

Research Article**Application of Nanopore whole-genome sequencing to analyze the genetic relatedness among *Salmonella* spp. isolates from a *bánh mì*-associated foodborne outbreak in Vietnam**Pham Ngoc Ha¹, Nguyen Phuong Thao¹, Ha Anh Thu¹, Do Duc Huy², Nguyen Thanh Trung^{1*}¹National Institute for Food Control, Hanoi, Vietnam²University of Science and Technology of Hanoi, Vietnam

(Received: 10 Jul 2026; Revised: 18 Sep 2026; Accepted: 27 Sep 2026)

Abstract

Salmonella spp. are among the major causative agents of foodborne illness. Conventional microbiological methods are effective for pathogen detection and isolation but have limited ability to assess the genetic relatedness among isolates recovered from different samples. This study applied Nanopore whole-genome sequencing to assess the genetic relatedness among *Salmonella* spp. isolates recovered during a *bánh mì*-associated foodborne outbreak in Vietnam. A total of 16 outbreak-associated samples, including food and clinical specimens, were tested for common microbiological hazards and bacterial toxins relevant to foodborne outbreak investigations. Ten of the 16 samples were positive for *Salmonella* spp. The recovered *Salmonella* isolates were subjected to whole-genome sequencing and bioinformatics analysis to assess read quality, genome assembly characteristics, serotype, multilocus sequence type, plasmid profile, and antimicrobial resistance-associated genes. All ten isolates were identified as *Salmonella enterica* serovar Enteritidis, belonged to sequence type ST11, and showed highly similar genetic profiles. Single nucleotide polymorphism (SNP) analysis revealed pairwise distances of 0 - 6 SNPs, with several food isolates showing identical SNP profiles to isolates recovered from clinical specimens. These findings indicate a close genetic relationship between the food and clinical isolates and provide genomic evidence supporting the epidemiological hypothesis of a common exposure source in the outbreak.

Keywords: *Salmonella* spp., foodborne outbreak, *bánh mì*, Nanopore sequencing, genetic relatedness, whole genome sequencing.

* Corresponding author: Nguyen Thanh Trung (E-mail: trungtn@nifc.gov.vn)
<https://doi.org/10.47866/2615-9252/vjfc.4754>

Copyright © 2026 The author(s). This article is licensed under [CC BY-NC 4.0](#)

Ứng dụng giải trình tự toàn bộ hệ gen bằng công nghệ Nanopore trong phân tích mối liên hệ di truyền giữa các chủng *Salmonella* spp. liên quan đến vụ ngộ độc thực phẩm do bánh mì tại Việt Nam

Phạm Ngọc Hà¹, Nguyễn Phương Thảo¹, Hà Anh Thư¹, Đỗ Đức Huy², Nguyễn Thành Trung¹

¹Viện Kiểm nghiệm an toàn vệ sinh thực phẩm quốc gia, Hà Nội, Việt Nam

²Trường Đại học Khoa học và Công nghệ Hà Nội, Hà Nội, Việt Nam

Tóm tắt

Salmonella spp. là một trong những tác nhân quan trọng gây ngộ độc thực phẩm, tuy nhiên, các phương pháp vi sinh truyền thống chủ yếu cho phép phát hiện tác nhân mà chưa đủ khả năng đánh giá mối liên hệ di truyền giữa các chủng phân lập từ các mẫu khác nhau. Nghiên cứu này ứng dụng giải trình tự toàn bộ hệ gen bằng công nghệ Nanopore nhằm hỗ trợ đánh giá mối liên hệ di truyền giữa các chủng *Salmonella* spp. trong một vụ ngộ độc thực phẩm liên quan đến bánh mì tại Việt Nam. Tổng số 16 mẫu thực phẩm liên quan đến vụ việc được xét nghiệm các chỉ tiêu vi sinh và độc tố thường gặp trong điều tra ngộ độc thực phẩm. Kết quả cho thấy 10/16 mẫu dương tính với *Salmonella* spp., trong khi phần lớn các chỉ tiêu khác âm tính hoặc ở mức thấp. Các chủng *Salmonella* phân lập từ các mẫu dương tính được giải trình tự toàn bộ hệ gen bằng công nghệ Nanopore và phân tích tin sinh học để xác định chất lượng dữ liệu đọc, đặc điểm lắp ráp hệ gen, kiểu huyết thanh, kiểu trình tự đa locus, plasmid và các gen liên quan đến kháng kháng sinh. Kết quả cho thấy 10 chủng phân lập đều được xác định là *Salmonella enterica* serovar Enteritidis, thuộc kiểu trình tự ST11 và có hồ sơ di truyền tương đồng. Phân tích đa hình nucleotide đơn (SNP) cho thấy khoảng cách giữa các chủng dao động từ 0 đến 6 SNP, trong đó một số chủng phân lập từ thực phẩm có kiểu SNP giống hệt các chủng phân lập từ bệnh phẩm. Các kết quả này cho thấy các chủng phân lập từ thực phẩm và bệnh phẩm có mối liên hệ di truyền chặt chẽ và cung cấp bằng chứng hệ gen hỗ trợ giả thuyết dịch tễ học về một nguồn phơi nhiễm chung trong vụ việc.

Từ khóa: *Salmonella* spp., ngộ độc thực phẩm, bánh mì, giải trình tự Nanopore, mối liên hệ di truyền, giải trình tự toàn bộ hệ gen

1. ĐẶT VẤN ĐỀ

Salmonella spp. là một trong những mầm bệnh vi khuẩn quan trọng liên quan đến ngộ độc thực phẩm trên toàn thế giới. Theo Tổ chức Y tế Thế giới (WHO), *Salmonella* nằm trong số các nguyên nhân hàng đầu gây bệnh tiêu chảy do thực phẩm trên toàn cầu, trong khi Trung tâm Kiểm soát và Phòng ngừa Dịch bệnh Hoa Kỳ xác định *Salmonella* là một trong những nguyên nhân hàng đầu gây bệnh, nhập viện và tử vong do thực phẩm [1], [2]. Tại Việt Nam, một số vụ ngộ độc thực phẩm gần đây cũng đã thu hút sự quan tâm đặc biệt đối với vai trò của *Salmonella* trong các sản phẩm bánh mì ăn liền. Năm 2023, vụ ngộ độc liên quan đến bánh mì tại Hội An đã ghi nhận 313 trường hợp nhập viện; quá trình điều tra phát hiện *Salmonella* trong một số nguyên liệu thực phẩm, bao gồm thịt lợn, rau và rau thơm [3]. Năm 2024, vụ ngộ độc khác diễn ra tại Long Khánh, tỉnh Đồng Nai, liên quan đến hơn 500 trường hợp nghi ngờ ngộ độc sau khi tiêu thụ bánh mì. Tại đây, các cập nhật chính thức báo cáo rằng 4/8 mẫu thực phẩm, bao gồm pate, thịt lợn chế biến, giò lụa và rau củ muối chua, dương tính với *Salmonella* [4]. Các vụ ngộ độc liên quan đến bánh mì ở miền Nam Việt Nam tiếp tục ghi nhận các mẫu bệnh phẩm lâm sàng dương tính với *Salmonella*, bao gồm các trường hợp tại Vũng Tàu và Đồng Tháp [5], [6]. Các sự cố an toàn thực phẩm trên cho thấy thực trạng nhiễm *Salmonella* trong nguyên liệu và bánh mì thành phẩm là vấn đề đáng quan tâm tại Việt Nam, qua đó đặt ra nhu cầu cấp thiết về các công cụ kiểm nghiệm phòng thí nghiệm không chỉ hỗ trợ phát hiện mầm bệnh mà còn phục vụ công tác truy xuất nguồn gốc trong quá trình điều tra ngộ độc.

Công tác điều tra ngộ độc thực phẩm đòi hỏi sự phối hợp đồng bộ giữa thông tin dịch tễ học, lịch sử tiếp xúc thực phẩm, điều tra truy vết và xét nghiệm kiểm nghiệm. Theo đó, việc phát hiện vi sinh vật gây bệnh trong thực phẩm nghi ngờ là cốt lõi, nhưng thường không đủ để chứng minh liệu các chủng phân lập từ các mẫu khác nhau có liên quan về mặt di truyền hay xuất phát từ một nguồn nhiễm chung. Trung tâm Kiểm soát và Phòng ngừa Dịch bệnh Hoa Kỳ (CDC) mô tả quy trình điều tra ngộ độc thực phẩm là một quá trình nhiều bước bao gồm phát hiện các chùm ca bệnh, xác định tác nhân gây bệnh, điều tra tiếp xúc thực phẩm, hoạt động

truy vết và xét nghiệm mẫu thực phẩm, môi trường hoặc lâm sàng [7]. Mặc dù nuôi cấy vi sinh truyền thống là nền tảng để phát hiện và phân lập *Salmonella* từ mẫu thực phẩm, phương pháp này cho thấy hạn chế đối với truy xuất nguồn gốc do không cung cấp thông tin ở mức độ hệ gen về mối liên hệ giữa các chủng. Do đó, giải trình tự toàn bộ hệ gen (WGS) đã trở thành công cụ quan trọng trong điều tra ngộ độc thực phẩm, cho phép so sánh các chủng vi khuẩn ở độ phân giải cao và có thể giúp phân biệt các chủng liên quan đến vụ ngộ độc với các chủng nền không liên quan. Cơ quan Quản lý Thực phẩm và Dược phẩm Hoa Kỳ (FDA) là tổ chức đã ứng dụng thành công WGS để so sánh các chủng phân lập từ bệnh nhân, thực phẩm và mẫu môi trường, qua đó hỗ trợ xác định ranh giới vụ ngộ độc, quy kết nguồn gốc và điều tra truy vết [8].

Hiện nay, công nghệ giải trình tự hệ gen đoạn dài Nanopore thu hút sự quan tâm ngày càng lớn trong lĩnh vực điều tra mầm bệnh truyền qua thực phẩm nhờ khả năng tạo ra các đoạn đọc dài, hỗ trợ quy trình thao tác nhanh chóng và cung cấp dữ liệu phù hợp cho việc lắp ráp hệ gen, dự đoán kiểu huyết thanh, phát hiện gen kháng kháng sinh, phân tích plasmid và đánh giá mối liên hệ di truyền. Quick và cộng sự đã chứng minh rằng giải trình tự Nanopore thời gian thực cho phép quản lý nhanh các trường hợp ngộ độc do *Salmonella* Enteritidis gây ra tại bệnh viện [9]. Li và cộng sự đã áp dụng cách tiếp cận tương tự vào năm 2022, truy xuất vụ ngộ độc do *Salmonella* Typhimurium gây ra tại Trung Quốc bằng kỹ thuật metagenomic [10]. Taylor và cộng sự đã áp dụng giải trình tự đoạn dài đa mẫu để thu hồi toàn bộ hệ gen và plasmid từ các mầm bệnh thực phẩm, bao gồm *Salmonella* và *Escherichia coli* O157 [11]. Trong một nghiên cứu sử dụng giải trình tự Nanopore đa mẫu, Wu và cộng sự đã dự đoán kiểu huyết thanh và gen kháng kháng sinh trên 69 kiểu huyết thanh *Salmonella* [12]. Rheman và cộng sự đã đặc trưng hóa *E. coli* và *Salmonella* kháng kháng sinh từ hải sản [13]. Bogaerts và cộng sự đã chỉ ra rằng công nghệ Nanopore ONT R10 cải tiến kết hợp với phân tích đa hình nucleotide đơn (SNP) cho kết quả tương đương với Illumina đối với phân tích vụ ngộ độc dựa trên SNP của các mầm bệnh vi khuẩn ngoài *Salmonella* [14]. Các nghiên cứu này cho thấy giải trình tự Nanopore có thể hỗ trợ định danh vi khuẩn nhanh, dự đoán kiểu huyết thanh, phân loại dựa trên định loại trình tự đa locus (MLST), phát hiện gen kháng kháng sinh, đặc trưng hóa plasmid và phân tích mối liên hệ di truyền.

Bánh mì là loại thực phẩm thường được cấu thành từ nhiều thành phần ăn liền khác nhau thay vì là một sản phẩm đồng nhất đơn lẻ. Các thành phần phổ biến bao gồm bánh mì bột mì, một hoặc nhiều loại nhân giàu protein như chả bông (ruốc), chả lụa, xúc xích hoặc thịt nướng, pate hoặc các loại sốt bơ, rau tươi hoặc rau chua (dưa góp) như dưa chuột, cà rốt, su hào, rau mùi, được bán lẻ trực tiếp tại điểm bán đường phố. Do sự nhiễm bẩn của một nguyên liệu thành phần có thể làm mất tính an toàn thực phẩm của sản phẩm hoàn chỉnh, cấu trúc dạng tổ hợp này là nguyên nhân cốt lõi để các cuộc điều tra vụ ngộ độc hiện tại tiến hành thu thập và kiểm nghiệm riêng lẻ từng nguyên liệu bánh mì cùng với các mẫu lâm sàng liên quan nhằm khu trú chính xác nguồn ô nhiễm.

Hiện nay, các công bố về việc ứng dụng giải trình tự WGS Nanopore trong điều tra ngộ độc thực phẩm và truy xuất nguồn gốc hệ gen đối với *Salmonella* tại Việt Nam vẫn còn hạn chế. Vì vậy, nghiên cứu này được thực hiện nhằm bước đầu ứng dụng công nghệ giải trình tự toàn bộ hệ gen Nanopore trong việc đặc trưng hóa các chủng *Salmonella* phân lập từ mẫu thực phẩm liên quan đến một vụ ngộ độc thực phẩm do bánh mì tại Việt Nam, với mục tiêu xác định kiểu huyết thanh, kiểu trình tự đa locus, hồ sơ plasmid, các gen kháng kháng sinh và mối liên hệ di truyền giữa các chủng liên quan đến vụ ngộ độc.

2. VẬT LIỆU VÀ PHƯƠNG PHÁP

2.1. Bối cảnh trường hợp ngộ độc thực phẩm và mô tả mẫu thử nghiệm

Vào tháng 4 năm 2026, một vụ ngộ độc thực phẩm nghi ngờ đã xảy ra tại một xã thuộc tỉnh Nghệ An, Việt Nam. Theo báo cáo điều tra ngộ độc ban đầu, sự việc có liên quan dịch tễ học đến việc tiêu thụ bánh mì và các nguyên liệu bán tại một cơ sở kinh doanh thức ăn đường phố. Các bữa ăn nghi ngờ diễn ra vào ngày 16 và 17 tháng 4 năm 2026. Ca bệnh đầu tiên ghi nhận triệu chứng lúc khoảng 20:00 ngày 16/04/2026 và ca cuối cùng khởi phát triệu chứng lúc khoảng 11:30 ngày 18/04/2026. Tổng số ghi nhận 32 ca mắc, không có trường hợp tử vong. Các triệu chứng lâm sàng chính bao gồm nôn mửa, đau bụng, tiêu chảy nhiều lần, mệt mỏi và sốt ở một số trường hợp với nhiệt độ ghi nhận từ 39 - 40°C.

Đội điều tra địa phương đã lập danh sách tuyến của tất cả 32 trường hợp, ghi nhận ngày/giờ ăn bữa ăn nghi ngờ, ngày/giờ khởi phát triệu chứng và ngày/giờ đến cơ sở y tế của từng trường hợp. Các bệnh nhân sinh sống rải rác ở nhiều xóm/thôn trong xã, phù hợp với nguồn phơi nhiễm gắn liền với một điểm bán chung thay vì một

hộ gia đình đơn lẻ. Toàn bộ 32 ca mắc đều khai thác có ăn cùng một loại bánh mì thập cẩm nhiều nhân, được chuẩn bị với sự kết hợp nhất quán gồm: pate, thịt nướng, chả lụa, xúc xích, ruốc (chà bông), su hào ngâm chua, cà rốt và sốt. Cơ sở kinh doanh liên quan là một quầy hàng đường phố đã bán khoảng 700 suất bánh mì, trong đó có khoảng 500 suất loại thập cẩm, trong giai đoạn từ 16 đến 17 tháng 4.

Các mẫu liên quan đến vụ ngộ độc được thu thập bởi đội điều tra địa phương và chuyển đến Viện Kiểm nghiệm an toàn vệ sinh thực phẩm quốc gia để xét nghiệm theo yêu cầu của cơ quan y tế tỉnh. Các mẫu nguyên liệu thực phẩm được tiếp nhận ngày 19/04/2026, các mẫu sốt và bệnh phẩm phân được tiếp nhận ngày 21/04/2026. Tổng cộng có 16 mẫu liên quan đến vụ ngộ độc được đưa vào nghiên cứu này, bao gồm ruốc, chả lụa, xúc xích, thịt nướng, su hào, cà rốt, sốt mayonnaise và các loại sốt khác. Bệnh phẩm phân cũng được gửi kèm trong điều tra bổ sung. Phân tích hệ gen trong nghiên cứu này tập trung vào các chủng *Salmonella* phân lập từ các mẫu dương tính với *Salmonella*.

Các mẫu thực phẩm gồm ruốc, chả lụa, xúc xích, thịt nướng và rau ngâm chua được thu thập trực tiếp bởi đoàn kiểm tra liên ngành tại cơ sở của người kinh doanh trong đợt kiểm tra tại chỗ ngày 17 - 18/04/2026. Mẫu bánh mì pate được thu thập riêng dưới dạng bánh mì hoàn chỉnh bán cho khách hàng chứ không tách thành các thành phần riêng lẻ. Mẫu sốt và mẫu phân được thu thập trong điều tra theo dõi và chính thức gửi đến Viện Kiểm nghiệm an toàn vệ sinh thực phẩm quốc gia bởi Sở Y tế tỉnh Nghệ An ngày 20 tháng 4 năm 2026. Để bảo đảm tính bảo mật, các định danh cá nhân của người bệnh và tên cơ sở kinh doanh thực phẩm không được đưa vào bản thảo này, thông tin mẫu chỉ được báo cáo bằng mã định danh phòng thí nghiệm.

2.2. Hóa chất và thiết bị

2.2.1. Hóa chất

Phương pháp nuôi cấy vi sinh truyền thống sử dụng môi trường tăng sinh Rappaport-Vassiliadis (RV) soya broth (Oxoid, Hoa Kỳ), canh thang Muller-Kauffmann tetrathionate-novobiocin - MKTTn (Oxoid, Hoa Kỳ), thạch Tryptone soya agar - TSA (Oxoid, Hoa Kỳ), thạch Xylose Lysine Deoxycholate - XLD (Oxoid, Hoa Kỳ), canh thang Brain heart infusion - BHI (Merck, Đức), Butterfield phosphate buffer (Merck, Đức), thạch mannitol Egg yolk polymyxin - MYP (Oxoid, Hoa Kỳ), thạch Modified charcoal cefoperazone deoxycholate agar - mCCDA (Oxoid, Hoa Kỳ), kit miễn dịch enzyme kiểm tra độc tố ruột tịt cầu, thạch Thạch trypton-mật-glucuronid - TBX (Oxoid, Hoa Kỳ) và thạch Baird-Parker - BP (Oxoid, Hoa Kỳ). Bộ kit GeneJET Genomic DNA Purification (Thermo Scientific, Hoa Kỳ) được sử dụng để tách chiết DNA tổng số của các chủng vi khuẩn phân lập.

2.2.2. Thiết bị

Thiết bị giải trình tự hệ gen bao gồm Oxford Nanopore Technologies PromethION 2 Solo (Oxford Nanopore Technologies, Anh), sử dụng Flow Cell R10.4.1 PromethION (FLO-PRO114M). Ngoài ra, quá trình chuẩn bị thư viện sử dụng các thiết bị như Quantus Fluorometer (Promega, Hoa Kỳ), NanoDrop ND-1000 UV-Vis Spectrophotometer (Thermo Fisher Scientific, Hoa Kỳ), tủ ấm Memmert IN110 (Mettler, Đức), tủ an toàn sinh học Esco AC2 (Esco, Singapore), micropipette đa kênh với dải thể tích thích hợp cùng các phụ kiện phòng thí nghiệm chuẩn khác.

2.3. Phương pháp nghiên cứu

2.3.1. Xét nghiệm vi sinh và độc tố vi sinh vật

Các mẫu liên quan đến vụ ngộ độc được xét nghiệm các mối nguy vi sinh và độc tố thường gặp liên quan đến ngộ độc thực phẩm. *Salmonella* spp. được phát hiện theo TCVN 10780-1:2017 và độc tố ruột tịt cầu theo ISO 19020:2017. Đối với các chỉ tiêu khác, *Bacillus cereus* được định lượng theo TCVN 4992:2005, *Campylobacter* spp. theo ISO 10272-2:2017, *E. coli* theo TCVN 7924-2:2008, và *Staphylococcus aureus* theo ISO 6888-1:2021.

2.3.2. Phân lập *Salmonella* và lựa chọn chủng

Đối với mẫu thử nghiệm cho kết quả dương tính với *Salmonella*, một khuẩn lạc nghi ngờ *Salmonella* có hình thái điển hình là khuẩn lạc tâm đen có vùng màu đỏ bao quanh trên thạch XLD được chọn để phân tích tiếp. Các khuẩn lạc được chọn được tinh sạch bằng cách cấy chuyển trên thạch TSA và được khẳng định là *Salmonella* spp. bằng các tiêu chí sinh hóa và huyết thanh quy định trong TCVN 10780-1:2017.

2.3.3. Tách chiết DNA và giải trình tự Nanopore

Mẫu DNA tổng số được tách từ các chủng *Salmonella* được tinh sạch bằng bộ kit GeneJET Genomic DNA Purification theo hướng dẫn của nhà sản xuất. Nồng độ DNA được đo bằng máy đo huỳnh quang Quantus, với nồng độ tối thiểu 36 ng/ μ L. Độ tinh khiết DNA được đánh giá bằng máy quang phổ NanoDrop; các mẫu có tỷ lệ OD_{260/280} từ 1,8 - 2,0 được chọn để giải trình tự hệ gen. Thư viện giải trình tự được chuẩn bị bằng Native Barcoding Kit 96 V14 (SQK-NBD114.96). Các thư viện đã gắn mã vạch được chuẩn hóa và gộp chung trước khi nạp lên flow cell. Quá trình giải trình tự được thực hiện trên thiết bị Oxford Nanopore Technologies PromethION 2 Solo sử dụng Flow Cell R10.4.1 PromethION (FLO-PRO114M). Thao tác mã hóa base và tách mã vạch được thực hiện bằng Dorado v2.0.0 với mô hình siêu chính xác R10.4.1/E8.2 400 bps dna_r10.4.1_e8.2_400bps_sup@v5.2.0 [15]. Các chuỗi đọc dữ liệu thô Nanopore của các chủng *Salmonella* được lưu trữ tại cơ sở dữ liệu NCBI Sequence Read Archive (SRA) với BioProject PRJNA1530768.

2.3.4. Phân tích tin sinh học

Dữ liệu đọc thô được mã hóa base và tách mã vạch bằng Dorado [15], các đoạn đọc theo từng mã vạch được lọc chất lượng và lắp ráp *de novo* bằng Flye [16] trong quy trình EPI2ME Labs wf-bacterial-genomes [17], sau đó được đánh bóng bằng Medaka [18], các bộ lắp ráp đã đánh bóng trải qua xác nhận phân loại và định týp huyết thanh/MLST *in silico* trong cùng quy trình, chú thích hệ gen được thực hiện bằng Bakta [19], các gen kháng kháng sinh thu được và đột biến điểm liên quan đến kháng được dự đoán bằng mô-đun ResFinder/PointFinder [20], các replicon plasmid, tính di động và gen AMR mang trên plasmid được xác định bằng MOB-suite [21], phân tích SNP lõi hệ gen, căn chỉnh và suy luận phát sinh loài được thực hiện bằng CSI Phylogeny [22].

Về chi tiết, các đoạn đọc đã được mã hóa base từ mỗi mã vạch được phân tích riêng bằng quy trình EPI2ME Labs wf-bacterial-genomes [17]. Chất lượng đoạn đọc được đánh giá bằng số lượng đoạn đọc, tổng số base, độ dài đoạn đọc trung bình và trung vị, giá trị N50, điểm Q trung bình và độ bao phủ ước tính do quy trình báo cáo. Các đoạn đọc vượt ngưỡng chất lượng Q > 10 được giữ lại cho phân tích tiếp theo. Lắp ráp hệ gen *de novo* được thực hiện trong quy trình bằng Flye [16], tiếp theo là đánh bóng bằng Medaka [18] để tạo ra các file FASTA contig cuối cùng. Quy trình được chạy bằng EPI2ME Labs wf-bacterial-genomes v2.0.2, tích hợp Flye v2.9.6 cho lắp ráp *de novo* và mô hình Medaka với tham số mặc định. Xác nhận loài và sàng lọc nhiễm bẩn được thực hiện bằng bước phân loại trong quy trình. Chỉ các bộ lắp ráp được phân loại rõ ràng là *Salmonella enterica* và không có bằng chứng về trình tự nhiễm bẩn đáng kể mới được giữ lại cho phân tích tiếp theo.

Chất lượng lắp ráp được đánh giá bằng số lượng contig, tổng chiều dài lắp ráp, contig lớn nhất, giá trị N50, hàm lượng GC và độ bao phủ lắp ráp trung bình ước tính. Tổng chiều dài lắp ráp được tính là tổng tất cả contig được giữ lại, trong khi độ bao phủ lắp ráp trung bình được ước tính từ các giá trị độ bao phủ cấp contig bằng cách tiếp cận có trọng số theo chiều dài. Các bộ lắp ráp được chấp nhận đặc trưng cho hệ gen khi chúng thể hiện độ bao phủ lắp ráp trung bình đầy đủ, ưu tiên ít nhất 30X, phân mảnh lắp ráp thấp, N50 contig lớn hơn 1 Mb, tổng chiều dài lắp ráp phù hợp với kích thước hệ gen kỳ vọng khoảng 4,6 - 4,9 Mb, và hàm lượng GC xấp xỉ 52%. Các bộ lắp ráp có độ bao phủ trung bình thấp hơn, số contig cao hơn hoặc giá trị N50 contig thấp hơn vẫn được giữ lại, nhưng kết quả tiếp theo của chúng khi so sánh dựa trên SNP sẽ được diễn giải thận trọng.

Các file FASTA contig sinh ra từ mỗi mã vạch được sử dụng cho các phân tích dựa trên hệ gen tiếp theo trong quy trình EPI2ME Labs [17]. Phân tích MLST và dự đoán kiểu huyết thanh *in silico* được thực hiện bằng các đầu ra tương ứng từ quy trình [17]. Chú thích hệ gen được thực hiện bằng Bakta [19]. Các gen liên quan đến kháng kháng sinh và đột biến điểm liên quan đến kháng được dự đoán bằng mô-đun ResFinder/PointFinder tích hợp trong quy trình EPI2ME Labs wf-bacterial-genomes [17], [20]. Các contig liên quan đến plasmid, replicon plasmid và yếu tố di động của plasmid được xác định bằng MOB-suite [21]. Việc xác định các gen kháng kháng sinh và đột biến điểm được thực hiện khi không chỉ định cơ sở dữ liệu tùy chỉnh, mô-đun quy trình này áp dụng cơ sở dữ liệu ResFinder v2.6.0 và cơ sở dữ liệu PointFinder mở rộng, cùng với ngưỡng phát hiện mặc định của công cụ là độ đồng nhất nucleotide tối thiểu 90% và độ bao phủ tối thiểu 60%. ResFinder/PointFinder phát hiện các gen và đột biến được liệt kê trong cơ sở dữ liệu tham chiếu đã được thẩm định của chúng và do đó có thể không nắm bắt các yếu tố kháng mới, phân kỳ cao hoặc chưa được đặc trưng. Hồ sơ kháng kháng sinh được diễn giải là tập hợp các gen và đột biến kháng được công nhận hiện tại mà quy trình phát hiện được, không đặc trưng hóa đầy đủ resistome của các chủng.

Để cho phép tái lập độc lập phân tích tin sinh học, các phiên bản công cụ và cơ sở dữ liệu sau đã được sử dụng: Bakta v1.12.0 với cơ sở dữ liệu Bakta đi kèm v5.1, ResFinder và PointFinder v4.7.2, như được đóng gói với quy trình EPI2ME wf-bacterial-genomes v2.0.2, cùng với cơ sở dữ liệu tham chiếu AMR/đột biến điểm tương ứng, và MOB-suite v3.1.9 với plasmid tham chiếu đi kèm.

Đối với phân tích mối liên hệ dựa trên SNP, các file FASTA contig cuối cùng được phân tích bằng nền tảng CSIPhylogeny của Center for Genomic Epidemiology [22]. CSIPhylogeny được sử dụng để gọi và lọc SNP, xác nhận các vị trí SNP, tạo căn chỉnh SNP chất lượng cao nối tiếp, tính khoảng cách SNP từng cặp và suy luận cây phát sinh loài dựa trên SNP. Kết quả QC cấp lắp ráp được xem xét cùng với kiểu huyết thanh, MLST, hồ sơ plasmid, hồ sơ kháng kháng sinh dự đoán và dữ liệu khoảng cách SNP để đánh giá mối liên hệ di truyền giữa các chủng phân lập từ các nguồn mẫu khác nhau. Việc phân tích SNP trong CSIPhylogeny được thực hiện bằng CSI Phylogeny v1.4 (Center for Genomic Epidemiology) với các tham số chất lượng mặc định của nền tảng: độ sâu tối thiểu tại vị trí SNP $\geq 10X$, độ sâu tương đối tối thiểu $\geq 10\%$, khoảng cách tối thiểu giữa các SNP ≥ 10 bp, chất lượng SNP tối thiểu ≥ 30 , Z-score tối thiểu 1,96, chất lượng ánh xạ đoạn đọc tối thiểu ≥ 25 , khoảng cách tối thiểu đến khoảng trống/đầu mút trình tự ≥ 10 bp. Bộ lắp ráp của chủng S03 được sử dụng làm hệ gen tham chiếu cho gọi SNP, được chọn dựa trên chất lượng lắp ráp: 4 contig, N50 contig khoảng 4,68 Mb, độ bao phủ ước tính khoảng 48X, và tính đại diện của kiểu gen ưu thế thu được trong vụ ngộ độc này [23]

3. KẾT QUẢ VÀ THẢO LUẬN

3.1. Các mẫu liên quan đến vụ ngộ độc và kết quả xét nghiệm truyền thống

Tổng số 16 mẫu liên quan đến vụ ngộ độc được phân tích và phân loại theo loại mẫu (**Bảng 1**). Dữ liệu về ngày nhận mẫu thể hiện trong **Bảng 1** phản ánh thứ tự mẫu đến phòng thí nghiệm và không đại diện cho nhóm có ý nghĩa sinh học hay dịch tễ. Không có so sánh nào giữa thời gian nhận mẫu được dự định và thời gian thực hiện. Các mẫu gửi gồm nguyên liệu liên quan đến bánh mì, sốt và mẫu phân thu thập trong điều tra vụ ngộ độc.

Bảng 1. Tóm tắt thông tin của các mẫu liên quan đến vụ ngộ độc và kết quả phân tích vi sinh truyền thống

Nhóm mẫu	Mô tả mẫu	Ngày nhận	<i>Salmonella</i> spp.	Phát hiện đáng chú ý khác
Nguyên liệu thực phẩm	Chà bông	19/4/2026	Dương tính	<i>B. cereus</i> : $4,0 \times 10^2$ CFU/g
	Giò lụa	19/4/2026	Âm tính	Không
	Xúc xích	19/4/2026	Dương tính	Không
	Thịt nướng	19/4/2026	Dương tính	Không
	Rau củ muối	19/4/2026	Dương tính	Không
	Sốt cà chua	19/4/2026	Âm tính	Không
	Bánh mì pate	19/4/2026	Dương tính	<i>E. coli</i> : $9,0 \times 10^1$ CFU/g
	Sốt mayonnaise 001	21/4/2026	Âm tính	Không
	Sốt mayonnaise 002	21/4/2026	Âm tính	Không
	Sốt Sài Gòn 003	21/4/2026	Âm tính	Không
Mẫu bệnh phẩm	Sốt Cốm 004	21/4/2026	Âm tính	Không
	Mẫu phân 1	21/4/2026	Dương tính	Không
	Mẫu phân 2	21/4/2026	Dương tính	Không
	Mẫu phân 3	21/4/2026	Dương tính	Không
	Mẫu phân 4	21/4/2026	Dương tính	Không
	Mẫu phân 5	21/4/2026	Dương tính	<i>S. aureus</i> : $2,1 \times 10^5$ CFU/g, độc tố ruột tụ cầu âm tính

Bảng 1 cho thấy *Salmonella* spp. được phát hiện trong 10/16 mẫu. Trong số 11 mẫu nguyên liệu thực phẩm, 5/11 phát hiện sự có mặt của *Salmonella* spp., bao gồm chà bông, xúc xích, thịt nướng, rau củ muối và bánh mì pate. Các mẫu nguyên liệu còn lại, bao gồm cả bốn mẫu sốt, đều cho kết quả không phát hiện vi khuẩn. Tất cả 5 mẫu bệnh phẩm lâm sàng đều phát hiện sự có mặt của *Salmonella* spp. Phần lớn các chỉ tiêu vi sinh và độc tố khác cho kết quả âm tính hoặc phát hiện ở mức thấp. *B. cereus* được phát hiện với nồng độ $4,0 \times 10^2$ CFU/g trong mẫu chà bông, *E. coli* với nồng độ $9,0 \times 10^1$ CFU/g được phát hiện trong mẫu bánh mì pate và *S.*

aureus được tìm thấy trong mẫu phân của bệnh nhân thu thập. Phân tích độc tố ruột tự cầu cho kết quả âm tính trong tất cả các mẫu được thử nghiệm. Kết quả trên chỉ ra rằng *Salmonella* spp. là mầm bệnh được phát hiện nhất quán nhất trong các mẫu liên quan đến vụ ngộ độc thực phẩm do bánh mì.

3.2. Kết quả giải trình tự Nanopore và chất lượng lắp ráp hệ gen

Các chủng *Salmonella* được khẳng định từ 10 mẫu dương tính với *Salmonella* được tách chiết DNA tổng số và chuẩn bị thư viện cho quá trình giải trình tự. Thông tin về chất lượng DNA đầu vào cho mỗi chủng được tóm tắt trong **Bảng 2**.

Bảng 2. Thông tin mẫu đầu vào, kết quả giải trình tự Nanopore và chất lượng lắp ráp hệ gen

Mẫu	Tên	SRA	Nồng độ (ng/μL)	OD _{260/280}	Đoạn đọc	Base (Gb)	Q _{TB}	N50 (kb)	Độ bao phủ (X)	Contig	Kích thước (Mb)	N50 contig (Mb)
27/1	S01	SRR40717781	91,3	1,83	24.519	0,158	20,2	11,2	32,1	3	4,77	4,68
27/3	S02	SRR40717780	180,5	1,89	120.601	1,3	22,5	21,1	273,6	4	4,78	4,68
27/4	S03	SRR40717779	88,0	1,86	36.218	0,236	20,3	11,1	48,2	4	4,78	4,68
27/5	S04	SRR40717778	77,0	1,83	36.849	0,217	20,1	11,0	44,1	3	4,77	4,68
27/7	S05	SRR40717777	108,0	1,84	28.927	0,193	20,3	11,9	39,2	4	4,78	4,68
28/5	S06	SRR40717776	91,8	1,80	30.508	0,207	20,4	11,4	42,1	3	4,77	4,68
28/6	S07	SRR40717775	175,6	1,79	265.458	1,3	20,8	8,41	305	3	4,77	4,68
28/7	S08	SRR40717774	85,3	1,64	10.635	0,0764	20,4	11,9	14,7	6	4,88	4,02
28/8	S09	SRR40717773	173,3	1,89	30.539	0,19	20,3	10,8	40,1	3	4,77	4,68
28/9	S10	SRR40717772	62,4	1,89	55.788	0,326	20,2	9,9	65,3	3	4,77	4,68

Mẫu DNA tách từ 10 chủng *Salmonella* được xác nhận có nồng độ từ 62,4 đến 180,5 ng/μL. Hầu hết các chủng có giá trị OD_{260/280} trong khoảng 1,8 - 2,0, ngoại trừ S08 có giá trị thấp hơn là 1,64. Giải trình tự Nanopore tạo ra từ 10.635 đến 265.458 đoạn đọc mỗi chủng, tương ứng với 0,0764 đến 1,3 Gb dữ liệu giải trình tự. Điểm Q trung bình dao động từ 20,1 đến 22,5, cho thấy chất lượng đoạn đọc nhất quán giữa các chủng. Hầu hết các bộ lắp ráp thể hiện chất lượng lắp ráp hệ gen ổn định, với kích thước lắp ráp khoảng 4,77 - 4,78 Mb, 3 - 4 contig và giá trị N50 contig khoảng 4,68 Mb. Những kết quả này cho thấy hệ gen thu được tương đối hoàn chỉnh và không bị phân mảnh nhiều. Tuy nhiên, chủng S08 thể hiện độ sâu giải trình tự thấp hơn, kích thước lắp ráp lớn hơn, nhiều contig hơn và N50 contig thấp hơn so với các chủng khác.

3.3. Định danh và xác định týp huyết thanh *Salmonella*

Kết quả định týp *in silico* cho thấy tất cả 10 chủng đều thuộc loài *Salmonella enterica* serovar Enteritidis, với công thức kháng nguyên 9:g,m:- và kiểu trình tự multilocus là ST11. Hồ sơ này được ghi nhận nhất quán ở các chủng phân lập từ cả mẫu thực phẩm và mẫu bệnh phẩm. Các giá trị định danh trung bình toàn cục (ANI) dao động từ 0,991 đến 0,992, xác nhận tất cả các chủng đều thuộc loài *Salmonella enterica* ở cấp độ loài; tuy nhiên, chỉ số này chưa cung cấp bằng chứng về mối liên hệ ở cấp độ chủng. Kiểu huyết thanh và hồ sơ MLST đồng nhất giữa các chủng là cơ sở cốt lõi để suy luận mối liên hệ di truyền chặt chẽ của chúng trong bối cảnh điều tra vụ ngộ độc (**Bảng 3**).

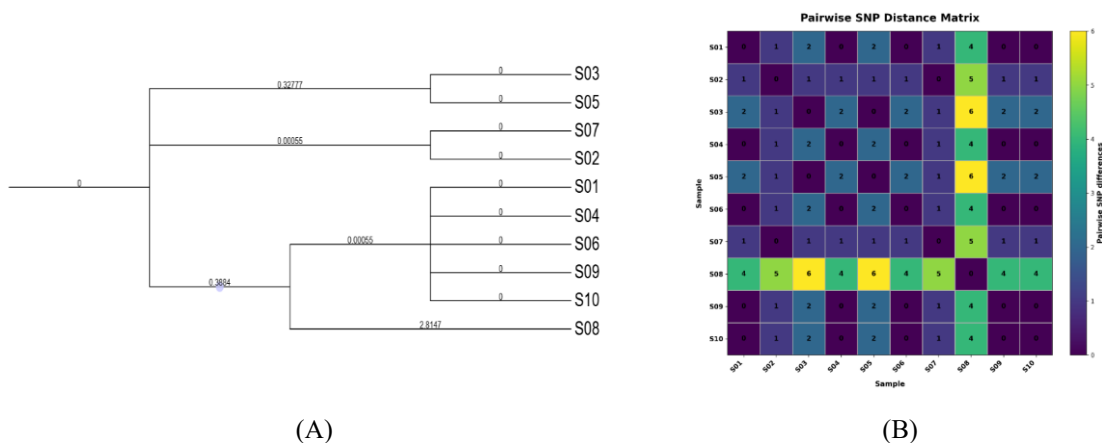
Ma trận khoảng cách SNP từng cặp cho thấy sự khác biệt di truyền rất thấp giữa 10 chủng *Salmonella*, dao động từ 0 đến 6 SNP (**Hình 1**). Một số chủng có hồ sơ SNP hoàn toàn giống hệt nhau, bao gồm các nhóm: S01/S04/S06/S09/S10, S03/S05 và S02/S07. Khi đối chiếu với nguồn gốc mẫu, *Salmonella* trong mẫu chà bông (S01) và rau củ muối (S04) không thể phân biệt về mặt di truyền với ba chủng phân lập từ mẫu phân (S06, S09 và S10), trong khi mẫu xúc xích (S02) không thể phân biệt với chủng phân lập từ mẫu phân thứ tư (S07). Vì có nhiều hơn một mẫu thực phẩm có đặc điểm hệ gen không thể phân biệt với các chủng phân lập từ bệnh phẩm, dữ liệu SNP độc lập chưa đủ cơ sở để xác định chính xác một nguyên liệu duy nhất là nguồn lây

nhễm. Khoảng cách SNP từng cặp giữa hầu hết các chủng còn lại cũng rất thấp, chỉ từ 1 đến 2 SNP, cho thấy mức độ liên hệ di truyền cực kỳ cao. Riêng chủng S08 có khoảng cách SNP từng cặp cao nhất so với các chủng còn lại (dao động từ 4 đến 6 SNP) và nằm trên một nhánh tách biệt trong cây phát sinh chủng loại dựa trên SNP. Tuy nhiên, mức độ khác biệt này vẫn rất nhỏ và chủng S08 vẫn mang chung kiểu huyết thanh cùng hồ sơ MLST với các chủng khác. Do chủng S08 có độ sâu giải trình tự thấp hơn và mức độ lắp ráp phân mảnh hơn, kết quả đối với chủng này cần được diễn giải một cách thận trọng. Nhìn chung, ma trận SNP và cây phát sinh chủng loại hoàn toàn phù hợp với kết quả định typ huyết thanh và MLST trình bày trong **Bảng 3**, qua đó cho thấy các chủng phân lập từ thực phẩm và từ phân đều thuộc cùng một cụm *Salmonella* Enteritidis ST11 có mối liên hệ chặt chẽ. Những phát hiện này củng cố vững chắc giả thuyết cho rằng các chủng này bắt nguồn từ cùng một nguồn lây nhiễm trong vụ ngộ độc.

Bảng 3. Dự đoán kiểu huyết thanh *in silico* và định typ trình tự đa locus (MLST) của các chủng *Salmonella*

Tên	Mẫu	Kiểu huyết thanh	Công thức kháng nguyên	Kiểu trình tự (ST)	ANI	Phân đoạn truy vấn
S01	Chà bông	Enteritidis	9:g,m:-	11	0,992	0,651
S02	Xúc xích	Enteritidis	9:g,m:-	11	0,992	0,651
S03	Thịt nướng	Enteritidis	9:g,m:-	11	0,992	0,651
S04	Rau củ muối	Enteritidis	9:g,m:-	11	0,992	0,651
S05	Bánh mì pâté	Enteritidis	9:g,m:-	11	0,992	0,651
S06	Mẫu phân 1	Enteritidis	9:g,m:-	11	0,992	0,651
S07	Mẫu phân 2	Enteritidis	9:g,m:-	11	0,992	0,651
S08	Mẫu phân 3	Enteritidis	9:g,m:-	11	0,991	0,632
S09	Mẫu phân 4	Enteritidis	9:g,m:-	11	0,992	0,651
S10	Mẫu phân 5	Enteritidis	9:g,m:-	11	0,991	0,632

ANI: giá trị định danh trung bình toàn cục



Hình 1. Cây phát sinh loài SNP lõi hệ gen (A) và ma trận khoảng cách SNP (B)

3.4. Hồ sơ plasmid

Tất cả 10 chủng *Salmonella* đều được phát hiện mang cùng một plasmid chính Plasmid_AB461, với các replicon IncFIB, IncFII và IncX1. Plasmid này được dự đoán là hợp bào và mang các dấu ấn đặc hiệu cho tính di động gồm dạng chuyển động sử dụng hệ thống lông bề mặt (MOBP) và dạng chuyển động lướt (MPF_F). Kích thước dự đoán của plasmid khá tương đồng giữa các chủng, dao động trong khoảng từ 93,6 đến 98,7 kb.

Các gen AMR liên quan đến plasmid đều được phát hiện nằm trên Plasmid_AB461 ở tất cả các chủng, bao gồm *blaTEM-1B*, *tet(A)*, *aph(6)-Id*, *aph(3'')-Ib* và *sul2*. Các gen này chịu trách nhiệm cho khả năng kháng các nhóm kháng sinh như beta-lactam, tetracycline, aminoglycoside và sulfonamide. Việc cùng xuất hiện của một loại plasmid mang gen kháng kháng sinh ở cả các chủng phân lập từ thực phẩm và từ bệnh phẩm đã cung cấp thêm bằng chứng củng cố mối quan hệ chặt chẽ giữa các chủng trong vụ ngộ độc.

Bên cạnh đó, chủng S08 mang thêm một plasmid dự đoán bổ sung là Plasmid_AA474 với thành phần replicon IncI-gamma/K1, tuy nhiên không phát hiện gen kháng kháng sinh nào trên plasmid bổ sung này. Do S08 có độ sâu giải trình tự thấp hơn và mức độ lắp ráp phân mảnh hơn so với các chủng còn lại, kết quả phát hiện plasmid bổ sung này cần được diễn giải một cách thận trọng. Hồ sơ chung của Plasmid_AB461 hoàn toàn phù hợp với các kết quả phân tích kiểu huyết thanh, MLST và SNP. Tuy nhiên, các gen kháng kháng sinh nằm trên Plasmid_AB461 (*blaTEM-1B*, *tet(A)*, *aph(6)-Id*, *aph(3'')-Ib* và *sul2*) thường là những gen phổ biến trên các plasmid hợp bào đang lưu hành rộng rãi ở vi khuẩn *Salmonella* cũng như họ vi khuẩn đường ruột *Enterobacteriaceae*. Do đó, nếu chỉ dựa vào sự hiện diện của các gen này thì chưa đủ cơ sở để khẳng định nguồn lây chung. Hồ sơ plasmid và gen kháng kháng sinh chung tốt nhất nên được xem là bằng chứng hỗ trợ, chứ không phải là bằng chứng độc lập để thiết lập mối liên hệ di truyền cốt lõi vốn đã được chứng minh rõ qua MLST và phân tích SNP. Ngoài ra, nghiên cứu này chưa thực hiện phân tích so sánh trình tự ở cấp độ plasmid như căn chỉnh từng cặp hay ước tính khoảng cách dựa trên thuật toán Mash, việc thực hiện các phân tích này trong tương lai sẽ cung cấp đánh giá trực tiếp và chính xác hơn về mối liên hệ giữa các plasmid thay vì chỉ dựa vào tóm tắt replicon và gen kháng kháng sinh trình bày trong **Bảng 4**.

Bảng 4. Thông tin replicon plasmid dự đoán, tính di động và gen AMR liên quan đến plasmid

Tên	ID plasmid	Replicon plasmid	Tính di động dự đoán/kiểu MOB	Kích thước contig plasmid	Gen AMR phát hiện trên contig plasmid
S01	Plasmid_AB461	IncFIB, IncFII, IncX1	Hợp bào, MOBP, MPF_F	93.665 bp	<i>blaTEM-1B</i> , <i>tet(A)</i> , <i>aph(6)-Id</i> , <i>aph(3'')-Ib</i> , <i>sul2</i>
S02	Plasmid_AB461	IncFIB, IncFII, IncX1	Hợp bào, MOBP, MPF_F	98.673 bp	<i>blaTEM-1B</i> , <i>tet(A)</i> , <i>aph(6)-Id</i> , <i>aph(3'')-Ib</i> , <i>sul2</i>
S03	Plasmid_AB461	IncFIB, IncFII, IncX1	Hợp bào, MOBP, MPF_F	98.612 bp	<i>blaTEM-1B</i> , <i>tet(A)</i> , <i>aph(6)-Id</i> , <i>aph(3'')-Ib</i> , <i>sul2</i>
S04	Plasmid_AB461	IncFIB, IncFII, IncX1	Hợp bào, MOBP, MPF_F	93.665 bp	<i>blaTEM-1B</i> , <i>tet(A)</i> , <i>aph(6)-Id</i> , <i>aph(3'')-Ib</i> , <i>sul2</i>
S05	Plasmid_AB461	IncFIB, IncFII, IncX1	Hợp bào, MOBP, MPF_F	98.608 bp	<i>blaTEM-1B</i> , <i>tet(A)</i> , <i>aph(6)-Id</i> , <i>aph(3'')-Ib</i> , <i>sul2</i>
S06	Plasmid_AB461	IncFIB, IncFII, IncX1	Hợp bào, MOBP, MPF_F	93.665 bp	<i>blaTEM-1B</i> , <i>tet(A)</i> , <i>aph(6)-Id</i> , <i>aph(3'')-Ib</i> , <i>sul2</i>
S07	Plasmid_AB461	IncFIB, IncFII, IncX1	Hợp bào, MOBP, MPF_F	93.665 bp	<i>blaTEM-1B</i> , <i>tet(A)</i> , <i>aph(6)-Id</i> , <i>aph(3'')-Ib</i> , <i>sul2</i>
S08	Plasmid_AA474	IncI-gamma/K1	Hợp bào, MOBP, MPF_I	94.458 bp	Không phát hiện
	Plasmid_AB461	IncFIB, IncFII, IncX1	Hợp bào, MOBP, MPF_F	93.667 bp	<i>blaTEM-1B</i> , <i>tet(A)</i> , <i>aph(6)-Id</i> , <i>aph(3'')-Ib</i> , <i>sul2</i>
S09	Plasmid_AB461	IncFIB, IncFII, IncX1	Hợp bào, MOBP, MPF_F	93.666 bp	<i>blaTEM-1B</i> , <i>tet(A)</i> , <i>aph(6)-Id</i> , <i>aph(3'')-Ib</i> , <i>sul2</i>
S10	Plasmid_AB461	IncFIB, IncFII, IncX1	Hợp bào, MOBP, MPF_F	93.665 bp	<i>blaTEM-1B</i> , <i>tet(A)</i> , <i>aph(6)-Id</i> , <i>aph(3'')-Ib</i> , <i>sul2</i>

3.5. Kiểu gen kháng kháng sinh

Kết quả dự đoán đặc điểm kháng kháng sinh được tóm tắt theo dấu hiệu kháng trong **Bảng 5**. Tất cả 10 chủng *Salmonella* đều mang cùng các gen và đột biến liên quan đến khả năng kháng kháng sinh, bao gồm *aac(6')-Iaa*, *aph(3'')-Ib*, *aph(6)-Id*, *blaTEM-1B*, *sul2*, *tet(A)* và *gyrA p.D87Y*. Các marker này liên quan đến khả năng kháng với các nhóm kháng sinh aminoglycoside, beta-lactam, sulfonamide, tetracycline và quinolone. Hầu hết các gen kháng thu được, bao gồm *aph(3'')-Ib*, *aph(6)-Id*, *blaTEM-1B*, *sul2* và *tet(A)*, nằm trên plasmid dự đoán chung Plasmid_AB461. Ngược lại, *aac(6')-Iaa* và đột biến liên quan đến kháng quinolone *gyrA p.D87Y* nằm trên nhiễm sắc thể. S08 cũng mang một plasmid IncI-gamma/K1 dự đoán bổ sung được báo cáo trong **Bảng 4**, nhưng không phát hiện gen AMR trên plasmid bổ sung này.

Bảng 5. Các dấu hiệu liên quan đến kháng kháng sinh, kiểu hình đề kháng dự đoán và vị trí trên hệ gen

Dấu hiệu kháng kháng sinh	Loại dấu hiệu	Nhóm kháng sinh	Kiểu hình đề kháng dự đoán	Vị trí hệ gen dự đoán	Chủng phát hiện
<i>aac(6')-Iaa</i>	Gen kháng	Aminoglycoside	Amikacin, tobramycin	Nhiễm sắc thể	S01 - S10
<i>aph(3'')-Ib</i>	Gen kháng	Aminoglycoside	Streptomycin	Plasmid_AB461	S01 - S10
<i>aph(6)-Id</i>	Gen kháng	Aminoglycoside	Streptomycin	Plasmid_AB461	S01 - S10
<i>blaTEM-1B</i>	Gen kháng	Beta-lactam	Amoxicillin, ampicillin, cephalothin, piperacillin, ticarcillin	Plasmid_AB461	S01 - S10
<i>sul2</i>	Gen kháng	Sulfonamide	Sulfamethoxazole	Plasmid_AB461	S01 - S10
<i>tet(A)</i>	Gen kháng	Tetracycline	Doxycycline, tetracycline	Plasmid_AB461	S01 - S10
<i>gyrA p.D87Y</i>	Đột biến điểm	Quinolone	Ciprofloxacin, acid nalidixic	Nhiễm sắc thể	S01 - S10

Hồ sơ dấu hiệu kháng kháng sinh hoàn toàn đồng nhất được ghi nhận ở tất cả các chủng phân lập từ thực phẩm và bệnh phẩm. Điều này hoàn toàn phù hợp với kết quả phân tích về plasmid, kiểu huyết thanh, MLST và SNP, qua đó củng cố vững chắc mối liên hệ di truyền chặt chẽ giữa các chủng trong vụ ngộ độc.

3.6. Bằng chứng hệ gen tích hợp cho chùm ca bệnh liên quan đến vụ ngộ độc

Kết quả vi sinh truyền thống cho thấy *Salmonella* spp. là mầm bệnh được phát hiện nhất quán nhất trong các mẫu liên quan đến vụ ngộ độc. *Salmonella* được phát hiện trong 10/16 mẫu, bao gồm năm mẫu thực phẩm và năm mẫu bệnh phẩm. Ngược lại, phần lớn các chỉ tiêu vi sinh và độc tố vi sinh vật khác cho kết quả âm tính hoặc phát hiện ở mức thấp. Mặc dù *S. aureus* xuất hiện ở mức cao trong một mẫu, kết quả xét nghiệm độc tố ruột âm tính, do đó không đủ cơ sở kết luận đây là vụ ngộ độc do tụ cầu.

Giải trình tự toàn bộ hệ gen cung cấp độ phân giải bổ sung vượt trội so với phương pháp nuôi cấy truyền thống. Tất cả 10 chủng *Salmonella* đều được phân loại là *Salmonella enterica* serovar Enteritidis, với công thức kháng nguyên 9:g,m:- và kiểu trình tự ST11. Phân tích dựa trên SNP cho thấy khoảng cách SNP từng cặp rất thấp giữa các chủng, dao động từ 0 đến 6 SNP, chỉ ra rằng các chủng phân lập từ thực phẩm và từ phân có mối liên hệ dịch tễ rất cao. Riêng chủng S08 có khoảng cách SNP cao hơn và chất lượng giải trình tự/lắp ráp kém hơn các chủng khác. Tuy nhiên, S08 vẫn chia sẻ chung kiểu huyết thanh, hồ sơ MLST và các đặc điểm hệ gen cốt lõi với các chủng còn lại.

Tất cả các chủng đều mang chung một plasmid chính được dự đoán là Plasmid_AB461, với các thành phần replicon gồm IncFIB, IncFII và IncX1. Các gen kháng kháng sinh liên quan đến plasmid, bao gồm *blaTEM-1B*, *tet(A)*, *aph(6)-Id*, *aph(3'')-Ib* và *sul2*, đều được phát hiện ở tất cả các chủng. Ngoài ra, nghiên cứu phát hiện tất cả các chủng cùng mang chung các dấu hiệu kháng kháng sinh trên nhiễm sắc thể, bao gồm *aac(6')-Iaa* và

đột biến *gyrA* p.D87Y. Chủng S08 được dự đoán mang thêm một plasmid IncI-gamma/K1 không mang bất kỳ gen kháng kháng sinh nào.

Nghiên cứu này đóng góp một số yếu tố mới cho công tác điều tra ngộ độc thực phẩm tại Việt Nam. Thứ nhất, đây là một trong những nghiên cứu đầu tiên ứng dụng công nghệ giải trình tự toàn bộ hệ gen đoạn dài Nanopore để đặc trưng hóa các chủng *Salmonella* từ một vụ ngộ độc liên quan đến bánh mì tại Việt Nam. Các vụ ngộ độc trong nước trước đây chủ yếu dựa vào nuôi cấy truyền thống mà không có sự đặc trưng hóa ở cấp độ hệ gen hoặc truy xuất nguồn gốc. Thứ hai, nghiên cứu đã tích hợp thành công việc đặc trưng hóa hệ gen bằng WGS đoạn dài (gồm định tít huyết thanh, MLST, hồ sơ plasmid, phát hiện gen kháng kháng sinh và phân tích mối liên hệ dựa trên SNP) với bằng chứng dịch tễ học thực địa và thanh tra an toàn thực phẩm. Đây là cách tiếp cận chưa từng được báo cáo trước đây đối với các vụ ngộ độc được ghi nhận tại Việt Nam. Các báo cáo trong nước đã công bố [3, 4, 5, 6] chỉ trình bày kết quả dựa trên nuôi cấy và/hoặc dịch tễ học thông thường, trong khi các nghiên cứu *Salmonella* dựa trên Nanopore trước đây [9, 10, 11, 12, 14] chủ yếu được thực hiện trong các nghiên cứu quốc tế và thiếu sự xác nhận dịch tễ học tại hiện trường. Thứ ba, nghiên cứu này chứng minh tính khả thi trong thực tiễn của việc ứng dụng giải trình tự Nanopore vào quy trình ứng phó khẩn cấp của phòng thí nghiệm tham chiếu an toàn thực phẩm quốc gia, qua đó phác thảo lộ trình khả thi để tích hợp giám sát hệ gen đoạn dài vào công tác điều tra ngộ độc thường quy tại Việt Nam.

Trong nghiên cứu này, giải trình tự Nanopore đã giúp nhanh chóng xác định một cụm dòng di truyền vô tính (*clonal cluster*) *Salmonella* Enteritidis ST11 và làm rõ cấu trúc Plasmid_AB461 đa replicon trên tất cả các chủng bằng dữ liệu đọc đoạn dài, hoàn toàn phù hợp với năng lực phân giải plasmid đã được ghi nhận của giải trình tự đoạn dài đối với mầm bệnh-thực phẩm [11]. Những phát hiện này cũng tương đồng với các báo cáo trước đây về tính khả thi của giải trình tự Nanopore trong việc đặc trưng hóa kịp thời các vụ ngộ độc do *Salmonella*, bao gồm vụ ngộ độc *Salmonella* Enteritidis liên quan đến bệnh viện được điều tra bằng giải trình tự Nanopore thời gian thực tại Vương quốc Anh [9] và các đánh giá trên nhiều mẫu với đa dạng kiểu huyết thanh [12]. Quá trình giải trình tự trong nghiên cứu này cũng hoàn thành với thời gian quay vòng nhanh và yêu cầu hạ tầng khiêm tốn [24], phù hợp với các báo cáo cho thấy cải tiến về hóa học flow cell và công nghệ mã hóa base đã thu hẹp đáng kể khoảng cách độ chính xác giữa giải trình tự Nanopore và Illumina trong phân tích dựa trên SNP [14]. Tuy nhiên, nghiên cứu không bao gồm tập dữ liệu giải trình tự đoạn ngắn song song, do đó, các lỗi đặc thù của công nghệ Nanopore còn lại có thể ảnh hưởng đến khoảng cách SNP từng cặp nhỏ nhất (từ 1 đến 2 SNP) và chưa thể loại trừ hoàn toàn.

Bản thân mối liên hệ di truyền không đồng nghĩa với xác nhận tuyệt đối về một nguồn lây chung. Theo hướng dẫn quốc tế và kinh nghiệm từ các cuộc điều tra ngộ độc thực phẩm dựa trên WGS trước đây, dữ liệu hệ gen cần phải được diễn giải song song với bằng chứng dịch tễ, môi trường và chuỗi cung ứng, không nên coi là bằng chứng độc lập duy nhất để xác định nguồn chung [25]. Khoảng cách SNP từng cặp quan sát được trong nghiên cứu này (từ 0 đến 6 SNP) nằm hoàn toàn trong khoảng giá trị thường được xem là chỉ báo cụm hệ gen trong các điều tra ngộ độc *Salmonella* trên thế giới (thường áp dụng ngưỡng cụm khoảng 5 - 10 SNP tùy thuộc vào kiểu huyết thanh và bối cảnh dịch tễ) [23]. Công tác điều tra hiện trường độc lập đã xác định một người bán thức ăn đường phố duy nhất là điểm nguồn có khả năng cao nhất, dựa trên danh sách tuyến trường hợp ghi nhận tất cả 32 bệnh nhân đều tiêu thụ sản phẩm bánh mì nhiều nhân từ người bán này ở nhiều thôn khác nhau. Quá trình kiểm tra tại chỗ cũng phát hiện nhiều điểm chưa đảm bảo an toàn thực phẩm tại cơ sở này. Những bằng chứng dịch tễ học thực địa này hoàn toàn phù hợp với kết quả phân tích hệ gen, ngay cả khi chưa thực hiện một phân tích bệnh chứng hay thuần tập chính thức. Khi kết hợp lại, bằng chứng từ hệ gen và dịch tễ học đã củng cố mạnh mẽ giả thuyết về nguồn lây chung liên quan đến người bán này.

Tổng hợp lại, việc phân lập lại *Salmonella* từ các mẫu liên quan đến vụ ngộ độc, kiểu huyết thanh và hồ sơ MLST hoàn toàn tương đồng, khoảng cách SNP thấp, cùng chung hồ sơ plasmid và gen kháng kháng sinh đã chỉ ra rằng các chủng thực phẩm và phân thuộc cùng một cụm *Salmonella* Enteritidis ST11 liên quan đến vụ ngộ độc. Kết hợp với bằng chứng dịch tễ học thực địa, những phát hiện di truyền này cung cấp sự hỗ trợ vững chắc cho nguồn lây chung trong vụ ngộ độc bánh mì. Như đã lưu ý, việc quy kết nguồn gốc chính xác đến một nguyên liệu cụ thể ở cấp độ chủng vẫn sẽ đòi hỏi các bước điều tra truy xuất nguồn gốc sâu hơn [25]. Các kết luận của nghiên cứu này cũng cần được nhìn nhận dưới một số hạn chế. Chỉ có một khuẩn lạc được chọn từ mỗi mẫu dương tính; điều này có thể chưa phản ánh toàn diện sự đa dạng hệ gen bên trong mẫu đó. Nghiên cứu thiếu sự so sánh với tập dữ liệu nền rộng hơn bao gồm các hệ gen *Salmonella* Enteritidis ST11

ngoài vụ ngộ độc, làm hạn chế khả năng phân biệt cụm đặc thù của vụ ngộ độc với sự đa dạng đang lưu hành rộng rãi của kiểu huyết thanh phổ biến này. Quá trình gọi SNP sử dụng bộ lắp ráp của chủng S03 (một trong các chủng của vụ ngộ độc) làm hệ gen tham chiếu thay vì một hệ gen tham chiếu đóng độc lập; đây là cách tiếp cận được chấp nhận cho các chủng có quan hệ họ hàng gần cùng kiểu huyết thanh và kiểu trình tự, nhưng có thể tạo ra độ lệch nhỏ liên quan đến tham chiếu trong khoảng cách SNP từng cặp được báo cáo [23]. Việc tích hợp chính thức dữ liệu cấp chủng với dữ liệu dịch tễ và truy xuất nguồn gốc chưa nằm trong phạm vi của nghiên cứu này. Thiếu tập dữ liệu giải trình tự đoạn ngắn chạy song song để đối chiếu trực tiếp kết quả từ Nanopore với tiêu chuẩn vàng hiện tại. Ở một khía cạnh khác, các kết quả dự đoán về kiểu huyết thanh, kiểu trình tự, plasmid và gen kháng kháng sinh trong nghiên cứu này được suy luận bằng phương pháp tin học, do đó sẽ cần được kiểm chứng thêm thông qua các phương pháp thực nghiệm độc lập trong tương lai, chẳng hạn như định tít huyết thanh truyền thống, thử nghiệm hợp bào plasmid hoặc kháng sinh đồ.

4. KẾT LUẬN

Xét nghiệm vi sinh truyền thống đã phát hiện *Salmonella* spp. trong nhiều mẫu thực phẩm liên quan đến vụ ngộ độc. Phương pháp giải trình tự toàn bộ hệ gen đoạn dài bằng công nghệ Nanopore cho thấy các chủng *Salmonella* phân lập được đều có chung kiểu huyết thanh, kiểu trình tự và hồ sơ di truyền có độ tương đồng rất cao. Khi kết hợp với kết quả điều tra dịch tễ cho thấy toàn bộ 32 bệnh nhân đều sử dụng sản phẩm bánh mì từ cùng một người bán, những bằng chứng hệ gen này đã củng cố mạnh mẽ giả thuyết về một nguồn lây chung. Tuy nhiên, việc xác định chính xác nguyên liệu nào là nguồn gốc gây bệnh cụ thể vẫn đòi hỏi phải tiến hành các bước điều tra truy xuất nguồn gốc sâu hơn [25]. Bên cạnh đó, giải trình tự Nanopore cũng cho phép phân giải cấu trúc plasmid ở cấp độ đoạn dài và có thời gian hoàn thành kết quả tương đối nhanh, tạo nền tảng thực tiễn vững chắc cho việc đặc trưng hóa hệ gen gần như theo thời gian thực để ứng phó kịp thời với các vụ ngộ độc [24]. Những phát hiện từ nghiên cứu này khẳng định giá trị thiết thực của việc kết hợp giữa phương pháp nuôi cấy truyền thống và giải trình tự toàn bộ hệ gen đoạn dài. Đây sẽ là một hướng tiếp cận toàn diện dựa trên bằng chứng khoa học, góp phần nâng cao hiệu quả công tác điều tra ngộ độc thực phẩm và truy xuất nguồn gốc tại Việt Nam.

DỮ LIỆU

Các đoạn đọc thô từ kết quả giải trình tự toàn bộ hệ gen bằng công nghệ Oxford Nanopore trong nghiên cứu này đã được lưu trữ tại cơ sở dữ liệu NCBI Sequence Read Archive (SRA) với mã BioProject PRJNA1530768. Các mã định danh SRA Run cụ thể cho từng chủng vi khuẩn bao gồm: S01, SRR40717781, S02, SRR40717780, S03, SRR40717779, S04, SRR40717778, S05, SRR40717777, S06, SRR40717776, S07, SRR40717775, S08, SRR40717774, S09, SRR40717773, và S10, SRR40717772.

TÀI LIỆU THAM KHẢO

- [1]. Centers for Disease Control and Prevention (CDC), "About *Salmonella* infection," October 4, 2024. [Online]. Available: <https://www.cdc.gov/salmonella/about/> [Accessed July 1, 2026].
- [2]. World Health Organization (WHO), "*Salmonella* (non-typhoidal)," February 20, 2018. [Online]. Available: [https://www.who.int/news-room/fact-sheets/detail/salmonella-\(non-typhoidal\)](https://www.who.int/news-room/fact-sheets/detail/salmonella-(non-typhoidal)) [Accessed July 1, 2026].
- [3]. Vietnamnet, "More than 310 people suffered from Phuong bread poisoning, 273 cases were hospitalized." [Online]. Available: <https://vietnamnet.vn/313-nguoi-ngo-doc-sau-an-banh-mi-phuong-2193143.html> [Accessed July 1, 2026].
- [4]. Vietnam Food Administration, "Update on the situation of food poisoning cases in Long Khanh City," May 8, 2024. [Online]. Available: <https://vfa.gov.vn/tin-tuc/cap-nhat-tinh-hinh-vu-ngo-doc-thuc-pham-tren-dia-ban-thanh-pho-long-khanh.html> [Accessed July 1, 2026].
- [5]. Sai Gon Giai Phong, "Suspected bread poisoning case in Dong Thap: 25/36 samples were positive for *Salmonella* sp.," June 20, 2026. [Online]. Available: <https://www.sggp.org.vn/vu-nghi-ngo-doc-banh-mi-o-dong-thap-2536-mau-benh-pham-duong-tinh-voi-salmonella-sp-post858348.html> [Accessed July 1, 2026].
- [6]. Nhan Dan, "Bread poisoning case in Vung Tau: More than 54% of patient samples are positive for *Salmonella* bacteria," March 9, 2026. [Online]. Available: <https://nhandan.vn/vu-ngo-doc-banh-mi-o-vung-tau-hon-54-mau-benh-nhan-duong-tinh-voi-vi-khuan-salmonella-post947250.html> [Accessed July 1, 2026].

- [7]. Centers for Disease Control and Prevention (CDC), “Multistate Foodborne Outbreaks: Investigation Steps,” November 21, 2025. [Online]. Available: <https://www.cdc.gov/foodborne-outbreaks/outbreak-basics/investigation-steps.html> [Accessed July 1, 2026].
- [8]. U.S. Food and Drug Administration (FDA), “Examples of How FDA Has Used Whole Genome Sequencing of Foodborne Pathogens For Regulatory Purposes.” [Online]. Available: <https://www.fda.gov/food/whole-genome-sequencing-wgs-program/examples-how-fda-has-used-whole-genome-sequencing-foodborne-pathogens-regulatory-purposes> [Accessed July 1, 2026].
- [9]. J. Quick, P. Ashton, S. Calus, *et al.*, “Rapid draft sequencing and real-time nanopore sequencing in a hospital outbreak of *Salmonella*,” *Genome Biology*, vol. 16, p. 114, 2015.
- [10]. Y. Lin, L. Yang, S. Qiu, *et al.*, “Rapid identification and source tracing of a *Salmonella* Typhimurium outbreak in China by metagenomic and whole-genome sequencing,” *Foodborne Pathogens and Disease*, vol. 19, no. 4, pp. 259–265, 2022.
- [11]. T. L. Taylor, J. D. Volkening, E. DeJesus, *et al.*, “Rapid, multiplexed, whole genome and plasmid sequencing of foodborne pathogens using long-read nanopore technology,” *Scientific Reports*, vol. 9, p. 16350, 2019.
- [12]. X. Wu, H. Luo, C. Ge, *et al.*, “Evaluation of multiplex nanopore sequencing for *Salmonella* serotype prediction and antimicrobial resistance gene and virulence gene detection,” *Frontiers in Microbiology*, vol. 13, p. 1073057, 2023.
- [13]. S. Rheman, S. Hossain, M. S. Sarker, *et al.*, “Nanopore sequencing for identification and characterization of antimicrobial-resistant *Escherichia coli* and *Salmonella* spp. from tilapia and shrimp sold at wet markets in Dhaka, Bangladesh,” *Frontiers in Microbiology*, vol. 15, p. 1329620, 2024.
- [14]. B. Bogaerts, A. Van den Bossche, B. Verhaegen, *et al.*, “Closing the gap: Oxford Nanopore Technologies R10 sequencing allows comparable results to Illumina sequencing for SNP-based outbreak investigation of bacterial pathogens,” *Journal of Clinical Microbiology*, vol. 62, no. 5, p. e01576-23, 2024.
- [15]. Oxford Nanopore Technologies, “Dorado,” August 9, 2026. [Online]. Available: <https://github.com/nanoporetech/dorado> [Accessed August 13, 2026].
- [16]. M. Kolmogorov, J. Yuan, Y. Lin, and P. A. Pevzner, “Assembly of long, error-prone reads using repeat graphs,” *Nature Biotechnology*, vol. 37, no. 5, pp. 540–546, 2019.
- [17]. EPI2ME, “wf-bacterial-genomes,” July 23, 2026. [Online]. Available: <https://github.com/epi2me-labs/wf-bacterial-genomes> [Accessed August 13, 2026].
- [18]. Oxford Nanopore Technologies, “Medaka,” August 7, 2026. [Online]. Available: <https://github.com/nanoporetech/medaka> [Accessed August 13, 2026].
- [19]. O. Schwengers, L. Jelonek, M. A. Dieckmann, S. Beyvers, J. Blom, and A. Goesmann, “Bakta: rapid and standardized annotation of bacterial genomes via alignment-free sequence identification,” *Microbial Genomics*, vol. 7, no. 11, p. 000685, 2021.
- [20]. V. Bortolaia, R. S. Kaas, E. Ruppe, *et al.*, “ResFinder 4.0 for predictions of phenotypes from genotypes,” *The Journal of Antimicrobial Chemotherapy*, vol. 75, no. 12, pp. 3491–3500, 2020.
- [21]. J. Robertson and J. H. E. Nash, “MOB-suite: software tools for clustering, reconstruction and typing of plasmids from draft assemblies,” *Microbial Genomics*, vol. 4, no. 8, p. e000206, 2018.
- [22]. R. S. Kaas, P. Leekitcharoenphon, F. M. Aarestrup, and O. Lund, “Solving the Problem of Comparing Whole Bacterial Genomes across Different Sequencing Platforms,” *PLoS One*, vol. 9, no. 8, p. e104984, 2014.
- [23]. Saltykova, V. Wuyts, W. Mattheus, *et al.*, “Comparison of SNP-based subtyping workflows for bacterial isolates using WGS data, applied to *Salmonella enterica* serotype Typhimurium and serotype 1,4,[5],12:i:-,” *PLoS One*, vol. 13, no. 2, p. e0192504, 2018.
- [24]. H. H. Mostafa, “An evolution of Nanopore next-generation sequencing technology: implications for medical microbiology and public health,” *Journal of Clinical Microbiology*, vol. 62, no. 5, p. e00246-24, 2024.
- [25]. C. Nadon, I. Van Walle, P. Gerner-Smidt, *et al.*, “PulseNet International: Vision for the implementation of whole genome sequencing (WGS) for global food-borne disease surveillance,” *Euro Surveillance*, vol. 22, no. 23, p. 30544, 2017