

論文 / 著書情報  
Article / Book Information

題目(和文)	高精度ゲノム解析手法の確立による大腸菌O157の大規模比較研究
Title(English)	
著者(和文)	吉村大
Author(English)	Dai Yoshimura
出典(和文)	学位:博士(理学), 学位授与機関:東京工業大学, 報告番号:甲第11224号, 授与年月日:2019年6月30日, 学位の種別:課程博士, 審査員:伊藤 武彦,和地 正明,立花 和則,二階堂 雅人,山田 拓司
Citation(English)	Degree:Doctor (Science), Conferring organization: Tokyo Institute of Technology, Report number:甲第11224号, Conferred date:2019/6/30, Degree Type:Course doctor, Examiner:,,,,,
学位種別(和文)	博士論文
Category(English)	Doctoral Thesis
種別(和文)	論文要旨
Type(English)	Summary

# 論文要旨

## THESIS SUMMARY

系・コース： Department of, Graduate major in	生命理工学 生命理工学	系 コース	申請学位 (専攻分野)： 博士 Academic Degree Requested Doctor of	( 理学 )
学生氏名： Student's Name	吉村 大		指導教員 (主)： Academic Supervisor(main)	伊藤 武彦
			指導教員 (副)： Academic Supervisor(sub)	

### 要旨 (和文 2000 字程度)

Thesis Summary (approx.2000 Japanese Characters )

腸管出血性大腸菌 (EHEC) は下痢や出血性腸炎のほか溶血性尿毒症症候群 (HUS) などの重毒な合併症を引き起こす大腸菌であり、先進国においても数多くの有症事例を引き起こしている。大腸菌は細胞壁の抗原性 (O 抗原) を基にした分類が広く普及しているが、中でも 0157 は EHEC に占める割合が最も高い O 抗原であり、その対策は公衆衛生上極めて重要な課題となっている。0157 は株により毒性が異なることが知られており、系統によって重症化率に差があることも提唱されているが、裏付けとなる高精度な系統解析事例は十分でなく、その重症化率の違いの原因も詳細には明らかになっていない。本研究では国内のヒトやウシから分離された 2500 株超の 0157 について全ゲノム比較解析を実施し、その系統関係や重症化因子を明らかにすることを目的とした。

全ゲノム比較解析において、株間の一塩基多型 (SNP) の検出やゲノムアセンブリは解析の基盤となる工程であり、配列と表現型との関連解析や系統解析などの下流の解析に重大な影響を与える。0157 の比較解析を実施するにあたりこれらの 2 工程を実施する既存ツールの性能に問題があることが明らかとなったため、本研究ではその問題点を解決する新規ツールの開発を併せて実施し、開発した新規ツールを用いて 0157 の比較解析を実施した。

まず最初に既存の SNP 検出ツールについてシミュレーションによるベンチマークを行ったところ、リードの参照配列へのマッピング情報を用いるマッピングベースのツールは検出精度が低く、リードをアセンブリして作成したコンティグと参照配列とのアライメント情報を用いるアセンブリベースのツールは検出感度が低いという問題点が明らかとなった。そこでマッピングとアセンブリの双方の情報を利用する新規株間 SNP 検出パイプライン BactSNP を新規開発したところ、BactSNP は同様のベンチマークにおいて高精度かつ高感度な株間 SNP の検出を実現した。

次に既存のゲノムアセンブラについて、完全長ゲノム配列が既知である菌株のシーケンズデータから元のゲノム配列を再現した場合の完成度を評価した。その結果、真核生物向けに開発されたゲノムアセンブラ Platanus は他のツールと比較して高精度にゲノムを再現可能であるものの、再構築された配列の連続度は他のツールに劣るものとなった。そこでバクテリアのゲノムサイズが小さいことを利用し、リードのペア情報を用いて配列同士をつなぎ合わせる工程と、つなぎ合わせた配列間のギャップを埋める工程とを繰り返す機能を Platanus に付加し Platanus\_B を開発したところ、Platanus\_B は高精度かつ高連続度なゲノムの再構築を実現した。

最後に、新規に開発した上述のツール BactSNP・Platanus\_B を利用し、大腸菌 0157 の比較解析を行った。まず BactSNP で検出した株間 SNP を基に系統解析を実施し、分離地域・宿主・毒性と系統との関連を調査したところ、国内の 0157 が地域によって隔てられることなく拡散していること、ウシからヒトへの感染が一部の系統だけでなく様々な系統で生じていること、HUS 患者由来株が一部の系統に有意に高頻度に存在することなどが明らかとなった。次に HUS の発症と主要毒性因子志賀毒素 (Stx) との関連性について調査を行った。その結果、特定の配列型の Stx ファージを有する株が高頻度に HUS 患者に由来することを示した結果が得られた。また Platanus\_B で再構築したゲノム配列を用いたパンゲノム解析からは、国内の 0157 の遺伝子レパートリが極めて多様であることが明らかとなった。その一因であるプラスミドについて replicon 領域の相同性に基づく分類を行った結果、近縁系統内でも多様なプラスミドが頻繁に挿入・欠失していることが明らかとなった。

本論文ではバクテリアの比較ゲノム研究において基盤となる株間 SNP の検出・ゲノムアセンブルという 2 つの工程について高精度な解析手法を確立し、さらにそれらを 0157 の大規模比較ゲノム解析に適用した結果、その拡散様式や強毒化因子、遺伝的多様性について重要な知見を明らかにした。

備考：論文要旨は、和文 2000 字と英文 300 語を 1 部ずつ提出するか、もしくは英文 800 語を 1 部提出してください。

Note：Thesis Summary should be submitted in either a copy of 2000 Japanese Characters and 300 Words (English) or 1 copy of 800 Words (English).

注意：論文要旨は、東工大リサーチポータル (T2R2) にてインターネット公表されますので、公表可能な範囲の内容で作成してください。

(博士課程)  
Doctoral Program

## 論文要旨

THESIS SUMMARY

系・コース： Department of, Graduate major in	生命理工学 生命理工学	系 コース	申請学位 (専攻分野)： Academic Degree Requested	博士 Doctor of	(理学)
学生氏名： Student's Name	吉村 大		指導教員 (主)： Academic Supervisor(main)	伊藤武彦	
			指導教員 (副)： Academic Supervisor(sub)		

### 要旨 (英文 300 語程度)

Thesis Summary (approx.300 English Words)

Enterohemorrhagic escherichia coli (EHEC) causes bloody diarrhea and severe complications such as hemolytic uremic syndrome (HUS). O157 is the most frequently reported O serotype in EHEC; thus, the surveillance of O157 is important for the public health. In this study, I performed a comparative genomic analysis of more than 2,500 O157 isolates in Japan in order to reveal the phylogenetic relationship of O157 and the risk factors for severe disease caused by O157.

The detection of single nucleotide polymorphisms (SNPs) among isolates and the *de novo* assembly of genome sequences are the fundamental steps in these comparative genomic analyses. However, the performances of the existing tools for these analyses turned out to be not enough through this study. Therefore, I developed novel tools BactSNP and Platanus\_B which resolved the problems of the existing tools and applied them to the comparative study of O157.

The phylogenetic analysis based on the SNPs information revealed that O157 in Japan has spread across the borders of the regions and that the isolates from humans and cattle were phylogenetically mixed. One of the phylogenetic clusters contained significantly higher rate of isolates from HUS patients. The major virulence factor of EHEC is Shiga toxin (Stx), and one of its subtypes Stx2a is said to be a particularly high-risk factor for severe disease. The comparative analysis of Stx2a phage subtypes revealed that one of the subtypes was significantly associated with HUS patients. In addition, the pan-genome analysis based on the assembled genomes revealed the high level of genetic diversity even among the typical O157 in Japan. These insights about O157 are expected to largely contribute to the development of microbiology.

備考：論文要旨は、和文 2000 字と英文 300 語を 1 部ずつ提出するか、もしくは英文 800 語を 1 部提出してください。

Note: Thesis Summary should be submitted in either a copy of 2000 Japanese Characters and 300 Words (English) or 1 copy of 800 Words (English).

注意：論文要旨は、東工大リサーチリポジトリ(T2R2)にてインターネット公表されますので、公表可能な範囲の内容で作成してください。

Attention: Thesis Summary will be published on Tokyo Tech Research Repository Website (T2R2).