

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）
（総合）研究報告書

広域食中毒発生時の早期探知のための調査の迅速化及びゲノム解析技術を利用した調査法の確立に資する研究

研究分担者 中村 佳司 九州大学・大学院医学研究院・講師

研究要旨

本研究班では、腸管出血性大腸菌（EHEC）食中毒アラートシステムの構築、全ゲノム配列（WGS）データベースの拡充および解析パイプラインの構築、および離島でのコホート研究による発生原因の評価を目的としている。分担者は主にWGS解析パイプラインの検証とデータベースの拡充を担当し、臨床および動物分離菌株のWGSデータ収集および解析に従事した。

まず、研究代表者が構築したEHECのWGS解析パイプラインの評価・検証を行い、国内および海外の集団食中毒事例を用いた精査を通じて、同一クローン判定の基準として「10 SNPs以下」という閾値を策定した。この基準は海外事例の解析結果とも整合しており、国際的な妥当性を有する指標であることを確認した。次に、国内の臨床分離株および牛由来分離株のWGSデータを広範に収集してデータベースを拡充し、その成果の一部を公共データベースへの登録および国際誌での公表により公開した。技術面では、Nanoporeシーケンシングを用いたWGS解析の適用性を検証した。ショートリードベースの手法と比較してシーケンスエラーに由来するSNPs数の増加が認められたものの、高精度な完全長ゲノムを迅速に構築できる点において、実用的な有用性を確認した。また、次期研究への展開を見据え、対象を*Campylobacter jejuni*にまで広げ、全ゲノム配列の構築を完了させた。併せて、地域の衛生研究所職員を対象とした講演や動画教材の作成を通じ、ゲノム解析技術の普及と解析環境の整備を支援し、現場の実務的なニーズを把握することで、地域におけるWGS解析の社会実装に寄与した。

A. 研究目的

腸管出血性大腸菌（EHEC）による食中毒は、その届出数や重症度から、国内において最も警戒すべき疾患の一つとなっている。本感染症の事例調査における分子型別法としては、迅速性に優れた反復配列多型解析法（MLVA法）が主に用いられてきた。前身の研究班では、MLVA法による全国的な分子疫学的解析に加え、EHEC菌株の全ゲノム配列（WGS）のデータベース化を推進した。その結果、ゲノム情報を利用してEHEC集団感染を早期探知するための基本的な情報基盤が構築され、これを実用的な集団感染探知に活用できる段階に至っている。

分担者は、本研究の支柱の一つである「EHEC等の検査のためのWGS解析法の確立に関する検討とEHEC株のWGSデータベースの拡充」を担当している。本分担課題は、国際整合性の観点を考慮しつつ、研究代表者と連携し、食中毒探知の最前線である地方衛生研究所等で利用可能なゲノム解析パイプラインを構築することを目指している。具体的には、以下の3点を主要な目標とする。第一に、蓄積された国内分離株のゲノムデータに、集団発生事例に由来する国内および海外の菌株データを加え、研究代表者が開発したゲノム解析パイプラインによる検証を行う。この結果に基づき、国際的に整合性のある同一クローンの判定基準等を確立する。第二に、新たに分離されたEHEC国内株のWGSデータを継続的に取得し、前身の研究班から継承したWGS情報データベースをさらに拡充する。第三に、WGS解析から得られたデータを用いた、感染源や感染経路等に関する後方視的な疫学解析への寄与を目的として、主感染源となる食品由来株および動物由来株のデータ収集を可能とする方策を検討し、臨床分離株との比較や事例調査への適用を試みる。分担者は、特に動物（牛）由来株のWGSデータ収集を随時サポートし、必要に応じてWGS情報データベースに組み込んで解析を行う役割を担う。

B. 研究方法

1. EHEC集団感染探知のためのWGS解析パイプラインの検証と判定基準の策定

研究代表者が開発したゲノム解析パイプライン（SNPcaster）を用いて、集団感染感知に最適な一塩基多型（SNPs）の検出条件および同一クローンの判定閾値の検証を行った。まず、海外のEHEC集団発生事例（イタリア、フランス、アメリカの計5事例）のWGSデータをNCBIデータベースより取得し、解析対象の参照配列の設定および判定閾値の検討を行った。次に、国内で発生した集団食中毒事例に由来するEHEC菌株（39集団、計402菌株）を対象に、以下の4条件における解析結果を比較し、国内事例に適した解析基準を検討した。

i) 参照配列：EHEC O157:H7 Sakai株（リピート領域除外）、Gubbinsプログラム（組み換え領域除外）あり

- ii) 参照配列：大腸菌K-12 MG1655株（リピート配列除去なし）、Gubbinsあり
- iii) 参照配列：Sakai株（リピート領域除外）、Gubbinsなし
- iv) 参照配列：MG1655株（リピート配列除去なし）、Gubbinsなし

2. EHEC/EPEC菌株のゲノムデータ取得とデータベースの拡充

本研究では、解析対象とする菌株数に応じて、取得手法を変更した。まず、1年目においては、国内の臨床分離株および動物分離株（計381株）を対象に、イルミナ社NovaSeqを用いてショートリード配列を取得した。取得したリード配列からshovillプログラムを用いてドラフトゲノム配列を構築し、checkMプログラムによる品質評価を行った。さらに、ECTyperプログラムを用いた菌種推定および血清型別、ならびに志賀毒素（*stx*）遺伝子およびインチミン（*eae*）遺伝子の保有状況を解析した。2年目以降においては、動物（牛）由来株を中心に、ナノポアシーケンサ（MinION）を用いた完全長ゲノム配列の決定を行い、プラスミドを含む高精度なゲノム構造を構築した（2年目：EHEC/EPEC 43菌株、3年目：EHEC 36菌株およびカンピロバクター15菌株）。得られたゲノムデータのうち、学術的価値が高いEHEC/EPEC株については、公共データベースへの登録を行い、ゲノムリソースとしての公開を完了した。その他の配列については、研究代表者および分担者間でデータ共有を行い、今後の登録方針に基づき順次公開する計画とした。

3. Nanoporeリードを用いたWGS解析の配列精度および妥当性の検証

近年のNanoporeシーケンシングにおける配列精度の向上に着目し、ショートリードによる補完を行わないNanoporeリードのみによるゲノム解析への適用の可否を検証した。

まず、MinION Flow Cell R10.4.1およびDorado (Super accuracy mode) を用いたベースコールによりNanoporeリードを生成し、これを用いてHybracterパイプラインによりゲノム配列を構築する手順を適用した。まず、EHEC O157:H7 Sakai株を用いて公共データベース上の参照配列とNanoporeリードを用いて構築したゲノム配列（以下、Nanoporeゲノム配列）との差異を確認し、配列精度の予備的な検証を行った。次に、実際の事例への適用性を検証するため、沖縄県内のEHEC集団事例由来株（18株）を対象に解析を行った。Nanoporeゲノム配列に対し、iSeq100によるショートリードを用いた配列修正（polypolishおよびpypolcaを使用）を施し、修正前後の配列を比較することで、具体的にどのような箇所が修正されたかを分析した。さらに、得られたNanoporeゲノム配列を用いたWGS解析（SNP検出）の結果が、本研究課題で開発・検証されたショートリード配列を用いる解析プログラム（SNPcaster）の結果とどの程度乖離するかを検証した。検証には、構築済みのゲノム配列から直接SNPsを検出するin-houseプログラムを用い、菌株間の最大・最小・平均SNPs数および標準偏差を算出して比較した。

4. 地方衛生研究所におけるゲノム解析体制の構築支援

地方衛生研究所がWGS解析パイプラインを円滑に運用できるよう、九州地域の衛生研究所を中心に技術支援および教育活動を行った。

具体的には、沖縄県衛生環境研究所において、解析環境の構築をサポートするとともに、煩雑になりやすい菌株データ管理の効率化に向けた提案、および解析プログラムの具体的な操作方法や結果の解釈に関する技術指導を実施した。

また、九州地方の衛生研究所職員を対象に、病原細菌の全ゲノム解析の現状と基礎知識に関する講演会を開催した。さらに、より広範な施設への情報共有と普及を図るため、本講演内容を音声付きの動画教材として作成し、希望する施設へ配布した。あわせて、教材の有用性の評価および今後必要とされる解析技術に関するアンケート調査を実施し、現場のニーズ把握を試みた。

（倫理面への配慮）

該当しない

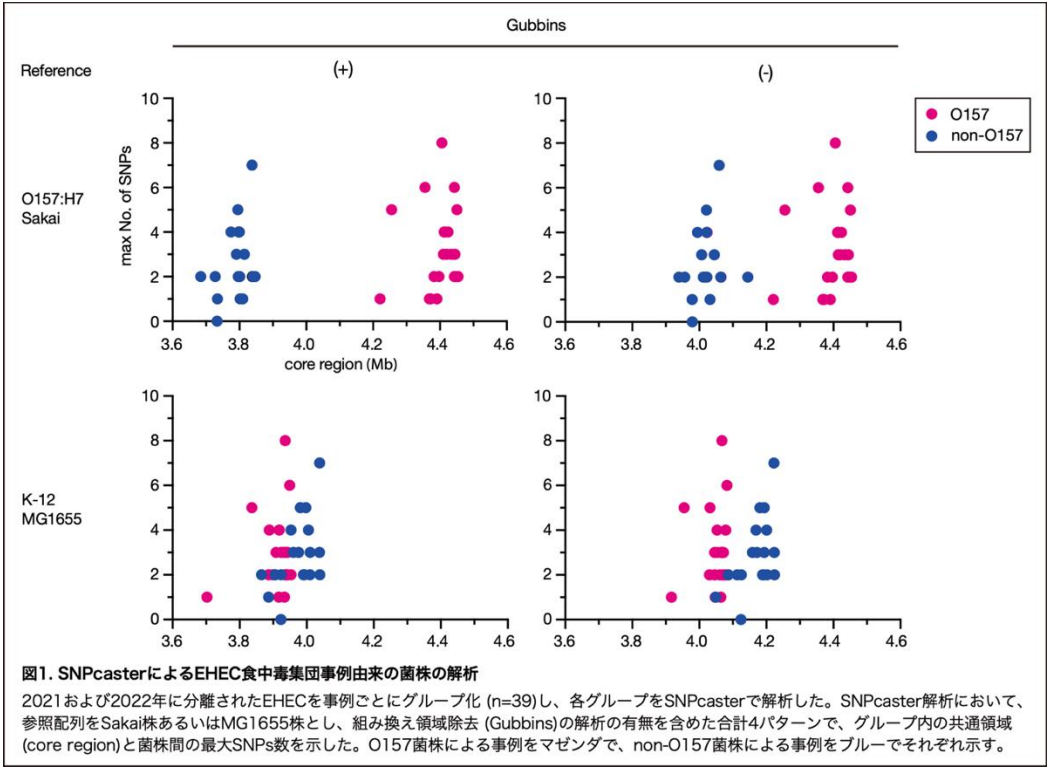
C. 研究結果

1. EHEC集団感染探知のためのWGS解析パイプラインの検証と判定基準の策定

研究代表者が開発したSNPcasterパイプラインを用い、国内の集団食中毒事例（39集団、計402菌株）を対象に解析条件の検討を行った（図1）。

解析の結果、参照配列にO157:H7 Sakai株を用いた場合、O157による事例ではcore regionが4.2～4.4 Mbであったのに対し、non-O157による事例では3.6～3.8 Mbとなり、血清型によってcore regionの範囲に明らかな差があることが判明した。一方で、参照配列をK-12 MG1655株に変更した場合には、O157およびnon-O157間でのcore regionの著しい差は見られなかった。また、組み換え領域を除去するGubbinsプログラムの影響を検証したところ、Sakai株を参照としたnon-O157事例において、core regionから約0.2 Mbの領域が除去されることが明らかとなった。これらの解析条件に基づき、国内の各事例における菌株間の最大SNPs数を算出したところ、いずれの解析パターンにおいても最大SNPs数が10を超える事例は存在しなかった。このことから、国内の食中毒事例において疫学的関連性のある菌株を判定する際の遺伝的閾値は、10 SNPs以下であることが示唆された。さらに、この基準の妥当性を検証するため、海外のEHEC集団発生事例を用いた解析を行った。イタリアおよびフランスのO26 EHEC事例（計3事例）では、菌株間の最大SNPs数は5 SNPs以下であった。また、アメリカのO157による事例においても、同一感染事例の原因と考えられる菌株同士の遺伝的距離は、一部の例外を除いて5 SNPs以内であった。これらの結果は、海外のEHEC集団感染事例においても、10 SNPs以下という基準が同一クローン判定の基準として十分に適

用可能であることを示しており、国内で得られた判定基準が国際的な整合性を有していることを示唆している。



2. EHEC/EPEC菌株のゲノムデータ取得とデータベースの拡充

年次ごとのWGSデータ取得実績およびその活用状況は以下の通りであり、3年間で臨床および動物由来の広範な菌株データを蓄積し、手法の高度化（ドラフト配列→完全長配列）に伴う詳細なゲノム解析基盤を構築した。

1年目には、2021年および2022年の国内集団感染事例由来のEHEC分離株226菌株、および離島コホート研究から分離された牛由来菌株155菌株について、NovaSeqを用いたショートリード配列の取得とドラフトゲノムの構築を完了した。これらのデータのうちEHEC分離株226菌株の大部分が、上述のSNPcasterを用いた解析パイプラインの検証および判定基準の策定において、重要な基礎データとして活用された。

2年目には、離島コホート研究により分離された牛由来EHECおよびEPEC計43菌株について、Nanoporeシーケンサを用いて完全長ゲノム配列を決定した。これらのうち、EHEC菌株は、ゲノムリソースとしての価値が高いことから、公共データベースへの登録を行うとともに、国際誌（Microbiology Resource Announcements）にてその成果を公表した。

3年目には、沖縄県内の集団事例由来のO157株18株、および離島コホート研究により分離された牛由来EHEC株18株の計36株について、Nanoporeシーケンサを用いて全ゲノム配列を構築した。あわせて、次期研究への展開を見据え、沖縄県内のカンピロバクター集団感染事例から分離された *Campylobacter jejuni* 15株についても、Nanoporeリードの取得および全ゲノム配列の構築を完了させた。

3. Nanoporeリードを用いたWGS解析の配列精度および妥当性の検証

Nanoporeリードのみを用いて構築したゲノム配列の精度および、それを用いたWGS解析の妥当性を検証した。

まず、EHEC O157:H7 Sakai株を用いた検証では、Nanoporeリードのみを用いて染色体および2つのプラスミド（pO157およびpOSAK1）のすべてを環状化することに成功した。構築した配列を公共データベース上の登録配列と比較したところ、ゲノムサイズはほぼ一致し、塩基の差異は染色体でミスマッチ2箇所および欠失・挿入7箇所、pO157でミスマッチ8箇所および欠失・挿入1箇所であり、pOSAK1については完全に一致していた（表1）。次に、沖縄県内のEHEC集団事例由来株（18株）を対象に、Nanoporeリードのみで構築したゲノム配列と、iSeq100のショートリードを用いて修正した後の配列を比較した。その結果、Nanoporeリードのみの配列では、ショートリードによる修正後にSNVsが0～11個（中央値2）、InDelsが3～25個（中央値7）の差異が認められた。さらに、構築したNanoporeゲノム配列を用いたWGS解析の結果を、本研究課題で開発したショートリードベースの解析プログラム（SNPcaster）の結果と比較した（表2）。in-houseプ

表1. ナノポアリードのみで構築したSakai株のゲノム配列のサイズと参照配列との違い

	染色体		pO157		pOSAK1	
	サイズ (bp)	SNV/InDel*数	サイズ (bp)	SNV/InDel数	サイズ (bp)	SNV/InDel数
参照配列†	5,498,578		92,721		3,306	
ナノポア	5,498,579	2/7	92,722	8/1	3,306	0/0

* SNV, Single Nucleotide Variation; InDel, Insertion and Deletion

† PacBioリード(ロングリード)とイルミナリードを用いて構築した配列 (Accession No. ASM886v2)

プログラムを用いて算出したNanoporeゲノム配列間のSNPs数は、最大11 SNPsとなり、SNPcasterによる解析結果（最大3 SNPs）やIlluminaリードを用いた解析結果（最大4 SNPs）と比較して、検出されるSNPs数が増加することが確認された。この結果から、Nanoporeゲノム配列をそのままWGS解析に用いた場合にSNPs数が増加することは、シーケンスエラーに由来するものと考えられた。

4. 地方衛生研究所におけるゲノム解析体制の構築支援

沖縄県衛生環境研究所への支援を通じて、データ管理の効率化に関する具体案の提示および解析プログラムの適切な運用方法に関する技術指導を完遂した。これにより、現場における解析環境の整備と、実務的な運用フローの適正化に寄与した。

九州地方の衛生研究所向けには、講演および動画教材の配布を通じて、ゲノム解析に関する基礎知識の普及を図った。配布した動画教材に対するアンケートの結果、教材の内容や形式について概ね高い評価を得たことが確認された。また、本調査を通じて、地方衛生研究所の職員が具体的にどのような知識や技術が必要としているかという実務的なニーズを詳細に把握することができ、今後の技術支援における優先課題を明確化することができた。

D. 考察

1. EHEC集団感染探知のためのWGS解析パイプラインの検証と判定基準の策定

本研究では、国内および海外の集団感染事例を包括的に解析した結果、同一クローン判定の閾値として10 SNPs以下という基準を導き出した。海外事例ではより厳格な5 SNPs以下という傾向が見られたものの、国内事例の結果と併せて検討すると、10 SNPsという閾値は国内・海外を問わず疫学的関連性を識別するための妥当な指標であり、国際的な整合性を有する基準であると考えられる。また、参照配列の選択やGubbinsによる組み換え領域の除去が、血清型によるcore regionの変動を抑え、解析の安定性を高めることが明らかとなった。このように、解析条件を最適化した上で10 SNPsという基準を適用することは、国内の集団感染事例を精度高く判定し、かつ国際的な解析基準との乖離を最小限に抑えるための極めて有効なアプローチであると言える。

2. EHEC/EPEC菌株のゲノムデータ取得とデータベースの拡充

3年間で臨床分離株および動物分離株の広範なゲノムデータを蓄積し、その一部を公共データベースへ登録・公表した。特に動物由来株の完全長ゲノム配列を蓄積したことは、これまで不足していたリソースを補完するものであり、今後の人獣共通感染症としての伝播経路解析や、菌株の多様性評価を行う上での重要な基盤になると考えられる。

3. Nanoporeリードを用いたWGS解析の配列精度および妥当性の検証

Nanoporeリードのみを用いてSakai株や集団事例由来株の高精度な完全長ゲノムを構築できたことは、NanoporeシーケンシングのWGS解析への有用性を示している。一方で、得られたゲノム配列を用いてSNPs数を算出したところ、ショートリードベースの解析（SNPcaster等）よりも数値が高くなる傾向が認められた。これはNanopore特有のシーケンスエラーに起因するものと考えられる。したがって、NanoporeゲノムをWGS解析に適用する場合、従来のショートリードベースの閾値（10 SNPs）をそのまま適用するのではなく、Nanopore特有のエラー率を考慮した新たな判定基準を策定する必要がある。

4. 地方衛生研究所におけるゲノム解析体制の構築支援

地方衛生研究所への技術支援および教育活動を通じて、解析プログラムの操作習熟のみならず、解析環境の整備やデータ管理の効率化が実務上の重要な課題であることが浮き彫りとなった。特に、動画教材などの教育的アプローチが概ね高く評価されたことは、専門的なゲノム解析技術を現場へ普及させる有効な手段になると考えられる。今後は、把握した実務的なニーズに基づき、より現場の運用に即した技術支援を行うことが、WGS解析の社会実装を加速させる鍵になると考えられる。

E. 結論

本研究では、腸管出血性大腸菌（EHEC）の集団感染を早期かつ正確に探知するため、WGS解析パイプラインの検証、データベースの拡充、および地方衛生研究所における解析体制の整備に取り組んだ。

まず、SNPcasterパイプラインの検証を通じて、菌株間のSNPs数が10以下である場合に同一クローンである可能性が高いという判定基準を策定し、実用的な解析基準を確立した。次に、臨床分離株および牛由来分離株のWGSデータを広範に収集してデータベースを拡充し、その成果を公共データベースへの登録および国際誌での公表を通じて公開した。また、Nanoporeシーケンシングを用いたWGS解析の適用性を検証し、その有用性を確認した。最後に、地方衛生研究所への技術指導および動画教材の配布を通じて、地域のゲノム解析能力の向上と、持続可能な情報共有体制の構築に寄与した。

以上の成果により、EHECの集団感染探知に向けた解析基準の明確化とデータ基盤の整備がなされ、地方衛生研究所におけるWGS解析の社会実装に向けた技術的基盤が構築された。

表2. EHEC 集団事例の菌株を用いた WGS 解析方法の比較

菌株間 SNPs	方法 i	方法 ii	方法 iii	方法 iv
最大	3	4	11	4
最小	0	0	0	0
平均	0.9	1.3	1.3	1.3
標準偏差	0.7	0.9	0.9	0.9

方法 i, Illumina リードを用いた SNPcaster

方法 ii, Illumina リードを用いた in-house プログラム

方法 iii, Nanopore リードを用いた in-house プログラム

方法 iv, Nanopore/Illumina リードを用いた in-house プログラム

F. 研究発表

1. 論文発表

Kakita T, Iyoda S, Taira H, Kudeken T, Akeda Y, Ohnishi M, Nakamura K. Closed genome sequences of 39 Shiga toxin-producing *Escherichia coli* strains isolated from healthy cattle in Okinawa, Japan. *Microbiology Resource Announcements*. 2025 Nov 13;14(11): e0081625.

2. 学会発表

1. 中村 佳司、柿田 徹也、大西 真、明田 幸宏. ナノポアリードのみを用いた全ゲノム配列比較による細菌検査方法の検討. 第 27 回腸管出血性大腸菌感染症研究会. 福岡, 2025. 本研究で構築したゲノム配列決定方法を用いて決定した腸管出血性大腸菌株のゲノム配列の精度に関して検討した結果を報告した。
2. 中村 佳司、柿田 徹也、大西 真、明田 幸宏. 完全長ゲノム配列の比較解析に基づく腸管出血性大腸菌感染症起因菌のゲノム変化の解明. 九州微生物研究フォーラム2025. 沖縄, 2025. 本研究で構築したゲノム配列決定方法を用いて、集団食中毒事例の原因菌の解析と伝播経路の推定を行った結果を報告した。

G. 知的所有権の取得状況

1. 特許取得

なし

2. 実用新案登録

なし

3. その他

なし