

研究課題

内在性ウイルス配列の探索による節足動物ウイルスの多様性および進化の解析

堀江 真行

大阪公立大学・大学院獣医学研究科

研究目的

これまでのヒトや家畜等の新興感染症の多くは、ヒト以外の動物を自然宿主とするウイルスによって引き起こされてきた。このような新興感染症を引き起こすウイルスは、顕著な被害が出たときにはじめて同定される。しかし、このような対応では感染源となる動物の特定などにおいて遅れが生じ、今後の新興感染症の発生時に被害が拡大するおそれがある。そのため、ウイルス感染症の制御には、どの動物に、どのようなウイルスが感染しているのかを事前に知っておくことが有用である。すなわち、ウイルスの多様性を徹底的に解明する必要がある。実際に、関連した種々のプロジェクトが世界中で展開されている。しかし、未だウイルスの探索は十分ではなく、地球上のウイルスの 99.99% 以上は未同定であると考えられている (Geoghegan et al., *Open Biol.* 2017)。

節足動物はジカウイルスやデングウイルスのような新興感染症から、日本脳炎や黄熱といった様々なウイルス感染症を媒介する。節足動物においても上記の通り、世界中においてウイルスの探索が行われている。しかしその探索は十分ではなく、さらなるウイルス探索が必要であると考えられている。

現在のウイルス探索法の主流はウイルス分離およびハイスループットシーケンズ解析である。両者ともに様々なメリット・デメリットが存在するが、どちらの手法においても、採材した個体に感染しているウイルスのみしか検出ができないという欠点がある。ウイルスの多様性の解明には、過去に感染してきたウイルスについて知ることが必須であるが、これらの手法では過去のウイルス感染について知ることはできない。

生物のゲノムにはウイルス由来の遺伝子配列が多数存在し、それらは内在性ウイルス様配列 (EVE: endogenous viral elements) と呼ばれる。EVE は祖先生物におけるウイルス感染を示す「ウイルスの分子化石」であり、EVE を解析することによって過去のウイルス感染に関する貴重な情報 (年代、宿主、遺伝子配列) が得られる (図 1)。申請者らはこれまでに、ボルナウイルスという RNA ウイルスをモデルとして、ゲノムデータベースに存在する EVE を網羅的に探索・解析し、脊椎動物における約 1 億年間のボルナウイルス感染履歴を明らかにした (Kawasaki et al., *PNAS* 2021)。これらの研究では、これまでに魚やへびにしか感染しないと思われていた系統のウイルスが、過去には霊長類やコウモリを含む様々な哺乳動物に感染していたことがわかり、現代においても哺乳動

物に未知のウイルスが存在することが示唆された。このように EVE を解析することによって、現存するウイルスの研究のみからは得られない貴重な知見が得られる。

本研究では節足動物、特にヒトや哺乳動物に感染症を媒介し得る蚊を中心とした節足動物におけるウイルス叢の解明を目的とし、EVE を利用して過去のウイルス感染の解明さらには分子系統解析等により有用な情報を得る。本研究では昨年度に行った新規 EVE 検出法のさらなる応用に加え、公共データベースのデータへと応用するための準備を行った。

研究内容と研究成果

はじめに、EVE 検出方法の改良とその検討を行った。これまでの EVE の検出方法の主流は、現代のウイルスのタンパク質のアミノ酸配列をクエリー、生物のゲノムの塩基配列をデータベースとして用いる tBLAST 検索であった (Katzourakis et al., PLoS Genet 2010)。tBLASTn は最も標準的な方法であるものの、フレームシフトが考慮されないため、変異によってフレームシフトが頻繁に起こっている EVE については十分に検出できない可能性がある。そのため、存在する EVE を見逃す可能性がある。FASTA は BLAST と同様に古くから使用されている配列類似性検索ソフトである (Pearson Curr Protoc Bioinformatics 2016)。FASTA パッケージでは tBLASTn と同様にアミノ酸配列をクエリーとした塩基配列データベース検索を行う tfastx および tfasty プログラムが存在する。特筆すべきは、tfastx および tfasty ではフレームシフトを考慮した検索を行うことができるということである (Pearson et al., Genomics 1997)。そのため、上記の tBLASTn の欠点を補い、今まで見逃していた EVE を検出し、より網羅的に過去のウイルス探索ができる可能性がある。実際、昨年度の検討では tfastx および tfasty は tBLASTn よりも効率的に EVE を検出できることを示した。そこで本年度は、さらなる tfastx および tfasty を用いたさらなる検討を行った。

本研究では既知の RNA ウイルスのアミノ酸配列をクエリーとした tfastx および tfasty 検索により、ヒトスジシマカ (*Aedes aegypti*) およびネッタシマカ (*Aedes albopictus*) のゲノムにおける EVE の探索を行った。多数のヒット配列が得られたものの、その数が膨大であり、現在、詳細なデータ解析を行っている。

さらに、上記の検索手法を公共データベースのハイスループットシーケンスデータへと応用を試みた。現在、様々な節足動物のゲノムが公開されているが、あくまでも代表配列であり、EVE の多型などは考慮されていない。そのため、公共データベースのハイスループットシーケンスのデータの tfastx および tfasty 検索を行うために、de novo アセンブリを行った。現在、NCBI SRA に

ある節足動物由来データのうち、17567 のハイスループットシーケンスのデータのアセンブルを終えた。今後、tfastx と tfasty による大規模検索を行う予定である。

成果発表

なし

今後の問題点

本研究では特にスパコンの混雑のために、当初想定した以上に検索と de novo アセンブリに時間がかかってしまった。また、tfastx や tfasty によって得られたデータについてもヒット数が多く手動での解析が困難であるため、検索後の検証についても何らかの自動化、あるいは半自動化するようなプログラムを作成する必要がある。