

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

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、1) 反復配列多型解析法（MLVA法）運用の効率化・迅速化・精度向上およびMLVA型別データと発生動向を利用した腸管出血性大腸菌（EHEC）食中毒アラート検出システムの構築、2) EHECの全ゲノム配列（WGS）のデータベースの拡充化によるMLVA法との優劣性の解明およびWGS解析パイプラインの構築、3) 離島でのコホート研究によるEHEC感染症発生とその原因の関連性の評価、を行うことを目的としている。分担者は、研究代表者とともに2)の課題を担当するとともに、3)の課題で行われる動物分離菌株の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を担っている。

本年度は、沖縄県で発生したEHEC感染症集団事例から分離された菌株（計18株）を用いて、Nanoporeシーケンサーから得られたリードのみで高精度なWGS解析が可能か検討した。また、前年度に引き続き、九州地方の複数の衛生研究所職員を対象とした講演会を実施し、あわせて資料を音声付き動画として作成し、不参加の衛生研究所へ共有した。WGSデータベースの拡充に関しては、前年度までに離島コホート研究により取得したウシ由来病原性大腸菌の完全長ゲノム配列を公共データベースに登録し、そのうちEHEC（計39株）の解析結果を国際誌にて報告した。さらに、R7年度に離島コホート研究により新たに分離されたウシ由来EHEC（計18株）および、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*（計15株）について、Nanoporeリードの取得および全ゲノム配列の構築を完了した。

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と相まって日本の食中毒として最も警戒が必要な食中毒となっている。本感染症の事例調査のため分子型別法として、反復配列多型解析法（MLVA法）が迅速性等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。前身の研究班では、MLVA法によるEHEC菌株の全国的な分子疫学的解析に加えて、EHEC菌株の全ゲノム配列（WGS）のデータベース化を行った。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でき、これを実際的なEHEC集団感染探知に利用できる段階にきている。分担者は、本研究の柱の1つである、「EHECなどの検査のための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て、EHEC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。このために、これまでに蓄積された国内分離株のゲノムデータに、EHEC感染症集団発生事例に由来する国内あるいは海外の菌株のデータを加えて、研究代表者のグループが開発したゲノム解析パイプラインによる解析を行い、このパイプラインの検証・解析結果を踏まえて、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ことを目指している。また、これと並行して、新たに分離されたEHEC国内株のWGSデータを取得することで、前身の研究班から培われてきたEHECのWGS情報データベースをさらに拡充することも目的となる。これらに加え、本研究班では、MLVA及びWGS解析から得られたデータを利用して感染源や感染経路等に関する後方視的な疫学解析を行う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であり、本分担者は、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、上記のWGS情報データベースに組み込んで解析する役割も担っている。

B. 研究方法

(1) EHECおよびカンピロバクター菌株のWGS決定

EHEC菌株（計36株：沖縄県集団事例由来O157株18株、R7年度離島コホート研究分離株18株）はLB培地（37℃、一晚培養）から、カンピロバクター菌株（18株）はBrucella broth（37℃、微好気、48時間培養）から、それぞれGenomic-Tip（Qiagen; 20/Gあるいは100/G）を用いてDNAを精製した。Rapid Sequencing Kit V14（Oxford

Nanopore Technologies; ONT)を用いてDNAライブラリを調製後、MinION Flow Cell R10.4.1 (ONT)によるNanoporeシーケンシングを行なった。取得データの塩基配列への変換（ベースコール）には、Dorado (ONT; Super accuracy mode)およびDNA modelを、菌株セットごとに以下のバージョンで利用した：沖縄県EHEC感染症集団事例O157菌18株, Dorado v.1.0.0 and DNA model v.5.2.0; カンピロバクター菌株, Dorado v.0.8.3 and DNA model v.5.0.0; R7年度離島コホート研究分離株18株, Dorado v.7.11.2 in MinKNOW v.25.09.16 and DNA model v.5.2.0。生成された塩基配列データ（Nanoporeリード）から、Hybracterパイプライン (Bouras G, *et al.* Microb Genom, 2024)を用いて、ゲノム配列を構築した。EHEC菌株については、必要に応じてNanoFilt v.2.8.0 (De Coster W, *et al.* Bioinformatics, 2018) プログラムを用いて、Quality score (Q score) >10およびリード長 >2-kbのどちらかあるいは両方を満たすリードのみを抽出し、配列構築に利用した。さらに、ゲノムの環状化が確認できない場合には、circlator (Hunt M, *et al.* Genome Biol., 2015) プログラムにより、環状化処理を行なった。沖縄県集団事例分離EHEC菌株のWGSゲノムに関しては、iSeq100 (Illumina)によるシーケンシングで得たショートリード配列（沖縄県衛生環境研究所より提供）を用いて、polypolish v.0.6.1 (Wick RR, *et al.* PLoS Comput Biol, 2022) およびpypolca v.0.4.0 (Bouras G, *et al.* Microb Genom, 2024) プログラムによる配列修正を行った。

各EHEC菌株の血清型およびSequence type (ST)を、ECTyper v.1.0.0 (Bessonov K, *et al.* Microb Genom, 2021) プログラムおよびmlst program v2.23.0 (<https://github.com/tseemann/mlst>; Achtmanのスキームを使用)によって決定するとともに、BLASTNに基づいた方法 (Nakamura K, *et al.* PLoS Pathog, 2021) によって、志賀毒素 (stx) 遺伝子およびインチミン (eae) 遺伝子の保有の有無を解析した。

(2) 沖縄県集団事例分離EHEC菌株を用いた比較ゲノム解析

方法 (1) により、各菌株において、① Nanoporeリードのみを用いた完全長ゲノム配列、② ①をIlluminaショートリードで修正した配列、③ IlluminaショートリードとShovill v1.1.0 (Seemann T, <https://github.com/tseemann/shovill>, 2016)プログラムを用いたドラフトゲノム配列、をそれぞれ構築した。これらの配列を用いて、以下の3つの解析を行った。

解析1: MUMmer (Kurtz S, *et al.* Genome Biol, 2004) にパッケージされているdnadiffプログラムを用いて、①と②の間の染色体配列間の差異を確認した。

解析2: 全18菌株の染色体配列間に存在するSingle Nucleotide Polymorphisms (SNPs) を、(i) Illuminaショートリードを用いたWGS解析パイプライン (SNPcaster)、(ii) ③のゲノムを用いた、in-houseプログラム (Ogura, *et al.* Microb Genom, 2017)、(iii) ①のゲノムを用いた、in-houseプログラム、(iv) ②のゲノムを用いた、in-houseプログラム、の方法でそれぞれ検出した。

解析3: 解析2の方法 (iv)で得られたSNPsを用いて、GrapeTree (Zhou Z, *et al.* Genome Res., 2018) プログラムにより、Minimum Spanning Tree (MST)を作成し、18菌株を染色体配列の違いに基づいて分類した。さらに、全18菌株のプラスミド配列間に存在するSNP/InDel (Insertion and Deletion) をdnadiffプログラムにより検出し、これを18菌株の詳細な分類に利用した。

(3) WGS解析に関する資料作成

九州地域の衛生研究所（福岡市、北九州市、福岡県、佐賀県）職員を対象として、病原細菌の全ゲノム解析の導入に向けた基礎知識および現状に関する講演を、福岡県保健環境研究所にて実施した。さらに、講演不参加の職員への周知・普及を図るため、本講演資料を音声付きの動画教材として作成した。具体的には、Microsoft PowerPoint for MacおよびOBS Studioを用いて音声付きスライドショーを録画し、iMovieでの編集とHandBrakeでの圧縮を経て、1本あたり10～20分（22～60 MB）の動画ファイルとして整備した。本教材の配布については、福岡県保健環境研究所の重村氏の協力を得て、九州地域の衛生研究所の希望者に提供した。あわせて、動画内容の分かりやすさや時間の適切性、動画教材の利用に関する意識、および今後動画化を希望するゲノム解析テーマについてアンケートを実施し、利用者のニーズとフィードバックを収集した。

(倫理面への配慮)

該当しない

C. 研究結果

(1) Nanoporeリードのみで構築したEHEC O157菌株の配列精度

前年度にEHEC O157 Sakai株を用いて実施したNanoporeリードのみによるゲノム配列（以下、Nanoporeゲノム配列）の精度検証を、本年度は対象株数を増やして詳細に検討した。多くの菌株では参照配列が存在しないため、高精度なIlluminaショートリードを用いてNanoporeゲノム配列を修正し、修正前後の配列を比較することで精度評価を行った。全18菌株におけるNanoporeゲノム配列と修正後配列との差は、SNVs (Single Nucleotide Variants)が0～11個、InDelsが3～25個であった（図1）。SNVsとInDelsの中央値はそれぞれ2および7であり、Sakai株での検証結果と一致していた。また、18菌株中14株でSNVsが0～2個、13株でInDelsが8個以下であったことから、現在のNanoporeリードによるO157染色体配列には、2 SNVsおよび8 InDels程度の修正箇所が含まれる可能性が高いことが示唆された。また、Nanoporeリードの合計データ量（69x～419x）とSNVs/InDels数の相関を解析したところ、100x以上のデータ量増加に伴うSNVs/InDels数の減少は確認されなかった。したがって、一定以上のリード量の増加はNanoporeゲノム配列の精度に影響しないと考えられた。

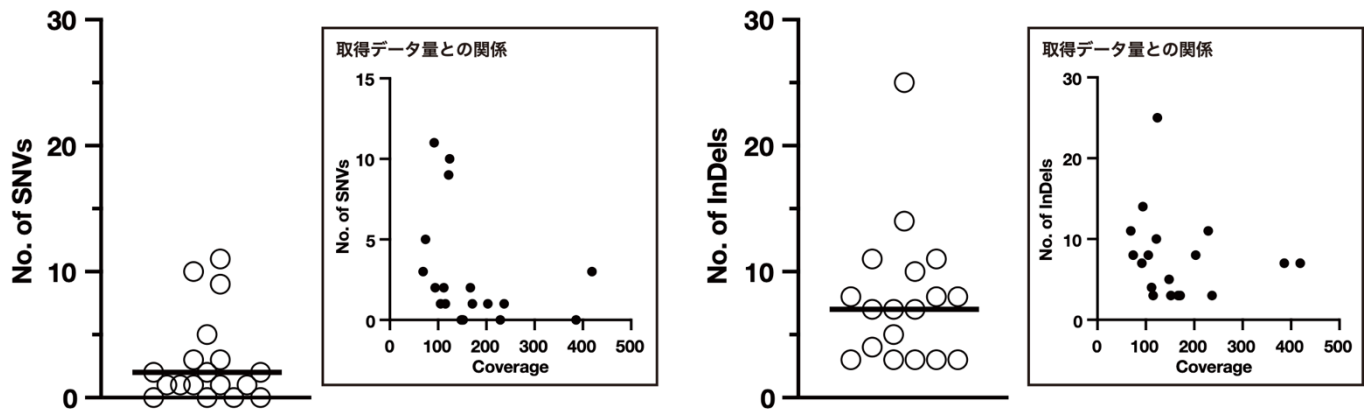


図 1. Nanopore ゲノム配列修正前後における SNVs および InDels の比較

EHEC O157 菌株（全 18 株）について、Nanopore リードのみで構築した染色体ゲノム配列とその配列をイルミナ配列で修正した配列を比較した。グループ化されたカラム散布図により、修正前後の塩基配列の差（SNVs あるいは InDels の数）の分布を示した。各ドットは個別の菌株を示し、横線は中央値を表す。また、SNVs/InDels の数と、ゲノム構築に使用した Nanopore リードのデータ量（genome coverage で表記）との関係をインセット内に示した。

（2）Nanoporeゲノム配列を用いたWGS解析の精度の検証

Nanoporeゲノム配列をWGS解析に利用した際、SNPcasterを用いた既存の解析結果とどの程度の乖離があるかを検証した。評価項目として、菌株間の最大・最小・平均 SNPs数およびSNPsの標準偏差を算出し、比較を行った（表1）。SNPcasterはIlluminaリードを用いた解析には対応しているが、Nanoporeリードには対応していない。そのため、本検証では、分担者自身の所属する研究室で開発され、既報となっているin-houseプログラムを用いて、構築済みの各菌株ゲノム配列からSNPsを検出し、SNPcasterの結果と比較した。

まず、沖縄県集団事例分離EHEC菌株をSNPcasterで解析したところ、最大3 SNPsの差異が認められた。一方、Illuminaリードを用いてin-houseプログラムで解析した結果は最大4 SNPsであり、最小・平均SNPs数および標準偏差についてもSNPcasterの結果と類似していた。したがって、これら2つのプログラムによる解析結果は同等であることが明らかとなった。

次に、Nanoporeゲノム配列をin-houseプログラムで解析したところ、最大11 SNPsの差異が認められ、SNPcasterおよびIlluminaリードを用いた解析結果と比較してSNPs数が増加した。一方で、Illuminaショートリードで修正した後のNanoporeゲノム配列を用いた場合には、最大4 SNPsとなり、その他の指標についてもIlluminaリードを用いたin-houseプログラムの結果と一致した。以上の結果から、Nanoporeゲノム配列をそのままWGS解析に用いた場合にSNPcasterよりも検出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。

表 1. 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

（3）EHEC感染症集団事例における起因菌の伝播経路推定

沖縄県EHEC感染症集団事例に由来するO157菌株の完全長ゲノム配列を用いて行ったこれまでのSNPs解析結果に基づいて、MSTを作成し、菌株間の遺伝的差異を確認した（図2A）。複数菌株に共通した染色体上の塩基変異は一箇所（*cheB*遺伝子上のA371G変異）のみであり、合計18菌株を7株（Group I）と11株（Group II）に区別した。さらに、全菌株に保存される3つのプラスミドのうち、1つのプラスミド上の塩基欠失がGroup IIの3菌株に共通して存在することが明らかとなった。この塩基欠失によって、Group IIの11株は、8株（Group II a）と3株（Group II b）に細分化された。

以上の塩基配列に基づく分類から、伝播経路は「① Group I → Group IIa → Group IIb」または「② Group IIb → Group IIa → Group I」のいずれかであると考えられた。ここに検体情報を照らし合わせ、今回のEHECの伝播経路は①であると推定した（図2B）。

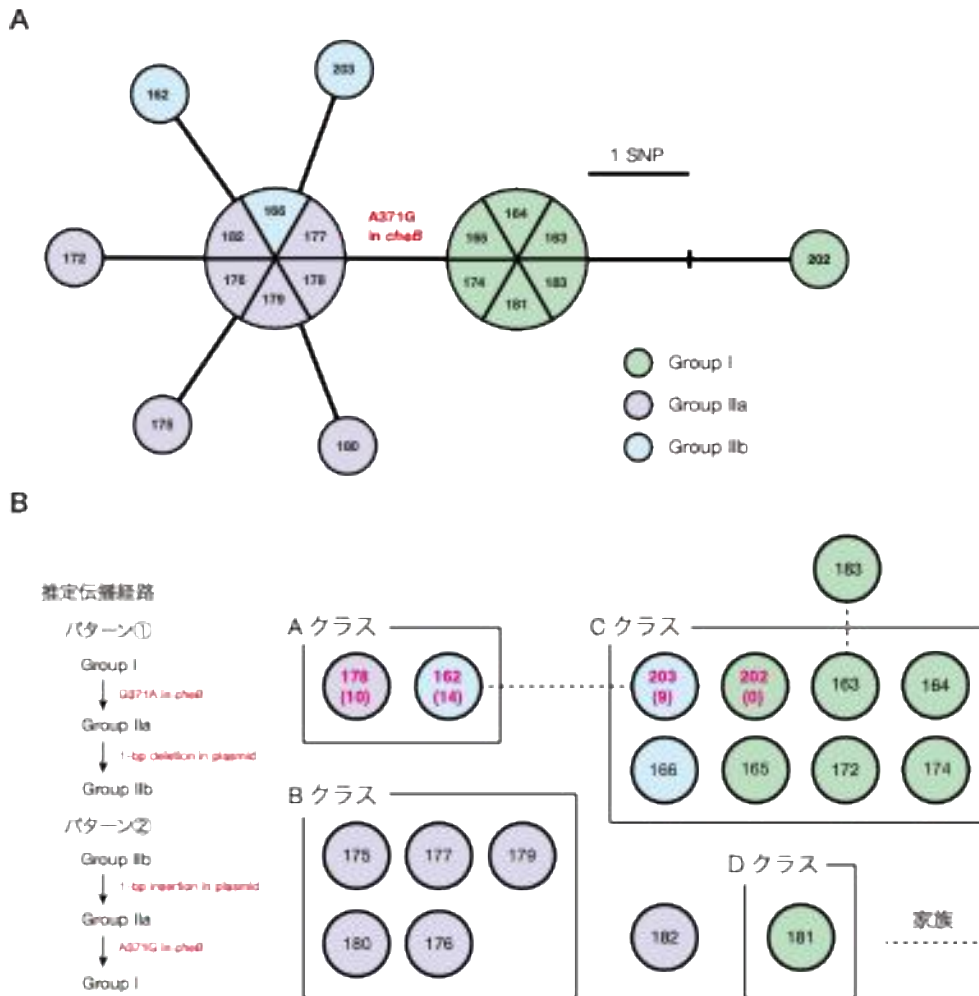


図 2. EHEC 感染症集団事例起因菌の伝播経路の推定

A. EHEC O157 菌株（全 18 株）の WGS 解析によって染色体上に検出された SNPs に基づく Minimum Spanning Tree (MST)。赤字で示した *cheB* 遺伝子上の 1 塩基変異とプラスミド上に特定された 1 塩基欠失により、18 菌株は緑・紫・青で示した 3 つのグループに分類された。

B. 菌株が分離された検体の情報。保育園のクラスと検体 ID を示す。有症者を赤字で示し、初発例からの経過日数を括弧内に示す。初発から次の発生までに 9 日間経過している状況から、パターン①の伝播の可能性が高いと推定された。

(4) WGSデータベースの拡充

離島コホート研究により前年度までに取得したウシ由来EPEC（病原性大腸菌）およびEHECのゲノム配列を公共データベースに登録した。登録にあたっては、沖縄県衛生環境研究所の職員に対し作業手順の指導を行い、地方衛生研究所におけるゲノム解析成果の公開を促進した。また、登録したEHEC（計39株）については、ゲノムリソースとしての価値が高いと判断し、Microbiology Resource Announcements誌にて公表した。

本年度の離島コホート研究で分離されたウシ由来EHECについてもWGS解析を行い、全18菌株の配列を決定した。血清型は15種類と多様であり、広範な系統のEHECが含まれていることが示唆された。なお、全菌株が *stx* 遺伝子を保有していた一方、*eae* 遺伝子陽性株は3株に留まった。

さらに、沖縄県内で発生した集団感染事例から分離された *Campylobacter jejuni* 計15株について、Nanoporeリードの取得およびWGS配列の構築を完了した。これらのゲノムデータは次期プロジェクトへ引き継ぎ、詳細な解析を行う予定である。

(5) 地方衛生研究所のゲノム解析環境の構築に関する支援活動

前年度に続き、九州地方の衛生研究所職員を対象とした研修会において、「病原細菌の全ゲノム解析方法の現状」と題して講演を行った。本年度は、参加可能な施設数の減少およびウェビナー開催における技術的制約があったため、より広範かつ効率的な情報共有を図るべく、講演内容を動画教材として作成し配布した。あわせて、教材の内容および今後解説を希望するゲノム解析技術に関するアンケートを実施した。その結果、教材は概ね高く評価され、地方衛生研究所の職員がゲノム解析に関してどのような知識や技術を必要としているのかという、具体的なニーズを把握することができた。

D. 考察

WGS解析の精度と実用性

本研究では、実際の食中毒事例から分離されたEHEC O157菌株を用い、Nanoporeゲノム配列がIlluminaリードを用いて構築される配列と同等の精度でWGS解析が可能であることを検証した。前年度の検証結果と同様に、Nanoporeゲノム配列に認められた差異の多くはIlluminaリードによる僅かな修正に留まっており、Nanoporeリードのみによるゲノム構築精度の高さが改めて示された。一方で、解析対象とした18菌株のゲノム配列は極めて類似していたものの、修正される塩基数には菌株間でばらつきが認められ、特にSNVs数において一定の差異が見られた。Nanoporeシーケンシングの精度は近年飛躍的に向上しているが、挿入・欠失(InDels)のエラーは依然として一定程度発生し、本解析においてもリード量の増加がこのエラー率の低減に寄与しないことが明らかとなった。これはNanoporeのシーケンシング原理に起因するものと考えられ、今後も不可避免なエラーとして残存する可能性がある。したがって、Nanoporeゲノム配列を利用する際は、一定数のInDelsが含まれることを前提に解析を行う必要がある。また、EHECゲノムにはプロフェージや挿入配列といった類似配列が複数箇所に存在するため、ショートリードを用いた修正(ポリッシング)が必ずしも正確に行われない可能性についても十分に留意すべきである。

Nanoporeシーケンシングの配列エラーはWGS解析の結果にも直接的に影響し、本解析では菌株間の最大SNPs数が11 SNPsとなり、SNPcasterによる解析結果を上回る結果となった。前年度までの研究により、食中毒事例において疫学的関連性のある菌株間をSNPcasterで解析した際の、遺伝的差異は10 SNPs以下であることが示されている。Nanoporeゲノム配列を用いた解析結果はこの閾値を僅かに上回っていたことから、この配列をWGS解析に適用する際には、従来の基準とは異なる「同一クローン判定基準」を策定するためのさらなる検証が必要であると考えられる。この一方で、Nanoporeリードによる完全長ゲノム配列の構築は、WGS解析に大きな利点をもたらした。具体的には、今回の食中毒事例における伝播経路の推定において、プラスミド配列の比較解析が不可欠であったが、これは完全長ゲノム配列を用いることで初めて可能となった。ショートリードベースのSNPcasterでは不可能な解析領域をカバーできる点は、Nanopore WGS解析の極めて有用な利点であると言える。

Nanoporeシーケンシングの運用上の課題

ゲノムデータベースの拡充を目的としたEHECおよびカンピロバクターのWGS解析において、ほぼ全ての菌株で完全長配列を構築することに成功した。特に、長鎖DNAの精製が可能なGenomic-Tipの利用が、安定した配列構築に寄与したと考えられる。しかし、Genomic-Tipの実務上の課題として、DNA精製工程における所要時間の不安定さが挙げられる。Genomic-Tipを用いたカラム精製では、菌株によってカラムへの透過速度が異なり、精製に多大な時間を要する場合がある。また、溶出後のDNAを適切なバッファーで再懸濁し、安定した濃度を得るための溶解プロセスにも時間を要する。以上の点から、DNA精製工程がNanopore解析全体のボトルネックとなっており、これは地方衛生研究所で実際にシーケンシングに従事している職員との共通認識でもある。今後、NanoporeによるWGS解析を迅速な検査体制として実装させるためには、解析時間の安定化および短縮を実現するためのDNA精製手法の最適化が急務であると思われる。

地方衛生研究所における解析環境の教育的支援

SNPcasterの導入により、地方衛生研究所におけるWGS解析の心理的・技術的なハードルは大幅に低下した。しかし、ツールの操作性と並行して、その背景にあるWGS解析の基礎知識を十分に習得している職員は必ずしも多くはなく、衛生研究所間においても知識の乖離が見られる。特に、数年単位で人事異動が発生する地方衛生研究所の組織特性から、解析技術および知識の継承は極めて困難な課題である。講演会や実技指導といった対面形式の教育は重要であるが、時間的・物理的な制約から頻繁に開催することは困難である。そこで本年度に実施した動画教材による情報共有は、時間や場所に縛られず、必要なタイミングで学習できる点で非常に有効であった。近年の動画制作ツールの普及と配信プラットフォームの整備により、こうした教材の作成と共有は比較的容易に実施可能となっている。動画教材を用いて、SNPcasterを運用するために最低限必要な基礎知識を効率的に伝達する仕組みを構築できれば、職員の入れ替わりがあっても解析精度を維持できる「安定した運用動線」を確保できると考えられる。

E. 結論

本研究により、Nanoporeシーケンシングを用いたWGS解析が、集団食中毒事例におけるEHEC菌株の解析および伝播経路の推定において十分に有用であることが明らかとなった。また、地方衛生研究所との連携による技術指導を通じて、取得したゲノム配列の公共データベースへの登録および国際誌への公表を実現し、地域の解析能力向上と学術的貢献を同時に達成した。また、離島コホート研究由来のEHECおよび集団食中毒事例由来のカンピロバクター菌株についてもWGS配列を決定し、公共データベースへの登録に向けた基盤を構築した。さらに、動画教材を活用した知識伝達モデルを提示し、九州地域の地方衛生研究所におけるWGS解析の普及と、持続可能な情報共有体制の構築に寄与した。

F. 健康危険情報

なし

(分担研究報告書には記入せずに、総括研究報告書にまとめて記入)

G. 研究発表

1. 論文発表

Kakita T, Iyoda S, Taira H, Kudaken 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 Reso

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。

H. 知的財産権の出願・登録状況（予定を含む。）

1. 特許取得
なし
2. 実用新案登録
なし
3. その他
なし