

氏名（本籍）	藤木 生聖（和歌山県）
学位の種類	博士（工学）
学位授与番号	甲第83号
学位授与日付	平成29年3月24日
専攻	システム工学専攻
学位論文題目	黒毛和種における枝肉成績向上のためのタンパク質発現量のデータマイニングに関する研究
学位論文審査委員	(主査) 教授 宗森 純 (副査) 教授 吉野 孝 准教授 吉廣 卓哉

## 論文内容の要旨

和牛とは、日本在来の肉用牛のことであり、高品質な牛肉を提供するために非常に重要とされている種雄牛の血統管理がなされている。そして、肉質の良い牛を生産するために重要な、優れた遺伝能力を得るために、種雄牛の育種改良や肥育方法の改善が行われてきた。現在、冷凍精子と用いることにより、優れた遺伝子を持つ種雄牛から高品質な牛肉が継続的に生産されている。つまり、和牛の質の良さを保証する指標となる種雄牛の血統は、継続的に高品質な牛肉を生産するために必要不可欠である貴重な遺伝子供給源といえる。

和牛や種雄牛の繁殖家にとって、新たな肉用牛や種雄牛を産むために種雄牛を選択すること（つまり、冷凍精子を選択すること）は最も重要な課題の一つである。なぜなら、新たに産まれる肉用牛や種雄牛の遺伝的な特徴（以降、血統と呼ぶ）は、選択した種雄牛から生産される牛肉の質と密接な関係があるからである。したがって、繁殖家は伝統的に、高品質な牛肉を生産するため、種雄牛の遺伝的な能力を表す育種価と呼ばれる値を利用し、育種価をもとに種雄牛の選択を行っている。一般的に、育種価はBLUP (Best Linear Unbiased Prediction)法を用いて計算される。BLUP法は、種雄牛から産まれた仔牛の枝肉成績や血統情報などを用いて育種価を推定するための統計的な予測手法である。和牛の繁殖家は優れた子孫を生産するために、育種価を参考にして優秀な種雄牛を選択することによって種雄牛の遺伝的能力の改良を行ってきた。

一方、近年では、血統や肉質に密接に関係する遺伝子やタンパク質の解析技術の向上により、様々な生命現象のメカニズムが急速に解明されつつある。これは、マイクロアレイや二次元電気泳動と呼ばれる、膨大な数の遺伝子やタンパク質の発現量を一度に測定できる実験手法が確立し、これらの高処理な実験手法の恩恵によって、計算機を用いた網羅的な解析が可能になったためである。もし、高品質な牛肉を生産する仕組みが明らかになれば、例えば、(i) 複雑な生命現象を支えるタンパク質間の相互作用、(ii) 血統とタンパク質の発現の相関関係、(iii) タンパク質の発現量が肉質に影響を与える仕組みが明らかになれば、これまでの繁殖農家の勘や経験などによる牛肉の生産に代わって、科学的な根拠に基づいた肉質を改善するための効果的な方法論の開発が可能になると考えられる。

以上の背景から、本研究では、牛肉の肉質向上を目指すための基礎研究として、背景で述べた (i) , (ii) , (iii) を推定するためのデータマイニングアルゴリズムを開発する。具体的には、黒毛和種の血統データ、タンパク質発現量データ、枝肉成績データから、(i) について、3タンパク質間の組合せ的な相互作用の推定、(ii) について、血統とタンパク質発現量の相関関係の推定、(iii) について、タンパク質の発現量と表現型（表出する生物個体の特徴であり、肉質も表現型である）の間の因果関係の推定の3つの研究を行った。

まず、3タンパク質間の組合せ的な相互作用の推定として、3つのタンパク質A、B、Cの組合せにおいてAとBが同時に発現しているときにCの発現量に与える影響が、AやBがそれぞれ単体でCの発現量に与える影響に対してはるかに大きい場合をタンパク質間相互作用のモデルを想定し、3タンパク質間の組合せ的な相互作用を、条件付き確率とz値による統計的指標を用いて推定する手法を提案した。評価として、提案手法を実際の発現量データに適用した結果、今回適用した発現量データには、組合せ的な相互作用の関係にあるタンパク質の組み合わせが多く含まれていることが示唆された。

また、組合せ的な相互作用の推定手法における計算時間を改善するための高速化を行った。組合せ的な相互作用の推定手法では、統計的指標として組合せ的な作用が存在しない場合の相互作用の確率分布を用いる必要があった。そして、この統計的指標をコンピュータシミュレーションによって作成していたため、計算時間が膨大になるという問題があった。これを解決するため、シミュレーションによる統計的指標の確率分布の作成方法（既存手法）に代わって、近似的な方法で統計的指標の確率分布を作成する高速化手法を提案した。提案手法の評価として、提案手法と既存手法で作成した確率分布の作成時間と分布の精度を確認した結果、提案手法は既存手法に対して、確率分布の精度をおとさず、高速化できていることがわかった。

次に、血統とタンパク質発現量の相関関係の推定として、発現量が和牛の血統によって制御されるタンパク質の発見を試みた。血統によって制御されるタンパク質とは、例えば、血統によってタンパク質が制御されている場合、2頭の牛の血統が良く似ていれば、2頭の牛の細胞組織中のタンパク質発現量は似たような値をとり、血統が似ていなければ、タンパク質発現量が必ずしも似ているわけではない。つまり、発現量が血統に依存していれば、血統が近いほど発現量の分散が小さいく、依存していなければ血統が近くても発現量の分散は小さいとは限らない。このような関係を発見するため、血統の近さとその時の発現量の分散を計算することで、発現量が血統に依存するタンパク質を推定する手法を提案した。評価として、ブランド牛の血統、遺伝的要因、およびタンパク質発現量の特徴を残した現実的なデータモデルを設計し、このモデルに基づいて人工データを作成した。人工データに提案手法を適用した結果、提案手法は、血統に発現量が依存するタンパク質を抽出できることが示唆された。

さらに、タンパク質発現量と表現型間の相互作用の推定として、2つのタンパク質の発現量を二軸とする二次元平面上において、表現型値が似た値をもつサンプルが局所的に集中する領域を発見することにより、タンパク質発現量と表現型の因果関係を推定する手法を提案した。提案手法は、領域の形がいびつでなく、かつ領域内のサンプルの表現型値の分散が小さい領域を抽出することで、タンパク質と表現型間の関係を推定することが可能である。評価として、提案手法の性能を評価するため、抽出する領域の形が常に円形となる領域抽出手法（単純拡張手法）を実装し、提案手法と単純拡張手法を実データに適用した。評価実験の結果、提案手法は単純拡張手法より良い領域が抽出できることが確認できた。

本研究において提案した手法は全て、生命現象を支えるタンパク質に着目し、血統と枝肉成績の関係をタンパク質の作用による現象として明らかにする試みであり、枝肉成績を改善するための効果的な方法論の開発に貢献できる可能性がある。

## 論文審査の結果の要旨

本論文では、ブランド牛の肉質改善を目標として、血統、肉質、及びタンパク質の発現量データを対象としたデータ解析アルゴリズムを提案した。第一に、2つのタンパク質と肉質の間の複合的な相互作用を統計的方法論に基づいてモデル化し、2タンパク質の発現の相乗効果を検出する手法を提案した。第二に、ブランド牛の血統を血統ベクトルとして新規にモデル化し、血統によって発現量が制御されるタンパク質を特定する手法を提案した。第三に、肉質と2つのタンパク質の間の複合的な相互作用を、発現量に近いサンプル集合に基づいて検出する手法を提案した。本論文は、肉質との関係が通常用いられる遺伝子よりも密接な、タンパク質を扱うデータ解析手法の開発に取り組み、タンパク質と肉質の間の新規な関係性を発見する手法を開発し、その有効性を明確に示した。一定程度の学術的価値が認められ、博士（工学）論文として評価できる。

## 最終試験の結果の要旨

公聴会は平成29年2月10日にA203講義室で約1時間30分にわたって行われ、研究発表及び質疑応答が各45分間なされた。最終試験では、主に以下の事項について質疑応答がなされ、一部不明確な部分もあったがおおむね明確な回答を得ることができた。主な内容は、まず、各提案手法の、馬、豚、鳥、羊等の他の家畜における研究との関連について質疑があり、他研究の具体的な言及はなかったものの、ブランド牛の交配方法が特殊なため、他の家畜との比較は難しい旨の回答があった。次に、各提案手法や提案モデルが現実に即した妥当なものである根拠についての質疑があり、各手法が想定する相互作用の実例が示されることで、各提案手法が扱う関係性が現実的であることが説明された。以上のことから、3名の審査委員が合議した結果、藤木生聖は、大学院博士後期課程修了者として、博士（工学）の学位を授与するに十分であると判断した。