

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

プロジェクトディレクター: 城石俊彦

1. プロジェクト体制

(1) 表現型数値計測システムの開発

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] 高田豊行、春島嘉章、倉田のり

[国立情報学研究所] 北本朝展、藤山秋佐夫

[統計数理研究所] 江口真透、藤澤洋徳、川崎能典

[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

[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

詳細は、サブテーマ参照

(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、種村正美、田村義保、土谷隆、藤澤洋徳

[国立遺伝学研究所] 小出剛、倉田のり

[東京大学] 原尚幸

[九州大学] 二宮嘉行

[大分県立看護科学大学] 坂口隆之

詳細は、サブテーマ参照

2. 今年度の研究概要

2-1. 前年度の研究成果

プロジェクト全体課題「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」の研究推進のため、(1)「表現型数値計測システムの開発」、(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」の二つのサブテーマに分けて融合研究を推進した。(1)では、主として以下の研究成果があった。a. マウス脂肪領域判定について、従来の ActScan 法を改良した ActCuts 法を開発した。b. マウスゲノム全体で 30 箇所以上のエネルギー代謝関連表現型の QTL を検出した。c. 野生イネの 1 年草と多年草の生育形質情報と遺伝子多型および発現情報を整備した。d. ロバスト推定による遺伝子

発現と SNP の同時検出法である SNEP 法の改良を行った。(2)としては、主に以下の研究成果があった。a. 生物形態についての新たな記述子として、TFD (P 型フーリエ記述子) より優れた protrusion フーリエ記述子 (PFD) を提案した。b. マウス社会性行動の評価系として開発した二つの手法を用いて、実際のマウスの時系列移動行動データを得た。c. マウスの自発活動について、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手法を考案した。d. 致死遺伝子相互作用による遺伝子座間の分離の独立性検定量の際に、ヘテロ型がホモ型の致死率の半分になる場合を考慮して遺伝的モデルを作製した。e. 表現型の外れ値に対してロバストな QTL 手法を開発した。また、QTL 解析結果に重大な影響を与える個体を特定する方法の検討を行った。

2-2. 今年度の研究目標

今年度は最終年度であるため、これまでの課題を継続して推進し完成させる。また、それらの成果発表を行うことを目標とする。各サブテーマの主な研究目標は下記の通りである。(1)「表現型数値計測システムの開発」では、a. 新規プログラムである ActCuts 法を用いた測定により、従来は困難であった複雑な部位での内臓脂肪と皮下脂肪の抽出精度の向上を目指す。b. 前年度に導入した SNP 検出システムを利用して、エネルギー代謝関連 QTL 近傍の染色体組換え部位を同定し連鎖地図を完成する。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し、SNEP 法によって解析可能かどうかを検証する。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) と比較検証する。(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、a. 生物形状の数値化のためのさらに新たなフーリエ記述子の可能性を探索する。b. 数理的解析により得られたマウス社会性の評価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、解析手法の総合的な検証を行う。c. 多変量時系列モデルによる自発活動性の特徴抽出の手法を 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検討する。d. 生殖隔離障壁を引き起こす真の遺伝子座相互作用を、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。e. 昨年度開発したロバスト QTL 法について、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めるよう、データに応じてロバスト QTL 法を選択する方法を考案する。また、QTL 解析結果の影響診断では、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。国際シンポジウムを開催して、これらの研究成果を発表する。

2-3. 研究推進の考え方

サブテーマ毎に以下の考え方で推進する。(1)「表現型数値計測システムの開発」では、a. 複雑な形状の脂肪組織の分離・定量が可能となり、処理対象の臓器の範囲がさらに広がるように留意する。さらに、測定については造影剤なども利用した抽出精度の向上に配慮

する。b. 前年度に導入した SNP タイピング系を活用して表現型を規定する QTL の責任領域をさらに限定するよう連鎖解析を一段と推し進める。c. アレイ解析によるデータと形態測定から得られる量的形質データを組み合わせ、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 解析に向けて、問題点の洗い出しを行う。d. SNEP 法によって塩基変異の多い（進化速度の速い）遺伝子についても、発現量多型の真の値を検出できるか検証する。次世代シーケンサーによる発現解析や塩基多型の検出を試みて、マイクロアレイによる解析法との比較を行う。(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、a. 統数研メンバーが中心になり形状を数値化するための方法を考え、遺伝研のメンバーが中心になり、形状と遺伝子の関連を調べるための QTL 解析を行う。b. 実データに基づいた動物の社会性の評価結果について、その有用性、問題点、改善点などを検討することに留意する。d. 局所定常自己回帰モデルや時変係数自己回帰モデル等を、これまで研究してきた枠組みに援用することで、自発活動の日周期の短期的変動の特徴抽出が可能かどうか検証する。e. QTL 解析のロバスト法は、影響解析と相補的な適用が可能である。二つを比較し、お互いの利点を生かした解析手順の開発にも配慮する。

2-4. 期待される効果

以下、サブテーマ毎の効果を列記する。(1)「表現型数値計測システムの開発」では、a. 生体構造の多様性の遺伝要因探索への貢献が期待できる。また、開発したソフトウェアはヒト疾患研究への応用も期待できる。b. 本手法は Synteny 情報をはじめとしたゲノム基盤情報と組み合わせることで、ヒトゲノム機能の解明にも応用できる。c. イネ系統間のゲノム、遺伝子、形態レベルで見られる多型の関連性を統合的に理解でき、有用形質を支配する未知遺伝子の探索も可能になる。d. ゲノム配列情報が未知の近縁種の遺伝的多型をゲノム全体に渡り検出可能になり、個々の遺伝子発現量をほぼ正確に把握できるようになる。従来マイクロアレイ解析で偽陽性や偽陰性として検出されていたデータの大幅な修正が期待できる。(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、a. 動植物の形態による系統分類に適用することができる。また、形態を支配するゲノム機能解析への道を開く可能性がある。b. 社会行動は人社会においても重要性が高く、防犯監視カメラによる自動解析などに応用できる可能性があり、将来的に社会への貢献も大きい。c. 活動周期成分の数理的解明はヒト精神疾患の理解に大きく貢献するものと期待できる。また、複雑な周期成分を数理的に理解する手法の確立は、周期性を示す他の現象に応用できる。d. 比較的小規模集団で真の遺伝子座間相互作用を検出できれば、種分化に働く生殖的隔離機構の研究に大きな貢献となる。e. ロバスト法と影響診断の観点をデータ解析に組み込むことで、より信頼度の高い統計解析手順が得られる。多次元の表現型データに対する QTL 解析は、近い将来の需要が見込まれる eQTL 解析を見越した準備研究としても期待できる。

3. サブテーマの構成

サブテーマ名称：(1) 表現型数値計測システムの開発

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、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、久保貴彦、小出 剛、杉本大樹

[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

[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

研究目標

動植物のモデル生物系統の表現型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の体系的な収集を行う。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を開発する。

サブテーマ名称：(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、種村正美、田村義保、土谷隆、藤澤洋徳、Nurul
Haque Mollah、奥田将己、Dou Xiaoling、田中英希

[国立遺伝学研究所] 杉本大樹、小出剛、岡彩子、城石俊彦、田村勝、春島嘉章、高田豊
行、前野哲輝、細谷正樹、倉田のり

[東京大学] 原尚幸

[九州大学] 二宮嘉行

[大分県立看護科学大学] 坂口隆之

研究目標

生物系統の表現形質多様性や遺伝子多型に関するデータを解析するための統計的手法の研究を行う。さらに、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。

4. 今年度の予算金額及びその内訳

費 目	金 額	主な用途
人件費	49,740 千円	研究員7名、非常勤職員8名雇用費、RA1名 マウス飼育管理者1名
物件費	97,842 千円	MassARRAY 解析用サーバー、ネットワークスキャナ、PC マイクロアレイ、シーケンス試薬、システム管理ソフト マウス飼育費、書籍 国際シンポジウム招待演者旅費 (600 千円 x 4 名) 成果発表旅費 (国内 4 件、海外 1 件)、研究打合せ旅費 専門知識の提供 学会等参加費、英文論文添削、国際シンポジウム経費
備品費	16,242 千円	
消耗品費	64,000 千円	
旅 費	10,300 千円	
謝 金	800 千円	
その他	6,500 千円	
合 計		<u>147,582 千円</u>

< 1品もしくは1組が500万円以上の物件 >

① SEQUENOM 社製 MassARRAY 解析用サーバー 11,342 千円

5. その他

なし

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

サブテーマ名：表現型数値計測システムの開発

1. プロジェクト体制

研究代表者

[国立遺伝学研究所系統生物研究センター] 城石俊彦

共同研究者

[国立遺伝学研究所] 倉田のり、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、久保貴彦、小出 剛、杉本大樹

[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

[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

2. 今年度の研究計画

2-1. 前年度の研究成果

課題 a. 3D 画像による生体再構築と生体内部の自動計測法の開発（北本、藤山、佐藤、高田、中谷、阿部、城石、田村（勝））： マウス個体を用いた脂肪領域判定について、撮影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 法を開発した。従来の ActScan 法と比較して、皮下脂肪と内臓脂肪の判定に関して 10%程以上の精度向上を確認した。

課題 b. マウスエネルギー代謝関連表現型計測法の開発（高田、前野、城石）： 日本産亜種由来 JF1 系統と汎用近交系統 B6 の F2 世代個体を用い、雌雄計 171 匹のエネルギー代謝関連表現型の収集および QTL 解析を行った。その結果、性染色体を除く全ゲノムの 30 か所以上に表記表現型に関する高い LOD 値 (2.5 以上) を示す領域を検出することができた。

課題 c. イネ穂形質・ストレス耐性の系統データ収集と遺伝子発現データとの相関解析（倉田、春島、堀内、久保、江口、藤澤）： イネ栽培種の祖先種である野生イネ *O. rufipogon* の 1 年生と多年生の複数系統で、10 項目以上の形質、生育形質を調査・整備した。また、遺伝子多型、発現多型解析用サンプルを収集し、一部のマイクロアレイデータを得た。

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、小出、杉本）： ロバスト推定による遺伝子発現及び塩基多型の同時検出法 (SNEP: Simultaneous detection of nucleotide and expression polymorphisms)

を更に改良して良い結果を得た。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った。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た。

2-2. 今年度の研究目標

- 課題 a. 3D 画像による生体再構築と生体内部の自動計測法の開発（北本、藤山、佐藤、高田、中谷、阿部、城石、田村（勝））：ActCuts法を用いた測定により、従来は困難であった複雑な部位での内臓脂肪と皮下脂肪の抽出精度の向上を目指す。内臓脂肪と皮下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腹膜の内外の脂肪は抽出できるが、実際のデータでは腹膜の位置自体も予測する必要があった。この問題を克服するため、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、これらのネットワークの抽出法の検討を行う。
- 課題 b. マウスエネルギー代謝関連表現型計測法の開発（高田、前野、城石）：エネルギー代謝関連表現型の QTL 解析で高い LOD 値を示す領域からの遺伝子探索を開始する。前年度に導入した SNP 検出システムを利用して染色体組換え部位の詳細な位置の探索を行う。さらに、肥満に関する QTL 解析に関して複数遺伝子の交互作用に関与するゲノム領域の抽出を行う。
- 課題 c. イネ穂形質・ストレス耐性の系統データ収集と遺伝子発現データとの相関解析（倉田、春島、堀内、久保、江口、藤澤）：野生イネの複数系統の発現データをマイクロアレイにより取得し、SNEP 法にて解析可能かどうか検討する。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 及び形質データを QTL 解析に適用し、問題点の洗い出しを行う。
- 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）：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試み、簡便で確実な統計的方法論の導入の可能性について考察する。

2-3. 研究推進の考え方

- 課題 a. 今年度新たに開発したActCuts法は、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手法である。処理対象の臓器の範囲がさらに広げるため、この性能をさらに向上させ、複雑な形状の脂肪組織を詳細に分離・定量する方向をめざす。測定については造影剤の利用による抽出精度の向上も検討する。
- 課題 b. エネルギー代謝に関連した表現型データに関しては、F2個体の全ゲノムを対象とし

た粗い遺伝子型判定が終了した。平成21年度は、SNP検出システムを利用してマーカー数を増大して解像度を上げた連鎖解析を実施し、QTL領域をさらに狭めることをめざす。

課題 c. これまでは解析集団の準備に時間を要したが、既に形質データも発現解析用サンプルも得られているので、SNEP法によって遺伝子の発現量多型及び塩基配列多型の検出を試みる。アレイ解析データと形態測定から得られる量的形質データを組み合わせ、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解析を試み、問題点の洗い出しを行う。

課題 d. SNEP法は低発現遺伝子上の塩基配列多型の検出には適さない、また塩基配列多型を多く含む遺伝子については、塩基配列多型の検出、発現量の推定共に適していない。genome DNA ハイブリおよび mRNA 発現量など複数のパラメーターとの重回帰分析を用い、mRNA ハイブリ時の SNP によるシグナル減少量の推定を試みる。probe の塩基変異の多い（進化速度の速い）遺伝子についても、発現量多型の真の値を検出できるか検証する。塩基配列未知の野生イネについて、次世代シーケンサーによる遺伝子発現量測定や塩基多型の検出を行い、これまで行ってきたマイクロアレイによる解析法との比較を行う。

2-4. 期待される効果等

課題 a. 新たに開発した ActCuts 法を利用して、画像データからの表現型数値化を高精度で高速に、かつ大量に行うことが可能となれば、複雑な形状の計測、特に本プロジェクトが目的としている生体構造の遺伝制御に関連したゲノム機能探索に大きな貢献が期待できる。また、成果として得られるアルゴリズムを実装したソフトウェアはヒト疾患研究分野への応用も期待できる。

課題 b. 遺伝的距離の離れたマウス系統を使用した表現型収集と複数遺伝子座の統計遺伝解析により、様々な量的形質を広範囲かつ体系的に解析することができる。さらに、Synteny情報をはじめとしたゲノム基盤情報と組み合わせることで、ヒトゲノム機能の解明にも応用できる。

課題 c. SNEP法をゲノム配列未知の多数の野生および栽培イネ系統に適用し、ゲノムDNA多型の検出と遺伝子発現量測定がほぼできるようになった。これにより、イネ系統間のゲノム、遺伝子、形態レベルで見られる多型の関連性を統合的に解析できるようになり、有用形質を支配する未知遺伝子の探索も可能になる。また、多様性生成にかかわる遺伝子群の発見も期待できる。

課題 d. ゲノム配列情報が未知の近縁種やそこから派生した生物系統の遺伝的多型をゲノム全体に渡り検出可能になり、個々の遺伝子発現量をほぼ正確に把握できるようになる。従来のマイクロアレイ解析で偽陽性や偽陰性として検出されていたデータの大幅な修正が期待でき、多因子が関与する遺伝子発現ネットワークの解明や、複雑

に発現調節を受ける遺伝子群の同定など高次生命現象に関与するゲノム機能の解明に貢献できる。次世代シーケンサーによる解析法と比較を行い、生物の多様性を理解する実験手法の最適化が期待される。

3. 今年度の予算金額及びその内訳

費 目	金額	主な用途
人件費	41,435 千円	研究員 5 名、非常勤職員 6 名雇用費、 マウス飼育管理者 1 名
物件費	82,142 千円	MassARRAY 解析用サーバー マイクロアレイ、シーケンス試薬、システム管理ソフト マウス飼育費 国際シンポジウム招待演者旅費 (600 千円 x 4 名) 成果発表旅費 (国内 4 件、海外 1 件)、研究打合せ旅費 専門知識の提供 学会等参加費、英文論文添削、国際シンポジウム経費
備品費	11,342 千円	
消耗品費	60,000 千円	
旅 費	6,200 千円	
謝 金	800 千円	
その他	3,800 千円	
合 計		<u>123,577 千円</u>

< 1 品もしくは 1 組が 5 0 0 万円以上の物件 >

① SEQUENOM 社製 MassARRAY 解析用サーバー 11,342 千円

4. その他

なし

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

サブテーマ名：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

1. プロジェクト体制

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, 種村正美, 田村義保, 土谷隆, 藤澤洋徳, Nurul Haque Mollah, 奥田将己, Dou Xiaoling, 田中英希

[国立遺伝学研究所] 杉本大樹, 小出剛, 岡彩子, 城石俊彦, 田村勝, 春島嘉章, 高田豊行, 前野哲輝, 細谷正樹, 倉田のり

[東京大学] 原 尚幸

[九州大学] 二宮嘉行

[大分県立看護科学大学] 坂口隆之

2. 今年度の研究計画

2-1. 前年度の研究成果

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 (田村, 江口, 城石, Mollah, 田中, 細谷): 新たな記述子として protrusion フーリエ記述子 (PFD) を提案し、PFD 使用時の方が TFD (P 型フーリエ記述子) より優れていることを確かめた。この PFD を用いて、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った。本解析の目的は形態形成に対して影響力の大きい遺伝子の探索等であった。

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, 杉本): 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方法である。これら 2 つの手法を用いて実際に解析するために、マウス 24 系統 461 ペアについての時系列移動行動データを得た。

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 (川崎, 小出, 杉本): 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. 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，二宮）： 遺伝子座相互作用を検出するために遺伝子座間の分離の独立性検定量を行うと、遺伝子座の組合せ数が膨大になるため重大な多重性を引き起こす。この多重性の調整行い閾値を見積もると、イネの高密度連鎖地図を作成した集団では、全ての相互作用を示すピークが閾値以下となる。しかし、致死遺伝子相互作用の単純な遺伝的モデルをたて独立性の検定統計量等との対応関係を明らかにしたところ、独立性検定量が閾値以下でもモデルと一致するものは真の相互作用であることが、再現性実験及び正逆戻し交雑集団によって確認された。そこで、遺伝的モデルをより一般的なものとするために、致死率が 100%ではない場合、ヘテロ型がホモ型の致死率の半分になる場合も考慮し作製した。

課題 e. QTL 解析法の最適化（栗木，江口，藤澤，坂口，Mollah, Dou, 城石，前野，高田，岡，小出，杉本）： ロバスト QTL についても幾つかの検討を加えた。特に公開されたデータにおいて表現形の観測値に外れ値を入れて解析をした。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された手法は非常に安定した性能を示すことが示された。カーネル法による QTL 解析に適用を検討した。また、QTL 解析における影響診断、すなわち解析結果に大きな影響を与える個体を特定する方法を検討した。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、グラフィカルに指示する方法の開発を行った。

2-2. 今年度の研究目標

今年度は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。

2-3. 研究推進の考え方

課題 a. 昨年度と同様に、統数研のメンバーが中心になり、形状を数値化するための方法を考える。遺伝研のメンバーが中心になり、形状と遺伝子の関連を調べるための QTL 解析を行う。

課題 b. これまでに、マウス時系列移動行動データから社会性を自動的に評価するための方法の確立を目指して研究を進めてきた。これまでの成果を用いて、実際の 24 系統 461 ペアという大量のマウス時系列移動行動データの解析を行う。そこで得られた社会性の評価結果は、人が実際にビデオ画像を用いて行った観察法による結果と比較し、その有用性、問題点、改善点などを検討することになる。

課題 c. H19 年度までに、マウス自発活動の日周期に着目した特徴付けを研究してきた。系統ごとにどの程度までの日周期が顕著かはある程度把握できているので、その周りの短期的変動での特徴抽出が残る重要課題である。個体差、活動・休止の違いを表現すべく、局所定常自己回帰モデル、時変係数自己回帰モデル等を、これまで研究してきた枠組みに援用することで限界突破を狙う。

課題 d. 閾値以下の独立性検定量を示す相互作用であっても遺伝モデルに適合する統計量を示す相互作用は真の相互作用である可能性が高い。殆どの致死率 100%の遺伝モデルでは、当初実験に用いた集団サイズでも閾値以上の独立性検定量を示すことが期待される。そこで、遺伝モデルも致死率が 100%ではない場合を考慮した一般的なモデルを構築した。しかしモデルの一般化によって、いずれのモデルが妥当かは容易に判定出来ない問題が生ずる。そこで、実験により検出した相互作用のピークがどの遺伝的モデルに当てはまるかを統計モデルの立場で検討し、真の相互作用の検出を目指す。

課題 e. Mollah が 6 月に帰国してもメール、電話を通して研究を継続するように進める。と

ここで QTL 解析のロバスト法は、影響解析と相補的な適用が可能である。この意味で現在並行的に研究が進められておるロバスト法と影響解析の双方を比較し、お互いの利点を生かした解析手順の開発も目指す。

2-4. 期待される効果等

- 課題 a. 形状を解析するための方法が実用化できると、マウスの下顎だけでなく、動植物の形態による分類に用いることができる。また、形態を支配するゲノム解析の道も開く。
- 課題 b. 一般的に行動の評価が難しい中で、社会行動は特にその評価が難しい形質である。その理由は、1 個体だけではなく、複数個体間の行動を評価する必要があるためである。本研究は、そのような社会性を自動解析により評価しようとするものであり、関連する分野に新たな解析手法をもたらすものと期待できる。また、社会行動は人社会においても重要性が高く、防犯監視カメラによる自動解析などに応用できる可能性があり、将来的に社会への貢献も大きいものと期待できる。
- 課題 c. 活動周期の異常は様々なヒト機能性精神障害でみられることから、本研究で目指す周期成分の数理的解明はヒト精神疾患の理解に大きく貢献するものと期待できる。また、複雑な周期成分を数理的に理解する手法の確立は、様々な周期性を示す現象に応用できると考えられる。
- 課題 d. 「種」を隔てる遺伝子座相互作用を統計的に有意に検出するには、大きな集団を用いて解析する必要があり時間と費用がかかる。遺伝的相互作用モデルを考慮することにより比較的小規模の集団で真の相互作用を検出する事ができれば、種分化に重要な役割を演ずる生殖的隔離機構の研究に大きな寄与を果たすと考えられる。
- 課題 e. QTL 解析において、その結論が少数の個体によって決定づけられる場合はめずらしくない。また遺伝子型データ、表現型データの外れ値も多い。古典的な QTL 解析はほぼ完成しているように思われているが、ロバスト法、影響診断の観点をデータ解析に組み込むことによって、より信頼度の高い統計解析手順が得られると期待される。また、QTL 解析において必要となるグラフィカルモデルは、連続変量（表現型）と離散変量（遺伝子型）の混在したものであるが、そのようなグラフィカルモデルは従来あまり研究されていない。また多次元の表現型データに対する QTL 解析は、近い将来の需要が見込まれる eQTL 解析を見越した準備研究と位置づけることができる。

3. 今年度の予算金額及びその内訳

費 目	金額	主な用途
人件費	8,305 千円	ポスドク 2 名, 研究補助 2 名, R A1 名
物件費	15,700 千円	
備品費	4,900 千円	ネットワークスキャナ, P C
消耗品費	4,000 千円	計算機消耗品, ソフトウェア, 書籍
旅 費	4,100 千円	学会参加, 研究者招聘
その他	2,700 千円	謝金, 論文掲載料, 英文校正
合 計		<u>24,005 千円</u>

< 1 品もしくは 1 組が 5 0 0 万円以上の物件 >

なし

4. その他

なし