

機関番号：14301

研究種目：基盤研究(C)

研究期間：2008～2010

課題番号：20580303

研究課題名(和文) 染色体セグメントモデルによる遺伝的評価システムの開発

研究課題名(英文) Developing the computing procedures and systems for genetic evaluation of resource animals using the chromosome segment model

研究代表者

祝前 博明 (IWASAKI HIROAKI)

京都大学・大学院農学研究科・教授

研究者番号：00109042

研究成果の概要(和文)：本研究では、染色体セグメントモデルを用いた遺伝的評価法(マーカーアシスト BLUP 法)ならびに大量の SNP 型情報の利用を想定した遺伝的評価法における演算手法の検討を行った。複数の演算手法について検討を加え、シミュレートしたデータならびにマウスおよび牛の実データを用いて、それらの手法の有効性を調べた。染色体セグメントモデルを用いたマーカーアシスト BLUP 法の演算アルゴリズムを確立し、理論とシミュレーションによる検討の結果、本手法の有効性が示された。また、SNP 型情報を利用する複数の遺伝的評価法の演算手法について検討を加えた結果、多数の SNP 効果を説明変数とするベイズ法などの手法は、育種価推定の正確度の点で、ポリジーンモデルによる混合モデル法よりも有望であることが示唆された。

研究成果の概要(英文)：In this study, a computing procedure and system for genetic evaluation of resource animals using the chromosome segment model was considered. Those for the genetic evaluation using a large amount of SNP data were also investigated considering several approaches. Using simulated and actual data of hetero-stock mice and cattle, validities of those computing procedures were examined. Based on the theory and simulation study, the genetic evaluation of resource animals using the chromosome segment model was shown to be useful. Bayesian methods fitting SNP effects appeared to be relatively accurate and promising, in terms of estimation of total genetic merit, compared with the mixed model method using the polygene model.

交付決定額

(金額単位：円)

|        | 直接経費      | 間接経費      | 合計        |
|--------|-----------|-----------|-----------|
| 2008年度 | 1,800,000 | 540,000   | 2,340,000 |
| 2009年度 | 1,100,000 | 330,000   | 1,430,000 |
| 2010年度 | 700,000   | 210,000   | 910,000   |
| 年度     |           |           |           |
| 年度     |           |           |           |
| 総計     | 3,600,000 | 1,080,000 | 4,680,000 |

研究分野：農学

科研費の分科・細目：分科：畜産学・獣医学 細目：応用動物科学

キーワード：ゲノム情報科学

1. 研究開始当初の背景

(1) 多数の量的形質遺伝子座(QTL)によって支配されている量的経済形質の遺伝的

改良を推進していく上では、選抜対象個体の遺伝能力を的確に予測することが第一義的に重要である。現行の遺伝的評価法では、

形質データと血統データとを利用し、いわゆる「ポリジーンモデル」を用いた最良線形不偏予測 (BLUP) 法による能力予測が行われている。

(2) 研究開始当初の段階では、ゲノムの全域にわたるマイクロサテライトのような DNA マーカーの連鎖地図が作成され、DNA マーカー情報をも利用したゲノム育種およびマーカーアシスト選抜 (MAS) のための種々の遺伝的評価法すなわちマーカーアシスト BLUP 法の開発が進められていた。しかし、その際、典型的な遺伝モデルである「単一 QTL モデル」や「複数 QTL モデル」では、QTL の染色体上の位置が既知である必要があり、通常の QTL マッピング法では、QTL 位置の信頼区間が広く、大きな推定誤差を伴うことから、MAS に実用する上での制約条件となっていた。

(3) 最近では事情が大きく変わってきている。すなわち、牛、豚、鶏などをはじめとして、数万～数十万の一塩基多型 (SNP) の遺伝子型情報の利用が可能となり、ゲノム育種における最先端の課題は、大量の SNP 型情報を利用した遺伝的評価法の開発・改良へと急速に変化してきている。

## 2. 研究の目的

(1) 申請者らが、近年、構築した「染色体セグメントモデル」と呼ばれる新たな遺伝モデルの基礎理論を用いて、個体の遺伝子型値の汎用型予測システムを確立する。ここでは、従来の「ポリジーンモデル」と申請者らの「染色体セグメントモデル」とを理論的に統合し、育種現場で実際に利用することのできる、「汎用染色体セグメントモデルによるマーカーアシスト BLUP 法」の確立を目指す。また、その具体的演算プログラムを完成させる。

(2) ゲノムの全域にわたる大量の SNP データの利用を想定し、それらを用いて個体の遺伝能力を予測するための複数の解析手法とそれらの計算アルゴリズムについて検討を加える。また、モンテカルロシミュレーションによるデータならびにマウスおよび和牛の実データを用いて、それらの手法の有効性を調べる。なお、大量の SNP 型情報の利用状況下での関連データでは、SNP 遺伝子型やハプロタイプの総数すなわち説明変数の総数が膨大な数にのぼる一方、非説明変数である分析対象形質の測定値を備えた個体の総数ははるかに少数となる。したがって、この種の数値解析問題に取り組む上では、いわゆる  $p > n$  問題に対応する優れた解析手法についての演算シ

テムの開発が不可欠となる。

## 3. 研究の方法

1) すべての個体が DNA マーカー情報を完備している状況を想定し、染色体セグメントモデルによる遺伝的評価システム、すなわち全個体の量的形質に関する相加的遺伝子型値の最良予測システムを構築した。染色体セグメントモデルに関して、IBD 行列 (同祖性確率行列) の形成法、IBD 行列の逆行列の計算法について理論的に再整理し、数値積分手法を応用した実用的方法を検討した。また、染色体セグメントモデルによる混合線形モデルを構成し、遺伝分散・環境分散が既知である場合の混合モデル方程式と染色体セグメント (MQTL) 効果および残りのポリジーン (RQTL) 効果の予測式を導出した。以上の結果を用いて、Fortran90/95 言語による演算プログラムを作成し、遺伝的評価のための演算システムを構築するとともに、シミュレーションにより、方法の有効性を調べた。

2) 大量の SNP 型情報を想定した場合における遺伝的評価法として、ベイズ推定法に属する複数の手法およびカーネル多変量解析に関連した手法を取り上げ、具体的な演算手法を構築した。ベイズ推定法に属するアプローチについては、(1) 各 SNP 座における遺伝分散がそれぞれ異なると仮定した遺伝モデルを想定したアプローチ (いわゆるベイズ A 法)、(2) 多数の DNA マーカー座のうちの大半における遺伝分散を 0 と仮定したアプローチ (いわゆるベイズ B 法)、(3) 各 DNA マーカー座における遺伝分散がすべて等しいと仮定した遺伝モデルによるアプローチ、(4) ベイジアン Lasso アプローチ、(5) SNP 群を相対的に効果の小さい群と効果の大きい群とに大別して取り扱ったアプローチなどを取り上げた。カーネル多変量解析に関するアプローチでは、ガウス・カーネルなどのいくつかのカーネルを用いた。

これら種々の方法について、用いるべき事前分布などを決めるとともに、Fortran90/95 言語による演算システムを構築した。ついで、モンテカルロシミュレーションにより、これらの方法の有効性について検討した。さらに、実データとして、英国 Wellcome Trust センターのヘテロストックマウスの関連データベースから、約 1,900 個体の測定値・血統情報と 11,037 の SNP の情報の分析用データセットを作成し、解析に供した。また、黒毛和種去勢肥育牛 673 頭の枝肉重量の記録と 39,422 個の SNP 座の遺伝子型データを含む分析用データセットを作成して、解析に供した。

#### 4. 研究成果

(1) 表 1 は、20cM のフランキンクマーカーク区間に 2cM 間隔で九つの QTL (MQTL) を含む遺伝モデルを設定し、ランダム交配による 5 世代をシミュレートした場合の各 BLUP 法による育種価予測値の相対的正確度の結果である。

表1. マーカー区間に9つのQTLsが存在する状況(9番目の18cMに位置するQTLが遺伝分散の20%を説明)における各BLUP法の相対的正確度(マーカーアレル数は4)<sup>1)</sup>

| MABLUPモデル              | トータル育種価 | MQTL効果 | RQTLs効果 |
|------------------------|---------|--------|---------|
| QM-6 <sup>2)</sup>     | 103.38  | 123.92 | 94.46   |
| QM-10 <sup>2)</sup>    | 103.82  | 125.93 | 94.27   |
| QM-14 <sup>2)</sup>    | 103.95  | 127.47 | 93.83   |
| QM-18 <sup>2)</sup>    | 103.75  | 128.63 | 93.16   |
| QM-22 <sup>2)</sup>    | 103.02  | 126.57 | 93.10   |
| QM-26 <sup>2)</sup>    | 102.81  | 123.86 | 93.96   |
| HM <sup>3)</sup>       | 101.29  | 111.73 | 96.86   |
| SM <sup>4)</sup>       | 103.89  | 126.73 | 94.06   |
| IBD-BLUP <sup>5)</sup> | 105.40  | 138.97 | 90.81   |

<sup>1)</sup> 値は従来のマーカー情報を用いないBLUPに対する相対値で%表示している

<sup>2)</sup> QM-X: 仮想QTL位置をXcMとした単一QTLモデル

<sup>3)</sup> マーカーハプロタイプモデル

<sup>4)</sup> 染色体セグメントモデル

<sup>5)</sup> マークされたQTLsについての真の平均IBD情報を利用したモデル

本シミュレーションでは、MQTL は相加的遺伝分散の 30%を説明し、18cM の位置に相加的遺伝分散の 20%を説明するメジャーQTL が位置するとした。したがって、マーカーによってマークされた MQTL 以外の QTL (RQTL) によって説明される相加的遺伝分散の割合は 70%であった。マーカーのアレル数は 4 とした。

まず、マークされた QTLs についての真の平均 IBD 情報を利用したモデルによる BLUP 法 (IBD-BLUP) における MQTL 効果およびトータル育種価の予測値の正確度は、あらかじめ期待されたように最も高かった。しかし、このようなアプローチは、シミュレーションにおいてのみ可能であり、現実の遺伝的評価では、真の平均 IBD 情報を用いることは不可能である。一方、仮想 QTL 位置を X cM とした単一 QTL モデルによる複数の BLUP 法 (QM-X) では、X が真の値から隔たるにつれて正確度が低下し、X がマーカー区間外に想定されたときには明らかに正確度の低下が認められた。したがって、正確度の低下の程度は真の QTL 位置と仮想位置との差に依存していた。マーカー区間内に単一 QTL 位置を仮定しない「染色体セグメントモデルによる BLUP 法」のトータル育種価の予測値の正確度は、IBD-BLUP 法の場合に比べると劣るものの、QM-X 法での最も高い値にほぼ匹敵していた。また、染色体セグメントモデルによる場合には、MQTL 効果および RQTL 効果の予測値の正確度はバランスがとられていた。

シミュレーションでの設定条件を種々に変化させた検討も行ったが、真の QTL 位置を仮定しない染色体セグメントモデルによる BLUP 法の正確度は、真の QTL 位置

がマーカー区間外の場合でも相対的に優れていた。

(2) SNP 情報を利用した各種のベイズ法の演算手法とプログラムは、良好に機能し、演算の体系が確立された。表 2 は、ヘテロストックマウスのデータを用いて交差確認を行い、6 週齢体重、体長、アルカリホスファターゼ活性および高比重リポタンパク質の 4 形質を取り上げて、各種のベイズ法などによる育種価予測値と表現型値との相関をみたものである。

表2. 交差確認における育種価予測値と表現型値との相関

| 形質 <sup>1)</sup> | モデル <sup>2)</sup> | 方法 <sup>3)</sup> |       |       |       |       |       |
|------------------|-------------------|------------------|-------|-------|-------|-------|-------|
|                  |                   | ベイズA             | ベイズB  | ベイズC  | ベイズD  | G法    | A法    |
| Weight           | M                 | 0.417            | 0.408 | 0.422 | 0.440 |       |       |
|                  | M&Pa              | 0.597            | 0.592 | 0.591 | 0.589 | 0.418 | 0.559 |
| BL               | M                 | 0.228            | 0.234 | 0.232 | 0.249 |       |       |
|                  | M&Pa              | 0.241            | 0.242 | 0.236 | 0.265 | 0.229 | 0.180 |
| ALP              | M                 | 0.655            | 0.647 | 0.637 | 0.648 |       |       |
|                  | M&Pa              | 0.660            | 0.656 | 0.641 | 0.645 | 0.640 | 0.489 |
| HDL              | M                 | 0.583            | 0.572 | 0.533 | 0.555 |       |       |
|                  | M&Pa              | 0.584            | 0.575 | 0.538 | 0.555 | 0.529 | 0.453 |

<sup>1)</sup>Weight, 6週齢体重; BL, 体長; ALP, アルカリホスファターゼ活性; HDL, 高比重リポタンパク質

<sup>2)</sup> Pa : ポリジーン効果(相加的血縁)  
M : SNPマーカー効果の総和  
M&Pa: ポリジーン効果(相加的血縁) + SNPマーカー効果の総和

<sup>3)</sup>G法, ゲノム関係行列を利用したBLUP法による育種価予測値  
A法, 相加的血縁行列を利用したBLUP法による育種価予測値

形質によって、相関のレベルには差異が認められたものの、SNP 情報を利用した各種のベイズ法間に相関の大きな差異は認められなかった。メジャージーンによる制御が明らかとなっているアルカリホスファターゼ活性および高比重リポタンパク質では、ベイズ法における相関は、通常の相加的血縁行列を用いた BLUP 法における相関に比べて明らかに高く、ベイズ法による育種価推定法の有効性が確認された。

(3) 表 3 は、50K SNP データを利用した黒毛和種肥育牛の枝肉重量の遺伝的パラメータ推定の結果である。

黒毛和種去勢肥育牛の枝肉重量および SNP 情報を用いて、線形回帰分析による LOD スコアの高い SNP 81 個(P<0.0001) から 4,216 個(P<0.05)を選択した。選択 SNP およびポリジーン効果を取り上げたモデルにより、ベイズ法を用いて分散成分の推定を行った。ここでは、血統情報に基づく相加的血縁行列の代わりに、すべての SNP データを利用して作成したゲノム関係行列を用いた。LOD スコアの高い SNP 数を増加させるに従い、残差分散はわずかに減少し、また SNP により説明される遺伝分散は増加した。一方、LOD スコアの低

い SNP ( $P > 0.1$ ; 32,590 個)により説明される遺伝分散は、ほぼ 0 であった。取り上げた SNP 数が少ない場合は、ポリジーン効果のみを取り上げたモデルによる遺伝率推定値よりも若干低かったが、SNP 数を増加させるにともない、同等の値となった。したがって、SNP 情報に基づくゲノム関係行列を用いる方法も有望な方法の一つであると考えられた。

表3. 50K SNPデータを用いた肥育牛枝肉重量に関する遺伝的パラメータの推定値

| 分散成分            | GRM    | 有意確率             |                   |
|-----------------|--------|------------------|-------------------|
|                 |        | P<0.05           | P>0.1             |
| 残差分散            | ALL    | 922.8<br>(113.4) | 1048.2<br>(140.8) |
|                 | SELECT | 917.7<br>(113.1) | 963.0<br>(106.5)  |
| 遺伝分散<br>(SNP)   | ALL    | 726.5<br>(139.2) | 1.0<br>(0.8)      |
|                 | SELECT | 749.9<br>(140.9) | 1.1<br>(0.8)      |
| 遺伝分散<br>(ポリジーン) | ALL    | 194.4<br>(101.6) | 916.1<br>(180.9)  |
|                 | SELECT | 179.2<br>(89.8)  | 848.5<br>(153.1)  |
| 遺伝分散<br>(トータル)  | ALL    | 920.9            | 917.1             |
|                 | SELECT | 929.1            | 849.6             |
| 遺伝率             | ALL    | 0.499            | 0.467             |
|                 | SELECT | 0.503            | 0.469             |

括弧内の値は標準誤差

## 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)

- ① H. Iwaisaki, A. Arakawa, H. Matsuda, Genomic selection: some relevant topics, Japanese Journal of Animal Breeding Research, 査読無, Vol.18, 2010, pp.17-49.

[学会発表] (計 4 件)

- ① T. Watanabe, A. Arakawa, S. Nishimura, T. Yamada, Y. Sugimoto, H. Iwaisaki, Estimation of variance components for carcass traits in Japanese Black cattle using 50K SNP genotype data, 32<sup>nd</sup> Conference for the International Society for Animal Genetics, Edinburgh, UK, 2010. 7. 28, p2025.
- ② 板東大三郎、荒川愛作、谷口幸雄、山田宜永、祝前博明、ゲノミック選抜における選択的タイピングの有効性の検討、日本畜産学会第 112 回大会、明治大学駿河台キャンパス、2010 年 3 月 29 日、p. 55.

- ③ 渡邊敏夫、荒川愛作、西村翔太、山田宜永、杉本喜憲、祝前博明、50K SNP データを利用した黒毛和種枝肉重量の分散の推定、日本畜産学会第 112 回大会、明治大学駿河台キャンパス、2010 年 3 月 29 日、p. 57.
- ④ 荒川愛作、板東大三郎、谷口幸雄、山田宜永、祝前博明、マウスのデータを用いたゲノミック評価法-MIXTURE-の評価、日本畜産学会第 112 回大会、明治大学駿河台キャンパス、2010 年 3 月 29 日、p. 64.

## 6. 研究組織

### (1) 研究代表者

祝前 博明 (IWAISAKI HIROAKI)

京都大学・農学研究科・教授

研究者番号：00109042

### (2) 研究分担者

( )

研究者番号：

### (3) 連携研究者

( )

研究者番号：