

食中毒調査の迅速化・高度化及び広域食中毒発生時の早期探知等に資する研究
分担課題 腸管出血性大腸菌等の検査法(全ゲノム解析)の開発
研究分担者 林 哲也 九州大学大学院医学研究院・教授

研究要旨

腸管出血性大腸菌（EHEC）感染症の事例調査のために各種の分子型別法が開発されてきたが、国内では、反復配列多型解析法（MLVA 法）がその迅速性・精微性から主に用いられてきた。本研究班では、MLVA に関して蓄積されてきたデータの検証、型が一致した場合等の評価法の整理、集団事例における有効性の評価、地方衛生研究所における利用促進のための精度管理手法の確立を行うとともに、EHEC 等の検査への全ゲノム解析（WGS）の適用に関する検討を行う。本分担者は、他の分担者とともに後者の課題を担当するとともに、並行して行われる食品及び動物分離菌株の WGS データの収集を随時サポートし、WGS データベースに組み込んで解析する役割も担っている。本年度は、国内で WGS 解析を利用した事例調査を効率的に実施するために求められる解析手法とデータベースの必要要件等を明らかにするため、諸外国における WGS の導入状況、活用状況、解析パイプラインやデータベース整備等の調査を行い、我が国の現状を踏まえた活用法について検討した。また、これまでに蓄積されている国内分離株の WGS データ（本分担者が取得している未発表データを含む）を収集・整理するとともに、公共の WGS 情報データベース（主に NCBI/EMBL/DBJ データベースと ENTEROBASE）から WGS データを収集し、プロトタイプとなるデータベースの構築を開始した。さらに、本分担者や本研究班代表者らが既に収集している分離株の WGS データ取得を開始し、新規性が高く参照配列になりうる株等については、ロングリードシーケンシングを併用して、完全長あるいは完全長に近い配列を取得した。

A. 研究目的

腸管出血性大腸菌（EHEC）感染症の事例調査のために各種の分子型別法が開発されてきたが、反復配列多型解析法（MLVA 法）が迅速性、精微性に優れていることから、国内では主に MLVA 法を用いた解析が行われている。本研究班の目的の 1 つは、MLVA については、蓄積されてきたデータの検証、型が一致した場合等の評価法の整理、集団事例における有効性の評価、地方衛生研究所における利用促進のための精度管理手法の確立である。本研究班のもう一つの目的は「腸管出血性大腸菌等の検査への全ゲノム解析（WGS）の適用に関する検討」であり、本分担者は、この課題を他の分担者とともに担当する。本課題は、国際整合性の観点から実施するものであり、まず海外で展開されている解析手法とデータベース等を

検証し、国内の現状を踏まえた上で、効率的に事例調査を実施するための解析手法とデータベースの必要要件を明らかにすることを旨とする。また、これまでに蓄積された国内分離株のゲノムデータと新たに取得する国内株のデータの整理と解析を行うとともに、海外株のデータも収集する。さらに、これらの検証・解析結果を踏まえて、地方衛生研究所等でも利用可能な解析パイプライン、国際的に整合性のあるデータフォーマットや同一クローンの判定基準等を確立する。本研究班では、これらの 2 課題と関連して、MLVA 及び WGS 解析から得られたデータを利用して感染源や経路経路等に関する後方視的な疫学解析手法を検証することも目的としている。そのため、主感染源である食品からの分離菌株の WGS データ収集を可能とする方策の検討と臨床分離株との比較、同様

に動物由来株のデータ収集を可能とする方策の検討と事例調査への利用を試みる。本分担者は、これらの食品及び動物からの分離菌株の WGS データの収集を随時サポートし、上記の WGS 情報データベースに組み込んで解析する。

B. 研究方法

(1) 国内で WGS 解析を利用した事例調査を効率的に実施するために求められる解析手法とデータベースの必要要件等を明らかにするため、本年度は、諸外国における WGS の導入状況、活用状況、解析パイプラインやデータベース整備等の調査を行い、我が国の現状を踏まえた活用法について検討した。具体的には、英国、他のヨーロッパ諸国、米国、カナダ等で用いられている WGS 解析手法(データの取得方法、解析法と解釈の基準、データベースの整備状況、データフォーマット等も含む)を、論文情報等を基に調査した。データの取得方法や解析法に関しては、WGS の適用を早くから進めている英国 Public Health England の状況を中心に調査したが、同一クローンの判定基準などのデータ解釈に関しては、解析法によって異なる場合もあり、国際的にもコンセンサスが得られていないため、各国の状況あるいは主要機関での状況を調査する必要があった。

(2) これまでに蓄積されている国内分離株の WGS データを収集・整理し、データベースの整備作業を開始した(プロトタイプのデータベース構築)。これらのデータは、分担者が中心となって算出したものが多いが、それ以外のデータは公共の WGS 情報データベース(主に NCBI/EMBL/DBJ のゲノム情報データベースと ENTEROBASE)から収集した。また、分担者が取得している未発表データもデータベースに加えた。さらに、分担者や本研究班の代表者らが既に収集している分離菌株の解析(WGS 情報の取得)を開始した。配列情報の取得には、基本的にはイルミナのシーケンサ(サンプル数に応じて MiSeq, HiSeq, NovaSeq のいずれかを使用し、ショートリード配列を取得)を利用したが、必要に応じて(新規性が高く参照配列になりうる株等)、ナノポア MInION を用いたロングリードシーケンシングを併用して、完全長あるいは完全長に近い配列を取得した。

(倫理面への配慮)
該当しない

C. 研究結果

(1) 諸外国における WGS の導入状況、活用状況、解析パイプラインやデータベース整備等の調査と我が国の現状を踏まえた活用法についての検討:

世界各国でのシーケンシングの主流は、イルミナのシーケンサ (MiSeq, NextSeq, HiSeq, NovaSeq) を用いたショートリード配列の取得であるが、最近では MGI 社のシーケンサを利用したショートリード配列の取得も始まっている。ロングリードシーケンサの利用は研究面にとどまっている。解析パイプラインは研究機関によって様々であり、WGS の取得に関しては、i) 参照配列に対するリード配列のマッピング、ii) リード配列のアセンブリのいずれかである。後者で使用されるアセンブラーは SPAdes と Velvet の使用が多いが、CLC のアセンブラーや Platanus なども使用されている。WGS を用いた菌株の実際の解析に関しては、i) コアゲノム配列に基づく系統解析、ii) コア遺伝子の配列に基づく系統解析、iii) コアゲノム配列に基づく wgMLST (Multi Locus Sequence Typing) のいずれかが使用されている。また、集団事例に関連するクローンの判定基準としては、現時点では、コアゲノムまたはコア遺伝子の SNP 距離 (例えば、5 SNPs) を採用している報告が多いが、解析対象によって異なり、国際的なコンセンサスはない。

(2) 国内・国外分離株の WGS データの収集と解析および基本的な WGS データベースの構築:

主要 EHEC 血清群に関して、本分担者がこれまでに取得したデータ、新規に取得したデータ、公共データベースから取得したデータ(基本的に海外株のデータ)の収集を進めた。現時点での収集数は、0157 は 4,515 株(国内は 2,525 株)、026 は 540 株(国内は 314 株)、0145 は 246 株(国内は 88 株)、0103 は 979 株(国内は 103 株)、0121 は 638 株(国内は 211 株)、0165 は 96 株(国内は 71 株)となっている。最も重要な 0157 に関しては、別プロジェクトで取得した 2005 年、2010 年、2015 年の国内分離株の大部分を網羅する 2,219 株と牛由来の 260 株が含まれる。また、0157 の中で特に病原性が高いと推察されている Clade 8 株に関して、特に重点的な収集・整備と解析を行い、511 株(国内は 150 株)の情報を整備するとともに、Clade 8 内の亜系統を代表する 18 株を選び、ロングリードシーケンシングを行った。現在、ショートリード配列とのハイブリッドアセンブリの作業中である。

D. 考察

(1) 諸外国における WGS の導入状況、活用状況、解析パイプラインやデータベース整備等の調査と我が国の現状を踏まえた活用法についての検討:

WGS 解析の主流は、イルミナのシーケンサを用いたショートリード配列の取得であるが、解析パイプラインは各国あるいは

研究機関によって様々であり、集団事例に関連するクローンの判定基準についても国際的なコンセンサスはない。地方衛生研究所などでは、コア遺伝子セットの配列に基づくwgMLSTが比較的取り入れやすい手法かもしれないが、一般のMLSTのようなSTの整備が行われていないため、国際的なシステムが構築される必要がある。そのため、影響力の大きい米国CDCの方針などに注意して継続的に調査する必要があると思われる。

(2) 国内・国外分離株のWGSデータの収集と解析および基本的なWGSデータベースの構築：

上記のように、主要EHEC血清群のうち、O157、O26、O145、O103、O121、O165のWGSデータの収集・整理を行ったが、公共データベースへの登録が急速に進んでおり、継続的な収集とデータベースのアップデートが必要と思われる。また、公共データベース間でのデータ重複の問題等があるため、収集したデータに関しても、追加の整理作業が必要である。さらに、データベースのスリム化を図るために近縁クローンなどの重複を一定の基準で除外することも検討する必要がある。これらの作業と同時に、参照配列となりうる代表株に関しては、ロングリードシーケンサを用いて完全長あるいはそれに近い高精度配列を準備することも重要であると思われる。この作業においては、EHECのゲノムには複雑な繰り返し配列が存在するため、ショートリード配列とのハイブリッドアセンブリのプログラム等について検討する必要がある。

E. 結論

諸外国の状況に関しては、シーケンシングの主流がイルミナのシーケンサを用いたショートリード配列の取得である一方、解析パイプラインは各国あるいは研究機関によって様々であり、集団事例に関連するクローンの判定基準についても国際的なコンセンサスはないことが明らかになった。また、主要EHEC血清群のうち、O157、O26、O145、O103、O121、O165の国内・国外分離株のWGSデータの収集・整理を行ったが、データ登録が急速に進んでいることから継続的な収集とデータベースのアップデートが必要であり、収集したデータ

に関しても、近縁株の整理によるデータベースのスリム化等、追加の整理作業が必要である。また、参照配列となりうる代表株に関しては、完全長あるいはそれに近い高精度配列を準備することも重要である。

F. 健康危険情報 なし

G. 研究発表

1. 論文発表

なし

2. 学会発表

(1) 林哲也:腸管出血性大腸菌のゲノム解析:次世代シーケンサを用いた感染症と病原体の解析の例として(教育講演).

第90回日本感染症学会西日本地方会学術集会・第63回日本感染症学会中日本地方会学術集会・第68回日本化学療法学会西日本支部総会(合同学術集会), 2020年11月5~7日, 福岡.

(2) 中村佳司, 小椋義俊, 後藤恭宏, 林哲也:プロファージ内プロファージによる大腸菌への3型分泌エフェクターと志賀毒素遺伝子の蓄積メカニズム.

第15回日本ゲノム微生物学会年会, 2021年3月4~6日, 福岡.

(3) 宮田達弥, 小椋義俊, 中村佳司, 後藤恭宏, 吉村大, 伊豫田淳, 伊藤武彦, 大西真, 林哲也:EHEC O157 clade 8のゲノム多様性とStx2とStx2ファージのバリエーション. 第94回日本細菌学会総会, 2021年3月23~25日, 岡山.

(4) 中村佳司, 小椋義俊, 後藤恭宏, 林哲也:プロファージ内プロファージによる大腸菌への3型分泌エフェクターと志賀毒素遺伝子の蓄積メカニズム.

第94回日本細菌学会総会, 2021年3月23~25日, 岡山.

(5) 矢野文悟, 中村佳司, 谷口愛樹, 後藤恭宏, 小椋義俊, 林哲也:腸管出血性大腸菌O26:H11におけるStxファージの遺伝的多様性とダイナミクス.

第94回日本細菌学会総会, 2021年3月23~25日, 岡山.