

プロジェクト名： 統計・情報技術を駆使したゲノム多型と表現型多様性の 連関解析システムの開発〔略称：生物多様性解析〕

プロジェクトディレクター： 城石俊彦

〔1〕 研究計画・研究内容について

(1) 目的・目標

生物の自然集団から独自に採集・育成してきたマウス系統やイネ系統が示す豊富な生物多様性を対象に、統計数理技術と情報処理技術を駆使してゲノム多型と数値計測化された表現型を関連づけて遺伝子機能と遺伝子パスウェーを体系的に抽出する解析システムを開発する。上記解析システムを活用して、大学等の外部機関との共同研究により動植物のゲノム機能と生物多様性を生み出した進化メカニズムの解明を行う。

(2) 必要性・重要性（緊急性）

近年、ゲノム科学の爆発的な進展により生物の DNA 配列情報について大量の実験・観測データが産生されている。また、生物の表現形質の測定技術の進歩もめざましい。しかし、大量の実験・観測データが産出されたものの、生命現象は個別の事象が相互に関係しあい複雑なネットワークシステムを作っていることから、その総体としての理解はデータ量が多くなった分だけ極めて困難な状態である。この状況にあって大量データを「情報とシステム」というキーワードに基づいて、ゲノム多型と表現型多様性を統計・情報学的に関連づけて体系的に生命現象を理解しようという本プロジェクトの意義は大きく、喫緊の課題である。

(3) 期待される成果等（学問的効果、社会的効果、改善効果等）

生物の自然集団が保有する表現型多様性とゲノム多型を関連づけることは、生物多様性研究の中で個性の解明の基盤となり、極めてユニークなものである。本プロジェクトの成果を大学等の研究に波及することにより、個々の大学では不可能であった体系的なゲノム機能の解明が促進できる。さらに、モデル生物のゲノム機能が解明されることにより、長期的には生物環境問題、身近には健康医療問題や食料問題など、現在人類が直面している懸案事項の解決にも直結すると期待される。

(4) 独創性・新規性等

野生マウスやイネなどのモデル生物系統のゲノム多型性と表現形質関連データを対象に、高度な統計・情報処理技術と最新のゲノム解析手法を駆使することにより様々な表現形質の遺伝的要因まで明らかにすることは、生命科学を進めている多くの大学が囑望するところであるが、これらの解析システムを単一の大学で開発し実施することは、人材や研究資源の点で困難である。その点で、研究資源、遺伝実験技術、統計・情報解析技術が全て揃った大学共同利用機関である情報・システム研究機構に設置された新領域融合研究センターで本プロジェクトを実施する意義は極めて大きい。

(5) これまでの取り組み内容の概要及び実績

生物の形態や動物行動等の表現形質の数値計測技術を開発し、さらに測定精度の向上と機能拡張を進めてきた。また、計測されたデータを統計遺伝解析に適用するため表現型および遺伝子型データの収集を行った。さらに、遺伝子発現及び塩基多型の同時検出のための統計手法を開発した。ゲノム多型と表現型を関連づける統計技術としては、生殖的隔離に関与する遺伝子座相互作用の多重性調整についての

研究に進展があった。また、QTL解析法の最適化として多変量の連続変量である表現型と多変量の離散変量である遺伝子型を関連づける統計手法について開発を行った。これらの研究活動によって、いずれも当初の目的を果たしている。

(6) 国内外における関連分野の学術研究の動向

モデル生物のゲノム多型と表現型多様性を関連づけることにより、個別の遺伝子の生物機能を解明しようという研究プロジェクトは国内外の多くの大学や研究機関等でも進められているが、本プロジェクトのように、統計・情報技術自体を体系的に開発して、幅広い生命現象に適用できるような解析プラットフォームを構築しようという研究は希である。特に、植物と動物の両方のモデル生物系統を含めて、解析系を構築しようという研究プロジェクトは、国外でも例がない。

[2] 研究計画

(1) 全体計画

| テーマ | 16年度 予備研究 | 17年度 プロジェクト 初年度 | 18年度 | 19年度 中間評価 | 20年度 | 21年度 |
|-----------------------------|-----------------------------------|-----------------------|------|--------------|------|------|
| 全体 | ↔ モデル生物のゲノム・表現型多様性の統計解析システムの開発 | ↔ | | ↔ | | |
| 表現型数値計測システムの開発 | | ← | | | | → |
| 生物多様性データの統計的モデリングと解析システムの開発 | | ← | | | | → |

(2) 年度計画

平成16年度（予備研究）

「モデル生物のゲノム・表現型多様性の統計的データ解析システムの開発」という研究課題で予備研究を行う。統計数理研究所と国立遺伝学研究所の二研究所間で以下の打合わせを行い、融合研究の実施内容について協議する。全体会議（2004年12月27日）：国立遺伝学研究所において、融合研究全体の基本方針を協議する。特に、マウス・イネの表現型、遺伝子間相互作用の数理解析手法について具体的な研究方針の検討を行う。個別会議（2005年1月26日）：統計数理研究所において、イネのマイクロアレイによる遺伝子発現解析と生殖的隔離障壁の研究について協議する。

平成17年度（プロジェクト開始）

サブテーマ(1)

形態多様性解析として、マウス下顎骨の画像データについてP型フーリエ記述子を用いた形態数値化

とそれを用いた主成分分析が系統間の多様性解析に有効であることを検討する。X線 CT 値による分析において、マウス内蔵脂肪と皮下脂肪を自動的に判別して各脂肪量を定量化するためのソフトウェアの開発に着手する。マウス行動パターンの内、社会行動と自発活動の日周期変動について、客観的な数値計測化と統計モデルによるシミュレーションを行う。遺伝的距離の大きな生物系統において、一方の系統のゲノム DNA のプローブセットを用いたマイクロアレイの統計解析についての検討を行い、SNP 由来の見せかけのシグナル強度を判定するための方法論を検討する。

サブテーマ(2)

- 課題 a. P 型フーリエ記述子を用いてマウス下顎骨の形を数値化し、得られた数値データの多変量解析を実施する。
- 課題 b. マウス 2 系統の活動性データから日内の短時間での変化の周期性を見つけ出すため、点過程モデルにより解析を進める。また、社会行動については動きの方向ベクトルに関する尤度推定を行うことで社会性の数値化を進める。また、2 個体間の距離の分布に基づく社会行動の定量的評価法の開発に向けて、ヒストグラムによる予備的解析を行う。
- 課題 c. 個々のマウス活動性データに点過程モデルを適用し、得られた強度関数の主成分分析から、系統特有の日周期パターンを探索する。
- 課題 d. 配偶体または接合体内の異なる遺伝子座の生殖的隔離障壁間の相互作用検出のため、マーカー分離の独立検定を行う。独立性カイ 2 乗検定統計量を格子点上の自由度 4 の確率場と捉え、マーカーの連鎖による相関構造を確定し、調整 p 値の計算をシミュレーションにより行う。
- 課題 e. ケージ内での活動性とオープンフィールド活動性 (OFA) に関わる行動要素をより明らかにするために、F2 集団のケージ内活動性について、従来の総活動性 (THA) に加え、時間的要素である活動時間 (AT) と量的要素である平均活動量 (AA) に分離した測度を用いた解析を行う。それぞれの形質について、広義の遺伝率 (H2) を調べ、その高い遺伝性を検討する。

平成 18 年度

サブテーマ(1)

- 課題 a. マウス皮下・内臓脂肪を自動測定し数値化するシステムを構築するため、(B6 x JF1) F2 世代の複数個体の CT 画像データの取得と、それらを用いた各組織の CT 値の画像毎のばらつき抽出と、皮下・内臓脂肪を自動測定するため腹筋線の特徴点選定を行う。
- 課題 b. マウス肥満に関連する表現型に関する責任遺伝子座を検出するため、(MSM x B6) F2 交配世代を用いた QTL 解析を行う。また、新規に (JF1 x B6) F2 交配世代個体について、エネルギー代謝関連表現型を中心としたパラメータの計測を行う。
- 課題 c. イネについて、野生イネ・コアコレクション 46 系統で穂形質 (穂型、穎花長、葯長、粒重、1 穂穎花数、稔性) を解析する。各形質の相関関係の検定を行い、今後のアソシエーション解析に適した形質と系統の検討を行う。
- 課題 d. イネとマウスの二つの生物種において、SNP に配慮したマイクロアレイによる遺伝子発現の統計解析手法開発のための実験データを生産する。また、アフィメトリクス社製 GeneChip を用いて、シグナル強度の比較から SNP を検出するために開発したプログラムを検証するため、このプログラムにより検出された SNP と、塩基配列比較から検出された SNP (すでに DB 化した) の比較検討を行う。

サブテーマ(2)

- 課題 a. P 型フーリエ記述子を用いてマウス下顎骨の形を数値化し、得られた数値データの多変量解析を実施するためのソフトウェアを整備する。
- 課題 b. マウスの自発活動性や社会性などの行動を時系列に沿って効果的に解析するシステムの確立を進める。活動性に関しては、点過程で活動をモデル化し、活動頻度を強度関数により推定する手法を用いて各系統に共通の日周期を抽出することを目指す。社会行動については、2次元 **unit vector chain** を用いた解析と 2 個体間の距離情報を用いた解析を進めることで、社会行動の評価を目指す。
- 課題 c. 30 分ごとの活動に集約した上で、系統ごとの日周期プラス個別変動に活動量を分解し、日周期の系統差・短期的変動の系統特性を定量的に分析する。
- 課題 d. 独立性カイ 2 乗検定統計量の自由度 1 の 4 個の独立なコンポーネントへ分解と、配偶体または接合体内の異なる遺伝子座の生殖的隔離障壁間の相互作用の遺伝的モデルとの対応関係を明らかにする。また、カイ 2 乗確率場の最大値分布の逐次解析理論による解析式を導出し、実用的な多重性調整 p 値の計算法を検討する。
- 課題 e. これらの F2 個体群の測度を用いた **composite interval mapping** を行い、THA、AT、AA のそれぞれについて関与する QTL 遺伝子座 (**Hylaq1**、**Hylaq2**、**Hylaq3**) を同定する。更に、これら遺伝子座の AT および AA に対する独立の効果を詳細に調べるため、**R/qtl** による共変量を用いた **QTL interval mapping** を行い、それぞれの行動要素間の関連を明らかにする。

平成 19 年度 (中間評価)

サブテーマ(1)

- 課題 a. マウス F2 (**JF1/B6**) 個体の BMI により分類した多様な CT 画像データを脂肪量自動抽出に必要な腹筋線検出のための各種モデルに当てはめ、最適となるモデルの選定を行う。マウスの全身 CT 画像から腹筋線および 2 種類の脂肪組織 (内臓脂肪と皮下脂肪) の自動抽出を行う画像処理プログラム (**ActScan**) を作成する。
- 課題 b. **JF1 - B6** 系統間のエネルギー代謝関連表現型差異を規定する遺伝的要因を探索するため、両系統より作出した F2 世代の当該表現型を中心としたパラメータの計測を充実させる。さらに全ゲノムスキャンによる各種表現型を規定する責任領域の抽出を行う。
- 課題 c. 野生イネ系統の集団を用いて穂形質特性についての集団内系統別形質データを収集し、さらに系統別の遺伝子発現をマイクロアレイデータとして取得する。特に種内で形質が大きく 2 つのグループに分かれた *O. rufipogon* に **target** を絞って、形質データを収集する。
- 課題 d. (i) イネゲノム SNP とハイブリシグナル強度差との相関解析の結果をもとに、SNP 推定プログラムの最適化を行い、論文を作成する。
(ii) 日本産野生由来近交系統のゲノム情報から SNP により影響を受けているプローブの抽出をさらに推し進める。B6 系統、日本産野生由来近交系統の発現データを使用し、プローブのシグナル値から **SNP effect** を検出するアルゴリズムによる遺伝子発現差異の検証を行う。マウス SNP 推定アルゴリズムにより、B6 - MSM 系統間の SNP 情報とシグナル強度の関連を検証し、論文を作成する。

サブテーマ(2)

- 課題 a. H18 年度に作成したデータ、開発してソフトウェアを用いて、下顎形態の数値化を行う。形状の数値化のための新たな方法を研究する。また、マウスの形態の解析のプロジェクトに参

加する。B6 マウスと MSM マウスの下顎形態の特徴を決定している因子を探索する。田村たちが行った P 型フーリエ記述子を特徴量としてパターン認識をアダブーストで行うとシャープな特徴量が発見された。この特徴量を量的形質とする QTL 解析を行い、関連する遺伝子座位をマップする。

- 課題 b. 複数の動物個体が示す社会行動は、生存において重要な意味を持つが、その定量化が難しく解析を困難にしている。ここでは、2 個体のマウスを同時にオープンフィールドにおいた際の 10 分間の社会行動を時系列のトラッキングデータに基づき、統計モデルで特徴づける解析を進める。社会的行動が異なる B6 および MSM の 2 近交系統の特徴を、移動行動のベクトル間の相互作用をもとに尤度推定値を算出することで数値化する。特に H19 年度は移動行動からベクトルを算出する方法に改良を加えて相互作用のパラメータを尤度推定することで、B6 と MSM の 2 近交系統の社会行動の違いを明確に捉えることを可能にする。また 2 個体のマウスのビデオ画像をトラッキングデータと同期させて表示させ、観測者がマウスの状態を目視により効率的に記録することができるソフトウェアを開発する。このソフトウェアを用いてマウスの行動の時系列パターンにマルコフモデルを当てはめて予備的解析を行う。
- 課題 c. マウス活動量データ（赤外センサー交差回数時系列）に、三角関数列を説明変数とするポアソン回帰モデルを適用し、24 時間周期から 2 分周期まで全ての可能な周波数の有意性を検出する方法を提示する。
- 課題 d. 相互作用の真偽判定の妥当性を吟味するため、以前構築したイネの高密度連鎖地図と同じ組合せ親による別集団で連鎖地図作製を行う。
- 課題 e. C57BL/6 系統と KJR 系統は顕著に異なった活動性を示す。この活動量の違いに関わる遺伝的基盤を解明する目的で、F2 集団を用いて時間的要因とスピードの要因を表現型とする QTL 解析を行い複数の遺伝子座を特定する。その結果に基づき、Structural Equation Model (SEM) 解析を行い、各遺伝子座と各活動量指標との関係を構成するシステムを明らかにする。

平成 20 年度

サブテーマ(1)

- 課題 a. マウス脂肪領域判定について、画像のピクセル間に定義したネットワーク形状の解析に基づく ActCuts 法を開発する。ActScan 法と比較して皮下脂肪と内臓脂肪の判定に関して精度向上を確認する。
- 課題 b. JF1 系統と B6 系統の F2 世代個体を用いたエネルギー代謝関連形質の QTL 解析を行い、全ゲノムで有意に高い LOD 値 (2.5 以上) を示す QTL を検出する。
- 課題 c. 野生イネ 46 系統で穂形質情報を収集する。野生イネ *O. rufipogon* 由来系統の形質情報調査・整備する。また、遺伝子多型、発現多型解析用サンプルを収集し、マイクロアレイデータを一部測定する。
- 課題 d. 前年度に開発した発現及び塩基多型の同時検出法 (SNEP: Simultaneous detection of nucleotide and expression polymorphisms) を改良した結果、検出率や擬陽性などのパフォーマンスが改善され、platform strain でない場合にも対応可能となった。この結果を受けて、イネの genome DNA のアレイへのハイブリを用いた塩基多型検出頻度の最適化を行う。mRNA のハイブリで SNP を含むプローブについて、SNP によるシグナル減少量の推定を様々な観測量の重回帰分析量から推定することを試みる。

サブテーマ(2)

- 課題 a. 部分形状の統計解析で部分形状の数値記述に用いられてきた開曲線記述子である接線フーリエ記述子 (TFD) は、曲線上の隣接点間の位置の差分系列のフーリエ変換でありノイズを増幅する傾向がある。そこで新たな記述子として protrusion フーリエ記述子 (PFD) を検討する。次に、2 系統を起源とする F2 世代 (計 207 匹) のマウスを用い下顎骨の部分形状の遺伝解析を行う。異なる部位に対し形に影響する遺伝子が大きく異なること、さらに記述子からの曲線形状の視覚化により、各々の遺伝子の形に対する影響の仕方がそれぞれ異なることなどを検討する。
- 課題 b. (1) 基礎的データ解析: 2 個体のマウスを同時にオープンフィールドに置いた際の 10 分間の社会行動時系列のトラッキングデータを抽出し、それらのデータに基づいて 2 個体の行動を統計数理モデルで特徴づける解析を進める。(2) 方向データ解析: マウス 2 個体の社会行動トラッキングデータから、2 個体同士の社会行動時系列を移動のベクトル系列に変換し、方向相互作用のパラメータを尤度推定することで移動行動の特徴を数値化する。今年度は、移動行動からベクトルを算出するアルゴリズムに改良を加え、新しく得られたマウス 24 系統 461 ペアに対する予備的解析を行う。(3) マルコフモデルによる解析: マウス 2 個体の画像データから目視によりマウスの状態を 5 つに分類した状態時系列を解析し、この時系列にマルコフ行列をあてはめて得られる遷移行列の各要素の重みつき和によって、適切な行動の社会性の尺度が構築できることを確認する。
- 課題 c. マウスの自発行動を記録したデータは非負整数値による計数時系列で極めてゼロが多く、上下の非対称性が強い。興味があるのは活動期の短期的変動の特徴であるので、1 分値で計数 0 は欠測として非ゼロのデータを対数・巾関数等で線形化し、欠測値を許す線形ガウス型カルマンフィルタで、日周期まわりの自己回帰モデルをあてはめることにより、特徴的な活動周波数を推定する方法を研究する。
- 課題 d. 致死遺伝子相互作用の遺伝的モデルについて浸透度等を考慮した一般化、正逆戻し交雑集団を使った真の相互作用の検証を行う。また正逆戻し交雑集団から観測された多数の検定統計量データについて、過去に開発された多重性調整を適用する。
- 課題 e. ロバスト QTL についても幾つかの検討を加える。従来の方法は外れ値の影響を大きく受けることがあるが提案した手法が非常に安定した性能を示すことを確認する。また、QTL 解析における影響診断の方法を検討する。単にロッドスコアのピークの大きさだけでなく、ロッドスコアが形状 (単峰であるか二峰であるかなど) に影響を与える個体を、グラフィカルに図示する方法の開発を行う。

平成 21 年度

サブテーマ(1)

- 課題 a. 腹筋線自動抽出アルゴリズムを実装したソフトウェアの完成と公開を行う。完成版ソフトウェアを使用した各種脂肪組織の測定とデータマイニングを行い、その有効性を検討する。
- 課題 b. MSM - B6 系統ならびに JF1 - B6 系統間のエネルギー代謝関連表現型を規定する原因遺伝子、エピスタシスに関与する遺伝子群の探索を行う。有力な候補遺伝子についてはノックアウトマウスやトランスジェニックマウス作製などの逆遺伝学的実験手法による検証を行う。
- 課題 c. 野生イネ複数系統の発現データをマイクロアレイにより取得、前年度までに課題 d にて開発したシステムにて解析可能かどうか検討する。さらに、これらの結果に基づいて構造化されていない自然集団における e - QTL 及び形質データの QTL 解析の具体的な問題点の洗い出

しを行う。

課題 d. SNP が多く含まれる遺伝子群の発現量の推定法を開発し、他の解析方法 (RT - PCR 法など) により検証する。野生イネの一部系統については、発現遺伝子およびゲノム塩基配列について、次世代シーケンサーによる読み取りと解析を行い、より簡便かつ確実な統計的方法論の導入について考察する。

サブテーマ(2)

最終年度であるため、昨年度の課題 a-e を継続して行い、成果発表を行う。

課題 a. 形状の数値化のための新たなフーリエ記述子について研究する。前年度までに開発した protrusion フーリエ記述子 (PFD) と新たな記述子を用いて、遺伝子の交互作用についての研究を進める。

課題 b. これまでに確立した二つの社会行動評価法を用いて、実際の 24 系統 461 ペアの時系列移動行動データから実際のマウス社会性の評価を行う。こうして数理的解析により得られた社会性評価の結果と、実際のビデオ観察法による行動項目の定量結果とを比較し、これらの解析手法の総合的な検討を行う。

課題 c. H20 年度に開発したプロトタイプを B6 と MSM の全データに適用し、系統共通の基調的活動周期が存在するかどうかを調べる。一本の時不変係数自己回帰モデルではあてはまりが不十分である可能性もあるので、個体の異質性をうまく捉えるモデリング方法について研究する。個体差をうまくコントロールするモデルを発見し、系統ごとの特徴抽出を行うことが目標である。

課題 d. 生殖隔離障壁を引き起こす遺伝子座相互作用の検索において、真の相互作用を再現性実験や別集団による確認なしに検出する手法を開発する。またマーカー間隔が均一でないことをモデルに取り込んだ大偏差型の近似公式を検討する。

課題 e. 昨年度、開発されたロバスト QTL 法について更に発展させる。特に外れ値の影響にロバストでありながらデータの情報を効率良く取り込むために、データに応じてロバスト QTL 法を選択する方法を考える。これらの成果を持って、このプロジェクトで得られたデータに実際に適用する。また、影響診断については、単一マーカー分析だけでなく、区間マッピング法によって得られたロードスコアについて、その大きさや形状に影響を与える個体を特定する方法を開発する。また近年高次元の表現型データに関する QTL 解析の需要が高まっている。eQTL 解析は、その最たるものである。多変量の連続変量である表現型と、多変量の離散変量である遺伝子型を関連づける QTL 解析のためのグラフィカルモデルの推測、モデル比較の方法を検討する。

[3] 研究推進・実施体制

サブテーマ(1) : 「表現型数値計測システムの開発と実験データの収集」

研究代表者

[国立遺伝学研究所] 城石俊彦

共同研究者

[国立遺伝学研究所] 倉田のり、伊藤幸博、春島嘉章、田村 勝、高田豊行、前野哲輝、岡 (木曾) 彩子、久保貴彦、永口 貢、堀内陽子、小出 剛、梅森十三

[国立情報学研究所] 北本朝展、藤山秋佐夫、佐藤真一

[統計数理研究所] 江口真透、池田思朗、藤澤洋徳、川崎能典、坂口隆之、Nurul Haque Mollah

[東京大学 大学院新領域創成科学研究科 情報生命科学専攻] 中谷明弘、中原達聡、廣田晃士
[長浜バイオ大学 バイオサイエンス学科 生命情報科学コース] 阿部貴志

サブテーマ(2)「生物多様性データの統計的モデリングと解析システムの開発」

研究代表者

[統計数理研究所] 栗木 哲

共同研究者

[統計数理研究所] 江口真透、奥田将己、川崎能典、坂口隆之、田中英希、種村正美、田村義保、土谷隆、藤澤洋徳、Dou Xiaoling、Nurul Haque Mollah

[国立遺伝学研究所] 梅森十三、小出剛、岡彩子、城石俊彦、杉本大樹、高田豊行、田村 勝、春嶋嘉章、前野哲輝、細谷正樹

[東京大学] 原 尚幸

[九州大学] 二宮嘉行

[4] 研究の進捗状況

(1) 第1期の研究進捗及び課題

[研究進捗]

平成16年度（予備研究）

「モデル生物のゲノム・表現型多様性の統計的データ解析システムの開発」という研究課題で予備研究をスタートした。統計数理研究所と国立遺伝学研究所の二研究所間で以下の打合わせを行い、融合研究の実施内容について協議した。全体会議（2004年12月27日）：国立遺伝学研究所において、融合研究全体の基本方針を協議した。特に、マウス・イネの表現型、遺伝子間相互作用の数理解析手法について具体的な研究方針の検討を行った。個別会議（2005年1月26日）：統計数理研究所において、イネのマイクロアレイによる遺伝子発現解析と生殖的隔離障壁の研究について協議した。

平成17年度（プロジェクト開始）

サブテーマ1

課題 a. マウスを使用して CT 画像から内臓脂肪、皮下脂肪を分離するソフトウェアの開発を開始した。

- ・マウス下顎骨の画像データについて P 型フーリエ記述子を用いた形態数値化とそれを用いた主成分分析が系統間の多様性解析に有効であることを示した。X 線 CT 値による分析において、マウス内臓脂肪と皮下脂肪を自動的に判別して各脂肪量を定量化するためのソフトウェアの開発に着手した。
- ・マウスの社会行動と自発活動の日周期変動について数理モデルを設定した。
- ・遺伝的距離の大きな二つの生物系統において、一方の系統のゲノム DNA のプローブセットを用いたマイクロアレイの統計解析について検討し、SNP 由来の見せかけのシグナル強度を判定する方法論を検討した。
- ・イネ生殖隔離障壁を引き起こす遺伝子座間相互作用検出のため、検定統計量の相関構造の特定とそれに基づいた多重性の調整法について検討した。
- ・新しい QTL 解析の開発とその有効性の検証を行うためのマウス F2 交配による実験データの生産を行った。

サブテーマ 2

- 課題 a - 1. P 型フーリエ記述子を用いてマウス下顎骨の部分形状の統計解析を行った。
- 課題 a - 2. マイクロアレイのシグナル値の標準化について統計的な観点から幾つかの既存の方法について比較検討をし、これからのデータ解析の準備をした。表現形に分けたグループの間の遺伝子発現の相違が際立つ遺伝子を取り出すための統計方法について考察した。Significant Analysis of Microarray (SAM) 及び Robust Microarray Analysis (RMA) が有用だと分かった。
- 課題 b. マウスの社会行動を解析するために、オープンフィールド内に複数個体を入れて自由に行動させた際の行動をトラッキングし、その時系列データから動きの社会性を判定する手法の開発を進めた。マウス 2 個体の動きを方向ベクトルに細かく分け、その方向ベクトルに関する尤度推定を行うことでマウスの社会性を数値化する試みを進めた。また、2 個体間の距離の分布に基づく社会行動の定量的評価法の開発に向けて、ヒストグラムによる予備的解析を行った。
- 課題 c. マウスのホームケージ内自発活動データ（赤外センサー交差回数時系列）に関して、これまで解析されてこなかった日内の短時間での変化の周期性を見つけ出すため、微少区間内にイベントが起きる確率をモデル化する点過程モデルにより解析を進めた。
- 課題 d. 統計量の相関構造を確定し、多重性調整 p 値をシミュレーションによって求めた。

平成 18 年度

サブテーマ 1

- 課題 a. 汎用実験用マウスと日本産野生由来マウスの交配から得られた F2 世代を使用して、肥満、痩せなど多様な個体を用いた CT 画像セットの取得を推進した。
- 課題 b. エネルギー代謝関連表現型を効果的かつ効率的に検出するため、サンプル採取条件を均一化した表現型収集システムに則った実験データの収集を開始した。
- 課題 c. 16species の野生イネ 46 系統の穂などの複数の形質について形質相関を解析した。また、イネおよびマウスの Affymetrics 社製の GeneChip 上のプローブセットについて、遺伝的距離の大きなイネおよびマウス各々の系統間での SNP 情報を整備した。マイクロアレイデータから SNP を検出するソフトウェアを開発し、既知 SNP と実データとの比較によりソフトウェアの有効性の検証を行った。
- 課題 d. イネ生殖隔離障壁の相互作用についてのカイ二乗検定統計量を配偶体及び接合体内の相互作用の各コンポーネントへ分解した遺伝モデルを作成し、観測値がこの遺伝的モデルと一致することを確認した。

サブテーマ 2

- 課題 a - 1. P 型フーリエ記述子を用いて 2 次元部分形状を多変量解析するためのソフトウェアの整備を行った。
- 課題 a - 2. 田研のイネのヤボニカ種とインデカ種のハイブリ化のマイクロアレイのデータについて SAM、及び RMA について適用した。（これは藤澤たちが行っている SFP 補正法が完成されれば補正後のマイクロアレイ解析に役立つ。）
- 課題 b. 2 個体のマウスをオープンフィールドにおいた 10 分間の社会行動パターンのトラッキングデータに基づき、2 個体の行動を方向データ解析および距離解析によって特徴づける解析を進めた。B6 および MSM の 2 近交系統では顕著に社会行動が異なるが、その特徴を移動行

動のベクトル間の相互作用をもとに尤度推定値を算出することで数値化した。平成 18 年度はデータからのベクトル設定についての検討を進めた。また、2 個体の距離分布を精密に求めることで社会性を特徴付ける解析も進めた。これらの数理解析と同時に、観察により人間が主観的に定量化したマウスの社会性の度合いを、トラッキングデータから抽出したいいくつかの特徴量を用いて説明することを試みた。

- 課題 c. マウス活動量データについて、MSM と B6（夜間のみ・30 分集約）のデータに対し線形ガウス型加法分解モデルをあてはめ、「両種で日周期が異なる」という仮説がモデル選択により支持されることを確認した。更に、推定した種固有日周期周りの個別変動を自己回帰モデルで推定したところ、B6 には顕著な活動周波数が見られたが、MSM には見られず、今後 resolution を上げて解析すべきであるとの結論に至った。
- 課題 d. 当該年度とその翌年は非線形再生理論の方法で、多重性調整 p 値の近似公式を与えた。とくに H18 は自由度 1 のカイ 2 乗分布に関する公式を求めた。

平成 19 年度

サブテーマ 1

- 課題 a. CT 画像から内臓脂肪、皮下脂肪を分離するソフトウェア (ActScan) を開発し、性能評価を行った。解析に使用した全個体について、90%以上の精度で脂肪組織を抽出することができた。
- 課題 b. 肥満関連形質を中心として体質に関わる表現型データを収集した。F2 個体で全ゲノム領域を対象としたタイピングを行い、数種の肥満関連 QTL をゲノム上に位置付けた。
- 課題 c. 栽培イネ祖先種である *Oryza rufipogon* 36 系統に絞って穂形質と分けつ能について調査した。
- 課題 d. マイクロアレイデータに基づいた SNP 検出プログラムを改良し、塩基配列未知の生物系統において SNP を考慮した遺伝子発現解析系を開発した。

サブテーマ 2

- 課題 a - 1. 新しい開曲線記述子、Protrusion フーリエ記述子を考案した。
- 課題 a - 2. マウスの形態の解析のプロジェクトに参加する。B6 マウスと MSM マウスの下顎形態の特徴を決定している因子を探索した。田村たちが行った P 型フーリエ子の特徴量としてパターン認識をアダプストで行うとシャープな特徴量が発見された。この特徴量を量的形質とする QTL 解析を行った結果 13 番染色体に QTL 遺伝子形が示唆されている。
- 課題 b. 2 次元 unit vector chain を用いた解析では、マウス 2 個体の動きの時系列に対応して 2 個体間の行動に角度を割り当てる方法にさらに改良を加えた。また、位置座標の時系列データから、マウス 2 個体の動きを各時刻ごとの系列パターンとする画像を作成した。また、マウス社会行動の画像データを観察者が目視観察し、2 個体の状態を（無関心、追跡、匂いかぎ、攻撃、攻撃後の追跡）の 5 つの状態のいずれかとして判断し、記録するシステムを開発した。このシステムを用いて、54 個の画像データについて、目視により状態を識別した時系列データを得た。この状態の時系列に対して、マルコフモデルをあてはめ遷移行列を推定した。
- 課題 c. マウス活動量データに三角関数列を説明変数とするポアソン回帰モデルを適用し、24 時間周期から 2 分周期まで全ての可能な周波数の有意性を検出する方法を提示した。この悉皆的な周期性分析の結果、マウス活動量データの日周期を表すのに高周波成分が有意となることは殆どなく、2 時間から 6 時間程度の低周波成分が顕著であることがわかった。

- 課題 d. 前年度の結果を多自由度の場合について拡張した。
(19年度より生殖隔離についてはサブテーマ 1 に統合) 生殖隔離における遺伝子間相互作用の検出に関する多重性調整の実用性をシミュレーションにより検討し、4つの方法を試みた。

平成 20 年度

サブテーマ 1

- 課題 a. 平成 19 年度までに開発したソフトウェアと比較して CT 画像から内臓脂肪、皮下脂肪をさらに精度よく抽出できるアルゴリズムを開発した。
- 課題 b. C57BL/6J およびモロシヌス亜種による QTL 解析でエピスタシスが検出可能か検討した。また SNP 検出によるタイピングシステムの整備のためのプライマーの設計を進めた。
- 課題 c. 20 年度は特に穂形質（穂型、穎花長、葯長、粒重、1 穂穎花数、稔性）データを蓄積、整備し、またこの中の代表的系統で、穂、シュートなどの遺伝子発現量の比較実験を行った。
- 課題 d. マイクロアレイデータ解析用に開発した発現及び塩基多型の同時検出法（SNEP: Simultaneous detection of nucleotide and expression polymorphisms）を改良し、検出率や擬陽性などのパフォーマンスを実験的な手法も含め改善をおこなった。

サブテーマ 2

- 課題 a - 1. Protrusion フーリエ記述子を用いてマウス下顎骨の部分形状遺伝解析を行った。
- 課題 a - 2. マウスなどの形態の特徴抽出において従来は主成分分析などの教師なし学習で行っているが昨年度、試験的に行ったアダプテストによる教師あり学習のほうがシャープな特徴が抽出できることが分かった。この統計方法を城石たちの形態データについてより広範に適用したい。
- 課題 b. これまでに、マウス複数個体のフィールド内での動きから社会性を評価する方法として二つの解析方法の確立に目処がたった。一つは、2 個体の時系列移動行動のベクトル間相互作用をもとに社会性の評価を行う方法である。もう一つは、マウスの時系列移動行動データから各時点での 2 個体の状態を無関心、追跡等の 3~5 種類に分類し、得られた「状態の時系列」にマルコフモデルをあてはめて、そのマルコフ行列を用いて社会性の尺度を構成するという方法である。
- 課題 c. マウスの自発行動を記録したデータは非負整数値による計数時系列で極めてゼロが多く、上下の非対称性が強い。興味があるのは活動期の短期的変動の特徴であるので、1 分値で計数 0 は欠測として非ゼロのデータを対数・巾関数等で線形化し、欠測値を許す線形ガウス型カルマンフィルタで、日周期まわりの自己回帰モデルをあてはめることにより、特徴的な活動周波数を推定する方法を研究した。
- 課題 d. チューブ法による近似公式を与えた。
- 課題 e. (平成 20 年度より開始)
単一マーカー分析に基づき、ロッドスコアの形状、特に凸性に影響を与える個体特定の研究を行った。

平成 21 年度

サブテーマ 1

- 課題 a. これまでに開発したアルゴリズムに内臓脂肪、皮下脂肪よりさらに細部にわたって脂肪組織の抽出を行うアルゴリズムの開発を進めた。

- 課題 b. 研究期間に行った C57BL/6J、MSM/Ms および JF1/Ms による QTL 解析の総括を行うとともに、SNP 検出によるタイピングシステムの整備を進めた。
- 課題 c. 野生イネ穂形質等と遺伝子多型の関連を見るため、形質差のある系統どうしの交配を行い、次世代を準備した。また、さらに遠縁のゲノム種間でゲノム DNA, mRNA を用いたマイクロアレイデータの比較解析を行ったが、高雑差が大きすぎてマイクロアレイによる解析は困難と思われた。
- 課題 d. マイクロアレイデータ解析を用いて、発現及び塩基多型の同時検出率を高める手法と考え方に関する論文を作成し、公表した。

サブテーマ 2

- 課題 a - 1. 速度フーリエ記述子を考案し、それを用いてマウス下顎骨の部分形状の遺伝解析を行った。
- 課題 a - 2. イネ、マウスの多様な亜種を決定する発現遺伝子の発見に標準化、クラスターリング、パターン認識の組み合わせで総合的なアルゴリズムを構築した。特にクラスターリングはロバストな階層クラスターリングを開発した。
- 課題 b. 2 個体の時系列移動行動のベクトル間相互作用をもとに社会性の評価を行う方法については、アルゴリズムの最終的な確立とそれに基づく解析結果を提出し、新しい解析方法としての有用性が示された。マルコフモデルを用いた社会性評価については、隠れマルコフモデルにより、2 匹のマウスが社会行動を行っているか否かを自動判別できるようになった。530 個のデータにおいて、自動的に社会的行動を起こしている時間の比率を推定することが可能となり、この観点からの各コンソミック系統の社会性の強さを明らかにすることができた。
- 課題 c. 計数 0 の対数値を $-1/e$ で補間した対数強度系列から周期成分を取り除き、分散が時変モデルに従う AR 過程をあてはめ、短期的な活動周期性を抽出する方法を研究した。少なくとも B6 系統に関しては、確率的ボラティリティモデルの採用により、マウス活動性の短期的周期性は格段に推定が安定することがわかった。
- 課題 d. 戻し交配データについて同じ解析を行った。また論文執筆のための追加のシミュレーション実験を行った。
- 課題 e. 区間マッピング法のピークの高さ位置に対する影響関数を求め、区間マッピング法のロッドスコアの形状に影響を与える個体を特定する方法の研究を行った。

(2) 第 2 期中期目標及び中期計画との関連性

1. 研究機構の教育研究等の室の向上に関する目標、2. 共同利用・共同研究に関する目標の内、(新領域融合研究センター)「多種・大量の地球科学や生命科学などのデータ及び知見を、モデリング技術や情報技術と結合させ、新しい研究分野を目指して融合研究を一層推進する。」、および対応する中期計画に関連する。

[5] サブテーマの研究報告

サブテーマ(1)：「表現型数値計測システムの開発と実験データの収集」

研究代表者：[国立遺伝学研究所] 城石俊彦

1. 研究目標

サブテーマ(1)研究目標

生命システムの理解には、生物多様性を客観的に記載する数値計測技術と得られた表現型としての計測数値をゲノム多型に関連づけるための統計解析手法の融合が必要である。そこで、国立遺伝学研究所

が保有する遺伝的多様性に富んだマウスやイネ等の多数のモデル生物系統の多様性を客観的に評価するために、表現形質の数値計測システムの開発を行う。さらに計測された数値データを統計解析に適用するため、生物系統や遺伝的交配個体からの表現型データ収集においても拡充する。イネ・マウスの形態多様性、脂肪分布などの生体内部の形質についての多様性データは特に重視する。また、マイクロアレイデータからの SNP 推定プログラムや推定された SNP に配慮した遺伝子発現解析のための統計解析システムについてもその完成を目指す。

この研究目標の達成のために、本年度は次の 4 つの課題を設定した。

課題 a. 3D 画像による体脂肪計測法の開発とデータマイニング

課題 b. マウスエネルギー代謝関連表現型計測法の開発と実験データの収集

課題 c. イネストレス耐性・穂形質の系統データ収集と遺伝子発現データとの相関解析法の開発

課題 d. マイクロアレイによるイネ・マウス遺伝子発現データの収集と解析

2. 年度研究計画 [2] の研究計画に記載した。

3. 研究経費の推移

平成 17 年度実績： 157,820 千円 (プロジェクト全体予算)

平成 18 年度実績： 135,170 千円

平成 19 年度実績： 94,210 千円

平成 20 年度実績： 111,750 千円

平成 21 年度実績： 128,000 千円

4. 研究実施体制 [3] 研究推進・実施体制に記載した。

5. 研究成果

(1) 第 1 期 (平成 17~21 年度) の研究成果のまとめ

サブテーマ 1. 「表現型数値計測システムの開発と実験データの収集」

課題 a. 3D 画像による体脂肪計測法の開発とデータマイニング (北本、藤山、佐藤、高田、前野、城石、中谷、中原、廣田、阿部)

[研究の目的]

本研究の目的は、生体イメージングから正確な表現型収集を行うための手法、特にマウスの全身 CT (Computed Tomography : コンピュータ断層撮影) 画像から内臓および皮下脂肪、さらには脂肪組織以外の臓器を自動的に分離し、正確で客観的な数値データを算出するための手法の開発について検討する。生体内部の臓器や構造を測定するためのアルゴリズム開発についても推進する。生体内部の非侵襲的計測、いわゆる生体イメージングからの表現型測定には、CT、MRI (magnetic resonance imaging : 核磁気共鳴画像法)、超音波検査などから得られた情報をいかに忠実に再構築するかがカギとなる。CT 画像については、装置の高性能化に伴い、現在では多数のスライス画像を短時間で取得することが可能となっている。そこで、大量に得られた画像データから対象臓器を正確に抽出し、客観的な数値データに基づく測定をおこなう手法の開発は非常に重要であり、この解析手法の構築は本研究が目的としている CT 画像を使用したマウス表現型の収集を高速に行うこと以外にも、医療分野における診断の高速化、定量化などを目的とした診断支援に応用できると考えられ、非常に有効かつ汎用性がある。この目的を達成するためには、腹筋線の自動認識が必須である。腹筋線認識については、形態テンプレートを構成するための特徴点の選定と、画像データからの特徴点の数値情報の抽出を最適に行うためのアルゴリズム

ムの構築を進める必要がある。そこで我々はマウス全身を対象として撮影した CT 画像から脂肪組織（皮下・内臓脂肪）を自動測定し数値化するシステムを構築するため、複数個体の CT 画像を用いて各組織の CT 値の抽出と画像毎の CT 値のばらつきを想定し、ノイズに強いアルゴリズムを実装した脂肪組織の自動分離計測ソフトウェアの開発を進めている。プロジェクト前半では、2 種類の脂肪組織を分離するために、主に脂肪組織および腹筋線を対象として形態テンプレートを作成していたが、最終的には処理対象の臓器の範囲を広げて、抽出する情報量と精度を向上させる。複数の臓器を扱うことによって、脂肪組織のみでは判別が難しい部分での抽出精度の向上や、各臓器周辺の脂肪組織の分布情報の抽出が期待できる。

【平成 17～21 年度の進捗】

マウス個体を用いた CT 撮影画像を使用した脂肪種判定について、プロジェクト期間内に先ず、ActScan 法を開発し良好な判定結果が得られることを確認した。次に、ピクセル間に定義したネットワークの形状の解析に基づく ActCuts 法を開発した。これにより、これまでに開発した ActScan 法と比較して皮下脂肪と内臓脂肪の判定に関して概ね 10%以上の精度向上を確認した。また、臓器ごとの CT 値の分布を調べ脂肪領域を抽出するための CT 値の閾値の検討を行った。

自動解析システム ActCuts 法の開発

CT 画像のピクセル間に定義したネットワーク形状の解析に基づく ActCuts 法と呼ばれる手法を開発した。ActCuts 法は、前年度までに開発した ActScan 法を補完および拡張したものである。ActCuts 法は、まず CT 画像から脂肪組織に相当する CT 値をもつピクセルを抽出し、画像上で隣接したピクセル同士を辺で結ぶことによって、CT 画像全体を覆うグリッド状のピクセル間ネットワークを生成する。この際に、CT 値が類似し合ったピクセル群がセグメント化される性質に注目して内臓脂肪と皮下脂肪の分離を行う。両脂肪が腹膜によって完全に分離される理想的な状況では、連結した部分ネットワークを抽出することによって、腹膜の内外の脂肪を抽出できるが、実際のデータでは腹膜の位置自体も予測する必要がある。また、腹膜が画像上で断続的になっており、両脂肪が腹膜によって完全に分離されていないこともある。そこで、ActCuts 法では、明らかに腹膜の中と外に位置するピクセルペア群を初期解から抽出し、このピクセルペア群間でネットワークが最も「細く」なっている位置（隘路・くびれ）を腹膜として抽出する。初期解として、前年度までに開発した ActScan 法による解、或いは、新規に開発した GUI によって人間が対話的に与える解を用いることができる。

図 1 は、ピクセル間ネットワークに基づいた脂肪種判定アルゴリズムの概略を示したものである。まず、隣接した脂肪ピクセル同士をグリッド状に接続する。次に、2 つの仮想的なノード S と T を用意し、各脂肪ピクセルを、初期解で内臓脂肪であれば S に、皮下脂肪であれば T に接続する。その上で、ノード S から T に向かうネットワークフロー（水色矢印）を計算し、流れの幅が最も狭くなる部分を 2 つの脂肪の境界線として抽出する（青線）。境界線のどちら側にあるかに従って、脂肪種を判定する（内臓脂肪を赤色、皮下脂肪を緑色で表示）。ピクセル間ネットワークを隣接したスライス間にもまたがって 3 次的に定義することで、単独のスライスでは判定できない領域も体軸に沿った前後関係から情報を補完して判定精度を改善するこ

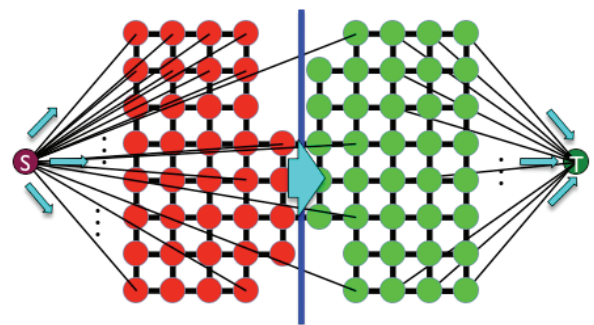


図 1、ピクセル間ネットワークに基づいた脂肪種判定アルゴリズムの概略

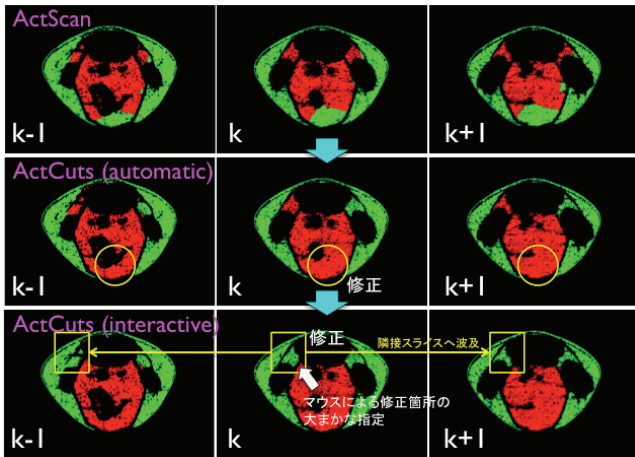


図 2、各種アルゴリズムによる CT スライスの領域判別結果

とが可能になっている。
 図 2 は、ある個体の連続する 3 スライス ($k-1$ 、 k 、 $k+1$) での判別結果を示したものである（実際の判別処理は 1 個体分の全スライスに対して行われる）。ここでは、内臓脂肪を赤色で、皮下脂肪を緑色で示している。上段は、前年度までに開発した ActScan 法による判別結果である。自由に変形する曲線をフィッティングすることによって腹膜の認識を行うが、この例では、腹膜以外の境界線を誤抽出している。中段は、本年度に開発した ActCuts 法による判別結果である。上段に示した ActScan 法によって得られた結果を初期解として用いている。丸印で示した領域の判別結果が修正されている。下段は、中段に示した結果から、対話的な初期解の追加と ActCuts 法の適用を繰り返して得た判別結果である。特定のスライス上での修正は、ピクセル間ネットワークを介して隣接するスライスにも波及するようになっている。

上述の 3 手法の判別結果を比較したものを図 3 に示す。ActScan 単独では良好な判別精度（別途生成した「手作業」による正解との比較）が得られていなかった 28 個体を選択して、全体の判別精度を対話的な ActCuts 法で得られた結果と比較したものである。横軸は個体 ($1 \sim 28 \times 3$) を示し、縦軸は判別精度（対話的な ActCuts 法との相対評価）を示している。自動的な ActCuts 法は、対話的な ActCuts 法の概ね 95~99% 程度の精度を達成していること、また、ActScan 法に比べて 10% 以上向上していることが確認できた。

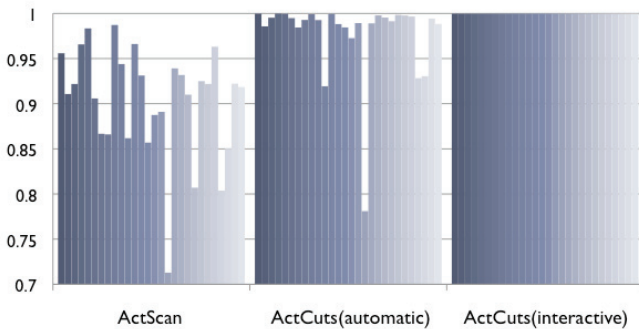


図 3、各種アルゴリズムによる CT スライスの領域判別結果比較

域ほど CT 値が高い）。上段は、各臓器での CT 値の頻度分布（割合）を示している。上段の縦軸は CT 値、横軸は中下段での位置、色によって CT 値の頻度（赤いほど頻度が高い）を示している。例えば、内臓脂肪は -250 以下の CT 値となっていることが分かる（上段）。一方、皮下脂肪は内臓脂肪に比べて均一ではなく（中段）、-125 以下の CT 値となっていることが分かる（中段）。また、褐色脂肪組織は、特異的な CT 値となっているが、その他の臓器に関しては、殆ど差異がな

これまでの開発した ActScan 法と ActCuts 法は互いに補完関係にあり、対話的な ActCuts 法も含めて良好な結果を達成している。但し、当初に予定した、脂肪以外の臓器の情報の加味は未完成的なままになっている。一つ目の原因として、臓器間での CT 値の差異が予想以上に少なかったことが挙げられる。図 4 は、下段に示す臓器の CT 値の分布を示したものである。中段は、特定の断面における CT 値の分布を色で示している（赤い領

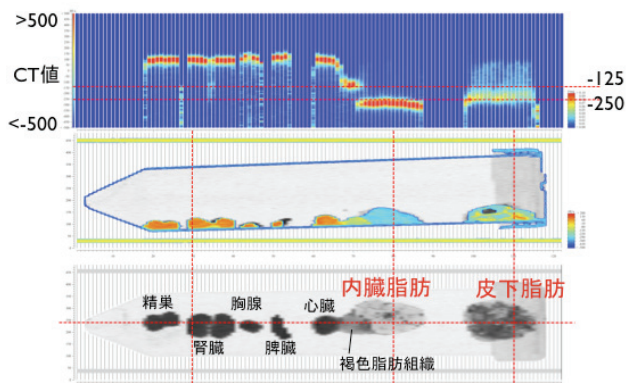


図 4、各種臓器の CT 値の分類

いため、CT 値の分布のみでは抽出できないことが分かる（上段）。

このため、脂肪以外のより多くの臓器の情報を加味するためには、各臓器の位置的な情報を用いた抽出処理が必要になる。図 5 は、1 個体分のスライスを統合することによって得たマウス全体像を特定の 2 平面で仮想的に切断したものである。横軸はスライス番号を示しており、実際のスライス位置を縦線で表示している。各ピクセルは、CT 値に従って着色しており（赤い領域ほど CT 値が高い）、また、隣り合うスライスの CT 値を用いてスライス間領域の CT 値を予測および補完して、擬似的に解像度を高めて表示している。これまでは、主に体軸に垂直な断面による解析を行ってきたが、体軸に平行な断面を用いた解析の可能性を示しており、多様な角度からの解析の必要性を示唆しているものと考えられる。

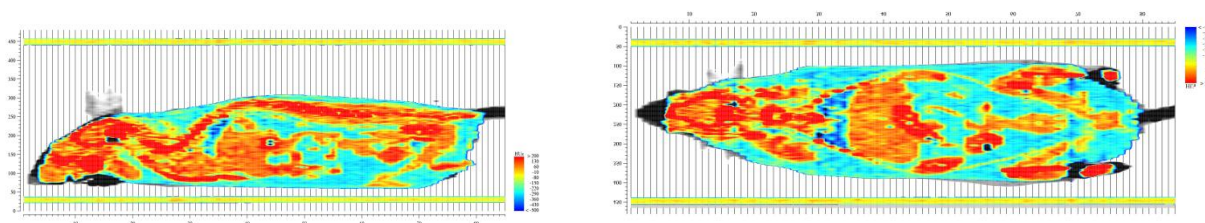


図 5、スライスを統合することによって得たマウス全体図

マウス体内構造モデルの構築

マウスの体内に存在する臓器や構造に基づいた特徴点の位置関係を記述したモデルを構築した。体内構造モデルは、プログラミング言語の手続きとして記述されており（従って、特定サンプルのデータに注釈を付したのではない）、任意のサンプルの CT データに適用することによって、特徴点群を自動的に抽出し、それらの特徴点間の位置関係を出力することができる。これによって、脂肪種の判定処理に、皮膚や骨格や肺などの他の臓器の形状を加味できるようになった。図 6、は体内構造モデルを胸部(A)および腰部(B)のスライスに適用し、その結果の一部を図化したものである。皮膚近傍は桃色に着色されている。加えて、胸部で肋骨が環状配置（赤線の多角形）されている位置関係や、腰部で骨盤や大腿骨の位置関係を抽出し、これらの位置関係に基づいて、皮下脂肪が存在する領域（水色）を絞り込んでいる。また、胸部では肋骨周囲の骨格筋と心臓を構成する筋肉が分離されている。これらの処理は人手の介入を経ずに自動で行っている。これらの結果に基づいて、脂肪種の判定を行った結果が図 7 である。体内構造モデルは、サンプル毎の変異に対してロバストである必要がある。現在は、主として比較的変異が少ない同一の近交系（C57BL/6）内の個体群を対象としているが、F₂ 個体群を用いた QTL 解析を

図 6、体内構造モデルの胸部および腰部スライスへの適用結果

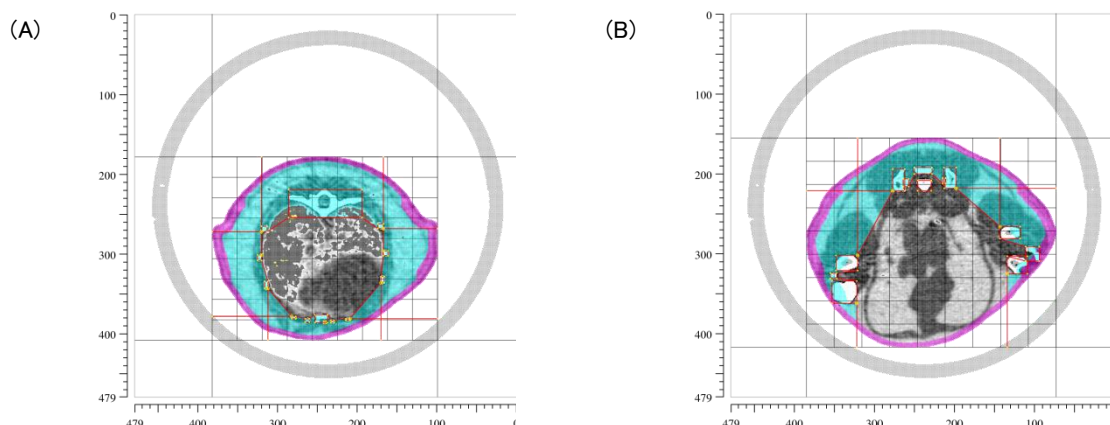
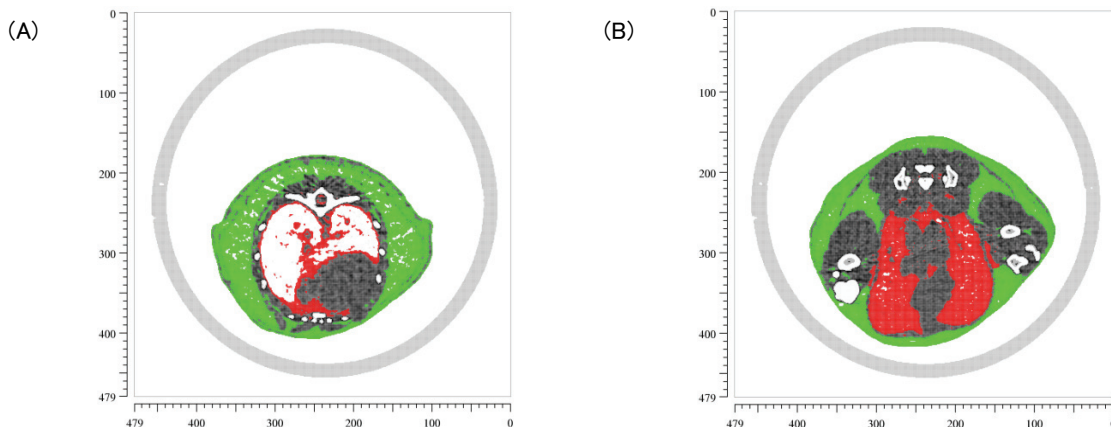


図7、体内構造モデルの適用結果を用いた脂肪種の判定結果



始めとするケースとコントロールを対象とする解析においては、対象個体群内での体内構造のバリエーションが大きくなるため、絶対的な数値に依存した情報ではなく、体内での位置関係といったトポロジー情報（位置関係）が体内構造モデルの記述に不可欠である。

データ処理パイプラインの構築

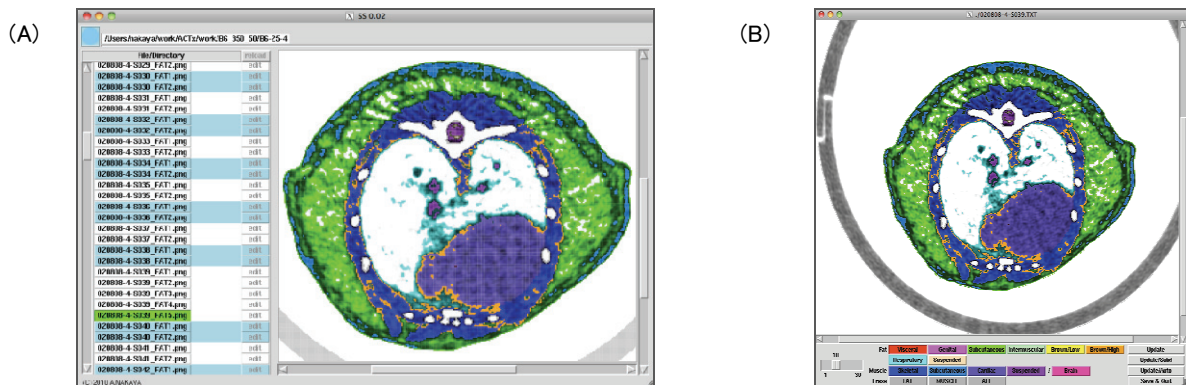
一連の処理を自動処理できる解析パイプラインを構築した。パイプラインは、以下の複数のステージから構成される：(1)骨の認識、(2)体内空気（肺・気管や消化管内）の認識、(3)スライスの体内での位置判定、(4)皮膚の認識、(5)体内モデルの適用、(6)動的輪郭法による脂肪種分類、(7)グラフカットによる脂肪種分類の修正、(8)筋肉の抽出と分類。また、矢状断面や水平断面データの生成も行う。スライス位置の判定(3)は、各スライス内で、骨に該当するピクセルが構成する平面上での連結成分（骨体の断面に相当）の数や、体内空気の量（肺部に多量に存在）等に基づいて判定を行う。筋肉の抽出と分類(8)は、筋肉と腹腔内の臓器や生殖器との境界の判別を行う（筋肉とこれらの臓器の CT 値は同一範囲内なので形状に基づいた判別処理が必要）。また、胸部では、肋骨や肩胛骨周囲の筋肉と心臓の分離も試みている。これらの処理は、CT スキャナから得られた個体データを格納したディレクトリを指定すれば、自動的に実行することが可能である。各ステージは独立のプログラムとなっており、前ステージまでに生成されたデータを後続のステージが入力データとして使用する仕様になっている（パイプライン実行）。各ステージの処理をモジュール化することにより、処理の追加が容易になっている他、ステージ間で、人間による修正等も行えるようになっている。さらに、各ステージでの各スライスの処理は、同一サンプルの他スライスの結果に依存しないモードも備えるようにもしている。これによって、複数のスライスの処理を互いに独立に行うことが可能になり、複数プロセッサを用いて並列処理が行える（例えば、100 スライスの処理を 100 プロセッサを用いて並列に実行可能）。

CT データ管理システム

データ処理パイプラインが出力するデータを管理するシステムを構築した。現在は、図 8 に示すような CT データブラウザ(A)と CT データエディタ(B)から構成されている。CT データブラウザは、パイプラインの途中結果も含めてデータが閲覧可能なようになっている。また、各データのアイコンをクリックすると CT データエディタを起動することができ、データを手動で修正できる（修正後にパイプラインの後続ステージを実行して最終結果に反映できる）。CT データエディタは、各ピクセルのラベル付け（分類）を行う。画面上ではラベル種は色で表示される。その際、各ラベルには CT 値の範囲が設定されており、ラベルを指定した上で、マウスやペンタブレットを用いてある領域の塗り潰し処理を行うと、

そのラベルに設定された CT 値の範囲内のピクセルのみにラベル付け (着色) できるようになっている。特定の CT 値範囲を持つ臓器のラベル指定はこの仕組みで比較的容易にラベル付けが可能である。例えば、脂肪と筋肉が混在した領域で脂肪を抽出したい場合などが該当する (大雑把に塗り潰した後に該当する CT 値のピクセルのみを選択する)。しかしながら、胸部の肋骨間の筋肉と心臓のように CT 値が同一の範囲にある臓器同士が隣接している場合には、CT 値の範囲を指定したのみではそれぞれをラベル付けすることは不可能であり、各ピクセルを手動で塗り潰してラベル付けを行うことになる。手動処理を用いた場合、結果の「らしさ」は高くなるが、客観性や再現性は失われる。また、処理効率に関しても、数百以上の個体群を扱うことは現実的ではない (1 スライス 1 分としても、1 個体 100 スライスとすると 1 時間半以上)。現状では、脂肪の分類以外の用途では、少数個体の高精度のラベル情報を生成して、「正解データ」として体内構造モデルにフィードバックするような用途に限定されている。CT データエディタによる修正作業は、自動的に履歴を蓄積して修正結果を任意のチェックポイントに戻すことが可能になっており、人間の試行錯誤を支援する仕組みになっている。

図 8. CT データ管理システム CT データブラウザ(A)と CT データエディタ(B)



今後の課題

プロジェクト期間内において行った成果をもとに、2 種類の脂肪組織を分離する以外にも、処理対象の臓器の範囲を広げて、抽出する情報量と精度を向上させる。複数の臓器を扱うことによって、脂肪組織のみでは判別が難しい部分での抽出精度の向上や、各臓器周辺の脂肪組織の分布情報の抽出が期待できる。また、CT 画像からの脂肪組織、あるいは特定の臓器などの抽出を完全自動化で行うための問題点は今後検討されるべきである。現段階では、完全自動化を目指す以外に修正作業の時間を短縮させるアルゴリズムに労力を掛ける必要性が高い。

課題 b. マウスエネルギー代謝関連表現型計測法の開発と実験データの収集 (高田、前野、城石)

【研究目的】

本テーマは複雑な生命システムを制御する遺伝子群とそれらの関与するネットワークを統計遺伝学的に検出するための基盤の整備を目的としている。具体的には、西ヨーロッパ産汎用実験系統である C57BL/6J (B6: *Mus musculus domesticus*)、およびこれとは遺伝的距離が離れた関係にある日本産野生由来系統の MSM/Ms (MSM: *Mus musculus molossinus*)、MSM/Ms と同亜種に属する JF1/Ms (JF1) を使用し、解剖学的、血液生化学的パラメータをはじめとしたエネルギー代謝関連表現型を経時的に計測する。最終的には、得られた数値データを使用し、両者のゲノム多様性を指標とした多段階の全ゲノムスキャンを行い、各種表現型の責任遺伝子領域とそれらが関与するエピスタシスの検出を行うことを

| | | B6MSM | | B6JF1 | |
|--|------------------|-------|--------|-------|--------|
| | | Male | Female | Male | Female |
| | Lod score 2 - 3 | 20 | 18 | 20 | 14 |
| | Lod score over 3 | 18 | 20 | 8 | 15 |
| | | 38/41 | 38/40 | 28/33 | 29/33 |
| | lod > 2 (%) | 92.68 | 95.00 | 87.88 | 90.91 |

図 3、本プロジェクトにおいて検出したエネルギー代謝関連 QTL の検出数

すべての項目について解析をおこなった結果、各形質には多因子により制御されていると考えられる遺伝子座が複数確認できる。これと比較して、例として示した毛色は一点の強い効果を示す遺伝子座が相加効果を示すエリアに雌雄に共通して確認できる。これらの QTL が相加的効果以外にも相乗的効果の存在が示唆されることから、今後はこれらの効果の真偽を統計学的に検証する必要がある。現在進行中の課題である B6 と MSM ならびに B6 と JF1 系統の F2 世代の解析に関しては、高密度の SNP 検出システムが整備でき次第、遺伝マーカー数を増やし、より密度の高いマッピングパネルを作成して QTL の候補領域を絞り込む。今年度までに実施した B6 - MSM、B6 - MSM のデータを使用して各 QTL 解析について同様の解析を行うとともに各形質間の相関解析を実施する。

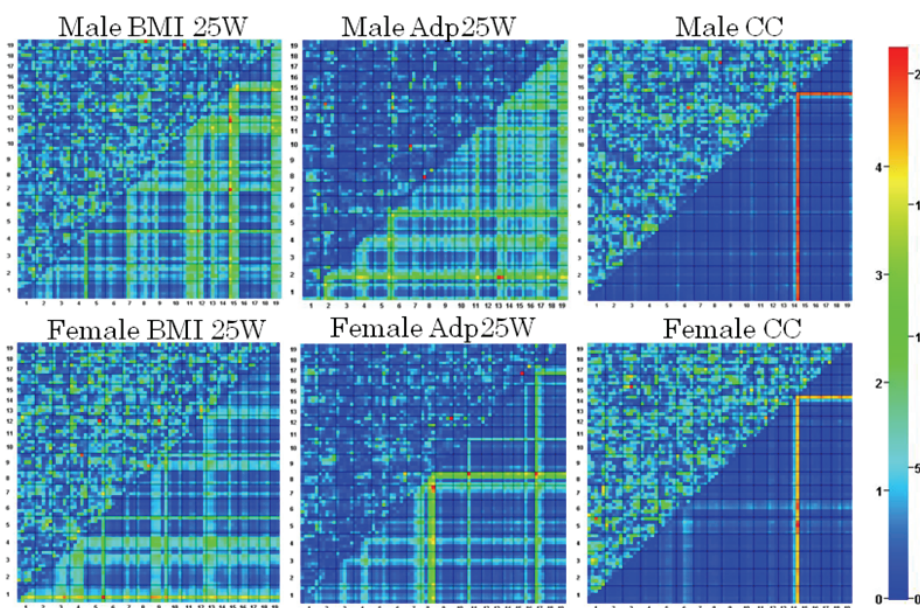


図 4、B6x JF1 F2 解析により示唆されるエネルギー代謝関連表現型の QTL。Two QTL genome scan の結果を示す。雌雄の 25 週令の BMI、アディポネクチン (Adp)、コントロールとして毛色 (CC) の結果を例として示す。各項目の左上は epistasis に関する LOD、右下は相加効果に関する LOD を示す。各図の下および左から染色体 1 番から順に 19 番までを示している。各形質には多因子により制御されていると考えられる遺伝子座が確認される。このほかにもエネルギー代謝に関する QTL 解析では複数の遺伝子座による効果と考えられる形質が複数観察された。

MassArray による SNP 検出システムの整備

QTL 解析を多段階で行うための SNP 検出システムを整備した。B6 および MSM 系統間の多型情報を利用して、全ゲノムを対象とした 1,855 か所に SNP 検出用のマーカーを設定した。最終年度において、マウスエネルギー代謝関連表現型を指標とした QTL 解析を高密度で行うための SNP 検出システムを整備しているが、マスアレイを使用した SNP 検出に対してパイロット試験によりデータを得ることができた。B6 および MSM 系統間の多型情報を利用して、全ゲノムを対象とした約 2,800 か所に SNP 検出用のマーカーを設定した。この中から、1,855 か所について SNP を検出するためのプライマーセットの合成を行った。B6 および MSM 系統、さらには両者の交配により得られた F1 個体を使用して、実際に解析を行った例を図 4、に示す。図には染色体一番のおよそシーケノム社製マスアレイシステムによる SNP

る有効な手法については未だ確立していない。これらの構造、発現多型の大量データの解析、いわゆる e-QTL と形質多様性のアソシエーション解析においては、特に融合研究としての特性を生かし取り組む。

【平成 17～21 年度の進捗】

まず 16 species の野生イネコアコレクション 46 系統の 18 形質を対象に、各形質の測定とデータ間での相関解析を行い、種間の相関関係と種内の相関を求めた。この結果、種の異なる系統間での比較は、ゲノム配列にも形質比較にも困難な点があると思われたので、栽培イネ祖先種である *Oryza rufipogon* 36 系統に絞って、穂形質と分けつ能についての予備調査を行った。さらにこの中から多年生と一年生の 2 群の 14 系統を選別し、各系統 3 個体のサンプルで、粒重、分けつ節位、葯長、柱頭長、子房長、穂型、分けつ数、草丈、種子稔性、発根・再生能の計 10 種類の量的形質についてのデータを得て、系統毎の特性を抽出した (図 1、2)。

図 1 に 14 系統における葯長を、図 2 に開花後の茎からの発根・再生力を数値で示した。個々の形質の解析から、多年生系統群、一年生系統群の間で柱頭長、種子稔性、穂の形 (着粒数と相関あり)、発根・再生力など多くの形質が、大きく 2 つのグループに分かれた。

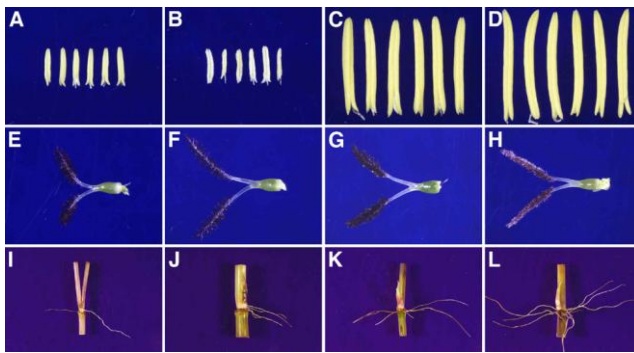


図 1、*O. rufipogon* の葯 (A-D)、雌蕊 (E-H)、挿木 (I-L)。1 年生型の W0106 (A, E, I)、W0610 (B, F, J)。多年生型の W1945 (C, G, K)、W1230 (D, H, L)。

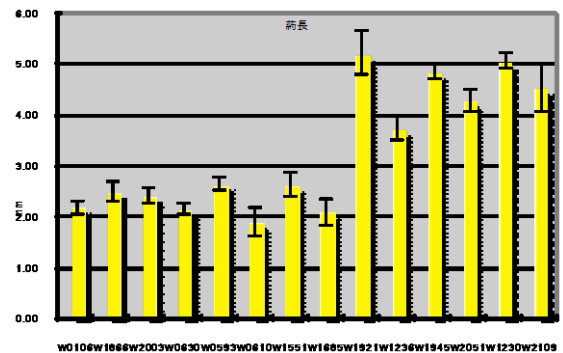


図 2、*O. rufipogon* 14 系統における葯長 (mm)。

これらの系統の中で一年生、多年生グループの特徴を持つそれぞれの代表的系統、W0106 と W1921 の交配を行い、F1 種子および F2 種子を得た。これらの種子から交配自殖系統や染色体置換系統の子孫を樹立し、アソシエーション解析と平行して QTL および e-QTL 解析用の系統としてこれらの樹立系統を利用する。同一素材を異なる方法で解析する事により、方法の長短が明らかになり、それぞれの利点を利用しながら、目的の遺伝因子により早く近づく方策を講じる。さらに、これらの系統では、形質と遺伝的多型とのアソシエーション解析を行うため、同時に葉や幼穂などのサンプルを収集し、まず両親のアレイ解析を行った。

この中で、遺伝子構造または遺伝子発現の差がマイクロアレイにより測定可能であるか検討した結果は以下のとおりであった。Affymetrix 社の GeneChip Rice Genome array は主に栽培イネ *O. sativa* の *japonica* 由来の転写産物を基に設計されている。アレイを用い発現多型や塩基多型を検出する手法は課題 d にて、ゲノム配列が既知の栽培イネの日本晴 (*japonica*) と 93-11 (*indica*) 間で開発してきた。栽培イネ *indica* の祖先種である野生イネ *O. rufipogon* ; W106 と、栽培イネとはゲノム種も異なる *O. punctata* の W1514 の 2 種、いずれも塩基配列が未知の系統を用いて、アレイで発現量多型や塩基配列多型が別々に検出できるか検討した。ゲノム DNA のハイブリダイゼーションにより系統間の塩基配列の差がどの程度シグナル強度に影響を与えるか、また mRNA ハイブリダイゼーションにより発現デー

タにその塩基配列構造差が影響するのか比較した（図 3）。対角線より右下は日本晴のシグナルが高い事を示す。W106 対日本晴のシグナル強度比較は、93-11 対日本晴のデータとゲノムハイブリダイゼーション, mRNA ハイブリダイゼーション共に良く似ていた。アレイ設計の基データとして使われた系統と遺伝的に近い *O. sativa* 内や *O. rufipogon* の系統では下記課題「マイクロアレイ・・統計解析システムの開発」で開発した手法を使い発現多型や塩基多型を検出する事が可能と考えられる。一方 BB ゲノム種 W1514 は、塩基配列が大きく異なることがゲノムハイブリダイゼーションのシグナル強度より明らかである。アレイが日本晴を基に設計されているので、日本晴よりもシグナル強度がかなり低いものが多いのは予想通りであるが、日本晴よりもシグナル強度が高いものも多数ある。これは、プローブを設定した配列が W1514 では重複しているためと考えられる。さらに、W1514 対日本晴の mRNA ハイブリダイゼーションにおけるシグナル強度比較も、他の 2 つとはかなり異なる。課題 d で開発した手法に加え、遺伝子の重複の程度を調べ重複遺伝子の発現量を推定する方法や、アレイのプローブ設定領域に多数の塩基置換が含まれる場合の遺伝子発現量を推定する改良等の課題が残る。

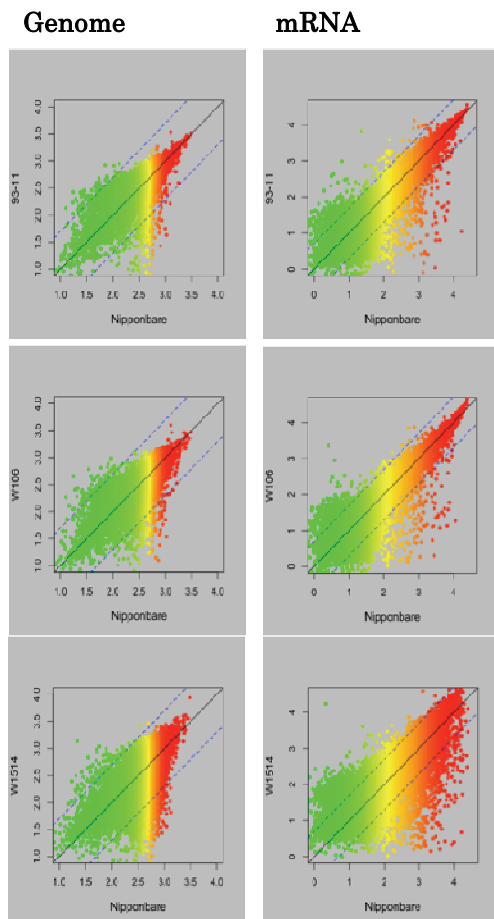


図 3、*O. sativa*, japonica (Nipponbare), indica (93-11), *O. rufipogon* (W106), *O. punctata* (W1514) 間の遺伝子発現比較。

今後の課題

上記課題と共に、発現遺伝子の塩基配列が良く調べられている系統では、次世代シーケンサーを用いて遺伝子の数塩基の差や、発現の差を直接観測する試みが行われており、ゲノムの塩基配列が既知の系統と遺伝的に離れた系統で、次世代シーケンサーでの発現遺伝子の解析を行う可能性を検討する必要がある。発現遺伝子の塩基配列多型と発現量多型を同時に調べることができるか、同時にゲノム塩基配列も決定する必要があるのか、アレイ解析を組み合わせた方が良いのかなど、課題 d と合わせて今後実際にデータを測定し、検討していく必要がある。

課題 d. マイクロアレイによるイネ・マウス遺伝子発現データの統計解析システムの開発（藤澤、江口、池田（思）、坂口、高田、岡（木曾）、城石、倉田、春島、堀内、久保）

【研究目的】

本課題では、多様な生物特性を理解するため、塩基配列レベルの多型および遺伝子発現量の多型を検出する方法の確立を目的とする。Affymetrix 社の GeneChip expression array はゲノム全体の遺伝子の発現量を比較する事を目的に作られた物だが、ゲノム全体に渡る塩基配列の差も見積もる事が出来、生物の多様性を研究する上で有用性が期待される実験手法である。しかしながら、アレイ設計と異なる系統の遺伝子発現と塩基多型を区別して検出することは容易ではない。GeneChip は遺伝子発現量の推定を PM プローブと MM プローブと呼ばれる 25mer のオリゴプローブの 11 組を 1 つのセットとして用いている。PM プローブは標的とする遺伝子の配列に対して完全に相補的な配列で、MM プローブ

は PM プローブの中央の塩基をその相補塩基に置換えた配列でバックグラウンド補正用に設計された物である。本課題は、まずこのセット内のプローブシグナルから、遺伝子の発現量の差と塩基配列の差を別々に見積もる方法の確立を目的とする。また、アレイによって検出される塩基多型 (SFP: single feature polymorphism) は、ミュータントや QTL の原因遺伝子の単離、集団遺伝学的解析に利用可能なので、ゲノムハイブリダイゼーションによる SFP の検出方法の検討を行い、方法の改良等による実用的利用の可能性を検討する。

【平成 17-21 年度の進捗】

遺伝子の発現量の差と塩基配列の差を別々に見積もる方法については、発現及び塩基多型の同時検出法 (SNEP: Simultaneous detection of nucleotide and expression polymorphisms) の開発を行った。SNEP はある遺伝子の塩基配列が品種間で同じで発現量が異なる時、PM プローブの対数シグナル値の品種間差がセット内で一定値であることに着目し、セット内で対数シグナル値の品種間差がはずれ値となるプローブを探す事により遺伝子の塩基多型を検出し、対数シグナル値の品種間差が 0 より有意に異なる時に品種間で発現量の多型を検出する方法である。はずれ値によるマスク効果は、尤度比検定によるセット内の他のはずれ値そのものの検出力の低下と発現量多型の検出力の低下を招くが、統計的ロバスト法を利用し克服する検定手法を開発、統計言語 R によるソフトウェアを作製した。

次に、アレイ実験の特徴である繰り返し実験が少ないが次元が多い検定に対し改良するため、ロバスト法にチューニングパラメーターを導入、更にアレイで 2 系統の発現遺伝子を比較し SNEP を行う時、両方の系統で塩基多型の検出が可能となる両側検定にも対応出来るよう改良した。本研究で開発した SNEP 法と尤度比検定による

SFP 検出効率の比較を大麦の 5 種類の組織の mRNA のアレイのデータを用い、既存の方法との比較を受信者動作特性曲線 (ROC curve: receiver operating characteristic curve) の比較によって行った。図 1 に示した通りどの組織のデータにおいても SNEP 法による優位性が示された。また、イネやマウスの mRNA のアレイデータについてもロバスト法のチューニングパラメーターを導入する事により、SFP 検出の更なる効率化が図られ、SNEP 法の論文を公表できた。

大麦などゲノムサイズの大きな種ではゲノムハイブリダイゼーションによる SFP を検出する事は出来ないが、アラビドプシス (~125 Mb) や出芽酵母 (~12 Mb) では、ゲノムハイブリダイゼーションで SFP を検出し、ミュータントや QTL の原因遺伝子の単離、集団遺伝学的解析に用いられている。イネ (~400 Mb) でゲノムハイブリダイゼーションによる SFP の検出を試みた。

イネのゲノムハイブリダイゼーションの実験条件の検討を行い最適条件を求めた。この実験条件で塩基配列が既知のジャポニカとインディカの 2 品種を使って SFP 検出を、SNEP 法を含め既存の検出方法、バックグラウンド補正法、ノーマライゼーション法により検討した。その結果、バックグラウンド補正なしで、検出法としては SAM (significant analysis of microarray) による SFP の検出が最適で

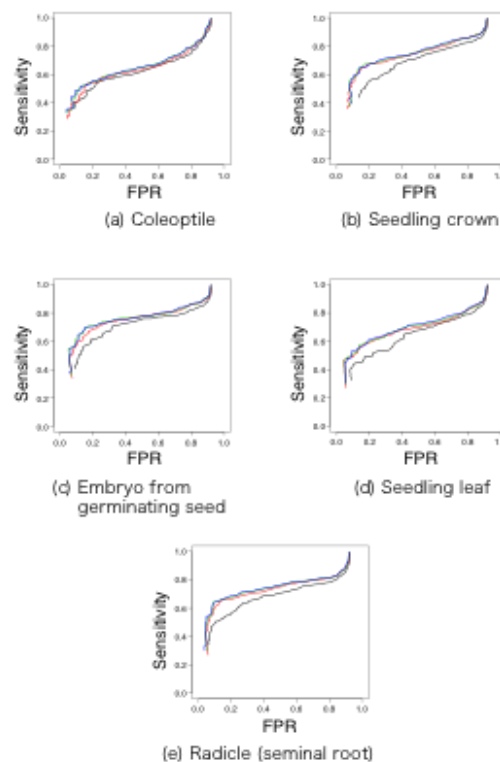


図1 ROC curve for SFP detection. The x- and y-axes are the FPR and sensitivity, respectively. The black, red, green and blue lines are based on LRT, SNEP ($\gamma = 0.2$), SNEP ($\gamma = 0.5$) and SNEP ($\gamma = 0.8$), respectively, using appropriate significance levels, which range from 10^{-26} to 1.

あり、27,365 個の真の SFP を検出できた。ゲノムハイブリダイゼーションによる SFP の検出率は 38.6% と低いが、解析対象とするプローブ数が多いため検出できた SFP の数は多い。同じ品種の組合せの young panicle の mRNA の SNEP 解析では、SFP 検出率は 53.6% と高いが、解析対象となる高発現遺伝子の数が少ないため、真の SFP が 3,577 個しか検出できない。イネで SFP の検出に限ればゲノムハイブリダイゼーションの方が数多くの SFP の検出が可能である。この結果は論文としてまとめ、公表した。

系統間の遺伝子発現量差を GeneChip にて推定する時、セット内の SFP の数が少ないときは SNEP によって正確に推定できるが、SFP の数が多いと推定する方法が無い。しかし、ゲノムや mRNA ハイブリダイゼーションによる SFP 検出が困難であることは、逆に塩基配列多型があっても多くのプローブではシグナル値に影響が無いことを意味する。セット内で SFP が多数を占める遺伝子についても、SFP によるシグナル減少量が少ないプローブを推定し、系統間の遺伝子発現量差を見積もる方法を開発した。

今後の課題

現在、新たに系統間の遺伝子発現量差をセット内の SFP の多さに関わらず推定する方法を開発したので、他の実験方法による確認と問題点の抽出を行っている所である。

① 成果物（知見・成果物・知的財産権等） （プログラム開発）

- 1) SNEP 法 ; Affymetrix 社のマイクロアレイを用い、塩基配列多型を同定し、同時に系統間の遺伝子発現量の差も検出する SNEP 法を開発した。SNEP 法によりゲノムサイズに関わり無く、遺伝子の塩基配列多型及び発現量の多型が検出できた。プログラムは <http://www.ism.ac.jp/~fujisawa/SNEP/> にて公開している。
- 2) ActScan 法 ; CT 画像から腹筋線を自動抽出するためのアルゴリズム。
- 3) ActCuts 法 ; CT 画像のピクセル間に定義したネットワーク形状の解析に基づく脂肪組織分布解析法。

② 成果発表等 〔学術論文〕

- 1 Miyabayashi, T., Nonomura, K., Morishima, H. and Kurata N. Genome size of twenty wild *Oryza* Species determined by flow cytometric and chromosome analyses. *Breeding Science* 57: 73-78. 2007
- 2 Nonomura, K., Morohoshi, A., Nakano, M., Eiguchi, M., Miyao, A., Hirochika, H. and Kurata, N. A germcell-specific gene of the *ARGONAUTE* family is essential for the progression of premeiotic mitosis and meiosis during sporogenesis in rice. *Plant Cell* 19:2583-2594. 2007.
- 3 Oka, A., Aoto, T., Totsuka, Y., Takahashi, R., Ueda, M., Mita, A., Sakurai-Yamatani, N., Yamamoto,

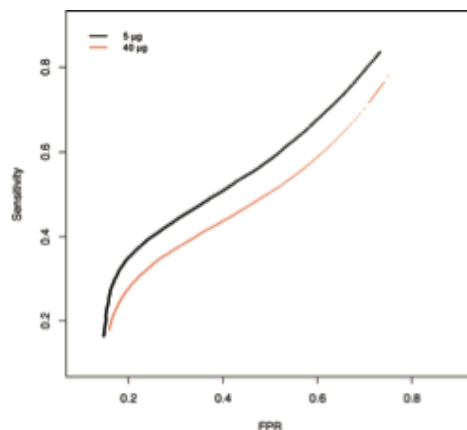


図 2 イネのゲノムハイブリダイゼーション時の添加する DNA 量による SFP 検出効率の差

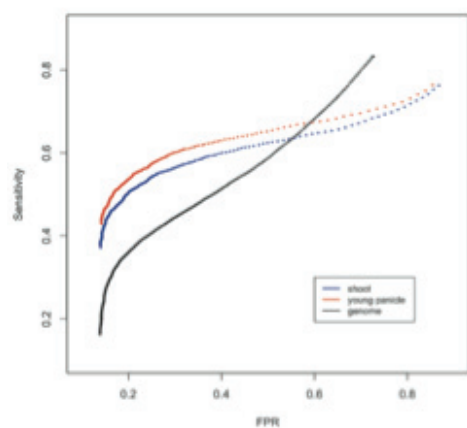


図 3 イネの mRNA とゲノムハイブリダイゼーションとの SFP 検出効率の差

- H., Kuriki, S., Takagi, N., Moriwaki, K. and Shiroishi, T. Disruption of Genetic Interaction Between Two Autosomal Regions and the X Chromosome Causes Reproductive Isolation Between Mouse Strains Derived From Different Subspecies. *Genetics* 175, 185-197 (2007).
- 4 Ito, N. and Kurata, N. Disruption of KNOX gene suppression in leaf by introducing its cDNA in rice. *Plant Science*. 174: 279-289. 2008.
 - 5 Thirumurugan, T., Ito, Y., Kubo, T., Serizawa, A. and Kurata, N. Identification, characterization and interaction of HAP family genes in rice. *Mol Gen Genomics*.279: 279-289. 2008.
 - 6 Suzuki, T., Eiguchi, M., Kumamaru T., Satoh, H., Matsusaka, H., Moriguchi, K. and Kurata, N. MNU-induced mutant pools and high performance TILLING enable finding of any gene mutation in rice. *Mol Gen Genomics*.279:213-223. 2008.
 - 7 Suwabe, K. Suzuki, G. Takahashi, H. Shiono, K. Endo, M. Yano, K. Fujita, M. Masuko, H. Saito, H. Fujioka, T. Kaneko, F. Kazama, T. Mizuta, Y. Separated transcriptomes of male gametophyte and tapetum in rice: validity of a laser microdissection (LM) microarray. *Plant Cell Physiol*.49:1407-1416. 2008.
 - 8 Hobo, T. Suwabe, K. Aya, K. Suzuki, G. Yano, K. Ishimizu, T. Fujita, M. Kikuchi, S. Hamada, K. Miyano, M. Fujioka, T. Kaneko, F. Kazama, T. Mizuta, Y. Takahashi, H. Shiono, K. Nakazono, M. Tsutsumi, N. Nagamura, Y. Kurata, N. Watanabe, M. Matsuoka, M. Various spatiotemporal expression profiles of anther-expressed genes in rice. *Plant Cell Physiol*.49:1417-1428. 2008.
 - 9 Takada, T., Mita, A., Maeno, A., Sakai, T., Shitara, H., Kikkawa, Y., Moriwaki, K., Yonekawa, H. and Shiroishi, T.: Mouse inter-subspecific consomic strains for genetic dissection of quantitative complex traits, *Genome Res*. 18, 500-508 (2008).
 - 10 Sakuraba Y, Kimura T, Masuya H, Noguchi H, Sezutsu H, Takahasi KR, Toyoda A, Fukumura R, Murata T, Sakaki Y, Yamamura M, Wakana S, Noda T, Shiroishi T, Gondo Y. Identification and characterization of new long conserved noncoding sequences in vertebrates. *Mamm Genome*. 2008 19(10-12): 703-712.
 - 11 Kaminuma E, Masuya H, Miura I, Motegi H, Takahasi KR, Nakazawa M, Matsui M, Gondo Y, Noda T, Shiroishi T, Wakana S, Toyoda T. Objective evaluation measures of genetic marker selection in large-scale SNP genotyping. *J Bioinform Comput Biol*. 2008 6(5): 905-17.
 - 12 Fujii T, Tamura M, Tanaka S, Kato Y, Yamamoto H, Mizushina Y, Shiroishi T. Gasdermin D (Gsdmd) is dispensable for mouse intestinal epithelium development. *Genesis*. 2008 46(8): 418-23.
 - 13 Liu YH, Takahashi A, Kitano T, Koide T, Shiroishi T, Moriwaki K, Saitou N. Mosaic genealogy of the *Mus musculus* genome revealed by 21 nuclear genes from its three subspecies. *Genes Genet Syst*. 2008 83(1): 77-88.
 - 14 Tanaka, E., Tamura, Y., Hosoya, M. and Shiroishi, T., "Protrusion Fourier Descriptor: Skeleton-based Representation of Open Curves", *Forma*, Vol. 23, 9-18, 2008.
 - 15 Tanaka, E. and Tamura, Y., "Skeleton-based Fourier Descriptor of Open Curves" , *Pacific Science Review*, Vol. 10, No. 10, 194-198, 2008.
 - 16 Takahashi A., Shiroishi T., Koide T.: Multigenic factors associated with a hydrocephalus-like phenotype found in inter-subspecific consomic mouse strains. *Mammalian Genome*, 29:333-338, 2008.
 - 17 Takahashi A., Nishi A., Ishii A., Shiroishi T., Koide T.: Systematic analysis of emotionality in consomic mouse strains established from C57BL/6J and wild-derived MSM/Ms. *Genes, Brain and*

Behavior. 7:849-858, 2008.

- 18 Kawakita, M., Eguchi, S., Boosting method for local learning in statistical pattern recognition. *Neural Computation* 20, 2792-2838 (Nov., 2008)
- 19 Fujisawa, H., Eguchi, S., Robust parameter estimation with a small bias against heavy contamination. *Journal of Multivariate Analysis* 99, 2053-2081 (Oct., 2008)
- 20 Takenouchi, T., Eguchi, S., Murata, N., Kanamori, T., Robust Boosting Algorithm Against Mislabeling in Multiclass Problems. *Neural Computation* 20, 1596-1630 (June, 2008)
- 21 Araki K, Takeda N, Yoshiki A, Obata Y, Nakagata N, Shiroishi T, Moriwaki K, Yamamura K. Establishment of germline-competent embryonic stem cell lines from the MSM/Ms strain. *Mamm Genome*. 2009 20(1): 14-20.
- 22 ※Umemori, J., Nishi, A., Lionikas, A., Sakaguchi, T., Kuriki, S., Blizard, D.A. and Koide, T.: QTL analyses of temporal and intensity components of home-cage activity in KJR and C57BL/6J strains, *BMC Genet*, 10:40 (2009) .
- 23 ※Fujisawa, H., Horiuchi, Y., Harushima, Y., Takada, T., Eguchi, S., Mochizuki, T., Sakaguchi, T., Shiroishi, T., Kurata, N. SNEP: Simultaneous detection of nucleotide and expression polymorphisms using Affymetrix GeneChip. *BMC Bioinformatics* 10, 131, 2009.
- 24 ※Tsuda, K., Ito, Y., Yamaki, S., Miyao, A., Hirochika, H., Kurata, N. Isolation and mapping of three rice mutants that showed ectopic expression of KNOX genes in leaves. *Plant Science* 177, 131-135, 2009.
- 25 ※Win, KT., Kubo, T., Miyazaki, Y., Doi, K., Yamagata, Y., Yoshimura, A. Identification of two loci causing F1 pollen sterility in inter-and intraspecific crosses of rice. *Breeding Science* 59, 411-418, 2009.
- 26 ※Takano-Kai N., Jiang H., Kubo T., Sweeney M., Matsumoto T., Kanamori H., Padhukasahasram B., Bustamante C., Yoshimura A., Doi K., McCouch S. Evolutionary History of GS3, a Gene Conferring Grain Length in Rice. *Genetics* 182, 1323-1343, 2009.
- 27 ※Kubo T. Yamagata Y, Eguchi M., Yoshimura A. A novel epistatic interaction at two loci causing hybrid male sterility in an inter-subspecific cross of rice (*Oryza sativa* L.). *Genes Genet. Syst.* 83, 443-453, 2009.
- 28 ※Takahashi, A., Tomihara, K., Shiroishi, T., and Koide, T.: Genetic mapping of social interaction behavior in B6/MSM consomic mouse strains, *Behavior Genet*, 40, 366-376 (2010).
- 29 ※Yamazaki, Y., Akashi, R., Banno, Y., Endo, T., Ezura, H., Fukami-Kobayashi, K., Inaba, K., Isa, T., Kamei, K., Kasai, F., Kobayashi, M., Kurata, N., Kusaba, M., Matuzawa, T., Mitani, S., Nakamura, T., Nakamura, Y., Nakatsuji, N., Naruse, K., Niki, H., Nitasaka, E., Obata, Y., Okamoto, H., Okuma, M., Sato, K., Serikawa, T., Shiroishi, T., Sugawara, H., Urushibara, H., Yamamoto, M., Yaoita, Y., Yoshiki, A., and Kohara, Y. NBRP databases: Databases of biological resources in Japan. *Nuc.Acid.Res.* D26-D32, 2010.
- 30 ※Fujii, S., Yamada, M., Fujita, M., Itabashi, E., Hamada, K., Yano, K., Kurata, N., Toriyama, K. Cytoplasmic-nuclear genomic barriers in rice pollen development revealed by comparison of global gene expression profiles among five independent cytoplasmic male sterile lines. *Plant Cell Physiol.* 51,0610-620, 2010.
- 31 ※Horiuchi, Y., Harushima, Y., Fujisawa, H., Mochizuki, T., Kawakita, M., Sakaguchi, T., Kurata N. A simple optimization can improve the performance of single feature polymorphism detection by

Affymetrix expression arrays. *BMC Genomics* 11, 315, 2010.

[会議録]

- 1 平成 17 年 8 月 19 日、統計数理研究所において平成 17 年度 新領域融合研究プロジェクト「生物多様性解析」第 1 回ワークショップを開催した。
- 2 倉田のり, 渡辺正夫: 国立遺伝学研究所研究集会「高等植物の受粉・受精形質 (雌雄間相互作用形質) を統御する遺伝子の分子遺伝学的解析」国立遺伝学研究所, 2007 年 11 月 16 日. 三島
- 3 渡辺正夫, 倉田のり 国立遺伝学研究所研究会「高等植物の生殖過程を制御する因子の多様性と生殖隔離」, 国立遺伝学研究所, 2008 年 11 月 5 日-6 日, 三島.
- 4 長戸康郎, 倉田のり 国立遺伝学研究所研究会「イネ分子遺伝学の新展開」, 国立遺伝学研究所, 2008 年 12 月 19-20 日, 三島
- 5 Mollah, Md. Nurul Haque, Eguchi, Shinto Robust Composite Interval Mapping for QTL Analysis by Minimum β -Divergence Method. *IEEE International Conference on Bioinformatics and Biomedicine 2008*, 115-120, 978-0-7695-3452-7, 2008.11
- 6 倉田のり 科学研究費特定領域研究「植物の生殖過程におけるゲノム障壁」班会議, 2009 年 7 月 7 日-8 日, 東京.
- 7 長戸康郎, 倉田のり 「イネ研究の多様性と展望」 国立遺伝学研究所研究会, 2009 年 10 月 30 日-31 日, 三島.
- 8 磯貝彰、倉田のり、渡辺正夫「International Symposium of Cell-Cell Communication in Plant Reproduction: from pollination to fertilization」、学術創成研究「植物自家不和合性」、特定領域研究「植物ゲノム障壁」、若手研究(S)「アブラナ科自家不和合性」共催国際シンポジウム, 2010 年 3 月 11 日-12 日, 奈良.
- 9 倉田のり 「新領域融合研究センター」第 II 期プロジェクト調査研究計画会議, 2010 年 1 月 16 日, 三島.

[解説・総説]

- 1 Kurata, N. Chromosome and genome evolution in rice. In “Rice Biology in the Genomics Era.” Ed. Hirano, Hirai, Sano and Sasaki. Springer, p235-243. 2007
- 2 Ohtsubo, H., Tsuchimoto, S., Xu, J-H., Cheng, C., Kuroda, M. Y., Kurata, N. and Ohtsubo, E. Rice retroposon, *p-SINE*, and its use for classification and Identification of *Oryza* species. In “Rice Biology in the Genomics Era.” Ed. Hirano, Hirai, Sano and Sasaki. Springer, p277-289. 2007
- 3 城石俊彦 マウス亜種間コンソミック系統によるゲノム機能解析.p905-911 学術月報 特集「遺伝学」平成 19 年 11 月号
- 4 栗木哲「QTL 解析の統計モデルと検定の多重性調整」, 21 世紀の統計科学, II, 小西貞則, 国友直人 (編), 東京大学出版会, 東京, 315-356, 2008.7
- 5 Koide T, Catanesi CI, Nishi A, Shiroishi T, Kasai S, Ikeda K, Takahashi A. Systematic mapping of pain-related QTL using consomic mouse strains: Advantage of using wild-derived strains. *Brain Res J.* in press.
- 6 奥野員敏, 倉田のり 動植物育種の到達点と目標 育種学研究, 11,107-126, 日本育種学会, 中西印刷, 2009

<会議発表等>

[招待講演]

- 1 Shiroishi, T. Genome anatomy of C57BL/6 and exploration of SNPs for energy metabolism. EUMODIC First Annual Meeting. 2007. 2.22-23, Barcelona, Spain.
- 2 Shiroishi, T. Functional Genomics of Complex Traits Based on Mouse Inter-Subspecific Differences. 第23回国際生物学賞シンポジウム 2007年10月29日, 京都.
- 3 Nonomura, K., Nakano, M., Eiguchi, M., Miyao, A., Hirochika, H., Kurata, N. Rice meiosis and its relation to small RNA-mediated gene silencing. EMBO world workshop, 8th European Meiosis Meeting. Hayama, 2007, September 14
- 4 Kurata, N., Suzuki, T., Kumamaru, T., Satoh, H. High Performance Rice Mutant Screening by using modified TILLING and MNU-induced mutant pools. The 5th International Symposium of Rice Functional Genomics. Tsukuba, 2007, October 15
- 5 城石俊彦：マウスゲノム多型に基づく多因子表現型の解析．第4回生命資源研究・支援センター国際シンポジウム．熊本、2008年2月15日
- 6 城石俊彦：ゲノム解読から見えてきた実験用マウスの起源：新たな実験用マウス系統の開発と利用．第4回国免疫フォーラム特別講演．松山 2008年7月14日
- 7 Eguchi, Shinto. Information divergence geometry and its application to machine learning. The 1st MSJ-SI, Probabilistic Approach to Geometry, Kyoto, Japan.2008.8.4.
- 8 城石俊彦：マウス亜種間コンソミック系統によるゲノム機能解析．第44回高血圧関連疾患モデル学会学術総会，2008年11月22日，出雲．
- 9 城石俊彦：糖尿病研究における forward genetics アプローチ．第1回疾患モデルシンポジウム，2008年12月3日，東京．
- 10 江口 真透．タンパク質構造と進化と情報幾何．数理研短期共同研究集会「離散力学系の分子細胞生物学への応用数理」．京都，日本．2009.1.8
- 11 Kurata, N., Mizuta, Y., Kubo, T., Hrushima, Y., Genome evolution and reproductive barriers in rice. 6th International Rice Genetics symposium, Manila, Philippine, 2009,11.09
- 12 城石俊彦：マウス亜種間コンソミック系統を用いた多因子形質の遺伝解剖．第24回日本糖尿病・肥満動物学会年，2010.1.23. 大阪
- 13 Kurata, N., Mizuta, Y., Kubo, T., Ohyanagi, H., Horiuchi, Y., Harushima, Y. Genome barriers in the reproductive process of rice. International Symposium of Cell-Cell Communication in Plant Reproduction, Nara, Japan.2010.3.12

[一般講演]

- 1 Oka, A., Takada, Y., Koseki, H., Mita, A., Moriwaki, K., Shiroishi, T. Meiotic arrest in males of the B6.ChrXMSM consomic strain. 21th International Mammalian Genome Conference, 2007. 10.28-11.1, Kyoto
- 2 Takada, T., Ebata, T., Narita, T., Shin-I, T., Abe, K., Sakaki, Y., Toyoda, A., Sagai T., Mita, A., Moriwaki, K., Kohara, Y. and Shiroishi, T. Whole genome shotgun sequencing of M. M. Molossinus-derived MSM/Ms and detection of vast amount of SNPs against C57BL/6. 21th International Mammalian Genome Conference, 2007. 10.28-11.1, Kyoto
- 3 Shiroishi, T., Takada, T., Maeno, A., Mita, A., Moriwaki, K., Yonekawa, H. Functional genomics of complex traits with inter-subspecific consomic strains. 21th International Mammalian Genome

Conference, 2007. 10.28-11.1, Kyoto

- 4 Oka, A. Disruption of genetic interaction between two autosomal regions and the X chromosome causes reproductive isolation between mouse strains derived from different subspecies. The Fifth Okazaki Biology Conference: "Speciation and Adaptation" 2007. 3.11-16, Kakegawa
- 5 Fujisawa, H., Horiuchi, Y., Harushima, Y., Eguchi, S., Mochizuki, T., Sakaguchi, T., Kurata, N. A unified method for detecting single feature polymorphisms and gene expression level differences. 1st joint research meeting of ISM and ISSAS. 2007. 11.30, Tokyo
- 6 高田豊行, 三田晃彦, 前野哲輝, 森脇和郎, 米川博通, 城石俊彦. 「マウス亜種間コンソミック系統群を用いたエネルギー代謝関連形質の遺伝解析」日本遺伝学会第79回大会 2007年9月19-21日, 岡山.
- 7 高田豊行 「B6.MSM コンソミック系統を用いた代謝関連表現型解析」国立遺伝学研究所研究会「多因子疾患の遺伝解析: 糖尿病・メタボリックシンドロームを中心に」2007年3月16-17日, 三島.
- 8 水多陽子, 春島嘉章, 倉田のり 「イネ亜種間交雑における生殖的隔離障壁遺伝子のポジショナルクローニング」2007年イネ分子遺伝学ワークショップ, 名古屋大学, 名古屋, 2007年7月3日.
- 9 春島嘉章, 栗木哲, 水多陽子, 藤澤洋徳, 倉田のり 「イネ F₂ 集団における生殖的隔離障壁の相互作用の検出」第79回日本遺伝学会, 岡山大学, 2007年9月21日, 岡山.
- 10 久保貴彦, 吉村淳, 倉田のり 「イネ雑種花粉で相互作用する2遺伝子座に起因する生殖的隔離」第79回日本遺伝学会, 岡山大学, 2007年9月20日, 岡山.
- 11 水多陽子, 春島嘉章, 倉田のり 「イネ雑種花粉で相互作用する2遺伝子座に起因する生殖的隔離」第79回日本遺伝学会, 岡山大学, 2007年9月20日, 岡山.
- 12 春島嘉章, 倉田のり 「栽培イネの第3染色体の雄性配偶体型生殖的隔離障壁と相互作用する第6染色体の雌性親遺伝子のポジショナルクローニング」日本育種学会112回講演会, 山形大学, 2007年9月22日, 鶴岡.
- 13 水多陽子, 春島嘉章, 倉田のり 「イネ雑種花粉で作用する生殖的隔離障壁遺伝子のポジショナルクローニング」日本育種学会112回講演会, 山形大学, 2007年9月23日, 鶴岡.
- 14 堀内陽子, 藤澤洋徳, 川喜田雅則, 望月孝子, 春島嘉章, 坂口隆之, 江口真透, 倉田のり 「Rice Genome Array を用いた SFP 検出手法の開発」日本育種学会112回講演会, 山形大学, 2007年9月23日, 鶴岡.
- 15 板橋悦子, 藤田雅丈, 倉田のり, 鳥山欽哉 「BT 型細胞質雄性不稔イネの花粉発達に関する核遺伝子の発現解析」日本育種学会112回講演会, 山形大学, 2007年9月23日, 鶴岡.
- 16 水多陽子, 春島嘉章, 倉田のり 「イネ雑種花粉で作用する生殖的隔離障壁の単離・機能解析」特定領域研究『植物ゲノム障壁』第1回若手の会, 茂庭荘, 2007年11月5日, 仙台.
- 17 堀内陽子, 藤澤洋徳, 川喜田雅則, 望月孝子, 春島嘉章, 坂口隆之, 江口真透, 倉田のり 「Affymetrix Rice Genome Array を用いた SFP 検出手法の開発」第30回日本分子生物学会年会・第80回日本生化学会大会 合同大会, パシフィコ横浜, 2007年12月13日, 横浜.
- 18 藤澤洋徳 「SNP を考慮した遺伝子発現データ解析」統計サマーセミナー, 2007年8月7日, 鹿児島.
- 19 藤澤洋徳, 堀内陽子, 春島嘉章, 江口真透, 望月孝子, 坂口隆之, 倉田のり 「SNP を考慮した遺伝子発現データ解析」統計関連学会連合大会, 2007年9月7日, 神戸
- 20 Fujisawa, H., Horiuchi, Y., Harushima, Y., Eguchi, S., Mochizuki, T., Sakaguchi, T., Kurata, N. A unified method for detecting single feature polymorphisms and gene expression level differences. Pasific Symposium on Biocomputing, 2008.1.6, Hawaii, USA
- 21 Nakahara, T. and Nakaya, A. ActScan: an Image Analysis Tool for Automatic Classification of Mouse Body Fat Japan-Taiwan Young Researchers Conference on Computational and Systems Biology. 2008, 3.9-11. National Tsing Hua University, Taiwan

- 22 藤澤洋徳, 堀内陽子, 春島嘉章, 江口真透, 望月孝子, 坂口隆之, 倉田のり「SNP を考慮に入れた遺伝子発現データ解析」シンポジウム「バイオインフォマティクスおよび経時観察データの解析」2008年2月8日, 広島.
- 23 Horiuch Y., Harushima Y., Mochizuki T., Fujisawa H., Eguchi., Kawakita M., Kurata N. Detection of nucleotide and expression polymorphisms between rice stains using affymetrix rice genome array. XX International congress of Genetics, Berlin, Germany, July 12-17, 2008.
- 24 Harushima Y., Kuriki S., Mizuta Y., Kurata N. Detection of pairs of interactive reproductive barriers in rice. XX International congress of Genetics, Berlin, Germany, July 12-17, 2008.
- 25 Mizuta Y., Harushima Y., Kurata N. Positional cloning of a pair of interactive genes causing reproductive barrier in the hybrid pollen of rice. XX International congress of Genetics, Berlin, Germany, July 12-17, 2008.
- 26 Harushima Y., Yano M., Kurata N. Identification of areproductive barrier working in the process of pollen competition in rice. XX International congress of Genetics, Berlin, Germany, July 12-17, 2008.
- 27 Tsuda K., Ito Y., Miyao A., Hirochika H., Kurata N. Identification and analysis of rice mutants misexpressing NKOX genes in leaves. XX International congress of Genetics, Berlin, Germany, July 12-17, 2008.
- 28 Takada, T., Ebata, T., Shin-I, T., Narita, T., Abe, K., Sakaki, Y., Toyoda, A., Obata, Y., Moriwaki, K., Kohara, Y., Shiroishi, T. Two wave intersubspecific introgression built up genome framework of the classical laboratory mouse strains. 22th International Mammalian Genome Conference, 2008. 11.2-5, Prague, Czech Republic.
- 29 Takada, T., Mita, A., Maeno, A., Moriwaki, K., Yonekawa, H., Shiroishi, T. Mouse inter-subspecific consomic strains uncovers additive and non-additive genetic effects on complex traits. 22th International Mammalian Genome Conference, 2008. 11.2-5, Prague, Czech Republic.
- 30 Takada, T., Mita, A., Moriwaki, K., Yonekawa, H., Shiroishi, T. Mouse inter-subspecific consomic strains for the study of energy metabolism-related traits. CBI Annual Meeting 2008 International Symposium, 2008. 10.22-24, Tokyo, Japan.
- 31 Oka, A., Takagi, N., Moriwaki, K., and Shiroishi, T. Genetic Study of the Reproductive Isolation between Two Mouse Subspecies, *Mus musculus domesticus* and *M. m. molossinus* XX International congress of Genetics, Berlin, Germany, July 12-17, 2008.
- 32 Tanaka, E. and Tamura, Y. “Skeleton-based Fourier Descriptor of Open Curves” APCOM2008, Kokushikan University, September 3, 2008
- 33 Takahashi, A., Sugimoto, H., Kimura, S., Tomihara, K., Tsuchiya, T., Kakihara, S., Tanemura, M., Shiroishi, T., Koide, T.: Complex genetic architecture of social interaction and aggressive behavior clarified using consomic strains derived from MSM and C57BL/6. *22nd International Mammalian Genome Conference*, Prague, November 2-5, 2008.
- 34 Takahashi, A., Tomihara, K, Shiroishi, T., Koide, T.: Genetic analysis of inter-male aggression using consomic mouse strains established from C57BL/6J and MSM. *10th Annual Meeting of the International Behavioural and Neural Genetics Society*, Portland, May 5-9, 2008.
- 35 Md. Nurul H.M., Eguch, S., Robust QTL Analysis by the Minimum β -Divergence Method. International Association for Statistical Computing 2008, Yokohama, 2008.12.6
- 36 Pritchard, M., Eguchi, S. Finding Optimal Gene Set for Classification from Multiple Predictive

Gene Sets. International Association for Statistical Computing, Yokohama, 2008.12.6

- 37 米田典央, 倉田のり, 野々村賢一「減数分裂第一分裂前期に特異的な染色体挙動の観察」 イネ遺伝学・分子生物学ワークショップ 2008, 九州大学, 2008年7月5日, 福岡.
- 38 水多陽子, 春島嘉章, 倉田のり「イネ亜種間交雑で生殖的隔離障壁となる重複遺伝子の解析」日本育種学会 114回講演会, 滋賀県立大学, 2008年10月11日, 彦根.
- 39 津田勝利, 伊藤幸博, 宮尾安藝雄, 廣近洋彦, 倉田のり「イネのシュート形成における極長鎖脂肪酸(LCFA)の機能解析」日本育種学会 114回講演会, 滋賀県立大学, 2008年10月11日, 彦根.
- 40 久保貴彦, 水多陽子, 新濱充, 春島嘉章, 倉田のり「イネ生殖隔離機構の解析から見えてくるもの」国立遺伝学研究所研究会「高等植物の生殖過程を制御する因子の多様性と生殖隔離」, 国立遺伝学研究所, 2008年11月5日, 三島.
- 41 望月孝子, 菊池俊介, 濱田和輝, 加藤大貴, 大木伸彦, 藤田雅丈, 堀内陽子, 倉田のり, 矢野健太郎「OryzeExpress: イネのゲノム・アノテーションとオミックス統合データベース」第31回日本分子生物学会年会、第81回日本生化学会大会 合同大会, 神戸ポートアイランド, 2008年12月10日, 神戸.
- 42 津田勝利, 伊藤幸博, 倉田のり「イネにおける KNOX 遺伝子を介した SAM の維持および葉の分化の研究」国立遺伝学研究所研究会「イネ分子遺伝学の新展開」, 国立遺伝学研究所, 2008年12月20日, 三島.
- 43 高田豊行, 三田晃彦, 森脇和郎, 米川博通, 城石俊彦: マウス亜種間コンソミック系統群を用いた肥満関連表現型の遺伝解析. 日本遺伝学会, 2008, 9.3-5, 名古屋.
- 44 高田豊行, 三田晃彦, 森脇和郎, 米川博通, 城石俊彦: マウス亜種間コンソミック系統群を用いたエネルギー代謝関連形質の遺伝解析. 第55回日本実験動物学会総会, 2008, 5.15-17, 仙台.
- 45 岡彩子, 高田幸, 古関明彦, 森脇和郎, 城石俊彦: マウス亜種間における生殖隔離と減数分裂期のチェックポイント機構との関連. 日本遺伝学会, 2008, 9.3-5, 名古屋.
- 46 田中英希, 田村義保「新しい開曲線記述子を用いた2次元部分形状の定量的評価」日本計算機統計学会第22回大会, 秋田文化会館, 平成20年5月23日
- 47 江口真透. バイオインフォマティクスにおける統計的課題について. 科研費研究集会「高次元データの統計解析」博多. 2008.11.21
- 48 小森理, 江口真透. 1クラスラベルに注目したブースティング. 統計関連学会連合大会, 横浜. 2008.9.8
- 49 プリチャード真理, 江口真透. マイクロアレイにおける遺伝子選択と判別能力の関係. 統計関連学会連合大会, 横浜. 2008.9.9
- 50 Kuriki, S., Harushima, Y., Fujisawa, H., and Kurata, N., Multiplicity Adjustments in Detecting Reproductive Barriers caused by Loci Interactions, BIRS Workshop 09w5040 Random Fields and Stochastic Geometry, 2009.2.26, Banff, Canada
- 51 高田豊行: 日本産野生マウス由来系統の遺伝的多様性を基盤とした機能ゲノム学. かずさ DNA 研究所ワークショップ 「遺伝的多様性に基づく有用遺伝子座の探索」, 2009, 3.25, 千葉.
- 52 高田豊行, 三田晃彦, 森脇和郎, 米川博通, 城石俊彦: マウス亜種間コンソミック系統群を用いた多因子肥満症の遺伝解析. 第56回日本実験動物学会総会, 2009, 5.14-16, 大宮.
- 53 梅森十三, 近藤亮太, 宇野毅明, 湯浅茂樹, 小出剛: 遺伝的不適合による神経発達異常第56回日本実験動物学会, 2009, 5.14-16日、大宮
- 54 Sugimoto, H., Takahashi, A., Shiroishi, T., Koide, T.: Molecular and genetic analysis for aggressive behavior using B6-MSM consomic mouse strains. 11th Annual Meeting of the International Behavioural and Neurogenetics Society, Dresden, June 4-8, 2009.
- 55 高田豊行, 江端俊伸, 小原雄治, 城石俊彦. NIG Mouse Functional Genomics Database. 文科省統合デ

- データベースプロジェクト「データベースが拓くこれからのライフサイエンス」2009, 6.12, 東京
- 56 Takada, T. Making a database to link mouse genome diversity and phenotype Symposium in International Phenome Integration Meeting RIKEN/InterPhenome/CASIMIR, 2009. 7.12-13, Shiga, Japan.
- 57 田邊彰, 石井亜矢子, 城石俊彦, 高橋阿貴, 小出剛: コンソミックマウス系統を用いた不安様行動の高解像度マッピング. 第32回日本神経科学大会, 2009, 9.16-18, 名古屋.
- 58 高田豊行, 三田晃彦, 森脇和郎, 米川博通, 城石俊彦: マウス亜種間コンソミック系統を用いたエネルギー代謝多様性の遺伝解析. 日本遺伝学会, 2009, 9.16-18, 松本.
- 59 石井亜矢子, 西明紀, 城石俊彦, 小出剛: B6-MSM コンソミックマウス系統を用いた自発活動性の遺伝学的解析. 第32回日本神経科学大会, 2009, 9.16-18, 名古屋.
- 60 梅森十三, 宇野毅明, 湯浅茂樹, 小出剛: 遺伝的不適合マウスは脊髄オリゴデンドロサイトの成熟が異常である. 第32回日本神経科学大会 2009, 9.16-18日, 名古屋.
- 61 杉本大樹, 高橋阿貴, 城石俊彦, 小出剛: マウス攻撃行動における遺伝的基盤の解析. 第32回日本神経科学大会, 2009, 9.16-18, 名古屋.
- 62 Umemori, J., Uno, T., Yuasa, S., Koide, T.: Decreased expression of myelination associated genes in genetic incompatibility mice, Genic mice. 23rd International Mammalian Genome Conference, La Jolla Torrey Pines, California, November 1-4, 2009.
- 63 Takada T., Mita A., Ebata T., Shin-I T., Toyoda A., Fujiyama A., Kohara Y., Obata Y., Moriwaki K., Yonekawa H., Shiroishi T. Genetic dissection of age-associated phenotypes using intersubspecific consomic strains. 23th International Mammalian Genome Conference, 2009. 11.1-4, La Jolla, USA.
- 64 梅森十三, 湯浅茂樹, 宇野毅明, 小出剛: 遺伝的不適合マウスではレトロトランスポゾン IAP の発現が上昇している. 第32回日本分子生物学会年会, 2009, 12.9-12. 横浜.
- 65 杉本大樹, 菊水健史, 城石俊彦, 小出剛: Different patterns of ultrasonic vocalizations of wild-derived mouse strains from laboratory strains. 第32回日本分子生物学会年会, 2009, 12.9-12, 横浜.
- 66 田邊彰, 石井亜矢子, 城石俊彦, 高橋阿貴, 小出剛: 野生マウスに特徴的な不安様行動の遺伝学的解析. 第32回日本分子生物学会年会, 2009, 12.9-12, 横浜.
- 67 Fujisawa, H., Horiuchi, Y., Harushima, Y., Takada, T., Eguchi, S., Mochizuki, T., Sakaguchi, T., Kawakita, M., Shiroishi, T., Kurata, N., Statistical Analysis of Affymetrix GeneChip Data, TRIC Symposium "Statistical and Systems Genetics", Mishima, Japan.2009.10.22
- 68 Harushima, Y., Kuriki, S., Fujisawa, H., Kurata, N., Detection of interactive pairs of reproductive barriers using F2 population, TRIC Symposium "Statistical and Systems Genetics", Mishima, Japan.2009.10.23
- 69 Kurata, N., Wild Rice Resource Project in Japan Its Perspectives. 6th International Rice Genetics symposium, Manila, Philippine, 2009.11.17
- 70 Win, KT., Kubo, T., Miyazaki, Y., Doi, K., Yamagata, Y., Yasui, H., Yoshimura, A., Molecular mapping of two loci conferring F1 pollen sterility in inter- and intraspecific crosses of rice. 6th International Rice Genetics symposium, Manila, Philippine, 2009.11.17
- 71 Mizuta, Y., Harushima, Y., Kurata, N. Analysis of a pair of genes, *DOPPELGÄNGER (DPL)1* and *DPL2* responsible for reproductive isolation between two rice subspecies. International Symposium of Cell-Cell Communication in Plant Reproduction, Nara, Japan.2010.3.13

<著書等>

高田豊行, 城石俊彦 「マウスの比較ゲノム, 日本産マウス系統のゲノム解読とコンソミック系統の樹立」細胞工学別冊 細胞工学別冊, 比較ゲノム学から読み解く生命システム pp. 97-104 藤山秋佐夫 監修 秀潤社, 東京 (2007)

Koide T, Catanesi CI, Nishi A, Shiroishi T, Kasai S, Ikeda K, Takahashi A. Advantage of using wild-derived mouse strains for a variety of pain-related studies: Genetic diversity and new genetic tools. In: *Acute Pain* (Columbus F, ed), Nova Science Publishers, New York, in press.

Eguchi, S. Information Divergence Geometry and the Application to Statistical Machine Learning. Eds. Frank Emmert-Streib and Matthias Dehmer 'Information Theory and Statistical Learning', Springer, New York, 10.1007/978-0-387-84816-7_13

倉田のり, 野々村賢一 バイオリソース&データベース活用術(細胞工学別冊)、秀潤社 190-192(2009)

倉田のり, 野々村賢一(2009) バイオリソース&データベース活用術(細胞工学別冊)、秀潤社 190-192

<受賞>

岡彩子 日本遺伝学会第80回大会 Best Paper 賞, 2008年9月3-5日,名古屋.

水多陽子, 春島嘉章, 倉田のり 日本遺伝学会第81回大会 Best Paper 賞, 2009年9月16日, 長野.

水多陽子, 春島嘉章, 倉田のり (2010) 第117回講演会日本育種学会優秀発表賞, 2010年3月26日, 京都.

濱田和輝, 山本直樹, 望月孝子, 諏訪部圭太, 倉田のり, 矢野健太郎 (2010) 第117回講演会日本育種学会優秀発表賞, 2010年3月26日, 京都.

サブテーマ2:「生物多様性データの統計的モデリングと解析システムの開発」

研究代表者 栗木哲

1. 研究目標

現代遺伝学研究の目的の一つは、生物個体の多様性をゲノム多型に結びつけて理解することである。その目的のためには、生物の多様性とゲノム多型の双方を計量化し、それらから必要な情報を抽出し、連関を探ることが必要となる。ところで遺伝現象はそれ自体確率的な事象であり、そのためそのモデル化において統計モデルを用いることはきわめて有効である。このことは遺伝学と統計学はそれらの初期において未分化であったことから伺い知ることができる。そのような背景を踏まえて、本サブテーマでは遺伝学研究所で得られたデータを解析するための統計的手法の研究を行った。データ解析の結果として遺伝学の分野で重要な知見を得るとともに、開発した統計解析の方法論をより広い分野へ適用するための普遍化を行うことが本サブテーマの研究目標である。

この研究目標の達成のために、本サブテーマは次の5つの課題を設定した。

課題 a. フーリエ記述子を用いた形態多様性データの統計的モデリング

課題 b. マウス社会行動の統計的モデリング

課題 c. 多変量時系列モデルによるマウス自発行動の特徴抽出

課題 d. 生殖隔離障壁に関わるエピスタシスの統計学的検出と多重性調整

課題 e. QTL 解析法の最適化

2. 年度研究計画 [2] の研究計画に記載した。

3. 研究経費の推移

平成 17 年度実績： 157,820 千円 (プロジェクト全体予算)

平成 18 年度実績： 18,800 千円

平成 19 年度実績： 24,560 千円

平成 20 年度実績： 25,000 千円

平成 21 年度実績： 20,820 千円

4. 研究実施体制 [3] 研究推進・実施体制 に記載した。

5. 研究成果

課題 a. フーリエ記述子を用いた形態多様性データの統計的モデリング (その 1) (田村、田中、

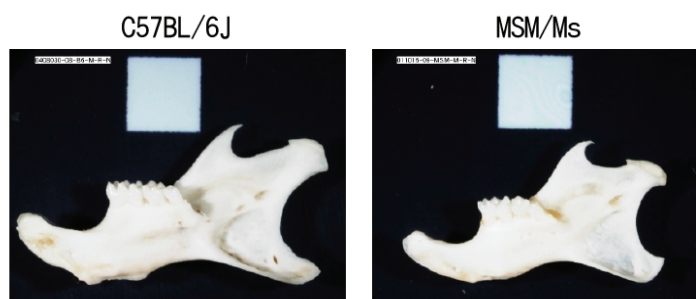


図 1 . MSM/Ms 系統と C57BL/6J 系統の下顎形態

遺伝的に異なるマウス近交系統である MSM/Ms 系統 (MSM 系統、*Mus musculus mollosinus*) および C57BL/6J 系統 (B6 系統、*Mus musculus domesticus*) (図 1 参照)、これら二系統間で F2 世代を作成し、その下顎骨の形態差を指標とした Quantitative Trait Loci (QTL) 解析を試みた。F2 世代の下顎骨は、MSM、B6 両系統の特徴を併せ持つ多様性に富んだ形態を示しており、これらを定量的に評価するには従来の計測手法では不十分な可能性がある。

そこで、形態的特徴の差をより詳細に比較するための解析手法として P 型フーリエ記述子の適用を試みた。P 型フーリエ記述子とは、曲線を長さの等しい線分からなる折れ線として近似し、隣接する線分間の角度変化 (偏角) を求め、その累積で定義される全曲率関数を基にしている。さらに、図形の回転、拡大、縮小に対して不変であるという性質を持つ。この手法では輪郭の部分形状 (開曲線) を定量的に評価することが可能であることから、下顎骨を筋突起、下顎突起などの部位ごとに切り分けて解析することが可能となる。実際に MSM、B6 両系統の筋突起について P 型フーリエ記述子を適用し、主成分分析を行った結果を図 2 に示す。解析結果をみると第一主成分において二系統間の違いが明確に表れている。

なお、P 型フーリエ記述子を用いた研究手法の普及のために、ホームページを立ち上げた (URL <http://www.k4.dion.ne.jp/%7Ejudgepic/>)。本ホームページでは、植物の器官の 2 次元における形状を定量的に数値化するために開発されたソフトウェア JudgePic を走らせることができる。実際に運用するとき、細胞、骨などの 2 次元デジタル画像であれば、理論的にすべて形状を定量化することが可能である。JudgePic が利用する原理は、フーリエ (P-type, G-type) 変換である。任意の 2 次元の形状の輪郭線を抽出し、指定された次元のフーリエ記述子に自動変換される。特徴は開曲線 (P-type Fourier の

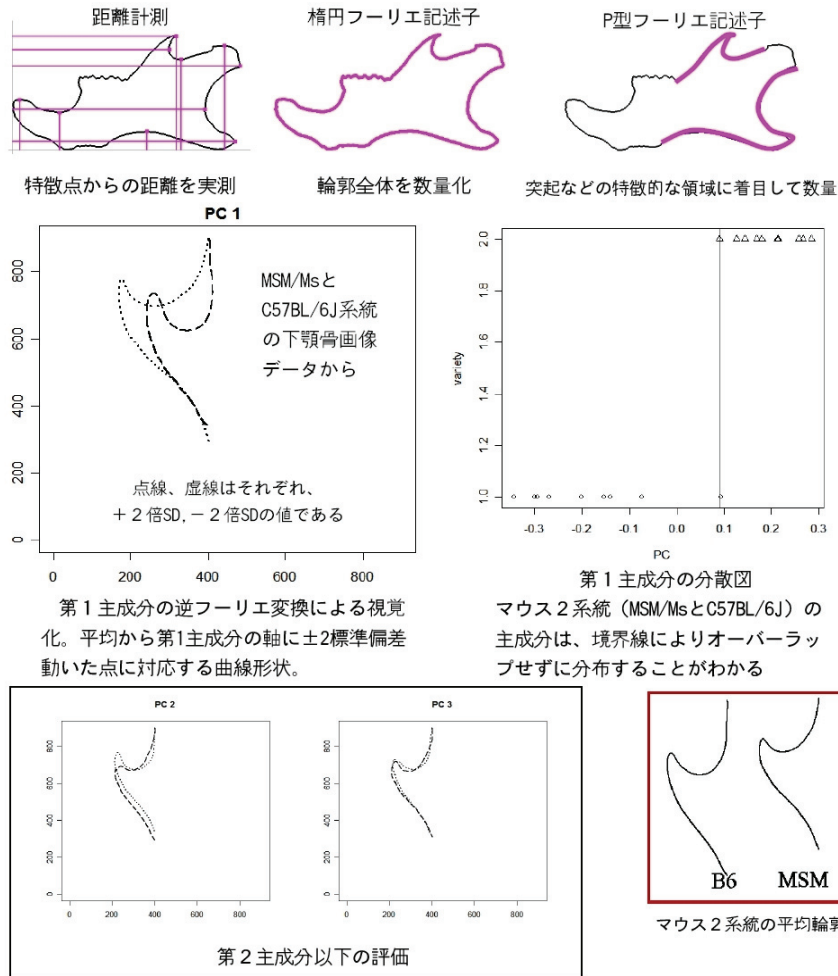


図 2 . P 型フーリエ記述子を用いた形態計測と解析結果

み)にも有効であることである。そのために、形状だけではなく曲線の定量分析にも使える。R 言語で開発された `rfourier.r` と一緒に使えば、形状の変化を観察することもできる。

80 年代前半に考案された P 型フーリエ記述子は、これまで手書き文字認識における特徴抽出や 2 次元部分形状の統計解析に用いられ一定の有用性が示されていたため、平成 17 年度は、マウス下顎骨の部分形状の解析に、この P 型フーリエ記述子を用いた。しかし、P 型フーリエ記述子は、位置関数 (xy 座標) の差分系列のフーリエ係数であり、この差分系列は、一般に連続な周期関数ではない。不連続な関数のフーリエ変換では、ギブス現象のため、曲線形状の全情報を低次元に縮約することができず、特徴抽出がうまくいかないことがある。そこで、この問題を克服するための、新たな曲線形状の記述方法として、Protrusion フーリエ記述子を考案した。この新たな手法では、まず、与えられた曲線 (図 3(a)) に対し、スケルトン-垂線グラフ (図 3(b)) を求め、このグラフを利用して、曲線上の各位置における、左右からの突き出し度合い (protrusion, 図 3(c)) を算出する。そして、始点から終点までの、この突き出し度合いの系列のフーリエ係数として、Protrusion フーリエ記述子を定義する。Protrusion フーリエ記述子は、並進、回転、スケールングに対し不変である。また、曲線の方向が急に変わる部分において、位置の差分系列では不連続が生じるが、protrusion の系列では連続になるので、このような形状を含む部分形状解析では、P 型フーリエ記述子に比べ、新手法がやや優れていることが分かった。

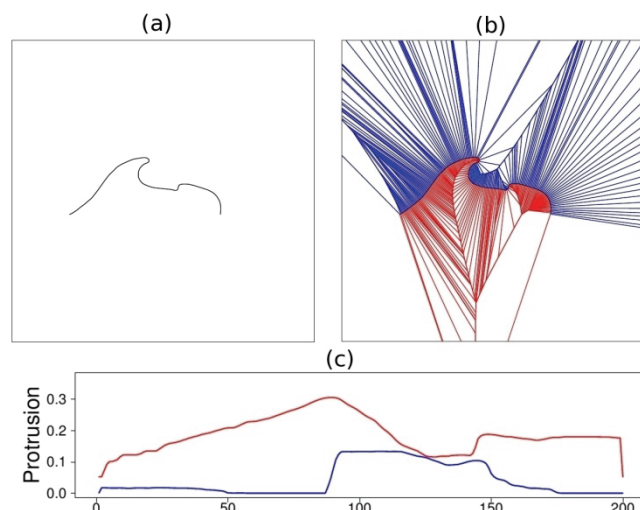


図3. (a) 与えられた開曲線、(b) その曲線に対するスケルトン-垂線グラフ、(c) 曲線の左右からの突出度合い (protrusion) の系列。左と右からの突出の系列はそれぞれ青と赤のグラフで示されている。

しかし、Protrusion の系列も、位置の差分系列と同様、始点と終点の 2 端点間における不連続の問題を解決していないことや、記述子から曲線形状を再構成する手法が複雑である点から、開曲線記述子の決定版とはなり得なかった。形の統計解析において、注目する形状変動の成分を視覚化して解釈するために、どうしても、記述子からの曲線形状の再構成の計算は簡単であることが望ましい。そこで、また別の開曲線記述子、速度フーリエ記述子 (VFD) を考案した。VFD は、P 型フーリエ記述子の自然な改良版である。並進、スケーリングに対して不変であり、回転に対する規格化が容易である。また、記述子からの曲線形状の再構成は、P 型フーリエ記述子と同様、逆フーリエ変換すればよいだけという簡単なものである。VFD 法では、まず、与えられた曲線に対し、その曲線に沿った始点から終点に向かう最小エネルギー運動 (図 4(a)) を求める。そして、その運動における速度関数 (位置の差分の系列、図 4(b)) のフーリエ係数として、VFD を定義する。速度関数は、どのような形状の曲線に対しても、原理的に連続な周期関数であり、VFD 法は、曲線形状の全情報を低次元に縮約する能力に優れている (図 4(c))。

そしてさらに、P 型フーリエ記述子を用いて行ったときと同様のマウス下顎骨の部分形状解析を、VFD を用いて行った。MSM/Ms と C57BL/6J の 2 系統を起源とする F2 世代の 207 個体 (図 5(a)) のマウス下顎骨から、特定の部分形状を表す開曲線を切り出し、VFD で特徴抽出を行った。また、全てのマウスについて、染色体上の特定の 122 箇所において遺伝型を調べた (図 5(b))。次に、VFD で表された形のデータに対し主成分分析を行い、全形状変動を独立な成分に分離した。最後に、染色体上の各位置どうしの交合効果を調べるため、染色体上の位置の各ペアについて、そのペアの遺伝型でグループ分けしながら、分散分析を用いて、そのペアの遺伝型によるグループ分けの、形に対する影響力の度合いを調べた (図 5(c))。同じ部位においても、異なる形状変動に対し、それぞれ全く異なる遺伝子のペアが影響力を及ぼしている様子が分かった。

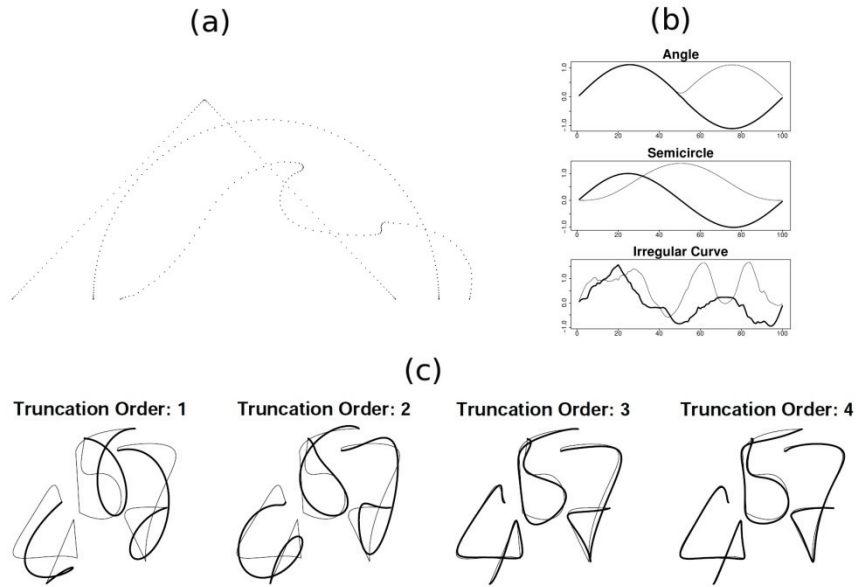


図4. (a) 3つの開曲線。左から Angle, Semicircle, Irregular curve. 点は曲線に沿った最小エネルギー運動における等時間間隔の位置を示す。(b) 3曲線に対する速度関数、細線は実数、太線は虚数成分を示す。速度関数は連続な周期関数になる。(c) 任意次数で打ちきられたVFDの部分集合から再構成された曲線。細線は元の曲線。

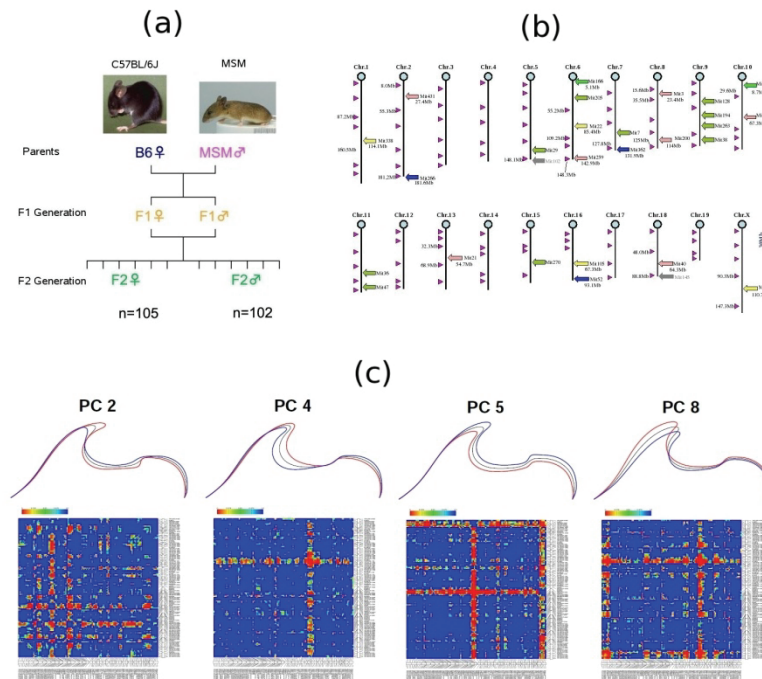


図5. (a) 実験に用いた F2 世代のマウスの起源とする系統。(b) 遺伝型データ。遺伝型をしらべた染色体上の位置。(c) 各主成分の形の変動の視覚化(上)とその形の変動が持つ、染色体上の位置の各ペアについての有意性のマップ(下)。

課題 a. フーリエ記述子を用いた形態多様性データの統計的モデリング (その 2)

生物の形態を決定する遺伝子を発見するために高次元の形態データと離散値の遺伝子型データの関連を探る問題に統計的機械学習の方法論を開発した。このためマウスの形態の解析のプロジェクトに参加し B6 マウスと MSM マウスの下顎形態の特徴を決定している因子を探索することを目指した。特に田村グループが行った P 型フーリエ子の特徴量としてパタン認識をアダブーストで行い従来の主成分による方法との比較を行った。田村グループで開発されたマウス下顎の形態計測のデータもとづいて B6 と MSM の系統による特徴抽出のための学習アルゴリズムを開発した。これによってシャープな特徴量が発見された。この抽出された変量から成るスコアを量的形質として捉えてと城石グループによって獲得された遺伝子形データに対して QTL 解析を行った。アダブーストを含む幾つかの方法についてもロバストネスなどの統計的検討を加えた。従来は形態計測の多次元データを主成分分析によって得た主成分を量的形質として QTL 解析が行われていた。提案したアダブースト方法では従来法の主成分分析で得られない幾つかの利点が明らかにされている。例えば主成分分析では形態計測のデータの全ての成分の線型結合によって主成分に還元されるが、実際のどの成分が効いたかはよく分からないことが多い。多くの場合は線型係数の絶対値を見るだけでは良く分からない。アダブースト法は形態計測のデータの中から少数の成分を選択するので形態のどの特徴が効いたかは直接に理解することが可能である。下図 A,B,C,D は主成分スコアとアダブーストスコアを量的形質とした QTL 解析の結果である。図 A,C,D において 13 番染色体に特異に関連する遺伝子形が示唆されるが図 B ではこの傾向が現われてない。これはアダブーストスコアによる特徴抽出の方が非常に少数の特徴量しか使っていないにも関わらず性能が高いと考えられる。

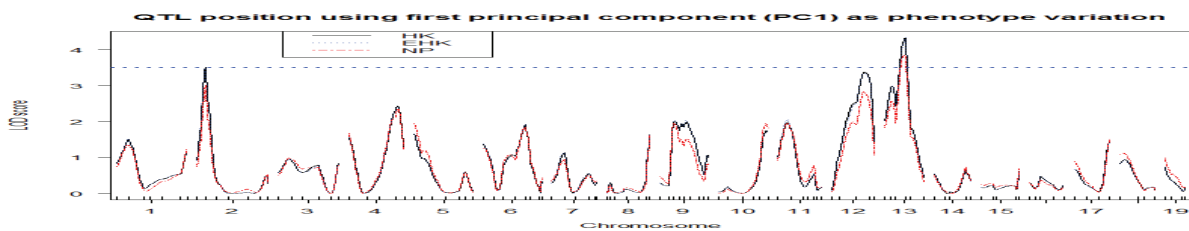


図 A: 第 1 主成分スコアによるロッドスコア(オス・メス)

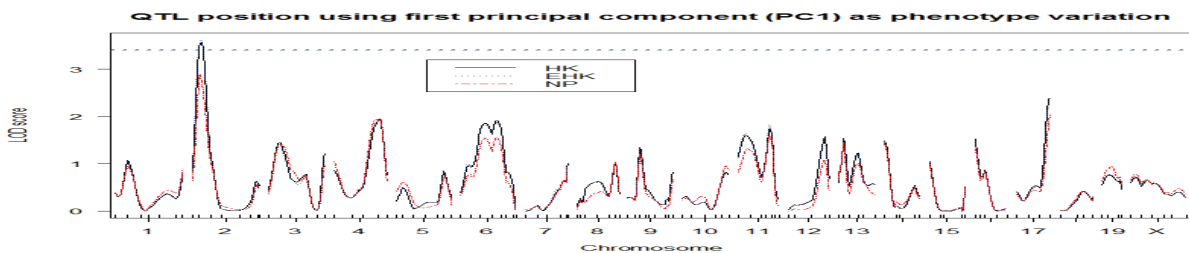


図 B: 第 1 主成分スコアによるによるロッドスコア (メス)

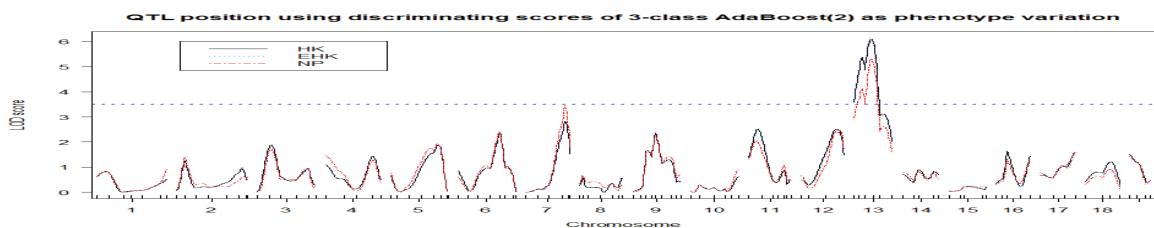


図 C: アダブーストスコアによるロッドスコア(オス・メス)

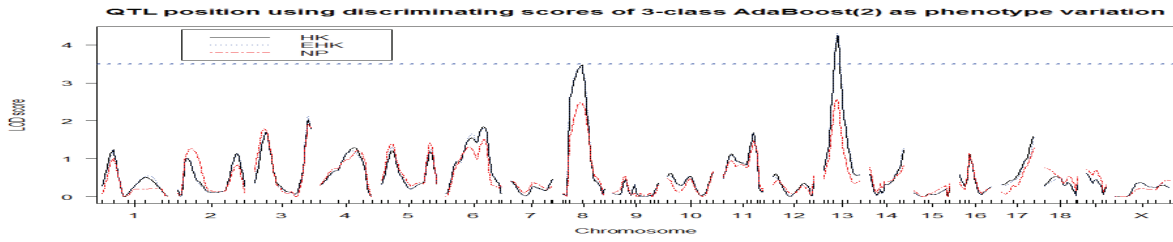
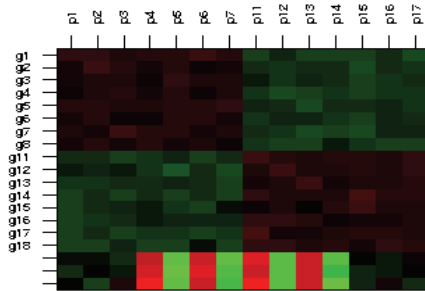


図 D: アダブーストスコア-ロッドスコア(メス)

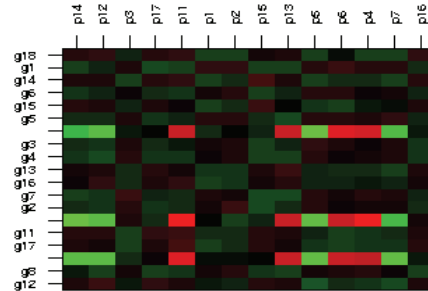
QTL 解析を統計的性能について研究した。特に量的形質データに外れ値がある場合にはロバストな方法が望まれる。従来法は最尤法に基づいてやられているのでロバストでないことが報告されている。江口・モーラのグループは最小ベータ・ダイバージェンスに基づく方法を考え、提案を検討している。いろいろな数値実験を加えてそのアルゴリズムの安定性について検証している。

次にクラスタリングについて研究した。幾つかあるクラスタリングの手法の中で相関係数や分散分析に基づく階層クラスタリングはクラスター数が未知であるとき広く用いられている。データ次元が高いときも比較的安定した解析が得られる。しかしながら、原理的にはモーメントを使うので外れ値に対して脆弱性が指摘されている。この研究は、その弱点を改良する方法を最小ベータ・ダイバージェンスに基づいて行った。既に QTL 解析でも分かったように外れ値の混入に対して良好な性能を持つことが確認された。下図は外れ値を加えた場合の従来法と提案法の比較を行ったものである。

(a) Original Clusters (S) in Presence of Outliers

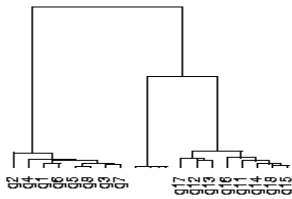


(b) Random Allocation (X*)

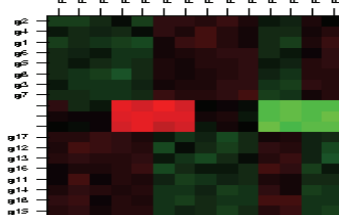


(A) 外れ値の混入

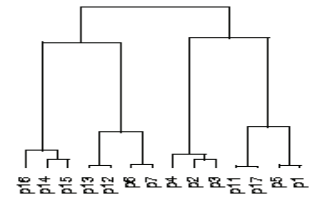
(c) Row-wise Dendrogram by HC



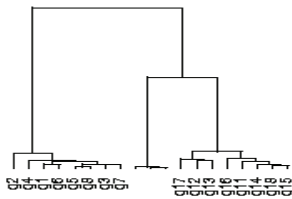
(d) Hierarchical Clustering (HC)



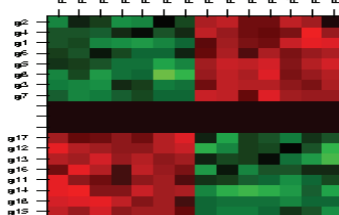
(e) Column-wise Dendrogram by HC



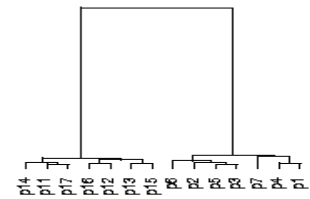
(f) Row-wise Dendrogram by RHC



(g) Robust Hierarchical Clustering (RHC)



(h) Column-wise Dendrogram by RHC



(B) 従来法と提案法の比較

この研究で分かったことはこのロバスト性だけでなく、最近、研究されている相補的クラスタリングについても密接な関連があることが分かった。これについては今後、研究を深めて実用化につなげたい。

課題 b. マウス社会行動の統計的モデリング

(H.17) マウスの社会行動を解析するために、オープンフィールド内に複数個体を入れて自由に行動させた際の行動をトラッキングし、その時系列データから動きの社会性を判定する手法の開発を進めた。マウス 2 個体の動きを方向ベクトルに細かく分け、その方向ベクトルに関する尤度推定を行うことでマウスの社会性をある程度数値化する事に成功している。また、2 個体間の距離の分布に基づく社会行動の定量的評価法の開発に向けて、ヒストグラムによる予備的解析を行った。

(H.18) 社会的動物であるマウスにとって、同種他個体を認識し、コミュニケーションを行うことは非常に重要である。この社会性には個体差があり、遺伝的寄与が示唆されているが、その遺伝的基盤については解明が進んでいない。その原因は、動物が示す社会性を定量化する効果的な手法が確立されていないことが大きな原因の一つであると考えられる。そこで、本研究では、マウス 2 個体が示す社会行動を定量化する手法の確立を目指した研究を行った。

テストに用いる 2 個体は同リターの同性の個体であり、9~10 週齢時に個別飼育し、テストは個別飼育から約 10 日後に行った。テストでは、2 個体のマウスをオープンフィールド (60×60×40 cm) に入れ、10 分間自由に探索させ、2 個体の行動をビデオで記録した (右図参照)。その後、コンピュータ上で Image SI (O'hara, Co., Ltd) を用いて画像解析を行い、2 個体の接触時間、接触回数、3 count/1 sec での各個体の位置情報を算出した。



Social Interaction Test

まず、一般的実験用系統である C57BL/6J (B6) と、野生の行動特性を多く残した野生由来系統である MSM について、社会行動の違いについて検討を行った。社会的接触の長さについては、MSM が雌雄ともに顕著に長い社会的接触を示すことが明らかとなり、それは B6 と比較して 3 倍近く長かった。一方、接触回数には B6 と MSM の間に差がなかったことから、MSM は 1 度接触すると長時間一緒にいる傾向にあることが分かった。

(H.19) マウス画像データ (同一の遺伝的特性を有する 2 匹のマウスを正方形のケージに入れて、10 分間行動を観察したもの) を観察者が目視観察し、2 個体の状態を (無関心、追跡、匂いかぎ、攻撃、攻撃後の追跡) の 5 つの状態のいずれかとして判断し、記録するシステムを開発した (図 1)。このシステムは、画面と 2 個体の距離等の特徴量が同期して表示され、再生速度を調節したり、巻き戻しなども

できるものである。観察者が記録した状態は、状態ごとに色分けされ、特徴量の時系列グラフ上に重ねて表示される。このシステムを用いて、昨年より解析している 54 個の画像データについて、目視により、状態を識別した時系列データを得た。

この状態の時系列に対して、マルコフモデルをあてはめ、遷移行列を推定した。各画像の遷移行列の要素のうち、頻度が高い状態 (無関心、追跡、匂いかぎ) にかかわる 9 個の要素を説明変数とし、各画像データを目視観察して “総合的な印象から人間が採点した社会性の度合い (8 人の被験者の平均)” を被説明変数とし

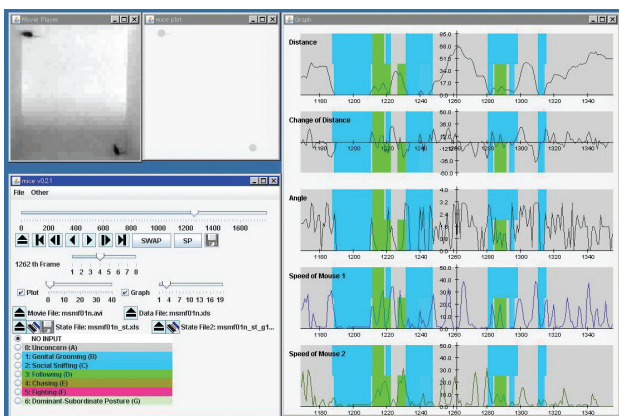


図 1

て線形回帰を行ったところ、昨年度よりもずっと良い結果が得られた（図2）。（昨年度は、(i) 2匹のマウスの距離分布から測定した B6（社会性が低いとされる典型例）と MSM（社会性が高いとされる典型例）の混合比率、そして(ii) 2匹のマウスの速度の成す角度が 0 度に近い状態の比率、(iii) 2匹のマウスの速度の成す角度が 180 度に近い状態の比率、の 3 つの特徴量を用いた）参考までに、昨年度の結果を図3に示す。

図2は、マウス2個体の状態を(無関心、追跡、匂いかぎ、攻撃、攻撃後の追跡)に切り分け、得られた時系列からマルコフ行列を推定し、その線形結合より社会性の尺度を構成することで、妥当性を持って行動の社会性を定量化できることを示唆している。このようにして定量化された社会性と遺伝的形質の関係を明らかにすることが、今後の研究課題である。現在は目視観察で状態の識別を行っているが、この識別を隠れマルコフモデルで自動的に行うことも重要な研究課題である。

また、2次元 unit vector chain を用いた解析では、マウス2個体の動きの時系列に対応して2個体間の行動に角度を割り当てる方法にさらに改良を加えた。また、位置座標の時系列データから、マウス2個体の動きを各時刻ごとの系列パターンとする画像を作成した。それによって、任意の時刻における2個体の動きとその時刻に割り当てられた角度と容易に対応させることが可能になった。その結果、今年度に改良を加えた角度割り当ての方法によって、マウスの行動によく呼応した角度割り当てができていたことが確認できた。今年度に開発した方法を用いて B6 と MSM のそれぞれ雄・雌ごとの2個体の社会行動についてデータ解析を行った。図4は、方向相互作用パラメータの推定値のボックス・プロット（今の場合、それぞれ10個の推定値の分布状況を表す）である。図の左側から順に B6 の雌、雄、MSM

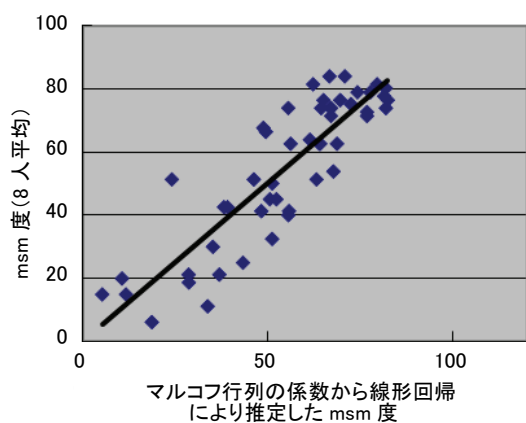


図2

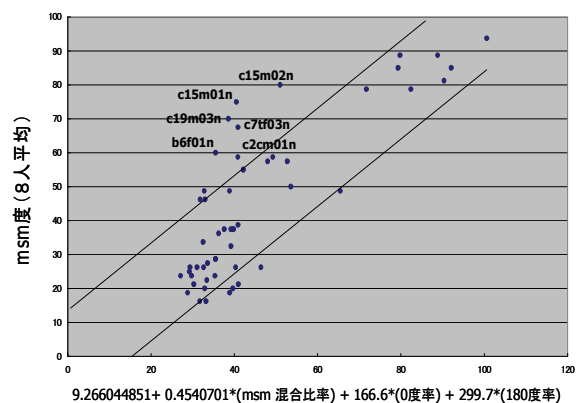


図3

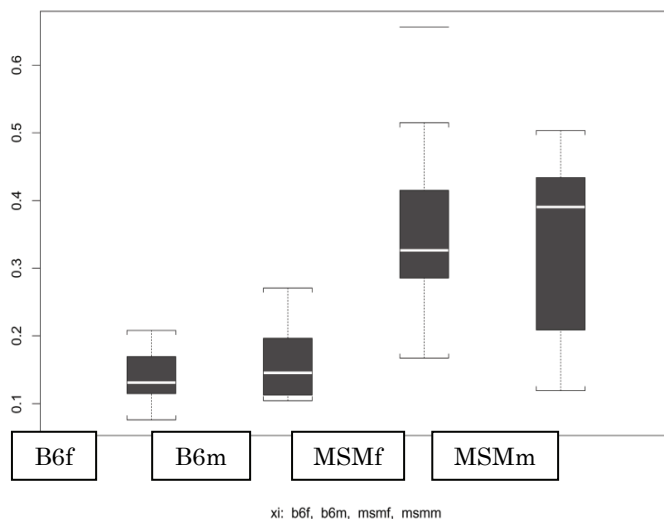


図4

の雌、雄の結果を示してある。

この図から、すべての場合においてパラメータの推定値は正であることから、観察された 10 分間を通した 2 個体の行動はいずれの場合も誘因的行動をとっていると見なされるが、B6 と MSM の間の社会行動に顕著な違いが現れていることが見て取れる。また、性別による行動の違いも見る事ができる。

以上から、このような解析の有用性が示されたが、今後は各時刻ごとのパラメータ推定を行うこと等によって、より正確な分析法が確立できるものと考えられる。

(H.20) 複数の動物個体が示す社会行動は、その一般的な定量化が難しく、解析が著しく困難である。本研究では、2 個体のマウスを同時にオープンフィールドに置いた際の 10 分間の社会行動の画像データから、自動トラッキングにより、2 個体の時系列位置座標データを抽出し、それらのデータに基づいて 2 個体の行動を統計数理モデルで特徴づける解析を進めた。20 年度は新たに 50 ペアについて、画像の目視観察法により行動項目分析時系列データを取得した。また、マウス 24 系統、461 ペアのオープンフィールド社会行動について、動画データとその時系列移動座標データを抽出し、基礎解析を行った。

このデータを用いて、マルコフ係数によるマウスの社会性の定量化についてより詳細に検討を進めた。

これは、2 個体のマウスの画像時系列を 5 つの状態に分解してその状態時系列に基づくマルコフ行列を構成し、人間による目視観察による各画像の社会性の得点 (8 人の観察者の得点の平均; より詳しくは B6 と MSM の画像を見たのちに、コンソミックマウスによる画像が B6 に近い MSM に近いかを得点化したもの (B6 と MSM をその行動パターンより社会性が低いタイプと高いタイプの象徴例と見なしている)) から社会性を定量化しようというものである。クロスバリデーションによって推定の安定性を確認した。また、AIC により、要因としてとりあげるマルコフ係数の個数を減らすことも試みた。目視による社会性得点を横軸、マルコフ行列による社会性推定得点を縦軸として描いたものが図 1 である。

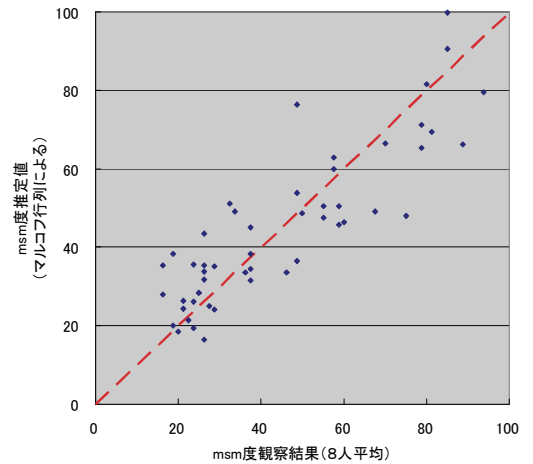


図 1

また、図 2 に各画像データについて目視得点の箱ひげ図とマルコフ行列による得点推定値の関係を示す。これらの結果より概ねマルコフ行列による推定値が一定の妥当性を有すると考えている。特に、図 1 から、概ね推定値の方が 60 点以上のものであれば、社会性が高いと判定できるのではないかと考えられる。MSM 度が高いデータについて、目視観察による結果よりも推定値が低くなる傾向があるが、この点は今後モデルを改良することで改善したい。来年度は各状態の自動推定を隠れマルコフモデルにより行い、411 ペアの画像データを自動解析する。

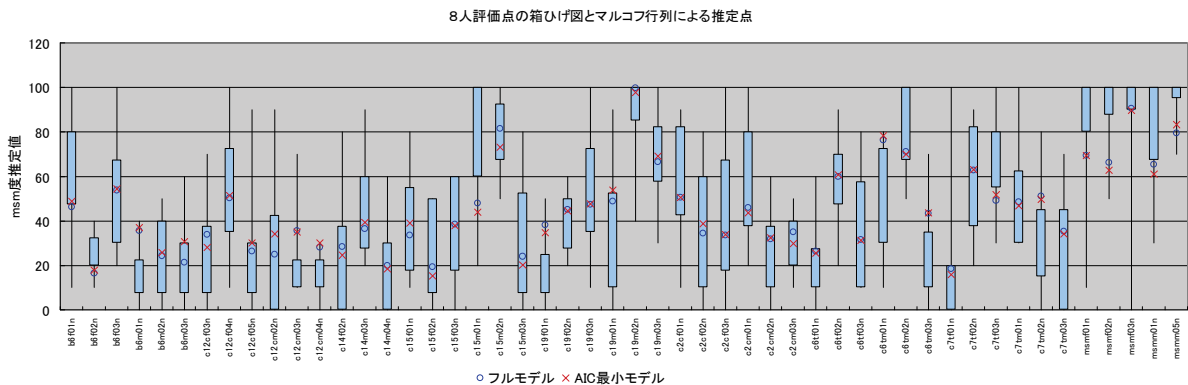


図 2

また、2次元 unit vector chain による解析では、マウス 2 個体の行動の時系列に対応して 2 個体間の角度を割り当てるアルゴリズムを一層改良した。図 4(a)、(b)は改良したアルゴリズムによって割り当てられた角度の時系列の例である。図 4(a)および(b)はそれぞれ B6 雌および MSM 雄に対応する。

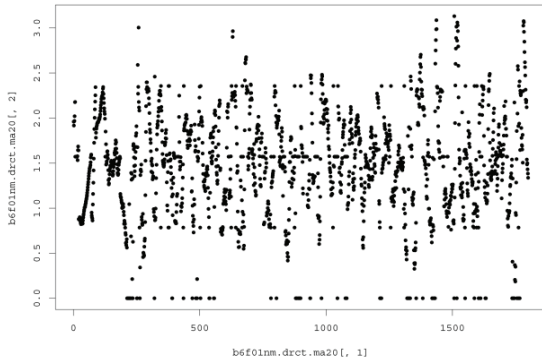


図 4(a)

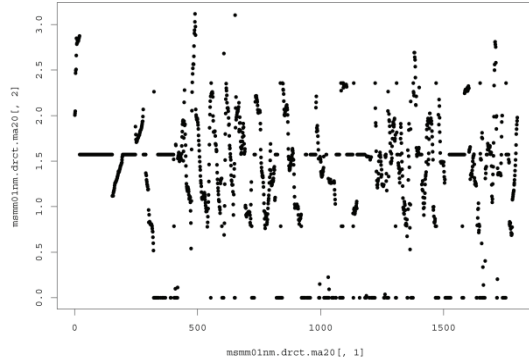


図 4(b)

得られた角度時系列から B6 および MSM の雄・雌ごとに方向相互作用パラメータの推定値を求めることによって、B6 と MSM の間の社会行動に明確な違いが現れること、また、性別による行動の違いも定量化できることが分かった。今年度新しく整理されたマウス 24 系統 461 ペアに対する社会行動トラッキングデータについても予備的解析を行った。

(H.21) 最終年度として、これまでに進めてきた社会行動の統計的モデリングに関して、解析方法の確立とデータの最終的な解析を進めた。

マウス社会行動に関するマルコフモデルによる解析では、50 個体ペアの目視観測データを学習データとして隠れマルコフモデル (HMM) の学習を行い、各時点で「社会的行動を行っている状態にあるか否か」が自動判別可能となった。実際、学習データについては図 1 のように、社会行動比率の目視観測値と HMM による推定値の間に明確な直線関係が得られた。そこで、(学習データも含む) 530 個の全個体ペアデータについて社会行動の比率を算出したところ、図 2 のような結果が得られた。コンソミック系統ごとに顕著な違いが見られる。また、HMM で高い社会行動比率を示した個体ペアデータについては、実際に画像で確認してみると、そのような傾向がみられ、推定結果の妥当性が裏付けられた。

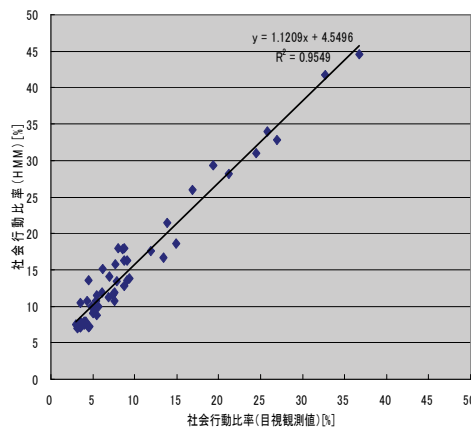


図 1. 目視観察と HMM による社会行動比率比較

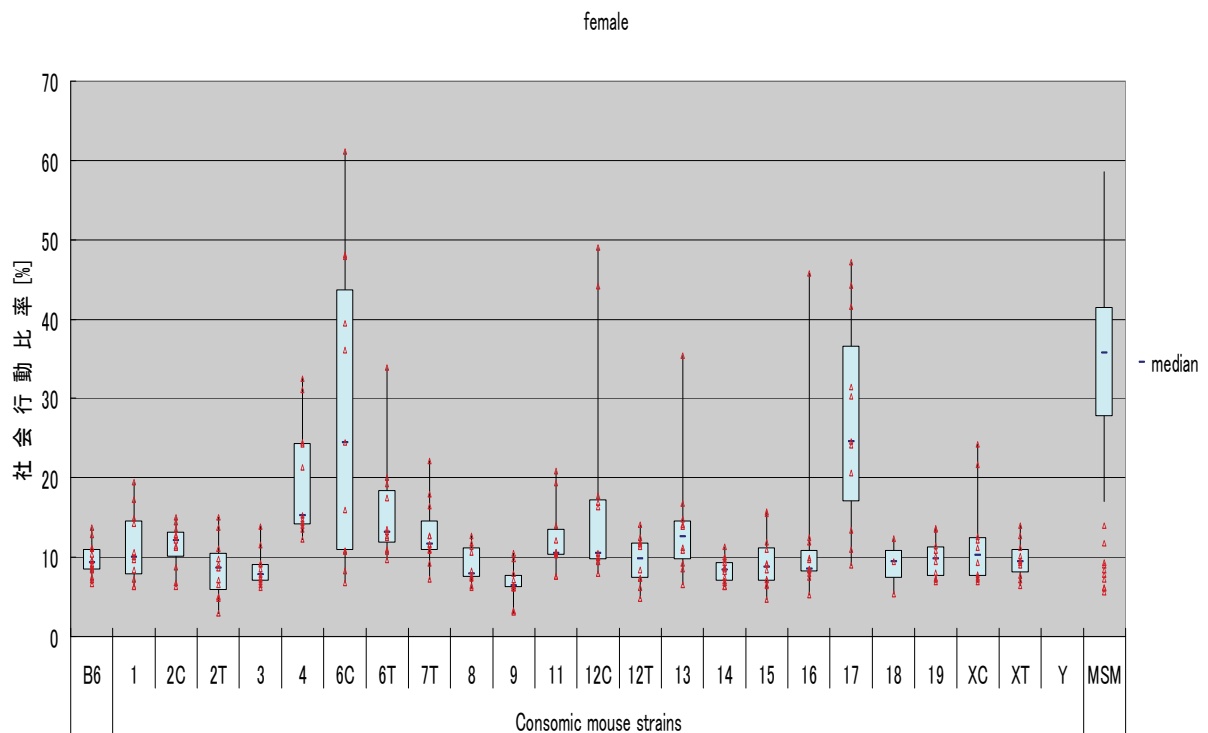
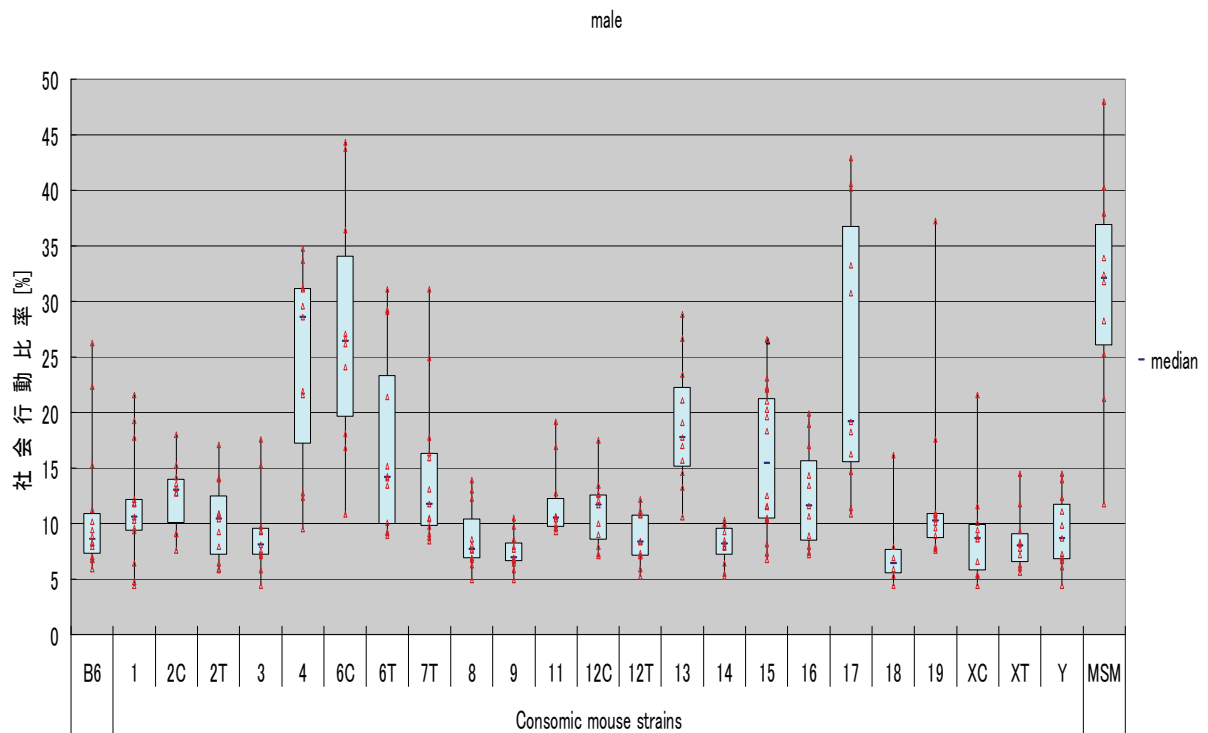
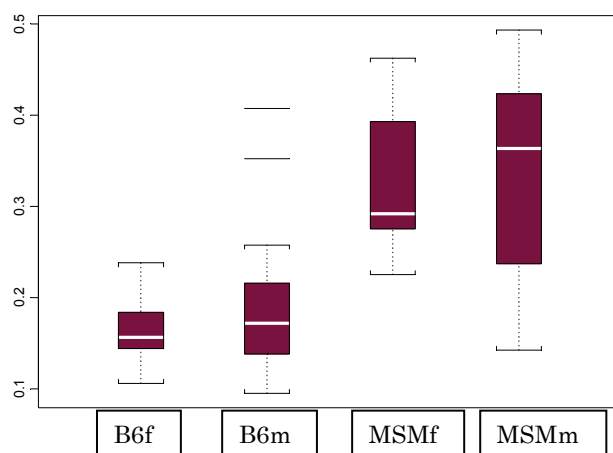


図 2. HMM により推定されたコンソミック株ごとの社会行動比率(青横線: 中央値、赤△: 各個体ペア推定値)

マウスの社会行動の 2 次元 unit vector chain による解析においては、これまでに加えた手法の改良に基づいて、マウス 24 系統 461 ペアに対する方向相互作用パラメータの尤度推定を行った。われわれの結果は、マウスの社会的接触による定量的解析の結果とうまく符合する面があると同時に、後者では得られない興味深い側面があることを示している。図には B6 と MSM の雌雄に対する最終結果 (box plot) のみを与えた。



課題 c. 多変量時系列モデルによるマウス自発行動の特徴抽出

(H.17) マウスのホームケージ内自発活動に関して、これまではある一定期間内での総活動量や日周期での活動量変化として解析する事が一般的であった。ここでは、これまで解析されてこなかった日内の短時間での変化の周期性を見つけ出すため、微小区間内にイベントが起きる確率をモデル化する点過程モデルにより解析を進めた。

(H.18) 遺伝的に異なったマウス系統である MSM と C57BL/6 (B6) を用いて、その活動性について解析系の確立を目指した。解析に用いたマウスは、照明を 8:00 から 20:00 までを明期、20:00 から 8:00 までを暗期としてコントロールし、室温 $22 \pm 2^\circ\text{C}$ 、湿度 $50 \pm 10\%$ に保った飼育室内で維持した。解析には MSM と C57BL/6 の生後 8~15 週齢の雌雄を解析に用いた。各マウスはテストの前に個別に飼育ケージに移し 1 日間飼育した。その後、ACTIVITY SENSOR (O'hara, Co., Ltd) の測定ケージにそれぞれ移し、4 日間連続でその活動量を自動計測した。最初の 1 日間は馴化期間として解析からは除外した。各個体の自発活動量は 1 分毎に表す時系列データとして記録した。

今年度は、MSM と B6 とで日周期のパターンに違いがあると言えるかどうかを検証することから始めた。その出発点として、1 分値データを 30 分に集約した上で、暗期のデータだけを取り出して接続した。これにより、個体ごとに時点数 72 の時系列データが得られる。

日周期パターンの違いは、モデル比較(モデル選択)によって議論する。比較するモデルを分かりやすく説明すると、ひとつは、MSM と B6 では全く異なる活動周期性を持つという仮説に対応するモデルで、

$$\text{個体活動量} = \text{個体別係数} \times \text{MSM 固有周期性} + \text{ノイズ}$$

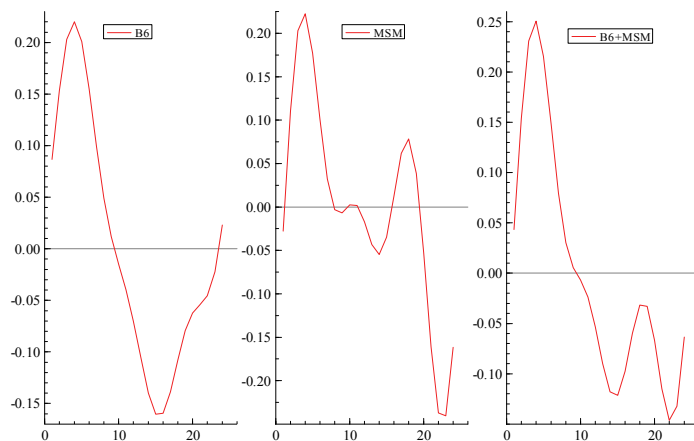
$$\text{個体活動量} = \text{個体別係数} \times \text{B6 固有周期性} + \text{ノイズ}$$

とでも表すべきものである。ここで周期性は複数の \sin , \cos で表現し、最短 4 時間の周期まで含める。一方、MSM と B6 とで周期性に違いがないのであれば、データを分けて推定しても混ぜて推定しても、結果に有意な差が生じないはずである。この「MSM と B6 とで周期に違いがない」という仮説に対応するモデルは、

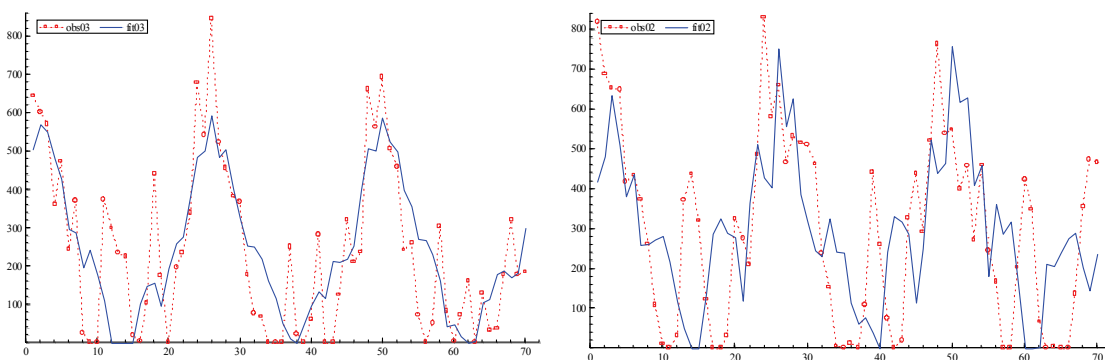
$$\text{個体活動量} = \text{個体別係数} \times \text{全共通周期性} + \text{ノイズ}$$

と書くことができよう。前者のほうが多くのパラメータを使用しているモデルであるが、データへの適合度の改善がそれに見合っているかどうかを情報量規準で比較する。種ごとに固有の周期性を仮定した

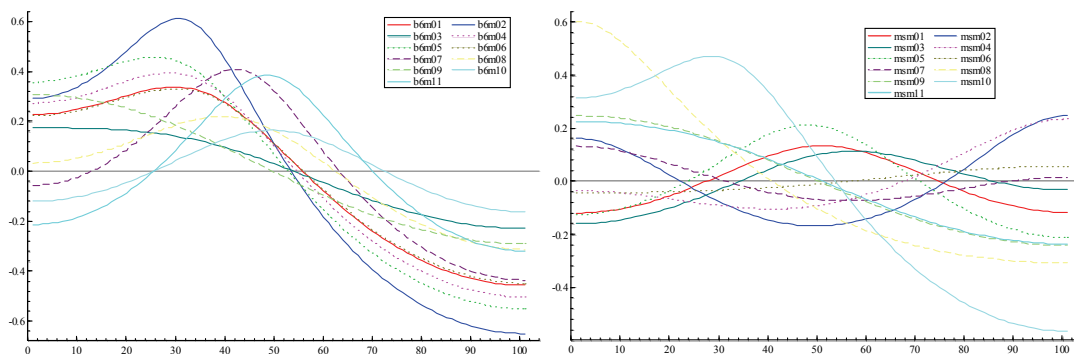
モデルの AIC は 8344.5、MSM と B6 とで共通周期性を仮定したモデルの AIC は 8403.6 となり、最小 AIC 規準から前者が支持され、MSM と B6 では日周期のパターンが異なることが確認された。そのことは、推定された周期関数を描画した右図からも確認できる。この図は、左から B6 単独、MSM 単独、両種混合での周期成分の推定結果を示し、横軸は午後 8 時から翌朝 8 時まで、30 分区分間にして 24 区間 (=12 時間) である。



次の問題として、推定された日周期まわりの変動に、MSM ないし B6 特有の短期的な変動特性がないかどうかを検証を試みた。下の図は B6 (左) と MSM (右) とで適当に個体を選び、推定された周期性 (青実線) の上に実際の活動量 (赤点線) を重ね書きしたものである。青線からの乖離部分に対して自己回帰モデル (AR モデル) をあてはめ、そこから導かれる周波数特性に MSM ないし B6 共通の特性が見えれば、種の特徴付けに役立つ情報と言える。



実際に推定した AR モデルの係数から導かれる規格化済スペクトルを B6 (左) と MSM (右) とで描画したのが下の図である。スペクトルにピークがあれば、その周波数帯に時系列データの変動が集中していることを表している。B6 に関しては周波数にバラツキはあるものの、1.5~3 時間のところに緩やかなピークが観察される。一方 MSM に関しては明確な傾向を見いだすことは難しい。MSM は原データを見ると、暗期に活動が集中する一方、細かい時間間隔での活動・休止は頻繁であるように思われることから、ここで時間分解能を粗くしたことでそれらが見えなくなっている可能性がある。



(H.19) 平成 19 年度は、再び 1 分値データの解析を行う立場に立ち返り、マウス活動量データに三角関数列を説明変数とするポアソン回帰モデルを適用し、24 時間周期から 2 分周期まで全ての可能な周波数の有意性を検出する方法を提示した。この悉皆的な周期性分析の結果、マウス活動量データの日周期を表すのに高周波成分が有意となることは殆どなく、2 時間から 6 時間程度の低周波成分が顕著であることがわかった。

累積標本スペクトル分布が 0.9 になるのに平均的にどこまで短い周期の貢献を必要とするか、という観点から結果を記すと、B6 雄で約 1 時間 20 分まで、B6 雌で約 2 時間まで、MSM 雄で約 2 時間まで、MSM 雌で約 6 時間までという結果を得た。ただし、こうした特徴量は個体差も大きく、カレンダータイムの意味での周期性のみから系統の特徴付けを行うほどには明瞭な特徴とは言えないとの結論に達した。

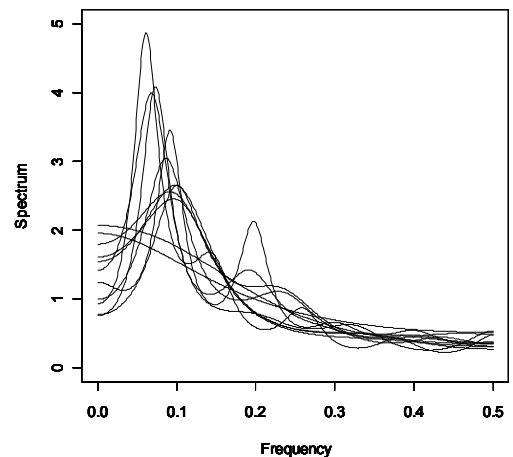
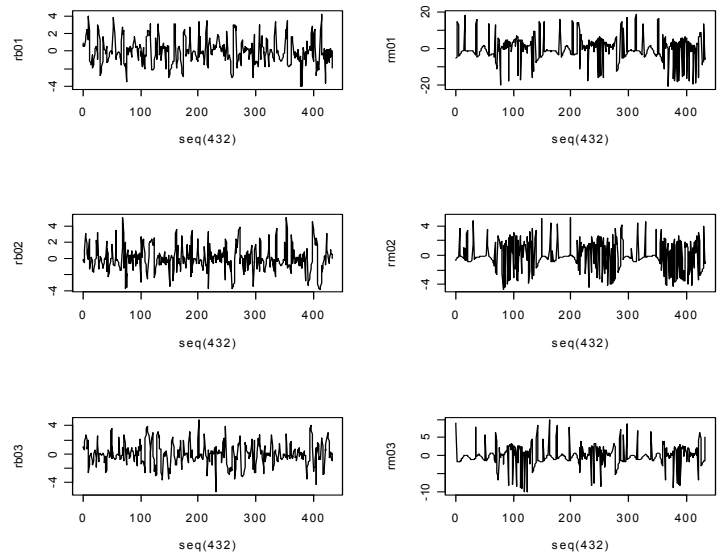
(H.20) マウスの自発行動を記録したデータは非負整数値による計数時系列で極めてゼロが多く、上下の非対称性が強い。興味があるのは活動期の短期的変動の特徴であるので、1 分値で計数 0 は欠測として非ゼロのデータを対数・巾関数等で線形化し、欠測値を許す線形ガウス型カルマンフィルタで、日周期まわりの自己回帰モデルをあてはめることにより、特徴的な活動周波数を推定する方法を研究した。

この結果、B6 (右の図で左側 3 枚のパネル) については比較的安定した短期的変動成分が抽出できているものの、MSM

(右側 3 枚のパネル) では明期と暗期の差が大きく、分散のモデル化が必要であることが示された。

(H.21) 計数 0 の対数値を $-1/e$ (ここで e は自然対数の底) で補間した対数強度系列から周期成分を取り除き、分散が時変モデルに従う AR 過程をあてはめ、短期的な活動周期性を抽出する方法を研究した。少なくとも B6 系統に関しては、確率的ボラティリティモデルの採用により、マウス活動性の短期的周期性は格段に推定が安定することがわかった。

右図は B6 雄 11 個体から推定された、周期成分除去後の短期的変動部分の分散変動を調整した上で得られた活動周期性のスペクトルである。活動周期性のピークは概ね 0.05 から 0.1 の間、すなわち時間にして 1 時間 40 分から 3 時間 20 分の間に分布していることが観察される。一方、MSM に関しては B6 ほどの明瞭な傾向は見られず、更なるモデル化と要因分解が必要との結論に至った。



課題 d. 生殖隔離障壁に関わるエピスタシスの統計学的検出と多重性調整

【研究の背景と目標】

生物学的「種」は「互いに交配可能な自然集団で、他のその様な集団から生殖的に隔離されている」と定義される。そのような種を分ける遺伝的しくみを生殖的隔離障壁と総称する。生殖的隔離機構の解明の重要性は古くから指摘されているが、生殖的隔離障壁となる遺伝子が同定され、その機構解明が進められている例は動植物を通じ少ない。Harushima らは、イネの F2 集団 (186 個体) において、異なる遺伝子座間で、特定の遺伝子型組合せにより配偶体又は接合体で選抜を受けた可能性があると考え、マーカーの遺伝子型分離の独立性検定を行うことにより、遺伝子座間の相互作用として生殖的隔離機構の検出が可能ではないかと考えた。そこで、12 本の染色体上の異なる座位のマーカー間 (1055×1055) で、遺伝子型分離の独立性についてカイ 2 乗検定を行った。その結果を図 (後述) にプロットすると、第 9 染色体と第 12 染色体間にみられるカイ 2 乗値 33.6 を最大のピークとして $p < 0.001$ を示すピークは 27 か所検出された。しかしながら、この交雑組合せの F1 稔性は高く、本当に相互作用によって特定の遺伝子型の組合せ個体が失われ、分離の独立性が失われたのかどうかは疑問であった。そこで、同じ組合せで別の F2 集団 (約 300 個体) について前の集団でピークを示したマーカー間で遺伝子型の独立性検定の再現性を確かめた。すると、多くのピークの再現性は得られなかった。本研究では、上記解析手法の統計学的妥当性を検討するとともに、生殖的隔離に関与する遺伝子座間の相互作用を検出する手法の検討を行った。

【研究成果】

最初に Harushima らが本研究に先立って行った、相互作用検出のためのデータ解析について説明する。イネの代表的な品種である Nipponbare (Japonica 種) と Kasalath (Indica 種) を交配した F1 を自殖して得られた F2 集団の 186 個体について、マーカー座の遺伝子型がデータとして得られている。イネの染色体 12 本 (全長 1650cM) に座数 1055 のマーカーが張られている。マーカーは共優性であるので、マーカー座 i とマーカー座 j の組み合わせについて、観測される遺伝子型の頻度は 3×3 表にまとめられる。この 3×3 表の独立性のカイ 2 乗統計量 (自由度 4) を T_{ij} とおく。この値が大きいとき、座間 i, j に相互作用があると判定することができる。 T_{ij} の等高線図を図に示す。

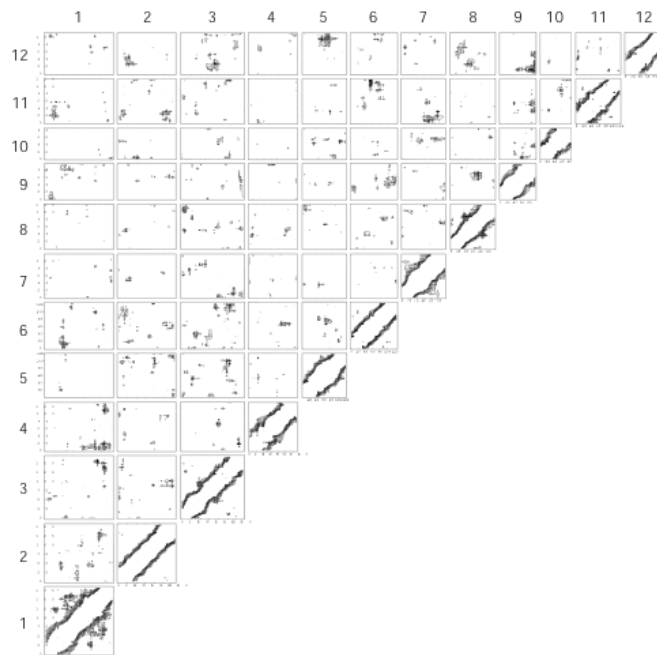


図. 3×3 表のカイ 2 乗値

先に述べたように、図における最大値は第 9 染色体と第 12 染色体間にみられるカイ 2 乗値 33.6 (これは自由度 4 のカイ 2 乗分布で $p = 0.89 \times 10^{-7}$ に相当) であるが、再現性がなく見せかけのピークの可能性が高いと思われる。このように一般に検定を多数繰り返して行う多重検定においては、見せかけの発見 (偽陽性) が多いことが問題となる。多重性を考慮した多重性調整 p 値は、

$$\text{adjusted } p\text{-value} = \Pr(\max_{i,j} T_{ij}(r.v.) \geq \max_{i,j} T_{ij}(\text{data}) | H_0)$$

と定義するものである。この式の計算のためには、相互作用が存在しないという帰無仮説 H_0 の下での T_{ij} の同時分布が必要である。我々の問題では、マーカー座が連鎖するため、各検定統計量 T_{ij} は独立ではないことに注意する。連鎖のモデルとしてホールデンの距離関数を仮定すると次が示される。

i, j を異なる染色体上のマーカー座とする。相互作用が存在しないという帰無仮説の下での T_{ij} の分布について、個体数 n が大きいときにすべての i, j に関する同時分布の意味で以下がなりたつ。

$$T_{ij} \cong T_{ij}^{(1)} + T_{ij}^{(2)} + T_{ij}^{(3)} + T_{ij}^{(4)}$$

この 4 つのコンポーネントは漸近的に独立に分布する。それぞれは自由度 1 のカイ 2 乗分布に従うが互いに相関をもつ。各コンポーネントは $T_{ij}^{(k)} = (z_{ij}^{(k)})^2$ 、ここで $z_{ij}^{(k)}$ は平均 0, 分散 1, 相関構造

$$\text{Cov}(z_{ij}^{(k)}, z_{i'j'}^{(k)}) = \exp\{-\rho_k |d_i - d_{i'}| - \lambda_k |e_j - e_{j'}|\} \delta_{kk}$$

を持つ正規変量と記述することができる。ただし d_i, e_j は遺伝子座 i, j の染色体上の位置 (単位は M)、また $\delta_{kk} = 1 (k = k'), 0 (k \neq k'), (\rho_k, \lambda_k) = (2, 2) (k=1), (2, 4), (k=2), (4, 2) (k=3), (4, 4) (k=4)$ である。異なる染色体上に位置する 2 つのマーカー座の遺伝的距離は無限大なので、 T_{ij} と $T_{i'j'}$ は $(i, j) = (i', j')$ でない限り独立である。

この結果を用いて多重性調整 p 値を 3 通りの方法 (a)シミュレーションによる方法、(b)非線形再生理論による方法、(c)チューブ法による方法、によって求めることができる。これを順に説明する。

(a) シミュレーションによる方法: $k=1, 2, 3, 4$ の場合について、係数 $\alpha_i = \exp(-\rho_i |d_i - d_{i-1}|)$, $\beta_j = \exp(-\lambda_j |e_j - e_{j-1}|)$ を用いて、ホワイトノイズ $\varepsilon_{ij} \sim N(0, 1)$ i.i.d. から、以下の漸化式によって逐次的に乱数 z_{ij} を生成する。

$$\begin{aligned} z_{11} &= \varepsilon_{11} \\ z_{i1} &= \alpha_i z_{i-1,1} + \sqrt{1 - \alpha_i^2} \varepsilon_{i1} \\ z_{1j} &= \beta_j z_{1,j-1} + \sqrt{1 - \beta_j^2} \varepsilon_{1j} \\ z_{ij} &= \alpha_i z_{i-1,j} + \beta_j z_{i,j-1} + \sqrt{1 - \alpha_i^2} \sqrt{1 - \beta_j^2} \varepsilon_{ij} \end{aligned}$$

4 つの場合のそれぞれを $z_{ij}^{(k)}$ ($k=1, 2, 3, 4$) とおき、 $W_{ij} = (z_{ij}^{(1)})^2 + (z_{ij}^{(2)})^2 + (z_{ij}^{(3)})^2 + (z_{ij}^{(4)})^2$ とおく。この乱数生成過程を 10000 回程度繰り返し、 $\max_{i,j} W_{ij}$ の経験分布 (ヒストグラム) を描くことによって多重性調整 p 値の近似値を得ることができる。

(b) 非線形再生理論による方法: 非線形再生理論は、D. Siegmund などによって開発された最大値分布の近似に関する方法論である。我々の問題については次の公式を得ることができる。

染色体の一つは全長 L (M) でその上には m 個のマーカーが等間隔 Δ (M) で配置されているとする。もう一方は全長 L' (M) でマーカー m' 個が等間隔 Δ' (M) で配置されているとする。このとき b が大きいとき、近似式

$$\Pr(\max_{i,j} T_{ij} \geq b^2) \approx 48LL'C(b)g_8(b^2)$$

がなりたつ。ここで g_8 は自由度 8 のカイ乗分布の密度関数、また $\bar{\rho}(u) = 2(u_1^2 + u_2^2) + 4(u_3^2 + u_4^2)$,

$\bar{\rho}'(u) = 2(u_1^2 + u_3^2) + 4(u_2^2 + u_4^2)$ とおくと、

$$C(b) = \int_{S^{4-1}} du \bar{\rho}(u) \bar{\rho}'(u) \nu(b\sqrt{2\bar{\rho}(u)\Delta}) \nu(b\sqrt{2\bar{\rho}'(u)\Delta'})$$

ただし

$$\nu(x) = 2x^{-2} \exp\left\{-2 \sum_{n=1}^{\infty} n^{-1} \Phi\left(-\frac{1}{2}x\sqrt{n}\right)\right\} \quad (x > 0)$$

($\Phi(\cdot)$ は標準正規分布の分布関数)。現実にはマーカーが等間隔ということは考えられないが、不等間隔の場合であってもマーカー間隔の平均値を Δ あるいは Δ' とおくとよい近似式を与えることが別の数値実験によって確認されている。

(c) チューブ法による方法：チューブ法は滑らかなサンプルパスを持つ連続な確率過程、確率場の最大値分布を求める方法論であり、数学的には積分幾何学に基づくものである。本来 T_{ij} は離散的な集合の上で定義されているものであるが、これを連続的な確率過程に埋め込むことによってチューブ法を適用することができる。我々の問題については次のような公式を得ることができる。

$\alpha_i = \exp(-2|d_i - d_{i-1}|)$, $\beta_j = \exp(-2|e_j - e_{j-1}|)$ とおく。

$$\Pr(\max_{i,j} T_{ij} \geq b^2) \leq 2C \sum_i \cos^{-1} \alpha_i \sum_j \cos^{-1} \beta_j \times g_6(b^2)$$

ここで $C=0.946$, g_6 は自由度 6 のカイ乗分布の密度関数である。

データ解析の結果は以下の通りであった。(a)のシミュレーションによると、最大値 33.6 の調整 p 値は 0.068 であり、5%有意でもないことが分かる。これはピークの最大値が偽陽性であったことと整合している。ただしこの計算には、統計数理研のスーパーコンピュータ (SGI Altix3700) で 14 日 8 時間 (延べ時間) 必要とした。また(b)と(c)の方法による p 値はそれぞれ、0.104, 0.122 であった。

Harushima らは、F2 実験とは別に、同じイネを対象にして戻し交配 (BC) 実験で同様のデータを採集している。BC の場合は、マーカー遺伝子型は 2 値のため、カイ 2 乗統計量は自由度 1 である。この実験における最大カイ 2 乗値は 39.01 であり、多重性調整 p 値は 2.88×10^{-6} ((b)による)、 7.23×10^{-6} ((c)による) となり、多重性調整後も高度に有意であった。この遺伝子座ペアは、真の生殖隔離障壁と考えることができる。

課題 e. QTL 解析法の最適化

【研究目的】

個体の形質を規定する遺伝子 QTL を探索するための QTL 解析では、ロッドスコアの極大点を探索することによって QTL の位置を推測する。しかし、同じ染色体に 2 つ以上のロッドスコアの極大値が存在するときに、その 2 つの極大値が 2 つの QTL がもたらすものか、あるいは単なる統計誤差によるものかどうかは遺伝学的に重要であり、慎重に吟味する必要がある。そのためには、ロッドスコアの形状に影響を与える個体を検出し、その個体のデータが信頼できるものか、何らかの誤りをふくむものかどうかを再検討する必要がある。この目的のため従来は、ロッドスコアの極大点付近で遺伝子型が変化するような個体の洗い出しを行っていた。しかしながら、この方法は非効率であるばかりでなく、表現型データを参照していないという意味で不十分な方法である。本研究ではロッドスコアの形状に影響する個体を系統型に検出する方法の開発を目的としている。

【研究成果】

マウスの肥満の原因となる遺伝子を探索するため、肥満と関連性が深いとされる血中アディポネクチンに着目し、標準的マウス近交系である B6 と、日本産亜種由来の MSM 系統の F2 雑種 170 個体に対して解析を行う。各マウスの 20 対の染色体における 119 箇所の遺伝子型が三つの値 (A, H, B) として観測されている。ここでは、血中アディポネクチンの対数を目的変数とし、遺伝子型の加法効果、優性効果ならび性別を変数とする線形モデルを仮定し、QTL 解析 (単一マーカー分析) を行う。得られたロッドスコアを図 1 に示す。

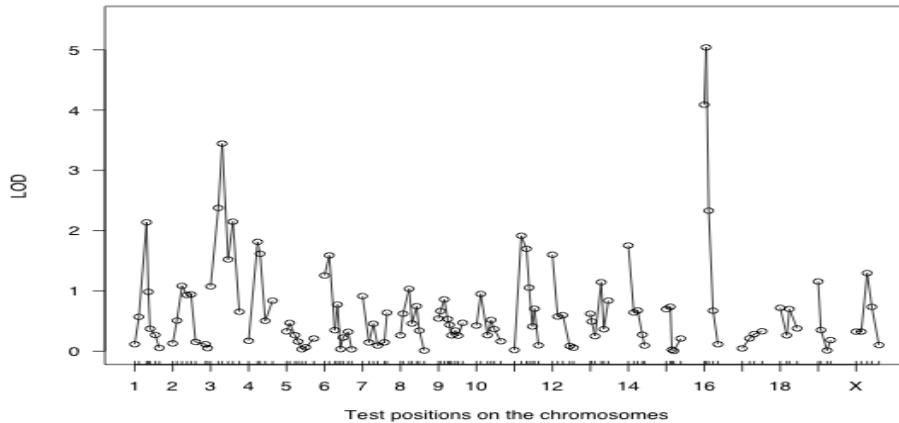


図 1. ロッドスコア

図 1 において、第 3 染色体上にロッドスコアの極大値が二つ存在することが見て取れる (双峰)。このような曲線の形が、少数の外れ値によるものかどうかを調べたい。そのために、複数の遺伝子座におけるロッドスコアの標本影響関数を求め、ロッドスコア曲線の形に影響を与える個体のグループを検出する。第 i 座のロッドスコアの $1/n$ 倍は経験分布関数の汎関数であるので、その標本影響関数 SIF_i が定義できる。興味のある遺伝子座 $i = (i_1, \dots, i_k)$ に対してロッドスコア曲線の変動に影響する個体を検出する。

複数の統計量の影響分析の方法としては Cook (1996) の方法がよく知られている。これは、モデルの適合度に影響を与えるパラメータの摂動方向を探索するものであるが、QTL 解析のようにモデルの当てはめによらず統計的検定を解析の目的とする場合はその考え方を直接適用できない。ここでは、複数の遺伝子座におけるロッドスコアの値 (ベクトル値) に対する標本影響関数行列 (Li, et al. 1997) を求め、主成分分析と類似の考え方によってロッドスコアに影響を与える個体のグループを検出する方法を考える。

ロッドスコアの、QTL がどこにも存在しないという帰無仮説の下での漸近共分散行列を Σ とおく (詳細略)。ロッドスコアの形状の変動を考えるためには、行列を用いて平均の変動を除いて影響する個体を見つけ出す。ここでは、 P を平均値を引き去る射影行列とし

$$(SIF_1, \dots, SIF_k)P(P'\Sigma P)^{-1}P'(SIF_1, \dots, SIF_k)'$$

に対して固有値解析を行う。このとき、行列の固有ベクトルはロッドスコア曲線の何らかの形状的な特性を意味するものと考えられ、固有ベクトルに大きな値をもつ個体はその特性に対する影響集合と判断できる。結果を図 2 に示す。

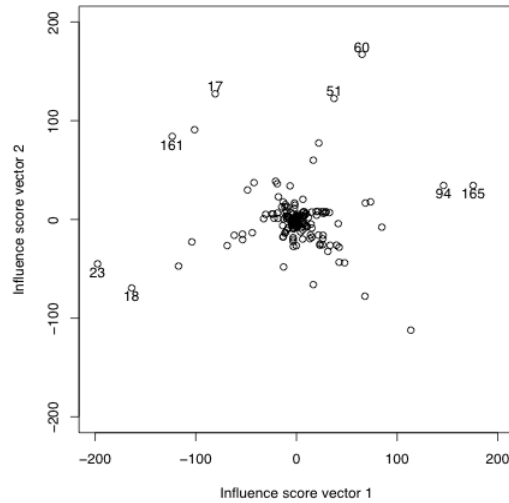


図 2. 影響関数の固有値解析

図 2 において第 1 軸方向に大きな絶対値をとる個体 23 と 60 および第 2 軸方向に大きな値をとる個体 51 と 60 を除いて得られたロッドスコア曲線はそれぞれ図 3 と図 4 のように示される。これらの図から図 2 の横軸（第 1 固有ベクトル）は全体の傾きを表し、縦軸（第 2 固有ベクトル）は凸性または谷の有無に対応することがわかる。

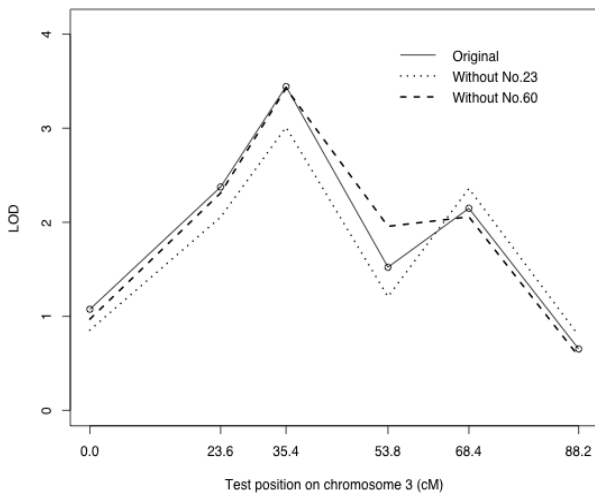


図 3. 個体 No.23,60 を除いて得られたロッドスコア

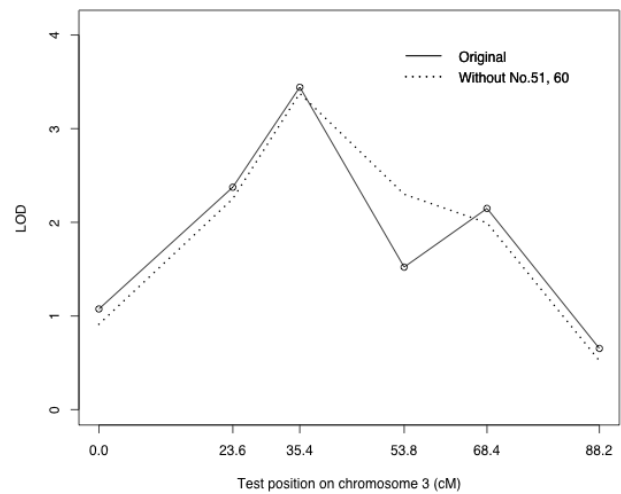


図 4. 個体 No.51, 60 を除いて得られたロッドスコア (第 1~3 染色体)

また、区間マッピング法によって得られるロッドスコアの曲線について、その高さや位置に対する標本影響関数を導出した。図 5, 6 の実線は、区間マッピング法を用いて描いたロッドスコアである。図 5 の点線は、ロッドスコアの高さに関する影響関数の値（絶対値）に影響を与える個体を取り除いたときにロッドスコアがどう変化するかを見たものである。図 6 は位置に関する影響関数の値が大きな個体を取り除いたときの変化を観察したものである。影響関数が正しくピークの位置に関する感度をとらえていることが図から見て取れる。

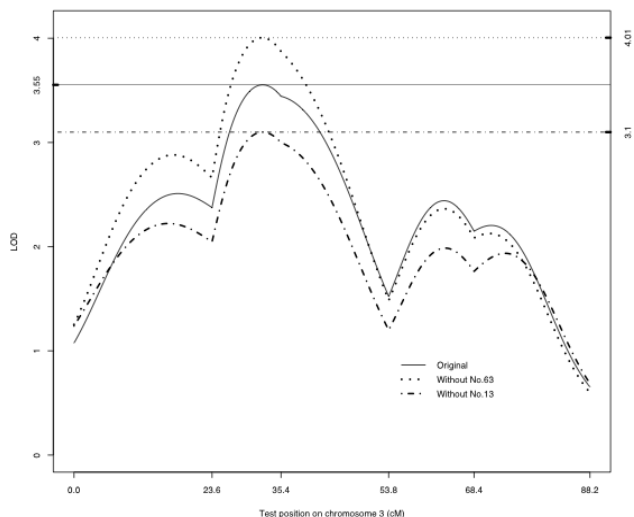


図 5. 個体 No.63, 13 を除いて得られたロッドスコア (区間マッピング法)

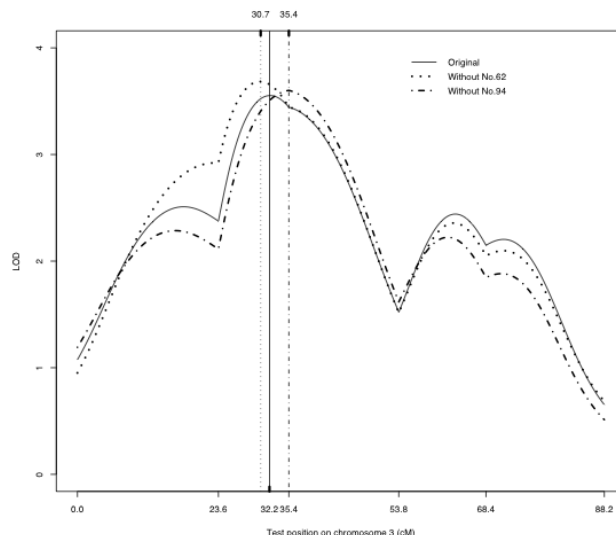


図 6. 個体 No.62, 94 を除いて得られたロッドスコア (区間マッピング法)

(まとめ) ロッドスコア曲線の形状に影響を与える個体の集合を特定する方法を提案した。この方法によってロッドスコア曲線の傾き、凸性及び平均移動などの特性に影響するそれぞれの個体集合を検出できた。

① 成果物 (知見・成果物・知的財産権等)

サブテーマ 2

(プログラム・ソフトウェア開発)

- ・ P 型フーリエ記述子に関するソフトウェア (<http://www.ism.ac.jp/~tamura/judge.html>) マウス下顎骨の輪郭線を統計モデリングにより記載するソフト。
- ・ マウス行動観測用ソフトウェア
機能：マウスの動画像を目視観察する時にマウスの状態を効率良く記録するためのソフト。早送り・巻戻し等が可能。

② 成果発表等

<論文発表>

[学術論文]

- 1 江口真透、ゲノムデータ解析のための統計的方法を目指して、特集:予測と発見、統計数理 54, 375-403 (2006).
- 2 Fushiki, T., Fujisawa, H. and Eguchi, S., Identification of biomarkers from mass spectrometry data using a "common" peak approach, BMC Bioinformatics 7, 358 (2006).
- 3 Eguchi, S. and Yoshioka, K. Maximum regularized likelihood estimation of finite mixtures with a structural model. Communications in Statistics - Theory and Methods, 39: 8, 1498 - 1510
- 4 Fujisawa, H., Isomura, H., Eguchi, S., Ushijima, M., Miyata, S., Miki, Y., Matsuura, M. Identifying haplotype block structure by using ancestor-derived model. J. Human Genetics 52 (9), 738-746, 2007.

- 5 Kanamori, T., Takenouchi, T., Eguchi, S. Murata, N. Robust loss functions for boosting. *Neural Computation* 19, 2183-2244, 2007
- 6 Oka, A., Aoto, T., Totsuka, Y., Takahashi, R., Ueda, M., Mita, A., Sakurai-Yamatani, N., Yamamoto, H., Kuriki, S., Takagi, N., Moriwaki, K. and Shiroishi, T. Disruption of Genetic Interaction Between Two Autosomal Regions and the X Chromosome Causes Reproductive Isolation Between Mouse Strains Derived From Different Subspecies. *Genetics* 175, 185-197 (2007).
- 7 Mollah, M. N. H., Minami, M. and Eguchi, S. Robust prewhitening for ICA by minimizing beta-divergence and its application to FastICA. *Neural Processing Letters* 25, 91-110, 2007.
- 8 Ninomiya, S. and Fujisawa, H. A conservative test for multiple comparison based on highly correlated test statistics, *Biometrics*, 63, 1135-1142, 2007.
- 9 Takenouchi, T., Eguchi, S., Murata, N. and Kanamori, T. Robust boosting algorithm against mislabeling in multi-class problems. To appear in *Neural Computation*, 2007.
- 10 Takenouchi, T., Ushijima, M. and Eguchi, S. GroupAdaBoost: accurate prediction and selection of important genes. *IPSJ Transactions on Bioinformatics* 3, 1-8, 2007.
- 11 Ushijima, M., Miyata, S., Eguchi, S., Kawakita, M., Yoshimoto, M., Iwase, T., Akiyama, F., Sakamoto, G., Nagasaki, K., Miki, Y., Noda, T., Hoshikawa Y. and Matsuura, M. Common peak approach using mass spectrometry data sets for predicting the effects of anticancer drugs on breast cancer. *Cancer Informatics* 3, 285-293, 2007.
- 12 Fujisawa, Hironori., Eguchi, Shinto., Robust parameter estimation with a small bias against heavy contamination. *Journal of Multivariate Analysis* 99, 2053-2081 (Oct., 2008)
- 13 Hobo, T. Suwabe, K. Aya, K. Suzuki, G. Yano, K. Ishimizu, T. Fujita, M. Kikuchi, S. Hamada, K. Miyano, M. Fujioka, T. Kaneko, F. Kazama, T. Mizuta, Y. Takahashi, H. Shiono, K. Nakazono, M. Tsutsumi, N. Nagamura, Y. Kurata, N. Watanabe, M. Matsuoka, M. Various spatiotemporal expression profiles of anther-expressed genes in rice. *Plant Cell Physiol.*49:1417-1428. 2008.
- 14 Ito, N. and Kurata, N. Disruption of KNOX gene suppression in leaf by introducing its cDNA in rice. *Plant Science*. 174: 279-289. 2008.
- 15 Kawakita, Masanori., Eguchi, Shinto., Boosting method for local learning in statistical pattern recognition. *Neural Computation* 20, 2792-2838 (Nov., 2008)
- 16 Suwabe, K. Suzuki, G. Takahashi, H. Shiono, K. Endo, M. Yano, K. Fujita, M. Masuko, H. Saito, H. Fujioka, T. Kaneko, F. Kazama, T. Mizuta, Y. Separated transcriptomes of male gametophyte and tapetum in rice: validity of a laser microdissection (LM) microarray. *Plant Cell Physiol.*49:1407-1416. 2008.
- 17 Takada, T., Mita, A., Maeno, A., Sakai, T., Shitara, H., Kikkawa, Y., Moriwaki, K., Yonekawa, H. and Shiroishi, T.: Mouse inter-subspecific consomic strains for genetic dissection of quantitative complex traits, *Genome Res*. 18, 500-508 (2008).
- 18 Suzuki, T., Eiguchi, M., Kumamaru T., Satoh, H., Matsusaka, H., Moriguchi, K. and Kurata, N. MNU-induced mutant pools and high performance TILLING enable finding of any gene mutation in rice. *Mol Gen Genomics.*279:213-223. 2008.
- 19 Takahashi A., Nishi A., Ishii A., Shiroishi T., Koide T.: Systematic analysis of emotionality in consomic mouse strains established from C57BL/6J and wild-derived MSM/Ms. *Genes, Brain and Behavior*. 7:849-858, 2008.
- 20 Takahashi A., Shiroishi T., Koide T.: Multigenic factors associated with a hydrocephalus-like

- phenotype found in inter-subspecific consomic mouse strains. *Mammalian Genome*, 29:333-338, 2008.
- 21 Takenouchi, Takashi., Eguchi, Shinto., Murata, Noboru., Kanamori, Takafumi., Robust Boosting Algorithm Against Mislabeling in Multiclass Problems. *Neural Computation* 20, 1596-1630 (June, 2008)
 - 22 Tanaka, E. and Tamura, Y., "Skeleton-based Fourier Descriptor of Open Curves" , *Pacific Science Review*, Vol. 10, No. 10, 194-198, 2008.
 - 23 Tanaka, E., Tamura, Y., Hosoya, M. and Shiroishi, T., "Protrusion Fourier Descriptor: Skeleton-based Representation of Open Curves", *Forma*, Vol. 23, 9-18, 2008.
 - 24 Thirumurugan, T., Ito, Y., Kubo, T., Serizawa, A. and Kurata, N. Identification, characterization and interaction of HAP family genes in rice. *Mol Gen Genomics*.279: 279-289. 2008.
 - 25 Zheng, Z., Iwata, H., Hirata, H. and Tamura, Y. Quantitative evaluation of the degree of sprout leaf bending of rice cultivars using P-type Fourier descriptors and principal component analysis, *Euphytica*, 2008, DOI 10.1007/s10681-007-9642-9 (to appear)(online 版 <http://www.springerlink.com/content/513u046251521246/>)
 - 26 Huang, S-Y., Yeh, Y-R. and Eguchi, S. Robust kernel principal component analysis, *Neural Computation*, 21, 3, (2009) 179-3213.
 - 27 プリチャード真理, 江口真透. 関連遺伝子セットの多重解の存在. *日本統計学会誌* 38 巻 シリーズ J 2 号, 199-212 (2009)
 - 28 Takahashi A, Tomihara K, Shiroishi T, Koide T. Genetic mapping of social interaction behavior in B6/MSM consomic mouse strains. *Behavior Genetics* 40: 366-376.2009
 - 29 Umemori J, Nishi A, Lionikas A, Sakaguchi T, Kuriki S, Blizard DA and Koide T. QTL analyses of temporal and intensity components of home-cage activity in KJR and C57BL/6J strains. *BMC Genetics* 10:40, 2009.
 - 30 Mollah, N. H., Minami, M. and Eguchi, S. Robust extraction of local structures by the minimum beta-divergence method, *Neural Networks*, 23, 2 (2010) 226-23
 - 31 Copas, J. and Eguchi, S. Likelihood for statistically equivalent models. *J. Royal Statistical Society B*, 72, 2 (2010) 193-217.
 - 32 Dowse H., Umemori J., Koide T. Ultradian Components in the Locomotor Activity Rhythms of the Genetically Normal Mouse, *Mus musculus*. *J Exp. Biol.* (印刷中)
 - 33 Kubota T, Miyake K, Hirasawa T, Nagai K, Koide T. Novel etiological and therapeutic strategies for neurodiseases: Epigenetic understanding of gene-environment interactions. *J Pharmacol Sci* (印刷中).
 - 34 Tanaka N., Waki K., Kaneda H., Suzuki T., Yamada I., Furuse T, Kobayashi K., Motegi H., Toki H., Inoue M., Minowa O., Noda N., Takao K., Miyakawa T., Takahashi A., Koide T., Wakana S., Masuya H. SDOP-DB: A comparative standardised-protocol database for mouse phenotypic analyses. *Bioinformatics* (in press)

[会議録]

Kawasaki, Y., Koide, T., Nishi, A., and Umemori, J., Bayesian analysis of locomotor activities of wild mice strains, Valencia/ISBA 8th World Meeting on Bayesian Statistics, Universita de Valencia, Spain, 2006, p.145.

Mollah, Md. Nurul Haque, Eguchi, Shinto Robust Composite Interval Mapping for QTL Analysis by Minimum β -Divergence Method. IEEE International Conference on Bioinformatics and Biomedicine 2008, 115-120, 978-0-7695-3452-7, 2008.11

長戸康郎, 倉田のり 国立遺伝学研究所研究会「イネ分子遺伝学の新展開」, 国立遺伝学研究所, 2008年12月19-20日, 三島

渡辺正夫, 倉田のり 国立遺伝学研究所研究会「高等植物の生殖過程を制御する因子の多様性と生殖隔離」, 国立遺伝学研究所, 2008年11月5日-6日, 三島.

[解説・総説]

江口真透, リスク科学のための人工知能的方法と統計的思考. 人工知能学会誌 22巻5号特集「定量的リスク科学」656-662, 2007年9月.

江口真透, 統計的パターン認識: 線型判別からアダブーストまで. 日本化学会情報化学部会誌 25, 68, 2007.

倉田のり, 春島嘉章, イネゲノムと生殖隔離, 細胞工学別冊 植物細胞工学シリーズ 23 「植物の進化」, pp. 97-101, 清水健太郎, 長谷部光泰監修, 秀潤社, 東京 (2007).

栗木哲「QTL解析の統計モデルと検定の多重性調整」, 21世紀の統計科学, II, 小西貞則, 国友直人(編), 東京大学出版会, 東京, 315-356, 2008.7

Koide T, Catanesi CI, Nishi A, Shiroishi T, Kasai S, Ikeda K, Takahashi A. Systematic mapping of pain-related QTL using consomic mouse strains: Advantage of using wild-derived strains. Brain Res J. in press.

[招待講演]

- 1 Eguchi, S., Statistical pattern recognition and genome data analysis, Institute of Statistical Science, Academia Sinica, 2006年12月16日.
- 2 春島嘉章, イネ生殖的隔離の遺伝解析, 日本進化学会, 2006年大会, 2006年8月30日.
- 3 春島嘉章, 水多陽子, 栗木哲, 藤澤洋徳, 倉田のり, イネ生殖的隔離の遺伝解析, 日本分子生物学会 2006 フォーラム, 2006年12月6日.
- 4 Eguchi, S., Statistical learning theory and bioinformatics, The 4th International Symposium on Ubiquitous Knowledge Network Environment 2007, Sapporo Convention Center, Sapporo, Japan, 2007年3月7日.
- 5 Eguchi, S. Boosting learning approach to association studies in bioinformatics, First joint meeting between Institute of Statistical Science, November 29, 2007, Academia Sinica, Taiwan and the Institute of Statistical Mathematics.
- 6 Eguchi, S. Boosting methods for association studies in bioinformatics, International Conference on Multiple Decision Theory, Statistical Inference and Applications, December 28, 2007, Taipei, Republic of China.
- 7 Harushima, Y., Kuriki, S., Mizuta, Y., and Kurata, N., Detection of pairs of interactive reproductive barriers within gametophyte or zygote, The Fifth Okazaki biology conference: "Speciation and Adaptation", March 13th, 2007.
- 8 Kuriki, S. Multiplicity Adjustments in Detecting Reproductive Barriers caused by Loci Interactions, First joint meeting between Institute of Statistical Science, Academia Sinica, Taiwan and the Institute of Statistical Mathematics, November 29, 2007, Tokyo.
- 9 Eguchi, Shinto. Information divergence geometry and its application to machine learning. The 1st

MSJ-SI, Probabilistic Approach to Geometry, Kyoto, Japan.2008.8.4

- 10 江口真透. 表現形の相関研究：パターン認識の先に何が見えるか. 春季統計学会, 2008年3月1日.
- 11 Eguchi, S. Boosting true positive and false positive rates for pattern recognition (Invited) Institute of Mathematical Statistics Asia Pacific Rim 1st Meeting, Seoul, South Korea, 2009.06.29
- 12 Eguchi, S. Information divergence geometry and its application to machine learning. The Fifth Statistics and Machine Learning Workshop, Tainan, Taiwan, 2009.04.28
- 13 Eguchi, S. Projective Tsallis Entropy and its Application to Robust Statistics. (Invited talk) Mathematical Aspects of Generalized Entropies and their Applications. Kyoto, Japan, 2009.07.08
- 14 江口真透. タンパク質構造と進化と情報幾何. 数理研短期共同研究集会「離散力学系の分子細胞生物学への応用数理」. 京都, 日本. 2009.1.8
- 15 小出剛, 「マウスにおけるコミュニケーションとその遺伝的要因ー行動遺伝学を通して」 金沢大学創基150年記念「講演会・シンポジウム」「社会認識」学際脳科学シンポジウム平成21年11月5日
- 16 Mollah, N. H., Eguchi, S. Robust QTL analysis by minimum beta-divergence method. (Invited talk) TRIC Symposium "Statistical and Systems Genetics", Mishima, Japan, 2009.10.22

[一般講演]

- 1 江口真透, アレイデータ解析からの統計学の新しい方向, 統計学関連学会, 2006年9月6日.
- 2 春島嘉章, 栗木哲, 水多陽子, 藤澤洋徳, 倉田のり, 配偶体内または接合体内の異なる遺伝子座間の相互作用による生殖的隔離障壁の検出, 日本育種学会第110回講演会, 愛媛大学, 松山, 2006年9月22日.
- 3 堀内陽子, 春島嘉章, 川喜田雅則, 望月孝子, 江口真透, 倉田のり, Affymetrix Rice Genome arrayを用いたイネ遺伝子発現量検出における塩基配列差の補正と適用, 日本育種学会第110回講演会, 愛媛大学, 松山, 2006年9月23日.
- 4 Kawasaki, Y., Koide, T., Nishi, A., and Umemori, J., Bayesian analysis of locomotor activities of wild mice strains, Valencia/ISBA 8th World Meeting on Bayesian Statistics, 2006年6月3日, Grand Hotel Bali, Benidorm, Spain.
- 5 多陽子, 春島嘉章, 倉田のり, イネ「日本晴」と「Kasalath」の雑種花粉で相互作用する2遺伝子座に起因する生殖的隔離, 日本育種学会第110回講演会, 愛媛大学, 松山, 2006年9月22日.
- 6 江口真透, 竹之内高志, マルチクラスの判別問題におけるミスラベルのモデリング, 統計関連学会, 2007年9月7日
- 7 春島嘉章, 倉田のり「栽培イネの第3染色体の雄性配偶体型生殖的隔離障壁と相互作用する第6染色体の雌性親遺伝子のポジショナルクローニング」日本育種学会112回講演会, 山形大学, 2007年9月22日, 鶴岡.
- 8 春島嘉章, 栗木哲, 水多陽子, 藤澤洋徳, 倉田のり「イネ F2 集団における生殖的隔離障壁の相互作用の検出」第79回日本遺伝学会, 岡山大学, 2007年9月21日, 岡山.
- 9 堀内陽子, 藤澤洋徳, 川喜田雅則, 望月孝子, 春島嘉章, 坂口隆之, 江口真透, 倉田のり「Rice Genome Arrayを用いた SFP 検出手法の開発」日本育種学会112回講演会, 山形大学, 2007年9月23日, 鶴岡.
- 10 堀内陽子, 藤澤洋徳, 川喜田雅則, 望月孝子, 春島嘉章, 坂口隆之, 江口真透, 倉田のり「Affymetrix Rice Genome Arrayを用いた SFP 検出手法の開発」第30回日本分子生物学会年会・第80回日本生化学会大会 合同大会, パシフィコ横浜, 2007年12月13日, 横浜.
- 11 Kawakita, M., Fujisawa, H., Eguchi, S., Ushijima, M., Ishikawa, Y., and Matsuura, M., A novel variable selection method in gene expression profiling for the classification of the distinct two histological types of high-grade neuroendocrine lung carcinomas, Pacific Symposium of

- Biocomputing 2007, Grand Wailea, Maui, Hawaii, 2007年1月7日.
- 12 Umemori J., Nishi A., Takahashi A., Kawasaki Y., Lionikas A., Blizard D., Koide, T. QTL analysis of differences of activities in home-cage and open-field between KJR and B6 mouse strains. 6th Annual Meeting of the Complex Trait Consortium, May 26-29, 2007, Braunschweig, Germany.
 - 13 川崎能典, 多変量計数時系列データのモデリング, 統計数理セミナー, 2007年4月18日, 統計数理研究所, 東京.
 - 14 栗木哲, 直積型の相関構造を持つカイ2乗確率場の最大値の分布, 日本数学会2007年度年会統計数学分科会, 2007年3月29日, 埼玉大学, さいたま市(講演アブストラクト, 89-90).
 - 15 栗木哲, 格子点上カイ2乗確率場の最大値分布の近似とその連鎖解析への応用, 科研費研究集会「統計的モデリングの方法と理論」, 一橋大学, 2007年11月27日, 東京.
 - 16 栗木哲, 生殖的隔離に関わるエピスタシス(相互作用)検出と多重性調整, 融合研究シンポジウム「地球と生命の新パラダイム創造への挑戦」ハイライト講演, 2007年11月29日, 東京.
 - 17 栗木哲, 春島嘉章, 藤澤洋徳, 倉田のり, 遺伝子座間の相互作用による生殖的隔離障壁の検出と多重性調整, 統計関連学会連合大会, 神戸大学, 2007年9月6日, 神戸.
 - 18 水多陽子, 春島嘉章, 倉田のり「イネ亜種間交雑における生殖的隔離障壁遺伝子のポジショナルクローニング」2007年イネ分子遺伝学ワークショップ, 名古屋大学, 2007年7月3日, 名古屋.
 - 19 水多陽子, 春島嘉章, 倉田のり「イネ雑種花粉で相互作用する2遺伝子座に起因する生殖的隔離」第79回日本遺伝学会, 岡山大学, 2007年9月20日, 岡山.
 - 20 水多陽子, 春島嘉章, 倉田のり「イネ雑種花粉で作用する生殖的隔離障壁遺伝子のポジショナルクローニング」日本育種学会112回講演会, 山形大学, 2007年9月23日, 鶴岡.
 - 21 水多陽子, 春島嘉章, 倉田のり「イネ雑種花粉で作用する生殖的隔離障壁の単離・機能解析」特定領域研究「植物ゲノム障壁」第1回若手の会, 茂庭荘, 2007年11月5日, 仙台.
 - 22 江口真透, 観察データの推測の限界—揺らぎモデルアプローチ—広島大学理学部, 科研費研究集会「統計的データ解析手法の評価と開発」2008年1月16日, 広島.
 - 23 江口真透. バイオインフォマティクスにおける統計的課題について. 科研費研究集会「高次元データの統計解析」博多. 2008.11.21
 - 24 Harushima Y., Kuriki S., Mizuta Y., Kurata N. Detection of pairs of interactive reproductive barriers in rice. XX International congress of Genetics, Berlin, Germany, July 12-17, 2008.
 - 25 Harushima Y., Yano M., Kurata N. Identification of areproductive barrier working in the process of pollen competition in rice. XX International congress of Genetics, Berlin, Germany, July 12-17, 2008.
 - 26 Horiuch Y., Harushima Y., Mochizuki T., Fujisawa H., Eguchi., Kawakita M., Kurata N. Detection of nucleotide and expression polymorphisms between rice stains using affymetrix rice genome array. XX International congress of Genetics, Berlin, Germany, July 12-17, 2008.
 - 27 小森理, 江口真透. 1クラスラベルに注目したブースティング. 統計関連学会連合大会, 横浜. 2008.9.8
 - 28 久保貴彦, 水多陽子, 新濱充, 春島嘉章, 倉田のり「イネ生殖隔離機構の解析から見えてくるもの」国立遺伝学研究所研究会「高等植物の生殖過程を制御する因子の多様性と生殖隔離」, 国立遺伝学研究所, 2008年11月5日, 三島.
 - 29 Md. Nurul Haque Mollah., Eguch, Shinto., Robust QTL Analysis by the Minimum β -Divergence Method. International Association for Statistical Computing 2008, Yokohama, 2008.12.6
 - 30 Mizuta Y., Harushima Y., Kurata N. Positional cloning of a pair of interactive genes causing reproductive barrier in the hybrid pollen of rice. XX International congress of Genetics, Berlin,

Germany, July 12-17, 2008.

- 31 水多陽子, 春島嘉章, 倉田のり「イネ亜種間交雑で生殖的隔離障壁となる重複遺伝子の解析」日本育種学会 114 回講演会, 滋賀県立大学, 2008 年 10 月 11 日, 彦根.
- 32 望月孝子, 菊池俊介, 濱田和輝, 加藤大貴, 大木伸彦, 藤田雅丈, 堀内陽子, 倉田のり, 矢野健太郎「OryzeExpress : イネのゲノム・アノテーションとオミックス統合データベース」第 31 回日本分子生物学会年会, 第 81 回日本生化学会大会 合同大会, 神戸ポートアイランド, 2008 年 12 月 10 日, 神戸.
- 33 Pritchard, Mari., Eguchi, Shinto. Finding Optimal Gene Set for Classification from Multiple Predictive Gene Sets. International Association for Statistical Computing, Yokohama, 2008.12.6
- 34 プリチャード 真理, 江口 真透. マイクロアレイにおける遺伝子選択と判別能力の関係. 統計関連学会連合大会, 横浜. 2008.9.9
- 35 Takahashi, A., Sugimoto, H., Kimura, S., Tomihara, K., Tsuchiya, T., Kakihara, S., Tanemura, M., Shiroishi, T., Koide, T.: Complex genetic architecture of social interaction and aggressive behavior clarified using consomic strains derived from MSM and C57BL/6. 22nd International Mammalian Genome Conference, Prague, November 2-5, 2008. (ポスター)
- 36 Takahashi, A., Tomihara, K., Shiroishi, T., Koide, T.: Genetic analysis of inter-male aggression using consomic mouse strains established from C57BL/6J and MSM. 10th Annual Meeting of the International Behavioural and Neural Genetics Society, Portland, May 5-9, 2008.
- 37 Tanaka, E. and Tamura, Y. "Skeleton-based Fourier Descriptor of Open Curves" APCOM2008, Kokushikan University, September 3, 2008
- 38 田中英希, 田村義保「新しい開曲線記述子を用いた 2 次元部分形状の定量的評価」日本計算機統計学会 第 22 回大会, 秋田文化会館, 平成 20 年 5 月 23 日
- 39 Tsuda K., Ito Y., Miyao A., Hirochika H., Kurata N. Identification and analysis of rice mutants misexpressing NKOX genes in leaves. XX International congress of Genetics, Berlin, Germany, July 12-17, 2008.
- 40 津田勝利, 伊藤幸博, 倉田のり「イネにおける KNOX 遺伝子を介した SAM の維持および葉の分化の研究」国立遺伝学研究所研究会「イネ分子遺伝学の新展開」, 国立遺伝学研究所, 2008 年 12 月 20 日, 三島.
- 41 津田勝利, 伊藤幸博, 宮尾安藝雄, 廣近洋彦, 倉田のり「イネのシュート形成における極長鎖脂肪酸(LCFA)の機能解析」日本育種学会 114 回講演会, 滋賀県立大学, 2008 年 10 月 11 日, 彦根.
- 42 米田典央, 倉田のり, 野々村賢一「減数分裂第一分裂前期に特異的な染色体挙動の観察」イネ遺伝学・分子生物学ワークショップ 2008, 九州大学, 2008 年 7 月 5 日, 福岡.
- 43 石井亜矢子, 西明紀, 城石俊彦, 小出剛 : B6-MSM コンソミックマウス系統を用いた自発活動性の遺伝学的解析 (口頭発表) 第 32 回日本神経科学大会、2009 年 9 月 16-18 日 名古屋
- 44 小森理, 江口真透. 3 次自然スプラインに基づく pAUCBoost, 一般講演, 統計関連学会連合大会, 京都, 日本, 2009.09.06
- 45 江口真透, 小森理. t 値を最大にする判別関数のブースティング, 統計関連学会連合大会, 京都, 日本, 2009.09.07
- 46 Kuriki, S., Harushima, Y., Fujisawa, H., and Kurata, N., Multiplicity Adjustments in Detecting Reproductive Barriers caused by Loci Interactions, BIRS Workshop 09w5040 Random Fields and Stochastic Geometry, 2009.2.26, Banff, Canada
- 47 栗木哲, 前野哲輝, 城石俊彦, 「QTL 解析における影響分析」, Dou Xiaoling, 日本計算機統計学会 第 23 回大会, 福岡, 2009.05.16.

- 48 Sugimoto, H., Takahashi, A., Shiroishi, T., Koide, T.: Molecular and genetic analysis for aggressive behavior using B6-MSM consomic mouse strains. 11th Annual Meeting of the International Behavioural and Neurogenetics Society, Dresden, June 4-8, 2009.
- 49 杉本大樹, 高橋阿貴, 城石俊彦, 小出剛: マウス攻撃行動における遺伝的基盤の解析 (ポスター発表) 第 32 回日本神経科学大会 2009 年 9 月 16-18 日 名古屋
- 50 杉本大樹, 菊水健史, 城石俊彦, 小出剛: Different patterns of ultrasonic vocalizations of wild-derived mouse strains from laboratory strains (ポスター発表) 第 32 回日本分子生物学会年会, 2009 年 12 月 9-12 日 横浜
- 51 田邊彰, 石井亜矢子, 城石俊彦, 高橋阿貴, 小出剛: コンソミックマウス系統を用いた不安様行動の高解像度マッピング (ポスター発表) 第 32 回日本神経科学大会, 2009 年 9 月 16-18 日 名古屋
- 52 田邊彰, 石井亜矢子, 城石俊彦, 高橋阿貴, 小出剛 (ポスター発表) 第 32 回日本分子生物学会年会, 2009 年 12 月 9-12 日 横浜
- 53 Tanaka, E., Tamura, Y., Shiroishi, T., Hosoya, M., "Partial shape analysis of mouse mandible using velocity Fourier descriptor", TRIC Symposium "Statistical and Systems Genetics", National Institute of Genetics, Mishima, Japan, October 22-23, 2009.
- 54 田中英希, 田村義保, 城石俊彦, 細谷理樹, 速度フーリエ記述子を用いた開曲線の解析。2009 年度統計関連学会連合大会。2009 年 9 月 7 日。
- 55 Umemori, J., Uno, T., Yuasa, S., Koide, T.: Decreased expression of myelination associated genes in genetic incompatibility mice, Genic mice. 23rd International Mammalian Genome Conference, La Jolla Torrey Pines, California, November 1-4, 2009.
- 56 梅森十三, 近藤亮太, 宇野毅明, 湯浅茂樹, 小出剛: 遺伝的不適合による神経発達異常 (口頭発表) 第 56 回日本実験動物学会, 2009 年 5 月 14-16 日、大宮
- 57 梅森十三, 宇野毅明, 湯浅茂樹, 小出剛: 遺伝的不適合マウスは脊髄オリゴデンドロサイトの成熟が異常である (口頭発表) 第 32 回日本神経科学大会 2009 年 9 月 16-18 日 名古屋
- 58 梅森十三, 湯浅茂樹, 宇野毅明, 小出剛: 遺伝的不適合マウスではレトロトランスポゾン IAP の発現が上昇している (ワークショップ口頭発表) 第 32 回日本分子生物学会年会, 2009 年 12 月 9-12 日 横浜
- 59 Xiaoling Dou, Satoshi Kuriki, Akiteru Maeno, and Toshihiko Shiroishi, "Influence Analysis in QTL Detection", The 57th Session of the International Statistics Institute, Durban, South Africa, August 16-22, 2009.
- 60 Xiaoling Dou, 栗木哲, 前野哲輝, 高田豊行, 城石俊彦, 「区間マッピング法における影響分析」, 日本計算機統計学会第 23 回シンポジウム, 札幌, 2009.11.08

< 著書等 >

Koide T, Catanesi CI, Nishi A, Shiroishi T, Kasai S, Ikeda K, Takahashi A. Advantage of using wild-derived mouse strains for a variety of pain-related studies: Genetic diversity and new genetic tools. In: Acute Pain (Columbus F, ed), Nova Science Publishers, New York, in press.

小出剛 「マウス実験の基礎知識」 編集 オーム社 2009 年 3 月出版 ISBN: 978-4-274-50217-0 (執筆・編集)

Eguchi, Shinto. Information Divergence Geometry and the Application to Statistical Machine Learning. Eds. Frank Emmert-Streib and Matthias Dehmer 'Information Theory and Statistical Learning', Springer, New York, 10.1007/978-0-387-84816-7_13

