

科学研究費助成事業 研究成果報告書

平成 30 年 6 月 21 日現在

機関番号：82657

研究種目：基盤研究(B) (一般)

研究期間：2014～2017

課題番号：26280105

研究課題名(和文) 抗原変異遺伝子群の進化メカニズムの解明

研究課題名(英文) Evolutionary mechanism of antigenic variation genes

研究代表者

五斗 進 (Goto, Susumu)

大学共同利用機関法人情報・システム研究機構(機構本部施設等)・データサイエンス共同利用基盤施設・教授

研究者番号：40263149

交付決定額(研究期間全体)：(直接経費) 11,700,000円

研究成果の概要(和文)：様々な感染症を引き起こす病原微生物が、ヒトをはじめとする宿主の免疫系から逃れるための仕組みの一つである抗原変異という現象に着目し、それが関与するタンパク質の進化メカニズムをゲノム情報から明らかにするためのデータベースを構築した。特にマラリア原虫に対する情報について関連遺伝子を網羅的に収集し、種間比較などができるようにした。また、それらの進化的な関係を解析するためのパイプラインを構築し、人類進化の解析や物質生産に関わる生合成経路の進化解析に応用した。

研究成果の概要(英文)：One of the mechanisms for microbial parasites to evade host immune systems is antigenic variation, which alters target antigens of the host immune attack to escape from their recognition. We have developed a database named varDB by accumulating information on proteins regarding antigenic variations and analyzed their amino acid and nucleotide sequences to elucidate their evolutionary mechanisms. Mainly focusing on the malaria parasites, Plasmodium species, whose complete genome sequences are available, we have collected related sequence information based on the literature search and ortholog detection, to make the interspecies comparisons available. We have also developed a computational pipeline for analyzing their evolutionary relationships and applied it to human evolution and biosynthetic pathway evolution analyses.

研究分野：バイオインフォマティクス

キーワード：抗原変異 比較ゲノム 遺伝子ファミリー 進化 病原微生物

1. 研究開始当初の背景

インフルエンザ、AIDS、マラリア、結核などの感染症は、現在も世界中で猛威を振るっている。例えば、マラリアでは30秒に一人の子供が命を落としており、特に貧困地域における経済発展を妨げる原因となっている。最も毒性の強いマラリア原虫である *Plasmodium falciparum* は発展途上国における公衆衛生問題の一つにもなっている[1]。また、日本でも新型インフルエンザへの対応など感染症対策は重要な課題の一つである。これらの病原微生物やウイルスは、ヒトをはじめとする宿主に感染するために、高度に発達した宿主の免疫系と戦う必要がある。宿主の免疫系に対抗するために病原微生物が獲得した仕組みの一つが抗原変異である。抗原変異とは、抗体に認識される細胞表面抗原を多様化するために、株ごと、個体ごと、生活環境ごとに細胞表面タンパク質を入れ替える仕組みである[2]。

これまでに抗原変異に関与する多くの遺伝子ファミリーが同定されてきた(表1にゲノムが決定された原虫での例を示す)。しかし、それらがどのようにして多様性を獲得してきたか、また、どのようにして発現するタンパク質を入れ替えているかについての詳細なメカニズムは vsg などの一部を除いては明らかになっていない。

表1

生物種	遺伝子ファミリー	遺伝子数	関連感染症
<i>Plasmodium falciparum</i>	var	67	ヒトマラリア
	rifin/stevor	226	
<i>Plasmodium vivax</i>	vir	297	霊長類マラリア
<i>Plasmodium knowlesi</i>	kir	63	
<i>Plasmodium berghei</i>	bir	131	げっ歯類マラリア
<i>Plasmodium chabaudi</i>	cir	195	
<i>Babesia bovis</i>	ves	140	バベシア症
<i>Trypanosoma brucei</i>	vsg	214	アフリカ睡眠病
<i>Giardia lamblia</i>	vsp	390	ジアルジア症

ゲノムプロジェクトの進展により、表1の生物種以外にも抗原変異の仕組みを持つ生物種のゲノムが多数決定されている。しかしながら、抗原変異遺伝子群は表1に示すように多くのパラログ遺伝子を含む遺伝子ファミリーを構成しているため、オーソログ同定は簡単ではない。また、同定できたとしても、多数からなる各遺伝子にどのような機能の違いがあるのかは詳細には明らかになっていないのが現状である。

本研究では、病原微生物の比較ゲノム解析により、抗原変異の進化・発現メカニズムの解明を目指す。具体的には、抗原変異遺伝子ファミリーの系統解析、遺伝子発現制御配列モチーフの解析、ファミリーに属する遺伝子のゲノム上の位置・発現パターン・系統プロファイルとの関連解析をするとともに、解析アルゴリズムを開発する。

2. 研究の目的

ゲノムが決定された生物種の遺伝子配列情報を機能と関連付けている KEGG データベース[3] および既知の抗原変異遺伝子の配列情報を収集した varDB[4,5] を用いた配列解析を中心とした方法を用いて、以下の3点を明らかにする。

- (1) 抗原変異遺伝子ファミリーは生物種によってそのメンバー数や配列の多様性が異なる。そこで、生物種間による抗原変異の仕組みの違いを見るために、感染宿主の重症度や生物種の異なる複数種のマラリア原虫から抗原変異遺伝子ファミリーを選び、これらの系統関係を配列の保存領域と可変領域に分けて解析することによって、機能やその分化過程を明らかにする。
- (2) 抗原変異遺伝子は同じファミリー内の遺伝子の発現パターンが多様化しているため、その発現メカニズムの解明は重要である。そこで、(1)で解析した遺伝子ファミリーについて、(a)ゲノム上の上流塩基配列モチーフ、(b)コドン利用頻度の違いを調べ、そのパターンと実際の発現パターンを比較することによって、発現制御機構の解明を目指す。
- (3) 上記2つの解析結果を元にして、同じ遺伝子ファミリーに属する遺伝子同士がどのように多様化し、どのように発現制御されているかを遺伝子間のネットワークとして表現する枠組みを開発し、抗原変異の進化過程の解明を目指す。

上記の解析を行うと同時に、解析で必要となった新規データなどの情報を用いて varDB を更新・改良する。

3. 研究の方法

- (1) 本研究では、まずマラリア原虫の抗原変異遺伝子ファミリーの一つである *Plasmodium interspersed repeats (pir)* に着目し、そのメンバー遺伝子の配列を保存領域、抗原変異部位の可変領域などに分けて、詳細な解析を行うことにより、本遺伝子ファミリーの系統関係を明らかにする。さらに、ファミリーメンバーの数の分布から、pir 遺伝子ファミリーの祖先型と、このファミリーが本来持っていた機能の手がかりを探索する。

また、pir 遺伝子ファミリー解析で適用する手法は、他の抗原変異遺伝子ファミリーにも適用可能である。特に、マラリア原虫と同じアピコンプレクサ原虫であるタイレリアやトキソプラズマはゲノムも決定されており、同様の遺伝子ファミ

リーを持つので、これらの生物種に対しても、抗原変異遺伝子ファミリーの系統関係と機能分化過程を明らかにする。

(2) 抗原変異遺伝子ファミリーの発現機構解明のための上流塩基配列モチーフ解析

pir 遺伝子ファミリーなど抗原変異に関わる遺伝子は、一度に一つの遺伝子のみを発現するという特徴を持つため、ファミリー内の遺伝子発現を制御している何らかの機構が存在するはずである。すでに varDB と KEGG で各遺伝子とゲノム上の位置は対応付けされており、各遺伝子の上流配列や非コード領域配列を取得することは、比較的容易であるので、これらの配列に対して、各遺伝子の発現を特徴づけるモチーフを探索する。モチーフ探索には MEME プログラム (<http://meme.sdsc.edu/>) を適用する予定であるが、それに加えて、平成 26 年度の系統解析の結果や公開されているマイクロアレイの発現データを組み合わせることにより、精度を上げる方法を開発する。

(3) 抗原変異遺伝子ファミリーの発現機構解明のためのコドン利用頻度解析

抗原変異遺伝子ファミリーの多様性はアミノ酸配列レベルだけでなく、塩基配列レベルで大きく異なっている。特に可変部位の中でもコドン利用頻度の違いが発現や機能の違いに関連している場合がある。そこで、この仮説が抗原変異遺伝子ファミリーにも当てはまるかどうかを解析する。本解析では、アミノ酸配列や塩基配列のマルチプルアライメントをコドンベースのアライメントに変換する PAL2NAL プログラム (<http://www.bork.embl.de/pal2nal/>) と塩基配列のマルチプルアライメント時にアミノ酸に翻訳した配列を考慮する RevTrans プログラム (<http://www.cbs.dtu.dk/services/RevTrans-2.0/>) を主に使用する予定であるが、必要に応じて新規に解析プログラムを開発する。

(4) 抗原変異遺伝子ファミリーの発現機構解明のためのネットワーク解析

同じ遺伝子ファミリーに属する遺伝子がどのように多様化し、どのように発現制御されているかを遺伝子間のネットワークで表現する枠組みを開発して、上記の 2 つの配列解析結果と平成 25 年度の系統解析結果を可視化することによって、抗原変異の進化について解析する。具体的には、遺伝子ファミリーメンバーの多様性を作り出す仕組みである、遺伝子重複 (gene duplication) と遺伝子変換 (gene conversion) に着目する。特に遺伝子変換では、同じ発現制御を受ける遺伝子群が

多様性を持つ原因ともなりうるため、その影響を評価する仕組みを開発する。

- (5) varDB と KEGG を通じた研究成果の公開を随時するとともに、データの更新やインタフェースの改良も進めることにより、広く世界中の研究者の研究促進も図る。また、varDB に登録されていない抗原変異遺伝子ファミリーのデータ収集を行い、varDB データ登録パイプラインにしたがって、新規データを登録する。

4. 研究成果

抗原変異遺伝子ファミリーの中からマラリア原虫など複数の近縁生物種で保存されているものに着目して系統解析を行い、その機能分化についての手がかりを探索した。抗原変異遺伝子ファミリーは生物種によってそのメンバー数や配列の多様性が異なる。そこで、生物種間による抗原変異の仕組みの違いを調べるために、複数種のマラリア原虫から抗原変異遺伝子ファミリーを選び、これらの系統関係を解析した。

全ゲノムが決定され、KEGG データベースに登録されている病原微生物に着目し、他のグループによってすでに解析が進められているマラリア原虫の *pir* 遺伝子ファミリーについて、生物種を増やしてデータを整備し、保存部位について解析を行った。具体的には、ヒトに感染する *Plasmodium falciparum* の 3 つの株に加え、霊長類に感染するマラリア原虫 3 種と齧歯類に感染するマラリア原虫 3 種の *pir* 遺伝子ファミリーについて、KEGG で計算されたオーソログ、パラログ情報を元に候補遺伝子を抽出し、既知のアミノ酸配列モチーフの有無を元にその関係をまとめた。

同様の手法は、他の遺伝子ファミリーについても適用することができるため、マラリア原虫と同じアピコンプレクサ原虫である *Theileria equi* (KEGG では *Babesia equi*) が持つ equi merozoit antigen (EMA) 遺伝子ファミリーなど 5 つの遺伝子ファミリーについても KEGG データベースからオーソログ遺伝子を抽出し、保存モチーフを探索した。これらの結果は varDB に反映した。また、研究期間中に新たに四日熱マラリア原虫と卵形マラリア原虫のゲノム論文が報告されたため、抗原変異遺伝子群および関連情報を GeneDB から取得し varDB に反映した。

病原性原生生物の抗原変異遺伝子の上流塩基配列とコドン利用頻度を解析し、発現機構の解明を目指した。*pir* 遺伝子ファミリーなどの抗原変異に関わる遺伝子ファミリーは、メンバーの遺伝子を一度に一つだけ発現す

るという特徴を持つため、ファミリーの中の遺伝子発現を制御している何らかの機構が存在しているはずである。収集した遺伝子ファミリーについて、本プロジェクトで開発している varDB と KEGG データベースとの対応を取ることで、KEGG に登録されているゲノム配列情報から各遺伝子の上流配列、コード領域配列、非コード領域配列を取得した。上流塩基配列については MEME プログラムを用いて共通モチーフの探索を試みたが、遺伝子タイプごとに評価する必要があることが判明した。今後、系統樹解析の結果などと比較してサブファミリー毎に解析することを検討する。

コドン利用頻度解析については、本研究期間内で解析を終えることができなかった。一方で、新たな抗原変異遺伝子ファミリーの探索を試みた。具体的には KEGG Ortholog Cluster (OC) データベースに登録されているマルチ遺伝子ファミリークラスタから原生生物種が単独の種または少数の種でのみパラログ遺伝子を増やしているものを抽出した。その結果、21 のマルチ遺伝子ファミリークラスタについて、既知の抗原変異遺伝子を含む 190 のクラスタが得られた。各遺伝子ファミリーに代表的なクラスタについては系統解析を行い、特定の生物種でのみ特異的にパラログが生成するものが複数存在することが確認された。既知の抗原変異遺伝子と類似したモチーフを持つものも見つかったため、それらについては varDB に反映した。今後はそれらの進化と機能についてさらに解析する予定である。

最終年度には、研究代表者の異動に伴い、varDB のインタフェースを刷新するとともに新規にテストサイト (<http://vardb.dbcls.jp/>) を構築した。将来的には vardb.org に統合する予定である。またテストサイトには追加で収集した *Plasmodium* 属の抗原変異遺伝子配列とそのアライメント情報、系統解析の結果を反映した。上流配列解析を含めた比較ゲノム解析ができるように、ライフサイエンス統合データベースセンターで開発されている TogoGenome の API を利用し、ゲノム上の指定した位置の塩基配列を所得する仕組みを varDB に組み込み、抗原変異遺伝子の上流配列を取得できるようにした。さらに、抗原変異遺伝子ファミリーの発現機構解明のためのネットワーク解析を進めた。ネットワーク可視化のツールについては Cytoscape など既存のツールを使用することが適当であることが判明したので見送った。

本研究の発展として、抗原変異遺伝子自体の解析に加えて、系統解析手法の機能予測への応用、および、主要宿主であるヒトの遺伝子変異に関する解析で進展が得られた。前者に

関しては、マルチ遺伝子ファミリーの一つである、Ⅱ型ポリケチド合成酵素 (PKS) を、ゲノムが決定された生物種から網羅的に取得するとともに、機能既知の遺伝子を文献から網羅的に集めた。パラログ遺伝子群の機能分類を促進するために、Ⅱ型 PKS の反応分類方法を確立し、系統解析結果と対応付けた (発表論文)。その結果、系統関係だけでは対応付けできない機能が明らかになったため、立体構造情報から得られた相補的に変異する配列上の位置を用いた機能予測法を開発した (発表論文)。また、後者に関しては、ハプロタイプの情報を集めた国際的なデータベースである HapMap からアフリカ由来、ヨーロッパ由来、アジア由来の人種間の差を計算し、古い年代に確立したハプロタイプブロックを抽出するためのパイプラインを開発した。得られたハプロタイプブロックに保存されている遺伝子群とマラリア感染の関係の解析に応用した (発表論文)。

<引用文献>

- Hay, S. I., et al. *Lancet Infect. Dis.* (2004) 4:327-336.
Craig, A. and Scherf, A., eds. *Antigenic Variation* (2003). Academic Press.
Kanehisa, M., et al. *Nucleic Acids Res.* (2000) 40:D108-114.
Diez, D., et al. *Acta Trop.* (2010) 114:144-151.
Hayes, C. N., et al. *Bioinformatics* (2008) 24:2564-2565.

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3件)

Onuki, R., Yamaguchi, R., Shibuya, T., Kanehisa, M. and Goto, S.; Revealing phenotype-associated functional differences by genome-wide scan of ancient haplotype blocks. *PLOS ONE*, 12:e0176530 (2017).
DOI: 10.1371/journal.pone.0176530

Shimizu, Y., Ogata, H. and Goto, S.; Discriminating the reaction types of plant type III polyketide synthases. *Bioinformatics*, 33:1937-1943 (2017).
DOI: 10.1093/bioinformatics/btx112

Shimizu, Y., Ogata, H. and Goto, S.; Type III polyketide synthases: functional classification and phylogenomics. *ChemBioChem* 18:50-65 (2017).

DOI: 10.1002/cbic.201600522

〔その他〕

ホームページ等

抗原変異遺伝子ファミリーデータベース

varDB、<http://vardb.org/>

現在、再構築中の開発版は

<https://vardb.dbcls.jp/>

6．研究組織

(1)研究代表者

五斗 進 (GOTO, Susumu)

情報・システム研究機構・データサイエンス

共同利用基盤施設・教授

研究者番号：40263149