

科学研究費助成事業 研究成果報告書

平成 28 年 6 月 10 日現在

機関番号：12608

研究種目：若手研究(B)

研究期間：2013～2015

課題番号：25870379

研究課題名(和文)多様な昆虫ゲノム比較へ向けた情報学的基盤形成

研究課題名(英文)Development of informatics infrastructure for the comparison of various insect genomes

研究代表者

小寺 正明(Masaaki, Kotera)

東京工業大学・大学院生命理工学研究科・講師

研究者番号：90643669

交付決定額(研究期間全体)：(直接経費) 3,500,000円

研究成果の概要(和文)：本研究では、昆虫ゲノム情報およびその他の遺伝子情報の収集と、鱗翅目昆虫の新規ゲノム解読を並行して行った。既知ゲノム情報としては、すでに世界各国の研究者により解読され公開されている多様な昆虫ゲノムや、タンパク質配列などの情報を収集した。新規ゲノム解読では、アゲハチョウ科の鱗翅目昆虫を複数種剪定し、生きた昆虫の採集、ゲノムDNA抽出、次世代シーケンサーMiSeqによる新規ゲノム解読、そしてゲノムデータのアセンブルを行った。また、これらの昆虫種を含む多様な昆虫種について、系統分類や、食草、忌避物質、誘引物質などの情報も収集した。

研究成果の概要(英文)：In this study we collected genomes and other genomic informations of insects, as well as novel sequencing of butterfly insects. As the known genome information, we collected various insect genomes and proteins in public that have been sequenced by many researchers worldwide. As the novel genome sequencing, we selected a number of swallowtail butterflies, captured them alive and extracted their DNA, subjected to the next generation sequencer, and assembled. We also collected various information of insects, including taxonomy, edible plants, attractant and repellents.

研究分野：バイオインフォマティクス

キーワード：昆虫ゲノム

1. 研究開始当初の背景

昆虫は生態学的にも経済的にも極めて重要な動物群であり、いくつかの昆虫ゲノムがこれまで解析されてきた。幅広い昆虫種の総遺伝情報(ゲノム)を提供することは、長年の昆虫研究で明らかにされた知見をさらに発展させ、大きな社会的利益を生み出すのに不可欠であると研究代表者は考えた。たとえばデンマークでは昆虫ゲノムセンターが設立され、昆虫ゲノミクス研究が進められていた。

研究代表者は研究開始当初、京都大学化学研究所の金久・五斗研究室の助教であった。当研究室では、世界中の研究者により解読された多様な生物種のゲノム情報を収集し、遺伝子オーソログ(共通祖先に由来し機能が同一と考えられる遺伝子群)情報を手作業でアノテーション(注釈付け)する KOALA や、自動アノテーションツール KAAS など、独自のアノテーション環境が整備され KEGG データベースとして公開・運営されていた。ところが系統分類全体を考えると、遺伝子情報が解読された昆虫は少数の属や科に偏っていた。ショウジョウバエ *Drosophila melanogaster* はモデル生物であり、他の 12 種の *Drosophila* 属ゲノムは *D. melanogaster* ゲノムの解釈を助けるために読まれた。またドクチョウ 12 種の遺伝子情報は集中的に読まれているがひとつの科に集中していた。他の昆虫種の遺伝子情報は散在し、昆虫だけを横断的に検索できる状態ではなかった。

昆虫は近縁種間でハイブリッドを形成する例が多く知られており、完全ゲノム解読は難しいとされている。しかし、ドラフトゲノムや遺伝子オーソログ程度の情報であっても、多数の昆虫種を整理し種横断的な検索を可能になれば昆虫の多様性を理解するにあたって利点が大きいと考えた。例えば昆虫は多様な化学コミュニケーションを行うが、化学受容体は一次配列の類似性に乏しく、他種生物の既知遺伝子を手掛かりに利用しがたいという問題点があった。特に味覚受容体は発現量が少ないなどの問題も加わり実験的手法による探索は困難を極め、大部分はゲノムから情報学的に検出されている。ドラフトゲノムや遺伝子オーソログ情報があれば、このような遺伝子についても発見しうることが期待され、非モデル生物の研究を促進すると研究代表者は考えた。

2. 研究の目的

日本は昆虫研究が盛んであり研究者も多く、蓄積した標本数や知見そして人的資源も比較的豊富であるが、分子生物学、特に次世代シーケンサーを用いた研究との連携が必ずしも進んでいるとは言えない状況にあった。そこで研究代表者は、まず、ゲノムが読まれていない昆虫種のうち、昆虫研究のデ

ータが蓄積している一連のグループのドラフトゲノムおよびトランスクリプトームを集中的に解読することで、昆虫ゲノム情報の分類系統的偏りを少なくすると同時に、これまでの昆虫研究の蓄積を活用できないかと考えた。具体的には、日本に生息しており、植生との関係がよく知られており、個体サイズも比較的大きめでゲノム抽出が行いやすいアゲハチョウ(ナミアゲハ、キアゲハなど)の鱗翅目昆虫(チョウ目昆虫)が、このアイデアを実行するのに適したグループであると考えた。

本研究では新型シーケンサー-Miseq を用いて、これら昆虫種から新規にゲノム・トランスクリプトーム解読を行い、既に公開されている昆虫も含めた遺伝子オーソログのライブラリーを作成するとともに、国内外の昆虫研究で蓄積された昆虫の植生や生態などの有用な情報を広く収集し、これらの情報を相互に関連付けしたデータベースを整備し公開することで、昆虫学の膨大な知識に発展させることを目的とする。

3. 研究の方法

本研究では、新規ゲノム・トランスクリプトーム解読と、既知ゲノム・トランスクリプトーム情報の収集を並行して行なった。また、系統分類、食草および関連化合物の情報の収集も広く行なった。

[昆虫標本の選定と採集]

当時ゲノムが読まれていなかった昆虫種のうち、比較的採集が容易で、昆虫研究のデータが蓄積しており、系統的にも重要な種としてナミアゲハおよびキアゲハを選定し採集した。

[新型シーケンサーによる新規ゲノム・トランスクリプトーム解読]

JT 生命誌研究館で導入されている新型シーケンサーである illumina 社の MiSeq を用いてゲノムおよびトランスクリプトームを解読した。いずれも成虫を用い、トランスクリプトームは、食草認識に重要とされていた前脚と、比較として触覚からの抽出を行った。

[新規ゲノム・トランスクリプトームのアセンブリおよび遺伝子予測]

ゲノムアセンブリとして Platanus、トランスクリプトームアセンブリとして Trinity をそれぞれ用いてアセンブルを行った。得られた配列情報に対し Smith-Waterman アルゴリズムによる類似度比較計算を行い、遺伝子予測および推定機能のアノテーションを行った。

また、研究代表者以外の世界各国の研究者らが新型シーケンサーで解読した遺伝子情報の公共レポジトリである Sequence

Read Archive (SRA) から昆虫のトランスクリプトーム情報を網羅的に取得し、同様に Trinity を用いたアセンブル、Smith-Waterman アルゴリズムによる類似度比較計算、推定機能のアノテーションを行った。

[昆虫遺伝子ファミリーのライブラリー作成とデータベース化]

新規ゲノム・トランスクリプトーム解析により得られた遺伝子情報と、米国 NCBI から得られた既に世界各国の研究者により解読され公開されている昆虫タンパク質情報をもとに、Smith-Waterman アルゴリズムによる類似性比較を行い、機能未知・既知にかかわらず昆虫種横断的な遺伝子ファミリーライブラリーを作成した。また、どの種がどの遺伝子ファミリーを持っているかという系統プロファイル情報を作成し、次に示す「遺伝子関連以外の情報」と対応付けデータベース化した。

[遺伝子関連以外の情報の収集]

鱗翅目昆虫を中心に、草食昆虫のその特異的な食草との関係を日本で出版されている種々の文献から情報収集を行い、昆虫種と植物種（あるいは属）の二項関係データとして整理した。

また、植物化合物の情報を KNApSACk データベースから、昆虫化合物の情報を Pherobase から収集し、昆虫・食草の二項関係データを統合してネットワーク解析を行った。

4. 研究成果

ナミアゲハから得られたゲノム DNA 配列 (paired-end, platanus_trim でトリミング後のサイズ 9.5G と 8.1G) を Platanus でアセンブルした結果、総延長 284,686,382 塩基 (最大長 120,232 塩基、N50 値 2,128) のコンティグを得た。同様に、前脚からのトランスクリプトーム RNA 配列、および触覚からのトランスクリプトーム配列を Trinity によりアセンブルした結果、それぞれ 25450 本、26097 本のタンパク質配列を得た。

食草選択に重要と考えられている前脚の発現データと、触覚の発現データを比較した結果、chemosensory タンパク質や gustatory receptor タンパク質と高い相同性を示す mRNA の種類の数は大きな変化を示さなかったが、orodant receptor タンパク質と高い相同性を示す mRNA の種類の数は触覚のほうが多く検出された。

引き続き、これら新規に解読した RNA-seq データと、SRA 公共データリポジトリに公開されている多様な昆虫種の RNA-seq とを統合して、Trinity によるアセンブル、Smith-Waterman アルゴリズムを用いた簡易アノテーションおよびクラスタリングを

行い、鱗翅目昆虫の遺伝子オースログとその食草選択との関連性を解析した。食草選択と植物中の化合物との関係において最も詳しく研究されている例の一つであるカラシ油配糖体について、その代謝に関係することが知られているミロシナーゼ酵素はゲノム中に存在が確認されたがトランスクリプトーム中ではいずれも発現が確認されなかった。また、カラシ油配糖体を無毒化するのに用いることが知られているシロチヨウ特異的 NSP タンパク質はゲノム中にもトランスクリプトーム中にも検出されなかった。同様の解析を、SRA に登録されている多種の鱗翅目昆虫に対して行ない (この中にシロチヨウは含まれていない) どの鱗翅目昆虫にも NSP が含まれていないことが確認された。

続いて、収集した鱗翅目昆虫とその食草の関係からネットワーク解析を行い、昆虫の系統分類と食草植物の系統分類から、特定の昆虫ファミリーに特異的な植物化合物の抽出を行った。結果として、シロチヨウ科の食草特異的な化合物として種々のカラシ油配糖体およびフラボノイド化合物が得られた。同様に、シジミチヨウ科の食草特異的な化合物として genistein が、アゲハチヨウ科の食草特異的な化合物として skimmianine が見出された。これらの食草化合物の代謝に用いられる酵素としてシトクロム P450 が考えられ、今後はその発現変化について解析予定である。

これらの新規シーケンシングから得られたタンパク質配列データ、SRA から得られたタンパク質配列データおよびアノテーション結果、そしてその昆虫種、その食草や化合物のデータは InsectInDB <http://insect-plant.org> としてまとめてあり、現時点で論文執筆中である。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文](計 0 件)

[学会発表](計 26 件)

1. Masaaki Kotera, Katsuhisa Ozaki, and Ai Muto. "Integrated database of insect genomes and ortholog genes, chemical interaction networks between insects and plants, and prediction of plant metabolic pathways." The 24th Annual International Conference on Intelligent Systems for Molecular Biology (ISMB) 2016, Orland, USA. (2016年7月発表予定)
2. Katsuhisa Ozaki, Masasuke Ryuda, Masaaki Kotera, Ai Muto, Hiroshi

- Yoshikawa. "Evolution of gustatory sensing mechanism for host selection in swallowtail butterflies." International Symposium on Olfaction and Taste, ISOT, 2016, Tokyo.
3. 小寺正明、武藤愛、尾崎克久. 「データ統合とオミックス解析による昆虫・植物間相互作用解析と昆虫代謝予測」日本昆虫学会第 76 回大会・第 60 回日本応用動物昆虫学会大会合同大会 2016 大阪
 4. 尾崎克久、小寺正明、武藤愛、吉川寛. 「アゲハチョウの食草選択に関わる化学感受容体遺伝子の種間比較」日本昆虫学会第 76 回大会・第 60 回日本応用動物昆虫学会大会合同大会 2016 大阪
 5. 尾崎克久、小寺正明、武藤愛、吉川寛. 「NGS を用いたアゲハチョウの食草認識に関わる遺伝子群の網羅的探索と比較」第 38 回日本分子生物学会年会・第 88 回日本生化学回大会合同大会 2015 神戸
 6. 小寺正明. 「化学情報学と生命情報学の融合による群集生態学」第 38 回ケモインフォマティクス討論会 2015 東京
 7. 尾崎克久、龍田勝輔、武藤愛、小寺正明、吉川寛. 「アゲハチョウの食草選択に関わる化学感受容体遺伝子」日本進化学会 2015 東京
 8. 小寺正明、武藤愛、尾崎克久. 「化学生態学ネットワークの理解のための昆虫ゲノム・植物メタボローム解析」NGS 現場の会 第 4 回研究会 2015 つくば
 9. 武藤愛、尾崎克久、小寺正明. 「Sequence Read Archive からの昆虫の植物化合物代謝遺伝子の探索」NGS 現場の会 第 4 回研究会 2015 つくば
 10. 尾崎克久、武藤愛、小寺正明. 「昆虫化学感受容体遺伝子を見つけるための NGS 利用と Trinity アセンブルの工夫」NGS 現場の会 第 4 回研究会 2015 つくば
 11. 尾崎克久、武藤愛、小寺正明、吉川 寛. 「NGS を使ったアゲハチョウの食草選択に関わる化学感受容体遺伝子の探索」第 59 回日本応用動物昆虫学会大会 2015 山形
 12. Ai Muto, Katsuhisa Ozaki, Masaaki Kotera. "Development of insect orthologue search database" GIW / ISCB-Asia 2014 Tokyo
 13. 尾崎克久、小寺正明、武藤愛. 「ゲノム解読と RNA-seq によるアゲハチョウの食草選択に関わる化学感受容体遺伝子候補の探索」第 37 回日本分子生物学会 2014 横浜
 14. 武藤愛、尾崎克久、小寺正明. 「鱗翅目昆虫とその食草情報のデータベース化」生命医薬情報学連合大会 2014 仙台
 15. 尾崎克久、龍田勝輔、小寺正明、武藤愛、吉川寛. "Evolution of gustatory sensing mechanism for host recognition in swallowtail butterflies" 第 16 回日本進化学会 2014 大阪府高槻市
 16. 武藤愛、尾崎克久、小寺正明. 「鱗翅目昆虫食草データベースの構築」第 16 回日本進化学会 2014 大阪府高槻市
 17. 小寺正明. 「昆虫植物間相互作用データベースとマルチオミックス基盤」生命情報若手の会 2014 神戸
 18. Masaaki Kotera, Ai Muto, Katsuhisa Ozaki. "Genome informatics platform of lepidopteran insects". The 22nd Annual International Conference on Intelligent Systems for Molecular Biology (ISMB 2014), Boston, USA, 2014.
 19. Masaaki Kotera, Katsuhisa Ozaki, Ai Muto, Susumu Goto. "Development of a Cross-species Search Database for Insect Genes." 第 58 回応用動物昆虫学会大会 2014 高知
 20. 尾崎克久、龍田勝輔、武藤愛、小寺正明、吉澤靖貴、吉川寛. 「アゲハチョウ食性進化の仕組み」第 58 回応用動物昆虫学会大会 2014 高知
 21. 武藤愛、尾崎克久、守屋勇樹、五斗進、小寺正明. 「アゲハチョウ科昆虫ドラフトゲノムの解読」第 58 回応用動物昆虫学会大会 2014 高知
 22. 小寺正明. 「昆虫種横断遺伝子検索データベースの開発」第 36 回日本分子生物学会年会 2013 神戸
 23. Ai Muto, Katsuhisa Ozaki, Susumu Goto Masaaki Kotera. "Development of lepidopteran insect ortholog database." The 24th International Conference on Genome Informatics (GIW2013), Dec 16-18, 2013, Singapore.
 24. Ai Muto, Katsuhisa Ozaki, Susumu Goto Masaaki Kotera. "Construction of an ortholog database of lepidopteran insects." The 36th Annual Meeting of the Molecular Biology Society of Japan (MBSJ2013), Dec 3-6, 2013, Kobe, Japan.

25. Ai Muto, Katsuhisa Ozaki, Susumu Goto Masaaki Kotera. "Development of a Cross-species Search Database for Insect Genes." 2013 Annual Convention of Japanese Society for Bioinformatics (JSBi2013) / Joint Conference on Informatics in Biology, Medicine and Pharmacology, Oct 28-31, 2013, Tokyo, Japan.
26. Masaaki Kotera, Ai Muto. "Lepidopteran insects orthologue database." NGS 現場の会 第3 研究会 2013 神戸

〔図書〕(計 0 件)

〔産業財産権〕

出願状況(計 0 件)

取得状況(計 0 件)

〔その他〕

ホームページ等

InsectInDB <http://insect-plant.org>

6. 研究組織

(1)研究代表者

小寺 正明

東京工業大学・大学院生命理工学研究科・
講師

研究者番号：90643669

(2)研究分担者 なし

(3)連携研究者 なし