlm.nih.gov/geo/>) における GSE10246 から得たマウスの 91 種の細胞および組織における発現データに基づき 23,856 遺伝子の組織特異性に関するスコアを算出するとともに、これらの遺伝子の機能カテゴリーの情報等を Entrez Gene (<http://www.ncbi.nlm.nih.gov/gene>) から取得し統合した。

また、これらの情報に基づいて抽出した候

補遺伝子を用いた量的形質の表現型値の予測性能について検証するために、統合したデータとは独立な ArrayExpress (<http://www.ebi.ac.uk/arrayexpress/>) における E-MTAB-88 データセットを用いた。ここでは、約 200 個体のヘテロストックマウスに関する約 1 万 2 千箇所の SNP 型および肝臓における約 2 万の遺伝子発現値を用いて、血中総コレステロールおよび血中グルコース濃度の量的形質の表現型値の予測を実施した。

(2) ウシについて、Sequence Read Archive (<http://www.ncbi.nlm.nih.gov/sra>) の SRP002394 から得た Bovine Gene Atlas の網羅的遺伝子発現データに対して最新のゲノム情報 (UMD3.1) を基にパイオインフォマテイクス解析を実施し、15 の組織における約 2 万遺伝子の発現に関する組織特異性スコアを算出した。また、候補遺伝子探索ツールの一つとして、同じ Bovine Gene Atlas の網羅的遺伝子発現データに基づいて、遺伝子発現量の相関情報から RNA 共発現ネットワークを構築した。さらに、GeneOntology (<http://geneontology.org/>) における遺伝子の機能情報および CattleQTLdb (<http://www.animalgenome.org/cgi-bin/QTLdb/BT/index>) における既報の QTL 情報を統合した候補遺伝子探索システムを構築した。

黒毛和種直接検定牛を対象とした複数の飼料利用性関連形質のゲノムワイド関連解析を実施し、その結果に基づいて SNP 共関連ネットワークを推定した。ここで推定されたネットワークの信頼性を高める上で、RNA 共発現ネットワークとの共通ネットワークが有用であるかについて検討を加えた。

4. 研究成果

(1) 説明変数に遺伝子発現量単独または遺伝子発現量および SNP 型を同時に用いる場合を取り上げ、5 分割交差確認により表現型値の予測の正確度について検討を加えた。なお、統計的手法としては、ベイジアンラッソ法を用い、遺伝子発現データに関しては、算出した組織特異性が上位の 3,000 遺伝子に絞り込んだ場合も実施した。

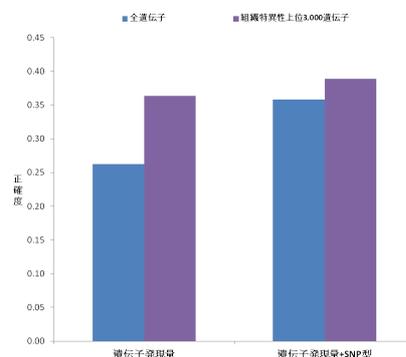


図1 血中総コレステロールの正確度

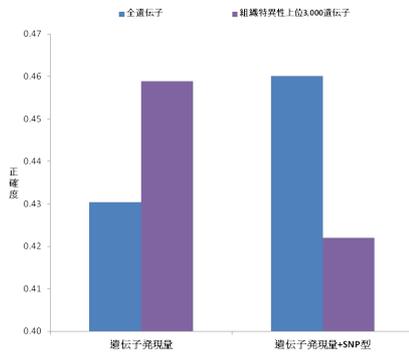


図2 血中グルコース濃度の正確度

図1は血中総コレステロールの正確度を示す。ここでは、遺伝子発現量を単独で説明変数として用いた場合および遺伝子発現量およびSNP型を同時に用いた場合の何れにおいても、組織特異遺伝子のみを用いることで予測性能の上昇が認められた。

一方で、血中グルコース濃度の正確度は図2に示すとおりである。遺伝子発現量を単独で説明変数として用いた場合はコレステロールと同様に組織特異性が上位の3,000遺伝子のみを用いることでより高い正確度が得られたが、遺伝子発現量およびSNP型を同時に用いた場合については、組織特異遺伝子のみを用いることで、全遺伝子を用いる場合に比べて正確度が低下する結果となった。このように、SNP情報を併用した場合に遺伝子選択の影響が異なったのは、各情報間の冗長性のためであると考えられたが、総じて組織特異性に基づく候補遺伝子の選択は、量的形質の表現型値の予測において、有用である可能性があると考えられた。

(2) ウシの15の組織における約2万遺伝子に関する網羅的発現データに基づいて構築したRNA共発現ネットワークが、量的形質に関する遺伝子ネットワークの信頼性を高める情報となり得るかについて、ここでは黒毛和種の飼料利用性を例とした検討を行った。

まず、黒毛和種直検牛の検定7形質のゲノムワイド関連解析の結果に基づいて推定されたSNP共関連ネットワークは、544遺伝子間の16,526の相互作用により構成された。このネットワークにおいて、有意に多く認められた遺伝子機能を有する遺伝子に限定し、それらの間の相互作用に基づくサブネットワークを抽出した(図3)。このサブネットワークにおいて、PPP3CAは報酬的側面から摂食行動を促進するドーパミン作動性ニューロンの変性に関与し、NMUR2はマウスにおいて飼料摂取および体重の制御に係ることが先行研究において確認されている。そのため、ここで推定されたSNP共関連情報のみに基づく遺伝子ネットワークについては、飼料利用性の背景にある生物学的過程を捉えていると考えられた。

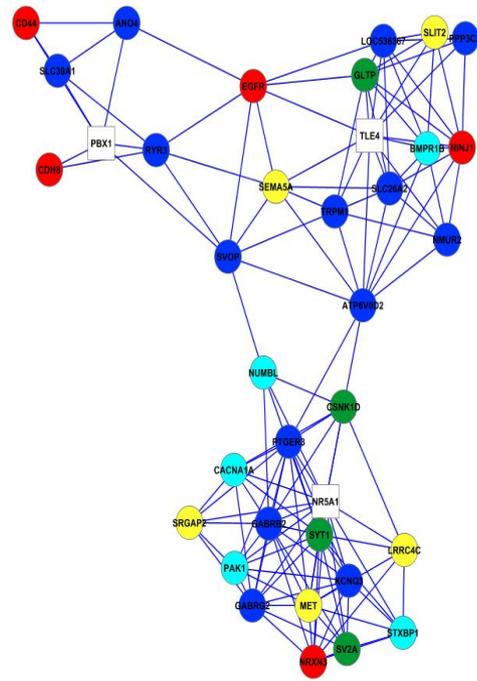


図3 機能カテゴリに属する遺伝子との間の相互作用に基づくサブネットワーク。四角形：転写因子、青色：イオンチャネル、黄色：軸索誘導、赤色：細胞接着、水色：神経分化、緑色：脳を示す。

次いで、SNP共関連ネットワーク上の遺伝子について、RNA共発現情報に基づくネットワークを推定し、さらに2つのネットワークの共通ネットワークを抽出した(図4)。推定された各ネットワークについて、生物学的ネットワークの信頼性の指標としてスケールフリー性(多くの遺伝子と接続する少数の遺伝子が存在するというネットワークの特性)を調べるとともに、データベース内の既知情報に基づくネットワークとの類似性について検討を加えた。なお、スケールフリー性の高いネットワークは、接続する遺伝子の数(次数)の分布がべき分布に従うことが期待されるが、データの適合を対数変換した次数のその確率に対する回帰の決定係数により評価した。

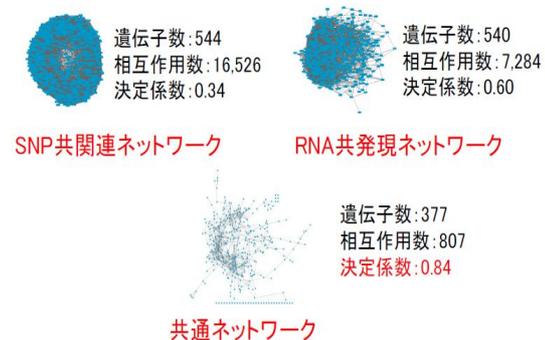


図4 SNP共関連ネットワークとRNA共発現ネットワークの共通ネットワーク

松田 洋和 (MATSUDA, Hirokazu)
京都大学・大学院農学研究科・助教
研究者番号：30597810

各ネットワークのスケールフリー性を評価した結果、共通ネットワークでは決定係数が 0.84 であることから高いスケールフリー性が示唆される一方、SNP 共関連および RNA 共発現ネットワークの決定係数はそれぞれ 0.34 および 0.60 であり、明確なスケールフリー性は認められなかった(図 4)。また、共通ネットワーク上の遺伝子について、接続する遺伝子の数が多い遺伝子に着目すると、飼料利用性との関連が知られているアンドロゲン受容体による転写を抑制する転写共役因子 URI1 や、ブタの発育に関与する可能性が指摘されている ETV6 などが認められ、対象となる量的形質に関する候補遺伝子を効率的に抽出することができた。データベースを用いた検証の結果、SNP 共関連ネットワークおよび RNA 共発現ネットワークのみならず共通ネットワークにおいても、既知のネットワークとの類似性が認められた。SNP 共関連ネットワークに RNA 共発現ネットワークの情報を組み合わせることにより、既知情報を反映したスケールフリー性の高いネットワークを推定することができると考えられた。ただし、推定されたネットワークが黒毛和種の飼料利用性にどのように関与しているかの詳細については、さらに検討が必要である。

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 1 件)

Yu Takagi, Hirokazu Matsuda, Yukio Taniguchi, Hiroaki Iwaisaki, Predicting the Phenotypic Values of Physiological Traits Using SNP Genotype and Gene Expression Data in Mice, PLOS ONE、査読有、9 巻、2014、e115532
<http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0115532> / DOI: 10.1371/journal.pone.0115532

[学会発表](計 2 件)

遠藤聡子、松田洋和、谷口幸雄、勝田智博、渡邊敏夫、杉本喜憲、祝前博明、黒毛和種直接検定牛を対象とした飼料利用性に関する遺伝子ネットワーク推定の試み、日本畜産学会第 119 回大会、2015.3.29、宇都宮大学(栃木県宇都宮市)
岡田大瑚、遠藤聡子、松田洋和、小川伸一郎、谷口幸雄、勝田智博、渡邊敏夫、杉本喜憲、祝前博明、SNP 共関連および RNA 共発現の情報を利用した遺伝子ネットワークの推定、日本畜産学会第 121 回大会、2016.3.29、日本獣医生命科学大学(東京都武蔵野市)

6. 研究組織

(1)研究代表者