

分 担 研 究 報 告 書

病原大腸菌の検出指標遺伝子および病原性発現解析

大岡 唯祐

令和4年度 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）

食中毒原因細菌の検査法の整備のための研究

研究代表者 工藤由起子 国立医薬品食品衛生研究所

分担研究報告書

病原大腸菌の検出指標遺伝子および病原性発現解析

研究分担者 大岡唯祐 鹿児島大学 大学院医歯学総合研究科

研究要旨

astA 遺伝子（腸管凝集付着性大腸菌耐熱性毒素をコード）保有大腸菌を中心とした種々の病原性大腸菌について“効果的な検出法の確立”および“病原因子の組み合わせによる重症化リスクの解明”を目的として研究を進めた。R4年度は、昨年度にドラフトゲノム配列を取得した *astA* 遺伝子陽性 31 株について、*astA* 遺伝子のゲノム局在を明らかにするため、完全長ゲノム配列の取得を試みた。また、NCBI 等公共データベース登録株から同定された完全長ゲノム配列決定大腸菌株（約 9,000 株）の中から同定した 713 株の *astA* 遺伝子陽性株について、本解析で同定した 31 株と併せて解析し、35 種類の *astA* 遺伝子バリエントを同定した。これら 35 種類のバリエント（V1-V35）について詳細な配列解析を実施した結果、保有株数の多い主要な *astA* 遺伝子バリエントは V22 (218 株)、prototype (145 株)、V6 (68 株)、V27 (46 株)、V12 (36 株)、V7 (12 株) であった。また、それらとは別に *astA* 遺伝子の断片を持つ株が 349 株存在した。各バリエントの intactness についての解析から、V7, V15, V22, V33 は premature stop codon などにより機能しない可能性が示唆された。これらの結果、機能する可能性があり、多数の株に保有されるバリエントは prototype、V6、V27 の 3 種類と考えられた。これとは別に、*astA* 遺伝子がコードされている挿入配列 IS1414 の intactness に関する解析から、prototype, V2, V7, V31, V32 がコードされた IS1414 は構造的に intact であることが判明し、これらのバリエントが今後、ゲノム上でのコピー数増加に関わる可能性が示唆された。これらのことから、今後、注意すべき *astA* 遺伝子バリエントは prototype、V6、V27 と考えられた。今後はこれらのバリエントについて、機能解析を実施し、検査現場における重要性を確認する予定である。

A. 研究目的

近年、病原大腸菌を原因とする食中毒が多発しており、令和2年には、学校給食を喫食した小中学生の児童生徒等 2,529 人の患者をとまなう *astA* 遺伝子保有大腸菌による大規模食中毒が発生した。*astA* 遺伝子保有大腸菌による食中毒は毎年発生が続いており、患者が 100 人を超える事例も多く、食中毒予防対策が必要とされている。また、腸管凝集付着性大腸菌 [EAEC] (凝集付着性因子: *aggR* 遺伝子等の保有株) や腸管病原性大腸菌 [EPEC] (細胞への局在付着性因子: *eae* 遺伝子等の保有株) による食中毒の発生も続いている。これらの病原大腸菌による食中毒においては、原因食品が不明であることが多く、感染源 (食材や食品) が明らかになることはまれである。これらの病原大腸菌の食品等の検査法は国内外で確立されておらず、一般的な大腸菌の検査法を用いて実施されることが多く、効率的な検査法が実施されていないことが危惧される。食中毒細菌の食品汚染菌数レベルは低いことがいわれており、その食中毒細菌に適した検査法が原因食品究明には重要である。

本研究では、ゲノム情報を基に各病原型に保存性の高い病原関連遺伝子を同定する。その際、既知の病原因子だけでなく、新規因子の同

定も試みる。病原因子の発現解析および機能解析をおこない、検出指標に適した病原関連遺伝子を既知・新規因子を含め検討する。以上のような研究を進めることにより、遺伝子検出法を主とした病原性大腸菌の効果的検査法の開発を目指す。

本年度は、R3 年度ドラフトゲノム配列を取得した 31 株および公開データベース上に登録された *astA* 遺伝子保有完全長ゲノム配列決定株 713 株について、*astA* 遺伝子の特徴など大規模ゲノム比較解析を行い、検出系構築のための疫学マーカーとなりうる遺伝子選定と病原機構を解析することを目的とした。

B. 研究方法

(1) 鹿児島大保有 *astA* 遺伝子保有 31 株の完全長ゲノム配列解析

昨年度ドラフトゲノム配列を取得した 31 株について、ドラフトゲノム配列では *astA* 遺伝子の局在など詳細な解析を行えないことが判明したため、MinION (Nanopore) を用いた完全長ゲノム配列の取得を試みた。具体的には、NucleoBond HMW DNA (Takara Bio) を用いて長鎖ゲノム DNA を精製し、短鎖 DNA を除去したのち、MinION で解析した。得られたロングリード配列を MiSeq リード配列とともに unicycler-v0.4.8 でハ

株 (28%) 存在し、1 株で最高 11 コピー保有する株が見られたが、他のバリエーションでは、最高でも 2 コピー保有する株が 1 株あるいは 2 株存在するのみであった (表 2)。

各バリエーションが存在する挿入配列 IS1414 の構造: prototype, V2, V7, V31, V32 の 5 種類のバリエーションをコードする IS1414 は intact な構造であった。それ以外の 30 種類は IS の転移酵素をコードする遺伝子が欠損しているため機能しない、あるいは IS の 5' 末端や 3' 末端が欠失していた (図 2)。

(3) *astA* 遺伝子保有株の進化系統解析と *astA* 遺伝子バリエーションの分布

Roary による解析からコア遺伝子が 1,600 個同定され、アライメントの結果、informative SNP は 42,536 個同定された。その結果を基に作成した系統樹に *astA* 遺伝子バリエーションの分布をマップした。その結果、バリエーションの多くは大腸菌の進化系統 (A, B1, B2, C, D, E) に偏りなく散在していることがわかった (図 3)。また、5' 末端領域が完全に欠失している *astA* 遺伝子は進化系統 A と E に多くみられた (図 4)。

(4) 異なる系統に存在する同一 *astA* 遺伝子バリエーション周辺のゲノム構造比較

異なる系統に同一の *astA* 遺伝子バリエーションが検出され、当該バリエーションがプラスミド上にコードされている場合、プラスミドを介した水平伝播が想定されるため、その可能性を確認する目的で該当株のプラスミド構造を比較した。一例として、バリエーション V24 は異なる系統に属する 7 株 (図 5) が保有し、全てプラスミド上にコードされていることが判明している。そのため、これら 7 株において *astA* 遺伝子をコードしているプラスミドの構造比較を行った (図 6)。その結果、①のグループに属する 5 株のプラスミドは非常に類似性が高いことが判明したが、①②③のグループ間では部分的な類似性しか検出されなかった。

(5) 主要 *astA* 遺伝子バリエーションの機能解析

35 種類の *astA* 遺伝子バリエーションがコードするアミノ酸配列を比較した結果、prototype を含む 7 種類は同じ配列であったが、それ以外は配列多様性を認めた (図 7)。このことから、集団感染事例由来株で保有頻度が高い、また、全株の中でも比較的保有頻度の高い prototype, V6, V27 に焦点を当て、その病原性を検討することとした。本年度は、高発現させ

た状態で細胞障害性や細胞侵入性など病原性の有無を調べるため、これらバリエントを T7 誘導性発現プラスミドである pETBA ベクターにクローニングし、そのタンパク質発現を確認した。

(倫理面への配慮)

該当しない。

D. 考察

本年度実施した 744 株の *astA* 遺伝子保有大腸菌株に関する 35 種類の遺伝子バリエントの詳細な解析から、①保有株数の多い主要なバリエントは V22 (218 株)、prototype (145 株)、V6 (68 株)、V27 (46 株)、V12 (36 株)、V7 (12 株) であること、②バリエント V7, V15, V22, V33 は欠損しており、機能しないと考えられること、③各バリエントの局在は prototype を除き、プラスミドあるいは染色体のいずれかであること、④ prototype を除いて、1 株あたりにマルチコピー存在するバリエントはほとんどないこと、⑤5' 末端が完全に欠損している *astA* 遺伝子を保有する株を含めるとほとんどの大腸菌系統に *astA* 遺伝子保有株が存在することが明らかになった。これらのことから、*astA* 遺伝子を保有する株を特異的に検出する疫学マーカーとな

る遺伝子を同定するのは困難であること、また、単に全ての *astA* 遺伝子バリエントを同定する現在用いられている PCR 等の検査法では、機能しないバリエントも検出してしまうため、実際に病原性に寄与しているバリエントのみを識別出来る検査法が必要であると考えられる。加えて、⑥ prototype, V2, V7, V31, V32 の 5 種類のバリエントをコードする挿入配列 IS1414 は intact な構造であるがそれ以外は欠損しているため転移能を持たないと考えられることが判明した。このことから、今後、転移等により菌株間で伝播していく可能性が高いバリエントは限定されることも明らかとなった。

今後は、構造的に intact なバリエントの中で高い保有率を示す prototype など主要なバリエントを中心として、病原性への関与メカニズムについて機能解析を進め、どのバリエントが臨床的に重要であるかを明らかにし、それらのバリエントを特異的に検出できる検出系の構築に繋げる必要がある。

E. 結論

astA 遺伝子陽性大腸菌に関して、全ゲノム情報を基に、保存性が高く

検出指標に適した病原関連遺伝子の同定とそれを利用した効果的検査法の開発を目指して解析を進めた。今年度は、昨年度に同定した 35 種類の *astA* 遺伝子バリエーションについて、その分布やコピー数、ゲノム局在、転移能などの特性解明を中心に解析を実施し、その結果、重要なバリエーションとそうでないものを識別する必要があることを明らかにした。また、*astA* 遺伝子が大腸菌進化系統のほぼ全てに分布していることが明らかとなったことから、*astA* 遺伝子保有株を検出する疫学マーカーとなる遺伝子の同定は困難であることが示唆された。本年度、重要と思われるバリエーションについては、遺伝子クローニングを行うまでしか進まなかったため、次年度はこれら *astA* 遺伝子の高発現株を用いて、その病原機構への関与について機能解析を行い、その研究成果を学会および国際雑誌等で報告することを目指す。

F. 健康危険情報

国民に至急知らせた方がよい情報に該当するものはない。

G. 研究発表

(誌上発表)

なし

(学会等発表)

大岡唯祐、後藤恭宏、林哲也、西順一郎. 大腸菌における *astA* 遺伝子バリエーションの同定およびその分布. 第 96 回日本細菌学会総会. 令和 5 年 3 月 17 日. 兵庫

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

図1 *astA* 遺伝子バリエントの intactness

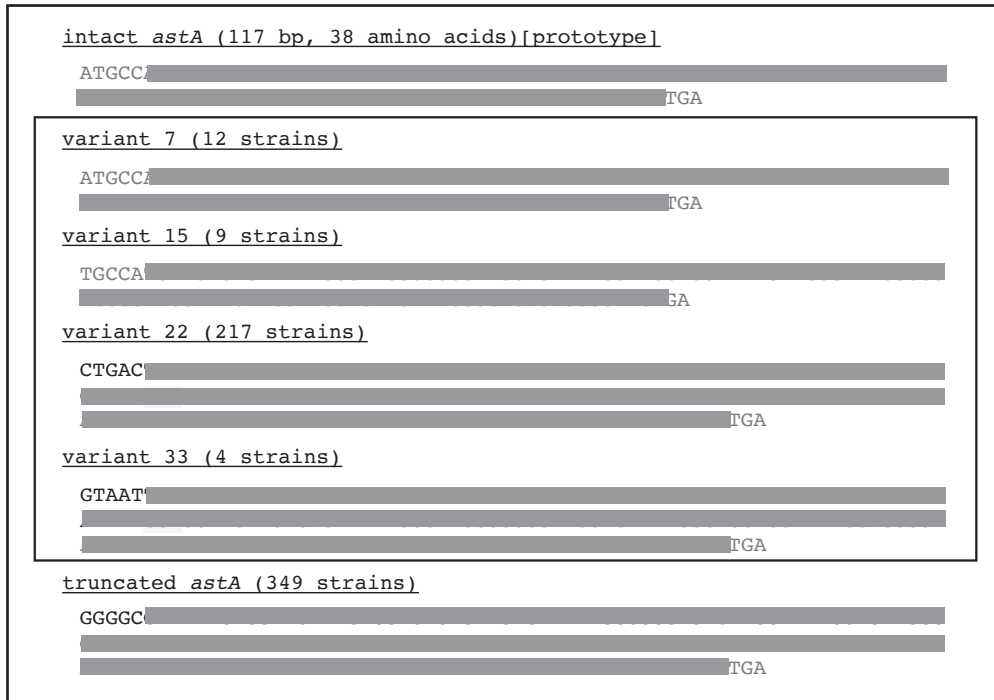


図2 *astA* 遺伝子バリエントをコードする挿入配列 IS1414 の構造比較

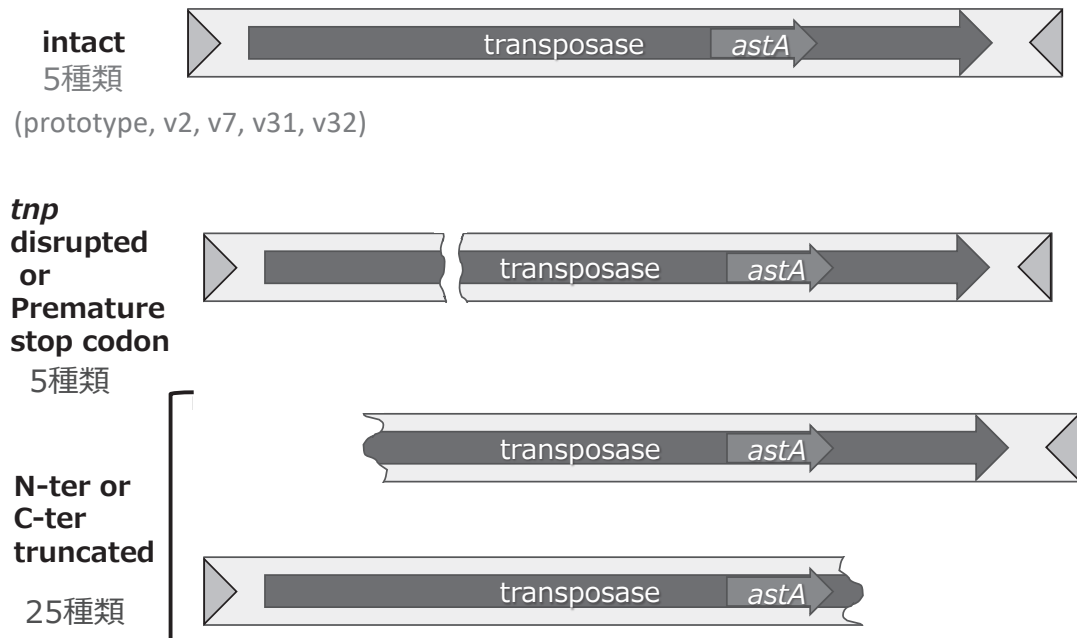


図3 大腸菌進化系統における *astA* 遺伝子バリエントの分布

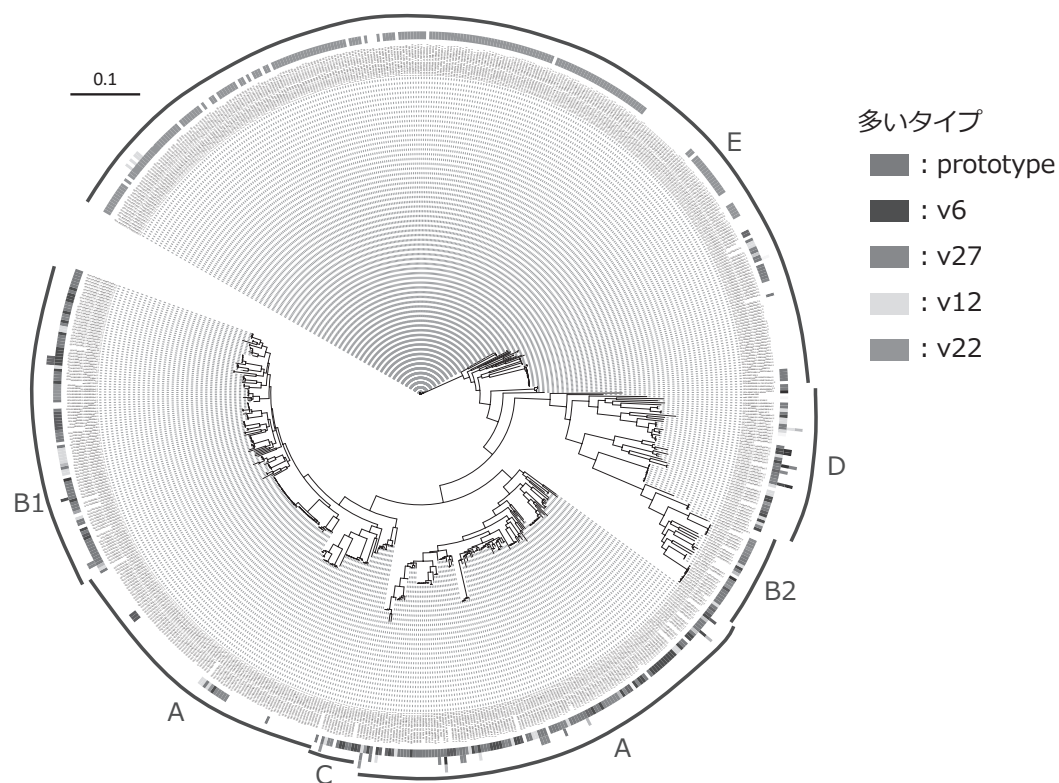


図4 大腸菌進化系統における 5'末端が完全に欠失している *astA* 遺伝子の分布

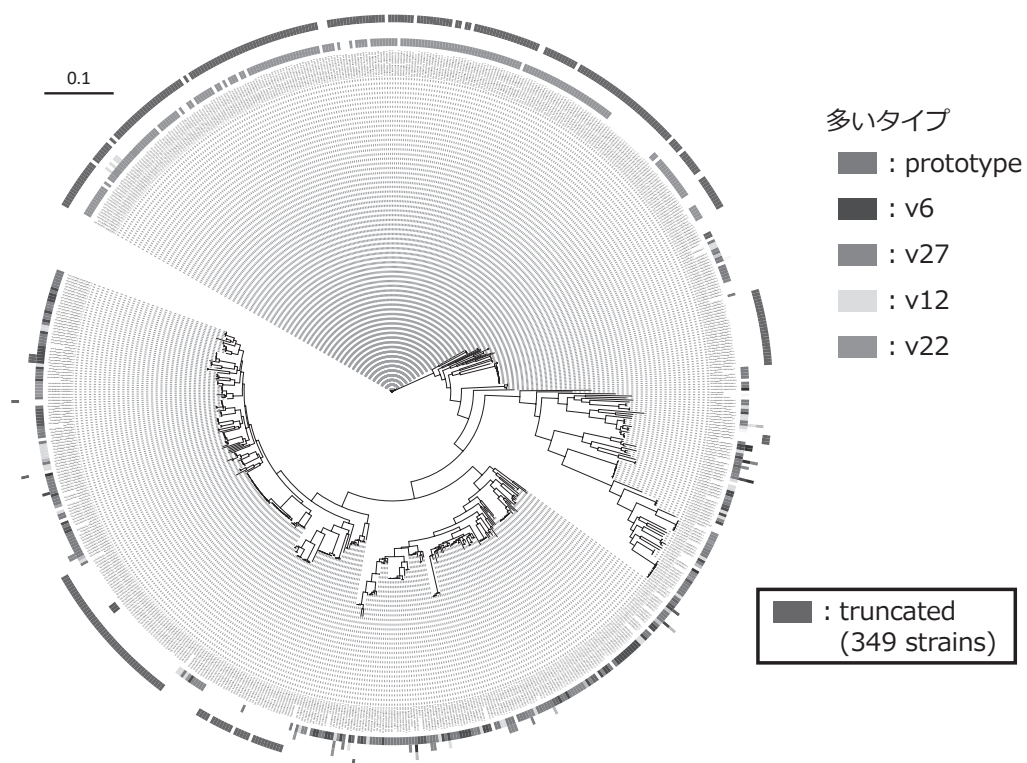


図5 大腸菌進化系統におけるバリエント V24 (7株) の分布

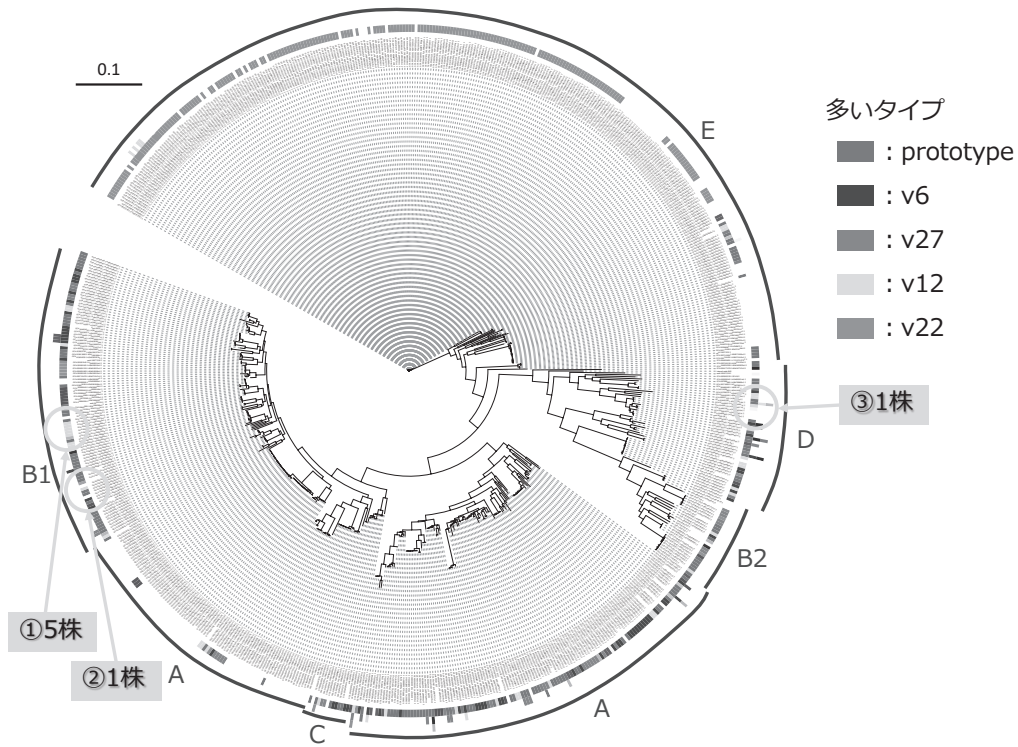


図6 バリエント V24 保有株 (7株) の astA 遺伝子をコードするプラスミドの比較

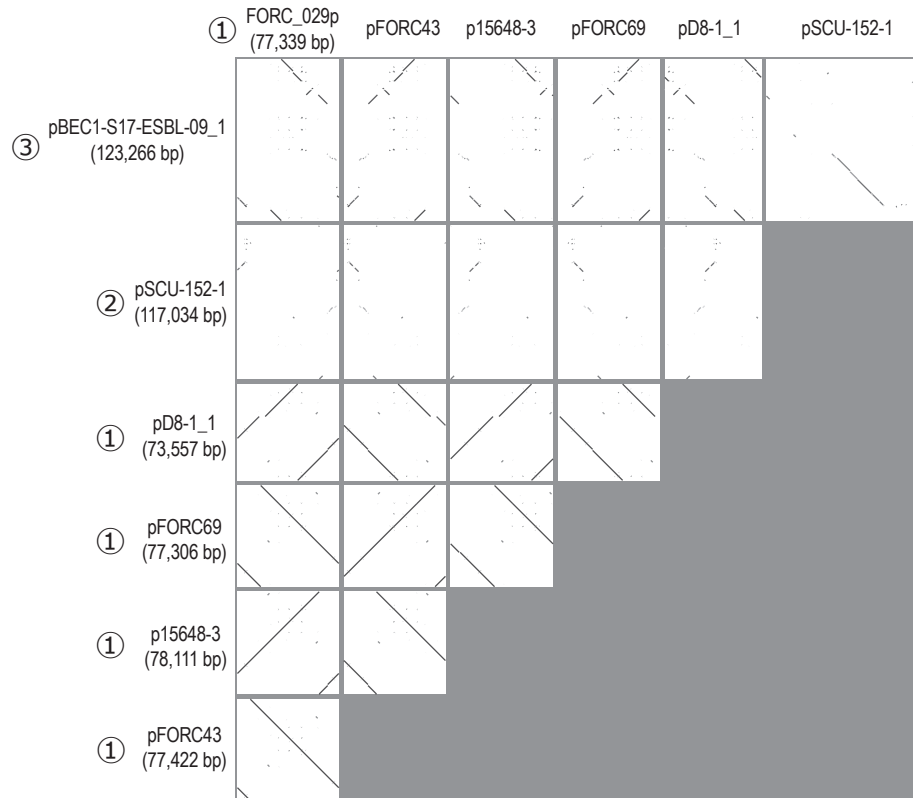


図 7 35 種類の *astA* 遺伝子バリエントのアミノ酸配列比較

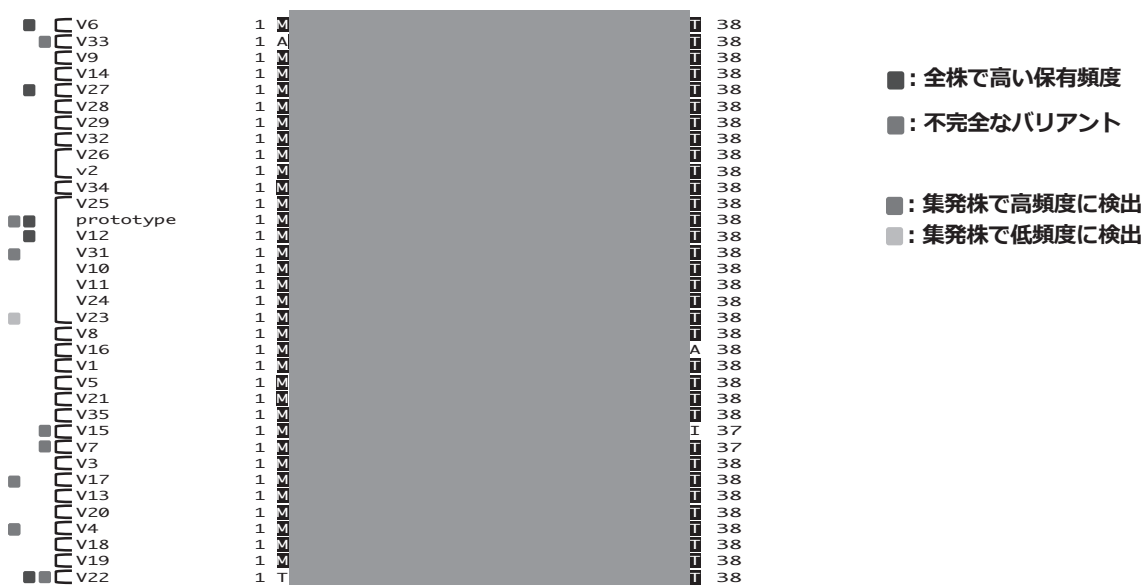


表 1 鹿兒島大保有 *astA* 遺伝子陽性 31 株のゲノム解析

strain name	No. of pair reads	sequence coverage	No. of scaffolds (>300 bp)	total length (bp) (>300 bp)	MinION summary	segments(plasmid)	chromosome	plasmid
K9228	711,916	64	464	5,512,493	incomplete	114(2)	(5,411,025)	P1(7,110), P2(1,552), P3(1,542)
K9291	748,626	72	238	5,036,673	incomplete	40(3)	(4,857,237)	P1(82,496), P2(57,930), P3(6,991)
K9331	726,842	63	69	4,838,693	incomplete	5(1)	(4,753,093)	90,597
K9339	677,580	56	102	5,131,569	complete	3(2)	(4,969,156)	P1(157,892), P2(4,237)
K9510	863,364	78	202	4,722,680	complete	4(3)	4,633,118	P1(73,912), P2(12,323), P3(4,286), P4(1,506)
K9531	653,133	56	189	5,229,048	ND			
K9550	836,644	72	182	5,219,273	complete	4(3)	5,073,586	P1(84,605), P2(80,948), P3(1,563)
K9559	821,188	70	102	5,266,830	complete	6(5)	5,007,961	P1(162,983), P2(81,415), P3(3,257), P4(2,873), P5(1,622)
K9867	726,685	60	75	5,254,186	complete	2(1)	5,140,777	P1(118,404)
K9910	650,356	53	122	5,378,142	incomplete	36(1,1[34])	5,098,104	P1(114,332), P2(145,344)
K10091	430,407	38	198	5,120,336	ND			
K10101	961,440	90	144	4,959,481	complete	4(3)	4,704,007	P1(136,087), P2(113,511), P3(4,072)
K10309	800,943	62	139	5,179,214	complete	3(2)	4,973,135	P1(194,881), P2(8,598)
K10443	861,337	68	282	5,486,044	incomplete	33 (?)	(5,041,144)	8 plasmids
K10593	796,607	62	176	5,412,244	complete	4(3)	5,281,511	P1(122,613), P2(8,4739), P3(1,551)
K10737	763,669	66	67	5,163,106	incomplete	20	(5,189,509)	-
K10876	935,021	68	299	6,059,351	complete	11(10)	5,651,975	P1(123,757), P2(92,955), P3(82,373), P4(74,020), P5(8,910), P6(6,773), P6(5,588), P7(2,312), P8(2,096), P9(1,552)
K10877	786,642	61	156	5,234,515	incomplete	14(3)	(5,070,516)	P1(143,174), P2(6,880), P3(4,051)
K10987	800,648	65	107	5,075,336	complete	4(3)	4,935,575	P1(134,342), P2(7,175), P3(2,750)
K10989	693,342	62	64	5,008,825	complete	4(3)	4,911,639	P1(61,597), P2(31,764), P3(3,373)
K11244	702,871	61	106	5,115,318	complete	3(2)	4,856,975	P1(204,139), P2(73,876)
K11407	611,343	51	130	5,369,318	incomplete	17(2)	(5,085,938)	P1(169,559), P2(90,876)
K11421	414,786	39	77	4,929,648	complete	3(2)	4,812,164	P1(108,823), P2(64,402)
K11469	1,034,279	96	116	4,936,622	complete	2(1)	4,803,443	P1(145,610)
K11627	758,454	64	192	5,064,133	complete	2(1)	4,872,138	P2(145,092)
K11797	732,321	61	62	5,160,203	incomplete	17	(5,167,532)	-
K11974	619,202	53	85	5,142,029	complete	4(3)	4,944,009	P1(151,495), P2(47,686), P3(2,257)
K12120	741,686	65	584	5,378,463	incomplete	58(?)	4,801,605	12 plasmids
K12196	802,784	67	99	5,081,272	complete	4(3)	4,884,269	P1(113,228), P2(84,629), P3(5,687)
K12240	788,968	66	108	5,087,851	complete	4(3)	4,765,535	P2(155,152), P2(117,468), P3(48,425)
K12343	841,346	73	26	4,790,472	incomplete	13(1)	(4,802,616)	P1(3,597)

表2 744株における *astA* 遺伝子バリエントの保有株数・局在・コピー数

	number of strains	localization			not identified	number of strains with multicopies	maximum copies
		chromosome	plasmid	(both)			
● prototype	145	76	76	11	4	41	11
1	0	0	0	0	0	0	0
● 2	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0
4	1	0	1	0	0	0	1
5	0	0	0	0	0	0	0
6	68	67	1	1	1	1	3
● 7	12	12	0	0	0	0	1
8	1	1	0	0	0	0	1
9	5	0	5	0	0	0	1
10	1	1	0	0	0	0	1
11	2	0	2	0	0	0	1
12	36	36	0	0	0	0	1
13	2	2	0	0	0	0	1
14	9	9	0	0	0	0	1
● 15	9	0	9	0	0	1	2
16	2	1	1	0	0	0	1
17	2	0	2	0	0	0	1
18	3	0	3	0	0	2	2
19	2	0	2	0	0	0	1
20	1	0	1	0	0	0	1
21	4	1	3	0	0	1	2
● 22	218	216	2	0	0	1	2
23	2	0	2	0	0	2	2
24	7	0	7	0	0	0	1
25	4	0	4	0	0	0	1
26	1	0	1	0	0	0	1
27	46	45	1	0	0	0	1
28	2	2	0	0	0	0	1
29	9	9	0	0	0	0	1
30	0	0	0	0	0	0	0
● 31	1	1	0	0	0	0	0
● 32	1	1	0	0	0	0	0
● 33	5	5	0	0	0	0	1
34	1	0	0	0	1	0	0
35	1	0	1	0	0	0	0

● : 不完全な*astA*遺伝子バリエント □ : 保有株が多い*astA*遺伝子バリエント ● : IS1414がintact (転移可)