

厚生労働科学研究費補助金（健康安全・危機管理対策総合研究事業）

公衆浴場におけるレジオネラ症対策に資する検査・消毒方法等の
衛生管理手法の開発のための研究
令和元年度分担研究報告書

MLVA タイピングの確立とゲノム分子疫学との比較解析

研究代表者	前川純子	国立感染症研究所 細菌第一部
研究分担者	中西典子	神戸市環境保健研究所 感染症部
研究協力者	野本竜平	神戸市環境保健研究所 感染症部
研究協力者	水本嗣郎	静岡県環境衛生科学研究所
研究協力者	高橋直人	静岡市環境保健研究所

研究要旨：MLVA 法は、その特性として、安定性・迅速性・比較の容易性から、利便性の高い分子タイピング法となっている。そこで、本研究では、これまでの研究から見出された問題点を解決するために、全ゲノム配列を用いた系統解析を取り入れることで、SBT と MLVA のタイピングの妥当性評価を行い、より最適な MLVA 領域の検討を行うことで、遺伝子型別方法としての MLVA タイピングを確立し、汎用性を高めることを目的とする。

今年度は、Miseq および MinION を用いて、12 株の完全長配列を決定し、計算上推測されるフラグメントの大きさから外れる MLVA 領域について詳細に解析し特徴づけを行った。その結果、フラグメントが検出されない菌株の中には、リピート領域は存在しているが、primer のミスマッチにより増幅されない MLVA 領域の存在が明らかになった。一方で、Intermediate-size として扱っていた MLVA 領域は、リピート領域が始まる上流の長さが短くなっていた。さらに、最後のリピートの長さも短くなっていることが明らかとなった。

さらに、PFGE、MLVA、SBT で相違が見いだされた 1 つの集団事例に由来する株について、ゲノム系統解析の結果と各分子疫学解析手法の解析結果との比較検討を行った。その結果、PFGE、SBT、MLVA のタイピングでは、SNP 数 300 個程度の菌株間を clonal complex として認識されていると考えられた。また、コアゲノム SNPs 解析では、SBT、PFGE それぞれの手法の結果を反映した系統樹となっていた。また、Miseq リードデータやコンティグ配列から ST を決める方法を確立し、2927 個の ST のプロファイルのデータベースに更新した。

A. 研究目的

感染源の特定には、レジオネラ症患者からの分離株と感染源と推定される環境分離株の遺伝子型を比較し、遺伝子型の一致

を確認する必要がある。その際に用いられる方法として主流になっているのが、パルスフィールドゲル電気泳動法（PFGE）や世界的に普及している SBT（Sequence

based typing) 法である。SBT法は、7つの遺伝子 (*flaA*, *pliE*, *asd*, *mip*, *mompS*, *proA*, *neuA*) のシーケンスを行い、その塩基配列により型別を行う手法である。しかしながら、これら従来法は、多検体処理の煩雑さ、時間、予算を要することが課題となっていた。近年、細菌の遺伝子型別解析としてMLVA法がよく用いられている。MLVA法は、その特性として、安定性・迅速性・比較の容易性から、利便性の高い分子タイピング法となっている。そこで、他の細菌の遺伝子型別解析にも利用されているMLVA法を*L. pneumophila*において導入することで、それら従来法の課題を克服できることが期待される。これまでに、約800株の*L. pneumophila*を用いて、従来法であるPFGEやSBTと比較し、MLVAタイピングの有用性について評価を行ってきた。その結果、MLVA型は、SBTのタイピングと同等の菌株識別能力を有することが明らかになった。その一方で、遺伝子型別の手法間の相違点も見出された。また、他自治体間との比較の際にも、フラグメントの大きさがずれる点やMLVA領域によってリポート数換算の際に判断に迷う点等いくつかの課題が見出され、汎用性の高いタイピングとしてのMLVAを確立するためには、プロトコル整備の必要性が示唆された。

今年度は、MiseqおよびMinIONを用いて、12株の完全長配列を決定し、計算上推測されるフラグメントの大きさから外れるMLVA領域について詳細に解析し特徴づけを行った。さらに、PFGE、MLVA、SBT

で相違が見いだされた1つの集団事例に由来する菌株について、ゲノム分子疫学を用いて各分子疫学解析手法の解析結果との比較検討を行った。また、Miseqリードデータやコンティグ配列からSTを決める方法を確立し、データベースを更新した。

B. 研究方法

①菌株：完全長配列決定の菌株は、リポート数が0になる株やIntermediate-sizeの株12株を用いた（表1）。

集団事例Aに由来する解析で用いた菌株は、表2の通りである。

②MLVA：Sobralら¹⁾によって報告された12領域（Lpms01, Lpms03, Lpms13, Lpms19, Lpms31, Lpms33, Lpms34, Lpms35, Lpms38, Lpms39, Lpms40, Lpms44）を用いた。蛍光標識したプライマーを用いて、4領域を1セットとした3種類のmultiplex PCR-A（Lpms01, Lpms31, Lpms33, Lpms35）、PCR-B（Lpms03, Lpms13, Lpms19, Lpms34）、PCR-C（Lpms38, Lpms39, Lpms40, Lpms44）とした。PCR反応は、QIAGEN Multiplexを用いた。PCR条件は、95°C15分後に95°C30秒、60°C1分、72°C70秒を35サイクル行った。50倍希釈したPCR産物1μlをサイズマーカー0.25μl（GeneScan 1200 LIZ Size Standard（PCR-AとPCR-B）、GeneScan 600 LIZ Size Standard（PCR-C）とHi-Di Formamide（ABI）10μlに混合し、95°Cで3分加熱後、氷中条件で2分間急冷した。その後、AB3500 Genetic

Analyzerにてフラグメント解析を行った。得られたデータは GeneMapper Ver. 4 (Applied Biosystems) を用いて、フラグメントサイズおよびリピート数を測定した。得られた MLVA 型による株間の類縁関係を明らかにするために、BioNumerics Ver7.6 を用いて、Minimum spanning tree (MST) を作成した²⁾。

③ ゲノム解析: QIAseqFX (QIAGEN) を用いて DNA ライブラリを調製し、Illumina Miseq system を用いてリードデータを取得した。a5-Miseq でアセンブリし、PROKKA でアノテーションを行った。全ゲノム配列による系統解析は kSNP3 を用いて実施した^{3,4)}。

完全長ゲノム配列は、MinION (Nanopore 社) から得られたロングリードと Miseq のショートリードデータを Unicycler により Hybrid assembly して決定した。

C. 研究結果

(1) Miseq リードデータから ST を決める方法の確立およびデータベースの更新

L. pneumophila の ST は、EWGLI (The European Working Group for Legionella Infections) により運用されている website に 7 遺伝子の必要な部分の配列情報を照合することで決定するが、ドラフトゲノムのコンティグ配列から当該遺伝子配列部分のみを正確に抽出するのは効率的ではないため、リードデータをそのまま当該遺伝子配列にマッピングすることで直

接 ST が決定できるパイプラインの構築を目指した。

前年度では、公開されている配列情報から 7 遺伝子 (*flaA*, *pliE*, *asd*, *mip*, *mompS*, *proA*, *neuA*) の SBT データベースを作成した。そのデータベースに対して、Miseq で得られたリードデータを、SRST2 を用いてマッピングを行い、得られたアリアルプロファイルを既知の ST データベースと照合し、ST を決定するパイプラインを構築した。但し、*mompS* についてはゲノム上にタンデムに存在するオルソログ領域のリードがマッピングされ、アリアルが正しくアサインされないケースが稀に存在するという課題があった。

今年度は、*mompS2* の配列にのみマッピングできる pipeline を動かすことができるようにし、7 遺伝子すべてに対応可能となった。また、データベースに登録されている 2927 個の ST のプロファイルのデータベースに更新した。サンガー法で決定した配列でも利用可能となっており、一時的に SBT の web サイトが使用できない状態になっている場合の有用な解析ツールとなった。

(2) MLVA 領域の特徴づけ

計算上推測されるフラグメントの大きさから外れる MLVA 型を示した菌株の完全長ゲノム配列を取得し、MLVA 領域周辺構造を解析した。その結果、フラグメントが検出されない菌株の中には、リピート領域は存在しているが、primer のミスマッチにより増幅されない MLVA 領

域 (Lpms31、Lmps01、Lmps13、Lpms39) の存在が明らかになった。

一方で、Intermediate-size として扱っていた Lpms31 領域は、リピート領域が始まる上流の長さが Philadelphia 株と比較して短くなっていた。さらに、最後のリピートの長さも短くなっていた。Lpms03 については、リピート数 7.5 を示した菌株は非特異的な領域が増幅している可能性が示唆された。

(3) 集団事例におけるゲノム分子疫学を用いた PFGE、MLVA、SBT の遺伝子型別の比較

PFGE、MLVA、SBT の相違点が見いだされた集団事例について、ゲノム系統解析を用いて各分子疫学解析手法の解析結果との比較検討を行った。

PFGE において、患者由来株と F20-1 (⑥レーン) と F25 (⑦レーン) 株以外の浴槽水由来株および拭き取り由来株の PFGE パターンは完全に一致した (図 1 A)。拭き取り由来株である F20-1 と F25 は他の菌株と 3 バンド違いであったが、UPGMA による近似性は 86.9% となった。

PFGE を実施した SG1 の 13 菌株の ST は、ST59 とデータベースに存在しない新しい ST の二種類に分かれたが、それらは *neuA* の 1 遺伝子違いであった。ST59 を示した 2 株は同じ浴槽の浴槽水と拭き取り由来であった。MLVA 型は、ST59 の 1 株が 1 領域違いであったが、他の 12 株の MLVA 型は一致した。

環境から分離された SG1 ではない F9-2、

F20-2 の ST、MLVA 型は他の菌株とは異なっていた。

これら 15 株のコアゲノム SNPs の系統解析を行った (図 1 B)。F9-2、F20-2 は、他の菌株と比較して SNP 数が 10000 個以上となり、他の菌株とは系統的に離れた位置にカテゴライズされた。患者由来株は系統的に近い位置に集約された。患者由来株と SNP 数 20 個以内の株が拭き取り由来 F9-1 と F10 であった。患者由来株と同一 ST だが、PFGE で 3 バンド違いだった F20-1 と F25 株は、患者由来株との SNP 数約 300 個であり、PFGE の結果を反映していた。また、ST59 の 2 株も患者由来株との SNP 数は約 300 個であり、SBT の結果を反映した系統樹となった。

D. 考察

MLVA 周辺構造の解析から、フラグメントが増幅されない菌株の中には、リピート領域は存在しているが、primer のミスマッチにより MLVA 領域が増幅されていないことが明らかとなり、primer の改良が必要であると考えられる。一方で、MLVA 領域の中には、Intermediate-size として扱っていた MLVA 領域があったが、リピート領域が始まる前の長さが短くなっており、また、最後のリピートの長さも短くなっていた。このような株は、計算通りのリピート数をとる菌株とは区別する必要があるが、計算通りのリピート数をとる菌株の中にリピート領域が始まる前の長さが異なるような株が存在するののかについては、菌株数を増やし検討

する必要がある。

集団事例における全ゲノム系統解析から、患者由来株と近縁な 2 株 (SNP 数が 20 個以内) は同一の浴槽の拭き取りから分離されており、その浴槽が直接の感染源となっていた可能性が考えられた。コアゲノム SNPs 解析では、SBT、PFGE それぞれの手法の結果を反映した系統樹となっていた。また、この結果から、PFGE、SBT、MLVA のタイピングでは SNP 数 300 個程度の菌株間を clonal complex として認識されていると考えられた。

さらに、Miseq リードデータやコンティグ配列から ST を決める方法を確立し、2927 個の ST のプロファイルのデータベースに更新した。今後も定期的に ST プロファイルを更新する予定である。一時的に SBT の web サイトが使用できない状態になっているが、サンガー法で決定した配列でも利用可能となっており、迅速で有用な解析ツールとなった。

E. 結論

MLVA 領域の中には、primer の再考が必要な領域、Intermediate-size として扱う必要がある MLVA 領域が明らかとなった。

さらに、集団事例におけるコアゲノム SNPs 解析から詳細な菌株間の関連性が明らかとなった。PFGE、SBT、MLVA のタイピングでは、相同性の高い株として認識されるが、コアゲノム SNPs 解析ではそれぞれの手法の結果を反映した系統樹となっていた。集団発生事例等個々の菌株の同一性を詳細に確定させたい場合は、可

能であればゲノム解析を実施することが望ましいが、コスト等の面で難しい場合は PFGE、MLVA、SBT を併用して分子疫学解析を実施することが必要である。

また、Miseq リードデータやコンティグ配列から ST を決める方法を確立し、2927 個の ST のプロファイルのデータベースに更新した。

謝辞

今回解析した分離株を分与くださった菊地孝司・小堀すみえ (さいたま市健康科学研究センター)、井上浩章 (アクアス筑波総合研究所)、渡辺祐子 (神奈川県衛生研究所) (敬称略) の諸氏に感謝いたします。

F. 参考文献

- 1) Sobral D, Le Cann P, Gerard A, Jarraud S, Lebeau B, Loisy-Hamon F, Vergnaud G, Pourcel C. 2011. High-throughput typing method to identify a non-outbreak-involved *Legionella pneumophila* strain colonizing the entire water supply system in the town of Rennes, France. *Appl Environ Microbiol.* 77:6899-6907.
- 2) 中西典子ら, MLVA 法における *Legionella pneumophila* の遺伝学的特徴, 厚生労働科学研究費補助金 (健康安全・危機管理対策総合研究事業) 「公衆浴場等施設の衛生管理におけるレジオネラ症対策に関する研究」平成 28~30 年度総合研究報告書, 研究代表者: 前川 純子, 37-46,

2019

- 3) Raphael BH, Baker DJ, Nazarian E, Lapierre P, Bopp D, Kozak-Muiznieks NA, Morrison SS, Lucas CE, Mercante JW, Musser KA, Winchell JM., 2016. Genomic Resolution of Outbreak-Associated *Legionella pneumophila* Serogroup 1 Isolates from New York State. *Appl Environ Microbiol.* 82:3582-3590.
- 4) Nakanishi N, Nomoto R, Tanaka S, Arikawa K, Iwamoto T., 2019. Analysis of Genetic Characterization and Clonality of *Legionella pneumophila* Isolated from Cooling Towers in Japan. *Int J Environ Res Public Health.* 16. doi: 10.3390/ijerph16091664.

H. 知的財産権の出願・登録状況
なし

G. 研究発表
なし

表1: 完全長配列決定の菌株

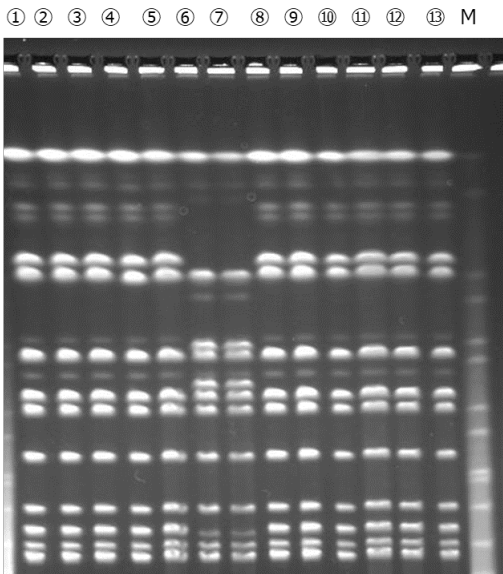
strain ID	SG	ST	Lpms31	Lpms01	Lpms35	Lpms33	Lpms34	Lpms13	Lpms19	Lpms03	Lpms40	Lpms38	Lpms39	Lpms44
KL06-287	1	1	9.5	7	17	4	2	10	4	7	5	3	14	7
KL10-682	1	1008	9.5	7	17	4	2	10	4	7	5	3	14	7
KL05-209	13	2603	6.5	7	3	2	1	0	4	7.5	0	0	7	7
KL10-685	UT	2603	6.5	7	3	2	1	0	4	7.5	0	0	7	7
KL09-603	1	154	0	0	8	2	0	3	0	7.5	0	3	11	0
KL12-819	7	1422	9.5	7	17	4	2	10	4	7	5	0	14	7
KL09-578	13	2603	6.5	7	3	2	1	0	4	7.5	0	0	0	7
KL09-548	1	598	0	8	8	2	0	0	4	0	0	3	10	0
KL10-679	1	1334	0	0	11	1	0	3	0	7.5	0	0	12	0
NIIB3710	1	507	13.5	7	25	4	3	5	4	7	5	3	20	9
NIIB3758	1	507	16.5	7	25	4	3	5	4	7	5	19	20	9
NIIB3150	1	22	15.5	8	26	5	3	11	4	7	5	19	14	9

表2：集団事例Aにおける菌株のSTとMLVA型

菌株番号	分離場所	SG	PFGE	ST	allele番号										MLVA型									
					<i>flaA</i>	<i>pile</i>	<i>asd</i>	<i>mip</i>	<i>mompS</i>	<i>proA</i>	<i>neuA</i>	Lpm31	Lpm01	Lpm35	Lpm33	Lpm34	Lpm13	Lpm19	Lpm03	Lpm40	Lpm38	Lpm39	Lpm44	
B69	男2浴槽水	SG1	①	59	7	6	17	3	13	11	11	15	8	13	4	1	10	5	8	4	3	10	9	
F9-1	男4-1ふさ取り	SG1	②	NT	7	6	17	3	13	11	6	15	8	13	4	1	10	5	8	4	3	10	9	
F9-2	男4-1ふさ取り	SG5	-	1417	8	6	34	9	2	8	209	15	8	8	3	1	8	0	7	4	0	10	0	
F10	男4-2ふさ取り	SG1	③	NT	7	6	17	3	13	11	6	15	8	13	4	1	10	5	8	4	3	10	9	
F14	女4ふさ取り	SG1	④	NT	7	6	17	3	13	11	6	15	8	13	4	1	10	5	8	4	3	10	9	
F19	男2ふさ取り	SG1	⑤	59	7	6	17	3	13	11	11	15	8	13	4	1	7	5	8	4	3	10	9	
F20-1	男1ふさ取り	SG1	⑥	NT	7	6	17	3	13	11	6	15	8	13	4	1	10	5	8	4	3	10	9	
F20-2	男1ふさ取り	SG8	-	-	2	3	9	10	2	1	-	17.5	7	18	4	3	10	4	7	5	0	22	7	
F25	男8ふさ取り	SG1	⑦	NT	7	6	17	3	13	11	6	15	8	13	4	1	10	5	8	4	3	10	9	
SK1	患者2	SG1	⑧	NT	7	6	17	3	13	11	6	15	8	13	4	1	10	5	8	4	3	10	9	
SK2	患者3	SG1	⑨	NT	7	6	17	3	13	11	6	15	8	13	4	1	10	5	8	4	3	10	9	
S1ぬ	患者1	SG1	⑩	NT	7	6	17	3	13	11	6	15	8	13	4	1	10	5	8	4	3	10	9	
S5ぬ	患者4	SG1	⑪	NT	7	6	17	3	13	11	6	15	8	13	4	1	10	5	8	4	3	10	9	
S6ぬ	患者5	SG1	⑫	NT	7	6	17	3	13	11	6	15	8	13	4	1	10	5	8	4	3	10	9	
S9ぬ	患者7	SG1	⑬	NT	7	6	17	3	13	11	6	15	8	13	4	1	10	5	8	4	3	10	9	

*NT：データベースに存在しない新しいST

(A)



(B)

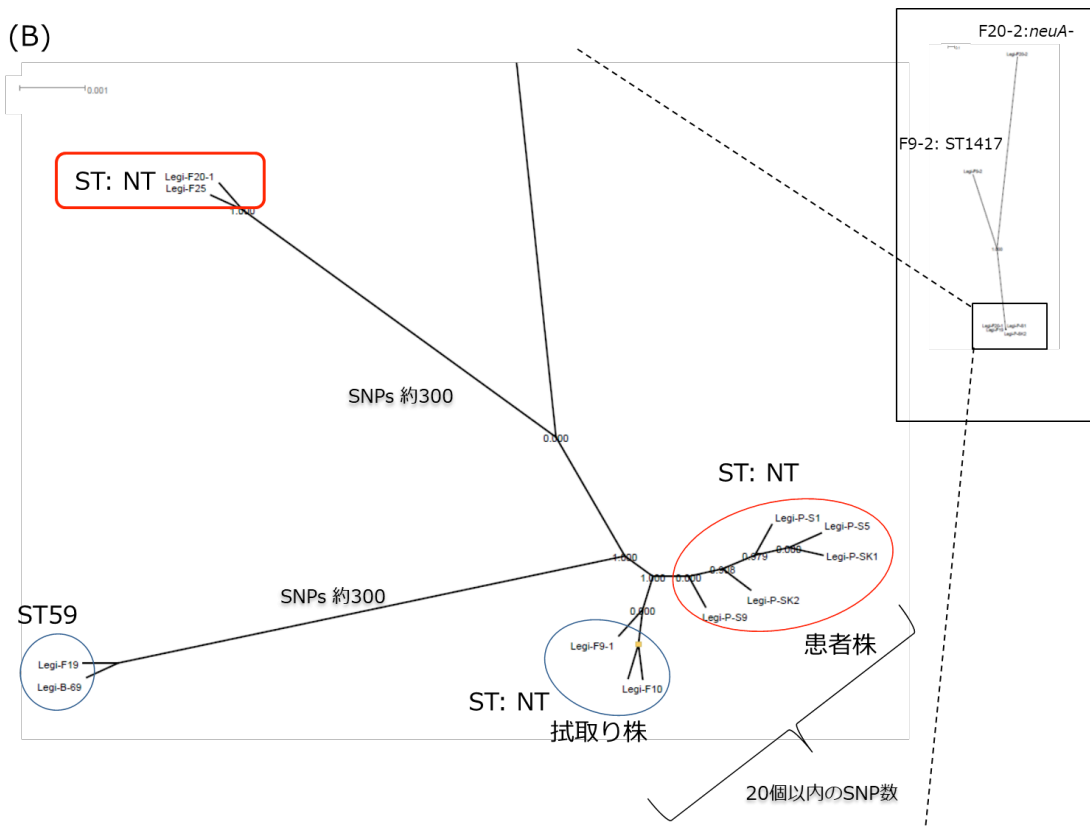


図 1 : (A)集団事例AのPFGE。レーン毎の菌株No.は表2の通り。
(B)コアゲノムSNPsに基づいた系統樹