

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〔生物多様性解析〕

プロジェクトディレクター： 城石俊彦

1. 研究目標

複雑系としての生命システムの理解を深めるために、遺伝的分化を遂げた生物個体の多様性の比較解析が有力な方法論として浮かび上がってきている。そのためには、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計測値の違いをゲノム多型に結びつけるための統計手法の融合が必要である。また、研究材料となる生物遺伝資源としては、長い進化的時代を経て多数の遺伝子に塩基置換が蓄積し、多様な表現形質を示す生物系統の利用が有効である。国立遺伝学研究所は、古くから世界各地の自然集団から独自に採集し育成してきた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を保有している。本提案プロジェクトでは、これらの生物系統の持つ表現形質の多様性を可能な限り客観的に数値計測化し、それらのデータを対象にして統計数理研究所が培ってきた統計解析技術と国立情報学研究所の情報処理技術を活用することにより、数値計測化された表現型多様性とゲノム多型を関連づけて遺伝子（ゲノム）機能と遺伝子間ネットワークを解明することを目標とする。

2. 研究概要

生命科学における動植物の代表的なモデル生物であるマウスやイネが示す生物多様性を対象に、統計数理技術と情報処理技術を活用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解析するシステムを開発する。また、大学等の外部機関との共同研究により、個々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。

3. 年度計画

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		←				→

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)

プロジェクト全体の研究目標を参加者が共有し、研究方法などを協議するためのワークショップを開催して情報交換を行った。以下の主な研究成果があった。

- ・ 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。

平成18年度

表現型多様性の数値化のための統計手法の開発を推進し、それらを用いた生物系統からの表現型データの生産と収集を推進した。以下の実質的な成果があった。

- ・ 複数の系統のマウス下顎骨の画像データから二値化画像を作成し、P型フーリエ記述子を用いた数値化と数値データの多変量解析のためのソフトウェアを整備した。また、CT画像からの脂肪組織領域自動抽出のためのデータを収集した。
- ・ マウス自発活動量データと社会行動パターンのトラッキングデータを収集し、系統による活動日周期の違いと社会行動の特徴付けについて統計解析を行った。数値化したマウス社会性の度合いを人間の観察により説明する手法を検討した。
- ・ マウス肥満関連形質とイネの穂形質と分けつ能について多数の系統のデータを収集した。また、(B6 x MSM)F2 マウス交配世代について肥満関連形質データを生産し、QTL解析を実施し複数の

QTL を検出した。

- ・ イネおよびマウスの Affymetrix 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とマウスの系統間での 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 のハイブリシグナル強度の実データを用いたソフトウェアの有効性の検証を行った。
- ・ 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、再現性のある観測値がこの遺伝的モデルと一致することを確認した。また、カイ二乗確率場の最大値分布について、逐次理論による多重性調整 p 値の解析式を導出した。マウス生殖隔離の原因となる X 染色体置換系統雄の繁殖力低下についての QTL 解析を実施し、原因遺伝子が X, 1 番, 1 1 番染色体の複数の遺伝子間不適合であることを明らかにした。

平成19年度(中間評価)

表現型多様性の数値計測手法の確立に努め、生物多様性解明のため多数の生物系統から数値計測された表現型データと実験交配個体の遺伝子型データの整備を継続した。プロジェクトの中間時点でのとりまとめを行った。

- ・ 形態多様性について新規アルゴリズムや3次元データの数値化についても検討した。
- ・ 系統間の実験交配個体の表現型解析と遺伝子型解析を行い、QTL 等の統計解析のための基盤情報を整備した。従来の QTL 解析手法の改良を試み、それによる量的形質の統計遺伝解析を行った。
- ・ マウスのエネルギー代謝や行動特性に関係した各形質間と多数の QTLs の相関構造（因果関係）について SEM 等を用いて抽出する統計手法の開発を試みた。
- ・ 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。
- ・ 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た。

平成20年度

サブテーマ(1)「表現形質の数値計測システムの開発」では、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表現型多様性を客観的に評価する数値計測システムの開発を継続して実施する。特に、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の拡充に注力する。表現形質の数値計測のために開発したプログラムについては、測定精度の向上と機能拡張を計り、表現型抽出力の向上を目指す。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収集を一層拡充させる。特に全ゲノムを対象とした遺伝子型判定の基盤整備に重点を置く。マイクロアレイ解析のためのロバスト推定による発現及び塩基多型の同時検出法については、開発し

たプログラムによる応用的研究をさらに推進する。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」では、各生物系統の特徴抽出を統計学的に行うための研究を一層推進する。特に、形態多様性データの統計的モデリングには新たな手法の導入を試みる。また、動物の移動行動の統計モデリングでは、ベクトル間相互作用など様々なパターン分類を応用したシミュレーションによる数理モデルの検証を行う。生殖的隔離に関与する遺伝子座相互作用の多重性に関しては、調整後の閾値以下の真偽判別を真の相互作用検出に応用する。QTL 解析法の最適化については、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モデルの検討や、カーネル法の適用を検討する。

平成21年度

表現型多様性の客観的な数値計測化システムと遺伝実験データの統計解析手法を確立して、研究コミュニティに広く公開する。また、開発したシステムを利用して表現型多様性を制御する責任遺伝子(群)を探索する。生物多様性に関わる遺伝子間相互作用(ネットワーク)について、本プロジェクトで開発した統計解析手法で解明する。自然集団での生物多様性についての表現型データベースを構築して公開する。以上の確立された統計手法をマウスやイネのさまざまな形質の解析に適用し、その有効性を検証する。5年間のプロジェクトの成果を基にして国内外の研究者に広く呼びかけて公開シンポジウムを企画する。

平成22年度以降の展開

生物形態、エネルギー代謝、植物の生育特性、動物行動などの高次生命機能に係わる遺伝子群や遺伝子パスウェイの解析システムを開発することが最終目標である。この研究の延長線上には、個体レベルでの様々な表現型(ゲノム機能)を遺伝子機能のネットワークとして捉えようという Systems Genetics への発展を想定している。信頼性と再現性の高い遺伝実験データからの統計・情報解析によるデータマイニング(情報抽出)によって、実体に近いシステムズバイオロジーへの展開が期待できる。さらに、自然集団での生物多様性にも目を向けることにより、進化過程を配慮した遺伝子ネットワークのより深い理解に至ることが可能であり、その方向へと研究を展開していく計画である。

4. 研究費の推移

平成17年度実績： 157,820 千円

平成18年度実績： 135,170 千円

平成19年度見込： 118,770 千円

5. 平成19年度の研究推進体制

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 久保貴彦 春島嘉章 田村 勝 高田豊行
前野哲輝 岡(木曾)彩子 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

[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 川崎能典 坂口隆之 種村正美 田村義保 土谷 隆 藤澤洋徳
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 田村 勝 春嶋嘉章 高田豊行
前野哲輝 細谷正樹

[九州大学] 二宮嘉行

6. 平成19年度の研究進捗

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, 阿部)

【研究の目的】

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について検討する。X線 CT 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。

上記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の構築を進める必要がある。我々はマウス全身を対象とした 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、複数個体を用いて撮影した 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。

【平成 19 年度の進捗】

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）を作成した。単一のプロセッサを使用して、1 個体当たり 10 分以内で処理可能であり、手動による抽出結果と比較して概ね 90%以上の抽出精度が達成できることを確認した。

自動解析システム ActScan :

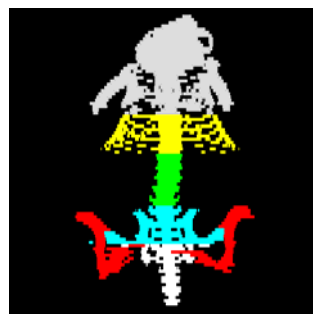
本研究で開発している手法は、ActScan (Active Contour and Tube for CTscan) と呼ばれるプログラムとして実装している。腹膜の形状の再構築および脂肪種の判定に関連して、これまでに以下のような機能を実現している。

1. 各CT画像がマウス全身のどの部位かを判定する機能（部位判定）
2. 各CT画像での腹膜の切断形状を抽出する機能（切断面抽出）
3. 切断形状から腹膜の全体形状を再構成する機能（全体形状再構成）
4. 腹膜形状を用いて内臓脂肪と皮下脂肪を分離する機能（脂肪種判定）
5. 脂肪種の分類精度を評価する機能（精度評価） ※正解データが別途必要

これら一連の機能は、マウス 1 個体分の CT 画像 (70~100 枚) の格納位置を指定することによって、半自動的にパイプライン実行が可能である。各機能の概要は以下の通りである。

(部位判定)

各CT画像がマウス全身のどの部位に該当するかを判定する。現在は、骨格の形状を指標として部位判定を行っている。各CT画像から骨に該当するピクセルを抽出し、それらを統合して骨格の全体像を再構築し、その特徴点を抽出することによって部位判定を行う。右図はある個体の部位判定の結果（色で表示）を骨格に重ねて画像化したものである。



(切断面抽出)

各CT画像では腹膜の切断面は、滑らかな閉曲線となり、尚且つ、その内外で画素値勾配が大きくなるという特徴がある。これら 2 つの特徴を加味したエネルギー関数が (局所的に) 最適化するように、初期閉曲線を少しずつ変形させて抽出したい閉曲線にフィットさせることによって、切断面を抽出する (active contour)。

(全体形状再構成)

各CT画像から得られた切断面を連結させて腹膜の初期全体形状を構成する。体軸方向の滑らかさを反映したエネルギー関数を最適化し、各切断面の形状を微調整することによって、腹膜の全体形状を再構成する(右図)(active tube)。

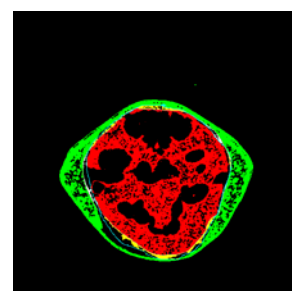


(脂肪種判定)

各CT画像で、脂肪に該当するピクセルが切断面の内外のいずれかに位置するかによって脂肪種を判定し、脂肪種ごとにピクセル数の集計を行う。右図はある個体の腹部の脂肪種判定の結果を画像化したものである(赤:内臓脂肪/緑:皮下脂肪)。

(精度評価)

任意に選択した8個体で脂肪種判定を手作業で行って「正解」を作成し、ActScanによる自動判定結果と比較した。脂肪に該当するピクセルのうち、両判定結果で脂肪種が一致したものの割合を判定精度の指標として用いる。90%以上の判定精度を達成していることを確認した。下表は各個体の部位ごと(腹部と腰部)および全身に関する判定精度である。



個体	BMI	腹部	腰部(腹側)	腰部(尾側)	全身
A	40-	0.98	0.92	0.98	0.96
B	40-	0.92	0.89	0.88	0.90
C	37-40	0.99	0.98	0.90	0.96
D	30-37	0.97	0.96	0.97	0.97
E	30-37	0.98	0.97	0.98	0.98
F	27-30	0.98	0.94	0.94	0.95
G	27-30	0.98	0.96	0.83	0.93
H	20-27	0.99	0.95	0.93	0.95

今後の課題

これまでに、一定の判定精度を達成するに至ったが、精度評価用途の「正解」データの準備が容易ではなく、十分な評価が行えているとは必ずしも言えない。多数の個体を用いて精度の評価を行うためには、本来自動化しようとしている判定作業を手作業で行わざるを得ないという状況に陥る。これを踏まえれば、自動判定の精度向上のみではなく、自動判定した結果をインタラクティブに確認・修正する作業の効率化も重要であると考えられる。判定結果の蓄積が進めば、そこから得られる情報を自動判定プログラムで使用する形態テンプレートに反映させることも可能となる。また、現在は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るが、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必要がある。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。

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】

本研究は遺伝的に離れたマウス系統を使用して、エネルギー代謝に関連する表現型を様々な角度から計測し実験データを収集、肥満ならびにエネルギー代謝を制御する遺伝子群を統計遺伝学的に検出するための実験データ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) を使用し、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*)、および同亜種に属する JF1/Ms (JF1) を使用し、解剖学的、血液生化学的パラメータを計測する。特に、平成 18 年度より新たに開始した JF1 と B6 およびこれらの系統より作出した F2 世代については、CT (Computed Tomography) による脂肪蓄積量の測定をはじめとした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。JF1 は MSM と同様に *Mus musculus molossinus* に属し B6 系統と亜種の関係にあるが、この系統の成立過程が MSM と異なるため、複数の表現型で MSM とは異なる形質が存在すると考えている。特に我々がターゲットとしているエネルギー代謝関連形質については、MSM では検出できない表現型の検出が期待できる。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全ゲノムスキャンを行い、サブテーマ(2)により推進中の QTL 解析による各種表現型の責任遺伝子領域とエピスタシスの検出を行うことを目的としている。

【平成 19 年度の進捗】

平成 19 年度は、JF1, B6 さらには両者から作製した F2 世代の肥満関連形質を中心とした表現型データの拡充を行い、粗い全ゲノムスキャンを使用した既存のソフト (J/qtl: <http://research.jax.org/faculty/churchill/software/Jqtl/index.html>) による QTL 解析を行った。特に、JF1 と B6 系統を使用して作製した F2 世代のエネルギー代謝関連表現型の収集については、時間軸に沿った表現型の検討を行うため、生後より 10 週令から 25 週令における各種表現型の測定を行った (図 1)。マウスは離乳より雄は単飼、雌は最大 5 匹までの群飼に

表現型	
10 週令	【体重・体長・尾長・頭胴長・血糖値・脂肪蓄積量 (CT 計測)】
15 週令	【体重】
20 週令	【体重】
25 週令	【体重・体長・尾長・頭胴長・血糖値・脂肪蓄積量 (CT 計測)・血液生化学・アディポサイトカイン・インスリン】

図 1. 本課題における表現型収集 (JF1, B6, F2)

より維持した。10 週令に達した個体は、体重、体長、血糖値を測定、さらに CT による撮影を行った。10 週令による測定を終えた個体は 5 週間毎に体重を測定、25 週令に達した個体は再び CT による撮影を行った。25 週令の個体は血液を採取し ELISA 法によるインスリン、レプチン、アディ

ポネクチンの測定を行った。図2に表現型の一例として、10週令と25週令における雌雄の体重と頭胴長について、B6、JF1およびF2を比較したものを示す。図に示されるように、JF1はB6と比較した場合、雌雄とも体重が少ない

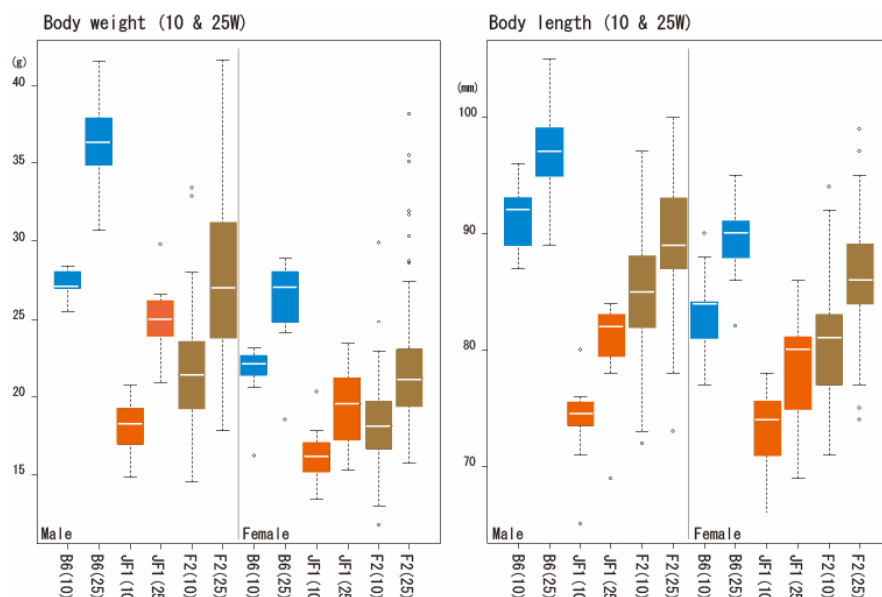


図2. B6、JF1およびF2個体群の体重および頭胴長の分布。10週令と25週令の測定結果を示す。

あるいは頭胴長が短い傾向にある。しかしながら、両者から作製したF2個体群はB6、JF1の特徴を反映した広い分布を示す。特に雌個体群の25週令における体重の分布については、B6、JF1のレンジから外れて増加傾向を示す個体が観察されることも判明した。

体重、頭胴長以外の表現型については、前年度の報告から明らかなように、B6およびJF1の両系統には脂肪蓄積の部位的な差異に基づく明らかな形態的差異が存在し、この違いは、これまでに行われてきた体重や体長の比較では検出することができなかった脂肪蓄積などに関与する遺伝的システムを抽出することが可能になると考えられる。とくに、CT画像については、課題aに示したごとく、F2世代、雌雄各100個体以上のCT画像の取得がすでに終了している。今後は課題aにより測定した各部位の脂肪蓄積についても比較する予定である。

これまでに本課題で収集した様々なデータおよび現在CT画像データから収集中の脂肪量のデータを使用して、肥満ならびにエネルギー代謝を制御する遺伝子群を統計遺伝学的に検出するため、本年度はSNPマーカーを指標とした全ゲノムマッピングを行う前段階の解析として、全ゲノムを

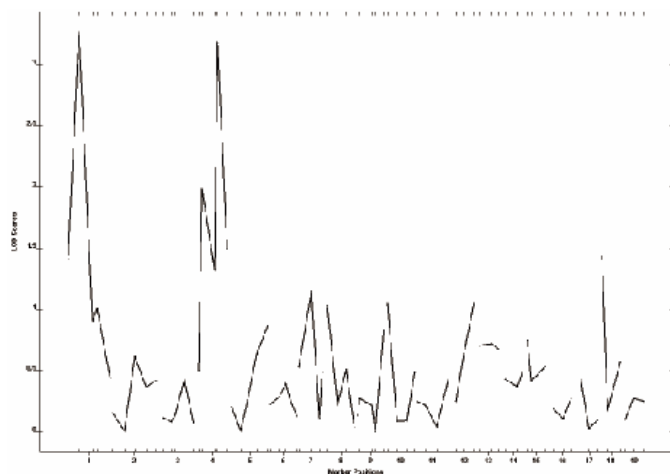


図3. 雌個体の25週令時の体重に関するQTL

対象としたマイクロサテライトマーカーによる粗いジェノタイピングを行った。前年度においてB6-JF1については約80種のマイクロサテライトマーカーが使用可能であることを確認しているため、ジェノタイピングにはこれらを使用した。タイピング結果はJ/qt1により各表現型を規定する染色体の位置情報の推定を行った。染色体1番と4番にLODの高いQTLが存在することが示唆された(図3)。さらに複

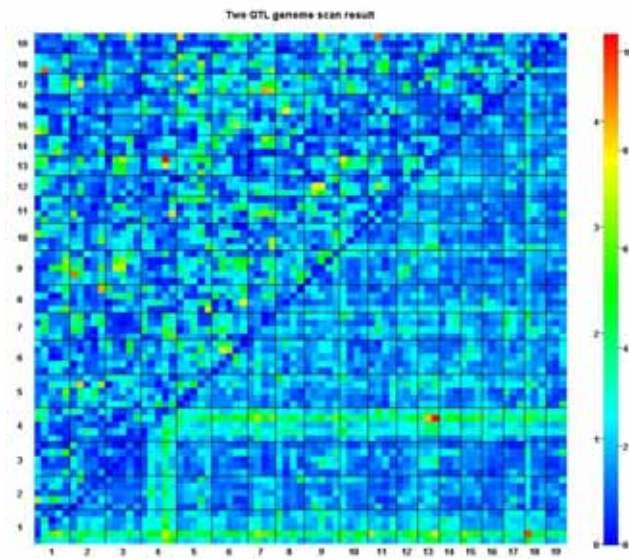


図 4. Two QTL genome scan の結果

数の QTL が関与する表現型変化についても解析するため、Two QTL genome scan についても行った(図 4)。左上にエピスタシス LOD, 右下に Additive LOD を示す。B6 と JF1 から作製された F2 群の雌の 25 週令の体重については、図 3 に示した高い LOD をしめすそれぞれの QTL について、染色体 4 番は 13 番, 1 番は 18 番に相互作用が存在することが示唆された。これらの QTL が相加的效果以外にも相乗的效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

必要がある。そこで、すべての測定項目について同様の解析を行い、効果のある QTL とそれらの相互作用を検証する。今後は、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、サブテーマ 2 と連携して、各形質間の相関解析を実施する。

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(倉田, 春島, 堀内, 永口, 久保)

【研究目的】

本課題では、多様な生育特性を示すイネ栽培系統や近縁野生イネの形質多様性と、塩基配列レベルの遺伝子構造多型および遺伝子発現多様性との相関を抽出し、個々の形質に関与する遺伝的要因を発見するため、統計的な処理法や解析集団、解析形質の適正な選択法の確立を目指している。このようなアソシエーション解析においては、遺伝的多型としての大量の SNP (single nucleotide polymorphism) 検出が必要であるが、イネ配列情報についての SFP (single-feature polymorphism) のデータは未だわずかしか収集されておらず、ほとんどが栽培イネ系統間の SNP だけである。野生イネ系統など、遺伝的多様性の高い系統間での遺伝子発現の見積もりは、SFP の影響による発現量の補正を余儀なくされると予想される。よって、この問題に関しては、別課題「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」の成果を利用する。野生イネの多様な形質の解析については、どの形質、どの系統群に焦点を絞るかの判定を行い、絞り込んだ系統での形質データ取得と遺伝子発現データの解析を行う。イネの形質として穂の器官の形状および生育特性等を中心に系統毎の形質を計測し、もう一方では全ゲノムにおける系統間での SNP データの抽出と、遺伝子発現データの取得と補正を行い、真の発現量を見

積もる。データ取得までは統計数理研究所との共同体制は必要ではないが、20年度からはアソシエーション解析等の統計的処理を融合研究として取り組む必要がある。

【平成19年度の進捗】

18年度までに、16 species の野生イネコアコレクション46系統の18形質を計測したデータを用いて、各形質データ間での相関を解析し、種間の相関関係と種内の相関を求めた。これらの結果、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。この結果、特に種子稔性、葯長などは、分類学上で言われている多年生と一年生の分類特性と相関が高い事がわかった。19年度はさらに各系統間で多様性を示す形質について、アソシエーション解析に向けたデータ蓄積を行った。

(1) まず多年生と一年生の2群から14系統を選別し、各系統3個体ずつから3-6のサンプルリングを行い、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。

(2) 個別の形質として図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。実際の形質写真を図3に示した。

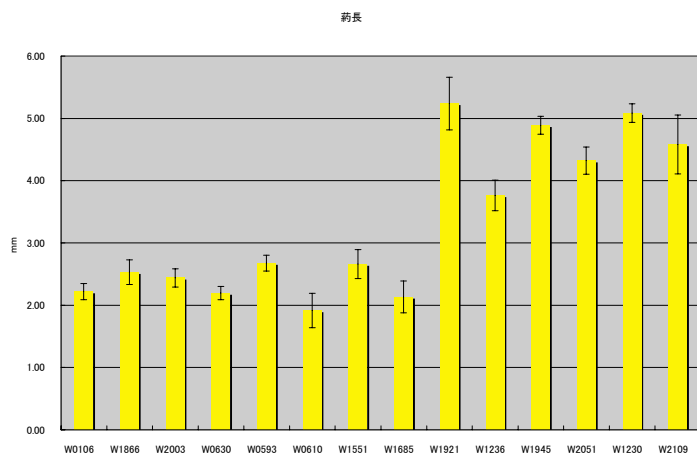


図1, *O. rufipogon* 14系統における葯長 (mm)。各系統XX個体からのYYサンプル

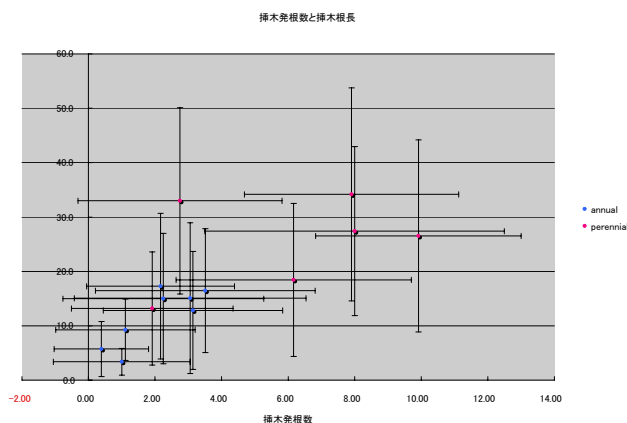


図2, *O. rufipogon*における出穂後茎からの発根数および発根長の系統間による変異。
annual: 1年生系統。
perennial: 多年生系統。

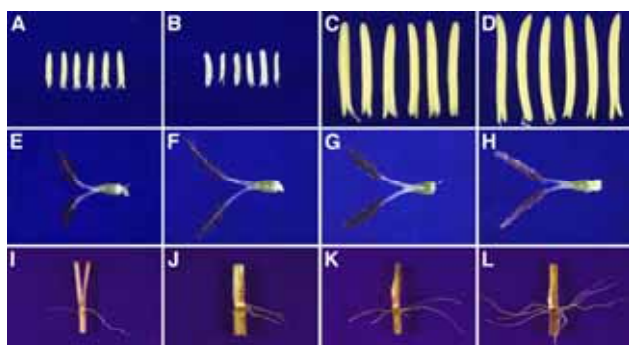


図3. *O. rufipogon* の葇 (A-D), 雌蕊 (E-H), 挿木 (I-L)。1年生型のW0106 (A, E, I), W0610 (B, F, J)。多年生型のW1945 (C, G, K), W1230 (D, H, L)。

1年生型の葇は多年生型の葇にくらべて短い。柱頭や子房の大きさは1年生型と多年生型でそれほど差がない。多年生型では挿木からの発根数が多く、各根はよく伸長する。

(3) イネの生産力に直結する種子数、穂型についての計測値を表1に示す。一次枝梗, 二次分枝, 小穂数などの個別形質が系統により異なっており, いくつかのタイプに分かれると考えられる。しかし, これら一次枝梗の数, 一次枝梗あたりの穎花数およびその分散は, 一年生多年生の分類特性とは相関が低いと思われた。今後はさらに穂長, 粒重など分解可能な特性を加え, それぞれの特性について系統間の主成分分析を行う。それぞれの主成分あるいは主成分間のコンビネーションがどのSNPと相関するかを解析して行く。一次枝梗, 二次分枝, 小穂数などの特性がどのような遺伝的要因に制御されるのか, 生産性への寄与率はどの特性がどのように大きいのか, どのように相関して多様性が出現するのか, 興味ある問題である。

表1 *O. rufipogon* の種における1次枝梗着生節位別の2次分枝数

Rank	Accession	Origin	Ann/Per	1次枝梗着生節位(穂首から)												
				1節	2節	3節	4節	5節	6節	7節	8節	9節	10節	11節	12節	
-	Nipponbare	-	-	7	8	8	8	8	8	8	7	7	7	6	6	
-	Kasalath	-	-	10	11	11	11	11	11	11	10	9	10	9		
(A)	1	W0106	India	Annual	4	5	5	5	5							
	1	W1866	Thailand	Annual	5	6	6	6	6	5	5					
	1	W2003	India	Annual	5	6	5	6	5	5						
	2	W0630	Burma	Annual	4	5	5	5	4							
	2	W1807	Sri Lanka	Annual	6	6	6	6	6	6	6	6				
	3	W0137	India	Annual	6	6	7	7	7	7	7	7				
	3	W0593	Malaya	Annual	退	5	6	7	7	7	7					
(B)	3	W0610	Burma	Annual	6	6	6	7	6	6	7	7	6	6		
	3	W1551	Thailand	Annual	5	5	6	6	6	6	5					
	3	W1681	India	Annual	4	4	5	5	5	6	6					
	3	W1685	India	Annual	退	4	5	5	5	5						
	3	W1690	Thailand	Annual	退	5	6	5	6	5	5	5	4			
	3	W1852	Thailand	Annual	6	6	6	6	6	7	7	6	6	6		
	3	W1865	Thailand	Annual	5	6	6	6	6	6	5	5				
	1	W0120	India	Perennial	退	5	6	5	6	6	5					
	1	W1294	Philippines	Perennial	5	6	6	6	6	6	6					
	1	W1921	Thailand	Perennial	6	7	7	7	7	7	6	6	5	4	3	2
	2	W1236	Australian New Guinea	Perennial	7	7	7	7	7	7	6	6				
(C)	2	W1945	China	Perennial	6	6	7	6	6	6	5					
	2	W2051	Bangladesh	Perennial	6	5	5	5	5	4	4	4				
	2	W2078	Australia	Perennial	5	5	5	5	5	4						
(D)	3	W1230	Dutch New Guinea	Perennial	7	7	6	6	6	6	5	5	5	4		
	3	W1235	Dutch New Guinea	Perennial	4	6	6	6	5	5	5	5	4	4		
	3	W1238	Dutch New Guinea	Perennial	5	5	5	6	6	5	5	4				
	3	W1666	India	Perennial	6	6	5	6	6	6	6	5	5	5	4	
	3	W2109	Australia	Perennial	退	5	5	5	5	4	4	3	2			
	3	W1239	Dutch New Guinea	-	5	6	6	6	6	5	5	5	5	4		

退: 1節(穂首節)の1次枝梗が退化している

(4) 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。今後, 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ていく。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できれば、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する事により、目的の遺伝因子により早く近づける可能性があると考えている。

(5) 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個別別に収集し、アレイ解析を進行中である。

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収集と解析 (藤澤, 江口, 坂口, 倉田, 春島, 堀内, 城石, 高田)

【研究目的】

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マイクロアレイであるが、ゲノム全体に渡る塩基配列の差をも見積もる事ができ、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、既存の研究は、塩基配列の多型の検出にのみ着目するか、塩基配列の多型を無視し発現量の推定を行うものが殆どである。GeneChip は一般に 1 つの遺伝子の発現量を推定するのに 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 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ものである。本課題では、これら PM11 個あるいは PM+MM 22 個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事が出来る手法を開発する事を目的とする。

ある遺伝子の塩基配列が系統や種間で同じで発現量だけが異なる時、PM プローブの対数シグナル値の系統間差がセット内で一定値であることに着目した。セット内で対数シグナル値の系統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、対数シグナル値の系統間差が 0 より有意に異なる時に系統間で発現量の多型を検出できる。はずれ値によるマスク効果で尤度比検定では、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した。この手法を SNEP (Simultaneous detection of nucleotide and expression polymorphisms) として、統計言語 R によるソフトウェアを開発した。本年度は、SNEP 法を実際のイネやマウスの GeneChip 開発の基となった品種 (イネ: 日本晴, マウス: B6) と塩基配列が既知の別の系統 (イネ: 93-11, マウス: MSM/Ms) の mRNA ハイブリダイゼーションのシグナルに適用し、開発した手法で、プローブ標的の塩基配列多型 (SFP: single feature polymorphism) 検出の効率の検証を行った。

【平成19年度の進捗】

はじめに、SNEP法のSFP検出効率の特徴を調べるため、実際の塩基配列データを基に合成データを作製した。MMプローブは中央の一塩基を相補塩基に置換しているため、PMプローブと完全にマッチする遺伝子をハイブリダイゼーションした時は、MMプローブのシグナルは理想的なSFPのシグナルとなると考えた。

日本晴と93-11間で塩基配列多型が無く、他の遺伝子とクロスハイブリする事の無いsingle copyのプローブのみで構成される10,421セットを選び、セット中のプローブをランダムに1個選び93-11のPMシグナル値をMMのものと同じデータで合成した。この合成データにはセット中1個のSFPが存在する事になる。日本晴と93-11の遺伝子発現量(セット内の対数シグナルの中央値)

で場合分けし、閾値 $\alpha = 10^{-2}$ 時のSFPの検出のSensitivityとFalse positive ratio (FPR)を示す(Table 1)。遺伝の発現量が高い時は効率よくSFPを検出するが、発現レベルの低い遺伝子ではSFPの検出力が下がる事を示している。日本晴93-11共に高発現する

遺伝子について、セット中にSFPが複数存在する合成データのSFPの検出率を調べた(Table 2)。セット中にSFPが4個以上あると検出率は急に落ちる事が分かった。

イネの実際のデータにSNEP法を適用し、閾値 $\alpha = 10^{-2}$ 時の日本晴と93-11の塩基配列より予測されるSFPの検出率は、合成データの時と同じように高発現する遺伝子については48.6%と良いが、発現レベルの低い遺伝子については7.1%と悪かった。高発現遺伝子でもSFPがセット中1個存在する時の合成データに比べ検出率は低い、これはプローブ標的の塩基配列の多型位置と、セット中のSFPの数のためである。下にセット中のSFPの数とSFPの検出効率を示す(Table 3)。セット中に1個SFPが予想される時のSensitivityは71.4%と高いがSFPの数が増えるに従い検出率は下がる。セット中にSFPの数が1個予想される時の塩基多型のプローブ中の位置とSensitivityとの関係を調べた所、中央の3塩基に多型が有る時のSensitivityは94%と合成データと良い一致を見た。これはMMプローブシグナルが理想的なSFPシグナルとなる事を示している。SFPを使ったLDマッピングやeQTLを調べる時、塩基配列差がシグナル強度にあまり影響を与え

Table 1. The performance of SFP detection for the synthetic data.

$x_{med}^{(j)} \setminus x_{med}^{(i)}$	≤ 2	2-2.5	2.5-3	$3 \leq$
# of probe sets				
≤ 2	1010	172	2	3
2-2.5	303	3521	174	14
2.5-3	8	493	2003	205
$3 \leq$	4	39	450	2020
Judged ^a				
≤ 2	271	61	5	7
2-2.5	155	2,020	211	21
2.5-3	2	452	2,205	294
$3 \leq$	0	20	497	2,466
Sensitivity				
≤ 2	10.7%	18.0%	100.0%	66.7%
2-2.5	18.8%	40.6%	78.2%	78.6%
2.5-3	0.0%	68.8%	91.9%	94.6%
$3 \leq$	0.0%	33.3%	93.1%	98.0%
FPR				
≤ 2	60.1%	49.2%	60.0%	71.4%
2-2.5	63.2%	31.6%	35.5%	47.6%
2.5-3	100.0%	25.0%	16.6%	34.0%
$3 \leq$	- %	35.0%	15.7%	19.7%

$x_{med}^{(i)}, x_{med}^{(j)}$: median-log-intensities for Nipponbare and 93-11.
^a The number of probes judged as SFP.

Table 2. The performance of SFP detection is effected by the number of SFP probes in a probe set for the highly expressed genes in the synthetic data.

	# of SFP probes in a probe set					
	0	1	2	3	4	5
Expected ^a	0	4678	9356	14034	18712	23390
Judged ^a	1328	5462	9518	12957	9692	2888
Sensitivity		94.8%	93.9%	89.2%	51.1%	12.3%
FPR		18.8%	7.7%	3.4%	1.4%	0.5%

^a The number of probes expected or judged as SFP.

Table 3. The performance of SFP detection is effected by the number of SFP probes in a probe set for the highly expressed genes in the raw rice mRNA data.

	# of SFP probes in a probe set											total	
	0	1	2	3	4	5	6	7	8	9	10		11
# of probe sets	4659	1052	513	202	94	54	25	24	8	9	7	40	6687
Expected ^a	0	1052	1026	606	376	270	150	168	64	81	70	440	4303
Judged ^a	1325	967	752	369	169	81	25	19	3	2	2	28	3742
Sensitivity		71.4%	66.4%	57.1%	43.4%	28.2%	15.3%	10.1%	4.7%	2.5%	2.9%	6.4%	48.6%
FPR		22.3%	9.4%	6.2%	3.4%	6.2%	8.0%	10.5%	0.0%	0.0%	0.0%	0.0%	44.1%

^a The number of probes expected or judged as SFP.

ず SFP 検出の Sensitivity が低いのはあまり問題ではないが、誤って SFP と判定する影響は大きい。そこで、閾値を変え、SFP 検出効率を調べたのが図 1 である。閾値を厳しくすると、Sensitivity も下がるが、急速に FPR が下がり、 $\alpha = 10^{-8}$ より厳しくすると、Sensitivity と FPR は同じように下がる。従って、閾値 $\alpha = 10^{-8}$ が SFP 検出に最適な値と考えられ、検出率 34%, FPR 14%であった。この結果は今まで報告された中でも良い効率で SFP を検出できる事を示している。

マウスの B6 と MSM/Ms の mRNA ついてもイネと同様に塩基配列の差から SFP と予想されるプローブの SNEP 法による SFP の検出効率を調べた。MM プローブシグナルは理想的な SFP シグナルとなる。Rice GeneChip のプローブ配列のプラットフォーム系統である日本晴の mRNA のハイブリダイゼーションシグナルは、セットの対数シグナルの中央値が 2.5 以上の時、PM プローブと MM プローブの 74%は有意に差がある。Mouse GeneChip のプラットフォーム系統 B6 の mRNA のハイブリダイゼーションシグナルは、セットの対数シグナルの中央値が 2.6 以上の時、イネと同程度に PM と MM の差を示す。そこで、B6 と MSM/Ms とともにセットの対数シグナルの中央値が 2.6 以上の遺伝子 3,126 個について SFP の検出を SNEP 法で行った(図 1)。マウスはイネの 6 倍のゲノムサイズを持つにもかかわらず、イネとほぼ同じ結果 ($\alpha = 10^{-8}$ で Sensitivity 30%, FPR 14%) を得た。これは、SNEP 法がより複雑なゲノムを持つものに対しても適応可能である事を示している。

小さなゲノムサイズの種では、SFP の検出にゲノム DNA をランダムプライマーで直接増幅ラベルしてハイブリダイゼー

ションを行った報告がある。イネでもゲノムハイブリダイゼーションにより SFP を検出できる条件を検討した。Affymetrix のシステムでは実験条件は殆ど変更する事が出来ないが、ハイブリダイゼーション時に添加するラベル DNA 量は変更出来る。MM プローブシグナルが仮想 SFP になる事から、プラットフォーム系統である日本

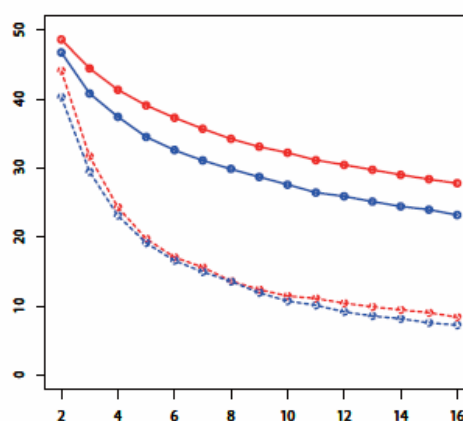


図 1. The sensitivity and FPR of SFP detection among the highly expressed genes for the raw data varies with the significance level. The red and blue lines correspond to rice and mice, respectively. The x-axis is the minus log₁₀ value of significance level. The solid and dashed lines correspond to the sensitivity and FPR, respectively.

晴ゲノム DNA を用い、ラベル DNA の添付量を変えて、PM プローブシグナルと MM プローブシグナルの差が最も開く条件を検討した。添加ラベル DNA 量を 1, 5, 40 μ g とした時、シグナル強度が PM > MM となるプローブの割合は 73.2%, 79.1%, 70.3%と 5 μ g を添加した時が最も良好な結果を得た。

SNEP 法はゲノムの大きさに関わり無く確かな SFP を検出し、SFP の影響を考慮し遺伝子の発現量差を見積もる事が出来る現時点で最も良い方法だが、高発現遺伝子の SFP しか検出できない、遺伝的変異の多い (SFP がセット中に多数含まれる) 時は、SFP の検出力も発現量差の推定も難しいという欠点がある。今回のイネでは対象とする遺伝子の 37%で、マウスでは 6.9%でしか SFP を検出できなかった。

イネについてはより多くの SFP を検出するためゲノムハイブリダイゼーションによる検出の最適化が求められる。来年度は、今年度の結果を踏まえ多数の遺伝的変異を含むような遺伝的に遠縁の種でも SFP を検出できるよう、SFP 検出方法の比較検討を行う。更に、遺伝的変異の多い遺伝子についての発現量を見積もる方法の比較検討を行う。マウスについても、SNEP 法を用いて、異なる亜種系統間のヘテロ接合体、さらに、ゲノム情報未知の野生由来マウス系統のマイクロアレイ解析に応用するための最適化法を検討する。

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

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(田村, 江口, 城石, Mollah, 田中, 細谷)

課題 a-1. フーリエ記述子を用いたマウス形態(下顎骨等)多様性データの数値計測システム(城石, 細谷, 田村, 田中)

【研究目的】

動植物の育種学、遺伝学、古生物学、進化生物学などでは、生物の形の時間的変化や、細かい環境条件による多様性、地域による多様性、形態と遺伝子との対応などについて、図形を数値表現し、それを用いて統計解析を行いたい場合がある。閉曲線であらわされた 2 次元図形の統計解析には、楕円フーリエ記述子などが多く用いられ、有用性が示されている。閉曲線で表された図形の一部の特徴だけを比較して統計解析を行うためには、統計解析にふさわしい開曲線の数値表現方法が必要である。線形性を仮定している多変量解析を用いるためには、局所的な形の単純な変動に対して、データがなるべく線形にならぶことが望ましい。

【平成 19 年度の成果】

平成 19 年度は、統計解析を行うためにふさわしい開曲線の記述方法を提案し、それをネズミの下顎骨のデータに適用した。ネズミの下顎骨の形には、亜種による特徴の違いがよく表われることは、よく知られている。中でも亜種により形の多様性がよく表われる部分である、筋突起(上側の突起)と下顎突起(下側の突起)の形と、遺伝子の働きとの関係を調べることで、生物の形が遺伝子によりどのように決められるのかを解明することを目的とした。

データは、B6 と MSM という 2 つの亜種、その 2 つの亜種を掛け合わせた F1 世代、F1 世代どうしをかけあわせた F2 世代、からの合計 266 個体のネズミの下顎骨の画像である (図 1)。ネズミの下顎骨は平坦で白いので、黒い床の上に置いて真上から写真にとり、画像を二値化すれば、エッジライン(アウトライン)を得られる。この二値化画像から、筋突起と下顎突起の部分のみを切りだし、それぞれのエッジラインからフリーマンチェーンコードで、点の座標の系列を取りだした。

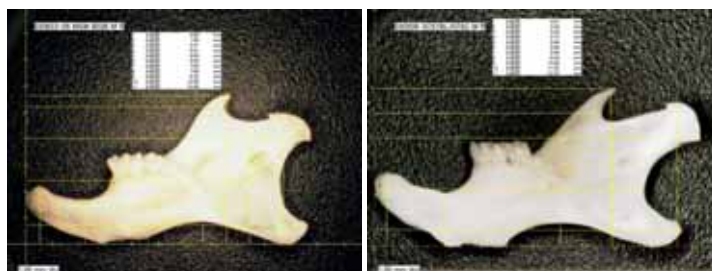


図 1. ネズミの下顎骨。右がMSM (日本産亜種由来系統), 左がB6 (西ヨーロッパ産亜種由来系統)

方法の説明

これまでに統計解析に用いられ一定の有用性が示されている開曲線の記述方法としては、P型フーリエ記述子がある。曲線を長さの等しい辺から成る折れ線で近似したとき、曲線は方向の系列と考えることができる。そして、方向(大きさ1の複素数で表す)の系列のフーリエ変換がP型フーリエ記述子である。方向の系列に不連続点(方向が急に変わる部分)がある場合、特にそこでの方向の変化の度合いが大きい場合、P型フーリエでは、高次周波数成分の寄与が大きくなり、低い次数に情報を縮約できない。そのような場合でも、より低い次数に情報を縮約できる開曲線の新たな記述方法として以下のようなアルゴリズムを考えた(図2)。まず、P型フーリエ記述子と同様に、曲線は等間隔に隔たった順序付けられたn個の点の座標の系列であるとする。本手法では、曲線と交わずに曲線上の各々の点と接することのできる最大の円の半径を、その曲線上の点に対応付けていく。曲線上のある点からみた接円は、曲線を境に両側にそれぞれあるが、それらを区別する。したがって、曲線上の各点に2つの最大接円の半径の情報が対応付けられ、曲線を2つの最大接円の半径の系列 $\{r_i, i=1, \dots, n\}$ として表すことができる。さらに半径を曲線の全長で規格化した後、半径 r_i の変換 $s_i = \exp(-10 \times r_i)$ の系列をフーリエ変換する。これを提案する開曲線の形の記述子とする。 $\{s_i, i=1, \dots, n\}$ は曲線の方向が急に化する部分でも不連続にならず、その他の箇所でも不連続点が生じたとしても変化の度合いが小さいので、多くの場合P型フーリエ記述子よりも低い次数のフーリエ係数で同じ曲線の形を表現できることが分かった。

```

Represent open curve by a set of  $N$  ordered equally-spaced points.
 $\Omega = \{p_1, \dots, p_N; \|p_{i+1} - p_i\| = \|p_i - p_{i-1}\|, i = 2, \dots, N - 1\}$  (ordered point set)
 $\omega = \{1, \dots, N\}$  (index set)
 $r_i^a = 0, s_i^a = 0$  for  $a = 0, 1$  (rightside or leftside), for  $i = 1, \dots, N$ 
For any triple  $\{i, j, k \in \omega; i < j < k\}$ 
  OUTPUT= 1
   $b = (p_j - p_i) \cdot (p_k - p_j)$  (dot product)
  If  $b > 0$  then  $a = 0$ . (rightside)
  If  $b < 0$  then  $a = 1$ , (leftside)
  If  $b = 0$  then continue.
  Calculate circumcenter  $c$  and circumradius  $r = \|p_i - c\|$ 
  For  $l \in \omega, l \neq i, j, k$ 
    If  $\|p_l - c\| < r$ , then OUTPUT= 0, break.
  End For
  If OUTPUT= 1 then
     $s = \exp(-Ar)$ , ( $A$  should be set beforehand.)
    If  $r_i^a < r$  then  $r_i^a = r, s_i^a = s$ .
    If  $r_j^a < r$  then  $r_j^a = r, s_j^a = s$ .
    If  $r_k^a < r$  then  $r_k^a = r, s_k^a = s$ .
  End If
End For.
Output Fourier transform of  $s_i^0, s_i^1, i = 1, \dots, N$ 

```

図2. 提案手法のアルゴリズム

解析結果

B6とMSMという2つの亜種のネズミ（B6が16個体，MSMが20個体）の下顎骨のアウトラインから，下側の突起と上側の突起の部分だけを切り出した（図3）。図3の下の段は，曲線上の点と，そこに対応付けられた最大接円の中心とを直線で結んだものである。最大接円の中心どうしを結んだ曲線を，図形の局所対称軸，局所中心軸，閉曲線で表される図形の場合，リッジカーブ，スケルトンなどとよぶことがある。この図から，最大接円の半径が意味するのは，曲線から法線方向に伸ばした直線が局所対称軸にぶつかるまでの距離であることが分かる。図4は $\{s_i, i = 1, \dots, n\}$ のグラフであるが，このグラフの形は2種類の亜種のそれぞれの形態差をとらえているように見える（図4）。この $\{s_i, i = 1, \dots, n\}$ のグラフをフーリエ変換を用いて統計解析を行った。

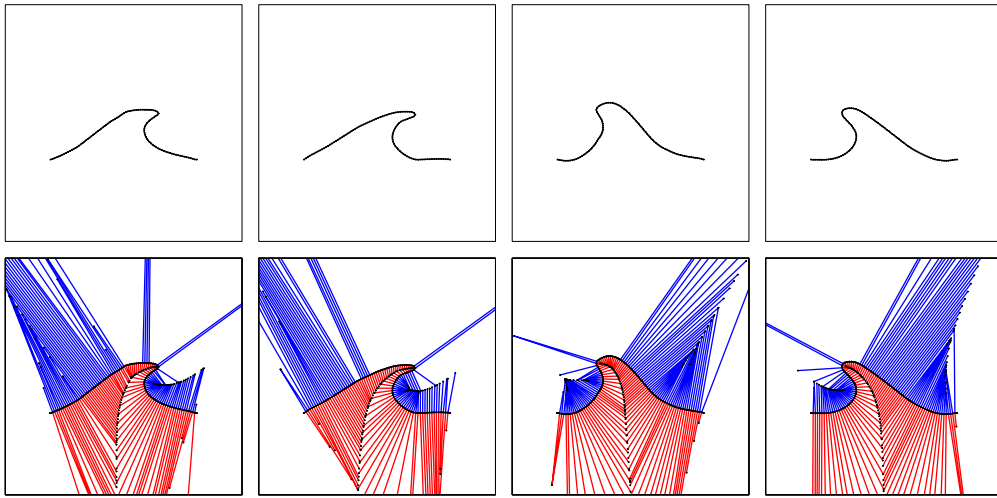


図3. 左からMSMの上側の突起, B6の上側の突起, MSMの下側の突起, B6の上側の突起

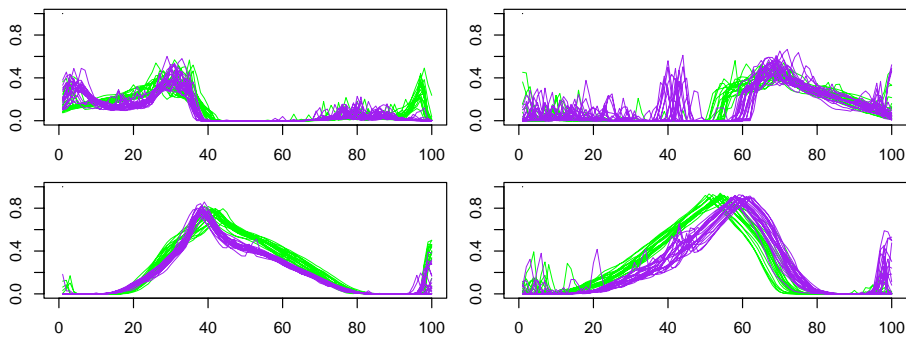


図4. $\{s_i, i=1, \dots, n\}$ のグラフ。左の2図が下側の突起, 右の2図が上側の突起。
 図2 青で表された側の系列が上の段, 赤で表された側の系列が下の段。紫がMSM, 緑がB6。

B6, MSM, 2つの亜種を掛け合わせたF1世代, F1世代どうしをかけあわせたF2世代, からの合計266個体について, 本提案手法で数値化し, 主成分分析をおこなった(図5の左と中央の列)。主成分分析に用いたフーリエ係数の最大次数は, P型フーリエでは32次まで(128次元ベクトル), 提案手法のフーリエ表現では16次まで($16 \times 4 + 2 = 66$ 次元ベクトル)。P型フーリエ記述子を用いた場合の上側突起についての主成分分析結果では, データが湾曲した形に分布している(図5右上), 本提案手法を用いた場合ではそうはならなかった(図5右下)。さらに, B6とMSMのデータのみを用いて正準判別分析を行った(図5右列)。正準判別には, 寄与率で切って, 筋突起のP型フーリエではPC10まで, 下顎突起のP型フーリエではPC13まで, 筋突起の提案手法ではPC8まで, 下顎突起の提案手法ではPC12までの主成分を用いた。さらにこのときの正準係数を用いてF1世代, F2世代の全データの, 上側, 下側突起についての正準得点を求めた。下側突起の正準得点を縦軸に, 上側突起の正準得点を横軸にとり, F1, F2世代も含めた全データをプロットしたのが図5の右の列である。F2世代はB6とMSMの間の領域に分布し,

F 2 世代の下側の突起の形の特徴はMSMから、上側の突起の形の特徴はB 6 から受け継いでいることがわかる。

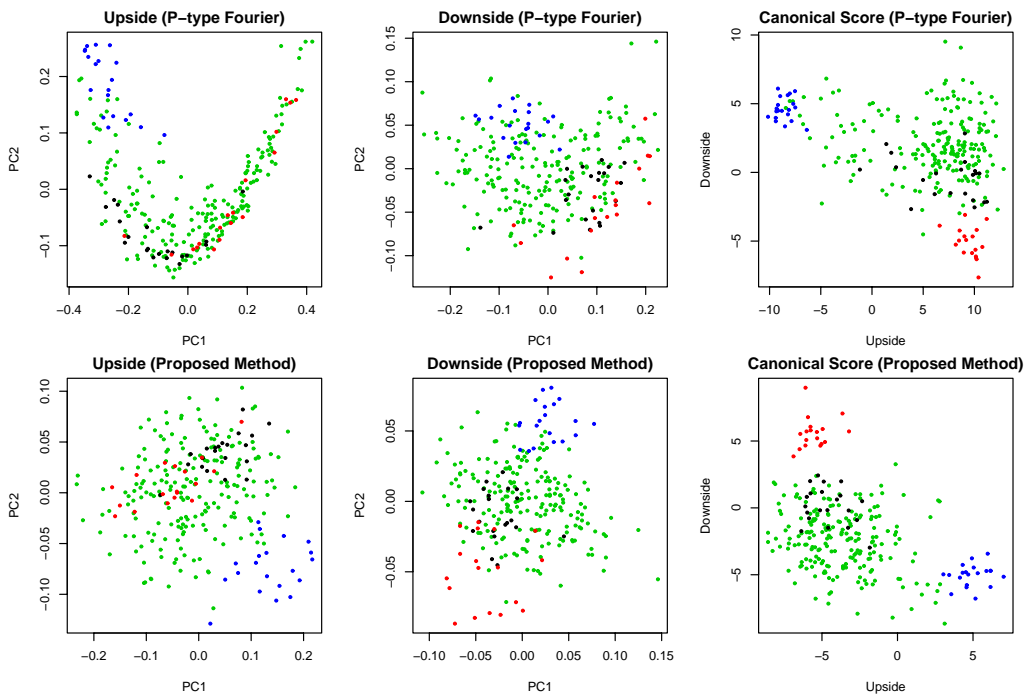


図5 P型フーリエ記述子を用いた場合(左上, 中央上)と, 提案手法のフーリエ表現を用いた場合(左下, 中央下)の主成分分析結果。左2図は筋突起, 中央2図は下顎突起についての結果。右2図は, B6とMSMのみで(上位の主成分を用いて)正準判別を行い, そのときの正準係数を用いてF1, F2のデータも合わせて, 下顎突起と筋突起についての正準得点を求めプロットしたもの。赤はB6, 青はMSM, 黒はF1, 緑はF2。

まとめ

方向が急に大きく変わる部分を含む曲線で, その部分の曲線上における位置に大きなばらつきがあるようなデータについて, P型フーリエ記述子で表わした場合, データは湾曲した形で分布することがある。このような場合, 線形性を仮定した多変量解析には, 本提案手法の方がふさわしいと考える。

課題 a-2. マウスの形態のパターン解析 (Mollah, 江口, 田村, 城石, 細谷)

【研究目的】

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連関を探る問題に統計的機械学習の方法論を開発する。

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す。特に田村グループが行った P 型フーリエ子の特微量としてパター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。

【平成19年度の成果】

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。この従来方途の比較を行い共同論文の作成中である。私たちが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。

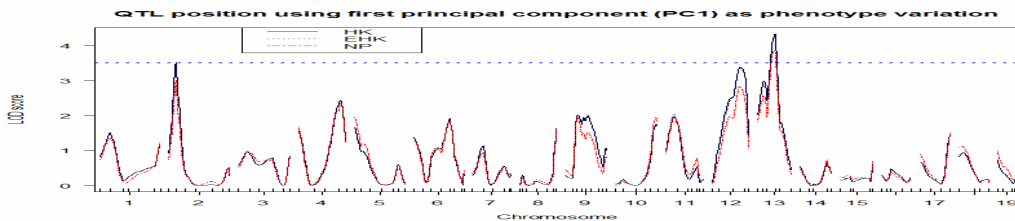


図 A. 第1主成分スコアによるロッドスコア (オス・メス)

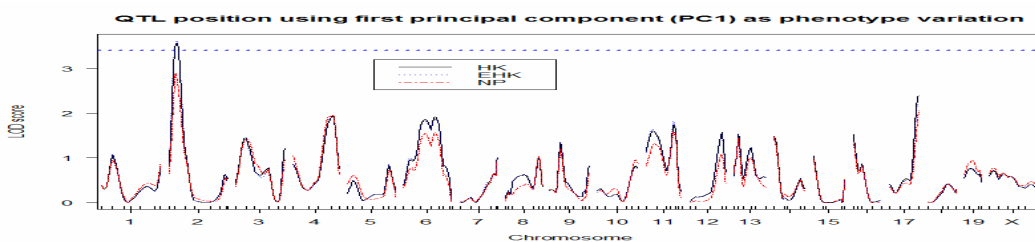


図 B. 第1主成分スコアによるによるロッドスコア (メス)

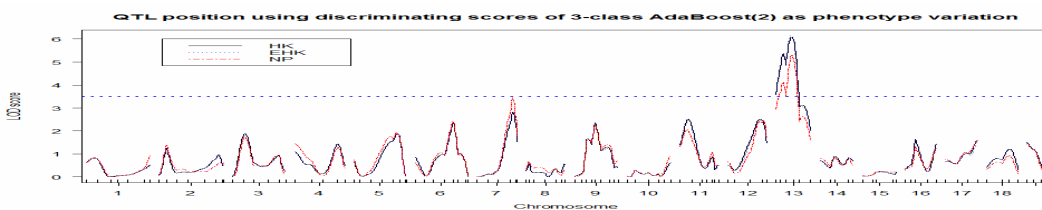
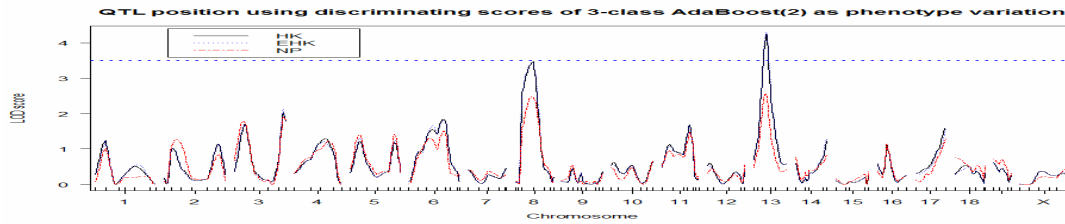


図 C. アダブーストスコアによるロッドスコア (オス・メス)



図D. アダブーストスコアーロードスコアー (メス)

QTL 解析を統計的性能について研究している。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。

課題 b. マウス社会行動の統計的モデリング (種村, 土谷, 小出, 梅森)

【研究目的】

集団中で連続的な多様性を示す高次機能の理解は、体質や性格などの個人差をもたらす機構の解明という点からも現在注目されている研究分野の一つである。その高次機能の中で、行動表現型は形としては残らないため、その計測記録システムおよび解析には困難が伴う。また、行動は時間軸に沿った発現をするため、その定量的な評価は難しい。本研究では、マウス行動を定量的に計測し、得られたデータを数理的な手法を用いて解析して、将来の研究において発展性のある行動成分を抽出し、その形質の背後にある特性をモデル化により明らかにすることを目的とする。

【平成19年度の成果】

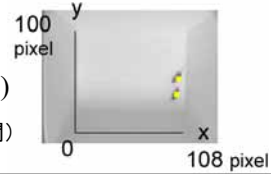
[小出] 動物は社会構造の中で生活しており、その行動における社会性は生存戦略の上で重要な要素となっている。マウスはグループで生活し、雄はそのテリトリーも持つなど顕著な社会行動を示すことから、動物の社会性を解析する上では優れた実験材料である。これまで、B6 および MSM の 2 近交系統では顕著に社会行動が異なることが分かっていたが、その定量化が難しく解析を困難にしていた。そこで、2 個体のマウスを同時にオープンフィールドにおいた際の 10 分間の社会行動をビデオ記録し、その 2 個体の時系列トラッキングデータを抽出した。

同じ系統の同性の2個体をオープンフィールドに入れ
10分間自由に探索させる。
装置上方からビデオで撮影し、画像解析を行う。

解析ソフトで抽出されるもの

- ・合計接触時間、接触回数
- ・2個体の位置(x,y)情報 (3/1sec)

(接触=2個体が3pixel以下に近づいた時間)



これをもとに、合計接触時間と接触回数の尺度について解析した結果、社会的接触時間、回数だけでは社会的行動の定量化は難しく、新たな手法による社会行動の評価が必要であることが分かった。

そこで、得られた動画データと時系列の座標軸データをもとに、統計・数理的な解析を進めた。

[土谷]本年度は、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

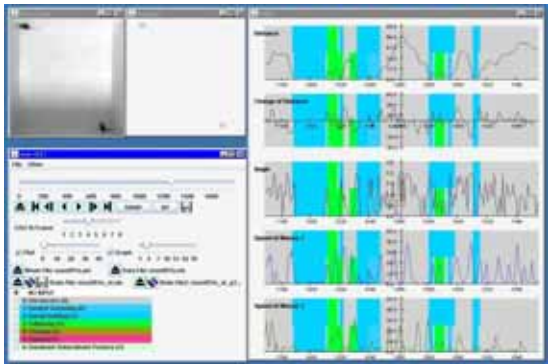


図 1

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。

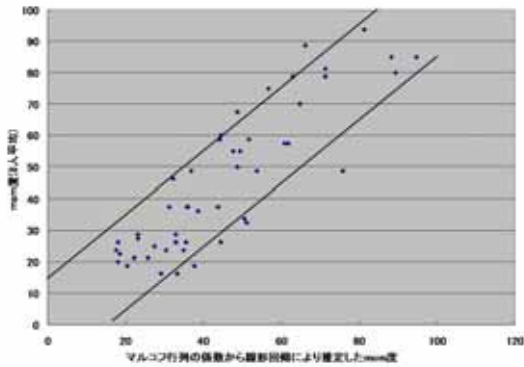


図 2

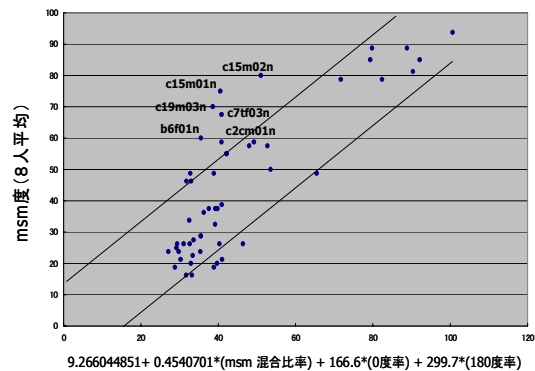


図 3

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遷移行列を推定し, 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。

[種村] 2次元 unit vector chain を用いた解析 (平成18年度研究成果報告書 p.38 参照) 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思えるが, 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ことができる。

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。

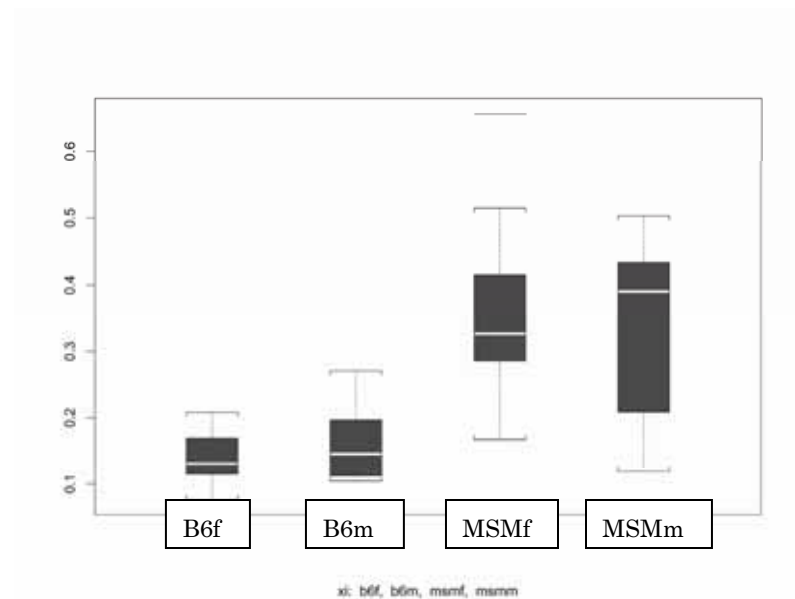


図 4

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 (川崎, 小出, 梅森)

【研究目的】

遺伝研において開発維持している多様な野生由来マウス系統は、高次機能の多様性を解析するための有用なリソースとなる。その高次機能の中で、行動表現型は形としては残らないため、その計測記録システムおよび解析には困難が伴う。また、行動は時間軸に沿った発現をするため、その定量的な評価は難しい。本研究では、マウス行動を定量的に計測し、得られたデータを数理的な手法を用いて解析して、将来の研究において発展性のある行動成分を抽出し、その形質の背後にある特性をモデル化により明らかにすることを目的とする。そのため、統計数理研究所と共同で、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。

【平成19年度の成果】

自発活動性は、動物にとってテリトリーの確保や餌の探索といった生存に深く関わる重要な行動形質の一つである。その活動には周期性が見られ、概日周期についてはすでによく知られている。しかし、動物は活動期においても一様に活動しているわけではなく短い時間で活動と休息を繰り返しているように見える。このような超日周期については、まだほとんど研究が進んでいない。その原因は、超日周期を解析するための有効な解析手法がこれまで確立されていなかったためである。そこで、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。

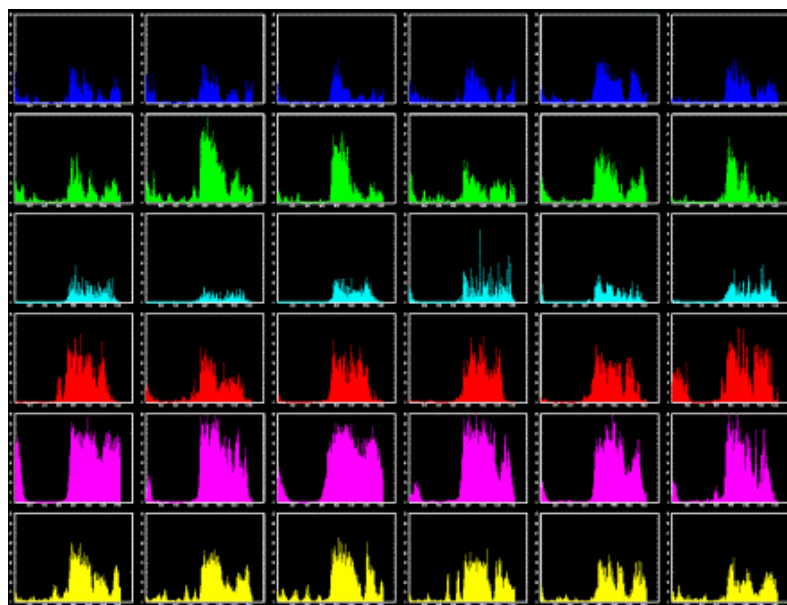
解析に用いたマウスは、照明を 8:00 から 20:00 までを明期、20:00 から 8:00 までを暗期としてコントロールし、室温 $22 \pm 2^\circ\text{C}$ 、湿度 $50 \pm 10\%$ に保った飼育室内で維持した。解析には MSM と C57BL/6 の生後 8~15 週齢の雌雄を解析に用いた。各マウスはテストの前に個別に飼育ケージに移し 1 日間飼育した。その後、ACTIVITY SENSOR (O' hara, Co., Ltd) の測定ケージにそれぞれ移し、4 日間連続でその活動量を自動計測した。最初の 1 日間は馴化期間として解析からは除外した。各個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

H18年度の研究では、種ごとに共通の周期性が存在することを仮定して、1分値を集計した上で話を進めたが、実際には個体差が大きいと思われるので、共通の周期性があるかどうかは先験的に仮定することには問題があるかもしれない。従ってH19年度は、短い周期までも含めた個別の周期性モデリングを1分値データに基づいて徹底的に行い、その結果から共通周期性がどれだけサポートされるのか、安定性を検証した。ここで用いたデータは、C57BL/6のオス(B6_M)53個体、メス(B6_F)個体、MSMのオス(MSM_M)11個体、メス(MSM_F)9個体、KJR 16個体、BLG 17個体である。

採用した統計モデルはポアソン時系列回帰モデルを用いた非定常ポアソンモデルである。すなわち、応答変数は計数なのでポアソン分布に従って発生すると仮定し、分布の強度(生起度)関数が時間とともに変わるモデルを用いる。強度関数は、有限個のフーリエ級数を指数変換した形で定式化する。元の時系列が1分値データであるので、24時間周期から最短2分周期までの周期成分を、推定されたフーリエ係数の有意性から判断する。時間帯ごとにばらつきが大きさが異なる可能性は十分あるので、分散不均一性があっても一致性のある標準誤差推定値としてHuber-White型のロバスト推定量を用いる。

このモデルの推定には数値的困難が伴う。最初に最も大きなモデルを推定しようとする、往々にして説明変数間に共線性が生じ、最適化に失敗するからである。従って、周期の長い方から短い方に向かって順番にsinとcosの組を1組ずつ追加し、その都度有意性をチェックする。あるステージで有意性検定により棄却された周期成分は、以降モデルに再度追加されることはない。このようにして得られた強度関数

の推定値の例が、右の図に示されている。行方向は1行目から6行目が先に述べたB6_MからBLGまでの種別に対応しており、列方向は任意の6個体の結果を順に並べたものである。横軸は時間で、朝の8時から翌朝の8時までである。横軸のほぼ中央が夜8時であり、暗期(活動期)の始まりに対応している。強度関数の縦軸の

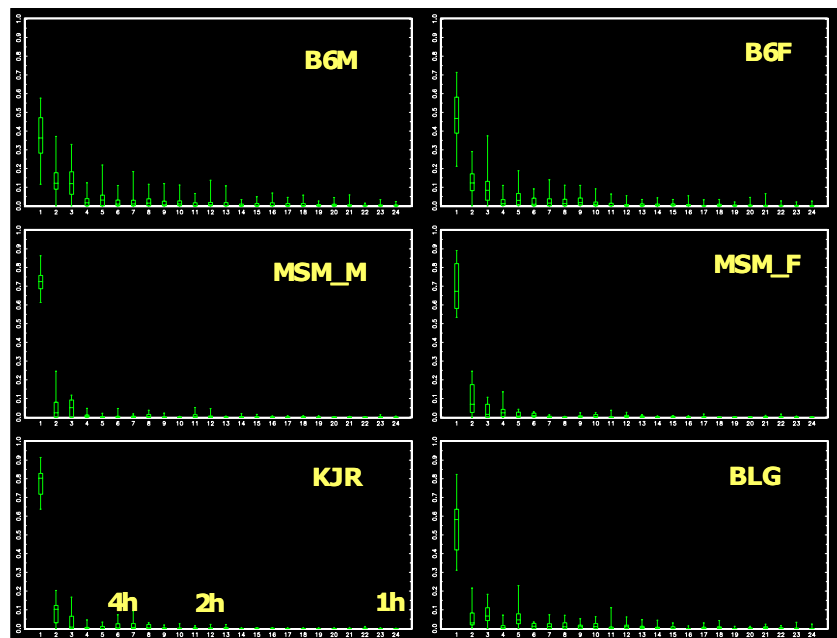


スケールは統一してあるので、面積の大きい個体ほど活動量が多いことを示している。

フーリエ係数の推定値(ただし有意なものばかりの絞る)が得られれば、そこからマウス各個体の行動周期性が線スペクトルとして得られる。換言すれば、どのような周期の波がどの程度の(平均的な)強さで分布しているかをフーリエ周波数上の離散分布として観察することができる。

分析の結果明らかになったことは、カレンダータイムの意味での周期性に関しては、高周波の貢献は非常に少ないということである。これをわかりやすく示すため、スペクトル分布の累積値が0.9になるまでに、各品種系統でどれだけのフーリエ級数を要するかを平均的に算出してみた。例えばB6_Mでは平均で21項目まで、これは時間にして1時間20分程度の周期までは貢献があるが、それより短い周期成分は殆ど影響していないことを示している。以下同様に、B6_Fでは13項(約2時間まで)、MSM_Mでは12項(約2時間まで)、MSM_Fでは4項(約6時間まで、但し個体数が非常に限られていることに注意)、KJRでは3項(約8時間まで)、BLGでは14項(約1時間40分まで)という結果が得られた。

推定された線スペクトルのボックスプロットが右の図である。活動量の個体差を排除するため、線スペクトルの分布は合計が1になるよう規格化してある。B6のように個体数が多いものと、MSMのように少ないものがあることは比較の際に注意が必要である。この図から、推定された日周期のパターンと周期成分



の特性が系統内でどの程度安定しているかを観察することができる。多くの場合、周期成分の貢献は低周波に集中していることを考え、図は最も短くて1時間周期の所までを描いている。一番左端のボックスは24時間周期に対応している。横軸(周期)のおおまかな目安は、左下のKJRの欄に記した、4時間(4h)等の記述を参考にしていきたい。

ここから以下の観察を得る。B6 系統は、MSM 系統に比べて、変動要因が比較的多くの周波数に渡っている。特に B6M の場合、平均的には 1 時間周期近くまでなだらかな分布になっている。その一方で、MSM 系統は 6 時間周期までの占める割合が多く、特に 24 時間周期の割合が圧倒的である。しかし、総じて個体間変動は大きく、特定周波数におけるパワーだけから、両系統を峻別できるほどの特徴ある周波数は見あたらない。また、B6 のスペクトル分布は、データと強度関数の形状から自然であると言える。というのは、B6 は明期(休息期)でもしばしば活動と休止を繰り返

返しているが、MSMは暗期に活動が集中しているからである。

今年度の研究結果で、非定常ポアソンとしてモデリングする部分はある程度突き詰めたと考えている。カレンダータイムの意味での周期性から各系統を峻別するほどの特徴量を抽出することは難しい、という結論は、やや残念な結果であった。しかし、このことは逆に、日周期のパターンはほどほどに表現しておいて、もっとローカルなところでの行動の挙動を系統ごとにモデル化することに集中してよいということを示している。これはH18年度の方針に立ち返って、その続きの研究を進めることに相当する。今年度も、観測時系列と推定した強度関数との差を適当に規格化して正規近似し、それに時系列モデルのあてはめることは試みたが、良好な結果は得られなかった。これを受けてH20年度は、マルコフ切り替えモデルや局所定常モデル等によるアプローチを計画している。

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（栗木，藤澤，倉田，春島，岡，城石）

【研究目的】

生物学的「種」は、互いに交配可能な自然集団で他の集団から生殖的に隔離されているものと定義される。その「種」を分ける遺伝的しくみを生殖的隔離機構（障壁）という。生殖的隔離障壁によって、交雑が阻害されたり、雑種が出来なかつたり、雑種が出来てもその後代の生存率が落ちるなどして遺伝子の自由な交換が妨げられ、結果として「種」の遺伝的特徴が維持される。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。研究を困難にしている要因の一つは、生殖的隔離が相互作用によってもたらされることである。

本研究課題の目的は生殖的隔離をもたらす相互作用を実験的に検出することである。実験的に相互作用を検出する上で、以下の4種類に生殖的隔離障壁の相互作用を分類することができる。

- (1) 対立遺伝子座間の相互作用。
- (2) 配偶体または接合体内の異なる遺伝子座間の相互作用。
- (3) 配偶体または接合体の遺伝子座と親の遺伝子座との相互作用。
- (4) 配偶体または接合体のゲノム遺伝子座と細胞質との相互作用。

本課題では、雑種の自殖集団 (F_2) におけるマーカー遺伝子型の分離データを用い、(2)の相互作用つまり、異なる遺伝子座が生殖的隔離障壁として作用し、特定の遺伝子型組合せを持つ配偶体または接合体が選択を受けたか否かの検出することを目的とする。

栽培イネ (*Oryza sativa*) には、ジャポニカ型とインディカ型があり、本来栽培イネは自殖することによって後代を残すが、両者間の交雑も可能でその雑種は組合せによって様々な稔性を示し、生殖的隔離が観測できる。ジャポニカ型日本晴とインディカ型 Kasalath との雑種 (F_1) の自殖後代 (F_2) 186 個体の集団を用い、イネ高密度連鎖地図が作成されており、そのマーカーの遺伝子型分離頻度から日本晴と Kasalath 間の生殖的隔離がマップされている。この生殖的隔離が上

記の(2)によってもたらされる場合は、異なる遺伝子座間で、遺伝子型分離の独立性検定を行うことにより、生殖的隔離障壁間の相互作用の検出が可能であると考えた。

配偶体及び接合体内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性検定を行った。連鎖しているマーカー間だけではなく多くの相互作用を示すと考えられるピークが認められた。しかし相互作用検出のため独立性検定を全てのマーカー間で行うことは、組合せ数が膨大になるため重大な多重性を引き起こすことが想定される。実際再現実験によると、そのピークの多くは偽陽性であった。(統計解析における多重性調整の方法は、昨年度までの報告書を参照。)本年度は、マーカーの遺伝子型分離頻度から生殖的隔離障壁のマッピングの再現性を確認するとともに、相互作用検定量の再現性を確認するため、イネ高密度連鎖地図を作製した同じ組合せ親による別のF₂集団を用い連鎖地図を作製した。

【平成19年度の成果】

高密度連鎖地図作成に用いた RFLP マーカーはサザンハイブリダイゼーションによって遺伝型を調べるため、各個体の DNA 製精度が高い事が求められると共に遺伝型決定に必要な DNA 量も PCR 法の 100 倍以上必要である。そこで、ゲノム全体をカバーする共優性の PCR マーカーの開発を 10cM 間隔で全ゲノムをカバーする事を目標に行った。全塩基配列が既知のジャポニカ型日本晴とインディカ型 93-11 の insertion/deletion 情報を TIGR Rice Browser (http://www.tigr.org/tigr-scripts/osal_web/gbrowse/rice/)より得、それを挟む primer 設定し日本晴と 93-11 の塩基配列に対し Electric PCR (Schuler, 1996)を用いて STS 化を確認し、実際に日本晴と Kasalath 間でも PCR 増幅フラグメントサイズに多型が見られるものをマーカーとした。

高密度連鎖地図を作成した 186 個体の F₂ 集団とは別の 286 個体の F₂ 集団を用い、新たに開発した 123 個の PCR マーカーを含む 166 個の共優性マーカーを用い連鎖地図を作成した。新たに作製された連鎖地図にそってマーカーの遺伝型の頻度をプロットし高密度連鎖地図の場合と比較した(図 1)。

本年度は各マーカーの遺伝型の精査が終了していないため、遺伝型の読み間違い等により連鎖地図が以前の物に比べ長めになっている染色体もあるが、設定した PCR マーカーは想定された所にマップされほぼ均等にマーカーを作製する事が出来た。遺伝型分離比がメンデル則の期待値から外れる領域は高密度連鎖地図とほぼ一致するが、第 3, 11 染色体長腕端付近, 第 8, 9 染色体短腕端付近に差が認められた。

来年度はマーカー遺伝型を精査し、連鎖地図の完成度を高めた後、回帰分析で分離頻度をゆがめる生殖的隔離障壁のマッピングの再現性を確認するとともに相互作用検定量の再現性を確認し、相互作用による独立性検定量のピークとランダムに生じたピークとの判別することが可能どうかを検討する。

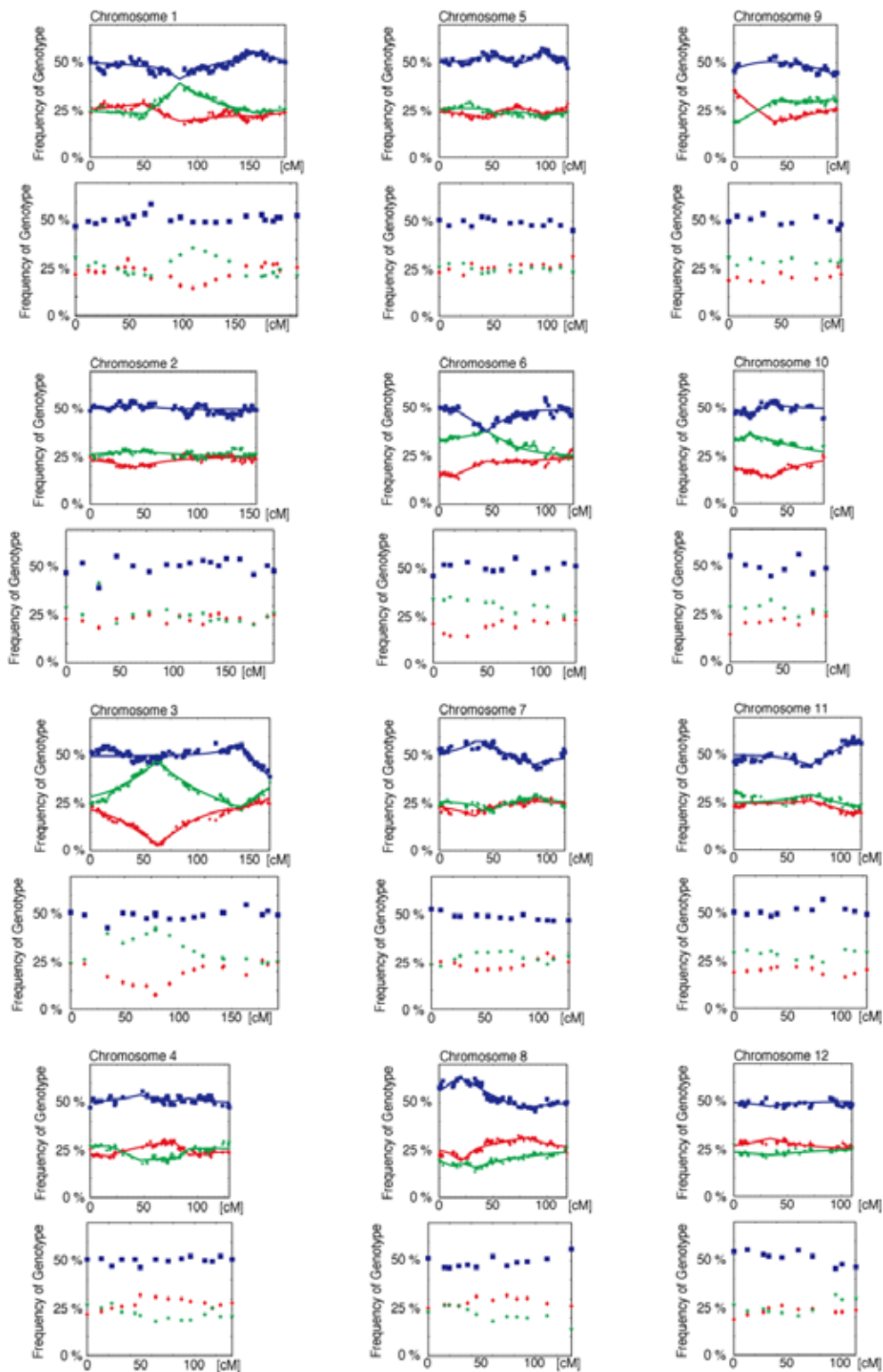


図1. 連鎖地図マーカーの遺伝型分離頻度

連鎖地図にそってマーカー遺伝型の頻度をプロット。各染色体左側が短腕，上側が高密度連鎖地図，下側が今回作製した連鎖地図。赤が日本晴ホモ型，緑が Kasalath ホモ型，青がヘテロ型

課題 e. QTL 解析法の最適化 (梅森, 小出, 栗木, 江口, 藤澤, 坂口, Mollah, 城石, 前野, 高田, 岡, 小出, 梅森)

【研究目的】

動物の行動形質は多因子により影響を受ける量的形質であるが, その行動は様々な行動要素から構成されていると考えられる。これらの行動要素にはそれぞれ異なった遺伝的要因が作用するため, 行動の遺伝学的解析, 特に QTL 解析は困難を伴う。このような問題に対応するには, 行動形質を構成する要素を効率よく確実に抽出し, それぞれの要素について, 的確な QTL 解析法を用いて研究を進めることが不可欠である。本研究では, マウスの行動データを例題とし, 従来の QTL 解析と SEM 解析 (Structural Equations Model) を組み合わせることにより, 行動形質と遺伝視座の連関を効率よく探索する方法論の開発を試みた。

【平成 19 年度の成果】

今回我々は, マウスの飼育ケージ内自発活動性について, 自発活動性を構成する要素を抽出し, それらについて QTL 解析を行った。同時に, 飼育ケージとは異なった場面である, 新規な環境のオープンフィールドにおける活動についても QTL 解析し, 飼育ケージ内自発活動性との関係を解析した。

マウスは, 馴化したケージ内で, 明暗周期に従い主に夜間に活動する。この活動性はケージ内自発活動性と呼ばれ, マウス系統間で大きく異なることが知られている。野生由来系統である KJR は代表的な実験用系統である C57BL/6J (B6) に比べてケージ内で高い活動性を示す。逆に KJR は, 新規な場面であるオープンフィールドにおいては, B6 に比べて低い活動性を示す。

これらケージ内での活動性とオープンフィールド活動性 (OFA) の系統差を決める遺伝的な基盤を明らかにするため, 高活動性系統である KJR と低活動性である C57BL/6J の F2 集団を用いた量的形質遺伝子座 (QTL) 連鎖解析を行った。この際, ケージ内での活動性については, 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) に分離した測度を同時に用いた (表 1)。それぞれの形質について, 広義の遺伝率 (H^2) を調べた結果, 何れの形質も 0.5 以上の高い値を示したことから, これらの形質は高い遺伝性を示すことがわかった。

表 1. KJR, B6, (B6xKJR)F1 および (B6xKJR)F2 のケージ内活動性とオープンフィールド活動性の比較

	Average value				H^2
	C57BL/6J	KJR	(B6xKJR)F ₁	(B6xKJR)F ₂	
THA	46.5 (8.9)	122.3 (15.9)	120.5 (19.1)	98.0 (31.0)	0.62
AT	19.6 (1.8)	22.1 (0.9)	25.7 (2.1)	23.5 (3.1)	0.53
AA	23.6 (3.2)	55.8 (6.2)	46.9 (5.6)	41.2 (10.6)	0.72
OFA	15.2 (3.2)	7.7 (1.3)	16.7 (1.5)	12.4 (6.3)	0.77

これらの F2 個体群の測度を用いた composite interval mapping を行った結果、総活動量を支配する有意な QTL である、*hyper locomotive activity related QTL1 (Hylaq1)* と *Hylaq2* を第 2 番染色体上に、第 10 番染色体のテロメア側に *Hylaq3* を検出した (図 1 および図 2)。

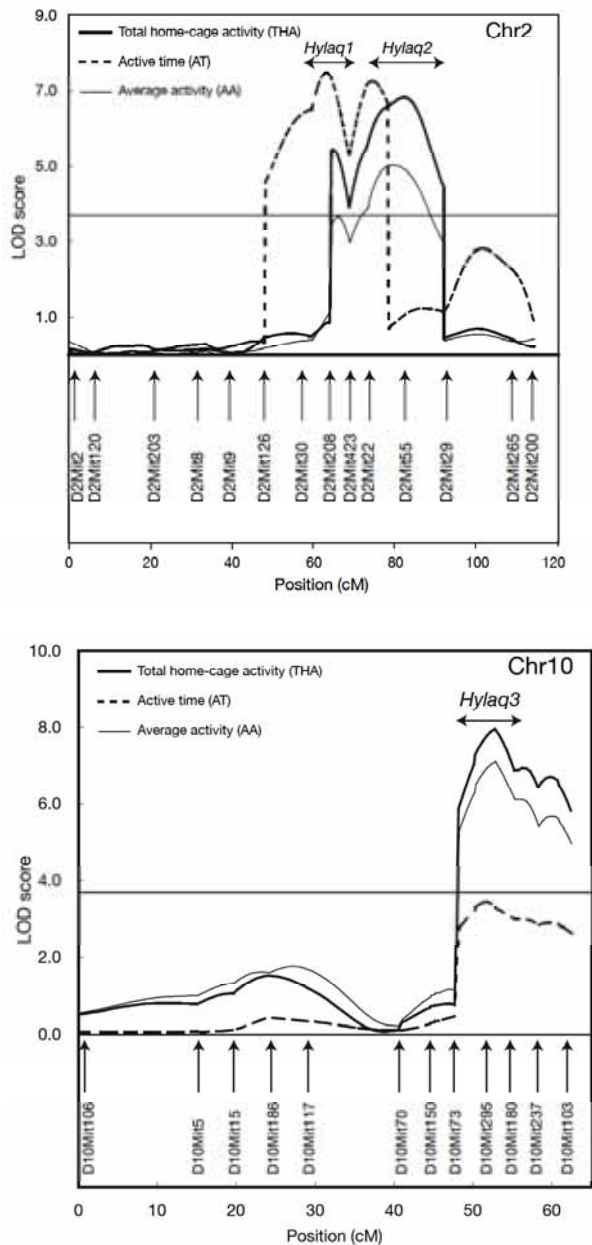


図 1. composite interval mapping による QTL の検出

さらに、composite interval mapping の結果、*Hylaq1* と *Hylaq2* は AT に、*Hylaq3* は AA に強く関与していることが示唆された。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/qt1 による共変量を用いた QTL interval mapping を行った (図 3)。さらに、得られた結果を用いて、構造方程式モデル (SEM) 解析を行い、検出された QTL と自発運動性の 3 つの測度との関係性の図示を試みた。

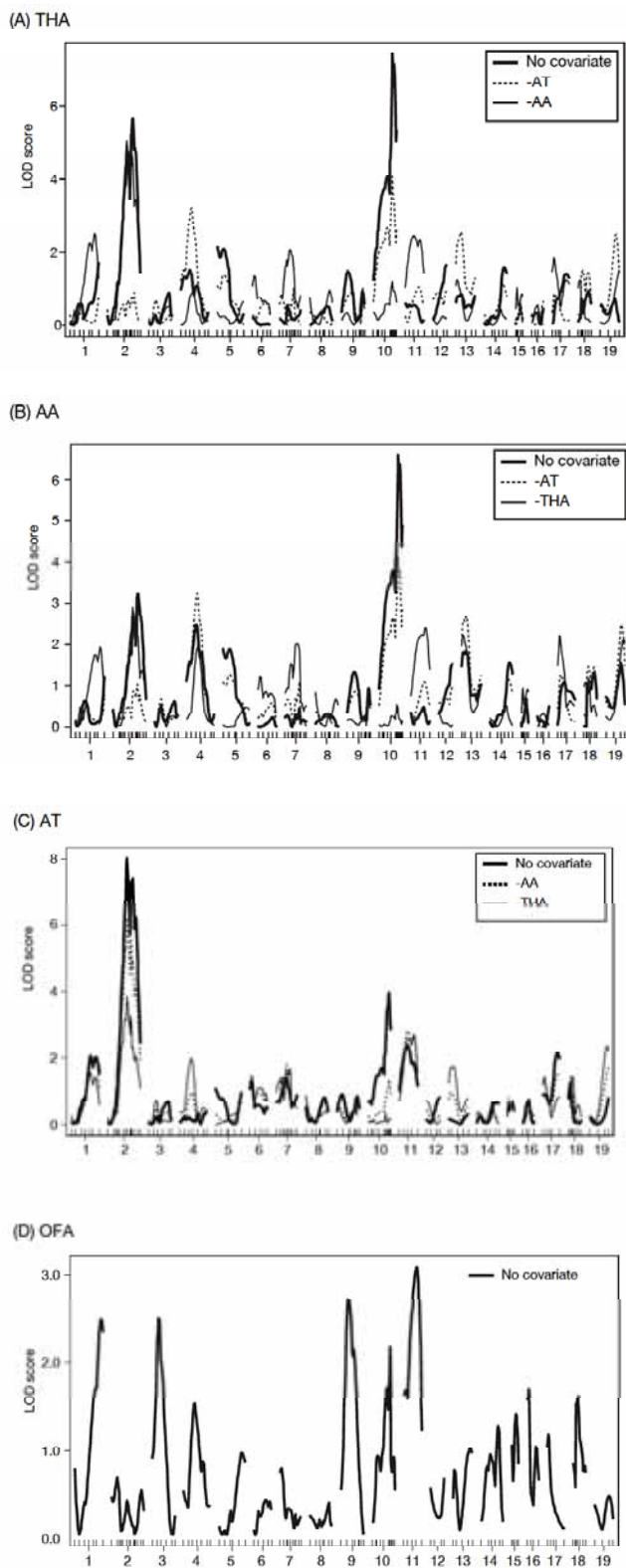


図2. 共変量を用いた interval mapping

(A) 共変量を用いないTHA (太線), ATを共変量とした場合のTHA (点線)およびAAを共変量とした場合のTHA (細線), (B) 共変量を用いないAA (太線), ATを共変量とした場合のAA (点線)およびTHAを共変量とした場合

の AA(細線), (C) 共変量を用いない AT (太線), AA を共変量とした場合の AT (点線)および THA を共変量とした場合の AT(細線)

SEM 解析は, Li らの方法を元に, 1) 初期パス図の作成, 2) パスの省略と統計的評価, 3) パスの追加と統計的評価を行い最終的なパス図を作成した (図 3)。これらの統計量については, 表 2 に示した。しかし, このパス図では, AA を制御する遺伝子座が無いことや THA から AA や AT にパスが出ているなど, 共変量を用いた interval mapping との矛盾がいくつか見受けられた。これは, $THA = AT \times AA$ という関数関係にある変数がパス図内に同時に存在することによるものと考えられる。従って, THA の代わりに, 以下の式で求められる残差 (RES) を用いて再度 SEM 解析を行い, より適切なパス図の作成を試みた (図 4)。

$$THA = a + b \times AA + c \times AT + RES$$

(a, b および c は最小 2 乗法によって求められる定数)

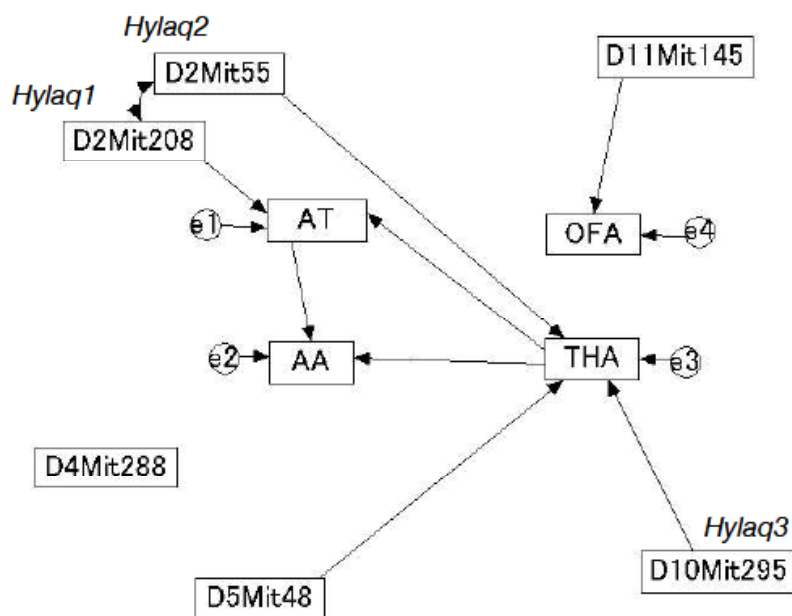


図 3. SEM 解析により得られた最終パス図

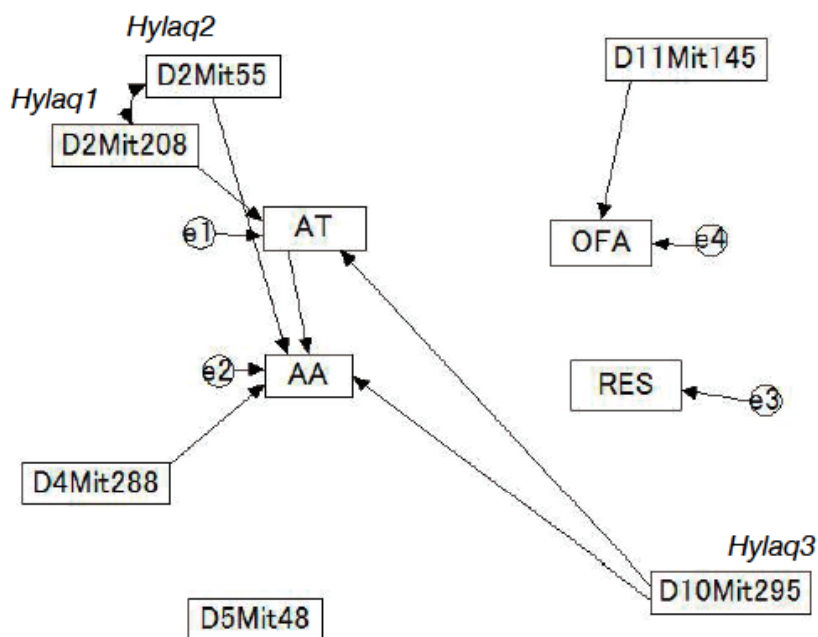


図 4. THA の代わりに RES を用いた SEM 解析により得られた最終パス図

SEM 解析の結果、*Hylaq1* と *Hylaq2* はそれぞれ、AT と AA に関連していることが示唆された。これは、ケージ内での自発活動性は、時間的要素と量的要素に分けられ、それぞれが異なる遺伝子座で支配されていることを示唆している。オープンフィールド活性に関連した suggestive な QTL は第 11 番染色体上に検出されたが(図 2)、上述の *Hylaq* 遺伝子座とは別の場所にあった。この結果から、馴化したホームケージ内での活動性は、オープンフィールドのような新規場面での活動性とは異なった遺伝的基盤を持つことが示唆された。

表 2. モデル適合のための 5 つの統計量

Model	AIC	chi-square (Df)	CFI	RMSEA	standardized residual
1	1136.179	1066.179 (30)	0.353	0.351	0.594
2	102.359	44.359 (36)	0.995	0.029	0.173
3	110.531	42.531 (31)	0.969	0.036	0.176
4	100.422	44.422 (37)	0.980	0.027	0.176

Model 1: THA を用いた初期パス図, Model 2: THA を用いた最終パス図, Model 3: RES を用いた初期パス図, Model 4: RES を用いた最終パス図. Df: 自由度. 全ての統計量は AMOS 7.0 で計算

した。

本研究の一部は Arimantas Lionikas, David Blizard 両博士 (Center for Developmental and Health Genetics, Pennsylvania State University) との共同研究である。

7. 平成19年度の研究成果

(1) 成果物 (知見・成果物・知的財産権等)

サブテーマの報告書に記載した。

(2) 成果発表等

<論文発表>

[学術論文]

1. 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.
2. Ammiraju, J. S. S., Luo, M., Goicoechea, J. L., Wang, W., Kudrna, D., Muller C., Talag, J., Kim, H., Sisneros, N. B., Blackmon, B., Fang, E., Tomkins, J. B., Brar, D., MacKilp, D., McCouch, S., Kurata, N., Lambert, G., Galbraith, D. W., Arumuganathan, K., Rao, K., Walling, J. G., Gill, N., Yu, Y., SanMiguel, P., Soderlund, C., Jackson S., Wing, R. A. The *Oryza* bacterial artificial chromosome library resource: Construction and analysis of 12 deep-coverage large-insert BAC libraries that represent the 10 genomes types of the genus *Oryza*. *Genome Research*. 16:140-147. 2007.
3. 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.
4. Schaller, G. E., Doi, K., Hwang, I., Kieber, J. J., Khurana, J. P., Kurata, N., Mizuno, T., Pareek, A., Shiu, S. H., Wu, P., Yip, W. K. Nomenclature for two-component signaling elements of rice. *Plant Physiol*. 143: 555-557. 2007.
5. 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.
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. 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.
8. Ito, N. and Kurata, N. Disruption of KNOX gene suppression in leaf by introducing its cDNA in rice. *Plant Science*. 174: 279-289. 2008.
 9. 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) (2007) 738-746.
 10. Kanamori, T., Takenouchi, T., Eguchi, S. Murata, N. Robust loss functions for boosting. *Neural Computation* 19 (2007) 2183-2244.
 11. 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 (2007) 91-110.
 12. Takenouchi, T., Ushijima, M. and Eguchi, S. GroupAdaBoost: accurate prediction and selection of important genes. *IPSI Transactions on Bioinformatics* 3 (2007), 1-8.
 13. Ninomiya, S. and Fujisawa, H. A conservative test for multiple comparison based on highly correlated test statistics, *Biometrics*, 63 (2007), 1135-1142.
 14. Takenouchi, T., Eguchi, S., Murata, N. and Kanamori, T. Robust boosting algorithm against mislabeling in multi-class problems. To appear in *Neural Computation*, 2007.
 15. 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 (2007) 285-293.
 16. 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/>)

[会議録]

1. 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

[解説・総説]

1. 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)。
2. 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)。
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

[研究ノート]

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

[その他]

<会議発表等>

[招待講演]

1. 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11月21-22日, 京都。
2. Shiroishi, T. Genome anatomy of C57BL/6 and exploration of SNPs for energy metabolism. EUMODIC First Annual Meeting, 2007. 2.22-23, Barcelona, Spain.
3. 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回国際免疫フォーラム特別講演。松山 2008年7月14日
4. 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。
5. 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
6. Eguchi, S. Boosting learning approach to association studies in bioinformatics, First joint meeting between Institute of Statistical Science, Academia Sinica, Taiwan and the Institute of Statistical Mathematics, 2007年11月29日
7. 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.
8. 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
9. 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", Tokyo, 2007年11月29日

〔一般講演〕

1. 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
2. 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
3. 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
4. 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
5. 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
6. 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
7. 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
8. 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
9. 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.2007.9.13-18. Hayama, Kanagawa, Japan
10. 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。
11. 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。
12. 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。

13. 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。
14. 上田健治, 豊澤恵子, 高橋幸子, 宮尾安藝雄, 廣近洋彦, 野々村賢一, 倉田のり, 井上正保. イネ花粉突然変異体 Tos0216 の解析. 日本植物学会第 71 回大会 2007. 9 月 6-9 日. 東京
15. 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. 2007. 9 月. 13-18 日. Hayama, Kanagawa, Japan
16. 米田典央, 野々村賢一, 倉田のり. イネ Pot1 ホモログの単離と解析. 日本遺伝学会第 79 回大会 2007 9 月 19-21 日. 岡山
17. 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. 日本遺伝学会第 79 回大会 2007 9 月 19-21 日. 岡山
18. 久保貴彦, 吉村淳, 倉田のり. イネの交雑後代に見出された F2 雌性不稔の遺伝機構. 日本遺伝学会第 79 回大会 2007. 9 月 19-21 日. 岡山
19. 藤田雅丈, 堀内陽子, 上田弥生, 水多陽子, 倉田のり. イネの生殖過程を通じた遺伝子発現プロファイリング. 日本遺伝学会第 79 回大会 2007 9 月 19-21 日. 岡山
20. 津田勝利, 伊藤幸博, 倉田のり. イネの KNOX 遺伝子 OSH1 の発現制御機構の解析. 日本遺伝学会第 79 回大会 2007 年 9 月 19-21 日. 岡山
21. 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害の相互作用の検出. 日本遺伝学会第 79 回大会 2007 年 9 月 19-21 日. 岡山
22. 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, 倉田のり. Affymetrix Rice Genome Array を用いた SFP 検出手法の開発. 日本育種学会第 112 回講演会 2007 年 9 月 22-23 日. 鶴岡
23. 津田勝利, 伊藤幸博, 宮尾安藝雄, 廣近洋彦, 倉田のり. KNOX 遺伝子を葉で異所的に発現するイネ突然変異体の解析. 日本育種学会第 112 回講演会 2007 年 9 月 22-23 日. 鶴岡
24. KNOX 遺伝子を葉で異所的に発現するイネ突然変異体の解析 2. 日本育種学会第 112 回講演会 2007. 9 月 22-23 日. 鶴岡
25. 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. 日本育種学会第 112 回講演会 2007 年 9 月 22-23 日. 鶴岡市
26. 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. 日本育種学会第 112 回講演会 2007 年 9 月 22-23 日. 鶴岡
27. 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關与する核遺伝子の発現解析. 日本育種学会第 112 回講演会 2007 年 9 月 22-23 日. 鶴岡市
28. 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。
29. 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 2007.12月11-15日。横浜
 30. 津田勝利，伊藤幸博，宮尾安藝雄，廣近洋彦，倉田のり。KNOX 遺伝子を葉で異所的に発現するイネ突然変異体の解析(1)。第49回日本植物生理学会年会 2008.3月20-22日。札幌
 31. 伊藤幸博，津田勝利，倉田のり KNOX 遺伝子を葉で異所的に発現するイネ突然変異体の解析(2) 第49回日本植物生理学会年会 2008.3月20-22日。札幌
 32. 石川亮，永口貢，池田陽子，倉田のり，木下哲。イネの胚乳における生殖隔離機構とゲノムインプリンティング。第49回日本植物生理学会年会 2008.3月20-22日。札幌
 33. 望月孝子，倉田のり，矢野健太郎。OryzaExpress:イネのゲノム・アノテーションと遺伝子発現の統合データベース。第49回日本植物生理学会年会 2008.3月20-22日。札幌
 34. 藤田雅丈，望月孝子，堀内陽子，水多陽子，上田弥生，春島嘉章，倉田のり。Affymetrix マイクロアレイプローブ再定義とイネ生殖過程の遺伝子発現解析。日本育種学会第113回講演会 2008.3月27-29日。川崎
 35. 矢野健太郎，大木信彦，望月孝子，藤田雅丈，堀内陽子，谷坂隆俊，倉田のり。OryzaExpress: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。日本育種学会第113回講演会 2008.3月27-29日。川崎市
 36. 上田弥生，野々村賢一，藤田雅丈，堀内陽子，倉田のり。生殖細胞の初期発生が異常になるイネ mel1 突然変異体を用いたマイクロアレイ解析。日本育種学会第113回講演会 2008.3月.27-29日。川崎
 37. 石川亮，永口貢，池田陽子，倉田のり，木下哲。イネの胚乳における生殖隔離機構とゲノムインプリンティング。日本育種学会第113回講演会 2008.3月27-29日。川崎
 38. 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。
 39. 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。
 40. 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。
 41. 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
 42. 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
 43. 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，2007年9月6日

44. 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」, 東京, 2007 年 11 月 27 日
45. 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, 東京, 2007 年 11 月 29 日

<著書等>

1. [執筆] 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)
2. [執筆] 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.
3. [執筆] 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.

<受賞>

- (3) その他の成果発表