

高解像度の遺伝子解析手法を用いた 琵琶湖・淀川水系における病原細菌の一斉検出

京都大学大学院 工学研究科 中西智宏

1. はじめに

近年の温暖化や気象の極端化を鑑みれば、水道水源における微生物リスク管理の重要性が増している。飲料水の微生物リスクを評価するためには、定量的微生物リスク評価 (Quantitative Microbial Risk Assessment; QMRA) 手法が用いられる。QMRA ではまず対象とする病原体を決定し、水源等での濃度を測定/推定したあとに感染確率の推定を行う。しかし、水源には多種多様な病原体が存在しており、始めに対象微生物を絞ることで本当に重要な病原体を見逃してしまう可能性がある。

一方、これまで琵琶湖・淀川水系においては *Campylobacter*¹⁾ や *Legionella*²⁾ といった細菌類、薬剤耐性菌³⁾、ウイルス^{4,5)}、原虫類⁶⁾が測定され、個々の病原体の分布調査や QMRA が行われてきた。しかし、「琵琶湖・淀川水系には潜在的にどれだけの種類の病原微生物が存在するのか？」ということは未だ明らかになっておらず、重要な病原体についての情報も乏しい。合理的な微生物リスク管理のためには、制御すべき病原微生物の多様性を正しく捉え、流域での分布特性を把握しておくことが望ましい。

上記を踏まえ、本研究では琵琶湖・淀川水系において病原細菌を網羅的に検出し、病原細菌の多様性や分布特性を把握することを目的とした。ここで、微生物の一斉検出には次世代シーケンサーを用いた網羅的な遺伝子解析が有用だが、細菌の病原性の有無を判断するには種レベルまでの高い分類解像度を備えた分析技術が必要である。そこで、本研究では 16S rRNA 遺伝子(16S rDNA)全長を解読対象としたメタバーコーディング手法に着目し、まずは適切な分析条件を決定した。その後、決定した分析条件を桂川～淀川水系の環境水に適用し、病原細菌の一斉検出を試みた。

2. 研究方法

2.1 標準試料に対する PCR 反応条件の決定

2.1.1 標準試料

PCR 条件の決定にあたっては、15 種の既知の細菌種が既知濃度で混合された標準試料を用いた。具体的には、独立行政法人 製品評価技術基盤機構 (NBRC) から標準 DNA (DNA-Mock-001) を購入した。その細菌種の組成と混合割合を表 1 に示す。この標準試料を鋳型として、2.1.2 に示す様々な反応条件で PCR 増幅を行った。

2.1.2 PCR 増幅とライブラリ調製、シーケンシング

上記試料を適宜希釈した後、LongAmp Hot Start Taq DNA Polymerase (New England BioLab) と 16S Barcoding Kit 1-24 (SQK-16S024; Oxford Nanopore Technologies 社, 以下 ONT 社) を用いて、真正細菌由来の 16S rDNA 全長を PCR 増幅した。増幅には Thermal Cycler Dice Gradient (タカラバイオ社) を使用した。アニーリング温度と伸長時間は増幅産物の細菌組成に影響しうる重要なパラメータであると判断し、アニーリング温度を 50, 55, 60 度、伸長時間を 1, 2, 3 分に設定し、25 サイクルの増幅を行った。反応液組成やその他の反応条件は原則 Barcoding kit 付

表 1 標準試料に含まれる細菌種の学名と 16S rDNA コピー数の存在割合

| 学名 | NBRC 番号 | 16S rDNA の存在割合 (理論値) (%) |
|---|---------|-----------------------------|
| <i>Bacillus subtilis subsp. subtilis</i> | 13719T | 10.5 |
| <i>Bifidobacterium pseudocatenulatum</i> | 113353 | 5.3 |
| <i>Clostridium butyricum</i> | 13949T | 11.6 |
| <i>Corynebacterium striatum</i> | 15291T | 4.2 |
| <i>Cutibacterium acnes subsp. acnes</i> | 107605T | 3.2 |
| <i>Lactobacillus delbrueckii subsp. delbrueckii</i> | 3202T | 8.4 |
| <i>Staphylococcus epidermidis</i> | 100911T | 6.3 |
| <i>Streptococcus mutans</i> | 13955T | 5.3 |
| <i>Acinetobacter radioresistens</i> | 102413T | 6.3 |
| <i>Bacteroides uniformis</i> | 113350 | 4.2 |
| <i>Enterocloster clostridioformis</i> | 113352 | 5.3 |
| <i>Comamonas terrigena</i> | 13299T | 7.4 |
| <i>Escherichia coli</i> (K-12株) | 3301 | 7.4 |
| <i>Parabacteroides distasonis</i> | 113806 | 7.4 |
| <i>Pseudomonas putida</i> | 14164T | 7.4 |

属のプロトコルに従った。増幅産物をビーズ精製し、2 本鎖 DNA 濃度を適宜調整した後全サンプルをプーリングしたものを Flow Cell R9.4.1(ONT 社)にロードし、MinION Mk1B(ONT 社)を用いて配列データを取得した。取得リード数はサンプル当たり 45 万~60 万リードであった。

2.1.3 配列データの解析

ONT 社の配布する解析ソフト Guppy ver. 5.0.16 の guppy_basecaller でベースコールした後、guppy_barcode を用いてバーコード配列に応じて検体ごとに配列データを仕分け、プライマー配列を除去した。その後、NanoCLUST⁷⁾という解析ソフトを用いてリードの品質フィルタリング、リードの 5mer 頻度データの次元削減 (UMAP 法)、HDBSCAN によるリードのクラスタリング、各クラスターにおけるリードのエラー補正を経て、主たる細菌種由来の代表配列を推定した。得られた代表配列を NCBI の Refseq データベースに対して BLAST 検索し、配列類似性が最上位にヒットした細菌種の分類群情報を取得し、種レベルまでの微生物名を割り当てた。なお、NanoCLUST の解析パラメータはデフォルトの値を用いた。

以上のデータ解析を PCR 条件ごとに行い、標準試料に含まれる細菌種の検出精度や感度、細菌組成の妥当性を評価し、最も適切な PCR 条件を選定した。

2.2 琵琶湖・淀川水系への手法の適用

調査対象地点を図 1 に示す。桂川の保津峡地点と宮前橋地点、さらに三川合流後の枚方大橋を対象地点とし、河川水を約 5L 程度採水した。また、宮前橋地点よりも上流に位置する鳥羽水環境保全センター放流口で下水処理後の放流水も採水した。採水時期は 2021 年 8 月~12 月とし、計 22 サンプルを採取した。

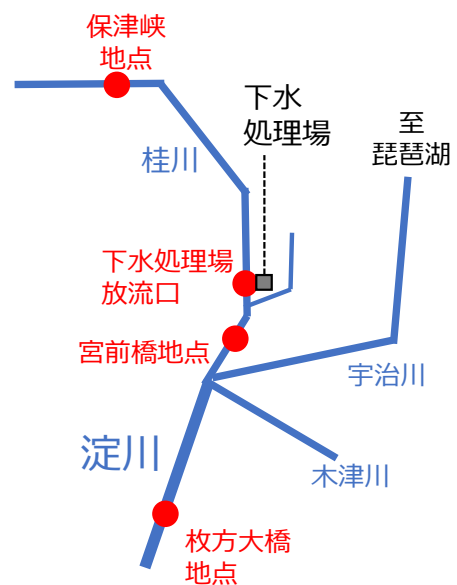


図 1 採水地点

採水した試料は実験室に持ち帰り、孔径 0.2 μm のメンブレンフィルターを用いてろ過濃縮後、PowerSoil kit (Qiagen 社)でろ紙から核酸を抽出した。この核酸抽出液を鋳型として、2.1 で決定した PCR 条件を用いながら同様の方法でメタバーコーディング解析を行った。最後に、NanoCLUST によって同定された全細菌種の中から病原種のみを抽出した。具体的には、既報^{8,9)}や日本細菌学会による病原細菌リスト、BacDive データベース¹⁰⁾を参考に全 1409 種からなる病原種リストを作成し、同定結果を照合することで病原種を抽出した。

3. 結果と考察

3.1 標準試料に対する PCR 反応条件の決定

アニーリング温度 50, 55, 60 度、伸長時間 1, 2, 3 分のもとで PCR 増幅し、標準試料中の細菌種を属レベルで検出した結果を表 2 に、種レベルで検出した結果を表 3 にまとめる。例えばアニーリング温度 50 度、伸長時間 2 分の条件では、標準試料に含まれる 15 種のうち 14 種が属レベルまで正しく検出でき、偽陰性となった種は 1 種、誤検出（偽陽性）された種は 1 種であった(表 2)。また同じ PCR 条件での種レベルでの同定結果(表 3)を見れば、15 種のうち 13 種を種レベルまで正しく検出でき、2 種が偽陰性、6 種が偽陽性となった。適合率（検出された種のうち本当に標準試料に存在するものの割合）と再現率（標準試料に含まれる細菌種のうち実際に検出できたものの割合）、F 値（適合率と再現率の調和平均）に着目すれば、「50 度-2 分」の条件で属レベルの検出は最も良好であり、また種レベルにおいても F 値が 0.76 と最も良好（＝適合率と再現率のバランスをうまくとっている）であった。

表 2 属レベルでの同定精度と感度

| アニーリング 温度(°C) | 伸長時間 (min) | 検出された/されなかった属の数 | | | 適合率 (-) | 再現率 (-) | F値 (-) |
|------------------|---------------|-----------------|------|------|------------|------------|-----------|
| | | 真陽性数 | 偽陽性数 | 偽陰性数 | | | |
| 55 | 1 | 13 | 2 | 2 | 0.87 | 0.87 | 0.87 |
| 55 | 2 | 12 | 2 | 3 | 0.86 | 0.80 | 0.83 |
| 55 | 3 | 13 | 2 | 2 | 0.87 | 0.87 | 0.87 |
| 50 | 2 | 14 | 1 | 1 | 0.93 | 0.93 | 0.93 |
| 60 | 2 | 12 | 1 | 3 | 0.92 | 0.80 | 0.86 |

表 3 種レベルでの同定精度と感度

| アニーリング 温度(°C) | 伸長時間 (min) | 検出された/されなかった種の数 | | | 適合率 (-) | 再現率 (-) | F値 (-) |
|------------------|---------------|-----------------|------|------|------------|------------|-----------|
| | | 真陽性数 | 偽陽性数 | 偽陰性数 | | | |
| 55 | 1 | 12 | 5 | 3 | 0.71 | 0.80 | 0.75 |
| 55 | 2 | 11 | 5 | 4 | 0.69 | 0.73 | 0.71 |
| 55 | 3 | 13 | 6 | 2 | 0.68 | 0.87 | 0.76 |
| 50 | 2 | 13 | 6 | 2 | 0.68 | 0.87 | 0.76 |
| 60 | 2 | 11 | 4 | 4 | 0.73 | 0.73 | 0.73 |

続いて、図 2 に各条件での細菌種の相対存在割合（属レベルで表示）を示す。図中左側の「理論値」は標準試料のメーカー公表値であり、積み上げグラフがこれに近いほど真の群集構造を正しく推定できていることを意味する。Bray-Curtis 非類似度によって標準試料の組成との違いを評価すると、「50 度-2 分」の条件で 0.28 と最も小さく、他の条件よりも正しく組成を推定できていることが確認された。

以上より、検出精度と感度のみならず細菌組成の正確さという点でもアニーリング温度 50 度、伸長時間 2 分という条件が最も良好な検出結果であったため、この条件を環境試料に適用することとした。

次に、表 4 に「50 度-2 分」の条件で検出された細菌種と真の細菌種の対応関係を表 3 に示す。大半の種は正しく検出・同定できており、真の 16S rDNA 配列に対して 99%以上の類似性で配列を推定できているものの、*Bifidobacterium* 由来の 16S rDNA 配列は全く検出されなかった。また、大腸菌(*Escherichia coli*)や *Staphylococcus*、*Lactobacillus* 由来の 16S rDNA 配列は、検出はされたものの実際とは異なる分類群情報が割り当てられ、誤検出を招いていた。特に大腸菌は属内での 16S rRNA 遺伝子の相同性が高く、たとえ 16S rDNA 全長を解読しても種の違いを正しく判別できない可能性が考えられた。このように一部の細菌種に関しては本分析手法の検出精度・感度に限界があるものの、大半の細菌種については当初の目的通り「種レベル」までの検出・同定が可能であることが確認でき、なおかつ適切な PCR 条件を決定することができた。

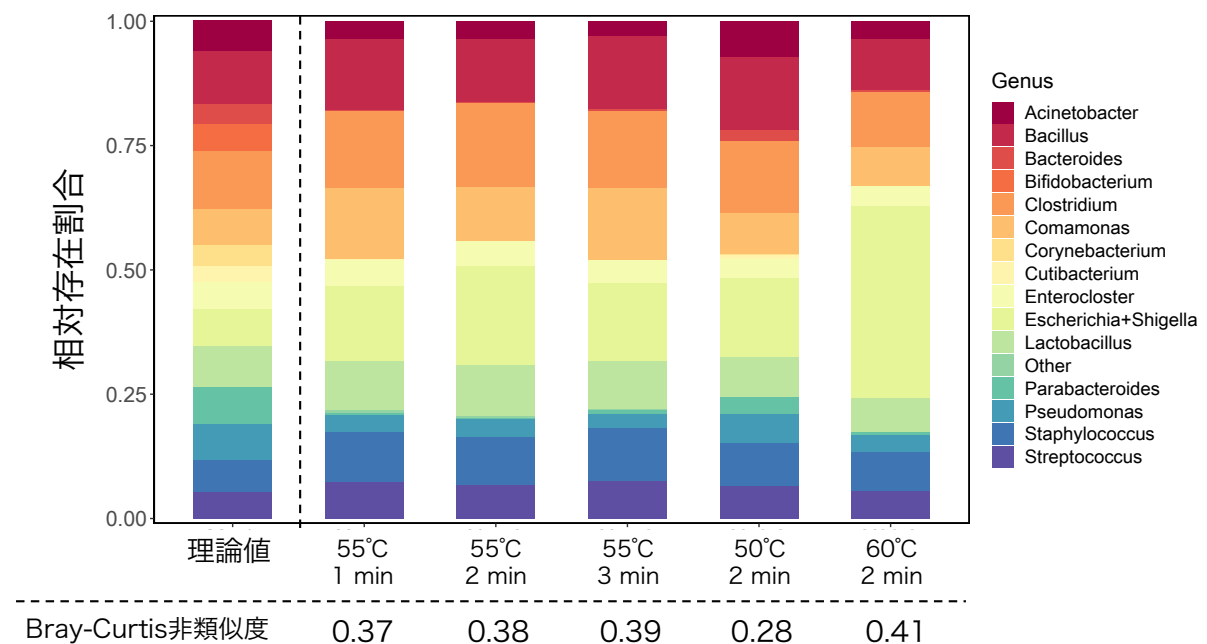


図 2 各 PCR 条件で得られた細菌組成と標準試料との非類似度

表 4 真の細菌種に対する同定結果と配列類似性
(PCR 条件: アニーリング温度 50 度、伸長時間 2 分)

| 真の細菌種 | 検出された細菌種 | 真の配列に対する類似性 (%) |
|---|--|-----------------|
| <i>Acinetobacter radioresistens</i> | <i>Acinetobacter radioresistens</i> | 99.8 |
| <i>Bacillus subtilis subsp. subtilis</i> | <i>Bacillus subtilis</i> | 100 |
| <i>Bacteroides uniformis</i> | <i>Bacteroides uniformis</i> | 99.8~99.9 |
| <i>Bifidobacterium pseudocatenulatum</i> | 不検出 | - |
| <i>Clostridium butyricum</i> | <i>Clostridium butyricum</i> | 100 |
| <i>Comamonas terrigena</i> | <i>Comamonas terrigena</i> | 100 |
| <i>Corynebacterium striatum</i> | <i>Corynebacterium striatum</i> | 99.5 |
| <i>Cutibacterium acnes subsp. acnes</i> | <i>Cutibacterium acnes</i> | 100 |
| <i>Enterocloster clostridioformis</i> | <i>Enterocloster clostridioformis</i> | 99.0 |
| <i>Escherichia coli</i> (K-12株) | <i>Escherichia fergusonii</i> <i>Escherichia marmotae</i> <i>Shigella boydii</i> <i>Shigella flexneri</i> | - |
| <i>Lactobacillus delbrueckii subsp. delbrueckii</i> | <i>Lactobacillus acidophilus</i> <i>Lactobacillus delbrueckii</i> | 99.9 |
| <i>Parabacteroides distasonis</i> | <i>Parabacteroides distasonis</i> | 97.9~99.4 |
| <i>Pseudomonas putida</i> | <i>Pseudomonas putida</i> | 99.9 |
| <i>Staphylococcus epidermidis</i> | <i>Staphylococcus caprae</i> <i>Staphylococcus epidermidis</i> | 99.9 |
| <i>Streptococcus mutans</i> | <i>Streptococcus mutans</i> | 100 |

3.2 琵琶湖・淀川水系への手法の適用

3.2.1 検出された病原種と同定結果の妥当性

図 3 に環境サンプルから抽出された病原種と、データベース上でヒットした参照配列との配列類似性をまとめる。下水処理場放流水を含む桂川～淀川流域 4 地点で採取された全 22 試料から、合計 47 属 78 種が病原種として抽出された。ただし、図 3 に示すように、これらの半分程度は参照配列との配列類似性が 99%を下回っており、抽出された病原種の多くが誤検出である可能性が考えられた。環境サンプルでは標準試料に比べてはるかに細菌の多様性が高いため、シーケンシングのエラー補正やクラスターの判別が困難となったことがその原因として考えられ、データ解析方法等の改善が望ましい。以上のように検出精度には課題が残るものの、16S rDNA 全長が病原種と「ある程度」類似した配列を多く検出できている、とも考えることができる。これより、水中の病原細菌を（多少の誤検出も許容しながら）スクリーニングする上では、本分析手法は依然として有用であると考えられる。

これらは桂川上流～下流や下水処理放流水において高頻度で検出されており、流域内の位置関係による明確な分布の違いは見られなかった。また、動物の糞便等を由来とする病原細菌としては、*Campylobacter*, *Clostridium*, *Enterobacter*, *Enterococcus*, *Klebsiella*, *Streptococcus* などが挙げられた。これらは下水処理放流水で比較的、高頻度かつ高い相対割合で検出されていた。さらに図 5 に検出された病原種のを地点別に示す。検出種数の中央値は、最上流の保津峡地点で 4 種、下水処理放流水で 14.5 種、宮前橋地点で 7.5 種、枚方大橋地点で 7 種となり、下水処理放流水で病原細菌の多様性が高いことがわかった。また、下水処理水の流入前後（保津峡と宮前橋/枚方大橋）で比較すれば、流入前の地点（宮前橋）で病原種数が小さかった。これより、桂川の流下過程で下水処理水が流入することによって病原細菌の汚染を受け、病原種の多様性が増している可能性が示された。

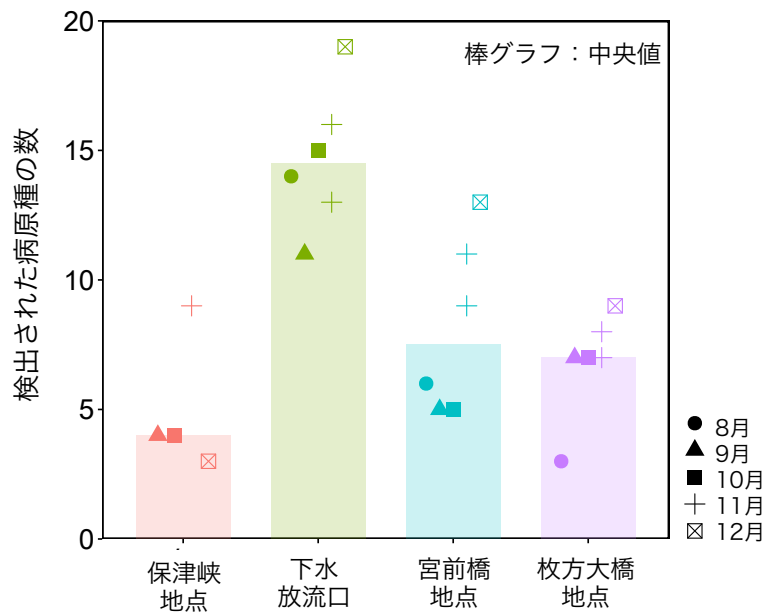


図 5 検出された病原細菌の種数に関する地点別比較

4. おわりに

本研究では琵琶湖・淀川水系における病原細菌を一斉検出可能とするために、まず種レベルまでの分類解像度を備えた細菌のメタバーコーディング手法を整備した。標準試料を用いた検討の結果、PCR におけるアニーリング温度 50 度、伸長時間 2 分の条件が種レベルでの検出精度と感度、細菌組成の正確さの観点から最適であった。さらにこの手法を桂川～淀川水系における河川水や下水処理放流水に適用した結果 47 属 78 種の病原種が検出され、下水処理放流水において病原細菌の多様性が特に高いこと、放流水の流入により河川水中の病原細菌の種数が増加していることが明らかとなった。

なお、環境水からの検出結果には偽陽性の可能性が高い病原種も見られたため、分析手法を今後改善することが望ましい。また、今回の河川水の採水回数には限界があり、病原細菌の詳細な分布特性（例えば地点ごとに特有の病原細菌種など）を明らかにするには至らなかったことは今後の課題である。

参考文献

- 1) 浅田安廣：淀川流域における感染性病原微生物の実態把握と遊泳等による感染リスク推定モデルの構築, 平成 26 年度水質保全研究助成 成果報告書.
- 2) 枝川 亜希子：淀川水系に生息するレジオネラ属菌と宿主となる自由生活性アメーバに関する研究, 平成 30 年度水質保全研究助成 成果報告書.
- 3) 東 剛志：臨床領域で対策が求められている薬剤耐性菌の河川環境中における動態, 平成 31 年度水質保全研究助成 成果報告書.
- 4) 浅田安廣：淀川流域における病原ウイルスの存在把握ならびに健康影響評価, 平成 28 年度水質保全研究助成 成果報告書.
- 5) 浅田安廣：淀川流域におけるアデノウイルス血清型の網羅的探索に基づいた水系感染による健康影響の推定, 平成 29 年度水質保全研究助成 成果報告書.
- 6) 大西義博・大瀧博文：大容量の淀川水系河川水からの病原性原虫類の検出, 平成 28 年度水質保全研究助成 成果報告書.
- 7) Rodríguez-Pérez, H., Ciuffreda, L. and Flores, C : NanoCLUST: a species-level analysis of 16S rRNA nanopore sequencing data, *Bioinformatics*, Vol. 37, No. 11, pp. 1600– 1601, 2021.
- 8) Fang, T., Cui, Q., Huang, Y., Dong, P., Wang, H., Liu, W.-T. and Ye, Q. : Distribution comparison and risk assessment of free-floating and particle-attached bacterial pathogens in urban recreational water: Implications for water quality management, *Science of the Total Environment*, Vol. 613, pp. 428–438, 2018.
- 9) Taylor, L. H., Latham, S. M. and Woolhouse, M. E. : Risk factors for human disease emergence, *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, Vol. 356, No. 1411, pp. 983–989, 2001.
- 10) Reimer, L. C., Vetcininova, A., Carbasse, J. S., Söhngen, C., Gleim, D., Ebeling, C., and Overmann, J.: Bac Dive in 2019: Bacterial phenotypic data for high-throughput biodiversity analysis, *Nucleic acids research*, 47(D1), D631-D636, 2019.