

厚生労働科学研究費補助金（食品の安全確保推進研究事業）
分担研究報告書

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、前年度までに引き続き、我が国で WGS を利用したサーベイランスと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、欧米諸国における WGS の活用状況、使用する解析パイプラインやデータベース等の調査を文献情報を基に行うとともに、我が国での活用法について検討した。これと並行して、前年度までに収集している国内分離株の WGS データの整理や再シーケンスを進めるとともに、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、公共データベースからの WGS データの追加収集を行った。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、O165、O103、O157 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。また、O145 に関しては、WGS を利用して Stx1 フェージの多様性に関する詳細な解析も行った。

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きた。本研究班の目的の 1 つは、MLVA に関する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、効率的に事例調査やサーベイランスを実施するための解析手法とデータベースの必要要件等を明らかにすることを目指している。また、これと並行して、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、海外株のデータを公共データベース（DB）から収集して 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おり、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とサーベイランス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、文献情報を基に欧米諸国（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、解析法と解析の基準、利用している DB、データフォーマット等を調査した。

(2) 前年度までに引き続き、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/DDBJ と ENTEROBASE) からデータを追加収集した。WGS 情報の取得には、基本的にはイルミナ社のシーケンサ（サンプル数に応じて MiSeq または NovaSeq 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、ONT 社の 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。0145 と 0121 については、前年度までに収集した WGS データに加え、それぞれ *lac* オペロンと *iee* 遺伝子や IS との関係及び Stx1 フェージ領域の個別シーケンスを行い、その多様性等を解析した。

(倫理面への配慮)

該当しない

C. 研究結果

(1) 欧米諸国における 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。欧米では中国製 (MGI 社) のシーケンサの利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

ている。

解析パイプラインは、前年度までに報告したように、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状況も変化がない。国内（分担者も含む）では、Platanus-b もアセンブリに使用されており、近縁菌株の SNP 解析には BactSNP 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 のいずれかが使用できるが、ii) と iii) が主流である。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) を採用していることが多いが、解析対象によって異なり、国際的なコンセンサスは依然として得られていない。この点は、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、まだ国際的なコンセンサスは得られていない。したがって、国際的なコンセンサスに基づいて 国内で使用する解析パイプラインを決定することは難しい。

(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 DB 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。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、新たな収集は行わなかったが、0145 に関しては、既に収集した WGS データに加えて、Stx1 フェージ領域の個別シーケンスを新たに行い、Stx1 フェージ領域の詳細な解析から 0145 における Stx1 フェージの多様性を明らかにし、論文発表した。0121 についても、前年度までに収集した WGS データに加え、*lac* オペロンと *iee* 遺伝子や IS の関係を解析し、ラクトース利用能の多様性と多様化メカニズムを明らかにし、論文発表した。

0165:H25 に関しては、近縁株 (0172 :

H25 EHEC を含む)を加えて、新たに clonal complex 119 (CC119)としてデータを収集整理し、197 株 (国内は 77 株) のゲノム情報 DB を構築した。また、その系統解析などを行って、グローバルな集団構造とその中での日本分離株の分布を明らかにし、論文発表した。

0103:H2 に関しては、前年度の解析で、系統の大きくなる 2 つの系統が存在すること、また近縁株を含めた系統解析の過程で、近縁株を含めた clonal complex として解析する必要があることが判明したため、主要系統については、ST17 complex として、0 抗原型が異なる近縁株も含めた WGS データの収集整理を行なった。具体的には、主要系統の 0103:H2 に加えて、40 の血清型/69 の ST に属する 6,091 株の WGS 情報を収集した (国内株は 199 株)。現在、これらの株のうち、クオリティーの悪いものや菌株情報が不十分なものを除いた 5,051 株の WGS を使って大規模な系統解析を行っている。ST2307 に代表されるマイナー系統に関しては、少数の海外株 (ベルギー株など) の WGS 情報を追加収集して 080:H2 を中心とした小規模な比較解析のみを行い、論文発表した。

0157:H7 のうち、特に病原性が高いと推察されている clade 8 株に関しては、前年度までに収集した 511 株 (国内株は 150: 亜系統を代表する 18 株については完全長配列を取得) の詳細な解析から、グローバルな集団構造に加えて、clade 8 の中で Stx2 高産生性で重症化頻度の高い亜系統を同定するなどの知見を得て、論文発表した。0157:H7 全体に関しては、前年度に引き続いてグローバルな大規模 WGS 情報 DB の構築を進め、7,290 株 (国内は 3,275 株) からなる DB を構築した。このうち、分担者が以前にシーケンスした国内分離株の中には、当時のライブラリー調整キットの性能の関係でクオリティーの低いものが多く含まれていたため、960 株を再シーケンスし、DB に組み込んだ。現在、これらの株の WGS を使って大規模な系統解析を行っている。

D. 考察

(1) 欧米諸国における 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り、集団事例に関連するクローンの判定基準についても未だに国際的なコンセンサスは得られていない。したがって、国内で使用すべき解析パイプラインに関しては、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は難しいと考えられ、WGS の幅広い活用を我が国で開始するためには独自のパイプラインを構築する必要があると思われる。この点では、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である。この場合、英国等で使用されている遺伝子セットと共通の大腸菌コア遺伝子セットも使用すれば、他の病原性大腸菌にも同じ解析パイプラインが利用できる。

(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、0157、026、0145、0103、0121、0165 の WGS データの収集・整理を進めているが、海外からの公共 DB への登録が急速に進んでいること、また 0103 や 0165 のように、血清型は異なるが極めて近縁の EHEC が存在することが明らかになったことから、0157 と 0103 では非常に大規模な WGS 情報 DB を構築することになった。他の血清型についても、幅広い継続的な収集と DB のアップデートが必要である。また、0103 や 0165 以外の血清型に関しても、近縁株を含めた DB に変更することを検討する必要がある。さらに、データ重複やクオリティーの問題、株情報の不備等の問題があるため、収集したデータに関しても、整理作業を行っていく必要がある。同時に、DB 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べきである。参照配列あるいは landmark となりうる株に関しては、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。このような解析に加えて、国内株については Stx の産生量や臨床情報が利用できるという我が国の強みを生かして、0157:H7 clade 8 で行ったように、高リスク系統の同定やその発生動向変化などの解析も進めるべきである。

E. 結論

欧米諸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使用する解析パイプライン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、WGSの幅広い活用を我が国で開始するためには独自のパイプラインを構築する必要がある。

WGS情報DBの構築に関しては、主要EHEC血清群のうち、O157、O103、O165の国内・国外分離株のWGSデータの収集・整理を進め、海外株を中心に大幅にDBを拡張した。O103、O165については、O抗原型が異なる近縁株も含めたDBを構築した。

F. 健康危険情報 なし

G. 研究発表

1. 論文発表

(1) Yano B., Taniguchi I., Gotoh Y., Hayashi T., & Nakamura K.: Dynamic changes in Shiga toxin (Stx) 1 transducing phage throughout the evolution of O26:H11 Stx-producing *Escherichia coli*, *Scientific Reports*, 13(1):4935, 2023.

(2) Nakamura K., Seto K., Lee K., Ooka T., Gotoh Y., Taniguchi I., Ogura Y., Jacques Georges Mainil, Denis Piérard, Harata T., Etoh Y., Ueda A., Hamasaki M., Isobe H., Kimata K., Narimatsu H., Yatsuyanagi J., Ohnishi M., Iyoda S., Hayashi T.: Global population structure, genomic diversity and carbohydrate fermentation characteristics of clonal complex 119 (CC119), an understudied Shiga toxin-producing *E. coli* (STEC) lineage including O165:H25 and O172:H25, *Microbial Genomics*, 9:000959, 2023.

(3) Miyata T., Taniguchi I., Nakamura K., Gotoh Y., Yoshimura D., Itoh T., Hirai S., Yokoyama E., Ohnishi M., Iyoda S., Ogura Y., Hayashi T.: Alteration of a Shiga toxin-encoding phage associated with a change in toxin production level and disease severity in *Escherichia coli*. *Microbial Genomics*, 9:0009935, 2023.

(4) Ikeda R., Nakamura K., Saulmont M., Habets A., Duprez JN, Korsak N., Hayashi T., Thiry D., Mainil JG.: *Escherichia coli* O80 in Healthy Cattle: Absence of Shiga toxinogenic and Enteropathogenic *E. coli* O80: H2 and (Phylo) Genomics of Non-Clonal Complex 165 *E. coli* O80, *Microorganisms*, 11(2):230, 2023.

(5) Nakamura K., Seto, K., Isobe J., Taniguchi I., Gotoh Y., Hayashi T.: Insertion Sequence (IS)-Excision Enhancer (IEE)-Mediated IS Excision from the lacZ Gene Restores the Lactose Utilization Defect of Shiga Toxin-Producing *Escherichia coli* O121:H19 Strains and Is Responsible for Their Delayed Lactose Utilization Phenotype, *Applied Environmental Microbiology*, 88(16): e0076022, 2022.

2. 学会発表

(1) 林哲也, 次世代シーケンサを活用した病原細菌と細菌感染症研究. 特別講演:第54回日本小児感染症学会総会・学術集会, 2022.11.5, 福岡.

(2) 宮田達弥, 谷口愛樹, 中村佳司, 後藤恭宏, 平井晋一郎, 横山栄二, 大西真, 伊豫田淳, 小椋義俊, 林哲也, STEC O157:H7 clade 8 の世界的な集団構造と Stx2 と Stx2a ファージのバリエーション. 第96回日本細菌学会総会, 2023.3.16~18, 姫路.

(3) 中村佳司, 瀬戸和子, 李謙一, 後藤恭宏, 伊豫田淳, 林哲也, 隠れた志賀毒素産生性大腸菌系統, Clonal complex 119(CC119)の集団構造と CC119 菌株の糖代謝. 第96回日本細菌学会総会, 2023.3.16~18, 姫路.

(4) 宮田達弥, 谷口愛樹, 中村佳司, 後藤恭宏, 吉村大, 伊藤武彦, 平井晋一郎, 横山栄二, 大西真, 伊豫田淳, 小椋義俊, 林哲也, STEC O157:H7 clade 8 のグローバルな集団構造および系統・Stx2a phage subtype と Stx2 産生量・重症度の関連性. 第24回EHEC研究会, 2022.10.13-14, 川崎.

(5) 矢野文悟, 中村佳司, 谷口愛樹, 後藤恭宏, 小椋義俊, 林哲也, 志賀毒素産生性大腸菌 O26:H11 ST21 系統における Stx1 ファージゲノム多様性と Stx1 産生性. 第24回EHEC研究会, 2022.10.13-14, 川崎.

(6) 西田留梨子, 中村佳司, 谷口愛樹, 伊豫田淳, 林哲也, STEC O121:H19のグローバルな集団構造と主要病原因子をコードする遺伝子の進化. 第24回EHEC研究会, 2022.10.13-14, 川崎.

(7) 中村佳司, 勢戸和子, 磯部順子, 林哲也, Insertion sequence (IS)-excision enhancer (IEE)を介したISの切り出しによるSTEC O121:H19の乳糖分解性の変化における乳糖の関与. 第24回EHEC研究会, 2022.10.13-14, 川崎.

