

厚生労働行政推進調査事業費補助金（健康安全・危機管理対策総合推進事業）
分担研究報告書

次世代シーケンサーに関わる各種検査マニュアルの作成：
疫学調査における NGS 解析の利用に関する研究

研究分担者 久保田 寛顕 東京都健康安全研究センター 課長代理
野本 竜平 神戸市健康科学研究所 副部長

研究要旨

集団食中毒をはじめ病原性細菌に起因する集団感染が発生した場合、地方衛生研究所ではパルスフィールド電気泳動（PFGE）による系統解析を行い、その結果をもって保健所等と連携し、事例究明と収束に向けた対応に当たってきた。今後、PFGE に代わり、一塩基単位の解像度で全ゲノム比較を行うことが可能な次世代シーケンサー（NGS）を用いた菌の比較が求められる。本研究課題では系統解析法としての NGS の活用方法を検討するとともに、その手法を統一化することを目的とした精度保証とマニュアル作りを行い、全国の地方衛生研究所を対象とした研修に落とし込む。特に本課題では NGS により取得したリードデータを解析し、菌株間の近縁関係を数値化する部分を方向づけることに焦点を当てる。

まず、PFGE の代替法という観点から、同一サンプルに対して PFGE と NGS の二法による分析を実施し、直接的な比較検証を行った。*Acinetobacter baumannii* 14 株（うち、13 株が Sequence Type [ST] 208、1 株が ST1316）をテストサンプルとしたところ、同一 ST である 13 株間ではバンドの相違数が 5 本以内、一塩基変異（SNV）数が 35 以内という、両法ともに近縁という結果が得られつつも、それらの数値には明確な相関が見られなかった。したがって、少なくとも全ての菌種に対して一様に PFGE における疫学的解釈を単純に NGS に置き換えることは難しく、集団感染とみなす SNV の線引きについては各菌種において検討する必要があると考えられた。さらに、NGS ではマッピング解析に用いるリファレンスの選定や、遠縁株を含むか否かという条件によっても検出される SNV の数が影響されることを確認した。これらのことは、いずれもリファレンスならびに供試株が共通に保有し、解析対象となるコアゲノム領域の増減によるものであり、リファレンス自身が供試株に対して遠縁な場合や、供試株の中に遠縁な株が含まれた場合にコアゲノム数が減少し、結果として検出 SNV 数の減少につながっていた。

今年度、NGS 解析の注意点として、リファレンスの選定、遠縁株の除外というポイントが明確になった。今後はこれらに加えて低品質のデータを除くといった工程を経た上で、マニュアル整備へと移る。その際は、本研究班における他の個別課題や、地衛研ネットワークの中で行われる研修などに参加するなどし、全国地衛研の職員との情報共有を行っていくことを考えている。

A. 研究目的

集団食中毒をはじめとする病原体による集団感染が発生した際、全国の地方衛生研究所（以下、地衛研）では本庁、保健所等との連携のもと、事例の究明と収束に向けた対応にあたる。この際、地衛研では主としてヒトや環境サンプルなどの分析を行い、科学的データを与えることをもって疫学調査をより強固なものとする役割を担う。特に腸管出血性大腸菌など、細菌が原因となる集団感染が起きた際は、分離株の一致、不一致を見る方法として、これまでパルスフィールド電気泳動（PFGE）法がゴールドスタンダードとして用いられてきた。しかし、PFGE 法に用いる標準的な泳動装置が今後製造終了になり、使用不能となることが知られている。このため、集団感染を追跡する上で、PFGE 法に代わる方法の策定が喫緊の課題である。

そこで、本研究課題では PFGE の代替法として次世代シーケンサー（NGS）の活用を検討する。従来、菌型を比較する方法として PFGE は最も解像度の良い方法であったが、NGS の登場により一塩基単位での比較が可能となり、PFGE と同等か、それ以上の解像度で集団感染を追跡できるようになってきた。その一方で、NGS は機器から出力される膨大なデジタルデータをバイオインフォマティクスにより処理して結果としていく必要がある。また、疫学調査との比較として一定の方法論（表 1）がコンセンサスを得ている PFGE に比べると、NGS では一塩基変異（SNV）の有無による菌の一致、不一致を線引きする基準がいまだ曖昧な状況にある。

表 1 PFGE における解釈
(Tenover et al., 1995, J. Clin. Microbiol.)

相違バンド数	近縁関係	疫学的解釈
0 本	区別不能	集団感染
2 から 3 本	近縁	集団感染の可能性が高い
4 から 6 本	近縁の可能性	集団感染かもしれない
7 本以上	遠縁	非集団感染

そこで、NGS を実務的な集団感染対応へ活用するためにも解析手法の統一化やその精度を保証する仕方を整理するための検討を行っていく。なお、次世代シーケンサーにより生データを取得する工程、すなわちウェットの作業については地衛研ゲノム解析ワーキンググループにおける統一的マニュアルの作成と重複する。さらに、実際に行行政依頼に対して保健所等に結果を発行する形式に関しては各自治体に適した形があると考えられる。したがって、本研究課題ではその間をつなぐデータ解析部分を対象とする（図 1）。

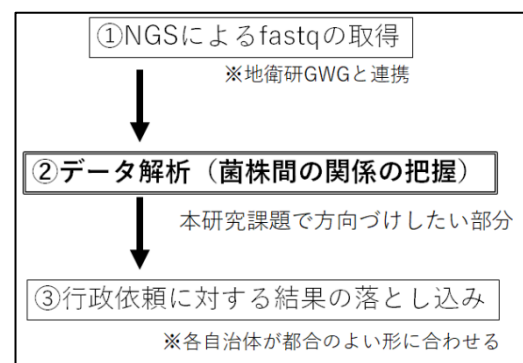


図 1 地衛研における NGS 対応

B. 研究方法

本研究課題は PFGE の代替法としての NGS の活用方法を探るものであるため、まずは同一サンプルに対して PFGE と NGS の分析を行い、出力されたデータについての直接的な比較検討に着手した。加えて、一塩基多型（SNP）解析のマニュアル化を行う上で重要と考えられる、解析結果に対して影響を与えうる要因を特定するために同一データに対して解析条件を変化させ、その結果を評価することとした。

まず、テストサンプルとして、過去に分離された *Acinetobacter baumannii* 14 株（うち、13 株は ST208 であり、残り 1 株は ST1316）について、両法にて分析を行った後、相関の有無の判断や、NGS 解析におけるリファレンス配列の違い等に起因する解析結果の変動について評価を行った。

1) PFGE

米国 CDC の方法に従って菌液からプラグを作成し、制限酵素 *ApaI* にて処理した後、CHEF Mapper システム (バイオラッド) にて電気泳動を行った。EtBr にて染色して得られた泳動パターンから、目視にて異なるバンドの本数を判定した。

2) NGS

菌株から抽出したゲノム DNA を用いて Nextera XT Library Preparation Kit (イルミナ) にてライブラリ調製を行った後、MiSeq (イルミナ) にて 300bp × 2 のペアエンドのリードデータを得た。次に、得られたリードデータを BactSNP を用いてリファレンスに対するマッピング解析により SNP コールを実施した。その後 Snippy によりコア領域の抽出とコアゲノムアライメントを取得した。取得したアライメントファイルから Gubbins を用いることで組換え領域を推定し、その個所を除去した SNP セットを取得した。マッピング時のリファレンスには、国際塩基配列データベース上に登録されている *A. baumannii* の Type Strain である ATCC19606 株 (以下、標準株)、または ST208 である XH386 株 (以下、ST208 株) のゲノムデータを用いた。

3) データ解析結果の検証

Snippy によるアライメント時、ならびに Gubbins 処理時の出力ログファイルを用い、計算に使用された SNP の数を評価した。

(倫理面への配慮)

本研究で用いた臨床分離株は、分離元とは連結不能な匿名化がされていることに加えて、本研究課題では一切の個人情報を用いず、個人が特定されることはない。また、東京都健康安全研究センター倫理審査委員会により、承認を受けているものである (6 健研健第 906 号)

C. 研究結果

1) PFGE における相違バンド数と SNV 数の比較

NGS で標準株をリファレンスとした場合の結果を図 2 に示した。検討に使用した 14 株の中で ST が異なる一株についてはそのほかの株に対して PFGE では 10 本以上のバンドが異なっており、また、SNV の数も 20 万を超えていたため、図には示していない。

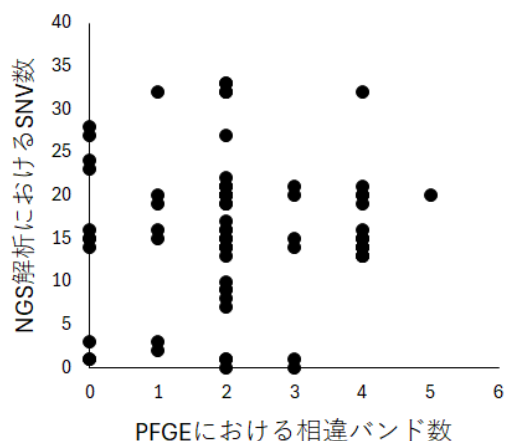


図 2 PFGE と NGS の比較

また、リファレンスを標準株から ST208 株に変更した場合、ならびに解析対象から ST1316 の一株を除いた場合に算出された SNV 数の変化を図 3 に示した。比較しやすいように、標準株と同じ SNV 数だった場合の理論直線を同時に示している。

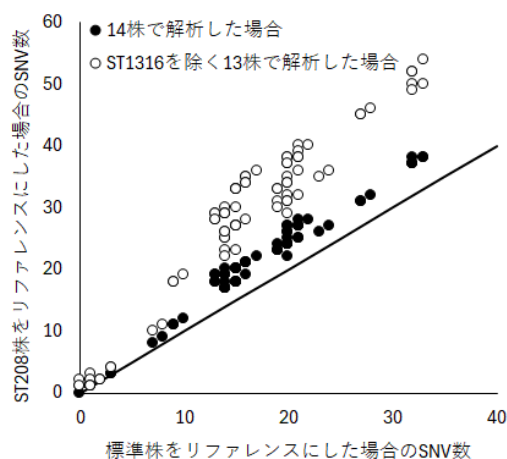


図 3 リファレンス、解析株数による変化

2) 解析対象となった SNP 数の検証結果

今回の解析でリファレンスやサンプルセットを変更した際のコアゲノム (全株に共通の領域) の変化を図 4 に、またそれぞれ

の解析での総 SNV 数と組換え領域内として除去された SNV 数を表 2 に示した。

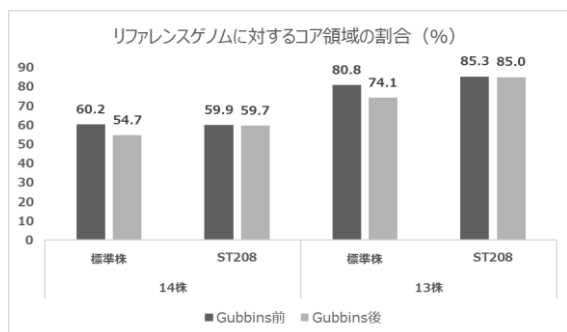


図 4. 各解析によるコア領域の変化

表 2 各解析における SNV 数

解析株数	14 株		13 株	
	標準株	ST208	標準株	ST208
総 SNV 数 (Gubbins 前)	242,951	228,239	55,036	648
除去された SNV 数	23,950 (9.9%)	924 (0.4%)	17,392 (31.6%)	478 (73.8%)

リファレンスゲノムに対するコアゲノム領域の割合は 14 株で解析した場合は 60% 前後であったのに対し、異なる ST の 1 株を除いた 13 株の解析では 80~85% となった。一方でリファレンスを変更した場合はコアゲノム領域の割合に与える影響は軽微であった。

また、各解析における総 SNV 数は 14 株で解析した場合はリファレンスに関わらず 20 万個を超える SNV が検出された一方で、13 株での解析では標準株をリファレンスにした場合は約 55,000 個、同一 ST をリファレンスにした場合は約 650 個の SNV となった。Gubbins による組換え領域除去を実施すると、14 株、13 株どちらの解析においても、リファレンスが標準株の場合は 17,000~23,000 個程度の SNV が除去され、リファレンスを ST208 にした場合は数百個の SNV のみが除去された。

D. 考察

当初は PFGE と NGS との結果の間に一定の相関があり、相違バンド数から SNV 数に差し替えることによって PFGE での疫学的解釈を NGS での解析に利用する可能性

も想定していた。しかし、少なくとも同一 ST となった *A. baumannii* 株については二法による相関は小さかったことから、一概にそうした図式は成り立たないと考えられた。ただし、本試行で用いた *A. baumannii* が特に二法の相関が小さい菌種だった可能性も否定できないことや、ST208 が国際流行株として知られる特定の系統だったことから、特異的な結果を生んだ可能性も否定できない。少なくとも、あらゆる菌種に対して菌種の一致、不一致を決める統一的な SNV の線引きを設定することは難しいと考えられた。

また、算出される SNV 数がリファレンスとする配列や、解析に含む株によっても変動することが確認された。コアゲノム領域については、リファレンス配列も含めた解析配列全体で判定されるため、今回の解析において異なる ST が混入していた 14 株の解析では大きくコアゲノム領域が減少することとなった。コアゲノム領域外にある SNV は解析には含まれないため、解析の分解能を向上させるためには遺伝系統的に離れた株は解析から排除することが望ましいと考えられる。また、組換え領域の除去に関してはリファレンスの影響が大きかった。これは Gubbins というソフトの仕様上、一株のみに検出される SNV では組換え領域として判定されないという事が寄与している。今回の解析では遺伝系統的に離れた株は 1 株だったため、その株を除去するかどうかは組換え領域の解析ではあまり影響せず、リファレンスを変更することにより近接 SNV がどれだけ検出されるかどうかの方が影響が大きかったと考えられる。

E. 結論

集団感染を追跡するツールとして PFGE から NGS へと移行するにあたり、まず SNV 数による線引きを決めていくことが急務であるが、それと同時に解析方法についても一定の道筋を提案することが重要である。本研究課題の一年目の結果から導かれる、統一する必要がある、解析結果に影響を及ぼすと考えられる部分はリファレンスの選定と、解析対象に含む株であった。

その一方で、今回は異なる ST という遠縁株が含まれたことによる解析結果の変化が大きかったが、例えば近縁であっても低品質のリードデータが含まれると、その株のみならず全体の解析結果に悪影響を及ぼす可能性があるため、二年目は早々にこのことについても検証を進める。

そして、そうした解析の注意点が絞られ次第、解析マニュアルの作成に移るとともに、地衛研職員との情報共有を図っていく。マニュアル化に際しては、リードデータから SNV 数を導く部分はオープンソース、あるいは市販のソフトウェアの使用方法に準拠することになると想定されるため、本研究課題ではそのようなソフトウェアから出力される結果からリファレンスを選定して再解析を行う方法や、遠縁の株を除くといった自動化されない、使用者の判断に委ねられる部分についてマニュアルを整備していくことを考えている。さらに、他の個別課題における研修や、地衛研ネットワークの中で行われる講習などに一部相乗りするような形で周知も考慮していく。

F. 健康危険情報

なし

G. 研究発表

1. 論文発表

- 1) Mitobe M, Kubota H, Kobayashi K, Miyake H, Takano M, Mizushima D, Gatanaga H, Oka S, Suzuki J, Sadamasu K. Clustering of Polymorphic Membrane Protein E Clade in *Chlamydia trachomatis* Lineages from Men Who Have Sex with Men. *Emerg Infect Dis.* 30(10):2183-2187. 2024.
- 2) Uchitani Y, Okuno R, Ariyoshi T, Kubota H, Suzuki J, Sadamasu K. Genetic characteristics of invasive pneumococcal disease-derived *Streptococcus pneumoniae* of serogroup 24 isolated in Tokyo, Japan. *J Infect Chemother.* 202. 31(1):102484. 2025.
- 3) Ando N, Mizushima D, Takano M, Mitobe M, Kobayashi K, Kubota H, Miyake H, Suzuki J, Sadamasu K, Aoki T, Watanabe K, Oka S, Gatanaga H. Prolonged sitafloxacin and doxycycline combination regimen for treating infections by highly resistant *Mycoplasma genitalium*. *J Antimicrob Chemother.* 2025. 80(1):247-253. 2025.
- 4) Kubota H, Okuno R, Kenri T, Ucitani Y, Ariyoshi T, Yoshida I, Kobayashi, K, Mitobe M, Suzuki J, Sadamasu K. Multiplex Amplicon Sequencing for the Comprehensive Genotyping of *Mycoplasma pneumoniae*. *Microbiol. Spectr.* In press.
- 5) Nakamura K, Taniguchi I, Gotoh Y, Isobe J, Kimata K, Igawa Y, Kitahashi T, Takahashi Y, Nomoto R, Iwabuchi K, Morimoto Y, Iyoda S, Hayashi T. Diversity of Shiga toxin transducing phages in *Escherichia coli* O145:H28 and the different Shiga toxin 2 production levels associated with short- or long-tailed phages. *Front Microbiol.* 15:1453887. 2024.
- 6) Nomoto R, Nakanishi N, Komatsu S, Matsui M, Suzuki S. Complete genome sequences of *Escherichia coli* KA0011 clinical isolate used as a quality control strain of carbapenem susceptibility testing in Japan. *Microbiol Resour Announc.* 13:e0015524. 2024.
- 7) Nomoto R, Osawa K, Kinoshita S, Kitagawa K, Nakanishi N, Sarassari R, Raharjo D, Fujisawa M, Kuntaman K, Shirakawa T. Metagenome and Resistome Analysis of Beta-Lactam-Resistant Bacteria Isolated from River Waters in Surabaya, Indonesia. *Microorganisms.* 12:199. 2024.
- 8) 長島真美, 藤原卓士, 村田ゆかり, 長谷川乃映瑠, 有吉司, 上原さとみ, 小西典子, 門間千枝, 浅倉弘幸, 林真輝, 熊谷遼太, 原田幸子, 永野美由紀, 山崎貴子, 河上麻美代, 糟谷文, 矢尾板優, 黒木絢士郎, 天野有紗, 北村有里恵, 磯貝まや, 小杉知宏, 加来英美子, 鈴木薫, 南須原亮, 田部井由紀子, 小川

麻萌, 中村三琴, 後藤千恵, 村内このみ, 古谷実, 原田順子, 川合由華, 小池浩二, 灘岡陽子, 中村貴恵子, 久保田寛顕, 安中めぐみ, 吉田勲, 小林甲斐, 水戸部森歌, 奥野ルミ, 内谷友美, 田淵優里, 森功次, 高橋由美, 村上昂, 和田紀乃, 河村真保, 赤瀬悟, 三関詞久, 中里彩乃, 小野明日香, 浅山睦子, 新井麻里, 鈴木愛, 岡田若葉, 古田菜摘, 齊木大, 前田雅子, 尾畑浩魅, 西野由香里, 井田美樹, 宗村佳子, 千葉隆司, 新開敬行, 横山敬子, 三宅啓文, 鈴木淳, 貞升健志. 健康危機管理に関連する微生物の分子疫学解析と検査法の開発に関する研究. 東京健安研セ年報.

75. 27-40. 2024.

9) 神田真軌, 小林甲斐, 西野由香里, 大場由実, 中島崇行, 栗田翔, 久保田寛顕, 有吉司, 神門幸大, 吉田勲, 安中めぐみ, 水戸部森歌, 長谷川乃映瑠, 尾畑浩魅, 畠山薫, 福井理恵, 市川めぐみ, 三橋華子, 黒田寿美代, 山崎華恵子, 井田美樹, 小林真紀子, 上原さとみ, 高橋由美, 高林晃子, 浅山睦子, 吉川聡一, 林洋, 松島陽子, 小池裕, 林もも香, 永野智恵子, 富澤早苗, 八巻ゆみこ, 上條恭子, 山本和興, 高田朋美, 小鍛治好恵, 渡邊趣衣, 志良堂裕子, 大澤佳浩, 小山彩音, 野口舞子, 大貫文, 岡優香, 山崎貴子, 相原三菜子, 渡邊喜美代, 鈴木淳也, 小田智子, 鈴木淳, 横山敬子, 大塚健治, 横山知子, 新藤哲也, 中嶋順一, 木下輝昭, 貞升健志, 笹本剛生, 猪又明子. ワンヘルスアプローチに基づく抗微生物薬と薬剤耐性微生物の実態把握に関する研究. 東京健安研セ年報. 75.

41-57. 2024.

10) 安中めぐみ, 長谷川乃映瑠, 久保田寛顕, 吉田勲, 鈴木淳, 貞升健志. 結核菌検査におけるゲノム解析法に関する検討. 東京健安研セ年報. 75. 115-118. 2024.

2. 学会発表

1) 小林甲斐, 久保田寛顕, 水戸部森歌, 有吉司, 鈴木淳, 貞升健志. 染色体上の *blactx-M-14* の増幅が認められたイヌおよびネコ由来基質特異性拡張型 β -ラクタマーゼ産生大腸菌 *sequence type 38* の分子疫学的性状. 第36回日本臨床微生物学会総会・学術総会. 2025.

2) 内谷友美, 奥野ルミ, 有吉司, 吉田勲, 久保田寛顕, 鈴木淳, 貞升健志. 侵襲性肺炎球菌感染症由来株の血清型と血清群 24 の分子疫学解析. 第36回日本臨床微生物学会総会・学術総会. 2025.

3) 有吉司, 内谷友美, 奥野ルミ, 吉田勲, 久保田寛顕, 鈴木淳, 貞升健志. 都内医療機関から分離された *bla_{NDM}* 搭載 *IncX3* プラスミド保有腸内細菌目細菌の分子疫学解析. 第36回日本臨床微生物学会総会・学術総会. 2025.

H. 知的財産権の出願・登録状況

1. 特許取得

なし

2. 実用新案登録

なし

3. その他

なし