

中山人間科学振興財団活動報告書 2025 年度

国際交流助成（海外渡航）

研究テーマ

新興感染症 *Escherichia marmotae* の分子疫学解析

氏名

新井 沙倉
Sakura Arai

所属

国立医薬品食品衛生研究所 衛生微生物部
National Institute of Health Sciences, Division of Microbiology

【背景と目的】

細菌による感染予防のためには、細菌の自然界（環境や動物）での分布や生態を把握することが重要である。かつて大腸菌と近縁であるため cryptic clade V とされていたグループが新たに *Escherichia marmotae* と命名された。この一部の菌株は、ヒトへ侵襲性の感染症を引き起こす新興感染症だが（文献1）、本菌は近縁の大腸菌との区別が難しく（文献2）、ヒトへの感染ルートも明らかになっていない。そこで本研究では、本菌の先行研究が実施されているニュージーランド Massay 大学 Hopkirk research institute にて、本菌のゲノム情報（文献3）をバイオインフォマティクス解析することで多様な国から分離された本菌の遺伝学的特徴を明らかにする。さらに、本菌の病原因子の一部は大腸菌と共通することが明らかとなっているため、各種病原因子の分布状況についても明らかにすることで、ヒトへ感染性を示す病原性株の環境や動物での分布を明らかにする。以上より、ヒトの *E. marmotae* 感染症の制御に資する基礎知見を得る。

【方法】

1. ゲノム配列データの収集

本研究では、Assembly 手法の違いによるデータのばらつきを最大限抑えるために、Assembly 前の Illumina 社の装置で読まれた全ゲノム配列の生データ (Read) のみを収集した。Hopkirk research institute にて事前に取得されていたニュージーランド由来の 213 株に加え、データベースに登録されている SRR、ERR、DRR 番号が付与された全ゲノムシーケンス (WGS) の Read を取得した。

2. Nullarbor 解析と分析対象菌株の選定

当初収集した全ゲノム配列を対象に、Nullarbor2 解析を実施した。本解析は、細菌の全ゲノムシーケンスデータを自動解析するためのバイオインフォマティクス・パイプラインである。Read の品質評価とトリミング、ゲノムアセンブリ、遺伝子アノテーション、薬剤耐性、病原遺伝子の検出および系統解析結果を一度の解析で得ることが可能である。この解析の結果、低クオリティのデータ、重複して登録されていたデータなどを除外した。

3. 菌種の同定試験

2. を経て得られた fasta ファイルを用いて、供試菌株が *E. marmotae* であることを確かめるため、EzClermont、ECtyper および Clermont Typing の3種類の解析を実施した。

4. コアゲノム Multilocus Sequence Typing (cgMLST) 解析

微生物の全ゲノムデータの中から、その菌種に共通する Core Genome だけ

を選定し、株間での多様性を解析する cgMLST 解析を実施した。cgMLST 解析にてどの遺伝子がコアゲノムのセットであるか定める Scheme が *E. marmotae* ではまだ確立されていなかったため、大腸菌用の Scheme を用いた場合と、本解析にて用いた全菌株を基に新たに自家調整した Scheme の 2 種類を用いて解析し、その結果を Tanglegram を用いて比較した。

5. コアゲノム SNP (cgSNP) 解析

2-2. Nullarbor 解析より得られた cgSNP 情報を元に、iTOL ソフトウェアを用いて系統樹を作製した。系統樹上に株の名称、由来、血清型、シーケンスタイプ (ST) を表示することで、系統が何に基づいて分岐しているかを明示した。

6. 個別遺伝子解析

検出された病原遺伝子の遺伝子配列の相同性解析のために、Geneious ソフトウェアを用いた。また、プラスミドの解析では、RFplasmid、PlasmidFinder、MOB-suite の 3 種類の解析を実施した。

【結果と考察】

1. 解析データの情報

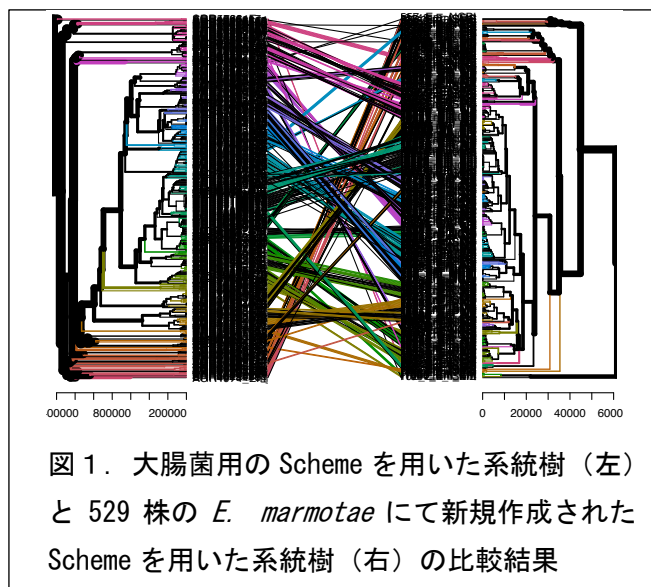
Nullarbor 解析後、データを整理したところ、529 株の WGS を用いて解析することが決定した。この 529 株は 1991 年から 2025 年までに 20 カ国から分離された菌株であり、ヒト由来株が 66 株、動物由来株が 216 株、環境由来株が 197 株、その他由来株が 50 株であった。

2. 菌種の同定

EzClermont 解析の結果、529 全株は *E. marmotae* の同定結果として正しい U/cryptic に型別された。ECTyper 解析の結果、529 全株は *Escherichia marmotae* に型別された。一方、Clermont Typing 解析の結果、529 株中 522 株は *E. marmotae* の同定結果として正しい cladeV に型別されたが、7 株は Non *Escherichia* に型別された。この 7 株のデータを Nullarbor 解析結果の 1 つである MLST 結果と併せて解析すると、7 株のうち 6 株は Sequence type (ST) 11819、1 株は ST5260 に属していた。ST11819 は本研究にて供試した全ての ST11819 が Non *Escherichia* に型別されており、後述する cgMLST および cgSNP 解析の結果、他の菌株と極端に離れた系統には属していないため、ST11819 の特徴として本解析では Non *Escherichia* という結果が得られるのだと考えられた。一方、ST5260 は本研究で供試した ST5260 が 56 株存在するものの、1 株だけが Non *Escherichia* と型別されたため、Assembly の誤差でこのような結果が得られたと考えられた。

3. cgMLST 解析

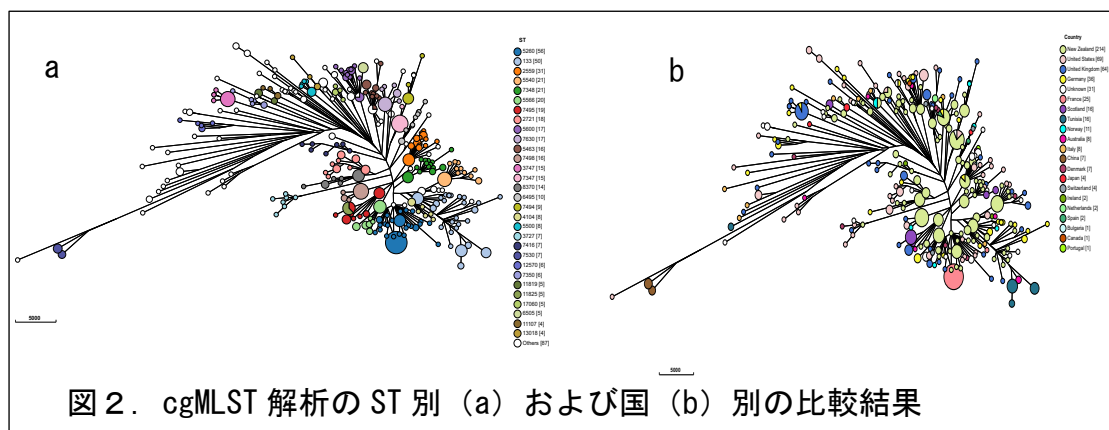
大腸菌用の Scheme を用いた場合と、本解析にて用いた全菌株を基に新たに自家調整した Scheme の 2 種類を Tanglegram を用いて比較したところ 2 者は大きく異なっていた (図 1)。



*E. marmotae*は大腸菌とは別種であり、大腸菌の Scheme に当てはめて解析することは本菌の遺伝学的特徴を考慮できないことが危惧されたため、本研究にて新規作製した Scheme が本菌を解析対象とした新たな Scheme として適切であると考えられた。

また、この新たな Scheme で作製された系統樹へ菌株が属する ST ごとに色分けすると、系統が ST ごとに分岐している

像が示された (図 2 a)。さらに、由来の国別に色分けすると、ニュージーランド由来株は多様な系統に分布していることが示された (図 2 b)。



4. cgSNP 解析

cgSNP 解析の系統樹に円の外側から血清型、国 (赤字がニュージーランド)、株の由来、ST 番号、そして血清型をラベルした (図 3)。さらに主要 ST は円の内側に扇形の色をつけた (図 3)。その結果、cgMLST の解析と同様に、cgSNP 解析の結果も ST ごとに系統が別れており、ニュージーランド由来株は多様な ST に分布していることが示された。

5. 病原遺伝子解析

大腸菌の病原遺伝子を参考配列に、どのような遺伝子を保有するかを

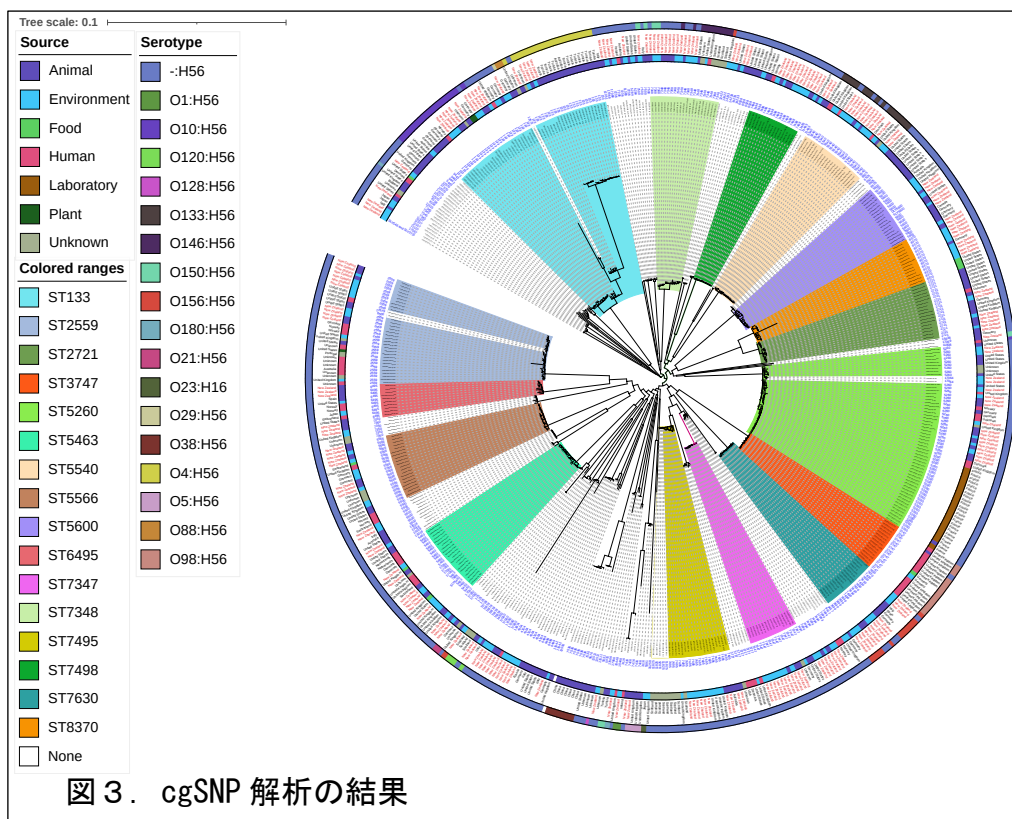


図 3. cgSNP 解析の結果

Nullarbor2 解析の一部である Abricate によって求めたところ、*eae* 遺伝子および *bfp* 遺伝子が検出された。

eae 遺伝子は、腸管に定着するための因子をコードしており、この遺伝子が陽性の *E. coli* ではヒトに下痢を引き起こす。今回この *eae* 遺伝子が検出されたのが ST3747 の 15 株のみであり、15 株間で配列の相同性が 99.9% 以上保存されていたため、*eae* 遺伝子はこの系統で保存されていると推定された。また、この 15 株の由来を見ると、12 株はヒト由来株であり、本研究に供試されたヒト由来株 66 株中 12 株 (18%) がこの系統に集中していた。そのため、*E. marmotae* ST3747 は、他の ST と比較して病原性が高い可能性が示唆された。この ST3747 はヒトの他にも、環境中の堆積物、動物、食品からも検出されているため、今後注意が必要である。

次に、*bfp* 遺伝子は、bundle-forming pili と呼ばれる特殊な線毛をコードする遺伝子であり、宿主細胞へ接着するのに重要な病原因子である。この *bfp* 遺伝子の解析結果では、ST5260、ST7350 および ST8370 に属する菌株がそれぞれ 1 株ずつ合計 3 株保有していた。*bfp* 遺伝子は *bfpA* や *bfpB* など複数遺伝子によるオペロンを形成し、プラスミドによって運ばれるため、3 株のプラスミド解析を実施した。その結果、*bfp* 遺伝子がコードされる contig は、RFplasmid ではいずれも Plasmid と判定され、PlasmidFinder では Plasmid または染色体と判定され、MOB-suite では Plasmid と判定されなかった。この結

果から、プラスミドの解析では、複数の解析手法を組み合わせ、その結果を総合して判定することが必要であることが示された。これら 3 株はいずれもニュージーランド以外の国で分離された株であり、Long read sequence 配列がデータベース上に存在しなかったため、追加解析は不可能であった。しかし、線毛の本体をコードする *bfpA* の配列長を確認すると、582-591 bp と 3 株とも異なっていたため、系統に関係なく、プラスミドを介した *bfp* 遺伝子の獲得が *E. marmotae* で起こっていることが推察された。

【結論】

大腸菌の近縁種である *E. marmotae* の大規模全ゲノム解析を実施し、本菌の系統学的な多様性を明らかにした。個別の遺伝子に着目した解析では、一部の ST は *eae* 遺伝子を有しヒトへの病原性が高いことが示唆される結果が得られた。本菌は、新興感染症のためヒトではまだ正確に同定されず報告数が限られている一方で、環境や野生動物をはじめとする多様な動物からの分離が世界中から報告されている。今後、どのようなルートでヒトが感染するのかを明らかにすることや、病原遺伝子や薬剤耐性遺伝子保有株の分布をさらに明らかにすることで、ヒトや動物への影響を正しく推測する必要があると考えられた。

【謝辞】

中山人間科学振興財団からのご支援により、ニュージーランドにおける研究活動および国際共同研究を実施することができました。心より感謝申し上げます。また、留学先においてご指導をいただいた Massey 大学 Patrick J. Biggs 教授、AgResearch の Adrian L. Cookson 先生、mEpiLab および Hopkirk Research Institute の皆様、そして国立医薬品食品衛生研究所の衛生微生物部の先生方に深く感謝申し上げます。

【参考文献】

1. Tanya Sinha et al., Unrecognised: isolation of *Escherichia marmotae* from clinical urine sample, phenotypically similar to *Escherichia coli*. Pathology. 2024 Jun;56(4):577-578. doi: 10.1016/j.pathol.2023.08.015.
2. Audun Sivertsen et al., *Escherichia marmotae* - a human pathogen easily misidentified as *Escherichia coli*. Microbiology Spectrum 2022 Apr 27;10(2):e0203521. doi: 10.1128/spectrum.02035-21.
3. Patrick J. Biggs et al., Draft genome sequences of *Escherichia* spp. isolates from New Zealand environmental sources. Microbiology Resource Announcements 2024 Mar 12;13(3):e0100723. doi: 10.1128/mra.01007-23.