

表5 C.coli 株の塩基置換数

塩基置換数	Cc-Gunma-24	Cc-Gunma-21	Cc-Gunma-22	Cc-Gunma-23
Cc-Gunma-20	1335	1412	1412	1412
Cc-Gunma-24		87	87	87
Cc-Gunma-21			0	0
Cc-Gunma-22				0

地方衛生研究所の連携による食品由来病原微生物の網羅的ゲノム解析を基盤とする
新たな食品の安全確保対策に関する研究
－食品分離株及び臨床分離株のゲノム解析－

研究分担者 佐多徹太郎（富山県衛生研究所）
研究協力者 綿引 正則（富山県衛生研究所）
研究協力者 磯部 順子（富山県衛生研究所）
研究協力者 木全 恵子（富山県衛生研究所）
研究協力者 清水美和子（富山県衛生研究所）
研究協力者 範本 志保（富山県衛生研究所）
研究協力者 三井千恵子（富山県衛生研究所）
研究協力者 金谷 潤一（富山県衛生研究所）

研究要旨 大量の塩基配列データが得られる次世代シーケンサー（NGS）の公衆衛生分野、特に地方衛生研究所（地衛研）での活用方法について検討した。食中毒原因菌として重要なサルモネラ属菌分離株 36 株、カンピロバクター分離株 20 株についてゲノム解析を行い、比較ゲノム解析手法を用いて解析した。さらに大量の細菌ゲノムの解析を行うことができるかどうか検証するため、*Salmonella Infantis* 77 株のゲノム配列を取得し、遺伝的背景と保有遺伝子の状況の相関を調査した。しかし、遺伝的背景と保有遺伝子の関係は明らかにならなかった。また、食中毒事例で分離された *Campylobacter lari* のゲノム解析を実施し、病原性に関連する遺伝子の探索を行い、53 遺伝子が検出された。このように NGS の有用性は明らかにしてきたが、得られた配列はドラフト配列である。このドラフト配列を定量的に理解するために、既に完全配列が明らかになっている株を解析から、ドラフト配列の定量的な考察を試みた。その結果、コーディング領域の約 85% が一致した。結論として、NGS データは、高精度疫学マーカーとしての SNP の活用、比較ゲノム学的手法を利用した分離株の性状解析について有用であることが分かった。また、これらの結果は、地衛研での NGS 技術の普及を考慮して、Windows 環境の利用を念頭に入れて検討した結果であり、高い専門性が必要とされていたゲノム解析手法が、地衛研でも十分活用できるということを示しており、さらに公衆衛生学分野での活用に応用できると思われる。

A. 研究目的

近年の食品流通システムの変化により、先進国では、病原微生物に汚染された流通食品が原因となる大規模かつ広域化した集団発生がしばしば報告されている。例えば腸管出血性大腸菌（EHEC）の食中毒の発生は重大な健康危機をもたらしている。その中で、地方衛生研究所（地衛研）は、法律にもとづき、集団発生時に食品あるいは臨床検体から病原体検索を行い、分離された病原体の解析から、原因究明に資する科学的根拠を提供する業務を行っている。しかし、原因となる病原体が検出できず、原因不明となる事例も少なくない。一方で、一度に大量の塩基配列データを簡単に得ることが可能となる次世代シーケンサー（NGS）が利用できるようになり、これまで原因が不明であった事例でも原因究明のための技術として非常に有用であることが知られるようになってきた。

しかしながら、NGS の利用は、これまで一部の研究機関の利用に限られていた。その理由は、当初、装置は非常に高額であり、装置を稼働させるための専用試薬も高額であることによる。さらに膨大な DNA 配列の情報処理を行う技術も、地衛研ではほとんど経験もなく、NGS を用いた解析はほとんど考えられていなかった。

2011 年に NGS のポータブル型機器が発売され、また、この NGS で出力される膨大な塩基配列データを処理する Windows ベースのソフトウェアが利用可能となりつつあった。このような時期に、富山県を中心として、多数の重症患者（うち死亡 5 名）を含む食中毒患者 181 名の EHEC による集団食中毒事例が発生し、検査機能の充実を図るという理由で、地衛研として初めて富山県衛生研究所に NGS の一つである、MiSeq が設置された。

近年では、人や様々な食品、物が地球規模で動

いており、これに合わせて感染症や食中毒もグローバル化し、複雑化している。従って、地衛研の業務も世界的な視野に立って、遂行することが求められるようになり、新規な病原体や複雑化する感染症、食中毒が発生したとき、迅速な原因究明に NGS の活用が期待される。

従って、本研究では、MiSeq を活用して、臨床及び食品から分離した食中毒菌のゲノム解析を行い、そのデータを利用して、病原性に関わる遺伝子の探索と保有状況、薬剤耐性遺伝子の探索と保有状況、及びその表現形質との比較、さらにゲノム上の SNP を用いた詳細な系統解析の可能性を評価し、地衛研での MiSeq の活用、NGS 技術の有効性を評価することを目的とする。

B. 研究方法

B-1. 供試菌株

本研究で使用した供試菌株は以下の通りである。

1. *Salmonella Typhimurium/Enteritidis*

1995 年から 2013 年にかけて、分離された血清型 Typhimurium および Enteritidis をおのおの 18 株、合計 36 株。各血清型では、食品及び臨床検体（主に便）から分離された株として、9 株ずつ選定した。

2. *Campylobacter jejuni/coli*

シプロフロキサシン耐性を中心に、*C.jejuni* 15

3. *Salmonella Infantis*

富山県で 2000 年～2013 年にかけて分離された臨床分離株、37 株、と市販鶏肉から分離された株、40 株である。

4. *Campylobacter lari*

平成 20 年に発生した集団食中毒事例で、共通食材としての鯨肉と複数の患者便から分離された *C.lari* 株。本事例から分離された *C.lari* 株はいずれも PFGE パターンは一致していた。本研究では、鯨肉から分離された *C.lari* 分離株をゲノム解析対象株 (Toyama07) とした。

5. 腸管出血性大腸菌 O157:H7 株

既に完全なゲノム配列が報告されていた堺株 (O157 堺株、大阪大学微生物研究所から入手) を用いた。

B-2. 供試菌のゲノム DNA の抽出と解析

供試菌は、トリプチケースソイ寒天培地 (TSA)、あるいは血液寒天培地 (*Campylobacter*) 上で 1 = 2 日培養した新鮮コロニーから、Gentra Puregene Yeast/Bact. Kit (キアゲン) を用いて、添付プロトコールに従って、ゲノム DNA を調製した。調製した DNA は、Qubit® 2.0 Fluorometer (ライフテクノ

ロジー社) を用いて、DNA 濃度を測定した。

B-3. DNA サンプル調製

MiSeq で解析するサンプルは、Nextera XT DNA サンプル調製キット (イルミナ) を用いて、調製した (これを「DNA ライブラリー」と称する)。合成した DNA の一部を、2%Nusieve3:1 アガロース (Takara Bio) で電気泳動して、合成反応がうまく進行しているかどうか確認した。また、DNA 分画する場合は、1%アガロース電気泳動を行い、電気泳動後、臭化エチジウムで染色し、LED ランプ照射下、アガロースを切り出し、Montage DNA Gel Extraction Kit (ミロポア) で、溶出後、液を濃縮した。塩基配列 (リード) の取得は、その一部を MiSeq Reagent Kit を用いて、MiSeq を用いた。

B-4. MiSeq データの情報処理

MiSeq からの出力データは、fastq 形式とし、CLC Genomics Workbench (CLC) を用いて、アセンブルを行ない、アセンブルしたデータは、fasta 形式で出力した。Single Nucleotide Polymorphism (SNP) は、CLC、Microsoft Excel を用いて抽出し、系統解析は、MEGA を使用した (Tamura, K., et al. Mol.Biol.Evol., 2011, 28:2731)。

アセンブルしたデータは、RAST (<http://rast.nmpdr.org/>) を使用して、注釈づけを行った。また、比較ゲノム学手法である Blast atlas は、Gview Server (<https://server.gview.ca/>) を利用した。薬剤耐性遺伝子の検索は、ResFinder (<https://cge.cbs.dtu.dk/services/ResFinder/>) を用いた。病原性因子の探索は、Virulence Searcher (http://www.hpa-bioinfotools.org.uk/pese/virfactfind_small.html) を利用した。

また、比較ゲノム解析には、が完全に終了している株をレファレンス株として、*Salmonella Typhimurium* LT2 (AE006468), *S. Enteritidis* (NC_011294) 及び *Campylobacter jejuni* NTCC 11168 (GIB00027CH01) を用いた。

blast 検索は、公的データベースを利用する場合は、DDBJ の Web サービスを、In-house データベースを利用する場合は、NCBI blast+ を Windows PC にインストールして実施した。

S. Infantis 77 株のゲノム配列から得られた CDS の比較には、RAST により得た CDS 情報をデータベースソフトである、File Maker Pro (ファイルメーカー) を用いて解析した。

B-5. 分離株の薬剤感受性試験

サルモネラ属菌の薬剤感受性試験には薬剤として、クロラムフェニコール(CP)、テトラサイクリン(TC)、ストレプトマイシン(SM)、カナマイシン(KM)、ゲンタマイシン(GM)、ナリジクス酸(NA)、セファロチン(CET)、アンピシリン(ABPC)、ミノサイクリン(MINO)を使用した。また、カンピロバクター属菌には、ナリジクス酸(NA)、オフロキサシン(OFLX)、ノルフロキサシン(NFLX)、シプロフロキサシン(CPFX)、エンロフロキサシン(ERFX)、エリスロマイシン(EM)、テトラサイクリン(TC)、ホスホマイシン(FOM)、ゲンタマイシン(GM)、ストレプトマイシン(SM)、アンピシリン(ABPC)、アンピシリン・スルバクタム(ABPC/SBT)、セファロチン(CF)、ピペラシリン(PIPC)、ピペラシリン・タゾバクタム(PIPC/TAZ)、セフォタキシム(CTX)を使用した。

(倫理面への配慮)

本研究は、臨床情報のない分離菌株を用いるものであり、解析対象にはヒトゲノム配列情報等は含まれていないため、倫理上の問題は発生しない。

C. 研究結果

C-1. *S. Typhimurium* および *S. Enteritidis* のゲノム解析

C-1-1. ゲノムワイド SNP の抽出と分子疫学マーカーとしての有用性

S. Enteritidis および *S. Typhimurium*、それぞれ 18 株の配列をアセンブルし、SNP を抽出し、18 株の SNP 情報を統合して、最終的に、タンパク質をコードする領域のみ、18 種類の株に相同部位の SNP を並べ、*S. Enteritidis* 18 株では、301bp、*S. Typhimurium* では、2,438 塩基の SNP の配列で系統樹を作成した (図 1～3)。

C-1-2. 比較ゲノム学的手法による MiSeq データの利用：Blast Atlas 解析

CLC のアセンブルにより得られた各株のコンテイングは、すでに報告されている同種のゲノム解析株 (レファレンス株) と得られたコード領域 (CDS) 比較することにより、株の特徴を CDS の有無や相同性の違いを知ることで推定が可能になる。そこで、Gview Server の Blast Atlas を利用して、*S. Enteritidis* および *S. Typhimurium* 分離株、それぞれ 18 株の得られた配列上の CDS をレファレンス株に存在するすべての CDS に対して、Blast 解析を実施し、全体の CDS の有無や相同性の違いを解析し

た。

図 4 は *S. Enteritidis* 18 株、図 5 は *S. Typhimurium* 18 株の、Blast Atlas 解析の結果を示す。図の内側にレファレンス株を配置し、外側に向かい、食品由来 9 株、続いて臨床由来 9 株の順で、相同 CDS を配置した。図 4 の *S. Enteritidis* 18 株の CDS は、レファレンス株と比較して、その相同性は高く、ほぼ均一で、臨床分離株と食品分離株で特徴的な構造の違いは見られなかった。しかし、*S. Typhimurium* 18 株は、一部に株間で多型を示す座位 3 か所 (A, B 及び C) が存在していた。この多型を示す座位の機能は、座位 A は、サルモネラが嫌氣的な生育時に炭素源としてアラントイン窒化に関する遺伝子群の有無に関していた。また、B, C については、いずれもプロフェージ遺伝子の差に起因していた。

C-2. *Campylobacter jejuni/coli* 分離株のゲノム解析

C-2-1. ResFinder を用いた薬剤耐性遺伝子の保有状況

20 株のドラフトゲノム配列を、薬剤耐性遺伝子の検索 Web サービスである ResFinder を用いて、検索した。結果については、表 1 の右側の列に記載した。その結果、20 株のうち 8 株から tet(O) が検出されたが、感受性試験で TC 耐性であった 7 株のうち、2 株 (strains Cj07 および Cj19) では tet(10) 陰性であった。また、逆に tet(10) 陽性で感性を示す株 (Cj20) が存在しており、一部で遺伝子型と表現型が合致していなかった。アミノグルコシド耐性株であった Cc01 は、耐性遺伝子として、aac(6')-aph(2'') が検出され、遺伝子型と表現型が合致していた。さらに、今回、検討した 20 株のうち、Cj07 を除く、すべての株で blaOXA-61 が検出された。この遺伝子の保持と β ラクタム耐性の表現型は必ずしも一致していなかった。

今回、解析した *Campylobacter* 属菌は、キノロン耐性菌を主に選択している。キノロン耐性機構は、DNA gyrase や topoisomerase の点変異に起因するアミノ酸置換により耐性を獲得することが知られており (図 9)、20 株のうち、複数の株でキノロン耐性決定領域 (QRDR) のアミノ酸置換を検出したが、一部の耐性菌では、検出されなかった。このような遺伝子変異による耐性機構では、薬剤耐性遺伝子のデータベースでは検索できない (表 1)。キノロン耐性や、多剤耐性が付与する場合、膜タンパクである薬剤排出ポンプの変異が関係することが知られている。そこで、*Campylobacter* 属菌の

薬剤耐性に関与する *cmeRABC* オペロンについて、20 株のゲノム配列から抽出し、レファレンス株と Blast 解析を実施した (図 10)。ここでは、Blast 検索する際のカットオフ値として、70%と 98%を用いて、検索した。その結果、*C.coli* については、レファレンス株として、*C.jejuni* を用いているため、相同性は低かったが、*C.jejuni* とは相同性は高く、キノロン耐性との関係については、明らかな傾向は示せなかった。

C-2-2. 比較ゲノム学的手法による MiSeq データの利用 : Blast Atlas 解析

先のサルモネラゲノム解析で用いた、Gview Server の Blast Atlas を、*Campylobacter jejuni/coli* 分離株に適用した (図 8)。Reference 株には、*Campylobacter jejuni* NTCC 11168 (GIB00027CH01) を用いた。

この結果、株特異的な多型が CDS 上にいくつか存在することが明らかになった。今回は、特に reference 株との比較から、3 つの座位 (A, B 及び C) に注目した。座位 A は、Capsular polysaccharide (CPS)-associated regions、座位 B は、Lipooligosaccharide (LOS)-associated regions、及び座位 C は I 型制限修飾酵素システムに該当していた。

C-3. Salmonella Infantis のゲノム解析

C-1-1. S. Infantis 77 株の MiSeq による解析データの品質評価

S.Infantis 77 株を MiSeq で解析した。取得した配列は、CLC でアセンブルを実施した。得られた塩基配列の品質を評価するためにその概要を表 1 に示した。今回 MiSeq から得られた配列は、カバー率 (全リード塩基数/コンティグ長の塩基数)、最大 169、最小 33、平均 100 であった。アセンブル後のコンティグ数は、最大 567、最小 64、平均 138 となり、得られた情報としては、若干、バラツキがあるものの、これ以降の解析に耐えうる値であると判断した。

C-3-2. S. Infantis 77 株の SNP データによる系統解析

S. Infantis 77 株の MiSeq による解析リードを、完全ゲノム配列が明らかになっていた *S. Infantis* CM001274 のゲノム配列にマッピングし、SNP を抽出し、77 株の SNP 情報を統合して、最終的に、タンパク質をコードする領域のみとして、104,586 塩基の SNP を得た。これらの配列から系統樹を作

成した (図 1、2)。この結果、臨床分離株 17 株からなる大きなクラスター (C5) が検出された。また、臨床分離株 13 株と食品分離株 36 株の混在した大きなクラスターを形成した (IC-1)。

C-3-3. 比較ゲノム学的手法による MiSeq データの利用の試み—クラスター特異的遺伝子の探索

図 2 の系統樹で研修された、クラスター C5 は、食品分離株の混在がなく特徴的なグループであった。そこで今回は、複数株のゲノム解析の試みとして、臨床分離株に特異的な遺伝子 (CDS) が存在するかどうかを検討した。

方法としては、解析するクラスター内の各 *S.Infantis* ゲノム配列について、アセンブル後、RAST により自動アノテーションを行った。そこで、得られたすべての Genbank ファイル (拡張子 =gbk) を用いて、クラスター毎に GViewServer の Pangenome analysis を実施し、そのクラスター全体で検出された CDS を統合した Pangenome 配列を得た。今回の解析では、クラスター C5 (17 株)、クラスター IC-1 (49 株)、及び IC-1 に含まれる食品由来株 (36 株) の Pangenome 配列を得た。これら 3 つの Pangenome 配列の gbk ファイルを作成し、GViewServer の Signature Analysis を選択し、クラスター C5 と IC-1 にともに存在し、IC-1 の食品由来株に存在しない、臨床分離株に特異的な CDS の検出を試みた。その結果、56 の候補遺伝子を得た。この 56 遺伝子の内訳は、機能が付与された 22 遺伝子と機能未知遺伝子 (hypothetical protein)、34 遺伝子であった (表 2)。

この 56 遺伝子が、臨床分離株特異的な遺伝子かどうかを検証した。方法は、今回、配列解析した *S.Infantis* 77 株のアセンブル後の配列を、RAST で自動アノテーションを行い、得られた CDS の塩基配列の fasta ファイルのデータをすべて統合した、*S.Infantis* 77 株の CDS データベースを作成した。次に、この CDS データベースに対して、先の 56 遺伝子をクエリー配列として、blast 検索を実施した。その結果を図 3 に示した。この図は、系統樹に示された株について、56 遺伝子の有無を示したマトリックスとして示した。このマトリックスの結果からは、臨床分離株特異的な遺伝子の特定には至らなかった。

C-4. Campylobacter lari 分離株のゲノム解析

C-4-1. Campylobacter lari 分離株の MiSeq データの取得

本食中毒事例は、平成 20 年に富山県で発生したもので、2 つの飲食店で提供された料理に含まれていた鯨肉が共通食であり、患者数 61 名を数えた事例であり、鯨肉残品と患者便から分離された *C.lari* は、PFGE で一致した。そこで、今回、鯨肉から分離された *C.lari* 株 (Toyama 07) のゲノム解析を実施した。

MiSeq からの解析リードのアセンブルした後の統計数字は、表 3 に示した。今回の解析では、コンティグ数 38、カバー率約 400、ゲノムサイズ 1,617,445 塩基のドラフト配列を得た。

C-4-2. *C.lari* Toyama07 の他のカンピロバクター属とのゲノム配列の比較

カンピロバクター属のゲノム解析は既に複数の株で実施されており、入手可能であった。*C.lari* については、RM2100 株、LMG11760 株の配列が既に報告されている。また、*C.jejuni*、*C.coli* 及び *C.fetus* の配列が報告されており、各 2 株ずつ選び、今回、解析した *C.lari* Toyama07 株の配列と比較した。

ゲノム配列の比較は、*C.lari* RM2100 株の配列を参照配列として、GViewServer の Blast atlas (図 4) 及び Pangenome Analysis (Pangenomic blast atlas) を行った (図 5)。

Blast atlas の結果から、*C.lari* Toyama07 株は、報告されている *C.lari* のゲノム配列と類似していた。また、参照株と異なるゲノム配列上の GAP が検出されたが、この GAP の遺伝子は、細胞表面の LPS の多型及びフェージ遺伝子の有無に関係していた。また、Pangenome Analysis の結果から、*C.jejuni* と *C.coli* は、ゲノム配列の 2/3 はほぼ共通なゲノム配列であるが、*C.lari* と *C.fetus* のゲノム配列は互いに相同性が低いことがわかる。また、*C.lari* Toyama07 株に特異的な配列があることが判明した。

C-4-3. 病原性遺伝子の検索

病原性遺伝子の検索は、Virulence Searcher を用いた。Web から、得られた *C.lari* Toyama07 のドラフト配列をアップロードし、解析してもらったところ、53 遺伝子 (図 7) が病原性遺伝子として検出された。これらの遺伝子は、もっとも解析の進んでいる *C.jejuni* から報告されている、定着因子である、CadF、JipA、MOMP、PEBA、毒素遺伝子 CdtABC、侵入 CiaB について、今回の *C.lari* Toyama07 株では、CadF、MOMP、CdtABC および CiaB がゲノムデータから検出された (表 4)。

C-3. 既知ゲノム解析株の MiSeq から得られたゲノム配列と報告配列との比較

O157 堺株のゲノム配列を MiSeq で解析した。MiSeq の 2 回のランにより得られた解析リードを個別にアセンブルしたコンティグ SK01、SK02、及びこの 2 つのランデータを統合したコンティグ SK12 を得た。これらのアセンブル後の統計データを表 5 に示した。この結果、SK01、SK02、および SK12 のシーケンスのカバー率は、それぞれ 43.5、50.9、及び 94.1 であり、MiSeq より得られた解析リードの品質を評価することが出来るこれらの統計データから、ドラフト配列の定量的な解析に耐えうる品質であると判断した。そこで、これらのコンティグを用いて、蛋白コーディング領域 (CDS) を抽出し、既知配列と比較した。

O157 堺株は、既に報告されている (Hayashi, T., et al. *DNA Res.*, 8:11, 2001)。報告では、O157 堺株は、2 つのプラスミドを保有しているため、今回解析した配列中には、プラスミド由来の配列が含まれているが、SK01、SK02、および SK12 は、それぞれコンティグ数は、203、181、および 182 であり (表 6)、そのうち、スキヤフォールド領域は (ペアエンドで解析していることから、配列が不明のまま、長さのみ情報としたリードが存在しているコンティグのこと)、それぞれ 17、19、および 20 か所検出された (データ未提示)。先の論文では、大きさ 150 塩基対以上の CDS の数は、5,447 個として、報告されているが、今回、MiSeq で得られた解析リードのアノテーションと比較できるように、報告されている O157 堺株の完全ゲノム配列を改めて、RAST によるアノテーション情報を得たところ、プラスミドの CDS と合わせると、総計で、5,310 個の CDS が検出された (これを以後、「Sakai」と呼ぶ)。この値は、Hayashi らの報告数より若干少ない。この差は、RAST の CDS の抽出法の問題もあるが、アノテーションについては、最終的には、ひとつひとつの CDS を詳細に検討することが必要となるが、今回、ほぼ、同じ数字であること、完全な配列は得ていないことから、この後は、RAST で得られた 5,310 の CDS をそのまま用いることとした。そして、MiSeq により得られた O157 堺株の配列である SK01、SK02 および SK12 も同じく RAST を用いて、アノテーションを行い、比較した。

得られた CDS 情報は、File Maker 上で比較した。実際には、Sakai と SK01、SK02 および SK12 それぞれ比較し、共通に存在する CDS、あるいはどちらか一方にのみ存在する CDS を数え、その結果を図 7 に示した。

D. 考察

NGSを用いて、食中毒対策に資する活用法を探る試みは、①NGSであるMiSeqから、目的にあった配列リードを安定に生産する、②出力データから、必要な情報を抽出する、2段階が必要である。そこで先ず、①を確立するために、Nextera XT DNAライブラリーの作成法を検討し、次に、その配列リードを、目的のあった方法で解析できるかどうかを検討した。イルミナ社のMiSeqが普及し、それを加工、解析するソフトが充実してきたことにより、簡単にできるようになってきているが、MiSeqから出力されるデータはあまりにも大きく、また、ドラフト配列であるため、どこまで利用活用できるか検討が必要であった。

*S. Enteritidis*_18株及び*S. Typhimurium*18株のゲノム解析データから、以下のことが確認できた。①最終的に抽出されたSNP数は、株により異なっている。今回は、*S. Enteritidis*_18株は301bps、*S. Typhimurium*18株は2,438bpであり、その差は8倍程度である(図2、3)。この違いの理由は、不明であるが、菌種の生態学的な違い(宿主の違いなど)を反映している可能性があり、Blast atlas解析により、ゲノム全体を俯瞰しても、前者は後者よりも保存性が高いように見える(図4と5)。また、今回の解析では、サルモネラ属菌が、人獣共通に存在する細菌であり、動物や環境から分離された菌種と臨床分離株とはなにか特徴的な違いがゲノム配列(特にコード領域の違い)に存在しているかどうか、期待したが、今回解析した*S. Enteritidis*と*S. Typhimurium*からは、その違いは見いだせなかった。しかしながら、*S. Typhimurium*については、染色体ゲノム上に3つの大きな多型が存在していた(図5)。3つのうち、2つはプロフェージに関する領域の多型であり、*S. Enteritidis*よりもフェージのサイクリングが活発であることが予想される。また、もう一つは、嫌気状態での生育に関する遺伝子の有無に関する領域であり、この遺伝子群の有無は、動物の消化管内で増殖する性質を持つ、サルモネラにとって生存に対して有利に働いている可能性があると考えられた。それだけ、*S. Typhimurium*はヒトと動物の間で生き残るすべを持っているのかもしれない。

カンピロバクターについては、キノロン耐性を示す*C. jejuni*及び*C. coli*を対象に解析した。近年のキノロン耐性株の増加について、ゲノム配列からそのメカニズムがどうなっているかを推定することが可能であるかどうか、今回の解析で最も知りたい点であった。細菌の抗菌薬耐性のメカニ

ズムは大別すると、菌がもともと持っている耐性である自然耐性と、プラスミドなどを介して、耐性を獲得する機構が存在することが知られている。ドラフトゲノムであっても、ゲノム上の繰返し配列(rRNAやtRNAなどの遺伝子)やフェージ由来の配列など、アセンブルが困難なものはMiSeqのドラフト配列から解析は困難であるが、通常のCDSであれば、十分な解析に耐えうるデータとなる可能性は高い。そこで、インターネット上で特に獲得耐性に関連する薬剤耐性遺伝子のデータベースがあり、解析したデータをアップロードするだけで、解析したゲノム中にどのような耐性遺伝子が存在しているかを検索してくれるResFinderというサイトがある。その結果は、遺伝子の存在の有無と耐性に関する表現型とは必ずしも一致しないことが確認された(表1)。このような違いは、プラスミド媒介性か、染色体性かによっても異なると思われるし、また耐性遺伝子の安定性にも影響すると思われる。たとえば、今回解析したカンピロバクターの場合には、*tet(O)*と

*aac(6')-aph(2'')*については、プラスミド媒介性である可能性が高いとも割れるが、20株のうち19株で検出された*blaOXA-61*は、その安定性から、染色体性であると推定される。しかし、半分程度は発現していないが、抗菌薬のストレスの存在する環境中では、発現する遺伝子をもつ菌株が選択される可能性もあり、遺伝子プールとして重要な役割を持っていると思われる。

さらにカンピロバクターにおいても、Blast Atlas解析を実施した(図8)。その結果、主要な3つの多型領域が存在していることが分かった。そのうち、2つは、Capsular polysaccharide合成と

Lipooligosaccharide(LOS)合成に関連する領域であり(図12,13)、いずれも菌体表面の抗原性決定領域であり、宿主である動物の消化管内で生存するために有利となるように、菌体表面の抗原性を変化させ、宿主の攻撃を受けない菌が選択的に生き残った結果であるということが出来る。また、もう一つは、I型制限修飾酵素の有無に関する領域であり、*C. coli*は持っていないが、*C. jejuni*については、持っている株と持っていない株が存在していた

(データ未提示)。この遺伝子の機能は、外から侵入してくるDNAの制限に関することであり、一般に*C. coli*は*C. jejuni*より、薬剤耐性能として、多剤化しやすいといわれているが、この現象を説明する証拠となるかは不明であるが、興味深い事実である。

*S. Infantis*77株の解析については、臨床分離株と食品分離株を含めて、SNPを抽出し、SNPによる系統関係を明らかにした。その結果、臨床分離株

17 株の特徴的なクラスターが検出された。この原因としては、我々が選択した食品由来株は、市販の鶏肉がほとんどであり、分離元としては偏っており、今回、検出された臨床分離株由来のクラスターの感染源は、鶏肉ではない別な感染源の可能性が高いと思われた。

さらに、*S. Infantis* の人に対する病原性に関する情報を今回の 77 株を用いて、抽出できるかどうかについて検討した。臨床分離株と食品分離株のゲノム配列上の遺伝子を比較 (GViewSever) し、臨床株に特異的な候補遺伝子として、56 遺伝子を選択した。今回、この 56 遺伝子の評価については、*S. Infantis* 77 株から得られた全 CDS をデータベース化し、その保有状況を調べることで実施した。もし、臨床分離株特異的な遺伝子であれば、その保有分布については、臨床分離株にのみ検出されるはずである。しかし、今回、そのような結果にはならなかった (図 3)。今後、改めて、抽出方法等を検討し、評価することが必要である。

食中毒事例で分離された *C. lari* のゲノム解析に関しては、病原遺伝子を検出する Web サービスを利用することが出来たので、ある程度の品質の配列リードが得られれば、簡単に病原遺伝子のスクリーニングは可能となる。そして、その細菌の病原性遺伝子が検出されれば、起因菌判断が難しいときに、起因菌とするかどうかの判断に役に立つ情報になると思われる。今回の結果では、*C. lari* のこれまでの報告と矛盾しない結果であった。しかし、一部で *C. jejuni* で保有している因子が *C. lari* で検出できなかったことから、このことが、所謂「ドラフト配列」を用いた結果、見逃しているのか、あるいは、*C. lari* による病態が比較的 *C. jejuni* よりも軽症傾向にあることと関連があるのか、興味もたれる。

今年度の *S. Infantis* 77 株のゲノム解析と食中毒事例で分離された *C. lari* のゲノム解析は、MiSeq で取得したリードのみで実施したもので、完全な配列でない、所謂「ドラフト配列」である。通常、完全なゲノム配列を得るためには、ドラフト配列を元にして、「finishing」という工程が必要であり、通常は、膨大なコストと時間が必要となる。その意味で、地衛研での finishing 作業は、事実上、不可能であり、ドラフト配列を利用した活用方法が主流となると思われる。その意味で、SNP による系統解析は十分利用価値が高いことは、既にこれまでの検討から明らかになっている。しかし、病原性に関する検討には、完全ゲノム配列が理想ではあるが、「ドラフト配列」でもある程度は可能であることは、先の *C. lari* のゲノム解析でも示唆されている。従って、この「ドラフト配列」がど

の程度のものであるか、定量的に理解することを試みた。その結果は、図 7 に示した。既に完全なゲノム配列が報告されている腸管出血性大腸菌 O157 堺株を MiSeq でシーケンスし、報告配列、MiSeq で取得した配列を比較することで「ドラフト配列」の評価を試みた。その結果、解析リードのカバー率が約 50 と 100 で比較したところ、コンティグ数、Web サービスによる自動アノテーションの結果から、カバレッジ 50 程度でも充分、ドラフト配列として、活用できることが示された。また、報告された配列と MiSeq から得られた配列から、CDS を比較したところ、83%~87%で一致していることが示された。この数字をどう考えるかについては、Hayashi らの論文の報告と較べてみると、O157 堺株の全 CDS は 5,447 であり、そのうち、大腸菌 K-12 のコアゲノムとしては、3,729 個であり、全体の 70%を占めており、残りは、株特異的な遺伝子となり、病原性に関与する遺伝子が含まれると推定される。このようなことから、83%~87%という数字が示す、一致度は、MiSeq の解析結果としては順当なところであろうと推定される。

E. 結論

NGS の一つである MiSeq を用いた解析は、細菌のドラフトゲノムデータであっても、高精度分子疫学的マーカーとしての SNP の利用価値が高いこと、さらに完全な配列が明らかなゲノム株を reference 株として全コード領域を対象とした Blast 解析を実施すること (Blast Atlas 解析) により、病原性や株の特徴などゲノムワイドな視点から知ることが可能であり、公衆衛生分野で利用価値は高いと思われた。

また、Windows PC 環境においても、インターネット環境の整備と、NGS 解析が出来るソフトウェアの導入、Windows 環境で動作するソフトウェアの導入、Web サービスの利用で、専門性の高い Linux 等を用いなくとも、充分解析が可能であることが示された。

さらに、公的データベースで実施している NGS データの解析支援プログラムを利用することにより、予算的に厳しくても、インターネット環境があれば、Windows 環境でも安価に NGS 解析が出来る可能性がある。このような状況から、地衛研において、MiSeq 等の NGS を所持に関係なく、膨大な NGS データを利用できる環境が整いつつあるなかで、NGS の活用について、検討することは、それほど高いハードルではないと考えられる。

F. 健康危機情報

なし

G. 研究発表

1. 論文発表

なし

2. 学会発表

- ① 肉の刺身を原因とする食中毒事例から分離されたカンピロバクター・ラリのゲノム解析と病原因子の検出
磯部順子、金谷潤一、木全恵子、三井千恵子、
範本志保、佐多徹太郎、綿引正則
第49回腸炎ビブリオシンポジウム（東京都渋谷区、平成27年10月16日）

H. 知的財産権の出願・登録状況

なし

I. 参考資料

以下、巻末に参考資料として追加した。

1. Windows PC 環境における有用な NGS データ解析用ソフト、アプリケーションのリスト

ここで示したソフト等は、一例を示しただけであり、他にも多くのソフトがある。予算や解析目的に沿ったソフトを導入し、評価しながら使うことが重要である。

2. Windows PC で活用することが有用と思われる

ソフトの導入例（チュートリアル）

ゲノム解析のために使用頻度が高い、あるいは、汎用性の高いソフトを導入のためのチュートリアルを参考資料として紹介する。ソフトの正式な使い方は、それぞれ入手して確認していただきたいが、ここでは、Windows PC 環境の導入のために必要なポイントをまとめ、失敗しないで導入するための参考としていただきたい。

- GView Server の使い方
- NCBI Blast+ の Windows 環境での利用について

3. MiSeq を使用する場合のコスト（試算結果例）

MiSeq を使って、ゲノム解析する場合、試薬が高いので、失敗をしないようにすることが重要である。我々は、複数の細菌のドラフト配列を得ているが、最適なクラスターを形成するために、DNA ライブラリーの定量が非常に重要なステップとなる。複数の株を同時に解析するとどうしても、カバー率のばらつきが出るので、たくさんの株を解析する場合には、同じ DNA ライブラリーを2回ランすることとしている。1回目のランで、クラスター形成のばらつきを判定し、2回目のランでは、そのばらつきを解消するように DNA ライブラリーのインプット量を調整する。この方法は失敗が少なくなるので、推奨したい。ここではこの方法を用いたときの消耗品、試薬のコスト計算をしたので、紹介する。

表 1

*Campylobacter jejuni/coli*分離株のゲノム解析
 富山県分離株の薬剤耐性機構の解析: 解析株の特徴

株選定基準: ニューキノロン耐性(4 *coli*+12 *jejuni*)、感性(1 *coli*+3 *jejuni*)

No	Strain	Source	Penner type	NA	OFLX	NFLX	CPFX	ERFX	EM	TC	FOM	GM	SM	ABPC	ABPC/SBT	CF	PIPC	PIPC/TAZ	CTX	ResFinder		
																				tet(0)	aac(6)-aph(2 ^{III})	blaOXA-61
1	Cc01	ヒト	/	R	R	R	R	R	R	R	R	R	R	I	S	R	R	S	I	tet(0)	aac(6)-aph(2 ^{III})	blaOXA-61
2	Cc02	牛レバー	/	R	R	R	R	R	S	R	S	S	S	S	S	R	S	S	S	tet(0)		blaOXA-61
3	Cc03	ヒト	/	R	R	R	R	R	S	R	R	S	S	R	I	R	R	I	I	tet(0)		blaOXA-61
4	Cc04	ヒト	/	S	S	S	S	S	S	S	R	S	S	R	S	R	R	S	I			blaOXA-61
5	Cc05	ヒト	/	R	R	R	R	R	S	S	R	S	S	R	I	R	R	S	S			blaOXA-61
6	Cj06	ヒト	G	R	R	R	R	R	S	R	R	S	S	R	S	R	R	S	S	tet(0)		blaOXA-61
7	Cj07	ヒト	Y	R	R	R	R	R	S	R	R	S	S	R	S	R	R	S	S			
8	Cj08	ヒト	B	R	R	R	R	R	S	R	R	S	S	S	S	R	S	S	S	tet(0)		blaOXA-61
9	Cj09	ヒト	B	R	R	R	R	R	S	S	R	S	S	S	S	R	S	S	S			blaOXA-61
10	Cj10	ヒト	D	R	R	R	R	R	S	S	I	S	S	S	S	R	S	S	S			blaOXA-61
11	Cj11	ヒト	UT	R	R	R	R	R	S	S	R	S	S	S	S	R	S	S	S			blaOXA-61
12	Cj12	ヒト	G	S	S	S	S	S	S	S	R	S	S	S	S	R	S	S	S			blaOXA-61
13	Cj13	ヒト	L	S	S	S	S	S	S	S	R	S	S	S	S	R	S	S	I			blaOXA-61
14	Cj14	ヒト	UT	R	R	R	R	R	S	R	R	S	S	S	S	R	S	S	S	tet(0)		blaOXA-61
15	Cj15	ヒト	UT	R	R	R	R	R	S	S	R	S	S	S	S	R	S	S	S			blaOXA-61
16	Cj16	ヒト	UT	R	R	R	S	R	S	I	I	S	S	S	S	R	S	S	S			blaOXA-61
17	Cj17	ヒト	UT	R	S	S	S	S	S	I	I	S	S	R	S	S	R	S	S			blaOXA-61
18	Cj18	ヒト	UT	R	R	R	R	R	S	R	R	S	S	S	S	R	S	S	S	tet(0)		blaOXA-61
19	Cj19	ヒト	B	R	R	R	R	R	S	R	R	S	S	S	S	R	S	S	S			blaOXA-61
20	Cj20	ヒト	D	R	R	R	R	R	S	S	S	S	S	S	S	R	R	S	I	tet(0)		blaOXA-61

表 2. *S. Infantis* 77 株のゲノム配列データのアセンブル後の概要

Salmonella Infantis 77 株

Miseq Reagent Kit v3 (600 cycle)

		平均値	最大値	最小値
No. of reads		2,047,210	3,835,990	603,348
Total base		490,074,777	825,754,204	158,906,599
Coverage		100	169	33
Contig	Total base	4,891,271	5,158,601	4,588,724
	Count	138	567	64
	Maximam	1,051,615	1,241,747	480,288
	N50	290,813	521,185	159,289

表3. 比較ゲノム学的手法を用いて、臨床分離 *S. Infantis* 株に特異的に存在することが期待された 56 の候補遺伝子

No.	Seq_ID	Hit Product
1	pj039-01	hypothetical protein
2	pj039-02	hypothetical protein
3	pj039-03	hypothetical protein
4	pj039-04	hypothetical protein
5	pj039-05	hypothetical protein
6	pj039-06	HTH-type transcriptional regulator hdfR
7	pj039-07	SanA protein
8	pj039-08	Cytidine deaminase
9	pj039-09	hypothetical protein
10	pj039-10	hypothetical protein
11	pj039-11	hypothetical protein
12	pj039-12	Ferredoxin
13	pj039-13	Mobile element protein
14	pj039-14	hypothetical protein
15	pj039-15	hypothetical protein
16	pj039-16	hypothetical protein
17	pj039-17	ClpB protein
18	pj039-18	hypothetical protein
19	pj039-19	hypothetical protein
20	pj039-20	Glutamate racemase
21	pj039-21	Glutamate racemase
22	pj039-22	Vitamin B12 receptor BtuB
23	pj039-23	hypothetical protein
24	pj039-24	hypothetical protein
25	pj039-25	5-deoxy-glucuronate isomerase
26	pj039-26	Methylmalonate-semialdehyde dehydrogenase
27	pj039-27	Cellulose synthase catalytic subunit
28	pj039-28	Cellulose synthase, putative
29	pj039-29	Putative cytoplasmic protein
30	pj039-30	hypothetical protein
31	pj039-31	hypothetical protein
32	pj039-32	Preprotein translocase subunit SecE
33	pj039-33	Transcription antitermination protein NusG
34	pj039-34	Cytochrome c heme lyase subunit CcmL / CcmH
35	pj039-35	hypothetical protein
36	pj039-36	hypothetical protein
37	pj039-37	hypothetical protein
38	pj039-38	hypothetical protein
39	pj039-39	Colicin V production protein
40	pj039-40	hypothetical protein
41	pj039-41	hypothetical protein
42	pj039-42	hypothetical protein
43	pj039-43	hypothetical protein
44	pj039-44	hypothetical protein
45	pj039-45	hypothetical protein
46	pj039-46	hypothetical protein
47	pj039-47	Polymyxin resistance protein AmC
48	pj039-48	hypothetical protein
49	pj039-49	hypothetical protein
50	pj039-50	hypothetical protein
51	pj039-51	putative cytoplasmic protein
52	pj039-52	Type III secretion protein SsaB
53	pj039-53	hypothetical protein
54	pj039-54	hypothetical protein
55	pj039-55	hypothetical protein
56	pj039-56	Mobile element protein

表4. 食中毒事例で分離された *C.lari* 株のゲノム配列データのアセンブル後の概要

	カウント数	平均長(bp)	全塩基数(bp)
リード数	2,311,897	277.25	641,005,075
Contigs数	38	42,564	1,617,445 (396 coverages)

表5. 食中毒事例で分離された *C.lari* 株のゲノム配列から推定した主要な病原因子の保有状況

	<i>C.lari</i> Toyama 07	<i>C.lari</i> RM2100	<i>C. Jejuni</i> All strains
定着因子	CadF	+	+
	JipA	-	+
	MOMP	+	+
	PEBA	-	+
毒素	CdtABC	+	+
	CiaB	+	+

表 6. 腸管出血性大腸菌 O157 堺株のゲノム解析結果の概要

SK01		Summary statistics	Count	average length	Total bases
Contig measurements (including scaffolded regions)		Reads	802,464	296	237,427,329
N75	52,302	Matched	788,226	296	233,224,139
N50	124,190	Not Matched	14,238	295	4,203,190
N25	235,262	Contigs	203	26,429	5,365,229
Minimum	500	Reads in pairs	660,940	487	
Maximum	314,054	broken paired reads	127,286	272	
average	26,430				Coverage=43.5
Count	203				

SK02		Summary statistics	Count	average length	Total bases
Contig measurements (including scaffolded regions)		Reads	941,258	296	278,710,760
N75	84,323	Matched	927,039	296	274,484,347
N50	146,270	Not Matched	14,219	297	4,226,413
N25	224,569	Contigs	183	29,450	5,389,464
Minimum	512	Reads in pairs	786,502	490	
Maximum	374,972	broken paired reads	140,537	271	
average	29,451				Coverage=50.9
Count	183				

SK1.2 (=SK01+SK02)		Summary statistics	Count	average length	Total bases
Contig measurements (including scaffolded regions)		Reads	1,743,722	296	516,138,089
N75	73,375	Matched	1,716,252	296	507,995,197
N50	148,380	Not Matched	27,470	296	8,142,892
N25	327,284	Contigs	182	29,670	5,399,993
Minimum	537	Reads in pairs	1,442,975	488	
Maximum	374,974	broken paired reads	273,278	272	
average	29,670				Coverage=94.1
Count	182				

表 7 O157 堺株の配列の評価

	EHEC O157 Sakai from Public site				EHEC O157 Sakai by MiSeq		
	Chromo some	pO157	pOSAK1	Total	SK01	SK02	SK12
Length of sequence (bp)	5,498,450	92,721	3,306	5,594,477	5,365,229	5,389,464	5,399,993
Contigs	1	1	1	3	203	183	182
ORF (=CDS, >150)	5,361	83	3	5,447			
RNA	110	0	0	110			

Hayashi, T., et al. *DNA Res.*, 8:11(2001)

	after Re-annotation by RAST				after annotations by RAST		
ORF (=CDS, >150)	5,208	99	3	5,310	5,091	5,190	5,138
RNA	124	0	0	124	107	137	107

de novo assembly

リードをレファレンス配列(アノテーション情報あり)にマッピング

(ない場合は RAST で取得)

<http://rast.nmpdr.org/>

SNP 検出

解析株の SNP 情報を統合

CDS の SNP のみ抽出

16S rRNA、prophage、pathogenicity island
にある SNP は除去

系統樹作成

CLC genomic workbench

エクセル

MEGA

図1 解析株のゲノム配列から Single Nucleotide Polymorphism (SNP) の抽出と系統樹作成工程

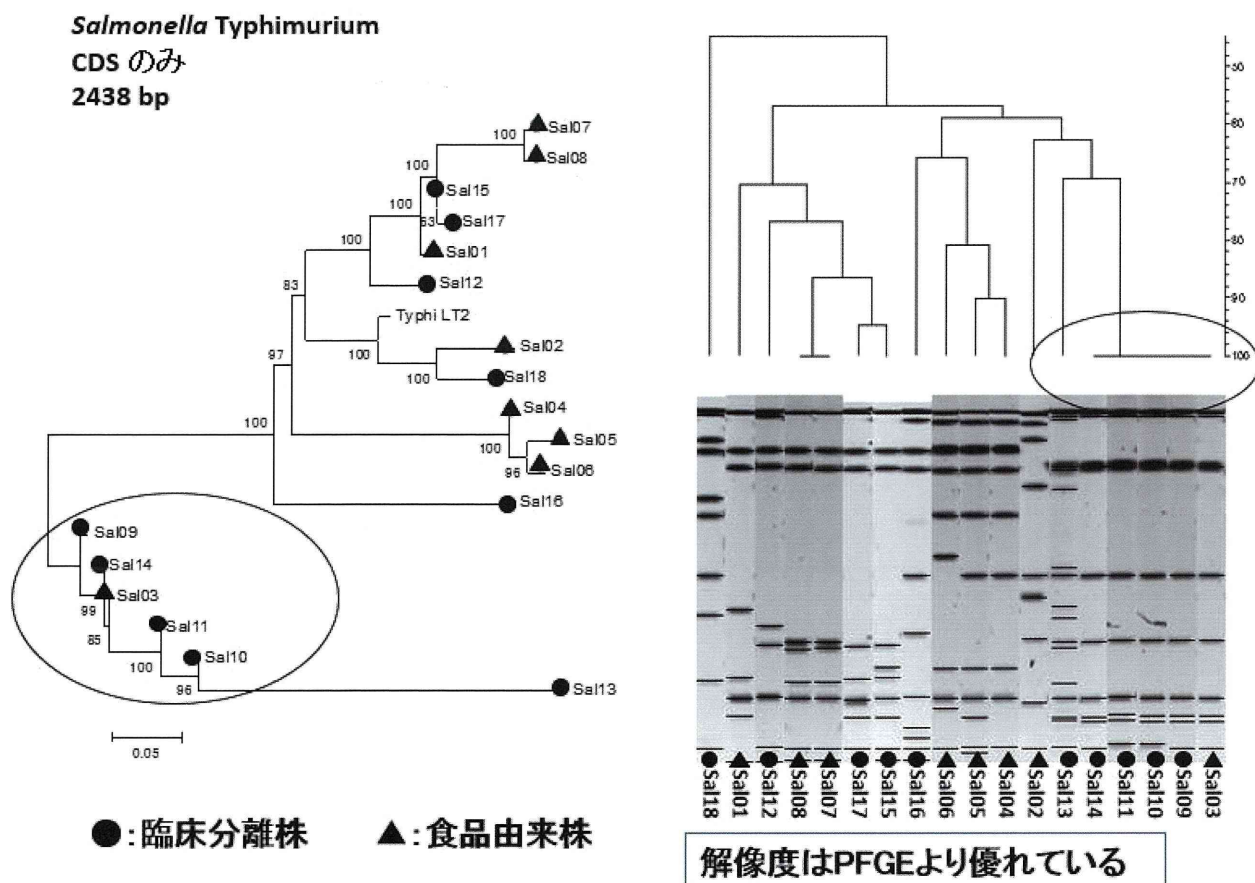


図2. *Salmonella Typhimurium* 分離 18 株の PFGE とゲノム SNP による系統樹の比較

Salmonella Enteritidis
CDSのみ
301 bp

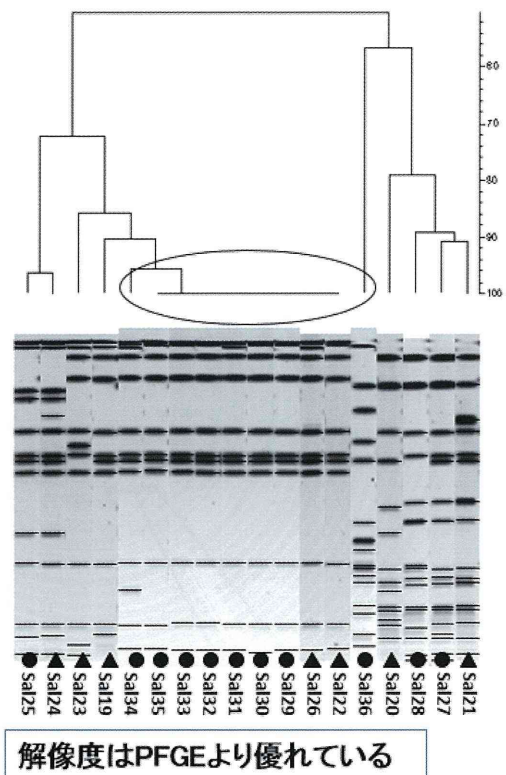
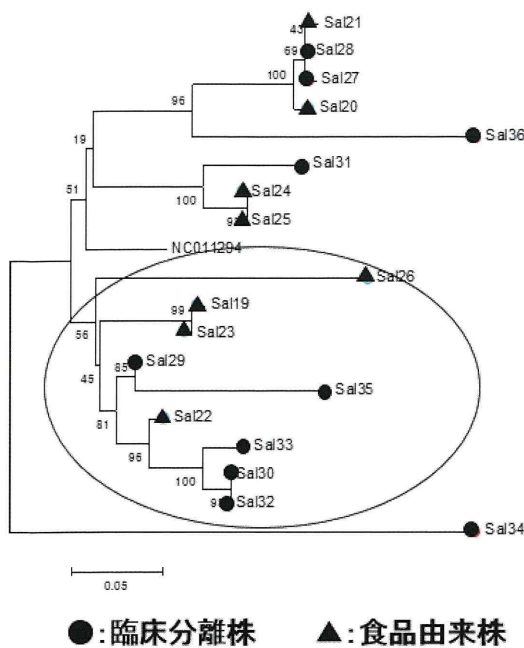
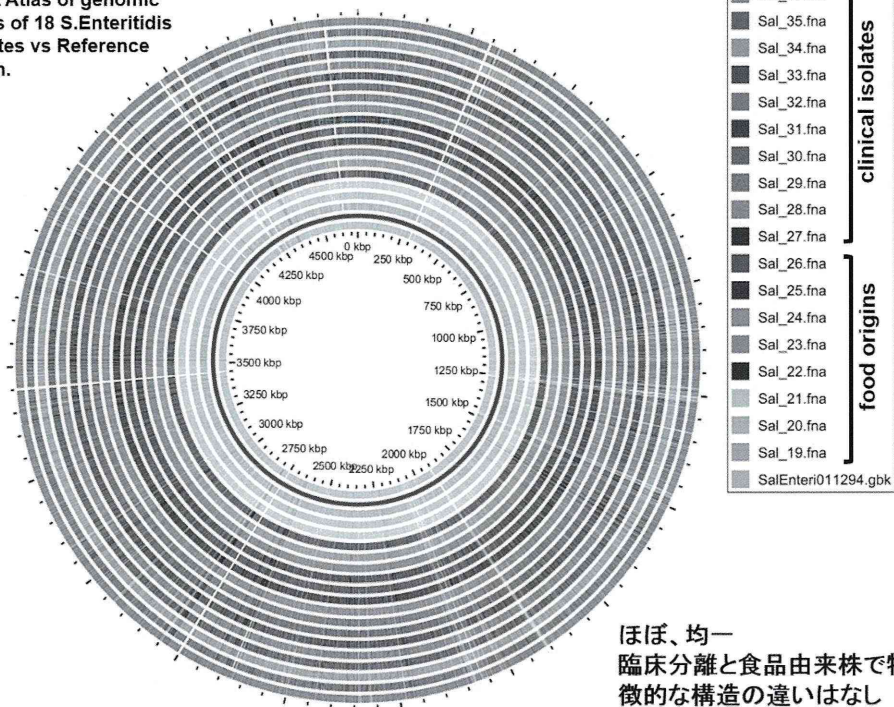


図3. *Salmonella* Enteritidis 分離 18 株の PFGE とゲノム SNP による系統樹の比較

Blast Atlas of genomic CDSs of 18 *S. Enteritidis* isolates vs Reference strain.



ほぼ、均一
臨床分離と食品由来株で特
徴的な構造の違いはなし

図4. *S. Enteritidis* 18 株の Blast Atlas 解析

Blast Atlas of genomic CDSs
of 18 *S. Typhimurium* isolates
vs Reference strain.

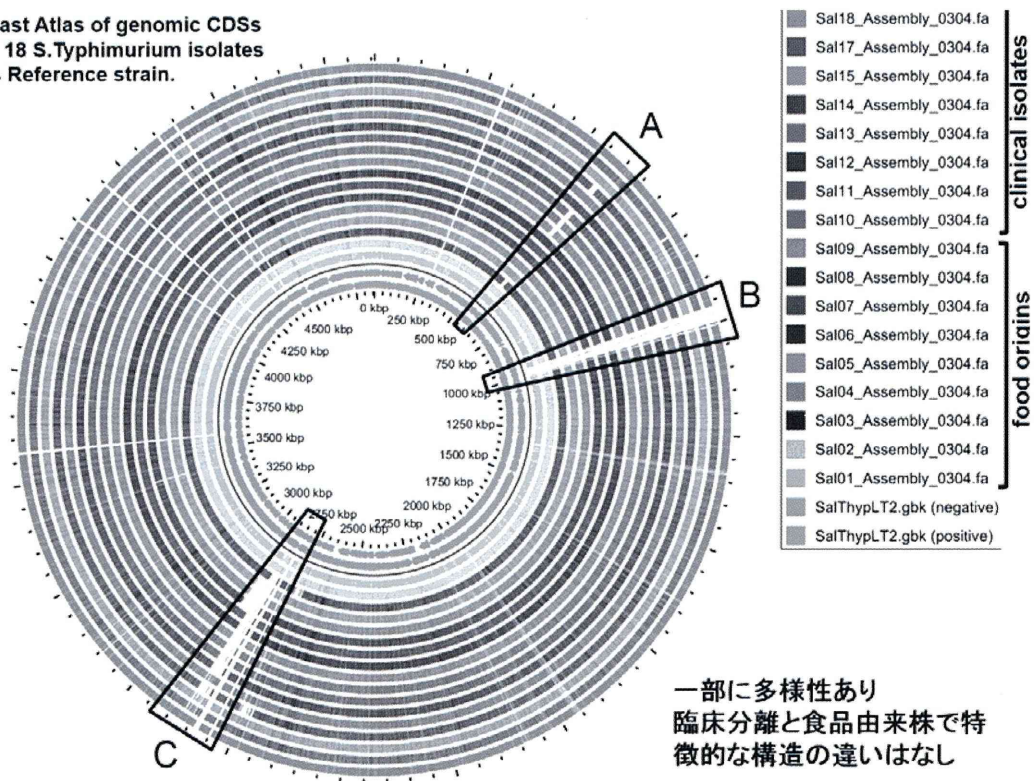


図 5. *S. Typhimurium* 18 株の Blast Atlas 解析

Campylobacter jejuni/coli 分離株のゲノム解析 富山県分離株の薬剤耐性機構の解析

カンピロの薬剤耐性メカニズム(ref: BioMed Res. Int., 2013:12)

背景

動物由来の食品から薬剤耐性:

フルオロキノロン耐性、マクロライド、アミノグルコシド、 β ラクタム耐性が最近、問題

自然耐性: ペニシリン、セファロスポリン、トリメトプリム、リファンピシン、バンコマイシン

Mechanism

-Resistance to Quinolones.

DNA gyrase, topoisomerase IVの amino acid substitution (QRDR, quinolone resistance-determining region)

CmeABC multidrug efflux pump (fluoroquinolones and macrolides): 自然耐性と獲得耐性に関与

-Resistance to Tetracyclines.

tet(O): カンピロバクターに共通、プラスミド媒介性が多い

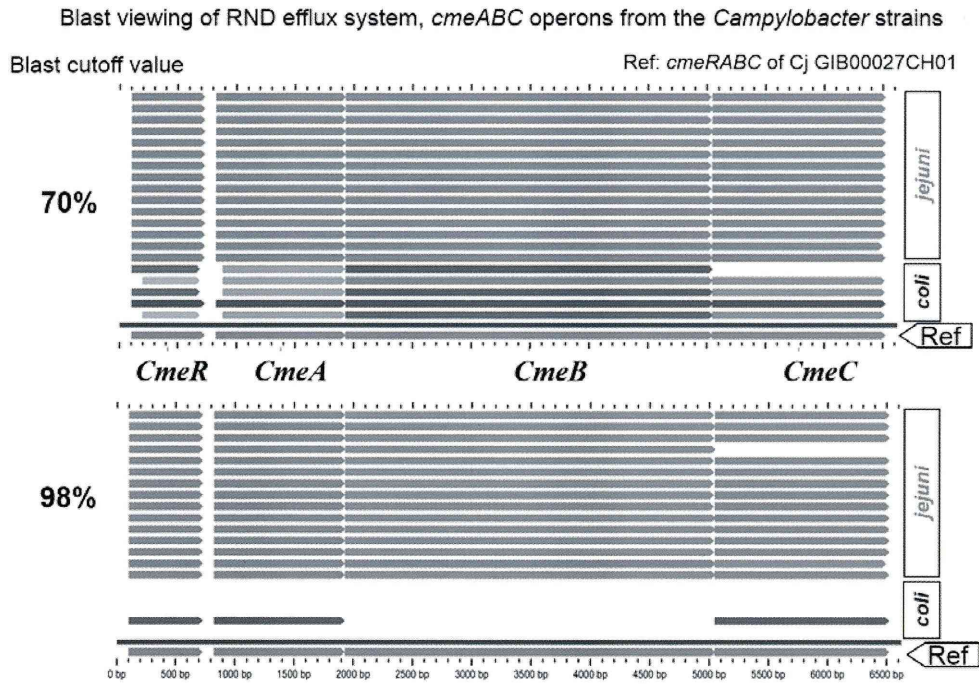
-Resistance to Macrolides.

Interrupt protein synthesis. 23S nuc no. 2058 and 2059=key contact sites for macrolides binding.

Resistance phenotype, modification of the rRNA target binding (mutation,

CmeABC multidrug efflux pump

図 6. *Campylobacter* 属菌の薬剤耐性機構 (これまでの報告)



*cmeRABC*の多型構造とキノロン耐性の相関については、もう少しデータの検証が必要である。

図 7. *Campylobacter* 属菌の *cmeABC* オペロンの Blast 解析

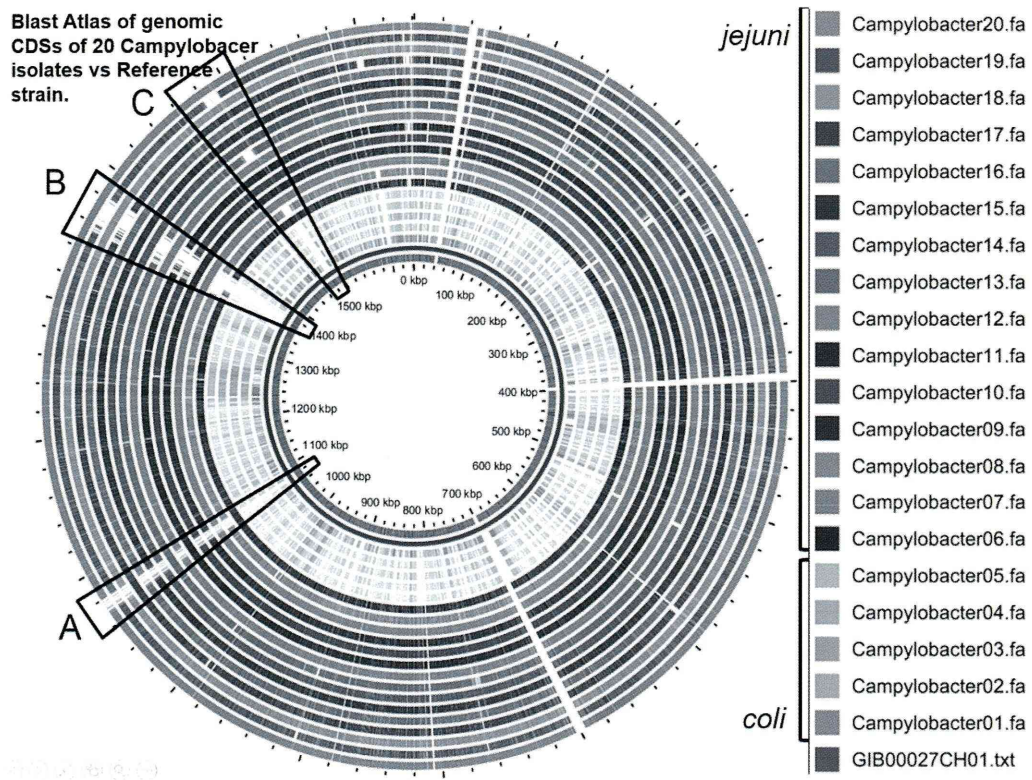


図 8. *Campylobacter* 属菌 20 株の Blast Atlas 解析

Salmonella Infantis
CDS のみ
104,586 bp

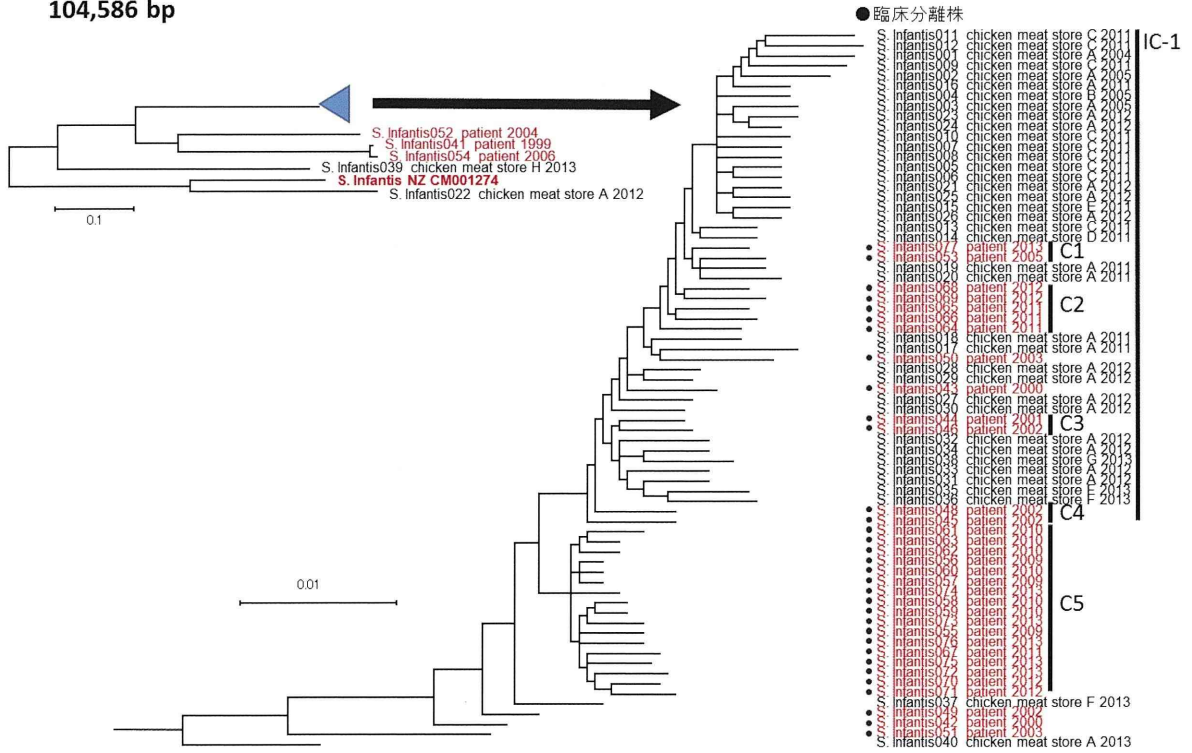


図 9. *S. Infantis* 77 株からの Single Nucleotide Polymorphism (SNP) による系統樹

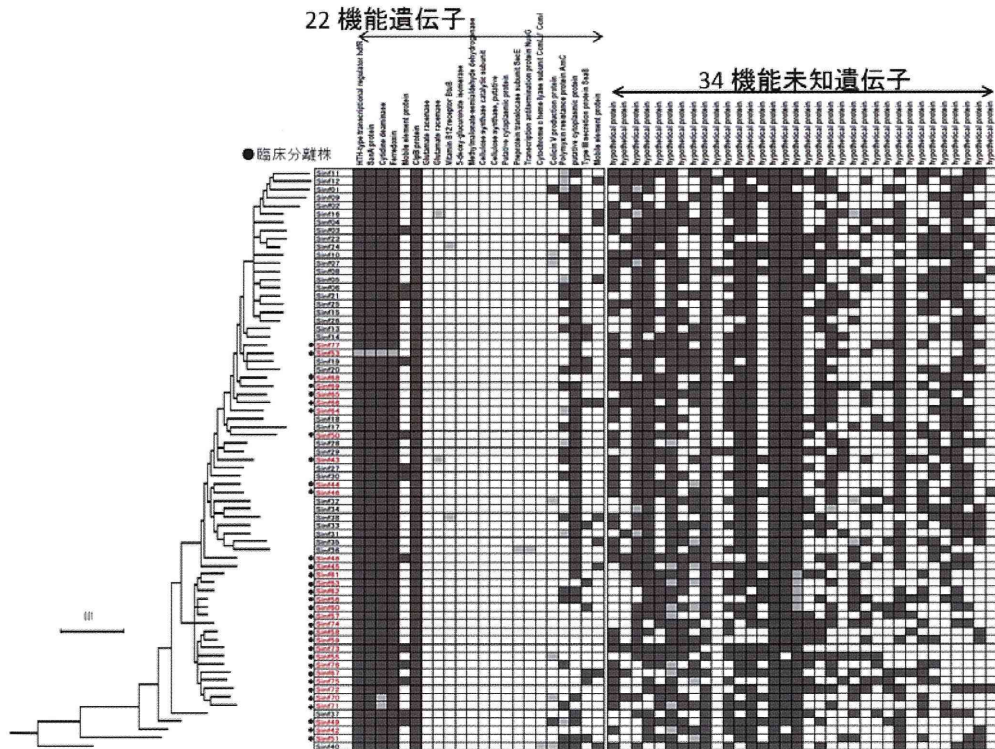


図 10. 臨床分離株特異的と推定された 56 遺伝子の確認

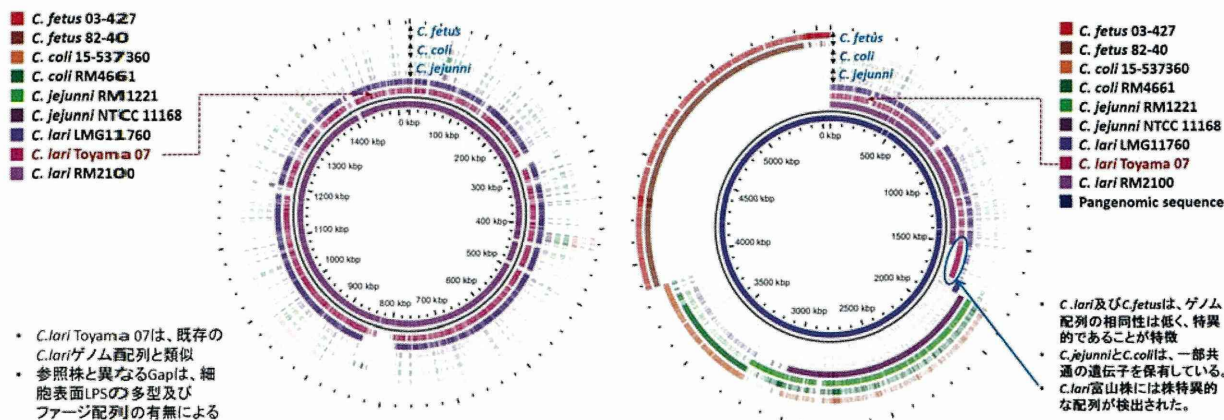


図 1 1 . ゲノム解析した *C.lari* 分離株の Blast atlas 結果

図 1 2 . ゲノム解析した *C.lari* 分離株の Pangenome blast atlas 結果

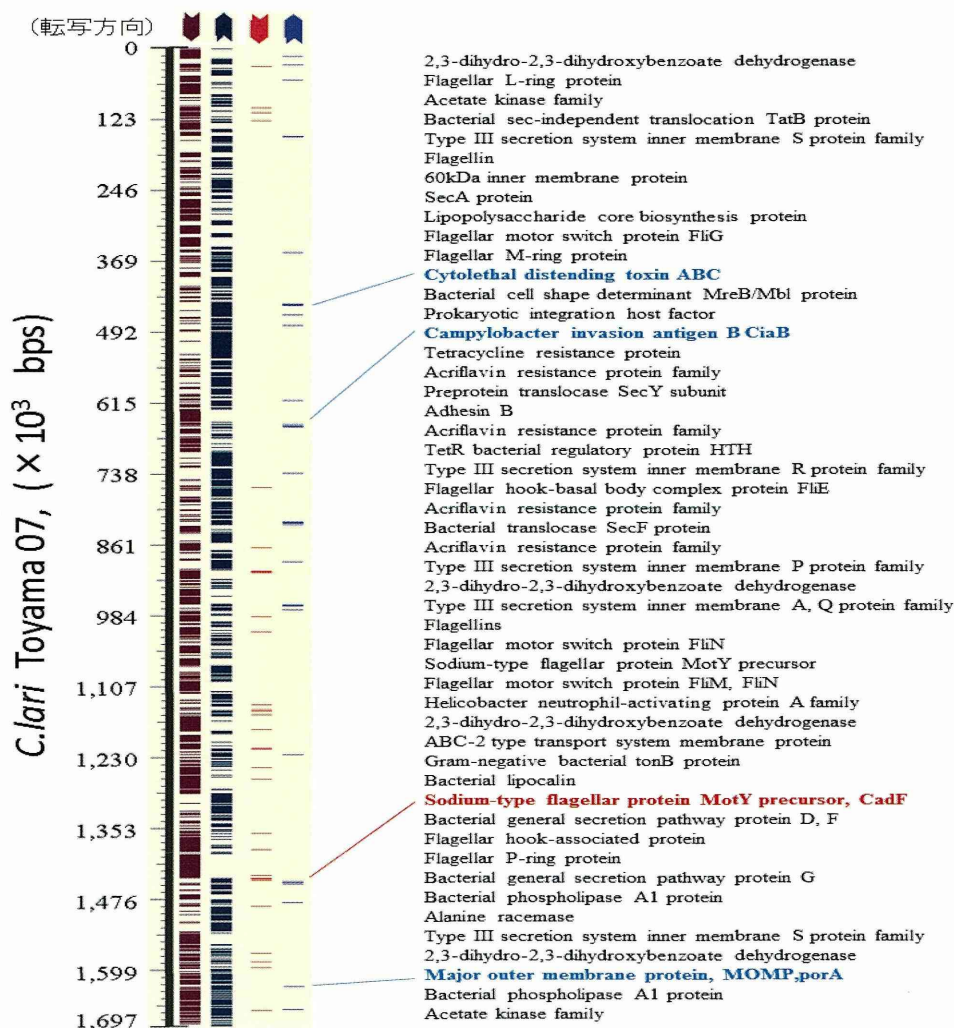
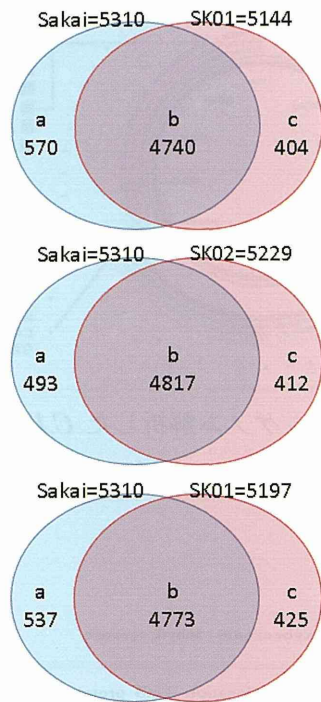


図 1 3 . *C.lari* Toyama07 の病原因子の検索結果

左図は、*C.lari* Toyama07 のゲノム上の遺伝子のマッピングしたもので、Virulence Searcher (Web サービス)を利用して検出された 53 遺伝子のゲノム上の位置

CDS, >150bp



a, Sakaiのみ検出
 b, SakaiとSKに共通
 c, SKのみ検出

Comparison of RAST Annotation Data between Public Sequence and MiSeq Draft Data

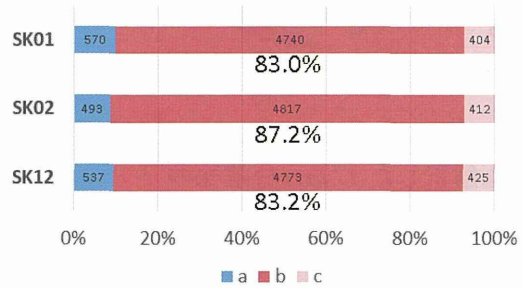


図 1 4. 腸管出血性大腸菌 O157 堺株を用いた MiSeq データのドラフト配列の定量的な理解
 ゲノム配列既知の O157 堺株を MiSeq でゲノム配列を得て、報告されている配列と比較して、
 CDS の一致度を評価した。

I. 参考資料 1. WindowsPC環境における有用なNGS解析のためのソフト、アプリケーションのリスト

用途	ソフト名	説明	PC	利用形態	必要条件等	コスト	URL	コメント
NGS用総合解析	CLC genomic workbench	NGSデータを様々な解析が可能。De novo assemble、各種マッピング等可能。また、ゲノム解析にしばしば利用されるLinux系のソフトで利用できるファイル形式を読み込んだり、書き出しなど可能である。このソフトがあれば、ほぼ、なんとかなる。但し、高額である。	Windows	独立		有料	フィルジェン株式会社 http://www.filgen.jp/	公的データベースや、論文等でも引用されている。
NGS	DDBJ Read Annotation Pipeline	Web環境が必要だが、de novo assemble等が可能。	Windows	ウェブ		無料	https://p.ddbj.nig.ac.jp/pipeline/Login.do	登録必要/登録しているが、使用経験なし。
自動アノテーション	RAST	アップロードされたゲノム配列について、遺伝子データベースに対して、blast解析を実施し、蛋白質をコードする領域等を推定し、注釈するサービス。解析結果は、様々な様式で入手可能で、細菌のゲノム解析では有用な解析サービスである。	Windows	ウェブ	firefox	無料	http://rast.nmpdr.org/	登録必要/スタートゴドンの選択が時々、異なっており、注意が必要であるが、様々な様式のファイルがダウンロードできるので、非常に有用なサービス。
自動アノテーション	MiGAP	DDBJ japanが提供している細菌ゲノム配列の注釈サービス。RASTと同じような目的で利用されるWebサービスである。	Windows	ウェブ		無料	http://www.migap.org/	登録必要/使用頻度は高くありません。
比較ゲノム解析結果の比較	GviewServer	細菌ゲノムの比較解析を実施してくれるサービス。複数のゲノムを環状、線状形式で、ゲノム上のすべての遺伝子を対象として、視覚化できる。	Windows	ウェブ	firefox	無料	https://server.gview.ca/	登録必要/使用頻度は高い。
ゲノムの可視化	Gview	genebank形式のファイルを見やすい図として表示できる。環状、線状の図として表示し、CDSの色分け等、カスタマイズできる。また、Genbank形式やgffファイルを読み込ませ、より複雑な図として、カスタマイズできる。GVewServerで得られたblast atlasの結果を表示する場合も使われるソフトである。	Windows	独立	Java	無料	https://www.gview.ca/wiki/GView/WebHome	使用頻度は高い。
手動アノテーション	Artemis	塩基配列の特徴を見やすく表示してくれる。また、遺伝子のアノテーションを行うことができる。EMBL及びGENBANK形式の配列データを読み込むことができる。	Windows	独立	Java	無料	http://www.sanger.ac.uk/science/tools/artemis	使用頻度は高い。
比較ゲノム解析	ACT	Artemisをベースにゲノムワイドな配列を比較し、その結果を表示する。その結果から、大きな領域での逆位、欠失、挿入などを視覚化してくれる。	Windows	独立	Java	無料	http://www.sanger.ac.uk/science/tools/artemis-comparison-tool-act	大きなGAPや逆位等が視覚化できる。
ABIシーケンサーデータのAssemble	Sequencher	ABIシーケンサーで解析した配列の確認、Assembleソフト	Windows	独立		有料	タカラバイオ	MiSeqデータではなく、ABIシーケンサーからのデータの解析のため、使用頻度は高い。