A novel approach for detecting *Salmonella enterica* strains frequently attributed to human illness—development and validation of the highly pathogenic *Salmonella* (HPS) multiplex PCR assay

Dayna M. Harhay*, Kerry D. Brader, Tatum S. Katz,
Gregory P. Harhay, James L. Bono, Joseph M. Bosilevac and
Tommy L. Wheeler

Roman L. Hruska, US Meat Animal Research Center, Meat Safety and Quality Research Unit, USDA ARS, Clay Center, NE, United States

Introduction: Non-typhoidal *Salmonella enterica* (NTS) are leading bacterial agents of foodborne illnesses and a global concern for human health. While there are over 2,600 different serovars of NTS, epidemiological data suggests that certain serovars are better at causing disease than others, resulting in the majority of reported human illnesses in the United States. To improve food safety, there is a need to rapidly detect these more pathogenic serovars to facilitate their removal from the food supply.

Methods: Addressing this need, we conducted a comparative analysis of 23 closed *Salmonella* genomic sequences of five serotypes. The analysis pinpointed eight genes (*sseK2*, *sseK3*, *gtgA/gogA*, *avrA*, *lpfB*, *SspH2*, *spvD*, and *invA*) that in combination, identify 7 of the 10 leading *Salmonella* serovars attributed to human illnesses in the US each year (i.e., Serovars of Concern or SoC). A multiplex PCR assay was developed to detect the presence of these genes, with strains amplifying five or more targets designated Highly Pathogenic *Salmonella*, or HPS. The utility of the resulting HPS assay for identifying SoC was examined *in silico*, using BLAST to determine the distribution of gene targets among closed *Salmonella* genome sequences in GenBank (n = 2,192 representing 148 serotypes) and by assaying 1,303 *Salmonella* (69 serotypes), isolated from FSIS regulatory samples.

Results and discussion: Comparison of serotypes identified by the assay as HPS, with those identified as SoC, produced an Area Under the Curve (AUC) of 92.2% with a specificity of 96% and a positive predictive value of 97.4%, indicating the HPS assay has strong ability to identify SoC. The data presented lay the groundwork for development of rapid commercial assays for the detection of SoC.

KEYWORDS

non-typhoidal *Salmonella enterica*, highly pathogenic *Salmonella* (HPS), multiplex PCR, secreted effector, fimbriae, serovars of concern (SoC)

1 Introduction

Non-typhoidal *Salmonella enterica* (NTS) are the leading cause of foodborne illness attributed to bacteria in the United States, Canada, and Mexico (Scallan et al., 2011; National Enteric Surveillance Program, 2022; Godínez-Oviedo et al., 2020) and the second most common causative agent among reported zoonotic infections in the European Union (EU) (EFSA and ECDC, 2021). NTS are also the primary cause of deaths attributed to foodborne disease around the world [59,771 in 2017 (Stanaway et al., 2019)]. The success of NTS as human pathogens is attributed to their genotypic and phenotypic diversity. Composed of over 2,600 serotypes (Issenhuth-Jeanjean et al., 2014), some NTS are adapted to particular hosts or environments (Merkushova et al., 2023), while others are generalists, capable of persisting in a variety of hosts or environments (Waldner et al., 2012; Jechalke et al., 2019). In production agriculture, certain NTS serovars are noted for causing clinical illnesses in animals, but more often they are harbored as part of the animal's normal flora and do not cause disease (Kuria, 2024). As a result, it is not uncommon to isolate *Salmonella* from healthy animals on farm or at harvest (Brichta-Harhay et al., 2008; Kunze et al., 2008; Loneragan et al., 2012; Schmidt et al., 2012; Gutema et al., 2019).

Comparisons of NTS serotypes isolated from animals in pre- and post-harvest settings have revealed a disparity, in that those commonly isolated from animals and meat products do not necessarily reflect those most attributed to meat and poultry related outbreaks and illnesses (Bosilevac et al., 2009; Brichta-Harhay et al., 2011; Schmidt et al., 2012; Cheng et al., 2019). An example of this can be seen with NTS isolated from ground beef, where surveys show serotypes Montevideo and Anatum are predominantly isolated, while serotypes Newport and Typhimurium are observed much less frequently (Bosilevac et al., 2009). Despite being rarely observed as contaminants, Newport and Typhimurium have been attributed to 17 ground beef related outbreaks (Centers for Disease Control and Prevention (CDC), National Outbreak Reporting System (NORS) data 2009–2018), resulting in 930 illnesses and one death, while Montevideo has been attributed to just two ground beef related outbreaks over the same time period, resulting in 124 illnesses and no deaths, and Anatum has not been attributed to any ground beef related outbreaks (Centers for Disease Control and Prevention, 2023). Data on serotypes attributed to ground beef outbreaks was obtained from CDC's NORS dashboard; accessed 8/31/2023. That site is no longer available as it has become part of the new BEAM dashboard.

Historically however, all NTS have been regarded as human pathogens, irrespective of serotype, because of their ability to invade intestinal epithelial cells and potential for causing systemic disease (Groisman and Ochman, 1997; Hensel, 2000). Yet, the past 30 years of surveillance data have also demonstrated that certain serovars consistently contribute to most human illnesses year over year (Collins et al., 2022), with ~10% of serotypes isolated each year (43/442) attributed to ~90% of human illness isolates characterized (n = 45,515) (Centers for Disease Control and Prevention, 2022). This observation has fueled the idea that to improve human health outcomes with regard to *Salmonella* infections, the food safety community should focus control efforts by specifically targeting serovars that pose a greater risk to human health (i.e., Serovars of Concern, or SoC) rather than attempting to broadly manage all *Salmonella* contamination (United States Department of Agriculture Food Safety and Inspection Service, 2024). Along these lines, two recent studies have described methods for analyzing *Salmonella* outbreak data, to identify which serovars attributed to meat and poultry contribute to most human illnesses in the US, year over year (Katz et al., 2024; Marshall et al., 2024). The study by Katz et al., identified a list of 21 SoC, while the study by Marshall et al., described 13 serovars associated with a high burden of outbreak illnesses. The resulting serovar lists overlapped substantially suggesting there may be genetic similarities between these more pathogenic versions of NTS.

The increase in *Salmonella* genome sequence data in recent years has provided the opportunity to characterize, through comparative genomics, the NTS virulence gene repertoire and several studies have provided summaries to this effect (den Bakker et al., 2011; Worley et al., 2018; Rakov et al., 2019; Wang et al., 2020). Furthermore, Fenske and colleagues recently reported an analysis of the distribution of 182 virulence genes among high and low illness incidence serovars and found that NTS associated with a higher incidence of human illness appear to possess a common repertoire of virulence genes, that likely enable them to more readily cause disease, possibly at a lower infectious dose (Fenske et al., 2023). These seminal investigations have provided a broad view of the NTS virulence gene landscape and insight into the differences within and between serovars that influence host pathogen interactions. What has been lacking to date, however, is the identification of a limited NTS virulence gene set that can be harnessed for the rapid detection of the more pathogenic versions of *Salmonella* (i.e., Highly Pathogenic *Salmonella* or HPS).

In Europe, poultry production control programs targeting the leading *Salmonella* serovars (Enteritidis, Typhimurium and the monophasic version of Typhimurium, I,4,[5],12:i:-; also known as SE/ST serovars) have been implemented[1] with the goal of reducing human illnesses by restricting the presence of these *Salmonella* serovars. Accordingly, real-time PCR assays have been developed for the identification of these serovars (Maurischat et al., 2015; Zhengwei et al., 2023) and commercial versions of rapid assays for identifying SE/ST are available. However, these leading serovars account for only ~35% of the US human illness isolates characterized by CDC each year (CDC BEAM Dashboard,[2] accessed September 20, 2024). Furthermore, studies show this targeted serovar approach has stalled out in recent years, and there is a perceived need for a more proactive screening approach based on the presence of particular virulence genes that contribute to overall pathogenicity of *Salmonella*, irrespective of serovar, as opposed to a reactive approach of targeting particular serovars, based on outbreak data or trends in serovar prevalence in food commodities (Petrin et al., 2023).

To assist with this effort, we conducted a comparative genomics analysis of 23 closed genome sequences of *Salmonella* serovars that are frequently associated with human illnesses (serovars Typhimurium, Newport and Dublin) with serovars that are less frequently associated with human illness (Montevideo and Anatum), that had been isolated from cattle and beef products or humans. Analysis of this test set of carefully selected closed genome sequences led to the identification of 8 virulence genes (*sseK2*, *sseK3*, *gtgA/gogA*, *avrA*, *lpfB*, *SspH2*, *spvD*, and *invA*) that tended to be present in the genomes of serovars more commonly associated with human illness. Here were report on the utility of these 8 virulence genes for identifying *Salmonella* SoC. A multiplex PCR assay targeting the detection of these 8 genes was developed (designated the HPS assay) and evaluated using 1,303 *Salmonella* strains

---

1   https://www.fao.org/faolex/results/details/en/c/LEX-FAOC040332/

2   https://www.cdc.gov/ncezid/dfwed/BEAM-dashborad.html

TABLE 1 Description of serotype, genbank accession numbers and sequence types (ST) of genomes used in the Mauve analysis, and heat map of the percent homology (dark blue highest and white lowest) of the HPS gene targets present in each strain, in comparison with *S. typhimurium* SL1344.

Genbank Accession Number	Strain ID	ST	Serotype	HPS Assay gene target							
				invA	*avrA*	*sspH2*	*gtgA*	*lpfB*	*sseK2*	*sseK3*	*spvD*
NC_016810	SL1344	ST19	Typhimurium	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
NC_003197	LT2	ST19	Typhimurium	99.95	99.67	100.00	100.00	100.00	100.00	null	100.0
CP014970	1808	ST19	Typhimurium	99.95	99.67	100.00	100.00	100.00	100.00	null	100.0
CP014979	1878	ST19	Typhimurium	99.95	99.67	100.00	100.00	100.00	100.00	null	null
CP014975	1906	ST19	Typhimurium	99.95	99.67	100.00	100.00	100.00	100.00	null	null
CP014977	1896	ST19	Typhimurium	99.95	99.67	99.96	99.85	100.0	99.91	null	null
CP001363	14028S	ST19	Typhimurium	99.95	99.56	100.0	100.0	100.0	100.0	100.0	100
CP014981	1880	ST19	Typhimurium	99.95	99.56	99.958	100.0	100.0	100.0	null	null
CP007559	1934	ST31	Newport	99.32	99.56	91.0	99.0	99.9	99.8	null	null
CP025232	1925	ST132	Newport	99.32	99.56	90.9	99.9	99.9	99.8	null	null
CP025237	1928	ST45	Newport	99.32	99.45	90.99	100.0	99.86	99.71	null	null
CP025241	1929	ST45	Newport	99.32	99.45	90.99	100.0	99.86	99.71	null	null
CP025248	1932	ST5	Newport	99.03	99.01	90.87	100.0	99.86	98.57	100.0	null
CP007216	1927	ST118	Newport	99.03	99.01	null	98.11	99.86	98.57	null	null
CP032379	69807	ST10	Dublin	99.27	99.34	96.80	99.56	99.86	null	100.0	100.0
CP032446	69840	ST10	Dublin	99.27	99.34	96.80	99.56	99.86	null	100.0	100.0
CP032449	69838	ST10	Dublin	99.27	99.34	96.75	99.56	99.86	null	100.0	100.0
CP014620	1676	ST2167	Anatum	99.17	99.34	99.70	null	null	100.0	null	null
CP014707	1735	ST64	Anatum	99.17	99.34	99.70	null	null	100.0	null	null
CP017977	1918	ST316	Montevideo	98.98	99.23	null	null	null	null	null	null
CP017978	1919	ST4	Montevideo	99.03	null	null	null	null	null	null	null
CP017974	1914	ST81	Montevideo	99.08	null	null	null	null	null	null	null
CP007540	1921	ST138	Montevideo	99.42	null	null	null	null	null	null	null

across 69 different serotypes. Further, the DNA sequence of each target amplicon was subjected to BLAST analysis (BLAST: Basic Local Alignment Search Tool, n.d.) to examine the distribution of hits among 2,192 closed *Salmonella* genome sequences of 148 serovars present in the NIH sequence database, GenBank. The overall performance of the HPS assay for identifying *Salmonella* SoC was evaluated and its strengths and limitations are presented and discussed.

2 Materials and methods

2.1 HPS target identification and multiplex PCR assay development

Comparative genomic analysis of 23 *Salmonella enterica* ssp. *enterica* closed genome sequences (listed in Table 1) was conducted using Geneious v.11.1[3] Mauve plugin; progressiveMauve algorithm, with the following settings: Automatically calculate seed weight; Match seed weight = 15; Minimum LCB Score = 30,000; Compute Locally Collinear Blocks (LCBs); Full alignment; Gapped Aligner = MUSCLE 3.6 (Darling et al., 2004). Prior to Mauve analysis, care was taken to ensure that all linearized genome sequences had the same +1 position as that of the *S. typhimurium* SL1344 origin of replication, identified using OriFinder (Gao and Zhang, 2008). Plasmid sequences (n = 16) were grouped by size and analyzed using Mauve. The largest of the resulting 187 'locally

collinear blocks' (LCBs) (n = 42 LCB ranging from 10 to 784 kbp) were visually inspected for differential regions of homology. Genes that demonstrated greatest homology among serovars commonly attributed to human illness were analyzed with Geneious Primer Design, which is a modified version of Primer3 v.2.3.7, with the goal of producing PCR amplicons ranging in size from 844 bp to 275 bp (amplicon sizes listed in Table 2). The resulting primer sequences, genes targeted, binding locations, and annealing temperatures are listed in Table 2. Nucleotide BLAST analysis of predicted amplicons based on the sequences present in *S. typhimurium* strain SL1344, was conducted using BLAST+ 2.8.0. The results of this *in silico* analysis are summarized for the top 55 serovars in Table 3 (full list summarized in Supplementary Table S1) with the percent homology cutoff used to determine gene presence or absence indicated, and the distribution of genes targeted in 2,192 closed *Salmonella* genome sequences of 148 serotypes present in GenBank (data accessed September 12, 2024).

The protocol for conducting the HPS multiplex PCR assay can be found in Protocols.IO (Harhay et al., 2024), but briefly described, it contained the following in a 25 μL reaction volume: 2.5 U Bullseye H.S. Taq Polymerase (MidSci, BE225108; 5 U/ul), 1 X Buffer II (10 X buffer with 15 mM MgCl$_2$), 0.24 mM dNTPs (NEB, N0447L), primers at a final concentration ranging from 0.16 to 0.25 μM, PCR grade water to bring volume to 20 μL, and 5 μL of template DNA. *Salmonella* isolates to be tested with the HPS assay were struck for isolation onto either XLD or Tryptic Soy Agar (TSA) plates, which were incubated at 37°C overnight. Isolated colonies were then sub-cultured to Trypticase Soy Broth (TSB) and incubated at 37°C, static for 18 to 20 h. From the overnight cultures, 5 μL was transferred into 200 μL of prepared BAX System lysis reagent (Hygiena, Camarillo, CA, United States) and lysed

3 www.geneious.com

TABLE 2 Description of genes targeted, location on chromosome or in plasmid, role in pathogenesis if known, primers used to identify each gene, amplicon size, and PMID for references describing gene function.

Gene	HPS Assay ID	SL1344_ Locus ID	Gene size (bp)	Location	Product	Function	Primer name	HPS primer sequences (5'–3')	Amplicon size (bp)	Reference (PMID)
sseK2	HPS-6	RS10980	1,047	chromosome	Secreted effector. Arginine glycosyltransferase.	Enhances bacterial survival in host by avoiding clearance mechanisms.	HPS-6-F	GCACGTTTTAATGCCGCTTTT	844	32131463
							HPS-6-R	GATGCTCGCTGCGGTTAAC		
sseK3	HPS-1	RS10000	1,008	GIFSY-1 or ST64B	NleB homologue, secreted effector shown to inhibit TNF-a and NF-kB host signalling by preventing translocation of immune regulator NF-kB to the host cell nucleus	Interference with host innate immune response; enhances bacterial survival and persistence in host; avoiding host clearance mechanisms	HPS-1-F	TGCCTTCATCAGGTAGTGCAA	739	28069818; 32131463; 32974215; 29449376
							HPS-6-R	GATGCTCGCTGCGGTTAAC		
avrA	HPS-3	RS14835	906	SPI-1*	yopJ homologue, secreted effector acetyltransferase; disrupts host innate immune signalling by targeting NF-kB.	Hijacks host epithelial cells, enhancing survival; reverses host inflammatory response; blocks recruitment of phagocytic leukocytes	HPS-3-F	CGAGTTTATCGCCTCAGCCT	663	30025209; 32842467; 21899703
							HPS-3-R	TCGTTGTRCGCCTTGAGTAT		
lpfB	HPS-5	RS18760	699	chromosome	PapD-like chaperone.	Long Polar fimbriae (lpf) promotes binding and colonization of M cells located at Peyer's patch.	HPS-5-F	GCTGGCGTTGATTGCTCAAA	557	33614531; 28630073
							HPS-5-R	TCAATGAGCCTTTTGCCGGA		
spvD	HPS-4	RS24010	651	sPV plasmid	Secreted effector. Inhibits NF k-B signalling pathway. Targets p65.	Suppresses host immune response and contributes to systemic/ invasive infection.	HPS-4-F	TGGTAGTGCGTCATCCCAAG	452	27232334; 27789710
							HPS-4-R	GAGCTATGTAGTTCTCTGCGCT		
sspH2	HPS-7	RS11515	2,367	SPI-12*	Secreted effector. E3-ubiquitin ligase	Alters host innate immune response and enhances bacterial survival in host cells.	HPS-7-F	ACTGTCTGAACGTACTTTGCAG	381	30049795; 33735297
							HPS-7-R	CGGTCCTCGCAGCTTGAG		
gtgA	HPS-2	RS05020	687	GIFSY-2	pipA (SPI-5) and gogA (GIFSY-2) homologues, secreted effectors – zinc metaloprotease that targets and cleaves NF-kB transcription factors p65, RelB, cRel	Inhibits host innate immune response and increases ability to infect host	HPS-2-F	CGCCTCCAACATGATGGTCT	338	30049795; 33735297
							HPS-2-R	TCCAGGTTGAGGGCAATCAC		
invA	invA	RS14990	2058	SPI-1	part of Type III Secretion System apparatus	Contributes to invasion of gut epithelial cells.	INVA-F	TATCGCCACGTTCGGGCAA	275	1528198; 1624429
							INVA-R	TCGCACCGTCAAAGGAACC		

Locations designated with an asterisk (*) indicate that the gene listed is not always present in that location, due to diveristy in pathogenicity island gene content within and between Salmonella serovars.

TABLE 3 Summary of the in silico analysis of the distribution of HPS genes among 2,192 closed *Salmonella* genome sequences of 148 serotypes present in genbank and the anticipated HPSi, in comparison with serovars observed with HPS analysis of 1,303 *Salmonella* isolated from 1,244 positive enrichment samples obtained from FSIS, ranked from greatest to least contribution to human illness (with gray color scale (darkest to lightest) indicating highest to lowest percent identified by commodity type), as determined by analysing human illness isolate data reported by CDC over a five year window, from 2018 to 2022.

	SoC/StW rank	Avg % human illness/yr	% FSIS isolates (n=1303)	🐄 n = 203	🐔 n = 579	🐖 n = 445	🦃 n = 76	Serovar	invA 98.2-100	avrA 98.3-100	lpfB 97.8-100	sspH2 98.2-100	sseK2 98.2-100	gtgA 98.2-100	spvD 99-100	sseK3 99-100	HPS index
Serotypes of Concern (SoC) – High burden of illness	1	15.36	8.37	2.5	**16.9**	0.9	2.6	Enteritidis	323	311	310	309	1	308	84	87	8
	2	12.36	0.92	4.4		0.4	1.3	Newport	95	91	92	90	90	74		11	7
	3	8.68	7.52	3.4	**11.6**	3.6	10.5	Typhimurium	347	306	294	291	288	283	125	99	8
	4	6.46	**0.00**					**Javiana**	8	8				6			3
	5	5.64	5.22	2.5	0.5	**13.5**		I 4,[5],12:i:-	137	137	138	133	138	134	17	16	8
	6	4.26	20.34	7.4	**33.3**	11.0	10.5	Infantis	149	148	146	147	148				5
	7	2.96	0.23			0.7		Saintpaul	25	24	25	25	25	18			6
	8	2.74	**0.00**					**Oranienburg**	6								1
	9	2.72	0.23		0.3	0.2		Braenderup	6	6	6	6	6	1			6
	10	2.4	2.76	**14.3**		0.9	3.9	Montevideo	32	10							2
	11	2.4	1.53	4.4		2.0	2.6	Muenchen	23	22	22	21	21	1			6
	12	2.18	0.54		1.2			Thompson	16	16	16	16		12			5
	13	2.14	0.08	0.5				Mississippi *	*			*				*	3
	14	1.22	**0.00**					Bareilly	32	32	29	31	31	30			6
	15	1.18	**0.00**					**Sandiego**	3								1
	16	1.1	0.08	0.5				Poona	4	3							2
	17	1.06	0.08	0.5				Paratyphi B	23		23	23	23				4
	18	1	0.54				9.2	Hadar	32	29	28	29	29	27			6
Serotypes to Watch (StW) – moderate to Low burden of illness	19	0.9	1.61	3.0		3.1	1.3	Agona	41	39	38						3
	20	0.88	**0.00**					**Typhi**	161			161					2
	21	0.82	1.00	1.0		2.5		Berta	3	3	3	3	3				5
	22	0.78	**0.00**					Rubislaw	6	2							2
	23	0.74	6.06	**11.8**	0.5	**11.7**		Anatum	43	41		41	41	3			5
	24	0.72	0.23	0.5		0.4		Panama	3	3							2
	25	0.66	**0.00**					Agbeni*	*	*				*			3
	26	0.58	0.84	4.9		0.2		Dublin	31	31	31	31		31	33	31	7
	27	0.54	**0.00**					Hartford	2	2	2	2	2				5
	28	0.52	0.46	0.5	0.2	0.9		Heidelberg	55	55	55	55	55				5
	29	0.52	0.54			1.3	1.3	Litchfield	1	1	1	1	1				5
	30	0.48	0.46	0.5		1.1		Manhattan	1	1	1	1	1	1			6
	31	0.48	**0.00**					Stanley	3	3	3			1			4
	32	0.46	**0.00**					Norwich*	*	*	*	*					4
	33	0.44	0.84	1.0	0.3	1.6		Mbandaka	18	18	17		17				4
	34	0.42	**0.00**					Miami	1	1				1			3
	35	0.42	1.46	1.5	0.2	0.9	**14.5**	Reading	13	10	3	3	3				5
	36	0.4	2.00	3.9		3.4	3.9	Uganda	12	12		12					3
	37	0.45	**0.00**					IV 50:Z4,Z23:- *	*								1
	38	0.36	1.61			1.6	**18.4**	Senftenberg	28	26	30		29				4
	39	0.34	0.23			0.7		Bovismorbificans	6	6							2
	40	0.28	0.61	1.0		1.3		Adelaide	2								1
Serotypes observed in FSIS enrichments, but not commonly attributed to human illnesses			12.97	1.5	**28.3**	0.4		Kentucky	61	57	57		36				4
			2.99	1.0	3.8	1.3	**11.8**	Schwarzengrund	34	32							2
			1.53	0.5	1.9	1.8		Johannesburg	6	6			1				3
			2.61	2.5		6.5		Derby	18	18	5						4
			2.38			7.0		Eko*	*	*			*			*	4
			0.92	4.4		0.7		Muenster	10								1
			0.92	3.9		0.9		Give	7	7							2
			1.53	1.0		3.8	1.3	London	**12**	**12**	**12**	**12**		7			5
			0.69	3.0		0.7		Cerro	5	5			4				3
			0.54			1.6		Worthington	**7**	**7**	**7**		**7**				4
			0.61	1.0		1.3		Brandenburg	8				2				2
			1.00			2.9		Ohio	4								1
			0.31			0.9		Soerenga	1	1	1						3
			0.08			0.2		Choleraesuis	6	1	5	5		5	6		6
			0.15		0.3			Alachua	1		1		1				3
								Total	1871	1543	1401	1448	1012	946	265	244	

Red and blue color scales correspond with SoC/StW, and HPSi designations in Figure 1.

according to manufacturer's protocol (incubation at 37°C for 20 min followed by 95°C for 10 min).

Reaction mixtures were subjected to a modified version of the touch down and extension thermal cycler amplification program previously described [(Paton and Paton, 1998); with specific modification as described in Harhay et al. (2024)]. *S. typhimurium* strain ATCC 14028 produces amplicons of the correct size for all 8 gene targets and was used as a control and DNA ladder for gel electrophoresis as shown in Harhay et al. (2024). Amplicons were visualized by agarose gel electrophoresis and staining with ethidium bromide (Sambrook and Russell, 2006; Lee et al., 2012).

2.2 *Salmonella* Isolation from FSIS positive primary enrichments, serotyping, and HPS analysis

Salmonella-positive primary enrichment broth samples (n = 1,244) were received from three USDA FSIS laboratories (~415 samples per location) over the course of 1 year, from June 2021 to May 2022 (Supplementary Table S2). Samples were stored as glycerol stocks (16% vol/vol glycerol) at −20°C and shipped to USMARC frozen on dry ice. Upon arrival, samples were thawed at room temperature (~30 min), vortexed, and then 1 mL from each sample was subjected

to immunomagnetic separation (IMS) using 20 microliters (20 μL) of Dynal anti-*Salmonella* Dynabeads™ (Applied Biosystems, Foster City, CA, United States) as previously described (Bosilevac et al., 2009). Recovered IMS beads were transferred to 3 mL of RVS medium and incubated at 42°C for 18–22 h.

Post incubation, RVS secondary enrichments were streaked for isolation onto XLD agar (Oxoid Thermo Fisher, Hampshire, United Kingdom) and incubated at 37°C for 20 h. Typical *Salmonella* colonies (pink with a black center and clear outer ring, or rarely, clear colonies with no evidence of H$_2$S production) were picked (up to eight putative *Salmonella* per sample) for downstream serovar characterization and HPS genotype profiling with the HPS assay. Two putative *Salmonella* isolates from each sample were cultured in TSB at 37°C for 18–20 h. These cultures were used to make two different DNA lysates, a BAX lysate as described above, and a proteinase-K lysate, made by combining 5 μL of TSB culture with 150 μL of PK-lysis buffer containing 1X Tris-EDTA, pH 8.0 (145 μL) and 5 μL proteinase K (20 mg/mL) per reaction, followed by incubation at 95°C for 15 min. Lysates were stored at −20°C until used as DNA template in three different multiplex PCR reactions to determine *Salmonella* molecular serovar., as previously described (Echeita et al., 2002; Herrera-León et al., 2004, 2007; Harhay et al., 2024). A subset of isolates (~10%) were selected for serovar confirmation using traditional O and H antisera agglutination typing methods. The serotypes and HPS genotypes identified for 1,303 unique isolates are summarized by commodity in Supplementary Table S3.

2.3 Examination of *Salmonella* incidence data to identify serotypes most often associated with human illness in the US

To assess the sensitivity and specificity of the HPS Assay for identifying *Salmonella* of greater concern for human health, it was necessary to develop a framework for determining which serotypes fell into this category. While recent studies have presented methods for identifying *Salmonella* serovars of concern (SoC) using analyses of outbreak data (Katz et al., 2024; Marshall et al., 2024), we endeavored to present a list of SoC based on analysis of *Salmonella* isolate data characterized by the CDC, and representative of serotypes associated with both sporadic and outbreak illnesses. This analysis entailed using *Salmonella* serotype data for human-derived isolates collected in the United States from 2018 to 2022, for which we examined the association between serotype and salmonellosis incidence (inferred from the number of isolates CDC characterized and reported) by year. The data assessed included *Salmonella* serotype information for isolates identified in PulseNet, NARMS, NORS (*Salmonella* outbreak data) and Epi Info, coordinated through the System for Enteric Disease Response, Investigation, and Coordination (SEDRIC).

For this analysis, BEAM Dashboard—Report Data (Centers for Disease Control and Prevention, 2022) were downloaded[4] on January 29, 2024. Analyses were conducted using the R statistical programming language version 4.2.1 (R Core Team, 2023). As illustrated in Supplementary Figure S1, *Salmonella* isolate data from 2018 to 2022

---

4 https://www.cdc.gov/ncezid/dfwed/BEAM-dashboard.html

were summarized by serotype to identify the total number of isolates characterized, as well as the number of isolates resulting from multistate outbreaks (MSOB), by year. The percent contribution to the total number of isolates characterized, was then calculated by year. These data were aggregated, resulting in 2,159 datapoints over the five-year window, ranging from 1 to 8,117 isolates per serotype identified. The top 13.7% of data points (ranging from 0.2 to 18.7% contribution per year; n = 297 datapoints and 93% of isolates characterized) were ranked from least to greatest, and a quartile analysis was performed to examine the distribution.

The top quartile (n = 75 data points) was further examined and serotypes (n = 40) that had at least one year with ≥ 10 MSOB isolates, and a total contribution of >0.2% over the five-year window were ranked from greatest to least. Those with an average contribution ≥ 1% (n = 18 serotypes and 76.7% of isolates) were classified as SoC, represent an increased burden of illness, and are summarized in Table 3. The remainder of the list (n = 22 serotypes and 9.7% of isolates) were classified as Serotypes to Watch (StW), as they represent a moderate burden of illness. The overlap between HPS genotype, serovar, SoC/StW, and abundance in the FSIS dataset, is illustrated in Figure 1.

2.4 Sensitivity and specificity of the HPS assay for identifying *Salmonella* serotypes of greater concern for human health
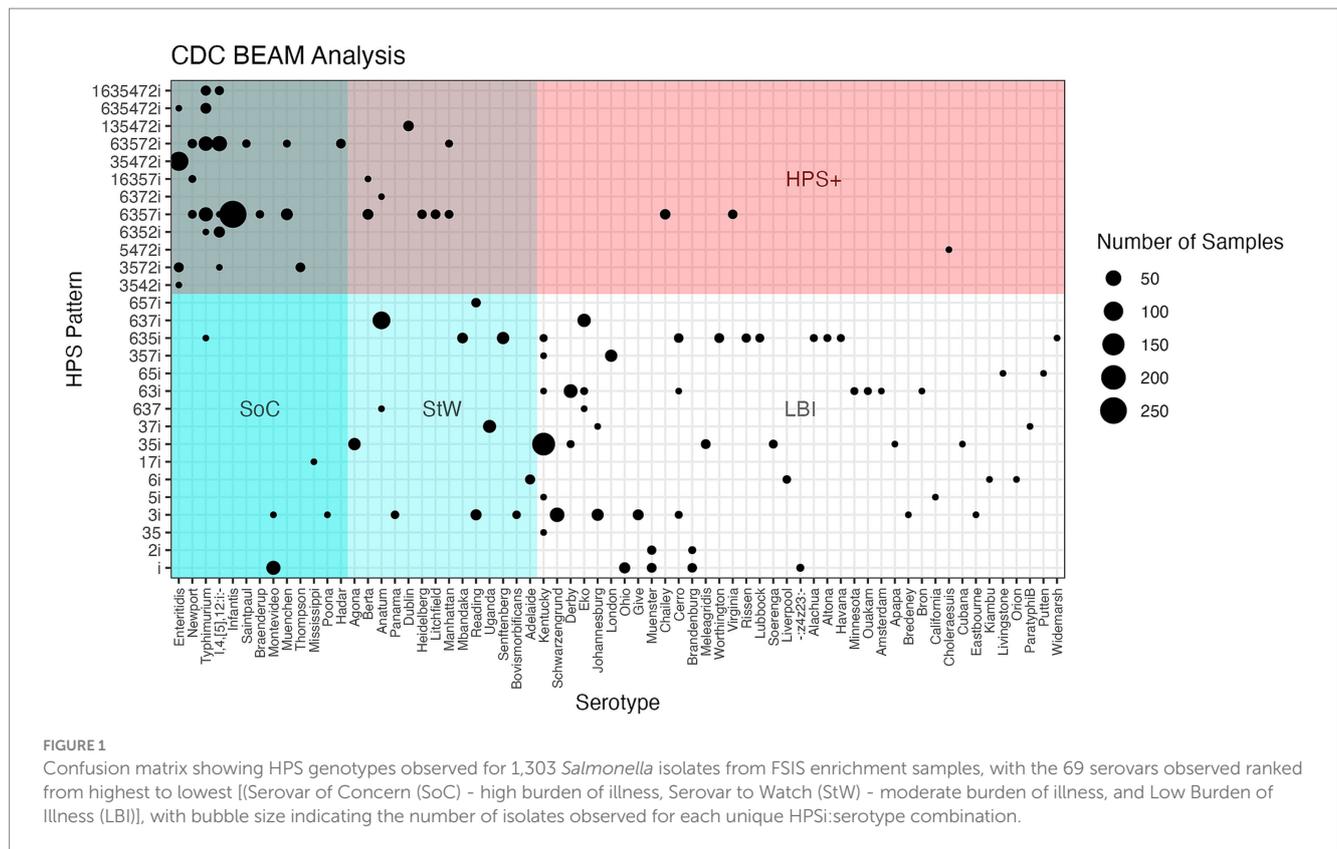
Serotypes with an HPSi ≥ 5 (HPS index, i.e., five or more gene targets amplified with the HPS assay) were compared with the list of SoC and StW summarized in Table 3, as well as the list of SoC generated by (Katz et al., 2024). Serotype:HPSi were compared with both SoC lists to ensure that a holistic group of important *Salmonella* serotypes was evaluated. The agreement between HPS and SoC definitions was calculated using confusion matrix statistics, specifically Area Under the Curve (AUC), sensitivity, and specificity (Altman and Bland, 1994; Kuhn, 2008) (Figure 2).

3 Results

3.1 Comparative genomics and BLAST analysis of genes targeted by the HPS assay

Comparative genomic analysis of 23 closed *Salmonella* genomes and 16 plasmids (Table 1) using Mauve, produced 187 LCB. LCB over 10kbp in size were visually inspected using the Geneious LCB viewer, for regions of greatest homology among serotypes with an increased association with human illness. In all, seven genes (*sseK2*, *sseK3*, *gtgA/gogA*, *avrA*, *lpfB*, *SspH2,* and *spvD*) were identified as potential targets based on their presence in the genome or plasmid sequences of *Salmonella* serotypes more commonly attributed to human outbreaks and illnesses. Notably, six of the seven genes identified were determined to encode secreted effector proteins that have been shown to alter host signal transduction pathways, and impede host innate immune response [*sseK, sseK3*—(Araujo-Garrido et al., 2020); *gtgA/gogA*—(Jennings et al., 2018; Takemura et al., 2021); *avrA*—(Labriola et al., 2018; Yin et al., 2020); *spvD*—(Grabe et al., 2016; Rolhion et al., 2016); *sspH2*—(Miao et al., 1999; De Meyer et al.,

FIGURE 1
Confusion matrix showing HPS genotypes observed for 1,303 *Salmonella* isolates from FSIS enrichment samples, with the 69 serovars observed ranked from highest to lowest [(Serovar of Concern (SoC) - high burden of illness, Serovar to Watch (StW) - moderate burden of illness, and Low Burden of Illness (LBI)], with bubble size indicating the number of isolates observed for each unique HPSi:serotype combination.

2021)]. One of the targets, the *spvD* gene, was found to be on the *spv* virulence plasmid (Guiney et al., 1995), while the other gene targets were located on the chromosome. The seventh gene (*lpfB*) encodes a chaperone protein, important in the synthesis of long polar fimbriae (lpf) which have been shown to promote binding and colonization of *Salmonella* to M (microfold) cells located in the Pyere's patch region of the ileum in the small intestine (Baümler and Heffron, 1995; Gonzales et al., 2017; Hansmeier et al., 2017). Primer sequences were selected in these genes, as well as primers previously identified for targeting the *Salmonella* specific portion of the *invA* gene [which has been shown to be present in >99.8% of *Salmonella* genome sequences (Rahn et al., 1992; Nucera et al., 2006)], such that a ladder of 8 amplicons ranging in size from 844 to 275 bp is produced and readily visualized with agarose gel electrophoresis (Table 2; Harhay et al., 2024).

BLAST analysis of the predicted amplicons (including primer sequences) showed them to be highly specific for *Salmonella enterica* subsp. *enterica*, with the exception of the *lpfB* amplicon which showed potential for cross-reactivity with a homolog (average amplicon identity of 90.9%) present in certain species of *Enterobacter* (specifically *hormaechei* and *asburiae* sequences present in genbank). The distribution of amplicons among 2,192 closed *Salmonella* genome sequences of 148 serotypes is summarized in Supplementary Table S1, and for the top 55 serovars (n = 1,871 sequences) attributed to human illnesses and or isolated from the FSIS enrichments in Table 3. This *in-silico* analysis confirmed the presence of HPS gene targets in serotypes Typhimurium, Newport and Dublin, but also showed differential presence in other serotypes commonly found in the CDCs list of top 20 *Salmonella* attributed to human illnesses (Centers for Disease Control and Prevention, 2018).

3.2 HPS assay results for *Salmonella* isolated from FSIS enrichments

To examine the utility of the HPS assay for identifying *Salmonella* SoC, 1,303 *Salmonella* were isolated from 1,244 positive enrichment samples provided by FSIS over the course of 1 year. These isolates were serotyped and their HPS genotype determined. The serotypes and percent observed from enrichments of each commodity are summarized in Supplementary Table S3 and Figure 1. Five serovars, namely Infantis (20.3%), Enteritidis (8.4%), Typhimurium (7.4%), Schwarzengrund (2.92%), and Reading (1.46%), were the only serovars found in all four meat commodities tested (Table 3). HPS analysis of the FSIS isolates showed 29 different HPS genotypes were observed, ranging from all 8 amplicons (scored as 1635472i observed for 14 Typhimurium and monophasic Typhimurium isolates) to just the amplicon for *invA* (scored as i; observed for 63 isolates of serovars Montevideo, Brandenburg, Muenster, −:z4,z23:-, and Ohio).

Overall, approximately 50% of the isolates from FSIS enrichments showed HPSi ≤ 4 amplicons, while the other 50% showed HPSi ≥ 5. The most common HPS genotype observed was 6357i (29% of isolates) found in 13 serotypes, of which S. infantis was the greatest contributor (264 of 1,303 *Salmonella* isolated). Remarkably, 3 isolates were observed to not amplify the *invA* target (two with HPS genotypes of 637, an Anatum and an Eko isolate from different pork samples; and one isolate with an HPS gene pattern of 35, a Kentucky isolate from a chicken enrichment), although not confirmed with sequence analysis if this was due to the absence of the *invA* gene or differences in primer binding sites. Importantly, analysis of HPSi by serotype showed that seven of the top 10 serovars on the CDC's list of Top 20 *Salmonella* (Centers for Disease Control and Prevention, 2018) had HPSi ≥ 5.

This observation led to a broader examination of HPSi among serovars present in CDC isolate data over a 5-year window (2018–2022).

3.3 Analysis of BEAM dashboard isolate data to define a list of SoC

To identify *Salmonella* serovars attributed to the greatest number of illnesses each year, BEAM Dashboard data (from 2018 to 2022) were downloaded and the contribution of 102,384 isolates to human illnesses and multistate outbreaks was assessed. This analysis produced a list of 40 serotypes, attributed to 86.3% of all isolates (a proxy for illnesses) characterized by CDC in that time window (Table 3). The top 18 serotypes of that list were attributed to 76.7% of isolates and given their prominent and repeated attribution to human illness, were classified as SoC. The second half of the list, classified as StW, comprised 22 serovars with moderate but repeated contribution to human illnesses (9.7% of all isolates characterized).

3.4 Comparison of SoC/StW classification with serotype:HPSi profiles observed for FSIS isolates

Having identified a list of SoC/StW from CDC isolate data, we applied this ranking to the serotypes isolated from the FSIS samples and used the HPS assay results of these isolates to examine the association between HPSi and SoC/StW. For the 1,303 *Salmonella* strains isolated and serotyped, 864 (66.3%) were serotypes classified as SoC/StW. The remaining 439 (33.7%) were serotypes categorized as being of lesser concern for human health, representing a low burden of illness (LBI). Examining the overlap between SoC/StW and HPSi ≥ 5 showed that 97% (635/652) of isolates with HPSi ≥ 5 were serotypes on the SoC/StW list, again demonstrating a high concordance. Illustrating this further, four of the top five serovars on the SoC list (Enteritidis, Newport, Typhimurium, and monophasic Typhimurium) all had HPSi ≥7. Conversely, 64.8% of isolates with HPSi ≤ 4 (422/651) were serotypes identified as LBI.

Sensitivity and specificity analyses also showed the HPS assay to have moderate-high sensitivity (73.5%) and a very high positive predictive value (97.4%) for isolates with an HPSi ≥ 5 being on the SoC/StW list, demonstrating a very strong ability of the HPS assay to identify SoC/StW. Specificity was found to be 96.3% with a negative predictive value of 64.8%, revealing a slight trend of the HPS assay to miss some serotypes on the SoC/StW lists. There were 13 serovars in this group which was dominated by Anatum (n = 78), Montevideo (n = 36), Uganda (n = 26), Agona, and Senftenberg each with 21 isolates, and Reading (n = 19). Taken together, an HPSi finding of 5 or greater is very strong, albeit slightly conservative evidence of an isolate being a SoC/StW. Figure 1 illustrates the distribution and abundance of FSIS serotype:HPS genotype, with serovars ranked by classification of SoC/StW, or LBI on the x-axis.

Finally, the AUC was calculated to examine the ability of HPSi ≥ 5 to correctly identify SoC/StW. AUC was calculated for two different lists of SoC: the list generated by Katz et al., based on analysis of outbreak data attributed to meat commodities (Katz et al., 2024), and the SoC/StW list generated here using CDC isolate data as a proxy for human illnesses. As shown in Figure 2, HPSi ≥ 5 and both SoC
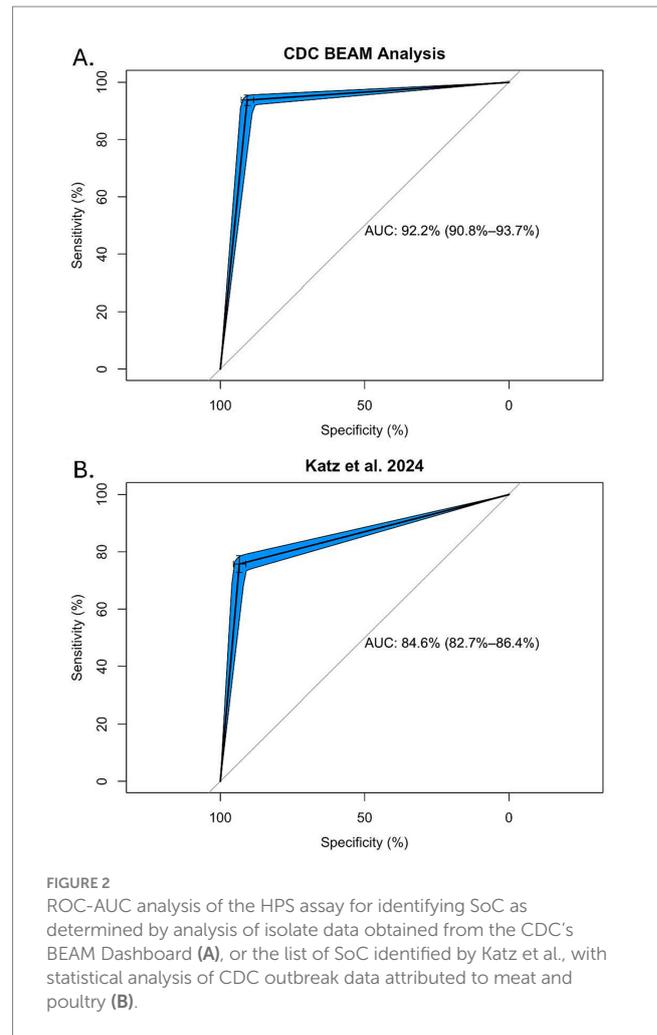

FIGURE 2
ROC-AUC analysis of the HPS assay for identifying SoC as determined by analysis of isolate data obtained from the CDC's BEAM Dashboard **(A)**, or the list of SoC identified by Katz et al., with statistical analysis of CDC outbreak data attributed to meat and poultry **(B)**.

definitions had high concordance, with an AUC of 92.2% for the CDC BEAM isolate SoC/StW list (Figure 2A), and 84.6% for the list of 21 SoC identified from outbreak analyses (Figure 2B).

4 Discussion

Historically, differences in *Salmonella* pathogenicity have been under appreciated and it was generally accepted that all serovars were equally capable of making humans sick. However, genomic sequencing efforts over the past two decades and subsequent comparative genomic analyses have revealed substantial differences in virulence gene content both among and within serovars (den Bakker et al., 2011; Suez et al., 2013; Worley et al., 2018; Cheng et al., 2019; Rakov et al., 2019; Fenske et al., 2023). These findings, in combination with an increased understanding that a relatively small subset of serovars contribute to the majority of illnesses, have resulted in a shift in the approach being considered for controlling *Salmonella*, from prevalence-based methods to techniques that assess both contamination level and pathogenicity level of the contaminating serovar [United States Department of Agriculture Food Safety and Inspection Service, 2022; National Advisory Committee on Microbiological Criteria in Foods (NACMCF), 2024], with the goal of removing products contaminated with higher levels of *Salmonella* or with serovars of greater significance for human health (i.e., SoC).

With this new approach, FSIS aims to improve public health by reducing SoC in the products they regulate, which the Interagency Food Safety Analytics Collaboration (IFSAC) estimates contribute to 38% of foodborne *Salmonella* illnesses annually (United States Department of Agriculture Food Safety and Inspection Service, 2020). It is anticipated that successful implementation of this approach will help the U.S. reach the Healthy People 2030 goal of reducing *Salmonella* infection incidence from 15.3 cases per 100,000 to 11.5 per 100,000 (Office of Disease Prevention and Health Promotion, 2023). For this to work however, new testing methods are needed for the rapid identification of more pathogenic versions of *Salmonella*. While general knowledge of the differences in *Salmonella* virulence gene content has increased over the past decade, specific information on a limited set of virulence genes that could be harnessed for rapid identification of more pathogenic versions of *Salmonella* has been lacking. Here we have described a multiplex PCR assay that targets 8 virulence genes and facilitates the identification of highly pathogenic *Salmonella* (HPS).

In silico determination of the presence of these 8 genes in 2,129 closed genome sequences in genbank showed that 7 of the top 10 SoC had an HPSi ≥ 5, with 4 of the top 5 serovars having HPSi = 7 or 8 (Table 3 and Supplementary Table S1). Multiplex PCR assays of 1,303 FSIS isolates showed half had an HPSi ≥ 5, with 97% (635/652) of those identified as SoC. Collectively, 19 serovars on the SoC/StW list had HPSi ≥ 5, and contributed on average 63.4% of the total isolates characterized by CDC each year. Sensitivity and specificity analyses showed the HPS assay has a very strong ability to identify SoC/StW (specificity = 96.3%, sensitivity = 73.5%), with most misclassifications being isolates with HPSi ≤4 that were SoC/StW. There were 13 serotypes in this group which was dominated by Anatum (n = 78), Montevideo (n = 36), Uganda (n = 26), Agona and Senftenberg each with 21 isolates, and Reading (n = 19). This finding highlights a limitation of the HPS assay and suggests these 13 serotypes either possess alternative virulence genes, yet to be identified, that allow them to cause widespread illness and require further investigation, or their presence on the SoC/StW list is due to their increased prevalence or increased contamination levels in food products that are frequently consumed raw, such as fruits and vegetables. In keeping with the latter possibility, serovars Anatum, Montevideo and Agona have been attributed to several large multi-state outbreaks connected with contaminated fruits, seeds, or dry food products such as cereals (Centers for Disease Control and Prevention (CDC), 2010; Miller et al., 2019; Hoffmann et al., 2020; Coulombe and Tamber, 2022). To decrease illnesses attributed to these types of vectors, methods for rapid detection of *Salmonella* contamination levels will be needed, as contamination level and distribution in products that commonly are consumed raw, is likely a greater contributor to illnesses than the presence of a particular cadre of virulence genes.

Salmonella SoC with HPSi ≤ 4, that were notably absent or rarely detected in FSIS enrichments were also observed and included serovars Javiana, Oranienburg, Mississippi, Bareilly, Sandiego, Poona, and Paratyphi B. Their absence in the FSIS dataset suggests that food animal production is not a common source of these serovars, and that the risk to public health comes from other sources. Consistent with this notion, these serovars have been associated with outbreaks in fresh produce, dairy products, shell eggs, well-water, and contact with animals (Kumar et al., 2018; Mukherjee et al., 2019; McGeoch et al., 2024; Mitchell et al., 2024). As mentioned above, there is a need for further investigation to identify targets for the rapid detection of these serovars, as collectively they are attributed to 17% of human illness isolates.

Screening new and legacy *Salmonella* WGS data in the NCBI pathogen detection browser for genes that contribute to *Salmonella* pathogenicity in the human host presents the opportunity for predictive modeling and pathogen control (Fenske et al., 2023; Petrin et al., 2023). An example of this was the finding in our *in silico* BLAST analysis, of 87 Enteritidis genome sequences that showed evidence of containing the *sseK3* gene. None of the 109 FSIS Enteritidis strains assayed were observed to amplify this gene. Examining the origin of the strains sequenced showed they were isolated primarily from Asia and Canada (60 and 24%, respectively). Given that *sseK3* has been shown to be carried on GIFSY-1 or ST64B prophage, this may represent a new or emerging lineage of Enteritidis that has acquired a new prophage. Recognizing the value of this *in silico* screening approach for identifying emerging pathogens, NCBI has recently added the genes targeted by the HPS assay to their AMRFinder+ reference gene database [(Feldgarden et al., 2021); https://ftp.ncbi.nlm.nih.gov/pathogen/Antimicrobial_resistance/Data/2024-05-02.2/changes.txt] so that all new *Salmonella* genome sequences in the Pathogen Detection Database[5] are screened for these virulence genes.

The new approaches being adopted for targeting and identification of NTS have stirred much conversation on how the *Salmonella* scientific community moves forward regarding typing. Do we continue to serotype *Salmonella*? (a legacy typing method with low resolution, high-labor and high-cost that is complicated by the availability of reagents and the effect of genetic drift of the proteins targeted, and by the existence of polyphyletic serovars with different pathogenicities). Or do we adopt a SNP based typing system such as cgMLST to understand *Salmonella* relatedness and conduct traceback investigations (Yoshida et al., 2016; Zhou et al., 2021; Achtman et al., 2022)? Or, given the increased knowledge gained by mining whole genome sequence data of hundreds of thousands of *Salmonella* isolates, do we develop a typing scheme based on the carriage of particular virulence genes (a vgMLST?) to help us rapidly identify SoC or emerging serovars that have acquired new tools, allowing them to expand the niches they occupy and perhaps become better pathogens? At present, there is a tug of war between "old and new" typing methods, and what evolves will likely be a combination that links the two. But regardless of how we proceed with characterizing and typing *Salmonella*, the data presented lay the foundation for molecular assays that will facilitate rapid detection of SoC and improve human health by reducing exposure to this foodborne foe.

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.

Author contributions

Funding

Acknowledgments

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.

Generative AI statement

The authors declare that no Gen AI was used in the creation of this manuscript.

Publisher's note

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2024.1504621/full#supplementary-material

References

Achtman, M., Zhou, Z., Charlesworth, J., and Baxter, L. (2022). EnteroBase: hierarchical clustering of 100,000s of bacterial genomes into species/subspecies and populations. *Philos. Trans. R. Soc. B* 377:20210240. doi: 10.1098/rstb.2021.0240

Altman, D. G., and Bland, J. M. (1994). Diagnostic tests 1: sensitivity and specificity. *BMJ* 308:1552. doi: 10.1136/bmj.308.6943.1552

Araujo-Garrido, J. L., Bernal-Bayard, J., and Ramos-Morales, F. (2020). Type III secretion effectors with arginine n-glycosyltransferase activity. *Microorganisms* 8:357. doi: 10.3390/microorganisms8030357

Baümler, A. J., and Heffron, F. (1995). Identification and sequence analysis of lpfABCDE, a putative fimbrial operon of *Salmonella typhimurium*. *J. Bacteriol.* 177, 2087–2097. doi: 10.1128/JB.177.8.2087-2097.1995

BLAST: Basic Local Alignment Search Tool (n.d.). Available at: https://blast.ncbi.nlm.nih.gov/Blast.cgi (accessed September 26, 2024).

Bosilevac, J. M., Guerini, M. N., Kalchayanand, N., and Koohmaraie, M. (2009). Prevalence and characterization of salmonellae in commercial ground beef in the United States. *Appl. Environ. Microbiol.* 75, 1892–1900. doi: 10.1128/AEM.02530-08

Brichta-Harhay, D. M., Arthur, T. M., Bosilevac, J. M., Kalchayanand, N., Shackelford, S. D., Wheeler, T. L., et al. (2011). Diversity of multidrug-resistant *Salmonella enterica* strains associated with cattle at harvest in the United States. *Appl. Environ. Microbiol.* 77, 1783–1796. doi: 10.1128/AEM.01885-10

Brichta-Harhay, D. M., Guerini, M. N., Arthur, T. M., Bosilevac, J. M., Kalchayanand, N., Shackelford, S. D., et al. (2008). *Salmonella* and *Escherichia coli* O157:H7 contamination on hides and carcasses of cull cattle presented for slaughter in the United States: an evaluation of prevalence and bacterial loads by Immunomagnetic separation and direct plating methods. *Appl. Environ. Microbiol.* 74, 6289–6297. doi: 10.1128/AEM.00700-08

Centers for Disease Control and Prevention (CDC) (2010). Salmonella montevideo infections associated with salami products made with contaminated imported black and red pepper --- United States, July 2009-April 2010. *MMWR. Morbidity and mortality weekly report*. 59, 1647–1650.

Centers for Disease Control and Prevention (2018). National Salmonella Surveillance Annual Report, 2016. *CDC*. Atlanta, Georgia: US Department of Health and Human Services. 1–87. Available at: https://www.census.gov/geo/pdfs/maps-data/maps/.

Centers for Disease Control and Prevention (2022). BEAM dashboard. Available at: https://www.cdc.gov/ncezid/dfwed/BEAM-dashboard.html (Accessed January 29, 2024).

Centers for Disease Control and Prevention (2023). National Outbreak Reporting System. Available at: https://www.cdc.gov/norsdashboard/ (Accessed August 31, 2023).

Cheng, R. A., Eade, C. R., and Wiedmann, M. (2019). Embracing diversity: differences in virulence mechanisms, disease severity, and host adaptations contribute to the success of nontyphoidal salmonella as a foodborne pathogen. *Front. Microbiol.* 10:1368. doi: 10.3389/fmicb.2019.01368

Collins, J. P., Shah, H. J., Lowell Weller, D., Ray, L. C., Smith, K., McGuire, S., et al. (2022). Preliminary incidence and trends of infections caused by pathogens transmitted commonly through food - foodborne diseases active surveillance network, 10 U.S. Sites, 2016-2021. *Morbidity Mortality Weekly Report*. 71, 1260–1264. doi: 10.15585/mmwr.mm7140a2

Coulombe, G., and Tamber, S. (2022). Salmonella enterica Outbreaks Linked to the Consumption of Tahini and Tahini-Based Products. *Microorganisms*. 10:2299. doi: 10.3390/microorganisms10112299

Darling, A. C. E., Mau, B., Blattner, F. R., and Perna, N. T. (2004). Mauve: multiple alignment of conserved genomic sequence with rearrangements. *Genome Res.* 14, 1394–1403. doi: 10.1101/gr.2289704

De Meyer, M., Fijalkowski, I., Jonckheere, V., De Sutter, D., Eyckerman, S., and Van Damme, P. (2021). Capturing salmonella SspH2 host targets in virus-like particles. *Front. Med.* 8:725072. doi: 10.3389/fmed.2021.725072

den Bakker, H. C., Moreno Switt, A. I., Govoni, G., Cummings, C. A., Ranieri, M. L., Degoricija, L., et al. (2011). Genome sequencing reveals diversification of virulence factor content and possible host adaptation in distinct subpopulations of *Salmonella enterica. BMC Genomics* 12:425. doi: 10.1186/1471-2164-12-425

Echeita, M. A., Herrera, S., Garaizar, J., and Usera, M. A. (2002). Multiplex PCR-based detection and identification of the most common salmonella second-phase flagellar antigens. *Res. Microbiol.* 153, 107–113. doi: 10.1016/S0923-2508(01)01295-5

EFSA and ECDC (2021). The European Union one health 2019 zoonoses report. *EFSA J.* 19:286. doi: 10.2903/j.efsa.2021.6406

Feldgarden, M., Brover, V., Gonzalez-Escalona, N., Frye, J. G., Haendiges, J., Haft, D. H., et al. (2021). AMRFinderPlus and the reference gene Catalog facilitate examination of the genomic links among antimicrobial resistance, stress response, and virulence. *Sci. Rep.* 11:12728. doi: 10.1038/s41598-021-91456-0

Fenske, G. J., Pouzou, J. G., Pouillot, R., Taylor, D. D., Costard, S., and Zagmutt, F. J. (2023). The genomic and epidemiological virulence patterns of *Salmonella enterica* serovars in the United States. *PLoS One* 18:e0294624. doi: 10.1371/journal.pone.0294624

Gao, F., and Zhang, C. T. (2008). Ori-finder: a web-based system for finding oriCs in unannotated bacterial genomes. *BMC Bioinformatics* 9:79. doi: 10.1186/1471-2105-9-79

Godínez-Oviedo, A., Tamplin, M. L., Bowman, J. P., and Hernández-Iturriaga, M. (2020). *Salmonella enterica* in Mexico 2000-2017: epidemiology, antimicrobial resistance, and prevalence in food. *Foodborne Pathog. Dis.* 17, 98–118. doi: 10.1089/fpd.2019.2627

Gonzales, A. M., Wilde, S., and Roland, K. L. (2017). New insights into the roles of long polar fimbriae and stg fimbriae in salmonella interactions with enterocytes and M cells. *Infect. Immun.* 85:9. doi: 10.1128/IAI.00172-17

Grabe, G. J., Zhang, Y., Przydacz, M., Rolhion, N., Yang, Y., Pruneda, J. N., et al. (2016). The salmonella effector SpvD is a cysteine hydrolase with a Serovar-specific polymorphism influencing catalytic activity, suppression of immune responses, and bacterial virulence. *J. Biol. Chem.* 291, 25853–25863. doi: 10.1074/JBC.M116.752782

Groisman, E. A., and Ochman, H. (1997). How salmonella became a pathogen. *Trends Microbiol.* 5, 343–349. doi: 10.1016/S0966-842X(97)01099-8

Guiney, D. G., Fang, F. C., Krause, M., Libby, S., Buchmeier, N. A., Fierer, J., et al. (1995). Biology and clinical significance of virulence plasmids in salmonella Serovars. *Clin. Infect. Dis.* 21, S146–S151. doi: 10.1093/clinids/21.Supplement_2.S146

Gutema, F. D., Agga, G. E., Abdi, R. D., De Zutter, L., Duchateau, L., and Gabriël, S. (2019). Prevalence and serotype diversity of salmonella in apparently healthy cattle: systematic review and meta-analysis of published studies, 2000-2017. *Front. Vet. Sci.* 6:102. doi: 10.3389/fvets.2019.00102

Hansmeier, N., Miskiewicz, K., Elpers, L., Liss, V., Hensel, M., and Sterzenbach, T. (2017). Functional expression of the entire adhesiome of *Salmonella enterica* serotype Typhimurium. *Sci. Rep.* 71, 7, 1–12. doi: 10.1038/s41598-017-10598-2

Harhay, D., Brader, K. D., Bono, J., Harhay, G. P., Bosilevac, M., and Katz, T. S. (2024). Protocol for Identifying Highly Pathogenic Salmonella Using the HPS Multiplex PCR Assay. doi: 10.17504/protocols.io.6qpvr3on3vmk/v1

Hensel, M. (2000). Salmonella pathogenicity island 2. *Mol. Microbiol.* 36, 1015–1023. doi: 10.1046/j.1365-2958.2000.01935.x

Herrera-León, S., McQuiston, J. R., Usera, M. A., Fields, P. I., Garaizar, J., and Echeita, M. A. (2004). Multiplex PCR for distinguishing the most common phase-1 flagellar antigens of salmonella spp. *J. Clin. Microbiol.* 42, 2581–2586. doi: 10.1128/JCM.42.6.2581-2586.2004

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

Hoffmann, M., Miller, J., Melka, D., Allard, M. W., Brown, E. W., and Pettengill, J. B. (2020). Temporal dynamics of salmonella enterica subsp. *enterica* serovar agona isolates from a recurrent multistate outbreak. *Front. Microbiol.* 11:478. doi: 10.3389/fmicb.2020.00478

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

Jechalke, S., Schierstaedt, J., Becker, M., Flemer, B., Grosch, R., Smalla, K., et al. (2019). Salmonella establishment in agricultural soil and colonization of crop plants depend on soil type and plant species. *Front. Microbiol.* 10:967. doi: 10.3389/fmicb.2019.00967

Jennings, E., Esposito, D., Rittinger, K., and Thurston, T. L. M. (2018). Structure–function analyses of the bacterial zinc metalloprotease effector protein GtgA uncover key residues required for deactivating NF-κB. *J. Biol. Chem.* 293, 15316–15329. doi: 10.1074/JBC.RA118.004255

Katz, T. S., Harhay, D. M., Schmidt, J. W., and Wheeler, T. L. (2024). Identifying a list of salmonella serotypes of concern to target for reducing risk of salmonellosis. *Front. Microbiol.* 15:1307563. doi: 10.3389/fmicb.2024.1307563

Kuhn, M. (2008). Building predictive models in R using the caret package. *J. Stat. Softw.* 28:5. doi: 10.18637/jss.v028.i05

Kumar, G., Williams, R. C., Sriranganathan, N., Boyer, R. R., and Eifert, J. D. (2018). Survival of tomato outbreak associated *salmonella* serotypes in soil and water and the role of biofilms in abiotic surface attachment. *Foodborne Pathog. Dis.* 15, 548–553. doi: 10.1089/fpd.2017.2416

Kunze, D. J., Loneragan, G. H., Platt, T. M., Miller, M. F., Besser, T. E., Koohmaraie, M., et al. (2008). *Salmonella enterica* burden in harvest-ready cattle populations from the southern high plains of the United States. *Appl. Environ. Microbiol.* 74, 345–351. doi: 10.1128/AEM.02076-07

Kuria, J. K. N. (2024). "Salmonellosis in food and companion animals and its public health importance" in Salmonella – perspectives for low-cost prevention, control and treatment. eds. H. Huang and S. Naushad (London: Intechopen).

Labriola, J. M., Zhou, Y., and Nagar, B. (2018). Structural analysis of the bacterial effector AvrA identifies a critical helix involved in substrate recognition. *Biochemistry* 57, 4985–4996. doi: 10.1021/acs.biochem.8b00512

Lee, P. Y., Costumbrado, J., Hsu, C. Y., and Kim, Y. H. (2012). Agarose gel electrophoresis for the separation of DNA fragments. *J. Vis. Exp.* doi: 10.3791/3923-v

Loneragan, G. H., Thomson, D. U., McCarthy, R. M., Webb, H. E., Daniels, A. E., Edrington, T. S., et al. (2012). Salmonella diversity and burden in cows on and culled from dairy farms in the Texas high plains. *Foodborne Pathog. Dis.* 9, 549–555. doi: 10.1089/fpd.2011.1069

Marshall, K. E., Cui, Z., Gleason, B. L., Hartley, C., Wise, M. E., Bruce, B. B., et al. (2024). An approach to describe salmonella serotypes of concern for outbreaks: using burden and trajectory of outbreak-related illnesses associated with meat and poultry. *J. Food Prot.* 87:100331. doi: 10.1016/j.jfp.2024.100331

Maurischat, S., Baumann, B., Martin, A., and Malorny, B. (2015). Rapid detection and specific differentiation of *Salmonella enterica* subsp. enterica Enteritidis, Typhimurium and its monophasic variant 4,[5],12:i:- by real-time multiplex PCR. *Int. J. Food Microbiol.* 193, 8–14. doi: 10.1016/j.ijfoodmicro.2014.10.004

McGeoch, L. J., Hoban, A., Sawyer, C., Rabie, H., Painset, A., Browning, L., et al. (2024). *Salmonella* Saintpaul outbreak associated with cantaloupe consumption, the United Kingdom and Portugal, September to November 2023. *Epidemiol. Infect.* 152:e78. doi: 10.1017/S0950268824000670

Merkushova, A. V., Shikov, A. E., Nizhnikov, A. A., and Antonets, K. S. (2023). For someone, you are the whole world: host-specificity of *Salmonella enterica. Int. J. Mol. Sci.* 24:13670. doi: 10.3390/ijms241813670

Miao, E. A., Scherer, C. A., Tsolis, R. M., Kingsley, R. A., Adams, L. G., Bäumler, A. J., et al. (1999). *Salmonella typhimurium* leucine-rich repeat proteins are targeted to the SPI1 and SPI2 type III secretion systems. *Mol. Microbiol.* 34, 850–864. doi: 10.1046/J.1365-2958.1999.01651.X

Miller, D., Yao, K., Whitney, B., Sanchez Leon, M., Tallent, S., and Hoffmann, M. (2019). Complete genome sequences of two salmonella enterica subsp. enterica serovar anatum strains isolated from papaya. *Microbiology resource announcements.* 8:e01071-19. doi: 10.1128/MRA.01071-19

Mitchell, M. R., Kirchner, M., Schneider, B., McClure, M., Neil, K. P., Madad, A., et al. (2024). Multistate outbreak of salmonella Oranienburg infections linked to bulb onions imported from Mexico – United States, 2021. *Food Control* 160:110325. doi: 10.1016/j.foodcont.2024.110325

Mukherjee, N., Nolan, V. G., Dunn, J. R., and Banerjee, P. (2019). Sources of human infection by *Salmonella enterica* serotype Javiana: a systematic review. *PLoS One* 14:e0222108. doi: 10.1371/journal.pone.0222108

National Advisory Committee on Microbiological Criteria in Foods (NACMCF) (2024). Response to questions posed by the Food Safety and Inspection Service: enhancing salmonella control in poultry products. *J. Food Prot.* 87:100168. doi: 10.1016/j.jfp.2023.100168

National Enteric Surveillance Program (2022). Available at: https://www.canada.ca/en/public-health/services/publications/food-nutrition/national-enteric-surveillance-program-annual-summary-2022.html

Nucera, D. M., Maddox, C. W., Hoien-Dalen, P., and Weigel, R. M. (2006). Comparison of API 20E and invA PCR for identification of *Salmonella enterica* isolates from swine production units. *J. Clin. Microbiol.* 44, 3388–3390. doi: 10.1128/JCM.00972-06

Office of Disease Prevention and Health Promotion (2023). Reduce infections caused by salmonella — FS-04 – Healthy People 2030 | health.gov. Available at: https://health.gov/healthypeople/objectives-and-data/browse-objectives/foodborne-illness/reduce-infections-caused-salmonella-fs-04 (accessed September 29, 2024).

Paton, A. W., and Paton, J. C. (1998). Detection and characterization of Shiga Toxigenic *escherichia coli* by using multiplex PCR assays for stx1, stx2, eaeA, Enterohemorrhagic E. Coli hlyA, rfbO111, and rfbO157. *J. Clin. Microbiol.* 36, 598–602. doi: 10.1128/JCM.36.2.598-602.1998

Petrin, S., Wijnands, L., Benincà, E., Mughini-Gras, L., Delfgou-van Asch, E. H. M., Villa, L., et al. (2023). Assessing phenotypic virulence of *Salmonella enterica* across serovars and sources. *Front. Microbiol.* 14:1184387. doi: 10.3389/fmicb.2023.1184387

R Core Team (2023). R: A language and environment for statistical computing. Available at: www.R-project.org (Accessed June 23, 2022).

Rahn, K., De Grandis, S. A., Clarke, R. C., Mcewen, S. A., Galin, 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

Rakov, A. V., Mastriani, E., Liu, S. L., and Schifferli, D. M. (2019). Association of Salmonella virulence factor alleles with intestinal and invasive serovars. *BMC Genomics* 20:429. doi: 10.1186/s12864-019-5809-8

Rolhion, N., Furniss, R. C. D., Grabe, G., Ryan, A., Liu, M., Matthews, S. A., et al. (2016). Inhibition of nuclear transport of NF-κB p65 by the salmonella type III secretion system effector SpvD. *PLoS Pathog.* 12:e1005653. doi: 10.1371/JOURNAL.PPAT.1005653

Sambrook, J., and Russell, D. W. (2006). Detection of DNA in agarose gels. *Cold Spring Harb. Protoc.* 2006:pdb.prot4022. doi: 10.1101/pdb.prot4022

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

Schmidt, J. W., Brichta-Harhay, D. M., Kalchayanand, N., Bosilevac, J. M., Shackelford, S. D., Wheeler, T. L., et al. (2012). Prevalence, enumeration, serotypes, and antimicrobial resistance phenotypes of *Salmonella enterica* isolates from carcasses at two large United States pork processing plants. *Appl. Environ. Microbiol.* 78, 2716–2726. doi: 10.1128/AEM.07015-11

Stanaway, J. D., Parisi, A., Sarkar, K., Blacker, B. F., Reiner, R. C., Hay, S. I., et al. (2019). The global burden of non-typhoidal salmonella invasive disease: a systematic analysis for the global burden of disease study 2017. *Lancet Infect. Dis.* 19, 1312–1324. doi: 10.1016/S1473-3099(19)30418-9

Suez, J., Porwollik, S., Dagan, A., Marzel, A., Schorr, Y. I., Desai, P. T., et al. (2013). Virulence gene profiling and pathogenicity characterization of non-Typhoidal salmonella accounted for invasive disease in humans. *PLoS One* 8:e58449. doi: 10.1371/journal.pone.0058449

Takemura, M., Haneda, T., Idei, H., Miki, T., and Okada, N. (2021). A salmonella type III effector, PipA, works in a different manner than the PipA family effectors GogA and GtgA. *PLoS One* 16:e0248975. doi: 10.1371/JOURNAL.PONE.0248975

United States Department of Agriculture Food Safety and Inspection Service (2020). Roadmap to reducing *salmonella*. Available at: https://www.fsis.usda.gov/sites/default/files/media_file/2020-12/FSISRoadmaptoReducingSalmonella.pdf (Accessed September 26, 2024).

United States Department of Agriculture Food Safety and Inspection Service (2022). Proposed regulatory framework to reduce *salmonella* illnesses attributable to poultry. Available at: https://www.fsis.usda.gov/sites/default/files/media_file/documents/FINAL-Salmonella-Framework-10112022-508-edited.pdf (Accessed September 26, 2024).

United States Department of Agriculture Food Safety and Inspection Service. (2024) *Salmonella* framework for raw poultry products. Federal Register, Vol. 89(152):64678–64748. Docket No. FSIS-2023-0028. Available at: https://www.fsis.usda.gov/sites/default/files/media_file/documents/FSIS-2023-0028.pdf (Accessed September 26, 2024).

Waldner, L. L., MacKenzie, K. D., Köster, W., and White, A. P. (2012). From exit to entry: long-term survival and transmission of salmonella. *Pathogens* 1, 128–155. doi: 10.3390/pathogens1020128

Wang, M., Qazi, I. H., Wang, L., Zhou, G., and Han, H. (2020). Salmonella virulence and immune escape. *Microorganisms* 8:407. doi: 10.3390/microorganisms8030407

Worley, J., Meng, J., Allard, M. W., Brown, E. W., and Timme, R. E. (2018). *Salmonella enterica* phylogeny based on whole-genome sequencing reveals two new clades and novel patterns of horizontally acquired genetic elements. *MBio* 9:e0230318. doi: 10.1128/mBio.02303-18

Yin, C., Liu, Z., Xian, H., Jiao, Y., Yuan, Y., Li, Y., et al. (2020). AvrA exerts inhibition of NF-κB pathway in its Naïve salmonella serotype through suppression of p-JNK and Beclin-1 molecules. *Int. J. Mol. Sci.* 21:6063. doi: 10.3390/IJMS21176063

Yoshida, C., Kruczkiewicz, P., Laing, C., Lingohr, R., Gannon, V., Nash, J., 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

Zhengwei, F., Xiujuan Zhou, X., and Wang, X. S. (2023). Development of a 3-plex droplet digital PCR for identification and absolute quantification of Salmonella and its two important serovars in various food samples. *Food Control* 145:109465. doi: 10.1016/j.foodcont.2022.109465

Zhou, Z., Charlesworth, J., and Achtman, M. (2021). HierCC: a multi-level clustering scheme for population assignments based on core genome MLST. *Bioinformatics* 37, 3645–3646. doi: 10.1093/bioinformatics/btab234