

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。本研究班は、MLVA に関する蓄積データの検証や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などを行うとともに、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ことを目的としている。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、昨年度に引き続き、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、欧米における WGS の活用状況、使用解析パイプラインやデータベース等の調査を行うとともに、我が国での活用法について検討した。これと並行して、昨年度までに収集している国内分離株の WGS データの整理を進めるとともに、本研究班代表者らが既に収集している分離株の WGS データを追加取得し、公共データベース（NCBI/EMBL/DDBJ と ENTEROBASE）からの WGS データの追加収集も行った。さらに参照配列になりうる株については、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、O111 以外の主な血清型の EHEC の WGS 情報に関するデータベースの構築と更新を行なった。

### A. 研究目的

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 MLVA 法を用いた解析が主に行われている。本研究班の目的の 1 つは、MLVA についての蓄積データの検証や地方衛生研究所での利用促進のための精度管理手法の確立等であるが、もう一つの目的は「EHEC 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、この課題を担当している。本分担課題は、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目指している。これと並行して、これまで（DB）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して EHEC の WGS 情報 DB を構築し、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ことを目標としている。また、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こととなっているため、本分担者は、これらの食品・動物分離株の WGS データ収集を随時サポートし、必要に応じて、上記の WGS 情報 DB に組み込んで解析する役割も担っている。

## 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 DB の必要要件等を明らかにするため、昨年度に引き続き、諸外国（英国、他のヨーロッパ諸国、米国、カナダ）における WGS の活用状況、解析パイプラインや DB 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、WGS データの取得方法、解析法と解釈の基準、利用している DB、データフォーマット等を、論文情報を基に調査した。

(2) 昨年度に引き続き、既に蓄積されている国内分離株の WGS データの収集と整理を行なった。また、一部の血清型については、分担者あるいは本研究班の代表者らによって新たに WGS 情報を算出した。また公共 DB（NCBI/EMBL/DBJ と ENTEROBASE）からデータを追加収集した。WGS 情報の取得は、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。参照配列になりうる株等については、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。また、この解析を通じて、完全長配列取得のためのパイプラインについて比較検討を行った。

（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 の活用状況、解析パイプラインや DB 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：

シーケンシングの主流は、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ことには変化はない。中国製（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一部には見られるが、論文レベルでは使用が拡大している様子はない。ロングリードシーケンサの利用も研究面にとどまっており、一般のサーベイランスでの利用はまだ行われていない。

解析パイプラインは、昨年度に報告したように、研究機関によって様々であり、WGS の取得には、i) 参照配列に対するリードマッピング、ii) リード配列のアッセン

ブリのいずれかであり、後者で使用されるアセンブラには SPAdes と Velvet が多く使われているという現状も変化がない。ただし、小規模な解析では CLC のアセンブラが、また国内（分担者も含む）では Platanus も使用されている。WGS を用いた実際の解析では、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) コアゲノムあるいはコア遺伝子の配列に基づく cgMLST (core gene Multi Locus Sequence Typing) または wg (whole genome) MLST のいずれかが使用されている。集団事例に関連するクローンの判定基準については、コアゲノムまたはコア遺伝子の SNP 距離 (5 SNPs など) 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まだ得られていないのが現状である。この点は、各血清型で使用されるコアゲノムあるいはコア遺伝子についても同様で、こういった配列あるいは遺伝子を使うかについては、国際的なコンセンサスは得られていない。

国内で使用すべき解析パイプラインに関してはまだ検討段階である。

(2) 国内・国外分離株の WGS データの収集・解析とデータベース構築：

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、026 (540 株：国内株は 314)、0145 (246 株：国内株は 88)、0121 (638 株：国内株は 211 株) については、収集週に大きな変化はないが、0121 に関しては、このデータセットの解析結果を論文発表した。また、0145 に関しては、既に収集したデータの詳細な解析から、prophage-in-prophage という新たな現象を見出し論文発表した。

0157:H7、0103:H2、0165:H25 に関しては、0157:H7 では 7,146 株（国内は 3,176 株）、0103:H2 では 2,701 株（国内は 200 株）0165:H25 では 202 株（国内は 71 株）と、海外株を中心に大幅に情報を拡張した。0157:H7 に関しては、海外株の情報だけでなく、2020 年の国内分離株の大部分を網羅する 623 株の WGS 情報を研究代表者から分与してもらいデータベースに組み込んだ。0157 の中で特に病原性が高いと推察されている Clade 8 株に関しては、別途収集・整備と解析を行っており、昨年度までに収集した株 (511 株：国内株は 150)

の詳細な解析とともに、亜系統を代表する 18 株を選び、ロングリードを用いたハイブリッドアッセムブリにより完全長配列を取得した O103:H2 に関しても国外株の情報を大幅に充実させるとともに、完全長配列を追加取得した。また、これらの完全長配列決定の中で、ハイブリッドアッセムブリのための解析パイプラインの有効性と精度に関する比較検証を行い、MicroPIPE と PolyPolish (いずれも昨年度に発表) の組み合わせが現時点では最も優れていると判断した。

O103:H2 に関しては、近縁株を含めた系統解析の過程で、ST17 complex として、O 抗原型が異なる近縁株も含めた解析が必要であることが判明したため、上記の O103:H2 に加えて、40 の血清型/69 の ST に属する 3,387 株の WGS 情報を追加収集し、これらの株を含めた大規模解析を開始した。類似の解析は、他の血清型においても検討する必要があると思われる。

この他、収集した菌株の特性解析に必要な Stx2 定量法の開発も行い、論文発表した。

#### D. 考察

(1) 諸外国における WGS の活用状況、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り、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ることには変わりがないが、集団事例に関連するクローンの判定基準についても、各血清型で使用するコアゲノムあるいはコア遺伝子で、どういった配列や遺伝子を使うかについては、国際的なコンセンサスは得られていない。国内で使用すべき解析パイプラインに関してもまだ検討段階であるが、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ためには、国際的な動向を注視していく必要がある。分担者個人の現在の見解としては、コア遺伝子セットの配列に基づく cgMLST が地方衛生研究所などでは最も取り入れやすい手法と考える。

(2) 国内・国外分離株の WGS データの収集・解析とデータベース構築：

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・整理を進めているが、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である。また、データ重複やクオリティーの問題等があるため、収集したデータについても、整理作業を行っていく必要がある。昨年度の報告書にも記載したように、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が、まだその作業には至っていない。参照配列となりうる代表株に関しては、ロングリード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が、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、既存のハイブリッドアッセムブリのプログラムの有効性と精度が問題となる。この点に関しては MicroPIPE と PolyPolish の組み合わせが現時点では最も優れていると判断できる。

#### 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、集団事例に関連するクローンの判定基準や、各血清型で使用するコアゲノムあるいはコア遺伝子については国際的なコンセンサスは得られていないため、今後の海外での動向を注視していく必要がある。

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・整理を進め、O157:H7、O103:H2、O165:H25 については海外株を中心に大幅に情報を拡張した。O103:H2 に関しては、ST17 complex として、O 抗原型が異なる近縁株も含めた解析が必要であることが判明し、O103:H2 に加えて、40 の血清型/69 の ST に属する 3,387 株の WGS 情報を追加収集したが、他の血清型においても類似の解析を行う必要性の有無を検討する必要がある。

参照配列となりうる代表株に関しては、ロングリードを用いた完全長あるいはそれに近い高精度配列の取得を進め、その過程で、ハイブリッドアッセムブリプログラムの有効性と精度に関しては、MicroPIPE と PolyPolish の組み合わせが現時点では最も優れていると判断した。

#### F. 健康危険情報 なし

## G. 研究発表

### 1. 論文発表

(1) Nakamura K, Ogura Y, Gotoh Y, Hayashi T.: Prophages integrating into prophages: A mechanism to accumulate type III secretion effector genes and duplicate Shiga toxin-encoding prophages in *Escherichia coli*, PLoS Pathogens, 17(4):e1009073, 2021, doi: 10.1371/journal.ppat.1009073.

(2) Nakamura K, Tokuda C, Arimitsu H, Etoh Y, Hamasaki M, Deguchi Y, Taniguchi I, Gotoh Y, Ogura Y, Hayashi T.: Development of a homogeneous time-resolved FRET (HTRF) assay for the quantification of Shiga toxin 2 produced by *E. coli*. PeerJ, 9:e11871, 2021, doi: 10.7717/peerj.11871.

(3) Nishida R, Nakamura K, Taniguchi I, Murase K, Ooka T, Ogura Y, Gotoh Y, Itoh T, Toyoda A, Mainil JG, Piérard D, Seto K, Harada T, Isobe J, Kimata K, Etoh Y, Hamasaki M, Narimatsu H, Yatsuyanagi J, Kameyama M, Matsumoto Y, Nagai Y, Kawase J, Yokoyama E, Ishikawa K, Shiimoto T, Lee K, Kang D, Akashi K, Ohnishi M, Iyoda S, Hayashi T.: The global population structure and evolutionary history of the acquisition of major virulence factor-encoding genetic elements in Shiga toxin-producing *Escherichia coli* O121:H19, Microbial Genomics, 7(12):000716, 2021.

doi: 10.1099/mgen.0.000716.

### 2. 学会発表

(1) 林哲也(招待講演):ゲノム解析を基盤とした細菌学遺伝的多様性に関する研究, 第95回日本細菌学会総会, 2022 (R3) 年

3月30日, 東京, オンライン.

(2) 林哲也(特別講演):細菌感染症とゲノム解析について, 2021 (R3) 年 11 月 25 日, 令和 3 年度地方衛生研究所地域専門家会議(地域保健総合推進事業), 太宰府市(福岡県保健環境研究所).

(3) 林哲也(招待講演):次世代シーケンサ(NGS)の活用によって進展する細菌ゲノムの進化・多様性解析, 2021 (R3) 年 8 月 21 日, 2021 年第 41 回阿蘇シンポジウム, 阿蘇市.

(4) 林哲也(特別講演):大規模ゲノム解析から紐解く細菌の多様性, 2021 (R3) 年 6 月 12 日, 2021 年度日本生化学会九州支部例会, 久留米, オンライン.

(5) 矢野文悟, 谷口愛樹, 後藤恭宏, 小椋義俊, 林哲也, 中村佳司:腸管出血性大腸菌 O26:H11 ST21 系統における Stx1 ファージのゲノム多様性, 第 95 回日本細菌学会総会, 2022 (R3) 年 3 月 29-31 日, 東京, オンライン.

(6) 谷口愛樹, 中村佳司, 後藤恭宏, 李謙一, 大岡唯祐, 小椋義俊, 大西真, 伊豫田淳, 林哲也:大腸菌 O103:H2 国内分離株とデータベース株を用いた高精度系統解析及び完全長配列決定株のゲノム構造比較, 第 95 回日本細菌学会総会, 2022 (R3) 年 3 月 29-31 日, 東京, オンライン.

(7) 中村佳司, 瀬戸和子, 磯部順子, 林哲也:挿入配列の切り出しによる大腸菌 O121:H19 の乳糖分解性の再活性化, 第 95 回日本細菌学会総会, 2022 (R3) 年 3 月 29-31 日, 東京, オンライン.

(8) 宮田達弥, 谷口愛樹, 中村佳司, 後藤恭宏, 小椋義俊, 大西真, 伊豫田淳, 林哲也:完全長配列を用いた EHEC O157 c1 ade 8 のゲノムと Stx2 ファージの多様性解析, 第 95 回日本細菌学会総会, 2022 (R3) 年 3 月 29-31 日, 東京, オンライン.