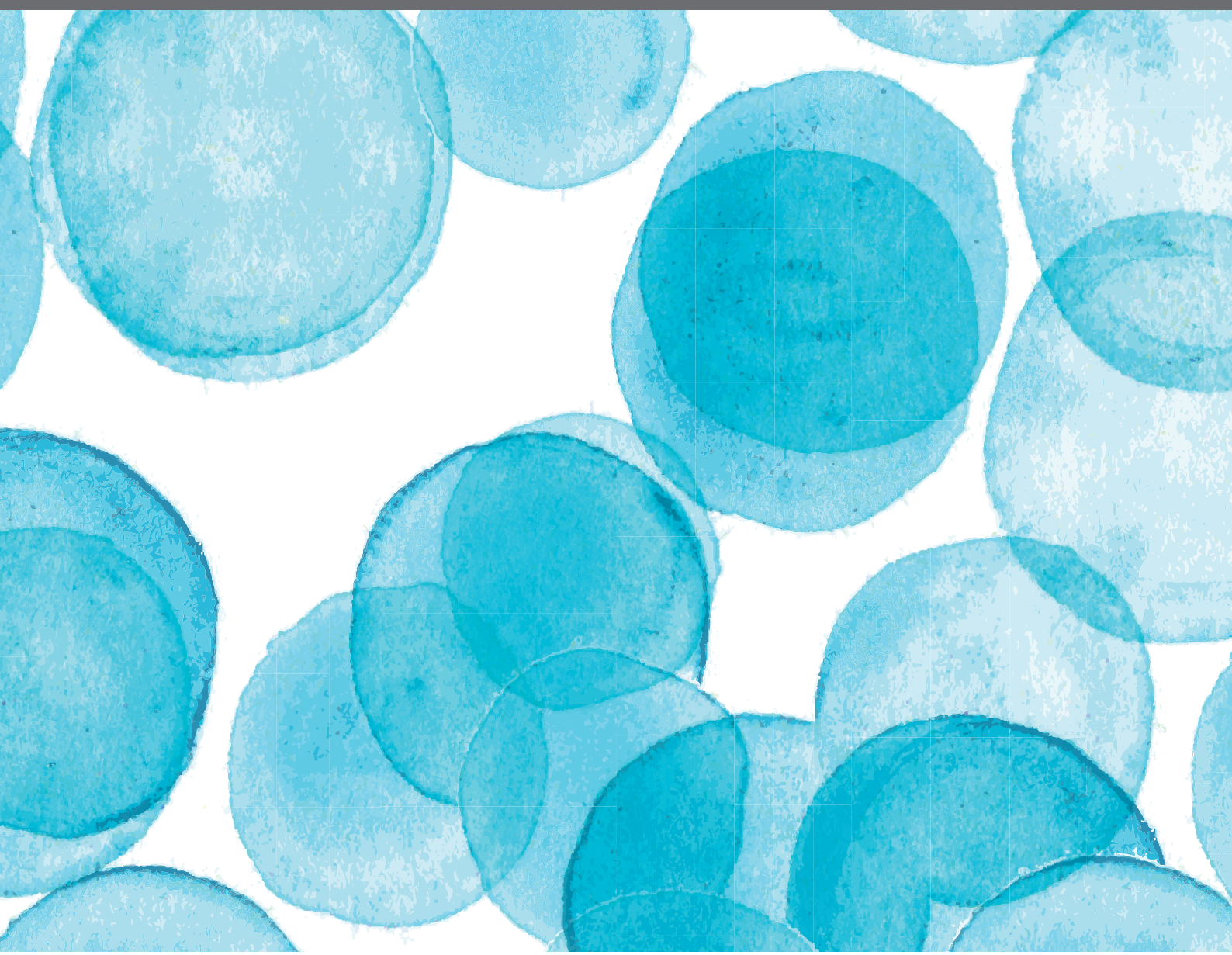
NEW ADVANCES IN IDENTIFICATION AND QUANTIFICATION OF FOODBORNE PATHOGENS

EDITED BY: Dario de Medici, David Rodriguez-Lazaro and Nigel Cook
PUBLISHED IN: Frontiers in Microbiology



frontiers Research Topics

frontiers

Frontiers eBook Copyright Statement

The copyright in the text of individual articles in this eBook is the property of their respective authors or their respective institutions or funders. The copyright in graphics and images within each article may be subject to copyright of other parties. In both cases this is subject to a license granted to Frontiers.

The compilation of articles constituting this eBook is the property of Frontiers.

Each article within this eBook, and the eBook itself, are published under the most recent version of the Creative Commons CC-BY licence. The version current at the date of publication of this eBook is CC-BY 4.0. If the CC-BY licence is updated, the licence granted by Frontiers is automatically updated to the new version.

When exercising any right under the CC-BY licence, Frontiers must be attributed as the original publisher of the article or eBook, as applicable.

Authors have the responsibility of ensuring that any graphics or other materials which are the property of others may be included in the CC-BY licence, but this should be checked before relying on the CC-BY licence to reproduce those materials. Any copyright notices relating to those materials must be complied with.

Copyright and source acknowledgement notices may not be removed and must be displayed in any copy, derivative work or partial copy which includes the elements in question.

All copyright, and all rights therein, are protected by national and international copyright laws. The above represents a summary only. For further information please read Frontiers' Conditions for Website Use and Copyright Statement, and the applicable CC-BY licence.

ISSN 1664-8714
ISBN 978-2-88971-924-2
DOI 10.3389/978-2-88971-924-2

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

NEW ADVANCES IN IDENTIFICATION AND QUANTIFICATION OF FOODBORNE PATHOGENS

Topic Editors:
Dario de Medici, National Institute of Health (ISS), Italy
David Rodriguez-Lazaro, University of Burgos, Spain
Nigel Cook, Jorvik Food & Environmental Virology Ltd, United Kingdom

Citation: De Medici, D., Rodriguez-Lazaro, D., Cook, N., eds. (2021). New Advances in Identification and Quantification of Foodborne Pathogens. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88971-924-2

Table of Contents

05 ***Editorial: New Advances in Identification and Quantification of Foodborne Pathogens***
Dario De Medici, David Rodriguez-Lazaro and Nigel Cook

08 ***Validation of Droplet Digital Polymerase Chain Reaction for* Salmonella *spp. Quantification***
Carolina Villamil, Martha Nancy Calderon, Maria Mercedes Arias and John Emerson Leguizamon

21 ***Prevalence, Serotype, Antibiotic Susceptibility, and Genotype of* Salmonella *in Eggs From Poultry Farms and Marketplaces in Yangling, Shaanxi Province, China***
Wei Li, Hao Li, Shujuan Zheng, Zewei Wang, Huanjing Sheng, Chunlei Shi, Xianming Shi, Qinya Niu and Baowei Yang

32 ***German-Wide Interlaboratory Study Compares Consistency, Accuracy and Reproducibility of Whole-Genome Short Read Sequencing***
Laura Uelze, Maria Borowiak, Markus Bönn, Erik Brinks, Carlus Deneke, Thomas Hankeln, Sylvia Kleta, Larissa Murr, Kerstin Stingl, Kathrin Szabo, Simon H. Tausch, Anne Wöhlke and Burkhard Malorny

46 ***Application of an Optimized Direct Lysis Method for Viral RNA Extraction Linking Contaminated Dates to Infection With Hepatitis A Virus***
Sheikh Md Rajiuddin, Sofie Elisabeth Midgley, Tenna Jensen, Luise Müller and Anna Charlotte Schultz

53 ***Development of PMAxx™-Based qPCR for the Quantification of Viable and Non-viable Load of* Salmonella *From Poultry Environment***
Jiawei Zhang, Samiullah Khan and Kapil K. Chousalkar

66 ***Methicillin-Resistant* Staphylococcus aureus *(MRSA) in Slaughter Houses and Meat Shops in Capital Territory of Pakistan During 2018–2019***
Asma Sadiq, Maroof Samad, Saddam, Nosheen Basharat, Shahid Ali, Roohullah, Zubaida Saad, Allah Nawaz Khan, Yasin Ahmad, Alam Khan and Jadoon Khan

78 ***Application of Short Pre-enrichment, and Double Chemistry Real-Time PCR, Combining Fluorescent Probes and an Intercalating Dye, for Same-Day Detection and Confirmation of* Salmonella *spp. and* Escherichia coli *O157 in Ground Beef and Chicken Samples***
Alejandro Garrido-Maestu, Sarah Azinheiro, Foteini Roumani, Joana Carvalho and Marta Prado

86 ***Proficiency Testing of Metagenomics-Based Detection of Food-Borne Pathogens Using a Complex Artificial Sequencing Dataset***
Dirk Höper, Josephine Grützke, Annika Brinkmann, Joël Mossong, Sébastien Matamoros, Richard J. Ellis, Carlus Deneke, Simon H. Tausch, Isabel Cuesta, Sara Monzón, Miguel Juliá, Thomas Nordahl Petersen, Rene S. Hendriksen, Sünje J. Pamp, Mikael Leijon, Mikhayil Hakhverdyan, Aaron M. Walsh, Paul D. Cotter, Lakshmi Chandrasekaran, Moon Y. F. Tay, Joergen Schlundt, Claudia Sala, Alessandra De Cesare, Andreas Nitsche, Martin Beer and Claudia Wylezich

97 ***Mammalian Cell-Based Immunoassay for Detection of Viable Bacterial Pathogens***

Luping Xu, Xingjian Bai, Shivendra Tenguria, Yi Liu, Rishi Drolia and Arun K. Bhunia

112 ***Multiplex PCR for the Identification of Pathogenic* Listeria *in* Flammulina velutipes *Plant Based on Novel Specific Targets Revealed by Pan-Genome Analysis***

Fan Li, Qinghua Ye, Moutong Chen, Jumei Zhang, Liang Xue, Juan Wang, Shi Wu, Haiyan Zeng, Qihui Gu, Youxiong Zhang, Xianhu Wei, Yu Ding and Qingping Wu

124 ***Identification of* Vibrio parahaemolyticus *and* Vibrio *spp. Specific Outer Membrane Proteins by Reverse Vaccinology and Surface Proteome***

Wenbin Wang, Jianxin Liu, Shanshan Guo, Lei Liu, Qianyun Yuan, Lei Guo and Saikun Pan

138 ***Evaluation of* Salmonella *Serotype Prediction With Multiplex Nanopore Sequencing***

Xingwen Wu, Hao Luo, Feng Xu, Chongtao Ge, Shaoting Li, Xiangyu Deng, Martin Wiedmann, Robert C. Baker, Abigail Stevenson, Guangtao Zhang and Silin Tang

frontiers in Microbiology

EDITORIAL
published: 01 November 2021
doi: 10.3389/fmicb.2021.783406

Editorial: New Advances in Identification and Quantification of Foodborne Pathogens

Dario De Medici[1]*, *David Rodriguez-Lazaro*[2,3] *and Nigel Cook*[4]

[1] *Istituto Superiore di Sanità, Rome, Italy,* [2] *Microbiology Division, Faculty of Sciences, University of Burgos, Burgos, Spain,* [3] *Centre for Emerging Pathogens and Global Health, University of Burgos, Burgos, Spain,* [4] *Jorvik Food and Environmental Virology Ltd., York, United Kingdom*

Keywords: food safety, molecular methods, rapid method development, foodborne pathogen detection, foodborne virus concentration, next generation (deep) sequencing (NGS), droplet PCR

Editorial on the Research Topic

New Advances in Identification and Quantification of Foodborne Pathogens

Microbiological analysis is an integral part of food safety quality control and monitoring systems. Recent outbreaks, associated with consumption of a variety of contaminated foods, have highlighted the difficulty in monitoring critical points of contamination along the food chain. The rapid and precise detection (and/or quantification) of foodborne pathogens, used appropriately along the food supply chain, represents a powerful tool to prevent outbreaks of foodborne disease. With this purpose in mind, and particularly for short shelf-life foods, it is necessary to increase the arsenal of available methods and develop innovative and time-effective assays which are able to complete the analytical process in as short time as possible. In this context, the major benefit of molecular methods, as alternatives to classical culture-based methods, is the significant reduction of the time taken for analysis. However, basic molecular approaches have also some critical weaknesses, such as an inability to discriminate between viable and dead cells (or infectious and non-infectious virus particles), and susceptibility to inhibitory influences.

In recent years, there have been several advances in molecular-based subtyping methods, particularly regarding the application of sequence-based techniques. Different molecular methods, including more recently next- generation-based approaches, are replacing traditional methods in identifying the source of outbreaks, improving trace-back studies, and elucidating the evolution of foodborne pathogens. However, all these techniques need further validation before being adopted internationally.

This Research Topic encompasses different aspects regarding the application of molecular methods for the detection, quantification, and identification of foodborne microorganisms, including the proposition of innovative solutions. The Special Issue comprises 12 original articles, contributed by 111 authors from 14 different countries.

SPECIFIC DETECTION OF VIABLE/INFECTIOUS FROM DEAD/NON-INFECTIOUS TARGETS

The major challenge of molecular methods is their inability to differentiate between dead/infectious and viable cells/non-infectious microorganisms (Zeng et al., 2016). An approach to address this is proposed within this issue.

OPEN ACCESS

Edited by:
Giovanna Suzzi,
University of Teramo, Italy

Reviewed by:
Arun K. Bhunia,
Purdue University, United States

****Correspondence:***
Dario De Medici
dario.demedici@iss.it

Specialty section:
This article was submitted to Food Microbiology, a section of the journal Frontiers in Microbiology

Received: *26 September 2021*
Accepted: *07 October 2021*
Published: *01 November 2021*

Citation:
De Medici D, Rodriguez-Lazaro D and Cook N (2021) Editorial: New Advances in Identification and Quantification of Foodborne Pathogens. Front. Microbiol. 12:783406. doi: 10.3389/fmicb.2021.783406

Zhang et al. optimized a commercial extraction kit, PMAxx-based qPCR for the quantification of viable *Salmonella* Enteritidis and *Salmonella* Typhimurium from poultry-production environments. The method was successfully optimized and applied to assess the survivability of *Salmonella* serovars in soil samples, incubated at 3 different temperatures (5, 25, and 37°C) over 6 weeks.

A very different approach is proposed by Xu et al., which utilizes pathogen interaction with host cells to detect only live microorganisms. They exploited the ability of viable enteric pathogens to adhere to intestinal cells, harnessing an antibody-based assay for specific detection of the adhered target pathogens. The Authors also proposed an innovative approach to prolong the shelf-life of cells (for at least 14 weeks) for possible point-of-need deployment and multi-sample testing on a single plate.

ADVANCE IN EXTRACTION METHODS FOR ENTERIC VIRUSES

Detection of enteric viruses in food products differs from detection of most foodborne bacteria, as viruses cannot be enriched in a food sample. The pre-PCR processing step, including sample preparation and nucleic acid extraction, is a critical step in foodborne virus detection. The quality of the extracted nucleic acids can strongly affect the efficiency of detection by a molecular method. In 2018 an outbreak of hepatitis A, caused by contaminated dates, was identified in Denmark by Rajiuddin et al. A new direct lysis method for the extraction of viral RNA from dates was developed and optimized to detect the virus, quantify the contamination level, and sequence the strain(s) detected.

SAME DAY DETECTION METHODS

The length of the enrichment time is a bottleneck for development of culture-based methods able to detect bacterial pathogens in the same day (Fachmann et al., 2017).

Garrido-Maestu et al. applied a rapid short enrichment step followed by an innovative unique multiplex real time PCR method, to mediate same-day detection and confirmation of *Salmonella* spp and *E. coli* O157. The method was also validated for stressed and dead cells inoculated into meat, demonstrating that it is capable of distinguishing thermally stressed bacteria, as well as avoiding false positive results due to the presence of dead bacteria.

Li et al. propose a multiplex PCR (mPCR) for rapid identification of *Listeria monocytogenes* and *Listeria ivanovii*, and non-pathogenic *Listeria* in edible mushroom (*Flammulina velutipes*) samples after a short enrichment culture step (4–12 h).

NEW MOLECULAR APPROACH: ABSOLUTE QUANTIFICATION AND NEXT GENERATION SEQUENCING

Real time PCR or qPCR has been widely used for the quantification of nucleic acids with high precision, but it can be inaccurate when targets are at very low concentrations. Droplet digital polymerase chain reaction (ddPCR), a relatively new and promising technology, has many advantages over qPCR, as it can provide greater sensitivity and accuracy. It can be used for ultrasensitive and absolute nucleic acid quantification without a standard curve (Peng et al., 2020).

Villamil et al. (2020) validated a method based on the application of a simplex and duplex droplet digital polymerase chain reaction (ddPCR)-based method, for the identification and quantification of *Salmonella*. The Authors demonstrated that the use of a multiplex detection, targetting *ttr*, *inv*A, *hil*A, *spa*Q, and siiA genes, provides more reliable quantification, particularly for some specific applications such as reference material characterization.

Recently, Next Generation Sequencing has become a powerful tool for epidemiological investigation, as it can define the causative agents of disease, trace the origin of the pathogen, and clarify routes of transmission. Within the framework of the European Union-funded project COMPARE [Collaborative Management Platform for detection and Analyses of (Re-) emerging and foodborne outbreaks in Europe], Hoper et al. present the results of a proficiency test to scrutinize the ability of 12 different laboratories to assess diagnostic metagenomics data, based on the use of an artificial dataset resembling shotgun sequencing of RNA from a sample of contaminated trout. Analysis of the results demonstrate that a reliable classification of the reads can be obtained by the software used.

The consistency, accuracy and reproducibility of next-generation short read sequencing, between different laboratories and platforms, has been evaluated by Uelze et al. The Authors report the results of a ring trial performed in private and public laboratories, sequencing DNA samples of three bacterial species (*Campylobacter jejuni*, *Listeria monocytogenes*, and *Salmonella enterica*), according to their routine in-house sequencing protocols. Four different types of Illumina and one Ion Torrent sequencing platform in ten different laboratories were involved in the study. The results show differences in data quality after assembling short reads into genome assemblies.

The bioinformatics pipelines of whole genome sequencing (WGS) data, generated by the long-read sequencing platform Oxford Nanopore Technologies (ONT), have been evaluated by Wu et al. Five isolates of *Salmonella* were analyzed, and the results demonstrated the reduction of the cost of ONT sequencing of a single isolate per flow cell, the achievement of high coverage of the genome, and a high accuracy of prediction within 1 day. This study provides food industries, food authorities, and regulators with the alternative of short-read sequencing platforms for serotype prediction for pathogen surveillance.

IMMUNODIAGNOSTIC ASSAY

The major economic threat to the marine aquaculture industry, with regard to consumer health, is the presence of *Vibrio* species in seafood.

Development of an accurate immunodiagnostic assay of pathogenic *Vibrio* species, using highly specific unique outer

membrane proteins (OMPs), is described by Wang et al. The Authors have applied a bioinformatic tools for screening OMPs candidates, to minimize target numbers in "*in vivo*" tests.

MONITORING

The Research Topic also includes two extensive surveillance studies in which the Authors apply molecular methods for characterizing the isolated strains. Very importantly, both papers have a particular focus on antibiotic-resistant bacteria (Aijuka and Buys, 2020).

The paper presented by Li et al. reports the results of a survey in which new molecular technologies have been applied for evaluating *Salmonella* presence in eggs in a specific region of China, with extensive characterization of isolates. The results of the work indicate that more attention should be paid to egg production, transportation, and storage to prevent foodborne outbreaks caused by *Salmonella*.

Finally, an extensive investigation of the prevalence of methicillin-resistant *Staphylococcus aureus* (MRSA) in slaughterhouses and meat shops in Pakistan has been presented by Sadiq et al. The authors collected a total of 300 samples. Fifty per cent of the analyzed samples were positive for *S. aureus* by phenotypic identification, and 96 isolates (63%) showed resistance to the antibiotic cefoxitin, known as a potential marker for detecting MRSA. All 150 isolates showed complete resistance to four antibiotics–neomycin, methicillin, ciprofloxacin, and tetracycline.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

ACKNOWLEDGMENTS

This collection of articles has been possible thanks to the collaboration of the reviewers (25 reviewers), who selflessly invested their time. We thank the 111 authors who entrusted us with their manuscript submissions.

REFERENCES

Aijuka, M., and Buys, E. M. (2020). Detection of extended-spectrum beta-lactamase cefotaxime resistance and virulence genes in Escherichia coli by duplex quantitative real-time PCR and melt curve analysis. *Lett. Appl. Microbiol.* 71, 54–60. doi: 10.1111/lam.13274

Fachmann, M. S. R., Löfström, C., Hoorfar, J., Hansen, F., Christensen, J., Mansdal, S., and Josefsen, M. H. (2017). Detection of *Salmonella* enterica in meat in less than 5 hours by a low-cost and noncomplex sample preparation method. *Appl. Environ. Microbiol.* 83, e03151–e03116. doi: 10.1128/AEM.03151-16

Peng, C., Zheng, M., Ding, L., Chen, X., Wang, X., Feng, X., et al. (2020). Accurate detection and evaluation of the gene-editing frequency in plants using droplet digital PCR. *Front. Plant Sci.* 14:610790. doi: 10.3389/fpls.2020.610790

Villamil, C., Calderon, M. N., Arias, M. M., and Leguizamon, J. E. (2020). Validation of droplet digital polymerase chain reaction for *Salmonella* spp quantification. *Front. Microbiol.* 11:1512. doi: 10.3389/fmicb.2020.01512

Zeng, D., Chen, Z., Jiang, Y., Xue, F., and Li, B. (2016). Advances and challenges in viability detection of foodborne pathogens. *Front. Microbiol.* 7:1833. doi: 10.3389/fmicb.2016.01833

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 De Medici, Rodriguez-Lazaro and Cook. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Microbiology

METHODS
published: 07 July 2020
doi: 10.3389/fmicb.2020.01512

Validation of Droplet Digital Polymerase Chain Reaction for *Salmonella* spp. Quantification

*Carolina Villamil[1], Martha Nancy Calderon[1], Maria Mercedes Arias[2] and John Emerson Leguizamon[2]**

[1] Departamento de Química, Universidad Nacional de Colombia, Bogota, Colombia, [2] Grupo de Metrología en Bioanálisis, Instituto Nacional de Metrología, Bogota, Colombia

Salmonellosis is a foodborne disease caused by *Salmonella* spp. Although cell culture is the gold standard for its identification, validated molecular methods are becoming an alternative, because of their rapidity, selectivity, and specificity. A simplex and duplex droplet digital polymerase chain reaction (ddPCR)-based method for the identification and quantification of *Salmonella* using *ttr*, *invA, hilA, spaQ,* and *siiA* gene sequences was validated. The method has high specificity, working interval between 8 and 8,000 cp/μL in ddPCR reaction, a limit of detection of 0.5 copies/μL, and precision ranging between 5 and 10% measured as a repeatability standard deviation. The relative standard measurement uncertainty was between 2 and 12%. This tool will improve food safety in national consumption products and will increase the competitiveness in agricultural product trade.

Keywords: droplet digital PCR, *Salmonella*, validation, food safety, bioanalysis

OPEN ACCESS

Edited by:
Dario De Medici,
Istituto Superiore di Sanità (ISS), Italy

Reviewed by:
David Dobnik,
National Institute of Biology (NIB), Slovenia
Barbara Bertasi,
Experimental Zooprophylactic Institute of Lombardy and Emilia Romagna (IZSLER), Italy

****Correspondence:***
John Emerson Leguizamon
jeleguizamon@inm.gov.co

Specialty section:
This article was submitted to Food Microbiology, a section of the journal Frontiers in Microbiology

Received: *09 March 2020*
Accepted: *10 June 2020*
Published: *07 July 2020*

Citation:
Villamil C, Calderon MN, Arias MM and Leguizamon JE (2020) Validation of Droplet Digital Polymerase Chain Reaction for Salmonella spp. Quantification. Front. Microbiol. 11:1512. doi: 10.3389/fmicb.2020.01512

INTRODUCTION

Salmonella spp. is one of the most important pathogens and the leading cause of foodborne diseases (World Health Organization, 2015). Hence, the development and validation of biometric tools, such as measurement methods and reference materials, are important for making decisions related to the national and global trade of food products (European Comission, 2007), as well as for public health surveillance, among others.

With recent advances in molecular techniques, several rapid methodologies for the detection and quantification of pathogens based on specific genes and proteins have been developed. Besides their rapidity, these methods provide lower limits of detection and better specificity (Law et al., 2014). In particular, the detection of *Salmonella* spp. using polymerase chain reaction (PCR) is based on the presence of single genes in a large number of different serotypes and unique copies within the genome. Among the most commonly used ones are *(i) invA*, which encodes for invA, a protein required for invasion of the bacterium to epithelial cells (Galan et al., 1992), *(ii) hilA* (hyperinvasive locus), which encodes for hilA, a transcriptional regulator of the OmpR/ToxR family that activates the expression of genes of the type III secretion system, required for bacterial invasion (Boddicker et al., 2003), *(iii)* the *ttr* locus, which is required for respiration with tetrathionate and survival of the bacteria after infection (Law et al., 2014), *(iv) spaQ*, a part of the inv/spa complex, which is required for the entry of the bacteria into non-phagocytic host cells (Kurowski et al., 2002), and *(v) siiA*, which produces a regulatory protein encoded by SP14, important for adhesion to epithelial cells during *Salmonella* invasion (Zhao et al., 2017).

Digital PCR is considered a potential primary method of measurement as it does not require standards of the same quantity to yield a measurement result (Gutierrez-Aguirre et al., 2015).

It is based on microfluidics technology, which allows the generation of multiple reaction partitions (1,000–10 million, depending on the platform) that work as individual reactions. Based on the positive or negative fraction and following the Poisson distribution, it is possible to determine the absolute concentration of the target of interest in terms of the number of copies per microliter (cp/μL) in the dPCR reaction (Huggett et al., 2013; Morisset et al., 2013; Magnusson and Örnemark, 2014). Besides, owing to the dilution and partitioning of the sample, the technique is less sensitive to inhibitors, which results in better precision and less uncertainty contribution (Somanath and Kerry, 2016). In this study, an in-house validation has been described for droplet digital PCR (ddPCR) for quantitating genomic DNA from *Salmonella* spp. using five different targets, run in two duplex and one simplex form. The performance characteristics evaluated were specificity, working interval, precision, limit of detection (LOD), limit of quantification (LOQ), and measurement uncertainty, contributing to the strengthening of monitoring process of *Salmonella* in agricultural products in Colombia.

MATERIALS AND METHODS

Bacterial Strains and Growth Conditions

Sixteen serotypes of *Salmonella enterica* subsp. *enterica* and other reference strains phylogenetically close to *Salmonella* spp., such as *Escherichia coli, Klebsiella* spp., and *Shigella* spp. (**Table 1**) were cultured in solid tryptone soya agar (TSA) or XLT4 (xylose, lysine, and tergitol 4) and liquid selective medium Rappaport Vassiliadis (RV). All strains were grown at 37°C for 12–18 h in liquid medium RV before DNA extraction.

TABLE 1 | Enterobacteria strains used in this study.

Strain description	Source
Salmonella enteritidis	ATCC® 13076
Salmonella typhimurium	Clinical isolate
Salmonella typhi	Clinical isolate
Salmonella derby	Clinical isolate
Salmonella dublin	Clinical isolate
Salmonella give	Clinical isolate
Salmonella saintpaul	Clinical isolate
Salmonella hadar	Clinical isolate
Salmonella infantis	Clinical isolate
Salmonella anatum	Clinical isolate
Salmonella panama	Clinical isolate
Salmonella typhimurium var5-	Clinical isolate
Salmonella javiana	Clinical isolate
Salmonella braenderup	Clinical isolate
Salmonella muenchen	Clinical isolate
Salmonella paratyphi A	Clinical isolate
Bacillus cereus	ATCC® 10876
Enterococcus faecalis	ATCC® 14506
Escherichia coli O104:H4	ATCC® BAA-2326
Escherichia coli O145:NM	ATCC® CDC99–3311
Escherichia coli O157:H7	ATCC® 35150
Proteus mirabilis	ATCC® 12453
Proteus vulgaris	ATCC® 33420
Shigella boydii	ATCC® 9207
Shigella sonnei	ATCC® 9290
Staphylococcus aureus	ATCC® 25923
Staphylococcus aureus	ATCC® 6538
Vibrio parahaemolyticus	ATCC® 17802

ATCC, Reference strain from the American Type Culture Collection.

DNA Extraction

A protocol equally efficient for both Gram-positive and Gram-negative bacteria was used with some modifications (Atashpaz et al., 2010). Briefly, 3 mL cell culture in the exponential phase was centrifuged at 5,000 rpm for 5 min at 4°C. Then, 800 μL lysis buffer [CTAB 2% (w/v), Tris−HCl 100 mM, NaCl 1.4 M, EDTA 20 mM, and LiCl 0.2% (w/v), pH 8.0] was added to the bacterial pellet. The samples were incubated at 65°C for 30 min for Gram-negative bacteria and 2 h for Gram-positive bacteria. Then, these were centrifuged at 10,000 rpm for 5 min at 4°C. The supernatant was transferred to a new tube where 1 volume of chloroform-isoamyl alcohol (24:1) was added to precipitate the proteins. The samples were gently mixed and centrifuged at 12,000 rpm for 8 min at 4°C. Two further extractions with chloroform-isoamyl alcohol (24:1) were performed on the aqueous phases for each sample. From the pooled supernatants, the DNA was precipitated by the addition of 1 volume of cold isopropanol and 100 μL of 5 M sodium acetate after overnight incubation at −20°C. Subsequently, the samples were centrifuged at 12,000 rpm for 15 min at 4°C and washed with 1 mL of 70% ethanol by inverting the tubes. Finally, once the ethanol had evaporated, the DNA was resuspended in 100 μL 1X TE buffer (Tris–HCl 10 mM, EDTA 1 mM pH 8.0) and stored at −20°C. The quality of the extracted genomic DNA was evaluated using UV spectrophotometry by measuring the absorbance ratios at 260 nm/280 nm and 260 nm/230 nm; DNA integrity was assessed using electrophoresis on 1% agarose gel.

Primers and Probes

The primers and probes (Biosearch Technologies, Petaluma, CA, United States, purified by HPLC) for each target are shown in **Table 2**.

Quantitative PCR (qPCR)

Quantitative PCR was used as a preliminary test for evaluating the reaction efficiency of primers and probes for each target in simplex mode (CFX 96 Deep well- BioRad, cat. 1855196). Six dilutions of *S. enteritidis* reference strain ATCC 13076 DNA in 1X TE buffer were prepared (100–0.01 ng/μL). The 20 μL reaction contained 1X iTaq^{TM} Universal Probes Supermix (BioRad cat. 1725131), 400 nM primers, 300 nM probes, and nuclease-free water. 1X TE was used as a non-template control (NTC). All samples were run in triplicate. The amplification cycle consisted of initial denaturation at 95°C for 10 min, followed by 40 cycles of denaturation at 95°C for 15 s and annealing-elongation at 60°C

TABLE 2 | Primers and probes selected for the analysis of *Salmonella spp.*

Gene	Name	Sequence 5′-3′	Amplicon sequence	Amplicon size (pb)	References
invA	invA_176F	CAACGTTTCCTGCGGTACTGT	CAACGTTTCCTGCGGTACTGTTAATTAC	119	González-Escalona et al., 2012
	invA_291R	CCCGAACGTGGCGATAATT	CACGCTCTTTCGTCTGGCATTATCG		
	invA_Tx_208 TXa-	FAM-CTCTTTCGTCTGGCATTATC GATCAGTACCA-BHQ1	ATCAGTACCAGTCGTCTTATCTT GATTGAAGCCGATGCCGGTGAAAT TATCGCCACGTTCGGGCAA		
ttr	ttr6F4287 – Directo	CTCACCAGGAGATTACAACATGG	CTCACCAGGAGATTACAACATGG	95	Rothrock et al., 2013;
	ttr4R4381 – Reverso	AGCTCAGACCAAAAGTGACCATC	CTAATTTAACCCGTCGTCAGTGGCTA		Ssemanda et al., 2018;
	ttr5p4336 – Sonda	CFO-CACCGACGGCGAGACCG ACTTT-BHQ1	AAAGTCGGTCTCGCCGTCGGTG GGATGGTCACTTTTGGTCTGAGCT		Boer et al., 2019
siiA	siiA Fw	ACGACTGGGATATGAACGGGGAA	ACGACTGGGATATGAACGGGGAATT	107	Ben Hassena et al., 2015
	siiA Rv	TCGTTGTACTTGATGCTGCGGAG	ATTTTAATGAAAGAGATTAAG		
	siiaA Tr	FAM-ATCCTGATGTAGTTATTGAC ATGAG-BHQ1*	AAGATATATCCTGATGTA GTTATTGACATGAGTGTTAACT CCGCAGCATCAAGTACAACGA		
hilA	hilA F	ACTGTACGGACAGGGCTAT*	ACTGTACGGACAGGGCTATCGGTT	129	McCabe et al., 2011;
	hilA R	AGA CTC TCG GAT TGA ACC TGA-3′	TAATCGTCCGGTCGTAGTGGTGT		Wang et al., 2016
	hilA Lc640	HEX-TCGTCCGGTCGTAGTGGTG TCTCC-BHQ1	CTCCGCCAGCGCCGCAACCTACGACT CATACATTGGCGATACTTCCTTTTCAGAT GCAGGATCAGGTTCAATCCGAGAGTCT		
spaQ	spaQ Fw	CCTGACGCCCGTAAGAGA	CCTGACGCCCGTAAGAGAGTAAAACTTA	113	Kurowski et al., 2002;
	spaQ Rv	GCAATTACAGGAACAGACGCT	CGCCATACCAGCCAGACAGTAAA		Ekiri et al., 2016
	spaQ P	FAM-TAAAACTTCGCCATACCAGC CAGACA-BHQ1	ACAAGCATAAACACACGCCA AGTAATTTAATGCCAAAAGG CAGCGTCTGTTCCTGTAATTGC		

*FAM: 6-Carboxyfluorescein, HEX: Hexachlorofluorescein, CFO: Cal Fluor Orange 560, equivalent to HEX dye. BHQ1: Black Hole Quencher 1. *Modified in this study from references.*

for 30 s, with a heating ramp of 2°C/s. Efficiency was evaluated using Eq. 1,

$$E = 10^{\frac{-1}{m}} - 1 \quad (1)$$

where E is the efficiency and m the slope of a linear regression between the amplification cycle and the log concentration of each sample. The acceptability criterion for PCR efficiency was 90–100%.

ddPCR

In order to establish the best conditions to perform the validation assay for every target sequence, the annealing temperature through a gradient from 55–63°C was evaluated. In addition, the primer concentration was also evaluated from 300 to 900 nM, and the ramp rate between 1 and 2°C/s. The aim was to establish general conditions to amplify the target sequences under the same experimental conditions. After optimization, the best amplification temperature was 60°C with a ramp rate of 2°C/s and 600 nM concentration of primers.

After optimization, according to MIQE guidelines (Huggett et al., 2013; **Supplementary Table S8**), ddPCR assay was performed in simplex (*spaQ*), and duplex (*hilA-siiA* and *invA*-ttr) form: *(i)* according to the number of total reactions, a stock master mix with 1X ddPCR™ Supermix for Probes (BioRad, CA, United States cat. 1863024), 600 nM primers, 300 nM probes, and nuclease-free water was prepared; *(ii)* for every triplicate 60 μL of the stock master mix was weighed, *(iii)* and finally 6 μL of DNA template (genomic DNA of *S. enteritidis,* measured gravimetrically) were added to complete 66 μL. Then to avoid pipetting bias, 21 μL PCR reaction mixture and 70 μL generator oil (BioRad cat. 183005) were loaded into an 8-well DG8 cartridge (BioRad cat. 1864008), for every replicate, in order to have the same dilution factor. Droplets were generated in a QX200 droplet generator, transferred to a 96-well plate (BioRad cat. 12001925), and sealed with a PX1 PCR plate sealer (BioRad cat. 1814000). Reactions were amplified in a CFX96 deep-well thermocycler (BioRad cat. 185–5196), and the cycling conditions were: 95°C for 10 min, 40 cycles at 95°C for 30 s, and 60°C for 1 min, with an overall ramp rate of 2°C/s. All the reactions were read in the QX200 droplet reader system (two detection channels FAM/EvaGreen and VIC/HEX) using the Quantasoft™ software V1.7 from BioRad.

Data Acquisition and Analysis

Droplets were classified as PCR-positive or PCR-negative according to a threshold fluorescence value set manually using the Quantasoft™ software V1.7. Technical reasons for excluding data from the analysis were: *(i)* the total number of droplets <12,000 (BioRad, 2018) and *(ii)* wells with amplitude different from those of other wells with the same target. This variance indicated poor droplet generation, possibly due to poor handling or mixing of samples, and is likely to generate erroneous concentration values. Reference material ERM-AD623 (White et al., 2015) was used as the amplification control for assessing the performance of the thermocycling parameters.

Data were exported to a spreadsheet to calculate the concentration of the sample in cp/μL in the ddPCR reaction, with a droplet volume of 0.819 ± 0.017 nL (k = 2; Eqs 2, 3), as long as the confidence limits were associated with Poisson distribution.

$$C_{\text{sample}} = \frac{\lambda}{v \times D} \quad (2)$$

where:

$$\lambda = -In\left(\frac{N}{P}\right) \quad (3)$$

C_{sample}: Sample concentration ($\frac{cp}{uL}$)
λ : Copies/partition
N: Number of negative partitions
P: Number of total partitions
v: Droplet volume
D: Total dilution of the sample from the master mix.

The droplet volume was calculated as the mean of values reported previously (Corbisier et al., 2015; Dagata et al., 2016) and the uncertainty was estimated based on a rectangular distribution, to encompass any influence of droplet volume variability and to avoid underestimation of measurement uncertainty (Košir et al., 2017; Emslie et al., 2019).

Method Performance Characteristics

The performance characteristics evaluated were specificity, working interval, equivalence between simplex and duplex assay, precision as repeatability and intermediate precision, the LOD, the LOQ, and measurement uncertainty.

RESULTS

qPCR Amplification Efficiency of DNA Targets

From the optimized extraction process, 1 mL of 4430 ng/μL of DNA was obtained, 1:4 dilution was prepared as a working solution (WS; see **Supplementary Figure S1**). Preliminary tests using qPCR were performed to determine whether all gene sequences amplified correctly and whether PCR inhibitors were present. The qPCR amplification efficiencies calculated for the five targets were between 90 and 99% (**Table 3**), indicating that inhibitors were not present in the starting genomic DNA. The dilutions and log concentration showed a good correlation, indicating the suitability of qPCR for the studied targets.

Specificity

The DNA of sixteen serotypes of *Salmonella enterica* subsp. *enterica* (**Table 1**) was amplified using the optimized ddPCR

TABLE 3 | Amplification efficiency for target sequences under study.

Gene	Slope	Efficiency (%)	R^2
invA	–3.346	99.0	1.000
ttr	–3.358	98.5	1.000
siiA	–3.383	97.5	1.000
hilA	–3.587	90.0	0.980
spaQ	–3.361	98.4	1.000

method. On the other hand, the DNA of four enterobacteria groups closely related and not related to the species under study were also evaluated. Each group contained 1 ng/μL DNA of each species, and they were classified as *(i)* Gram-positive bacteria group: *Staphylococcus aureus* subsp. *aureus, Bacillus cereus, and Enterococcus faecalis; (ii) Shigella* group*: Shigella sonnei, Shigella boydii*, and *(iii)* SHEC *Escherichia* Group: *E. coli* O104: H4, *E. coli* O145: NM, *E. coli* O157: H7. As a positive control for this assay, a DNA mixture of 1 ng/μL of four serotypes of *Salmonella* spp.*, (S. typhimurium, S. typhi, S. derby,* and *S. paratyphi* A.) was used.

The specificity of the selected sequences was evaluated initially using an *in silico* analysis. Then, sixteen serotypes of *Salmonella* spp. (**Table 4**) were evaluated using a ddPCR assay. The results indicated that a positive amplification response was generated for all *Salmonella* serotypes and that negative amplification results were generated for all Enterobacteria strains evaluated.

TABLE 4 | Specificity analysis.

Serovar	Results for all targets
Bacillus cereus	–
Enterococcus faecalis	–
Escherichia coli O104:H4	–
Escherichia coli O145:NM	–
Escherichia. coli O157:H7	–
Proteus mirabilis	–
Proteus vulgaris	–
Shigella boydii	–
Shigella sonnei	–
Staphylococcus aureus subsp. aureus	–
Staphylococcus aureus subsp. aureus	–
Vibrio parahaemolyticus	–
Salmonella typhimurium	+
Salmonella typhi	+
Salmonella derby	+
Salmonella enteritidis	+
Salmonella dublin	+
Salmonella give	+
Salmonella saintpaul	+
Salmonella hadar	+
Salmonella infantis	+
Salmonella anatum	+
Salmonella panama	+
Salmonella typhimurium var5-	+
Salmonella javiana	+
Salmonella braenderup	+
Salmonella muenchen	+
Salmonella paratyphi A	+

Working Interval

Using *S. enteritidis* ATCC 13076 DNA, five WS from 10 to 0.001 ng/μL (as nominal concentration), were prepared. Serial gravimetric dilution from WS to ddPCR master mix reaction were measured in triplicate on three different days (see **Supplementary Tables S1–S5**). The working interval was determined using least squares regression analysis of the linear relationship between each dilution and their corresponding concentration (cp/μL) using the ddPCR method (**Table 5**). Dilution value is used to analyze the working interval as the nominal concentration is just an estimate of the real concentration. An example is presented in **Figure 1**. Grubbs' test was performed to determine outliers on one *hilA* level replicate. The results for the NTC were negative.

The acceptance criteria for linearity were as follows: a slope statistically different from zero and correlation coefficient >0.99 (Deprez et al., 2012). The method showed excellent linearity between 1 and 8,000 cp/μL in reaction (except for *siiA*, the interval of which was from 8 to 8,000 cp/μL). Therefore, the working interval for the whole method was established as 8–8,000 cp/μ L.

As part of the validation process, the ability of the method to run in a duplex form (*invA-ttr* and *hilA- siiA*) was also evaluated through regression analysis. Both methods were compared over the working interval, using data of four concentration levels (8–8,000 cp/μL) in triplicate and on three different days (**Figure 2**; see **Supplementary Table S6**). A slope statistically equal to 1 and correlation coefficient >0.99 were set as acceptance criteria.

According to **Figure 2**, the results obtained did not differ when *invA* and *ttr* were measured independently or in combination. Identical results were obtained for *hilA* and *siiA* target sequences (see **Supplementary Figure S2**). This allowed simultaneous evaluation of four different targets in two duplex reactions with the same confidence.

In addition, the amplitude heat plots of 1D and 2D duplex reactions showed clear populations of droplet distribution in positive and negative partitions, minimizing misclassification of droplets (rain) and absence of cross-reaction (**Figure 3**). As DNA distribution follows a random pattern, in the 2D plot, droplets are cluster in four groups: *(i)* double negatives, *(ii)* FAM positives and HEX negatives, *(iii)* FAM negatives and HEX positives, or *(iv)* double positives; not all positive partitions are double-positive in the duplex assay (**Figure 3C**) due to some possible fragmentation process that affect DNA integrity (Furuta-Hanawa et al., 2019); **Table 6** shows the variation of clusters fraction with DNA concentration.

TABLE 5 | Regression analysis for each target sequence evaluated at 95% confidence.

Gene	Slope	Intercept	R^2
invA	1.0094	6.8378	0.9996
ttr	1.0200	6.8474	0.9990
spaQ	0.9749	6.2832	0.9985
hilA	0.9707	6.2528	0.9985
*siiA**	1.0063	6.3995	0.9996

**siiA shows good linearity up to fourth Log dilution.*

Precision

Five concentration levels covering 1–8,000 cp/μL in ddPCR reaction, in triplicate reactions on three different days (see **Supplementary Tables S1–S5**), relative repeatability standard deviation (Eq. 4), and relative intermediate standard deviation (Eq. 5), were calculated using one factor ANOVA were calculated

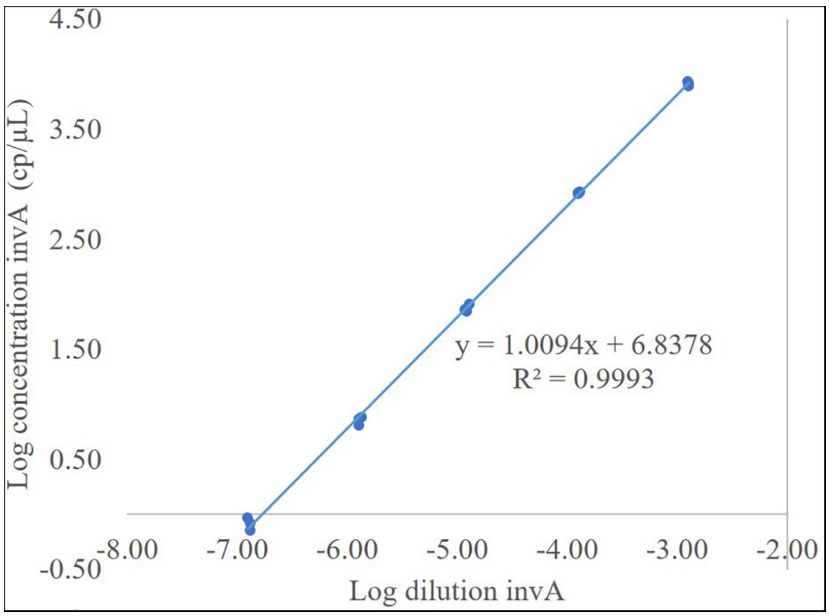


FIGURE 1 | The working interval for *invA*. Total DNA gravimetric dilution in Log scale vs. DNA concentration (cp/µL) in Log scale.

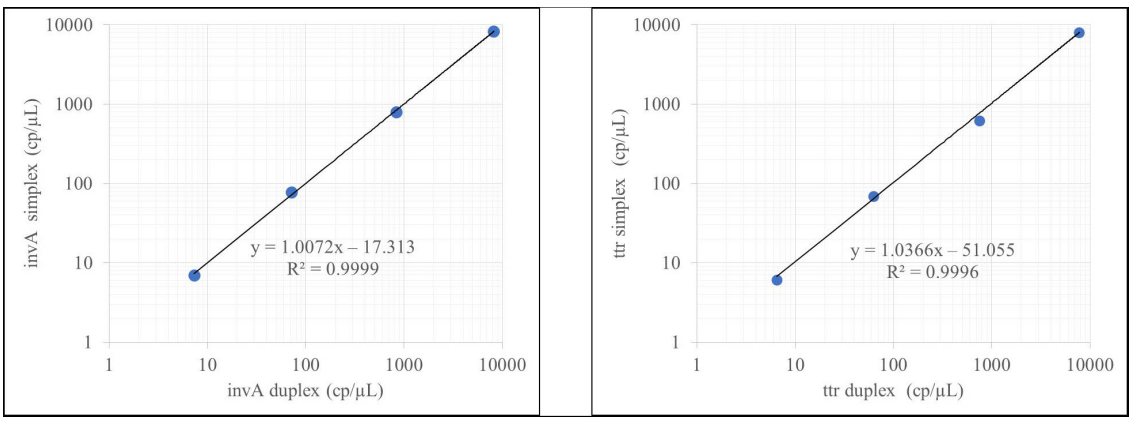


FIGURE 2 | Duplex evaluation for the *invA-ttr* assay.

using one-factor ANOVA (Deprez et al., 2016).

$$S_{repeat,rel} = \frac{\sqrt{MS_{witin\ run}}}{C_{sample,\ mean}} \quad (4)$$

$$S_{interm,\ rel} = \frac{\frac{\sqrt{MS_{between\ run} - MS_{\ within\ run}}}{n_{replicates}}}{C_{sample,\ mean}} \quad (5)$$

where

$MS_{witin\ run}$: Within run mean squares calculated using one-way ANOVA

$MS_{between\ run}$: Between run mean squares calculated using one-way ANOVA

$n_{replicates}$: Number of replicates per run

$C_{sample,\ mean}$: Averaged copy number concentration calculated over all runs.

Repeatability was below 5% for the three highest concentration levels and increased to 10% for the fourth one. However, for the lowest concentration (approximately 1 cp/µL in reaction), repeatability, was beyond 20% (**Figure 4A**), while intermediate precision was below 5% in almost the entire interval, except for *ttr,* and *siiA* (**Figure 4B**). When MS between <MS within, the $S_{interm,\ rel}$ was considered negligible with respect to $S_{repeat,rel}$.

LOQ

Limit of quantification represents the lowest copy number of the analyte that can be determined with acceptable performance.

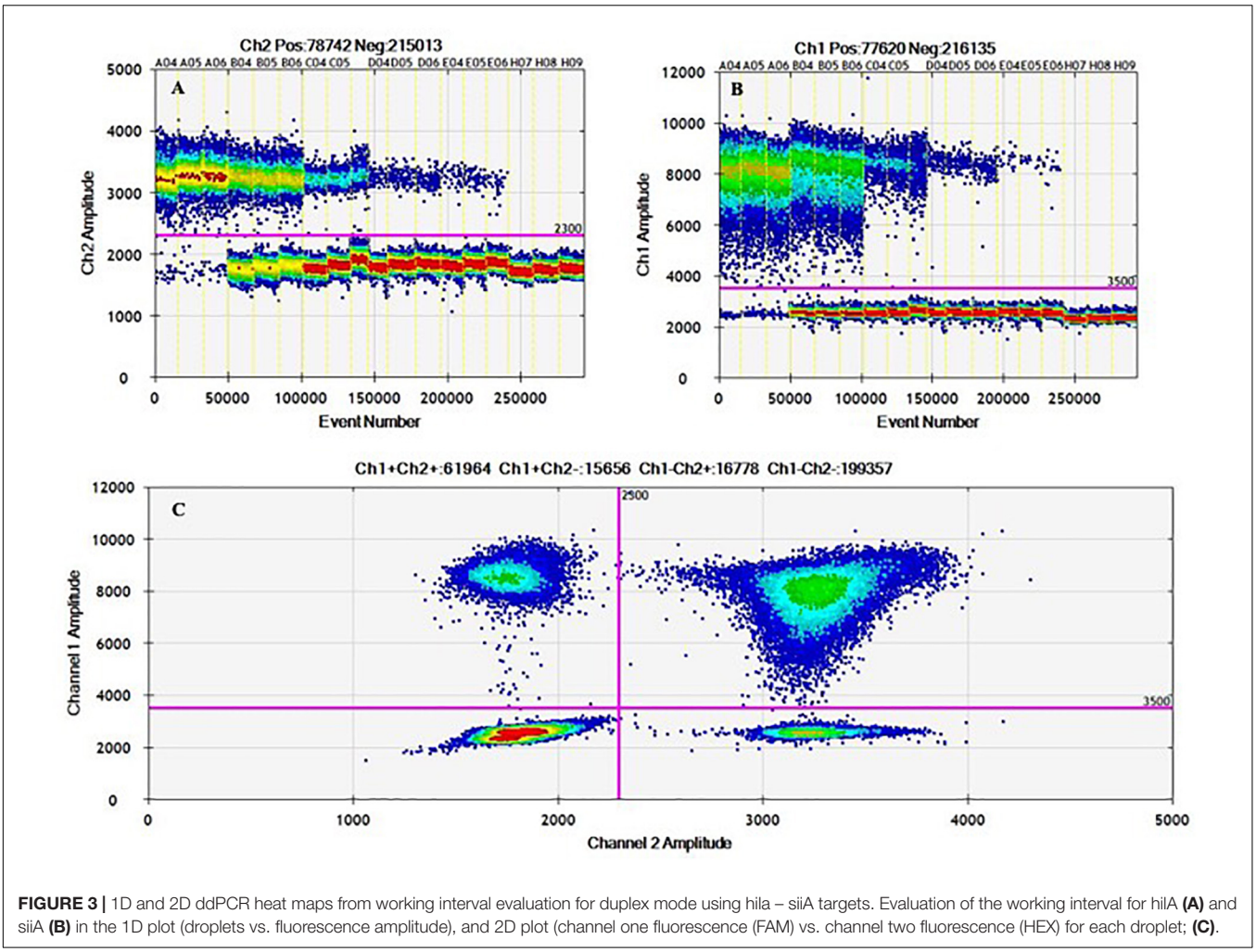


FIGURE 3 | 1D and 2D ddPCR heat maps from working interval evaluation for duplex mode using hila – siiA targets. Evaluation of the working interval for hilA **(A)** and siiA **(B)** in the 1D plot (droplets vs. fluorescence amplitude), and 2D plot (channel one fluorescence (FAM) vs. channel two fluorescence (HEX) for each droplet; **(C)**.

TABLE 6 | Relative fraction in respect to total partitions for every cluster in the duplex amplification for *hilA-siiA* targets.

Concentration level	*hilA+ siiA+* cluster	Relative fraction	*hilA+ siiA-* cluster	Relative fraction	*hilA- siiA+* cluster	Relative fraction	*hilA- siiA-* cluster	Relative fraction	Total partitions	Ratio *hilA/siiA*
1	50,162	99,6%	78	0,2%	125	0,2%	0	0,0%	50,365	0.93
2	11,582	22,5%	12,485	24,3%	13,290	25,8%	14,072	27,4%	51,429	0.96
3	215	0,5%	2,680	6,0%	2,878	6,4%	38,851	87,1%	44,624	0.93
4	5	0,0%	357	0,7%	340	0,7%	49,054	98,6%	49,756	1.05

We have defined the performance criteria as relative repeatability standard deviation <20%, according to the precision study. The higher values represent a significant increase in the measurement uncertainty. Thus, the LOQ based on the response of all the genes in this study was established as 8 cp/μL (**Figure 4A**).

LOD

Limit of detection is defined as the lowest analyte concentration that can be distinguished from zero with a specified level of confidence. LOD was determined as the lowest concentration level, where at least three positive droplets were present in all three replicates (Morisset et al., 2013), or equivalent in at least nine positive partitions in the pooled replicates. Amplifications were performed in a simplex mode for *spaQ* and duplex mode for *invA-ttr* and *hilA-siiA*. Five gravimetric serial dilutions (1–0.1 cp/μL in the ddPCR reaction) were prepared and run in triplicate. **Figure 5** presents the results in which each bar represents the total positive partitions obtained for three replicates compared to a total negative NTC. The estimated LOD was 0.5 cp/μL in the ddPCR reaction for three positive droplets (on average) in at least 12,000 total partitions.

Measurement Uncertainty

Based on precision data and considering the mathematical model, the measurement uncertainty estimation was established: first, all possible sources of uncertainty were considered (**Figure 6**), and

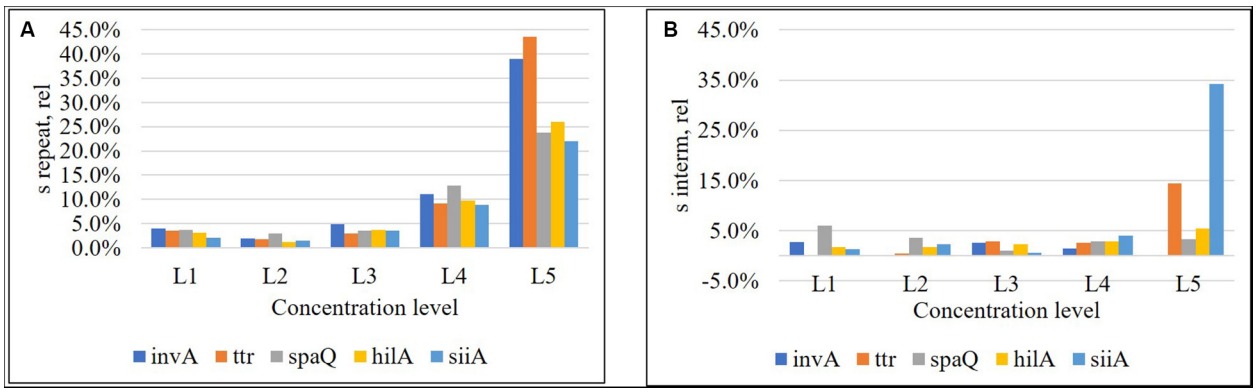


FIGURE 4 | Relative repeatability standard deviation **(A)** and relative intermediate standard deviation **(B)** of the method for each target sequence. Concentration levels were, L1: 8,000 cp/μL, L2: 800 cp/μL, L3: 80 cp/μL, L4: 8 cp/μL, and L5: 1 cp/μL.

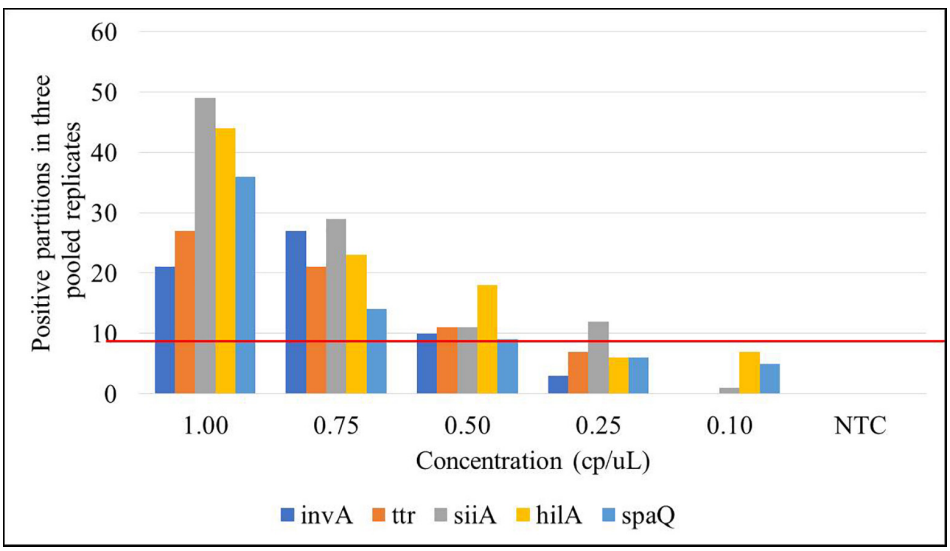


FIGURE 5 | Limit of detection (LOD) for the targets under study. Each bar represents the positive partitions over the three pooled replicates within the 5 concentration levels (L1: 8,000 cp/μL, L2: 800 cp/μL, L3: 80 cp/μL, L4: 8 cp/μL, and L5: 1 cp/μL) compared to the no template control (NTC). Red line represents the selected threshold: nine positive partitions.

they were grouped according to their relation and quantified. Then, all contributions were combined to have standard uncertainty (Eq. 6; Joint Committee for Guides in Metrology [JCGM], 2008).

$$u_{sample} = \sqrt{u^2_{Model} + u^2_{Precision}} \quad (6)$$

where

u^2_{Model} corresponds to the uncertainty contribution provided by the mathematical model (Eq. 2) of the measuring method and includes the uncertainty provided by the partitioning of the sample (λ factor), gravimetric dilution, and droplet volume (Eq. 7),

$$u_{Model} = C_{sample} \times \sqrt{\left(\frac{u_\lambda}{\lambda}\right)^2 + \left(\frac{u_D}{D}\right)^2 + \left(\frac{u_v}{v}\right)^2} \quad (7)$$

and $u^2_{Precision}$, which corresponds to the precision uncertainty provided by the total variation of each concentration level, was defined as the highest degree of dispersion obtained between replicates s^2_{repeat} and between different days s^2_{interm} (Eq. 8),

$$u_{Precision} = \sqrt{\frac{s^2_{repeat} + s^2_{interm}}{n}} \quad (8)$$

n: Days of measurement.

According to the validation data for all target sequences, the relative measurement uncertainty associated with the mathematical model varied between 1,6 and 9% (**Table 7**). The maximum contribution corresponded to λ factor (50–99%) at all levels (see **Supplementary Figure S3**). After combining the mathematical model and measurement uncertainty contributions (Eq. 6), the combined uncertainty was calculated for every

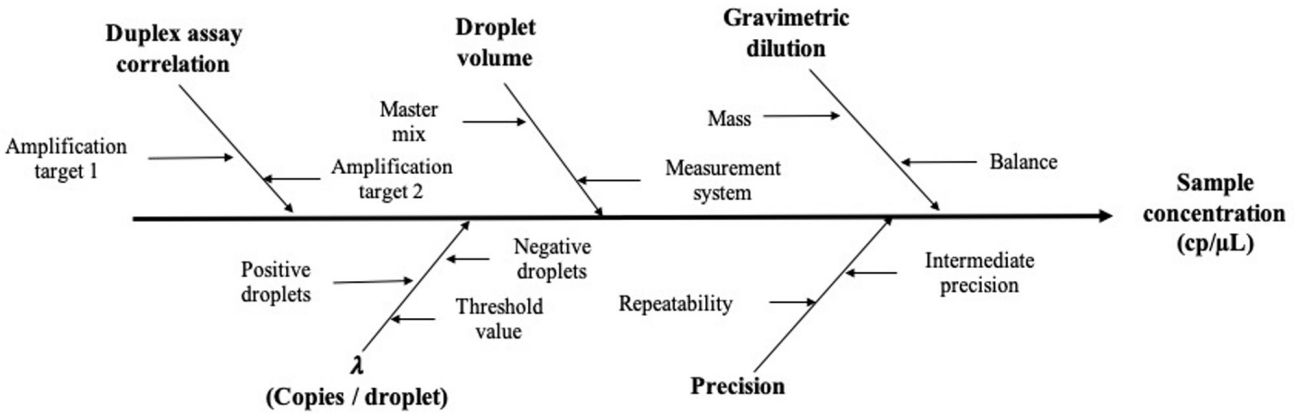


FIGURE 6 | Schematic representation of the main factors affecting the measurement value and its uncertainty.

TABLE 7 | Relative measurement uncertainty using the mathematical model for the five targets.

Level	Concentration (cp/uL)	Relative uncertainty (%)
1	8,000	3.6
2	800	1.6
3	80	3.8
4	8	9.0

Data correspond to an average value.

target in every concentration level (**Figure 7**). The relative standard uncertainty value ranged between 2 and 12%, while its lower value was obtained at the concentration of 800 cp/µL in the ddPCR reaction.

Multiplex Correlation Analysis

The correlation between the targets *invA-ttr* and *hilA-siiA* was evaluated to determine the contribution source to the uncertainty associated with covariance between target sequences in duplex assays. Considering that all target sequences are single copy, the correlation was evaluated, keeping the primers concentration constant for one target sequence (600 nM), while this was varied for the other one as 0, 150, 300, 450, 600, 750, and 900 nM. In all cases, the probe concentration was 300 nM. The slope and correlation coefficients were calculated and their significance analyzed using the student's *t*-test with a 95% confidence.

As the slopes were statistically equal to zero and the correlation were not significant (**Figure 8** and **Supplementary Figure S4**) for any of the target sequences, then, there is not a covariance contribution to the combined uncertainty (Joint Committee for Guides in Metrology [JCGM], 2008).

DISCUSSION

Digital PCR is a relatively new and promising technology, which is currently being used for detecting pathogens in food (Quan et al., 2018). The molecular quantification of specific genes in food samples is a useful tool for evaluating the quality and safety of food products in time-, effort-, and cost-effective manner to generate reliable results (Huber, 2007; Morisset et al., 2013). At the same time, digital PCR is considered a potential primary method of measurement in chemical metrology, allowing among its multiple applications, developing reference materials in bioanalysis especially at the clinical and industrial level (Whale et al., 2018).

The method reported here uses five targets of common use for the identification of *Salmonella* spp. using ddPCR. Two assays in duplex mode (*invA-ttr* and *hilA-siiA*) and one in simplex mode (*spaQ*) were used. The qPCR dilution curves showed that the DNA extraction and purification processes were suitable for obtaining DNA of good quality, which is necessary for validation. The amplification efficiency confirmed the absence of inhibitors that can affect the digital PCR validation process.

2D plots for duplex amplification (**Figure 3**), shows there is an apparent DNA integrity change, and some variation with DNA concentration. From **Table 6**, most positive droplets are double-positive for both target sequences (99.6%) at 8,000 cp/µL, but this proportion decrease with concentration level, passing to 22.5 and 0.5% in 2 log DNA concentrations (**Table 6** and **Supplementary Table S7**), indicating a possible fragmentation process in the DNA. The observed clusters relative fraction depends on the fraction of positive droplets. For a fraction higher than 90%, according to Eq. 3, there will be 2 or more copies per partition while for lower ratios, there will be lower occupancy degree of DNA molecules. Droplets with less amplitude for the double positive cluster could indicate an imbalanced amplification process between hilA and siiA DNA target sequences for a non-competing duplex reaction (Whale et al., 2016). However, the ration between hilA/siiA and invA/ttr target sequences keeps around 1, indicating that neither this fragmentation process, nor the imbalance amplification process are enough to affect the quantification in duplex form by digital PCR.

Although four targets (*invA, ttr, hilA,* and *spaQ*) showed good linearity over 5 log concentrations (covering 1 cp/µL, according to linear regression analysis), the working interval goes from 8,000 to 8 cp/µL in ddPCR reaction, where the lowest concentration corresponded to the LOQ, which was defined

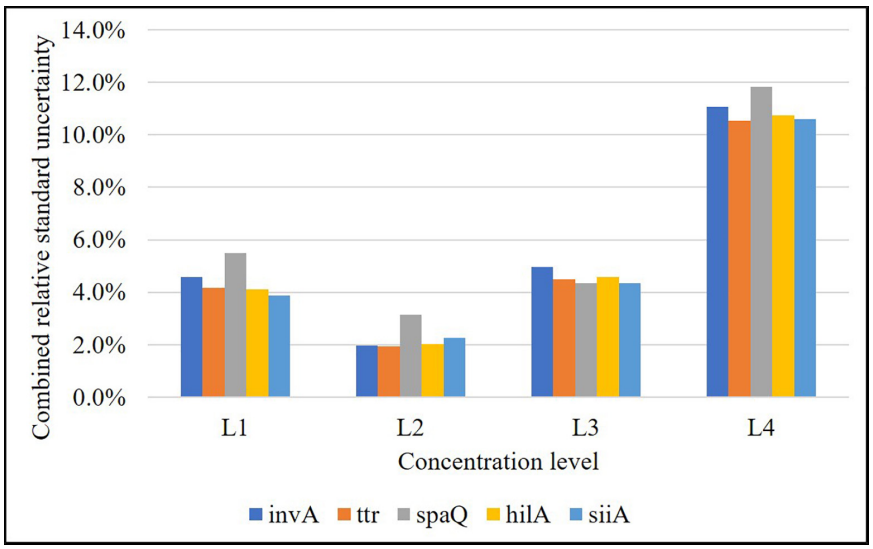


FIGURE 7 | Relative combined uncertainty for each target in each concentration level over the working range (L1: 8,000 cp/μL, L2: 800 cp/μL, L3: 80 cp/μL, and L4: 8 cp/μL in reaction). Level 2 is the one with the lowest uncertainty (approximately 2%).

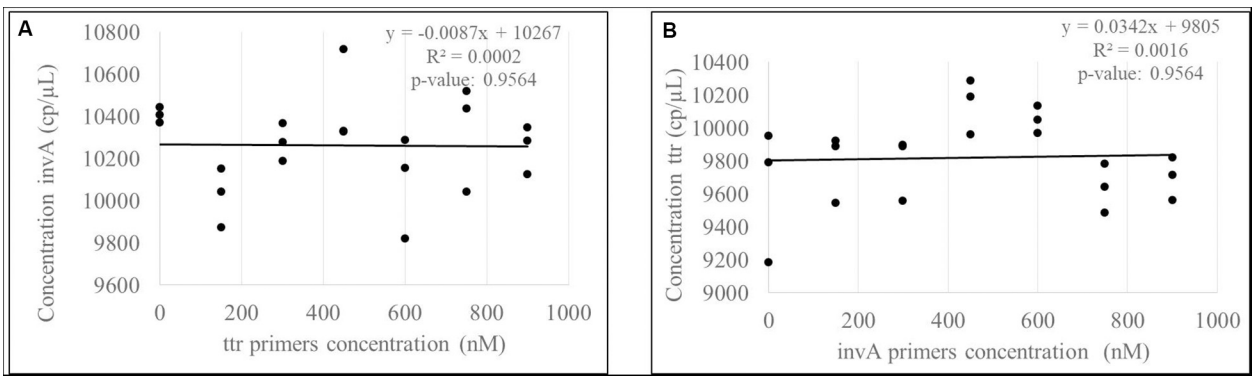


FIGURE 8 | Evaluation of correlation between *invA* **(A)** -*ttr* **(B)** multiplex amplification. According to *p*-values, there are not a significant correlation between targets.

based on precision criteria of maximum 20% as a relative repeatability standard deviation. Precision values higher than 20% (as they were obtained for the lowest levels), contributed more to combined measurement uncertainty.

The LOD, based on the number of positive partitions (at least three in each of the three replicates) and required to differentiate a positive sample of low concentration from the blank, was established as 0.5 cp/μL in reaction, which equals 5 cp/μL, approximately in the stock/sample solution (according to the gravimetric dilution). This would be very important if the objective is diagnostic because it will allow taking decisions in advance. After all, the target sequences used would be enough to differentiate a positive from a negative sample at a very low concentration. Nevertheless, this fact confirms one of the most critical characteristics of dPCR, its sensitivity. Theoretically, it will allow the detection of as few as 5 live/dead cells in the sample. However, the sensitivity would be determined by other factors, especially the sample homogeneity and the DNA extraction efficiency.

The specificity test on 16 different *Salmonella* serotypes and other related microorganisms, indicated that these sequences could be used together as potential biomarkers for the detection and quantification of the genus. These five genes are of great importance for detecting *Salmonella*. In total, 329 isolates from environmental and food samples, containing 126 serovars belonging to all subspecies of *Salmonella* were identified using *invA* (Malorny et al., 2007; Kasturi and Drgon, 2017). The *ttr* locus has also been used as an essential molecular marker for reducing false-positive results (Ssemanda et al., 2018; Boer et al., 2019).

As shown in **Figure 4**, the relative repeatability standard deviation of the assay varies from 2 to 12% and correlates with the uncertainty according to the mathematical model of the method. Concentration level 2 (800 cp/μL) presents the best precision and also has the lowest uncertainty associated

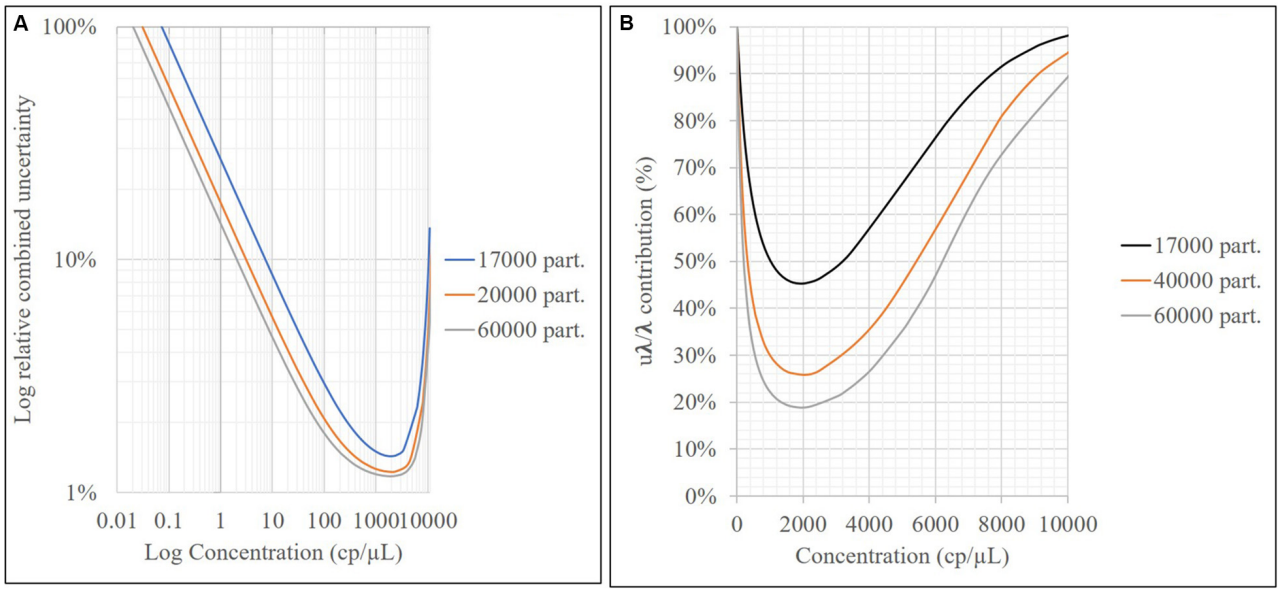


FIGURE 9 | Relative uncertainty associated with the model **(A)** and contribution of $u\lambda/\lambda$ to mathematical model uncertainty vs. the sample concentration (cp/µL) for different partition numbers **(B)**.

with the model, indicating that it is the best concentration to measure. The mathematical model contribution (Eq. 7) is a function of the number of positive or negative fractions. Subsequently, once the assay has been optimized, it is possible to assign a relative uncertainty component to the sample based on its concentration (**Figure 9A**). $u\lambda/\lambda$ is the predominant measurement uncertainty source within the mathematical model, ranging from 50 to 99%. The partition number can be increased to improve this estimation. In the QX200 platform (BioRad, 2018), the maximum partitions per well were 20,000; however, several wells in the Quantasoft™ software can be merged to increase the partition number. This will in turn reduce the $u\lambda/\lambda$ contribution (**Figure 9B**), with consequent reduction in combined measurement uncertainty.

Measurement uncertainty contribution was not associated with the correlation between target amplification in the duplex assays. This can be attributed to the primer sequence, which amplifies different regions inside the genome and does not inter-hybridize. Therefore, each primer can function independently.

CONCLUSION

After a literature review and *in silico* analysis of the primary target sequences used for the quantification of *Salmonella* spp., five different amplification targets, namely *invA, ttr, hilA, spaQ*, and *siiA*, were selected, which belonged to five genes involved in pathogenicity and metabolism of the bacteria.

According to our results, the method presented here is suitable for estimating the genomic DNA content of *Salmonella* spp. with five different sequences in simplex and duplex form, from 8 to 8,000 cp/µL, with a detection limit of 0.5 copies/µL in the ddPCR reaction, LOQ of 8 cp/µL, precision between 5 and 10%, and combined relative standard measurement uncertainty between 2.0 and 12% over the working interval considering all sources that contribute to it. The most important factor contributing to measurement uncertainty comes from the mathematical model, specifically from the λ factor. To decrease the measurement uncertainty, some wells (at least two) can be merged using the Quantasoft™ software, which will increase the partition number, especially at low concentrations close to the LOQ.

Nevertheless, the unique feature of this assay that enables the detection of five genes instead of one or two ensures higher accuracy and confidence in the results. While a single gene can indicate the presence or absence of *Salmonella* spp., adding other molecular targets provides tools for more reliable quantification, especially for some specific applications such as reference material characterization.

This study contributes to the strengthening of monitoring processes and the increase in competitiveness in the marketing of agricultural industry via developing of metrological assurance tools as validated measurement methods for the identification and quantification of *Salmonella* spp. which is one of the most critical foodborne disease agents. This tool can also be used for the assignment of a reference value for a reference material based on the different sequences evaluated.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

CV performed the experiments and wrote the original draft. MC was involved in conceptualization, investigation, and funding acquisition. MA was involved in validation, writing, reviewing, and editing. JL designed the methodology, wrote the original draft, and was involved in project administration. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by Departamento Administrativo de Ciencia, Tecnología e Innovación (Colciencias), (project 110174558815, contract 018:2017), and and División de Investigación Sede Bogotá, Universidad Nacional de Colombia (DIB; Call 2016–2018, project 37670).

ACKNOWLEDGMENTS

We acknowledge the Instituto Nacional de Salud (INS) for kindly providing *Salmonella* references and clinical strains.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2020.01512/full#supplementary-material

REFERENCES

Atashpaz, S., Khani, S., Barzegari, A., Barar, J., Vahed, S. Z., Azarbaijani, R., et al. (2010). A robust universal method for extraction of genomic DNA from bacterial species. *Microbiology* 79, 538–542. doi: 10.1134/S0026261710040168

Ben Hassena, A., Barkallah, M., Fendri, I., Grosset, N., Ben Neila, I., Gautier, M., et al. (2015). Real time PCR gene profiling and detection of *Salmonella* using a novel target: the siiA gene. *J. Microbiol. Methods* 109, 9–15. doi: 10.1016/j.mimet.2014.11.018

BioRad (2018). *Droplet Digital PCR Applications Guide*. Available online at: http://www.bio-rad.com/webroot/web/pdf/lsr/literature/Bulletin_6407.pdf (accessed April 18, 2020).

Boddicker, J. D., Knosp, B. M., and Jones, B. D. (2003). Transcription of the *Salmonella* invasion gene activator, hilA, requires HilD activation in the absence of negative regulators. *J. Bacteriol.* 185, 525–533. doi: 10.1128/JB.185.2.525

Boer, M. D., de Boer, R. F., Lameijer, A., Sterne, E., Skidmore, B., Suijkerbuijk, A. W. M., et al. (2019). Selenite enrichment broth to improve the sensitivity in molecular diagnostics of *Salmonella*. *J. Microbiol. Methods* 157, 59–64. doi: 10.1016/j.mimet.2018.12.018

Corbisier, P., Pinheiro, L., Mazoua, S., Kortekaas, A.-M., Chung, P. Y. J., Gerganova, T., et al. (2015). DNA copy number concentration measured by digital and droplet digital quantitative PCR using certified reference materials. *Anal. Bioanal. Chem.* 407, 1831–1840. doi: 10.1007/s00216-015-8458-z

Dagata, J. A., Farkas, N., and Kramer, J. A. (2016). *Method for Measuring The Volume Of Nominally 100 μM Diameter Spherical Water-In-Oil Emulsion Droplets*. Gaithersburg, MD: NIST.

Deprez, L., Corbisier, P., Kortekaas, A. M., Mazoua, S., Beaz Hidalgo, R., Trapmann, S., et al. (2016). Validation of a digital PCR method for quantification of DNA copy number concentrations by using a certified reference material. *Biomol. Detect. Q.* 9, 29–39. doi: 10.1016/j.bdq.2016.08.002

Deprez, L., Mazoua, S., Corbisier, P., Trapmann, S., Schimmel, H., White, H., et al. (2012). *The Certification Of The Copy Number Concentration Of Solutions Of Plasmid DNA Containing A BCR-ABL b3a2 Transcript Fragment*. Luxemburgo: Publications Office of the EU.

Ekiri, A. B., Long, M. T., and Hernandez, J. A. (2016). Diagnostic performance and application of a real-time PCR assay for the detection of *Salmonella* in fecal samples collected from hospitalized horses with or without signs of gastrointestinal tract disease. *Vet. J.* 208, 28–32. doi: 10.1016/j.tvjl.2015.11.011

Emslie, K. R., McLaughlin, J., Griffiths, K., Forbes-Smith, M., Pinheiro, L., and Burke, D. (2019). Droplet volume variability and impact on digital PCR copy number concentration measurements. *Anal. Chem.* 91, 4124–4131. doi: 10.1021/acs.analchem.8b05828

European Comission (2007). *Regulation (EC) No 2073/2005 Regarding Microbiological Criteria Specific To Food Products*. Madrid: Official Journal of the European Union.

Furuta-Hanawa, B., Yamaguchi, T., and Uchida, E. (2019). Two-dimensional droplet digital pcr as a tool for titration and integrity evaluation of recombinant adeno-associated viral vectors. *Hum. Gene. Ther. Methods.* 30, 127–136. doi: 10.1089/hgtb.2019.031

Galan, J. E., Ginocchio, C., and Costeas, P. (1992). Molecular and functional characterization of the *Salmonella* invasion gene inv a: homology of InvA to members of a new protein family. *J. Bacteriol.* 174, 4338–4349. doi: 10.1128/jb.174.13.4338-4349.1992

González-Escalona, N., Brown, E. W., and Zhang, G. (2012). Development and evaluation of a multiplex real-time PCR (qPCR) assay targeting ttrRSBCA locus and invA gene for accurate detection of *Salmonella* spp. in fresh produce and eggs. *Food Res. Int.* 48, 202–208. doi: 10.1016/j.foodres.2012.03.009

Gutierrez-Aguirre, I., Racki, N., Dreo, T., and Ravnikar, M. (2015). Droplet digital PCR for absolute quantification of pathogens. *Methods Mol. Biol.* 1302, 331–347. doi: 10.1007/978-1-4939-2620-6_24

Huber, L. (2007). *Validation and Qualification In Analytical Laboratories*. London: Informa Healthcare.

Huggett, J. F., Foy, C. A., Benes, V., Emslie, K., Garson, J. A., Haynes, R., et al. (2013). The digital MIQE guidelines: minimum information for publication of quantitative digital PCR experiments. *Clin. Chem.* 59, 892–902. doi: 10.1373/clinchem.2013.206375

Joint Committee for Guides in Metrology [JCGM] (2008). *ISO/IEC Guide 98-3. JCGM 100. Evaluation of Measurement Data — Guide to the Expression of Uncertainty in Measurement*. Available online at: https://www.bipm.org/en/publications/guides/#gum (accessed April 10, 2020).

Kasturi, K. N., and Drgon, T. (2017). Real-time PCR method for detection of *Salmonella* spp. in environmental samples. *Appl. Environ. Microbiol.* 83, 1–12. doi: 10.1128/aem.00644-17

Košir, A. B., Divieto, C., Pavšič, J., Pavarelli, E., Dobnik, D., Dreo, T., et al. (2017). Droplet volume variability as a critical factor for accuracy of absolute quantification using droplet digital PCR. *Anal. Bioanal. Chem.* 409, 6689–6697. doi: 10.1007/s00216-017-0625-y

Kurowski, B., Traub-dargatz, J. L., Morley, P. S., and Gentry-weeks, C. R. (2002). Detection of *Salmonella* spp. in fecal specimens by use of real-time polymerase chain reaction. *Am J. Vet. Res.* 63, 1265–1268. doi: 10.2460/ajvr.2002.63.1265

Law, J. W.-F., Ab Mutalib, N.-S., Chan, K.-G., and Lee, L.-H. (2014). Rapid methods for the detection of foodborne bacterial pathogens: principles, applications, advantages and limitations. *Front. Microbiol.* 5:770. doi: 10.3389/fmicb.2014.00770

Magnusson, B., and Örnemark, U. (2014). *Eurachem Guide: The Fitness for Purpose of Analytical Methods – A Laboratory Guide to Method Validation and Related Topics*. London: Eurachem Guide.

Malorny, B., Bunge, C., and Helmuth, R. (2007). A real-time PCR for the detection of *Salmonella enteritidis* in poultry meat and consumption eggs. *J. Microbiol. Methods* 70, 245–251. doi: 10.1016/j.mimet.2007.04.013

McCabe, E. M., Burgess, C. M., O'Regan, E., McGuinness, S., Barry, T., Fanning, S., et al. (2011). Development and evaluation of DNA and RNA real-time assays for food analysis using the hilA gene of *Salmonella enterica* subspecies enterica. *Food Microbiol.* 28, 447–456. doi: 10.1016/j.fm.2010.10.012

Morisset, D., Stebih, D., Milavec, M., Gruden, K., and Zel, J. (2013). Quantitative analysis of food and feed samples with droplet digital PCR. *PLoS One* 8:e062583. doi: 10.1371/journal.pone.0062583

Quan, P. L., Sauzade, M., and Brouzes, E. (2018). dPCR: a technology review. *Sensors* 18, 1271–1298. doi: 10.3390/s18041271

Rothrock, M. J., Hiett, K. L., Kiepper, B. H., Ingram, K., and Hinton, A. (2013). Quantification of zoonotic bacterial pathogens within commercial poultry processing water samples using droplet digital PCR. *Adv. Microbiol.* 3, 403–411. doi: 10.4236/aim.2013.35055

Somanath, B., and Kerry, R. (2016). Digital polymerase chain reaction for characterization of DNA reference materials. *BDQ* 10, 47–49. doi: 10.1016/j.bdq.2016.04.001

Ssemanda, J. N., Reij, M. W., van Middendorp, G., Bouw, E., van der Plaats, R., Franz, E., et al. (2018). Foodborne pathogens and their risk exposure factors associated with farm vegetables in Rwanda. *Food Control* 89, 86–96. doi: 10.1016/j.foodcont.2017.12.034

Wang, H., Jiang, Y., Liu, X., Qian, W., Xu, X., and Zhou, G. (2016). Behavior variability of *Salmonella enterica* isolates from meat-related sources. *LWT Food Sci. Technol.* 73, 375–382. doi: 10.1016/j.lwt.2016.06.036

Whale, A. S., Hugget, J. F., and Tzonev, Z. (2016). Fundamentals of multiplexing with digital PCR. *BDQ* 10, 15–23. doi: 10.1016/j.bdq.2016.05.002

Whale, A. S., Jones, G. M., Pavšič, J., Dreo, T., Redshaw, N., Akyürek, S., et al. (2018). Assessment of digital PCR as a primary reference measurement procedure to support advances in precision medicine. *Clin. Chem.* 64, 1296–1307. doi: 10.1373/clinchem.2017.285478

White, H., Deprez, L., Corbisier, P., Hall, V., Lin, F., Mazoua, S., et al. (2015). A certified plasmid reference material for the standardisation of BCR-ABL1 mRNA quantification by real-time quantitative PCR. *Leukemia* 29, 369–376. doi: 10.1038/leu.2014.217

World Health Organization (2015). *WHO Estimates Of The Global Burden Of Foodborne Diseases*. Geneva: WHO.

Zhao, Y., Jiang, X., Qu, Y., Pan, R., Pang, X., Jiang, Y., et al. (2017). *Salmonella* detection in powdered dairy products using a novel molecular tool. *J. Dairy Sci.* 100, 3480–3496. doi: 10.3168/jds.2016-12535

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Villamil, Calderon, Arias and Leguizamon. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 04 August 2020
doi: 10.3389/fmicb.2020.01482

Prevalence, Serotype, Antibiotic Susceptibility, and Genotype of *Salmonella* in Eggs From Poultry Farms and Marketplaces in Yangling, Shaanxi Province, China

Wei Li[1†], Hao Li[1†], Shujuan Zheng[1], Zewei Wang[1], Huanjing Sheng[1], Chunlei Shi[2], Xianming Shi[2], Qinya Niu[1] and Baowei Yang[1]*

[1] College of Food Science and Engineering, Northwest A&F University, Yangling, China, [2] MOST-USDA Joint Research Center for Food Safety, School of Agriculture and Biology, Shanghai Jiao Tong University, Shanghai, China

Poultry products such as eggs provide essential nutrients to the human body and thus play vital roles in the human food network. *Salmonella* is one of the most notorious foodborne pathogens and has been found to be prevalent in eggs. To better understand the characteristics of *Salmonella* in eggs, we investigated the prevalence of *Salmonella* spp. in 814 fresh eggs collected from poultry farms and retail marketplaces in Yangling, Shaanxi Province, China. The serotype, genotype, and antibiotic susceptibilities of 61 *Salmonella* isolates recovered from the eggs were analyzed. The average detection rate of *Salmonella*-positive eggs was 5.6%, with 6.6% of the eggs collected from poultry farms and 5.1% from marketplaces. Thirteen serotypes were identified from the 61 isolates, among which *Salmonella* Typhimurium (24.5%) and *Salmonella* Indiana (22.9%) were the most prevalent serotypes. Other dominant serotypes included *Salmonella* Thompson (13.1%) and *Salmonella* Enteritidis (11.4%), with the remaining nine serotypes detected at low rates (1.6–4.9%). All the *Salmonella* isolates tested were resistant to sulfisoxazole (100.0%). The majority (77.1%) of the isolates were resistant to nalidixic acid, amoxicillin-clavulanate, and ampicillin, while nearly two-thirds (63.9–68.9%) were resistant to trimethoprim–sulfamethoxazole, kanamycin, tetracyclines, and chloramphenicol. The rate of resistance to ciprofloxacin was 40.1%; the resistance rates to streptomycin, ceftiofur, and ceftriaxone ranged from 21.3 to 26.2%; and those to gentamicin, amikacin, and cefoxitin were relatively low (3.3–16.4%). Forty-nine (80.3%) *Salmonella* isolates exhibited resistance to multiple antibiotics, 20 (32.8%) of which were resistant to at least 10 antibiotics. Subtyping by pulse-field gel electrophoresis revealed a close genetic relatedness of *Salmonella* isolates from poultry farms, in striking contrast to the high diversity of the isolates obtained from marketplaces. Isolates of the same serotype always shared identical genotype and antibiotic resistance profiles, even the ones that were recovered from eggs sampled at different locations and times. These findings indicate that diverse *Salmonella* spp. with high rates of multidrug resistance are prevalent in fresh eggs in the study area. More attention should be paid to egg production, transportation, and storage to prevent foodborne outbreaks caused by *Salmonella*.

Keywords: antibiotic resistance, eggs, genotype, *Salmonella*, serotype

OPEN ACCESS

Edited by:
Dario De Medici,
Istituto Superiore di Sanità (ISS), Italy

Reviewed by:
Min Yue,
Zhejiang University, China
Alejandro Castillo,
Texas A&M University, United States

****Correspondence:***
Baowei Yang
ybw090925@163.com

† These authors have contributed equally to this work and share first authorship

Specialty section:
This article was submitted to Food Microbiology, a section of the journal Frontiers in Microbiology

Received: *17 March 2020*
Accepted: *08 June 2020*
Published: *04 August 2020*

Citation:
Li W, Li H, Zheng S, Wang Z, Sheng H, Shi C, Shi X, Niu Q and Yang B (2020) Prevalence, Serotype, Antibiotic Susceptibility, and Genotype of Salmonella in Eggs From Poultry Farms and Marketplaces in Yangling, Shaanxi Province, China. Front. Microbiol. 11:1482. doi: 10.3389/fmicb.2020.01482

INTRODUCTION

Salmonella spp. are notorious foodborne pathogens that can cause diarrhea in humans and animals (Cowen et al., 2016). Reportedly, there are approximately 9.38 million cases of *Salmonella* infections and 15,000 deaths from the infection worldwide every year (Patchanee et al., 2017). Among the various vehicles of *Salmonella*, poultry and poultry products are not only identified as the predominant reservoirs, but are also considered to be the primary sources of human salmonellosis based on evidence from epidemiological traceback investigations (Andino and Hanning, 2015; McWhorter and Chousalkar, 2019). A previous study has indicated that the egg white possesses unique physical and biochemical properties, creating a complex antimicrobial environment to resist antigens (Huang et al., 2019). However, egg white and egg yolk membranes are the major infection sites for *Salmonella*, which means that *Salmonella* still has the opportunity to survive in the resistant environment of eggs (Raspoet et al., 2019).

Currently, the outbreak of foodborne diseases caused by the consumption of eggs contaminated with *Salmonella* remains severe worldwide. According to reports from the United States Food and Drug Administration (FDA) and the Centers for Disease Control and Prevention (CDC), there were 52 foodborne outbreaks in Missouri in 2015 due to eggs contaminated by *Salmonella* Oranienburg. In 2016, the CDC again notified the FDA of eight clinical cases from three states, i.e., Illinois, Kansas, and Missouri, which were closely related to the hereditary strains of the *S.* Oranienburg outbreak in 2015 (FDA, 2019c). In March 2018, the FDA learned about 45 *Salmonella* Braenderup-infected consumers in 10 states, 11 of whom were hospitalized without death. The outbreak was tracked down to Rose Acre Farms' Hyde County farm in North Carolina and 207 million eggs were recalled around the United States (FDA, 2019a).

Recently, *Salmonella* was detected in eggs collected from 41 (63.5%) of 63 farms in Australia that underwent environmental sampling (Moffatt et al., 2017). In addition, 88.6% (124/140) of the eggs and all of the 19 farms sampled in the western region of Cameroon were found to be *Salmonella*-positive (Kouam et al., 2018). So far, multiple *Salmonella* serotypes including *Salmonella* Typhimurium, *Salmonella* Indiana, and *Salmonella* Enteritidis have been commonly detected from eggs, poultry, and poultry farm environments (Moffatt et al., 2017). To prevent infection by pathogenic bacteria and promote the growth of laying hens, a large number of antibiotics are used in the feeding process. A nationwide market survey indicated that the total amount of 36 antibiotics that are frequently detected in livestock farms, wastewater treatment plants, and environment settings reached 92,700 tons in China in 2015. Notably, the rate of antibiotic usage in poultry farms was as high as 19.6% (18,100/92,700) (Zhang et al., 2015).

β-Lactams, cephalosporins, and fluoroquinolones are the most frequently used antibiotics in the poultry industry (Fardsanei et al., 2017). Taking into account the antibiotics used in other fields, *Salmonella* has developed high resistance to a broad range of antibiotics, leading to increased healthcare costs and clinical treatment failure (Cui et al., 2016). Many studies have characterized *Salmonella* in poultry and poultry-associated foods globally, especially eggs (Moffatt et al., 2017; Kouam et al., 2018; Li et al., 2018; FDA, 2019a,b,c; Karimiazar et al., 2019; Kingsbury et al., 2019; Sanchez-Salazar et al., 2019). Evidence suggests that the genetic diversity of *Salmonella* isolated from various sources is abundant, whereas the isolates from the same poultry farms are closely related (Xie et al., 2019). Moreover, the occurrence and characteristics of *Salmonella* in poultry products including eggs always show dynamic variation during production, storage, and handling (Li et al., 2018). Therefore, it is of great significance to continuously monitor this group of pathogens in eggs to ensure food safety.

In this study, we characterized *Salmonella* isolates from fresh eggs in poultry farms and retail marketplaces in Yangling, Shaanxi Province, China. The objective of the study was to further explore the prevalence and transmission of *Salmonella* during egg production and sales links in order to prevent *Salmonella* outbreaks.

MATERIALS AND METHODS

Sample Collection

A total of 814 fresh eggs were collected in Yangling and its surrounding districts in Shaanxi Province, China, from mid-2013 to late 2014. Specifically, 304 eggs were collected from three different poultry farms. Farm X in Xiajiagou Village is a large modern laying hen farm with fully automated cage raising equipment, including equipment for heating, ventilation, water supply, feeding, egg collection, manure removal, cages, and lighting. In this farm, there are approximately 20,000 laying hens with a breeding density of 9–10/m^2. Farm C in Cuixigou Village is a semi-automated laying hen farm of a medium production scale, with basic equipment for heating, ventilation, cages, lighting, and water feeding. In this farm, there are approximately 9000 laying hens with a breeding density of 11–12/m^2. Farm F in Fuzeyuan Company is the smallest laying hen farm with basic equipment for heating, ventilation, cages, and lighting. In this farm, there are approximately 4000 laying hens with a breeding density of 9–10/m^2. Additionally, 510 eggs of seven brands were collected from seven supermarkets (eight retail outlets) and four wet markets (10 retail outlets).

At each sampling location, the eggs were sampled weekly in August, September, and October 2013 and in March, June, August, October, and November 2014. At each sampling time, three, six, or nine eggs (depending on the total number of eggs for sale) were collected at random in a supermarket or wet market. In each poultry farm, no more than 50 eggs were collected at random from egg-laying troughs. The egg samples were delivered to the Microbiology Laboratory in the College of Food Science and Engineering, Northwest A&F University (Yangling, Shaanxi Province China) for *Salmonella* isolation within 12 h of collection.

Isolation and Identification of *Salmonella*

Salmonella isolates were recovered from egg samples according to the method described by Yang et al. (2010) with some

TABLE 1 | Antibiotic concentration ranges and data interpretation for the susceptibility test of *Salmonella* isolates.

Antibiotic agent	Abbreviation	Antibiotic concentration range (μg/mL)	Breakpoint interpretive criteria (μg/mL)*		
			S	I	R
Aminoglycosides					
Amikacin	AMK	0.5–64	≤16	32	≥64
Gentamicin	GEN	0.25–16	≤4	8	≥16
Kanamycin	KAN	1–64	≤16	32	≥64
Streptomycin #	STR	4–64	≤32	NA	≥64
Penicillin β-lactamase inhibitor combinations					
Amoxicillin–clavulanic acid	AMC	0.25/0.12–32/16	≤8/4	16/8	≥32/16
Ampicillin	AMP	0.5–32	≤8	16	≥32
Cephalosporins (3rd and 4th generations)					
Cefoxitin	FOX	2–32	≤8	16	≥32
Ceftriaxone	CRO	0.03–4	≤8	16-32	≥64
Ceftiofur	TIO	0.25–8	≤8	16-32	≥64
Quinolone					
Nalidixic acid	NAL	1–32	≤16	NA	≥32
Fluoroquinolones					
Ciprofloxacin	CIP	0.004–4	≤0.06	0.12–0.5	1≥
Tetracyclines					
Tetracyclines	TET	0.5–16	≤4	8	≥16
Chloramphenicol					
Chloramphenicol	CHL	2–32	≤8	16	≥32
Folate pathway inhibitors					
Sulfisoxazole	FIS	8–512	≤256	NA	≥512
Trimethoprim-sulfamethoxazole	SXT	0.5/9.5-4/76	≤2/38	NA	≥4/76

**S, sensitive; I, intermediate resistance; and R, resistance. #No CLSI interpretative criteria are available for streptomycin; thus, provisional breakpoint from NARMS was used.*

modifications. Briefly, each egg (ca. 60 g, including the shell) was broken and uniformly mixed with 540 mL of buffered peptone water (BD, Cockeysville, MD, United States) in a homogenization bag, then homogenized for 2 min. Before handling the next egg, a new pair of germ-free gloves was used to avoid bacterial cross-contamination throughout the process. The mixed cultures were pre-incubated at 37°C with shaking at 100 rpm for 6 h. Subsequently, 10 mL cultures were transferred into 100 mL of tetrathionate broth (TT; BD) and 1 mL cultures were transferred to 100 mL of Rappaport–Vassiliadis broth (RV; BD). The inoculated TT and RV broths were incubated at 42°C with shaking at 100 rpm for 24 h. Afterward, TT cultures were streaked on to xylose lysine tergitol 4 agar (XLT4; BD) and RV cultures were streaked on to xylose lysine deoxycholate agar (XLD; BD). The inoculated XLT4 and XLD plates were incubated at 35°C for 48 h, and two putative colonies with a typical *Salmonella* phenotype were picked from each plate and subcultured on a fresh XLT4 plate for purification. The purified isolates were stabbed into one triple sugar iron (BD) slant and one urea-agar (BD) slant, respectively, followed by incubation at 35°C for 18–24 h to exclude *Escherichia coli* and *Proteus* species.

Presumptive *Salmonella* isolates were identified by polymerase chain reaction (PCR) using the primers *inv*AF (5′-GTGAAATTATCGCCACGTTCGGGCAA-3′) and *inv*AR (5′-TCATCGCACCGTCAAAGGAACC-3′) (Rahn et al., 1992). Polymerase chain reaction was carried out in a 25-μL reaction mixture containing 0.5 μM of each primer, 250 μM of dNTP, 1 × PCR buffer, 1.5 mM $MgCl_2$, 0.5 U of *Taq* DNA polymerase (TaKaRa, Dalian, China), and 5 μL of DNA template. All PCR reactions were performed on a MyCircle thermocycler (Bio-Rad, Hercules, CA, United States) with pre-denaturation at 95°C for 5 min, followed by 35 cycles of 95°C for 30 s, 64°C for 30 s, and 72°C for 30 s, and a final extension of 72°C for 5 min. Polymerase chain reaction products were stained with ethidium bromide and visualized under UV light after gel electrophoresis in 1% agarose (Kasturi and Drgon, 2017). *S.* Typhimurium LT2 was used as the positive control strain.

Serotyping of Somatic and Flagellar Antigens

After PCR identification, *Salmonella* isolates were serotyped in the Henan Center of Disease Control and Prevention (Zhengzhou, Henan Province, China). Somatic (O) and flagellar (H) antigens were determined by the slide agglutination method using *Salmonella*-specific hyperimmune sera (S&A Company, Bangkok, Thailand). The serotype was identified by the antigen form according to the Kauffman–White scheme (Ben Aissa et al., 2007). *S.* Typhimurium LT2 was used as the positive control strain.

Antibiotic Susceptibility Test

Table 1 lists the antibiotics used for the susceptibility test. The agar dilution method suggested by the Clinical and Laboratory

Standards Institute (CLSI, 2015) was used to determine the minimum inhibitory concentrations of the antibiotics to *Salmonella* isolates on Mueller-Hinton agar (Beijing Land Bridge Technology Co., Ltd, Beijing, China). The results of resistance, intermediate resistance, and susceptibility were interpreted via the guidelines established by the CLSI, except for streptomycin, the breakpoint of which was interpreted by that specified by the National Antimicrobial Resistance Monitoring System (U.S. Food and Drug Administration, 2013; FAO, 2018a,b). *E. coli* ATCC25922 and ATCC35218 and *Enterococcus faecalis* ATCC29212 were used as the quality control strains.

Subtyping by Pulse-Field Gel Electrophoresis

Pulse-field gel electrophoresis (PFGE) was used to determine the genetic relationship between *Salmonella* isolates according to the PulseNet protocol (Centers for Disease Control and Prevention [CDC], 2013). Briefly, *Salmonella* subcultures were grown on Luria–Bertani agar overnight at 37°C to reach an optical density of 0.5 and then embedded in SeaKem Gold agarose (Lonza, Rockland, ME, United States). Then culture plugs were lysed in cell lysis buffer with 100 μg mL^{-1} protease K (TaKaRa) by incubation at 55°C in a shaking water bath for 2 h. Subsequently, the lysed plugs were washed twice with sterile water and then washed four times with sterilized Tris–EDTA buffer. A slice was cut from the plug and digested using 50 U of *Xba*I (TaKaRa) by incubation in a 37°C water bath for 1.5–2 h. The digested DNA fragments in each slice were separated in 0.5 × Tris-borate-EDTA buffer at 14°C for 20 h using a Chef Mapper electrophoresis system (Bio-Rad) with pulse times of 2.16–63.8 s. The gel was stained using ethidium bromide (100 μg mL^{-1}), and the DNA bands were visualized via UV transillumination (Bio-Rad). Pulse-field gel electrophoresis results were manually analyzed using BioNumerics v3.0 (Applied Maths, Kortrijk, Belgium), and the extent of variability was determined using the Dice coefficient and clustering by the unweighted pair-group average method (Yang et al., 2014). *S.* Braenderup H9812 was used as the standard DNA control.

Statistical Analysis

Statistical analysis was performed using IBM SPSS Statistics v22 (IBM Corp., Armonk, NY, United States). Pearson's chi-square test was used to determine the differences in the detection rates of *Salmonella*-positive samples and *Salmonella* serotypes in eggs across different sampling locations (i.e., wet markets, supermarkets, and poultry farms) and times. The results were analyzed at the 5% ($\alpha = 0.05$) level to determine whether differences were significant. Data were edited using Microsoft Excel 2016 (Microsoft Corp., Redmond, WA, United States), with environmental variables (i.e., serotype, sampling location, and sampling time) and explanatory environmental variables (i.e., number and category of the antibiotics to which *Salmonella* exhibited resistance) organized into the form of "row representing samples and column representing variables." The edited data were imported into Canoco v5.0 (Microcomputer Power, Ithaca, NY, United States) for redundancy analysis. The rate of antibiotic-resistant *Salmonella* isolates between different serotypes was analyzed using R v3.4.4[1], and a heatmap was created using the R package "pheatmap" v3.0.1[2].

TABLE 2 | Prevalence of *Salmonella* in eggs collected from poultry farms and marketplaces at different sampling locations and times.

	Location or time (egg number)	Number (percentage) of *Salmonella*-positive eggs	Number (percentage) of *Salmonella* isolates
Source	Retail market (510)	26 (5.1)	40 (65.6)
	Poultry farm (304)	20 (6.6)	21 (34.4)
Retail	Supermarket (259)	15 (5.8)	22 (36.1)
market	Wet market (251)	11 (4.4)	18 (29.5)
Poultry	C (88)	19 (21.6)##	20 (32.8)
farm	X (108)	1 (0.9)	1 (1.6)
	F (108)	0	0
Sampling	2013 (134)	2 (1.5)	2 (3.3)
year	2014 (680)	44 (6.5)*	59 (96.7)
Sampling	Aug. 2013 (98)	2 (2.0)c	2
month	Spt. 2013 (12)	0c	0
	Oct. 2013 (24)	0c	0
	Mar. 2014 (66)	20 (30.3)a	33
	Jun. 2014 (175)	4 (2.3)c	5
	Aug. 2014 (135)	0c	0
	Oct. 2014 (196)	19 (9.7)b	20
	Nov. 2014 (108)	1 (0.9)c	1
Total	814	46 (5.6)	61 (100)

*##P < 0.01; *P < 0.05, P = 0.023. a,b,cValues (percentage) with different superscripted letters are significantly different between groups.*

RESULTS

Prevalence of *Salmonella*

Of the 814 eggs, 46 (5.6%) were detected to be *Salmonella*-positive (**Table 2**). The detection rate of *Salmonella*-positive eggs from poultry farms was slightly higher (6.6%, 20/304) than that from retail markets (5.1%, 26/510). Among the different marketplaces, the detection rate of *Salmonella*-positive eggs was slightly higher in supermarkets (5.8%, 15/259) than in wet markets (4.4%, 11/251). However, the differences in the detection rates of *Salmonella*-positive eggs were not significant among retail markets and poultry farms.

Among the three poultry farms, no *Salmonella*-positive eggs were detected from farm F, only one (0.9%) *Salmonella*-positive egg was detected in farm X, while 19 (21.6%) were detected in farm C. The detection rate of *Salmonella* in eggs from farm C was significantly ($P < 0.01$) higher than those from the other two poultry farms. During the investigation period, the detection rate of *Salmonella*-positive eggs was significantly higher ($P < 0.05$) in 2014 (6.5%, 44/680) than in 2013 (1.5%, 2/134). Additionally, the detection rates of *Salmonella*-positive eggs in March 2014 (30.3%, 20/60) and October 2014 (9.7%, 19/196)

[1]https://rstudio.com/

[2]https://www.rdocumentation.org/packages/pheatmap/versions/1.0.12

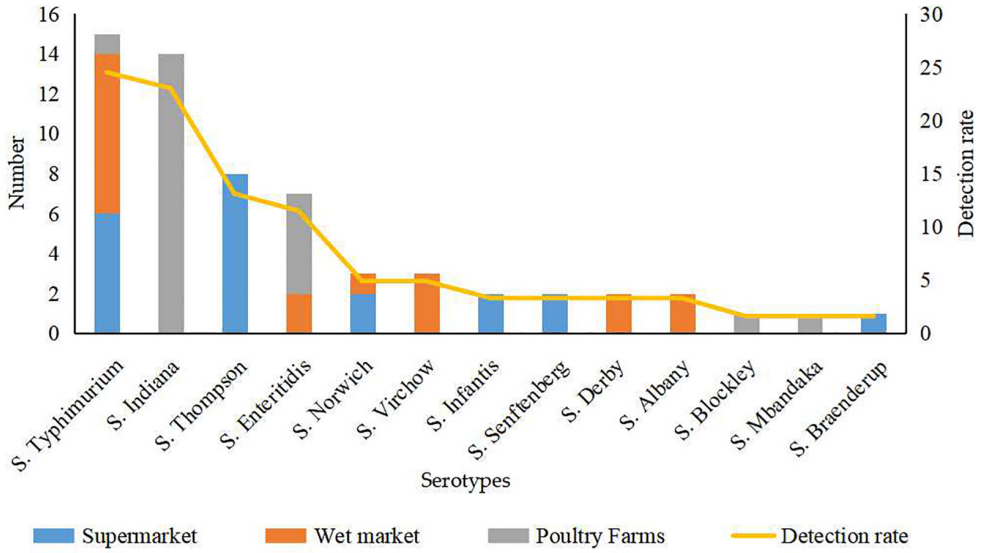


FIGURE 1 | Serotype of 61 *Salmonella* isolates from eggs from different sources.

were significantly ($P < 0.05$) higher than those from the other sampling times (**Table 2**).

In total, 61 *Salmonella* isolates were recovered from XLD and XLT4 plates derived from the 46 *Salmonella*-positive eggs. Among them, 40 (65.6%) isolates were obtained from eggs collected from retail markets and 21 (34.4%) were from eggs sampled from poultry farms (**Table 2**).

Serotype of *Salmonella*

Thirteen serotypes were identified from the 61 *Salmonella* isolates, among which *S.* Typhimurium (24.5%) and *S.* Indiana (22.9%) were the most commonly detected serotypes (**Figure 1**). The detection rates of *Salmonella* Thompson (13.1%) and *Salmonella* Enteritidis (11.4%) were moderately high. Additionally, *Salmonella* Norwich and *Salmonella* Virchow accounted for 4.9% of all the identified isolates each; *Salmonella* Derby, *Salmonella* Senftenberg, *Salmonella* Infantis, and *Salmonella* Albany accounted for 3.2% each; and *Salmonella* Blockley, *Salmonella* Mbandaka, and *S.* Braenderup were detected at the lowest rates (1.6% each).

The 39 *Salmonella* isolates recovered from retail eggs covered 10 serotypes, while the 22 isolates recovered from poultry farms covered five serotypes (**Figure 1**). Considering their source, all *S.* Indiana, *S.* Blockley, and *S.* Mbandaka isolates and the majority (71.4%) of *S.* Enteritidis isolates were from poultry farms. All *S.* Thompson, *S.* Infantis, *S.* Senftenberg, and *S.* Braenderup isolates and 40.0% of *S.* Typhimurium isolates were from supermarkets. All *S.* Virchow, *S.* Derby, and *S.* Albany isolates and more than half of *S.* Norwich (66.7.%) and *S.* Typhimurium (53.3%) isolates were from wet markets.

Antibiotic Susceptibility of *Salmonella*

All the 61 *Salmonella* isolates were resistant to sulfisoxazole (100%), and the rate of sulfisoxazole resistance was significantly ($P < 0.05$) higher than the rates of resistance to the other 13 antibiotics tested for (**Figure 2**). For example, 77.1% of the isolates were resistant to amoxicillin-clavulanate, ampicillin, and nalidixic acid, while nearly two-thirds (63.9–68.9%) of the tested isolates exhibited resistance to trimethoprim-sulfamethoxazole, kanamycin, tetracyclines, and chloramphenicol. Less than half (40.1%) of the isolates were resistant to ciprofloxacin; around one-fifth (21.3–26.2%) of the isolates were resistant to streptomycin, ceftiofur, and ceftriaxone; and the resistance rates to gentamicin, amikacin, and cefoxitin were relatively low (3.3–16.4%).

There were no significant differences in the resistance rates of *Salmonella* isolates to amoxicillin-clavulanate, ampicillin, and nalidixic acid; however, these resistance rates were significantly ($P < 0.05$) higher than the rates of resistance to the other antibiotics tested, except the resistance rate to sulfisoxazole. Additionally, significant ($P < 0.05$) differences were found in the resistance rates to amikacin, gentamicin, kanamycin, streptomycin, chloramphenicol, cefoxitin, ceftiofur, ceftriaxone, ciprofloxacin, and trimethoprim-sulfamethoxazole. Moreover, the resistance rates to aminoglycosides (amikacin, gentamicin, and kanamycin) and cephalosporins (cefoxitin, ceftriaxone, and ceftiofur) differed significantly ($P < 0.05$; **Figure 2**).

A total of 49 (80.3%) *Salmonella* isolates were resistant to more than three categories of antibiotics. Specifically, 10 (20.4%) isolates were resistant to 3–5 categories (3–8 species) of antibiotics and the remaining 39 (79.6%) isolates were resistant to 6–8 categories (8–12 species) of antibiotics. The *Salmonella* isolates that were resistant to the 15 antibiotics based on different serotypes could be grouped into two clusters (G1 and G2; **Figure 3**).

Isolates resistant to ceftriaxone, amikacin, ciprofloxacin, ceftiofur, nalidixic acid, kanamycin, chloramphenicol, and cefoxitin were grouped in cluster G1. Of these, isolates resistant

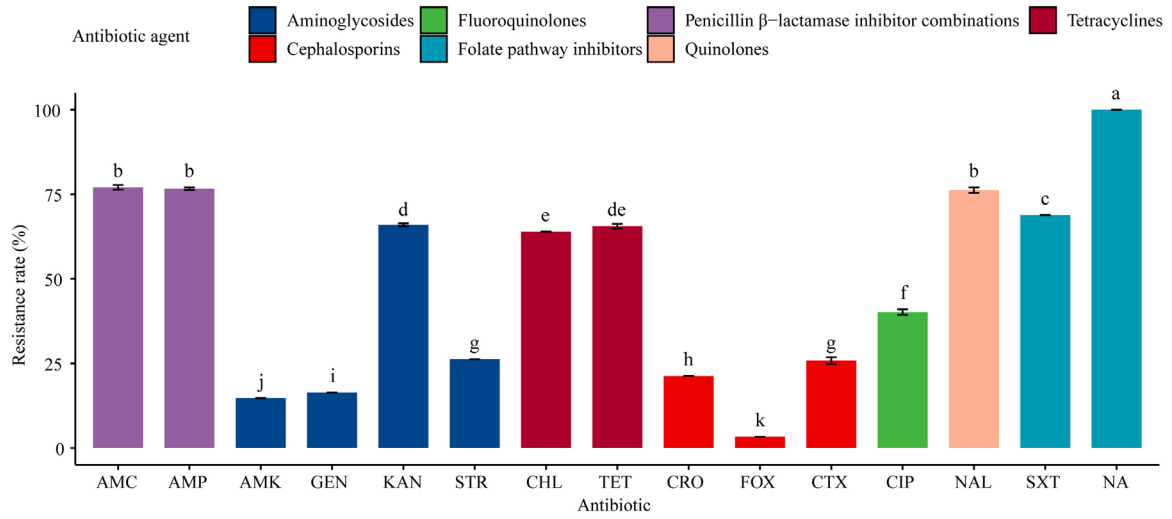


FIGURE 2 | Resistance of 61 *Salmonella* isolates to 15 antibiotics. **Table 1** has a list of abbreviations used in the abscissa. Different colors represent seven different antibiotic categories. Data are mean ± standard deviation (n = 3). Different letters used to label the bar chart denote significant differences found between the resistance rates of isolates to the corresponding antibiotic ($P < 0.05$).

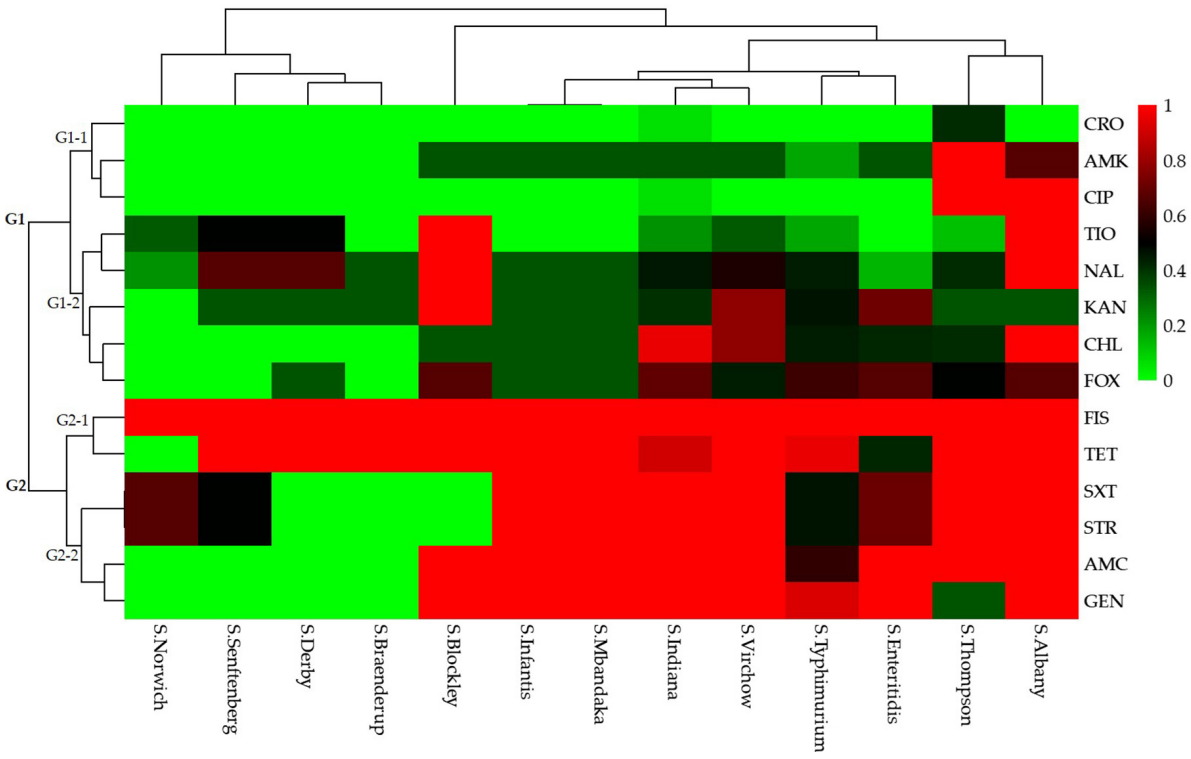


FIGURE 3 | Heatmap (with dendrograms) showing the antibiotic susceptibilities of 61 *Salmonella* isolates among 13 serotypes. Red squares (pixels) represent high-frequency values and green squares represent low-frequency values. **Table 1** presents a list of the abbreviations of the antibiotics.

to ceftriaxone, amikacin, and ciprofloxacin were grouped in a secondary cluster (G1-1), and isolates resistant to the other five antibiotics were grouped in another secondary cluster (G1-2). Isolates resistant to sulfisoxazole, tetracyclines, trimethoprim-sulfamethoxazole, streptomycin, amoxicillin-clavulanate, and gentamicin were grouped in cluster G2. In this cluster, isolates resistant to sulfisoxazole and tetracycline were grouped in a secondary cluster (G2-1), and those resistant to trimethoprim-sulfamethoxazole, streptomycin, amoxicillin-clavulanate, and gentamicin were grouped in another secondary cluster (G2-2).

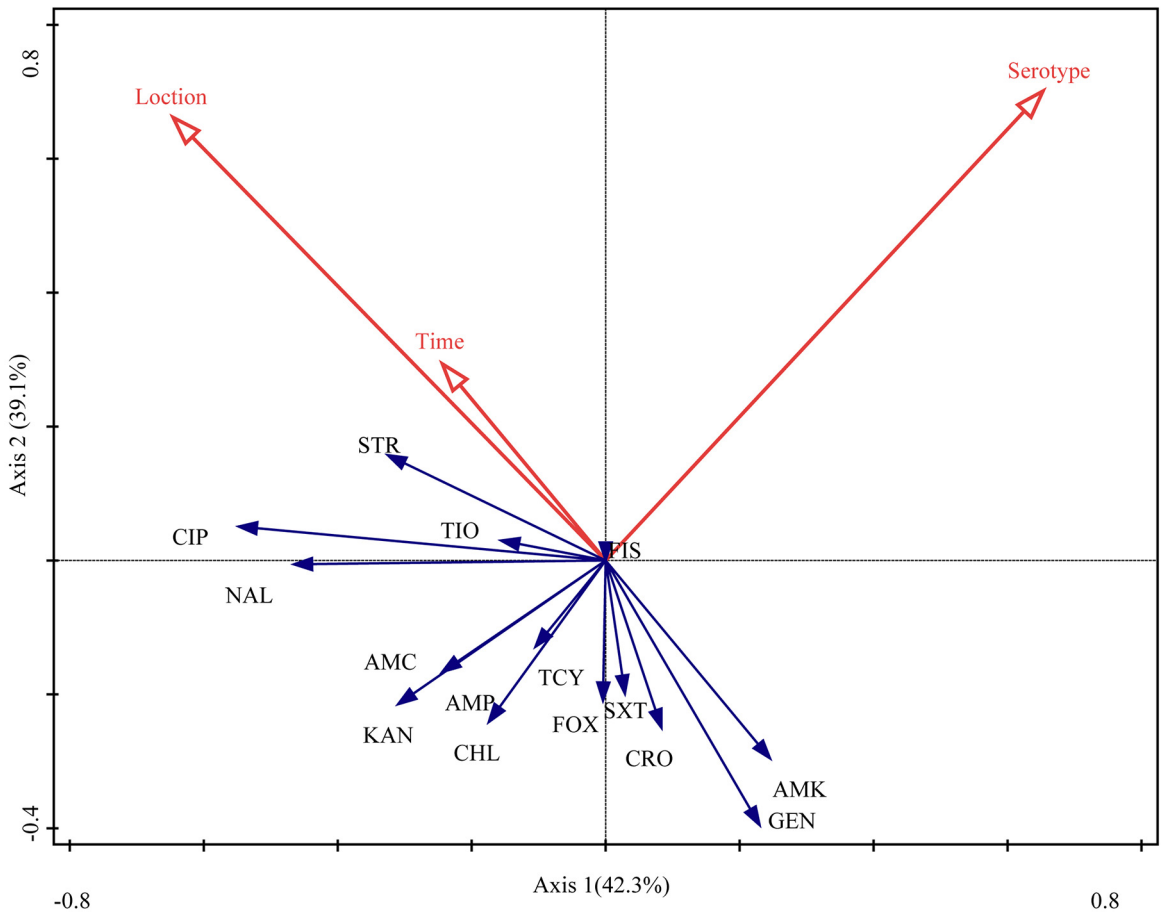


FIGURE 4 | Relationship between the antibiotic susceptibility of *Salmonella* isolates, their serotype, sampling location, and sampling time. **Table 1** presents a list of the abbreviations of the antibiotics. Blue arrow, species variable; red arrow, environmental variable. The cosine of the angle between two variables represents the correlation between them. For example, two variables that are nearly at a right angle to each other are almost uncorrelated.

The RDA biplot (**Figure 4**) indicated that the contribution of serotype to the antibiotic susceptibility of *Salmonella* isolates was the highest (42.3%, P = 0.004), followed by sampling location (39.1%, P = 0.018) and time (18.6%, P = 0.138). The serotype of the isolates was correlated with their susceptibility to most of the antibiotics tested for, except amoxicillin-clavulanate, ampicillin, kanamycin, and chloramphenicol. Additionally, sampling location and time were (closely) correlated with the susceptibility of the isolates to streptomycin, ciprofloxacin, ceftiofur, nalidixic acid, amoxicillin-clavulanate, ampicillin, kanamycin, tetracyclines, and chloramphenicol.

PFGE Subtype of *Salmonella*

When the *Salmonella* isolates were subtyped by PFGE, the genomic DNA of each isolate produced 13–21 bands with the typing rate of 100% (**Figure 5**). Isolates derived from the eggs collected in the same month always belonged to the same serotype, and they exhibited the same or similar antibiotic resistance phenotype, despite being recovered from different locations (C1-1, C2-3, and C5). In contrast, some of the isolates recovered from the eggs sampled from different locations belonged to various serotypes, yet they also showed identical PFGE profiles and antibiotic resistance phenotypes (C1-2 and C1-3). Moreover, some isolates recovered from the eggs collected at the same location and time belonged to the same serotype, but they showed distinct PFGE and antibiotic resistance profiles (C2-2, C2-4, and C2-5).

DISCUSSION

The total egg production reached 24,446 million metric tons in 2013 and 4539 billion in 2014 (FAO, 2018a,b). If contaminated eggs were produced by laying hens infected with *Salmonella*, it would be difficult to effectively eliminate such contamination through vaccination and other interventions. Therefore, not only could this lead to outbreaks of foodborne illness in humans, but chickens hatched from contaminated eggs may also have inherent defects (Barrow, 2007; Kouam et al., 2018). The detection and characterization of *Salmonella* spp. in eggs can provide useful information for the mitigation of socioeconomic losses caused by contamination with *Salmonella*. In the present

FIGURE 5 | *Xba*I-based pulse-field gel electrophoresis (PFGE) and antibiotic resistance profiles of *Salmonella* isolates from eggs.

study, the average detection rate of *Salmonella*-positive eggs in poultry farms, supermarkets, and wet markets in Yangling (5.6%) was higher than those reported in Shandong Province (2.1%, 49/2342) and Hebei Province (4.8%) (Li et al., 2018; Yang et al., 2019); however, it was much lower than the detection rates in some conventional farms in Sichuan Province (12.2%) and Jiangsu Province (17.4%, 160/920), China (Li et al., 2013; Hai et al., 2020). During the study period, a high detection rate of *Salmonella*-positive eggs in 2014 was mainly found in the samples collected in March and October, and all the *Salmonella*-positive samples in October were obtained from the medium-scale poultry farm C. Although these results indicate that *Salmonella* is still prevalent in the eggs, the tendency of reduced prevalence suggests that in recent years, China has achieved remarkable results in monitoring and controlling the contamination of eggs with pathogens including *Salmonella* from farm to table.

Based on our results, the detection rate of *Salmonella*-positive eggs in the production (6.6%) and sales (5.1%) links were similar. A previous investigation provided a detailed description of contamination during five breeder farm production stages (i.e., laying, hatching, rearing, brooding, and post-hatching), with a considerably high prevalence of egg samples containing *Salmonella* found at the laying (29.2%) and hatching (21.6%) stages (Fei et al., 2017). In the current study, a total of 304 eggs were collected from three different poultry farms of different scales in Yangling. Although the number of eggs collected from each farm was similar, the detection rate of *Salmonella*-positive eggs varied in a broad range across different farms. Except one from the large-scale farm X, all the remaining *Salmonella*-positive eggs were detected from medium-scale farm C, while no *Salmonella*-positive eggs were detected from the small-scale farm F. Farm C was a conventional poultry farm with a smaller production scale and a higher density of breeding hens compared to poultry farm X. According to our observations during sampling, raw eggs in farm C all rolled to the same egg tray and were harvested by hand. Moreover, there was no separate space between the breeding and egg collection areas in farm C. Thus, cross-contaminations may be a major factor responsible for the high prevalence of *Salmonella* in the eggs collected from this farm. Conversely, in the large-scale and modernized farm X, the production units were completely closed off, with feed supply and egg collection occurring through different conveyor belts; this could reduce the chance of cross-contamination caused by *Salmonella* present in the environment, especially in the feces of hens. Changes in the management of stocking density at the laying stage can influence *Salmonella* contamination (Gast et al., 2014). In addition, immunological changes in breeder chickens at the laying stage can increase the contamination rates (Johnston et al., 2012).

Salmonella Typhimurium and *S.* Enteritidis are the two most commonly identified serotypes and causative agents involved

in foodborne salmonellosis (Galis et al., 2013; Whiley and Ross, 2015). In the present study, 13 different serotypes were identified from the 61 *Salmonella* isolates from the eggs. The top four serotypes were *S.* Typhimurium, *S.* Indiana, *S.* Thompson, and *S.* Enteritidis. Generally, our results were consistent with the results of previous studies reporting that *S.* Typhimurium was the most prevalent serotype in eggs obtained from Penang in Malaysia and 50 poultry farms throughout Korea (Adzitey et al., 2012; Bae et al., 2013). Additionally, *S.* Indiana, *S.* Typhimurium, and *S.* Enteritidis were found to be the predominant *Salmonella* serovars in eggs sampled in Yangzhou, China (Li et al., 2017), while *S.* Enteritidis was the most prevalent serotype in eggs collected from Shandong, Shanghai, and Guangdong in China (Ni et al., 2018; Xie et al., 2019; Yang et al., 2019). Particularly, *S.* Enteritidis has always been a common foodborne pathogen associated with salmonellosis caused by consumption of contaminated eggs or egg products, and this is mainly due to its specific survival mechanism in egg white with the assistance of the outer membrane channel TolC (Huang et al., 2019; Raspoet et al., 2019). Here, although *S.* Enteritidis was not the most dominant serotype in all egg samples, over 70% of *S.* Enteritidis isolates were derived from poultry farms, and the remaining isolates were recovered from wet markets. It could be considered that *S.* Enteritidis remained to be the predominant serotype of *Salmonella* in contaminated eggs from poultry farms in the study area.

In total, 10 *Salmonella* serotypes were detected in the isolates derived from retail eggs, while five serotypes were identified from eggs collected from poultry farms. Both *S.* Typhimurium and *S.* Enteritidis were detected in the eggs from poultry farms and retail markets, whereas the other eight serotypes were detected in the eggs from retail markets only. Interestingly, *S.* Indiana and *S.* Blockley isolates that were commonly detected from poultry farms were absent in the eggs from marketplaces. These results indicate that the prevalent serotypes of *Salmonella* in eggs differed between retail markets and poultry farms. Previous studies have reported that the pathways of pathogen contamination could be influenced by the egg production, storage, and handling procedures; in other words, cross-contamination may occur during egg storage, transportation, and sales, although some *Salmonella* strains die in these periods (Namata et al., 2008; Food Drug Administration, 2009; Gast et al., 2014). In addition, *S.* Indiana accounted for a large proportion (23.0%; 15/61) of the *Salmonella* isolates from eggs in the current study, and it was frequently detected in poultry in China during 2006–2012 (Gong et al., 2017). However, this serotype has not been commonly documented in other countries. Since *Salmonella*, especially *S.* Typhimurium, could maintain a high survival rate on eggshells and *S.* Enteritidis within the egg, potential multiple outbreaks associated with these two serotypes in eggs could occur (McAuley et al., 2015). Therefore, apart from improving the hygienic conditions of egg production, efficient measures should be implemented to reduce exposure and surface contact of eggs during transportation and handling in order to prevent contamination by foodborne pathogens including *Salmonella*.

In poultry rearing, excessive antibiotics are commonly used for disease prevention and growth promotion, leading to the occurrence of antibiotic-resistant bacteria (Krishnasamy et al., 2015; Haskell et al., 2018). A considerable increase in multidrug-resistant (MDR) *Salmonella* has been observed, and the number of MDR *Salmonella* infections has increased worldwide over the past few years (Fardsanei et al., 2018). In the present study, all the 61 *Salmonella* isolates from eggs exhibited resistance to one or more antibiotic agents, and only three isolates were non-MDR. Notably, one-third of the isolates were resistant to at least 10 antibiotics. Similarly, Al et al. (2015) found that the rate of MDR *Salmonella* isolated from eggs and poultry products reached up to 86.5%, and *S.* Thompson isolates exhibited higher antibiotic resistance than isolates of other serotypes recovered from the same marketplace. Moreover, Sanchez-Salazar et al. (2019) obtained 31 *Salmonella* isolates from layer poultry farms in central Ecuador in 2017, 58.1% (18/31) of which were MDR. Taken together, these findings corroborate reports showing that MDR *Salmonella* is prevalent in eggs (Garedew et al., 2015; Bezerra et al., 2016; Taddese et al., 2019). In the current study, almost all *S.* Indiana and *S.* Enteritidis isolates derived from poultry farm C were MDR strains. This result might be attributed to excessive use and overdosage of antibiotics in poultry farm C associated with poor production environment and sanitary conditions.

According to the PFGE profiles, some *Salmonella* isolates within the same serotype were derived from the same location and/or time; these isolates had a close genetic relatedness while they shared the same or similar antibiotic resistance phenotypes. This result is consistent with a previous study of *Salmonella* contamination in layer poultry farms in Shandong and Hebei Provinces, China, that certain samples within henhouses and egg-collecting areas showed relatively high genetic similarities (Li et al., 2018). Here, almost all the *Salmonella* isolates recovered from poultry farms were identified to be the same serotype with similar PFGE and antibiotic resistance profiles. Likewise, the *Salmonella* isolated from different backyard eggs in Portugal displayed identical PFGE profiles, indicating that they belonged to a prevalent clone in the region (Ferreiraa et al., 2020). Moreover, some isolates of a certain serotype sampled from different locations and at different times exhibited the same or similar PFGE profiles and antibiotic resistance phenotypes. This result indicates that *Salmonella* isolates of these serotypes might have existed in the parturient canal of laying hens, poultry rearing environments, or transportation and storage systems for a long period of time. In contrast, some isolates of the same serotype showed diverse PFGE profiles and antibiotic resistance phenotypes. From a poultry slaughterhouse in Brazil, 40 *Salmonella* isolates obtained over a 20-week period showed diverse PFGE profiles in the same serotype, except *S.* Enteritidis, which could occur due to the low genetic diversity in this serovar (Dantas et al., 2020). Our results indicate that the prevalence of *Salmonella* in eggs could pose potential hazards for consumers and result in *Salmonella* outbreaks over a certain period of time.

This study indicated that *Salmonella* was prevalent in fresh eggs from poultry farms and retail marketplaces with diverse

serotypes, and the majority of the isolates were resistant to the multiple antibiotics tested. Some isolates of the same serotype were sampled from the same location and/or time, which shared identical or highly similar PFGE profiles and antibiotic resistance profiles. As eggs play a vital role in daily human life and can be easily contaminated by MDR *Salmonella*, it is crucial for local health departments to monitor the occurrence of *Salmonella* in eggs and prevent egg contamination via the food production and supply chains, including poultry farms and retail markets.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/supplementary material.

AUTHOR CONTRIBUTIONS

WL conceived the study and drafted the manuscript. WL, HL, SZ, ZW, and QN performed the experiments and collected the data. HS conducted data analysis. BY supervised the project, experimental design, data collection and analysis, and manuscript preparation. CS and XS guided the experimental design and assisted with manuscript revision. All authors contributed to manuscript revision and agreed to the published version of the manuscript.

FUNDING

This study was supported by the National Key R&D Program of China (2017YFC1600100).

REFERENCES

Adzitey, F., Rusul, G., and Huda, N. (2012). Prevalence and antibiotic resistance of *Salmonella serovars* in ducks, duck rearing and processing environments in Penang, Malaysia. *Food Res. Int.* 45, 947–952. doi: 10.1016/j.foodres.2011.02.051

Al, S., Hizlisoy, H., Ertas Onmaz, N., Yildirim, Y., and Gonulalan, Z. (2015). Occurrence and antimicrobial resistance of *Salmonella enterica* subsp. enterica serovars Typhimurium, Enteritidis, and Typhi isolated from chicken eggs and poultry products. *Turk. J. Vet. Anim. Sci.* 40, 737–743. doi: 10.3906/vet-1601-17

Andino, A., and Hanning, I. (2015). *Salmonella enterica*: survival, colonization, and virulence differences among serovars. *Sci. World J.* 2015:520179. doi: 10.1155/2015/520179

Bae, D. H., Dessie, H. K., Baek, H. J., Kim, S. G., Lee, H. S., and Lee, Y. J. (2013). Prevalence and characteristics of *Salmonella* spp. isolated from poultry slaughterhouses in Korea. *J. Vet. Med. Sci.* 75, 1193–1200. doi: 10.1292/jvms.13-0093

Barrow, P. A. (2007). *Salmonella* infections: immune and non-immune protection with vaccines. *Avian Pathol.* 36, 1–13. doi: 10.1080/03079450601113167

Ben Aissa, R., Al-Gallas, N., Troudi, H., Bethadj, N., and Belhadj, A. (2007). Trends in *Salmonella enterica* serotypes isolated from human, food, animal, and environment in Tunisia, 1994–2004. *J. Infect.* 55, 324–339. doi: 10.1016/j.jinf.2007.06.007

Bezerra, W. G. A., da Silva, I. N. G., Vasconcelos, R. H., Machado, D. N., Lopes, E. D., Lima, S. V. G., et al. (2016). Isolation and antimicrobial resistance of *Escherichia coli* and *Salmonella enterica* subsp. enterica in broiler chickens. *Acta Sci. Vet.* 44, 1–7. doi: 10.22456/1679-9216.80957

Centers for Disease Control and Prevention [CDC] (2013). *Standard Operating Procedure for PulseNet PFGE of Escherichia coli O157:H7, Escherichia coli Non-O157 (STEC), Salmonella Serotypes, Shigellasonnei and Shigellaflexneri, National Center for Infectious Diseases*. Atlanta, GA: Centers for Disease Control and Prevention.

CLSI (2015). *Clinical and Laboratory Standards Institute. Performance Standards for Antimicrobial Susceptibility Testing: Twentieth-third Informational Supplement M100-S23*. Wayne, PA: Clinical and Laboratory Standards Institute.

Cowen, P., Currier, R. W., and Steele, J. H. (2016). A short history of one health in the United States. *Vet. Herit.* 39, 1–15.

Cui, M. Q., Xie, M. Y., Qu, Z. N., Zhao, S. J., Wang, J. W., Wang, Y., et al. (2016). Prevalence and antimicrobial resistance of *Salmonella* isolated from an integrated broiler chicken supply chain in Qingdao China. *Food Control* 62, 270–276. doi: 10.1016/j.foodcont.2015.10.036

Dantas, S. T. A., Camargo, C. H., Tiba-Casas, M. R., Vivian, R. C., Pinto, J. P. A. N., Pantoja, J. C. F., et al. (2020). Environmental persistence and virulence of *Salmonella* spp. Isolated from a poultry slaughterhouse. *Food Res. Int.* 129:108835. doi: 10.1016/j.foodres.2019.108835

FAO (2018a). Available online at: http://faostat3.fao.org/browse/Q/QA/E (accessed June 6, 2018a).

FAO (2018b). Available online at: http://faostat3.fao.org/browse/Q/QL/E (accessed June 6, 2018b).

Fardsanei, F., Dallal, M. M. S., Douraghi, M., Memariani, H., Bakhshi, B., Salehi, T. Z., et al. (2018). Antimicrobial resistance, virulence genes and genetic relatedness of *Salmonella enterica* serotype *Enteritidis* isolates recovered from human gastroenteritis in Tehran, Iran. *J. Glob. Antimicrob. Resist.* 12, 220–226. doi: 10.1016/j.jgar.2017.10.005

Fardsanei, F., Dallal, M. M. S., Douraghi, M., Salehi, T. Z., Mahmoodi, M., Memariani, H., et al. (2017). Genetic diversity and virulence genes of *Salmonella enterica* subspecies enterica serotype Enteritidis isolated from meats and eggs. *Microb. Pathog.* 107, 451–456. doi: 10.1016/j.micpath.2017.04.026

FDA (2019a). *Investigated Multistate Outbreak of Salmonella Braenderup Linked to Shell Eggs from Rose Acre Farms*. Silver Spring, MD: FDA.

FDA (2019b). *Investigated Multistate Outbreak of Salmonella Enteritidis Linked to Shell Eggs from Gravel Ridge Farms*. Silver Spring, MD: FDA.

FDA (2019c). *Investigates Outbreak of Salmonella Oranienburg Linked to Shell Eggs*. Silver Spring, MD: FDA.

Fei, X., He, X., Guo, R. X., Yin, C., Geng, H. P., Wu, K. Y., et al. (2017). Analysis of prevalence and CRISPR typing reveals persistent antimicrobial-resistant *Salmonella* infection across chicken breeder farm production stages. *Food Control.* 77, 102–109. doi: 10.1016/j.foodcont.2017.01.023

Ferreiraa, V., Cardosoa, M. J., Magalhãesa, R., Maiab, R., Neaguc, C., Dumitraşcuc, L., et al. (2020). Occurrence of *Salmonella* spp. in eggs from backyard chicken flocks in Portugal and Romania - Results of a preliminary study. *Food Control.* 113:107180. doi: 10.1016/j.foodcont.2020.107180

Food Drug Administration (2009). Prevention of *Salmonella* enteritidis in shell eggs during production, storage, and transportation. Final rule. *Fed. Regist.* 74, 33030–33101.

Galis, A. M., Marcq, C., Marlier, D., Portetelle, D., Van, I., Beckers, Y., et al. (2013). Control of *Salmonella* contamination of shell eggs-Preharvest and postharvest methods: a review. *Compr. Rev. Food Sci. Food Saf.* 12, 155–182. doi: 10.1111/1541-4337.12007

Garedew, L., Hagos, Z., Addis, Z., Tesfaye, R., and Zegeye, B. (2015). Prevalence and antimicrobial susceptibility patterns of *Salmonella* isolates in association with hygienic status from butcher shops in Gondar town, Ethiopia. *Resist. Infect. Control.* 4:21. doi: 10.1186/s13756-015-0062-7

Gast, R. K., Guraya, R., Jones, D. R., and Anderson, K. E. (2014). Contamination of eggs by *Salmonella* enteritidis in experimentally infected laying hens housed in conventional or enriched cages. *Poult. Sci.* 93, 728–733. doi: 10.3382/ps.2012-02811

Gong, J., Kelly, P., and Wang, C. (2017). Prevalence and antimicrobial resistance of *Salmonella* 537 enterica serovar Indiana in China (1984-2016). *Zoonoses Public Health* 64, 239–251. doi: 10.1111/zph.12328

Hai, D., Yin, X. P., Lu, Z. X., Lv, F. X., Zhao, H. Z., and Bie, X. M. (2020). Occurrence, drug resistance, and virulence genes of *Salmonella* isolated from

chicken and eggs. *Food Control.* 113, 107109. doi: 10.1016/j.foodcont.2020.107109

Haskell, K. J., Schriever, S. R., Fonoimoana, K. D., Haws, B., Hair, B. B., Wienclaw, T. M., et al. (2018). Antibiotic resistance is lower in *Staphylococcus aureus* isolated from antibiotic-free raw meat as compared to conventional raw meat. *PLoS One* 13:e0206712. doi: 10.1371/journal.pone.0206712

Huang, X. Z., Zhou, X. J., Jia, B., Li, N., Jia, J. Y., He, M., et al. (2019). Transcriptional sequencing uncovers survival mechanisms of *Salmonella enterica* serovar Enteritidis in antibacterial egg white. *mSphere* 4:e00700-18. doi: 10.1128/mSphere.00700-18

Johnston, C. E., Hartley, C., Salisbury, A. M., and Wigley, P. (2012). Immunological changes at point-of lay increase susceptibility to *Salmonella enterica* serovar Enteritidis infection in vaccinated chickens. *PLoS One* 7:e48195. doi: 10.1371/journal.pone.0048195

Karimiazar, F., Soltanpour, M. S., Aminzare, M., and Hassanzadazar, H. (2019). Prevalence, genotyping, serotyping, and antibiotic resistance of isolated *Salmonella* strains from industrial and local eggs in Iran. *J. Food Saf.* 39:e12585. doi: 10.1111/jfs.12585

Kasturi, K. N., and Drgon, T. (2017). Real-Time PCR method for detection of *Salmonella* spp. in environmental samples. *Appl. Environ. Microbiol.* 83:e00644-17. doi: 10.1128/AEM.00644-17

Kingsbury, J. M., Thom, K., Erskine, H., Olsen, L., and Soboleva, T. (2019). Prevalence and genetic analysis of *Salmonella enterica* from a cross-sectional survey of the new zealand egg production environment. *J. Food Prot.* 82, 2201–2214. doi: 10.4315/0362-028x.jfp-19-159

Kouam, M. K., Biekop, M. H. F., Katte, B., and Teguia, A. (2018). *Salmonella* status of table eggs in commercial layer farms in Menoua Division, West region of Cameroon. *Food Control.* 85, 345–349. doi: 10.1016/j.foodcont.2017.09.037

Krishnasamy, V., Otte, J., and Silbergeld, E. (2015). Antimicrobial use in Chinese swine and broiler poultry production. *Antimicrob. Resist. Infect. Control.* 4:17. doi: 10.1186/s13756-015-0050-y

Li, Q. C., Yin, K. Q., Xie, X. L., Zhao, F., Xia, J., Chen, Y., et al. (2017). Detection and CRISPR subtyping of *Salmonella* spp. isolated from whole raw chickens in Yangzhou from China. *Food Control.* 82, 291–297. doi: 10.1016/j.foodcont.2017.07.008

Li, R. C., Lai, J., Wang, Y., Liu, S. L., Li, Y., Liu, K. Y., et al. (2013). Prevalence and characterization of *Salmonella* species isolated from pigs, ducks and chickens in Sichuan Province, China. *Int. J. Food Microbiol.* 163, 14–18. doi: 10.1016/j.ijfoodmicro.2013.01.020

Li, X. Z., Liu, L., Li, Q. L., Xu, G. Y., and Zheng, J. X. (2018). *Salmonella* contamination in layer farms in China: detection and genetic analysis. *J. Poult. Sci.* 55, 1–9. doi: 10.2141/jpsa.0160144

McAuley, C. M., Duffy, L. L., Subasinghe, N., Hogg, G., Coventry, J., and Fegan, N. (2015). *Salmonella* Typhimurium and *Salmonella* sofia: growth in and persistence on eggs under production and retail conditions. *Biomed. Res. Int.* 2015:914987. doi: 10.1155/2015/914987

McWhorter, A. R., and Chousalkar, K. K. (2019). From hatch to egg grading: monitoring of *Salmonella* shedding in free-range egg production systems. *Vet. Res.* 50:58. doi: 10.1186/s13567-019-0677-4

Moffatt, C. R. M., Musto, J., and Pingault, N. (2017). Recovery of *Salmonella enterica* from Australian layer and processing environments following outbreaks linked to eggs. *Foodborne Pathog. Dis.* 14, 478–482. doi: 10.1089/fpd.2016.2268

Namata, H., Meroc, E., Aerts, M., Faes, C., Abrahantes, J. C., Imberechts, H., et al. (2008). *Salmonella* in Belgian laying hens: an identification of risk factors. *Prev. Vet. Med.* 83, 323–336. doi: 10.1016/j.prevetmed.2007.09.002

Ni, P. E., Xu, Q., Yin, Y. J., Liu, D. L., Zhang, J. M., Wu, Q. P., et al. (2018). Prevalence and characterization of *Salmonella* serovars isolated from farm products in Shanghai. *Food Control.* 85, 269–275. doi: 10.1016/j.foodcont.2017.10.009

Patchanee, P., Eiamsam-ang, T., Vanaseang, J., Boonkhot, P., and Tadee, P. (2017). Determination of regional relationships among *Salmonella* spp. isolated from retail pork circulating in the Chiang Mai municipality area using a WGS data approach. *Int. J. Food Microbiol.* 254, 18–24. doi: 10.1016/j.ijfoodmicro.2017.05.006

Rahn, K., Degrandis, S. A., Clarke, R. C., Mcewen, S. A., Galan, J. E., Ginocchio, C., et al. (1992). Amplification of an *invA* gene sequence of *Salmonella*-typhimurium by polymerase chain-reaction as a specific method of detection of *Salmonella*. *Mol. Cell Probes* 6, 271–279. doi: 10.1016/0890-8508(92)90002-F

Raspoet, R., Eeckhaut, V., Vermeulen, K., De Smet, L., Wen, Y., Nishino, K., et al. (2019). The *Salmonella* enteritidis TolC outer membrane channel is essential for egg white survival. *Poult. Sci.* 98, 2281–2289. doi: 10.3382/ps/pey584

Sanchez-Salazar, E. M., Gudino, M. E., Sevillano, G., Zurita, J., Guerrero-Lopez, R., Jaramillo, K., et al. (2019). Antibiotic resistance of *Salmonella* strains from layer poultry farms in central Ecuador. *J. Appl. Microbiol.* 128, 1347–1354. doi: 10.1111/jam.14562

Taddese, D., Tolosa, T., Deresa, B., Lakow, M., Olani, A., and Shumi, E. (2019). Antibiograms and risk factors of *Salmonella* isolates from laying hens and eggs in Jimma Town, South Western Ethiopia. *BMC Res. Note* 12:472. doi: 10.1186/s13104-019-4516-5

U.S. Food and Drug Administration (2013). *NARMS Retail Meat Annual Report, 2011*. Rockville, MD: U.S. Food and Drug Administration.

Whiley, H., and Ross, K. (2015). *Salmonella* and eggs: from production to plate. *Int. J. Environ. Res. Public. Health.* 12, 2543–2556. doi: 10.3390/ijerph120302543

Xie, T. F., Wu, G., He, X. J., Lai, Z. Z., Zhang, H. T., and Zhao, J. (2019). Antimicrobial resistance and genetic diversity of *Salmonella enterica* from eggs. *Food Sci. Nutr.* 7, 2847–2853. doi: 10.1002/fsn3.1126

Yang, B. W., Cui, Y., Shi, C., Wang, J. Q., Xia, X. D., Xi, M. L., et al. (2014). Counts, serotypes, and antimicrobial resistance of *Salmonella* isolates on retail raw poultry in the people's republic of China. *J. Food Prot.* 77, 894–902. doi: 10.4315/0362-028X.JFP-13-439

Yang, B. W., Qu, D., Zhang, X. L., Shen, J. L., Cui, S. H., Shi, Y., et al. (2010). Prevalence and characterization of *Salmonella* serovars in retail meats of marketplace in Shaanxi, China. *Int. J. Food Microbiol.* 141, 63–72. doi: 10.1016/j.ijfoodmicro.2010.04.015

Yang, J., Ju, Z. J., Yang, Y., Zhao, X. N., Jiang, Z. Y., and Sun, S. H. (2019). Serotype, antimicrobial susceptibility and genotype profiles of *Salmonella* isolated from duck farms and a slaughterhouse in Shandong province, China. *BMC Microbiol.* 19:202. doi: 10.1186/s12866-019-1570-z

Zhang, Q. Q., Ying, G. G., Pan, C. G., Liu, Y. S., and Zhao, J. L. (2015). Comprehensive evaluation of antibiotics emission and fate in the river basins of China: source analysis, multimedia modeling, and linkage to bacterial resistance. *Environ. Sci. Technol.* 49, 6772–6782. doi: 10.1021/acs.est.5b00729

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Li, Li, Zheng, Wang, Sheng, Shi, Shi, Niu and Yang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 11 September 2020
doi: 10.3389/fmicb.2020.573972

German-Wide Interlaboratory Study Compares Consistency, Accuracy and Reproducibility of Whole-Genome Short Read Sequencing

*Laura Uelze[1], Maria Borowiak[1], Markus Bönn[2], Erik Brinks[3], Carlus Deneke[1], Thomas Hankeln[4], Sylvia Kleta[1], Larissa Murr[5], Kerstin Stingl[1], Kathrin Szabo[6], Simon H. Tausch[1], Anne Wöhlke[7] and Burkhard Malorny[1]**

[1] Department of Biological Safety, German Federal Institute for Risk Assessment (BfR), Berlin, Germany, [2] Landesamt für Verbraucherschutz Sachsen-Anhalt (LAV), Halle (Saale), Germany, [3] Department of Microbiology and Biotechnology, Max Rubner-Institut (MRI), Kiel, Germany, [4] Institute of Organismic and Molecular Evolution, AG Molecular Genetics and Genome Analysis, Johannes Gutenberg Universität Mainz, Mainz, Germany, [5] Bavarian Health and Food Safety Authority (LGL), Oberschleißheim, Germany, [6] Department 5, Federal Office of Consumer Protection and Food Safety (BVL), Berlin, Germany, [7] Food and Veterinary Institute, Lower Saxony State Office for Consumer Protection and Food Safety (LAVES), Braunschweig, Germany

OPEN ACCESS

Edited by:
Dario De Medici,
National Institute of Health (ISS), Italy

Reviewed by:
Errol A. Strain,
United States Food and Drug Administration, United States
Heather A. Carleton,
Centers for Disease Control and Prevention (CDC), United States
Bo Segerman,
National Veterinary Institute, Sweden

***Correspondence:**
Burkhard Malorny
burkhard.malorny@bfr.bund.de

Specialty section:
This article was submitted to Food Microbiology, a section of the journal Frontiers in Microbiology

Received: 18 June 2020
Accepted: 14 August 2020
Published: 11 September 2020

Citation:
Uelze L, Borowiak M, Bönn M, Brinks E, Deneke C, Hankeln T, Kleta S, Murr L, Stingl K, Szabo K, Tausch SH, Wöhlke A and Malorny B (2020) German-Wide Interlaboratory Study Compares Consistency, Accuracy and Reproducibility of Whole-Genome Short Read Sequencing. Front. Microbiol. 11:573972. doi: 10.3389/fmicb.2020.573972

We compared the consistency, accuracy and reproducibility of next-generation short read sequencing between ten laboratories involved in food safety (research institutes, state laboratories, universities and companies) from Germany and Austria. Participants were asked to sequence six DNA samples of three bacterial species (*Campylobacter jejuni*, *Listeria monocytogenes* and *Salmonella enterica*) in duplicate, according to their routine in-house sequencing protocol. Four different types of Illumina sequencing platforms (MiSeq, NextSeq, iSeq, NovaSeq) and one Ion Torrent sequencing instrument (S5) were involved in the study. Sequence quality parameters were determined for all data sets and centrally compared between laboratories. SNP and cgMLST calling were performed to assess the reproducibility of sequence data collected for individual samples. Overall, we found Illumina short read data to be more accurate (higher base calling accuracy, fewer miss-assemblies) and consistent (little variability between independent sequencing runs within a laboratory) than Ion Torrent sequence data, with little variation between the different Illumina instruments. Two laboratories with Illumina instruments submitted sequence data with lower quality, probably due to the use of a library preparation kit, which shows difficulty in sequencing low GC genome regions. Differences in data quality were more evident after assembling short reads into genome assemblies, with Ion Torrent assemblies featuring a great number of allele differences to Illumina assemblies. Clonality of samples was confirmed through SNP calling, which proved to be a more suitable method for an integrated data analysis of Illumina and Ion Torrent data sets in this study.

Keywords: interlaboratory study, whole-genome sequencing, food safety, illumina, ion torrent

Abbreviations: BLAST, basic local alignment search tool; cgMLST, core genome multilocus sequence typing; DNA, deoxyribonucleic acid; GMI, global microbial identifier; MLST, multilocus sequence typing; NGS, next-generation sequencing; PT, proficiency testing; SNP, single-nucleotide polymorphism; ST, sequence type; wgMLST, whole-genome MLST; WGS, whole genome sequencing.

INTRODUCTION

Whole genome sequencing is a high resolution, high-throughput method for the molecular typing of bacteria. Through bioinformatic analysis of bacterial genome sequences, it is not only possible to identify bacteria on a species and sub-species level, but also to identify antimicrobial resistance and virulence genes. Further, it is possible through a variety of methods, such as variant calling, k-mer based, or gene-by-gene approaches, to determine the relatedness/clonality between bacterial isolates, making it the ideal tool for outbreak studies, routine surveillance and clinical diagnostics (Ronholm et al., 2016). Initially expensive and difficult to set up, the technology is becoming continuously more user-friendly and affordable (Uelze et al., 2020b). In recent years, funding provided through federal initiatives has enabled public health and food safety laboratories in Germany and worldwide to acquire sequencing platforms. A number of different sequencing technologies exist, each with their own upsides and shortcomings. For example, Illumina sequencing platforms generally produce relatively short paired-end sequencing reads with high accuracy, while the Ion Torrent technology outputs single-end reads with often greater read lengths, but higher error rates (Quail et al., 2012; Fox et al., 2014; Salipante et al., 2014; Kwong et al., 2015; Escalona et al., 2016). Which sequencing platform different laboratories choose to acquire is not only dependent on financial resources, but also on individual needs and routine applications, with throughput, error rates/error types, read lengths and run time as the main concerning parameters. This leads to an increased diversification of the sequencing community (Moran-Gilad et al., 2015), creating a natural competition between producers, which benefits users through an ongoing improvement of technology and equipment. However, diversification also hampers standardization and despite ongoing calls for the establishment of agreed minimal sequencing quality parameters, this process has been much delayed (Endrullat et al., 2016).

Increasingly, microbial disease surveillance systems are based on WGS data. For example, Pathogenwatch[1] is a global platform for genomic surveillance, which analyses genomic data submitted by users and conducts cgMLST clustering to monitor the spread of important bacterial pathogens. Similarly, the GenomeTrakr network (FDA) (Timme et al., 2019) uses whole-genome sequence data and performs cg/wgMLST and SNP calling to track food-borne pathogens integrated into NCBI Pathogen Detection[2].

Other large WGS surveillance programs include PulseNet (Tolar et al., 2019) run by the Centers for Disease Control and Prevention (CDC), as well as a genomic surveillance program established by Public Health England (Ashton et al., 2016).

In Germany, a network of Federal State Laboratories and Federal Research Institutions supports the investigation of food-borne outbreaks through traditional typing and WGS methods. All genomic surveillance systems have in common that a high quality and accuracy of the sequencing data is crucial for a robust and reliable data analysis.

Proficiency testing (PT) is an important external quality assessment tool to compare the ability and competency of individual laboratories to perform a method, whereas the aim of an interlaboratory study is to compare the performance of a method, when conducted by different collaborators. Several PT exercises with the focus on the sequencing of microbial pathogens have been published in recent years. In 2015, the GenomeTrakr network conducted a PT with 26 different US laboratories, which were instructed to sequence eight bacterial isolates according to a fixed protocol (Timme et al., 2018). In the same year, the GMI initiative conducted an extensive survey with the aim to assess requirements and implementation strategies of PTs for bacterial WGS (Moran-Gilad et al., 2015), followed by a series of global PT exercises[3]. In an interlaboratory exercise in 2016, five laboratories from three European countries (Denmark, Germany, the Netherlands) were asked to sequence 20 *Staphylococcus aureus* DNA samples according to a specific protocol and report cgMLST cluster types (Mellmann et al., 2017). In this study, we present the results of an interlaboratory study for short-read bacterial genome sequencing with ten participating laboratories from German-speaking countries initiated by the §64 German Food and Feed Code (LFGB) working group "NGS Bacterial Characterization" chaired by the Federal Office of Consumer Protection and Food Safety (BVL). The working group serves to validate and standardize WGS methods for pathogen characterization in the context of outbreak investigations. The interlaboratory study was carried out by the German Federal Institute of Risk Assessment (BfR) in 2019, with the aim to answer the question whether different WGS technology platforms provide comparable sequence data, taking into account the routine sequencing procedures established in these laboratories.

MATERIALS AND METHODS

Study Design

In the frame of the §64 LFGB working group "NGS Bacterial Characterization", we conducted a interlaboratory study for next-generation sequencing. Twelve teams participated in the study. Participants included four Federal Research Institutes (3 German, 1 Austrian), four German State Laboratories, one German university and three German companies.

Participants were provided with DNA samples (40–55 µl, 60–187 ng/µl) of six bacterial isolates (**Table 1**) (two of each *Campylobacter jejuni*, *Listeria monocytogenes* and *Salmonella enterica*), with the species of the sample visibly marked on the tube containing the sample DNA.

Participants were instructed to sequence the samples according to their standard in-house sequencing procedure. Where possible, participants were asked to sequence each isolate in two independent sequencing runs with two independent library preparation steps. No minimum quality criteria for the

[1]https://pathogen.watch

[2]https://www.ncbi.nlm.nih.gov/pathogens/

[3]https://www.globalmicrobialidentifier.org/Workgroups/GMI-Proficiency-Test-Reports

TABLE 1 | Strain characteristics of analyzed DNA samples (species, serovar, MLST, size and GC content) used in the interlaboratory study.

Sample	Species	Serovar	MLST	Size (Mbp)	GC (%)
19-RV1-P64-1	*Campylobacter jejuni*		4774	1.6	30.5
19-RV1-P64-2	*Campylobacter jejuni*		21	1.7	30.5
19-RV1-P64-3	*Listeria monocytogenes*	IIc	9	3.0	38.0
19-RV1-P64-4	*Listeria monocytogenes*	IIb	59	3.0	37.9
19-RV1-P64-5	*Salmonella enterica* subsp. *enterica*	Infantis	32	5.1	52.0
19-RV1-P64-6	*Salmonella enterica* subsp. *enterica*	Paratyphi B var. Java	28	4.8	52.2

resulting sequencing data were requested. Together with the samples, participants received a questionnaire to document their applied sequencing method. Participants were given 4 weeks to conduct the sequencing and report the resulting raw sequencing data. Sequencing data was exchanged through a cloud-based platform and data quality was centrally analyzed with open-source programs and in-house bioinformatic pipelines. Results of the sequencing data analysis were presented to the members of the §64 LFGB working group in November 2019. Following the meeting, ten participants agreed to a publication of the results of the interlaboratory study. Two participants declined a publication of their data due to a conflict of interest. Participants are anonymously identified with their laboratory code LC01 – LC10 assigned for this study.

Study Isolates, Cultivation and DNA Isolation

Detailed information to the samples is summarized in **Supplementary File S1**.

The samples 19-RV1-P64-1 and 19-RV1-P64-2 were obtained from *Campylobacter jejuni* isolates (MLST type 4774 and 21, respectively). *Campylobacter jejuni* were pre-cultured on Columbia blood agar, supplemented with 5% sheep blood (Oxoid, Wesel, Germany) for 24 h at 42°C under micro-aerobic atmosphere (5% O2; 10% CO2). A single colony was inoculated on a fresh Columbia blood agar plate for an additional 24 h. After incubation, bacterial cells were re-suspended in buffered peptone water (Merck, Darmstadt, Germany) to an OD600 of 2. Genomic DNA was extracted from this suspension with the PureLink Genomic DNA Mini Kit (Thermo Fisher Scientific, Dreieich, Germany) according to manual instructions.

The samples 19-RV1-P64-3 and 19-RV1-P64-4 were obtained from *Listeria monocytogenes* serovar IIc and serovar IIb, respectively. *Listeria monocytogenes* were cultured on sheep blood agar plates and incubated at 37°C over night. Genomic DNA was directly extracted from bacterial colonies using the QIAamp DNA Mini Kit (Qiagen, Hilden, Germany) following the manual instructions for gram-positive bacteria.

The samples 19-RV1-P64-5 and 19-RV1-P64-6 were obtained from *Salmonella enterica* subsp. *enterica* serovar Infantis and serovar Paratyphi B var. Java, respectively.

Salmonella enterica were cultivated on LB agar (Merck). A single colony was inoculated in 4 ml liquid LB and cultivated under shaking conditions (180–220 rpm) at 37°C for 16 h. Genomic DNA was extracted from 1 ml liquid cultures using the PureLink Genomic DNA Mini Kit (Thermo Fisher Scientific) according to manual instructions.

DNA quality of all samples was verified with Nanodrop and Qubit and samples were stored at 4°C before being express shipped to the participants in liquid form on ice.

PacBio Reference Sequences

As Pacific Biosciences (herein abbreviated as PacBio) sequencing was performed before the interlaboratory study started, DNA extractions used for PacBio sequencing differed from DNA extractions used for short read-sequencing. For *Campylobacter jejuni*, *Listeria monocytogenes* and *Salmonella enterica* the PureLink Genomic DNA Mini Kit (Invitrogen) was used for genomic DNA extraction.

PacBio sequences for samples 19-RV1-P64-1 to 19-RV1-P64-5 were obtained from GATC as described before (Borowiak et al., 2018).

Sample 19-RV1-P64-6 was sequenced in-house. Genomic DNA was sheared to approximately 10 kb using g-Tubes (Covaris, Brighton, United Kingdom) and library preparation was performed using the SMRTbell Template Prep Kit 1.0 and the Barcode Adapter Kit 8A (Pacific Biosciense, Menlo Park, CA, United States). Sequencing was performed on a PacBio Sequel instrument using the Sequel Binding Kit and Internal control Kit 3.0 and the Sequel Sequencing Kit 3.0 (PacBio). Long read data was assembled using the HGAP4 assembler.

Information to the PacBio sequences is summarized in **Supplementary File S1**.

Whole-Genome Short Read Sequencing

All ten participants followed their own in-house standard protocol for sequencing. Important sequencing parameters such as the type of library preparation and sequencing kits, as well as the type of sequencing instrument were documented with a questionnaire (the questionnaire template in German language is provided as **Supplementary File S2**). The results of the questionnaire are summarized in **Supplementary File S3**. All participants determined the DNA concentration prior to sequencing library preparation. Of ten participants, seven chose a enzymatic digest for DNA fragmentation, while three laboratories fragmented DNA through mechanical breakage. Over half of participants pooled sequence libraries relative to genome sizes and almost all (with the exception of laboratory LC01) included a control in the sequencing run (i.e., PhiX).

All participants, with the exception of laboratories LC02 and LC08, sequenced samples in duplicates. Duplicates were defined as one sample sequenced in two independent sequencing runs on the same sequencing instrument, henceforth identified as sequencing run A and sequencing run B. Participants LC01, LC03, LC04, LC05, LC06, LC07, LC09, LC10 contributed 12

whole-genome sequencing data sets (combined forward and reverse reads) each, while participant LC08 contributed 6 whole-genome sequencing data sets. In contrast, laboratory LC02 sequenced the complete sample set on three different sequencing instruments in single runs, henceforth identified as LC02_a (Illumina iSeq), LC02_b (Illumina MiSeq), LC02_c (Illumina NextSeq). Therefore, participant LC02 contributed 18 whole-genome sequencing data sets.

Together, 120 whole-genome sequencing data sets were available for analysis.

Taken the fact into consideration, that participant LC02 used three different sequencing instruments, a total of twelve individual sequencing instruments were included in the interlaboratory study: one Ion Torrent S5 instrument (Thermo Fisher Scientific), two iSeq, six MiSeq, two NextSeq and one NovaSeq instrument (all Illumina).

Assessment of Raw Sequencing Data Quality

The quality of the raw sequencing reads was assessed with fastp (Chen et al., 2018) with default parameters. Quality control parameters for each data set (forward and reverse reads for Illumina data, single reads for Ion Torrent), such as the number of total reads and the Q30 (both before filtering), were parsed from the resulting fastp json reports. The coverage depth was calculated as the sum of the length of all raw reads divided by the length of the respective PacBio reference sequence.

Short-Read Genome Assemblies

Untrimmed Ion Torrent reads were *de novo* assembled with SPAdes v3.13.1 (Nurk et al., 2013) with read correction. For SNP calling, Ion Torrent reads were trimmed using fastp v0.19.5 (Chen et al., 2018) with parameters *–cut_by_quality3 –cut_by_quality5 –cut_window_size 4 –cut_mean_quality 30*.

Raw Illumina reads were trimmed and *de novo* assembled with our in-house developed Aquamis pipeline[4] which implements fastp (Chen et al., 2018) for trimming and shovill (based on SPAdes)[5] for assembly. Unlike SPAdes, shovill automatically down samples reads to a coverage depth of 100× prior to assembling.

Assessment of Genome Assembly Quality and Bacterial Characterization

Quality of the genome assemblies was assessed with QUAST v5.0.2[6] without a reference. Quality parameters such as number of contigs, length of largest contig and N50 were parsed from the QUAST report for each assembly.

Based on the genome assemblies (including the PacBio reference sequences), bacterial characterization was conducted with our in-house developed Bakcharak pipeline[7] which implements among other tools, ABRicate for antimicrobial resistance and virulence factor screening[8], and the PlasmidFinder database for plasmid detection (Carattoli et al., 2014), mlst[9], SISTR (Yoshida et al., 2016) for *in silico* Salmonella serotyping and Prokka (Seemann, 2014) for gene annotation.

CgMLST Allele Calling

CgMLST allele calling was conducted with our in-house developed chewieSnake pipeline[10] which implements chewBBACA (Silva et al., 2018). Only complete coding DNA sequences, with start and stop codon, according to the NCBI genetic code table 11, are identified as alleles by chewBBACA [with Prodigal 2.6.0 (Hyatt et al., 2010)]. The default of 0.6 was used as the minimum BLAST score ratio for defining locus similarity (*–bsr 0.6*). Furthermore alleles 20% larger or smaller then the average length for one locus were excluded (*–st 0.2*). CgMLST allele distance matrices were computed with grapetree (ignoring missing data in pairwise comparison).

CgMLST schemes for *Listeria monocytogenes* (Ruppitsch et al., 2015) were derived from the cgMLST.org nomenclature server[11]. CgMLST schemes for *Campylobacter jejuni* and *Salmonella enterica* were derived from the chewBBACA nomenclature server[12].

SNP Calling

SNP (single-nucleotide polymorphism) calling was conducted for each sample. Sequencing reads were trimmed prior to SNP calling. Assembled uncirculated PacBio sequences of the samples were used as reference sequences for SNP calling. SNP calling was conducted with our in-house developed snippySnake pipeline[13] which implements Snippy v4.1.0[14]. Within Snippy, variants are called with freebayes (Garrison and Marth, 2012). SnippySnake was run with the following parameters: *mapqual: 60; basequal: 13; mincov: 10; minfrac: 0; minqual: 100; maxsoft: 10*. Only substitutions are consider as SNPs (indels and complex variants are removed during a filtering step). Compound SNPs are broken into single SNPs.

RESULTS

Comparison of Quality of Sequencing Reads

One important parameter to assess the quality of sequencing reads is the phred quality score. Commonly a Q score of 30 is used, which indicates a base call accuracy of ≥99.9%. We compared the percentages of bases that have a quality score equal or larger to 30. The results visualized in **Figure 1** (see **Supplementary File S4** for exact numbers), show that on average ∼ 90% of Illumina bases have a Q score ≥ Q30.

[4]https://gitlab.com/bfr_bioinformatics/AQUAMIS/

[5]https://github.com/tseemann/shovill

[6]https://github.com/ablab/quast

[7]https://gitlab.com/bfr_bioinformatics/bakcharak

[8]https://github.com/tseemann/abricate

[9]https://github.com/tseemann/mlst

[10]https://gitlab.com/bfr_bioinformatics/chewieSnake

[11]https://www.cgmlst.org/

[12]http://chewbbaca.online/

[13]https://gitlab.com/bfr_bioinformatics/snippy-snake

[14]https://github.com/tseemann/snippy

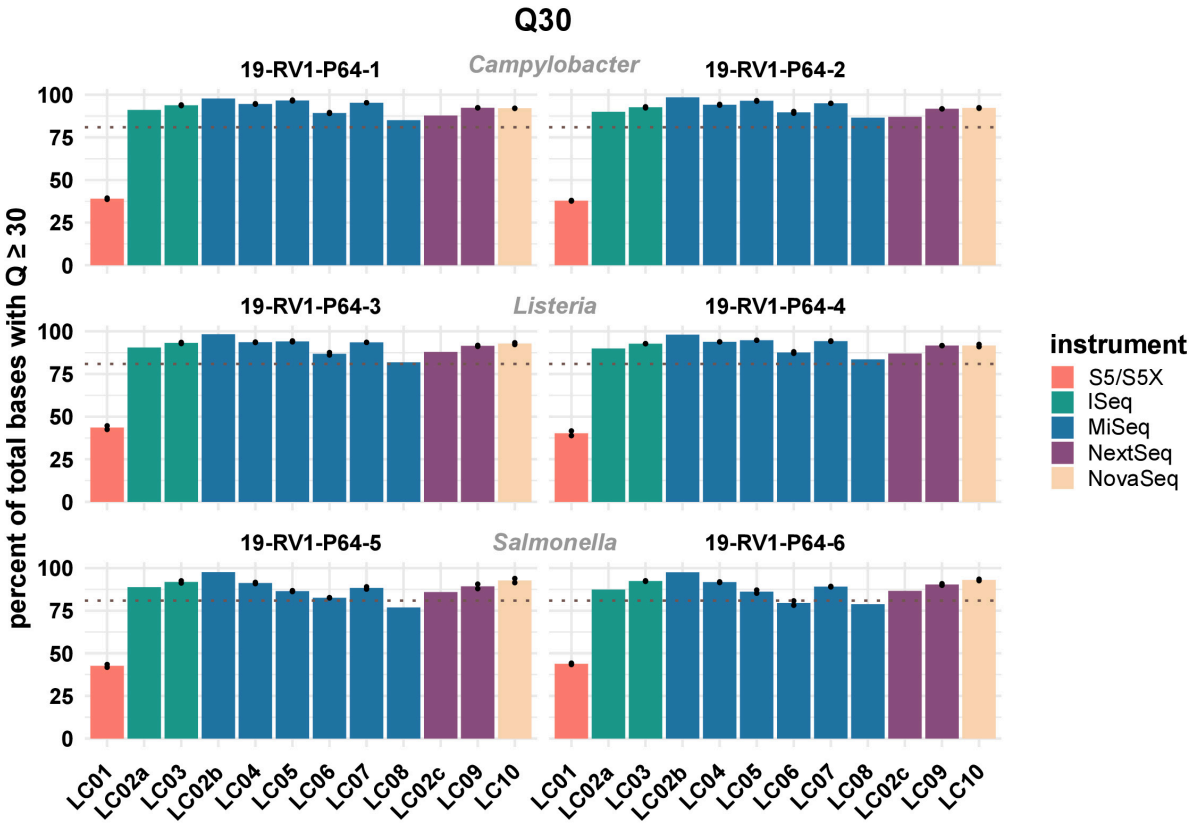


FIGURE 1 | The bar plot shows the mean percentage of total bases with a phred score above or equal to Q30 grouped by laboratories and samples. Line-connected points indicate the variance between sequencing runs (run A/run B), with the exception of laboratories LC02 and LC08 (single sequencing run). Fill colors identify the sequencing instrument. The species of the samples is indicated. The dotted line marks a Q30 of 80%.

For Ion Torrent, ∼40% of bases achieve a Q score ≥ Q30. However, since base calling and quality prediction algorithms for Illumina and Ion Torrent are different, Ion Torrent reads are usually assessed with a Q score of ≥Q20, hampering a direct comparison.

There is little variation within the Illumina instrument series (mean values: iSeq: 91.7%; MiSeq: 90.8%; NextSeq: 90.4%; NovaSeq: 92.4%), indicating that no particular instrument of the series out or under performs the others. In contrast, sequencing data with higher or lower quality scores was consistently associated with individual laboratories. Among the participants employing Illumina instruments, LC08 overall produced the lowest quality data (LC08 mean: 82.1%), while LC02_b produced the highest quality data (LC02_b mean: 97.9%), both with a MiSeq instrument. Interestingly, the same laboratory LC02, remained below the average for Illumina data when employing a NextSeq instrument (LC02_c mean: 87.1%). Of course, sequence quality might also depend on loading concentration and number of cycles used for sequencing. Quality scores remained largely consistent between runs. Equally, the type of bacterial species had little influence on sequencing data quality.

Sufficient coverage depth (in this study calculated as the total number of bases divided by the length of the PacBio reference) is an important requirement for successful downstream analysis, such as variant detection and assembly. However, up to now there is no widespread consensus for the recommended minimum coverage depth for bacterial WGS. In the accompanying questionnaire, participants stated that they intended to achieve coverage depths ranging from >20× to <300×, with most participants opting for a coverage depth of 60× to 70×. Actual coverage depths ranged from 26× (LC03, 19-RV1-P64-5, run A) to 1201× (LC10, 19-RV1-P64-1, run B), with most data sets featuring coverage depths from 75× to 196× ($Q_{0,25}$ and $Q_{0,75}$) as shown in **Figure 2**. With the exception of a small number of data sets (LC03: 19-RV1-P64-2, 19-RV1-P64-5, 19-RV1-P64-6; LC05: 19-RV1-P64-6, all run A), all data sets were well above a coverage depth of 30×. Coverage depths varied between laboratories, instruments and samples, as well as between sequencing runs. Laboratory LC10 produced data sets with very high coverage depths with an average of 736×. When coverage depths were normalized by assigning a coverage depth of 1 to sample 19-RV1-P64-1 for each laboratory, we found that coverage depths varied in a predictable manner in relation to the genome size of the sample as shown in **Figure 3**. Some participants chose to pool sequencing libraries relative to genome sizes of the samples, which in most cases ensured a more consistent sequencing depth across the samples (LC02_a, LC03, LC04, LC06). In comparison, participants that pooled sequencing libraries of all

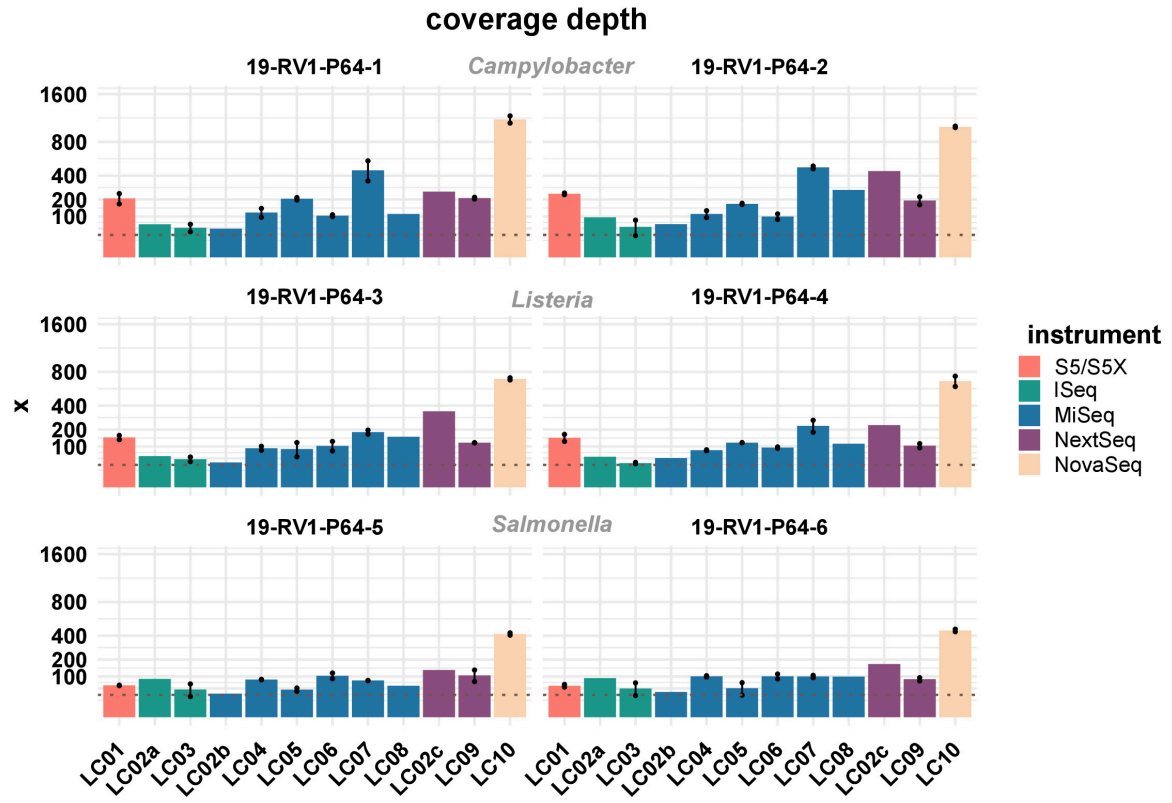


FIGURE 2 | The bar plot shows the mean coverage depth grouped by laboratories and samples. Line-connected points indicate the variance between sequencing runs (run A/run B), with the exception of laboratories LC02 and LC08 (single sequencing run). The coverage depth was defined as the sum of the length of all raw reads divided by the length of the respective PacBio reference sequence. Fill colors identify the sequencing instrument. The species of the samples is indicated. The *y*-axis is squared. The dotted line marks a coverage depth of 30×.

samples equally (LC01, LC05, LC07, LC08, LC10) obtained lower coverage depths for bacterial isolates with larger genome sizes (i.e., ∼4.9 Mbp for *Salmonella enterica*), and higher coverage depths for bacterial isolates with smaller genome sizes (i.e., ∼1.7 Mbp *Campylobacter jejuni*). However, in most cases pooling the DNA libraries relative to genome size only reduced the impact of the genome size effect, without eliminating it. Only laboratory LC06 achieved a high consistency across all samples.

Comparison of Quality of Genome Assemblies and Bacterial Characterization

The genome assemblies constructed from the short read data were assessed and the determined quality parameters are listed in **Supplementary File S4**. We found little variation in the lengths of the genome assemblies (sd values for the samples ranged from ∼3 Kbp to ∼11 Kbp). However, all short read assemblies were ∼36 to ∼66 Kbp shorter than their respective PacBio references, likely due to overlapping end regions in the PacBio sequences, which were not circularized prior to analysis.

Similarly, there was little variation for the calculated GC values (sd values for the samples ranged from 0.01 to 0.03%). Besides the length, the quality of genome assemblies is determined by the total number of contigs, and the size of the largest contig, with assemblies featuring fewer, larger contigs generally being more useful for downstream analyses. Both parameters are combined in the N50 value, which is defined as the length of the shortest contig in the set of largest contigs that together constitute at least half of the total assembly size. The N50 values for all assemblies are visualized in **Figure 4**. We found N50 values to be overall very similar for individual samples, regardless of which laboratory or instrument provided the sequencing data, with a few notable exceptions (i.e., LC06, LC08). In general, highest N50 values were obtained for *Listeria monocytogenes* (19-RV1-P64-3: ∼600 Kbp; 19-RV1-P64-4: ∼480 Kbp), followed by *Salmonella enterica* (19-RV1-P64-5: ∼200 Kbp; 19-RV1-P64-6: ∼340 Kbp), and *Campylobacter jejuni* (19-RV1-P64-1: ∼220 Kbp; 19-RV1-P64-2: ∼180 Kbp).

Assemblies of laboratories LC06 and LC08 consistently had much lower N50 values (also shown by a higher total number of contigs and shorter contigs lengths), compared to the other laboratories. For example, while the majority of assemblies achieved an N50 of ∼ 605 Kbp (±550 bp) for sample 19-RV1-P64-3, the N50 for assemblies of LC06 ranged around ∼ 256 Kbp, and the N50 for assemblies of laboratory LC08 was even lower (∼71 Kbp). Interestingly, no linear correlation was apparent

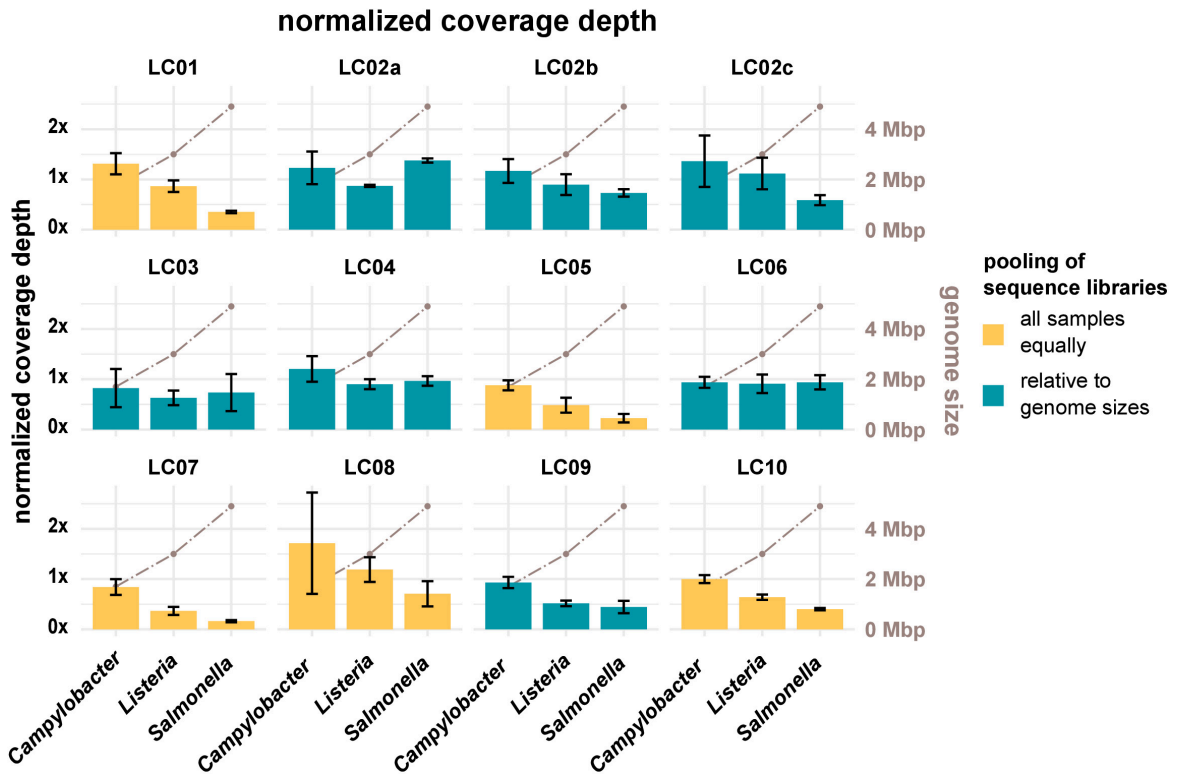


FIGURE 3 | The bar plot (left *y*-axis) shows the mean normalized coverage depth grouped by laboratories and species of the samples with error bar. The coverage depth was normalized for each laboratory to the coverage depth for sample 19-RV1-P64-1, sequencing run A, which was assigned a value of 1. The coverage depth was defined as the sum of the length of all raw reads divided by the length of the respective PacBio reference sequence. Fill colors identify, whether DNA libraries were pooled relative to genome sizes prior to sequencing or whether DNA libraries were pooled equally. The brown line graph in the background (right *y*-axis) indicates the average genome size of the species.

between the N50 value and the coverage depth as shown in **Figure 5**.

Coding frames in the genome assemblies were annotated to determine the MLST type, as well as identify resistance and virulence genes. In total, there was little variation for the total number of detected CDS (defined as a sequence containing a start and stop codon). The total number of CDSs varied by sample (19-RV1-P64-1: n = ∼1597; 19-RV1-P64-2: n = ∼1713; 19-RV1-P64-3: n = ∼2892; 19-RV1-P64-4: n = ∼2913; 19-RV1-P64-5: n = ∼4667; 19-RV1-P64-6: n = ∼4393) with a standard deviation of 8 to 15 coding frames (compare **Supplementary File S4**).

The Multilocus Sequence Type (MLST) was determined correctly for all data sets. The same plasmid markers could be detected from all short read genome assemblies. Two more plasmid markers (*Col8282_1* and *ColRNAI_1*) could be detected in the short read assemblies compared to the PacBio reference for 19-RV1-P64-6, likely due to the fact that small plasmids are often excluded from PacBio sequences (as read lengths are too short). In three cases, resistance genes detected in the PacBio references were not present in the short read assemblies: $bla_{OXA-184}$ in 19-RV1-P64-1, of laboratory LC06 (run A) and *aadA1* in 19-RV1-P64-6, of laboratory LC09 (both runs).

Although overall the same sets of virulence genes were detected from the short-read assemblies, there was some variation with assemblies from laboratories LC01, LC06 and LC08 often missing virulence genes (**Supplementary File S4**). For example, virulence factors *flaA* and *flaB* could not be detected in assemblies from laboratory LC01 for sample 19-RV1-P64-1. Contrary, *flaA* and *flaB* were generally not detected in assemblies for sample 19-RV1-P64-2, with the exception of both assemblies from laboratory LC01. In another example the genes *sopD2* and *sseK1* could not be detected from the assembly for sample 19-RV1-P64-5 from laboratory LC08. The absence of virulence and resistance genes is likely caused by assembly issues where genes are broken at contig borders.

Analysis of cgMLST Calling Results

CgMLST was conducted to compare the effect of differences in the genome assemblies on clustering. All cgMLST distance allele matrices are presented in **Supplementary File S5**. The cgMLST distance matrix for sample 19-RV1-P64-1 is visualized in **Figure 6**. CgMLST distance matrices for the six samples were overall very similar. In general, most assemblies had zero allele differences. However, assemblies constructed from Ion Torrent short read data (LC01) generally had a much higher number of allele differences, than those constructed from Illumina short reads. For easy comparison, we calculated the 'median cgMLST

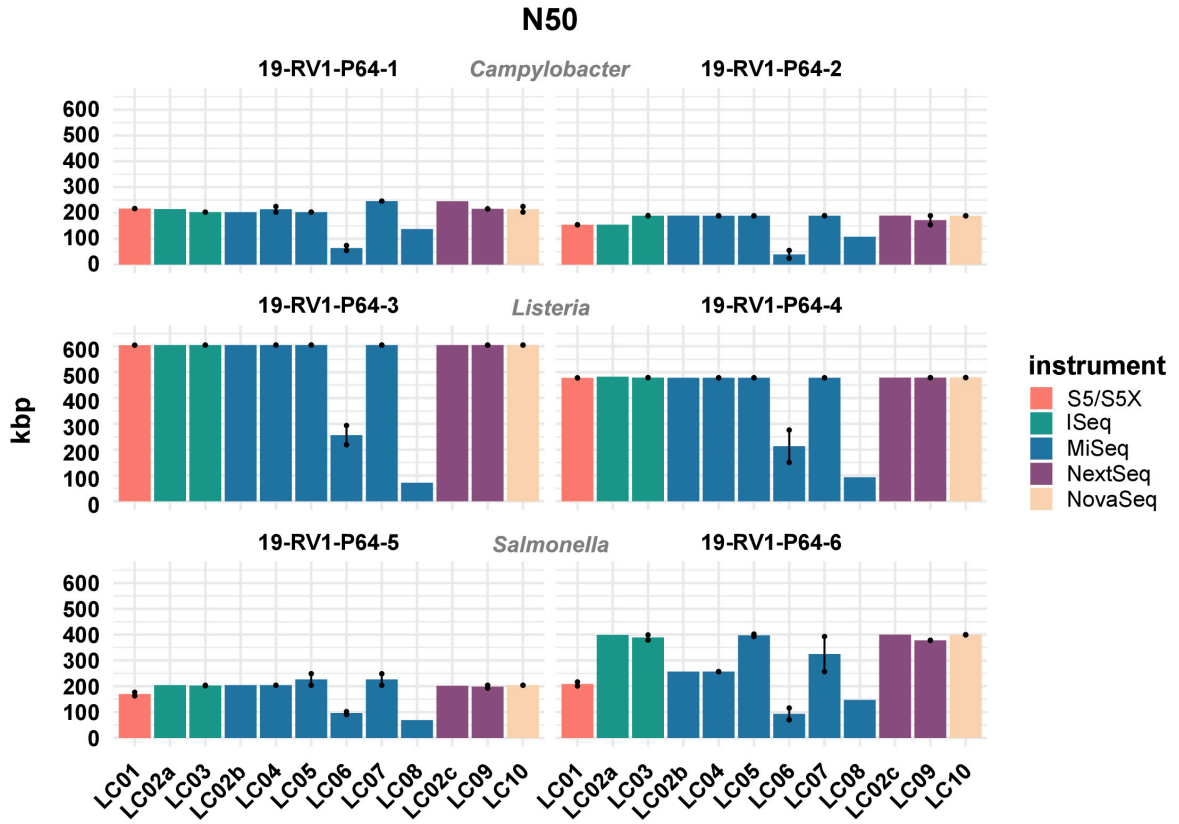


FIGURE 4 | The bar plot shows the mean N50 determined for the short-read genome assemblies grouped by laboratories and samples. Line-connected points indicate the variance between sequencing runs (run A/run B), with the exception of laboratories LC02 and LC08 (single sequencing run). Fill colors identify the sequencing instrument. The species of the samples is indicated.

distance' for each assembly, by computing the medium of all allele differences to a specific assembly (compare **Figure 6**).

Figure 7 shows the median cgMLST distance for all assemblies. As mentioned, the highest number of allele differences were calculated for the assemblies of laboratory LC01 (using an Ion Torrent instrument). However, allele differences for the Ion Torrent assemblies varied dependent on the species of the sample. The smallest number of cgMLST allele differences were obtained for *Listeria monocytogenes* (LC01: ~ 7.1), followed by *Campylobacter jejuni* (LC01: ~ 11.1) and *Salmonella enterica* (LC01: ~ 26.1). Illumina assembly generally had much lower allele differences. Median cgMLST allele differences for the assemblies of the laboratories LC02a, LC02b, LC02c, LC03, LC04, and LC010 were zero for all samples. Median allele differences for assemblies of the laboratories LC05, LC06, LC07, LC08, and LC09 were between zero and three, often slightly higher for laboratories LC05 and LC08. Interestingly, the assembly of sample 19-RV1-P64-6 produced in run A by laboratory LC05 featured a median number of 10 alleles, while the assembly of run B by laboratory LC05 had a median number of zero allele differences.

We further compared the effect of the assembly algorithm on the cgMLST calling by assembling trimmed Illumina reads with SPAdes (as opposed to shovill) prior to cgMLST calling. However, no significant difference was found in the number of allele differences (data not shown).

Analysis of SNP Calling Results

SNP calling was conducted to detect sequencing errors. The assembled PacBio sequences were used as reference sequences. All SNP distance allele matrices are presented in **Supplementary File S6**. No SNPs were detected within the data sets. Equally, all data sets featured zero SNPs to the reference sequence, with the exception of the PacBio reference for sample 19-RV1-P64-5, to which all data sets had 2 SNPs.

DISCUSSION

We conducted an interlaboratory study for the investigation of the reproducibility and consistency of bacterial whole-genome sequencing. Ten participants were instructed to sequence six DNA samples in duplicate according to their in-house standard procedure protocol. We were interested to see, how the quality of sequencing data varied across different sequencing instruments, library preparation kits, sequencing kits and individual expertise of the participating laboratories. Overall, we were able to

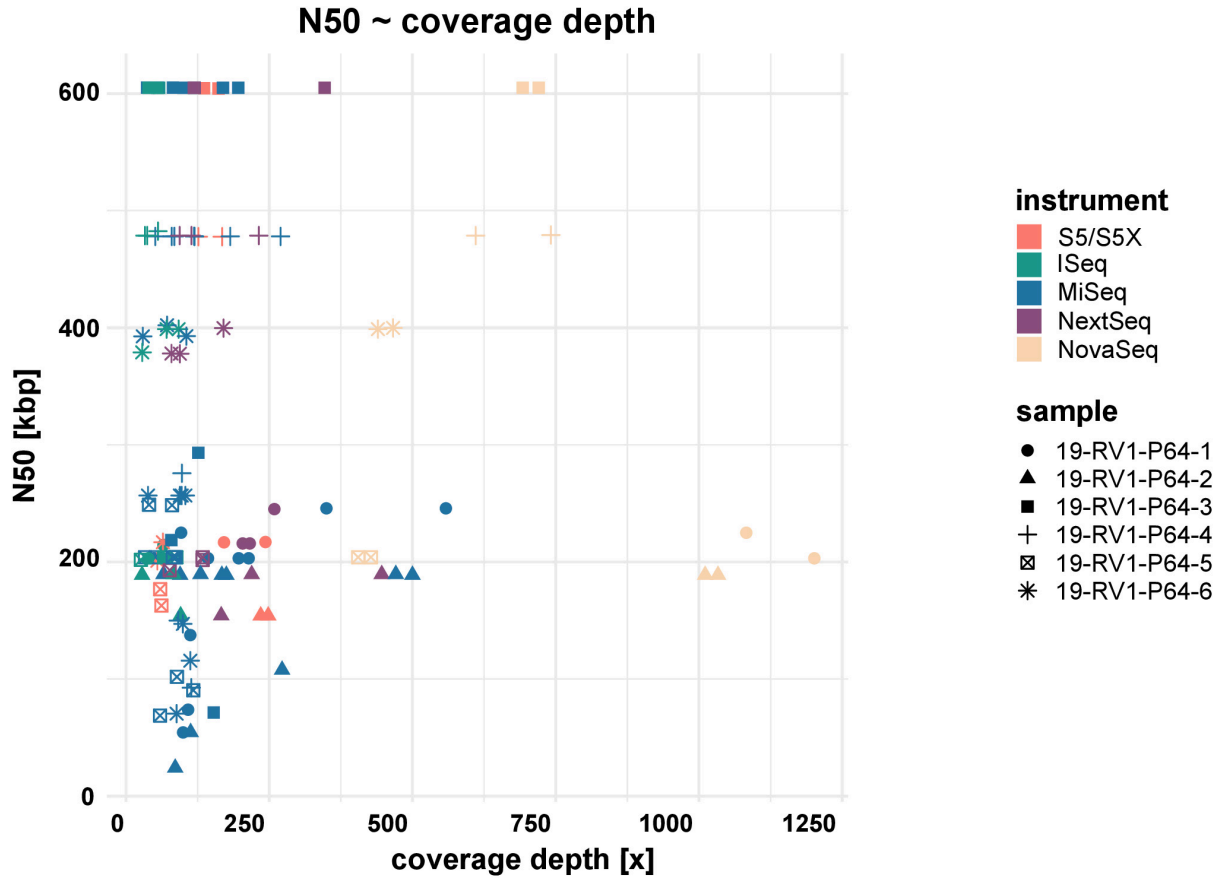


FIGURE 5 | The dot plot shows the correlation between N50 and coverage depth for the short-read genome assemblies/sequence data sets. Fill colors indicate the sequencing instrument. The shape of the dots identifies the sample.

compare 12 Illumina sequencing instruments and one Ion Torrent instrument.

It has been established that different sequencing technologies vary in their average error rates, with Ion Torrent data generally having higher error rates compared to Illumina (Quail et al., 2012; Fox et al., 2014; Salipante et al., 2014; Kwong et al., 2015; Escalona et al., 2016). Indeed, we assessed that Ion Torrent bases achieved much lower quality scores than Illumina bases. Interestingly, we found the four different Illumina sequencing instruments types involved in our study (iSeq, MiSeq, NextSeq, NovaSeq) to be very similar in terms of base quality, suggesting that the underlying sequencing technology is similar, despite the different color chemistry used.

There was a great variety in coverage depths that participants obtained for their data sets (ranging from 26× to 1200×). Although no widely accepted minimal coverage depth for bacterial whole-genome sequencing is established yet, most studies recommended coverage depths ranging from ≥30× to ≥50× (Chun et al., 2018). Positively, most data sets submitted by the participants in our study had coverage depths well above 30×, demonstrating that insufficient coverage depth is not usually a concern. However, coverage depths frequently fell short of the intended coverage depths stated by participants in the accompanying questionnaire, indicating that this parameter is not always well controlled for. For example, while laboratory LC02_b aimed for a coverage depth of ≥60×, the majority of data sets submitted by this laboratory had a much lower coverage depth (30–50×). Similarly, laboratory LC01, LC02a, LC05 and LC08 frequently obtained lower than intended coverage depths.

Resulting from experience and producer instructions, users generally know the number of reads/total bases that their sequencing instrument is capable of producing in one sequencing run. By pooling DNA libraries relative to genome sizes (provided the species of the isolates is known), users can influence the number of reads/bases and therefore the coverage depth for each isolate. As was shown in this study, participants that pooled DNA libraries prior to sequencing relative to genome sizes achieved more consistent coverage depths across the three species (e.g., LC06), while participants that pooled all DNA libraries equally, obtained sequencing data with predictable fluctuation in coverage depth (i.e., LC10), depending on the genome size of the organism.

Both, too low (problematic for variance calling/fragmented assembly) and too high (increased "noisiness" of the data since the number of sequencing errors increases with the read number/the assembly graph is too complex and cannot be resolved) coverage depths can have negative

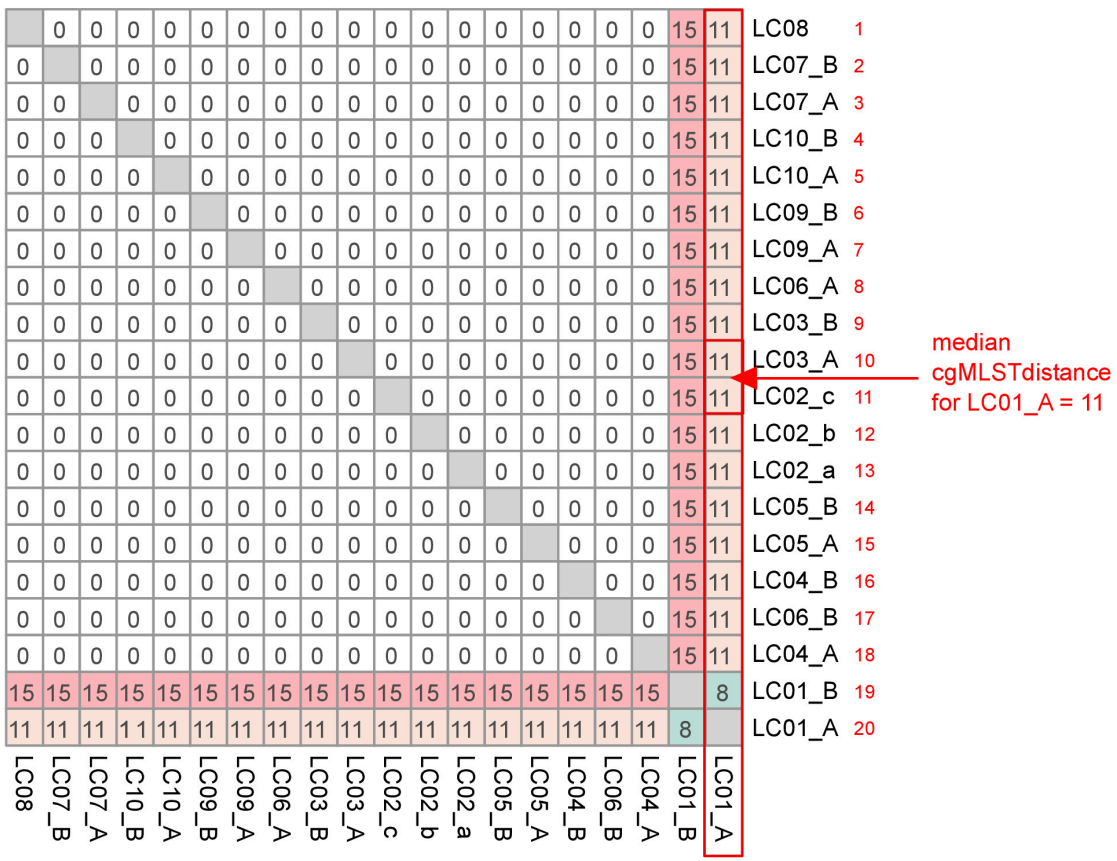


FIGURE 6 | The figure shows the cgMLST distance matrix for sample 19-RV1-P64-1. Laboratories (LC01-LC10) and respective sequencing runs (run A/run B) are identified. The red box, arrow and text demonstrate how the median cgMLST distance was determined.

effects on downstream analysis. For this reason, updated assembly algorithms, such as shovill, "down sample" data to a moderate coverage prior to assembly (e.g., shovill down samples to 100×). Indeed, we did not find a linear correlation between coverage depth and N50 (i.e., the very high coverage depths observed for some data sets had neither positive nor negative effects on assembly quality). Nevertheless, we recommend that sequencing laboratories pool DNA libraries by genome sizes prior to sequencing in order to produce sequencing data with consistent coverage depth for optimal downstream analysis. This has the additional benefit that smart pooling strategies decrease the sequencing costs, as a greater number of samples can be sequenced in one run.

We employed SNP calling for the detection of potential sequencing errors in the trimmed sequence reads, as well as for assessing the utility of a SNP calling approach for an integrated outbreak analysis with data from different sequencing platforms. Given that participants were provided with purified DNA samples, thereby eliminating the potential for the development of mutations during cultivation, any SNP potentially flags a sequencing error. Positively, we detected zero SNPs within the data sets. The fact that all data sets of sample 19-RV1-P64-5 differed in two SNPs from the respective PacBio reference, either points to a sequence error within the PacBio reference, or might indicate that the strain underwent mutations between the independent cultivations for short read and long read sequencing DNA isolation.

We further constructed *de novo* assemblies from the short read sequence data to assess the influence of variations in sequence data quality on assembly based downstream analyses. To eliminate assembler specific effects, we strove to construct all assemblies in an equal manner. Naturally, single-end Ion Torrent data requires different assembly algorithm, than those employed for paired-end Illumina data, which hampers a direct comparison.

Nevertheless, we found that all assemblies were overall very similar, with respect to assembly length, N50, GC and the number of CDSs, with a few notable exceptions. In particular, assemblies constructed from short read data of laboratories LC06 and LC08 (both using a MiSeq Illumina instrument) had much lower N50 values and a greater number of contigs, probably due to their use of the Nextera XT DNA Library Preparation Kit, which was shown to have a strong GC bias (Lan et al., 2015;

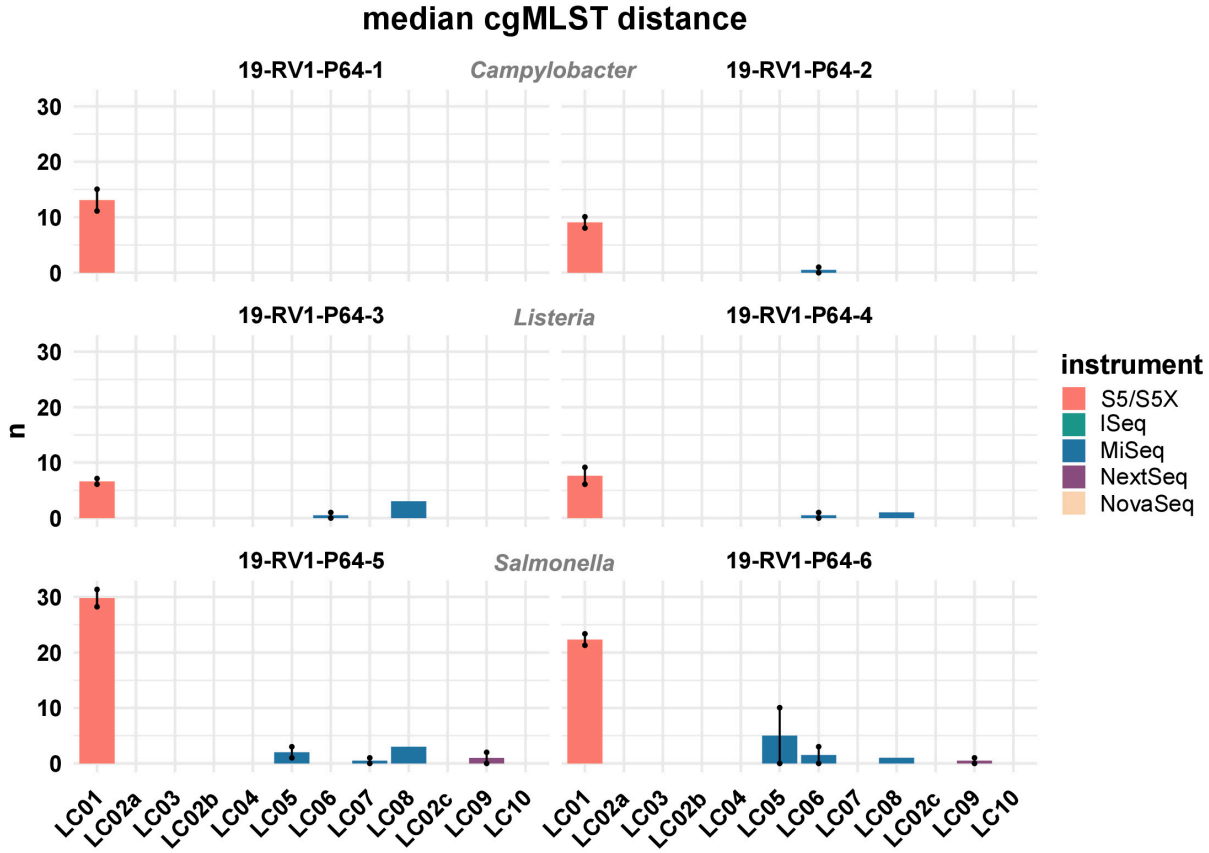


FIGURE 7 | The bar plot shows the mean median cgMLST distance grouped by laboratories and samples. Line-connected points indicate the variance between sequencing runs (run A/run B), with the exception of laboratories LC02 and LC08 (single sequencing run). Fill colors identify the sequencing instrument. The species of the samples is indicated.

Tyler et al., 2016; Grützke et al., 2019; Sato et al., 2019; Uelze et al., 2019; Browne et al., 2020) (also compare **Supplementary File S7**). This is a concern since a high number of contigs in a genome assembly may cause a fragmentation of genes at the contig borders, thereby affecting gene annotation and multilocus sequence typing. Furthermore we found that Ion Torrent assemblies differed from Illumina assemblies in length (slightly shorter), N50 (slightly lower), GC (slightly lower) and number of CDSs (slightly increased).

Complementary to SNP calling, we employed a cgMLST approach to compare genome assemblies in a simulated outbreak analysis. Noteworthy, cgMLST revealed a major distinction between Illumina and Ion Torrent data with assemblies constructed from Ion Torrent reads generally computing a much greater number of allele differences (Illumina: ∼ 0-3 allele differences, Ion Torrent: ∼10–30 allele differences). This increased number of allele differences is likely caused by frame shifts in the Ion Torrent assemblies which we verified exemplary for sample 19-RV1-P64-5 (**Supplementary File S8**). While the typical error type associated with Illumina reads are randomly distributed incorrect bases (substitution error) which do not cause frame shifts, Ion Torrent reads are prone to systematic insertions and deletions errors which lead to frame shifts in coding sequences (Buermans and den Dunnen, 2014; Escalona et al., 2016). Given that the cgMLST method employed in this study identifies coding frames based on their start and stop codons (as opposed to methods which implement a similarity based blastn search against a set of reference loci for allele identification), frame shifts will have a major effect on allele detection, thereby likely causing the observed increased number of allele differences. This is further supported by the low reproducibility of the Ion Torrent assemblies with up to 24 allele differences between two independent sequencing runs for the same sample.

The strong effect of frame shifts on allele differences can be prevented by removing alleles containing frame shifts during cgMLST calling. In chewBBACA, alleles with frame shifts can be filtered by excluding unusually large or small alleles during the size validation step (assuming that the frame shifts lead to a change in allele lengths). When cgMLST analysis was repeated with a strict allele length threshold for LC01 (Ion Torrent) and LC02a (Illumina iSeq), allele differences between Ion Torrent and Illumina assemblies could be reduced to close to zero (compare **Supplementary Files S9A**). However, allele sizes may vary naturally, making it difficult to determine a suitable allele length threshold. The biological variation of the allele length

depends furthermore on the details of the cgMLST scheme design and the species the scheme was created for. In addition, by applying a stringent allele length filter, a substantial number of alleles may be excluded from cgMLST calling, which in turn reduces the accuracy and discriminatory power of the results (compare **Supplementary File S9B**).

Similar to a frame shift filter/allele length filter in cgMLST, various filters exclude Ion Torrent typical indels, as well as heterozygous or low quality sites from SNP calling. Through SNP calling it was possible to correctly identify the clonality between data sets for the same sample (i.e., there were zero SNPs between the Illumina and the Ion Torrent data sets for all samples). SNP calling further has the advantage that no assembling step is required, for which currently no optimized assembly algorithm is available for Ion Torrent, thereby avoiding the introduction of assembly biases. Although we additionally assembled Illumina reads with SPAdes to increase the comparability to Ion Torrent assemblies (currently shovill is unable to assemble Ion Torrent reads), SPAdes remains inherently tailored for Illumina reads and cgMLST calling was not improved with all SPAdes assemblies.

Many surveillance platforms and programs perform cg/wgMLST for (pre-)clustering and SNP calling for a more detailed analysis (Uelze et al., 2020b). Based on our results, we recommend that users and developers be aware of the differences between Illumina and Ion Torrent data in combined outbreak studies. Stringent pre-filtering steps (such as frame shift filters/allele length during cgMLST calling and indel filtering during SNP calling) may be necessary to avoid erroneous clustering results, which otherwise could disrupt outbreak studies. However, further research is needed to investigate the trade off between stringent filtering and a decrease in resolution, as well as a loss of potentially significant biological information. Likely, this balance between a robust method and a preservation of the biological "truth" will need to be defined for each species separately, taking the particularities of each species (e.g., mutation frequencies) into account.

CONCLUSION

We found that seven of nine participants with Illumina sequencing instruments were able to obtain reproducible sequence data with consistent high quality. Two participants with Illumina instruments submitted data with lower quality, probably due to the use of a library preparation kit, which shows difficulty in sequencing low GC genome regions. Frame shifts in the Ion Torrent assemblies were evident during cgMLST calling, making SNP calling our preferred approach for an integrated outbreak analysis of Ion Torrent and Illumina data.

In the future, sequencing laboratories will continue to adapt and modify their laboratory protocols in order to optimize sequencing data quality, throughput and user-friendliness, while striving for the most cost and time-effective procedure. We welcome these efforts by innovative and thoughtful staff, which should not be unnecessarily hampered by overly rigid procedural protocols. Instead, a set of widely accepted, scientifically based and sensible minimal sequencing quality parameters, together with good standard practice protocols are urgently needed to ensure a consistent high quality of sequencing data for comparative data analysis.

Continuous interlaboratory testing, such as the one employed in this study and external PTs, will play an important role in ensuring that laboratories of the diverse public health setting adhere to these standards, while providing important feedback to participants on their competency level. Open or anonymous sharing of sequencing parameters allows an assessment of the utility of different sequencing approaches and helps to identify potential user issues. In the best case, interlaboratory studies promote knowledge and expertise sharing, enabling laboratories to adopt the sequencing procedures best suited for their unique setting, while simultaneously contributing to a standardization of the technology, which will greatly improve the efficacy of sequencing data for surveillance, outbreak analyses and comparative studies.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. Short read sequencing data for all data sets analysed in this study has been deposited in the NCBI database under the BioProject number PRJEB37768. PacBio assemblies used as references for SNP calling have been deposited under the BioProject number PRJNA638266.

AUTHOR CONTRIBUTIONS

LU, MBo, and BM designed the study. LU and MBö coordinated the interlaboratory study. LU and CD conducted the bioinformatic analysis and evaluation of the sequencing data quality. CD and ST developed the in-house bioinformatic pipelines used for analysis of the sequencing data. BM supervised the project. LU wrote the manuscript and created the figures. We thank MBö, EB, TH, SK, LM, KSt, KSz, and AW for constructive criticism of the manuscript. All authors read and approved the manuscript.

FUNDING

This work was supported by the German Federal Institute for Risk Assessment (BfR). The BfR has received financial support from the Federal Government for Laura Uelze on the basis of a resolution of the German Bundestag by the Federal Government and funded by the Ministry of Health within the framework of the project "Integrated genome-based surveillance of *Salmonella* (GenoSalmSurv)," decision ZMVI1-2518FSB709 of 26/11/2018.

ACKNOWLEDGMENTS

We thank all participating laboratories for their valuable contribution to the interlaboratory study. We are grateful for the continuous collaboration with the National Reference Laboratory for *Salmonella*, as well as the National Reference Laboratory for

Listeria monocytogenes and the National Reference Laboratory for *Campylobacter* who kindly provided us with the bacterial isolates and DNA samples. We also thank Adrian Prager (MRI), Sara Walter (LAVES), Sven Bikar and Tilmann Laufs (StarSEQ GmbH, Mainz), as well as Beatrice Baumann and Katharina Thomas (BfR) for their excellent support. This manuscript has been released as a pre-print at bioRxiv (Uelze et al., 2020a).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2020.573972/full#supplementary-material

FILE S1 | Information about the strains used for the interlaboratory study.

FILE S2 | The questionnaire template in German language.

FILE S3 | Summarized results from the questionnaire.

FILE S4 | Sequence quality parameters for all sample sets.

FILE S5 | CgMLST distance allele matrices for all samples.

FILE S6 | SNP distance allele matrices for all samples.

FILE S7 | Figures show the global GC-bias across the whole genome calculated using Benjamini's method (Benjamini and Speed, 2012) with the computeGCBias function of the deepTools package (Ramírez et al., 2016) for all sample sets. The function counts the number of reads per GC fraction and compares them to the expected GC profile, calculated by counting the number of DNA fragments per GC fraction in a reference genome. In an ideal experiment, the observed GC profile would match the expected profile, producing a flat line at 0. The fluctuations to both ends of the *x*-axis are due to the fact that only a small number of genome regions have extreme GC fractions.

FILE S8 | The table lists those cgMLST loci that differ between the Ion Torrent assembly (LC01) and the Illumina assemblies (LC02-LC10) for sample 19-RV1-P64 (run A). Furthermore the table displays alignment statistics (number of insertions, number of deletions, presence of indels, number of mismatches) for the Illumina allele sequences, which were mapped with Minimap2 (Li, 2018) against the Ion Torrent assembly (LC01), as well as against the Illumina assemblies (LC02-LC10).

FILE S9 | The figure shows the effect of a varying allele length threshold, based on comparing cgMLST results of LC02a (Illumina iSeq) to LC01 (Ion Torrent) for all samples and all sequencing runs. Shown is the number of remaining allele distances of the IonTorrent sample to the Illumina samples **(A)** and the number of removed loci **(B)** when applying an allele length filtering. The threshold is defined as the ratio of observed allele length of the sample compared to the median allele length for a given locus in the scheme. Thereby, the initially substantial allele difference can be largely reduced. However, the nearly complete elimination of the allele difference is only possible by a substantial reduction of the effective number of loci. The schemes contain 678, 1701 and 3255 loci for *Campylobacter*, *Listeria* and *Salmonella*, respectively.

REFERENCES

Ashton, P. M., Nair, S., Peters, T. M., Bale, J. A., Powell, D. G., Painset, A., et al. (2016). Identification of *Salmonella* for public health surveillance using whole genome sequencing. *PeerJ* 4:e1752. doi: 10.7717/peerj.1752

Benjamini, Y., and Speed, T. P. (2012). Summarizing and correcting the GC content bias in high-throughput sequencing. *Nucleic Acids Res.* 40:e72. doi: 10.1093/nar/gks001

Borowiak, M., Fischer, J., Baumann, B., Hammerl, J. A., Szabo, I., and Malorny, B. (2018). Complete genome sequence of a VIM-1-producing *Salmonella enterica* subsp. *enterica* serovar Infantis isolate derived from minced pork meat. *Genome Announc.* 6:e00327-18. doi: 10.1128/genomeA.00327-18

Browne, P. D., Nielsen, T. K., Kot, W., Aggerholm, A., Gilbert, M. T. P., Puetz, L., et al. (2020). GC bias affects genomic and metagenomic reconstructions, underrepresenting GC-poor organisms. *GigaScience* 9:giaa008. doi: 10.1093/gigascience/giaa008

Buermans, H. P. J., and den Dunnen, J. T. (2014). Next generation sequencing technology: advances and applications. *Biochim. Biophys. Acta* 1842, 1932–1941. doi: 10.1016/j.bbadis.2014.06.015

Carattoli, A., Zankari, E., García-Fernández, A., Voldby Larsen, M., Lund, O., Villa, L., et al. (2014). *In silico* detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing. *Antimicrob. Agents Chemother.* 58, 3895–3903. doi: 10.1128/AAC.02412-14

Chen, S., Zhou, Y., Chen, Y., and Gu, J. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34, i884–i890. doi: 10.1093/bioinformatics/bty560

Chun, J., Oren, A., Ventosa, A., Christensen, H., Arahal, D. R., da Costa, M. S., et al. (2018). Proposed minimal standards for the use of genome data for the taxonomy of prokaryotes. *Int. J. Syst. Evol. Microbiol.* 68, 461–466. doi: 10.1099/ijsem.0.002516

Endrullat, C., Glökler, J., Franke, P., and Frohme, M. (2016). Standardization and quality management in next-generation sequencing. *Appl. Transl. Genomics* 10, 2–9. doi: 10.1016/j.atg.2016.06.001

Escalona, M., Rocha, S., and Posada, D. (2016). A comparison of tools for the simulation of genomic next-generation sequencing data. *Nat. Rev. Genet.* 17, 459–469. doi: 10.1038/nrg.2016.57

Fox, E. J., Reid-Bayliss, K. S., Emond, M. J., and Loeb, L. A. (2014). Accuracy of next generation sequencing platforms. *J. Gener. Seq. Appl.* 1:1000106. doi: 10.4172/jngsa.1000106

Garrison, E., and Marth, G. (2012). Haplotype-based variant detection from short-read sequencing. *ArXiv* [Preprint]. Available online at: http://arxiv.org/abs/1207.3907 (accessed July 30, 2020).

Grützke, J., Malorny, B., Hammerl, J. A., Busch, A., Tausch, S. H., Tomaso, H., et al. (2019). Fishing in the soup – pathogen detection in food safety using metabarcoding and metagenomic sequencing. *Front. Microbiol.* 10:1805. doi: 10.3389/fmicb.2019.01805

Hyatt, D., Chen, G.-L., LoCascio, P. F., Land, M. L., Larimer, F. W., and Hauser, L. J. (2010). Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 11:119. doi: 10.1186/1471-2105-11-119

Kwong, J. C., Mccallum, N., Sintchenko, V., and Howden, B. P. (2015). Whole genome sequencing in clinical and public health microbiology. *Pathology* 47, 199–210. doi: 10.1097/PAT.0000000000000235

Lan, J. H., Yin, Y., Reed, E. F., Moua, K., Thomas, K., and Zhang, Q. (2015). Impact of three Illumina library construction methods on GC bias and HLA genotype calling. *Hum. Immunol.* 76, 166–175. doi: 10.1016/j.humimm.2014.12.016

Li, H. (2018). Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* 34, 3094–3100. doi: 10.1093/bioinformatics/bty191

Mellmann, A., Andersen, P. S., Bletz, S., Friedrich, A. W., Kohl, T. A., Lilje, B., et al. (2017). High interlaboratory reproducibility and accuracy of next-generation-sequencing-based bacterial genotyping in a ring trial. *J. Clin. Microbiol.* 55, 908–913. doi: 10.1128/JCM.02242-16

Moran-Gilad, J., Pedersen, S. K., Wolfgang, W. J., Pettengill, J., Strain, E., Hendriksen, R. S., et al. (2015). Proficiency testing for bacterial whole genome sequencing: an end-user survey of current capabilities, requirements and priorities. *BMC Infect. Dis.* 15:174. doi: 10.1186/s12879-015-0902-3

Nurk, S., Bankevich, A., Antipov, D., Gurevich, A., Korobeynikov, A., Lapidus, A., et al. (2013). "Assembling genomes and mini-metagenomes from highly chimeric reads," in *Research in Computational Molecular Biology*, eds M. Deng, R. Jiang, F. Sun, and X. Zhang (Berlin: Springer), 158–170. doi: 10.1007/978-3-642-37195-0_13

Quail, M., Smith, M. E., Coupland, P., Otto, T. D., Harris, S. R., Connor, T. R., et al. (2012). A tale of three next generation sequencing platforms: comparison of

Ion torrent, pacific biosciences and illumina MiSeq sequencers. *BMC Genomics* 13:341. doi: 10.1186/1471-2164-13-341

Ramírez, F., Ryan, D. P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A. S., et al. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* 44, W160–W165. doi: 10.1093/nar/gkw257

Ronholm, J., Nasheri, N., Petronella, N., and Pagotto, F. (2016). Navigating microbiological food safety in the era of whole-genome sequencing. *Clin. Microbiol. Rev.* 29, 837–857. doi: 10.1128/CMR.00056-16

Ruppitsch, W., Pietzka, A., Prior, K., Bletz, S., Fernandez, H. L., Allerberger, F., et al. (2015). Defining and evaluating a core genome multilocus sequence typing scheme for whole-genome sequence-based typing of *Listeria monocytogenes*. *J. Clin. Microbiol.* 53, 2869–2876. doi: 10.1128/JCM.01193-15

Salipante, S. J., Kawashima, T., Rosenthal, C., Hoogestraat, D. R., Cummings, L. A., Sengupta, D. J., et al. (2014). Performance comparison of Illumina and Ion Torrent next-generation sequencing platforms for 16S rRNA-based bacterial community profiling. *Appl. Environ. Microbiol.* 80, 7583–7591. doi: 10.1128/AEM.02206-14

Sato, M. P., Ogura, Y., Nakamura, K., Nishida, R., Gotoh, Y., Hayashi, M., et al. (2019). Comparison of the sequencing bias of currently available library preparation kits for Illumina sequencing of bacterial genomes and metagenomes. *DNA Res.* 26, 391–398. doi: 10.1093/dnares/dsz017

Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30, 2068–2069. doi: 10.1093/bioinformatics/btu153

Silva, M., Machado, M. P., Silva, D. N., Rossi, M., Moran-Gilad, J., Santos, S., et al. (2018). chewBBACA: a complete suite for gene-by-gene schema creation and strain identification. *Microb. Genomics* 4:e000166. doi: 10.1099/mgen.0.000166

Timme, R. E., Rand, H., Sanchez Leon, M., Hoffmann, M., Strain, E., Allard, M., et al. (2018). GenomeTrakr proficiency testing for foodborne pathogen surveillance: an exercise from 2015. *Microb. Genomics* 4:e000185. doi: 10.1099/mgen.0.000185

Timme, R. E., Sanchez Leon, M., and Allard, M. W. (2019). "Utilizing the Public GenomeTrakr Database for Foodborne Pathogen Traceback," in *Foodborne Bacterial Pathogens Methods in Molecular Biology*, ed. A. Bridier (New York, NY: Springer), 201–212. doi: 10.1007/978-1-4939-9000-9_17

Tolar, B., Joseph, L. A., Schroeder, M. N., Stroika, S., Ribot, E. M., Hise, K. B., et al. (2019). An overview of PulseNet USA databases. *Foodborne Pathog. Dis.* 16, 457–462. doi: 10.1089/fpd.2019.2637

Tyler, A. D., Christianson, S., Knox, N. C., Mabon, P., Wolfe, J., Van Domselaar, G., et al. (2016). Comparison of sample preparation methods used for the next-generation sequencing of *Mycobacterium tuberculosis*. *PLoS One* 11:e0148676. doi: 10.1371/journal.pone.0148676

Uelze, L., Borowiak, M., Brinks, E., Deneke, C., Stingl, K., Kleta, S., et al. (2020a). German-wide interlaboratory study compares consistency, accuracy and reproducibility of whole-genome short read sequencing. *bioRxiv* [Preprint]. doi: 10.1101/2020.04.22.054759

Uelze, L., Borowiak, M., Deneke, C., Szabó, I., Fischer, J., Tausch, S. H., et al. (2019). Performance and accuracy of four open-source tools for *in silico* serotyping of *Salmonella* spp. based on whole-genome short-read sequencing data. *Appl. Environ. Microbiol.* 86:e02265-19. doi: 10.1128/AEM.02265-19

Uelze, L., Grützke, J., Borowiak, M., Hammerl, J. A., Juraschek, K., Deneke, C., et al. (2020b). Typing methods based on whole genome sequencing data. *One Health Outlook* 2:3. doi: 10.1186/s42522-020-0010-1

Yoshida, C. E., Kruczkiewicz, P., Laing, C. R., Lingohr, E. J., Gannon, V. P. J., Nash, J. H. E., et al. (2016). The *Salmonella in silico* typing resource (SISTR): an open web-accessible tool for rapidly typing and subtyping draft *Salmonella* genome assemblies. *PLoS One* 11:e0147101. doi: 10.1371/journal.pone.0147101

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Uelze, Borowiak, Bönn, Brinks, Deneke, Hankeln, Kleta, Murr, Stingl, Szabo, Tausch, Wöhlke and Malorny. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 15 September 2020
doi: 10.3389/fmicb.2020.516445

Application of an Optimized Direct Lysis Method for Viral RNA Extraction Linking Contaminated Dates to Infection With Hepatitis A Virus

Sheikh Md Rajiuddin[1], Sofie Elisabeth Midgley[2], Tenna Jensen[3], Luise Müller[4] and Anna Charlotte Schultz[1†]*

[1] Division of Microbiology and Production, National Food Institute, Technical University of Denmark, Kongens Lyngby, Denmark, [2] Department of Virus and Specialist Microbiological Diagnostics, Statens Serum Institut, Copenhagen, Denmark, [3] Division for Food and Feed Safety, Danish Veterinary and Food Administration, Glostrup, Denmark, [4] Department of Infectious Disease Epidemiology and Prevention, Statens Serum Institut, Copenhagen, Denmark

OPEN ACCESS

Edited by:
Dario De Medici, National Institute of Health (ISS), Italy

Reviewed by:
Jesus L. Romalde, University of Santiago de Compostela, Spain
Sandra Martin-Latil, Agence Nationale de Sécurité Sanitaire de l'Alimentation, de l'Environnement et du Travail (ANSES), France

***Correspondence:**
Anna Charlotte Schultz
acsc@outlook.dk

†ORCID:
Anna Charlotte Schultz
orcid.org/0000-0003-1484-9120

Specialty section:
This article was submitted to Food Microbiology, a section of the journal Frontiers in Microbiology

Received: *30 November 2019*
Accepted: *24 August 2020*
Published: *15 September 2020*

Citation:
Rajiuddin SM, Midgley SE, Jensen T, Müller L and Schultz AC (2020) Application of an Optimized Direct Lysis Method for Viral RNA Extraction Linking Contaminated Dates to Infection With Hepatitis A Virus. Front. Microbiol. 11:516445. doi: 10.3389/fmicb.2020.516445

Consumption of dates has not been considered a common risk of hepatitis A virus (HAV) infection. In January 2018, an outbreak of hepatitis was identified with cases resident in all regions of Denmark. All the detected strains belonged to HAV genotype 3A. Epidemiological investigations through patients' interviews, case-control and trace-back studies pointed toward different batches of dates from a single producer as the vehicle of infection. Boxes of dates from suspected batches were collected from homes of patients and healthy families and analyzed using a recently reported optimized direct lysis method, consisting of simultaneous viral RNA elution and extraction from dates followed by purification of the nucleic acids. Extracts were analyzed for HAV and norovirus (NoV) RNA using RT-qPCR, while detected HAV were genotyped by Sanger sequencing. Among 20 nucleic acid extracts representing eight batches of dates, RNA of HAV (9.3×10^2 genome copies/g) and NoV genogroup (G)II (trace amounts) were detected in one batch, while NoV GII RNA (trace amounts) was detected in another. Average extraction efficiency of spiked process control murine norovirus was $20 \pm 13\%$ and the inhibitions of RT-qPCR detection of NoV GI, NoV GII, and HAV were 31 ± 34, 9 ± 9, and $3 \pm 7\%$, respectively. The HAV genome detected in the dates matched by sequence 100% to the HAV genotype 3A detected in stool samples from cases implicated in the outbreak. This confirmed, to our knowledge, for the first time a sequence link between HAV infection and consumption of contaminated dates, suggesting dates to be an important vehicle of HAV transmission.

Keywords: PCR, sequencing, outbreak, jaundice, foodborne, HAV

INTRODUCTION

Hepatitis A virus (HAV) is one of the causative agents of viral jaundice and transmits primarily via the fecal-oral route through person-to-person contact or through contaminated food and water (Hollinger and Martin, 2013). Due to the ability to resist drying, freezing, food preservatives and acidification, HAV is highly persistent in the environment and has frequently caused foodborne

disease outbreaks among susceptible populations worldwide (World Health Organization [WHO], 2017).

In recent years, consumption of HAV contaminated fruit imported from endemic areas has been identified as the source of several large and sometimes multistate outbreaks of HAV in Europe (EFSA, 2014; Severi et al., 2015). For example, HAV contaminated frozen berries caused 1,589 cases including two deaths in 13 EU/EEA countries during 2013–2014 (Severi et al., 2015), 71 cases in four Nordic countries including Denmark during 2012–2013 (Lassen et al., 2013) and 14 cases in the Netherlands in 2017 (Mollers et al., 2018). Consumption of semidried tomatoes has also been linked to outbreaks of HAV in Australia in 2009, and in the Netherlands and France in 2010 (Petrignani et al., 2010; Gallot et al., 2011; Donnan et al., 2012).

Attempts to detect and characterize HAV in suspected food products, identified by case control studies, to conduct a genetic comparison between HAV strains identified in clinical and food samples have often been unsuccessful, although it is essential to verify the origin of the foodborne disease (Petrignani et al., 2010; Gallot et al., 2011; Collier et al., 2014). One reason for this can be the un-availability of relevant food samples for laboratory analyses due to long lag time (2–6 weeks) between infection and onset of symptoms (EFSA, 2014). Other reasons may include uneven distribution of virus in suspected food, and low titer or purity of extracted viral RNA for un-inhibited HAV detection by reverse transcription (RT)-qPCR and characterization by sequencing (Schrader et al., 2012; EFSA, 2014).

An ISO standard for RT-qPCR detection of NoV and HAV in food matrices, including berries, ISO 15216-1, 2017 was recently validated (ISO, 2017). The method contains a protocol for virus and nucleic acid extraction based on elution of virus particles from food and capsid disruption using a chaotropic agent followed by nucleic acid adsorption to silica particles. The protocol includes quality control to identify false-negatives and prevent underestimation of viral load by evaluating the efficiency of viral RNA extraction using a virus process control, and amplification efficiency of target viral RNA by addition of an RNA control sequence to the RT-qPCR reaction. Altogether, the method is laborious and complicated, and needs to be simplified and optimized to increase viral recovery and purity of extracts (Bosch et al., 2018). In addition, the ISO standard is only validated for detection of NoV and HAV in the matrices described (soft fruit, leafy greens, stem and bulb vegetables, bottled water, mussels and oysters), excluding other types of potential vehicles of virus transmission, which limits the scope of the standard.

Perrin et al. (2015) proposed a direct viral lysis method from soft-fruit by immersion in lysis buffer, which enabled efficient detection of NoV surrogates on spiked raspberries, although it was less efficient in analyzing strawberries for the detection of spiked NoV (Bartsch et al., 2016). A slightly modified version in which inhibition could be overcome by filtering the RNA extract before molecular detection, confirmed the applicability of the method principle on samples of different berry types spiked with model viruses (Sun et al., 2019). However, the applicability of the direct lysis methods to detect foodborne viruses in naturally contaminated food samples remained to be demonstrated (Bosch et al., 2018).

To address the analytical challenge, we recently reported a further step toward improvement of the direct lysis method to allow for higher recovery of the process controls (mengovirus and murine norovirus, MNV) and less inhibited RT-PCR detection of the target viruses (NoV and HAV) in extracts from 13 different types of foods; fruits (including dates), vegetables, a compound food (seaweed salad) and seeds/nuts. The performance of the optimized direct lysis method was compared against a modified version of the ISO 15216-1 standard, and demonstrated successful detection of NoV in naturally contaminated strawberries and seaweed salad (Rajiuddin et al., 2020). However, the method's broadness in detecting natural contamination remains to be shown for other viral pathogens and food matrices (Rajiuddin et al., 2020).

In January 2018, a hepatitis A outbreak was identified by RT-qPCR detection of HAV RNA in stool samples from 19 of 31 cases showing clinical symptoms of hepatitis infection from all regions of Denmark. Sanger sequencing and phylogenetic analysis grouped all the detected strains to HAV genotype 3A. Epidemiological investigations through patient interviews as well as case-control and trace-back investigations pointed toward different batches of dates from a single producer as the common route of exposure and thus the suspected vehicle of infection (Müller et al., 2018).

This study aimed to analyze samples of dates suspected to be implicated in the identified hepatitis A outbreak using a modified direct lysis method for the extraction of HAV RNA. In case of successful HAV detection, the aim was further to quantify the contamination level and to sequence the strain(s) detected in the date sample(s) in order to establish a match to the strains detected in the patient samples. Finally, the study aimed to explore the presence of NoV in the date samples, to evaluate a possible fecal contamination, in the case of unsuccessful HAV detection.

MATERIALS AND METHODS

Date Samples

Ten boxes (box 1–10) of eight different batches (batch 1–8) of dates from the same brand and producer, were collected from the homes of four patients and from the homes of five families that contacted the Danish Veterinary and Food Administration (DVFA) offering the remaining dates for analyses in the wake of the preventive recall of the product. The dates were sent to the National Food Institute, Technical University of Denmark (DTU), for analysis for the presence of HAV and NoV.

Elution, Extraction and Purification of Viral RNA From Date Samples

Viral RNA was extracted from 2 × 25–35 g (1–2 pieces) samples of each box of dates. To quality assess the extraction performance and control for false-negatives, the dates were spiked with MNV (10^4 pfu or 10^5 RT-qPCR units), propagated and titrated by plaque assay or end-point RT-qPCR, respectively

(Rawsthorne et al., 2009; Hwang et al., 2014), and left to dry at room temperature for 30 min on a sterile bench.

Initially, box 1–2 were analyzed according to a viral RNA extraction method (Rajiuddin et al., 2020) structurally similar to the ISO 15216-1 (ISO, 2017) standard.

Subsequently, dates from box 1–10 were analyzed using a modified direct lysis method (Rajiuddin et al., 2020). Briefly, viral elution and simultaneous degradation of pectin and other potential RT-qPCR inhibitory substances, were conducted by inverting each sample 2–3 times in lysis buffer (10 ml) (NucliSENS®, bioMerieux, 280134) with the addition of pectinase (1 ml ∼3800 U) (Sigma, p2611 > 3,800 U/ml) and Plant RNA Isolation aid (400 µl) (Invitrogen, AM9690) followed by 10 min shaking at 200 rpm at room temperature. After incubation, the liquid buffer was separated from the solid date tissue by pipetting and centrifuged at 10,000 × *g* for 10 min at room temperature to pellet debris and small fruit particles. The supernatant (∼10.5–11.5 ml) was collected by pipetting and immediately used for nucleic acid extraction. Nucleic acids from all viral suspensions obtained by either method described above were extracted from the supernatant using the NucliSENS® miniMAG® system (BioMérieux, Herlev, Denmark) according to the manufactures instruction, except for using 140 µl of magnetic silica particles (BioMérieux, 280133) to catch released nucleic acid, and 100 µl of washing buffer three for the elution of purified RNA.

For the extracts obtained by the direct lysis method, a further purification of the nucleic acids was conducted using an OneStep PCR Inhibitor Removal Kit (Zymo Research, D6030) according to manufacturer's instruction. Briefly, the *Prep-Solution* (600 µl) was added into a Zymo-Spin™ III-HRC column placed in a *collection tube* beforehand and centrifuged at 8000 × *g* for 3 min. The spin column was transferred to a 1.5 ml microcentrifuge tube and 100 µl of previously extracted nucleic acid extracts from the date samples was transferred to the prepared column and centrifuged at 8000 × *g* for 1 min. The filtered RNA sample was preserved at 4°C for maximum 24 h or at −80°C for longer before being used for RT-qPCR analysis.

Detection and Quantification of Viral RNA

Each extract was analyzed in duplicate for quantitative presence of HAV, NoV GI, and GII, as well as of MNV by RT-qPCR in a 96-well plate (MicroAmp Fast Optical 96-Well Reaction Plate 0.1 ml, Applier Biosystems, 4306311) on an ABI StepOnePlus Real-Time PCR machine (Applied Biosystems, Naerum, Denmark).

RT-qPCR reactions were carried out in a total of 25.0 µl mixture containing 5.0 µl of nucleic acid extract and 20.0 µl of master mix, including 1 × RNA Ultrasense One-Step Quantitative RT-PCR System (Thermo Fisher Scientific, Hvidovre, Denmark), and primers and TaqMan probes (DNA Technology AS, Risskov, Denmark) for NoV GI and NoV GII (Le Guyader et al., 2009), HAV (Costafreda et al., 2006) or MNV (Rawsthorne et al., 2009), and reaction conditions described by Le Guyader et al. (2009).

The theoretical limit of detection (tLOD) of the method was calculated to 20 genome copy (GC)/25 g date or 0.8 GC/g date. This was done by taking into account the volume changes and assuming that at least one GC of viral target (HAV or NoV) must be present in a positive RT-qPCR reaction containing 5 µl tested of total 100 µl extract obtained by processing a sample of 25–35 g date.

Quantification was performed using standard curves generated from 10-fold dilution series of artificially constructed dsDNA of HAV HM-175, NoV GI.1b, and NoV GII.4 (kindly provided by Dr. Lowther, Cefas, United Kingdom), produced and applied according to ISO 15216-1 (ISO, 2017), and extracted RNA of MNV.

The extraction efficiencies of the date samples were calculated using the formula $100e^{-0.6978\,\Delta Ct}$ as previously described by Le Guyader et al. (2009), where ΔCt is the difference in Ct values of the MNV RNA recovered from spiked samples and from nuclease free water.

To assess the purity of the extracted nucleic acids, inhibition controls for the detection of target virus were performed on all extracts according to ISO 15216-1 (ISO, 2017), by adding 10^4 genome copies (GC) of NoV GI, NoV GII, and HAV transcripts (curtesy of Dr. Lowther CEFAS, United Kingdom) to separate wells containing 5.0 µl-volumes of each nucleic acid extract and to nuclease-free water.

The percentage of RT-qPCR inhibition was calculated as quotient of the GC amounts recovered of spiked RNA transcripts from the extracts (GC extract) and nuclease-free water (GC control) using the following formula: Inhibition (%) of target sequence = [1−(GC control/GC extract)] × 100. In cases of calculated negative inhibition of target sequence, lower inhibition in sample extract than in control, the inhibition was indicated as 0.0%.

Characterization of Detected HAV Sequences

HAV genotyping by Sanger sequencing was carried out as described previously (Lassen et al., 2013). Sequence assembly and genotyping was performed using BioNumerics v7.6 (Applied Maths) and phylogenetic analysis was carried out using MEGA 6.

RESULTS

Detection and Quantification of Viral RNA in Date Samples

Twenty nucleic acid extracts were obtained using the direct lysis method. HAV ($9.3 \times 10^2 \pm 6.5 \times 10^2$ GC/g) and NoV GII (estimated 8.1 ± 7.7 GC/g) were detected in both extracts of box 10, originating from batch 4, while NoV GII RNA (estimated 2.4 ± 0.5 GC/g) was detected in one of two extracts of box 6, originating from batch 1 (**Table 1**). Neither HAV nor NoV GII was detected in dates from the remaining eight boxes extracted with the direct lysis method (box 1–5 and 7–9) and NoV GI could not be detected in any of the samples.

The mean RNA extraction efficiency of spiked MNV was 20 ± 13% (range 10–37%) and the RT-qPCR inhibition of target

TABLE 1 | Detection of target viral RNA, extraction efficiencies and RT-qPCR inhibition in extracts from dates attained using the optimized direct lysis method and the modified ISO 15216-1 standard (ISO, 2017).

Box[1] ID	Batch ID	Best before date	Optimized direct lysis method[2]							Modified ISO 15216-1 method[2]					
			Weight of sample tested (g)	Extraction efficiency of MNV (%)		RT-qPCR inhibition[3] (%)			Detection of viral RNA[4]			Weight of sample tested (g)	Extraction efficiency of MNV (%)		RT-qPCR inhibition (%)		
				Undiluted	10x diluted	NoV GI	NoV GII	HAV	NoV GII	HAV			Undiluted	10x diluted	NoV GI	NoV GII	HAV
1	1	30.09.2018	25.4 ± 1.9	14.1 ± 4.9	32.0 ± 8.5	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	−−	−−		25.1 ± 0.1	0.5 ± 0.3	0.6 ± 0.6	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
2	2	30.03.2018	25.1 ± 7.0	26.2 ± 27.7	34.0 ± 37.9	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	−−	−−		25.3 ± 0.2	0.9 ± 0.3	1.0 ± 0.3	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
3	3	10.06.2018	25.5 ± 0.6	13.0 ± 4.0	17.4 ± 5.4	40.5 ± 33.0	5.2 ± 3.5	0.0 ± 0.0	−−	−−							
4	4	30.06.2018	21.9 ± 0.6	17.1 ± 3.2	35.5 ± 16.5	43.4 ± 45.5	11.0 ± 15.1	1.4 ± 2.0	−−	−−							
5	5	30.06.2018	23.7 ± 0.2	34.5 ± 8.3	55.9 ± 18.1	26.7 ± 11.0	5.6 ± 7.9	8.2 ± 9.5	−−	−−							
6	1	30.09.2018	21.3 ± 0.6	37.2 ± 9.6	39.2 ± 7.6	0.0 ± 0.0	3.6 ± 5.0	5.9 ± 8.3	+−	−−							
7	6	30.06.2018	25.2 ± 0.5	9.7 ± 3.0	42.7 ± 12.1	79.7 ± 14.2	21.5 ± 3.1	15.0 ± 21.3	−−	−−							
8	7	10.05.2018	33.0 ± 0.1	13.8 ± 9.6	40.8 ± 13.7	73.8 ± 18.6	16.3 ± 6.9	1.8 ± 2.5	−−	−−							
9	8	31.12.2017	24.2 ± 3.4	22.7 ± 2.3	27.2 ± 6.6	66.2 ± 16.8	15.4 ± 5.3	1.2 ± 1.7	−−	−−							
10	4	30.09.2018	33.1 ± 3.6	10.8 ± 6.5	7.7 ± 5.3	3.0 ± 5.2	8.4 ± 11.9	0.0 ± 0.0	++	++							
Mean values				19.9 ± 13.3	33.3 ± 19.2	30.6 ± 34.2	8.7 ± 8.9	3.1 ± 7.2					0.7 ± 0.4	0.8 ± 0.6	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0

Data are presented as the mean ± standard deviations (Mean ± SD). [1]All boxes of dates had been opened at the place of consumers, except for box 5. Original content of dates was 600 g's for box 3, 5, and 7 and 400 g's for the remaining boxes. [2]Two extractions were performed from each box of dates, and the detection of target viruses as well as the extraction efficiency and RT-qPCR inhibition were each determined by separate two RT-qPCR reactions per extraction. [3]In cases of calculated negative inhibition of target sequence, lower inhibition in sample extract than in control, the inhibition was indicated as 0.0%". [4]Each symbol represents positive (+) or negative (−) detection of viral RNA in one of two extract. NoV GI was not detected in any extractions using the optimized direct lysis method, neither was HAV, NoV GI, or NoV GII using the modified ISO 15216-1 method.

sequences, NoV GI, NoV GII, and HAV, were 31 ± 34% (range 0–80), 9 ± 9% (range 0–21), and 3 ± 7% (range 0–15), respectively (**Table 1**). Except for one of two extracts from each box 4 and 7–9 displaying 75–90% NoV GI RT-qPCR inhibition, all extraction efficiencies and percentages of inhibition in undiluted extracts obtained using the direct lysis method complied with the quality criteria of extraction efficiency (≥1%) and RT-qPCR inhibition (≤75%) described in the ISO 15216-1 (ISO, 2017), see **Table 1**.

Neither HAV nor NoV could be detected in four nucleic acid extracts processed from box 1 and 2 using the modified ISO 15216-1 method, resulting in mean 0.7 ± 0.4% MNV RNA extraction efficiency and 0 ± 0% RT-qPCR inhibitions for all target viruses (NoV GI, NoV GII, and HAV).

Sequence Analysis

A sequence of 1228 nucleotides covering the entire VP1 gene was obtained for strain DTU69S1 (deposited in GenBank, accession number MT222962) detected in the dates. Initial BLAST analysis against the SSI internal HAV database, and the NCBI GenBank nucleotide database, revealed it to be of genotype 3A, most closely related to outbreak strains and strains originating in Afghanistan and Pakistan. Phylogenetic analysis of 19 Danish outbreak strains and one strain from a patient from another country, and DTU69S1 showed a 100% nucleotide match between DTU69S1 and both the Danish case from whom the HAV positive date box was collected and another patient known to have consumed dates from a different batch (Müller et al., 2018).

DISCUSSION

Despite frequent outbreaks of hepatitis A suspected to be following consumption of contaminated foods, it has rarely been confirmed by detection of HAV in the implicated food (Petrignani et al., 2010; Gallot et al., 2011; Collier et al., 2014). Food products incriminated in such outbreaks are often imported from HAV endemic areas to European countries with susceptible populations at high risk of developing severe disease symptoms following consumption (World Health Organization [WHO], 2010; Lassen et al., 2013).

For the first time a HAV strain from samples of a batch of dates implicated in a foodborne HAV outbreak could be detected. The strain was quantified, characterized and genetically linked with a 100% match to the HAV sequence isolated from stool samples of one of the Danish cases included in the outbreak and from one other patient. The detection of HAV in nucleic acid extracts from date samples processed using the optimized and simple direct lysis method demonstrates the method's applicability for detection of HAV in naturally contaminated dates.

However, since extractions of the positive dates were not conducted using the modified ISO 15216-1 method (ISO, 2017), we cannot exclude that this could also have resulted in a similar outcome. This widely used EU standard method for detection of viruses on soft fruit can be inefficient in recovering RT-qPCR detectable viral RNA from different types of soft fruits (Bartsch

et al., 2016; Rajiuddin et al., 2020) including dates (Boxman et al., 2012). The reason is that food matrices may challenge efficient viral elution from the matrix surfaces and contain organic and inorganic substances that interfere with RT-qPCR detection of target sequences (Schrader et al., 2012; Perrin et al., 2015). Although with few samples, this study indicated a poor extraction efficiency (average <1%) of viral RNA from dates using the modified ISO 15216-1 method (ISO, 2017) compared to the direct lysis method (average 20%), without compromising on the degree of inhibition (average 0%) obtained for the four samples analyzed using both methods.

Previous studies using alternative variations of a direct lysis methods for extraction of spiked NoV (Bartsch et al., 2016) or model viruses have reported challenges with RT-qPCR inhibition, which could be overcome by dilution or filtration of extracts before RT-qPCR detection (Sun et al., 2019). In our study an inhibition >75% was observed for only NoV GI RT-qPCR detection in four sub-extractions from the dates, which thus failed the quality criteria described in the ISO 15216-1 (ISO, 2017), although extracts from the remaining subsamples and a 10-fold dilution of the inhibited subsamples did comply with this criteria. It is likely that the use of digital PCR (Coudray-Meunier et al., 2015) as genome detection and quantification method would overcome potential challenge of inhibition in sample extracts. However, this was not explored in this study.

In a Dutch surveillance study of dates including 185 samples, two batches of dates were tested positive for HAV and genotyped to 1A using a simplified method containing elution and RNA extraction compared to the ISO 15216-1 method (Boxman et al., 2012). However, despite suspicion of HAV illness among date consumers in the population, it was not possible to establish a link between the batches of dates tested and the infected patients.

Due to the estimated low contamination level in foods and the low efficiency of common methods to recover viral RNA, the characterization by sequencing of detected viral RNA in food materials may often be more difficult than in human stool or blood samples (EFSA, 2014). Previously, we reported that the optimized direct lysis method applied in this study was significantly more efficient in recovering RNA of spiked model viruses on a selection of types of fruits compared to the ISO standard (Rajiuddin et al., 2020). In this study, the optimized direct lysis method proved fit for purpose to extract HAV RNA that allowed RT-qPCR detection and quantification as well as sequence characterization from samples of naturally contaminated dates suspected to have caused a disease outbreak. The relatively high viral recovery and purity of extract, likely attributed to the fewer steps and the use of the OneStep PCR Inhibitor Removal Kit (McKee et al., 2015; Fraisse et al., 2017), may have complimented the successful sequencing in the optimized direct lysis method.

Approximately 10^3 GC/g of HAV and trace amounts of NoV GII (<10 GC/gRNA were detected in the positive date samples. Although detection of viral RNA do not necessarily reflect detection of infectious viral particles (de Roda Husman et al., 2009), the related illness amongst consumers of dates indicates that at least some of the detected HAV genomes originated from intact virions.

In HAV endemic countries, the routes of contamination and spread of HAV and NoV may be common, with sewage-polluted water as an example (FAO/WHO, 2008). Since HAV detection in foods implicated in disease outbreaks is often unsuccessful (Dentinger et al., 2001; Petrignani et al., 2010; Carvalho et al., 2012; Lassen et al., 2013), we analyzed the date samples for presence of NoV in addition to HAV, with the rationale that, should the detection of HAV be unsuccessful, detection of NoV may indicate fecal contamination. Indeed, NoV GII RNA was detected in dates from boxes with and without simultaneous detection of HAV and RNA. This suggests that if HAV detection fails during investigation of HAV suspected contaminated food samples, NoV may be used as indicator for human fecal contamination. With the lack of reported gastroenteritis with dates as the suspected vehicle, the detection of NoV RNA might represent degraded viruses or levels below the number that cause symptoms, which has been estimated to 18–1000 virus particles (Teunis et al., 2008). Due to the detected levels of NoV RNA being lower than required for successful sequencing and the lack of reported gastroenteritis suspected to be caused by dates, sequence characterization of the detected NoV RNA in the two boxes of dates was not attempted.

While dates from two boxes (box 10 and 6) representing two different batches (batch 4 and 1, respectively) tested positive for either both HAV and NoV RNA or only NoV RNA, the paired two boxes (box 1 and 4, respectively) originating from the same batches did not test positive for any virus. This indicates that viruses may be unevenly distributed in contaminated dates, as has been previously observed for viral contaminated raspberries (EFSA, 2014), semidried tomatoes (Donnan et al., 2012) and dates (Boxman et al., 2012). The uneven distribution of viral contaminated fruits underlines the importance of a sound sampling strategy, when conducting surveillance or outbreak studies of fruits suspected to be contaminated with viruses (Bosch et al., 2011).

The HAV genome detected in the dates matched by sequence 100% to the HAV genotype 3A detected in the stool sample collected from one case from whose household the specific date sample was obtained and to another non-related infected HAV patient who also had consumed dates from the same producer. This strongly indicates that these cases were infected after consumption of the contaminated dates, rather than the product was contaminated by the case wherefrom the dates were collected.

Although dates have not yet been described as a main risk factor for HAV outbreaks (EFSA, 2014), the present study suggest like the Dutch surveillance study (Boxman et al., 2012) that dates may prove to be an important vehicle of widespread, cross-border foodborne outbreaks of hepatitis.

In conclusion, use of an optimized direct lysis method for the extraction of viral RNA from dates suspected to be the source of a hepatitis A outbreak, allowed detection, quantification and characterization of a HAV strain matching a sequence detected among the outbreak cases. This establishes, to our knowledge, for the first time a direct sequence link between a hepatitis A infection and consumption of contaminated dates.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation, to any qualified researcher.

AUTHOR CONTRIBUTIONS

SR: writing – original draft preparation. AS: funding acquisition and supervision. TJ, SM, LM, and AS: resources. All authors: conceptualization and methodology of the study, formal analysis and investigation, and writing, reviewing, and editing of subsequent versions.

FUNDING

This work was supported by a Ph.D. scholarship to SR and supervised by the AS from the National Food Institute at the Technical University of Denmark.

ACKNOWLEDGMENTS

We thank Resadije Idrizi, National Food Institute, Technical University of Denmark and Kristine Vorborg, Virus Surveillance and Research, Statens Serum Institute, for technical assistance, as well as the Danish Veterinary and Food Administration and Technical University of Denmark for supporting this study.

REFERENCES

Bartsch, C., Szabo, K., Dinh-Thanh, M., Schrader, C., Trojnar, E., and Johne, R. (2016). Comparison and optimization of detection methods for noroviruses in frozen strawberries containing different amounts of RT-PCR inhibitors. *Food Microbiol.* 60, 124–130. doi: 10.1016/j.fm.2016.07.005

Bosch, A., Gkogka, E., Le Guyader, F. S., Loisy-Hamon, F., Lee, A., and van Lieshout, L. (2018). Foodborne viruses: detection, risk assessment, and control options in food processing. *Int. J. Food Microbiol.* 285, 110–128. doi: 10.1016/j.ijfoodmicro.2018.06.001

Bosch, A., Sánchez, G., Abbaszadegan, M., Carducci, A., Guix, S., and Le Guyader, F. S. (2011). Analytical methods for virus detection in water and food. *Food Anal. Methods* 4, 4–12. doi: 10.1007/s12161-010-9161-5

Boxman, I. L. A., te Loeke, N. A. J. M., Klunder, K., Hägele, G., and Jansen, C. C. C. (2012). Surveillance study of hepatitis A virus RNA on fig and date samples. *Appl. Environ. Microbiol.* 78, 878–879. doi: 10.1128/AEM.06574-11

Carvalho, C., Thomas, H., Balogun, K., Tedder, R., Pebody, R., Ramsay, M., et al. (2012). A possible outbreak of hepatitis A associated with semi-dried tomatoes, England, July-November 2011. *Euro Surveill.* 17, 4–7. doi: 10.2807/ese.17.06.20083-en

Collier, M. G., Khudyakov, Y. E., Selvage, D., Adams-Cameron, M., Epson, E., and Cronquist, A. (2014). Outbreak of hepatitis A in the USA associated with frozen pomegranate arils imported from Turkey: an epidemiological case study. *Lancet Infect. Dis.* 14, 976–981. doi: 10.1016/S1473-3099(14)70883-7

Costafreda, M. I., Bosch, A., and Pintó, R. M. (2006). Development, evaluation, and standardization of a real-time TaqMan reverse transcription-PCR assay for quantification of hepatitis A virus in clinical and shellfish samples. *Appl. Environ. Microbiol.* 72, 3846–3855. doi: 10.1128/AEM.02660-05

Coudray-Meunier, C., Fraisse, A., Martin-Latil, S., Guillier, L., Delannoy, S., Fach, P., et al. (2015). A comparative study of digital RT-PCR and RT-qPCR for quantification of Hepatitis A virus and Norovirus in lettuce and water samples. *Int. J. Food Microbiol.* 201, 17–26. doi: 10.1016/j.ijfoodmicro.2015.02.006

de Roda Husman, A. M., Lodder, W. J., Rutjes, S. A., Schijven, J. F., and Teunis, P. F. M. (2009). Long-term inactivation study of three enteroviruses in artificial surface and groundwaters, using PCR and cell culture. *Appl. Environ. Microbiol.* 75, 1050–1057. doi: 10.1128/AEM.01750-08

Dentinger, C. M., Bower, W. A., Nainan, O. V., Cotter, S. M., Myers, G., Dubusky, L. M., et al. (2001). An outbreak of hepatitis A associated with green onions. *J. Infect. Dis.* 183, 1273–1276. doi: 10.1086/319688

Donnan, E. J., Fielding, J. E., Gregory, J. E., Lalor, K., Rowe, S., and Goldsmith, P. (2012). A multistate outbreak of hepatitis A associated with semidried tomatoes in Australia, 2009. *Clin. Infect. Dis.* 54, 775–781. doi: 10.1093/cid/cir949

EFSA (2014). Tracing of food items in connection to the multinational hepatitis A virus outbreak in Europe. *EFSA J.* 12, 3821–4007. doi: 10.2903/j.efsa.2014.3821

FAO/WHO (2008). *Viruses in Food: Scientific Advice to support risk management activities. Meeting Notes*. Geneva: WHO.

Fraisse, A., Coudray-Meunier, C., Martin-Latil, S., Hennechart-Collette, C., Delannoy, S., Fach, P., et al. (2017). Digital RT-PCR method for hepatitis A virus and norovirus quantification in soft berries. *Int. J. Food Microbiol.* 243, 36–45. doi: 10.1016/j.ijfoodmicro.2016.11.022

Gallot, C., Grout, L., Roque-Afonso, A.-M., Couturier, E., Carrillo-Santisteve, P., and Pouey, J. (2011). Hepatitis A associated with semidried tomatoes, France, 2010. *Emerg. Infect. Dis.* 17, 566–567. doi: 10.3201/eid1703.101479

Hollinger, F. B., and Martin, A. (2013). "Hepatitis a virus," in *Fields Virology: Sixth Edition*, eds D. M. Knipe and P. Howley (Alphen aan den Rijn: Wolters Kluwer Health Adis), 551–577.

Hwang, S., Alhatlani, B., Arias, A., Caddy, S. L., Christodoulou, C., and Bragazzi Cunha, J. (2014). Murine Norovirus: propagation, quantification, and genetic manipulation. *Current Protocols in Microbiology* 33, 15K.2.1–15K.2.61.

ISO (2017). *ISO 15216-1:2017. Microbiology of the Food Chain — Horizontal Method for Determination of Hepatitis A Virus and Norovirus Using Real-Time RT-PCR — Part 1: Method for Quantification*. Bengaluru: ISO.

Lassen, G. S., Soborg, B., Midgley, S. E., Steens, A., Vold, L., and Stene-Johansen, K. (2013). Ongoing multi-strain food-borne hepatitis A outbreak with frozen berries as suspected vehicle: four Nordic countries affected, October 2012 to April 2013. *Euro Surveill.* 18:20467. doi: 10.2807/ese.18.17.20467-en

Le Guyader, F. S., Parnaudeau, S., Schaeffer, J., Bosch, A., Loisy, F., Pommepuy, M., et al. (2009). Detection and quantification of noroviruses in shellfish. *Appl. Environ. Microbiol.* 75, 618–624. doi: 10.1128/AEM.01507-08

McKee, A., Spear, S. F., and Pierson, T. W. (2015). The effect of dilution and the use of a post-extraction nucleic acid purification column on the accuracy, precision, and inhibition of environmental DNA samples. *Biol. Conserv.* 183, 70–76. doi: 10.1016/j.biocon.2014.11.031

Mollers, M., Boxman, I. L. A., Vennema, H., Slegers-Fitz-James, I. A., Brandwagt, D., Friesema, I. H., et al. (2018). Successful use of advertisement pictures to assist recall in a food-borne hepatitis A outbreak in the Netherlands, 2017. *Food Environ. Virol.* 10, 272–277. doi: 10.1007/s12560-018-9347-3

Müller, L., Raiser, S. G., Ethelberg, S., Vestergaard, H., Midgley, S. E., and Fischer, T. K. (2018). *No 11-2018: Hepatitis A virus outbreak due to Iranian dates. Epi-News, 11–2018*. Available online at: https://en.ssi.dk/news/epi-news/2018/no-11---2018 (accessed August 04, 2020).

Perrin, A., Loutreul, J., Boudaud, N., Bertrand, I., and Gantzer, C. (2015). Rapid, simple and efficient method for detection of viral genomes on raspberries. *J. Virol. Methods* 224, 95–101. doi: 10.1016/j.jviromet.2015.08.005

Petrignani, M., Harms, M., Verhoef, L., van Hunen, R., Swaan, C., and van Steenbergen, J. (2010). Update: a food-borne outbreak of hepatitis A in the Netherlands related to semi-dried tomatoes in oil, January-February 2010. *Euro Surveill.* 15, 1–4. doi: 10.2807/ese.15.20.19572-en

Rajiuddin, S. M., Jensen, T., Hansen, T. B., and Schultz, A. C. (2020). An optimised direct lysis method for viral RNA extraction, and detection of foodborne viruses on fruits, and vegetables. *Food Environ. Virol.* 12, 226–239. doi: 10.1007/s12560-020-09437-x

Rawsthorne, H., Phister, T. G., and Jaykus, L.-A. (2009). development of a fluorescent in situ method for visualization of enteric viruses.

Appl. Environ. Microbiol. 75, 7822–7827. doi: 10.1128/AEM.01986-09

Schrader, C., Schielke, A., Ellerbroek, L., and Johne, R. (2012). PCR inhibitors - occurrence, properties and removal. *J. Appl. Microbiol.* 113, 1014–1026. doi: 10.1111/j.1365-2672.2012.05384.x

Severi, E., Verhoef, L., Thornton, L., Guzman-Herrador, B. R., Faber, M., and Sundqvist, L. (2015). Large and prolonged food-borne multistate hepatitis A outbreak in Europe associated with consumption of frozen berries, 2013 to 2014. *Eurosurveillance* 20:21192. doi: 10.2807/1560-7917.ES2015.20.29.21192

Sun, B., Bosch, A., and Myrmel, M. (2019). Extended direct lysis method for virus detection on berries including droplet digital RT-PCR or real time RT-PCR with reduced influence from inhibitors. *J. Virol. Methods* 271:113638. doi: 10.1016/j.jviromet.2019.04.004

Teunis, P. F. M., Moe, C. L., Liu, P., Miller, E. S., Lindesmith, L., Baric, R. S., et al. (2008). Norwalk virus: how infectious is it? *J. Med. Virol.* 80, 1468–1476. doi: 10.1002/jmv.21237

World Health Organization [WHO] (2010). *The Global Prevalence of Hepatitis A Virus Infection and Susceptibility: A Systematic Review*. Geneva: World Heal. Organ.

World Health Organization [WHO] (2017). *Global Hepatitis Report, 2017*. Geneva: World Heal. Organ.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Rajiuddin, Midgley, Jensen, Müller and Schultz. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 22 September 2020
doi: 10.3389/fmicb.2020.581201

Development of PMAxx™-Based qPCR for the Quantification of Viable and Non-viable Load of *Salmonella* From Poultry Environment

*Jiawei Zhang, Samiullah Khan and Kapil K. Chousalkar**

School of Animal and Veterinary Sciences, The University of Adelaide, Roseworthy, SA, Australia

Determining the viable and non-viable load of foodborne pathogens in animal production can be useful in reducing the number of human outbreaks. In this study, we optimized a PMAxx™-based qPCR for quantifying viable and non-viable load of *Salmonella* from soil collected from free range poultry environment. The optimized nucleic acid extraction method resulted in a significantly higher ($P < 0.05$) yield and quality of DNA from the pure culture and *Salmonella* inoculated soil samples. The optimized primer for the amplification of the *invA* gene fragment showed high target specificity and a minimum detection limit of 10^2 viable *Salmonella* from soil samples. To test the optimized PMAxx™-based qPCR assay, soil obtained from a free range farm was inoculated with *Salmonella* Enteritidis or *Salmonella* Typhimurium, incubated at 5, 25, and 37°C over 6 weeks. The survivability of *Salmonella* Typhimurium was significantly higher than *Salmonella* Enteritidis. Both the serovars showed moisture level dependent survivability, which was significantly higher at 5°C compared with 25°C and 37°C. The PMAxx™-based qPCR was more sensitive in quantifying the viable load compared to the culture method used in the study. Data obtained in the current study demonstrated that the optimized PMAxx™-based qPCR is a suitable assay for quantification of a viable and non-viable load of *Salmonella* from poultry environment. The developed assay has applicability in poultry diagnostics for determining the load of important *Salmonella* serovars containing *invA*.

Keywords: viability assay, *Salmonella* contamination, poultry production, PMAxx-based qPCR, *Salmonella* load

OPEN ACCESS

Edited by:
David Rodriguez-Lazaro, University of Burgos, Spain

Reviewed by:
Eduardo Cesar Tondo, Federal University of Rio Grande do Sul, Brazil
Alejandro Garrido-Maestu, International Iberian Nanotechnology Laboratory (INL), Portugal

****Correspondence:***
Kapil K. Chousalkar
Kapil.chousalkar@adelaide.edu.au

Specialty section:
This article was submitted to Food Microbiology, a section of the journal Frontiers in Microbiology

Received: *08 July 2020*
Accepted: *31 August 2020*
Published: *22 September 2020*

Citation:
Zhang J, Khan S and Chousalkar KK (2020) Development of PMAxx™-Based qPCR for the Quantification of Viable and Non-viable Load of Salmonella From Poultry Environment. Front. Microbiol. 11:581201. doi: 10.3389/fmicb.2020.581201

INTRODUCTION

Food safety is an important aspect of human health and, therefore; state health departments are routinely tasked to trace back foodborne pathogens for their source of origin. Zoonotic pathogens, such as *Salmonella* and *Campylobacter,* often result in human gastroenteritis after the consumption of contaminated poultry products (Kärenlampi et al., 2007; Ford et al., 2018). Due to consumer demand, there has been an increase in free range poultry production over the years. In a free-range poultry production system, birds have access to the range area outside poultry shed. Chickens are the common host for multiple serovars of *Salmonella enterica,* which naturally colonize their gut. Therefore, infected chickens can shed *Salmonella* to their environment intermittently. In a free

range production system, the infected flock can contaminate range soil and result in the infection of the newly introduced flock. Therefore, it is vital to understand the viable and non-viable load of *Salmonella* in the poultry environment in relation to changes in temperature.

Soil type and temperature affect the survivability of *Salmonella* more than moisture level (Underthun et al., 2018). The final distribution of water and bacteria in the soil is influenced by soil texture, pH, temperature, and organic matter (Semenov et al., 2009). Temperature and water activity affect the survivability of *Salmonella* with the thermodynamic activity of water as the main factor in determining the survival. A study has shown that the incidence of *Salmonella* increases during warm seasons (Milazzo et al., 2016). Cross-contamination of the range area may pose a continuous threat to the health of the layer flocks. Therefore, understanding the viability of *Salmonella* in the soil is a crucial factor in devising strategies for its control in poultry production. In Australia, *Salmonella* Enteritidis is considered as an exotic pathogen in commercial poultry. Recently, there have been several *Salmonella* Enteritidis outbreaks on free range layer farms in Australia, which has prompted the mass culling of layer flocks. After mass culling, sheds are decontaminated, but it is often a challenge to decontaminate ranging areas.

Quantitative PCR (qPCR) can be used for quantifying pathogen load in clinical samples; however, primers targeting a specific region of the nucleic acids will also amplify DNA obtained from the non-viable load present in the population. Viability assays, such as live/dead cell viability assay, trypan blue, BacTiter-Glo^{TM} microbial cell viability assay (Promega, Australia), LIVE/DEAD BacLight bacterial viability kit (ThermoFisher Scientific, Australia) and propidium monoazide (PMA) (Banihashemi et al., 2012; Barbau-Piednoir et al., 2014; Jäger et al., 2018) or ethidium bromide monoazide (Nocker et al., 2006; Wang and Mustapha, 2010) based qPCR can be used for quantifying viable microbial load. PMA is a nucleic acid intercalating dye that is cell impermeant and therefore will diffuse only across the cell membrane of dead microbes. After binding to nucleic acids, upon photolysis, the dye covalently reacts with the microbial DNA/RNA thus inhibiting it from the amplification by PCR. Therefore, PMA-based qPCR can be a reliable assay for differentiating viable cells from non-viable population of microbes. However, given the complex nature of the environmental samples, none of them have been optimized for quantifying the viable and non-viable load of *Salmonella* from the environment. In the current study, PMAxx^{TM}-based qPCR was optimized for the quantification of *Salmonella* from poultry environment and in the subsequent *Salmonella* survivability experiment, its applicability was tested. Therefore, the three main objectives of the current study were: (A) Optimize method for PMAxx^{TM} based viability qPCR for *Salmonella* quantification from the soil. (B) Understand the effects of temperature on the survivability of *Salmonella* Typhimurium and *Salmonella* Enteritidis in soil. (C) Apply the optimized PMAxx^{TM} based qPCR to quantify the viable and non-viable load of *Salmonella* Typhimurium and *Salmonella* Enteritidis from soil samples incubated at various temperatures.

MATERIALS AND METHODS

Pure cultures of *Salmonella* Typhimurium phage type 9 or *Salmonella* Enteritidis (phage-type 7A obtained from Elizabeth McArthur Agriculture Institute, NSW, Australia) were streaked onto nutrient agar (NA; ThermoFisher Scientific, Australia) plates and incubated overnight at 37°C. A single colony was subcultured in Luria Bertani (LB; ThermoFisher Scientific, Australia) broth at 37°C in a shaking incubator until the OD_{600} was approximately 1, which was used for calculating the colony forming unit (CFU) per mL.

DNA Extraction From Pure Culture of *Salmonella* and *Salmonella* Inoculated Soil

The *Salmonella* DNA extraction from pure culture (n = 6) and soil (n = 6) was carried out using the QIAamp DNA Mini Kit (Qiagen, Australia) with some modifications in the protocol. Approximately 10^9 CFU/mL of *Salmonella* broth culture was centrifuged at 19,500 × g for 3 min and the pellet was resuspended in 100 μL phosphate buffered saline (PBS) into which 300 μL tissue lysis buffer (ATL) and half spoon of acids washed 106 μm beads (Sigma Aldrich, Australia) were added. The samples were homogenized (Bullet Blender, Next Advance, United States) for 2–3 min. The samples were heated at 70°C for 15 min with occasional mixing and then centrifuged for 1 min at 5,000 × g to remove the beads and undissolved materials. The supernatant was transferred to a 1.5 mL tube that contained 50 μL proteinase K (20 mg/mL). After a brief vortexing, 50 μL lysozyme (100 mg/mL) and 300 μL AL lysis buffer were added. The samples were briefly vortexed and incubated at 70°C for 15 min with occasional inversion. To avoid clump formation, the samples were vortexed at each step of adding the reagents. After the incubation, 300 μL ethanol (100%) was added and the lysates were passed through the spin column by centrifuging for 1 min at 6,000 × g and the process was repeated for the rest of the lysates. The DNA in the spin columns was washed with 500 μL of each of wash buffers I and II and eluted in 50 μL of buffer ATE as per the manufacturer's protocol. The extracted DNA was assessed for quality and concentration in a Nanodrop 1000 and stored at −20°C for further use.

The DNA extraction method from soil was optimized based on adding 9 mL of PBS to 1 g of soil. Briefly, 1 g of soil was weighed in 15 mL tubes and inoculated with 10^9 CFU/mL of *Salmonella* and the samples were mixed thoroughly. The tubes were centrifuged at 900 × g for 1 min to settle down the undissolved materials. The maximum supernatants were transferred to 15 mL tubes and centrifuged at full speed (4500 × g) for 15 min at 4°C. The bacterial pellet layer was gently resuspended in 100 μL of PBS and transferred to 1.5 mL tubes. The DNA was extracted as described previously. In a separate pilot study, to test the efficiency of the bacterial separation method from soil, the bacterial pellet was resuspended in PBS, serially diluted, plated onto xylose-lysine deoxycholate (XLD; ThermoFisher Scientific, Australia) agar media and incubated overnight at 37°C.

Optimization of PMAxx^{TM} Step for Distinguishing Between Viable and Non-viable Load

Salmonella culture in the LB broth was prepared as described previously. The broth culture samples were prepared as viable and non-viable *Salmonella*. For the non-viable culture preparation, *Salmonella* was exposed to 95°C for 5 min in a dry heat block. To confirm the non-viable *Salmonella*, 100 μL of the culture was plated onto XLD and incubated overnight at 37°C. For determining the optimum concentration of PMAxx^{TM} dye (Biotium, United States), 10 μL of the dye solution (2.5, 5, and 10 mM) was added into 1 mL of the *Salmonella* culture in LB broth (n = 6) or left as a control (n = 6). The PMAxx^{TM} treated samples were incubated in dark for 10 min on a rocker. After the incubation, the samples were exposed to light (500 volts) at a distance of 20 cm for 15 min with intermittent inversion. The samples were centrifuged at 19500 × g for 3 min, the bacterial pellets were suspended in 100 μL PBS and the DNA was extracted as described previously.

Optimization of Primers and Construction of Standard Curve

Primers were either designed using the NCBI primer software or sourced from literature. Primers were evaluated for their secondary structure characteristics in web based NetPrimer and Beacon Designer software. Short sequence primers (<300 bp) were avoided to reduce the chances of amplifying genomic DNA obtained from non-viable bacterial population in samples treated with PMAxx^{TM}. To broaden the scope of this study for detection of a range of *Salmonella enterica* serovars, primer amplifying the fragment of invasion A gene (*invA*; **Table 1**) was used for the PMAxx^{TM}-based qPCR. To determine the specificity and amplification efficiency of the *invA* primer, qPCR was performed on 10-fold serial dilutions (10^8 to 10^1) of *Salmonella* DNA extracted from a pure culture (n = 3 and repeated twice). To determine the minimum detection limit of the *invA* primer, qPCR was performed on DNA extracted pure culture and from spiked soil samples (n = 3 and repeated twice) with different dilutions of *Salmonella* (10^8 to 10^1). The qPCR products were assessed by melting curve analysis and then visualized on 2% gel electrophoresis to assess the specificity of the primer pair. For the quantification of *Salmonella* from soil samples, a standard curve was generated from a 10-fold serial dilutions (10^{-2} to 10^{-12}) of purified qPCR product (using the QIAquick PCR Purification Kit) of *invA* fragment and used as a reference to determine the load of *Salmonella*. Primer pairs for amplifying *ompA* gene were also evaluated for specificity through qPCR conducted on DNA obtained from pure cultures of *Salmonella* Typhimurium and *Escherichia coli* (**Table 1**).

Quantitative PCR

The qPCR was performed with the SensiFAST SYBR Hi-ROX Kit (Bioline, Australia) following the manufacturer's protocol. The master mix was prepared and applied to reaction wells through Corbett CAS1200 robot (Corbett Life Science, Australia). The final reaction volume of 20 μL contained 10 μL SensiFAST SYBR Hi-Rox mix, 1 μL of each of the forward and reverse primers (10 μM concentration), 6 μL RNase-free water and 2 μL of DNA template. The reaction was performed in duplicates in a Rotor-Gene Disc 100 (Qiagen, Australia) with a Rotor-Gene 6000 Thermocycler (Corbett Life Science, Australia). Each run contained no-template control for ruling out external contaminations. The three-step cycling conditions included an initial denaturation at 95°C for 3 min, then 40 cycles of denaturation at 95°C for 5 s and annealing and extension at 59°C (**Table 1**) and 72°C for 20 and 30 s, respectively. During qPCR cycles, the fluorescence detection was conducted at the end of each annealing step, and a melting curve analysis step (at a ramp from 55 to 95°C) was included to assess the specificity of the amplification.

Preparation of Soil Samples

Soil from a range area in a free-range production system was collected for assessing the survivability of *Salmonella* Typhimurium and *Salmonella* Enteritidis. To mimic the field conditions, the samples were inoculated with 5 × 10^9 CFU per 100 g of soil with moisture adjusted to around 15%. The moisture level was adjusted by adding autoclaved water into the soil sample containers and measuring the moisture content with a basic Moisture Analyser Model MJ33 (Mettler Toledo, Switzerland). Each treatment group contained three biological replicates and

TABLE 1 | Detail of the primers used in optimization of the PMAxx^{TM}-based qPCR assay.

Gene symbol	Primer sequence (5′-3′)	Accession no.	Product size (bp)	Annealing temperature (°C)	Amplification efficiency (%)
invA	F: AAACCTAAAACCAGCAAAGG R: TGTACCGTGGCATGTCTGAG	M90846.1	605	59	92
invA	F: TGGGGCGGAATATCATGACG R: AGGAAGGTACTGCCAGAGGT	M90846.1	1045	60	80
ompA	F: TACGCTGGTGCTAAACTGGGCT R: AGCGCGAGGTTTCACGTTGTCA	NC_003197.2	882	60	83
ompA	F: CAGTACCATGACACCGGCTT R: ATTCCAGACGGGTTGCGATT	NC_003197.2	385	60	90

*Specificity of each primer pairs was confirmed through melting curve analysis and gel electrophoresis. Sequence of the invA 605 bp was sourced from literature (Akiba et al., 2011). Among the listed primers in **Table 1**, invA having 605 bp sequence length had higher amplification efficiency and target specificity. Therefore, primer pair labeled as invA 605 bp was used in the quantification of viable and non-viable load of Salmonella Enteritidis and Salmonella Typhimurium through PMAxx^{TM}-based qPCR from soil samples incubated at 5, 25, and 37°C for up to 6 weeks.*

the experiment was repeated twice to confirm the applicability of the optimized method. The *Salmonella* inoculated soil samples were incubated at 5, 25, and 37°C for 6 weeks and the samples were processed at weekly intervals for quantifying the bacterial load through culture method and PMAxx^{TM}-based qPCR.

The moisture content and water activity of the individual soil samples were measured at each sampling time-point. Moisture content is a measure of the total amount of water in a sample. A basic moisture analyzer that consists of both weighing and heating and is used to determine the moisture content of a sample with the loss on drying principle. For measuring the moisture content, approximately 8 g of soil was spread onto a sample pan and measured by a Basic Moisture Analyzer (Model MJ33, Mettler Toledo, Switzerland) following the below formula:

$$\text{Moisture content } (0\ldots-100\%) = -\frac{\text{wet weight} - \text{dry weight}}{\text{wet weight}} \times 100\%$$

The values were recorded as loss in total moisture content.

Water activity (a_w) is a thermodynamic measurement of the energy of water in a sample. The value indicates how tightly water is bound, structurally or chemically, within a substance. The lower a sample's water activity, the more tightly bound that water is within the sample. It is measured using a water activity meter (range: 0.00–1.00 a_w) that assesses the partial vapor pressure of water in a sample and divides it by the standard state partial vapor pressure of water. To measure the water activity, approximately 3 g of soil was dispensed into a sample cup and measured by PawKit Water Activity Meter (Model P06760; AquaLab, Decagon Devices, United States).

Quantification of *Salmonella* by Culture Method and PMAxx^{TM}-Based qPCR

In order to understand the applicability of the optimized PMAxx^{TM}-based qPCR in the field conditions, soil samples inoculated with *Salmonella* Enteritidis or *Salmonella* Typhimurium, and incubated at different temperatures (5, 25, and 37°C) were processed at weekly intervals for the quantification of *Salmonella* load through culture method and qPCR. The culture data (obtained by direct plating) were presented as $\log_{10}$ CFU/g, while the PMAxx^{TM}-based qPCR data were presented as $\log_{10}$ *Salmonella* load for viable and non-viable counts per gram of soil samples.

At each sampling time-point, 1 g of the soil samples was mixed into 9 mL of buffered peptone water (BPW; ThermoFisher Scientific, Australia), serially diluted, plated onto XLD agar media and incubated overnight at 37°C. The *Salmonella* count was expressed as $\log_{10}$ CFU/g of soil sample. A sub-set of the samples was processed using an enrichment method. For the enrichment method, the BPW samples were incubated overnight at 37°C and a 100 μL was added into 9 mL of Rappaport-Vassiliadis soy peptone broth (RVS; ThermoFisher Scientific, Australia). The RVS samples were incubated overnight at 42°C to allow the selective growth of *Salmonella*. The incubated RVS samples were streaked on XLD plates for the confirmation of *Salmonella* in the soil samples. Positive *Salmonella* samples were scored as 1 and negative as 0 for the determination of the proportion of positive samples.

At each sampling time-point, 1 g of soil from individual samples was weighed into 15 mL tubes containing 9 mL PBS for DNA extraction and PMAxx^{TM}-based qPCR. The soil samples were centrifuged at 900 × *g* for 1 min at 4°C to remove the undissolved materials. The supernatants were centrifuged at 4500 × *g* for 15 min at 4°C and the bacterial pellets were resuspended in 1 mL PBS. The samples were treated with PMAxx^{TM} or left as control and processed for DNA extraction and qPCR as described previously. A standard curve was used to quantify the DNA copy number from the cycle quantitation (Cq) values. The qPCR data ($\log_{10}$ DNA copy number) were expressed as viable and non-viable load of *Salmonella* per gram of soil samples.

Statistical Analysis

Data for *Salmonella* CFU/g, the proportion of positive samples and viable and non-viable load/g were analyzed in StatView version 5.0.1.0 by taking *Salmonella* treatment and sampling time-point as main effects. Level of significance was determined at protected least significant difference (PLSD) < 0.05. Mean values ± SD were visualized in GraphPad Prism 8.

RESULTS

Optimization of PMAxx^{TM}-Based qPCR for the Quantification of Viable and Non-viable Load of *Salmonella*

The optimized method for DNA extraction from pure culture resulted in significantly ($P < 0.05$) higher yield and pure genomic DNA of *Salmonella*. Tested in Nanodrop-1000, 1×10^9 CFU/mL of *Salmonella* broth yielded an average of 101.38 ng/μL DNA with 260/280 and 260/230 ratios of 2.10 and 2.32, respectively (**Figure 1**). Primer optimization steps showed that primers for the *ompA* gene were non-specific and amplified the genomic DNA of *Escherichia coli* (data not provided). Long size sequence primers

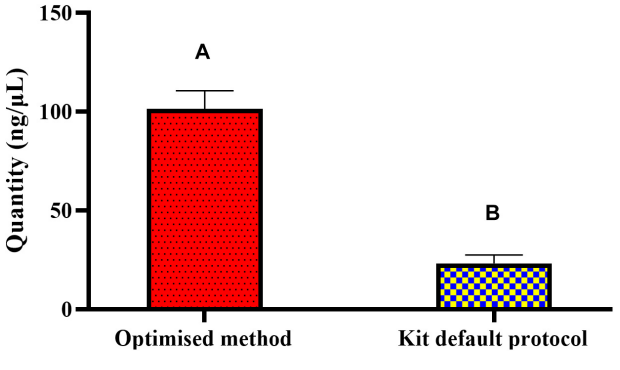


FIGURE 1 | DNA quantity (ng/μL) obtained through the optimized and the standard kit protocol methods from pure culture of *Salmonella*. The optimized extraction method resulted in significantly higher yield of DNA. Different superscripts (^{A,B}) across the bars show significant difference.

(∼1045 bp) designed for the amplification of the *invA* gene were less efficient (80%) compared to the medium size primers (∼605 bp). Therefore, the primers that amplified the 605 bp region of *invA* were used in the subsequent PMAxx^{TM}-based qPCR for quantifying viable and non-viable load of *Salmonella* Enteritidis and *Salmonella* Typhimurium from soil samples incubated at 5, 25, and 37°C for up to 6 weeks. Culture count data obtained from plating the isolated *Salmonella* from the inoculated soil showed no significant effect of the background materials. These data (4×10^8 CFU/mL) compared well with the dose (10^9 CFU/mL) used for 1 g of soil inoculation.

The qPCR results showed that the amplification efficiency of the *invA* primers in the 10-fold serial dilutions of DNA extracted from *Salmonella* was 92% with R^2 value as 0.99 (**Figure 2A**). The melting curve analysis of the amplicon showed a single peak, thereby confirming the specificity of the *invA* primers in the DNA obtained from *Salmonella* spiked soil samples (**Figure 2B**). The gel results confirmed that the minimum detection limit of *invA* in pure culture and soil samples spiked with 10-fold serial dilutions of viable *Salmonella* subsequently treated with PMAxx^{TM} was 10^1 and 10^2 *Salmonella* cells, respectively (data not shown).

To assess the inhibitory effects of PMAxx^{TM} on the target DNA amplification, qPCR was performed on DNA obtained from viable and non-viable *Salmonella* culture treated with 3 different concentrations of PMAxx^{TM}. The qPCR data (expressed as $\log_{10}$ *Salmonella* load) showed that the PMAxx^{TM} concentration did not significantly ($P > 0.05$) inhibit the amplification of DNA obtained from the viable *Salmonella* (**Figure 3**). However, PMAxx^{TM} concentrations significantly ($P < 0.05$) inhibited the

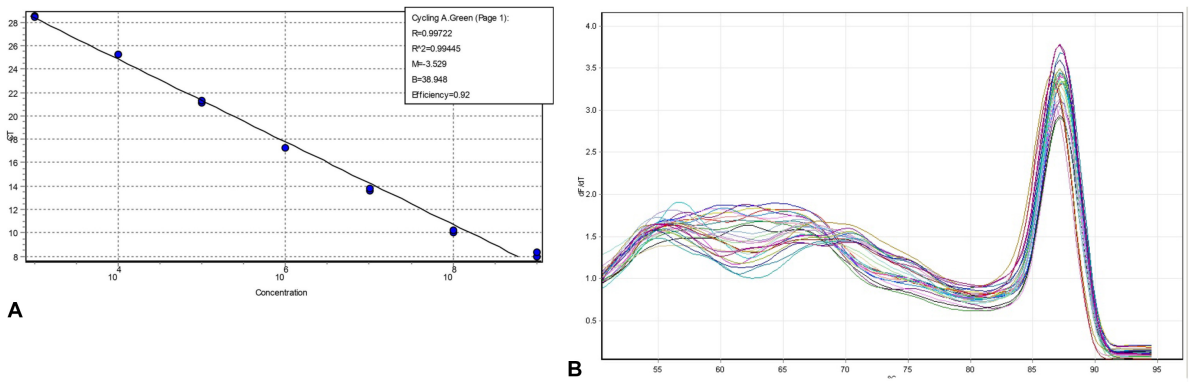


FIGURE 2 | Amplification efficiency and melting curve analyses of *invA* determined in 10-fold serial dilutions of DNA. **(A)** PCR performed on *Salmonella* DNA extracted from a pure culture with 1×10^9 CFU/mL. **(B)** Melting curve of *invA* (605 bp) in DNA obtained from 10-fold serially diluted (10^8 to 10^1 CFU/g) *Salmonella* spiked soil samples.

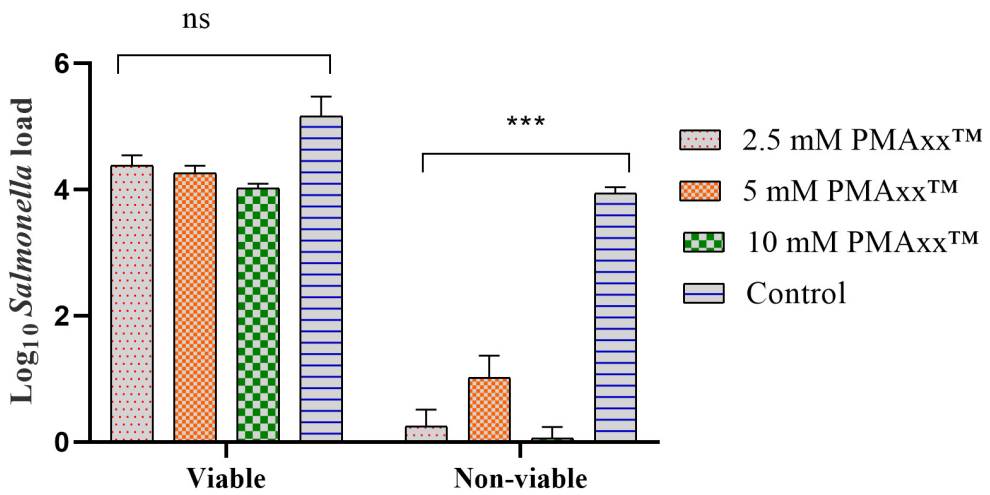


FIGURE 3 | Effects of different concentrations (2.5, 5, and 10 mM) of PMAxx^{TM} on the amplification of target DNA obtained from the viable and non-viable culture of *Salmonella*. The data showed that PMAxx^{TM} efficiently inhibited the amplification of the target DNA obtained only from the non-viable culture of *Salmonella*. This trial was repeated three times and an average of the data was presented. For obtaining non-viable cells, *Salmonella* in LB broth was heated at 95°C for 5 min, and a 100 μL plated on XLD agar and incubated overnight at 37°C confirmed that the heat treatment was effective in killing 100% of *Salmonella* in the broth. Asterisks (***) denote significant difference at $P < 0.0001$, while ns denotes non-significant difference.

amplification of DNA obtained from the non-viable culture of *Salmonella* (**Figure 3**). These data showed that 5 μL of 2.5 mM PMAxx™ per 500 μL of *Salmonella* culture was sufficient for use in PMAxx™-based qPCR assay to discriminate between the viable and non-viable *Salmonella* load. The significantly higher *Salmonella* load in the control samples of the non-viable group confirmed that PMAxx™ penetrated across the cell membrane of compromised cells only.

Survivability of *Salmonella* Assessed by Culture Method

The culture based method showed that the overall survivability of *Salmonella* was significantly higher for *Salmonella* Typhimurium compared with *Salmonella* Enteritidis (**Figure 4A**) despite the non-significant difference in moisture level and water activity between the two treatment groups (**Figures 4C,D**). Samples that were negative by direct plating and processed by the enrichment method showed higher survivability for *Salmonella* Typhimurium strain compared with *Salmonella* Enteritidis strain in soil samples incubated at 25 and 37°C (**Figure 4B**).

The survivability of *Salmonella* Enteritidis and *Salmonella* Typhimurium ($\log_{10}$ CFU) in soil was significantly affected by storage temperature, moisture level and water activity (**Figures 5A–D**). *Salmonella* Typhimurium survived at a significantly higher level than *Salmonella* Enteritidis at 5, 25, and 37°C (**Figure 5A**). Throughout the experiment, the survivability of both the serovars decreased significantly, whereas *Salmonella* Enteritidis could not be quantified by culture method from soil samples incubated at the 25 and 37°C from week 2 onward (**Figure 5A**). Furthermore, both the *Salmonella* Enteritidis and *Salmonella* Typhimurium in soil samples incubated at 5°C survived better than the samples incubated at 25 and 37°C. Therefore, the data confirmed that the *Salmonella* Typhimurium strain had a better capability to resist to temperature and moisture driven changes compared to the *Salmonella* Enteritidis strain.

To determine if there was relatively low viable *Salmonella* in the samples with no growth after direct plating, samples were further tested by the enrichment method. The results showed that *Salmonella* Enteritidis was not detectable in the week 5, and 6 soil samples incubated at 37°C, while *Salmonella* Typhimurium was detected until week 5 of incubation (**Figure 5B**). At 25°C, *Salmonella* Typhimurium was quantified from the soil samples until week 6 of incubation (**Figure 5A**). Therefore, the data confirmed that the culturability was driven by water activity and moisture level in the soil with a higher loss in moisture at 25 and 37°C (**Figures 5C,D**).

To investigate how moisture level (%) and water activity influenced the survivability of different *Salmonella* serovars, regression analysis between the moisture level and $\log_{10}$ CFU/g of soil samples was performed. Overall, for both the *Salmonella* Enteritidis and *Salmonella* Typhimurium treatment groups at all temperatures, the $\log_{10}$ CFU/g decreased with a decline in the moisture level (**Figures 6A–F**) and water activity (**Table 2**). For the samples incubated at 25 and 37°C, the decline of $\log_{10}$ CFU/g was highly associated with the decrease in the moisture level, while lower consistency of results was recorded at 5°C. Especially for *Salmonella* Enteritidis samples incubated at 5°C, there was a weak correlation between the $\log_{10}$ CFU/g and the moisture level.

Survivability of *Salmonella* Assessed by PMAxx™-Based qPCR

The optimized PMAxx™-based qPCR was used to assess the survivability of *Salmonella* serovars in soil samples incubated at three different temperatures. The overall viable and non-viable load in the *Salmonella* Enteritidis inoculated soil was significantly lower than *Salmonella* Typhimurium (**Figure 7A**). The viable load was significantly higher in soil samples incubated at 5°C compared to the samples incubated at 25 and 37°C (**Figure 7B**) due to moisture level driven changes.

The PMAxx™-based qPCR data for both the *Salmonella* Enteritidis and *Salmonella* Typhimurium treatment groups showed that the survivability of *Salmonella* decreased throughout incubation period (**Figures 8A,B**). The data at weekly intervals and incubation specific temperature showed some discrepancy in terms of higher load at week 2 compared with week 1 possibly due to the unequal distribution of *Salmonella* in the soil samples. The viable load of *Salmonella* Enteritidis and *Salmonella* Typhimurium in the samples incubated at 5°C was significantly higher than the samples incubated at 25 and 37°C (**Figures 8A,B**) due to moisture level driven changes. The viable load decreased with the increase in storage temperature and sample incubation period.

The CFU and viable load data obtained from different treatment groups were compared to understand a correlation between the culture method and the optimized PMAxx™-based qPCR. There was a significant positive correlation between the CFU and viable load data (**Figures 9A–F**). The decline in CFU was positively correlated with the qPCR data. The PMAxx™-based qPCR was more sensitive in quantifying the viable load of *Salmonella* serovars as compared to the culture method used in the study. For both the *Salmonella* Enteritidis and *Salmonella* Typhimurium samples incubated at 5°C, the correlation between CFU and the viable load was significantly stronger compared with the samples incubated at 25 and 37°C due to higher moisture content.

DISCUSSION

The main objective of this study was to optimize a PMAxx™-based qPCR for quantification of viable and non-viable *Salmonella* from environmental samples and to test the optimized assay on the artificially contaminated samples. The findings showed that the optimized qPCR can be accurately used for differentiating viable and non-viable *Salmonella* in soil samples collected from the free-range egg farm environment.

The QIAamp Fast DNA Stool Mini Kit (Qiagen, Australia) was used for the extraction of *Salmonella* DNA from pure cultures and *Salmonella* inoculated soil samples. As the manufacturer's protocol did not yield high quality and quantity of DNA and was not applicable for the inclusion of the PMA step

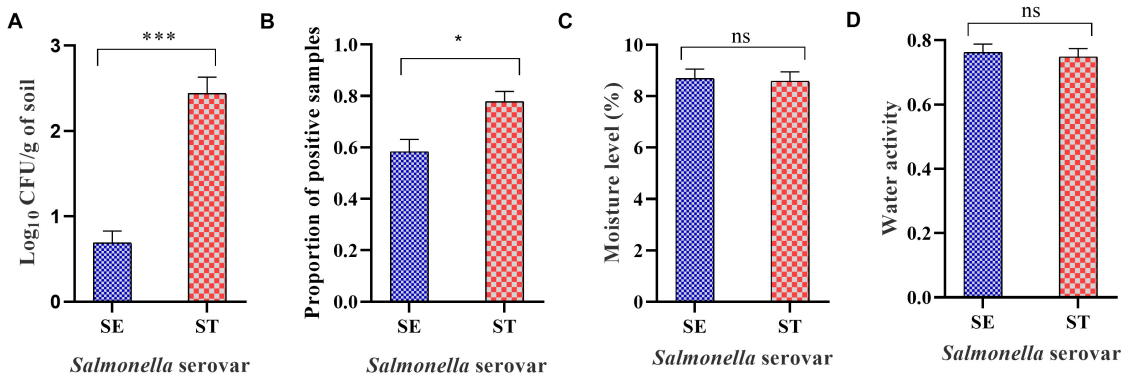


FIGURE 4 | Overall load of *Salmonella* Enteritidis (SE) and *Salmonella* Typhimurium (ST) in soil incubated at different temperatures and sampled at weekly intervals. The samples were processed through direct plating for *Salmonella* load determination. Samples that turned negative through the direct plating were enriched for the qualitative assessment of *Salmonella*. **(A)** Mean $\log_{10}$ CFU/g of SE and ST in soil over 6-week of incubation time period. **(B)** Mean proportion of *Salmonella* positive samples of SE and ST in soil incubated at different temperatures. **(C)** Mean moisture level of soil inoculated with SE or ST **(D)** Mean water activity (a_w) of soil inoculated with SE or ST. Asterisks (***) and (*) denote significant difference at $P < 0.0001$ and $P < 0.01$, respectively, while ns denotes non-significant difference.

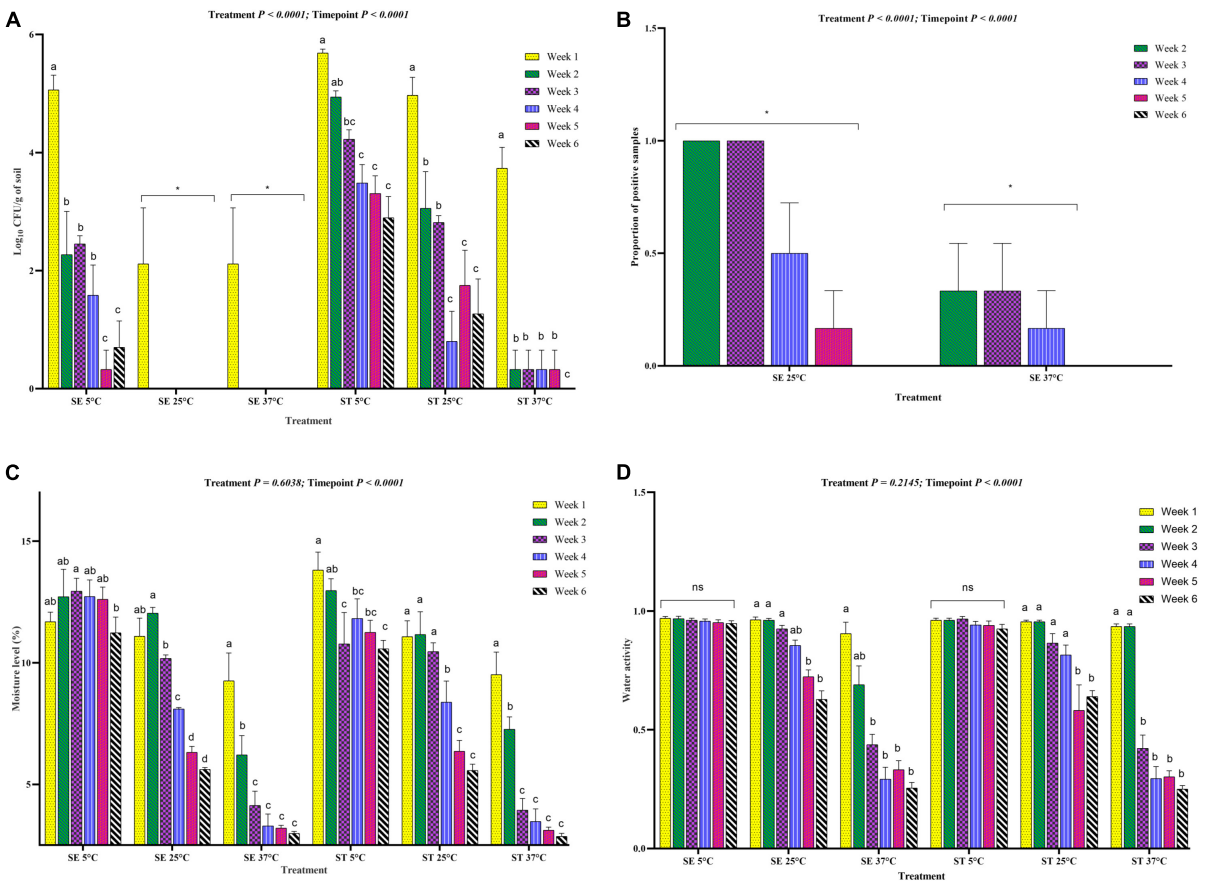


FIGURE 5 | Load of *Salmonella* Enteritidis (SE) and *Salmonella* Typhimurium (ST) in soil incubated at three different temperatures for 6 weeks. **(A)** Log_{10} CFU/g of soil of SE and ST affected by storage temperatures (5, 25, and 37°C). **(B)** Proportion of positive soil samples for SE and ST determined through the enrichment method. Only, the samples that did not show any *Salmonella* growth after direct plating were enriched. **(C)** Moisture level of the *Salmonella* inoculated soil samples incubated at 5, 25, and 37°C. **(D)** Water activity of the *Salmonella* inoculated soil samples incubated at 5, 25, and 37°C. The *Salmonella* inoculated soil samples were processed for *Salmonella* load through culture, *Salmonella* detection through enrichment method and moisture and water activity determination at weekly intervals for up to 6 weeks of incubation time period. Asterisk (*) denotes significant difference at $P < 0.01$, while ns denotes non-significant difference. Different superscripts (^{a,b,c,d}) show significant difference within each treatment group at each sampling timepoint.

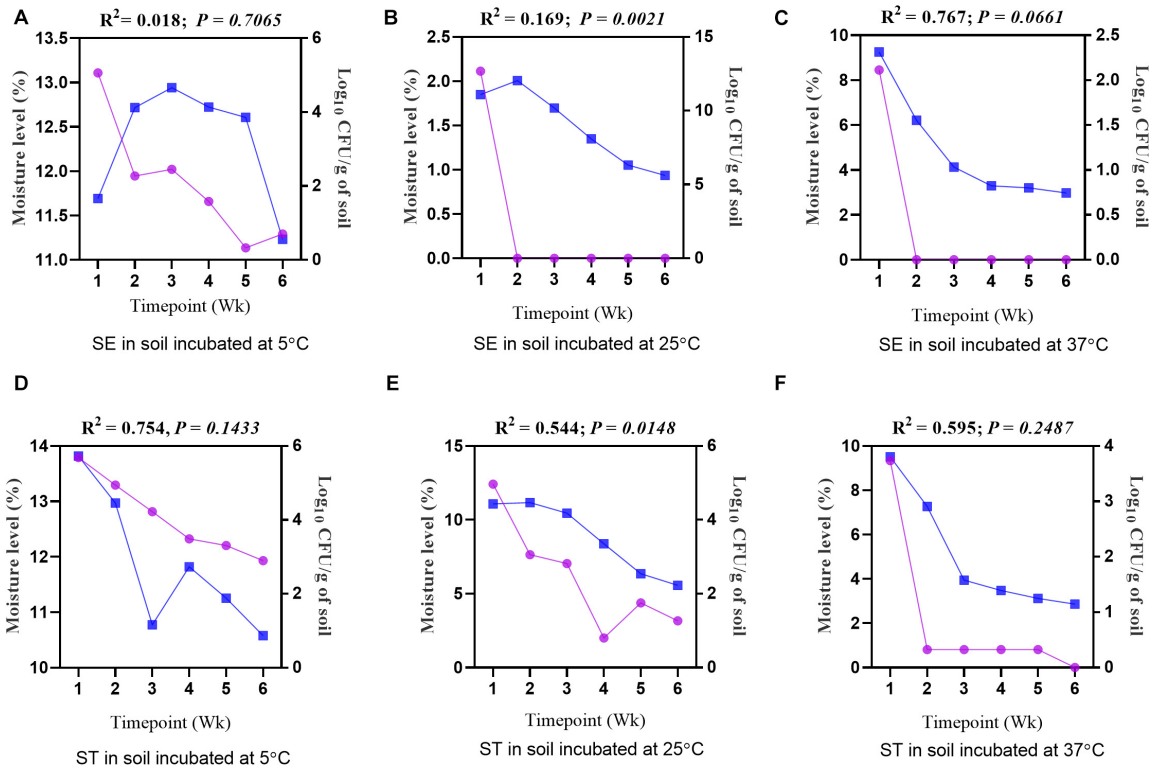


FIGURE 6 | Regression analysis showing a positive correlation between the moisture level (%) and the load of *Salmonella* Enteritidis (SE) and *Salmonella* Typhimurium (ST) in soil incubated at 5, 25, and 37°C. **(A)** Soil inoculated with SE and incubated at 5°C. **(B)** Soil inoculated with SE and incubated at 25°C. **(C)** Soil inoculated with SE and incubated at 37°C. **(D)** Soil inoculated with ST and incubated at 5°C. **(E)** Soil inoculated with ST and incubated at 25°C. **(F)** Soil inoculated with ST and incubated at 37°C. Blue lines show moisture level (%), while purple lines show $\log_{10}$ CFU/g of soil.

hence, the optimized method for DNA extraction was used in this study. The optimized method in the current study was based on microbe separation from the samples through gradient centrifugation. The qPCR results showed no effects of background materials in soil samples on amplification efficiency. This showed that the optimized method developed in this study removed the background materials during bacterial pelleting from soil samples. The *invA* primers used during this study showed that the primer pair was highly specific for target amplification with high efficiency (92%). The data also showed that its minimum detection limit was 10^2 *Salmonella* per gram of soil. Primer optimization for the amplification of *ompA* gene showed that these primers were non-specific by amplifying DNA obtained from *Escherichia coli*.

TABLE 2 | Correlation of water activity and *Salmonella* load ($\log_{10}$ CFU/g of soil).

Treatment	R^2	*P*-value
Salmonella Enteritidis in soil incubated at 4°C	0.474	0.0001
Salmonella Enteritidis in soil incubated at 25°C	0.181	0.0020
Salmonella Enteritidis in soil incubated at 37°C	0.633	0.0069
Salmonella Typhimurium in soil incubated at 4°C	0.707	0.0001
Salmonella Typhimurium in soil incubated at 25°C	0.461	0.0036
Salmonella Typhimurium in soil incubated at 37°C	0.333	0.0397

At each sampling time-point, water activity of the individual samples was measured by PawKit Water Activity Meter as per the manufacturer's instructions.

PMA has been used in distinguishing viable from non-viable cells of bacteria in pure culture (Nocker et al., 2006; Li and Chen, 2013) and from food/water matrices (Dinu and Bach, 2013; Kibbee and Örmeci, 2017; Li et al., 2017; Jäger et al., 2018). Similarly, PMA has also been used to differentiate live viruses from the dead population (Bellehumeur et al., 2015). However, the method has not been optimized for complex materials such as soil. Also, previous literature presented data as Cq values rather than viable and non-viable load (Li and Chen, 2013; Li et al., 2015). Therefore, the method optimized during this study is accurate and rapid that can be used to test the presence of viable *Salmonella* from complex soil matrices.

In the current study, the data showed that 10 μL of 2.5 mM PMAxx^{TM} into 1 mL of the *Salmonella* culture was sufficient to differentiate viable cells from the non-viable cells. The DNA extracted from the non-viable bacteria was minimally amplified by qPCR. Studies have reported different concentrations of PMA. For example, 10 μL of 2.5 mM PMA in 0.5 mL of culture samples was used (Liang et al., 2011). However, in the current study, there was no significant difference in the load of *Salmonella* obtained from the viable culture treated with 2.5, 5, and 10 mM PMAxx^{TM}. The concentration used in the current study is the recommended

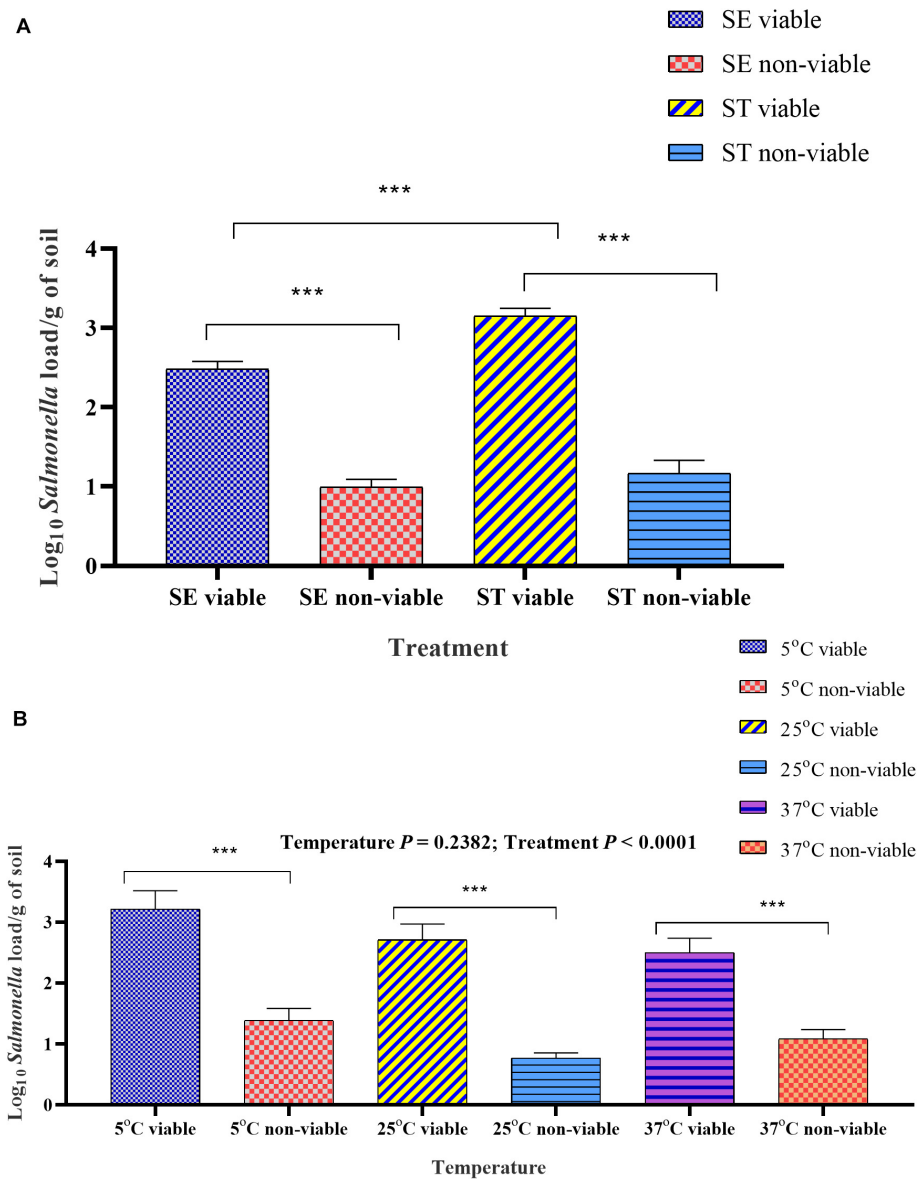


FIGURE 7 | Overall viable and non-viable load quantified by PMAxx^{TM}-based qPCR from *Salmonella* inoculated soil affected by temperature. **(A)** Mean $\log_{10}$ viable and non-viable load of *Salmonella* Enteritidis (SE) and *Salmonella* Typhimurium (ST) in soil samples incubated at 5, 25, and 37°C and processed at weekly intervals over 6-week of incubation time period. **(B)** Mean $\log_{10}$ viable and non-viable load of SE and ST affected by temperature. Samples were treated either with PMAxx^{TM} or left as control before the extraction of DNA. The DNA copy number was calculated through the standard curve and the data were presented as $\log_{10}$ *Salmonella* load. Asterisks (***) denote significant difference at $P < 0.0001$.

concentration from the manufacturer of the improved version of PMA, which is called PMAxx^{TM}. If used in the high throughput diagnostic assay platforms, the lower concentration of PMAxx^{TM} can save significant costs required for veterinary diagnostics. Data in the current study showed that the PMAxx^{TM} concentration had no significant inhibitory effects on the amplification of DNA obtained from viable *Salmonella*.

In the subsequent applicability steps of the PMAxx^{TM}-based qPCR, the CFU and qPCR data showed that the survivability of *Salmonella* Typhimurium and *Salmonella* Enteritidis was driven by incubation temperature, moisture level and water activity. The data showed that *Salmonella* Typhimurium had significantly higher capability to survive in soil compared with the *Salmonella* Enteritidis in the conditions applied during this study. However, it is worth to note that survivability of different strains of the same serovars of *Salmonella* may vary in different conditions. Significantly higher survivability at 5°C showed that both the *Salmonella* Enteritidis and *Salmonella* Typhimurium can remain viable for longer periods at 5°C in soil. The significant role of moisture and water activity in the survivability of *Salmonella*

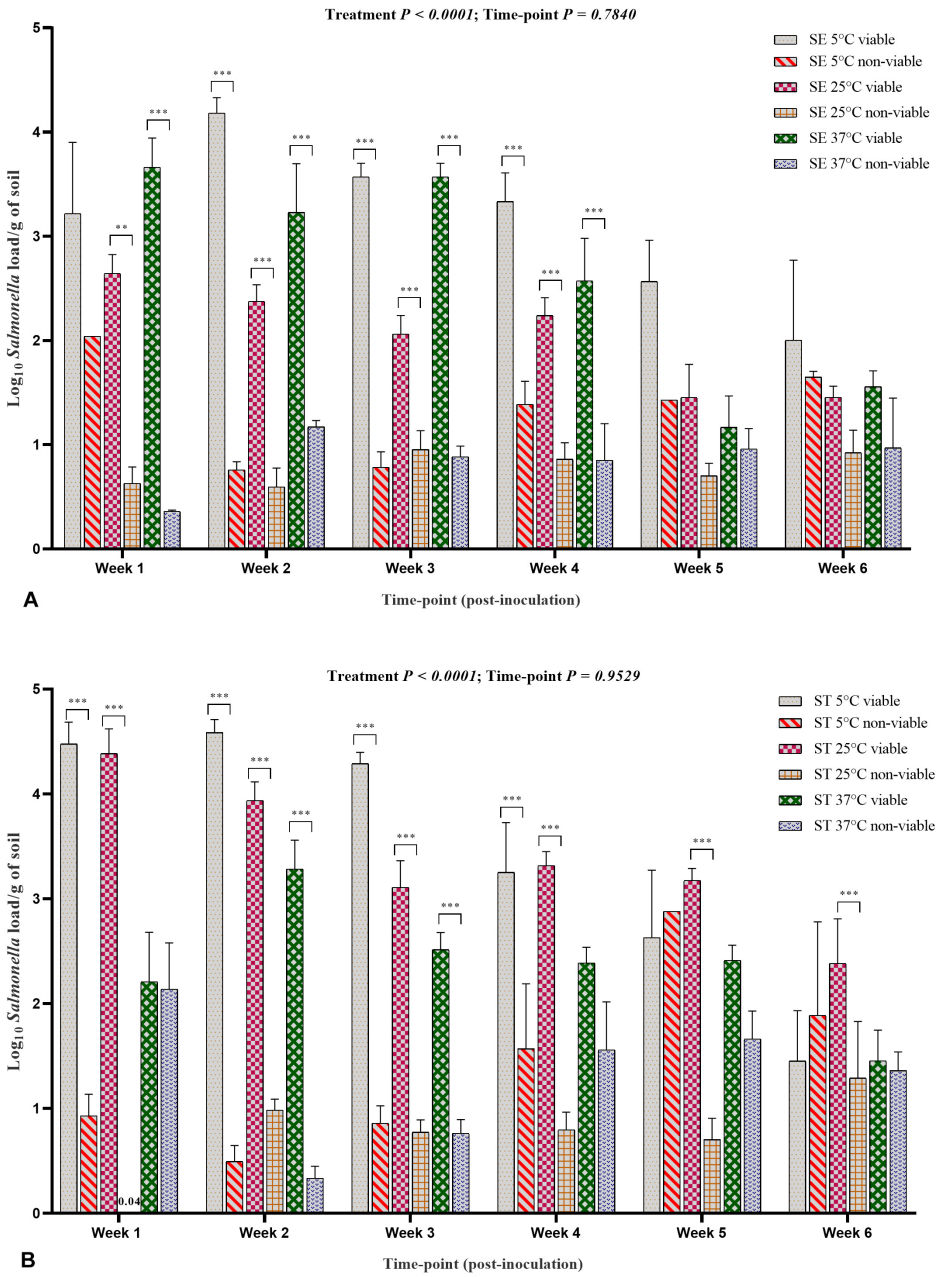


FIGURE 8 | Viable and non-viable load of *Salmonella* Enteritidis (SE) and *Salmonella* Typhimurium (ST) in soil samples. **(A)** SE load. **(B)** ST load. The samples were incubated at 5, 25, and 37°C and processed at weekly intervals for PMAxx^{TM}-based qPCR. Asterisks (***) show significant difference ($P < 0.0001$) between the viable and non-viable load of *Salmonella* affected by each incubation temperature at each sampling timepoint.

shows that the survival of *Salmonella* in the soil in dry and hot summer might be lower due to hot temperatures and low moisture levels. The decrease in the moisture level has been proven as a factor responsible for decline in viable count (Marsh et al., 1998). Nevertheless, our data showed that compared to the *Salmonella* Typhimurium, the *Salmonella* Enteritidis viable load declined at a significantly higher rate when the moisture level and water activity declined. The findings were consistent when the trial was repeated. A higher proportion of positive samples for *Salmonella* Typhimurium compared with *Salmonella* Enteritidis from the enriched samples further proved that the viability of *Salmonella* Enteritidis strain was more susceptible to changes in temperature, water activity and moisture level. A positive correlation between the CFU and PMAxx^{TM}-based qPCR data obtained in the current study is in agreement with the findings of Shah et al. (2019) who quantified the viable count of *Salmonella* Newport through a culture method and PMA-qPCR from heat treated poultry amended soil.

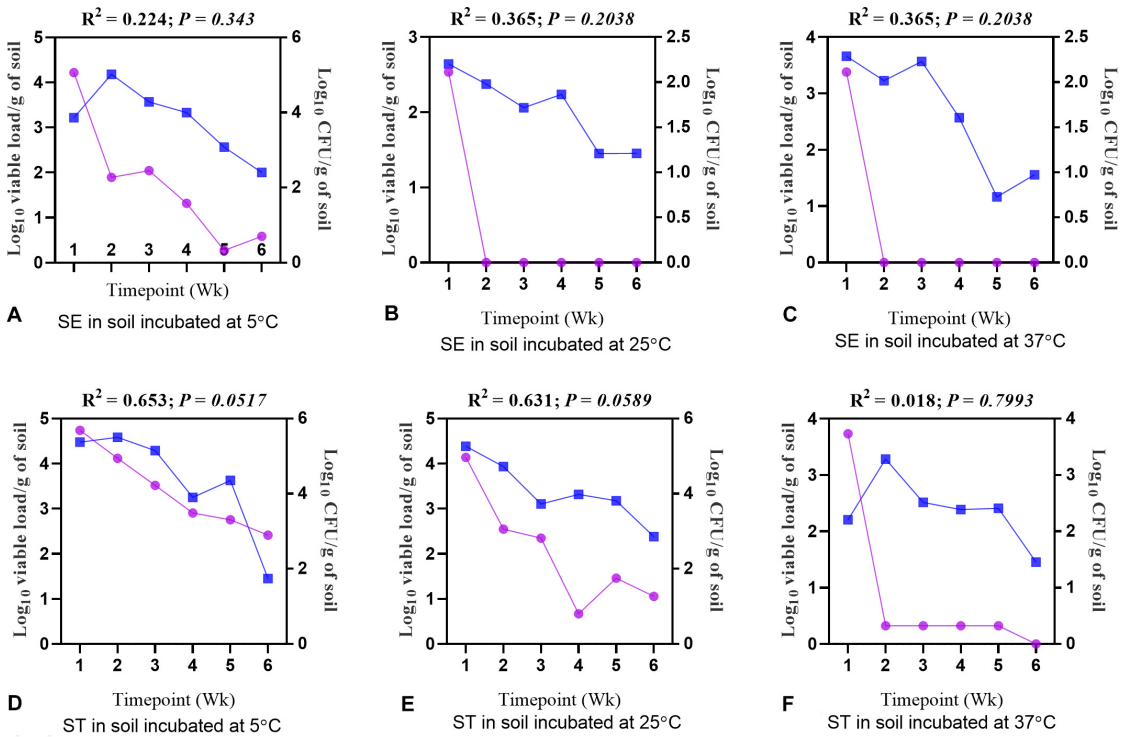


FIGURE 9 | Correlation between $\log_{10}$ CFU/g and $\log_{10}$ viable load of *Salmonella*. **(A)** Soil inoculated with *Salmonella* Enteritidis (SE) and incubated at 5°C. **(B)** Soil inoculated with SE and incubated at 25°C. **(C)** Soil inoculated with SE and incubated at 37°C. **(D)** Soil inoculated with *Salmonella* Typhimurium (ST) and incubated at 5°C. **(E)** Soil inoculated with ST and incubated at 25°C. **(F)** Soil inoculated with ST and incubated at 37°C. Blue lines represent $\log_{10}$ viable load, while purple lines represent $\log_{10}$ CFU per gram of soil.

A sharp decline in the survivability of both the serovars after 1 week of incubation at 37°C was due to a significantly higher loss of moisture and water activity at this temperature. In the current study, keeping the moisture level around 15% was to simulate real field conditions where the temperature in summer in South Australia can reach up to 45°C with moisture level quite low. A higher moisture content favors the survivability of bacteria (Cools et al., 2001), where *Salmonella* Newport can survive in soil for longer periods (Shah et al., 2019). The *Salmonella* Enteritidis strain used in this study was isolated from a commercial egg farm. Recently, several *Salmonella* Enteritidis infected commercial free range egg flocks were culled in Australia because *Salmonella* Enteritidis infection in poultry is categorized as exotic and notifiable. Our previous epidemiological investigation suggested that *Salmonella* Typhimurium was persistent in the range area of free range farms during winter (Gole et al., 2017). After *Salmonella* Enteritidis infection of the flocks, the poultry sheds are typically cleaned but, it is challenging to clean the ranging area of a free range poultry farm. The persistence of *Salmonella* Enteritidis or *Salmonella* Typhimurium in range area can result in reinfection of current flock or infection of the newly introduced flock. The assay developed in this study can be used for the rapid detection of *Salmonella* Enteritidis and *Salmonella* Typhimurium from soil samples and can provide useful information to the industry in making decisions on restocking flocks after mass culling. The data on the survival of *Salmonella* Typhimurium and *Salmonella* Enteritidis in soil samples are useful for mitigating the risks on farms and also for developing on farm biosecurity measures to avoid future on farm *Salmonella* outbreaks. The reduction of *Salmonella* in food animals and in a farm environment is essential to reduce human infection. This study was conducted on soil samples collected from a small number of egg farms located in the South Australia region, therefore; further studies are required to investigate the effects of different soil types on the survival of *Salmonella*. Soil type and temperature can influence the persistence of *Salmonella* (Garcia et al., 2010; Underthun et al., 2018). For the persistence of *Salmonella* in soil, humidity has been identified as an important meteorological factor (Hwang et al., 2020).

Salmonella can form biofilm when exposed to harsh environmental conditions (De Oliveira et al., 2014) such as soil. Biofilm is an assemblage of microorganism associated with the cell surface, which on the one hand can be the attachment to the surface of the materials in the poultry production system or on the other hand, the ability of the biofilm has been confirmed as one of the reasons that the foodborne pathogen can exist in the system continuously (Wang et al., 2013). Our study revealed that when the soil was inoculated with *Salmonella* and incubated at 5°C, *Salmonella* survived better compared with the samples incubated at higher temperature due to moisture level and water activity driven changes. This finding is in agreement

with the previous report (Garcia et al., 2010). The possible reason might be that the biofilm was formed as a barrier to protect bacteria in a harsh environment (De Oliveira et al., 2014), however; further studies are necessary to confirm this hypothesis. *Salmonella* can be exposed to extreme temperature fluctuations including freezing and thawing. Freezing and thawing cycle can significantly affect bacterial cell viability and induce viable but non-culturable (VBNC) state with compromised ability to form a biofilm (Rocard et al., 2018). In this experiment, *Salmonella* Enteritidis and *Salmonella* Typhimurium remained undetected by culture method from week 5 onward, but the qPCR assay was able to detect the viable bacteria from the soil at week 6. This suggests that the low moisture level and high temperature can also induce the VBNC state of *Salmonella*.

The PMAxx^{TM}-based qPCR data showed that the PMA penetrated through the compromised cell membrane of *Salmonella* thereby inhibiting the DNA amplification through qPCR. The positive correlation graphs between the CFU and viable count data at 5, 25, and 37°C for *Salmonella* Enteritidis and *Salmonella* Typhimurium showed that the optimized assay correlated well with the culture method. The viable counts obtained through the PMAxx^{TM}-based qPCR from the soil sampled at weekly intervals showed that the optimized qPCR was sensitive in detecting the load of *Salmonella* compared to the culture method.

In conclusion, the PMAxx^{TM}-based qPCR optimized in this study was sensitive in determining a load of *Salmonella* in soil samples. Both the culture based method and the optimized qPCR data showed that the *Salmonella* Typhimurium survived better in soil than *Salmonella* Enteritidis. Future studies would be able to use this optimized assay in conditions, such as the efficacy of sanitizers against *Salmonella* in poultry production. In poultry and food industries, the quantification of viable *Salmonella* from feces, dust, egg, meat, food products, or other complex mixture of inorganic materials could be tested through our optimized method.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

FUNDING

This study was supported by the School of Animal and Veterinary Sciences, The University of Adelaide.

ACKNOWLEDGMENTS

Help from Sonali Deshmukh and Qixing Lin during soil samples processing for moisture content and water activity measurements and DNA extraction is highly acknowledged. We would like to acknowledge Dr. Cheryl Jenkins, Elizabeth Macarthur Agricultural Institute, NSW Department of Primary Industries for providing *Salmonella* Enteritidis strain for this study.

REFERENCES

Akiba, M., Kusumoto, M., and Iwata, T. (2011). Rapid identification of *Salmonella enterica* serovars, typhimurium, choleraesuis, infantis, hadar, enteritidis, dublin and gallinarum, by multiplex PCR. *J. Microbiol. Methods* 85, 9–15. doi: 10.1016/j.mimet.2011.02.002

Banihashemi, A., Van Dyke, M. I., and Huck, P. M. (2012). Long-amplicon propidium monoazide-PCR enumeration assay to detect viable *Campylobacter* and *Salmonella*. *J. Appl. Microbiol.* 113, 863–873. doi: 10.1111/j.1365-2672.2012.05382.x

Barbau-Piednoir, E., Mahillon, J., Pillyser, J., Coucke, W., Roosens, N. H., and Botteldoorn, N. (2014). Evaluation of viability-qPCR detection system on viable and dead *Salmonella serovar* Enteritidis. *J. Microbiol. Methods* 103, 131–137. doi: 10.1016/j.mimet.2014.06.003

Bellehumeur, C., Boyle, B., Charette, S. J., Harel, J., L'homme, Y., Masson, L., et al. (2015). Propidium monoazide (PMA) and ethidium bromide monoazide (EMA) improve DNA array and high-throughput sequencing of porcine reproductive and respiratory syndrome virus identification. *J. Virol. Methods* 222, 182–191. doi: 10.1016/j.jviromet.2015.06.014

Cools, D., Merckx, R., Vlassak, K., and Verhaegen, J. (2001). Survival of *E. coli* and *Enterococcus* spp. derived from pig slurry in soils of different texture. *Appl. Soil Ecol.* 17, 53–62. doi: 10.1016/s0929-1393(00)00133-5

De Oliveira, D. C. V., Fernandes Junior, A., Kaneno, R., Silva, M. G., Araujo Junior, J. P., Silva, N. C. C., et al. (2014). Ability of *Salmonella* spp. to produce biofilm is dependent on temperature and surface material. *Foodborne Pathog. Dis.* 11, 478–483. doi: 10.1089/fpd.2013.1710

Dinu, L.-D., and Bach, S. (2013). Detection of viable but non-culturable *Escherichia coli* O157: H7 from vegetable samples using quantitative PCR with propidium monoazide and immunological assays. *Food Control* 31, 268–273. doi: 10.1016/j.foodcont.2012.10.020

Ford, L., Moffatt, C. R., Fearnley, E., Miller, M., Gregory, J., Sloan-Gardner, T. S., et al. (2018). The epidemiology of *Salmonella enterica* outbreaks in Australia, 2001–2016. *Front. Sustainab. Food Syst.* 2:86. doi: 10.3389/fsufs.2018.00086

Garcia, R., Baelum, J., Fredslund, L., Santorum, P., and Jacobsen, C. S. (2010). Influence of temperature and predation on survival of *Salmonella enterica* serovar Typhimurium and expression of invA in soil and manure-amended soil. *Appl. Environ. Microbiol.* 76, 5025–5031. doi: 10.1128/aem.00628-10

Gole, V. C., Woodhouse, R., Caraguel, C., Moyle, T., Rault, J. L., Sexton, M., et al. (2017). Dynamics of *Salmonella* shedding and welfare of hens in free-range egg production systems. *Appl. Environ. Microbiol.* 83:e03313-16.

Hwang, D., Rothrock, M. J. Jr., Pang, H., Guo, M., and Mishra, A. (2020). Predicting *Salmonella* prevalence associated with meteorological factors in pastured poultry farms in southeastern United States. *Sci. Total Environ.* 713:136359. doi: 10.1016/j.scitotenv.2019.136359

Jäger, T., Alexander, J., Kirchen, S., Dötsch, A., Wieland, A., Hiller, C., et al. (2018). Live-dead discrimination analysis, qPCR assessment for opportunistic pathogens, and population analysis at ozone wastewater treatment plants. *Environ. Pollut.* 232, 571–579. doi: 10.1016/j.envpol.2017.09.089

Kärenlampi, R., Rautelin, H., Schönberg-Norio, D., Paulin, L., and Hänninen, M.-L. (2007). Longitudinal study of Finnish *Campylobacter jejuni* and *C. coli* isolates from humans, using multilocus sequence typing, including comparison with epidemiological data and isolates from poultry and cattle. *Appl. Environ. Microbiol.* 73, 148–155. doi: 10.1128/aem.01488-06

Kibbee, R. J., and Örmeci, B. (2017). Development of a sensitive and false-positive free PMA-qPCR viability assay to quantify VBNC *Escherichia coli* and evaluate disinfection performance in wastewater effluent. *J. Microbiol. Methods* 132, 139–147. doi: 10.1016/j.mimet.2016.12.004

Li, B., and Chen, J.-Q. (2013). Development of a sensitive and specific qPCR assay in conjunction with propidium monoazide for enhanced detection of live *Salmonella* spp. in food. *BMC Microbiol.* 13:273. doi: 10.1186/1471-2180-13-273

Li, F., Li, F., Chen, B., Zhou, B., Yu, P., Yu, S., et al. (2017). Sextuplex PCR combined with immunomagnetic separation and PMA treatment for rapid detection and specific identification of viable *Salmonella* spp., *Salmonella enterica* serovars Paratyphi B, *Salmonella* Typhimurium, and *Salmonella* ENTERITIDIS in raw meat. *Food Control* 73, 587–594. doi: 10.1016/j.foodcont.2016.09.009

Li, H., Xin, H., and Li, S. F. Y. (2015). Multiplex PMA–qPCR Assay with internal amplification control for simultaneous detection of viable *Legionella pneumophila*, *Salmonella typhimurium*, and *Staphylococcus aureus* in environmental waters. *Environ. Sci. Technol.* 49, 14249–14256. doi: 10.1021/acs.est.5b03583

Liang, N., Dong, J., Luo, L., and Li, Y. (2011). Detection of viable *Salmonella* in lettuce by propidium monoazide real-time PCR. *J. Food Sci.* 76, M234–M237.

Marsh, P., Morris, N. Z., and Wellington, E. M. (1998). Quantitative molecular detection of *Salmonella typhimurium* in soil and demonstration of persistence of an active but non-culturable population. *FEMS Microbiol. Ecol.* 27, 351–363. doi: 10.1111/j.1574-6941.1998.tb00551.x

Milazzo, A., Giles, L. C., Zhang, Y., Koehler, A. P., Hiller, J. E., and Bi, P. (2016). The effect of temperature on different *Salmonella* serotypes during warm seasons in a Mediterranean climate city, Adelaide, Australia. *Epidemiol. Infect.* 144, 1231–1240. doi: 10.1017/s0950268815002587

Nocker, A., Cheung, C.-Y., and Camper, A. K. (2006). Comparison of propidium monoazide with ethidium monoazide for differentiation of live vs. dead bacteria by selective removal of DNA from dead cells. *J. Microbiol. Methods* 67, 310–320. doi: 10.1016/j.mimet.2006.04.015

Rocard, J. M., Asadishad, B., Samonte, P. R. V., Ghoshal, S., and Tufenkji, N. (2018). Natural freeze-thaw cycles may increase the risk associated with *Salmonella* contamination in surface and groundwater environments. *Water Res. X* 1:100005. doi: 10.1016/j.wroa.2018.10.002

Semenov, A. V., Van Overbeek, L., and Van Bruggen, A. H. C. (2009). Percolation and survival of *Escherichia coli* O157: H7 and *Salmonella enterica* serovar Typhimurium in soil amended with contaminated dairy manure or slurry. *Appl. Environ. Microbiol.* 75, 3206–3215. doi: 10.1128/aem.01791-08

Shah, M. K., Bradshaw, R., Nyarko, E., Handy, E. T., East, C., Millner, P. D., et al. (2019). *Salmonella enterica* in soils amended with heat-treated poultry pellets survived longer than bacteria in unamended soils and more readily transferred to and persisted on spinach. *Appl. Environ. Microbiol.* 85:e00334-19.

Underthun, K., De, J., Gutierrez, A., Silverberg, R., and Schneider, K. R. (2018). Survival of *Salmonella* and *Escherichia coli* in two different soil types at various moisture levels and temperatures. *J. Food Prot.* 81, 150–157. doi: 10.4315/0362-028x.jfp-17-226

Wang, H., Ding, S., Wang, G., Xu, X., and Zhou, G. (2013). In situ characterization and analysis of *Salmonella* biofilm formation under meat processing environments using a combined microscopic and spectroscopic approach. *Int. J. Food Microbiol.* 167, 293–302. doi: 10.1016/j.ijfoodmicro.2013.10.005

Wang, L., and Mustapha, A. (2010). EMA-real-time PCR as a reliable method for detection of viable *Salmonella* in chicken and eggs. *J. Food Sci.* 75, M134–M139.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Zhang, Khan and Chousalkar. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 28 September 2020
doi: 10.3389/fmicb.2020.577707

Methicillin-Resistant *Staphylococcus aureus* (MRSA) in Slaughter Houses and Meat Shops in Capital Territory of Pakistan During 2018–2019

Asma Sadiq[1], Maroof Samad[2,3], Saddam[4], Nosheen Basharat[1], Shahid Ali[5], Roohullah[5], Zubaida Saad[6], Allah Nawaz Khan[7], Yasin Ahmad[5], Alam Khan[5] and Jadoon Khan[1,5]**

[1] Molecular Virology Laboratory, Department of Biosciences, COMSATS University Islamabad, Islamabad, Pakistan, [2] COMWAVE Institute Islamabad, Islamabad, Pakistan, [3] Pakistan Institute of Medical Sciences (PIMS) Islamabad, Islamabad, Pakistan, [4] Department of Microbiology, Abdul Wali Khan University Mardan, Mardan, Pakistan, [5] Department of Microbiology, Quaid-i-Azam University (QAU), Islamabad, Pakistan, [6] Department of Zoology, Women University Mardan, Mardan, Pakistan, [7] Department of Life Sciences, Abasyn University, Islamabad, Pakistan

OPEN ACCESS

Edited by:
David Rodriguez-Lazaro, University of Burgos, Spain

Reviewed by:
Anusak Kerdsin, Kasetsart University, Thailand
Georgios Efthimiou, University of Hull, United Kingdom

***Correspondence:**
Asma Sadiq
asma_sadiq17@hotmail.com
Jadoon Khan
jadoonkhan@bs.qau.edu.pk

Specialty section:
This article was submitted to Food Microbiology, a section of the journal Frontiers in Microbiology

Received: *29 June 2020*
Accepted: *07 September 2020*
Published: *28 September 2020*

Citation:
Sadiq A, Samad M, Saddam, Basharat N, Ali S, Roohullah, Saad Z, Khan AN, Ahmad Y, Khan A and Khan J (2020) Methicillin-Resistant Staphylococcus aureus (MRSA) in Slaughter Houses and Meat Shops in Capital Territory of Pakistan During 2018–2019. Front. Microbiol. 11:577707. doi: 10.3389/fmicb.2020.577707

Livestock-associated methicillin-resistant *Staphylococcus aureus* (LA-MRSA) is a major concern in many parts of the world, including Pakistan. The aim of this study was to investigate the prevalence of MRSA in slaughterhouses and meat shops in Rawalpindi-Islamabad, Pakistan, 2018–2019. A total of 300 samples were collected: 40 from each of working area, tools (knives, hooks), butcher hands and beef, 30 from each of chicken and mutton, 20 from each of nasal and rectal swabs. *S. aureus* was phenotypically identified by performing gram staining and biochemical tests. 150 of the 300 samples were confirmed to be *S. aureus* by phenotypic identification. MRSA was identified among *S. aureus* positive isolates by performing disk diffusion test and by detecting *S. aureus*-specific genes such as 16s rRNA, *nuc*, *mecA*, *spa*, and *coa*. Out of 150 isolates 96 (63%) showed resistance to antibiotic cefoxitin, known as a potential marker for detecting MRSA. While all 150 isolates have shown complete resistance to the four antibiotics neomycin, methicillin, ciprofloxacin and tetracycline. The *nuc* and 16s rRNA genes were detected in all 150 *S. aureus*-positive isolates and 118 (79%) were confirmed to be MRSA through the detection of the *mecA* gene. MRSA prevalence was highest in chicken (23/30, 77%) followed by beef (25/40, 63%), mutton (15/30, 50%), knives (18/40, 45%), nasal swabs (7/20, 35%), working area (11/40, 28%), rectal swabs (5/20, 25%), hooks (7/40, 18%), and butcher hands (7/40, 18%). 50 MRSA-positive isolates were chosen to identify two virulence factors (*spa* and *coa* gene). Of the 50 MRSA isolates subject to *coa* and *spa* gene typing, 27 (54%) were positive for the *coa* gene and 18 (36%) were positive for the *spa* gene, respectively. To the best of our knowledge, this was the first study on the molecular identification of MRSA in meat samples from Pakistan. High prevalence of MRSA in meat samples demand for implementation of proper hygienic practices and procedures during the slaughtering, transport and marketing of meat and meat products in order to prevent the spread of these bacteria to the human population.

Keywords: LA-MRSA, phenotypic, antibiotic sensitivity, PCR, *mecA* gene 3

INTRODUCTION

Staphylococcus aureus is a Gram-positive, coagulase- positive pathogen belong to the family *Staphylococcaceae*. This is a spherical bacterium about 1 μm in diameter that forms grape-clusters (Lakhundi and Zhang, 2018). *Staphylococcus aureus* (*S. aureus*) is one of the most common microorganisms colonizing the nasal cavity of humans and different animal species (Al-Amery et al., 2019). This may also be found in external body surfaces as either commensal or pathogenic bacteria that can cause multiple infectious diseases (Weese and van Duijkeren, 2010). *S. aureus* has a variety of virulence factors and toxins, frequently responsible for many toxin-diseases including toxic shock syndrome, staphylococcal foodborne diseases (SFDs), and scalded skin syndrome (Okoli et al., 2018). *Staphylococcus* has the potential to establish resistance to broad-spectrum antibiotics in a short period, such as the β-lactam group of antibiotics, aminoglycosides and quinolones commonly used in clinical practice for the treatment of serious infections (Deurenberg et al., 2007). Methicillin-resistant *staphylococcus* (MRSA) strain was first identified in United Kingdom in 1961, continues to be a serious hospital concern for public health (Patricia Jevons, 1961; Lowy, 1998). The United States Centers for Disease Control and Prevention (CDC) reported in July 2002 the first *S. aureus* strain resistant to both vancomycin and methicillin (Graveland et al., 2011). Nonetheless, unusual strains appeared starting in the 1980s, leading to a global spread.

A distinct MRSA lineage, Clonal Complex (CC) 398, was first identified from pigs in The Netherlands and France in food producing animals in Europe, particularly in pig herds, turkeys, veal calves and broiler flocks (Vossenkuhl et al., 2014). The term "Livestock-associated MRSA" (LA-MRSA) has therefore been introduced, considering that livestock form a new and separate reservoir for MRSA (Larsen et al., 2012). In Asian countries, however, the ST9 sequence, a separate genetic lineage, is predominant among MRSA isolates from livestock (Güven Gökmen et al., 2018). Notably, outbreaks of LA-MRSA in hospitals and invasive infections of LA-MRSA in humans are growing (Lewis et al., 2008). Accordingly, LA-MRSA has become an important public health problem that needs close monitoring.

MRSA strains are resistant to all β-lactam antimicrobials by a penicillin-binding protein (PBP2a) that has a weak affinity to all β-lactams. The protein is encoded by the *mecA* gene, which resides on the lactams of a mobile genetic element called the staphylococcal chromosome cassette (SCCmec) (Ivbule et al., 2017). To classify the epidemiological characteristics of MRSA strains and, more importantly, to research the evolution and spread of disease clones, it is appropriate to employ relevant and reproducibility molecular methods with ample discriminative capacity to track changes in time. Most of the methods used for this are pulsed-gel electrophoresis (PFGE), Multilocus Sequence Typing (MLST), staphylococcal cassette chromosome mec (SCCmec), and staphylococcal protein A (*spa*) and staphylococcal *coa* (coagulase) gene (Cookson et al., 2007). Virulence factors of *S. aureus*, *coa*, and *spa* genes, respectively has been shown to be directly linked to pathogenesis and the magnitude of infection (Rathore et al., 2012). Both virulent genes are highly polymorphic and can provide critical information on strain variations.

The widespread use of antimicrobials in animal production is believed to promote the emergence and spread of MRSA because of selection pressure induced by antimicrobials (Pires et al., 2009). Overcrowding in animal husbandry and intensive animal trade can help the rapid spread of MRSA among the farm animals (Guo et al., 2018). LA-MRSA strains were also found on wholesale in raw meat including poultry, beef, veal, and pork (Alt et al., 2011). Recent studies have also found that LA-MRSA may colonize in multiple animals and associated workers (Rinsky et al., 2013). It indicates possible transmission of cross-contamination in the chain during slaughter and processing (Waters et al., 2011). Nevertheless, the scale of this transmission is not well-understood up to now.

Pakistan, as a developing country, suffers primarily from antibiotic resistance, which is a concern not only for Pakistan, but for the entire human/animal population (Ali et al., 2018). Weak steps to monitor infections as well as continued unregulated exposure of humans and animals to antibiotics have contributed to this enormous problem of MRSA development and transmission (Lakhundi and Zhang, 2018). This in effect restricts the treatment options for MRSA infections. Continuous monitoring for MRSA is therefore necessary in any setting by analyzing the characteristics, host specificity, and propagation paths of newer strains.

Detection of staphylococci in meat is often related to poor hygienic practices during processing, shipping, slicing, storage and point of sale by individuals involved in the production process. Therefore, the main objective of this study was to identify *S. aureus* phenotypically and molecularly through amplification of 16S ribosomal RNA, *mecA*, *nuc*, *spa*, and *coa* genes from selected slaughter houses and meat shops in twin cities Rawalpindi, Islamabad, Pakistan during 2018–2019.

MATERIALS AND METHODS

Ethical Approval

The ethical approval of the study was taken from ethical review committee of Sarhad University, Peshawar and Pakistan Institute of Medical Sciences (PIMS), Islamabad, Pakistan. The institutional committees approved the experiments carried out for the current research described in the "Materials and Methods" section of the manuscript. All procedures have been carried out in compliance with the relevant regulations and standards.

Sample Collection and Processing

Three hundred samples from various slaughterhouses and meat stores in different areas of Rawalpindi, Islamabad, Pakistan were collected during 2018–2019. The sample size for this study was calculated using prevalence formula in a software N-Query Advisory (STATCON Gmbh, Germany, Version 7.0) (Pourhoseingholi et al., 2013). A total of 300 samples were collected: 40 from each of working area, tools (knives, hooks), butcher hands and beef, 30 from each of chicken and mutton, 20 from each of nasal and rectal swabs. Sterile cotton swabs were first

TABLE 1 | List of primers used and reaction conditions for each gene amplified in this study.

List of primers			
Primer name	**Sequence (5′-3′)**	**Product size (bp)**	**References**
16S-F	5′-GTGCCAGCAGCCGCGGTAA-3′	876	Salisbury et al., 1997
16S-R	5′-AGACCCGGGAACGTATTCAC-3′		
mecA-F	5′-AAA ATC GAT GGT AAA GGT TGG C-3′	533	Murakami et al., 1991
mecA-R	5′AGT TCT GCA GTA CCG GAT TTG C3′		
nuc-F	5′-GCG ATT GAT GGT GAT ACG GTT-3′	270	Murakami et al., 1991
nuc-R	5′-AGC CAA GCC TTG ACG AAC TAA AGC-3′		
coa-F	5′-ATAGAGATGCTGGTACAGG-3′	680, 891	Hookey et al., 1998
coa-R	5′-GCTTCCGATTGTTCGATGC-3′		
spa-F (x-region)	5′-CAAGCACCAAAAGAGGAA-3′	100, 200	Bhati et al., 2016
spa-R (x-region)	5′CACCAGGTTTAACGACAT3′		

dipped in buffer peptone water (Oxoid, United Kingdom) rubbed horizontally and then vertically on the selected materials (Adugna et al., 2013). Then swabs were placed with proper labeling in airtight zip bags and stored at −20°C. Collected samples were processed through standard operating procedures. The raw meat samples (1 g) were first added to 10 ml tryptone water (10 g/l tryptone and 5 g/l NaCl) and properly mixed. This solution was used as inoculum for making serial dilutions. This makes the first dilution i.e., 10^{-1}. The dilutions 10^{-3}, 10^{-4}, and 10^{-5} were used for inoculation. The swab samples were placed in peptone water-containing falcon tubes and then centrifuged at 5,000 rpm until all material on swab is dissolved fully in peptone water.

Isolation and Phenotypic Identification of *S. aureus*

Samples were initially grown on selective medium mannitol salt agar (MSA) for the growth of *S. aureus* using cotton swabs. A sample of 100 μl was transferred to the Mannitol salt agar (Thermo Fisher Scientific, United States) medium and distributed uniformly. The plates were incubated at 37°C overnight. Isolated colonies that showed fermentation on MSA medium were subcultured again on the MSA in order to get pure cultures. Bacterial isolates were cultured on blood agar (Sigma-Aldrich, Germany) medium containing 5% heparin free sheep blood added to the blood agar base after autoclaving and cooling to 50°C. *S. aureus* have the ability to hemolyse sheep blood and shows alpha (α) hemolysis observed in the form of clear zones around colonies. Until observations, plates were incubated at 37°C for 24–48 h. *Staphylococci* have been detected phenotypically using standard methods specific to the detection of enzymes and certain biochemical processes. These techniques include culturing, gram staining, catalase test, coagulase and DNase test (O'Brien et al., 2012). Following the phenotypic study of *S. aureus*, isolates were stored in a nutrient broth with an addition of 20% glycerol to prevent any frost shock due to crystal formation in bacterial cells.

TABLE 2 | The prevalence of *S. aureus* among 300 samples from different sources.

Isolate ID	Source	Total number of samples (n)	*S. aureus* positivity rate n (%)
A	Knives	40	22 (55%)
B	Hooks	40	14 (35%)
C	Working Area	40	16 (40%)
D	Butcher hands	40	12 (30%)
E	Nasal swab	20	9 (45%)
F	Rectal swab	20	7 (35%)
G	Chicken	30	25 (83%)
H	Mutton	30	15 (50%)
I	Beef	40	30 (75%)

Antimicrobials Susceptibility Test

To confirm the antibiotic susceptibility of bacterial isolates to specific antibiotics, Kirby Bauer or disk diffusion test was used (Bauer et al., 1996). Mueller Hinton agar medium (MHA) was used for disk diffusion test (Oxoid Ltd., United Kingdom). The antibiotic disks that were selected for the antibiotic susceptibility test were (novobiocin 5 μg, cefoxitin 30 μg, neomycin 30 μg, methiciline 5 μg, amoxycilin 30 μg, erythromycin 15 μg, gentamycin 5 μg, vancomycin 30 μg, ciprofloxacin 5 μg, and tetracycline 30 μg), respectively. These disks were carefully placed on the Petri dishes to avoid any environmental contamination. The results were noted after overnight incubation at 37°C according to the Clinical and Laboratory Standards Institute (Clinical and Laboratory Standards Institute [CLSI], 2019).

PCR Detection of 16s rRNA, *nuc*, and *mecA* genes of *S. aureus*

A molecular analysis of phenotypically identified *S. aureus* isolates was performed by amplification of the *Staphylococci* 16s rRNA gene, the species-specific *nuc* gene and the *mecA*-resistant gene. The genomic DNA was extracted using the DNA extraction method CTAB (Cetyltrimethylammonium bromide) (Maristela Oliveira Lara et al., 2018). Before molecular analysis, the extracted DNA was analyzed using gel electrophoresis (1%). Molecular typing was carried out for one MRSA-like colony per positive sample. The total volume of reaction mixture was 20 μl that contained 1 μl each of forward and reverse primer, 10 μl of

TABLE 3 | Susceptibility to antimicrobials among 150 isolates of *S. aureus*.

S. No.	Antimicrobial agent	Disk content (μg)	Susceptible no. of isolates (%)	Resistant no. of isolates (%)	Intermediate no. of isolates (%)
1	Novobiocin	5	18 (12%)	78 (52%)	54 (36%)
2	Cefoxitin	30	30 (20%)	95 (63%)	25 (17%)
3	Neomycin	30	0	150 (100%)	0
4	Methicillin	5	0	150 (100%)	0
5	Amoxicillin	30	15 (10%)	105 (70%)	30 (20%)
6	Erythromycin	15	12 (8%)	114 (76%)	24 (16%)
7	Gentamycin	5	9 (6%)	123 (82%)	18 (12%)
8	Vancomycin	30	9 (6%)	120 (80%)	21 (14%)
9	Ciprofloxacin	5	0	150 (100%)	0
10	Tetracycline	30	0	150 (100%)	0

Wizepure 2X PCR master mix (Wiz biosolutions, South Korea), 7 μl of molecular grade water and 1 μl of template DNA. Gene specific primers were used as described by Murakami et al. (1991) and Salisbury et al. (1997) (**Table 1**). The PCR conditions set for16s rRNA and *mecA* gene amplification reaction were, an initial denaturation at 95°C for 5 min, 29 cycles of amplification (denaturation at 95°C for 30 s, annealing at 55°C for 30 s, extension at 72°C for 40 s and final extension at 72°C for 2 min). The PCR conditions set for *nuc* gene amplification reaction were, an initial denaturation 95°C for 5 min, 34 cycles of amplification (denaturation at 95°C for 30 s, annealing at 53°C for 30 s, extension 72°C for 30 s, final extension 72°C for 2 min). For product visualization, 8 μl of DNA sample was mixed with 2 μl of 6× loading dye. Then 1.5% of gel was prepared in 1× TAE buffer and 100-bp DNA ladder (Bioron, Cat. No. 304105, Germany) with a voltage of 100V for 20–25 min was used to validate the amplified product. After the gel electrophoresis was completed, the gel results were visualized using the Gel doc system (Thomas scientific, United States).

coa and *spa* Genes Typing of MRSA

The amplification of *coa* (coagulase) and x region of *Staphylococcus* protein A (*spa*) genes is carried out for 50 selected MRSA isolates based on the identification of *mecA* (resistant) gene. DNA extraction was performed using the QIAamp DNA Mini Kit (Qiagen, Germany), with manufacturer instructions. Gene-specific primers were used as described (Hookey et al., 1998; Bhati et al., 2016). The sequences of primers for *coa* and *spa* genes are given in **Table 1**. The total volume of the reaction mixture was 20 μl, including the PCR Master mix (Thermo Fisher Scientific, United States), distilled water, 2 μl of the DNA sample and 2 μl of each of forward and reverse primer. The PCR cycling conditions used for *coa* gene were: an initial denaturation at 94°C for 45 s, 30 cycles of amplification (denaturation at 94°C for 20 s, annealing at 57°C for 15 s, extension at 70°C for 15 s and final extension at 75°C for 2 min). The PCR cycling conditions used for *spa* gene were: 34 cycles of amplification (denaturation at 94°C for 60 s, annealing at 55°C for 60 s, extension at 70°C for 60 s and final extension at 72°C for 5 min). For product visualization, 8 μl of DNA sample was mixed with 2 μl of 6× loading dye. Then 1.2% of gel was prepared in 1× TAE buffer having ethidium bromide (0.5 μg/ml) and 100-bp DNA ladder (Bioron, Cat. No. 304105, Germany) was used to validate the amplified product. The voltage used was 80 V for 1 h. After the gel electrophoresis was completed, the results of the gel were visualized using the Gel doc system (Thomas scientific, United States).

RESULTS

Phenotypic Identification Results

A total of 300 samples were grown in the mannitol salt agar (MSA) medium. Of the 300 samples, 150 fermented mannitol and showed yellow colonies on MSA (**Supplementary Figure S1**). In order to obtain pure colonies, *S. aureus* positive samples were subcultured further on MSA and selected for other biochemical tests. All 150 samples were grown in the blood agar, and all showed beta-hemolysis activity (**Supplementary Figure S2**). Gram staining was carried out for all 150 isolates. Cluster of cocci, purple in color showed a definite pattern under the light microscope at the 100× lens (**Supplementary Figure S3**). Further catalase, coagulase and DNase tests were conducted on all 150 isolates. All of the tests showed positive results for *S. aureus* (**Supplementary Figures S4–S6**).

Prevalence of *S. aureus* Based on Phenotypic Identification Results

The prevalence of *S. aureus* differed among sample sources and specimen types based on the phenotypic findings. Out of a total of 300 samples, 150 samples were positive for *S. aureus* with an overall prevalence of 50%. The highest prevalence of *S. aureus* was observed for chicken (25/30, 83%) followed by beef (30/40, 75%), knives (22/40, 55%), mutton (15/30, 50%), nasal swabs (9/20, 45%), and working area (16/40, 40%). The least prevalence of *S. aureus* was found in rectal swabs (7/20, 35%), hooks (14/40, 35%), and butcher hands (12/40, 30%), respectively (**Table 2**).

Antimicrobial Resistance Results

One hundred and fifty *S. aureus* isolates were screened for different antimicrobials using a disk diffusion technique. **Table 3** shows the percentage of resistant, intermediate and susceptible isolates for 10 different antibiotic disks. The isolates were

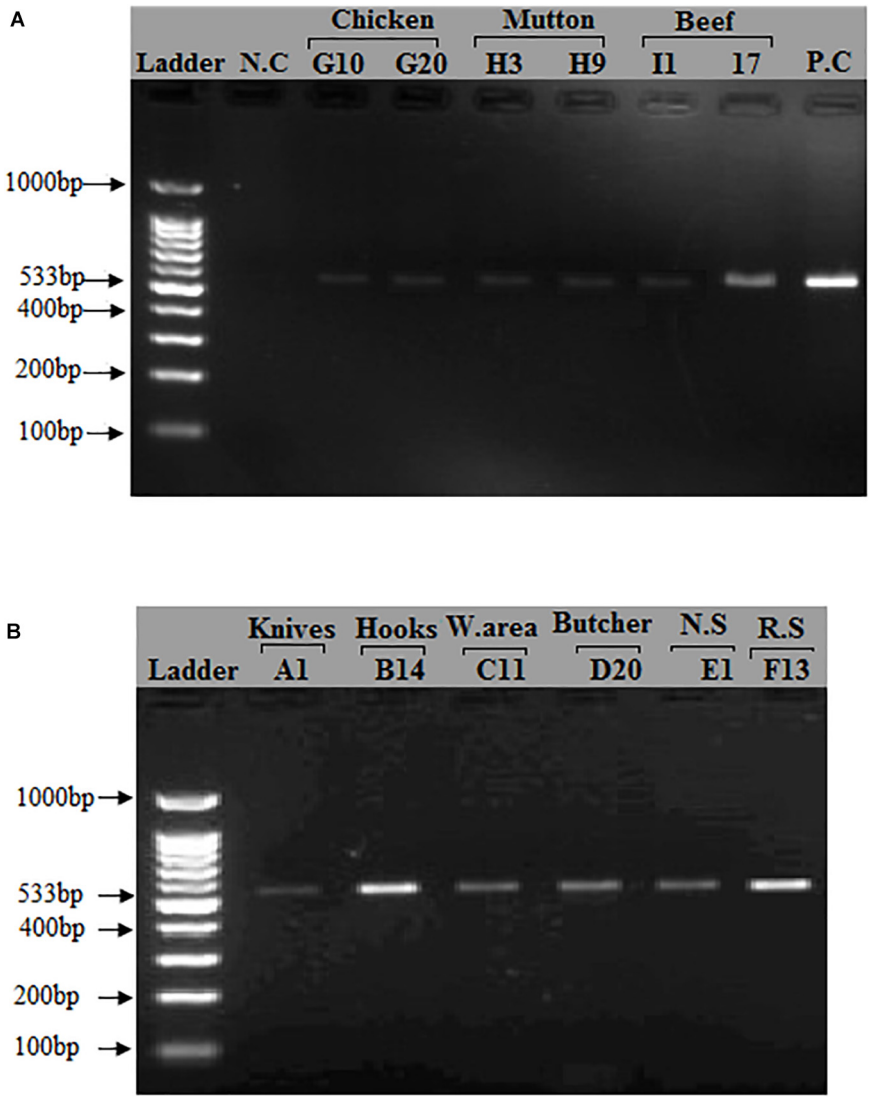


FIGURE 1 | Detection of mecA gene in *S. aureus* isolates: A band (533 bp) corresponding to amplified region of mecA gene. (A) First lane = 100 bp ladder, 2nd Lane = Negative control, Lane# 9 = Positive control (*S. aureus* ATCC 33591), Lane# 3–8 = PCR products of mecA gene from meat samples (Beef, Chicken, Mutton). (B) First lane = 100 bp ladder, Lane# 2–7 = PCR products of mecA gene from different sources (Knives, Hooks, Working area, Butcher hands, Nasal swabs, Rectal swabs, Beef, Chicken, and Mutton).

completely resistance to the neomycin, methicillin, ciprofloxacin and tetracycline and showed 63 and 52% resistance to cefoxitin and novobiocin, respectively. It was noted that 70, 76, 82, and 80% of the strains were also resistant to amoxycillin, erythromycin, gentamycin and vancomycin, respectively (**Supplementary Figure S7**). The antibiotic susceptibility and intermediate susceptibility trend differed among 10 drugs. All isolates showed 20, 12, 10, and 8% susceptibility to cefoxitin, novobiocin, amoxicillin and Erythromycin, respectively. While, gentamycin and vancomycin showed equal (6%) susceptibility. It was also observed that 36, 17, 20, 16, 12, and 14% of the isolates showed intermediate susceptibility toward novobiocin, cefoxitin, amoxicillin, erythromycin, gentamycin, and vancomycin, respectively.

Distribution of MRSA via Molecular Analysis

After phenotypic validation, all 150 isolates undergo molecular characterization to detect methicillin resistant *S. aureus* on the basis of 16S rRNA, *nuc*, *mecA*, *coa*, and *spa* genes amplification. The extracted DNA was run on agarose gel 1% and results were observed on gel doc (**Supplementary Figure S8**). All 150 *S. aureus* positive isolates were subjected to 16s rRNA, *nuc* gene and *mecA* typing. All isolates were tested positive for 16S rRNA gene and *nuc* genes (**Supplementary Figures S9, S10**). The *nuc* gene primer gave the PCR product equal to 270 bp. The 16S rRNA primers gave PCR products equal to 886 bp. The methicillin resistant *S. aureus* (MRSA) gene (*mec A*) was identified in (118/150, 79%) of the samples. The *mecA* gene

TABLE 4 | Comparison between *mecA* gene PCR results and cefoxitin disk diffusion test among 150 *S. aureus* positive isolates from different sources.

Isolate ID	Source	Total number of *S. aureus* positive isolates (n)	Cefoxitin resistant MRSA n (%)	MRSA positivity rate (*mecA* gene) n (%)
A	Knives	22	13 (33%)	18 (45%)
B	Hooks	14	7 (18%)	7 (18%)
C	Working Area	16	10 (25%)	11 (28%)
D	Butcher hands	12	5 (13%)	7 (18%)
E	Nasal swab	9	2 (10%)	7 (35%)
F	Rectal swab	7	2 (10%)	5 (25%)
G	Chicken	25	19 (63%)	23 (77%)
H	Mutton	15	12 (40%)	15 (50%)
I	Beef	30	25 (63%)	25 (63%)

primer gave the PCR product equal to 533 bp (**Figure 1**). The distribution of MRSA based on the detection of *mecA* and cefoxitin resistance pattern with respect to the specimen type is shown in **Table 4**. The highest prevalence of MRSA was detected in chicken (23/30, 77%) followed by beef (25/40, 63%), mutton (15/30, 50%), knives (18/40, 45%), nasal swabs (7/20, 35%), working area (11/40, 28%), rectal swabs (5/20, 25%), hooks (7/40, 18%), and butcher hands (7/40, 18%). 95 of the 150 isolates were resistant to cefoxitin and the *mecA* gene was found in all cefoxitin-resistant isolates. 30 out of 150 isolates were susceptible to cefoxitin. 23 out of 30 susceptible isolates showed presence of *mecA* gene (**Table 4**).

Results of *coa* and *spa* Genes Typing of MRSA

Of the 50 MRSA isolates subject to *coa* and *spa* gene typing, 27 were positive for the *coa* gene and 18 were positive for the *spa* gene, respectively. PCR amplification of the X region of the *spa* gene generated a single amplicon in each isolate. Two amplicons of different sizes (100 and 200 bp) were produced (**Figure 2**). One hundred bp was more common pattern in isolates numbers G3, G10, G15, G20, G23, H3, H9, H15, I1, I7, I10 and I13, respectively. While, 200 bp was more common pattern in isolate numbers A1, D20, E1, E2, F13 and F18, respectively (**Table 5**). PCR amplification of the *coa* gene yielded a single amplicon in each isolate. Two amplicons of different sizes (681 and 891 bp) were developed (**Figure 3**). Six hundred eighty bp was more common pattern in isolates numbers G1, G3, G10, G15, G20, G23, G27, H3, H9, H15, I1, I7, I10, I13 and I17, respectively. While, 891 bp was more common pattern in isolate numbers A1, A13, A35, D9, D20, E1, E2, E11, E14, F13, F18 and F20, respectively (**Table 5**). The distribution of *spa* gene in different sample types is as follows: knives (1/5, 20%), hooks (0/5, 0%), working Area (0/5, 0%), butcher hands (1/5, 20%), nasal swabs (2/5, 40%), rectal swabs (2/5, 40%), chicken (5/8, 63%), mutton (3/5, 60%), and beef (4/7, 57%). The distribution of *coa* gene in different sample types is as follows: knives (3/5, 60%), hooks (0/5, 0%), working Area (0/5, 0%), butcher hands (2/5, 40%), nasal swabs (4/5, 80%), rectal swabs (3/5, 60%), chicken (7/8, 88%), mutton (3/5, 60%), and beef (5/7, 71.%).

DISCUSSION

The emergence and persistent spread of drug-resistant bacteria has become one of the most daunting problems facing the world today (Hayat et al., 2020). Global antibiotic use in low- and middle-income countries increased by 65% from 2000 to 2015 (Dweba et al., 2019). There is evidence that overuse of antimicrobials in food animals leads to the production of drug-resistant bacterial infections in both animals and humans (Mohsin et al., 2019). Many antimicrobials used in veterinary medicine are also listed in the World Health Organization (WHO) catalog of vitally important antimicrobials in hospital settings (World Health Organization [WHO], 2019). Potential risk of spreading MRSA is through contamination of meat sellers, slaughter house workers and people associated with livestock and poultry.

The present study documented the prevalence of MRSA during 2018–2019 in slaughterhouses and meat shops in Rawalpindi-Islamabad, Pakistan. After phenotypic detection, 150 samples of a total of 300 samples were positive for *S. aureus*, with an overall prevalence of 50%. For confirmation of MRSA all 150 isolates were subjected to a disk diffusion test and further molecular analysis. Ninety-five out of one hundred fifty *S. aureus* isolates were considered MRSA by antimicrobial resistance pattern based on the frequency of resistance to cefoxitin (30 μg) which is known as a potential marker for detecting methicillin resistance (Fernandes et al., 2005; Jain et al., 2008; Pourmand et al., 2014). Similar results have been reported in other studies in China, Sudan, United Kingdom, Greece and Algeria (Sergelidis et al., 2015; Ibrahim et al., 2017; Chaalal et al., 2018; Wu et al., 2018; Anjum et al., 2019). All *S. aureus* strains isolated in the current study showed complete resistance to the neomycin, methicillin, ciprofloxacin and tetracycline. Such results are consistent with studies performed in Ethiopia and Nigeria, which document related patterns in neomycin, methicillin and tetracycline resistance (Iroha et al., 2011; Adugna et al., 2013). A similar study by Igbinosa et al. (2016) in Nigeria in 2016 showed 100% resistance of *S. aureus* isolates to methicillin. Resistance to novobiocin, amoxycillin, erythromycin, gentamycin and vancomycin was 52, 70, 76, 82, and 80%, respectively. A similar research conducted in Egypt revealed 78% of *S. aureus* isolates showed resistance to vancomycin and erythromycin, and 100% of the strains showed resistance to gentamycin (Osman et al., 2016). There are many studies in which *S. aureus* antibiotic-resistant strains have been identified in different food products (Hanson et al., 2011; Basanisi et al., 2017; Rong et al., 2017; Haskell et al., 2018; Pekana and Green, 2018; Wu et al., 2018).

The molecular confirmation of *S. aureus* and detection of methicillin resistant *S. aureus* (MRSA) were conducted via *PCR*. This technique has been used previously for the detection of *S. aureus* and methicillin resistant *S. aureus* (Vossenkuhl et al., 2014; Osman et al., 2016; Vaiyapuri et al., 2019). In this study,

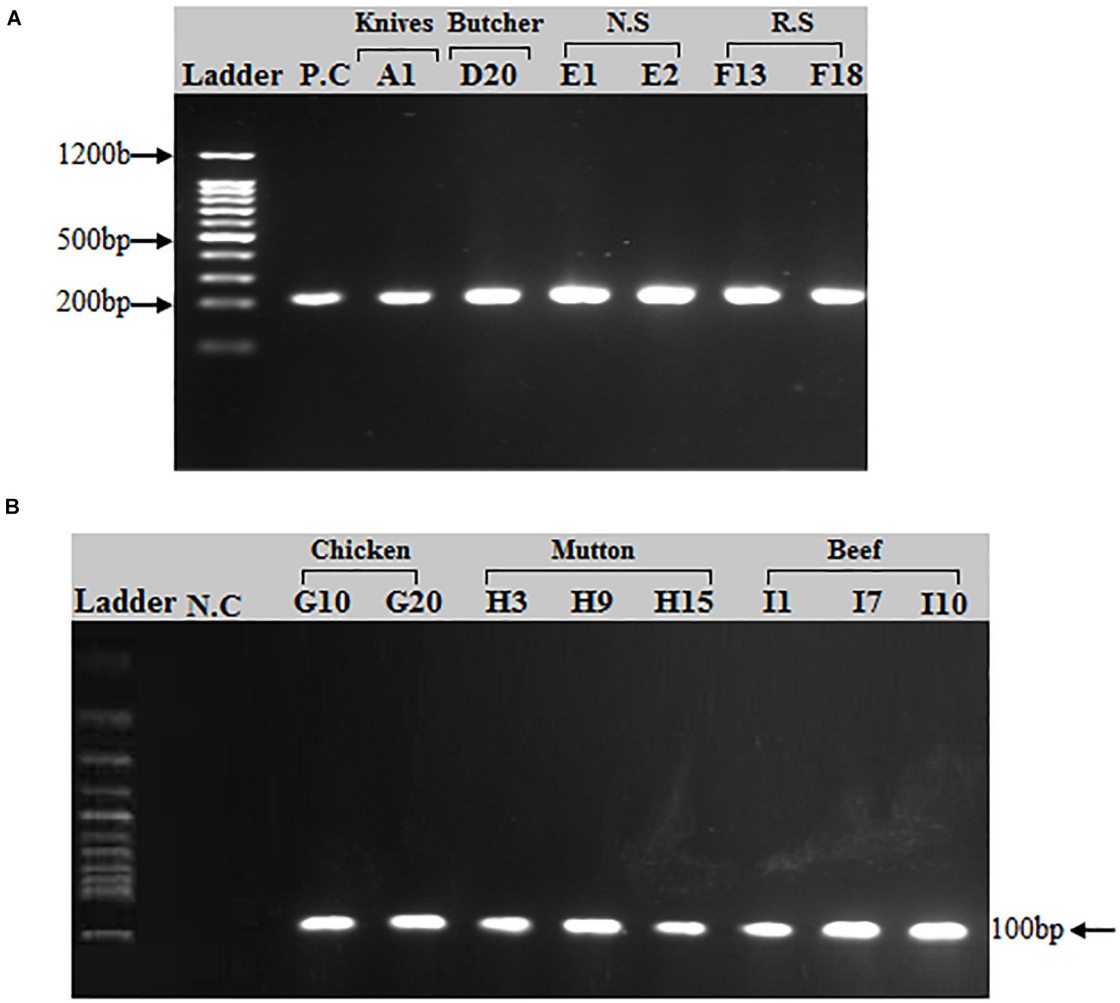


FIGURE 2 | PCR amplification of *spa* (*spa*-X) gene in *S. aureus* isolates: **(A)** A band (200 bp) corresponding to amplified region of *spa* gene. **(A)** First lane = 100 bp ladder, 2nd Lane = Positive control (*spa* gene), Lane# 3–8 = PCR products of *spa* X region from different sources (Knives, Butcher hands, Nasal swabs, Rectal swabs). **(B)** A band (100 bp) corresponding to amplified region of *spa* gene (X region). First lane = 100 bp ladder, 2nd Lane = Negative control, Lane# 3–10 = PCR products of *spa* X region from meat samples (beef, chicken, and mutton).

it was used for detection of *16S rRNA* gene that is specific for bacterial identification, *nuc* gene that confirms *Staphylococcus aureus* presence, *mecA* gene that codes for *protein binding protein 2a (PBP2a)* and is used for the detection of MRSA and for two virulence factors (*spa* gene and *coa* gene) of highly pathogenic *S. aureus*. All 150 isolates were detected positive for *nuc* and*16S rRNA* genes. Of the 150 Staphylococci phenotypically identified in our study only 118 (63%) *S. aureus* isolates were considered MRSA by molecular detection of *mecA* gene.

The *mecA* gene is highly conserved among staphylococcal strains and is being used as a potential biomarker for the detection of methicillin resistance *S. aureus* (MRSA) (Jain et al., 2008). In our study, MRSA prevalence based on *mecA* identification is highest in chicken (23/30, 77%) followed by beef (25/40, 63%), mutton (15/30, 50%), knives (18/40, 45%), nasal swabs (7/20, 35%), working area (11/40, 28%), rectal swabs (5/20, 25%), hooks (7/40, 18%), and butcher hands (7/40, 18%). The highest prevalence of MRSA was found in meat samples in our study, which is comparable to a study conducted in Georgia (Jackson et al., 2013) and higher than those reported in Egypt and Netherland, United States, Korea, and Canada, which showed a 14.5 and 2.5, 1.9, 0.5, and 24.8% prevalence, respectively, in meat samples (Diederen et al., 2007; Kim et al., 2015; Ge et al., 2017; Al-Amery et al., 2019). This high prevalence (79%) of MRSA in the current study may be the result of variations in sample size, sampling sites and raw meat samples from open markets processed through butcher tools and are treated by humans who may be a possible carrier of *staphylococci* isolates. Possible reasons for lower prevalence in some countries may be the sampling season and time, i.e., the samples were obtained in the winter season within 8 h of slaughter and in the early afternoon in order to reduce the risk of contamination. Total of 23 isolates that were susceptible to cefoxitin phenotypically showed presence of *mecA* gene, which is a unique property of these isolates that has not

TABLE 5 | *coa* and *spa* genes patterns of the 50 MRSA strains isolated from different sources.

				***spa* X region band patterns**			***coa* gene band patterns**	
Source	**Total number of isolates (n)**	**Isolate names**	***spa* + ve isolates**	**100 bp**	**200 bp**	***coa* + ve isolates**	**680 bp**	**891 bp**
Knives	5	A1, A13, A28, A35, A40	A1	–	1	A1, A13, A35	–	3
Hooks	5	B1, B2, B3, B4, B5	–ve	–ve	–ve	–ve	–ve	–ve
Working Area	5	C1, C3, C12, C27, C39	–ve	–ve	–ve	–ve	–ve	–ve
Butcher hands	5	D1, D4, D9, D20, D37	D20	–	1	D9, D20	–	2
Nasal swab	5	E1, E2, E6, E11, E14	E1, E2	–	2	E1, E2, E11, E14	–	4
Rectal swab	5	F2, F13, F18, F19, F20	F13, F18	–	2	F13, F18, F20	–	3
Chicken	8	G1, G3, G10, G15, G20, G23, G24, G27	G3, G10, G15, G20, G23	5	–	G1, G3, G10, G15, G20, G23, G27	7	–
Mutton	5	H3, H9, H15, H16, H30	H3, H9, H15	3	–	H3, H9, H15	3	–
Beef	7	I1, I7, I9, I10, 113, I17, I32	I1, I7, I10, I13	4	–	I1, I7, I10, I13, I17	5	–

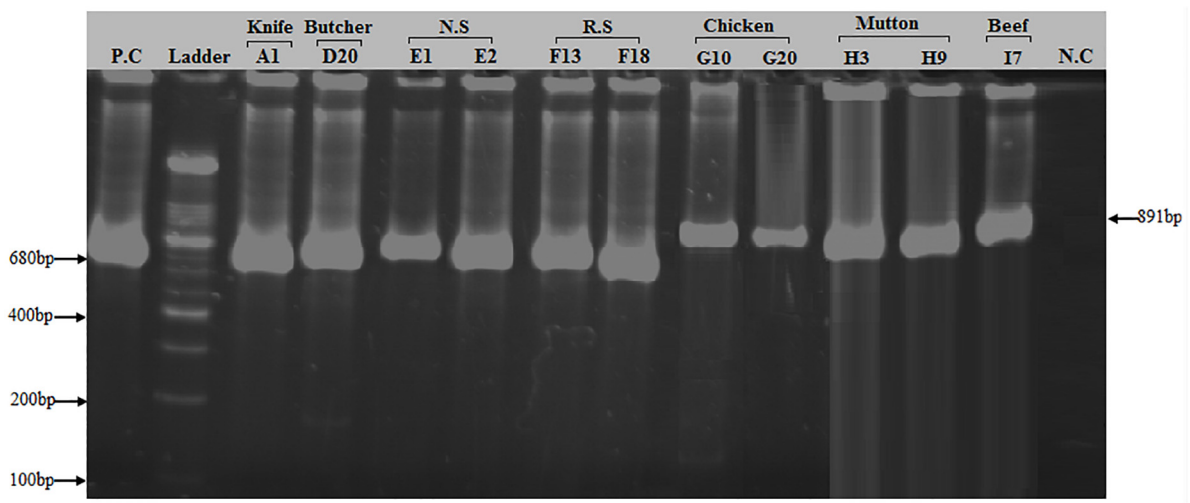


FIGURE 3 | PCR amplification of *coa* (Coagulase) gene in *S. aureus* isolates: Two bands of 680 and 891 bp corresponding to amplified region of *coa* gene. First lane = Positive control (*coa* gene), 2nd Lane = 100 bp ladder, Lane# 3–8 = PCR products of 680 bp size of *coa* gene from different sources (Knife, Butcher hands, Nasal swabs, Rectal swabs), Lane# 9–13 = PCR products of 891 bp size of *coa* gene from meat samples (beef, chicken, and mutton), Lane# 14 = Negative control.

been reported so far in research involving detection of methicillin resistant *S. aureus*.

Our current study findings demonstrated a congruence between phenotypic resistance and molecular typing. Methicillin resistance was identified in 95 isolates checked with cefoxitin disk diffusion technique, while 118 isolates had *mecA* gene. Generally speaking, our results are consistent with studies that indicated the detection of *staphylococci* strains which were *mecA* positive but susceptible to cefoxitin (Martineau et al., 2000; Lee et al., 2004; Osman et al., 2016). The possible explanation is that the resistance pattern of the phenotypic expression can differ depending on the temperature or osmolarity of the media used. It would probably make MRS susceptibility testing by conventional laboratory procedures difficult.

Fifty *mecA* positive samples out of one hundred eighteen were selected for *spa* and *coa* genes typing. The *spa* gene coding for the outer coat protein known as Protein A which is conserved between *S. aureus* strains (Okorie-Kanu et al., 2020). This gene provides sufficient short sequence repeat region (known as the X-region) containing variable number tandem repeats (VNTRs)

that are genetically heterogeneous and are used as a single-locus sequence typing target (SLST), commonly known as *spa* typing (Koreen et al., 2004). Protein A an antiphagocytic protein bound to the cell wall with its C-terminal end, the amino terminal end being free outside and binding with the Fc region of IgG (Rathore et al., 2012). In the present study, 18 of the 50 MRSA isolates were identified as positive for the *spa* gene by amplifying the X-region of protein A generating amplicons of two different sizes, 100 and 200 bp. The 200 bp was the most common band in knives, butcher hands, nasal and rectal swab samples, while the 100 bp band was prevalent in meat samples (Chicken, Beef and Mutton). A similar study conducted in Egypt showed *spa* gene segment size ranges from 100, 200, 280, and 290 bp after PCR amplification of MRSA isolates (Salem-Bekhit et al., 2010). A related studies conducted in India showed *spa* gene amplicon sizes after PCR in MRSA isolates using same primer set as (206, 243, 262, 277, 292, 306, and 339 bp), (280, 250, 240, 200, 190, 180, 170, 150, and 140 bp), and in Italy (253 bp), respectively (Casagrande Proietti et al., 2010; Khichar et al., 2014; Bhati et al., 2016). In the current study, MRSA strains in which the *spa* gene is absent, it is proposed that either the *spa* mutation has occurred or the *spa* gene appears to have been absent from these strains. Similarly, previous studies also identified *S. aureus* isolates without *spa* gene (Baum et al., 2009; Momtaz et al., 2010; Salem-Bekhit et al., 2010; Shakeri et al., 2010).

The amplification of the coagulase gene was regarded as a fast and precise method for typing *S. aureus*. Coagulase enzyme is a major virulent element that is secreted by all strains of *S. aureus*. Coagulase trigger the coagulation of plasma at the host and is identification marker for *S. aureus* infection (Himabindu et al., 2009). The heterogeneity among different strains of *S. aureus* is based on the region containing the 81 bp tandem repeats the 3′ coding region of the coagulase gene, which varies in the number of tandem repeats as well as in the position of the *AluI* and *HaeII* restriction sites between the different isolates (Afrough et al., 2013; Javid et al., 2018). In the present study, 27 of the 50 MRSA isolates were found to be positive for coa (Coagulase) gene producing segments of two different sizes, 680 and 891 bp. The 680 bp was the most common band in knives, butcher hands, nasal and rectal swab samples, while the 891 bp was prevalent in meat samples (Chicken, Beef and Mutton). A related studies conducted in India and United Kingdom showed coa gene amplicon sizes after PCR in MRSA isolates ranges from 510–1,000 bp using same primer sets (Hookey et al., 1998; Himabindu et al., 2009; Khichar et al., 2014; Javid et al., 2018).

This research added to the literature by contrasting the phenotypic and molecular characteristics of *S. aureus* in slaughter houses is and discovering the possible transmission modes of MRSA. However, this analysis also has possible drawbacks which cannot be overlooked. First of all, we found the methicillin resistant strains of *S. aureus* via the detection of a *mecA* gene and did not detect a novel *mecA* homolog *mecC* gene, which should be detected in future studies. Also there are no current surveillance programs which allow us to recommend a large scale of research to be carried out to cover the whole country, specifying a large number of sample sizes to be included for the majority of meat consumed in Pakistan.

CONCLUSION

The current study concludes that raw meat in the twin cities of Pakistan was contaminated with pathogenic methicillin-resistant *S. aureus* strains that were also resistant to clinically important antimicrobials, which is alarming for public health. This study is the first in Pakistan to report on the detection of *nuc*, *mecA*, *coa*, and *spa* genes positive MRSA. The findings of the present study will significantly add to existing knowledge of veterinary health research as well as food safety by providing proper education and training to slaughterhouse staff that would lead to the low *S. aureus* contamination by butchers, particularly in developing countries around the world. Furthermore the antibiotic resistance rates observed in the current study would highlight the importance of implementation of strict policies and strategies on the prudent use of antibiotics by the public as well as the farming sector. *S. aureus* in the current study were screened for two virulent genes (*coa* and *spa*), thus further in-depth genetic analysis covering the entire country is required including the detection of resistant genes (*blaZ, tetA, tetM, tetK, ermA, ermB, ermC, femA*, etc.), enterotoxins, pvl (*Sea, Seb, Sec, Sed, See*), *SCCmec* types (*SCCmecI-SCCmecIII, SCCmecIva, SCCmecIVb, SCCmecIVc, SCCmecIVd, SCCmecV*), *spa* typing and MLST (Multilocus sequence typing). There is also a need for continuous tracking, and the introduction of improved management methods inside the food chain to reduce contamination of food with MRSA and the eventual spread of the bacteria.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/**Supplementary Material** and nucleotide sequencing data is submitted in GenBank.

AUTHOR CONTRIBUTIONS

AS and JK contributed to the idea or design of the research. JK compiled the data. AS wrote down the first draft of the manuscript and wrote the subsequent revisions of the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This research work was funded by the Higher Education Commission (HEC) of Pakistan and COMSATS University, Islamabad, Pakistan and Quaid-i-Azam University, Islamabad. The funders were not involved in the design, review or interpretation of the data, the writing or the decision to submit for publication.

ACKNOWLEDGMENTS

We would like to thank the Higher education Commission of Pakistan, COMSATS University, Islamabad and Quaid-i-Azam University, Islamabad for providing funding for this project.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2020.577707/full#supplementary-material

Supplementary Figure 1 | Growth of *S. aureus* on mannitol salt agar. Growth of *Staphylococcus aureus* showing yellow colonies and mannitol fermentation on mannitol salt agar medium.

Supplementary Figure 2 | Growth of *S. aureus* on blood agar. *Staphylococcus aureus* showed beta hemolysis on blood agar medium.

Supplementary Figure 3 | Gram staining test for the identification of *S. aureus*. Clusters of gram positive cocci observed under microscope.

Supplementary Figure 4 | Catalase test for the identification of *S. aureus*. Formation of bubbles after colony mixing with H_2O_2 shows catalase Positive result.

Supplementary Figure 5 | Coagulase test for the identification of *S. aureus*. Formation of clumps after colony mixing with blood shows coagulase positive result.

Supplementary Figure 6 | DNase test for the identification of *S. aureus*. Clear zone around colonies shows DNase positive result after treating with 1 N HCL.

Supplementary Figure 7 | Results of antibiotic Resistance tests. Mueller-Hinton agar plates were seeded with *Staphylococcus aureus*. Ten antibiotic disks were placed on each plate. All plates were incubated at 37°C overnight. The diameter of each zone was measured in millimeters and evaluated for resistance using the comparative standard method.

Supplementary Figure 8 | Extracted genomic DNA. The extracted DNA for Sample#1–9 was analyzed through gel electrophoresis using 1% gel.

Supplementary Figure 9 | Detection of 16S rRNA gene in *S. aureus* isolates: A band (886 bp) corresponding the amplified region of 16S rRNA. First lane = Negative control, 2nd lane = 100 bp ladder, Lane 17 = Positive control, Lane# 3–16 = PCR products of 16S rRNA gene from different sources (Knives, Hooks, Working area, Butcher hands, Nasal swabs, Rectal swabs, Beef, Chicken, and Mutton).

Supplementary Figure 10 | Detection of *nuc* gene in *S. aureus* isolates: A band (270 bp) corresponding to amplified region of *nuc* gene. First lane = 100 bp ladder, 2nd Lane = Positive control (*S. aureus* ATCC29213), Lane 16th = Negative control, Lane# 3–15 = PCR products of *nuc* gene from different sources (Knives, Hooks, Working area, Butcher hands, Nasal swabs, Rectal swabs, Beef, Chicken, and Mutton).

REFERENCES

Adugna, F., Pal, M., Girmay, G., and Baynes, R. E. (2013). Prevalence and antibiogram assessment of *Staphylococcus aureus* in beef at municipal abattoir and butcher shops in Addis Ababa, Ethiopia. *Int. J. Food Sci.* 2018:5017685. doi: 10.1155/2018/5017685

Afrough, P., Pourmand, M. R., Sarajian, A. A., Saki, M., and Saremy, S. (2013). Molecular investigation of *Staphylococcus aureus*, coa and spa genes in Ahvaz hospitals, staff nose compared with patients clinical samples. *Jundishapur J. Microbiol.* 6:5377. doi: 10.5812/jjm.5377

Al-Amery, K., Elhariri, M., Elsayed, A., El-Moghazy, G., Elhelw, R., El-Mahallawy, H., et al. (2019). Vancomycin-resistant *Staphylococcus aureus* isolated from camel meat and slaughterhouse workers in Egypt. *Antimicrob. Resist. Infect. Control* 8:129.

Ali, M., Irtiqa, A., Mahrukh, F., and Tooba, A. (2018). Factors leading to acquired bacterial resistance due to antibiotics in Pakistan. *Curr. Trends Biotechnol. Microbiol.* 1, 1–7. doi: 10.32474/CTBM.2018.01.000101

Alt, K., Fetsch, A., Schroeter, A., Guerra, B., Hammerl, J. A., Hertwig, S., et al. (2011). Factors associated with the occurrence of MRSA CC398 in herds of fattening pigs in Germany. *BMC Vet. Res.* 7:69. doi: 10.1186/1746-6148-7-69

Anjum, M. F., Marco-Jimenez, F., Duncan, D., Marín, C., Smith, R. P., and Evans, S. J. (2019). Livestock-associated methicillin-resistant *Staphylococcus aureus* from animals and animal products in the UK. *Front. Microbiol.* 10:2136. doi: 10.3389/fmicb.2019.02136

Basanisi, M. G., La Bella, G., Nobili, G., Franconieri, I., and La Salandra, G. (2017). Genotyping of methicillin-resistant *Staphylococcus aureus* (MRSA) isolated from milk and dairy products in South Italy. *Food Microbiol.* 62, 141–146. doi: 10.1016/j.fm.2016.10.020

Bauer, A. W., Kirby, W. M., Sherris, J. C., and Turck, M. (1996). Antibiotic susceptibility testing by a standardized single disk method. *Am. J. Clin. Pathol.* 45, 493–496.

Baum, C., Haslinger-Löffler, B., Westh, H., Boye, K., Peters, G., Neumann, C., et al. (2009). Non-spa-typeable clinical *Staphylococcus aureus* strains are naturally occurring protein A mutants. *J. Clin. Microbiol.* 47, 3624–3629.

Bhati, T., Nathawat, P., Kumar Sharma, S., Yadav, R., Bishnoi, J., and Kumar Kataria, A. (2016). Polymorphism in spa gene of *Staphylococcus aureus* from bovine subclinical mastitis. *Vet. World* 9, 421–424.

Casagrande Proietti, P., Coppola, G., Bietta, A., Luisa Marenzoni, M., Hyatt, D. R., Coletti, M., et al. (2010). Characterization of genes encoding virulence determinants and toxins in *Staphylococcus aureus* from bovine milk in central Italy. *J. Vet. Med. Sci.* 72, 1443–1448. doi: 10.1292/jvms.10-0158

Chaalal, W., Chaalal, N., Bourafa, N., Kihal, M., Diene, S. M., and Rolain, J. M. (2018). Characterization of *Staphylococcus aureus* isolated from food products in western Algeria. *Foodborne Pathog. Dis.* 15, 353–360. doi: 10.1089/fpd.2017.2339

Clinical and Laboratory Standards Institute [CLSI] (2019). *Performance Standards for Antimicrobial Susceptibility Testing. CLSI supplement M100*, 29th Edn. Wayne, PA: Clinical and Laboratory Standards Institute.

Cookson, B. D., Robinson, D. A., Monk, A. B., Murchan, S., Deplano, A., De Ryck, R., et al. (2007). Evaluation of molecular typing methods in characterizing a European collection of epidemic methicillin-resistant *Staphylococcus aureus* strains: the HARMONY collection. *J. Clin. Microbiol.* 45, 1830–1837.

Deurenberg, R. H., Vink, C., Kalenic, S., Friedrich, A. W., Bruggeman, C. A., and Stobberingh, E. E. (2007). The molecular evolution of methicillin-resistant *Staphylococcus aureus*. *Clin. Microbiol. Infect.* 13, 222–235. doi: 10.1111/j.1469-0691.2006.01573.x

Diederen, B. M. W., van Loo, I. H. M., Woudenberg, J. H. C., Roosendaal, R., Verhulst, C., van Keulen, P. H. J., et al. (2007). "Low prevalence of non-typable Methicillin-resistant *Staphylococcus aureus* in meat products in The Netherlands," in *Proceedings of the 7th International Symposium on the Epidemiology & Control of Foodborne Pathogens in Pork*, (Ames, IA: Iowa State University), 398–401. doi: 10.31274/safepork-180809-13

Dweba, C. C., Zishiri, O. T., and El Zowalaty, M. E. (2019). Isolation and molecular identification of virulence, antimicrobial and heavy metal resistance genes in livestock-associated methicillin-resistant *Staphylococcus aureus*. *Pathogens* 8:79. doi: 10.3390/pathogens8020079

Fernandes, C. J., Fernandes, L. A., and Collignon, P. (2005). Cefoxitin resistance as a surrogate marker for the detection of methicillin-resistant *Staphylococcus aureus*. *J. Antimicrob. Chemother.* 55, 506–520. doi: 10.1093/jac/dki052

Ge, B., Mukherjee, S., Hsu, C. H., Davis, J. A., Tran, T. T. T., Yang, Q., et al. (2017). MRSA and multidrug-resistant *Staphylococcus aureus* in U.S. retail meats, 2010–2011. *Food Microbiol.* 62, 289–297. doi: 10.1016/j.fm.2016.10.029

Graveland, H., Duim, B., van Duijkeren, E., Heederik, D., and Wagenaar, J. A. (2011). Livestock-associated methicillin-resistant *Staphylococcus aureus* in animals and humans. *Int. J. Med. Microbiol.* 301, 630–634. doi: 10.1016/j.ijmm.2011.09.004

Guo, D., Liu, Y., Han, C., Chen, Z., and Ye, X. (2018). Phenotypic and molecular characteristics of methicillin-resistant and methicillin-susceptible

Staphylococcus aureus isolated from pigs: implication for livestock-association markers and vaccine strategies. *Infect. Drug Resist*. 11, 1299–1307. doi: 10.2147/IDR.S173624

Güven Gökmen, T., Kalayci, Y., Yaman, A., and Köksal, F. (2018). Molecular characterization of methicillin-resistant *Staphylococcus aureus* strains by spa typing and pulsed field gel electrophoresis methods. *BMC Microbiol.* 18:155. doi: 10.1186/s12866-018-1305-6

Hanson, B. M., Dressler, A. E., Harper, A. L., Scheibel, R. P., Wardyn, S. E., Roberts, L. K., et al. (2011). Prevalence of *Staphylococcus aureus* and methicillin-resistant *Staphylococcus aureus* (MRSA) on retail meat in Iowa. *J. Infect. Public Health* 4, 169–174. doi: 10.1016/j.jiph.2011.06.001

Haskell, K. J., Schriever, S. R., Fonoimoana, K. D., Haws, B., Hair, B. B., Wienclaw, T. M., et al. (2018). Antibiotic resistance is lower in *Staphylococcus aureus* isolated from antibiotic-free raw meat as compared to conventional raw meat. *PLoS One* 13:e0206712. doi: 10.1371/journal.pone.0206712

Hayat, K., Rosenthal, M., Gillani, A. H., Chang, J., Ji, W., Yang, C., et al. (2020). Perspective of key healthcare professionals on antimicrobial resistance and stewardship programs: a multicenter cross-sectional study from Pakistan. *Front. Pharmacol.* 10:1520. doi: 10.3389/fphar.2019.01520

Himabindu, M., Muthamilse, D. S., Bishi, D. K., and Verma, R. S. (2009). Molecular analysis of coagulase gene polymorphism in clinical isolates of methicillin resistant *Staphylococcus aureus* by restriction fragment length polymorphism based genotyping. *Am. J. Infect. Dis.* 5, 170–176. doi: 10.3844/ajidsp.2009.170.176

Hookey, J. V., Richardson, J. F., and Cookson, B. D. (1998). Molecular typing of *Staphylococcus aureus* based on PCR restriction fragment length polymorphism and DNA sequence analysis of the coagulase gene. *J. Clin. Microbiol.* 36, 1083–1089. doi: 10.1128/jcm.36.4.1083-1089.1998

Ibrahim, O. M. A., Bilal, N. E., Osman, O. F., and Magzoub, M. A. (2017). Assessment of methicillin resistant *Staphylococcus aureus* detection methods: analytical comparative study. *Pan Afr. Med. J.* 27:281. doi: 10.11604/pamj.2017.27.281.9016

Igbinosa, E. O., Beshiru, A., Akporehe, L. U., Oviasogie, F. E., and Igbinosa, O. O. (2016). Prevalence of methicillin-resistant *Staphylococcus aureus* and other *Staphylococcus* species in raw meat samples intended for human consumption in Benin City, Nigeria: implications for public health. *Int. J. Environ. Res. Public Health* 13:949. doi: 10.3390/ijerph13100949

Iroha, I. R., Ugbo, E. C., Ilang, D. C., and OjiA, E. (2011). Bacteria contamination of raw meat sold in Abakaliki, Ebonyi State Nigeria. *J. Public Health Epidemiol.* 3, 49–53.

Ivbule, M., Miklašević́s, E., Ćupane, L., Berziņa, L., Balinš, A., and Valdovska, A. (2017). Presence of methicillin-resistant *Staphylococcus aureus* in slaughterhouse environment, pigs, carcasses, and workers. *J. Vet. Res.* 61, 267–277.

Jackson, C. R., Davis, J. A., and Barrett, J. B. (2013). Prevalence and characterization of methicillin-resistant *staphylococcus aureus* isolates from retail meat and humans in Georgia. *J. Clin. Microbiol.* 51, 1199–1207.

Jain, A., Agarwal, A., and Verma, R. K. (2008). Cefoxitin disc diffusion test for detection of meticillin-resistant staphylococci. *J. Med. Microbiol.* 57, 957–961. doi: 10.1099/jmm.0.47152-0

Javid, F., Taku, A., Bhat, M. A., Badroo, G. A., Mudasir, M., Sofi, T. A., et al. (2018). Molecular typing of *Staphylococcus aureus* based on coagulase gene. *Vet. World* 11, 423–430.

Khichar, V., Kataria, A. K., and Sharma, R. (2014). Characterization of *Staphylococcus aureus* of cattle mastitis origin for two virulence-associated genes (coa and spa). *Comp. Clin. Pathol.* 23, 603–611.

Kim, Y. J., Oh, D. H., Song, B. R., Heo, E. J., Lim, J. S., Moon, J. S., et al. (2015). Molecular characterization, antibiotic resistance, and virulence factors of methicillin-resistant *staphylococcus aureus* strains isolated from imported and domestic meat in Korea. *Foodborne Pathog. Dis.* 12, 390–398. doi: 10.1089/fpd.2014.1885

Koreen, L., Ramaswamy, S. V., Graviss, E. A., Naidich, S., Musser, J. M., and Kreiswirth, B. N. (2004). spa typing method for discriminating among *Staphylococcus aureus* isolates: implications for use of a single marker to detect genetic micro- and macrovariation. *J. Clin. Microbiol.* 42, 792–799. doi: 10.1128/JCM.42.2.792-799.2004

Lakhundi, S., and Zhang, K. (2018). Methicillin-resistant *Staphylococcus aureus*: molecular characterization, evolution, and epidemiology. *Clin. Microbiol. Rev.* 31:e00020-18.

Larsen, J., Imanishi, M., Hinjoy, S., Tharavichitkul, P., Duangsong, K., Davis, M. F., et al. (2012). Methicillin-resistant *Staphylococcus aureus* ST9 in pigs in Thailand. *PLoS One* 7:e31245. doi: 10.1371/journal.pone.0031245

Lee, J. H., Jeong, J. M., Park, Y. H., Choi, S. S., Kim, Y. H., Chae, J. S., et al. (2004). Evaluation of the methicillin-resistant *Staphylococcus aureus* (MRSA) -screen latex agglutination test for detection of MRSA of animal origin. *J. Clin. Microbiol.* 42, 2780–2782. doi: 10.1128/JCM.42.6.2780-2782.2004

Lewis, H. C., Mølbak, K., Reese, C., Aarestrup, F. M., Selchau, M., Sørum, M., et al. (2008). Pigs as source of methicillin-resistant *Staphylococcus aureus* CC398 infections in humans, Denmark. *Emerg. Infect. Dis.* 14, 1383–1389. doi: 10.3201/eid1409.071576

Lowy, F. D. (1998). Medical progress: *Staphylococcus aureus* infections. *N. Engl. J. Med.* 339, 520–532. doi: 10.1056/NEJM199808203390806

Maristela Oliveira Lara, D., Oliveira Lara, M., Coaglio Lucas, T., Kalapothakis, E., Luis Thomasini, R., and Jorge Machado, C. (2018). Comparison of five methods of extraction of *Staphylococcus aureus* DNA for molecular detection by PCR. *Rev. Soc. Bras. Med. Trop.* 51, 528–532.

Martineau, F., Picard, F. J., Lansac, N., Ménard, C., Roy, P. H., Ouellette, M., et al. (2000). Correlation between the resistance genotype determined by multiplex PCR assays and the antibiotic susceptibility patterns of *Staphylococcus aureus* and *Staphylococcus* epidermidis. *Antimicrob. Agents Chemother.* 44, 231–238. doi: 10.1128/AAC.44.2.231-238.2000

Mohsin, M., Van Boeckel, T. P., Saleemi, M. K., Umair, M., Naseem, M. N., He, C., et al. (2019). Excessive use of medically important antimicrobials in food animals in Pakistan: a five-year surveillance survey. *Glob. Health Action* 12:1697541. doi: 10.1080/16549716.2019.1697541

Momtaz, H., Rahimi, E., and Tajbakhsh, E. (2010). Detection of some virulence factors in *Staphylococcus aureus* isolated from clinical and subclinical bovine mastitis in Iran. *Afr. J. Biotechnol.* 9, 3753–3758.

Murakami, K., Minamide, W., Wada, K., Nakamura, E., Teraoka, H., and Watanabe, S. (1991). Identification of methicillin-resistant strains of staphylococci by polymerase chain reaction. *J. Clin. Microbiol.* 29, 2240–2244. doi: 10.1128/jcm.29.10.2240-2244.1991

O'Brien, A. M., Hanson, B. M., Farina, S. A., Wu, J. Y., Simmering, J. E., Wardyn, S. E., et al. (2012). MRSA in conventional and alternative retail pork products. *PLoS One* 7:e30092. doi: 10.1371/journal.pone.0030092

Okoli, C. E., Njoga, E. O., Enem, S. I., Godwin, E. E., Nwanta, J. A., and Chah, K. F. (2018). Prevalence, toxigenic potential and antimicrobial susceptibility profile of *Staphylococcus* isolated from ready-to-eat meats. *Vet. World* 11, 1214–1221.

Okorie-Kanu, O. J., Anyanwu, M. U., Ezenduka, E. V., Mgbeahuruike, A. C., Thapaliya, D., Gerbig, G., et al. (2020). Molecular epidemiology, genetic diversity and antimicrobial resistance of *Staphylococcus aureus* isolated from chicken and pig carcasses, and carcass handlers. *PLoS One* 15:e0232913. doi: 10.1371/journal.pone.0232913

Osman, K. M., Amer, A. M., Badr, J. M., Helmy, N. M., Elhelw, R. A., Orabi, A., et al. (2016). Antimicrobial resistance, biofilm formation and mecA characterization of methicillin-susceptible *S. aureus* and Non-*S. aureus* of beef meat origin in Egypt. *Front. Microbiol.* 7:222. doi: 10.3389/fmicb.2016.00222

Patricia Jevons, M. (1961). "Celbenin" - resistant staphylococci. *Br. Med. J.* 1, 124–125.

Pekana, A., and Green, E. (2018). Antimicrobial resistance profiles of *Staphylococcus aureus* isolated from meat carcasses and bovine milk in abattoirs and dairy farms of the Eastern Cape, South Africa. *Int. J. Environ. Res. Public Health* 15:2223. doi: 10.3390/ijerph15102223

Pires, S. M., Evers, E. G., Van Pelt, W., Ayers, T., Scallan, E., Angulo, F. J., et al. (2009). Attributing the human disease burden of foodborne infections to specific sources. *Foodborne Pathog. Dis.* 6, 417–424. doi: 10.1089/fpd.2008.0208

Pourhoseingholi, M. A., Vahedi, M., and Rahimzadeh, M. (2013). Sample size calculation in medical studies. *Gastroenterol. Hepatol. Bed Bench* 6, 14–17. doi: 10.22037/ghfbb.v6i1.332

Pourmand, M. R., Hassanzadeh, S., Mashhadi, R., and Askari, E. (2014). Comparison of four diagnostic methods for detection of methicillin resistant *Staphylococcus aureus*. *Iran. J. Microbiol.* 6, 341–344.

Rathore, P., Kataria, A. K., Khichar, V., and Sharma, R. (2012). Polymorphism in COA and SPA virulence genes in *Staphylococcus aureus* of camel skin origin. *J. Camel Pract. Res.* 19, 129–134.

Rinsky, J. L., Nadimpalli, M., Wing, S., Hall, D., Baron, D., Price, L. B., et al. (2013). Livestock-associated methicillin and multidrug resistant *Staphylococcus aureus* is present among industrial, not antibiotic-free livestock operation workers in North Carolina. *PLoS One* 8:e67641. doi: 10.1371/journal.pone.0067641

Rong, D., Wu, Q., Xu, M., Zhang, J., and Yu, S. (2017). Prevalence, virulence genes, antimicrobial susceptibility, and genetic diversity of *Staphylococcus aureus* from retail aquatic products in China. *Front. Microbiol.* 8:714. doi: 10.3389/fmicb.2017.00714

Salem-Bekhit, M. M., Muharram, M. M., Alhosiny, I. M., and Hashim, M. E. S. Y. (2010). Molecular detection of genes encoding virulence determinants in *Staphylococcus aureus* strains isolated from bovine mastitis. *J. Appl. Sci. Res.* 6, 121–128.

Salisbury, S. M., Sabatini, L. M., and Spiegel, C. A. (1997). Identification of methicillin-resistant staphylococci by multiplex polymerase chain reaction assay. *Am. J. Clin. Pathol.* 107, 368–373. doi: 10.1093/ajcp/107.3.368

Sergelidis, D., Papadopoulos, T., Komodromos, D., Sergelidou, E., Lazou, T., Papagianni, M., et al. (2015). Isolation of methicillin-resistant *Staphylococcus aureus* from small ruminants and their meat at slaughter and retail level in Greece. *Lett. Appl. Microbiol.* 61, 498–503. doi: 10.1111/lam.12485

Shakeri, F., Shojai, A., Golalipour, M., Alang, S. R., Vaez, H., and Ghaemi, E. A. (2010). Spa diversity among MRSA and MSSA strains of *Staphylococcus aureus* in north of Iran. *Int. J. Microbiol.* 2010:351397. doi: 10.1155/2010/351397

Vaiyapuri, M., Joseph, T. C., Rao, B. M., Lalitha, K. V., and Prasad, M. M. (2019). Methicillin-resistant *Staphylococcus aureus* in seafood: prevalence, laboratory detection, clonal nature, and control in seafood Chain. *J. Food Sci.* 84, 3341–3351. doi: 10.1111/1750-3841.14915

Vossenkuhl, B., Rgen Brandt, J., Fetsch, A., Käbohrer, A., Kraushaar, B., Alt, K., et al. (2014). Comparison of spa types, SCCmec types and antimicrobial resistance profiles of MRSA isolated from turkeys at farm, slaughter and from retail meat indicates transmission along the production Chain. *PLoS One* 9:e96308. doi: 10.1371/journal.pone.0096308

Waters, A. E., Contente-Cuomo, T., Buchhagen, J., Liu, C. M., Watson, L., Pearce, K., et al. (2011). Multidrug-resistant *Staphylococcus aureus* in US meat and poultry. *Clin. Infect. Dis.* 52, 1227–1230. doi: 10.1093/cid/cir181

Weese, J. S., and van Duijkeren, E. (2010). Methicillin-resistant *Staphylococcus aureus* and *Staphylococcus pseudintermedius* in veterinary medicine. *Vet. Microbiol.* 140, 418–429. doi: 10.1016/j.vetmic.2009.01.039

World Health Organization [WHO], (2019). *WHO | Critically Important Antimicrobials for Human Medicine*, 6th Edn. Geneva: WHO.

Wu, S., Huang, J., Wu, Q., Zhang, J., Zhang, F., Yang, X., et al. (2018). *Staphylococcus aureus* isolated from retail meat and meat products in China: incidence, antibiotic resistance and genetic diversity. *Front. Microbiol.* 9:2767. doi: 10.3389/fmicb.2018.02767

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Sadiq, Samad, Saddam, Basharat, Ali, Roohullah, Saad, Khan, Ahmad, Khan and Khan. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 09 October 2020
doi: 10.3389/fmicb.2020.591041

Application of Short Pre-enrichment, and Double Chemistry Real-Time PCR, Combining Fluorescent Probes and an Intercalating Dye, for Same-Day Detection and Confirmation of *Salmonella* spp. and *Escherichia coli* O157 in Ground Beef and Chicken Samples

Alejandro Garrido-Maestu[1], Sarah Azinheiro[1,2], Foteini Roumani[1,2], Joana Carvalho[1,2] and Marta Prado[1]*

[1]*Food Quality and Safety Research Group, International Iberian Nanotechnology Laboratory, Braga, Portugal,* [2]*Department of Analytical Chemistry, Nutrition and Food Science, School of Veterinary Sciences, University of Santiago de Compostela, Santiago de Compostela, Spain*

Molecular methods, particularly those based on real-time PCR (qPCR), have become a popular approach to detect pathogens in food samples. This technique may take advantage of hydrolysis fluorescent probes for increased specificity. Even though suitable, this approach loses the capacity of performing result confirmation by melt curve analysis. In the current study, we developed an alternative approach, combining fluorescent probes along with an intercalating dye (SYBR Green) in order to simultaneously detect, and confirm the result, of two foodborne pathogens (*Salmonella* spp. and *Escherichia coli* O157). This new approach named double chemistry qPCR was combined with a short pre-enrichment in order to obtain a multiplex "same-day" detection method for the selected pathogens. The evaluation of the novel method in spiked food samples (ground beef and chicken breast) obtained values of relative sensitivity, specificity, and accuracy higher than 95%, and Cohen's kappa of 0.92, with a Limit of Detection$_{95}$ below 5 cfu/25 g, demonstrating its reliability. In addition to this, the method was challenged by inoculating heat-stressed bacteria as well as dead ones. It was observed that it was also possible to detect stressed bacteria with an initial inoculation level below 10 cfu/25 g. Also, it was noticed that high initial concentration of either pathogen (higher than 10^4 cfu/25 g) was needed in order to generate false positive results due to the presence of dead bacteria, thus the method presents potential for its application in the specific detection of live microorganisms.

Keywords: melt curve analysis, *Salmonella* spp., *Escherichia coli* O157, intercalating dye, same-day detection, hydrolysis probe

OPEN ACCESS

Edited by:
Dario De Medici,
National Institute of Health (ISS), Italy

Reviewed by:
Joseph M. Bosilevac,
US Meat Animal Research Center (USDA-ARS), United States
Hengyi Xu,
Nanchang University, China

****Correspondence:***
Alejandro Garrido-Maestu
alejandro.garrido@inl.int;
a.garridomaestu@gmail.com

Specialty section:
This article was submitted to Food Microbiology, a section of the journal Frontiers in Microbiology

Received: *03 August 2020*
Accepted: *22 September 2020*
Published: *09 October 2020*

Citation:
Garrido-Maestu A, Azinheiro S, Roumani F, Carvalho J and Prado M (2020) Application of Short Pre-enrichment, and Double Chemistry Real-Time PCR, Combining Fluorescent Probes and an Intercalating Dye, for Same-Day Detection and Confirmation of Salmonella spp. and Escherichia coli O157 in Ground Beef and Chicken Samples. Front. Microbiol. 11:591041. doi: 10.3389/fmicb.2020.591041

INTRODUCTION

Molecular methods, particularly those based on DNA amplification, have become a very popular approach when fast and accurate results are needed in the food industry. Real-time PCR (qPCR) is one of the most widely accepted techniques, due to its high sensitivity, specificity, and capacity to detect the amplified fragments in real time without the need of additional manipulation (Kokkinos et al., 2014; Chapela et al., 2015). Also, it can take advantage of fluorescent probes for improved specificity. Out of these, the most commonly implemented are hydrolysis probes, due to their relatively simple design and reduced cost. A limitation of these probes is that melt curve analysis cannot be performed as the probe is degraded in the process of amplification. This issue may be overcome by selecting a hybridization probe instead, like the molecular beacons (MBs) or "adjacent probes" (Marras et al., 2006), but their design, particularly in the case of the MB, tends to be more complex as not only the specific target sequence has to be considered, but also the stem and the secondary structure formed (Zheng et al., 2015), and the cost is higher. Additionally, in the melting values obtained only correspond to that of the probes, and not to the complete amplified fragment (O'Grady et al., 2008; Chakravorty et al., 2010).

Nowadays, foodborne pathogens continue to be a major threat for human health. Among them, two of the most commonly reported bacterial pathogens are *Salmonella* spp. and Shiga Toxin-Producing *Escherichia coli* (STEC). According to the European Food Safety Authority, in 2017 a total of 91,662 confirmed salmonellosis cases were reported in Europe, as well as 6,073 STEC infections, being serogroup O157 the most commonly reported (31.9%). Overall the hospitalization rate was around 40%, what further highlights the health and economic impact of these pathogens (EFSA and ECDC, 2018).

Current reference methods for the detection of these bacteria in food samples are culture-based. Among these, the most extended ones are those described by the ISO and the FDA (ISO, 2002, 2003; Andrews et al., 2011; Feng et al., 2011). It is worth to mention that the FDA also provides a molecular-based method for the detection of *E. coli* O157:H7. All these methods have demonstrated to be highly reliable, and in the particular case of *E. coli*, novel approaches were included in the method such as immunomagnetic separation to concentrate specific serotypes, or implementation of qPCR for accurate detection, but lacks the capacity of result confirmation by melt curve as commented above. It may also be noted that all the mentioned reference methods begin with a sample enrichment/pre-enrichment what generates an overall delay of the results regardless the implementation of qPCR or not. A high number rapid methods have been reported, many based on PCR/qPCR, and even included in multi-center validation trials, but they still rely on sample enrichment to increase the concentration of the target bacteria to detectable numbers, what in the best case scenario makes them next-day detection approaches (Abdulmawjood et al., 2003, 2004; Malorny et al., 2003, 2004; Cheng et al., 2009, 2015). In this sense, recent studies have reported that an appropriate sample pre-treatment can significantly reduce the time of analysis by directly tacking the major bottleneck, the enrichment step. In the study published by Fachmann et al. (2017), they indicated that a short pre-enrichment (SpEn; 3 h) could allow for a sensitive detection of *Salmonella* spp. in meat samples. Likewise, Garrido-Maestu et al. (2020) followed a similar approach and also demonstrated that the methodology could be applied, with minor modifications, for the rapid detection of *E. coli* O157 in meat and salad samples. These studies provide real same-day detection of the pathogens, but have only targeted one single bacterial species, thus presenting a lower throughput.

In the current study, the development of a same-day detection methodology based on SpEn for the simultaneous detection of *Salmonella* spp. and *E. coli* O157 was developed. Additionally, the multiplex detection of the bacteria was performed by a novel qPCR approach named double chemistry (DC-qPCR) as it combines hydrolysis probes along with an intercalating dye (SYBR Green) in order to simultaneously detect and confirm the results.

MATERIALS AND METHODS

Bacterial Strains and Culture Media

Escherichia coli O157 WDCM 00014 and *Salmonella* Typhimurium WDCM 00031 (World Data Centre for Microorganisms) were selected as the reference strains for spiking experiments. Fresh cultures of both bacteria were prepared inoculating one single colony in 4 mL of Nutrient Broth (NB, Biokar diagnostics S.A., France) and incubated overnight at 37°C. After incubation, 100-fold serial dilutions were performed and plated on Tryptic Soy Agar (TSA, Biokar diagnostics S.A., France). The plates were incubated overnight at 37°C and counted to obtain a reference value of viable bacteria.

The sample SpEn was performed in buffered peptone water (BPW, Biokar diagnostics S.A., France) supplemented with 0.4% (w/v) of Tween 80 (Sigma-Aldrich, St. Louis, United States). The confirmation of the results obtained by the molecular method was performed by a culture-based approach. To this end, for the confirmation of *Salmonella* spp. Rappaport Vassiliadis Soya (RVS, Biokar diagnostics S.A., France), xylose lysine deoxycolate (XLD, Biokar diagnostics S.A., France), and ChromAgar™ *Salmonella* Plus (CHROMagar Microbiology, Paris, France) were used. Regarding *E. coli* O157, the media used were modified Tryptic Soy Broth with novobiocin (mTSBn, Biokar diagnostics S.A., France), Sorbitol MacConkey with Cefixime and Tellurite (CT-SMAC, Sigma-Aldrich, St. Louis, United States), and ChromAgar™ O157 (CHROMagar Microbiology, Paris, France). Details about the procedure followed are provided below in M&M 2.6.

Short Pre-enrichment

Two different food samples were included in the current study, ground beef and chicken breast. The protocol of analysis followed was adapted from Garrido-Maestu et al. (2020). Briefly, 25 g were mixed with 25 mL of BPW, with 0.4% of Tween 80, pre-warmed at 37°C in a filter bag for stomacher, the matrix

TABLE 1 | Primers and probes.

	Sequence (5'→3')	Modification	Reference
ttr-P3F	GGC TAA TTT AAC CCG TCG TCA G	-	
ttr-P3R	GTT TCG CCA CAT CAC GGT AGC	-	(Garrido-Maestu et al., 2018b)
ttr-P3P	AAG TCG GTC TCG CCG TCG GTG	NED/MGBNFQ	
O157-rfbE-F	TCA ACA GTC TTG TAC AAG TCC AC	-	
O157-rfbE-R	ACT GGC CTT GTT TCG ATG AG	-	(Garrido-Maestu et al., 2020)
O157-rfbE-P	ACT AGG ACC//GCA GAG GAA AGA GAG GAA	FAM/ZEN/IABkFQ	
NC-IAC-F	AGT TGC ACA CAG TTA GTT CGA G	-	
NC-IAC-R	TGG AGT GCT GGA CGA TTT GAA G	-	(Garrido-Maestu et al., 2019)
IAC-P	AGT GGC GGT//GAC ACT GTT GAC CT	YY/ZEN/IABkFQ	(Garrido-Maestu et al., 2018a)

YY, Yakima Yellow; IABkFQ, Iowa Black®FQ; ZEN (secondary, internal quencher) are trademarks from IDT.

was hand-massaged and incubated at 37°C for 3 h with constant agitation (200 rpm). After the 3 h, the enriched sample was recovered from the filter side to remove large food particles, and centrifuged at 8,960 *g* for 10 min, the supernatant was discarded, and the pellet was resuspended in 45 mL of protease buffer, and incubated horizontally at 37°C for 10 min at 200 rpm. After digestion, the samples were centrifuged as described above. The pellet was resuspended in 45 mL of washing buffer. This was followed by a new centrifugation step at 8,960 *g* for 10 min. The new pellet was recovered in 1.5 mL of washing buffer, transferred to a clean tube, and centrifuged at 11,000 *g* for 5 min. Finally, the bacterial pellet was rinsed with 1 mL of phosphate buffered saline (PBS), and centrifuged once more at 11,000 *g* for 5 min. The clean pellet was used for DNA extraction.

DNA Extraction

The protocol described by Garrido-Maestu et al. (2020) was followed. The bacterial pellet recovered after the SpEn was resuspended in 200 µL of 6% Chelex®100 (w/v; Bio-Rad Laboratories, Inc., United States), supplemented with 25 µL of Proteinase K (Macherey-Nagel, Düren, Germany), incubated at 56°C for 15 min, and this was followed by a thermal lysis at 99°C for 10 min. Both incubation steps were performed with constant agitation at 1,400 rpm in a Thermomixer comfort (Eppendorf AG, Germany). Finally, the samples were centrifuged at 11,000 *g* at 4°C for 5 min, and the supernatants were transferred to new clean tubes and sterile tubes for storage at 4°C until analysis.

Multiplex DC-qPCR

The detection of both bacterial pathogens was performed by multiplex qPCR. To this end, the detection of *Salmonella* spp. was performed targeting the *ttr* gene with the primers (200 nM) and probe (150 nM) described by Garrido-Maestu et al. (2017), regarding *E. coli* O157, and the *rfbE* gene (Garrido-Maestu et al., 2020) was targeted (500 nM primers and 250 nM probe). In addition to these, a non-competitive internal amplification control (IAC) was also included to avoid false negative results due to reaction inhibition (100 nM primers and probe, and 7×10^3 copies of IAC DNA; Garrido-Maestu et al., 2019). All primers and probes were purchased from Integrated DNA Technologies (Integrated DNA Technologies Inc., Leuven, Belgium), and are provided in **Table 1**.

The qPCR was performed in a final reaction volume of 20 µL, composed of 10 µL of TaqMan®Fast Advanced Master Mix supplier (Applied Biosystems™, Foster City, CA, USA), with the primer/probe concentrations specified above, 1 µL of 10X Sybr Geen I dissolved in DMSO (SG, Invitrogen™, Carlsbad, CA, United States), and 3 µL of template DNA, and the remaining volume was filled with sterile milliQ water.

The thermal profile consisted of a uracil-DNA glycosylase (UDG) treatment for 2 min at 50°C. This was followed by 2 min at 95°C for polymerase activation, and then 40 cycles of dissociation at 95°C for 1 s, and annealing-extension at 61°C for 20 s. This was continued by a melt curve stage, where the temperature was increased to 95°C for 1 s, decreased to 70°C for 20 s, and increased back to 95°C at a rate of 0.1°C/s. The analysis was performed in a QuantStudio 5 real-time PCR system with the QuantStudio™ Design and Analysis Software v1.4.3 (Applied Biosystems™, Foster City, CA, United States). Positive samples will be considered those providing amplification of the corresponding fluorophore, and with a melting peak at the expected temperature. In the same way, negative samples will lack amplification of the specific fluorophores and/or will not present the expected melting peak, in addition of having amplification for the NC-IAC with its own specific melting peak.

Correct performance of the novel approach was confirmed evaluating the amplification efficiency and the dynamic range of the multiplex reaction using 10-fold dilutions of pure bacterial DNA mixed in equal concentrations. Three biological replicates with three technical replicates for each concentration were analyzed (total of nine data for each point). The amplification efficiency was calculated with the following formula: $E = 10^{-1/b-1}$, where "*E*" is the efficiency, and "b" the slope of the curve obtained (González-Escalona et al., 2009).

Confirmation Procedure

A total of 100 µL of the SpEn were transferred to 10 mL of RVS and mTSBn, which were incubated at 42 and 37°C overnight respectively. After selective enrichment, a loopful of RVS was streaked on XLD and ChromAgar™ *Salmonella*. The mTSBn was streaked on CT-SMAC and ChromAgar™ O157. All plates were incubated overnight at 37°C, and screened the following day for typical colonies of the corresponding pathogen.

Evaluation of the Methodology

The evaluation of the proposed method was performed in spiked food samples. The 95% Limit of Detection (LOD_{95}) was determined. This was followed by the evaluation of its fit-for-purpose attending to its relative sensitivity, specificity, and accuracy (SE, SP, and AC), as well as its positive and negative predictive values (PPV/NPV) and its Cohen's kappa (*k*). The determination of all this parameters is detailed below.

Ninety-Five Percent Limit of Detection

The LOD_{95} was calculated according to Wilrich and Wilrich (2009). Thirty-two samples were spiked, 16 with each pathogen. The samples were divided in four groups of four samples, and each group was inoculated with decreasing concentrations of the corresponding pathogen, with the goal of reaching a concentration with positive and negative results (starting from an inoculation level between 10 and 10 × 10 cfu/25 g, down to a range between 1 and 10 cfu/25 g). The spiking was performed from fresh cultures prepared as detailed in M&M 2.1, as well as the serial dilutions and plating to determine the reference values of viable bacteria.

Fit-for-Purpose

Additional samples were inoculated with different concentrations in order to evaluate the performance of the method. These were classified attending to the obtained and expected result as positive/negative agreements if they match the expected result (PA/NA) or deviations if they did not match (PD/ND). These data obtained were used for the calculation of the relative sensitivity, specificity, and accuracy (SE, SP, and AC), along with the PPV/NPV and the Cohen's *k*, as described elsewhere (Tomas et al., 2009; Anderson et al., 2011; NordVaL, 2017).

Extended Study

Stressed Bacteria

In order to determine the performance of the methodology in a more realistic scenario, two sets of four samples were inoculated with thermally stressed bacteria. These were generated diluting 1/100 in PBS a fresh culture, prepared as described in M&M 2.1, and then heat-treat the dilution at 60°C for 10 min with constant agitation (1,000 rpm). The stressed bacteria were 100-fold serially diluted and inoculated at a final concentration below 10 cfu/25 g. A set of four-spiked samples were stored refrigerated (4–8°C) for 24 h, and another four were stored for 48 h. On the corresponding day, all samples were analyzed following the methodology described above.

Dead Bacteria

A final test was conducted in order to determine the effect of dead target bacteria in the final result. To do so, both target species were inactivated. The protocol followed consisted on taking 1 mL of a fresh culture of the corresponding microorganism, prepared as described in M&M 2.1. This was centrifuged at 16,000 *g* for 5 min. The supernatant was eliminated, and the bacterial pellet was resuspended in 1 mL of isopropanol 70% and heated at 92°C for 15 min. Finally, the dead bacterial suspension was centrifuged again at 16,000 *g* for 5 min, the isopropanol was removed, and the pellet was resuspended in 1 mL of PBS. A loopful was streaked on TSA and incubated at 37°C for 24 h to assure correct inactivation of both pathogens. The dead bacteria were used to inoculate four pairs of samples, eight in total, with increasing concentrations ranging from 10^3 to 10^6 cfu/25 g. The results obtained were introduced in the LOD model of Wilrich and Wilrich, in order to determine which was the minimum bacterial concentration needed to generate a false positive result due to the presence of dead microorganisms.

RESULTS

Determination of the Amplification Efficiency of the Multiplex DC-qPCR

The calculation of the multiplex DC-qPCR efficiency obtained values of 92.9 and 94.8% for *ttr* and *rfbE*, respectively. Additionally, the R^2 was determined to be 0.997 and 0.998. These results are graphically depicted in **Figure 1A**. In **Figures 1B–D**, the amplification plots obtained for *ttr*, *rfbE* and SG can be observed, and their corresponding melting curves are depicted in **Figure 1E**. As can be observed, the dynamic range for both targets corresponds to six decimal dilutions, from 12.8 ng/μL to 0.13 pg/μL. The melting temperature was experimentally determined to be 77.7 ± 0.6°C and 73.0 ± 0.5°C for *ttr* and *rfbE*, respectively.

Evaluation of the Methodology

Determination of the LOD_{95} for the DC-qPCR Combined With SpEn

The LOD_{95} was statistically calculated with the model described by Wilrich and Wilrich (2009). Similar values were obtained for both pathogens, 3.7 and 3.4 cfu/25 g for *Salmonella* spp. and *E. coli* O157, respectively, resulting in a combined LOD_{95} of 3.6 cfu/25 g for the method.

Fit-for-Purpose

The fit-for-purpose was determined in a total of 78 samples, 48 ground beef and 30 chicken breast. All parameters got values higher than 89% as only one ND was obtained for each pathogen. Both deviations were obtained in chicken samples. The results were confirmed by a culture-based approach, but it is worth to highlight that, while it was possible to isolate typical colonies of *Salmonella* spp. from XLD and ChromAgar *Salmonella*, in the case of *E. coli* O157, whenever high background microflora was present in the food samples, it was not possible to perform the confirmation with CT-SMAC, and only ChromAgar O157 allowed to isolate typical colonies. The specific values obtained, for each pathogen, as well as the combined results for the overall method, are summarized in **Table 2**.

Extended Study

Stressed Bacteria

In order to obtain some insights about the capacity of the novel methodology to detect the pathogens of interest in a

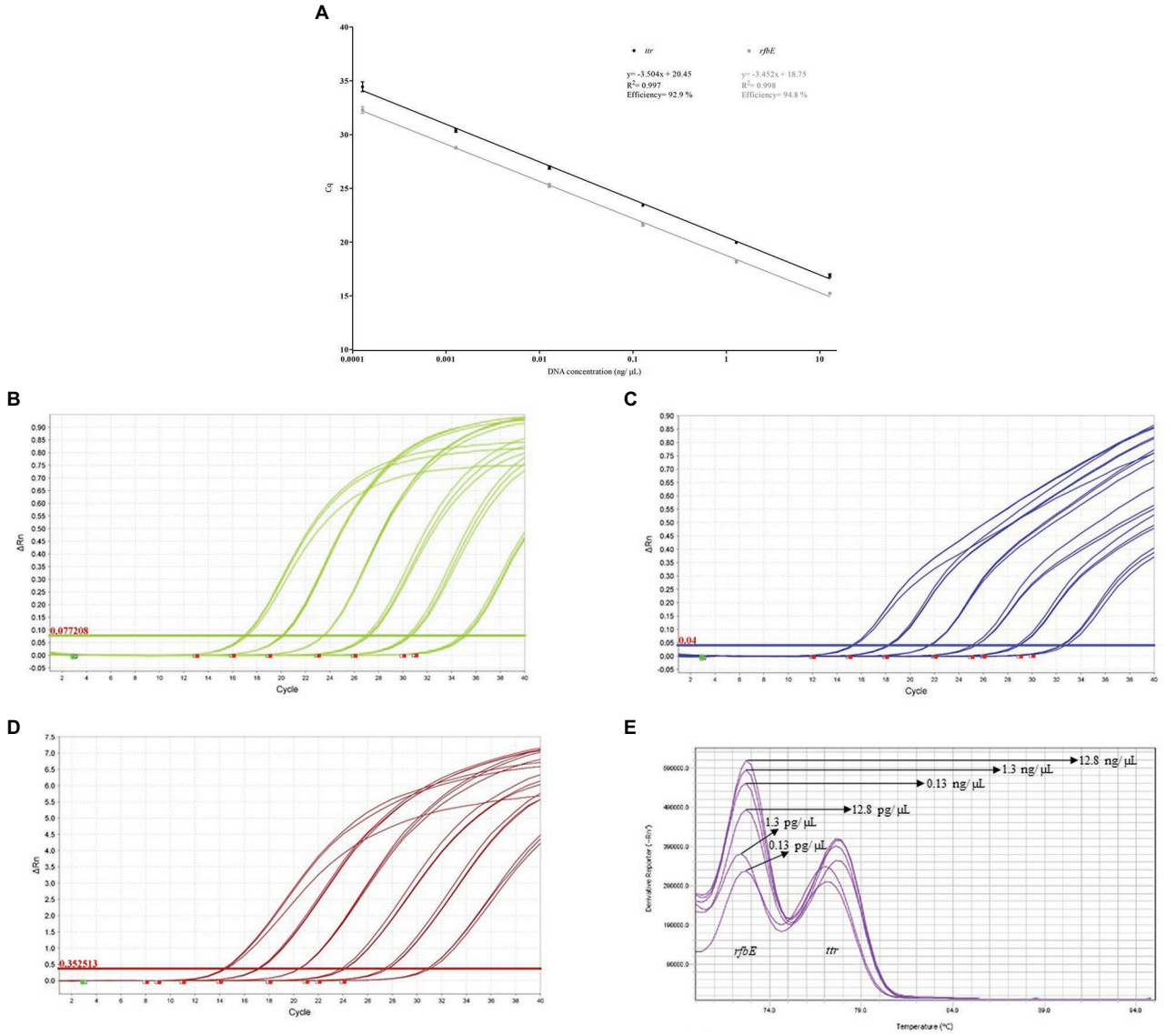


FIGURE 1 | (A) Multiplex double chemistry qPCR (DC-qPCR) amplification efficiency. **(B)** *ttr* amplification plots. **(C)** *rfbE* amplification plots. **(D)** SG amplification plots. **(E)** Melt curves generated after multiplex DR-qPCR.

TABLE 2 | Summary of the fitness-for-purpose for each pathogen and the combined method.

Pathogen	*N*°	PD	NA	ND	PA	SE	SP	AC	PPV	NPV	κ	LOD_{95} (cfu/25 g)
Salmonella spp.	39	0	8	1	30	97	100	97	100	89	0.92	3.7
Escherichia coli O157	39	0	8	1	30	97	100	97	100	89	0.92	3.4
Combined results	78	0	16	2	60	97	100	97	100	89	0.92	3.6

*N, total number of samples; PA, positive agreement; PD, positive deviation; NA, negative agreement; ND, negative deviation; SE, relative sensitivity; SP, relative specificity; AC, relative accuracy; PPV, positive predictive value; NPV, negative predictive value; k, Cohen's kappa, interpretation: 0.81–1.00 almost complete concordance according to Altman (1991) and Anderson et al. (2011). *The 78 samples corresponded to 48 ground beef and 30 chicken breast.*

real scenario, both bacteria were thermally stressed, and the spiked samples were stored under refrigeration for 24 and 48 h prior to analysis (four samples for each time point). The inoculation level for both pathogens was very low, 8 and 6 cfu/25 g of *E. coli* O157 and *Salmonella* spp., respectively. After 24 h of storage, three out of the four spiked samples were positive (75%). After 48 h, the number of positive samples for *Salmonella* spp. was still the same (three out of four, thus 75%), while only two were positive for *E. coli* O157 (50%).

Dead Bacteria

The novel method was also tested spiking some samples, eight in total, with different concentrations of dead bacteria (10^3–10^6 cfu/25 g) in order to determine the effect of inactivated target pathogens on the molecular method. Applying the statistical approach described by Wilrich and Wilrich (2009) for the LOD, it was calculated that 3.9×10^5 and 2.9×10^5 cfu/25 g of dead *Salmonella* spp. and *E. coli* O157 must be present in the original samples in order to obtain a PD linked to the presence of DNA from dead bacteria.

DISCUSSION

In the current study, a novel method combining short pre-enrichment (SpEn) in order to have a same-day detection, along with a novel detection approach named double chemistry qPCR (DC-qPCR) was developed and evaluated. The SpEn has been reported to provide good results for the pathogens included in the current study, *Salmonella* spp. and *E. coli* O157 (Fachmann et al., 2017; Garrido-Maestu et al., 2020), but no method was reported to attempt the simultaneous detection of both bacteria. In the current study, taking our previous findings dealing with the short-enrichment of *E. coli* O157 as a starting point, we have proceeded to tackle this issue by modifying the enrichment broth selected in order to select one more suited for both microorganisms. Regarding the DC-qPCR, we took advantage of the higher specificity provided by the hydrolysis probes, along with the capacity of SYBR Green I to bind double-stranded DNA and so to later perform melt curve analysis to confirm the results obtained. The combination of fluorescent probes with intercalating dyes has been previously described, but never tested for the detection of foodborne pathogens (Lind et al., 2006; Nagy et al., 2016).

Due to the fact that the final detection relies on DC-qPCR, first the effect of combining the different target primers and probes, along with those from a NC-IAC to rule out ND due to reaction inhibition, was assessed. The primers and probes selected to perform the multiplex detection of both pathogens, along with the NC-IAC, had been previously designed and tested by our research team. All three targets demonstrated to provide good results for their intended application in the corresponding original studies (detection of *Salmonella* spp., *E. coli* O157, and identify reaction inhibition) but were not tested in a multiplex format (Garrido-Maestu et al., 2018a,b, 2019, 2020). For this reason, we proceeded to re-evaluate them in this format. The multiplex amplification efficiency calculated for both targets was within range reported as acceptable (80–110%). Additionally, as expected, the implementation of the NC-IAC did not significantly affect, neither the amplification efficiency, nor the dynamic range covered for both targets. Finally, it was also confirmed that the peaks generated for the melt curve analysis were clearly distinguishable, thus suitable for the confirmation of each target (close to 5°C difference). Thus the DC-qPCR demonstrated to be a suitable alternative to simply detecting foodborne pathogens based on hydrolysis probes, as can provide the added value of melting curve confirmation of the results.

Due to the fact that the food industry in constantly seeking for novel and more rapid methods to detect different pathogens, the novel detection approach was combined with SpEn, which has already been described to allow for same-day detection of *Salmonella* spp. and *E. coli* O157 in different types of foodstuffs (Fachmann et al., 2017; Garrido-Maestu et al., 2020). The combined DC-qPCR with SpEn successfully allowed to accomplish the detection of both pathogens in one single working day, as after the initial pre-enrichment, the sample processing can be completed in roughly 1 h including the bacterial concentration, food left over proteolysis, sample clean-up, and DNA extraction. Additionally, economic cost of the sample treatment is relatively low as can be performed with a centrifuge, industrial proteases and the DNA extraction is performed by simple thermal lysis approach (Chelex). Furthermore, the results obtained in the process of evaluation (LOD_{95}, the five quality parameters evaluated as well as the Cohen's *k*) compared favorably to those previously described performing simplex pathogen detection. It is worth to highlight that the values calculated for SE (97%) and *k* (0.92) fulfill the requirements stated in the NordVal regulation for the validation of alternative methods (NordVaL, 2017). Overall, only two NDs were detected, one for each pathogen and both corresponded to chicken samples. The deviation related with *E. coli* O157 may be related with the fact that the inoculation level was very close to that of the LOD_{95} (calculated to be 5 cfu/25 g), while in the case of *Salmonella* spp., it was noticed that the sample causing the deviation got amplification by the specific probe, but classified as negative due to low Tm (below 77°C). This may have been originated by leftover contaminants from the actual food sample (chicken) or from the chemicals/reagents used in the process of sample pre-treatment and DNA extraction (e.g., the surfactant).

The last part of the evaluation study for the DC-qPCR combined with the SpEn is consisted in the determination of the capacity of the method to recover and detected stressed microorganisms, as well as to evaluate the capacity of dead bacteria to generate PD. In the particular case of stressed pathogens, it was observed that the number of deviations for *E. coli* O157 increases along with storage time, while for *Salmonella* spp. the number of positive and negative samples remained the same. Even though a relative viability of these pathogens has been reported when stored refrigerated, or even after stressing conditions in different food products (Park et al., 1970; Anassib et al., 2003; Barrera et al., 2007), it is worth to highlight that in the current study, the bacteria were submitted

to two different stressing conditions, first heat and later cold storage, this treatment combined with the low initial concentration (<10 cfu/25 g) may be behind the reduced detection of stressed bacteria.

Finally, the effect of dead bacteria on the DC-qPCR was also determined as PD due to the presence of dead microorganisms has been commonly regarded as a major limitation for the implementation of these types of methodologies in the food industry (Postollec et al., 2011; Cangelosi and Meschke, 2014). It was statistically calculated that more than 10^5 cfu of dead bacteria must be present in the original sample, in order to generate a PD with a 95% confidence, and more than 10^4 for a 50% chance. It can be observed by these figures, food samples should be highly contaminated with dead bacteria in order to generate PD due to them, even though additional studies should be conducted, there are strong evidences to assume that only live pathogens will be detected following the described methodology. To the best of our knowledge, no other study has estimated the effect of dead bacteria in SpEn, but previous studies have investigated the capacity of other sample pre-treatments to remove free DNA. According to the studies of Mann et al. (2014) and Mayrl et al. (2009), application of the matrix lysis procedure allows to reduce 5 log the amount of free DNA, thus greatly reducing the possible interferences in the detection methodology. In the same way, the studies published by D'Urso et al. (2009) and Garrido-Maestu et al. (2018b) described a pre-treatment combined with filtration to remove dead microorganisms, and reported to be fully effective up to 10^4–10^5 cfu/25 g. A study from Wang et al. (2013), indicated that 10^6 cells, or higher should be present in ground beef to generate positive result by qPCR due to dead microorganism. As it can be observed by the figures provided, the described method allowed to effectively detect viable bacteria in the same range of other previously published studies.

CONCLUSION

In the present study, a novel qPCR approach combining the specificity of the hydrolysis probes, along with the capacity of the intercalating dyes for result confirmation, was developed. This technique was successfully combined with a SpEn approach for the multiplex detection of *Salmonella* spp. and *E. coli* O157. The final method was successfully evaluated on inoculated chicken and ground beef samples, and demonstrated capable of detecting thermally stressed bacteria, as well as avoiding false positive results due to the presence of dead bacteria.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

AG-M envisioned the study, supervised the work, analyzed the data, and wrote the original draft of the manuscript. SA, JC, and FR conducted the experimental part. MP helped in the data analysis and revised the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the Seafood Age project, which was co-financed by the Interreg Atlantic Area Program (EAPA_758/2018) though the European Development Fund (ERDF), by the project Nanotechnology Based Functional Solutions (NORTE-01-0145-FEDER-000019), supported by Norte Portugal Regional Operational Program (NORTE2020), under the PORTUGAL 2020 Partnership Agreement, through the ERDF. SA was financed by Ph.D. grant of the Fundação para a Ciência e a Tecnologia (SFRH/BD/140396/2018).

ACKNOWLEDGMENTS

The authors would like to thank Novozymes for kindly providing the Alcalase and Neutrase used in the present study to prepare the protease buffer.

REFERENCES

Abdulmawjood, A., Bulte, M., Cook, N., Roth, S., Schonenbrucher, H., and Hoorfar, J. (2003). Toward an international standard for PCR-based detection of *Escherichia coli* O157–part 1. Assay development and multi-center validation. *J. Microbiol. Methods* 55, 775–786. doi: 10.1016/j.mimet.2003.08.012

Abdulmawjood, A., Bulte, M., Roth, S., Schonenbrucher, H., Cook, N., Heuvelink, A. E., et al. (2004). Development, validation, and standardization of polymerase chain reaction-based detection of *E. coli* O157. *J. AOAC Int.* 87, 596–603. doi: 10.1093/jaoac/87.3.596

Altman, D. G. (1991). *Practical statistics for medical research*. New York: Chapman and Hall.

Anassib, T., El-Din, M. Z., and El-Sharoud, W. M. (2003). Viability of *Salmonella enterica* subsp. *enterica* during the preparation and cold storage of Egyptian soft cheeses and ice-cream. *Int. J. Dairy Technol.* 56, 30–34. doi: 10.1046/j.1471-0307.2003.00072.x

Anderson, A., Pietsch, K., Zucker, R., Mayr, A., Muller-Hohe, E., Messelhausser, U., et al. (2011). Validation of a duplex real-time PCR for the detection of *Salmonella* spp. in different food products. *Food Anal. Methods* 4, 259–267. doi: 10.1007/s12161-010-9142-8

Andrews, W. H., Jacobson, A., and Hammack, T. (2011). Chapter 5: *Salmonella*. Available from: https://www.fda.gov/food/laboratory-methods-food/bam-chapter-5-salmonella

Barrera, O., Rodríguez-Calleja, J. M., Santos, J. A., Otero, A., and García-López, M. L. (2007). Effect of different storage conditions on *E. coli* O157:H7 and the indigenous bacterial microflora on lamb meat. *Int. J. Food Microbiol.* 115, 244–251. doi: 10.1016/j.ijfoodmicro.2006.10.053

Cangelosi, G. A., and Meschke, J. S. (2014). Dead or alive: molecular assessment of microbial viability. *Appl. Environ. Microbiol.* 80, 5884–5891. doi: 10.1128/AEM.01763-14

Chakravorty, S., Aladegbami, B., Burday, M., Levi, M., Marras, S. A. E., Shah, D., et al. (2010). Rapid universal identification of bacterial pathogens from clinical cultures by using a novel sloppy molecular beacon melting temperature signature technique. *J. Clin. Microbiol.* 48, 258–267. doi: 10.1128/JCM.01725-09

Chapela, M., Garrido-Maestu, A., and Cabado, A. G. (2015). Detection of foodborne pathogens by qPCR: a practical approach for food industry applications. *Cogent Food Agric.* 1, 1–19. doi: 10.1080/23311932.2015.1013771

Cheng, C. -M., Doran, T., Lin, W., Chen, K. -S., Williams-Hill, D., and Pamboukian, R. (2015). Interlaboratory validation for a real-time PCR *Salmonella* detection method using the ABI 7500 FAST real-time PCR system. *J. Food Prot.* 78, 1119–1124. doi: 10.4315/0362-028X.JFP-14-244

Cheng, C. M., Van, K. T., Lin, W., and Ruby, R. M. (2009). Interlaboratory validation of a real-time PCR 24-hour rapid method for detection of *Salmonella* in foods. *J. Food Prot.* 72, 945–951. doi: 10.4315/0362-028x-72.5.945

D'Urso, O. F., Poltronieri, P., Marsigliante, S., Storelli, C., Hernandez, M., and Rodriguez-Lazaro, D. (2009). A filtration-based real-time PCR method for the quantitative detection of viable *Salmonella enterica* and *Listeria monocytogenes* in food samples. *Food Microbiol.* 26, 311–316. doi: 10.1016/j.fm.2008.12.006

EFSA and ECDC (2018). The European Union summary report on trends and sources of zoonoses, zoonotic agents and food-borne outbreaks in 2017. *EFSA J.* 16:262. doi: 10.2903/j.efsa.2018.5500

Fachmann, M. S. R., Löfström, C., Hoorfar, J., Hansen, F., Christensen, J., Mansdal, S., et al. (2017). Detection of *Salmonella enterica* in meat in less than 5 hours by a low-cost and noncomplex sample preparation method. *Appl. Environ. Microbiol.* 83, e03151–e03216. doi: 10.1128/AEM.03151-16

Feng, P., Weagant, S. D., and Jinneman, K. (2011). "Chapter 5: Diarrheagenic *Escherichia coli*." Available from: https://www.fda.gov/food/laboratory-methods-food/bam-chapter-4a-diarrheagenic-escherichia-coli

Garrido-Maestu, A., Azinheiro, S., Carvalho, J., Fuciños, P., and Prado, M. (2018a). Development and evaluation of loop-mediated isothermal amplification, and recombinase polymerase amplification methodologies, for the detection of *Listeria monocytogenes* in ready-to-eat food samples. *Food Control* 86, 27–34. doi: 10.1016/j.foodcont.2017.11.006

Garrido-Maestu, A., Azinheiro, S., Carvalho, J., Fuciños, P., and Prado, M. (2020). Optimized sample treatment, combined with real-time PCR, for same-day detection of *E. coli* O157 in ground beef and leafy greens. *Food Control* 108:106790. doi: 10.1016/j.foodcont.2019.106790

Garrido-Maestu, A., Azinheiro, S., Carvalho, J., and Prado, M. (2018b). Rapid and sensitive detection of viable *Listeria monocytogenes* in food products by a filtration-based protocol and qPCR. *Food Microbiol.* 73, 253–263. doi: 10.1016/j.fm.2018.02.004

Garrido-Maestu, A., Azinheiro, S., Carvalho, J., and Prado, M. (2019). Combination of immunomagnetic separation and real-time recombinase polymerase amplification (IMS-qRPA) for specific detection of *Listeria monocytogenes* in smoked salmon samples. *J. Food Sci.* 84, 1881–1887. doi: 10.1111/1750-3841.14662 (in press).

Garrido-Maestu, A., Fuciños, P., Azinheiro, S., Carvalho, J., and Prado, M. (2017). Systematic loop-mediated isothermal amplification assays for rapid detection and characterization of *Salmonella* spp., Enteritidis and Typhimurium in food samples. *Food Control* 80, 297–306. doi: 10.1016/j.foodcont.2017.05.011

González-Escalona, N., Hammack, T. S., Russell, M., Jacobson, A. P., De Jesús, A. J., Brown, E. W., et al. (2009). Detection of live *Salmonella* sp. cells in produce by a taqman-based quantitative reverse transcriptase real-time PCR targeting invA mRNA. *Appl. Environ. Microbiol.* 75, 3714–3720. doi: 10.1128/AEM.02686-08

ISO (2002). Microbiology of food and animal feeding stuffs--Horizontal method for the detection of *Escherichia coli* O157. ISO 16654:2001, 22.

ISO (2003). Microbiology of food and animal feeding stuffs--Horizontal method for the detection of *Salmonella* spp. ISO 6579:2002.

Kokkinos, P. A., Ziros, P. G., Bellou, M., and Vantarakis, A. (2014). Loop-mediated isothermal amplification (LAMP) for the detection of *Salmonella* in food. *Food Anal. Methods* 7, 512–526. doi: 10.1007/s12161-013-9748-8

Lind, K., Ståhlberg, A., Zoric, N., and Kubista, M. (2006). Combining sequence-specific probes and DNA binding dyes in real-time PCR for specific nucleic acid quantification and melting curve analysis. *Biotechniques* 40, 315–319. doi: 10.2144/000112101

Malorny, B., Cook, N., D'Agostino, M., De Medici, D., Croci, L., Abdulmawjood, A., et al. (2004). Multicenter validation of PCR-based method for detection of *Salmonella* in chicken and pig samples. *J. AOAC Int.* 87, 861–866. doi: 10.1093/jaoac/87.4.861

Malorny, B., Hoorfar, J., Bunge, C., and Helmuth, R. (2003). Multicenter validation of the analytical accuracy of *Salmonella* PCR: towards an international standard. *Appl. Environ. Microbiol.* 69, 290–296. doi: 10.1128/aem.69.1.290-296.2003

Mann, E., Pommer, K., Mester, P., Wagner, M., and Rossmanith, P. (2014). Quantification of gram-positive bacteria: adaptation and evaluation of a preparation strategy using high amounts of clinical tissue. *BMC Vet. Res.* 10:53. doi: 10.1186/1746-6148-10-53

Marras, S. A. E., Tyagi, S., and Kramer, F. R. (2006). Real-time assays with molecular beacons and other fluorescent nucleic acid hybridization probes. *Clin. Chim. Acta* 363, 48–60. doi: 10.1016/j.cccn.2005.04.037

Mayrl, E., Roeder, B., Mester, P., Wagner, M., and Rossmanith, P. (2009). Broad range evaluation of the matrix solubilization (matrix lysis) strategy for direct enumeration of foodborne pathogens by nucleic acids technologies. *J. Food Prot.* 72, 1225–1233. doi: 10.4315/0362-028x-72.6.1225

Nagy, A., Černíková, L., Vitásková, E., Křivda, V., Dán, Á., Dirbáková, Z., et al. (2016). MeltMan: optimization, evaluation, and universal application of a qPCR system integrating the TaqMan qPCR and melting analysis into a single assay. *PLoS One* 11:e0151204. doi: 10.1371/journal.pone.0151204

NordVaL (2017). NordVal International Protocol for the validation of microbiological alternative (proprietary) methods against a reference method. 1–29.

O'Grady, J., Sedano-Balbás, S., Maher, M., Smith, T., and Barry, T. (2008). Rapid real-time PCR detection of *Listeria monocytogenes* in enriched food samples based on the ssrA gene, a novel diagnostic target. *Food Microbiol.* 25, 75–84. doi: 10.1016/j.fm.2007.07.007

Park, H. S., Marth, E. H., and Olson, N. F. (1970). Survival of *Salmonella* typhimurium in cold-pack-cheese food during refrigerated storage. *J. Milk Food Technol.* 33, 383–388. doi: 10.4315/0022-2747-33.12.383

Postollec, F., Falentin, H., Pavan, S., Combrisson, J., and Sohier, D. (2011). Recent advances in quantitative PCR (qPCR) applications in food microbiology. *Food Microbiol.* 28, 848–861. doi: 10.1016/j.fm.2011.02.008

Tomas, D., Rodrigo, A., Hernandez, M., and Ferrus, M. A. (2009). Validation of real-time PCR and enzyme-linked fluorescent assay-based methods for detection of *Salmonella* spp. in chicken feces samples. *Food Anal. Methods* 2, 180–189. doi: 10.1007/s12161-009-9082-3

Wang, R., Schmidt, J. W., Arthur, T. M., and Bosilevac, J. M. (2013). The physiologic state of *Escherichia coli* O157:H7 does not affect its detection in two commercial real-time PCR-based tests. *Food Microbiol.* 33, 205–212. doi: 10.1016/j.fm.2012.09.018

Wilrich, C., and Wilrich, P. T. (2009). Estimation of the pod function and the LOD of a qualitative microbiological measurement method. *J. AOAC Int.* 92, 1763–1772. doi: 10.1093/jaoac/92.6.1763

Zheng, J., Yang, R., Shi, M., Wu, C., Fang, X., Li, Y., et al. (2015). Rationally designed molecular beacons for bioanalytical and biomedical applications. *Chem. Soc. Rev.* 44, 3036–3055. doi: 10.1039/c5cs00020c

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Garrido-Maestu, Azinheiro, Roumani, Carvalho and Prado. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 04 November 2020
doi: 10.3389/fmicb.2020.575377

Proficiency Testing of Metagenomics-Based Detection of Food-Borne Pathogens Using a Complex Artificial Sequencing Dataset

Dirk Höper[1], Josephine Grützke[2], Annika Brinkmann[3], Joël Mossong[4], Sébastien Matamoros[5], Richard J. Ellis[6], Carlus Deneke[2], Simon H. Tausch[2], Isabel Cuesta[7], Sara Monzón[7], Miguel Juliá[7], Thomas Nordahl Petersen[8], Rene S. Hendriksen[8], Sünje J. Pamp[8], Mikael Leijon[9], Mikhayil Hakhverdyan[9], Aaron M. Walsh[10], Paul D. Cotter[10], Lakshmi Chandrasekaran[11], Moon Y. F. Tay[11], Joergen Schlundt[11], Claudia Sala[12], Alessandra De Cesare[13], Andreas Nitsche[3], Martin Beer[1] and Claudia Wylezich[1]*

[1] Institute of Diagnostic Virology, Friedrich-Loeffler-Institut, Greifswald-Insel Riems, Germany, [2] Department of Biological Safety, German Federal Institute for Risk Assessment, Berlin, Germany, [3] Centre for Biological Threats and Special Pathogens, Robert Koch Institute, Berlin, Germany, [4] Département de Microbiologie, Laboratoire National de Santé, Dudelange, Luxembourg, [5] Department of Medical Microbiology, Amsterdam UMC University of Amsterdam, Amsterdam, Netherlands, [6] Animal and Plant Health Agency, Addlestone, United Kingdom, [7] Bioinformatics Unit, Institute of Health Carlos III (ISCIII), Madrid, Spain, [8] Research Group for Genomic Epidemiology, National Food Institute, Technical University of Denmark, Lyngby, Denmark, [9] Department of Microbiology, National Veterinary Institute (SVA), Uppsala, Sweden, [10] Teagasc Food Research Centre, APC Microbiome Ireland and Vistamilk, Moorepark, Ireland, [11] Nanyang Technological University Food Technology Centre (NAFTEC), Nanyang Technological University (NTU), Singapore, Singapore, [12] Department of Physics and Astronomy, University of Bologna, Bologna, Italy, [13] Department of Veterinary Medical Sciences, University of Bologna, Bologna, Italy

OPEN ACCESS

Edited by:
David Rodriguez-Lazaro, University of Burgos, Spain

Reviewed by:
Laurent Dacheux, Institut Pasteur, France
Samantha Chandranath Karunarathna, Kunming Institute of Botany, China

***Correspondence:**
Dirk Höper
Dirk.Hoeper@fli.de

Specialty section:
This article was submitted to Food Microbiology, a section of the journal Frontiers in Microbiology

Received: *23 June 2020*
Accepted: *13 October 2020*
Published: *04 November 2020*

Citation:
Höper D, Grützke J, Brinkmann A, Mossong J, Matamoros S, Ellis RJ, Deneke C, Tausch SH, Cuesta I, Monzón S, Juliá M, Petersen TN, Hendriksen RS, Pamp SJ, Leijon M, Hakhverdyan M, Walsh AM, Cotter PD, Chandrasekaran L, Tay MYF, Schlundt J, Sala C, De Cesare A, Nitsche A, Beer M and Wylezich C (2020) Proficiency Testing of Metagenomics-Based Detection of Food-Borne Pathogens Using a Complex Artificial Sequencing Dataset. Front. Microbiol. 11:575377. doi: 10.3389/fmicb.2020.575377

Metagenomics-based high-throughput sequencing (HTS) enables comprehensive detection of all species comprised in a sample with a single assay and is becoming a standard method for outbreak investigation. However, unlike real-time PCR or serological assays, HTS datasets generated for pathogen detection do not easily provide yes/no answers. Rather, results of the taxonomic read assignment need to be assessed by trained personnel to gain information thereof. Proficiency tests are important instruments of validation, harmonization, and standardization. Within the European Union funded project COMPARE [COllaborative Management Platform for detection and Analyses of (Re-) emerging and foodborne outbreaks in Europe], we conducted a proficiency test to scrutinize the ability to assess diagnostic metagenomics data. An artificial dataset resembling shotgun sequencing of RNA from a sample of contaminated trout was provided to 12 participants with the request to provide a table with per-read taxonomic assignments at species level and a report with a summary and assessment of their findings, considering different categories like pathogen, background, or contaminations. Analysis of the read assignment tables showed that the software used reliably classified the reads taxonomically overall. However, usage of incomplete reference databases or inappropriate data pre-processing caused difficulties. From the combination of the participants' reports with their read assignments, we conclude that, although most species were detected, a number of

important taxa were not or not correctly categorized. This implies that knowledge of and awareness for potentially dangerous species and contaminations need to be improved, hence, capacity building for the interpretation of diagnostic metagenomics datasets is necessary.

Keywords: background contamination, diagnostic assessment, high-throughput sequencing, metagenomics, pathogen, proficiency test, training

INTRODUCTION

Approaches for the investigation of food-borne outbreaks regarding pathogen characterization, source attribution and risk assessment need to be precise, fast and independent from slow and biased cultivation techniques. Metagenomics-based high-throughput sequencing (HTS) is becoming a standard method for outbreak investigations of non-culturable, difficult-to-culture or slow-growing microorganisms (Koutsoumanis et al., 2019) yet protocols and analysis pipelines need to be standardized for routine use. In addition, training in result assessment and interpretation is needed for unexperienced users to be applicable as gold standard.

Problems with the analysis and the diagnostic assessment of HTS datasets may occur in several sample processing steps including sequencing, and during bioinformatics analysis. Beside the contamination of a sample during sampling and sample processing, microbial DNA can be introduced within the reagents during the preparation of sequencing libraries (Salter et al., 2014). Therefore, the specific reagent background should be known in ideal circumstances (Kirstahler et al., 2018; Wylezich et al., 2018) and at the very least, should be taken into consideration when assessing the taxa found in metagenomics datasets. Cross-contamination of multiplexed libraries is possible due to adapter swapping (Sinha et al., 2017) or carry-over between runs (Illumina, 2013; Höper et al., 2016). Furthermore, the interpreter of data should be aware of possible false-positives detected due to contaminated genomes and insufficiently curated databases (e.g., Kirstahler et al., 2018). All these points are very important when interpreting metagenomic datasets in search for possible pathogens that may be less abundant in terms of sequencing reads.

Some ring trials and proficiency tests have previously been conducted to push forward the standardization of HTS approaches and their implementation into clinical diagnostic routine and diagnosis reporting systems. Metagenomics-based proficiency tests for pathogen detection have often been focused on virus detection (Brinkmann et al., 2019; Junier et al., 2019; Zamperin et al., 2019). These studies highlighted that recognition of viruses that exhibit high mutation rates can cause some difficulties and might only be possible by inclusion of protein-based alignment approaches (Brinkmann et al., 2019). In addition, the undetected pathogen of relevance can also be a bacterial or parasite pathogen, which could be comprised in the sample but be masked by the sample background. Thus, interpreters need to train their capability recognizing the different categories of contaminations.

To some extent, the wet lab procedures for metagenomics-based pathogen detection were already standardized within the European Union funded COMPARE-network.[1] Within this framework, a first metagenomics-based virus detection proficiency test with separate parts addressing bioinformatics only or sample processing combined with bioinformatics and result assessment, respectively, was conducted (Brinkmann et al., 2019). Due to the importance of foods in the transmission of zoonotic agents, a second proficiency test for metagenomics-based pathogen detection in food was organized in 2018 within the COMPARE network. Again, this proficiency test consisted of two independent parts addressing either the sample processing (wet lab) or bioinformatics combined with result assessment (dry lab). In the laboratory part (wet lab), the participants had to generate a metagenomic sequence dataset from a piece of smoked salmon that was spiked with a complex mock community (bacteria, fungi, eukaryotic parasite, and virus). The wet lab part was analyzed and is published separately in a companion manuscript (Sala et al., Unpublished). In the present study, the bioinformatics and assessment part (dry lab), the participants had to analyze a synthetic dataset and assess their obtained result. While here the analysis of the same dataset by all participants ensured the comparability of the submitted results, in the wet lab part, the comparison of the wet lab workflows was ensured by the centralized sequence data analysis. Moreover, participants were free to decide in which of the two independent parts they participated.

Whereas in the first COMPARE virus proficiency test (Brinkmann et al., 2019) the quality of the data analysis software was in the focus, in the present proficiency test we focused on the participants assessment of the results obtained by the software. Hence, the purpose was (i) testing the interpretation of results obtained from the software analysis by the participants, i.e., the recognition of potentially dangerous species and (ii) the awareness of artifacts occurring in the sample processing and sequencing.

MATERIALS AND METHODS

Organization of the Dry Lab Proficiency Test

Like the COMPARE virus proficiency test (Brinkmann et al., 2019), the food metagenomics proficiency test dry lab part was initiated within the COMPARE network, and

[1] https://www.compare-europe.eu/library/protocols-and-sops

arranged by the network partner Friedrich-Loeffler-Institut between April and June 2018. The participants of the dry lab part of the food metagenomics proficiency test received a synthetically generated complex metagenomics sequencing dataset that they had to analyze and assess their obtained results regarding potentially present pathogens, sequencing artifacts, potential lab contaminations, and other information deemed relevant for assessment.

Participants of the Dry Lab Part

Twelve participants applied for the COMPARE food metagenomics dry lab proficiency test and completed the survey. Participants were registered from Denmark (n = 1), Germany (n = 3), Luxembourg (n = 1), Ireland (n = 1), Italy (n = 1), Netherlands (n = 1), Singapore (n = 1), Spain (n = 1), Sweden (n = 1), and United Kingdom (n = 1). The 12 participants represented 12 different institutes or organizations. Information about the participants' background is given in **Table 2**.

Creating a Synthetic Metagenomics Dataset

The synthetic metagenomics dataset was created using ART_Illumina, Q Version 2.5.8 (Huang et al., 2012) and a number of EST (expressed sequence tag) and coding sequence data retrieved from the NCBI database (see **Table 1**). For details of the dataset creation, please refer to **Supplementary Materials and Methods**.

Requirements to Participants

In parallel with the dry lab synthetic dataset, an email (see email 1 in the **Supplementary Materials and Methods**) was sent to all participants that included the necessary instructions for the data and assessment preparation and delivery. The requirements were later further specified (see email 2 in the **Supplementary Materials and Methods**). Briefly, the participants were asked to use their routine bioinformatics for taxonomic read classification and to submit a table in which the taxonomic classification for all reads was recorded in two columns ("Read accession" and "Species"). Hereinafter, this table is called "read assignment table." Furthermore, they were asked to prepare a report as Word file or pdf containing their summary and assessment of the read classifications. They were asked to especially consider potentially present pathogens, sequencing artifacts, possible sequencing lab contaminations, and other facts they deemed important. This document is called "summary and assessment file" in the following.

The participants had about 50 days (25 April–15 June 2018) for conducting a bioinformatics analysis with the dataset provided and for assessing the results from obtained data until the delivery of their results.

Data Analyses

The submitted read assignment tables were analyzed using R and RStudio (R v3.6.2, R Core Team, 2019; RStudio v1.2.5033).[2] For an assessment of the taxonomic classifications reported in the participants' read assignment tables, sensitivity, specificity, correct classification rate, positive predictive value, and negative predictive value were calculated from the read-to-taxon assignments. For further details of these calculations, please see **Supplementary Materials and Methods**. For the final rating of the participants' assessments, the expected assessments were defined as shown in **Table 1**. The submitted assessments were rated in five different classes: (i) species detected and rated by the participant as expected, (ii) species detected and rated more serious than expected, (iii) species detected but assessed less critical than expected, (iv) species detected but not assessed, and (v) species not detected (according to the submitted read assignment table).

[2]https://rstudio.com/

Data Availability

The synthetic metagenomics dataset simulating a contaminated trout was submitted to European Nucleotide Archive and is accessible under the study accession number PRJEB37463. The read-accessions in this file comprise the database identifiers of the used reference sequence and the suffix "-fpt2018" with a continuous numbering per input sequence, for instance "lcl|NC_026023.1_cds_YP_009113336.1_1-fpt2018105".

RESULTS AND DISCUSSION

The aim of the present proficiency test was in particular to test the interpretation of results obtained from metagenomics sequencing datasets based on the software analysis performed by the participants. For this purpose, the detection of potentially dangerous species was most important but also the awareness and recognition of artifacts that may occur during the wet lab sample processing and sequencing.

General Considerations of the Provided Files

All participants except one (P10) uploaded the requested read assignment table with the read to species assignments output by their applied software pipeline, eight of which in due time, and three with a delay. All participants sent in a summary and assessment file of their obtained results (see **Supplementary Material**).

The format requested for the preparation of the read assignment tables (see above and in **Supplementary Materials and Methods**) turned out to be not clear for all participants. Six participants sent in incomplete tables, only reporting assignments for a small fraction of reads (see **Figure 1**). Moreover, only five participants sent in their read assignment tables in the requested format. In the remaining cases, participants reported more details with strain or isolate names or incomplete assignments only to a level between superkingdom and genus. Especially the users of Kraken did not report in the requested format but used the Kraken format including the full taxonomic path up to the most specific level reported. A similar observation was recently made by the organizers of another ring trial, who stated that the participants did not in all cases use official scientific names. They emphasized the importance of using a standardized set of species names based on NCBI taxonomy for reporting (Junier et al., 2019).

TABLE 1 | Composition of the simulated sequence dataset.

Species	Super-Kingdom	Category	Rationale	Number Reads
Bacteroides fragilis	Bacteria	Opportunistic pathogen	Analogous to wet-lab proficiency test mock community	20,000
Burkholderia pseudomallei	Bacteria	Pathogen	Burkholderiaceae also found in real smoked salmon sample	8,000
Escherichia coli	Bacteria	Pathogen	Analogous to wet-lab proficiency test mock community, also found in real smoked salmon sample	80,000
Salmonella enterica	Bacteria	Pathogen	Analogous to wet-lab proficiency test mock community	125,000
Fusobacterium nucleatum	Bacteria	Opportunistic pathogen	Analogous to wet-lab proficiency test mock community	40,000
Lactobacillus acidophilus	Bacteria	Background	Food additive	175,000
Lactobacillus delbrueckii	Bacteria	Background		25,000
Listeria monocytogenes	Bacteria	Pathogen	Frequently found as food contamination	10,000
Mycobacterium colombiense	Bacteria	Opportunistic pathogen	Mycobacteria with increasing impact as food spoilage	2,000
Pseudomonas libanensis	Bacteria	Background	*P. libanensis* previously detected in food; Pseudomonadaceae also found in real smoked salmon sample	200
Anisakis berlandi	Eukaryota	Pathogen	Analogous to *Cryptosporidium parvum* from wet-lab proficiency test mock community	1,212
Anisakis brevispiculata	Eukaryota	Pathogen		246
Anisakis paggiae	Eukaryota	Pathogen		248
Anisakis pegreffii	Eukaryota	Pathogen		2,051
Anisakis physeteris	Eukaryota	Pathogen		262
Anisakis simplex	Eukaryota	Pathogen		6,044
Anisakis typica	Eukaryota	Pathogen		247
Aspergillus flavus	Eukaryota	Opportunistic pathogen	Analogous to *Saccharomyces cerevisiae* from wet-lab proficiency test mock community, toxin producer	5,000
Danio rerio	Eukaryota	Database misclassification		4,310
Brugia malayi	Eukaryota			64
Caenorhabditis remanei	Eukaryota			193
Scomber japonicus	Eukaryota			181
Oncorhynchus mykiss	Eukaryota	Food	*O. mykiss* EST data as host background	9,451,675
African swine fever virus (ASFV)	Viruses	Run contamination	DNA virus; barcode mis-alignment/index swapping on Illumina MiSeq	15
Norwalk virus	Viruses	Pathogen	RNA virus; typical food contaminant	946
Escherichia virus phiX174 (phiX174)	Viruses	Run contamination	Barcode mis-alignment/index swapping on Illumina MiSeq	735
Aspergillus foetidus dsRNA mycovirus (AfdsV)	Viruses	Background	Model for virus of the contaminant *Aspergillus flavus*	107
Sum				9,958,736

The summary and assessment files of four participants (P2, P4, P5, and P10) only contained the requested summary table without an assessment. The remaining eight participants provided both the requested summary table and an assessment of the reported results.

Insights From Analysis of the Read Assignment Tables

Of the uploaded 11 read assignment tables, only five (from P1, P3, P4, P8, and P11) contained an assignment for all reads of the dataset; the remaining six contained assignments for approximately 3–18% of the reads (**Figure 1**). The reasons not to report assignments for the missing reads were (i) filtered low quality (P2, P6, P7, and P9), (ii) filtered eukaryotic sequences (P2, P5, P6, and P9), or (iii) no justification (P12). In two cases (P1 and P3; see **Figure 1**) the majority of the reads (approximately 95%) was reported "unclassified." In one case (P1), this was partly due to incomplete classification to higher-level taxonomic entities only. Three participants (P4, P8, and P11; **Figure 1**) reported classifications for the majority of the reads and only in these cases the overall composition resembled the actual one (**Figure 1**).

The provided read assignments were used to calculate the key characteristics of the classifications. Namely, specificity, correct classification rate, and negative predictive values (**Supplementary Figures 1–3**), as well as positive predictive value (**Figure 2**) and sensitivity (**Figure 3**) were determined. While the specificity, correct classification rate (except for host sequences),

TABLE 2 | Participants, their background and applied data processing and important comments from their summaries.

Participants	Participants' sector	Data processing workflow	Participants comments
P1	Food	– Quality assessment w/FastQC (Wingett and Andrews, 2018) – Quality trimming w/fastp (Chen et al., 2018) – Taxonomic classification w/Kraken (custom database and MiniKraken DB; Wood and Salzberg, 2014) – Additionally pathoLive analysis for classification of viral reads (Tausch et al., 2018)	– FastQC revealed bases of bad quality at the beginning of the reads. Therefore the reads were trimmed – kraken analysis with custom database resulted in many false-positive results; therefore, results were confirmed with BLASTn (Boratyn et al., 2012).
P2	Human	– Quality assessment w/FastQC (Wingett and Andrews, 2018) – Quality trimming w/Trimmomatic (Bolger et al., 2014) – Taxonomic classification based on mapping and assembly w/Pikavirus (in-house in-development tool at https://github.com/BU-ISCIII/PikaVirus) – Taxonomic classification based on mapping and assembly w/oases (Schulz et al., 2012) – Taxonomic classification based on rRNA clustering w/MeTRS (Cottier et al., 2018) – Taxonomic classification based on protein identity analysis w/Kaiju (Menzel et al., 2016)	– trimming parameters: nucleotides at 3′ with phred quality <10 or average quality ≤15 (window size 4), removal of reads shorter 50 bp – trimming dropped 8137323 sequences (81.71%) – unusual bad quality 5′ end was observed at the 25 firsts bases
P3	Human	– Taxonomic classification w/Kraken (Wood and Salzberg, 2014) as implemented on Galaxy public server	– Norovirus GV (murine norovirus, not a human pathogen) Unlikely to be on food sample – Hepatitis C virus (human pathogen, but route of transmission is via blood), highly unlikely to be found on food sample, and contamination with human blood?
P4	Food	– Quality assessment w/FastQC (Wingett and Andrews, 2018) – Quality/Adapter trimming w/BBduk v. 36.49 (https://sourceforge.net/projects/bbmap/) – Taxonomic classification w/MGmapper (Petersen et al., 2017)	– mapped against a phiX174 reference sequence to remove potential control library reads – Certain contaminations can be difficult to identify without control samples. For example, certain microorganisms may be part of the natural microbiome of fish or could have been introduced during sample handling and processing.
P5	Veterinary	– Quality assessment w/FastQC (Wingett and Andrews, 2018) – Quality/Adapter trimming w/Trimmomatic (Bolger et al., 2014) – Host sequence removal w/BWA-MEM (Li, 2013) – Taxonomic classification w/Kraken (MiniKraken database; ref. Wood and Salzberg, 2014)	– Filtered for minimum read count (threshold 500 reads) – species for which there were less than 1,000 reads would need further confirmation before release of the information – PhiX carry over from the sequencing lab
P6	Human	– Quality assessment w/FastQC (Wingett and Andrews, 2018) – Quality trimming w/Trimmomatic (Bolger et al., 2014) – Host sequence removal w/BBmap (id threshold 0.65; https://sourceforge.net/projects/bbmap/) – Taxonomic classification w/Kraken (database version 13/10/2017; Wood and Salzberg, 2014)	– FastQC result: No adapters detected
P7	Human	– Quality/Adapter trimming w/Trimmomatic (Bolger et al., 2014) – Host sequence removal w/Bowtie2 (Langmead and Salzberg, 2012) – Taxonomic classification w/MALT (Herbig et al., 2016), DIAMOND (Buchfink et al., 2015), MEGAN (Huson et al., 2016), custom database with refseq viruses, bacterial, fungi, and protists	– Pathogen of importance is Norovirus GV – Abundance of a cloning vector could be an artifact of sequencing reagents and preparation
P8	Veterinary	– Quality trimming w/RIEMS (Scheuch et al., 2015) – Taxonomic classification w/RIEMS (Scheuch et al., 2015); ncbi nt	– The Caliciviridae/Norwalk virus reads indicate the presence of noroviruses in the sample. This is the most important entero-pathogenic virus in the analyzed sample
P9	Food	– Species-level classification w/Kraken (Wood and Salzberg, 2014) – Multi-locus sequence types (MLSTs) reconstructed w/MetaMLST (Zolfo et al., 2017) – Strain-level identification w/PanPhlAn (Scholz et al., 2016)	None
P10	Food	– Quality assessment w/FastQC (Wingett and Andrews, 2018) – Quality trimming w/cutadapt (part of MGmapper processing; (Martin, 2011) – Taxonomic classification w/MGmapper (Petersen et al., 2017)	– no hits w/default settings, re-analysis w/adjusted parameters (max mismatch ratio = 0.15, min read count = 20) – sequence GC content measured by FastQC is reported as failure

(Continued)

TABLE 2 | Continued

Participants	Participants' sector	Data processing workflow	Participants comments
P11	Veterinary	– Host sequences were removed by blasting (BLASTn) against a database created from the *Oncorhynchus mykiss* isolate Swanson WGS data (NCBI acc. MSJN00000000.1) using an E-value cutoff 1E-100 – Taxonomic classification w/carried out by blasting (BLASTn) the remaining reads against an NCBI nt database using TimeLogic® DeCypher® server (Active Motif Inc., Carlsbad, CA, United States) with an E-value cutoff of 1E-5. The assignment of sequences to species were carried out by an in-house Python script using the nucl_gb.accessions2taxid (accession to taxid) and names.dmp (taxid to scientific names) files available from the resources at NCBI	– We included Murine norovirus in the table despite it is not human pathogen
P12	Veterinary/Food	– Quality assessment w/FastQC (Wingett and Andrews, 2018) – Quality/Adapter trimming w/cutadapt (part of MGmapper processing) (Martin, 2011) Taxonomic classification w/MGmapper (Databases: Bacteria, Bacteria_draft, Human Microbiome, Virus, Fungi, Protozoa,and MetaHitAssembly) (Petersen et al., 2017)	None

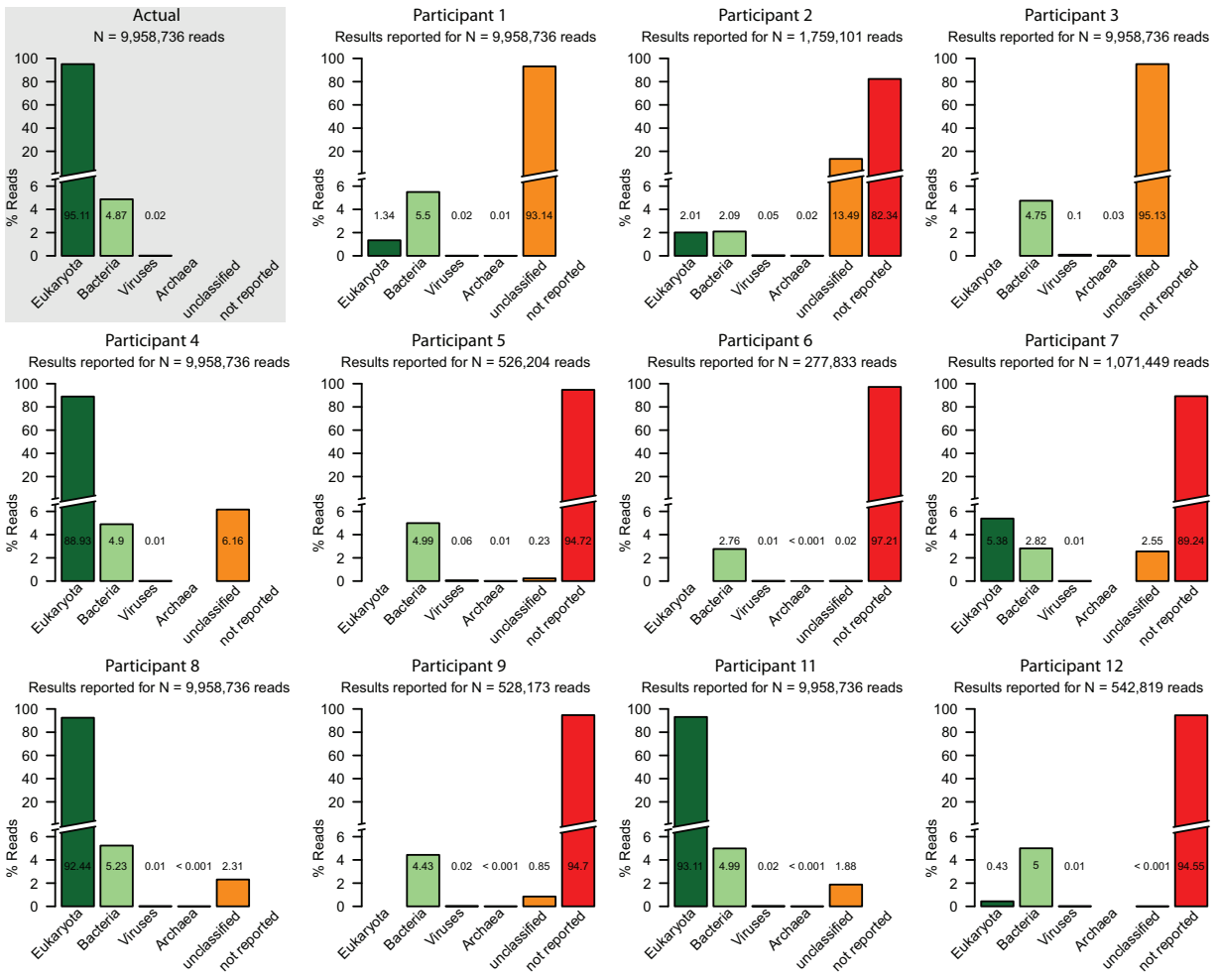


FIGURE 1 | Summary of the read assignments at the superkingdom-level for 11 participants that provided the requested read assignment table with the read to species assignments (except P10). Only five read assignment tables (P1, P3, P4, P8, and P11) contained an assignment for all reads of the dataset. Only the compositions reported by P4, P8, and P11 fit the known actual composition (actual; upper left) of the dataset.

and negative predictive value (except for host sequences) were in all analyses high, the positive predictive value (**Figure 2**) and especially the sensitivity (**Figure 3**) were in some cases insufficient. Both the positive predictive value and the sensitivity appear to be compromised by either the use of incomplete databases for the taxonomic classification and/or by improper

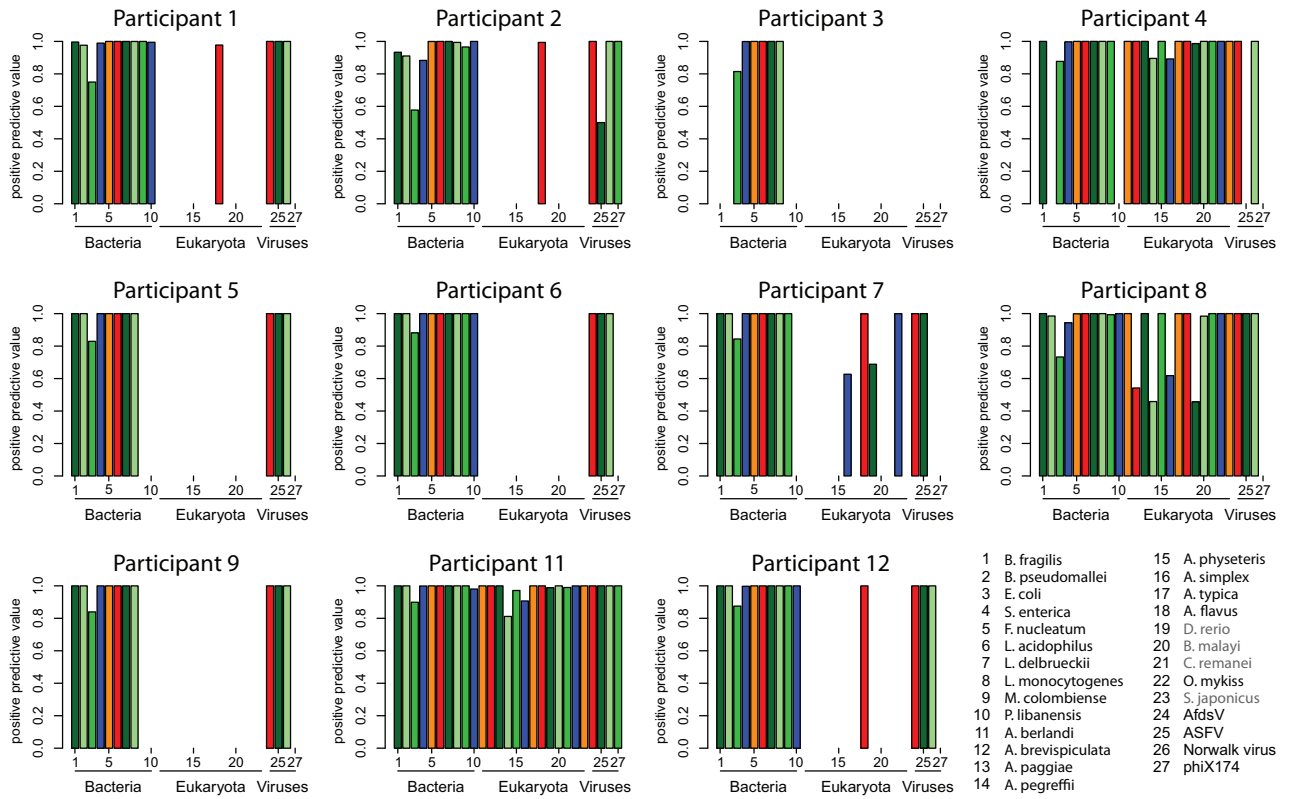


FIGURE 2 | Positive predictive values of read assignments calculated from the complete read set, calculated based on the species assignments. Sequences of the taxa labeled gray (*Brugia malayi*, *Caenorhabditis remanei*, *Danio rerio*, and *Scomber japonicus*) were downloaded unintentionally as part of the *Anisakis* sequence dataset.

pre-processing of the dataset (compare data analysis workflows and comments in **Table 2**). In three cases (P2, P6, and P7), pre-processing using software default or otherwise accepted parameters appears to have removed reads by chance, because the sensitivity is at the same low level for all detected species, meaning that this was not a species-specific effect (see results for participants P2, P6, and P7 in **Figure 3**). The high rate of discarded reads could have prevailed the respective participants to have a closer look at the reason for this phenomenon and modify the software settings depending on the dataset quality. The notion that this was not a species-specific effect is also emphasized by the high specificity mentioned above and by the unaltered sensitivity in case of calculation at the genus, instead the species level (see **Figure 3B**). If it was a species-specific effect, this should result in an improved sensitivity because then reads that are classified to closely related species should improve the result, as can indeed be seen by comparison of results obtained for participants P8 and P11 (compare **Figure 3B**). This was possibly the case with *Pseudomonas libanensis* in the results of participants P3, P4, P7, and P9 (**Figure 3B**).

Despite the partly suboptimal results calculated from the submitted complete classifications, **Figure 4A** clearly demonstrates that in most cases the majority of the comprised species were detected by the used software. A marked exception were those species that were not deliberately included in the dataset, namely *Scomber japonicus*, *Brugia malayi*, *Danio rerio*, and *Caenorhabditis remanei*. These were only detected by three participants (P4, P8, and P11) who used (nearly) complete databases for the taxonomic classification. All other participants reported to have used custom databases or the MiniKraken database that also comprises only selected sequences (Wood and Salzberg, 2014). This database effect is also highlighted by the fact that at least seven and five participants, respectively, failed to detect *Anisakis simplex* and *Aspergillus flavus*, two eukaryotic taxa. This emphasizes the impact of the database used for taxonomic classification to obtain a comprehensive classification. This result is in contrast to the results of another proficiency test (Junier et al., 2019), where the impact of the database was negligible compared with the influence of the applied algorithms.

Insights From Evaluations of Participants' Summary and Assessment Files

In the present proficiency test, the assessment of the results turned out to be the most critical part. Overall, the quality of the reports was varying. Although explicitly requested, the results of the taxonomic binning were not in all cases assessed regarding the requested categories (potentially present pathogens, sequencing artifacts, possible sequencing lab contaminations, and other important facts). Looking at the summary and assessment files, there was no correlation between the overall quality of the

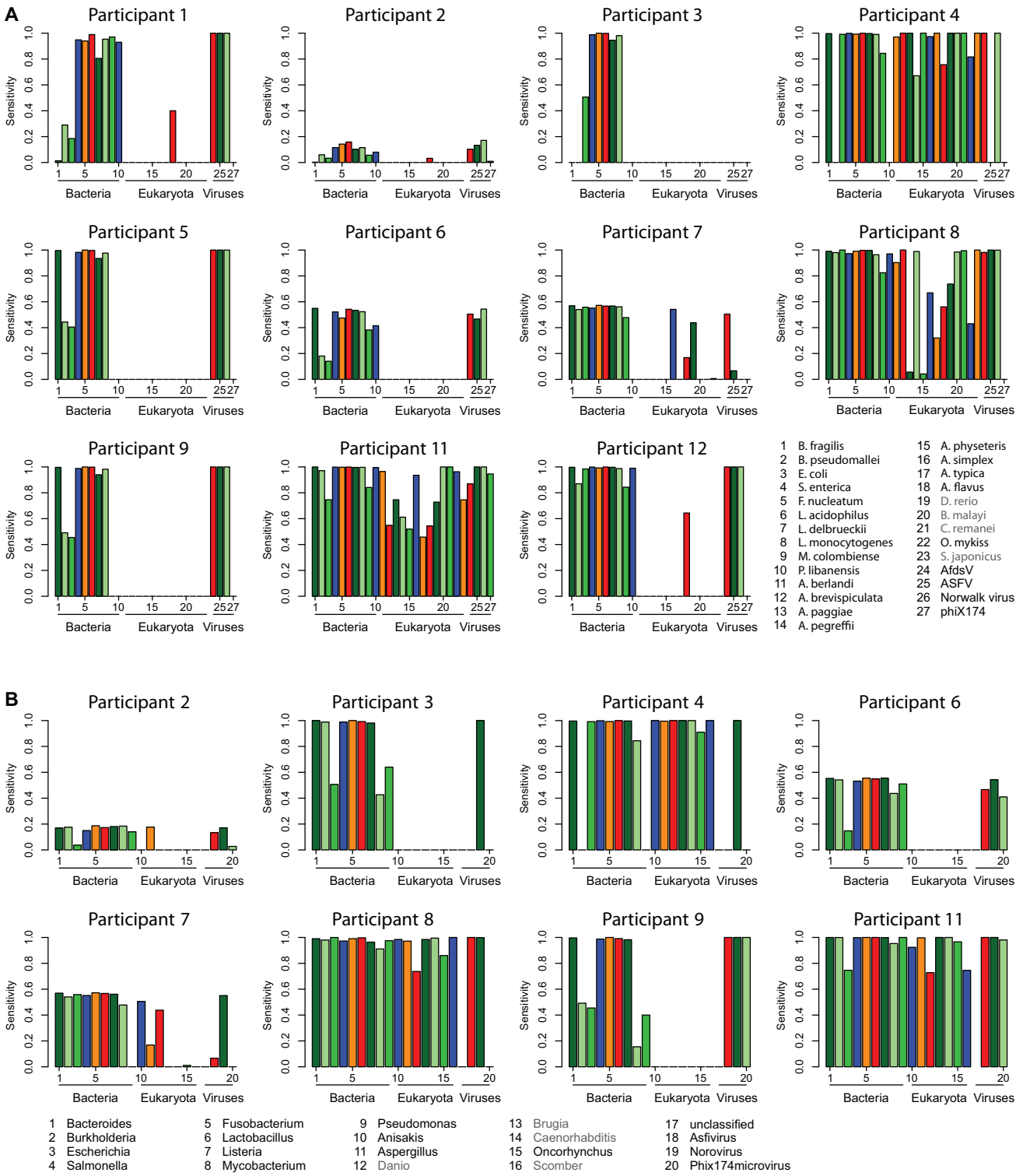


FIGURE 3 | Sensitivity of read assignments calculated from the complete read set. **(A)** Sensitivities calculated based on the species assignments. **(B)** Sensitivities calculated based on the genus assignments. Sequences of the taxa labeled gray (*Brugia malayi*, *Caenorhabditis remanei*, *Danio rerio*, and *Scomber japonicus*) were downloaded unintentionally as part of the *Anisakis* sequence dataset.

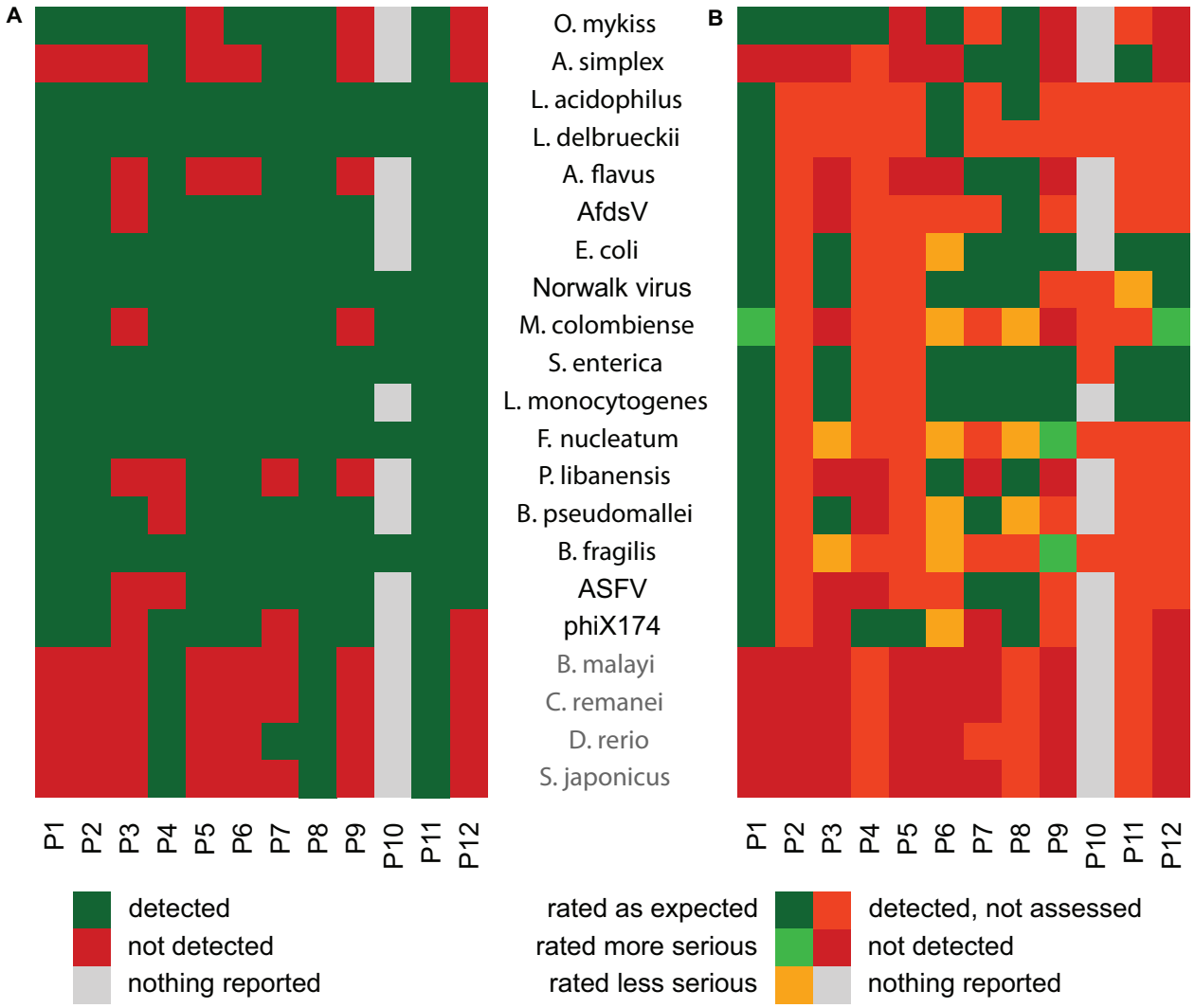


FIGURE 4 | Summary of the read-assignment tables **(A)**, and assessment and interpretation of the assignments **(B)**. Heatmap showing the positive (comprised species detected; green) and negative (comprised species NOT detected; red) results of the software analyses. The results shown for participants are based on their uploaded read assignment tables, except for P10, for which the results are derived from their summary table and assessment **(A)**. Heatmap summarizing assessments of the detected species by the participants **(B)**. Sequences of the taxa labeled gray (*Brugia malayi*, *Caenorhabditis remanei*, *Danio rerio*, and *Scomber japonicus*) were downloaded unintentionally as part of the *Anisakis* sequence dataset.

assessment and the background of the participant (compare **Table 2**). Though most species were detected (**Figure 4A**), the overall result was suboptimal (see summary in **Figure 4B**). For this assessment, the participants' assessments were compared with the expected categorization of the respective species (compare **Table 1**). The reasons for the observed deviations may be diverse, located both at the technical level and at the individual experiences of the personnel.

At the technical level, unsuitable parametrization of the analysis may be a possible cause for missing important species from the result the diagnostician gets for the assessment, i.e., arbitrary thresholds for taking detected species into consideration. For instance, thresholds set for read numbers assigned to a single species can prevent detection, as was the case with P5 (reporting a minimum read number of 500 for each individual species) failing to recognize the African swine fever virus (15 reads) and the Aspergillus foetidus dsRNA mycovirus (107 reads). Moreover, minimum genome coverage of detected species, or too stringent cut-offs for the identity of reads with reference sequences may prevent species from appearing in the table of the detected species. Another technical issue was the use of incomplete databases for taxonomic binning (see above).

With regard to the individual experience of the personnel, a number of different reasons could be considered. In two cases (P3 and P11), the assessment of Norovirus was based on the detected viral strain, which was closest to murine strains and therefore assessed as "no human pathogen". However, indications of potential pathogens should always be followed

up, because of the possibility of detecting a modified or novel pathogen with only weak relationship with the closest known relative in the database. In other cases, assessment of the software output by personnel not trained for this task (e.g., assessment by bioinformaticians instead of microbiologists, physicians, or veterinarians), insufficient awareness of the impact of certain species due to insufficient training, or maybe unsatisfactory consideration due to time constraints may have caused the result.

CONCLUSION

The dry-lab part of this ring-trial showed that despite the abovementioned shortcomings in some analyses (namely usage of incomplete databases or unsuitable data pre-processing), overall the used software appears to have matured over the last years to allow for the correct identification of the majority of organisms represented in a metagenomics dataset. However, for a truly beneficial effect of diagnostic metagenomics for the detection of potentially present pathogens, it is especially necessary to put more effort into the training for the assessment and interpretation of the results delivered by the different software pipelines for the analysis of metagenomics data.

Two additional points should be stressed. First, in this proficiency test dataset, we included African swine fever virus and Escherichia virus phiX174 sequences as within run contaminants, which only three and four participants, respectively, correctly assessed. Noteworthy, the same effect frequently occurs in real sequencing runs [Illumina, 2013 (between runs), Sinha et al., 2017; Illumina, 2017 (within run)]. Therefore, knowledge of the content of samples from the same and previous runs might be necessary to take into account, information that was not available for this proficiency test. Most importantly, however, awareness to the problem needs to be raised. Second, the interaction between the different sectors (human, veterinary, and food) and the disciplines within these, e.g., virology, bacteriology, parasitology, needs to be strengthened. This must necessarily include enhanced awareness of the pathogens of importance for other sectors and disciplines, including reporting to the respective colleagues in case relevant pathogens are detected.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: data are available from https://www.ebi.ac.uk/ebisearch/overview.ebi/about under accession PRJEB37463.

AUTHOR CONTRIBUTIONS

DH and CW conceived the study, analyzed data, and wrote the manuscript. DH created the artificial dataset. JG, AB, JM, SMa, RE, CD, ST, IC, SMo, MJ, TP, RH, SP, ML, MH, AW, PC, LC, MT, JS, CS, AD, AN, and MB analyzed the artificial dataset and assessed the results. All authors revised and approved the manuscript.

FUNDING

This research was supported by the EU funded project COMPARE (Grant Agreement No. 643476). JG was supported by a grant of the German Federal Ministry of Education within the framework of the project Ess-B.A.R. (FKZ 13N13982).

ACKNOWLEDGMENTS

We thank Simone M. Cacciò for helpful discussions regarding eukaryotic microorganisms.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2020.575377/full#supplementary-material

Supplementary Figure 1 | Specificities of read assignments calculated from the complete read set based on the species assignments. Sequences of the taxa labeled gray (*Brugia malayi*, *Caenorhabditis remanei*, *Danio rerio*, and *Scomber japonicus*) were downloaded unintentionally as part of the *Anisakis* sequence dataset.

Supplementary Figure 2 | Correct classification rates of read assignments calculated from the complete read set based on the species assignments. Sequences of the taxa labeled gray (*Caenorhabditis remanei*, *Brugia malayi*, *Danio rerio*, and *Scomber japonicus*) were downloaded unintentionally as part of the *Anisakis* sequence dataset.

Supplementary Figure 3 | Negative predictive values of read assignments calculated from the complete read set based on the species assignments. Sequences of the taxa labeled gray (*Caenorhabditis remanei*, *Brugia malayi*, *Danio rerio*, and *Scomber japonicus*) were downloaded unintentionally as part of the *Anisakis* sequence dataset.

Supplementary Materials and Methods | Emails with requirements to participants and details on the calculation of sensitivity, specificity, correct classification rate, positive predictive value, and negative predictive value are presented.

Summary and Assessment of the Participants | All summary and assessment files sent in by the participants are included.

REFERENCES

Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. doi: 10.1093/bioinformatics/btu170

Boratyn, G. M., Schäffer, A. A., Agarwala, R., Altschul, S. F., Lipman, D. J., and Madden, T. L. (2012). Domain enhanced lookup time accelerated BLAST. *Biol. Direct.* 7:12. doi: 10.1186/1745-6150-7-12

Brinkmann, A., Andrusch, A., Belka, A., Wylezich, C., Höper, D., Pohlmann, A., et al. (2019). Proficiency testing of virus diagnostics based on bioinformatics analysis of simulated in silico high-throughput sequencing datasets. *J. Clin. Microbiol.* 57:e00466-19. doi: 10.1128/JCM.00466-19

Buchfink, B., Xie, C., and Huson, D. H. (2015). Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* 12, 59–60. doi: 10.1038/nmeth.3176

Chen, S., Zhou, Y., Chen, Y., and Gu, J. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34, i884–i890. doi: 10.1093/bioinformatics/bty560

Cottier, F., Srinivasan, K. G., Yurieva, M., Liao, W., Poidinger, M., Zolezzi, F., et al. (2018). Advantages of meta-total RNA sequencing (MeTRS) over shotgun metagenomics and amplicon-based sequencing in the profiling of complex microbial communities. *NPJ Biofilms Microbiomes*. 4:2. doi: 10.1038/s41522-017-0046-x

Herbig, A., Maixner, F., Bos, K. I., Zink, A., Krause, J., and Huson, D. H. (2016). MALT: fast alignment and analysis of metagenomic DNA sequence data applied to the Tyrolean Iceman. *bioRxiv* [Preprint]. doi: 10.1101/050559

Höper, D., Mettenleiter, T. C., and Beer, M. (2016). Metagenomic approaches to identifying infectious agents. *Rev. Sci. Tech.* 35, 83–93. doi: 10.20506/rst.35.1.2419

Huang, W., Li, L., Myers, J. R., and Marth, G. T. (2012). ART: a next-generation sequencing read simulator. *Bioinformatics* 28, 593–594. doi: 10.1093/bioinformatics/btr708

Huson, D. H., Beier, S., Flade, I., Górska, A., El-Hadidi, M., Mitra, S., et al. (2016). MEGAN community edition - Interactive exploration and analysis of large-scale microbiome sequencing data. *PLoS Comput. Biol.* 12:e1004957. doi: 10.1371/journal.pcbi.1004957

Illumina (2013). *Reducing Run-To-Run Carryover on the MiSeq Using Dilute Sodium Hypochlorite Solution*. Illumina: San Diego, CA.

Illumina (2017). *NextSeq™ 550Dx Instrument. Document # 1000000043133 v00*. Illumina: San Diego, CA.

Junier, T., Huber, M., Schmutz, S., Kufner, V., Zagordi, O., Neuenschwander, S., et al. (2019). Viral metagenomics in the clinical realm: lessons learned from a Swiss-wide ring trial. *Genes* 655:10. doi: 10.3390/genes10090655

Kirstahler, P., Bjerrum, S. S., Friis-Møller, A., la Cour, M., Aarestrup, F. M., Westh, H., et al. (2018). Genomics-based identification of microorganisms in human ocular body fluid. *Sci. Rep.* 8:4126. doi: 10.1038/s41598-018-22416-4

Koutsoumanis, K., Allende, A., Alvarez-Ordoñez, A., Bolton, D., Bover-Cid, S., Chemaly, M., et al. (2019). Scientific Opinion on the whole genome sequencing and metagenomics for outbreak investigation, source attribution and risk assessment of food-borne microorganisms. *EFSA J.* 17:5898. doi: 10.2903/j.efsa.2019.5898

Langmead, B., and Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359. doi: 10.1038/nmeth.1923

Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv* [Preprint]. Available online at: https://arxiv.org/abs/1303.3997 (accessed April 25, 2018).

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* 17, 10–12. doi: 10.14806/ej.17.1.200

Menzel, P., Ng, K. L., and Krogh, A. (2016). Fast and sensitive taxonomic classification for metagenomics with Kaiju. *Nat. Commun.* 7:11257. doi: 10.1038/ncomms11257

Petersen, T. N., Lukjancenko, O., Thomsen, M., Sperotto, M. M., Lund, O., Aarestrup, F. M., et al. (2017). MGmapper: reference based mapping and taxonomy annotation of metagenomics sequence reads. *PLoS One* 12:e0176469. doi: 10.1371/journal.pone.0176469

R Core Team (2019). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Salter, S. J., Cox, M. J., Turek, E. M., Calus, S. T., Cookson, W. O., Moffatt, M. F., et al. (2014). Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biol.* 12:87. doi: 10.1186/s12915-014-0087-z

Scheuch, M., Höper, D., and Beer, M. (2015). RIEMS: a software pipeline for sensitive and comprehensive taxonomic classification of reads from metagenomics datasets. *BMC Bioinf.* 16:69. doi: 10.1186/s12859-015-0503-6

Scholz, M., Ward, D., Pasolli, E., Tolio, T., Zolfo, M., Asnicar, F., et al. (2016). Strain-level microbial epidemiology and population genomics from shotgun metagenomics. *Nat. Methods* 13, 435–438. doi: 10.1038/nmeth.3802

Schulz, M. H., Zerbino, D. R., Vingron, M., and Birney, E. (2012). Oases: robust de novo RNA-seq assembly across the dynamic range of expression levels. *Bioinformatics* 28, 1086–1092. doi: 10.1093/bioinformatics/bts094

Sinha, R., Stanley, G., Gulati, G. S., Ezran, C., Travaglini, K. J., Wie, E., et al. (2017). Index switching causes "spreading-of-signal" among multiplexed samples in Illumina HiSeq 4000 DNA Sequencing. *BioRxiv* [Preprint]. Available online at: https://www.biorxiv.org/content/early/2017/04/09/125724 (accessed June 23, 2020).

Tausch, S. H., Loka, T. P., Schulze, J. M., Andrusch, A., Klenner, J., Dabrowski, P. W., et al. (2018). PathoLive - Real time pathogen identification from metagenomic Illumina datasets. *BioRxiv* [Preprint]. doi: 10.1101/402370

Wingett, S. W., and Andrews, S. (2018). FastQ screen: a tool for multi-genome mapping and quality control. *F1000Research* 7:1338. doi: 10.12688/f1000research.15931.2

Wood, D. E., and Salzberg, S. L. (2014). Kraken: ultrafast metagenomic sequence classification using exact alignments. *Genome Biol.* 15:R46. doi: 10.1186/gb-2014-15-3-r46

Wylezich, C., Papa, A., Beer, M., and Höper, D. (2018). A versatile sample processing workflow for diagnostic metagenomics. *Sci. Rep.* 8:13108. doi: 10.1038/s41598-018-31496-1

Zamperin, G., Lucas, P., Cano, I., Ryder, D., Abbadi, M., Stone, D., et al. (2019). Sequencing of animal viruses: quality data assurance for NGS bioinformatics. *Virol. J.* 16:140. doi: 10.1186/s12985-019-1223-8

Zolfo, M., Tett, A., Jousson, O., Donati, C., and Segata, N. (2017). MetaMLST: multi-locus strain-level bacterial typing from metagenomic samples. *Nucleic Acids Res.* 45:e7. doi: 10.1093/nar/gkw837

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Höper, Grützke, Brinkmann, Mossong, Matamoros, Ellis, Deneke, Tausch, Cuesta, Monzón, Juliá, Petersen, Hendriksen, Pamp, Leijon, Hakhverdyan, Walsh, Cotter, Chandrasekaran, Tay, Schlundt, Sala, De Cesare, Nitsche, Beer and Wylezich. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Microbiology

METHODS
published: 23 November 2020
doi: 10.3389/fmicb.2020.575615

Mammalian Cell-Based Immunoassay for Detection of Viable Bacterial Pathogens

Luping Xu[1†], Xingjian Bai[1†], Shivendra Tenguria[1], Yi Liu[1], Rishi Drolia[1,2] and Arun K. Bhunia[1,2,3]*

[1] Molecular Food Microbiology Laboratory, Department of Food Science, Purdue University, West Lafayette, IN, United States, [2] Purdue Institute of Inflammation, Immunology and Infectious Disease, Purdue University, West Lafayette, IN, United States, [3] Department of Comparative Pathobiology, Purdue University, West Lafayette, IN, United States

Rapid detection of live pathogens is of paramount importance to ensure food safety. At present, nucleic acid-based polymerase chain reaction and antibody-based lateral flow assays are the primary methods of choice for rapid detection, but these are prone to interference from inhibitors, and resident microbes. Moreover, the positive results may neither assure virulence potential nor viability of the analyte. In contrast, the mammalian cell-based assay detects pathogen interaction with the host cells and is responsive to only live pathogens, but the short shelf-life of the mammalian cells is the major impediment for its widespread application. An innovative approach to prolong the shelf-life of mammalian cells by using formalin was undertaken. Formalin (4% formaldehyde)-fixed human ileocecal adenocarcinoma cell line, HCT-8 on 24-well tissue culture plates was used for the capture of viable pathogens while an antibody was used for specific detection. The specificity of the Mammalian Cell-based ImmunoAssay (MaCIA) was validated with *Salmonella enterica* serovar Enteritidis and Typhimurium as model pathogens and further confirmed against a panel of 15 *S.* Enteritidis strains, 8 *S*. Typhimurium, 11 other *Salmonella* serovars, and 14 non-*Salmonella* spp. The total detection time (sample-to-result) of MaCIA with artificially inoculated ground chicken, eggs, milk, and cake mix at 1–10 CFU/25 g was 16–21 h using a traditional enrichment set up but the detection time was shortened to 10–12 h using direct on-cell (MaCIA) enrichment. Formalin-fixed stable cell monolayers in MaCIA provide longer shelf-life (at least 14 weeks) for possible point-of-need deployment and multi-sample testing on a single plate.

Keywords: immunoassay, poultry, mammalian cells, *Salmonella*, detection, MaCIA, cell-based sensor, stress

OPEN ACCESS

Edited by:
Nigel Cook,
Jorvik Food & Environmental Virology Ltd., United Kingdom

Reviewed by:
Sandeep Tamber,
Health Canada, Canada
Hengyi Xu,
Nanchang University, China

***Correspondence:**
Arun K. Bhunia
bhunia@purdue.edu

[†] These authors have contributed equally to this work

Specialty section:
This article was submitted to Food Microbiology, a section of the journal Frontiers in Microbiology

Received: *23 June 2020*
Accepted: *29 October 2020*
Published: *23 November 2020*

Citation:
Xu L, Bai X, Tenguria S, Liu Y, Drolia R and Bhunia AK (2020) Mammalian Cell-Based Immunoassay for Detection of Viable Bacterial Pathogens. Front. Microbiol. 11:575615. doi: 10.3389/fmicb.2020.575615

INTRODUCTION

Pathogen interaction with the host cells is the crucial first step for initiating infection (Finlay and Falkow, 1997; Kline et al., 2009), and harnessing such interaction may yield a robust detection platform not only to assess pathogenic potential but also its viability. Mammalian cell-based biosensors (CBBs) exploit host-pathogen interactions including pathogen adhesion, activation of host cell signaling events, cell-cycle arrest, apoptosis, and/or cytotoxicity

(Banerjee and Bhunia, 2009). The ability to detect host-pathogen interaction makes CBB a functionality test, thus sets it apart from other conventional methods (Banerjee and Bhunia, 2009). A common approach to monitoring such interaction is to measure the cytotoxic effects of the analytes on mammalian cells. Gray et al. (2005) used lymphocyte (Ped-2E9)-based cytotoxicity assay to detect toxin produced by *Bacillus cereus*. Later, this cell line was used in a collagen-encapsulated 3-D platform to detect *Listeria monocytogenes* cells and its toxins (Banerjee et al., 2008) and several other toxin-secreting foodborne pathogens (Banerjee and Bhunia, 2010). Most recently, a 3-D Vero cell-platform was made to screen Shiga-toxin producing *Escherichia coli* (STEC) by measuring lactate dehydrogenase (LDH) release (To and Bhunia, 2019). Although these studies demonstrate the versatility of CBBs in detecting foodborne pathogens and toxins, the specificity of CBBs cannot be guaranteed when the detection solely relies on cytotoxicity measurement because cytosolic proteins/enzymes could be released from cells in response to more than one type of triggers. Furthermore, researchers have pointed out the short-comings of the practical applicability of CBBs due to the short shelf-life and the requirement for stringent growth conditions of mammalian cells outside a controlled laboratory environment (Bhunia et al., 1995; Banerjee et al., 2007; Banerjee and Bhunia, 2009; Ye et al., 2019). Thus, novel approaches for developing CBBs with higher specificity and longer shelf-life are in continued demand.

Pathogen detection is categorized into three basic types: culture-based, immunological, and nucleic acid-based (Bhunia, 2014; Lee et al., 2015; Bell et al., 2016; Schlaberg et al., 2017; Ricke et al., 2018; Rajapaksha et al., 2019). The detection time for the culture-based method is usually 4–7 days. Immunological and nucleic acid-based PCR methods are faster, but the inherent inability to assess the viability or the pathogenic potential of the target microorganisms is of concern (Bhunia, 2014; Kasturi and Drgon, 2017; Ricke et al., 2018). Moreover, these methods are often prone to interferences from sample inhibitors and resident microflora. Alternative detection methods that are faster, user-friendly, and accurate are in high demand (Bhunia, 2014). Therefore, CBBs have been proposed to serve as a reliable tool for the rapid screening of viable pathogens or active toxins in foods (Ngamwongsatit et al., 2008; Banerjee and Bhunia, 2009; Bhunia, 2011; Ye et al., 2019; To et al., 2020). However, maintaining the viability of mammalian cells outside the laboratory environment is a major challenge thus limits CBB's utility in routine foodborne pathogen testing (Bhunia et al., 1995; Banerjee et al., 2007; Banerjee and Bhunia, 2009; Ye et al., 2019).

In this study, we took an innovative approach and developed a shelf-stable Mammalian Cell-based ImmunoAssay (MaCIA) platform for the detection of live pathogenic bacteria. Shelf-life of MaCIA was prolonged by fixing the mammalian cells in formalin (4% formaldehyde) which is a common practice in histology and tissue imaging to preserve the cells by preventing protein degradation (Eltoum et al., 2001). Furthermore, instead of measuring cytotoxicity, we took advantage of the adhesion ability of enteric pathogens to the intestinal cells followed by antibody-based assay for specific detection of the adhered target pathogens. Adhesion to the epithelial cells is the crucial first step for enteric pathogens (Kline et al., 2009; Dos Reis and Horn, 2010). For example, *L. monocytogenes* binds to Hsp60 and E-cadherin on the epithelial cell surface through *Listeria* adhesion protein (LAP) and Internalin A (InlA), respectively to initiate adhesion, invasion, translocation, and systemic spread during the intestinal phase of infection (Drolia et al., 2018; Drolia and Bhunia, 2019). Enterohaemorrhagic *E. coli* employs intimin, fimbrial proteins, flagella, and autotransporter proteins to attach to the host cells at different stages of its life cycle during infection (McWilliams and Torres, 2014). Likewise, *Salmonella enterica* utilizes multiple fimbrial adhesins, such as type 1 fimbriae (T1F) and long polar fimbriae (Lpf), and several autotransporter adhesins, such as ShdA and MisL, to promote adhesion to D-mannose receptors on M cells in Peyer's Patches and assist colonization in the intestine (Bäumler et al., 1996; Wagner and Hensel, 2011; Bhunia, 2018; Kolenda et al., 2019). Therefore, detecting only adhered pathogens using antibodies is a rational approach. We chose human ileocecal adenocarcinoma cell line, HCT-8, as the target cells for building MaCIA platform on 24-well tissue culture plates. HCT-8 is one of the commonly used model cell lines to study the adhesion of enteric pathogens (McKee and O'Brien, 1995; Dibao-Dina et al., 2015; Hu and Wai, 2017). Unlike other cell lines used, HCT-8 cells can form a fully confluent monolayer in only 5 days.

The objective of this study was to develop a shelf-stable MaCIA platform for the rapid detection of viable bacterial pathogens and to validate its performance using *Salmonella enterica* serovar Enteritidis as a model foodborne pathogen.

Salmonella enterica is a major foodborne pathogen of global public health concern. Meat, poultry, eggs, nuts, fruits, and vegetables are common vehicles for *Salmonella* transmission. Each year, *Salmonella* infections contribute to 1.3 billion cases of gastroenteritis and 3 million deaths worldwide (Kirk et al., 2015) and 1.35 million cases, 26,500 hospitalizations, and 420 deaths in the United States (CDC, 2020). Among *Salmonella* serovars, *Salmonella enterica* serovar Enteritidis is one of the most prevalent serovars in the United States. The Centers for Disease Control and Prevention (CDC) has reported eight major outbreaks between 2006 and 2018 resulting in about 4,000 cases (CDC, 2018). In a survey of salmonellosis outbreaks (total 2,447) in the United States between 1998 and 2015, *S.* Enteritidis (29.1%) was reported to be the most common serovar followed by *S.* Typhimurium (12.6%), *S.* Newport (7.6%), and others (Snyder et al., 2019). The frequent occurrence of food-associated *S.* Enteritidis outbreaks with the high number of infections was the motivation for developing a mammalian cell-based functional bioassay for the detection of *S.* Enteritidis.

The initial study involved screening of MaCIA with a panel of food-associated bacterial cultures (**Table 1**) in confirming the specificity and the limit of detection (LOD) from artificially inoculated food samples. Next, the performance of MaCIA was validated using the US Department of Agriculture (USDA-FSIS, 2013) and the Food and Drug Administration (FDA, 2001) reference methods. "On-cell enrichment" and "one-step antibody probing methods" of MaCIA were also explored to reduce the

TABLE 1 | Specificity of mammalian cell-based immunoassay (MaCIA) platform tested against *Salmonella* and non-*Salmonella* spp.

Bacteria	CFU/Well	MaCIA Result*			
		mAb-2F11		mAb-F68C	
		Abs_{450nm} ± SD	Result	Abs_{450nm} ± SD	Result
***Salmonella enterica* serovars**					
Enteritidis PT21	2.0–13 × 10^7	0.95 ± 0.08	+	0.10 ± 0.01	-
Enteritidis 13ENT1344	2.9 × 10^7	1.13 ± 0.16	+	NT	NT
Enteritidis 13ENT1374	2.8–3.3 × 10^7	0.91 ± 0.15	+	0.09 ± 0.00	-
Enteritidis 13ENT1376	2.0 × 10^7	1.06 ± 0.03	+	NT	NT
Enteritidis 13ENT1375	3.1 × 10^7	1.07 ± 0.15	+	NT	NT
Enteritidis 13ENT1032	2.2 × 10^7	1.08 ± 0.25	+	NT	NT
Enteritidis PT1	2.8 × 10^7	1.19 ± 0.04	+	NT	NT
Enteritidis PT4	2.0 × 10^7	1.17 ± 0.06	+	NT	NT
Enteritidis PT6	1.8 × 10^7	1.41 ± 0.04	+	NT	NT
Enteritidis PT7	7.5 × 10^6	0.70 ± 0.06	+	NT	NT
Enteritidis PT8	1.4 × 10^7	1.42 ± 0.06	+	NT	NT
Enteritidis PT13a	1.5 × 10^7	0.74 ± 0.02	+	NT	NT
Enteritidis PT13	1.1 × 10^7	0.90 ± 0.06	+	NT	NT
Enteritidis PT14b	1.3 × 10^7	1.05 ± 0.04	+	NT	NT
Enteritidis PT28	1.1 × 10^7	0.53 ± 0.05	+	NT	NT
Typhimurium 13ENT906	6.7–8.8 × 10^7	0.33 ± 0.03	-	1.14 ± 0.06	+
Typhimurium NOS12	4.0 × 10^7	0.33 ± 0.03	-	0.98 ± 0.04	+
Typhimurium NOS3	3.3 × 10^8	NT	NT	0.80 ± 0.04	+
Typhimurium NOS10	1.3 × 10^8	NT	NT	0.90 ± 0.12	+
Typhimurium NOS2	6.7 × 10^8	NT	NT	0.73 ± 0.08	+
Typhimurium NOS4	6.7 × 10^8	NT	NT	0.89 ± 0.04	+
Typhimurium NOS1	3.3 × 10^7	NT	NT	0.84 ± 0.06	+
Typhimurium ST1	3.8 × 10^6	0.13 ± 0.02	-	NT	NT
Newport 13ENT1060	2.3–23 × 10^7	0.32 ± 0.04	-	0.13 ± 0.01	-
Braenderup 12ENT1138	6.3 × 10^7	0.33 ± 0.03	-	NT	NT
Agona 12ENT1356	2.7–13 × 10^8	0.32 ± 0.02	-	0.09 ± 0.01	-
Hadar 13ENT979	4.3 × 10^7	0.27 ± 0.02	-	NT	NT
Paratyphi 11J85	2.4 × 10^7	0.27 ± 0.05	-	NT	NT
Heidelberg 18ENT1418	4.0 × 10^7	0.29 ± 0.04	-	NT	NT
Saintpaul 13ENT1045	5.0 × 10^7	0.30 ± 0.04	-	NT	NT
Javiana 13ENT86F	0.4–2.7 × 10^8	0.38 ± 0.10	-	0.14 ± 0.07	-
Infantis 13ENT866	2.0 × 10^7	0.32 ± 0.02	-	NT	NT
Bareilly 12ENT1164	0.1–4.0 × 10^8	0.32 ± 0.09	-	NT	NT
Pullorum DUP-PVUII 1006	1.9 × 10^7	0.34 ± 0.04	-	NT	NT
Miscellaneous					
Listeria monocytogenes F4244	0.5–1.6 × 10^8	0.27 ± 0.03	-	0.13 ± 0.01	-
L. innocua F4248	5.0 × 10^7	0.27 ± 0.03	-	NT	NT
Escherichia coli O157:H7 EDL933	0.4–1.3 × 10^8	0.26 ± 0.03	-	0.08 ± 0.04	-
Hafnia alvei 18066	3.3–6.3 × 10^7	0.28 ± 0.03	-	0.15 ± 0.02	-
Citrobacter freundii ATCC8090	0.3–1.0 × 10^8	0.29 ± 0.02	-	0.13 ± 0.01	-
Citrobacter freundii ATCC43864	0.4–3.3 × 10^7	0.11 ± 0.00	-	0.11 ± 0.01	-
Citrobacter freundii ATCC3624	0.3–1.3 × 10^8	0.13 ± 0.01	-	0.12 ± 0.02	-
Serratia marcescens ATCC8100	0.6–5.3 × 10^7	0.33 ± 0.02	-	0.12 ± 0.01	-
Serratia marcescens ATCC43862	0.1–1.0 × 10^8	0.11 ± 0.01	-	0.13 ± 0.01	-
Serratia marcescens B-2544	0.6–3.3 × 10^7	0.13 ± 0.01	-	0.11 ± 0.02	-
Pseudomonas aeruginosa PRI99	2.25 × 10^7	0.24 ± 0.02	-	NT	NT
Proteus mirabilis B-3402	0.7–6.7 × 10^8	0.11 ± 0.01	-	0.11+0.01	-
Proteus vulgaris DUP-10086	0.4–4.0 × 10^8	0.11 ± 0.01	-	0.10 ± 0.02	-
Klebsiella pneumoniae B-41958	6.7 × 10^6	0.10 ± 0.01	-	0.13 ± 0.01	-

**Values are from four independent replicates; Results (±) are decided by comparing to the negative control in each experiment. Values that are significantly different ($P < 0.001$) from the negative control in each experiment are regarded as +; NT, not tested.*

assay steps and total detection time. Overall, the data showed that MaCIA could detect viable *S.* Enteritidis (1–10 CFU/25 g) in ground chicken, shelled eggs, whole milk, and cake mix using a traditional enrichment set up, but the detection time was shortened to 10–12 h using direct on-cell (MaCIA) enrichment. We also demonstrated the versatility of MaCIA by using a commercial anti-*Salmonella* reporter antibody for the detection of *S.* Typhimurium. Formalin-fixed cells in the MaCIA platform permits a longer shelf life (at least 14-week at 4°C), minimum on-site maintenance care, and a stable cell monolayer for point-of-need deployment.

RESULTS

Development of MaCIA (Mammalian Cell-Based ImmunoAssay) Platform

The MaCIA platform was built on a 24-well tissue culture plate, and it consisted of two steps: fixation of mammalian cells and immunoassay for specific detection of adherent target pathogens. We used the formalin-fixed HCT-8 cell line for *Salmonella* adhesion/capture (30 min) and anti-*S.* Enteritidis monoclonal antibody, mAb-2F11 (Masi and Zawistowski, 1995), or anti-*Salmonella* mAb-F68C (Thermo Fisher Scientific; 1.5 h), horseradish peroxidase (HRP)-conjugated second antibody and a substrate for color development (1.5 h). The mAb-2F11 is highly specific for *S.* Enteritidis (Masi and Zawistowski, 1995; Jaradat et al., 2004), and the Western blot analysis confirmed its specificity without showing any reaction with bands from whole-cell preparations of *L. monocytogenes*, *E. coli* O157:H7 or *Pseudomonas aeruginosa* (**Figure 1A**).

To fix mammalian cells on the MaCIA platform, HCT-8 cell monolayers in 24 well-plates were treated with a 4% formaldehyde solution for 10 min, followed by three sequential wash using phosphate-buffered saline (PBS, 0.1 M, pH 7). Initially, the performance of formalin-fixed MaCIA was compared with a live cell-based MaCIA platform to detect *S.* Enteritidis PT21 that was incubated for 30 min at 37°C. Remarkably, both assay configurations showed strong positive signals toward viable *S.* Enteritidis, which was significantly ($P < 0.0001$) higher than the equivalent amounts of dead *S.* Enteritidis cells (verified by plating) and the negative control (PBS) (**Figure 1B**).

The performance of MaCIA was also compared with traditional sandwich ELISA where mAb-2F11 was used as capture and anti-*Salmonella* pAb-3238 (Abdelhaseib et al., 2016) was used as the reporter. MaCIA gave positive results when tested with viable *S.* Enteritidis cells (1×10^8 CFU/mL), which is significantly higher ($P < 0.0001$) than that of the equivalent numbers of dead cells or the PBS control. On the other hand, both viable and dead *S.* Enteritidis cells showed positive signals with sandwich ELISA, though the signals for viable cells were slightly higher than those of the dead cells (**Figure 1C**). Furthermore, the total detection time (after addition of bacteria to the wells of assay plates) required for sandwich ELISA was 5.5 h, while 4 h for MaCIA (**Figure 1C**).

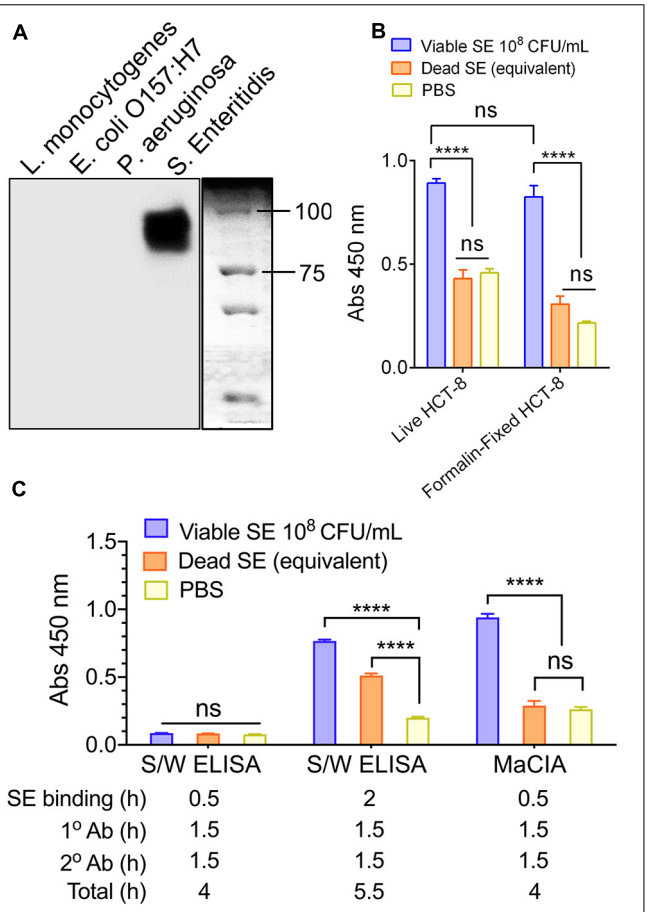


FIGURE 1 | Mammalian cell-based immunoassay (MaCIA) development. **(A)** Western blot showing the reaction of mAb-2F11 to *Salmonella enterica* serovar Enteritidis PT21 but not to *L. monocytogenes* F4244, *E. coli* O157:H7 EDL933 and *Pseudomonas aeruginosa* PRI99. **(B)** MaCIA analysis with live (Live HCT-8) and formalin-fixed (Formalin-fixed HCT-8) HCT-8 cell. **(C)** Comparison of MaCIA with sandwich ELISA (S/W ELISA). Error bars represent the standard error of the mean (SEM). ****$P < 0.0001$; ns, no significance. Cut-off for positive: $P < 0.001$. SE, *S.* enteritidis; Ab, antibody.

Specificity of the MaCIA Platform

Next, the specificity of the MaCIA was determined by testing a panel of 15 *S.* Enteritidis strains, eight *S.* Typhimurium strains, 11 other *Salmonella* serovars, and 14 non-*Salmonella* spp. at ~ 1×10^6 to 1×10^7 CFU/mL each. The data showed that MaCIA was highly specific toward all tested viable strains of *S.* Enteritidis or *S.* Typhimurium serovars depending on the reporter antibody used and the signals were significantly ($P < 0.001$) higher than the signals obtained for other *Salmonella* serovars or non-*Salmonella* species (**Figures 2A,B** and **Table 1**). Thus, any sample showing a significantly higher signal ($P < 0.001$) than the negative control was considered positive. Furthermore, samples containing live *S.* Enteritidis cells gave significantly ($P < 0.0001$) higher absorbance values (signals) than that of the values obtained for dead cells or the PBS control (**Figure 2B**). The specificity of MaCIA toward viable cells was not affected when tested against a mixture containing equal amounts of viable and dead

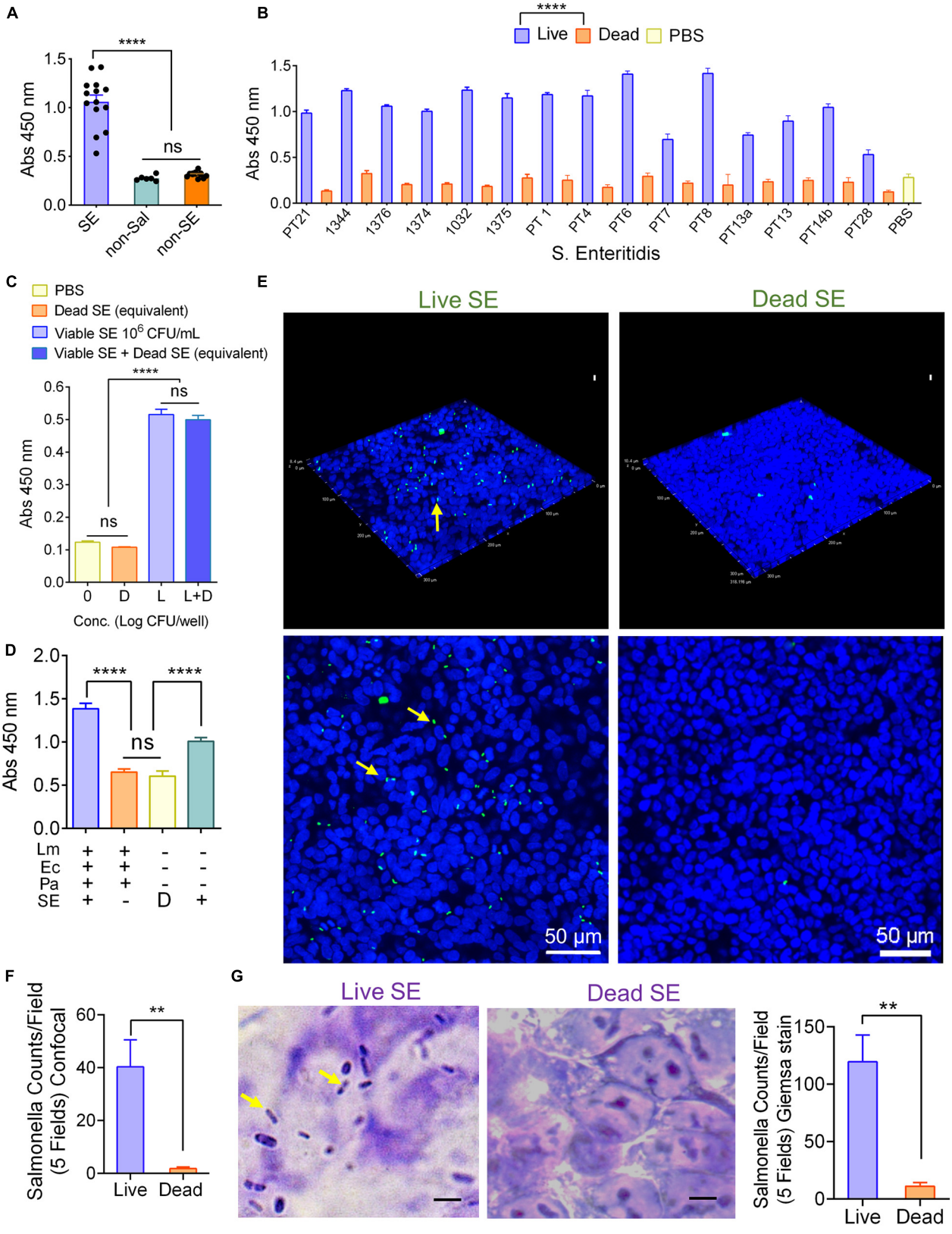


FIGURE 2 | Continued

FIGURE 2 | Mammalian cell-based immunoassay specificity. MaCIA reaction with 15 *Salmonella* Enteritidis strains (SE), 12 non-SE and 7 non-*Salmonella* bacteria **(A)**, with viable and dead *S.* Enteritidis serovars **(B)**, to viable *S.* Enteritidis in the presence of the equivalent amount of dead *S.* Enteritidis **(C)**, and *S.* Enteritidis PT21 in the presence of other bacteria (Lm, *L. monocytogenes* F4244; Ec, *E. coli* EDL933 and Pa, *Pseudomonas aeruginosa* PRI99). L: live SE; D: Dead SE **(D)**. Confocal image and Giemsa staining analyses of adhesion of live (Live SE) and dead (Dead SE) *S.* Enteritidis PT21 to formalin-fixed HCT-8 cells; **(E)** Z-stack of the scanned images, **(F)** total bacterial counts per five fields for confocal images. Blue: nucleus, green: *S.* Enteritidis, **(G)** Giemsa stained images showing adhesion of live (Live SE) but not dead (Dead SE) *S.* Enteritidis PT21 to formalin-fixed HCT-8 cells. Rod-shaped dark blue, *S.* Enteritidis (arrows); purple, nucleus; Bar graph showing bacterial counts per field from five fields. Error bars represent SEM. ****$P < 0.0001$; **$P < 0.01$; ns, no significance. Cut-off for positive: $P < 0.001$. Scale bar = 5 μm.

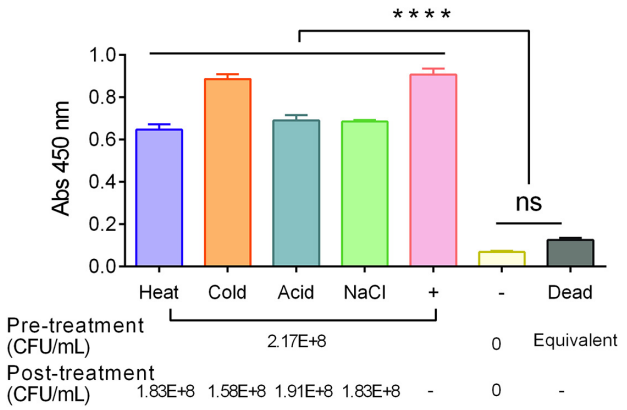


FIGURE 3 | Detection of stress-exposed *S.* Enteritidis PT21 using MaCIA. Bacteria were exposed to heat (45°C), cold (4°C), acidic pH (4.5) and NaCl (5.5%) for 3 h before analysis. +, Positive control (bacteria without any stress exposure); –, No bacteria; dead, heat-killed *S.* Enteritidis cells. Error bars represent SEM. ****$P < 0.0001$; ns, no significance.

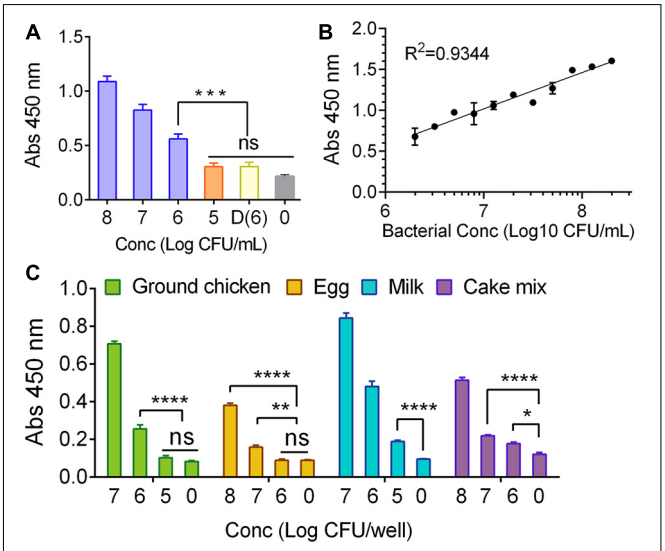


FIGURE 4 | Assay sensitivity for MaCIA. **(A)** Analysis of limit of detection (LOD) of MaCIA against *S.* Enteritidis PT21 at 1×10^5 CFU/mL to 1×10^8 CFU/ml suspended in PBS; and **(B)** corresponding correlation coefficient of absorbance and bacterial concentration. **(C)** Analyses of LOD of MaCIA when *S.* Enteritidis PT21 was suspended in different food matrices. 0, no bacteria; D(6), dead *S.* Enteritidis PT21 at 1×10^6 cells/ml. In all figures, samples with higher concentrations were also significantly ($P < 0.001$) different than the dead samples and negative control. Error bars represent SEM. ****$P < 0.0001$; ***$P < 0.001$; **$P < 0.01$; *$P < 0.05$; ns, no significance. Cut-off for positive: $P < 0.001$.

S. Enteritidis cells (**Figure 2C**), and non-*S.* Enteritidis bacteria (**Figure 2D**). Immuno-stained confocal images, the Z-stacking (three-dimensional image), and Giemsa stain images confirmed increased adhesion of viable *S.* Enteritidis cells to HCT-8 cells than that of the dead *S.* Enteritidis cells (**Figures 2E–G**). Confocal imaging further revealed the absence of non-specific binding of mAb-2F11 to the HCT-8 cell monolayer (**Figures 2E,F**). Furthermore, MaCIA successfully detected *S.* Enteritidis cells when exposed to various stressors for 3 h (Hahm and Bhunia, 2006) including cold (4°C), heat (45°C), acidic pH (4.5), and 5.5% NaCl (**Figure 3**).

Detection Sensitivity of MaCIA

To determine assay sensitivity, *S.* Enteritidis cells were serially diluted using either PBS or ground chicken suspended in buffered peptone water (BPW) and added to the wells containing formalin-fixed HCT-8 cell monolayers (30-min post-fixation). After a 30-min incubation at 37°C with test samples, the monolayers were washed, probed with mAb-2F11, and the color was developed. An initial bacterial concentration of 1×10^6 to 1×10^8 CFU/mL showed significantly ($P < 0.001$) higher signal than the wells containing 1×10^5 CFU/mL or dead cells (1×10^6 cells) suspended in PBS (**Figure 4A**) or ground chicken slurry (**Supplementary Figure 1**) and the absorbance values showed strong correlation ($R^2 = 0.9344$) with *S.* Enteritidis cell numbers (1×10^6 CFU/mL to 1×10^8 CFU/mL) (**Figure 4B**). Furthermore, MaCIA also showed a similar concentration-dependent rise in signals when bacteria were suspended in ground chicken, liquid egg, milk, and cake mix slurry (**Figure 4C**). However, the detection sensitivity varied depending on the food matrix tested. In milk, the detection limit was determined to be 1×10^5 CFU/mL while in ground chicken, 1×10^6 CFU/mL, in cake mix, 1×10^7 CFU/mL, and in egg, 1×10^8 CFU/mL (**Figure 4C**). These results indicate that assay sensitivity for MaCIA for the detection of *S.* Enteritidis varies from 1×10^5 CFU/mL to 1×10^8 CFU/mL depending on the food matrix tested.

Further Optimization of MaCIA

One-Step Antibody Probing Method

To shorten the detection time, we explored if a one-step antibody probing approach is feasible. Ground chicken was inoculated with *S.* Enteritidis at 6×10^2 CFU/25 g in a stomacher bag. After 10-h enrichment at 37°C, the

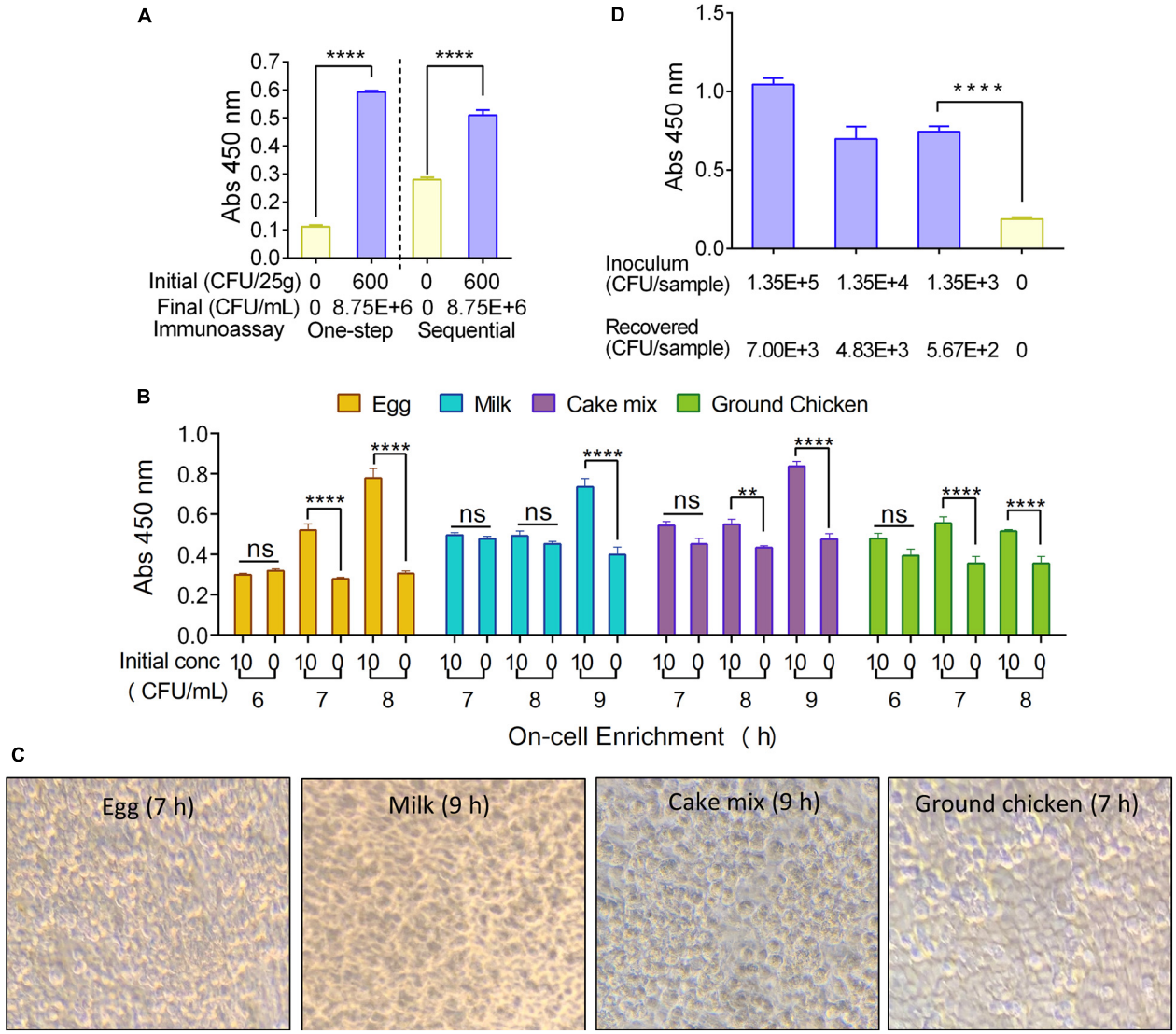


FIGURE 5 | Mammalian cell-based immunoassay assay optimization. **(A)** One-step antibody probing vs sequential antibody probing against a bacterial cell concentration of 8.75×10^6 CFU/ml of *S.* Enteritidis. **(B)** Analysis of time (h) required for positive MaCIA result during on-cell enrichment of *S.* Enteritidis PT21 (~10 CFU/mL) inoculated into different food products. **(C)** Light microscopic images of formalin-fixed HCT-8 cell monolayers after on-cell enrichment for 7–9 h. Magnification (400×). **(D)** MaCIA analysis of skin swab samples after on-cell enrichment (7 h). Samples with higher concentrations were also significantly ($P < 0.001$) different than the negative control. Error bars represent SEM. ****$P < 0.0001$; **$P < 0.01$; ns, no significance. Cut-off for positive: $P < 0.001$.

enriched chicken samples (1 mL) were added to the fixed HCT-8 cell monolayer for 30 min, followed by PBS wash (3 times). The cell monolayers were probed with an antibody cocktail that contained both primary (mAb-2F11) and secondary (anti-mouse HRP-conjugated IgG) antibodies, followed by the colorimetric substrate. Data showed that the signal obtained from the one-step antibody probing was comparable to the results when the sequential antibody probing method was used (**Figure 5A**). This experiment indicates that one-step antibody probing is equally effective as the sequential antibody probing method, thus shortening the assay time by 2.5 h.

On-Cell Food Sample Enrichment

Direct on-cell (MaCIA platform) enrichment of test samples was pursued to simplify the assay procedure and to reduce the sample handling steps. *S.* Enteritidis inoculated food suspensions (with an initial inoculation of 10 CFU/mL) were directly added to the wells (1 mL/well) containing formalin-fixed HCT-8 cell monolayers and incubated at 37°C. The assay was performed after 6, 7, 8, and 9-h on-cell sample enrichment followed by sequential antibody probing (3 h). After 7-h on-cell enrichment, both ground chicken and egg samples gave positive results while the whole milk and cake mix needed 9-h enrichment to give positive results when compared with

uninoculated food samples (**Figure 5B**). A similar result was obtained when the food samples were tested in a blinded fashion (**Supplementary Figure 2**). Total assay time (sample-to-result) for on-cell enrichment was estimated to be 10–12 h. Remarkably, the HCT-8 cell monolayers remained intact without any visible damage during on-cell enrichment (**Figure 5C**). Due to the limitation in the amount of sample volume (1 mL/well), that can be tested, the "on-cell enrichment" option is suitable only when the starting *S.* Enteritidis concentration is above 10 CFU/mL (2.5×10^3 CFU/25 g); hence it may not be suitable for routine testing of bulk-food samples that may contain < 100 CFU/g.

We then examined if the on-cell enrichment set up is suitable for testing surface swab samples. Skin swabs from inoculated chicken thigh parts (1.35×10^3 to 1.35×10^5 CFU/50 cm^2 at 4°C for 24 h) were resuspended in 1.1 mL of BPW, and 1 ml of each suspension was added to the wells of MaCIA. After 7-h on-cell enrichment followed by sequential immunoprobing (3 h), MaCIA generated significantly ($P < 0.0001$) higher signals than that of the values obtained from the negative control (swabbed suspension of the uninoculated sample) (**Figure 5D**). These data indicate that MaCIA is suitable for testing surface swab samples, and results can be obtained in less than 12 h.

Comparison of MaCIA With the USDA/FDA Detection Methods

To compare the performance of MaCIA with USDA/FDA detection methods, *S.* Enteritidis inoculated food samples (ground chicken, egg, milk and cake mix held at 4°C for 24 h) were also tested in parallel using the US Department of Agriculture (USDA-FSIS, 2013) or Food and Drug Administration (FDA, 2001) reference method.

Growth Kinetics of *S.* Enteritidis in Different Foods

Freshly grown (37°C, 18 h) *S.* Enteritidis PT21 culture was inoculated (<10 CFU/ml) into 25 g of each ground chicken, egg, whole milk, or cake mix in 225 mL BPW in a stomacher bag (Seward Inc., Bohemia, NY, United States) and held at 4°C for 24 h. Inoculated food samples were then incubated at 37°C and bacterial counts were determined every 2-h intervals until 18 h. The growth data of *S.* Enteritidis in all tested food samples were fitted with the Gompertz model to generate a growth curve (**Figure 6A**). The R^2 values of Gompertz fitted growth curves of *S.* Enteritidis PT21 in ground chicken, egg, whole milk, and cake mix were 0.99, 0.99, 0.96, and 0.99, respectively. Based on the Gompertz modeled growth curve equations, the lag phase duration (LPD) and exponential growth rate (EGR) were estimated to be 2.204–2.427 h and 0.767–0.934 log (CFU/mL)/h, respectively (**Supplementary Table 1**). Utilizing LPD, EGR, and the MaCIA detection limit data, we were able to estimate the required enrichment time for each food product, assuming the starting *S.* Enteritidis concentration is 1 CFU/25 g (**Supplementary Table 1**). The required enrichment time for ground chicken, egg, milk, and cake mix was estimated to be 14, 19, 16, and 16 h, respectively (**Supplementary Table 1**).

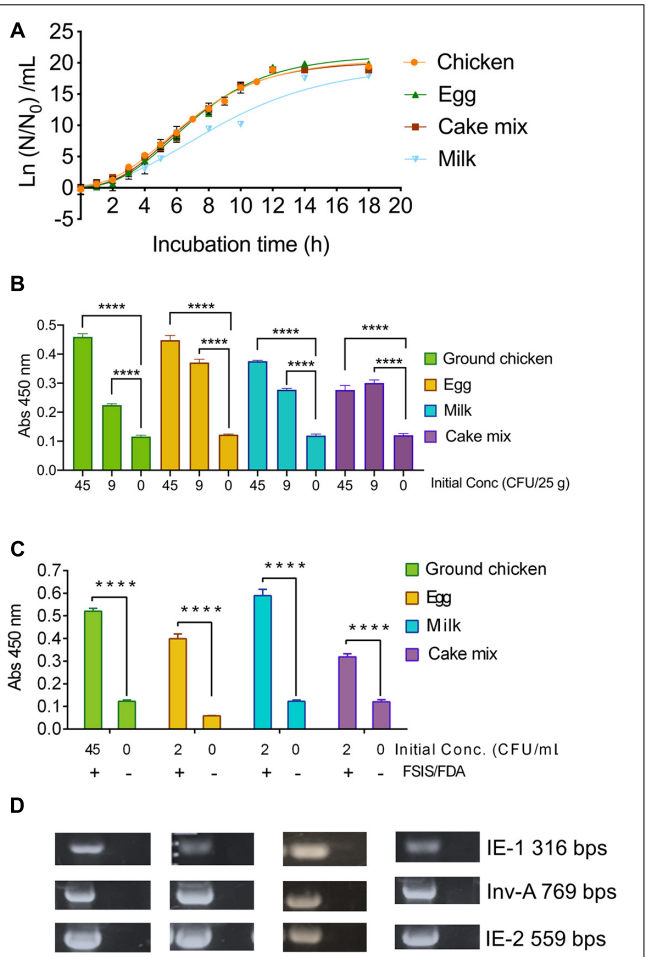


FIGURE 6 | Mammalian cell-based immunoassay validation with inoculated food samples. **(A)** Growth curve of *S.* Enteritidis PT21 in various food products suspended in buffered-peptone water (BPW) at 37°C. Before growth analysis, inoculated food samples were held at 4°C for 24 h. The best-fit curves for *Salmonella* growth in different foods were generated by using the Gompertz model. R^2 values of each fitted Gompertz curve are 0.99 (Chicken), 0.99 (Egg), 0.99 (cake mix) and 0.96 (Milk). N_0, initial *S.* Enteritidis concentration; N, *S.* Enteritidis concentration at the corresponding time point. MaCIA results of *S.* Enteritidis inoculated (at 0, 9, 45 CFU/25 g) **(B)** and at 0, 2, 45 CFU/mL **(C)** food samples after 14–19 h enrichment. **(D)** PCR confirmation of *S.* Enteritidis targeting *Salmonella* specific genes. Error bars represent SEM. ****$P < 0.0001$.

Sample-to-Result Time

To confirm the sample-to-result time, we inoculated the selected food samples with *S.* Enteritidis at 0, 9, or 45 CFU/25 g (**Figure 6B**) and 0, 2 or 45 CFU/mL (**Figure 6C**). After a specified enrichment period, we analyzed the samples using MaCIA. All *S.* Enteritidis-inoculated samples produced significantly higher signals ($P < 0.001$) than the uninoculated food samples (**Figures 6B,C**) even in the presence of background microflora (**Supplementary Figure 3**). The sample-to-result time was estimated to be 16–21 h. Analysis of food samples by the USDA-FSIS or FDA-BAM method followed by polymerase chain reaction (PCR) assay using three sets of primers targeting *invA*,

IE-1, and IE-2 genes (**Figure 6D**) confirmed the presence of *S.* Enteritidis in these food samples. Note, the USDA method needed 72 h, while the FDA method needed 72–168 h to confirm the presence of *Salmonella* in the inoculated food samples.

Formalin-Fixation Prolongs the Shelf-Life of MaCIA

The bottleneck for widespread use of cell-based sensors is its limited shelf-life. As we have demonstrated earlier (**Figure 1B**), the performance of MaCIA prepared with live HCT-8 cells is equally sensitive to the formalin-fixed HCT-8 cells (30 min after fixation). In this experiment, we investigated if the prolonged storage (4, 8, and 14 weeks at 4°C or 4 weeks at room temperature) of formalin-fixed HCT-8 cell would uphold MaCIA's performance. Data showed that formalin-fixed HCT-8 cells stored for 4–12 weeks at 4°C generated comparable results to that of live HCT-8 cells when tested with viable *S.* Enteritidis PT21 at a concentration of 1×10^7 CFU/ml and signals were significantly higher ($P < 0.0001$) than the signals generated by an equivalent amount of dead *S.* Enteritidis cells or the PBS control (no bacteria) (**Figure 7A**). The light microscopic photomicrographs further confirmed that the cell monolayer and the cellular morphology in formalin-fixed HCT-8 cells were unaffected after 14 weeks of storage at 4°C or even after bacterial exposure and the subsequent three PBS wash (**Figure 7B**). These results indicate that formalin fixation was able to prolong the shelf-life of HCT-8 cells up to 14 weeks without affecting their performance, thus showing a promising application of the MaCIA for point-of-need deployment.

DISCUSSION

The conventional culture-based detection methods (sample-to-result) take 4–7 days to obtain the results (FDA, 2001; USDA-FSIS, 2013; Bell et al., 2016), and the so-called rapid methods take at least 24–48 h (Bhunia, 2014; Lee et al., 2015; Ricke et al., 2018; Rajapaksha et al., 2019). This is a major inconvenience for the food industries since some foods have a limited shelf-life. Furthermore, holding of products until the microbiological safety assessment can also increase the cost of storage. Therefore, products are released into the supply chain even before obtaining test results. Such practice is very costly, resulting in more than hundreds of recalls each year and the loss of millions of pounds of food (Buzby et al., 2014; Elkhishin et al., 2017). Therefore, rapid, accurate, and user-friendly viable pathogen detection tools are in high demand to lower recalls, reduce food waste and financial loss, and prevent foodborne outbreaks.

Mammalian cell-based assays are highly attractive functional screening tools to assess the presence of viable pathogens or active toxins in near-real-time (Bhunia, 2011, 2014; To et al., 2020). CBB monitors host-hazard interaction (Banerjee and Bhunia, 2009); therefore, non-pathogenic, non-hazardous, dead, or non-toxic agents do not yield false results. However, the major drawback is its short self-life, i.e., the mammalian cells may not survive on the sensor platform for a prolonged period without the proper growth conditions. Mammalian cells have stringent requirements for specialized growth media and growth conditions for survival, such as temperature and CO_2-controlled humidified environment. Limited self-life of cells is a monumental deterrent for CBB's widespread application affecting its deployment for point-of-need use. To overcome the limitation, we employed formalin (4% formaldehyde) to preserve the functionality of the mammalian cells. We used the human ileocecal cell line, HCT-8, as our model cell line, which maintained its functionality after formalin-fixation, at least for 14 weeks at 4°C. The fixed HCT-8 cells showed selective interaction with viable or even stress-exposed *Salmonella*, while dead cells had negligible or no interaction at all (**Figures 1–3**). Further specificity of the assay was accomplished by immunoprobing the adhered bacterial cells using a specific antibody. The MaCIA was found to be highly specific for the detection of *S.* Enteritidis or *S.* Typhimurium without showing any cross-reaction with other *Salmonella* serovars or non-*Salmonella* species tested. The assay was further validated for its ability to detect *S.* Enteritidis in inoculated ground chicken, egg, whole milk, and cake mix in the presence of background natural microflora. A brief sample enrichment step also allows the resuscitation of stressed or injured cells before detection (Wu, 2008).

In the MaCIA platform, HCT-8 cells were used as a capture element instead of an antibody, which is traditionally used in a sandwich ELISA. In this study, HCT-8 cells out-performed the antibody (**Figure 1C**), and 30 min incubation was sufficient for optimal capture of viable bacteria by HCT-8 cells (Jaradat and Bhunia, 2003; Barrila et al., 2017) while 2 h was needed for sandwich ELISA. Improved bacterial capture by HCT-8 is attributed to the formation of a three-dimensional structure by mammalian cell monolayer (**Figure 2E**), creating a larger surface area for bacteria to bind. Furthermore, HCT-8 cell possesses surface receptor molecules for specific interaction with *Salmonella* adhesion factors. *S.* Enteritidis utilize type 1 fimbria to recognize and bind to high-mannose oligosaccharides, which are carried by various glycoproteins on the host cell surface (Kolenda et al., 2019). Long polar fimbriae also mediate adhesion of *Salmonella* to Peyer's patches on the host cell (Bäumler et al., 1996). Besides, MaCIA was able to differentiate viable cells from dead *Salmonella* cells while sandwich immunoassay was unable. Lack of adhesion of dead *Salmonella* to HCT-8 may be due to the loss or denaturation of bacterial adhesins (**Figure 2G**). While in sandwich ELISA, bacterial surface antigens from dead cells were still able to bind the capture-antibody. MaCIA also showed strong signals when tested with stress-exposed *S.* Enteritidis cells suggesting a brief stress exposure (3 h) does not affect bacterial ability to interact with the HCT-8 cells (**Figure 3**) while such exposure caused a 20–48% reduction in ELISA signal for *Salmonella* in a previous study (Hahm and Bhunia, 2006).

The sensitivity of MaCIA was found to be about 1×10^6 CFU/mL to 1×10^7 CFU/mL, which is in agreement with a typical ELISA where antibodies serve as the capture molecule (Mansfield and Forsythe, 2000; Eriksson and Aspan, 2007) or ELISA with bacteriophage as a recognition molecule (Galikowska et al., 2011). However, MaCIA has the potential to outperform ELISA in some aspects, due to its ability to differentiate viable from dead bacteria. Viable pathogens that can adhere and invade

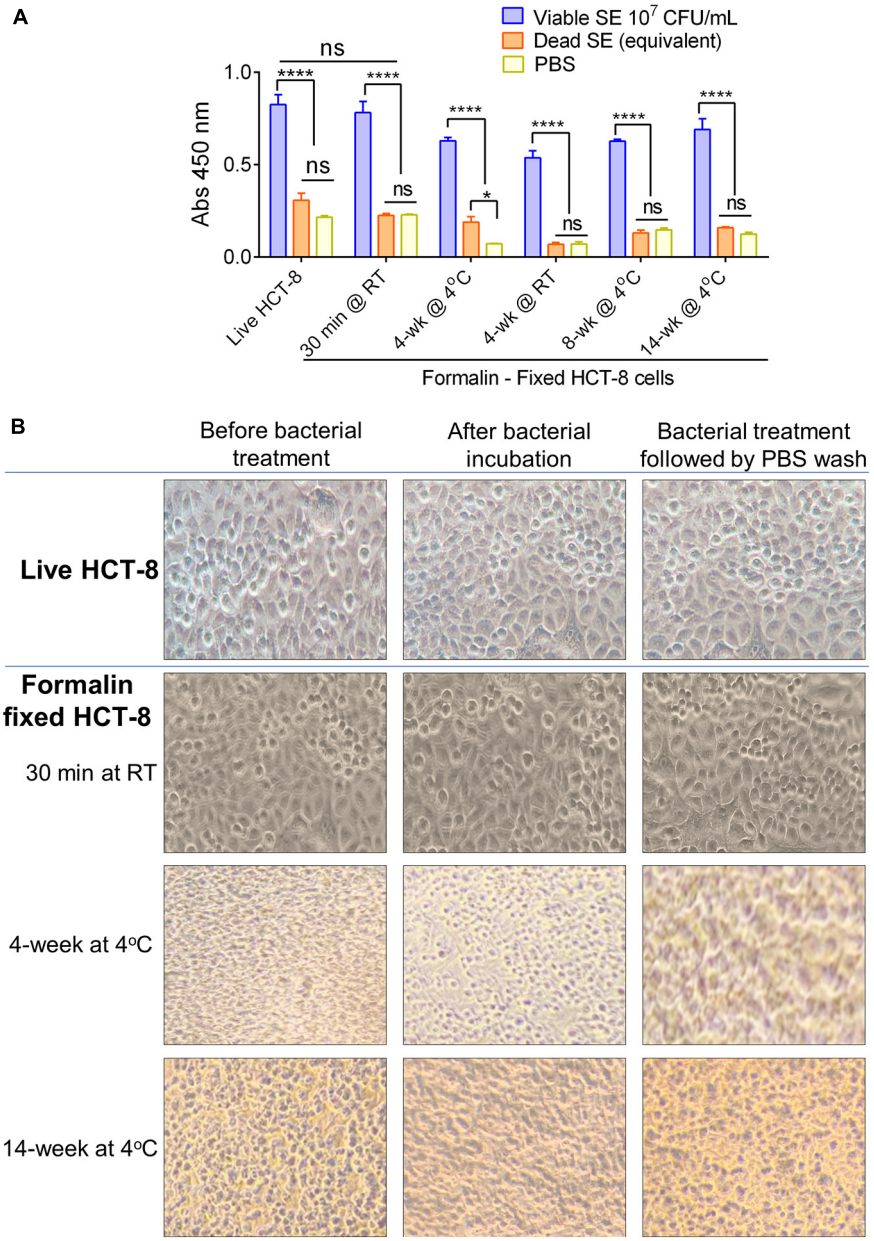


FIGURE 7 | Performance of MaCIA after prolonged storage. **(A)** Comparison of MaCIA signals (absorbance reading) of *S*. Enteritidis cells (1×10^7 cells/ml) originating from live HCT-8 and formalin-fixed HCT-8 cells stored at 4°C for 30 min to 14 weeks. **(B)** Light microscopic analysis of cell morphology of formalin-fixed HCT-8 cells stored up to 14 weeks. Panels showing intact cell morphology before bacterial treatment, after treatment, and after PBS wash. Magnification, 400×. Error bars represent SEM. ****$P < 0.0001$; ns, no significance.

into intestinal cells are of food safety concerns. MaCIA is a better choice over ELISA for the food industry when viable pathogens in food are the target. False-positive results generated by either ELISA or PCR due to the presence of non-viable pathogens could lead to unnecessary recalls, food waste, and economic loss. On the other hand, assays with higher sensitivity may be useful for detecting samples with low bacterial concentration, but enrichment is considered a necessary step to ensure accuracy (Bhunia, 2014). Assuming a 25 g sample contains 1 CFU of bacteria unless one performs a test on the entire 25 g sample, there is a high possibility that one would not be able to accurately detect the bacteria even with a sensor that has the sensitivity to detect 1 CFU. So, the sensitivity of an assay not only depends on the limit of detection but also on the sample size. Therefore, we proposed to perform MaCIA in concert with the traditional enrichment step, to offer a more reliable and accurate testing result.

The assay sensitivity was also affected by the food matrices tested. Ground chicken, raw eggs, whole milk, and cake mix

were chosen since these products were implicated in *Salmonella* outbreaks, and they also represent foods with high protein, fat, or carbohydrate contents. In milk, the detection limit for *S.* Enteritidis was 1×10^5 CFU/mL while in ground chicken, 1×10^6 CFU/mL, in cake mix, 1×10^7 CFU/mL, and in egg, 1×10^8 CFU/mL (**Figure 4C**). Among the foods tested, eggs exhibited the highest interference while milk had the least. Egg contains about 13 g protein and 11 g fat per 100 g while whole milk contains only 3.15 g of protein and 3.25 g of fat per 100 g (Kuang et al., 2018).

Mammalian cell-based immunoassay is highly specific for *S.* Enteritidis and did not show any non-specific reaction with other *Salmonella* serovars, non-*Salmonella* organisms, or natural microflora present in uninoculated food samples. The specificity of MaCIA is attributed to the specificity of the reporter antibody, mAb-2F11 used, that binds the O-antigen (LPS) on the surface of *S.* Enteritidis (Masi and Zawistowski, 1995; Jaradat et al., 2004). The advantage of the MaCIA platform is that the specificity depends on the primary reporter antibody used. We have demonstrated that using a commercial anti-*Salmonella* mAb-F68C (Thermo-Fisher) as a reporter antibody, *Salmonella* enterica serovar Typhimurium can be detected on the MaCIA platform (**Table 1**). These results indicate that the MaCIA platform is versatile and can be adapted for a different target pathogen using an appropriate antibody.

The accuracy of MaCIA for *S.* Enteritidis was also confirmed by comparing the results with the reference methods, such as the FDA-BAM, USDA-FSIS, and PCR (**Figure 6D**). The three primer sets that were used in PCR (**Supplementary Table 3**) target *IE-1, IE-2* in *S.* Enteritidis, and *InvA* in *S.* Enteritidis and *S.* Typhimurium (Fratamico and Strobaugh, 1998; Wang and Yeh, 2002; Paião et al., 2013), which again confirm the accuracy of MaCIA for its ability to detect *S.* Enteritidis from spiked food samples.

The major advancement of the MaCIA is its extended shelf-life, at least for 14 weeks, that was achieved through formalin-fixation of HCT-8 cells. Formalin is routinely used to preserve tissues and cells and it protects protein from denaturation (Eltoum et al., 2001). Therefore, receptor molecules on formalin-fixed HCT-8 cells, remained active and enabled viable *Salmonella* binding without diminishing MaCIA's performance. Previously, many attempts have been made to extend the shelf-life (functionality) of cells in CBB; however, none were satisfactory. Bhunia et al. (1995) used ultra-low temperature (freezing at −80 and −196°C) to extend the shelf-life of cells (up to 8 weeks) before performing the cytotoxicity assay for *L. monocytogenes*. However, the major drawback was the generation of high background signal originating from freeze-injured or dead mammalian cells. Banerjee et al. (2007) used modified growth conditions that included 5% fetal calf serum without any exogenous CO_2 and was able to extend the viability of the lymphocyte cell line for 6–7 days at room temperature. Curtis et al. (2009) used an automated media delivery system integrated with a thermoelectric controller to keep endothelial cells healthy up to 16 weeks. More recently, Jiang et al. (2018) used a screen-printed hydrogel-encapsulated rat basophilic leukemia mast cell-based electrochemical sensor for the detection of quorum sensing molecules for fish spoilage and the sensor-generated stable signal for 10 days. However, these attempts required incorporating mammalian cells in a specially designed external device to ensure the success of detection.

CONCLUSION

In conclusion, the present study demonstrates that MaCIA is a highly specific functional cell-based assay coupled with an immunoassay for the rapid and specific detection of the viable target pathogen. *S.* Enteritidis was used as a model pathogen which was successfully detected from food samples (ground chicken, shelled egg, whole milk, and cake mix) in 16–21 h using a conventional sample enrichment set up. The assay time (sample-to-result) was shortened to 10–12 h when an on-cell (on the MaCIA platform) sample enrichment was used. Thus, MaCIA could serve as a universal platform for other pathogens provided an appropriate cell line and a pathogen-specific antibody is used. The extended shelf-life of mammalian cells made MaCIA an attractive screening tool for point-of-need deployment. Furthermore, the MaCIA platform (24-well tissue culture plate) is suitable for testing at least 10 samples (plus positive and negative controls) in duplicate on a single plate thus reducing overall cost per sample testing.

MATERIALS AND METHODS

Mammalian Cell Culture

HCT-8 cell line (ATCC, Manassas, VA, United States) was maintained in Dulbecco's modified Eagles medium (DMEM; Thermo Fisher Scientific) with 10% fetal bovine serum (FBS; Bio-Techne Sales Corp, Minneapolis, MN, United States) at 37°C with 5% CO_2 in cell culture flasks (T25). For all experiments, HCT-8 cells were seeded in 24-well tissue culture plates (Fisher Scientific) at a density of 5×10^4 cell/mL/well. Media were replaced on day 4 and a final cell density of 2×10^5 cell/mL (monolayer) was achieved on day 5. Cell monolayers were washed twice with PBS (0.1 M, pH 7.0) and used immediately (Live HCT-8 cell assay) or exposed to 4% formaldehyde (Polysciences Inc., Warrington, PA) of 500 μL/well and incubated at room temperature for 10 min (Formalin-fixed HCT-8). Formaldehyde solution was removed and the cell monolayers were washed three times with PBS. Formalin-fixed cells were stored in 1 mL PBS/well for 14 weeks at 4°C or until use.

Bacterial Culture and Growth Media

Bacterial strains (**Table 1**) were stored as 10% glycerol stocks at −80°C. To revive frozen cultures, each strain was streaked onto tryptic soy agar (TSA; Thermo Fisher Scientific, Rochester, NY, United States) plate and incubated at 37°C for 18 h to obtain pure colonies. A single colony of each strain was inoculated and propagated in tryptic soy broth containing 0.5% yeast extract (TSBYE; Thermo Fisher Scientific) at 37°C for 18 h with shaking at 120 rpm.

Development and Specificity of MaCIA

HCT-8 cell monolayers were prepared and maintained as described above in 24-well plates. Overnight grown bacterial cultures (**Table 1**) were diluted in PBS to achieve a cell concentration of 1×10^7 CFU/ml. To obtain dead cells, cell suspensions were treated with heat (80°C for 10 min) or formaldehyde (4% for 10 min) and plated on TSA to ensure bacterial inactivation. One milliliter of bacterial cell suspensions was added into each well containing HCT-8 cells and incubated for 30 min at 37°C (Jaradat and Bhunia, 2003; Barrila et al., 2017). Cell monolayers (live or formalin-fixed) were washed 2–3 times with PBS gently and sequentially probed with either mAb-2F11 (3.06 μg/mL) (Jaradat et al., 2004) or mAb-F68C (0.2 μg/mL; Catalog # MA1-7443; Thermo Fisher Scientific) as primary antibodies, and anti-mouse HRP-conjugated IgG (0.1 μg/mL; Cell-Signaling, Danvers, MA, United States) as secondary antibodies for 1.5 h each at room temperature. Both antibodies were suspended in PBS containing 3% bovine serum albumin (BSA; Sigma-Aldrich). For one-step antibody probing, both mAb-2F11 and anti-mouse HRP conjugated secondary antibodies were mixed in PBS containing 3% BSA and incubated for 1.5 h. Cell monolayers were washed 3 times with PBS and the color was developed by adding 500 μl/well substrate solution (o-phenylenediamine, OPD) containing hydrogen-peroxide; Sigma-Aldrich). The oxidative coupling of OPD to 2,3-diaminophenazine, an orange-brown substance, was catalyzed by HRP at room temperature in the dark for 10 min. The intensity of the colored product was measured using a microplate reader (BioTek, Winooski, VT, United States) at a wavelength of 450 nm.

Sandwich ELISA

High-affinity (4HBX) ELISA plates (Thermo Fisher Scientific) were coated with mAb-2F11 for 2 h at 37°C, followed by 3 times wash with PBS-T (PBS containing 0.01% Tween-20). Freshly prepared BSA-PBS solution (1 mg/mL) was used for blocking at 4°C overnight, followed by 2 × PBS-T wash. Freshly prepared viable or formalin-inactivated cells of *S.* Enteritidis (1×10^8 cells/100 μl) were added to each well and incubated at 37°C for 30 min or 2 h. Anti- *Salmonella* pAb-3288 (2.86 μg/mL) used as a reporter (Abdelhaseib et al., 2016) and an HRP-conjugated anti-rabbit antibody (0.25 μg/mL) as the secondary antibody. After 3 washes with PBS-T, the OPD substrate was added and the absorbance (450 nm) was measured.

Western Blot

The whole-cell lysate of *L. monocytogenes* F4244, *P. aeruginosa* PRI99, *E. coli* EDL933, and *S.* Enteritidis PT21 overnight cultures (5 mL each) was prepared by sonication (Branson, Danbury, CT, United States). Bacterial samples were sonicated in an ice bucket (three 10 s cycles at 30-s intervals) and centrifuged for 10 min at 14,000 rpm (Eppendorf) at 4°C to separate the soluble fraction (supernatant) from the bacterial debris (pellet). The protein concentration was determined by the BCA method (Thermo Fisher Scientific). Equal amounts of proteins were separated on SDS-PAGE gel (10% polyacrylamide) and electro-transferred to polyvinylidene difluoride (PVDF) membrane (Fisher Scientific) (Singh et al., 2016; Drolia et al., 2018). Primary and secondary antibodies were diluted as above. Membranes were first probed with mAb-2F11 at 4°C overnight, and then with anti-mouse HRP conjugated antibody at room temperature for 1.5 h. LumiGLO reagent (Cell-Signaling Technology) was used to visualize the bands using the Chemi-Doc XRS system (Bio-Rad).

Immunofluorescence and Giemsa Staining

After exposure of formalin-fixed HCT-8 cell monolayers to viable or dead *S.* Enteritidis (1×10^8 cells/ml) for 30-min, the wells of the chambered slides (Fisher Scientific) were washed with PBS to remove unattached bacterial cells (as above). After immunoprobing with mAb-2F11, the monolayers were washed and probed with Alexa Fluor 488 conjugated anti-mouse antibody for 1.5 h at room temperature in the dark, followed by three PBS wash. Note, antibody concentrations used were the same as above. The monolayers were counterstained with DAPI (500 ng/mL; Cell-Signaling) for nuclear staining and the slides were mounted using an antifade reagent (Cell-Signaling). Images were acquired using the Nikon A1R confocal microscope with a Plano AP VC oil immersion objective (Drolia et al., 2018) and were processed with the Nikon Elements software at the Purdue Bindley Bioscience Imaging Facility.

For Giemsa staining, the formalin-fixed HCT-8 cell monolayers were exposed to viable or dead *S.* Enteritidis cells as above, air-dried, and immersed in Giemsa staining solution for 20 min. Giemsa staining solution was prepared using a 20-fold dilution of the KaryoMAX Giemsa staining solution (Thermo-Fisher) in deionized water. The slides were examined under a Leica DAS Microscope at the magnification of 1,000×.

Sensitivity of MaCIA

HCT-8 cell monolayers were prepared and maintained as described above in 24-well tissue culture plates. Overnight grown fresh *S.* Enteritidis PT21 culture was serially diluted to obtain 1×10^8 CFU/mL to 1×10^4 CFU/mL using PBS or homogenized 25 g food samples (**Supplementary Table 2**) in 225 mL BPW (Becton Dickinson, Sparks, MD, United States). One milliliter of each diluted sample was added onto HCT-8 cell monolayer and was incubated at 37°C for 30 min. The remaining steps were the same as above.

Detection of Stressed Cells Using MaCIA

Freshly prepared *S.* Enteritidis cells (2.17×10^8 CFU/ml) suspended in TSB were exposed to cold (4°C), heat (45°C), acidified TSB (pH, 4.5) and 5.5% NaCl for 3 h, as reported before (Hahm and Bhunia, 2006). Bacterial cells were washed with PBS and added onto the fixed HCT-8 monolayer for 30-min and probed with mAb-2F11 as above.

Salmonella Growth Kinetics Assessment

Overnight-grown *S.* Enteritidis PT21 cultures were serially diluted in PBS to achieve a concentration of 1×10^2 CFU/mL. One hundred microliters of the diluted culture were added into 25 g of each ground chicken, whole fat milk, liquid eggs, and

cake mix with 225 mL BPW and were incubated at 4°C for 24 h. The samples were then incubated at 37°C for 20 h with shaking at 120 rpm and enumerated on XLD (xylose lysine deoxycholate) agar plates (Remel, San Diego, CA) at every hour. *S.* Enteritidis counts in artificially inoculated samples at earlier stages of growth was determined by directly plating 1, 0.5, 0.1 mL of the sample on XLD plates with four repeats (1, 2, and 3 h); and *S.* Enteritidis counts from the later stages of growth (3 h and after) was obtained after serially diluting the samples in PBS. The growth of *S.* Enteritidis in food samples enriched using BPW was modeled using the Gompertz equation (Silk et al., 2002; Kim and Bhunia, 2008) through Prism software version 8.0. Lag-phase duration (LPD) and exponential growth rate (EGR) were calculated from the Gompertz model and were used to determine an enrichment time required for each food product to reach an optimum *S.* Enteritidis concentration required for detection by MaCIA, assuming the initial concentration was 1 CFU/25 g of sample.

Food Sample Testing With MaCIA and Validation With the FDA and USDA Methods

Food samples (ground poultry, milk, egg, or cake mix) were inoculated with variable concentrations of *S.* Enteritidis PT21. To simulate cold storage, inoculated foods were stored at 4°C for 24 h. Samples (25 g in 225 mL BPW) were then homogenized or pummeled using hands and incubated at 37°C for 14–19 h (**Supplementary Table 1**) with shaking at 120 rpm. One milliliter of enriched food sample was added into each well of MaCIA for 30 min, followed by immunoprobing as above.

For direct on-cell enrichment, the homogenized food suspensions (1 ml of each food sample) were dispensed into wells containing formalin-fixed HCT-8 cells (MaCIA) and incubated for 7–9 h. After the removal of food samples, wells were washed 3 times with PBS before immunoprobing and color development. *Salmonella* counts in enriched food samples (inoculated or uninoculated) were enumerated on XLD plates. The presence of background bacteria in uninoculated food samples was assessed on TSA plates after incubation at 37°C for 24 h. For the blind test, the inoculation of the samples was performed by XB, while the MaCIA test was done by LX in a blinded fashion without prior knowledge of samples that were inoculated with *Salmonella*.

Inoculated food samples were also analyzed by the FDA-BAM (FDA, 2001) or USDA-FSIS (2013) method as before. The ground chicken was processed according to the USDA-FSIS method, while shelled egg, whole milk, and cake mix were prepared based on the FDA-BAM. Twenty-five gram of each prepared sample was then enriched in 225 mL of BPW (ground chicken), trypticase soy broth (shelled egg), and lactose broth (whole milk and cake mix) at 37°C for 24 h followed by sequential enrichment in RV (Rappaport-Vassiliadis) broth and TT (tetrathionate) broth at 42°C for 24 h. Samples were then plated on selective BGS (Brilliant Green Agar with Sulfadiazine) or XLD agar plates to isolate colonies, which were further confirmed by PCR assay.

For PCR assay, DNA was extracted from the isolated colonies by the boiling method (Kim and Bhunia, 2008; Kim et al., 2015). The primer sequences and the putative product sizes for each amplicon are listed in **Supplementary Table 3** (Wang and Yeh, 2002; Paião et al., 2013). PCR reaction mixture (25 μL) contained 1 μL of DNA template, 0.2 μM of each primer, 2.5 mM $MgCl_2$, 200 μM of dNTP, 1 x GoTaq Flexi buffer of buffer and 1 U of GoTaq Flexi DNA polymerase (Promega) (Singh et al., 2014). The PCR amplification was performed in the Proflex PCR system with an initial denaturation at 94°C for 3 min, 35 amplification cycles consisting of 1 min of denaturation at 94°C, 1.5 min of annealing at 50°C, and 1.5 min of elongation at 72°C. DNA amplicons were analyzed using agarose gel (1.5%, wt/vol) electrophoresis containing ethidium bromide (0.5 μg of /mL).

Swab Sample Testing

Chicken thigh cuts (procured from a local grocery store) were inoculated with overnight grown *S.* Enteritidis PT21 at 1.35×10^3 to 1.35×10^5 CFU per 50 cm^2 evenly on the skin of chicken thighs. Inoculated samples were stored at 4°C for 24 h. BPW-soaked sterile rayon tipped swab applicators (Puritan, Guilford, ME, United States) were used to swab the chicken skin and were vortexed in 1.1 mL of BPW. One milliliter of the sample was added into each well of MaCIA and incubated at 37°C for 7 h for on-cell enrichment, followed by immunoprobing as above. The rest of the swabbed sample (0.1 mL) was used to enumerate *Salmonella* on XLD plates.

Statistical Analysis

All data were analyzed using GraphPad Prism software (San Diego, CA, United States). The unpaired *t*-test was used when comparing two datasets. Tukey's multiple comparison test was also used when comparing more than two datasets. All data were presented with mean ± standard error of the mean (SEM).

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

LX, XB, RD, and AB conceived and designed the experiments. LX, XB, ST, YL, and RD performed the experiments. LX, XB, and AB analyzed the data. LX and AB wrote the manuscript. All authors approved the final version.

FUNDING

This material is based upon work supported by the U.S. Department of Agriculture, Agricultural Research Service, under Agreement No. 59-8072-6-001, and the USDA National Institute of Food and Agriculture (Hatch accession no. 1016249). Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the view of the U.S. Department of Agriculture.

ACKNOWLEDGMENTS

We thank S. Lelièvre, B. M. Applegate, T. Kwok, and S. Chittiboyina for advice and D. Liu, Z. Tang, W. Lv, M. Samaddar for technical assistance.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2020.575615/full#supplementary-material

Supplementary Figure 1 | Detection sensitivity of MaCIA tested against the different concentrations of *S. enterica* serovar Enteritidis cells suspended in ground chicken slurry (in buffered peptone water). D, dead cells.

Supplementary Figure 2 | Experimental set-up of the blind test using an on-cell (MaCIA) enrichment method. **(A,B)** The checkerboard filled areas correspond with sample a; the diagonal stripes filled areas correspond with sample b; No pattern-filled area corresponds with negative control for each food product. The numbers in the table represent the concentration (CFU/mL) of the inoculant, *S.* Enteritidis PT21. **(C)** Blinded test using on-cell enrichment. Positive samples were inoculated with 25 CFU/mL cold-stored *S.* Enteritidis PT21. Neg: negative control. a, b: blind tested samples.

Supplementary Figure 3 | Tryptic soy agar (TSA) plates showing the presence of background bacterial populations from different food samples except for the eggs.

Supplementary Table 1 | Proposed enrichment time for different food products before testing with MaCIA.

Supplementary Table 2 | Total detection time required for each method.

Supplementary Table 3 | PCR primer sequences used.

REFERENCES

Abdelhaseib, M. U., Singh, A. K., Bailey, M., Singh, M., El-Khateib, T., and Bhunia, A. K. (2016). Fiber optic and light scattering sensors: complimentary approaches to rapid detection of *Salmonella enterica* in food samples. *Food Control* 61, 135–145. doi: 10.1016/j.foodcont.2015.09.031

Banerjee, P., and Bhunia, A. K. (2009). Mammalian cell-based biosensors for pathogens and toxins. *Trends Biotechnol.* 27, 179–188. doi: 10.1016/j.tibtech.2008.11.006

Banerjee, P., and Bhunia, A. K. (2010). Cell-based biosensor for rapid screening of pathogens and toxins. *Biosens. Bioelectron.* 26, 99–106. doi: 10.1016/j.bios.2010.05.020

Banerjee, P., Lenz, D., Robinson, J. P., Rickus, J. L., and Bhunia, A. K. (2008). A novel and simple cell-based detection system with a collagen-encapsulated B-lymphocyte cell line as a biosensor for rapid detection of pathogens and toxins. *Lab. Invest.* 88, 196–206. doi: 10.1038/labinvest.3700703

Banerjee, P., Morgan, M. T., Rickus, J. L., Ragheb, K., Corvalan, C., Robinson, J. P., et al. (2007). Hybridoma Ped-2E9 cells cultured under modified conditions can sensitively detect *Listeria monocytogenes* and *Bacillus cereus*. *Appl. Microbiol. Biotechnol.* 73, 1423–1434. doi: 10.1007/s00253-006-0622-0

Barrila, J., Yang, J., Crabbé, A., Sarker, S. F., Liu, Y., Ott, C. M., et al. (2017). Three-dimensional organotypic co-culture model of intestinal epithelial cells and macrophages to study *Salmonella enterica* colonization patterns. *Npj Microgravity* 3:10.

Bäumler, A. J., Tsolis, R. M., and Heffron, F. (1996). The lpf fimbrial operon mediates adhesion of *Salmonella* typhimurium to murine Peyer's patches. *Proc. Nat. Acad. Sci. U.S.A.* 93, 279–283. doi: 10.1073/pnas.93.1.279

Bell, R. L., Jarvis, K. G., Ottesen, A. R., Mcfarland, M. A., and Brown, E. W. (2016). Recent and emerging innovations in *Salmonella* detection: a food and environmental perspective. *Microb. Biotechnol.* 9, 279–292. doi: 10.1111/1751-7915.12359

Bhunia, A. K. (2011). Rapid pathogen screening tools for food safety. *Food Technol.* 65, 38–43.

Bhunia, A. K. (2014). One day to one hour: how quickly can foodborne pathogens be detected? *Future Microbiol.* 9, 935–946. doi: 10.2217/fmb.14.61

Bhunia, A. K. (2018). "Salmonella enterica," in *Foodborne Microbial Pathogens: Mechanisms and Pathogenesis*, ed. A. K. Bhunia (New York, NY: Springer), 271–287.

Bhunia, A. K., Westbrook, D. G., Story, R., and Johnson, M. G. (1995). Frozen stored murine hybridoma cells can be used to determine the virulence of *Listeria monocytogenes*. *J. Clin. Microbiol.* 33, 3349–3351. doi: 10.1128/jcm.33.12.3349-3351.1995

Buzby, J. C., Farah-Wells, H., and Hyman, J. (2014). *The Estimated Amount, Value, and Calories of Postharvest Food Losses at the Retail and Consumer Levels in the United States*. Washington, DC: USDA-ERS, 121. Economic Information Bulletin Number.

CDC (2018). *Salmonella-reports of selected Salmonella outbreak investigations*. Atlanta, GA: CDC.

CDC (2020). *Salmonella*. Atlanta, GA: CDC.

Curtis, T. M., Widder, M. W., Brennan, L. M., Schwager, S. J., Van Der Schalie, W. H., Fey, J., et al. (2009). A portable cell-based impedance sensor for toxicity testing of drinking water. *Lab Chip* 9, 2176–2183. doi: 10.1039/b901314h

Dibao-Dina, A., Follet, J., Ibrahim, M., Vlandas, A., and Senez, V. (2015). Electrical impedance sensor for quantitative monitoring of infection processes on HCT-8 cells by the waterborne parasite *Cryptosporidium*. *Biosens. Bioelectron.* 66, 69–76. doi: 10.1016/j.bios.2014.11.009

Dos Reis, R. S., and Horn, F. (2010). Enteropathogenic *Escherichia coli*, Samonella, Shigella and Yersinia: cellular aspects of host-bacteria interactions in enteric diseases. *Gut Pathog.* 2:8. doi: 10.1186/1757-4749-2-8

Drolia, R., and Bhunia, A. K. (2019). Crossing the intestinal barrier via *Listeria adhesion* protein and internalin A. *Trends Microbiol.* 27, 408–425. doi: 10.1016/j.tim.2018.12.007

Drolia, R., Tenguria, S., Durkes, A. C., Turner, J. R., and Bhunia, A. K. (2018). *Listeria adhesion* protein induces intestinal epithelial barrier dysfunction for bacterial translocation. *Cell Host & Microbe* 23, 470–484. doi: 10.1016/j.chom.2018.03.004

Elkhishin, M. T., Gooneratne, R., and Hussain, M. A. (2017). Microbial safety of foods in the supply chain and food security. *Adv. Food Technol. Nutr. Sci. Open J.* 3, 22–32.

Eltoum, I., Fredenburgh, J., and Grizzle, W. E. (2001). Advanced Concepts in Fixation: 1. Effects of fixation on immunohistochemistry, reversibility of fixation and recovery of proteins, nucleic acids, and other molecules from fixed and processed tissues. 2. Developmental methods of fixation. *J. Histotechnol.* 24, 201–210. doi: 10.1179/his.2001.24.3.201

Eriksson, E., and Aspan, A. (2007). Comparison of culture, ELISA and PCR techniques for *Salmonella* detection in faecal samples for cattle, pig and poultry. *BMC Vet. Res.* 3:21. doi: 10.1186/1746-6148-3-21

FDA (2001). *Bacteriological Analytical Manual Online*, 8th Edn. Arlington, VA: AOAC International.

Finlay, B. B., and Falkow, S. (1997). Common themes in microbial pathogenicity revisited. *Microbiol. Mol. Biol. Rev.* 61, 136–169. doi: 10.1128/.61.2.136-169.1997

Fratamico, P. M., and Strobaugh, T. P. (1998). Simultaneous detection of *Salmonella* spp and *Escherichia coli* O157:H7 by multiplex PCR. *J. Indust. Microbiol. Biotechnol.* 21, 92–98.

Galikowska, E., Kunikowska, D., Tokarska-Pietrzak, E., Dziadziuszko, H., Łoś, J. M., Golec, P., et al. (2011). Specific detection of *Salmonella enterica* and *Escherichia coli* strains by using ELISA with bacteriophages as recognition agents. *Eur. J. Clin. Microbiol. Infect. Dis.* 30, 1067–1073. doi: 10.1007/s10096-011-1193-2

Gray, K. M., Banada, P. P., O'neal, E., and Bhunia, A. K. (2005). Rapid Ped-2E9 cell-based cytotoxicity analysis and genotyping of *Bacillus* species. *J. Clin. Microbiol.* 43, 5865–5872. doi: 10.1128/jcm.43.12.5865-5872.2005

Hahm, B. K., and Bhunia, A. K. (2006). Effect of environmental stresses on antibody-based detection of *Escherichia coli* O157:H7, *Salmonella enterica*

serotype *Enteritidis* and *Listeria monocytogenes*. *J. Appl. Microbiol.* 100, 1017–1027. doi: 10.1111/j.1365-2672.2006.02814.x

Hu, L., and Wai, T. T. (2017). Comparing invasive effects of five foodborne bacterial pathogens in human embryonic intestine 407 cells and human ileocecum HCT-8 cells. *Asian Pacific J. Trop. Biomed.* 7, 937–944. doi: 10.1016/j.apjtb.2017.09.004

Jaradat, Z. W., and Bhunia, A. K. (2003). Adhesion, invasion and translocation characteristics of *Listeria monocytogenes* serotypes in Caco-2 cell and mouse models. *Appl. Environ. Microbiol.* 69, 3640–3645. doi: 10.1128/aem.69.6.3640-3645.2003

Jaradat, Z. W., Bzikot, J. H., Zawistowski, J., and Bhunia, A. K. (2004). Optimization of a rapid dot-blot immunoassay for detection of *Salmonella enterica* serovar Enteritidis in poultry products and environmental samples. *Food Micrrobiol.* 21, 761–769. doi: 10.1016/j.fm.2004.01.010

Jiang, D., Liu, Y., Jiang, H., Rao, S., Fang, W., Wu, M., et al. (2018). A novel screen-printed mast cell-based electrochemical sensor for detecting spoilage bacterial quorum signaling molecules (N-acyl-homoserine-lactones) in freshwater fish. *Biosens. Bioelectron.* 102, 396–402. doi: 10.1016/j.bios.2017.11.040

Kasturi, K. N., and Drgon, T. (2017). Real-time PCR Method for detection of *Salmonella* spp. *in environmental samples*. *Appl. Environ. Microbiol.* 83: e00644-17.

Kim, H., and Bhunia, A. K. (2008). SEL, a selective enrichment broth for simultaneous growth of *Salmonella enterica*, *Escherichia coli* O157:H7, and *Listeria monocytogenes*. *Appl. Environ. Microbiol.* 74, 4853–4866. doi: 10.1128/aem.02756-07

Kim, K. P., Singh, A. K., Bai, X., Leprun, L., and Bhunia, A. K. (2015). Novel PCR assays complement laser biosensor-based method and facilitate *Listeria* species detection from food. *Sensors* 15, 22672–22691. doi: 10.3390/s150922672

Kirk, M. D., Pires, S. M., Black, R. E., Caipo, M., Crump, J. A., Devleesschauwer, B., et al. (2015). World Health Organization estimates of the global and regional disease burden of 22 foodborne bacterial, protozoal, and viral diseases, 2010: a data synthesis. *PLoS Med.* 12:e1001921. doi: 10.1371/journal.pmed.1001921

Kline, K. A., Fälker, S., Dahlberg, S., Normark, S., and Henriques-Normark, B. (2009). Bacterial adhesins in host-microbe interactions. *Cell Host Microbe* 5, 580–592. doi: 10.1016/j.chom.2009.05.011

Kolenda, R., Ugorski, M., and Grzymajło, K. (2019). Everything you always wanted to know about *Salmonella* type 1 fimbriae, but were afraid to ask. *Front. Microbiol.* 10:1017. doi: 10.3389/fmicb.2019.01017

Kuang, H., Yang, F., Zhang, Y., Wang, T., and Chen, G. (2018). The impact of egg nutrient composition and its consumption on cholesterol homeostasis. *Cholesterol* 2018:6303810.

Lee, K.-M., Runyon, M., Herrman, T. J., Phillips, R., and Hsieh, J. (2015). Review of *Salmonella* detection and identification methods: aspects of rapid emergency response and food safety. *Food Control* 47, 264–276. doi: 10.1016/j.foodcont.2014.07.011

Mansfield, L. P., and Forsythe, S. J. (2000). The detection of *Salmonella* using a combined immunomagnetic separation and ELISA end-detection procedure. *Lett. Appl. Microbiol.* 31, 279–283. doi: 10.1046/j.1472-765x.2000.00811.x

Masi, A., and Zawistowski, J. (1995). Detection of live and heat-treated *Salmonella* Enteritidis by a D1-serospecific anti-lipopolysaccharide O-9 monoclonal antibody. *Food Ag Immunol.* 7, 351–363. doi: 10.1080/09540109509354895

McKee, M. L., and O'Brien, A. D. (1995). Investigation of enterohemorrhagic *Escherichia coli* O157: H7 adherence characteristics and invasion potential reveals a new attachment pattern shared by intestinal E. *coli*. *Infect. Immun.* 63, 2070–2074. doi: 10.1128/iai.63.5.2070-2074.1995

McWilliams, B. D., and Torres, A. G. (2014). Enterohaemorrhagic *Escherichia coli* adhesins. *Microbiol. Spectrum*. 2:EHEC-0003.

Ngamwongsatit, P., Banada, P. P., Panbangred, W., and Bhunia, A. K. (2008). WST-1-based cell cytotoxicity assay as a substitute for MTT-based assay for rapid detection of toxigenic *Bacillus* species using CHO cell line. *J. Microbiol. Methods* 73, 211–215. doi: 10.1016/j.mimet.2008.03.002

Paião, F. G., Arisitides, L. G. A., Murate, L. S., Vilas-Bôas, G. T., Vilas-Boas, L. A., and Shimokomaki, M. (2013). Detection of *Salmonella* spp, Salmonella Enteritidis and Typhimurium in naturally infected broiler chickens by a multiplex PCR-based assay. *Braz. J. Microbiol.* 44, 37–42. doi: 10.1590/s1517-83822013005000002

Rajapaksha, P., Elbourne, A., Gangadoo, S., Brown, R., Cozzolino, D., and Chapman, J. (2019). A review of methods for the detection of pathogenic microorganisms. *Analyst* 144, 396–411. doi: 10.1039/c8an01488d

Ricke, S. C., Kim, S. A., Shi, Z., and Park, S. H. (2018). Molecular-based identification and detection of *Salmonella* in food production systems: current perspectives. *J. Appl. Microbiol.* 125, 313–327. doi: 10.1111/jam.13888

Schlaberg, R., Chiu, C. Y., Miller, S., Procop, G. W., Weinstock, G., Professional Practice, C., et al. (2017). Validation of metagenomic next-generation sequencing tests for universal pathogen detection. *Arch. Pathol. Lab. Med.* 141, 776–786. doi: 10.5858/arpa.2016-0539-ra

Silk, T. M., Roth, T. M. T., and Donnelly, C. W. (2002). Comparison of growth kinetics for healthy and heat-injured *Listeria monocytogenes* in eight enrichment broths. *J. Food Prot.* 65, 1333–1337. doi: 10.4315/0362-028x-65.8.1333

Singh, A. K., Leprun, L., Drolia, R., Bai, X., Kim, H., Aroonnual, A., et al. (2016). Virulence gene-associated mutant bacterial colonies generate differentiating two-dimensional laser scatter fingerprints. *Appl. Environ. Microbiol.* 82, 3256–3268. doi: 10.1128/aem.04129-15

Singh, A. K., Bettasso, A. M., Bae, E., Rajwa, B., Dundar, M. M., Forster, M. D., et al. (2014). Laser optical sensor, a label-free on-plate *Salmonella enterica* colony detection tool. *mBio* 5:e01019-13.

Snyder, T. R., Boktor, S. W., and M'ikanatha, N. M. (2019). Salmonellosis outbreaks by food vehicle, serotype, season, and geographical location. United States, 1998 to 2015. *J. Food Prot.* 82, 1191–1199. doi: 10.4315/0362-028x.jfp-18-494

To, C., Banerjee, P., and Bhunia, A. K. (2020). "Cell-based biosensor for rapid screening of pathogens and toxins," in *Handbook of Cell Biosensors*, ed. G. Thouand (Cham: Springer International Publishing), 1–16. doi: 10.1007/978-3-319-47405-2_102-1

To, C. Z., and Bhunia, A. K. (2019). Three dimensional Vero cell-platform for rapid and sensitive screening of Shiga-toxin producing *Escherichia coli*. *Front. Microbiol.* 10:949. doi: 10.3389/fmicb.2019.00949

USDA-FSIS (2013). Isolation and Identification of Salmonella from Meat, Poultry, Pasteurized Egg and Catfish Products. Method Number MLG 4.06. Avaliable at: http://www.fsis.usda.gov/wps/wcm/connect/700c05fe-06a2-492a-a6e1-3357f7701f52/MLG-4.pdf?MOD=AJPERES (accessed November 5, 2019).

Wagner, C., and Hensel, M. (2011). "Adhesive mechanisms of *Salmonella enterica*," in *Bacterial Adhesion*, eds D. Linke and A. Goldman (Berlin: Springer), 17–34. doi: 10.1007/978-94-007-0940-9_2

Wang, S. J., and Yeh, D. B. (2002). Designing of polymerase chain reaction primers for the detection of *Salmonella* Enteritidis in foods and faecal samples. *Lett. Appl. Microbiol.* 34, 422–427. doi: 10.1046/j.1472-765x.2002.01114.x

Wu, V. C. H. (2008). A review of microbial injury and recovery methods in food. *Food Microbiol.* 25, 735–744. doi: 10.1016/j.fm.2008.04.011

Ye, Y., Guo, H., and Sun, X. (2019). Recent progress on cell-based biosensors for analysis of food safety and quality control. *Biosens. Bioelectron.* 126, 389–404. doi: 10.1016/j.bios.2018.10.039

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Xu, Bai, Tenguria, Liu, Drolia and Bhunia. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 15 January 2021
doi: 10.3389/fmicb.2020.634255

Multiplex PCR for the Identification of Pathogenic *Listeria* in *Flammulina velutipes* Plant Based on Novel Specific Targets Revealed by Pan-Genome Analysis

Fan Li[1,2†], Qinghua Ye[1†], Moutong Chen[1†], Jumei Zhang[1], Liang Xue[1], Juan Wang[3], Shi Wu[1], Haiyan Zeng[1], Qihui Gu[1], Youxiong Zhang[1], Xianhu Wei[1], Yu Ding[4] and Qingping Wu[1]**

[1] Guangdong Provincial Key Laboratory of Microbial Safety and Health, State Key Laboratory of Applied Microbiology Southern China, Guangdong Institute of Microbiology, Guangdong Academy of Sciences, Guangzhou, China, [2] School of Biology and Biological Engineering, South China University of Technology, Guangzhou, China, [3] College of Food Science, South China Agricultural University, Guangzhou, China, [4] Department of Food Science and Technology, Jinan University, Guangzhou, China

OPEN ACCESS

Edited by:
Dario De Medici,
National Institute of Health (ISS), Italy

Reviewed by:
Reza Ranjbar,
Baqiyatallah University of Medical Sciences, Iran
Manuela Tamburro,
University of Molise, Italy

***Correspondence:**
Qingping Wu
wuqp203@163.com
Yu Ding
dingyu@jnu.edu.cn

[†] These authors have contributed equally to this work

Specialty section:
This article was submitted to Food Microbiology, a section of the journal Frontiers in Microbiology

Received: 27 November 2020
Accepted: 16 December 2020
Published: 15 January 2021

Citation:
Li F, Ye Q, Chen M, Zhang J, Xue L, Wang J, Wu S, Zeng H, Gu Q, Zhang Y, Wei X, Ding Y and Wu Q (2021) Multiplex PCR for the Identification of Pathogenic Listeria in Flammulina velutipes Plant Based on Novel Specific Targets Revealed by Pan-Genome Analysis. Front. Microbiol. 11:634255. doi: 10.3389/fmicb.2020.634255

Listeria spp. is an important foodborne disease agent, often found in the fresh mushroom (*Flammulina velutipes*) and its production environment. The aim of this study was to develop multiplex PCR for rapid identification of *Listeria monocytogenes* and *Listeria ivanovii*, and nonpathogenic *Listeria* in *F. velutipes* plants. Pan-genome analysis was first used to identify five novel *Listeria*-specific targets: one for the *Listeria* genus, one for *L. monocytogenes*, and three for *L. ivanovii*. Primers for the novel targets were highly specific in individual reactions. The detection limits were 10^3–10^4 CFU/mL, meeting the requirements of molecular detection. A mPCR assay for the identification of pathogenic *Listeria*, with primers targeting the novel genes specific for *Listeria* genus (*LMOSLCC2755_0944*), *L. monocytogenes* (*LMOSLCC2755_0090*), and *L. ivanovii* (*queT_1*) was then designed. The assay specificity was robustly verified by analyzing nonpathogenic *Listeria* and non-*Listeria* spp. strains. The determined detection limits were 2.0×10^3 CFU/mL for *L. monocytogenes* and 3.4×10^3 CFU/mL for *L. ivanovii*, for pure culture analysis. Further, the assay detected 7.6×10^4 to 7.6×10^0 CFU/10 g of pathogenic *Listeria* spiked into *F. velutipes* samples following 4–12 h enrichment. The assay feasibility was evaluated by comparing with a traditional culture-based method, by analyzing 129 samples collected from different *F. velutipes* plants. The prevalence of *Listeria* spp. and *L. monocytogenes* was 58.1% and 41.1%, respectively. The calculated κ factors for *Listeria* spp., *L. monocytogenes*, and *L. ivanovii* were 0.97, 0.97, and 1, respectively. The results of the novel mPCR assay were highly consistent with those of the culture-based method. The new assay thus will allow rapid, specific, and accurate detection and monitoring of pathogenic *Listeria* in food and its production environment.

Keywords: novel target gene, *Listeria*, pan-genome analysis, multiplex PCR, mushroom

INTRODUCTION

Listeria is a genus of gram-positive facultative anaerobic bacilli that is widely distributed in food and natural environments. The genus *Listeria* includes many species: *Listeria monocytogenes*, *Listeria innocua*, *Listeria ivanovii*, *Listeria seeligeri*, *Listeria welshimeri*, *Listeria grayi*, *Listeria rocourtiae*, *Listeria fleischmannii*, *and Listeria marthii*, etc. (Vitullo et al., 2013). Among them, *L. monocytogenes* and *L. ivanovii* are human and animal pathogens (Snapir et al., 2006; Jagadeesan et al., 2019; Lee et al., 2020). These bacteria can cause severe listeriosis, leading to spontaneous abortion, neonatal sepsis, and meningoencephalitis, and the post-infection mortality rate ranges from 20 to 30% (Radoshevich and Cossart, 2018; Ranjbar and Halaji, 2018), which places this pathogen among the most frequent causes of death from foodborne illnesses (Tamburro et al., 2015b).

Listeria spp. can contaminate food at multiple stages of production and distribution because it is ubiquitous in the environment and tolerates harsh environmental conditions (Tamburro et al., 2015a). Edible mushrooms are potential vehicles for the transmission of pathogenic *Listeria* (Chen et al., 2014, 2018). For example, our previous study found that the prevalence of *Listeria* spp. and *L. monocytogenes* in *F. velutipes* production plants is 52.5% (155/295) and 18.6% (55/295), respectively (Chen et al., 2014). However, currently, the gold standard method for detection of *Listeria* in *F. velutipes* is the conventional culture method (Chen et al., 2014; Tao et al., 2017), which is labor-intensive, expensive, and time-consuming. It is therefore important to develop rapid and accurate diagnostic techniques or tools for the detection and distinguishing of pathogenic *Listeria* in *F. velutipes*, to facilitate introduction of appropriate intervention measures in the *F. velutipes* production chain to reduce the risk of *Listeria* contamination.

Polymerase chain reaction (PCR) has been widely employed as a rapid and specific method for the detection of *Listeria* in a variety of foods and processing environments because of its high specificity, sensitivity, time-saving and easy operation (Norton et al., 2000; Ryu et al., 2013; Rosimin et al., 2016; Sheng et al., 2018). However, most of the reported PCR-based methods for the identification and characterization of *Listeria* target the bacterial virulence genes (Doijad et al., 2011; Mao et al., 2016; Li et al., 2017; Tamburro et al., 2018, 2019), or 16S and 23S rRNA genes (Czajka et al., 1993; Hellberg et al., 2013), which have a limited number of targets, and some genes cannot accurately identify target bacteria. For example, avirulent *L. monocytogenes* strains may lack one or more virulence genes due to mutations or low gene expression under certain conditions (Nishibori et al., 1995). In addition, the highly conserved sequence of the 16S rRNA gene may fail to distinguish between *Listeria* species, especially *L. monocytogenes*, *L. innocua*, and *L. welshimeri* (Hellberg et al., 2013; Tao et al., 2017). Therefore, mining of target genes with high species-specificity is vital for improved accuracy, specificity, and efficiency of pathogen detection.

At present (as of 27 May 2019), a large number of sequenced genomes of *Listeria* strains are available from the Genome Database of the National Center for Biotechnology Information (NCBI)[1]. Using a wealth of genome data, pan-genome analysis has become a representative discipline for studying an entire repertoire of gene families in genomes of a clade of pathogenic bacteria, which provides not only the whole set of genes shared among *Listeria* species but also can be applied for inter-species differentiation analysis to mine species-specific genes (Kuenne et al., 2013; Kim et al., 2020). In the current study, we aimed to devise a novel mPCR assay for accurate detection and identification of pathogenic *Listeria*. The pan-genome analysis was performed to screen the genus- and species-specific genes. Using the identified genes, a mPCR method was developed to simultaneously detect *L. monocytogenes*, *L. ivanovii*, and nonpathogenic *Listeria*. Then the assay was used to evaluate the prevalence of pathogenic *Listeria* (including *L. monocytogenes* and *L. ivanovii*) in samples from *F. velutipes* plants.

MATERIALS AND METHODS

Screening of Genus- and Species-Specific Novel Target Genes for *Listeria*

A total of 205 genome sequences of pathogenic bacteria were downloaded from the NCBI Genome Database (last accessed on 27 May 2019), including 165 *Listeria* sequences and 40 other common foodborne pathogens sequences. The specific information for the sequences is provided in **Supplementary Table S1**. The Pan-genome analysis was used to identify *Listeria* genus- and species-specific genes. Briefly, all nucleic acid sequences were annotated using Prokka v1.11 (Seemann, 2014). The output of Prokka was then used to construct a pan-genome analysis using Roary v3.11.2 (Page et al., 2015). A core genome was determined for each strain using a 99% cutoff, with a BLASTP identity cutoff of 85% (Chen et al., 2019; Pang et al., 2019). The absence/existence profile of all genes across strains was converted into a 0/1 matrix with local script. The *Listeria* genus-specific (or species-specific) genes were screened according to the following criteria: 100% presence in *Listeria* (or *L. monocytogenes*, or *L. ivanovii*) strains and no presence in non-*Listeria* (or non-target *Listeria* species) strains. Then these candidate targets were used as further screened against the nucleotide collection (nr/nt) databases using the online BLAST program[2] and PCR verification to ensure its specificity. At the same time, the specific targets reported in the previous studies [including *prs* specific for *Listeria* genus (Doumith et al., 2004), *actA* (Gouin et al., 1995; Zhou and Jiao, 2005), *hly* (Paziak-Domańska et al., 1999), and *inlA* (Jung et al., 2003) specific for *L. monocytogenes*, and *iactA* specific for *L. ivanovii* (Gouin et al., 1995)] were also analyzed in the 0/1 matrix to evaluate their absence/existence profile.

Bacterial Strains and Genomic DNA Extraction

A total of 126 bacteria strains were used in the study, which were obtained respectively from the Chinese Center for Disease

[1]https://www.ncbi.nlm.nih.gov/genome/

[2]https://blast.ncbi.nlm.nih.gov/Blast.cgi

Control and Prevention, the College of Food Science and Technology of Nanjing Agricultural University, the Guangdong Huankai Co., Ltd., the State Key Laboratory of Food Science and Technology of Nanchang University, and our laboratory (**Table 1** and **Supplementary Table S2**). For DNA extraction, all strains were inoculated in brain heart infusion (BHI) medium (Guangdong Huankai Co., Ltd., Guangzhou, China) and incubated for 8–12 h at 37 °C. DNA was extracted using DNeasy Blood and Tissue kit (Qiagen, Shanghai, China), according to the manufacturer's instructions. Extracted DNA was stored at –20 °C for PCR analysis.

Design and Validation of Genus- and Species-Specific Primers for *Listeria*

Primers for the candidate *Listeria* genus- and species-specific target genes were designed using the Oligo 7.0 software. Candidate primer sets were synthesized by Generay Biotech

TABLE 1 | *Listeria* spp. and other common foodborne pathogenic strains used in this study to verify the specificity of candidate targets.

Bacterial species	Serotype	Strain	Number of strains	Source*	PCR primer paris				
					SPP1	LM1	LIV1	LIV2	LIV3
L. monocytogenes	1/2a		21	a	+	+	–	–	–
			8	a		+			
	3a	[1]ATCC 51782	1	b	+	+			
	3a		2	b	+	+			
	1/2c		15	a	+	+	–	–	–
	4b	ATCC 19115	1	d	+	+	–	–	–
	4b	[2]CMCC 54007	1	d	+	+			
	4b		15	a	+	+			
	4d	ATCC 19117	1	c	+	+			
	4e	ATCC 19118	1	c	+	+			
	4ab	Murray B	1	c	+	+			
	1/2b		20	a	+	+	–	–	–
	3b		1	c	+	+			
	7	[3]SLCC 2428	1	c	+	+			
	4a	ATCC 19114	1	c	+	+			
	4c		12	a	+	+	–	–	–
L. innocua	6a	ATCC 33090	1	d	+	–	–	–	–
L. innocua			3	a	+	–			
L. ivanovii	5	ATCC 19119	1	e	+	–	+	+	+
L. seeligeri	1/2b	[4]CICC 21671	1	c	+	–	–	–	–
L. welshimeri	6b	ATCC 35897	1	e	+	–	–	–	–
L. grayi		ATCC 19120	1	d	+	–	–	–	–
Cronobacter sakazakii		ATCC 29544	1	d	–	–	–	–	–
Cronobacter sakazakii			1	a	–	–	–	–	–
Staphylococcus aureus		ATCC 25923	1	d	–	–	–	–	–
Staphylococcus aureus		ATCC 29213	1	d	–	–	–	–	–
Pseudomonas aeruginosa		ATCC 9027	1	d	–	–	–	–	–
Pseudomonas aeruginosa		ATCC 27853	1	d	–	–	–	–	–
Salmonella Enteritidis		CMCC 50335	1	d	–	–	–	–	–
Salmonella Typhimurium		ATCC 14028	1	d	–	–	–	–	–
Campylobacter jejuni		ATCC 33291	1	d	–	–	–	–	–
Escherichia coli		ATCC 25922	1	d	–	–	–	–	–
Shigella sonnei		CMCC(B) 51592	1	d	–	–	–	–	–
Bacillus cereus		ATCC 14579	1	a	–	–	–	–	–
Yersinia enterocolitica			2	d	–	–	–	–	–
Vibrio parahaemolyticus		ATCC 33847	1	d	–	–	–	–	–
Vibrio parahaemolyticus		ATCC 17802	1	d	–	–	–	–	–

**a, our laboratory; b, Chinese Center for Disease Control and Prevention, China; c, College of Food Science and Technology, Nanjing Agricultural University, China; d, Guangdong Huankai Co., Ltd., China; e, State Key Laboratory of Food Science and Technology, Nanchang University, China.*

[1]ATCC, American Type Culture Collection, United States.

[2]CMCC, China Medical Culture Collection, China.

[3]SLCC, Seeliger's Special Listeria Culture Collection, Germany.

[4]CICC, Center of Industrial Culture Collection, China. Result (±) indicate positive and negative signals.

Co., Ltd., Shanghai, China. Primer specificity was tested by PCR analysis of strains from the laboratory collection (**Table 1**). Detection limits were determined using serially (10-fold) diluted cell suspensions of fresh cultures of *L. monocytogenes* ATCC 19115 and *L. ivanovii* ATCC 19119. Viable cells counts were determined by plating of 100 μL of 10^{-6}, 10^{-7}, and 10^{-8} dilutions of bacterial cultures on trypticase soy-yeast extract agar plates (Guangdong Huankai Co., Ltd., Guangzhou, China) and incubating overnight at 37 °C. PCR was performed using 2 μL of genomic DNA extracted from 1 mL different dilutions of cell mixtures as a template, with genus or species-specific primers. An equal volume of sterile distilled water was used instead of the template as a negative control. PCR thermal cycling involved an initial denaturation step at 95 °C for 10 min; followed by 35 cycles at 95 °C for 30 s, 55 °C for 30 s, and 72 °C for 30 s; and a final extension step at 72 °C for 10 min. PCR products were evaluated by 1.5% agarose electrophoresis.

Multiplex PCR Conditions for Detection of Pathogenic *Listeria*

In the genus *Listeria*, only two species, *L. monocytogenes* and *L. ivanovii*, are human and animal pathogens. Therefore, a mPCR assay was devised by combining three specific primer sets targeting genes specific for the *Listeria* genus, *L. monocytogenes*, and *L. ivanovii* (**Table 2**). Multiplex PCR was performed in a volume of 25 μL; the reaction mixture contained 12.5 μL DreamTaq Green PCR Master Mix (2×), 0.1 mM dNTPs, 6 mM $MgCl_2$, 1.5 U TaKaRa Ex Taq, 1.6 μM for *Listeria* genus-specific primers targeting *LMOSLCC2755_0944*, 0.32 μM for *L. monocytogenes*-specific primers targeting *LMOSLCC2755_0090* and 0.32 μM for *L. ivanovii*-specific primers targeting *queT_1*, and 4 μL of template. The genomic DNA of *L. monocytogenes* and *L. ivanovii* was added as a template for the positive control, and an equal volume of sterile distilled water was used instead of the template as a negative control. Multiplex PCR thermal cycling involved an initial denaturation step at 95 °C for 10 min; followed by 35 cycles at 95 °C for 30 s, 58.1°C for 30 s, and 72 °C for 30 s; and a final extension step at 72 °C for 10 min. The mPCR products were analyzed after electrophoresis on 2.5% agarose gel under ultraviolet light illumination.

Evaluation of Specificity and Detection Limit of mPCR Assay

The specificity of mPCR with novel specific primers was verified using 33 bacteria strains (including 17 *Listeri*a strains and 16 non-*Listeria* strains) (**Supplementary Table S2**). Primer sets that amplified bands of the expected length from the corresponding strains of *Listeria* but not from non-*Listeria* strains were considered species-specific. For the detection limit evaluation, *Listeria* cultured overnight were diluted 10-fold in normal saline and mixed. Purified genomic DNA extracted from different dilutions of cell suspensions was used as the mPCR template (4 μL). An equal volume of sterile distilled water was included in the mPCR mixture instead of the template, as a negative control. The sensitivity of the PCR was determined by electrophoresis, as described in section "Multiplex PCR Conditions for Detection of Pathogenic *Listeria*."

Artificial Contamination Experiments

Artificial contamination experiments were performed as previously described (Chen et al., 2019). Briefly, *L. monocytogenes* ATCC 19115 and *L. ivanovii* ATCC 19119 were cultured overnight, and the cell concentration was determined by plating. *F. velutipes* samples (10 g) were mixed by 88 mL of the BHI medium, and then 2 mL *Listeria* mixtures were added with the final inoculum concentration ranging from 7.6×10^0 to 7.6×10^4 CFU/10 g. Next, the mixtures were incubated at 37 °C for 4 h, 6 h, 8 h, 10 h, or 12 h. Genomic DNA was extracted at the indicated time points from 1 mL of sample, and then analyzed by mPCR.

Interference Evaluation

To validate the accuracy and scope for interference in the mPCR assay, *L. monocytogenes* ATCC 19115, *L. ivanovii* ATCC19119, and three common pathogens (*Salmonella* Enteritidis CMCC 50335, *Staphylococcus aureus* ATCC 25923, and *Escherichia coli* ATCC 25922) were used. The strains were cultured in the BHI broth overnight and serially diluted (10-fold) with normal saline. The density of *Listeria* cells was adjusted to 10^6 CFU/mL. *Listeria* cultures were individually mixed with the interference testing strain in ratios of $1{:}10^2$, 1:10, 1:1, 10:1, and $10^2{:}1$. Genomic DNA was extracted from the mixtures. Meanwhile, the genomic DNA from *Listeria* cultures without interference strains was used as a template for the positive control. The scope of mPCR assay to overcome interference was evaluated by 2.5% agarose gel electrophoresis.

Detection of Pathogenic *Listeria* in *F. velutipes* Plants

To validate the detection ability of mPCR, a total of 129 samples were collected from *F. velutipes* plants in Guangdong Province (China). The sampled sites included composting sites (compost and sterile compost, n = 15), mycelium culture rooms (cultures and shelf surfaces, n = 18), mycelium stimulation rooms (mycelium stimulation machinery and floor, n = 18), fruiting body cultivation rooms (drains, shelf surfaces, and *F. velutipes*, n = 27), and harvesting rooms (packaging machinery surfaces, scales, conveyor belts, cutler surfaces, packaged *F. velutipes*, floor, drains, and workers, n = 51). The sites at the *F. velutipes* plants were randomly sampled, and the samples were transported on ice to the laboratory for immediate analysis. The conventional culture method was used to test for the presence of *Listeria* as previously described (Chen et al., 2019). Briefly, 25 g of each sample was randomly weighed-out, added to 225 mL of *Listeria* enrichment broth 1 (LB1, Guangdong Huankai

Co., Ltd., Guangzhou, China), and incubated at 30 °C for 24 h. After incubation, 100 μL of LB1 enrichment culture was transferred to 10 mL of *Listeria* enrichment broth 2 (LB2, Guangdong Huankai Co., Ltd., Guangzhou, China), and incubated at 30 °C for 24 h. A loopful (approximately 10 μL) of the LB1 enrichment culture was streaked onto *Listeria* selective agar plates (Guangdong Huankai Co., Ltd., Guangzhou, China), and incubated at 37 °C for 48 h. At least three presumptive colonies were selected for the identification of *Listeria* using the Microgen ID *Listeria* identification system (Microgen, Camberley, United Kingdom), according to the manufacturer's instructions. Meanwhile, 1 mL of LB1 enrichment culture was collected from each sample at 12 h. Genomic DNA was extracted from LB1 enrichment cultures for mPCR.

Data Analysis

The Kappa coefficient (κ) was used to evaluate the performance of the developed mPCR assay. Since the true pathogenic *Listeria* status in naturally contamination samples was unknown, this calculation method assumed that conventional culture analysis was the true value. The calculation method was performed as previously described (Tao et al., 2017).

RESULTS

Novel Target Genes Specific for the *Listeria* Genus and Species

In the current study, 205 genomic sequences (**Supplementary Table S1**) were included in pan-genome analysis to identify novel molecular targets for the detection and differentiation of *L. monocytogenes* and *L. ivanovii*, and nonpathogenic *Listeria* species. After filtering using pan-genome and PCR analysis, five novel *Listeria*-specific targets, including *LMOSLCC2755_0944* (165/165) specific for *Listeria* genus, *LMOSLCC2755_0090* (128/128) specific for *L. monocytogenes*, *NCTC12701_01099* (9/9), *queT_1* (9/9), and *gmuC_3* (9/9) specific for *L. ivanovii*, were uniquely present in all target strains, but not in non-target strains (**Tables 1**, **3** and **Supplementary Figure S1**). The *Listeria*-specific targets reported in the previous studies, including *prs*, *actA*, *hly*, *inlA*, and *iactA*, were used to specifically target strains and detected these targets in the current study, with 98.8% (163/165), 97.7% (125/128), 99.2% (127/128), 71.9% (92/128), and 55.6% (5/9) of target strains harboring the marker, respectively (**Table 3**).

As shown in **Table 2**, for the *Listeria* genus, only one gene (*LMOSLCC2755_0944*, encoding a non-heme iron-containing

TABLE 2 | Specific genes and primers for PCR identification of *Listeria* spp. and pathogenic *Listeria*.

Species	Name of target genes	Encoded protein	Primer set name	Sequences (5′–3′)	Product size (bp)
Listeria spp.	*LMOSLCC2755_0944*	Non-heme iron-containing ferritin	SPP1	CAGTAGACACAAAGGAAT GCTTTGAACATCCAGATA	427
L. monocytogenes	*LMOSLCC2755_0090*	Transcriptional regulator belonging to the MerR family	LM1	GCTTAATAACCCCTGACCG AATCCCAATCTTCCTAACCAC	260
L. ivanovii	*NCTC12701_01099*	Hypothetical protein	LIV1	GCTAAAAACGACATAGAGG CTCTTCTAACACAATCACT	264
	queT_1	Queuosine precursor ECF transporter S component QueT	LIV2	AGCCATCCAACTACGAAT GACCCACCGCCTATAAAA	144
	gmuC_3	PTS system oligo-beta-mannoside-specific EIIC component	LIV3	TATTTGGACCACCGCATTA TAGCAGGAACATCTTCACTT	452

TABLE 3 | Presence profile of novel *Listeria*-specific targets for target and non-target strains.

Species	Related gene	Presence profile		Source
		In target	In non-target	
Listeria spp.	*LMOSLCC2755_0944*	165 (100%)	0 (0)	Our study
	prs	163 (98.8%)	0 (0)	Doumith et al., 2004
L. monocytogenes	*LMOSLCC2755_0090*	128 (100%)	0 (0)	Our study
	actA	125 (97.7%)	0 (0)	Gouin et al., 1995; Zhou and Jiao, 2005
	hly	127 (99.2%)	0 (0)	Paziak-Domańska et al., 1999
	inlA	92 (71.9%)	0 (0)	Jung et al., 2003
L. ivanovii	*NCTC12701_01099*	9 (100%)	0 (0)	Our study
	queT_1	9 (100%)	0 (0)	Our study
	gmuC_3	9 (100%)	0 (0)	Our study
	iactA	5 (55.6%)	0 (0)	Gouin et al., 1995

ferritin) was identified as genus-specific. The four other genes were species-specific: one gene (*LMOSLCC2755_0090*, encoding a transcriptional regulator from the MerR family) was specific for *L. monocytogenes*, and three genes (*NCTC12701_01099*, encoding a hypothetical protein; *queT_1*, encoding a queuosine precursor ECF transporter S component QueT; and *gmuC_3*, encoding a PTS system oligo-beta-mannoside-specific EIIC component) were specific for *L. ivanovii*.

Selection of Novel Species-Specific Genes for mPCR

Before devising a mPCR assay, the sensitivity of primers for each target gene was evaluated in an individual PCR assay. The results are shown in **Supplementary Table S3** and **Figure 1**. The detection limit of these primer pairs was 10^3–10^4 CFU/mL. Considering the amplicon lengths for different genes, the sensitivity of species-specific primers, and the competition and inhibition of primer sets in the reaction system, a combination of suitable primer pairs (SPP1, LM1, and LIV2) targeting the *LMOSLCC2755_0944*, *LMOSLCC2755_0090*, and *queT_1* genes (respectively) was selected for the mPCR assay.

Evaluation of the Specificity and Sensitivity of the mPCR Assay for *Listeria* spp.

The results of the specificity of mPCR assay with novel specific primers are shown in **Supplementary Table S2**. The genus-specific target band of 427 bp was only obtained with *Listeria* spp. as the template. Further, species-specific bands were also obtained for the *L. monocytogenes* and *L. ivanovii* strains tested. No product bands were obtained with the 16 non-*Listeria* strains tested, and no cross-reactivity of the mPCR was observed.

To determine the detection limit of the novel assay, DNA was extracted from different dilutions of *Listeria* spp. cultures and used as the reaction template. Following mPCR detection, three specific fragments of 427 bp (SPP1), 260 bp (LM1), and 144 bp (LIV2) were obtained for cell concentrations of 10^7 to 10^3 CFU/mL (**Figure 2**). The analysis suggested that the

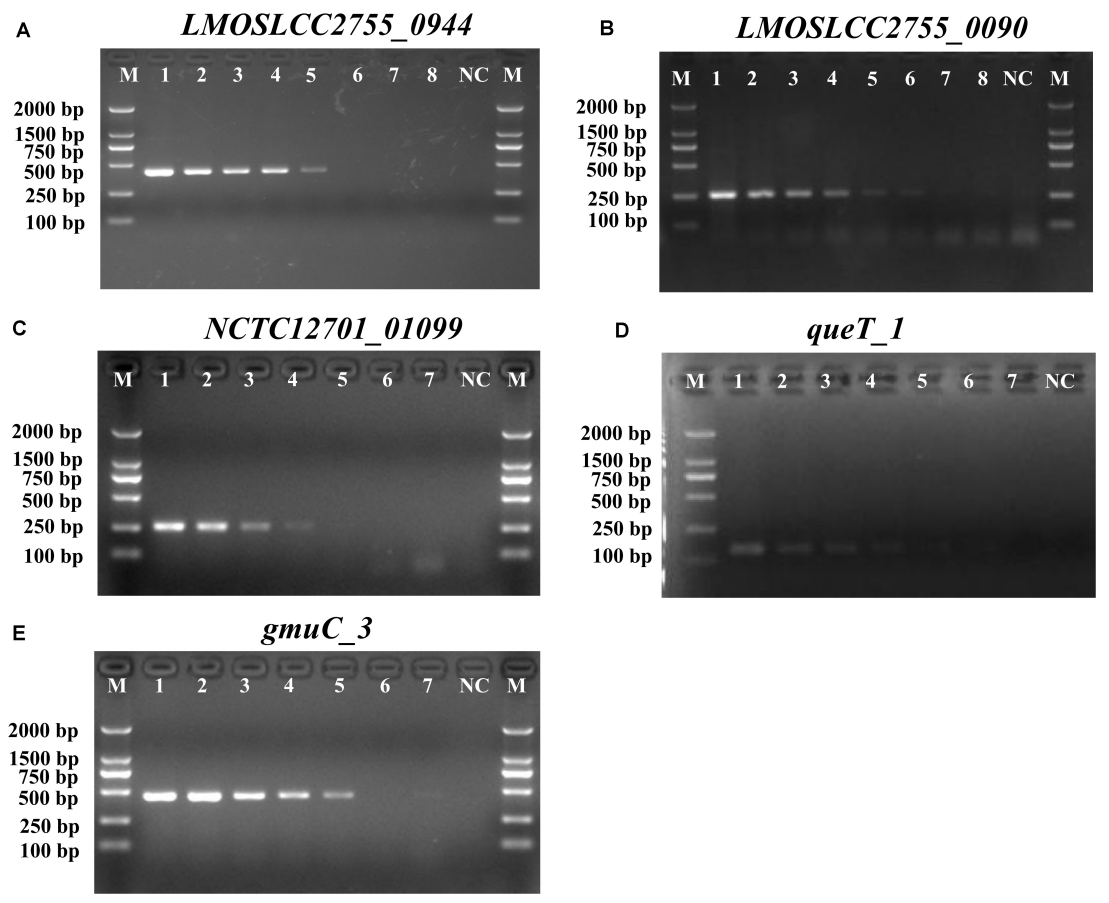


FIGURE 1 | PCR detection sensitivity of novel targets specific for *Listeria*. **(A)** *LMOSLCC2755_0944* specific for the *Listeria* spp.; **(B)** *LMOSLCC2755_0090* specific for *L. monocytogenes*; **(C)** *NCTC12701_01099*; **(D)** *queT_1*; and **(E)** *gmuC_3* specific for *L. ivanovii*. Lane M: DL2000 DNA standard marker; **(A)** lanes 1–8: the bacterial culture concentration per PCR assay from 10^7 to 10^0 CFU/mL; **(B)** lanes 1–8: the bacterial culture concentration per PCR assay from 10^8 to 10^1 CFU/mL; **(C–E)** lanes 1–7: the bacterial culture concentration per PCR assay from 10^7 to 10^1 CFU/mL; lane NC: negative control.

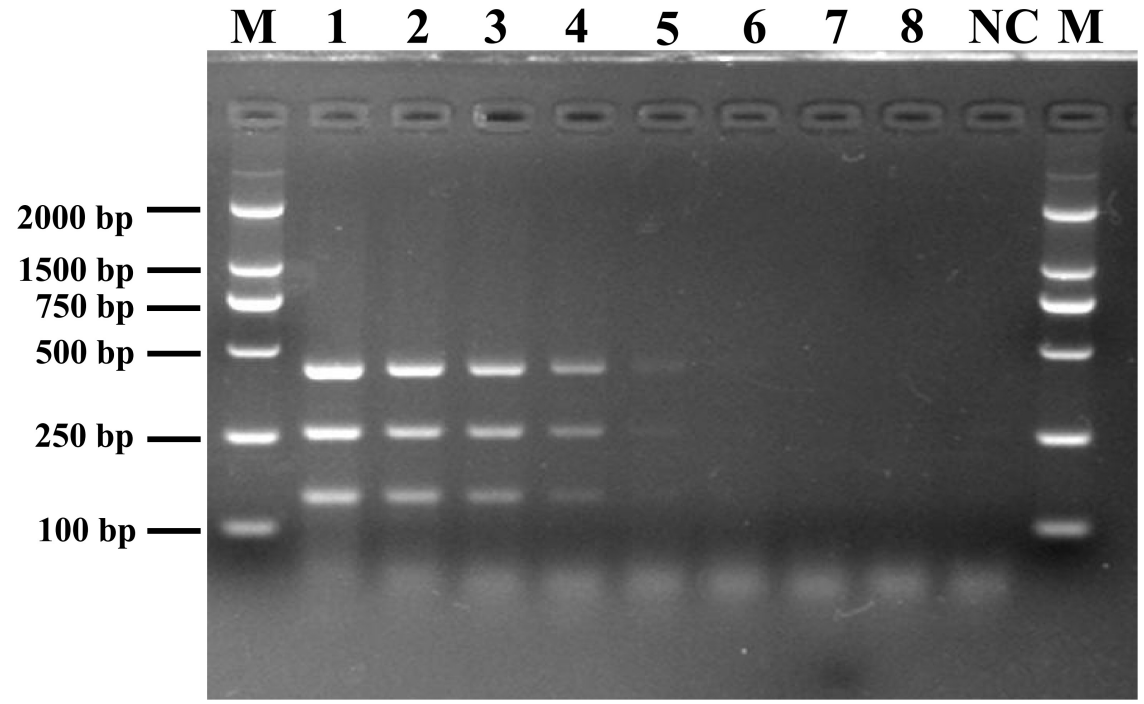


FIGURE 2 | Detection limit of mPCR for pure culture from pathogenic *Listeria*. Lane M: DL2000 DNA standard marker; lanes 1–8: the bacterial culture concentration per PCR assay from 5.4×10^7 to 5.4×10^0 CFU/mL; lane NC: negative control.

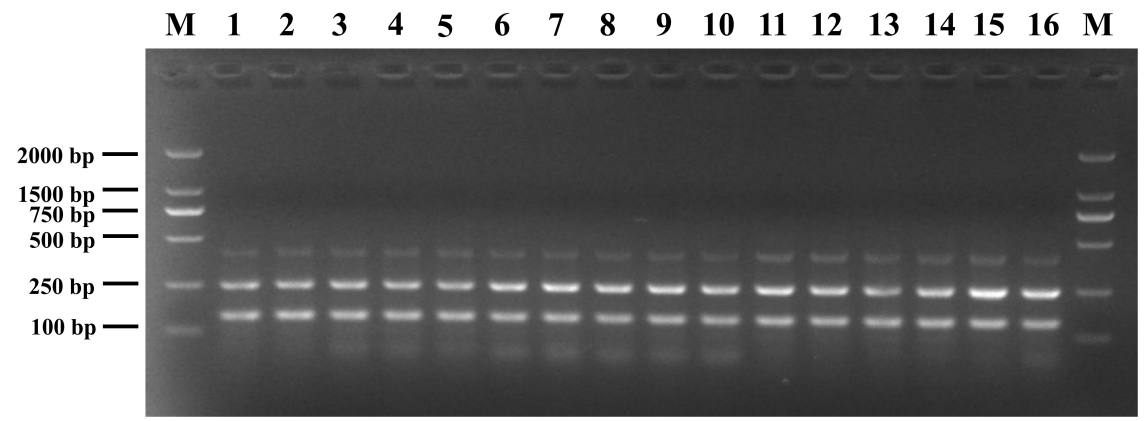


FIGURE 3 | Anti-interference tests of mPCR detection of pathogenic *Listeria* with *S.* Enteritidis CMCC 50335 (lanes 1–5), *S. aureus* ATCC 25923 (lanes 6–10), and *E. coli* ATCC 25922 (lanes 11–15) at different mixing proportion. Lane M: DL2000 DNA standard marker; lanes 1–5, 6–10, 11–15: the ratio of target bacteria to interfering bacteria at $1{:}10^2$, 1:10, 1:1, 10:1, and $10^2{:}1$; lane 16: positive control.

detection limits of the mPCR assay were 2.0×10^3 CFU/mL for *L. monocytogenes* and 3.4×10^3 CFU/mL for *L. ivanovii*.

Multiplex PCR Detection of *Listeria* in Artificially Contaminated *F. velutipes* Samples

F. velutipes samples were spiked with *Listeria* cells at the inoculum levels of 7.6×10^4, 7.6×10^3, 7.6×10^2, 7.6×10^1, and 7.6×10^0 CFU/10 g, and the enrichment cultures were followed for up to 12 h. As shown in **Supplementary Table S4**, the detection outcomes of mPCR varied depending on the enrichment time. The detection limit of mPCR after 4-h, 6-h, 8-h, 10-h, and 12-h enrichment was 7.6×10^4, 7.6×10^3, 7.6×10^2, 7.6×10^1, and 7.6×10^0 CFU/10 g *F. velutipes*, respectively (**Supplementary Table S4**).

Interference Testing of the mPCR Assay

The susceptibility of the mPCR assay to interference by non-target DNA was determined by mixing pathogenic *Listeria* and

non-*Listeria* strains in different ratios (**Figure 3**). Three clear bands were generated for mixtures of all strains tested, and the brightness of these bands was comparable with that obtained from analyzing pure pathogenic *Listeria* cultures. This indicated that the presence of *S.* Enteritidis CMCC 50335, *S. aureus* ATCC 25923, and *E. coli* ATCC 25922 did not interfere with the detection of pathogenic *Listeria*.

Application of the mPCR Assay for the Analysis of *F. velutipes* Plants Samples

To verify the practicality and effectiveness of the developed mPCR method, we next used the assay to detect *Listeria* in 129 non-spiked *F. velutipes* plants samples (**Supplementary Table S5**). In the analyzed samples, the prevalence of *Listeria* spp. and *L. monocytogenes* was 58.1% and 41.1%, respectively. The contaminated sites included the composting area, mycelium scraping equipment surfaces, floor, fruiting body cultivation room, and harvesting room. No pathogenic *Listeria* spp. were detected at the composting phase and mycelium culture room samples. The bacteria appeared to originate from the mycelium scraping machinery, which contaminated the product and downstream packaging equipment.

The results of positive identification of *Listeria* spp. by traditional methods in 75 samples also analyzed by mPCR are summarized in **Figure 4**. Among 174 strains identified as *Listeria* spp., 79 strains (45.4%) were identified as *L. monocytogenes* and 95 strains (54.6%) were identified as nonpathogenic *Listeria* (including 91 *L. innocua* and 4 *L. grayi*). In 129 *F. velutipes* plant samples, no *L. ivanovii, L. seeligeri*, and *L. welshimeri* were detected. A comparison of the mPCR outcomes with the results of conventional culture method is shown in **Table 4**. For detection of *Listeria* spp., the sensitivity, specificity, and efficacy of mPCR were 98.7%, 98.1%, and 98.5%, respectively. For detection of *L. monocytogenes*, the sensitivity, specificity, and efficacy of mPCR were 98.2%, 98.7%, and 98.5%, respectively. Only one false-positive and one false-negative result were obtained for 129 samples tested. The calculated κ factors for *Listeria* spp., *L. monocytogenes*, and *L. ivanovii* were 0.97, 0.97, and 1, respectively, indicating that the mPCR outcomes were highly consistent with those of the conventional culture method.

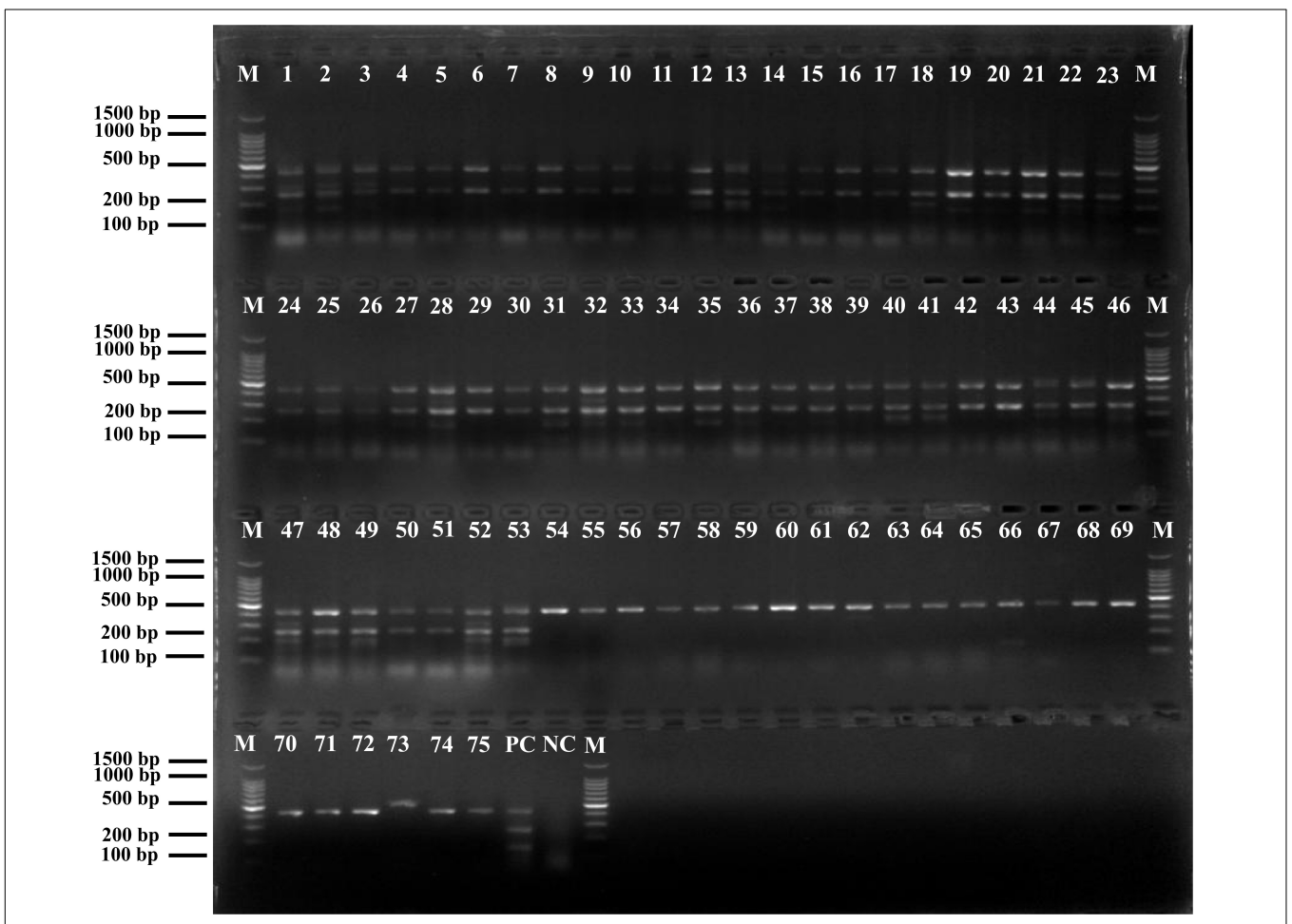


FIGURE 4 | Positive results of mPCR in natural *Flammulina velutipes* plants samples. Lane M: DL2000 DNA standard marker; lanes 1–53: positive samples containing *L. monocytogenes*; lanes 54–75: positive samples containing non-pathogenic *Listeria* spp.; lane PC: positive control; lane NC: negative control.

TABLE 4 | Statistical evaluation of mPCR assay for natural *Flammulina velutipes* plants samples comparison with conventional culture method.

Species	Samples number	Multiplex PCR results	Conventional culture method results		Sensitivity (%)	Specificity (%)	Efficiency (%)	Predictive value (%)		κ
			+	−				+	−	
Listeria spp.	129	+	75 (a)	1 (b)	98.7	98.1	98.5	98.7	98.1	0.97
		−	1 (c)	52 (d)						
L. monocytogenes		+	53 (a)	1 (b)	98.2	98.7	98.5	98.2	98.7	0.97
		−	1 (c)	74 (d)						
L. ivanovii		+	0 (a)	0 (b)	100	100	100	0	100	1
		−	0 (c)	129 (d)						

[a] *True positive;*
[b] *false positive;*
[c] *false negative;*
[d] *true negative.*

DISCUSSION

In the current study, we devised a novel mPCR method for detection of pathogenic *Listeria* spp. in food. The method focuses on new genus- and species-specific gene targets, and is highly specific and sensitive. It may inform the development of highly sensitive PCR approaches for the detection of *Listeria* in food samples and food processing plants.

Listeria spp. is widely found in food, especially in edible mushrooms, with the detection rate as high as 21.2% (Chen et al., 2018). Therefore, the development of PCR-based detection methods for species-specific classification would provide an independent means for confirming species identity (Ryu et al., 2013; Liu et al., 2015). The current PCR detection methods for *Listeria* species target virulence genes, or 16S rRNA and 23S rRNA genes. However, the lack of genes or mutations in *Listeria* strains can result in no identification or misidentification, posing a potential threat of food poisoning (Nishibori et al., 1995). As numerous complete microbial genome sequences become available with the development of sequencing technologies and bioinformatics, many researchers are committed to the search for novel specific gene targets, to replace current target genes that exhibit poor specificity (Tao et al., 2017; Sheng et al., 2018). Previously, specific target genes for *Listeria* were identified by DNA hybridization or using the BLAST program (Michel et al., 2004; Tao et al., 2017). Identification of a specific target by using DNA arrays involves the synthesis of probes, PCR amplification, array construction, and hybridization (Michel et al., 2004). The BLAST search method of specific target screening usually involves nucleotide sequence similarity comparisons with one or only few representative genomic sequences (Tao et al., 2017). With the rapid development of bioinformatics in recent years, using the pan-genome analysis to identify specific targets is more economical, convenient, and effective than other molecular target screening methods. In the current study, we identified gene targets specific for the *Listeria* genus and species via pan-genome analysis using a large number of genome sequences (n = 205). Through pan-genome and PCR analysis, five novel *Listeria*-specific targets were 100% specific for the targeted *Listeria* genomes and did not detect non-target *Listeria* and non-*Listeria* genomes. However, The *Listeria*-specific targets reported in the previous studies, including *prs*, *actA*, *hly*, *inlA*, and *iactA,* were present in 98.8%, 97.7%, 99.2%, 71.9%, and 55.6% of the target strains, respectively (**Table 3**). In addition, the detection limits of the corresponding primer pairs (10^3–10^4 CFU/mL) of these novel target genes were comparable with those of existing molecular detection targets (Chiang et al., 2012; Tao et al., 2017). Therefore, these *Listeria*-specific targets obtained by this approach have better specificity, and their sensitivity can meet the needs of existing molecular detection methods, so these novel targets can represent unique detection targets for monitoring of pathogenic *Listeria* in food and its production environment.

The presence of pathogenic *Listeria* strains in *F. velutipes* is an emerging public health hazard (Chen et al., 2018). Therefore, rapid detection of pathogenic *Listeria* is crucial for implementing optimal sterilization procedures for decontamination during the preparation of *F. velutipes* products. The mPCR assay developed in the current study combines three specific primer sets (SPP1, LM1, and LIV2) based on novel molecular markers (*LMOSLCC2755_0944*, *LMOSLCC2755_0090*, and *queT_1*, respectively) and allows simultaneous identification of pathogenic *Listeria* species. To increase the accuracy of *Listeria* identification, the *Listeria* genus-specific SPP1 primer set was also included. This enhances the specificity of the mPCR, especially when high concentrations of non-*Listeria* bacteria are present in a sample. The minimum detection limits of the assay were 2.0×10^3 CFU/mL for *L. monocytogenes* and 3.4×10^3 CFU/mL for *L. ivanovii* when pure enriched cultures were analyzed, which is comparable to those of mPCRs reported in previous studies (Chiang et al., 2012; Thapa et al., 2013; Sheng et al., 2018). In addition, the developed mPCR method detected pathogenic *Listeria* in artificially contaminated *F. velutipes* samples in the range of 7.6×10^4–7.6×10^0 CFU/10 g after 4–12 h of enrichment culture. These observations indicated that the devised mPCR assay could be used to detect pathogenic *Listeria* in samples more rapidly (The overall assay time, including 4–12 h pre-enrichment, DNA extraction and PCR assay, only took 5–17 h) than by using the standard culture method (4–7 days).

We also used the novel mPCR assay to analyze 129 samples from *F. velutipes* plants. The assay was highly sensitive, specific,

and efficient. High κ values of the assay (0.97 for *Listeria* spp., 0.97 for *L. monocytogenes*, and 1 for *L. ivanovii*) indicated good correspondence between the assay and a conventional culture method. In an earlier preliminary study, we found that the prevalence of *Listeria* spp. and *L. monocytogenes* in *F. velutipes* production plants is 52.5% and 18.6%, respectively (Chen et al., 2014). In the current study, we noted a relatively high incidence of *Listeria* contamination (58.1% for *Listeria* spp. and 41.1% for *L. monocytogenes*). A possible reason for such high incidence is that many samples were collected on the floor and drains in this study, and the cross-contamination of *L. monocytogenes* easily occurred in these collection sites during the production process of *F. velutipes*. Among 174 strains identified as *Listeria* spp., approximately 54.6% strains were nonpathogenic *Listeria* species (including 91 *L. innocua* and 4 *L. grayi*). *L. innocua* occurred more frequently than *L. monocytogenes* in the plant and related samples, which was consistent with previous findings (Chen et al., 2014). The presence of *L. innocua* may hamper the detectability of *L. monocytogenes*, by inhibiting the pathogen via excessive growth and production of inhibitory compounds (Cornu et al., 2002; Zitz et al., 2011). Therefore, the presence of nonpathogenic *Listeria*, especially *L. innocua*, may accompany *L. monocytogenes* contamination (Pagadala et al., 2012). Consequently, a positive band for the *Listeria* genus-specific target gene in the novel mPCR assay may be an indicator of the presence of *L. monocytogenes.* Finally, isolation of pathogenic *Listeria* strains from *F. velutipes* plants indicated a potential risk of contracting listeriosis when consuming this mushroom. In most cases, *L. monocytogenes* is inactivated by cooking and heating. Therefore, edible mushrooms should be fully heated before use and cross-contamination between foods should be avoided.

CONCLUSION

In conclusion, we used pan-genome analysis to identify five novel genus- and species-specific targets for *Listeria* detection. We then devised a mPCR assay based on these new targets for simultaneous detection of *L. monocytogenes*, *L. ivanovii*, and other *Listeria* spp. with high specificity and sensitivity. The results of the assay were consistent with the results of traditional culture methods. Hence, the developed assay can be applied for rapid screening and detection of pathogenic *Listeria* in foods and food processing environments, providing accurate results to inform effective monitoring measures to improve the microbiological safety of foods.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

QW, JZ, QY, and MC conceived and designed the experiments. FL, LX, and YZ performed the experiments. FL, JW, SW, and HZ analyzed the data. YD, QG, and XW contributed reagents, materials, and analysis tools. FL contributed to the writing of the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was funded by the National Key Research and Development Program of China (2018YFD0400900), the Key Research and Development Program of Guangdong Province, China (2018B020205001), and GDAS' Project of Science and Technology Development (2019GDASYL-0201001).

ACKNOWLEDGMENTS

The authors would like to thank to the College of Food Science and Technology of Nanjing Agricultural University, the Chinese Center for Disease Control and Prevention, and the State Key Laboratory of Food Science and Technology of Nanchang University for donating strains, which made this study carried out successfully.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2020.634255/full#supplementary-material

REFERENCES

Chen, M., Cheng, J., Pang, R., Zhang, J., Chen, Y., Zeng, H., et al. (2019). Rapid detection of *Listeria* monocytogenes sequence type 121 strains using a novel multiplex Pcr assay. textitLwt Food Sci. Technol. 116:108474. doi: 10.1016/j.lwt.2019.108474

Chen, M., Cheng, J., Wu, Q., Zhang, J., Chen, Y., Zeng, H.., et al. (2018). Prevalence, potential virulence, and genetic diversity of *Listeria monocytogenes* isolates from edible mushrooms in Chinese markets. *Front. Microbiol.* 9:1711. doi: 10.3389/fmicb.2018.01711

Chen, M., Wu, Q., Zhang, J., Guo, W., Wu, S., and Yang, X. (2014). Prevalence and Contamination Patterns of *Listeria monocytogenes* in *Flammulina velutipes* Plants. *Foodborne Pathog. Dis.* 11, 620–627. doi: 10.1089/fpd.2013.1727

Chiang, Y. C., Tsen, H. Y., Chen, H. Y., Chang, Y. H., Lin, C. K., Chen, C. Y.., et al. (2012). Multiplex PCR and a chromogenic DNA macroarray for the detection of *Listeria monocytogens*, *Staphylococcus aureus*, *Streptococcus agalactiae*, *Enterobacter sakazakii*, *Escherichia coli* O157:H7, *Vibrio parahaemolyticus*, *Salmonella* spp. and *Pseudomonas fluorescens* in milk and meat samples. *J. Microbiol. Methods* 88, 110–116. doi: 10.1016/j.mimet.2011.10.021

Cornu, M., Kalmokoff, M., and Flandrois, J.-P. (2002). Modelling the competitive growth of *Listeria monocytogenes* and *Listeria innocua* in enrichment broths. *Int. J. Food Microbiol.* 73, 261–274. doi: 10.1016/S0168-1605(01)00658-4

Czajka, J., Bsat, N., Piani, M., Russ, W., Sultana, K., Wiedmann, M.., et al. (1993). Differentiation of *Listeria monocytogenes* and *Listeria innocua* by 16S rRNA genes and intraspecies discrimination of *Listeria monocytogenes* strains by random amplified polymorphic DNA polymorphisms. *Appl. Environ. Microb.* 59, 304–308. doi: 10.1128/AEM.59.1.304-308.1993

Doijad, S., Barbuddhe, S., Garg, S., Kalekar, S., Rodrigues, J., D'Costa, D.., et al. (2011). Incidence and genetic variability of *Listeria* species from three milk processing plants. *Food Control* 22, 1900–1904. doi: 10.1016/j.foodcont.2011.05.001

Doumith, M., Buchrieser, C. P., Jacquet, C., and Martin, P. (2004). Differentiation of the major *Listeria monocytogenes* serovars by multiplex PCR. *J. Clin. Microbiol.* 42:3819. doi: 10.1128/JCM.42.8.3819-3822.2004

Gouin, E., Dehoux, P., Mengaud, J., Kocks, C., and Cossart, P. (1995). *iactA* of *Listeria ivanovii*, although distantly related to *Listeria monocytogenes actA*, restores actin tail formation in an *L. monocytogenes actA* mutant. *Infect. Immun.* 63, 2729–2737. doi: 10.1016/0167-5699(95)80157-X

Hellberg, R. S., Martin, K. G., Keys, A. L., Haney, C. J., Shen, Y., and Smiley, R. D. (2013). 16S rRNA partial gene sequencing for the differentiation and molecular subtyping of *Listeria* species. *Food Microbiol.* 36, 231–240. doi: 10.1016/j.fm.2013.06.001

Jagadeesan, B., Schmid, V. B., Kupski, B., McMahon, W., and Klijn, A. (2019). Detection of *Listeria* spp. and *L. monocytogenes* in pooled test portion samples of processed dairy products. *Int. J. Food Microbiol.* 289, 30–39. doi: 10.1016/j.ijfoodmicro.2018.08.017

Jung, Y. S., Frank, J. F., Brackett, R. E., and Chen, J. (2003). Polymerase chain reaction detection of *Listeria monocytogenes* on frankfurters using oligonucleotide primers targeting the genes encoding internalin AB. *J. Food Protect.* 66, 237–241. doi: 10.1016/S0260-8774(02)00271-6

Kim, Y., Gu, C., Kim, H. U., and Lee, S. Y. (2020). Current status of pan-genome analysis for pathogenic bacteria. *Curr. Opin. Biotech.* 63, 54–62. doi: 10.1016/j.copbio.2019.12.001

Kuenne, C., Billion, A., Mraheil, M. A., Strittmatter, A., Daniel, R., Goesmann, A.., et al. (2013). Reassessment of the *Listeria monocytogenes* pan-genome reveals dynamic integration hotspots and mobile genetic elements as major components of the accessory genome. *BMC Genom.* 14, 1–19. doi: 10.1186/1471-2164-14-47

Lee, Y., Yoon, Y., Seo, Y., Kim, S., Ha, J., Lee, J.., et al. (2020). Combined enrichment and quantitative polymerase chain reaction to improve sensitivity and reduce time of detection of *Listeria monocytogenes* in mushrooms. *Foodborne Pathog. Dis.* 17, 276–283. doi: 10.1089/fpd.2019.2688

Li, F., Li, B., Dang, H., Kang, Q., Yang, L., Wang, Y.., et al. (2017). Viable pathogens detection in fresh vegetables by quadruplex PCR. *LWT Food Sci. Technol.* 81, 306–313. doi: 10.1016/j.lwt.2017.03.064

Liu, H., Lu, L., Pan, Y., Sun, X., Hwang, C., Zhao, Y.., et al. (2015). Rapid detection and differentiation of *Listeria monocytogenes* and *Listeria* species in deli meats by a new multiplex PCR method. *Food Control* 52, 78–84. doi: 10.1016/j.foodcont.2014.12.017

Mao, Y., Huang, X., Xiong, S., Xu, H., Aguilar, Z. P., and Xiong, Y. (2016). Large-volume immunomagnetic separation combined with multiplex PCR assay for simultaneous detection of *Listeria monocytogenes* and *Listeria ivanovii* in lettuce. *Food Control* 59, 601–608. doi: 10.1016/j.foodcont.2015.06.048

Michel, Doumith, Christel, Cazalet, Natalie, and Simoes (2004). New aspects regarding evolution and virulence of *Listeria monocytogenes* revealed by comparative genomics and DNA arrays. *Infect. Immun.* 72, 1072–1083. doi: 10.1128/IAI.72.2.1072-1083.2004

Nishibori, T., Cooray, K., Xiong, H., Kawamura, I., Fujita, M., and Mitsuyama, M. (1995). Correlation between the presence of virulence-associated genes as determined by PCR and actual virulence to mice in various strains of *Listeria* spp. *Microbiol. Immun.* 39, 343–349. doi: 10.1111/j.1348-0421.1995.tb02211.x

Norton, D. M., McCAMEY, M., Boor, K. J., and Wiedmann, M. (2000). Application of the BAX for screening/genus *Listeria* polymerase chain reaction system for monitoring *Listeria* species in cold-smoked fish and in the smoked fish processing environment. *J. Food Protect.* 63, 343–346. doi: 10.4315/0362-028x-63.3.343

Pagadala, S., Parveen, S., Rippen, T., Luchansky, J. B., Call, J. E., Tamplin, M. L.., et al. (2012). Prevalence, characterization and sources of *Listeria monocytogenes* in blue crab (Callinectus sapidus) meat and blue crab processing plants. *Food Microbiol.* 31, 263–270. doi: 10.1016/j.fm.2012.03.015

Page, A. J., Cummins, C. A., Martin, H., Wong, V. K., Sandra, R., Holden, M. T. G.., et al. (2015). Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics* 22:22. doi: 10.1093/bioinformatics/btv421

Pang, R., Xie, T., Wu, Q., Li, Y., Lei, T., Zhang, J.., et al. (2019). Comparative genomic analysis reveals the potential risk of *Vibrio parahaemolyticus* isolated from ready-to-eat foods in China. *Front. Microbiol.* 10:186.

Paziak-DomaDomańska, B., Bogusławska, E., Wiêckowska-Szakiel, M., Kotłowski, R., Różalska, B., Chmiela, M.., et al. (1999). Evaluation of the API test, phosphatidylinositol-specific phospholipase C activity and PCR method in identification of *Listeria monocytogenes* in meat foods. *Fems Microbiol. Lett.* 171, 209–214. doi: 10.1111/j.1574-6968.1999.tb13434.x

Radoshevich, L., and Cossart, P. (2018). *Listeria monocytogenes* : towards a complete picture of its physiology and pathogenesis. *Nat. Rev. Microbiol.* 16, 32–46. doi: 10.1038/nrmicro.2017.126

Ranjbar, R., and Halaji, M. (2018). Epidemiology of *Listeria monocytogenes* prevalence in foods, animals and human origin from Iran: a systematic review and meta-analysis. *BMC Publ. Health* 18:1057. doi: 10.1186/s12889-018-5966-8

Rosimin, A. A., Kim, M. J., Joo, I. S., Suh, S. H., and Kim, K. S. (2016). Simultaneous detection of pathogenic *Listeria* including atypical *Listeria innocua* in vegetables by a quadruplex PCR method. *LWT Food Sci. Technol.* 69, 601–607. doi: 10.1016/j.lwt.2016.02.007

Ryu, J., Park, S. H., Yeom, Y. S., Shrivastav, A., Lee, S.-H., Kim, Y.-R.., et al. (2013). Simultaneous detection of *Listeria* species isolated from meat processed foods using multiplex PCR. *Food Control* 32, 659–664. doi: 10.1016/j.foodcont.2013.01.048

Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 14, 2068–2069. doi: 10.1093/bioinformatics/btu153

Sheng, J., Tao, T., Zhu, X., Bie, X., Lv, F., Zhao, H.., et al. (2018). A multiplex PCR detection method for milk based on novel primers specific for *Listeria monocytogenes* 1/2a serotype. *Food Control* 86, 183–190. doi: 10.1016/j.foodcont.2017.11.028

Snapir, Y. M., Vaisbein, E., and Nassar, F. (2006). Low virulence but potentially fatal outcome - *Listeria ivanovii*. *Eur. J. Intern. Med.* 17, 286–287. doi: 10.1016/j.ejim.2005.12.006

Tamburro, M., Ripabelli, G., Vitullo, M., Dallman, T. J., Pontello, M., Amar, C. F. L.., et al. (2015a). Gene expression in *Listeria monocytogenes* exposed to sublethal concentration of benzalkonium chloride. *Comp. Immunol. Microb.* 40, 31–39. doi: 10.1016/j.cimid.2015.03.004

Tamburro, M., Sammarco, M. L., Ammendolia, M. G., Fanelli, I., Minelli, F., and Ripabelli, G. (2015b). Evaluation of transcription levels of *inlA*, *inlB, hly*, *bsh* and *prfA* genes in *Listeria monocytogenes* strains using quantitative reverse-transcription PCR and ability of invasion into human CaCo-2 cells. *Fems Microbiol. Lett.* 362:fnv018. doi: 10.1093/femsle/fnv018

Tamburro, M., Sammarco, M. L., Fanelli, I., and Ripabelli, G. (2019). Characterization of *Listeria monocytogenes* serovar 1/2a, 1/2b, 1/2c and 4b by high resolution melting analysis for epidemiological investigations. *Int. J. Food Microbiol.* 310:108289. doi: 10.1016/j.ijfoodmicro.2019.108289

Tamburro, M., Sammarco, M., and Ripabelli, G. (2018). High resolution melting analysis for the characterization of lineage II *Listeria monocytogenes* serovars 1/2a and 1/2c based on single nucleotide polymorphisms identification within the *Listeria* Pathogenicity Island−1 and inlAB operon: a novel approach for epidemiological surveillance. *J. Appl. Microbiol.* 125, 1920–1937. doi: 10.1111/jam.14100

Tao, T., Chen, Q., Bie, X., Lu, F., and Lu, Z. (2017). Investigation on prevalence of *Listeria* spp. and *Listeria monocytogenes* in animal-derived foods by multiplex PCR assay targeting novel genes. *Food Control* 73, 704–711. doi: 10.1016/j.foodcont.2016.09.026

Thapa, S. P., Han, A. R., Cho, J. M., and Hur, J. H. (2013). Multiplex PCR and DNA array for the detection of *Bacillus cereus*, *Staphylococcus aureus*, *Listeria monocytogenes*, *Escherichia coli* O157:H7, and *Salmonella* spp. targeting virulence-related genes. *Ann. Microbiol.* 63, 725–731. doi: 10.1007/s13213-012-0526-4

Vitullo, M., Grant, K. A., Sammarco, M. L., Tamburro, M., Ripabelli, G., and Amar, C. F. L. (2013). Real-time PCRs assay for serogrouping *Listeria monocytogenes* and differentiation from other *Listeria* spp. *Mol. Cell. Probe.* 27, 68–70. doi: 10.1016/j.mcp.2012.10.001

Zhou, X., and Jiao, X. (2005). Polymerase chain reaction detection of *Listeria monocytogenes* using oligonucleotide primers targeting *actA* gene. *Food Control* 16, 125–130. doi: 10.1016/j.foodcont.2004.01.001

Zitz, U., Zunabovic, M., Domig, K. J., Wilrich, P.-T., and Kneifel, W. (2011). Reduced detectability of *Listeria monocytogenes* in the presence of *Listeria innocua*. *J. Food Protect.* 74, 1282–1287. doi: 10.4315/0362-028X.JFP-11-045

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Li, Ye, Chen, Zhang, Xue, Wang, Wu, Zeng, Gu, Zhang, Wei, Ding and Wu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 28 January 2021
doi: 10.3389/fmicb.2020.625315

Identification of *Vibrio parahaemolyticus* and *Vibrio* spp. Specific Outer Membrane Proteins by Reverse Vaccinology and Surface Proteome

Wenbin Wang[1,2,3], Jianxin Liu[1,2], Shanshan Guo[1,2], Lei Liu[1,2], Qianyun Yuan[1,2], Lei Guo[1,2,3] and Saikun Pan[1,2,3]**

[1] Jiangsu Key Laboratory of Marine Biotechnology, Jiangsu Ocean University, Lianyungang, China, [2] Co-Innovation Center of Jiangsu Marine Bio-industry Technology, Jiangsu Ocean University, Lianyungang, China, [3] Jiangsu Key Laboratory of Marine Bioresources and Environment, Jiangsu Ocean University, Lianyungang, China

OPEN ACCESS

Edited by:
David Rodriguez-Lazaro, University of Burgos, Spain

Reviewed by:
Xuanxian Peng, Sun Yat-sen University, China
Amjad Ali, National University of Sciences and Technology, Pakistan

***Correspondence:**
Wenbin Wang
wenbin66@jou.edu.cn
Saikun Pan
pskgx@jou.edu.cn

Specialty section:
This article was submitted to Food Microbiology, a section of the journal Frontiers in Microbiology

Received: 02 November 2020
Accepted: 18 December 2020
Published: 28 January 2021

Citation:
Wang W, Liu J, Guo S, Liu L, Yuan Q, Guo L and Pan S (2021) Identification of Vibrio parahaemolyticus and Vibrio spp. Specific Outer Membrane Proteins by Reverse Vaccinology and Surface Proteome. Front. Microbiol. 11:625315. doi: 10.3389/fmicb.2020.625315

The discovery of outer membrane proteins (OMPs) with desirable specificity and surface availability is a fundamental challenge to develop accurate immunodiagnostic assay and multivalent vaccine of pathogenic *Vibrio* species in food and aquaculture. Herein 101 OMPs were systemically screened from 4,831 non-redundant proteins of *Vibrio parahaemolyticus* by bioinformatical predication of signaling peptides, transmembrane (TM) α-helix, and subcellular location. The sequence homology analysis with 32 species of *Vibrio* spp. and all the non-*Vibrio* strains revealed that 15 OMPs were conserved in at least 23 *Vibrio* species, including BamA (VP2310), GspD (VP0133), TolC (VP0425), OmpK (VP2362), OmpW (VPA0096), LptD (VP0339), Pal (VP1061), flagellar L-ring protein (VP0782), flagellar protein MotY (VP2111), hypothetical protein (VP1713), fimbrial assembly protein (VP2746), VacJ lipoprotein (VP2214), agglutination protein (VP1634), and lipoprotein (VP1267), Chitobiase (VP0755); high adhesion probability of flgH, LptD, OmpK, and OmpW indicated they were potential multivalent *Vibrio* vaccine candidates. *V. parahaemolyticus* OMPs were found to share high homology with at least one or two *Vibrio* species, 19 OMPs including OmpA like protein (VPA073), CsuD (VPA1504), and MtrC (VP1220) were found relatively specific to *V. parahaemolyticus*. The surface proteomic study by enzymatical shaving the cells showed the capsular polysaccharides most likely limited the protease action, while the glycosidases improved the availability of OMPs to trypsin. The OmpA (VPA1186, VPA0248, VP0764), Omp (VPA0166), OmpU (VP2467), BamA (VP2310), TolC (VP0425), GspD (VP0133), OmpK (VP2362), lpp (VPA1469), Pal (VP1061), agglutination protein (VP1634), and putative iron (III) compound receptor (VPA1435) have better availability on the cell surface.

Keywords: *Vibrio parahaemolyticus*, outer membrane protein, identification, surface proteome, bioinformatics, shaving, reverse vaccinology

HIGHLIGHTS

- OMPs between *Vibrio species* had a close and complex genetic relationship.
- The capsule affected surface proteome analysis of *V. parahaemolyticus.*
- Differential OMPs of *V. parahaemolyticus* and *Vibrio* spp. were revealed.

INTRODUCTION

Vibrio spp. are halophilic bacteria that are widely distributed in seawater, offshore sediments, marine life, and seafood products (Hackbusch et al., 2020). Due to the high prevalence of *Vibrio* species in seafood, *Vibrio*-related food poisoning represents the leading cause of infectious diarrhea in coastal regions (Elmahdi et al., 2016). Acute attacks, abdominal pain, and watery symptoms are the main clinical symptoms (Guin et al., 2019). The most prevalent pathogenic *Vibrio* species in human infections include *Vibrio parahaemolyticus*, *Vibrio vulnificus*, and *Vibrio cholera*, among which *V. vulnificus* is more life threatening and *V. parahaemolyticus* exhibits a higher prevalence in sea food (Bonnin-Jusserand et al., 2019). Furthermore, the vibriosis caused by *Vibrio* species (e.g., *V. parahaemolyticus*, *Vibrio anguillarum*, *Vibrio alginolyticus*, *Vibrio ordalii*, and *Vibrio harveyi*) has been found to cause an infection in more than 50 economic fish species and is considered a major economic threat to the marine aquaculture industry (Yu et al., 2016).

Rapid diagnostic tests are promising to provide better surveillance in the aquaculture and food industries (Baker-Austin et al., 2018). Although there have been several reports on the preparation of diagnostic antibodies against *V. parahaemolyticus* or *Vibrio* species by whole cell antigens, flagella, and hemolysins, the specific diagnostic surface antigens are largely unknown (Xu et al., 2019). Also, vaccines are the preferential way to battle vibriosis in aquaculture for sustainability and food safety (Wang et al., 2020). Polyvalent vaccines that cover the main serotypes of the pathogen are preferred when compared with the monovalent vaccine that contains a single strain of a single antigen (Schlingmann et al., 2018). Furthermore, the multiple vaccine (combination vaccine) combined two or more vaccines have effectively protected groupers against multiple *Vibrio* and viral pathogens (Huang et al., 2012). The traditional strategy of *Vibrio* polyvalent vaccines is based on the whole-cell antigens of representative strains like *V. alginolyticus*, *V. parahaemolyticus* (Aly et al., 2020). The protection against homologous strains with these vaccines was very effective, but the protection of heterogeneous strains or multiple species were elusive. Outer membrane proteins (OMPs) are currently the main candidate antigens used in polyvalent vaccine studies for their essential function, surface exposure, and conservation among different strains (Pore and Chakrabarti, 2013). There have been some studies of the OmpK and OmpU as polyvalent vaccine candidates (Duperthuy et al., 2010; Li et al., 2010b), while many studies focus on the discovery of more multivalent and potent Omps for future vaccine development.

Based on the bacterial serum-based immunoproteomics, Li et al. (2010a) identified 33 OMPs of *V. parahaemolyticus* on the two-dimensional (2-DE) gel, and found that the OmpA and Pal proteins had a multiple protective effect for the tested strains of *Vibrio*, *Aeromonas*, and *Pseudomonas* in carp. They also revealed the polyvalent vaccines from *V. alginolyticus* by a heterogeneous antiserum-based immunoproteomics with bacterial immunization challenging method (Li et al., 2009). Li et al. (2014) in another group used the same strategy and found 10 immunogenic OMPs, including LptD, and OmpK. Comparative analysis of Omps that respond to fish and human plasma stress by differential sub-proteomic methodologies was also used to develop highly protective vaccine candidates of *E. tarda* (Wang et al., 2012). DNA shuffling of six OmpA gene from *V. alginolyticus*, *V. parahaemolyticus*, *Edwardsiella tarda*, and *Escherichia coli* was also reported to develop polyvalent vaccines against *V. alginolyticus* and *E. tarda* infections (Cheng et al., 2018). However, the multivalence of these proteins among the *Vibrio* species remains unclear and some conserved OMPs with a lower abundance may be missed due to competition of the immune system for whole-cell immunogen (Shinoy et al., 2013). Another proteomics strategy is the enzymatic shaving-based surface proteome analysis (Rodríguez-Ortega et al., 2006). It was successfully used to identify the surface-exposed proteins of Gram-positive group A *Streptococcus*, and widely applied to some Gram-negative bacteria (Solis and Cordwell, 2011; Olaya-Abril et al., 2014). No shaving-based surface proteome study of *Vibrio* has been reported.

A genomic and bioinformatic approach, also known as reverse vaccinology, has revolutionized antigen screening and vaccine development by adopting computerized screening of all protein sequences from the target pathogen (Dalsass et al., 2019). Pizza et al. (2000) first reported an analysis of the open reading frames of *Neisseria meningitidis* and intensive recombinant expression of the 570 potential OMPs (Giuliani et al., 2006). To improve the labor-intensive screening process, many bioinformatics tools have been used to analyze large quantities of genomic or protein sequences (Garcia-Angulo et al., 2014). Specially, to reveal broad-spectrum immunogenic targets and analyze the close-related species, the pangenome analysis for the core, accessory, and unique genes was introduced to reverse vaccinology (Zeng et al., 2017; Naz et al., 2019). To our knowledge, the OMPs of *V. anguillarum* have been screened by reverse vaccinology (Baliga et al., 2018). However, little is known about the immunogenic OMPs that are specific to *V. parahaemolyticus* or *Vibrio* spp. and represent promising diagnostic antigens and vaccine candidates.

In this study, the OMPs of *V. parahaemolyticus* were comprehensively screened from the 4,831 protein sequences with intergraded bioinformatics predication methods. The differential OMPs and vaccine candidates of *V. parahaemolyticus* and *Vibrio* spp. were obtained after the overall homology analysis with the other bacteria using the Basic local alignment search tool (BLAST,[1]). The possible vaccine candidates were further selected by the adhesion probability analysis. To experimentally study the OMPs of *V. parahaemolyticus*, the surface proteome was analyzed

[1]https://blast.ncbi.nlm.nih.gov/Blast.cgi

with enzymatical shaving (glycosidases, trypsin) of the intact cells and mass spectrometry identification of the related peptides.

MATERIALS AND METHODS

Screening of Signal Peptide-Positive Proteins From the Encoded Proteins of *V. parahaemolyticus* in the National Center for Biotechnology Information

The flow chart of OMP screening of *V. parahaemolyticus* is shown in **Figure 1**.

The 4,831 protein sequences encoded by nucleotides (accession numbers: BA000031.2 and BA000032.2) of the *V. parahaemolyticus* RIMD 2210633 strain were downloaded as FASTA files from the National Center for Biotechnology Information (NCBI) database. The signal peptide is involved in protein transportation and is located on the cell membrane. All of the 4,831 encoded proteins of *V. parahaemolyticus* were analyzed for the presence of the N-terminal signal peptide using SignalP 5.0 (Armenteros et al., 2019). The following parameters were used, and the maximum number of inputted proteins was 5,000. The organism group was Gram-negative and the output format was a short output.

Predicting the Transmembrane Helices in the Candidate Proteins

The number of highly hydrophobic TM helices in the proteins was negatively associated with the possibility of recombinant expression. The topology information revealed the location of aminol acids on the protein structure and the cell membrane. The transmembrane (TM) helix and topology of the candidate proteins were predicted using the online server TMHMM 2.0 (Erik et al., 1998) and HMMTOP (also known as Community College Transfer Opportunity Program [CCTOP]) (Tusnády and Simon, 2001). Due to the substantial challenge of multi-pass TM protein expression in the prokaryotic expression system, the proteins with less than two TM helixes were selected for further analysis.

Predicting the Subcellular Localization of the Candidate Proteins

The bacterial proteins are primarily located in the cytoplasm, the inner membrane (cytoplasmic membrane), periplasm, outer membrane, and extracellularly. The subcellular locations of the obtained proteins with less than two TM helixes were predicted by PSORTb, CELLO, and UniProt. PSORTb is a web-based tool that provides a link to the PSORT family of programs for predicting the subcellular localization (Yu et al., 2010). CELLO based on Support Vector Machine was used to detect specific amino acid compositions and motifs, thereby predicting the subcellular location of proteins (Yu et al., 2004). The UniProt database provides a comprehensive, high-quality, and freely accessible resource of protein information, including the location and topology of the mature proteins in the cell (The UniProt, 2018). To cross-check the results and improve the coverage and accuracy of prediction, proteins predicted as the OMP by more than one method were selected for further analysis.

Sequence Homology Analysis of the Candidate Proteins With Large Number of Strains

Basic local alignment search tool is an online tool used to identify the local similarity between sequences. This program is supported by the NCBI to compare the nucleotide or protein sequences with sequence databases and calculate the statistical significance of matches. BLAST can also be used to infer the functional and evolutionary relationships between sequences and help identify the protein families. The BLAST program was used to determine the homology with all the identified OMPs among *V. parahaemolyticus* serovars, as well as those present in the other 32 *Vibrio* spp. (*Vibrio cholerae*; *V. vulnificus*; *V. alginolyticus*; *Vibrio diabolicus*; *Vibrio antiquarius*; *Vibrio campbelli*; *V. harveyi*; *Vibrio owensii*; *Vibrio rotiferianus*; *Vibrio jasicida*; *Vibrio natriegens*; *V. alfacsensis*; *V. coralliilyticus*; *V. anguillarum*; *Vibrio splendidus*; *Vibrio mimicus*; *Vibrio cincinnatiensis*; *Vibrio metschnikovii*; *Vibrio gazogenes*; *Vibrio mediterranei*; *Vibrio diazotrophicus*; *Vibrio fluvialis*; *Vibrio furnissii*; *Vibrio proteolyticus*; *Vibrio nereis*; *Vibrio tubiashii*; *Vibrio nigripulchritudo*; *Vibrio aerogenes*; *Vibrio orientalis*; *V. ordalii*; *Vibrio aestuarianus*; and *Vibrio carchariae*), seven other marine bacteria (*Aliivibrio*; *Salinivibrio*; *Enterovibrio*; *Grimontia*; *Photobacterium*; *Plesiomonas*; and *Shewanella*), and six common bacteria (*E. coli*; *Aeromonas*; *Salmonella*; *Pseudomonas*; *Acinetobacter*; and *Yersinia*). The seven marine bacteria were selected because they were found to exhibit a relatively high sequence homology with the candidate proteins in an overall BLAST with non-*Vibrio* spp. strains.

Antigenicity and Adhesion Index Analysis of the Candidate Proteins

Antigenicity refers to the ability of an antigen to be recognized by the immune system. The potential antigenicity of the candidates was estimated through VaxiJen (Doytchinova and Flower, 2007) with a cut-off of 0.4. Adhesion-related OMPs are involved in the pathogenicity of the bacteria and have a greater priority for vaccine development. Vaxign (He et al., 2010) predicts the possible vaccine candidates according to various vaccine design criteria (e.g., sub-cellular location, adhesion, and trans-membrane helix). The predication of Vaxign programs has corresponding sensitivity and specificity values of 0.494 and 0.853 (Jaiswal et al., 2013). Proteins with a Vaxign index higher than 0.5 were predicted to be adhesion-related proteins and selected as vaccine candidates.

Surface Proteome of *V. parahaemolyticus* Analyzed by Trypsin Shaving and Mass Spectrometry

The surface proteome of *V. parahaemolyticus* was studied by "shaving" the cells with trypsin as described (Rodríguez-Ortega, 2018; Hornburg et al., 2019). Briefly, *V. parahaemolyticus* (CICC 21617) was cultured in Luria-Bertani broth with 3.0% NaCl

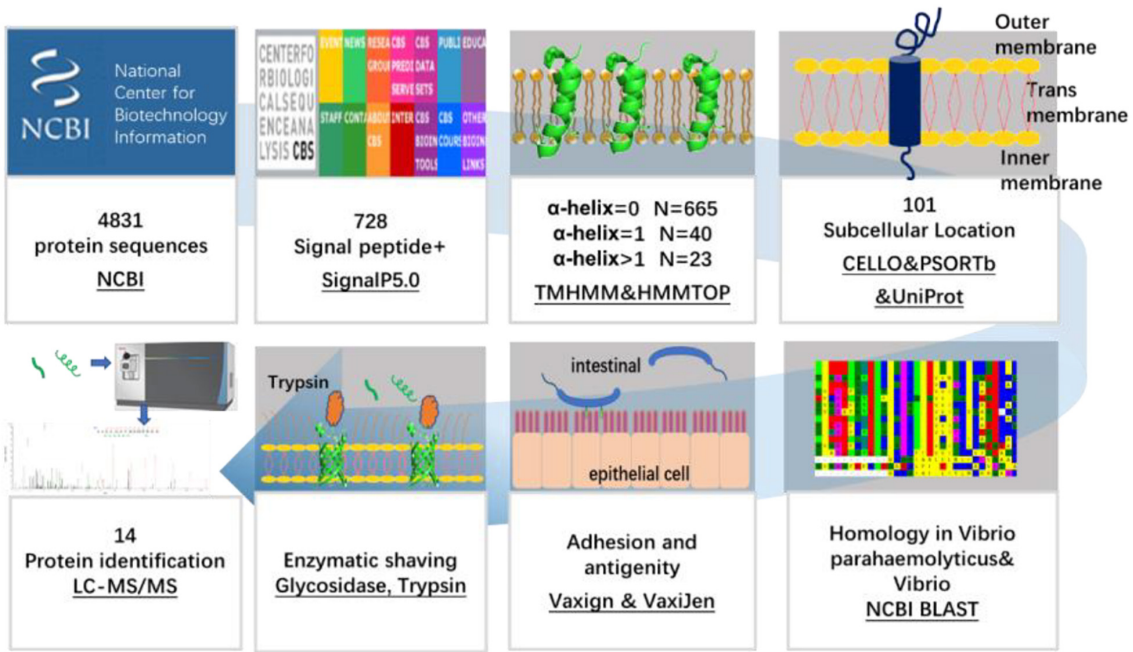


FIGURE 1 | Schematic diagram of a bioinformatical and proteomic approach for the identification of OMPs against *V. parahaemolyticus* and *Vibrio* spp.

at 37°C until the OD_{600} was 0.38, which corresponded to the mid-exponential phase. The culture (50 mL) was centrifuged at 3,500 g, 10 min, 4°C. The pellet was gently suspended in 5 mL 1% paraformaldehyde (PFA) (Aladdin Bio-Chem Technology, Shanghai, China) in PBS for 5 min and washed twice with PBS. One set of the sample was resuspended in 1 mL shaving buffer (1× PBS, 20% [w/v] sucrose, 10 mM DTT) and added with 20 μg sequencing grade trypsin (Sigma-Aldrich, St. Louis, MO, United States) in 0.5 mL 25 mM ammonium bicarbonate buffer. The treated samples were incubated at 37°C in water bath for 15 and 30 min. The control was added with the 0.5 mL 25 mM ammonium bicarbonate buffer and incubated at 37°C in water bath for 30 min. In a further study, the cells were first treated with endoglycosidase α-amylase (Macklin Biochemical Co., Ltd., Shanghai, China), cellulase (solarbio, Beijing, China), pullulanase (Shijiazhuang Ningnuo Trading Co., Ltd., Hebei, China), dextranase (G-CLONE, Biotech Co., Ltd., Beijing, China), and exoglycosidase starch glucosidase (Macklin Biochemical Co., Ltd., Shanghai, China) to reduce the capsule and lipopolysaccharide on the cell surface before "shaving" with trypsin. The cell pellets of 50 mL culture were resuspended in 5 mL phosphate buffer (pH 5.5) with final concentration of 100 U/mL of α-amylase (α-1, 4-linkage), cellulase (β-1, 4), pullulanase (α-1, 6), dextranase (β-1, 3; β-1, 4), and starch glucosidase (α-1, 4; α-1, 6; and α-1, 3). The samples were incubated at 50°C in water bath for 6 h. The content of saccharides in the supernatant was measured by phenol sulfuric acid method (Cuesta et al., 2003). Then the cells were centrifuged, washed by PBS, and treated with the trypsin by the same way as described above. After centrifugation at 8,000 g, 10 min, 4°C, the samples were filtered with a 0.22-μm pore-size filter and sent to Sangon Biotech (Shanghai) Co., Ltd. for the 30 min LC-MS/MS analysis. The MS/MS data was acquired when the precursor ion signal was greater than 120° and the charge number was +2∼+5. The database used for searching was the proteome reference database of *V. parahaemolyticus O3:K6 (strain RIMD 2210633)* in UniProt. Two trial of the experiments were conducted and three replicates of each sample were tested.

Structure Simulation of the Candidate Proteins

The three-dimensional protein structure was simulated by the online servers, I-TASSER (Yang et al., 2015), and TrRosetta (Yang et al., 2020). I-TASSER is an iterative template-based fragment assembly simulation server. TrRosetta builds the protein structure based on direct energy minimization with a restrained Rosetta. Briefly, the sequence of the candidate protein was inputted and submitted to the server with an automatic model, and a high-quality simulated protein structure is obtained when the C-score is more than −1.5 in I-TASSER, the C-score is greater than −1.5, and a Tm score more than 0.5 in TrRosetta (Xu and Zhang, 2010).

RESULTS

Bioinformatical Identification of *V. parahaemolyticus* Outer Membrane Proteins

In this study, screening of the 4,831 proteins encoded by *V. parahaemolyticus* RIMD 2210633 revealed 728 proteins that

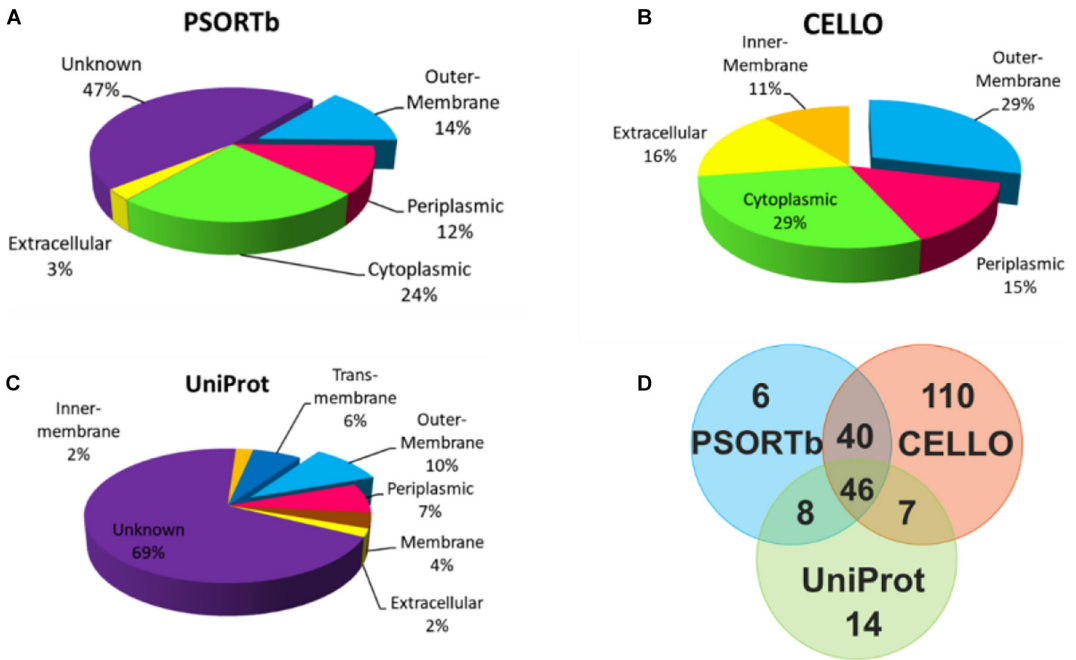


FIGURE 2 | Subcellular location analysis of the 705 proteins by PSORTb **(A)**, CELLO **(B)**, and UniProt **(C)**. **(D)** Mapping the OMPs predicted by three methods.

had a signal peptide sequence in the N-terminal region. Among them, 505 were general secretory pathway (Sec) signal peptides, 21 were twin-arginine translocation (TAT) signal peptides, and 202 were lipoprotein signal peptides. The outer membrane proteome of bacteria is dominated by a β-barrel with no TM α-helix or α-helical membrane proteins with one or more TM α-helix (Rollauer et al., 2015). TMHMM 2.0 and HMMTOP showed 665 proteins had no TM α-helix and 40 proteins had one TM α-helix, which indicated they were primarily β-barrel proteins. The β-barrel proteins are mainly porin proteins that occur only in the outer membranes of bacteria and are involved in metabolite and protein transport, osmotic pressure regulation, and numerous signaling processes (Wimley, 2003). The other 23 proteins were excluded for further analysis due to a substantial challenge of multi-pass TM protein expression in a prokaryotic expression system as reported by Serruto (Serruto et al., 2004).

To select OMPs that are primarily involved in the host environment interaction and exclude the cytoplasmic proteins (Sanober et al., 2017), the subcellular localization of the proteins was predicted using three parallel online approaches. Of the 705 *V. parahaemolyticus* proteins, PSORTb, CELLO, and UniProt revealed that 100 (14.2%), 203 (28.8%), and 75 (10.6%), respectively, were located in the outer membrane region (**Figure 2**). As shown in **Figure 2A** and **C**, 47 and 69% of the proteins were not predicted or annotated with clear subcellular location by PSORTb and UniProt. It's noteworthy that UniProt annotated 26 membrane proteins and 45 TM proteins, some of which may also be OMPs. To balance the prediction coverage and confidence of prediction, the 101 proteins (**Supplementary Table 1**) that were predicted as OMPs by at least two servers were selected for further analysis.

Homology and Adhesion of *V. parahaemolyticus* and *Vibrio* spp. Proteins

The sequence homology of 101 OMPs identified in *V. parahaemolyticus* RIMD 2210633 with heterogenous serotypes and isolates of *V. parahaemolyticus* was analyzed by BLAST in NCBI. The results showed that these OMPs are conserved (>98%) in *V. parahaemolyticus*. The BLAST of the identified OMPs with 32 other *Vibrio* species, seven closely related marine bacteria, and six other common bacteria showed that 70% of the OMPs (e.g., previously reported OmpA, OmpW, and OmpU) had relatively high but uneven sequence homology with other *Vibrio* species and marine bacteria, such as *Photobacterium*, *Shewanella*, *Aliivibrio*, *Salinivibrio*, *Enterovibrio*, and *Grimontia*. Interestingly, most of the OMPs were found to share a high sequence homology with nine *Vibrio* species, including *V. alginolyticus*, *V. diabolicus*, *V. antiquaries*, *V. campbelli*, *V. harveyi*, *V. owensii*, *V. rotiferianus*, *V. jasicida*, and *V. natriegens*, which indicated a close genetic relationship between these strains. Notably, *V. diabolicus* and *V. antiquaries* are isolated from a deep-sea hydrothermal vent environment, and their prevalence and pathogenicity in coastal systems and seafood samples remains unclear (Turner et al., 2018).

However, we did identify 19 OMPs that exhibited relatively good specificity for *V. parahaemolyticus*. As listed in **Table 1**, six OMPs, including OmpA-like domain-containing protein (VPA0731), outer membrane protein (VPA0316), AraC/XylS

TABLE 1 | The homology of differential OMPs of *V. parahaemolyticus* identified by bioinformatic tools.

Accession	Description	α -Helix	Adhesion	Sequence homology		
				V. parahaemolyticus	**32 Species of *Vibrio* spp.**	**non-*Vibrio* strains**
NP_800241.1	OmpA-like domain-containing protein	0	0.200	92% 100	74%, *V. harveyi* 74%, *V. diabolicus*	*N*
NP_798570.1	hypothetical protein	1	0.125	91% 280	<66%, nine species[a]	*N*
NP_800966.1	hypothetical protein	1	0.103	97% 430	<75%, nine species	*N*
NP_799826.1	Outer membrane protein	0	0.686	92% 70	<76%, nine species	*N*
NP_797233.1	AraC/XylS family transcription factor	1	0.250	99% 500	<81%, nine species	*N*
NP_800058.1	hypothetical protein	0	0.353	98% 73	<85%, nine species	*N*
NP_799595.1	hypothetical protein	0	0.325	99% 54	99%, *V. diabolicus*	*Photobacterium* 50%
NP_800886.1	hypothetical protein	0	0.586	97% 20	83%, *V. cholerae* 97%, *V. mimicus*	*Shewanella* et al. 57%
NP_799735.1	Putative efflux pump channel protein	0	0.565	97% 344	97%, *V. alginolyticus* 80%, *V. natriegens*	*Shewanella* et al. 55%
NP_797735.1	hypothetical protein	0	0.468	95% 287	87%, *V. natriegens*	*Photobacterium* 74%
NP_801014.1	CsuD protein	0	0.411	97% 444	87%, *V. alginolyticus*, *V. diabolicus, V. antiquaries*	*Enterovibrio* 60%
NP_797599.1	MtrC	0	0.430	97% 467	*96%, V. diabolicus, V. antiquaries*	*Enterovibrio* 52%
OAR40234.1	flagellin	0	0.393	98% 304	<81%, nine species	*Grimontia* et al. 47%
NP_797143.1	OmpA	0	0.311	97% 78	<88%, nine species	*Photobacterium* 50%
NP_799317.1	vitamin B12 receptor	0	0.608	97% 154	<89%, nine species	*Aeromonas* et al. 41%
NP_799752.1	OMP_b-brl domain-containing protein	0	0.611	95% 122	<89%, nine species	*Grimontia* 67%
NP_800976.1	TonB system receptor	0	0.508	97% 405	<89%, nine species	*Aliivibrio* et al. 59%

[a]*homology was only found with the nine species including V. alginolyticus, V. diabolicus, V. antiquaries, V. campbelli, V. harveyi, V. owensii, V. rotiferianus, V. jasicida, and V. natriegens.*

family transcription factor (VP0854), and three hypothetical proteins (VP2191, VPA1456, and VPA0548), shared negligible homology (<32%) with non-*Vibrio* spp. and relatively low homology (<85%) with two to nine *Vibrio* species. Eleven OMPs shared a low homology (38 – 74%) with several non-*Vibrio* strains, six of which (NP_799595.1, NP_800886.1, NP_799735.1, and NP_797735.1, CsuD protein, and MtrC) only had a high homology (83 – 99%) with one to three *Vibrio* species (*V. alginolyticus*, *V. diabolicus*, and *V.antiquaries*). The last five OMPs (flagellin, OmpA, NP_799317.1, NP_799752.1, and NP_800976.1) shared a sequence homology of less than 89% with the closely related nine *Vibrio* species. There are two differential OMPs that are not listed in **Table 1**. The long-chain fatty acid transport protein (NP_798592.1) was only found to display a high homology (93%) with *V. alginolyticus*, *V. diabolicus*, *V. antiquaries*, and *V. rotiferianus* and shares a 66% identify with *Photobacterium* and *Shewanella*. The outer membrane porin protein (NP_797387.1) is specific to the *V. parahaemolyticus*, *V. diabolicus*, *V. antiquaries*, *V. campbelli*, *V. harveyi*, and *V. jasicida* and was not found in the non-*Vibrio* spp. strains.

Table 2 shows that there were 15 OMPs that shared at least a 65% sequence homology across 23 or more species out of the 32 *Vibrio* species. These proteins are widely conserved in *Vibrio* spp. because they are involved in the fundamental structure or biological function, such as the flagellar motor protein complex (e.g., flgH1 and motY), new OMPs assembly (e.g., BamA), polysaccharide assembly (e.g., LptD), fimbriae assembly (NP_799125.1 and NP_796512.1), the TolC subfamily within the efflux systems (NP_798013.1, NP_798092.1, and NP_796804.1), maintaining the outer membrane integrity (VacJ), polysaccharide binding, and carbohydrate metabolic process (NP_797134.1). While the protective efficacy of proteins OmpW (Mao et al., 2007), OmpK (Li et al., 2010b), LptD (Zha et al., 2016), and Pal (Li et al., 2010a) has already been studied. The other OMPs that are conserved in more *Vibrio* species were newly identified in this study.

Figure 3A shows the relationship of the identified OMPs with the number of α-helix structure, *V. parahaemolyticus* differential OMPs, *Vibrio* spp. conserved OMPs. Most candidate OMPs had no TM α-helix, which indicated a dominant β-barrel structure. The analysis of protein antigenicity by VaxiJen showed that 96 out of the 101 proteins exhibited a higher score than the cut-off (>0.4), which indicated the identified OMPs were potentially antigenic proteins. Adhesion proteins help bacteria attach to the host cell receptors and promote colonization. Creating a vaccine and generating associated neutralizing antibodies against adhesion proteins is effective in preventing bacterial infections at an early stage (Wizemann et al., 1999). The results of the analysis by Vaxign showed that 43 OMPs had adhesion probability higher than 0.5 and were predicated to be related to adhesion (**Figure 3B**). Among these, eight OMPs (outer membrane protein NP_799826.1, hypothetical protein NP_800886.1, putative efflux pump channel protein NP_799735.1, vitamin B12 receptor NP_799317.1, OMP_b-brl domain-containing protein NP_799752.1, and TonB system receptor NP_800976.1) were more specific to *V. parahaemolyticus* (**Table 1**) and four OMPs (flgH, LptD, OmpK, and OmpW) were widely conserved in *Vibrio* spp. (**Table 2**).

Surface Proteins Identified by Enzymic Shaving and Mass Spectrometry

Trypsin was specific to hydrolyze carboxyl side of lysine and arginine residues in polypeptide chain and was widely used in proteomic studies (Olaya-Abril et al., 2014). To experimentally study the OMPs on the cell surface, trypsin was first used to enzymatically shave off the intact cells. The results of LC-MS/MS showed although little intracellular proteins were observed for the control sample without trypsin, there were no known OMPs identified in the experimental group after 30 min shaving except some intracellular proteins (**Table 3**). The identified BamD that annotated as OMP actually faced the periplasm side of the outer membrane (Han et al., 2016). *Vibrio* was known to express the capsule (K antigen) on the cell to adapt the environment and escape the phagocytosis (Pettis and Mukerji, 2020). This may limit the availability of OMPs to trypsin digestion. Many studies have shown that the capsule or lipopolysaccharides consist of α-1,3, α-1,2, α-1,4, β-1,4, β-1,3, α-1,5, β-1,5, α-1,6, and β-1,6 glycosidic bond (Chen et al., 2007). The endoglycosidases were known to cleavage the specific glycosidic bond inside the polysaccharide and have been reported to reduce the capsule level of *Streptococcus pneumoniae* (Middleton et al., 2018). With the combination of four endoglycosidase and one exoglycosidase in a one-step enzymatical digestion of the intact cells, the sugar content in the supernatant increased from 4.89 to 12.90 mg/100 g after 6 h (**Supplementary Figure 1**). This indicated the reduction of exopolysaccharides on the cell. The following enzymatical shaving of the intact cells by trypsin showed more intracellular proteins in both the positive and control group were observed, but peptides of some proteins were mainly observed in the experimental group and further analysis with UniProt confirmed there were 14 OMPs identified. These OMPs were OmpA (VPA1186, VPA0248, and VP0764), Omp (VPA0166), OmpU (VP2467), BamA (VP2310), TolC (VP0425), GspD (VP0133), OmpK (VP2362), lpp (VPA1469), Pal (VP1061), 5′-nucleotidase (VP0748), agglutination protein (VP1634), and putative iron (III) compound receptor (VPA1435).

Protein Structure of the Omps Identified by Both Reverse Vaccinology and Surface Proteome

The results in **Tables 2**, **3** revealed that OmpK, agglutination protein, Pal, TolC, GspD, and BamA were identified by the bioinformatical screening of the encoded protein database and enzymatical shaving the intact cells of *V. parahaemolyticus*. Protein simulation by I-TASSER and TrRosetta (**Figure 4**) showed that all these protein structures had an estimated TM score ≥0.5 and C-score more than −1.5. The tolC super family protein agglutination protein (VP1634) and Tolc (VP0425) had a typical β-barrel domain on the outer membrane and α-helical colied coli structure protrude deep into the periplasm (Zgurskaya et al., 2011). The GspD belong to the type II secretory system also had a similar structure which promote the secretion of

TABLE 2 | The homology of differential OMPs of *Vibrio* spp. identified by bioinformatic tools.

Accession	Description	α -Helix	Adhesion	Sequence homology		
				V. parahaemolyticus	**32 Species of *Vibrio* spp.**	**non-*Vibrio* strains[a]**
NP_798689.1	Outer membrane protein assembly factor YaeT BamA	0	0.439	98% 63	>70%, 31 species	53–78%
NP_796804.1	Outer membrane channel protein Tolc	0	0.427	91% 62	>70%, 31 species	51–69%
NP_799125.1	Fimbrial assembly protein	0	0.222	93% 106	>65%, 31 species	52–70%
NP_796512.1	General secretion pathway protein D	0	0.142	93% 49	>75%, 31 species	54–70%
NP_797161.1	Flagellar basal body L-ring protein	0	0.681	99% 33	>70%, 30 species	*Salinivibrio* 58%
NP_798092.1	Hypothetical protein	0	0.336	93% 160	>65%, 29 species	52–62%
NP_798593.1	VacJ lipoprotein	0	0.289	99% 99	>65%, 29 species	50–62%
NP_796718.1	LPS-assembly protein LptD	0	0.649	92%117	>60%, 29 species	44–48%
NP_798013.1	Agglutination protein tolC super family	0	0.411	99% 94	>70%, 28 species	51–68%
NP_798490.1	Sodium-type flagellar protein MotY	0	0.255	94% 60	>70%, 28 species	62–65%
SUQ25448.1	Outer membrane protein OmpK	0	0.763	79% 43	>65%, 27 species	65–80%
NP_797646.1	Lipoprotein	1	0.416	97%26	>70%, 25 species	40–70%
NP_797440.1	Peptidoglycan-associated lipoprotein Pal	0	0.333	99%24	>70%, 25 species	61–64%
NP_797134.1	Chitobiase	0	0.386	99% 380	>70%, 24 species	57–74%
NP_799606.1	Outer membrane protein OmpW	0	0.774	90%51	>65%, 23 species	*Grimontia* 71%

[a]*These non-Vibrio strains include Aliivibrio Sp., Salinivibrio Sp., Enterovibrio Sp.,Grimontia Sp., and Photobacterium Sp.*

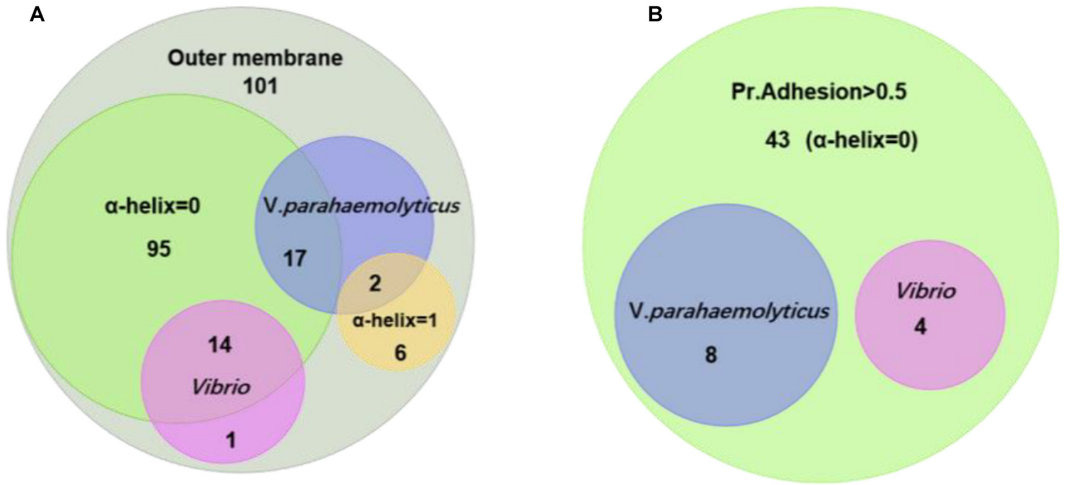


FIGURE 3 | Homology and adhesion. **(A)** The distribution and crossover of 101 outer membrane protein candidates. There are 15 proteins that are homologous in *Vibrio*; 19 proteins that are homologous in *V. parahaemolyticus*; 95 proteins that have no TM α-helix; and six proteins that have one TM α-helix. **(B)** There are 43 proteins that have a Pr. adhesion score of more than 0.5; four proteins are homologous in *Vibrio*; and eight proteins are homologous in *V. parahaemolyticus*.

protein. The BamA and OmpK structure had a common β-barrel structure of a porin protein, which are involved in protein assembly and receptor.

DISCUSSION

The novel OMPs as immunogens for preparing diagnostic antibodies and subunit vaccine against *V. parahaemolyticus* and *Vibrio* spp. are of interest in aquaculture and food safety. Traditional study of the major OMPs was unsystematic and time consuming. Based on bioinformatic tools with different filter thresholds and applications, reverse vaccinology processes a large quantity of genetic or proteomic data and minimize the targets in the following *in vivo* experiments. The surface proteome based on enzymatical shaving of the extracellular peptides of the cells and massive identification of the peptides provides an experimental insight of the surface available OMPs. The two comprehensive and complementary strategies were both adopted in this work to study the diagnostic surface antigens of *V. parahaemolyticus* and *Vibrio* species.

For the bioinformatic analysis, signal sequences presented at the N-terminal were first predicted because most bacterial OMPs have an N-terminal signal sequence that promotes protein secretion (Gao et al., 2016). Our result identified three different signal peptides. The Sec signal peptide directs unfolded protein translocation across the plasma membrane in prokaryotes (Tsirigotaki et al., 2017). TAT signal peptide actively translocate the folded proteins across the lipid bilayer of the membrane (Armenteros et al., 2019) and are generally longer and less hydrophobic (Berks, 2015). The lipoprotein signal peptides play an important role in the maturation of bacterial lipoprotein (Kitamura and Wolan, 2018), many of which are surface-exposed (Wilson and Bernstein, 2016). To cross-check the subcellular location of candidate proteins, three parallel prediction tools were used in this study. Among them, PSORTb was reported to be the most precise bacterial localization prediction tool available based on a report in 2010 (Yu et al., 2010). The results in **Figure 2D** showed that when CELLO and PSORTb, CELLO and UniProt, PSORTb and UniProt were used to cross-check, 86, 53, and 54 proteins were, respectively, predicted as OMPs. Parallel use of the three prediction tools and retaining the proteins predicted by two or three methods obtained more candidate OMPs (101) with better coverage and accuracy.

The homology analysis of the *V. parahaemolyticus* OMP candidates by BLASTp revealed a relatively conserved and complex genetic relationship with the other *Vibrio* species and some marine bacteria, which provide the insight of possible cross-reaction or cross-protection with these OMPs. Based on the overall homology analysis, 15 OMPs that are broadly conserved in the genus of *Vibrio* were identified. Four of them were predicted to be adhesion proteins and were potentially multivalent subunit vaccine candidates against vibriosis. Many of these conserved candidate proteins are involved in some significant function. For example, BamA participates in the delivery of extracellular membrane proteins and is the core component of the β-barrel assembly machinery BAM complex (Singh et al., 2017). Peptidoglycan-associated lipoprotein (Pal) stabilizes the outer membrane by providing a non-covalent biding through the peptidoglycan layer (Parsons et al., 2006). The general secretion pathway protein (GspD) in *V. cholerae* directly involves the production of rugose polysaccharide and the secretion of cholera toxin and hemolysin (Ali et al., 2000). The TolC family protein provides a channel for the cell to connect with the external environment during export and plays a role in the bacterial efflux pumps (Koronakis et al., 2000). The protein LptD is responsible for lipopolysaccharide transport and insertion into the outer membrane (Dong et al., 2014). The flagellar basal body L-ring protein (flgH), sodium-type flagellar protein motY, and fimbrial assembly protein involved in motility

TABLE 3 | The identified OMPs in *V. parahaemolyticus* O3:K6 (RIMD 2210633) by proteomic analysis.

Accession	Description	Function	Mw (kD)	Peptide (–[a])	Peptide (T[b])	Peptide (G[c])	Peptide (G + T[d])	Bioinformatics
NP_797127.1	5′-nucleotidase	Degradation for nutrition	62.175	–	–	1	1	+
NP_798741.1	OmpK	Receptor for vibriophage	29.877	–	–	–	2	+
NP_800979.1	Outer membrane lipoprotein (lpp)	Distance controls of membranes	8.671	–	–	–	1	+
NP_800945.1	Putative iron (III) compound receptor	Siderophore uptake transporter	77.059	–	–	–	1	+
NP_800696.1	OmpA	Porin activity	36.014	–	–	–	5	+
NP_799758.1	OmpA	Porin activity	35.553	–	–	–	9	+
NP_799676.1	Putative Omp	Porin activity	37.974	–	–	1	4	+
NP_798846.1	OmpU	Passive diffusion of small molecules	36.285	–	–	–	10	+
NP_798689.1	BamA	Omp assembly complex	90.053	–	–	–	3	+
NP_798013.1	Agglutination protein	Efflux transmembrane transporter	48.69	–	–	–	1	+
NP_797440.1	Pal	Division and cell integrity	18.713	–	–	1	3	+
NP_797143.1	OmpA	Porin activity	34.073	–	–	2	9	+
NP_796804.1	TolC	Efflux transmembrane transporter	47.983	–	–	–	8	+
NP_796512.1	GspD	Protein secretion by the T2S	73.317	–	–	–	1	+

[a]*Control group treated with only the shaving buffer of trypsin (1× PBS, 20% [w/v] sucrose, 10 mM DTT).*
[b]*Experimental group treated with trypsin in shaving buffer.*
[c]*Experimental group treated with glycosidases in phosphate buffer (pH 5.5) and the shaving buffer of trypsin.*
[d]*Experimental group treated with glycosidases in phosphate buffer (pH 5.5) and trypsin in shaving buffer.*

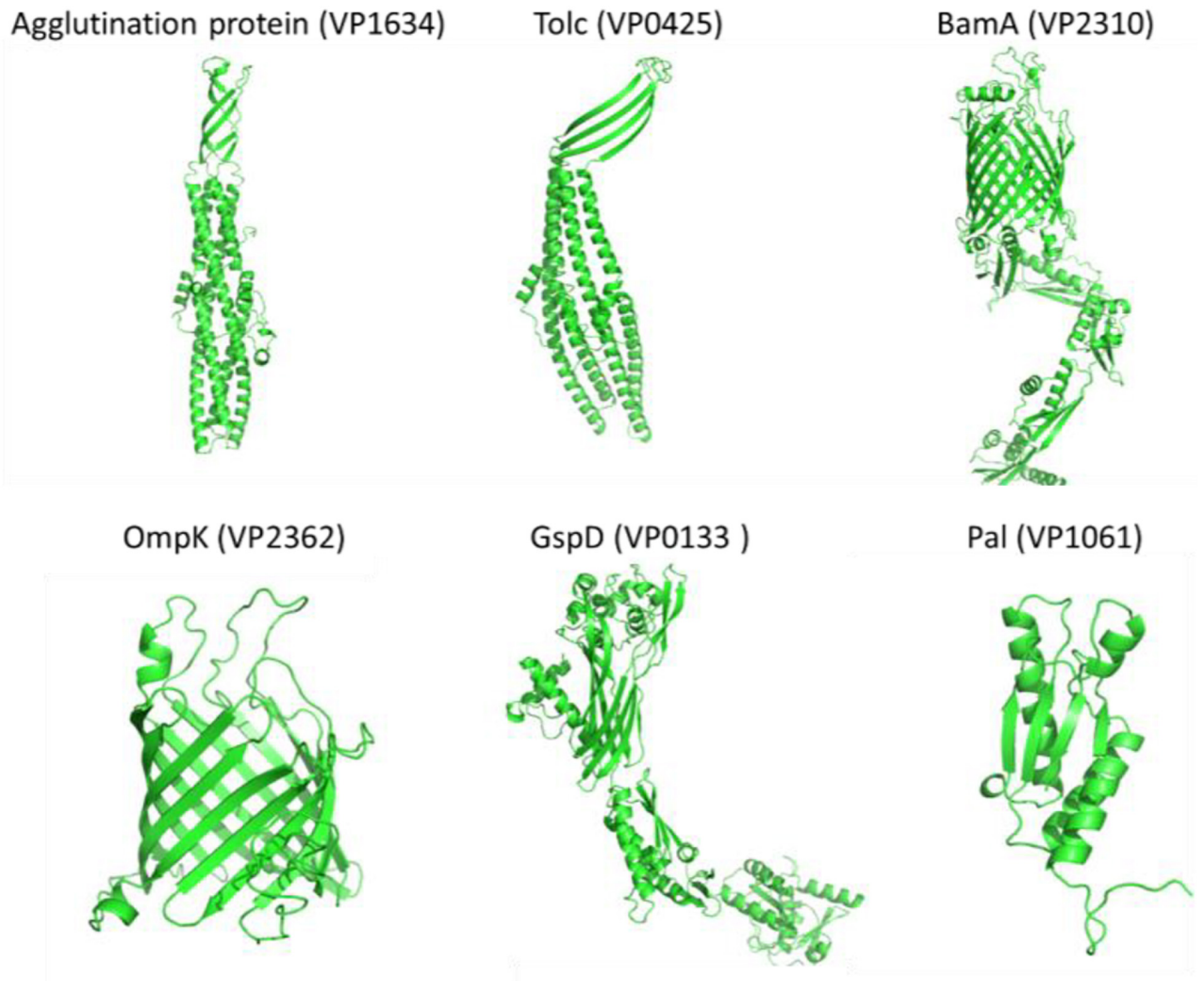


FIGURE 4 | The simulated protein structures of six widely conserved *Vibrio* Omps (Agglutination protein, Tolc, BamA, OmpK, GspD, and Pal) identified by both reverse vaccinology and surface proteomic analysis.

and adhesion (Rehman et al., 2019). The OmpK and OmpW are porin proteins that are involved in the transport of small hydrophobic molecules across the bacterial outer membrane (Hong et al., 2006).

Due to the complex homology, completely specific *V. parahaemolyticus* OMPs were not found, but 19 differential OMPs were obtained in this study. The CsuD protein and MtrC (Beliaev et al., 2001), respectively, has fimbrial usher porin activity and terminal Fe(III) reductase activity. Vitamin B12 receptor and TonB system receptor play a role in nutrition uptake and transportation (Köster, 2001). OmpA and OmpA like protein function as an adhesin and invasin, participate in biofilm formation, and serve as a bacteriophage receptor (Smith et al., 2007). The AraC-XylS family proteins are known as transcriptional proteins regulator (Tobes and Ramos, 2002). Further analysis of the structure, topology, and sequence homology of these OMPs may reveal the specific extracellular epitopes for *V. parahaemolyticus* and overcome the challenge of discovering diagnostic targets.

Previous immunoproteomic work performed by Li et al. (2010a) identified the putative iron (III) compound receptor (VPA1435), TolC (VP0425), and OmpA (VPA0764), conserved hypothetical protein (VP2850), BamA (VP2310), and Pal (VP1061) as the immunogenic OMPs of *V. parahaemolyticus*. Similar work described by another group identified the lipopolysaccharide-assembly protein, LptD, hypothetical protein (VP0802), Maltoporin (LamB), OmpA, OmpU, and OmpK, as well as hypothetical VP1243 and VP0966 (Li et al., 2014). Recently, Yonekita et al. (2020) reported that a diagnostic antibody prepared against a whole cell antigen was specific to 12 serotypes of *V. parahaemolyticus* and did not react with the other six *Vibrio* species. The antigen was identified to be outer membrane lipoprotein (NP_800979.1 and VPA1469). All of these proteins were predicted as OMPs in this study, except VP2850,

VP1243, and VP0966, which were only predicted by CELLO. Importantly, we confirmed the homology of these OMPs among the 32 species of *Vibrio* and closely related non-*Vibrio* strains at the sequence level. LptD, Pal, BamA, TolC, OmpK, and OmpW (**Table 2**) were conserved in at least 23 *Vibrio* species. The lipo protein VPA1469 was also found to have a high homology with *V. campbelli*, *V. harveyi V. owensii*, and *V. rotiferianus*. Apart from these Omps, we identified new OMPs that were conserved in *Vibrio* spp. or relatively specific to *V. parahaemolyticus* for development of better diagnostic tools and vaccines.

In our surface proteome study, direct shaving of PFA fixed *V. parahaemolyticus* cells resulted in no OMPs but some cytoplasmic proteins, which were also found in many previous surface proteome studies of Gram-negative strains (Fagerquist and Zaragoza, 2018). This is mainly because the soft cell membrane is more vulnerable to osmotic lysis (Olaya-Abril et al., 2014). Combination of four endoglycosidases and an exoglycosidase in a one-step digestion reduced capsule level of *V. parahaemolyticus* in this study. The following shaving with trypsin and analysis indicated 14 OMPs, also identified in the reverse vaccinology strategy, may have a better accessibility on the cell surface. Rodríguez-Ortega found the number of surface-exposed proteins for group A *Streptococcus* varied from strain to strain and it's mostly because of the different capsule content (Rodríguez-Ortega et al., 2006). Our results also suggested a steric hindrance effect of the capsular polysaccharide, which was widely expressed by *Vibrio* species, most likely masked the OMPs from enzymatical shaving of the intact cell. The mask of many OMPs by capsule may also limit its reaction with antibodies, which should be considered in immunodetection and vaccine development. The expression of capsule of *Vibrio* was regulated by temperature, divalent cation (Ca^{2+} and Mn^{2+}), oxygen content, and host environment (Pettis and Mukerji, 2020). Further study of the surface proteome of *V. parahaemolyticus* with minimal capsule expression will elucidate more stably expressed OMPs and the optimal enrich broth for the immunodetection of *V. parahaemolyticus* from various environmental samples.

In short, reverse vaccinology analysis of the surface proteins based on the protein database of *V. parahaemolyticus* in NCBI comprehensively predicted 101 OMPs and revealed 19 differential OMPs of *V. parahaemolyticus* and 15 conserved OMPs of *Vibrio* spp. Surface proteome study of *V. parahaemolyticus* by enzymatical shaving of the exopolysaccharide and proteins on the intact cell identified 14 OMPs, which mainly belong to the conserved OMPs of *Vibrio* spp. Future study of the capsule expression level, surface proteome of *V. parahaemolyticus* under various environments, and the interaction of *V. parahaemolyticus* with antibodies against these Omps will further reveal the diagnostic surface antigen or epitopes of *V. parahaemolyticus* and *Vibrio* spp.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/ **Supplementary Material**.

ETHICS STATEMENT

This article does not contain any studies with human participants or animal experiments.

AUTHOR CONTRIBUTIONS

WW and SP initiated the study. JL and WW mainly conducted the experiment, analyzed the data, and wrote the manuscript. JL, LL, SG, and QY participated the data analysis. LG revised the manuscript. All the authors read and approved the manuscript.

FUNDING

This work is supported by the National Natural Science Foundation of China (Grant No. 31701680), the Natural Science Foundation of Jiangsu Province of China (Grant No. SBK2017041308), the Open Project Program of Key Laboratory for Analytical Science of Food Safety and Biology, Ministry of Education (FS2003), Open-end Funds of Jiangsu Key Laboratory of Marine Biotechnology, Jiangsu Ocean University (HS2019003), Postgraduate Research and Practice Innovation Program of Jiangsu Province(KYCX19_2289), and also funded by the Priority Academic Program Development of Jiangsu Higher Education Institutions.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2020.625315/full#supplementary-material

REFERENCES

Ali, A., Johnson, J. A., Franco, A. A., Metzger, D. J., Connell, T. D., Morris, J. G., et al. (2000). Mutations in the extracellular protein secretion pathway genes (eps) interfere with rugose polysaccharide production in and motility of *Vibrio cholerae*. *Infect. Immun.* 68, 1967–1974. doi: 10.1128/iai.68.4.1967-1974.2000

Aly, S. M., Eissa, A. E., Elbanna, N. I., and Albutti, A. (2020). Efficiency of monovalent and polyvalent *Vibrio alginolyticus* and *Vibrio parahaemolyticus* vaccines on the immune response and protection in gilthead sea bream, *Sparus aurata* (L.) against vibriosis. *Fish Shellfish Immunol.* 2020:11. doi: 10.1016/j.fsi.2020.10.011

Armenteros, J. J. A., Tsirigos, K. D., Sønderby, C. K., Petersen, T. N., Winther, O., Brunak, S., et al. (2019). SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat. Biotechnol.* 37, 420–423. doi: 10.1038/s41587-019-0036-z

Baker-Austin, C., Oliver, J. D., Alam, M., Ali, A., Waldor, M. K., Qadri, F., et al. (2018). *Vibrio* spp. infections. *Nat. Rev. Dis. Primers* 4:8.

Baliga, P., Shekar, M., and Venugopal, M. N. (2018). Potential outer membrane protein candidates for vaccine development against the pathogen *Vibrio anguillarum*: A reverse vaccinology based identification. *Curr. Microbiol.* 75, 368–377. doi: 10.1007/s00284-017-1390-z

Beliaev, A. S., Saffarini, D. A., Mclaughlin, J. L., and Hunnicutt, D. (2001). MtrC, an outer membrane decahaem c cytochrome required for metal reduction in *Shewanella putrefaciens* MR-1. *Mol. Microbiol.* 39, 722–730. doi: 10.1046/j.1365-2958.2001.02257.x

Berks, C. B. (2015). The twin-arginine protein translocation pathway. *Annu. Rev. Biochem.* 84, 843–864. doi: 10.1146/annurev-biochem-060614-034251

Bonnin-Jusserand, M., Copin, S., Le Bris, C., Brauge, T., Gay, M., Brisabois, A., et al. (2019). *Vibrio* species involved in seafood-borne outbreaks (*Vibrio cholerae*, *V. parahaemolyticus* and *V. vulnificus*): Review of microbiological versus recent molecular detection methods in seafood products. *Crit. Rev. Food Sci. Nutr.* 59, 597–610. doi: 10.1080/10408398.2017.1384715

Chen, Y., Bystricky, P., Adeyeye, J., Panigrahi, P., Ali, A., Johnson, J. A., et al. (2007). The capsule polysaccharide structure and biogenesis for non-O1 *Vibrio cholerae* NRT36S: genes are embedded in the LPS region. *BMC Microbiol.* 7, 20–20. doi: 10.1186/1471-2180-7-20

Cheng, Z.-X., Chu, X., Wang, S.-N., Peng, X.-X., and Li, H. (2018). Six genes of ompA family shuffling for development of polyvalent vaccines against *Vibrio alginolyticus* and *Edwardsiella tarda*. *Fish Shellfish Immunol.* 75, 308–315. doi: 10.1016/j.fsi.2018.02.022

Cuesta, G., Suarez, N., Bessio, M. I., Ferreira, F., and Massaldi, H. (2003). Quantitative determination of pneumococcal capsular polysaccharide serotype 14 using a modification of phenol–sulfuric acid method. *J. Microbiol. Methods* 52, 69–73. doi: 10.1016/s0167-7012(02)00151-3

Dalsass, M., Brozzi, A., Medini, D., and Rappuoli, R. (2019). Comparison of open-source reverse vaccinology programs for bacterial vaccine antigen discovery. *Front. Immunol.* 10:113. doi: 10.3389/fimmu.2019.00113

Dong, H., Xiang, Q., Gu, Y., Wang, Z., Paterson, N. G., Stansfeld, P. J., et al. (2014). Structural basis for outer membrane lipopolysaccharide insertion. *Nature* 511, 52–56. doi: 10.1038/nature13464

Doytchinova, I. A., and Flower, D. R. (2007). VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC Bioinf.* 8:4. doi: 10.1186/1471-2105-8-4

Duperthuy, M., Binesse, J., Le Roux, F., Romestand, B., Caro, A., Got, P., et al. (2010). The major outer membrane protein OmpU of *Vibrio splendidus* contributes to host antimicrobial peptide resistance and is required for virulence in the oyster *Crassostrea gigas*. *Environ. Microbiol.* 12, 951–963. doi: 10.1111/j.1462-2920.2009.02138.x

Elmahdi, S., Dasilva, L. V., and Parveen, S. (2016). Antibiotic resistance of *Vibrio parahaemolyticus* and *Vibrio vulnificus* in various countries: a review. *Food Microbiol.* 57, 128–134. doi: 10.1016/j.fm.2016.02.008

Erik, L. L., Sonnhammer, Heijne, G. V., and Krogh, A. (1998). A hidden Markov model for predicting transmembrane helices in protein sequences. *ISMB* 6, 175–182.

Fagerquist, C. K., and Zaragoza, W. J. (2018). Proteolytic surface-shaving and serotype-dependent expression of SPI-1 invasion proteins in *Salmonella enterica* subspecies *enterica*. *Front. Nutr.* 5:124. doi: 10.3389/fnut.2018.00124

Gao, D., Luan, Y., Liang, Q., and Qi, Q. (2016). Exploring the N-terminal role of a heterologous protein in secreting out of *Escherichia coli*. *Biotechnol. Bioeng.* 113, 2561–2567. doi: 10.1002/bit.26028

Garcia-Angulo, V. A., Kalita, A., Kalita, M., Lozano, L., and Torres, A. G. (2014). Comparative genomics and immunoinformatics approach for the identification of vaccine candidates for enterohemorrhagic *Escherichia coli* O157:H7. *Infect. Immun.* 82, 2016–2026. doi: 10.1128/iai.01437-13

Giuliani, M. M., Adu-Bobie, J., Comanducci, M., Aricò, B., Savino, S., Santini, L., et al. (2006). A universal vaccine for serogroup B meningococcus. *PNAS* 103, 10834–10839.

Guin, S., Saravanan, M., Chowdhury, G., Pazhani, G. P., Ramamurthy, T., and Das, S. C. (2019). Pathogenic *Vibrio parahaemolyticus* indiarrhoeal patients, fish and aquatic environments and their potential for inter-source transmission. *Heliyon* 5:e01743. doi: 10.1016/j.heliyon.2019.e01743

Hackbusch, S., Wichels, A., Gimenez, L., Döpke, H., and Gerdts, G. (2020). Potentially human pathogenic *Vibrio* spp. in a coastal transect: Occurrence and multiple virulence factors. *Sci. Total Environ.* 707:136113. doi: 10.1016/j.scitotenv.2019.136113

Han, L., Zheng, J., Wang, Y., Yang, X., Liu, Y., Sun, C., et al. (2016). Structure of the BAM complex and its implications for biogenesis of outer-membrane proteins. *Nat. Struct. Mol. Biol.* 23, 192–196. doi: 10.1038/nsmb.3181

He, Y., Xiang, Z., and Mobley, H. T. (2010). Vaxign: the first web-based vaccine design program for reverse vaccinology and an application for vaccine development. *J. Biomed. Biotechnol.* 2010:297505.

Hong, H., Patel, D. R., Tamm, L. K., and Van Den Berg, B. (2006). The outer membrane protein OmpW forms an eight-stranded β-barrel with a hydrophobic channel. *J. Biol. Chem.* 281, 7568–7577. doi: 10.1074/jbc.m512365200

Hornburg, D., Kruse, T., Anderl, F., Daschkin, C., Semper, R. P., Klar, K., et al. (2019). A mass spectrometry guided approach for the identification of novel vaccine candidates in Gram-negative pathogens. *Sci. Rep.* 9:17401.

Huang, Z., Tang, J., Li, M., Fu, Y., Dong, C., Zhong, J. F., et al. (2012). Immunological evaluation of *Vibrio alginolyticus*, *Vibrio harveyi*, *Vibrio vulnificus* and infectious spleen and kidney necrosis virus (ISKNV) combined-vaccine efficacy in *Epinephelus coioides*. *Vet. Immunol. Immunopathol.* 150, 61–68. doi: 10.1016/j.vetimm.2012.08.008

Jaiswal, V., Chanumolu, S. K., Gupta, A., Chauhan, R. S., and Rout, C. (2013). Jenner-predict server: prediction of protein vaccine candidates (PVCs) in bacteria based on host-pathogen interactions. *BMC Bioinf.* 14:211. doi: 10.1186/1471-2105-14-211

Kitamura, S., and Wolan, D. W. (2018). Probing substrate recognition of bacterial lipoprotein signal peptidase using FRET reporters(Letter). *FEBS Lett.* 592, 2289–2296. doi: 10.1002/1873-3468.13155

Koronakis, V., Sharff, A., Koronakis, E., Luisi, B., and Hughes, C. (2000). Crystal structure of the bacterial membrane protein TolC central to multidrug efflux and protein export. *Nature* 405, 914–919. doi: 10.1038/35016007

Köster, W. (2001). ABC transporter-mediated uptake of iron, siderophores, heme and vitamin B12. *Res. Microbiol.* 152, 291–301. doi: 10.1016/s0923-2508(01)01200-1

Li, C., Ye, Z., Wen, L., Chen, R., Tian, L., Zhao, F., et al. (2014). Identification of a novel vaccine candidate by immunogenic screening of *Vibrio parahaemolyticus* outer membrane proteins. *Vaccine* 32, 6115–6121. doi: 10.1016/j.vaccine.2014.08.077

Li, H., Xiong, X.-P., Peng, B., Xu, C.-X., Ye, M.-Z., Yang, T.-C., et al. (2009). Identification of broad cross-protective immunogens using heterogeneous antiserum-based immunoproteomic approach. *J. Proteome Res.* 8, 4342–4349. doi: 10.1021/pr900439j

Li, H., Ye, M. Z., Peng, B., Wu, H. K., Xu, C. X., Xiong, X. P., et al. (2010a). Immunoproteomic identification of polyvalent vaccine candidates from *Vibrio parahaemolyticus* outer membrane proteins. *J. Proteome Res.* 9, 2573–2583. doi: 10.1021/pr1000219

Li, N., Yang, Z., Bai, J., Fu, X., Liu, L., Shi, C., et al. (2010b). A shared antigen among *Vibrio* species: Outer membrane protein-OmpK as a versatile vibriosis vaccine candidate in Orange-spotted grouper (*Epinephelus coioides*). *Fish Shellfish Immunol.* 28, 952–956. doi: 10.1016/j.fsi.2010.02.010

Mao, Z., Yu, L., You, Z., Wei, Y., and Liu, Y. (2007). Cloning, expression and immunogenicty analysis of five outer membrane proteins of *Vibrio parahaemolyticus*. *Fish Shellfish Immunol.* 23, 567–575. doi: 10.1016/j.fsi.2007.01.004

Middleton, D. R., Paschall, A. V., Duke, J. A., and Avci, F. Y. (2018). Enzymatic hydrolysis of pneumococcal capsular polysaccharide renders the bacterium vulnerable to host defense. *Infect. Immun.* 86, 00316–00318e.

Naz, K., Naz, A., Ashraf, S. T., Rizwan, M., Ahmad, J., Baumbach, J., et al. (2019). PanRV: Pangenome-reverse vaccinology approach for identifications of potential vaccine candidates in microbial pangenome. *BMC Bioinf.* 20:123. doi: 10.1186/s12859-019-2713-9

Olaya-Abril, A., Jiménez-Munguía, I., Gómez-Gascón, L., and Rodríguez-Ortega, M. J. (2014). Surfomics: Shaving live organisms for a fast proteomic identification of surface proteins. *J. Proteomics* 97, 164–176. doi: 10.1016/j.jprot.2013.03.035

Parsons, L. M., Lin, F., and Orban, J. (2006). Peptidoglycan recognition by Pal, an outer membrane lipoprotein. *Biochemistry* 45, 2122–2128. doi: 10.1021/bi052227i

Pettis, G. S., and Mukerji, A. S. (2020). Structure, function, and regulation of the essential virulence factor capsular polysaccharide of *Vibrio vulnificus*. *Int. J. Mol. Sci.* 21:3259. doi: 10.3390/ijms21093259

Pizza, M., Scarlato, V., Masignani, V., Giuliani, M. M., Aric, B., Comanducci, M., et al. (2000). Identification of vaccine candidates against serogroup B meningococcus by whole-genome sequencing. *Science* 287, 1816–1820. doi: 10.1126/science.287.5459.1816

Pore, D., and Chakrabarti, M. K. (2013). Outer membrane protein A (OmpA) from *Shigella flexneri* 2a: A promising subunit vaccine candidate. *Vaccine* 31, 3644–3650. doi: 10.1016/j.vaccine.2013.05.100

Rehman, T., Yin, L., Latif, M. B., Chen, J., Wang, K., Geng, Y., et al. (2019). Adhesive mechanism of different *Salmonella* fimbrial adhesins. *Microb. Pathog.* 137:103748. doi: 10.1016/j.micpath.2019.103748

Rodríguez-Ortega, M. J. (2018). ""Shaving" live bacterial cells with proteases for proteomic analysis of surface proteins," in *The Surfaceome: Methods and Protocols*, eds K. R. Boheler and R. L. Gundry (New York, NY: Springer), 21–29. doi: 10.1007/978-1-4939-7553-2_2

Rodríguez-Ortega, M. J., Norais, N., Bensi, G., Liberatori, S., Capo, S., Mora, M., et al. (2006). Characterization and identification of vaccine candidate proteins through analysis of the group A *Streptococcus* surface proteome. *Nat. Biotechnol.* 24, 191–197. doi: 10.1038/nbt1179

Rollauer, S. E., Sooreshjani, M. A., Noinaj, N., and Buchanan, S. K. (2015). Outer membrane protein biogenesis in Gram-negative bacteria. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 370:20150023.

Sanober, G., Ahmad, S., and Azam, S. S. (2017). Identification of plausible drug targets by investigating the druggable genome of MDR *Staphylococcus epidermidis*. *Gene Rep.* 7, 147–153. doi: 10.1016/j.genrep.2017.04.008

Schlingmann, B., Castiglia, K. R., Stobart, C. C., and Moore, M. L. (2018). Polyvalent vaccines: High-maintenance heroes. *PLoS Pathog.* 14:e1006904. doi: 10.1371/journal.ppat.1006904

Serruto, D., Adu-Bobie, J., Capecchi, B., Rappuoli, R., Pizza, M., and Masignani, V. (2004). Biotechnology and vaccines: application of functional genomics to *Neisseria meningitidis* and other bacterial pathogens. *J. Biotechnol.* 113, 15–32. doi: 10.1016/j.jbiotec.2004.03.024

Shinoy, M., Dennehy, R., Coleman, L., Carberry, S., Schaffer, K., Callaghan, M., et al. (2013). Immunoproteomic analysis of proteins expressed by two related pathogens, *Burkholderia multivorans* and *Burkholderia cenocepacia*, during Human Infection. *PLoS One* 8:e80796. doi: 10.1371/journal.pone.0080796

Singh, R., Capalash, N., and Sharma, P. (2017). Immunoprotective potential of BamA, the outer membrane protein assembly factor, against MDR *Acinetobacter baumannii*. *Sci. Rep.* 7:12411.

Smith, S. G. J., Mahon, V., Lambert, M. A., and Fagan, R. P. (2007). A molecular Swiss army knife: OmpA structure, function and expression. *FEMS Microbiol. Lett.* 273, 1–11. doi: 10.1111/j.1574-6968.2007.00778.x

Solis, N., and Cordwell, S. J. (2011). Current methodologies for proteomics of bacterial surface-exposed and cell envelope proteins. *Proteomics* 11, 3169–3189. doi: 10.1002/pmic.201000808

The UniProt (2018). UniProt: a worldwide hub of protein knowledge. *Nucl. Acids Res.* 47, D506–D515.

Tobes, R., and Ramos, J. L. (2002). AraC-XylS database: a family of positive transcriptional regulators in bacteria. *Nucl. Acids Res.* 30, 318–321. doi: 10.1093/nar/30.1.318

Tsirigotaki, A., De Geyter, J., Šoštarić, N., Economou, A., and Karamanou, S. (2017). Protein export through the bacterial Sec pathway. *Nat. Rev. Microbiol.* 15, 21–36. doi: 10.1038/nrmicro.2016.161

Turner, J. W., Tallman, J. J., Macias, A., Pinnell, L. J., Elledge, N. C., Nasr Azadani, D., et al. (2018). Comparative genomic analysis of *Vibrio diabolicus* and six taxonomic synonyms: A first look at the distribution and diversity of the expanded species. *Front. Microbiol.* 15:1893. doi: 10.3389/fmicb.2018.01893

Tusnády, G. E., and Simon, I. (2001). The HMMTOP transmembrane topology prediction server. *Bioinformatics* 17, 849–850. doi: 10.1093/bioinformatics/17.9.849

Wang, C., Liu, Y., Li, H., Xu, W.-J., Zhang, H., and Peng, X.-X. (2012). Identification of plasma-responsive outer membrane proteins and their vaccine potential in *Edwardsiella tarda* using proteomic approach. *J. Proteomics* 75, 1263–1275. doi: 10.1016/j.jprot.2011.11.001

Wang, Q., Ji, W., and Xu, Z. (2020). Current use and development of fish vaccines in China. *Fish Shellfish Immunol.* 96, 223–234. doi: 10.1016/j.fsi.2019.12.010

Wilson, M. M., and Bernstein, H. D. (2016). Surface-exposed lipoproteins: An emerging secretion phenomenon in Gram-negative bacteria. *Trends Microbiol.* 24, 198–208. doi: 10.1016/j.tim.2015.11.006

Wimley, W. C. (2003). The versatile β-barrel membrane protein. *Curr. Opin. Struct. Biol.* 13, 404–411. doi: 10.1016/s0959-440x(03)00099-x

Wizemann, T. M., Adamou, J. E., and Langermann, S. (1999). Adhesins as targets for vaccine development. *Emerg. Infect. Dis.* 5, 395–403. doi: 10.3201/eid0503.990310

Xu, C., Kuang, H., and Xu, L. (2019). "Pathogen Immunoassay in Food," in *Food Immunoassay*, eds H. Kuang, C. Xu, and L. Xu (Singapore: Springer), 255–319. doi: 10.1007/978-981-13-9034-0_8

Xu, J., and Zhang, Y. (2010). How significant is a protein structure similarity with TM-score = 0.5? *Bioinformatics* 26, 889–895. doi: 10.1093/bioinformatics/btq066

Yang, J., Anishchenko, I., Park, H., Peng, Z., Ovchinnikov, S., and Baker, D. (2020). Improved protein structure prediction using predicted interresidue orientations. *PNAS* 117, 1496–1503. doi: 10.1073/pnas.1914677117

Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J., and Zhang, Y. (2015). The I-TASSER Suite: protein structure and function prediction. *Nat. Methods* 12, 7–8. doi: 10.1038/nmeth.3213

Yonekita, T., Morishita, N., Arakawa, E., and Matsumoto, T. (2020). Development of a monoclonal antibody for specific detection of *Vibrio parahaemolyticus* and analysis of its antigen. *J. Microbiol. Methods* 173:105919. doi: 10.1016/j.mimet.2020.105919

Yu, C., Lin, C., and Hwang, J. (2004). Predicting subcellular localization of proteins for Gram-negative bacteria by support vector machines based on n-peptide compositions. *Protein Sci.* 13, 1402–1406. doi: 10.1110/ps.03479604

Yu, N. Y., Wagner, J. R., Laird, M. R., Melli, G., Rey, S., Lo, R., et al. (2010). PSORTb 3.0: Improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. *Bioinformatics* 26, 1608–1615. doi: 10.1093/bioinformatics/btq249

Yu, Q. Q., Niu, M. Y., Yu, M. Q., Liu, Y. H., Wang, D. P., and Shi, X. M. (2016). Prevalence and antimicrobial susceptibility of *Vibrio parahaemolyticus* isolated from retail shellfish in Shanghai. *Food Control* 60, 263–268. doi: 10.1016/j.foodcont.2015.08.005

Zeng, L., Wang, D., Hu, N., Zhu, Q., Chen, K., Dong, K., et al. (2017). A novel Pan-genome reverse vaccinology approach employing a negative-selection strategy for screening surface-exposed antigens against leptospirosis. *Front. Microbiol.* 14:396. doi: 10.3389/fmicb.2017.00396

Zgurskaya, H. I., Krishnamoorthy, G., Ntreh, A., and Lu, S. (2011). Mechanism and function of the outer membrane channel TolC in multidrug resistance and physiology of enterobacteria. *Front. Microbiol.* 2:189. doi: 10.3389/fmicb.2011.00189

Zha, Z., Li, C., Li, W., Ye, Z., and Pan, J. (2016). LptD is a promising vaccine antigen and potential immunotherapeutic target for protection against *Vibrio* species infection. *Sci. Rep.* 6:38577.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Wang, Liu, Guo, Liu, Yuan, Guo and Pan. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 10 March 2021
doi: 10.3389/fmicb.2021.637771

Evaluation of *Salmonella* Serotype Prediction With Multiplex Nanopore Sequencing

*Xingwen Wu[1], Hao Luo[1], Feng Xu[1], Chongtao Ge[1], Shaoting Li[2], Xiangyu Deng[2], Martin Wiedmann[3], Robert C. Baker[1], Abigail Stevenson[1], Guangtao Zhang[1] and Silin Tang[1]**

[1] Mars Global Food Safety Center, Beijing, China, [2] Center for Food Safety, University of Georgia, Griffin, GA, United States, [3] Department of Food Science, Cornell University, Ithaca, NY, United States

OPEN ACCESS

Edited by:
Nigel Cook,
Jorvik Food & Environmental Virology Ltd., United Kingdom

Reviewed by:
James Pettengill,
United States Food and Drug Administration, United States
Christopher John Grim,
United States Food and Drug Administration, United States

***Correspondence:**
Silin Tang
Silin.tang@effem.com

Specialty section:
This article was submitted to Food Microbiology, a section of the journal Frontiers in Microbiology

Received: 04 December 2020
Accepted: 08 February 2021
Published: 10 March 2021

Citation:
Wu X, Luo H, Xu F, Ge C, Li S, Deng X, Wiedmann M, Baker RC, Stevenson A, Zhang G and Tang S (2021) Evaluation of Salmonella Serotype Prediction With Multiplex Nanopore Sequencing. Front. Microbiol. 12:637771. doi: 10.3389/fmicb.2021.637771

The use of whole genome sequencing (WGS) data generated by the long-read sequencing platform Oxford Nanopore Technologies (ONT) has been shown to provide reliable results for *Salmonella* serotype prediction in a previous study. To further meet the needs of industry for accurate, rapid, and cost-efficient *Salmonella* confirmation and serotype classification, we evaluated the serotype prediction accuracy of using WGS data from multiplex ONT sequencing with three, four, five, seven, or ten *Salmonella* isolates (each isolate represented one *Salmonella* serotype) pooled in one R9.4.1 flow cell. Each multiplexing strategy was repeated with five flow cells, and the loaded samples were sequenced simultaneously in a GridION sequencer for 48 h. *In silico* serotype prediction was performed using both SeqSero2 (for raw reads and genome assemblies) and SISTR (for genome assemblies) software suites. An average of 10.63 Gbp of clean sequencing data was obtained per flow cell. We found that the unevenness of data yield among each multiplexed isolate was a major barrier for shortening sequencing time. Using genome assemblies, both SeqSero2 and SISTR accurately predicted all the multiplexed isolates under each multiplexing strategy when depth of genome coverage ≥50× for each isolate. We identified that cross-sample barcode assignment was a major cause of prediction errors when raw sequencing data were used for prediction. This study also demonstrated that, (i) sequence data generated by ONT multiplex sequencing can be used to simultaneously predict serotype for three to ten *Salmonella* isolates, (ii) with three to ten *Salmonella* isolates multiplexed, genome coverage at ≥50× per isolate was obtained within an average of 6 h of ONT multiplex sequencing, and (iii) with five isolates multiplexed, the cost per isolate might be reduced to 23% of that incurred with single ONT sequencing. This study is a starting point for future validation of multiplex ONT WGS as a cost-efficient and rapid *Salmonella* confirmation and serotype classification tool for the food industry.

Keywords: whole genome sequencing, Oxford Nanopore Technologies, multiplex, *Salmonella*, subtyping, foodborne pathogens, serotyping, food industry

INTRODUCTION

Salmonella is a genus of rod-shaped, Gram-negative bacteria which has imposed great risk on food production and public health (GMA, 2009; Scallan et al., 2011; Oh and Park, 2017; EFSA and ECDC, 2019a,b). As one of the most common foodborne pathogens, non-typhoid *Salmonella* contributes the second most cases of infections in the U.S. (Tack et al., 2020). Contamination often results in significant financial loss throughout the food supply chain, which makes effective surveillance and control of this pathogen necessary. Although there are only two species of *Salmonella*, over 2,600 serotypes have been reported to date (Issenhuth-Jeanjean et al., 2014). Thus, accurate and rapid identification of *Salmonella* serotypes is important for efficient source tracking during incident investigation and *Salmonella* surveillance.

The widely used White-Kauffmann-Le minor scheme defines *Salmonella* serotypes mainly by 46 somatic (O) and 114 flagella (H1/H2) antigens, and an antigenic formula with specific combination of O and H antigens is used to differentiate serotypes (Grimont and Weill, 2007). Traditional serotyping methods that use this scheme involve the use of specific antisera (Herikstad et al., 2002), with some serotypes further defined by biochemical characteristics. Reliable serotyping with traditional methods is time consuming, and requires careful maintenance of a large number of different antisera (Wattiau et al., 2011; Shi et al., 2015). Several molecular methods have been developed to overcome the drawbacks of traditional serotyping methods (Foley et al., 2007; Wattiau et al., 2011; Diep et al., 2019); for example, PCR (Herrera-León et al., 2007) and microarray-based subtyping methods (McQuiston et al., 2011) have been widely used. Due to increasing utilization of next generation sequencing (NGS), *in silico* subtyping through whole genome sequencing (WGS) data is gradually becoming mainstream. Several *in silico* methods have been developed to predict *Salmonella* serotypes from WGS sequencing data (Zhang et al., 2015; Ashton et al., 2016; Yoshida et al., 2016; Zhang et al., 2019), among which SISTR (Yoshida et al., 2016), SeqSero (Zhang et al., 2015), and its recent upgrade SeqSero2 (Zhang et al., 2019) have been substantively validated and shown to deliver accurate predictions (Diep et al., 2019; Banerji et al., 2020; Cooper et al., 2020). Such methods are now being accepted and used by regulators, industry, and academia for source attribution, outbreak investigation, surveillance, and research purposes (Didelot et al., 2016; EFSA Panel on Biological Hazards et al., 2019).

Illumina sequencing platforms[1], which use short read sequencing approaches have been used commonly to provide WGS data for surveillance of foodborne pathogens, including *Salmonella* (Allard et al., 2016; Ashton et al., 2016; CDC, 2006, 2018). While Illumina sequencing has advantages of low error rate and high throughput, its short reads have limited capability for the closed assembly of genomes. Oxford Nanopore Technologies (ONT) offers comparable sequencing platforms with extra-long reads, real time sequencing, and rapid processing time[2], but has higher rates of sequencing errors than Illumina (Fox et al., 2014; Rang et al., 2018). Raw ONT reads are long enough to be treated as "contigs" by SeqSero2, and this serotype prediction tool offers two options for ONT data: (i) raw ONT reads using SeqSero2 raw reads workflow (ii) raw ONT reads using SeqSero2 assembly workflow (Zhang et al., 2019). In SeqSero2 v1.1.2[3], a single nanopore workflow is available for both raw ONT reads and their genome assemblies. This workflow algorithmically unifies the processing of both ONT raw reads and ONT assemblies by taking advantage of the long lengths of ONT reads, which are usually similar to those of contigs assembled from short reads (i.e., Illumina reads). With the serotype prediction tool SISTR, ONT raw reads have to first be assembled to correct base call errors, which have been reported to compromise SISTR prediction results (Zhang et al., 2019). Our previous evaluation of ONT sequencing based on sequencing a single *Salmonella* strain per flow cell and data analyzed by both the SeqSero2 raw reads workflow and SISTR revealed that a total of 2 h of ONT sequencing data were sufficient for successful *Salmonella* serotyping (Xu et al., 2020). This represents a considerable time saving compared to short-read-sequencing-based approaches. The ONT platforms have the capability for simultaneously sequencing multiple strains by applying DNA indexing (multiplex sequencing) (Karamitros and Magiorkinis, 2018). Combined with the continuously growing sequencing data yield of ONT sequencers, multiplex sequencing allows large quantities of complexed samples to be sequenced in one run (Piper et al., 2019; Kennedy et al., 2020). Several studies have successfully performed species identification by Multi Locus Sequence Typing (MLST) analysis from multiplexed sequencing data using the multiplex sequencing system provided by ONT (Imai et al., 2020; Liou et al., 2020). However, mis-assignment or cross-contamination of barcodes was also observed (Xu et al., 2018).

This study aimed to evaluate the efficiency, accuracy, and economic value of serotype prediction from ONT sequencing data generated from multiplexing up to ten *Salmonella* isolates. A combination of common, rare, and difficult-to-differentiate isolates were selected to (i) investigate whether correct antigenic formulae are assigned to isolates, (ii) determine whether mis-assignment or cross-contamination occurs during the process, and (iii) find the best combination of isolate number for multiplexing and the total sequencing time for the most efficient and accurate serotype prediction.

MATERIALS AND METHODS

Bacterial Strains

Ten *Salmonella* isolates representing 10 different serotypes were assessed in the current study (**Table 1**). Seven isolates represented some of the most common serotypes, which included *Salmonella* serotype Typhimurium (including Typhimurium O5−), 4,[5],12:i:-, Paratyphi B (dt +), Enteritidis, Mbandaka, and

[1]https://www.illumina.com/systems/sequencing-platforms.html

[2]https://nanoporetech.com/products

[3]https://github.com/denglab/SeqSero2

TABLE 1 | *Salmonella* strains used for evaluating the performance of multiplex ONT sequencing for serotype prediction from whole genome sequencing data.

No.	Category	Strain ID	Recorded serotype[a]	Antigenic profile[b]	Source	Barcode no.[c]
1	Most common	FSL S5-0483	Enteritidis	1,9,12:g,m:-	Human	07
2		FSL R8-4405	Mbandaka	6,7,14:z10:e,n,z15	Human, clinical	02
3		FSL S5-0658	Senftenberg/Dessau	1,3,19:g,[s],t:-	Human	03
4	Common and difficult to differentiate serovars (genetically similar)	FSL R9-3346	Typhimurium	1,4,[5],12:i:1,2	Human, clinical	01
5		FSL R8-3714	Typhimurium ($O5^-$)	1,4,[5],12:i:1,2	Human, clinical	05
6		FSL S5-0580	4,[5],12:i:-	4,[5],12:i:-	Bovine	04
7	Common, and difficult to predict serovar correctly	FSL S5-0447	Paratyphi B (dt +)	1,4,[5],12:b:1,2	Human	08
8	Species and subspecies that are distinct from common *Salmonella* subspecies (i.e., *S. enterica* subspecies *enterica*	FSL R9-0514	*S. enterica* subsp. *Salamae*	55:z39:-	–	10
9		FSL R9-0518	*S. bongori*	66:z41:-	ATCC 43975	11
10	Rare	FSL S5-0549	Havana	1,13,23:f,g,[s]:-	Bovine	09

[a]*Food Safety Lab, Cornell University, provided the information of recorded serotype for each isolate, which was determined by conventional serotyping method with Kauffman-White scheme*

[b]*We listed the antigenic profiles of the strains in use based on the serotype information from "Antigenic Formulae of the Salmonella Serovars 2007 9th edition" (Grimont and Weill, 2007)*

[c]*The barcode No. 01–11 correspond to the barcode RB01–RB11 in the ONT Rapid Barcoding Sequencing kit (SQK-RBK004).*

Senftenberg. These were selected from i) the top 20 serotypes reported by the U.S. national *Salmonella* surveillance system (CDC, 2006, 2018) and ii) the top 20 serotypes causing human infections worldwide as reported to the WHO (Galanis et al., 2006). One isolate representing a rare serotype (serotype Havana) found in the food industry (information obtained by personal communication) was also included. Six isolates represented six serotypes that may be difficult to identify with molecular-level serotyping methods such as MLST and Pulsed-Field Gel Electrophoresis (PFGE). These six serotypes included (i) serotype Typhimurium, its $O5^-$ variant, and serotype 4,[5],12:i:-, which are difficult to differentiate by MLST and phylogenetic analysis due to their close relatedness (Ranieri et al., 2013); (ii) serotype Paratyphi B (dt +), which has been incorrectly predicted by other molecular subtyping methods such as PFGE (Bailey et al., 2002; Soyer et al., 2010; Zou et al., 2012), and (iii) one strain of *S. enterica* subspecies *salamae*, and one strain of *S. bongori*; these two strains represent subspecies and species distinct from the most common *Salmonella* subspecies (i.e., *S. enterica* subspecies *enterica*). Detailed isolate information, including all sequence data associated with a given isolate, can be found at www.foodmicrobetracker.com under the isolate ID (e.g., FSL S5-0393).

Genomic DNA Extraction

All *Salmonella* isolates were cultured on Trypticase Soy Agar at 37°C for 20∼22 h. Genomic DNA was extracted from single colonies using the QIAamp DNA mini kit (Qiagen, Hilden, Germany). The DNA quality and quantity were assessed with the NanoDrop 2000 (Thermo Fisher Scientific, Delaware, United States) for absorbance value (A) and the Qubit 3.0 fluorimeter (Life Technologies, Paisley, United Kingdom) for double strand DNA quantity, based on the guidance for qualification requirements for successful sequencing provided by ONT. The genomic DNA samples that met the following criteria were used for library construction: (i) A 260/280 between 1.8 and 1.9; (ii) A 260/230 between 2.0 and 2.2. The total input DNA was about 400 ng for each flow cell (FC).

Oxford Nanopore Library Preparation and Sequencing

The rapid Barcoding Sequencing kit (SQK-RBK004) was used according to the manufacturer's instructions. Libraries were multiplexed and sequenced with qualified FLO-MIN106 flow cells (R9.4.1, active pore number ≥800) for 48 h on GridION (Oxford Nanopore Technologies, Oxford, United Kingdom) following the workflow described in **Figure 1**. Basecalling was performed in real time using Guppy with a basecalling model modified for 6 mA dam/5 mC dcm and CpG, which was integrated in the MinKNOW software v3.5.40 installed on GridION. Ten barcodes were assigned to ten isolates (**Table 1**). Five multiplexing strategies were tested, including multiplexing three (barcode No. 01 ∼ No. 03), four (barcode No. 01 ∼ No. 04), five (barcode No. 01 ∼ No. 05), seven (barcode No. 01 ∼ No. 05 and No. 07 ∼ No. 08), or ten (barcode No. 01 ∼ No. 05 and No. 07 ∼ No. 11) isolates on one flow cell. Repeats of each multiplexing strategy were tested on five flow cells, which were sequenced simultaneously on one GridION.

Genomics Analysis and Data Distribution Analysis

Basecalled reads were demultiplexed using qcat v1.1.0[4], the argument "--min-score" was set to a default value of 60 for general analysis, and various values were manually defined for the

[4]https://github.com/nanoporetech/qcat

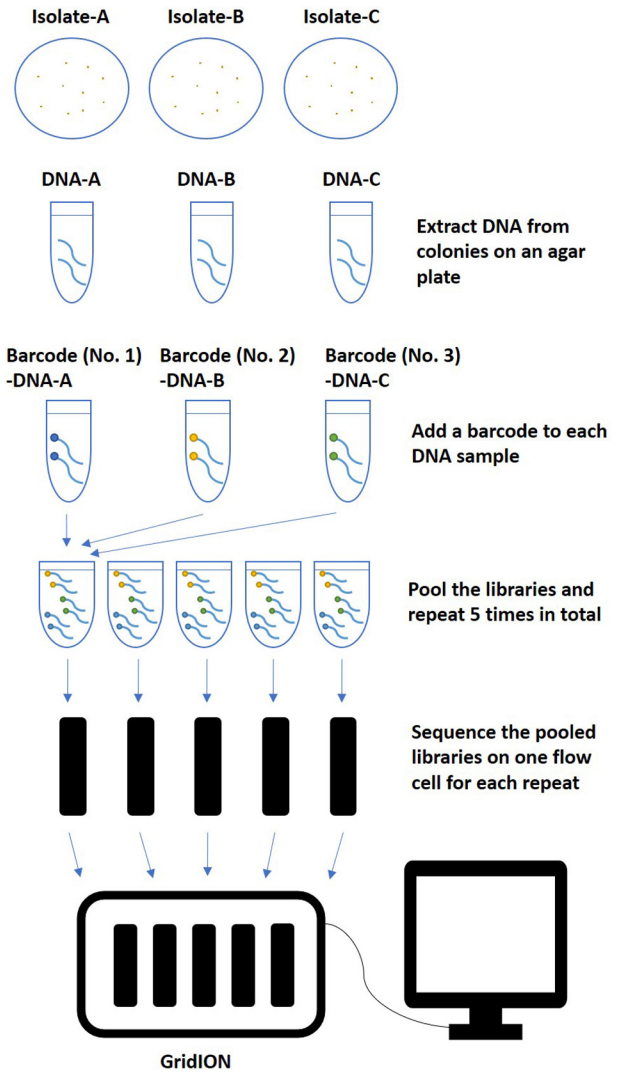


FIGURE 1 | An example of the workflow of multiplex ONT sequencing library preparation and sequencing. After extracted from colonies on an agar plate, *Salmonella* genome DNA of three isolates were multiplexed and sequenced in one flow cell with five repeats. The same workflow was also used for multiplexing four, five, seven, or ten isolates in one flow cell.

purposes of error analysis. Demultiplexed reads were classified based on the barcode qcat identified, reads of each barcode were applied to a modified genomic analysis workflow derived from our previous study (Xu et al., 2020). Porechop v0.2.3[5] was used to trim adaptor sequences, and the quality of trimmed data was assessed using NanoPlot v1.18.1 (De Coster et al., 2018). A total of approximately 500 Mbp of high-quality long reads (minimum length > 1,000 bp) were selected through Filtlong v0.2.0[6]. Filtered raw reads were subsequently applied to an initial serotype prediction by SeqSero2 v1.0.2[7] under a preset of arguments specifically designed for ONT raw data. Wtdbg2 v2.4 (Ruan and Li, 2020)[8] was then used for initial genome assembly followed by one round of correction through Racon v1.3.3 (Vaser et al., 2017)[9].

To assess the influence of barcode and isolate (as one barcode was always assigned to one isolate in this study, we combined these two factors into one as the barcode/isolate factor) on the distribution of data yield among multiplexed isolates in each flow cell, a one-way analysis of variance (ANOVA) was carried out to compare if the proportion of data yield of each barcoded isolate was significantly different from the other barcoded isolates multiplexed in the same flow cell. As the data yield per isolate was diverse within one flow cell, we calculated the coefficient of variation for the proportion of data yield per multiplexed isolate in each flow cell to compare the degree of diversity of data yield per multiplexed isolate between different multiplexing strategies. A one-way ANOVA was carried out to compare if this coefficient of variation for a given multiplexing strategy (e.g., three isolates) was significantly different from the other multiplexing strategies (e.g., four, five, seven, and ten isolates). Statistical significance for the ANOVA was assigned at $\alpha = 0.05$.

Serotype Prediction Analysis

The nanopore workflow available in SeqSero2 v1.1.2 (see text footnote 3) was used for genome assemblies from ONT reads for serotype prediction analysis. To avoid any possible impact of barcode cross-assignment on serotype prediction, the nanopore workflow of SeqSero2 v1.1.2 was not used on ONT raw reads for the purpose of serotype prediction. Instead, it was used to detect cross-assigned reads. ONT raw reads were also first assembled to correct basecall errors that were reported to compromise SISTR prediction (Zhang et al., 2019) and then used for further serotype prediction using SISTR_cmd (The *Salmonella in silico* Typing Resource Command-line Tool) v1.1.0 (Yoshida et al., 2016) under a default setting of arguments for assembled data. A brief data analysis pipeline for serotype prediction is shown in **Figure 2**.

Assessment of the Influence of Sequencing Time and Depth of Genome Coverage in Serotype Prediction

To assess the influence of sequencing depth and sequencing time on the accuracy of serotype prediction, serotype prediction analyses were carried out for each flow cell using different sizes of sequencing data collected as follows. The raw sequencing reads were arranged sequentially by their time of generation. For each flow cell, we first identified the isolate that accounted for the least amount of sequencing data, then we collected reads up to the time point when this isolate obtained the desired depth of genome coverage [average *Salmonella* genome size: 4.8 Mbp (McClelland et al., 2001; Parkhill et al., 2001)]. We defined this depth of genome coverage as $\text{Depth}_{\text{min}}$ for this flow cell. For example, in FC01 (flow cell 01), sequencing data yield for isolates with BC01 (barcode 01), BC02, and BC03 after 48 h of ONT sequencing accounted for 29.36%, 34.95%, and 27.85% of the total

[5]https://github.com/rrwick/Porechop
[6]https://github.com/rrwick/Filtlong
[7]https://github.com/denglab/SeqSero2/releases
[8]https://github.com/ruanjue/wtdbg2/releases/tag/v2.4
[9]https://github.com/isovic/racon/releases

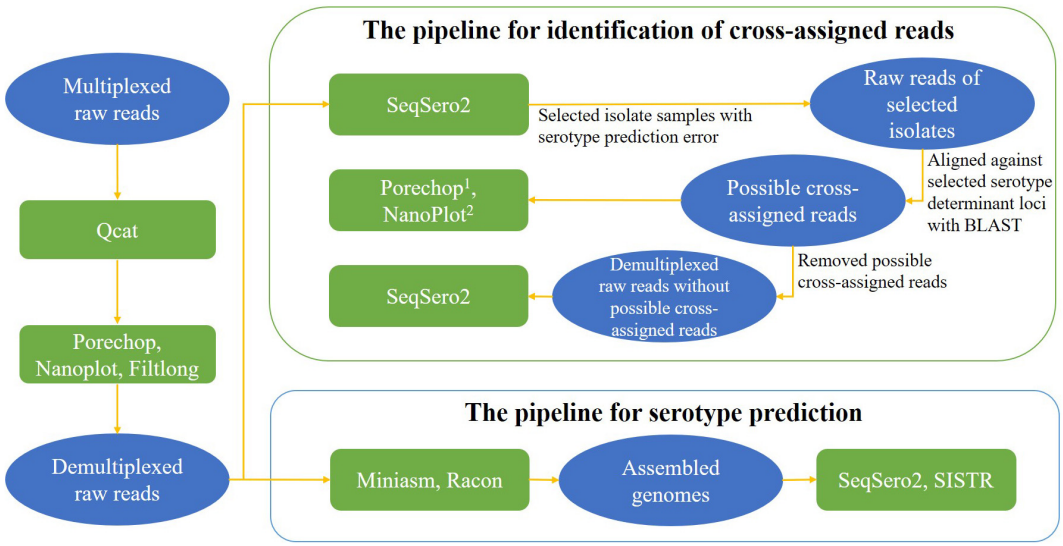


FIGURE 2 | Data analysis pipelines for serotype prediction and identification of cross-assigned reads. Contents in ovals present the type of input data for each analysis, contents in rectangles present the tools for each analysis. [1]Porechop was used to search for middle adaptors from the possible cross-assigned reads; [2]NanoPlot was used to assess the sequencing quality of the possible cross-assigned reads.

sequencing data, respectively. If the desired $Depth_{min}$ of FC01 for serotype prediction analysis was 15×, we collected sequencing data until the isolate with BC03 reached 15×depth of genome coverage (15 × 4.8 Mbp = 72 Mbp) since BC03 had the least share of sequencing data (27.85%) in this case.

Identification of Possible Cross-Assigned Reads and Influence of Cross-Assigned Reads on the Accuracy of Serotype Prediction

To further investigate if there were cross-assigned reads among multiplexed isolates on each flow cell, ONT raw reads was used as input of SeqSero2 to detect prediction errors caused by single ONT reads. Raw sequencing reads (at 99×depth of genome coverage) of the isolate that was predicted as a different serotype as recorded were aligned against selected serotype determinant loci using BLAST[10]. These antigen determinant loci were selected from the SeqSero2 antigen outputs that did not match the recorded antigen profile of this isolate. For example, the antigen profile prediction result using raw reads of isolate FSL R9-3346 from sample FC19-BC03 (recorded serotype: Senftenberg) showed its H2 antigen as "e,n,z15," while serotype Senftenberg did not have an H2 antigen. We therefore included all four alleles of the H2 antigen "e,n,z15" from SeqSero2 database to align against the raw reads of FC19-BC03 for identifying possible cross-assigned reads. We defined a read as a possible cross-assigned read if it showed BLAST identity ≥90% and coverage = 100% to a selected antigen determinant allele.

Porechop v0.2.3 (see text footnote 5) was used to further assess whether there was a barcode in the middle of these reads (we choose a -middle_threshold of 60). Sequencing quality of these putative cross-assigned reads was assessed using NanoPlot v1.18.1 (De Coster et al., 2018). To assess the influence of the possible cross-assigned reads on the accuracy of serotype prediction, serotype prediction analyses with SeqSero2 were performed again after removal of these reads for the aforementioned two isolates. A brief data analysis pipeline for identification of cross-assigned reads is shown in **Figure 2**.

Assessment of the Influence of Sequencing Quality and Barcode Quality on the Accuracy of Serotype Prediction

To further assess the influence of (i) sequencing quality and (ii) barcode quality on the accuracy of serotype prediction for the two isolates FSL R9-3346 and FSL S5-0658 from the two samples noted in 2.4.2, serotype prediction analyses with SeqSero2 were performed with the raw reads selected with the following settings as input: (i) sequencing quality score ≥12, 13, or 14 (selected by Filtlong v0.2.0), qcat barcode score ≥60 (default setting), and $Depth_{min}$ = 99×, as well as (ii) sequencing quality ≥7 (default setting), qcat barcode score ≥70, 80, or 90, and $Depth_{min}$ = 99×.

RESULTS AND DISCUSSION

Overview of Nanopore Sequencing Data and Assembly of *Salmonella* Genomes

The quality of raw sequencing data from the multiplex ONT sequencing experiments was analyzed by NanoPlot (version 1.18.1) (**Table 2**). An average of 10,629 Mbp of clean sequencing data per flow cell were obtained in 48 h from 25 experiments (across different numbers of multiplexed isolates). Data outputs of flow cells ranged from 5,650 to 17,998 Mbp, with the exception of FC17 which generated only 3,201 Mbp data; the mean read length and read length N50 varied much less (**Table 2**).

[10]https://blast.ncbi.nlm.nih.gov/Blast.cgi

TABLE 2 | Summary of quality statistics of multiplexed raw sequencing data.

Flow cell ID	Number of isolates multiplexed	Total clean data yield (Mbp)	Mean read length (bp)	Mean quality score	Number of reads	Read length N50
01	3	13,787	10,368	11.7	1,329,726	18,407
02	3	12,312	8,025	11.5	1,534,118	14,802
03	3	13,694	7,583	11.2	1,805,906	15,665
04	3	11,085	8,207	11.3	1,350,613	15,284
05	3	8,135	8,385	11.3	970,123	15,235
06	4	14,463	8,305	11.7	1,741,442	14,596
07	4	6,332	8,017	11.7	789,817	14,266
08	4	17,100	8,283	11.8	2,064,501	14,731
09	4	15,234	8,226	11.2	1,851,964	15,137
10	4	17,998	7,658	11.8	2,350,186	13,968
11	5	13,715	7,832	10.8	1,751,213	14,332
12	5	7,181	8,018	11.5	895,600	14,680
13	5	8,761	7,600	11.3	1,152,794	14,648
14	5	10,530	7,583	11.1	1,388,584	14,689
15	5	6,833	7,434	11.6	919,206	13,320
16	7	13,513	7,167	12.2	1,885,531	13,479
17	7	3,201	7,831	12.1	408,795	14,618
18	7	8,450	7,542	12.1	1,120,392	14,249
19	7	12,113	7,706	12	1,571,937	14,447
20	7	5,984	8,212	12.3	728,703	14,714
21	10	5,650	9,429	12.1	599,172	17,944
22	10	9,666	9,851	12	981,235	18,013
23	10	7,794	9,370	12.3	831,812	17,038
24	10	10,592	9,770	12.1	1,084,229	18,057
25	10	11,603	10,309	12.2	1,125,520	18,549
Average		10,629.0	8,348.4	11.7	1,289,324.8	15,394.7

Sequencing quality was shown to be highly consistent among 25 experiments, with mean quality scores for a given flow cell ranging from 10.8 to 12.3. The average quality score was 11.7.

Overview of Demultiplexed Sequencing Data

An average of 8.02% (N = 25) reads per flow cell failed to be assigned to any barcode after demultiplexing analysis of the sequencing data generated with 48 h of ONT sequencing; these reads were defined as Non-assigned reads. An average of 0.03% reads per flow cell were assigned to barcodes that were not included in the flow cell; these reads were defined as mis-assigned reads (**Figure 3** and **Supplementary Table 1**). Demultiplexed raw sequencing data of each multiplexed isolate in each flow cell from 48 h ONT sequencing (No. = 145) were submitted to NCBI–SRA (Accession number: PRJNA694442).

The unevenness of data yield among multiplexed isolates was a major barrier for shortening the total sequencing time for each flow cell. There was no significant difference among the data yields for different barcoded samples after 48 h of ONT sequencing when multiplexing three or four isolates (**Figures 4A,B** and **Supplementary Table 1**). However, ANOVA indicated a significant difference ($P < 0.05$) in data yields among isolates when multiplexing five, seven, or ten isolates (**Figures 4C–E** and **Supplementary Table 1**). For example, *post hoc* comparisons using the Tukey HSD test indicated that BC03 showed a significantly lower data yield ($P < 0.05$) than some other isolates when multiplexing five or seven isolates (**Figures 4C,D**), and BC07 showed a significantly lower data yield ($P < 0.05$) than some other isolates when multiplexing seven or ten isolates (**Figures 4D,E**).

We also found that that the level of unevenness of data yield of each multiplexed isolate is different among multiplexing strategies (**Figure 5**). Comparison of the coefficient of variation of the proportion of data yield per multiplexed isolate among multiplexing strategies showed that, the level of isolate-data-yield variation of multiplexing seven or ten isolates is significantly greater ($P < 0.05$) as compared to multiplexing three, four, or five isolates. For example, on FC01 (three isolates multiplexed), the isolate (BC02) with the highest data yield accounted for 34.95% of total data (at 1,004×depth of genome coverage), which was 1.25 times the data yield of the isolate (BC03) with the smallest data sharing proportion (BC03 with 27.85% data yield at 800 × depth of genome coverage). On the other hand, for FC22 (ten isolates multiplexed), the isolate with the highest data

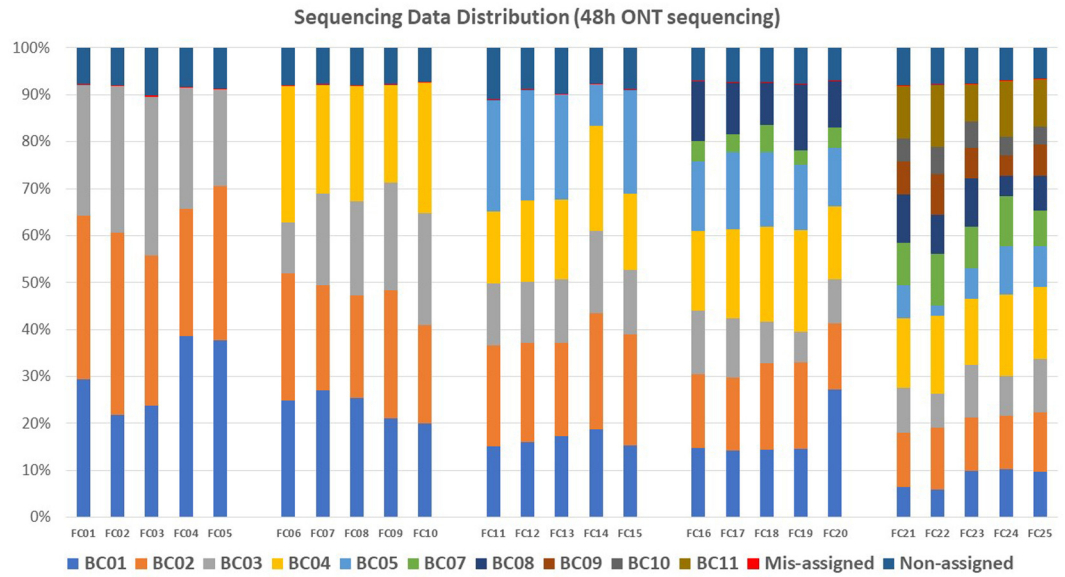


FIGURE 3 | Sequencing data distribution within each flow cell (FC) for each barcode (BC). Each color represents one barcode. Within each flow cell, sequencing reads that were not assigned to any barcode were defined as non-assigned reads, reads assigned to a barcode that was not included in the flow cell were defined as mis-assigned reads.

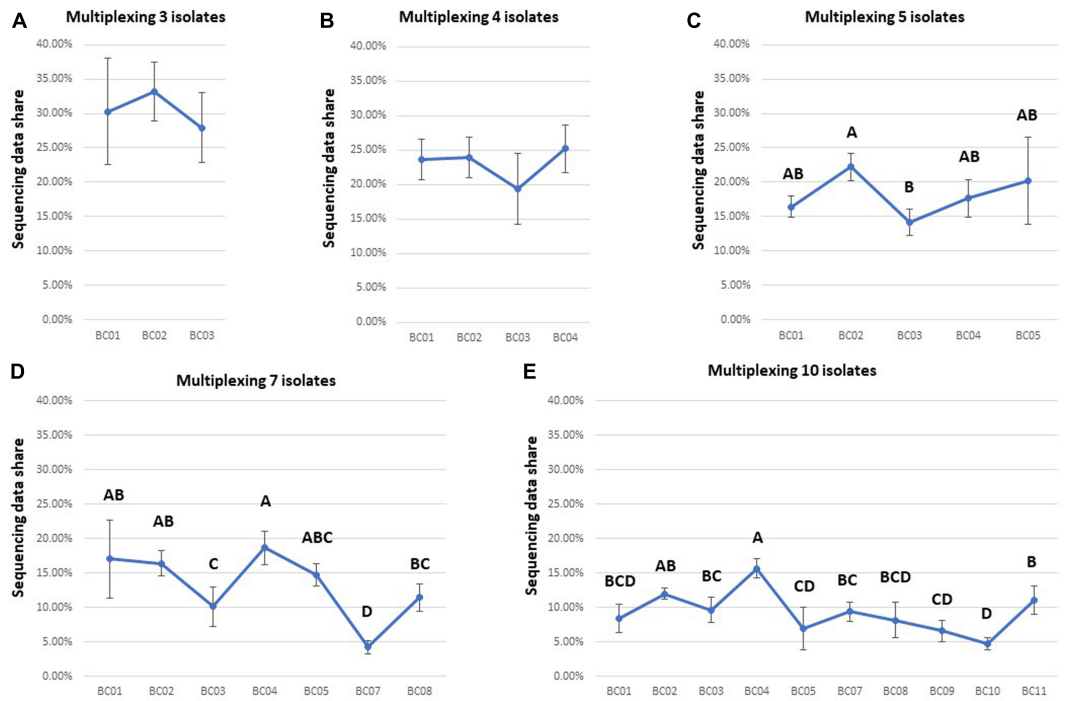


FIGURE 4 | Comparison of the data yield of each isolate multiplexed in one flow cell for each multiplexing strategy (sequencing time: 48 h). Based on the statistical analysis results of Tukey HSD test, barcodes with different letters were significantly different (P-value < 0.05) from each other in data yield. Ideally, data share of each isolate was expected to be 31, 23, 18, 13, and 9% in **(A–E)**.

yield (BC04) accounted for 16.72% of the total data (at 337 × depth of genome coverage); thus was 7.92 times the data yield of the isolate with the smallest data sharing proportion (BC05), which accounted for only 2.11% of the total data (at 42 × depth of genome coverage), after 48 h ONT sequencing.

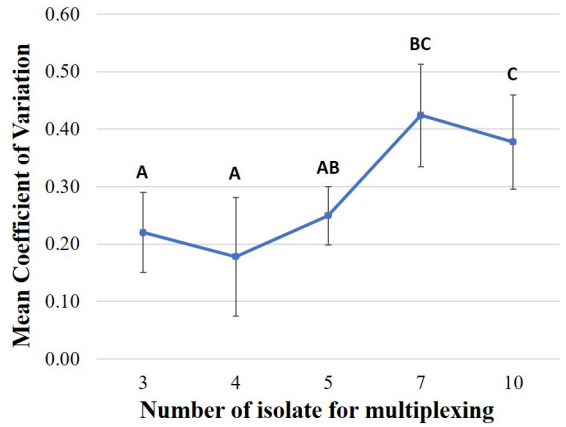


FIGURE 5 | Comparison of the coefficient of variation of the proportion of data yield per multiplexed isolate in one flow cell under each multiplexing strategy including multiplexing three, four, five, seven, or ten isolates (Sequencing time: 48 h). Based on the statistical analysis results of Tukey HSD test, multiplexing strategies with different letters were significantly different (*P*-value < 0.05) from each other in coefficient of variation.

Influence of Sequencing Time and Depth on Accuracy of *Salmonella* Serotype Prediction Using ONT Assembled Genomes

For flow cells multiplexing three, four, or five isolates, five levels of $Depth_{min}$ (15×, 30×, 50×, 75×, 99×) were used to perform serotype prediction analysis. Using genomes assembled from the raw reads with (i) sequencing quality ≥7 and (ii) qcat barcode score ≥60, both SeqSero2 and SISTR generated the same results at the serotype level as the recorded serotype of each isolate for all tested $Depth_{min}$ levels when multiplexing up to five isolates. The ONT sequencing time for each flow cell at different $Depth_{min}$ levels are shown below (**Table 3**).

For flow cells multiplexing seven or ten isolates, none of the flow cells achieved $Depth_{min}$ = 99 × within 48 h of ONT sequencing. Hence $Depth_{min}$ levels 15×, 30×, 50×, 75×, and/or the maximum $Depth_{min}$ of each flow cell were used to perform serotype prediction analysis. One flow cell multiplexing seven isolates (FC17) generated only 3.20 Gbp trimmed data after 48 h of sequencing, which was 70% lower than the average total sequencing data; this flow cell was not included in the serotype prediction analyses. When multiplexing seven isolates, earliest accurate serotype predictions for tested isolates were obtained using sequencing data at $Depth_{min}$ = 15×. When multiplexing ten isolates, SISTR identified isolate FSL S5-0483 (recorded serotype and antigen profile: Enteritidis, 1,9,12:g,m:-) as Gallinarum | Pullorum (predicted antigen profile: 1,9,12:-:-) in one (FC24) of the five replicates/flow cells, while SeqSero2 reported the correct serotype (with a depth of genome coverage of 41 × for this isolate in FC24). Earliest accurate serotype predictions for tested isolates were obtained using sequencing data at $Depth_{min}$ = 15 × with SeqSero2 alone, and at $Depth_{min}$ = 30 × with both SeqSero2 and SISTR.

TABLE 3 | ONT sequencing time for different levels of $Depth_{min}$.

$Depth_{min}$ level	Sequencing time (hour)					Average Sequencing Time (hour)
	Multiplexing three isolates					
	FC01	***FC02***	***FC03***	***FC04***	***FC05***	
99×	3.84	5.80	5.56	5.56	6.77	5.50
75×	2.93	4.41	4.25	4.25	5.09	4.19
50×	2.05	3.02	2.94	2.96	3.47	2.89
30×	1.31	1.93	1.90	1.90	2.19	1.85
15×	0.75	1.06	1.04	1.04	1.20	1.02
	Multiplexing four isolates					
	FC06	***FC07***	***FC08***	***FC09***	***FC10***	
99×	10.36	11.38	5.37	6.71	5.25	7.81
75×	7.81	8.43	4.10	5.10	3.99	5.89
50×	5.29	5.56	2.83	3.47	2.72	3.97
30×	3.27	3.34	1.82	2.14	1.72	2.46
15×	1.71	1.71	1.01	1.12	0.92	1.29
	Multiplexing five isolates					
	FC11	***FC12***	***FC13***	***FC14***	***FC15***	
99×	8.81	13.61	12.25	11.70	15.62	12.40
75×	6.66	9.84	9.01	8.71	11.35	9.11
50×	4.52	6.35	5.92	5.81	7.46	6.01
30×	2.79	3.69	3.62	3.58	4.50	3.64
15×	1.49	1.79	1.87	1.90	2.40	1.89
	Multiplexing seven isolates					
	FC16	***FC17***	***FC18***	***FC19***	***FC20***	
99×	33.35	–	>48	>48	>48	–
75×	21.75	–	24.09	41.82	>48	–
50×	13.62	–	12.99	20.37	31.18	19.54
30×	7.94	–	6.91	10.82	9.05	8.68
15×	3.96	–	3.35	5.16	4.00	4.12
	Multiplexing ten isolates					
	FC21	***FC22***	***FC23***	***FC24***	***FC25***	
99×	>48 h	>48 h	>48 h	>48 h	>48 h	–
75×	>48 h	>48 h	27.52	>48 h	31.01	–
50×	26.58	>48 h	13.20	20.81	17.06	–
30×	10.08	21.85	6.51	11.66	9.26	11.87
15×	4.13	9.54	3.16	5.60	4.41	5.37

While molecular serotyping does not typically identify serotype variants caused by ancillary O antigens, Typhimurium var. O5- (FSL R8-3714) was consistently (across flow cells) correctly identified by SeqSero2 [as SeqSero2 targets a mutation that causes the O5- phenotype (Hauser et al., 2011)]; SISTR, on the other hand, identified FSL R8-3714 as Typhimurium, without specifying the variant O5-.

In summary, for all multiplexing strategies tested in the current study, (i) with SeqSero2 alone, the earliest accurate serotype predictions were obtained using genomes assembled from sequence data at $Depth_{min}$ 15 × as input, generated from 1.02 to 5.47 h of ONT sequencing for multiplexing three to ten isolates; (ii) with SISTR alone, the earliest accurate serotype predictions were obtained using genomes assembled from sequence data $Depth_{min}$ at 30 × as input, generated from 1.85 to 11.87 h for multiplexing three to ten isolates (**Table 3**); (iii) both SISTR and SeqSero2 correctly predicted *Salmonella* serotypes from genomes assembled using multiplexed ONT raw data with 50 × genome coverage (please note that this is an actual coverage of a given isolate, not a $Depth_{min}$ indicating the lowest genome coverage that each one of the multiplexed isolates can reach on one flow cell).

Higher quality of the assembled genome was shown to be associated with higher sequencing depth for both the WGS and ONT platforms in our previous study (Xu et al., 2020), while here we demonstrated a genome coverage of 50× is generally sufficient for assembling genomes to support accurate serotyping. An average sequencing duration of approximately 6 h for multiplexing five isolates was sufficient to reach this genome coverage. The majority of previous evaluations of WGS-based *Salmonella* serotype prediction used Illumina data (Zhang et al., 2015, 2019; Yachison et al., 2017; Uelze et al., 2020), since Illumina platforms can generally yield high quality sequencing data. Currently, a minimum of 30 × genome coverage (≥10 kb reads) of ONT sequencing data is recommended for bacteria assembly[11].

Multiplexing several isolates inevitably caused variations in size of data allocation for each isolate, unlike sequencing a single isolate on a single flow cell. It is thus important to manage multiplex sequencing to allow each isolate to reach at least the genome coverage of 50 × for reliable serotype prediction, hence sequencing duration for multiplex runs will necessarily be longer than the sequencing duration required for a single isolate. Our results from pooling 3–10 isolates demonstrated that the unevenness of data yield among multiplexed isolates increased significantly and sequencing times required for accurate serotyping exceeded 19 h when multiplexing more than five isolates. Consequently, multiplexing more than five isolates may potentially undermine the benefit of the relatively short time required for ONT sequencing. Based on these observations, we conclude that multiplexing five isolates represents the optimum for obtaining the sequencing depth required for reliable serotype prediction within a reasonable time frame.

Identification of Possible Cross-Assigned Reads and Their Influence on the Accuracy of Serotype Prediction

One and two possible cross-assigned reads were identified for samples FC10-BC01 (isolate FSL R9-3346; serotype: Typhimurium) and FC19-BC03 (isolate FSL S5-0658; serotype: Senftenberg), respectively (**Supplementary Table 2**). Please note that the analysis method we used could only identify possible cross-assigned reads leading to serotype prediction errors. Therefore, it was highly possible that the cross-assigned reads we identified were just a small fraction of all types of existing cross-assigned reads. The mean read length and sequencing quality of these reads were 10.96 Kbp and 12.93 Kbp, respectively. We did not detect any middle adaptors among these reads. These possible cross-assigned reads each had a barcode score ≥95.8 at one end of the read (**Supplementary Table 2**), and < 90 at the other end, as determined by Porechop (data not shown), consistent with the fact that the ONT library preparation kit added barcode adaptors to only one end of each read. Serotype predictions with raw reads by SeqSero2 after removal of these possible cross-assigned reads, were consistent with recorded serotypes of these samples, suggesting that these reads were the cause of serotype prediction errors for these two isolates. In addition, we did not identify any possible cross-assigned reads with our screening criteria from the corresponding isolates on the other flow cells multiplexing the same number of isolates but showing correct serotype predictions. As the barcode score and quality score of these detected cross-assigned reads were quite high and as no chimeric reads were detected (no evidence of middle adaptor), we speculated that one of the major causes of barcode cross-assignment was contamination from free adapters after pooling the libraries. These reads possibly captured wrong barcodes after multiplexing during library preparation. We did not perform cleaning-up to remove short sequences (< 100 bp) after pooling, hence these free barcodes might have had opportunities to link to the DNA sequences from multiple isolates. Tyler et al. (2018) also found erroneously barcoded reads that standard filtering practices could not remove as they were of high quality, and suggested running only a single sample at a time on a flow cell to avoid contaminating reads. It has been reported that clean-up by a bead-based or gel purification step could remove free adapters[12], therefore a clean-up step after pooling the indexed libraries might be added in future studies. This may help alleviate index misassignment.

This study showed that cross-assigned reads might cause serotyping errors only when raw reads were directly used for serotype prediction. However, such errors may be avoided by assembling genomes prior to prediction. Barcode misassignment (including cross-assignment also known as index hopping) between multiplexed libraries is a recognized cause of misidentification (Kircher et al., 2012) for NGS. The Illumina sequencing technology has been reported to typically generate 0.1–2% barcode misassignment (see text footnote 12). Similarly, ONT sequencing has been reported to generate 0.02–0.3% barcode switching or misassignment to an unused barcode when using the 1D ligation sequencing kit (SQK-LSK108) and the native barcoding expansion kit (EXP-NBD103) (Tyler et al., 2018; Wick et al., 2018; Xu et al., 2018). In the current study, using the Rapid Barcoding Sequencing kit (SQK-RBK004), the index misassignment level was around 0.04% of the total sequencing data size (data not shown). Barcode misassignment levels for

[11] https://nanoporetech.com/sites/default/files/s3/literature/microbial-genome-assembly-workflow.pdf

[12] https://www.illumina.com/science/education/minimizing-index-hopping.html

multiplex sequencing have shown to be highly dependent on the library preparation kit used, as well as quality and handling of the library (see text footnote 12). As the mean read length of the 25 ONT flow cells tested in the current study was above 8 Kbp (**Table 2**) and the length of an antigen determent loci in *Salmonella* is usually between 0.1 and 5 Kbp (data not shown), a small number of ONT cross-assigned reads can possibly alter the serotype prediction result. High molecular weight DNA extraction methods are available for purification of DNA in the 50 Kbp to 1 Mbp + size range. Reads with greater mean length may lead to higher quality of genome assembly, yet their influence on the possible cross-assigned reads and accuracy of serotype prediction are unknown. With DNA extraction method generating regular read length of genome DNA (mean read length around 8 Kbp), we recommend using assembled genomes for serotype prediction through SeqSero2 with multiplex ONT sequencing data, if more rapid serotype prediction enabled by raw reads input [∼5 s/million bases using SeqSero2 nanopore workflow (Zhang et al., 2019)] is not the primary concern.

Influence of Sequencing Quality and Barcode Quality on the Accuracy of Serotype Prediction

Accurate serotype prediction of the two isolates noted above (FSL R9-3346, FC10-BC01, serotype Typhimurium, and isolate FSL S5-0658, FC19-BC03, serotype Senftenberg) was obtained when sequencing quality score was raised to ≥13 and qcat barcode score was set at ≥60 (default setting) using ONT raw reads as input of SeqSero2. However, this approach removed about 60% of the raw sequence data. Setting the quality score to ≥13 may have removed most of the error-causing reads from the data set, as the average sequencing quality score of the possible cross-assigned reads was 12.93. When the sequencing quality score was set to ≥14, more than 90% of the raw sequence data were lost. Consequently, the depth of genome coverage of some of the other isolates multiplexed in the same flow cell dropped to below 10×, which led to serotype prediction errors for these isolates due to low genome coverage.

When sequencing quality score was set at 7 (default setting) for filtering raw reads, and qcat barcode score was raised to ≥90 at the same time, accurate serotype prediction was obtained for these two samples by using raw reads with SeqSero2. However, these settings still resulted in loss of more than 90% of the raw sequence data, and reduction of the depth of genome coverage of some of the other isolates in the same flow cell to below 10×, again leading to serotype prediction errors for these isolates.

In summary, raising the quality score of raw reads to ≥13 improved serotype prediction accuracy for the isolates tested in the current study, while removing more than half of the total sequencing data. On the other hand, raising the barcode score of raw sequencing reads did improve prediction accuracy, but resulted in exclusion of 90% of the raw sequencing data, thus introducing errors due to lack of sequencing depth for some of the isolates multiplexed in the same flow cell. This approach could be used to avoid errors caused by cross-assigned reads when sequencing data depth permits, or where sequencing time length is not the primary concern. Xu et al. (2018) found that chimeric reads and low-barcode-score reads were the main causes of cross sample contamination in their data set. In contrast, we found the error-causing reads were of high barcode quality and without evidence of having internal adaptors. Removing all the reads with middle adaptors from the raw sequencing data did not alter the prediction results generated by using ONT raw reads (data not shown). This discrepancy in the cause of cross contamination may be attributed to the difference in library preparation and barcoding kits used in the current study and Xu's study (Xu et al., 2018).

Recommendation for Cost-Effective Multiplexing Strategy and Limitations of the Current Study

Multiplexing has the advantage of higher time and cost efficiencies compared to sequencing single isolates, particularly in a practical scenario where large numbers of samples are routinely analyzed. The sequencing kit (RBK004) used in this study has the capability of multiplexing up to 12 different isolates in one flow cell, which will yield considerable costs savings relative to sequencing a single isolate on a flow cell. Multiplexing three isolates could reduce consumable cost per isolate by 64% compared to the cost of sequencing a single isolate, and the cost reduction can be increased to 87% when ten isolates are multiplexed. However, the total sequencing time needs to be increased above the 1∼2 h required for single isolate sequencing,s

TABLE 4 | Summary of turnaround time and cost for each multiplexing strategy.

Multiplexing isolate number	1	3	4	5	7	10
Cost per isolate[a]	$910	$330	$255	$210	$159	$120
Average sequencing time to obtain at least 50 × depth of genome coverage for each multiplexed *Salmonella* isolate	<1.0 h	2.9 h	4.0 h	6.0 h	19.5 h	>25.1
Data analysis time	<1.0 h	<2.0 h	<2.0 h	<2.0 h	<3.0 h	<3.0 h
DNA extraction + quality control	3.5 h	3.5 h	3.5 h	4.0 h	4.5 h	4.5 h
Library construction	1.0 h	1.5 h	1.5 h	1.5 h	2.0 h	2.0 h
Total time	6.5 h	9.9 h	11.0 h	13.5 h	29.0 h	34.6 h

[a]*Cost per isolate was calculated based on the market price of DNA extraction and ONT consumables in China in Dec. 2019.*

so that sufficient data are obtained for each multiplexed isolate to allow for correct serotype prediction. Also, the unevenness of data yield between each multiplexed isolate increases as the number of isolates being multiplexed increases. The total number of multiplexed isolates, therefore, should be carefully considered. In this study 50 × depth of genome coverage (∼ 5 Gbp of raw data) could be obtained within an average of 6.0 h of ONT sequencing when multiplexing five isolates, while multiplexing seven isolates took an average of 19.5 h to reach equivalent depth (**Table 4**). Hence multiplexing seven to ten isolates resulted in only a small cost benefit compared with multiplexing five isolates (**Table 4**). Moreover, our previous study showed that both the data yield and quality started to decline after 12 h of sequencing on the flow cells of GridION, and the flow cells usually showed remarkably low numbers of active pores after 48 h of sequencing (Xu et al., 2020). These data suggest that multiplexing five isolates is likely to be more efficient overall, than multiplexing greater numbers of isolates. Total turnaround time is of critical importance in the practical application of serotype prediction in the food industry; multiplexing five isolates will allow serotyping results to be obtained within one day. Increasing the number of isolates being multiplexed resulted in increased sequencing time, and more total time for DNA extraction, library construction and data analysis (**Table 4**).

It is not known if certain strains of *Salmonella* would alter the accuracy of serotyping under the recommended settings, as only Ten serotypes were involved in the current study. Further validation with more *Salmonella* serotypes is necessary to operationalize the serotyping of *Salmonella* through multiplex ONT WGS, for example in the food industry or in public health.

CONCLUSION

This study demonstrated that accurate serotype prediction results could be obtained when multiplexing five or less *Salmonella* isolates with an average of 6 h of multiplex ONT sequencing, where each multiplexed isolate received at least 50 × depth of genome coverage of sequencing data after demultiplexing. Multiplexing up to five isolates in one flow cell is recommended to achieve high coverage of the genome and high accuracy of prediction within one day. Multiplexing five isolates results in a cost reduction to 23% of the cost of ONT sequencing of a single isolate per flow cell. Our study helps to identify the optimal combinations of isolate multiplexing number and sequencing time to achieve the most accurate, rapid, and cost-efficient *Salmonella* serotype prediction with multiplex ONT sequencing. This study also sets a starting point for the future validation of multiplex ONT WGS as a cost-efficient, rapid *Salmonella* confirmation, and serotype classification tool for the food industry.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://www.ncbi.nlm.nih.gov/, PRJNA694442.

AUTHOR CONTRIBUTIONS

ST and CG: conception and design of the work. XW, ST, and HL: data collection, data analysis, and drafting the article. CG, FX, GZ, SL, XD, MW, AS, and RB: critical revision of the article. All authors contributed to the article and approved the submitted version.

FUNDING

We self-funded by the Mars Global Food Safety Center, a non-competitive and non-profit research institute. The funders declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

ACKNOWLEDGMENTS

We thank Peter Markwell and Dr. Bala Ganesan for comments that greatly improved the manuscript. We also would like to thank Oxford Nanopore Technologies for supporting the establishment of ONT capability at the Mars Global Food Safety Center.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.637771/full#supplementary-material

REFERENCES

Allard, M. W., Strain, E., Melka, D., Bunning, K., Musser, S. M., Brown, E. W., et al. (2016). Practical value of food pathogen traceability through building a whole-genome sequencing network and database. *J. Clin. Microbiol.* 54, 1975–1983. doi: 10.1128/JCM.00081-16

Ashton, P. M., Nair, S., Peters, T. M., Bale, J. A., Powell, D. G., Painset, A., et al. (2016). Identification of *Salmonella* for public health surveillance using whole genome sequencing. *PeerJ* 4:e1752. doi: 10.7717/peerj.1752

Bailey, J. S., Fedorka-Cray, P. J., Stern, N. J., Craven, S. E., Cox, N. A., and Cosby, D. E. (2002). Serotyping and ribotyping of *Salmonella* using restriction enzyme *PvuII*. *J. Food Prot.* 65, 1005–1007. doi: 10.4315/0362-028x-65.6.1005

Banerji, S., Simon, S., Tille, A., Fruth, A., and Flieger, A. (2020). Genome-based *Salmonella* serotyping as the new gold standard. *Sci. Rep.* 10:4333. doi: 10.1038/s41598-020-61254-1

CDC (2006). *Salmonella* Surveillance: Annual Summary, 2006. Atlanta, GA: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention.

CDC (2018). *National Enteric Disease Surveillance — Salmonella* Annual Report, 2016. Atlanta, GA: US Centers for Disease Control and Prevention (CDC).

Cooper, A. L., Low, A. J., Koziol, A. G., Thomas, M. C., Leclair, D., Tamber, S., et al. (2020). Systematic evaluation of whole genome sequence-based predictions of *Salmonella* serotype and antimicrobial resistance. *Front. Microbiol.* 11:549. doi: 10.3389/fmicb.2020.00549

De Coster, W., D'Hert, S., Schultz, D. T., Cruts, M., and Van Broeckhoven, C. (2018). NanoPack: visualizing and processing long-read sequencing data. *Bioinformatics* 34, 2666–2669. doi: 10.1093/bioinformatics/bty149

Didelot, X., Walker, A. S., Peto, T. E., Crook, D. W., and Wilson, D. J. (2016). Within-host evolution of bacterial pathogens. *Nat. Rev. Microbiol.* 14, 150–162. doi: 10.1038/nrmicro.2015.13

Diep, B., Barretto, C., Portmann, A. C., Fournier, C., Karczmarek, A., Voets, G., et al. (2019). *Salmonella* serotyping; comparison of the traditional method to a microarray-based method and an *in silico* platform using whole genome sequencing data. *Front. Microbiol.* 10:2554. doi: 10.3389/fmicb.2019.02554

EFSA, and ECDC (2019a). The European Union One Health 2018 zoonoses report. *EFSA. J.* 17:e05926. doi: 10.2903/j.efsa.2019.5926

EFSA, and ECDC (2019b). The European Union summary report on antimicrobial resistance in zoonotic and indicator bacteria from humans, animals and food in 2017. *EFSA. J.* 17:e05598. doi: 10.2903/j.efsa.2019.5598

EFSA Panel on Biological Hazards, Koutsoumanis, K., Allende, A., Alvarez-Ordóñez, A., Bolton, D., Bover-Cid, S., et al. (2019). Whole genome sequencing and metagenomics for outbreak investigation, source attribution and risk assessment of food-borne microorganisms. *EFSA. J.* 17:e05898. doi: 10.2903/j.efsa.2019.5898

Foley, S. L., Zhao, S., and Walker, R. D. (2007). Comparison of molecular typing methods for the differentiation of *Salmonella* foodborne pathogens. *Foodborne Pathog. Dis.* 4, 253–276. doi: 10.1089/fpd.2007.0085

Fox, E. J., Reid-Bayliss, K. S., Emond, M. J., and Loeb, L. A. (2014). Accuracy of next generation sequencing platforms. *Next Gener. Seq. Appl.* 1:1000106. doi: 10.4172/jngsa.1000106

Galanis, E., Lo Fo Wong, D. M., Patrick, M. E., Binsztein, N., Cieslik, A., Chalermchikit, T., et al. (2006). Web-based surveillance and global *Salmonella* distribution, 2000-2002. *Emerg. Infect. Dis.* 12, 381–388. doi: 10.3201/eid1205.050854

GMA (2009). *Control of Salmonella in Low-Moisture Foods*. Rosslyn, VA: The Association of Food, Beverage and Consumer Products Companies (GMA).

Grimont, P., and Weill, F. (2007). *Antigenic Formulae of the Salmonella Serovars*, 9th Edn. Paris: WHO Collaborating Centre for Reference and Research on *Salmonella*.

Hauser, E., Tietze, E., Helmuth, R., and Malorny, B. (2011). Different mutations in the oafA gene lead to loss of O5-antigen expression in *Salmonella enterica* serovar typhimurium. *J. Appl. Microbiol.* 110, 248–253. doi: 10.1111/j.1365-2672.2010.04877.x

Herikstad, H., Motarjemi, Y., and Tauxe, R. V. (2002). *Salmonella* surveillance: a global survey of public health serotyping. *Epidemiol. Infect.* 129, 1–8. doi: 10.1017/s0950268802006842

Herrera-León, S., Ramiro, R., Arroyo, M., Díez, R., Usera, M. A., and Echeita, M. A. (2007). Blind comparison of traditional serotyping with three multiplex PCRs for the identification of *Salmonella* serotypes. *Res. Microbiol.* 158, 122–127. doi: 10.1016/j.resmic.2006.09.009

Imai, K., Nemoto, R., Kodana, M., Tarumoto, N., Sakai, J., Kawamura, T., et al. (2020). Rapid and accurate species identification of Mitis group *Streptococci* using the MinION nanopore sequencer. *Front. Cell. Infect. Microbiol.* 10:11. doi: 10.3389/fcimb.2020.00011

Issenhuth-Jeanjean, S., Roggentin, P., Mikoleit, M., Guibourdenche, M., de Pinna, E., Nair, S., et al. (2014). Supplement 2008-2010 (no. 48) to the White-Kauffmann-Le Minor scheme. *Res. Microbiol.* 165, 526–530. doi: 10.1016/j.resmic.2014.07.004

Karamitros, T., and Magiorkinis, G. (2018). Multiplexed targeted sequencing for Oxford Nanopore MinION: a detailed library preparation procedure. *Methods Mol. Biol.* 1712, 43–51. doi: 10.1007/978-1-4939-7514-3_4

Kennedy, S. R., Prost, S., Overcast, I., Rominger, A. J., Gillespie, R. G., and Krehenwinkel, H. (2020). High-throughput sequencing for community analysis: the promise of DNA barcoding to uncover diversity, relatedness, abundances and interactions in spider communities. *Dev. Genes Evol.* 230, 185–201. doi: 10.1007/s00427-020-00652-x

Kircher, M., Sawyer, S., and Meyer, M. (2012). Double indexing overcomes inaccuracies in multiplex sequencing on the Illumina platform. *Nucleic Acids Res.* 40:e3. doi: 10.1093/nar/gkr771

Liou, C. H., Wu, H. C., Liao, Y. C., Yang Lauderdale, T. L., Huang, I. W., and Chen, F. J. (2020). nanoMLST: accurate multilocus sequence typing using Oxford Nanopore Technologies MinION with a dual-barcode approach to multiplex large numbers of samples. *Microb. Genom.* 6:e000336. doi: 10.1099/mgen.0.000336

McClelland, M., Sanderson, K. E., Spieth, J., Clifton, S. W., Latreille, P., Courtney, L., et al. (2001). Complete genome sequence of *Salmonella enterica* serovar Typhimurium LT2. *Nature* 413, 852–856. doi: 10.1038/35101614

McQuiston, J. R., Waters, R. J., Dinsmore, B. A., Mikoleit, M. L., and Fields, P. I. (2011). Molecular determination of H antigens of *Salmonella* by use of a microsphere-based liquid array. *J. Clin. Microbiol.* 49, 565–573. doi: 10.1128/JCM.01323-10

Oh, J. H., and Park, M. K. (2017). Recent trends in *Salmonella* outbreaks and emerging technology for biocontrol of *Salmonella* using phages in foods: a review. *J. Microbiol. Biotechnol.* 27, 2075–2088. doi: 10.4014/jmb.1710.10049

Parkhill, J., Dougan, G., James, K. D., Thomson, N. R., Pickard, D., Wain, J., et al. (2001). Complete genome sequence of a multiple drug resistant *Salmonella enterica* serovar Typhi CT18. *Nature* 413, 848–852. doi: 10.1038/35101607

Piper, A. M., Batovska, J., Cogan, N. O. I., Weiss, J., Cunningham, J. P., Rodoni, B. C., et al. (2019). Prospects and challenges of implementing DNA metabarcoding for high-throughput insect surveillance. *Gigascience* 8:giz092. doi: 10.1093/gigascience/giz092

Rang, F. J., Kloosterman, W. P., and de Ridder, J. (2018). From squiggle to basepair: computational approaches for improving nanopore sequencing read accuracy. *Genome Biol.* 19:90. doi: 10.1186/s13059-018-1462-9

Ranieri, M. L., Shi, C., Moreno Switt, A. I., den Bakker, H. C., and Wiedmann, M. (2013). Comparison of typing methods with a new procedure based on sequence characterization for *Salmonella* serovar prediction. *J. Clin. Microbiol.* 51, 1786–1797. doi: 10.1128/JCM.03201-12

Ruan, J., and Li, H. (2020). Fast and accurate long-read assembly with wtdbg2. *Nat. Methods* 17, 155–158. doi: 10.1038/s41592-019-0669-3

Scallan, E., Hoekstra, R. M., Angulo, F. J., Tauxe, R. V., Widdowson, M. A., Roy, S. L., et al. (2011). Foodborne illness acquired in the United States–major pathogens. *Emerg. Infect. Dis.* 17, 7–15. doi: 10.3201/eid1701.p11101

Shi, C., Singh, P., Ranieri, M. L., Wiedmann, M., and Moreno Switt, A. I. (2015). Molecular methods for serovar determination of *Salmonella*. *Crit. Rev. Microbiol.* 41, 309–325. doi: 10.3109/1040841X.2013.837862

Soyer, Y., Alcaine, S. D., Schoonmaker-Bopp, D. J., Root, T. P., Warnick, L. D., McDonough, P. L., et al. (2010). Pulsed-field gel electrophoresis diversity of human and bovine clinical *Salmonella* isolates. *Foodborne Pathog. Dis.* 7, 707–717. doi: 10.1089/fpd.2009.0424

Tack, D. M., Ray, L., and Griffin, P. M. (2020). Preliminary incidence and trends of infections with pathogens transmitted commonly through food — foodborne diseases active surveillance network, 10 U.S. sites, 2016–2019. *MMWR Morb. Mortal. Wkly. Rep.* 69, 509–514. doi: 10.15585/mmwr.mm6917a1

Tyler, A. D., Mataseje, L., Urfano, C. J., Schmidt, L., Antonation, K. S., Mulvey, M. R., et al. (2018). Evaluation of Oxford Nanopore's MinION sequencing device for microbial whole genome sequencing applications. *Sci. Rep.* 8:10931. doi: 10.1038/s41598-018-29334-5

Uelze, L., Borowiak, M., Deneke, C., Szabó, I., Fischer, J., Tausch, S. H., et al. (2020). Performance and accuracy of four open-source tools for *in silico* serotyping of *Salmonella spp*. based on whole-genome short-read sequencing data. *Appl. Environ. Microbiol.* 86:e02265-19. doi: 10.1128/AEM.02265-19

Vaser, R., Soviæ, I., Nagarajan, N., and Šikiæ, M. (2017). Fast and accurate de novo genome assembly from long uncorrected reads. *Genome Res.* 27, 737–746. doi: 10.1101/gr.214270.116

Wattiau, P., Boland, C., and Bertrand, S. (2011). Methodologies for *Salmonella enterica* subsp. *enterica* subtyping: gold standards and alternatives. *Appl. Environ. Microbiol.* 77, 7877–7885. doi: 10.1128/AEM.05527-11

Wick, R. R., Judd, L. M., and Holt, K. E. (2018). Deepbinner: demultiplexing barcoded Oxford Nanopore reads with deep convolutional neural networks. *PLoS Comput. Biol.* 14:e1006583. doi: 10.1371/journal.pcbi.1006583

Xu, F., Ge, C., Luo, H., Li, S., Wiedmann, M., Deng, X., et al. (2020). Evaluation of real-time nanopore sequencing for *Salmonella* serotype prediction. *Food Microbiol.* 89:103452. doi: 10.1016/j.fm.2020.103452

Xu, Y., Lewandowski, K., Lumley, S., Pullan, S., Vipond, R., Carroll, M., et al. (2018). Detection of viral pathogens with multiplex Nanopore MinION sequencing: be careful with cross-talk. *Front. Microbiol.* 9:2225. doi: 10.3389/fmicb.2018.02225

Yachison, C. A., Yoshida, C., Robertson, J., Nash, J. H. E., Kruczkiewicz, P., Taboada, E. N., et al. (2017). The validation and implications of using whole genome sequencing as a replacement for traditional serotyping for a national *Salmonella* reference laboratory. *Front. Microbiol.* 8:1044. doi: 10.3389/fmicb.2017.01044

Yoshida, C., Gurnik, S., Ahmad, A., Blimkie, T., Murphy, S. A., Kropinski, A. M., et al. (2016). Evaluation of molecular methods for identification of *Salmonella* serovars. *J. Clin. Microbiol.* 54, 1992–1998. doi: 10.1128/JCM.00262-16

Zhang, S., den Bakker, H. C., Li, S., Chen, J., Dinsmore, B. A., Lane, C., et al. (2019). SeqSero2: rapid and improved *Salmonella* serotype determination using whole-genome sequencing data. *Appl. Environ. Microbiol.* 85, e01746-19.

Zhang, S., Yin, Y., Jones, M. B., Zhang, Z., Deatherage Kaiser, B. L., Dinsmore, B. A., et al. (2015). *Salmonella* serotype determination utilizing high-throughput genome sequencing data. *J. Clin. Microbiol.* 53, 1685–1692. doi: 10.1128/JCM.00323-15

Zou, W., Lin, W. J., Hise, K. B., Chen, H. C., Keys, C., and Chen, J. J. (2012). Prediction system for rapid identification of *Salmonella* serotypes based on pulsed-field gel electrophoresis fingerprints. *J. Clin. Microbiol.* 50, 1524–1532. doi: 10.1128/JCM.00111-12

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Wu, Luo, Xu, Ge, Li, Deng, Wiedmann, Baker, Stevenson, Zhang and Tang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership