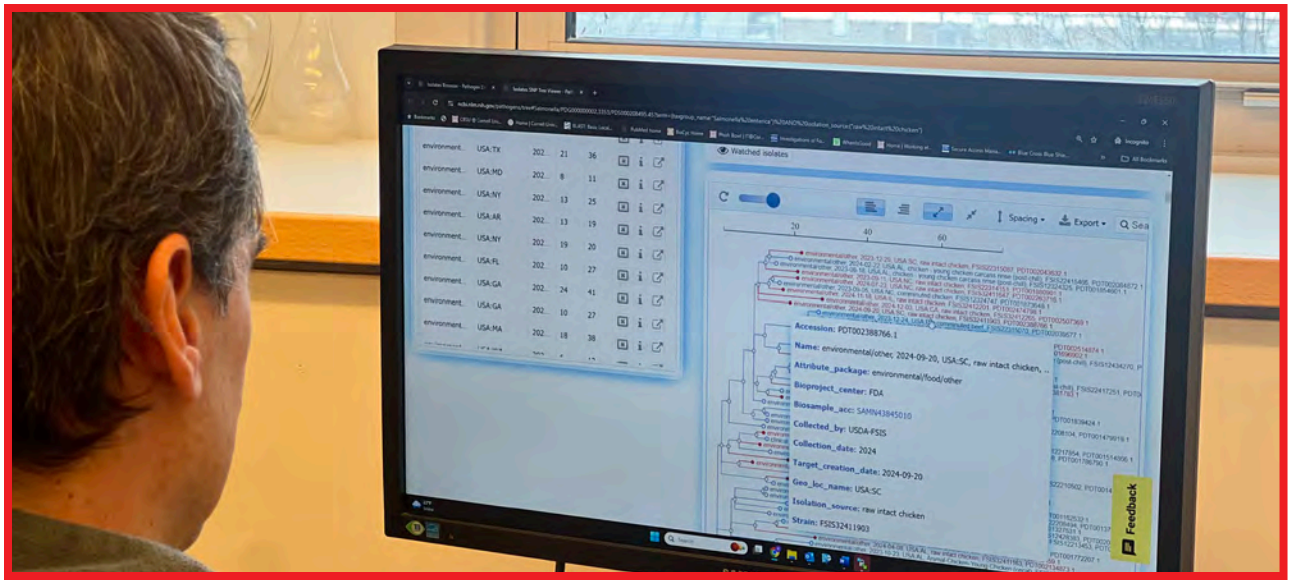
PEER-REVIEWED ARTICLE

Food Protection Trends, Vol 45, No. 2, p. 103–122
https://doi.org/10.4315/FPT-24-025

2900 100th Street, Suite 309, Des Moines, IA 50322-3855, USA

Shrinidhi S. Joshi,[1] Martin Wiedmann,[2] and Alexa R. Cohn[2]*

[1]Stop Foodborne Illness, Chicago, IL 60640, USA
[2]Dept. of Food Science, Cornell University, Ithaca, NY 14853, USA



Comparison of Serovar Reporting Frequency in the United States Department of Agriculture Food Safety and Inspection Service and the National Center for Biotechnology Information Pathogen Detection Databases of *Salmonella* Strains Isolated from Livestock

ABSTRACT

Salmonella is a significant threat to human health, causing an estimated 1.35 million illnesses each year in the United States. There is increasing consensus that regulatory strategies and industry efforts that target serovars of public health concern are essential to reduce human salmonellosis, and it is important to understand the data available to assess serovar distribution among food sources. We analyzed isolate data from 2015 to 2020 for 21 serovars common in food animals in public data sets available through the Food Safety and Inspection Service (FSIS) and National Center for Biotechnology Information Pathogen Detection database (NCBI PD). Following defined criteria, we obtained metadata from 7,812 and 12,248 *Salmonella* isolates on the NCBI and FSIS websites, respectively. Our analyses found significant differences in serovar distribution between (i) FSIS data and NCBI data contributed by non-FSIS sources and between (ii) different isolation sources for a commodity. Specifically, we found isolation patterns of certain serovars (e.g., *Salmonella* Infantis) coincided with reported outbreaks, and more serovars were overrepresented in the NCBI PD data set. Although our results suggest biases in *Salmonella* serovar distribution sets, we found consistent trends across data sets that indicate the value of public data sets for informing future subtype-specific *Salmonella* regulations and control efforts.

INTRODUCTION

Salmonella causes an estimated 1.35 million infections annually in the United States (*38*). The species *Salmonella enterica* includes six recognized subspecies, the most clinically relevant of which, subspecies enterica, includes >1,500 serovars (*22*). Although *Salmonella* infections typically cause mild illness (e.g., nausea, diarrhea, stomach cramps, and vomiting), some illnesses can become severe enough that they require hospitalization. Importantly, there is substantial evidence that not all *Salmonella* are equally likely to cause human disease, and some *Salmonella* subtypes (e.g., *Salmonella* Kentucky sequence type [ST]152, which is common in poultry in the United States) are well documented to have reduced likelihood to cause human disease (*3*).

*Author for correspondence: Email: arc326@cornell.edu

In an effort to decrease the prevalence of *Salmonella*, the U.S. government set a Healthy People 2020 objective to reduce the incidence of salmonellosis to 11.4 laboratory-confirmed infections per 100,000 people *(33)*. However, the United States failed to meet this goal by 2020, and, instead, the incidence of human illness due to *Salmonella* increased from the 2006 and 2008 baseline of 15 cases per 100,000 people to the latest 2016 to 2018 baseline of 15.3 cases per 100,000 people *(34)*. With a lack of meaningful progress, the federal government set the new Healthy People 2030 goal to reduce the incidence to 11.5 per 100,000 people. Specifically for raw meat and poultry, targeted approaches to decrease the prevalence of *Salmonella* serovars most likely to cause human disease represents one proposed strategy that could help effectively meet the new 2030 goal.

Although the number of annual outbreaks linked to poultry has not decreased, there has been some progress in reducing the overall frequency of *Salmonella*-positive poultry product contamination in the United States *(39, 40)*. Although this apparent discrepancy could be due to a number of factors, one important possibility is that certain serovars cause a disproportionate number of infections and that the prevalence reductions achieved may have focused on serovars that were common but represented limited public health concern (e.g., *Salmonella* serovar Kentucky) *(11)*. Hence, there is increasing consensus that treating all *Salmonella* serovars as representing an equal public health risk may not be successful in decreasing incidence of human salmonellosis infections, particularly those linked to raw meat and poultry. As an alternative, it has been proposed that risk-based U.S. Department of Agriculture Food Safety and Inspection Service (FSIS) regulations and standards should focus on targeting serovars of greatest public health importance *(12)*. Defining serovars of public health concerns is also important, as there is some evidence that vaccination of live animals may be able to reduce human infection caused by those serovars targeted by vaccination. Circumstantial evidence supporting this is provided by data that indicate that *Salmonella* Typhimurium human illness and product contamination have decreased over the past 20 years after the poultry industry began vaccinating flocks against *Salmonella* Typhimurium *(1, 14, 16)*. Although defining *Salmonella* serovars that differ in public health relevance and likelihood of causing human disease is clearly important for modern risk-based food safety systems, particularly those targeting *Salmonella* in raw meat and poultry, the scientific definition of serovars and subtypes that differ in virulence (including development of quantitative measures that can be used to compare the relative virulence of different subtypes) remains a challenge. It is, however, clear that the increasingly large whole genome sequencing (WGS) databases for *Salmonella* (e.g., in National Center for Biotechnology Information [NCBI] Pathogen Detection [PD]) provide one opportunity to better identify, define, and characterize virulence differences among *Salmonella* serovars and strains. Hence, we used two major data sources (i.e., NCBI PD and FSIS sampling reports) to identify the differences in counts of *Salmonella* serovars among data sources with the goals of (i) identifying potential biases and (ii) assessing whether these databases identify similar trends in *Salmonella* serovar prevalence by year and isolation source (e.g., feces versus ground chicken). Our results highlight how large data sets can help inform tracking of serovars prevalent in animal food sources.

MATERIALS AND METHODS

Selection of *Salmonella* serovars for inclusion in this study

We used FSI's *Salmonella* serovar quarterly sampling report for the 2020 fiscal year to determine which *Salmonella* serovars to include in our analysis *(43)*. For each food animal (i.e., cattle, chicken, turkey, swine), we tallied up the serovar isolate counts and chose the six serovars most common for a given food animal *(Table 1)*. Because some serovars were in the top six for more than one food animal, the final data set included 19 *Salmonella* serovars (Adelaide, Agona, Anatum, Cerro, Derby, Dublin, Enteritidis, Hadar, Heidelberg, I 4,[5],12:i:–, Infantis, Johannesburg, Kentucky, Montevideo, Muenchen, Muenster, Reading, Schwarzengrund, and Typhimurium). To ensure that the 2020 top serovars were not biased due to coronavirus disease 2019 disruptions, we repeated this process with FSIS's *Salmonella* serovar quarterly sampling report for the 2019 fiscal year *(Table 2)* *(41)*. The list based on the 2019 serovar report only differed from the list based on the 2020 serovar report by *Salmonella* serovars Cerro, Newport, and Uganda (*Salmonella* Cerro was the fifth most common serovar isolated from cattle in 2020 but was not in the top six serovars in 2019, *Salmonella* Newport was the sixth most common serovar isolated from cattle in 2019 but was not in the top six serovars in 2020, and *Salmonella* Uganda was the second most common serovar isolated from turkey in 2019 but was not in the top six turkey serovars in 2020). Due to a lack of major differences, we used the 2020 FSIS data to compile a list of serovars. *Salmonella* Javiana was added to our list as a control serovar, as it is not commonly found in foods of animal origin and is one of the top five most common serovars from human clinical infections *(7, 30)*. *Salmonella* Newport was also added to the list because it is a serovar of interest as one of the top five human pathogens; this yielded a total of 21 serovars that were analyzed. When reporting results, we reported observations on the top 10 serovars for a given animal host; this is possible as our final data set of 21 serovars included the top 10 serovars for each animal host.

NCBI data processing

We downloaded metadata from NCBI's PD database which, at the time (24 March 2022), included 429,048 *Salmonella enterica* isolates (ncbi.nlm.nih.gov/pathogen) *(31)*. We used the "filter" option to narrow our search

TABLE 1. Top *Salmonella* serovars from FSIS 2020 fiscal year

Salmonella serovar	No. of isolates (rank) among				
	Chicken	Cattle	Turkey	Swine	Total
Kentucky	1,118[a] (1)	0	0	0	1,118[a]
Enteritidis	1,064[a] (2)	0	0	0	1,064[a]
Infantis	794[a] (3)	0	38[a] (1)	141[a] (3)	973[a]
Schwarzengrund	327[a] (4)	0	0	0	327[a]
Typhimurium	194[a] (5)	8	20[a] (6)	10	232[a]
I 4,[5],12:i:–	13	2	13	178[a] (1)	206[a]
Anatum	0	35[a] (2)	6	147[a] (2)	188[a]
Johannesburg	37	2	0	79[a] (6)	118[a]
Derby	0	4	0	112[a] (4)	116[a]
Adelaide	0	0	0	105[a] (5)	105[a]
Montevideo	9	69[a] (1)	0	0	78[a]
Ohio	0	0	0	73	73
London	0	0	0	69	69
Heidelberg	64[a] (6)	0	0	0	64[a]
Uganda	0	0	15	44	59
Hadar	30	0	28[a] (4)	0	58[a]
Thompson	53	0	0	0	53
Muenchen	5	33 (3)[a]	9	0	47[a]
Rissen	0	0	0	44	44
Reading	0	3	38[a] (1)	0	41[a]
Agona	0	2	21[a] (5)	14	37[a]
Braenderup	30	0	0	0	30
Schwarzengrund	0	0	30[a] (3)	0	30[a]
Dublin	0	26 (4)[a]	0	0	26[a]
Cerro	0	23 (5)[a]	0	0	23[a]
Muenster	0	19 (6)[a]	0	0	19[a]
Senftenberg	0	0	19	0	19
Rough_O:r:1,5	18	0	0	0	18
Meleagridis	0	18	0	0	18
Newport	0	8	8	0	16
Worthington	0	0	0	11	11
Brandenburg	0	7	0	0	7
Mbandaka	0	5	0	0	5
Albany	0	0	4	0	4
Chailey	0	2	0	0	2
Give	0	2	0	0	2
IIIa [1],13,23:g,z51:–	0	0	2	0	2
					5,302

[a]To limit bias, the top six serovars for each animal source were selected. Duplicate serovars were removed. In each column, a indicate serovars that fell in the top six for that animal. In the total column, a indicates the final serovars selected.

TABLE 2. Top *Salmonella* serovars from FSIS 2019 fiscal year

Salmonella serovar	Chicken	Cattle	Turkey	Swine	Total
Kentucky	715[a] (1)	14	0	0	729[a]
Infantis	592[a] (2)	10	26[a] (4)	69[a] (2)	697[a]
Enteritidis	442[a] (3)	0	0	0	442[a]
Schwarzengrund	164[a] (4)	0	21[a] (5)	0	185[a]
Typhimurium	145[a] (5)	13	18	5	181[a]
Anatum	0	30[a] (2)	15	84[a] (1)	129[a]
Reading	0	0	118[a] (1)	7	125[a]
I 4,[5],12:i:–	35	0	2	47[a] (4)	84[a]
Montevideo	6	60[a] (1)	7	0	73[a]
Johannesburg	6	0	0	48[a] (3)	54[a]
Derby	0	4	0	45[a] (5)	49[a]
Muenchen	0	17[a] (4)	16	14	47[a]
Uganda	0	3	32[a] (2)	11	46[a]
Adelaide	0	0	0	40[a] (6)	40[a]
Agona	0	4	29[a] (3)	7	40[a]
Heidelberg	38[a] (6)	0	0	0	38[a]
Braenderup	31	0	0	0	31
Thompson	30	0	0	0	30
London	0	0	2	26	28
Cerro	0	15	0	6	21
Senftenberg	0	2	15	4	21
Dublin	0	19[a] (3)	0	0	19[a]
Hadar	0	0	19[a] (6)	0	19[a]
Muenster	0	17[a] (4)	2	0	19[a]
Ohio	0	0	0	17	17
Newport	0	16[a] (6)	0	0	16[a]
Rough_O:r:1,5	16	0	0	0	16
Mbandaka	0	11	0	0	11
Worthington	0	0	0	11	11
Litchfield	7	0	0	0	7
Meleagridis	0	7	0	0	7
Berta	0	0	2	4	6
Albany	0	0	5	0	5
4,[5],12:d:–	4	0	0	0	4
Eko	0	0	0	4	4
Give	0	4	0	0	4
6,7:g,m,s:e,n,z15	0	2	0	0	2
Blockley	0	2	0	0	2
					3,259

[a]To limit bias, the top six serovars for each animal source were selected. Duplicate serovars were removed. In each column, a indicates serovars that fell in the top six for that animal. In the total column, a indicates the final serovars selected.

results by location to only include isolates with "location" reported as "USA". Serovars were chosen based on the "computed type" category, which reports serovars predicted with SeqSero2 on the basis of the WGS data for a given isolate *(45)*. Finally, we filtered data by isolation source (e.g., "animal–cattle–dairy cow," "animal–chicken–young chicken," "bovine feces"). To select isolates from each of the four food animal sources included here (i.e., chicken, turkey, cattle, pigs), we used multiple search terms to ensure we had broad coverage (e.g., *Gallus gallus*, chicken, egg, hen for chicken) *(Table S1)*. After filtering, we exported the isolate metadata (organism group, Run #, strain, collection date, location, isolation source, isolation type, SNP cluster, computed type, serovars). After downloading metadata for the 21 serovars for each food animal, isolates that were collected before 2015 or after 2020 were removed. We removed these isolates because WGS became more widely implemented in surveillance systems in 2015 *(5)*. We also removed FSIS isolates by removing isolates that had the code FSIS within the strain name to prevent redundancy, as FSIS submits data to NCBI as well. NCBI isolates from feed were removed because they are not associated with final food products, and FSIS does not collect isolates from these sources. Because NCBI isolation sources are entered by individuals and are not uniform, we standardized the data by recoding the isolation sources *(Table S2)*. Terms that were similar, such as "beef feces," "cattle feces," and "*Bos taurus* feces," were all recoded (in this case to "cattle feces") to have a common name. Isolation sources that were vague, such as "cattle" or "dairy cow," were recoded (e.g., as "unspecified, cattle").

FSIS data processing

FSIS's laboratory sampling data from 2015 to 2020 were downloaded for cattle, chicken, and turkey *(44)*. Laboratory sampling data for swine were limited and only available starting in 2019; thus, swine data were not included in this study.

Statistical analyses

Statistical analyses were performed in R Studio *(36)*. We performed Fisher exact tests to determine over- and underrepresentation of serovars between NCBI and FSIS data sources using isolate count data from (i) FSIS's laboratory data and (ii) NCBI's PD data with FSIS submissions removed; this analysis was performed for cattle, chicken, and turkey. An odds ratio (OR) > 2 and $P < 0.05$ indicated overrepresentation in the NCBI database, while an OR < 0.5 and $P < 0.05$ indicated overrepresentation in the FSIS data set. Overrepresentation indicates that statistically, a serovar is more likely to be found in one database over the other (e.g., if an isolate is overrepresented in the NCBI data set, it is more likely to find that serovar in the NCBI data set than it would be to find it in the FSIS data set). In addition, separate chi-square tests were performed on isolation source data from FSIS's laboratory data (for cattle, chicken, and turkey) and NCBI's PD data (for cattle, chicken, swine, and turkey); this was followed by post hoc hypothesis testing with adjusted Pearson residuals if the chi-square test yielded a significant *P*-value ($P < 0.05$). For these analyses, samples where the isolation source was listed as a specific organ (e.g., heart, kidney, lymph node) were combined into a single isolation source category named "organs and tissues" *(Table S2)*. For NCBI data, isolation sources that were combined to create the category organs and tissues varied vastly; we thus also separated these data into each individual organ category and performed a chi-square test, followed by post hoc hypothesis testing with adjusted Pearson residuals, if the chi-square test yielded significant results. Analysis of swine isolation source data were only performed on the NCBI data because the collection time frame for FSIS's swine data (2016 to 2020) was different from the data for the other animal sources.

RESULTS

Serovar distribution for cattle isolates

The top 10 *Salmonella* serovars found in cattle between 2015 and 2020 were Anatum, Cerro, Dublin, Infantis, Kentucky, Montevideo, Muenster, Typhimurium, Newport, and Muenchen (*Fig. 1A* and *Table S3*). In the FSIS data set, the numbers of isolates representing *Salmonella* serovars Cerro, Muenchen, and Newport all increased in 2016 (increases of 26, 24, and 50% over 2015, respectively) and then decreased from 2017 to 2018 (decreases of 88, 76, and 100%). Similarly, the number of *Salmonella* Dublin isolates decreased by 94% from 2017 to 2018. Though numbers of *Salmonella* Montevideo isolates sharply dropped in 2018 (a 57% reduction of total isolates relative to 2017) and remained lower afterward, *Salmonella* Montevideo remained the most prevalent serovar among the FSIS cattle isolates throughout the study period, including from 2018 to 2020.

In the NCBI data set, there was a noticeable increase in numbers of isolates representing serovars Anatum (900%), Montevideo (374%), and Cerro (200%) between 2015 and 2016 (*Fig. 1B* and *Table S4*). Numbers of isolates representing serovars Anatum and Montevideo showed patterns that mimicked each other in that both spiked in 2016 as described above, and then decreased between 2016 and 2018 and increased in subsequent years. Interestingly, the number of *Salmonella* Cerro isolates on NCBI decreased by 91% from 2016 to 2020. Numbers of *Salmonella* Dublin isolates decreased by 48% in 2017 compared with 2016, but were consistently high from 2017 to 2019 as Dublin increased by 87% from 2017 to 2018 and then only saw a 1% decrease from 2018 to 2019. However, numbers of *Salmonella* Dublin isolates did decrease by 55% in 2020.

A comparison of serovar distributions among cattle isolates in the NCBI and FSIS database found that three and seven serovars are overrepresented in the NCBI and FSIS data, respectively, with (i) *Salmonella* serovars Dublin,

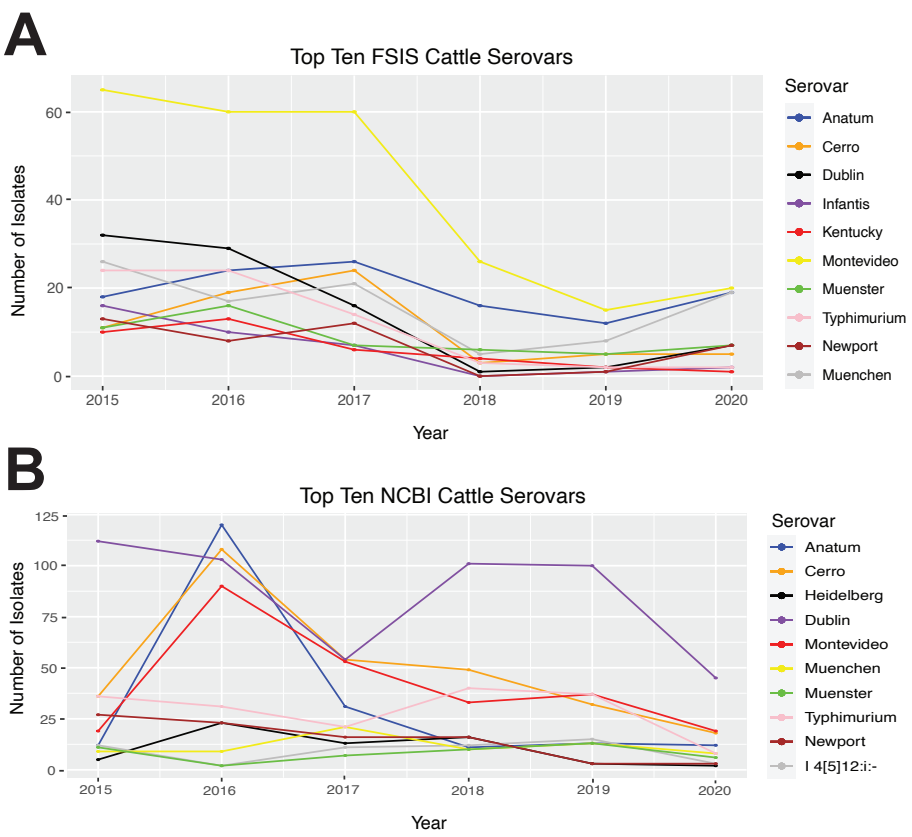


Figure 1. Top 10 cattle *Salmonella* serovars from 2015 to 2020. Line plots of *Salmonella* serovar trends over 6 years in cattle based on isolate counts from A, FSIS data and B, NCBI PD.

Cerro, and Heidelberg overrepresented in NCBI cattle data and (ii) *Salmonella* serovars Derby, Infantis, Johannesburg, Montevideo, Muenchen, Muenster, and Reading overrepresented in FSIS cattle data. *Salmonella* Heidelberg (OR = 10.5) was the serovar most highly overrepresented among the NCBI data, while *Salmonella* Johannesburg (OR = 0.0752) was the most highly overrepresented among the FSIS data *(Table 3)*.

A chi-square test was performed to better understand over- and underrepresentation of serovars in relation to isolation sources (e.g., feces, ground beef), followed by post hoc hypothesis testing with adjusted Pearson residuals for the chi-square tests that yielded a significant *P*-value *(Table S5)*. An overall chi-square test on the NCBI cattle data set was significant ($P < 2e\text{-}16$), indicating that serovars were not randomly distributed among isolation sources. More specifically, *Salmonella* Muenchen was highly overrepresented in ground beef ($r_i = 10.85$) *(Fig. 4A)*. *Salmonella* Cerro was highly overrepresented in cattle feces ($r_i = 13.42$) and hide ($r_i = 9.49$) but underrepresented in cattle organs and tissues ($r_i = -6.12$) and boneless beef ($r_i = -6.76$). *Salmonella* Dublin was overrepresented in cattle organs and tissues ($r_i = 14.72$) and boneless beef ($r_i = 7.01$), while underrepresented in cattle feces ($r_i = -15.07$) and ground beef ($r_i = -7.00$). *Salmonella* Infantis was highly overrepresented in ground beef ($r_i = 9.16$). *Salmonella* Typhimurium was overrepresented among unspecified cattle isolates ($r_i = 6.00$).

The associations between cattle serovars and certain organs and tissue isolates on the basis of data from NCBI are presented in *Fig. S1*. Our analyses found that *Salmonella* Anatum was most highly overrepresented in cattle lymph nodes ($r_i = 11.26$) and most highly underrepresented in cattle lungs ($r_i = -4.66$). *Salmonella* Derby was found to be overrepresented in stomach samples ($r_i = 11.95$). *Salmonella* Dublin was overrepresented among cattle lung ($r_i = 13.20$) and liver samples ($r_i = 6.02$), while it was underrepresented among cattle intestines and lymph nodes ($r_i = -9.33$ and -7.55, respectively). Similarly, *Salmonella* Heidelberg was overrepresented among cattle liver samples ($r_i = 6.27$).

Serovar distribution in chicken

In chicken, the top 10 *Salmonella* serovars were Enteritidis, Hadar, Heidelberg, Infantis, Kentucky, Montevideo, Schwarzengrund, Typhimurium, Johannesburg, and I 4,[5],12:i:– *(Fig. 2A* and *Table S3)*. Over the inclusion period (2015 to 2020), *Salmonella* Kentucky consistently was the most fre-

TABLE 3. Overrepresentation of specific serovars among cattle, chicken, and turkey in NCBI versus FSIS databases[a]

	Cattle		Chicken		Turkey	
Salmonella serovar	OR	*P*-value	OR	*P*-value	OR	*P*-value
Adelaide	0.33	0.207	Inf	0.277	—	—
Agona	0.776	0.192	2.84	8.54E-03	0.503	6.34E-05
Anatum	0.841	0.0921	1.96	0.0381	0.676	0.0622
Cerro	2.43	1.30E-11	3.74	6.84E-03	0	0.327
Derby	0.228	7.81E-04	Inf	5.57E-40	7.84	0.0108
Dublin	3.63	6.40E-31	Inf	0.0765	Inf	0.673
Enteritidis	1.49	0.399	0.605	1.13E-25	0.693	0.306
Hadar	0.495	0.404	3.12	2.33E-06	0.102	0.447
Heidelberg	10.5	1.96E-08	1.71	2.63E-09	2.15	2.93E-04
I 4,[5],12:i:–	1.09	0.410	1.04	0.451	1.18	0.177
Infantis	0.5	2.70E-03	0.852	5.77E-04	0.579	4.49E-04
Javiana	Inf	0.447	Inf	0.0212	0.729	0.527
Johannesburg	0.0752	2.78E-05	0.348	4.75E-05	0	0.327
Kentucky	0.624	0.0253	0.886	1.37E-03	3.21	2.55E-03
Montevideo	0.432	6.12E-17	1.16	0.289	1.24	0.280
Muenchen	0.337	1.46E-11	47.2	6.55E-123	2.30	3.02E-09
Muenster	0.453	8.01E-05	4.80	1.57E-03	0	1.29E-04
Newport	1.07	0.41	11.0	2.45E-19	0.797	0.307
Reading	0.358	0.0226	19.3	4.66E-10	0.795	5.11E-03
Schwarzengrund	0.628	0.171	0.272	1.43E-41	0.833	0.120
Typhimurium	1.27	0.0574	1.67	4.21E-19	1.32	0.0414

[a]On the basis of the results of the Fisher exact test, if $P < 0.05$. An OR > 2 indicated overrepresentation in the NCBI database, while an OR < 0.5 indicated overrepresentation in the FSIS data set for that serovar and animal source. If $P > 0.05$, the result was recorded as "—."

quently identified serovar among both NCBI and FSIS data. In the FSIS data set, numbers of *Salmonella* Enteritidis and *Salmonella* Kentucky isolates followed similar patterns until 2020. Both isolate numbers of serovars increased between 2015 and 2016 (a 1-year increase of 86 and 61%, respectively), increased further in 2017 (a 1-year increase of 11 and 4%, respectively), and decreased in 2018 (a decrease of 8 and 14%, respectively) and 2019 (a decrease of 6 and 5%, respectively). From 2019 to 2020, numbers of *Salmonella* Kentucky isolates in the FSIS data increased by 11%, while numbers of *Salmonella* Enteritidis isolates decreased by 11%. Notably, numbers of *Salmonella* Infantis isolates steadily increased by 826% from 2015 to 2020, and *Salmonella* Infantis became the second most frequently isolated serovar in chicken in 2020.

In the NCBI data set, numbers of isolates of *Salmonella* serovars Infantis, Typhimurium, and Kentucky followed similar patterns in that all three increased between 2016 and 2017 (400, 170, and 76% respectively), increased between 2018 and 2019 (449, 128, and 343%, respectively), and then decreased from 2019 to 2020 (−82, −78, and −85%, respectively) (*Fig. 2B* and *Table S4*). *Salmonella* Infantis was the least frequently identified serovar among the top 10 serovars in 2015 but became the second most frequently identified serovar in 2020, as also observed in the FSIS data discussed in the previous paragraph. Numbers of *Salmonella* Enteritidis isolates remained similar between 2015 and 2018 but increased by 88% between 2018 and 2019. Notably, for nearly all serovars, the total numbers of isolates decreased in 2020 and fell to between 0 and 100 isolates in 2020 (most likely a consequence of the coronavirus disease 2019 pandemic).

A comparison of counts of chicken serovars between NCBI and FSIS data sets found that seven and two serovars were

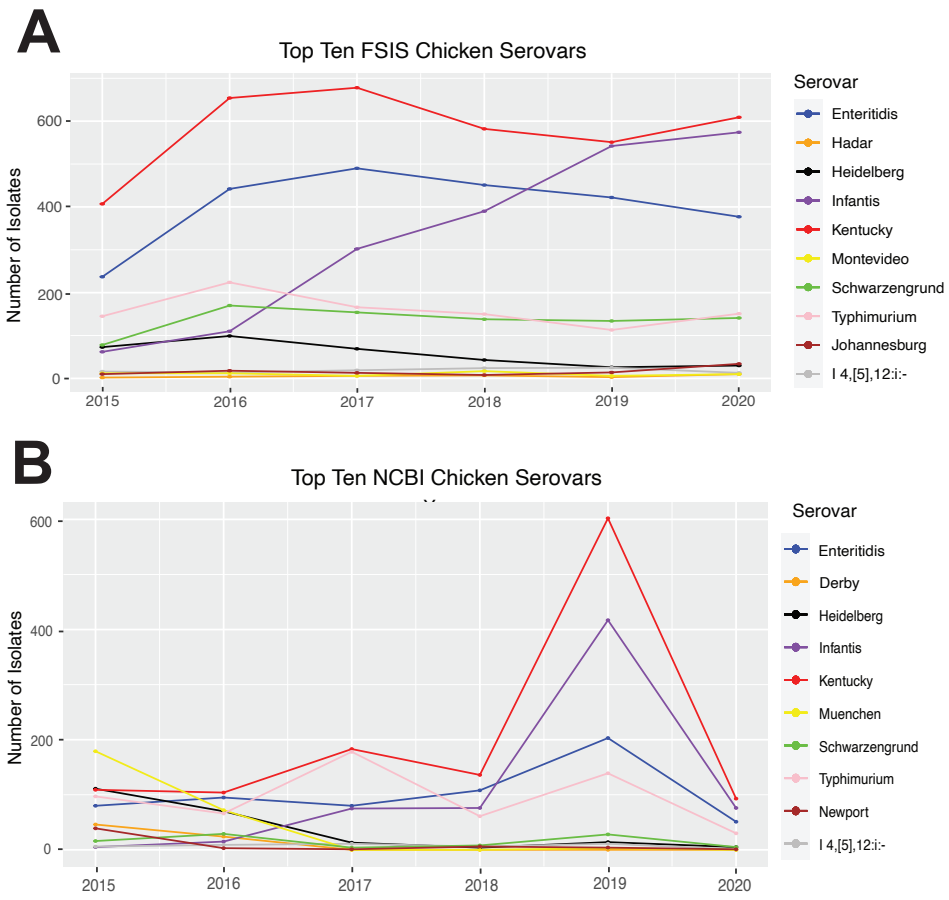


Figure 2. Top 10 chicken *Salmonella* serovars from 2015 to 2020. Line plots of *Salmonella* serovar trends over 6 years in chicken on the basis of isolate counts from A, FSIS data and B, NCBI PD.

overrepresented in the NCBI and FSIS data sets, respectively, with (i) *Salmonella* serovars Agona, Hadar, Newport, Cerro, Muenchen, Muenster, and Reading overrepresented in NCBI chicken data set and (ii) *Salmonella* serovars Johannesburg and Schwarzengrund overrepresented in the FSIS chicken data set compared with the NCBI chicken data. *Salmonella* Muenchen (OR = 47.2) was the serovar most highly overrepresented among the NCBI data, while *Salmonella* Schwarzengrund (OR = 0.272) was the most highly overrepresented among the FSIS data *(Table 3)*.

Chi-square results showed that serovar numbers differed significantly among isolation sources in both the NCBI (P < 2e-16) and FSIS (P < 2e-16) chicken data sets. In NCBI chicken data, *Salmonella* Heidelberg was overrepresented in chicken feces (r_i = 10.19) *(Fig. 4C)*. *Salmonella* Muenchen was highly overrepresented in ground chicken (r_i = 36.43), with *Salmonella* Heidelberg and *Salmonella* Derby also overrepresented in ground chicken (r_i = 16.64 and 12.28, respectively). Conversely, *Salmonella* serovars Enteritidis (r_i = −8.66), Infantis (r_i = −9.78), Kentucky (r_i = −14.60), and Typhimurium (r_i = −8.22) were underrepresented in ground chicken. *Salmonella* Muenchen was underrepresented in chicken breast (r_i = −9.51). On the other hand, in FSIS chicken data, *Salmonella* Enteritidis was overrepresented in raw intact chicken (r_i = 8.0) and underrepresented in young chicken (r_i = −11.91) *(Fig. 5)*. In ground chicken, *Salmonella* Infantis was highly overrepresented (r_i = 19.17), while *Salmonella* Kentucky was highly underrepresented (r_i = −14.77). However, *Salmonella* Kentucky was highly overrepresented in young chicken (ri = 20.99). *Salmonella* Schwarzengrund was overrepresented in raw intact chicken (r_i = 7.86) and underrepresented in ground and young chicken (r_i = −4.72 and −4.41, respectively). *Salmonella* Infantis was underrepresented in raw intact chicken and young chicken (r_i = −10.28 and −8.61, respectively). To a lesser degree, *Salmonella* Johannesburg was also underrepresented in young chicken (r_i = −4.62).

Serovar distribution in turkey

In turkey, the top 10 *Salmonella* serovars were Agona, Hadar, Heidelberg, Infantis, Muenchen, Anatum, Schwarzengrund, Typhimurium, Reading, and I 4,[5],12:i:– (*Fig.*

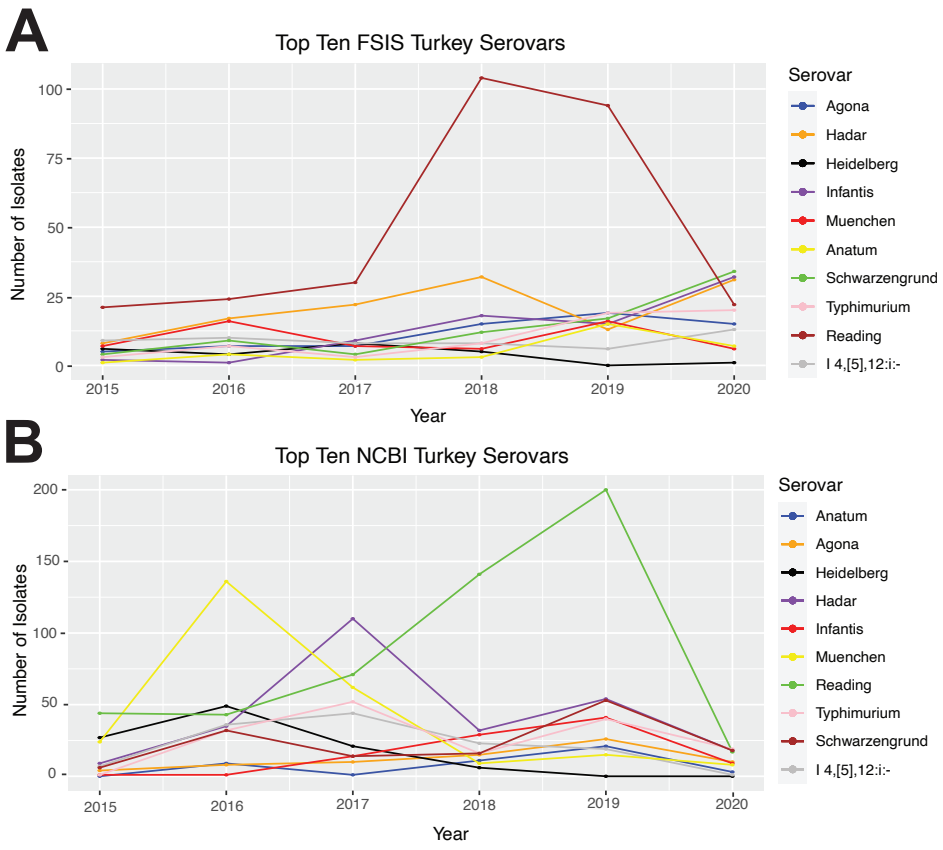


Figure 3. Top 10 turkey *Salmonella* serovars from 2015 to 2020. Line plots of *Salmonella* serovar trends over 6 years in turkey on the basis of isolate counts from A, FSIS data and B, NCBI PD.

3A and *Table S3*). In the FSIS data set, from 2015 to 2019, numbers of *Salmonella* Reading isolates were higher than all other serovars. From 2017 to 2018, there was a 247% increase in numbers of *Salmonella* Reading isolates. However, numbers of *Salmonella* Reading isolates decreased slightly by 10% from 2018 to 2019 and then further decreased by 77% between 2019 and 2020. *Salmonella* Schwarzengrund started off with low numbers of isolates at the beginning of the observation period, but numbers of isolates steadily increased, and Schwarzengrund became the most common serovar in turkey in 2020. In fact, between 2015 and 2020, *Salmonella* Schwarzengrund isolate numbers increased by 750%.

In the NCBI data set, *Salmonella* Reading isolate numbers were consistently high throughout the years, though there was an increase of 99% from 2017 to 2018 and an increase of 42% between 2018 and 2019, followed by a notable decrease of 92% in 2020 (*Fig. 3B* and *Table S4*). This drop in 2020 was observed in all turkey serovars. *Salmonella* Hadar isolate numbers increased by 214% between 2017 and 2018 and by 69% between 2018 and 2019. *Salmonella* Muenchen isolate numbers increased by 467% between 2015 and 2016 but decreased and remained in the low 0 to 25 isolate range in the following years.

A comparison of counts of turkey serovars between NCBI and FSIS data sets found that four and one serovars were overrepresented in the NCBI and FSIS data sets, with (i) *Salmonella* serovars Derby, Kentucky, Muenchen, and Heidelberg overrepresented in the NCBI data set and (ii) *Salmonella* Muenster overrepresented in the FSIS data set. *Salmonella* Derby (OR = 7.8) was the most significantly overrepresented NCBI turkey serovar when compared with FSIS turkey data set *(Table 3)*.

Chi-square results were only significant ($P < 0.05$) for the NCBI data set ($P < 2.2e\text{-}16$). In NCBI turkey data, *Salmonella* Kentucky was highly overrepresented in the turkey farm environment ($r_i = 17.73$) and underrepresented in ground turkey ($r_i = -8.51$) *(Fig. 4B)*. In addition, (i) *Salmonella* Heidelberg was highly overrepresented in turkey feces ($r_i = 12.68$); (ii) *Salmonella* Agona was overrepresented in turkey carcasses ($r_i = 10.37$); (iii) *Salmonella* Montevideo was overrepresented in young turkey isolates ($r_i = 9.32$); (iv) *Salmonella* Anatum was overrepresented in the turkey farm environment ($r_i = 9.02$); and (v) *Salmonella* Javiana was overrepresented in turkey carcasses ($r_i = 7.50$).

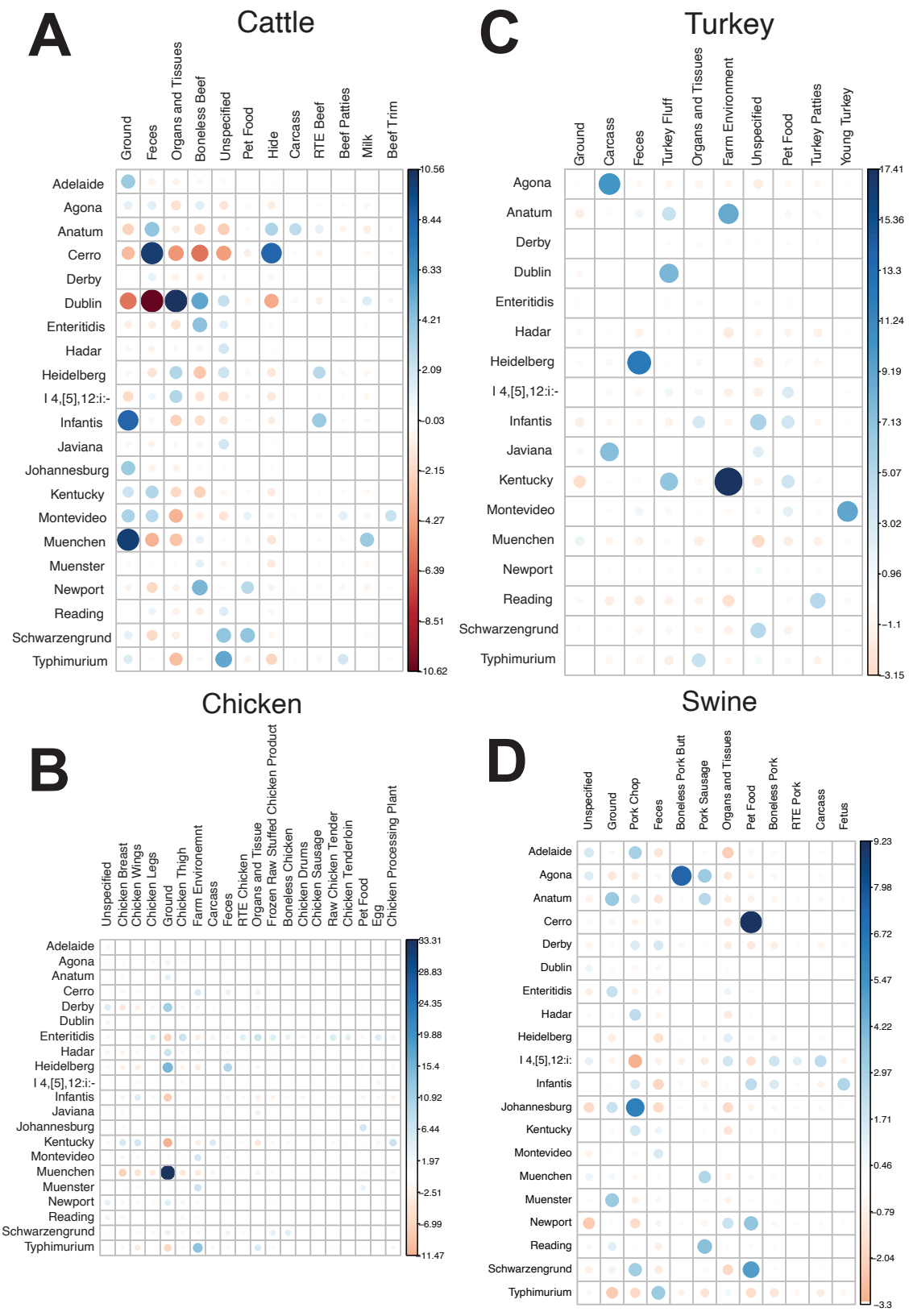


Figure 4. NCBI isolation source post hoc hypothesis testing with adjusted Pearson residual plot for cattle, chicken, turkey, and swine. Bubble plots displaying varying degrees of overrepresentation (blue) and underrepresentation (red) in the NCBI data of various serovars in certain isolation sources in cattle A, chicken B, turkey C, and swine D.

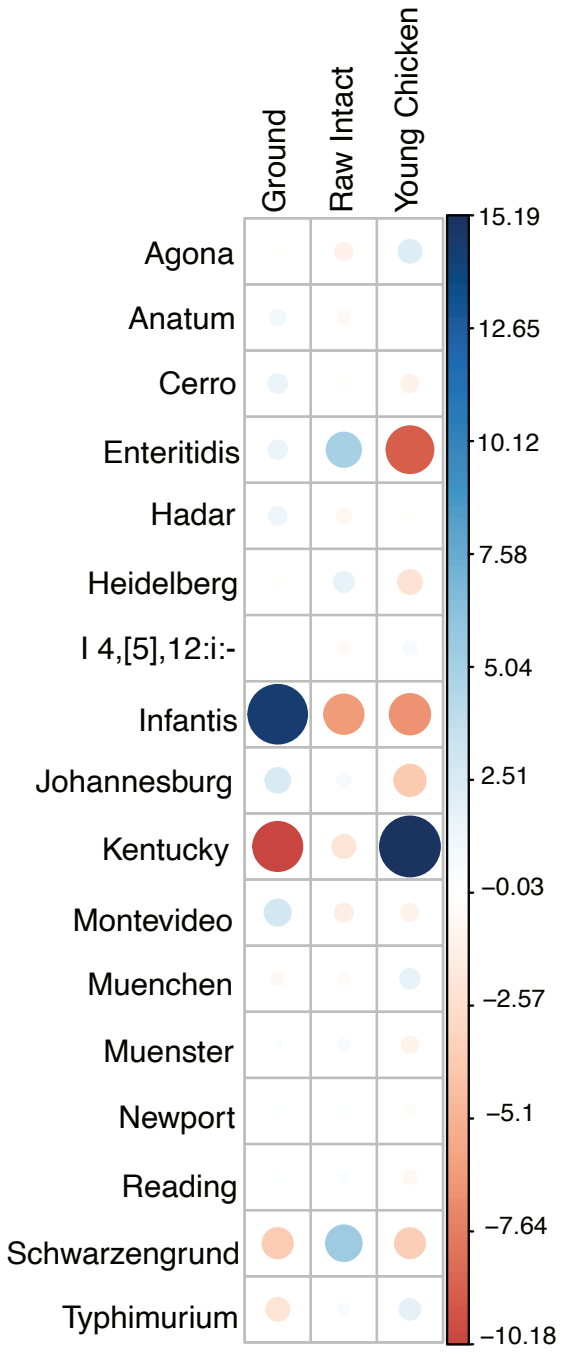


Figure 5. FSIS isolation source post hoc hypothesis testing with adjusted Pearson residual plot for chicken. Bubble plot displaying varying degrees of overrepresentation (blue) and underrepresentation (red) in the FSIS data of various serovars in certain chicken isolation sources.

Serovar distribution in swine

Due to recently implemented FSIS performance standards in swine, the collection dates for the FSIS swine data differed from the collection dates for FSIS cattle, chicken, and turkey data. We thus only analyzed the NCBI swine data set for temporal trends and overrepresentation of isolation sources. The top 10 swine-associated *Salmonella* serovars analyzed here included Adelaide, Agona, Anatum, Derby, Heidelberg, I 4,[5],12:i:–, Infantis, Johannesburg, Newport, and Typhimurium (*Fig. S2* and *Table S4*). During the inclusion period, *Salmonella* I 4,[5],12:i:– was the most frequently isolated serovar from swine, except in 2016, when numbers of *Salmonella* Typhimurium isolates were higher. There was a 62 and 72% decrease in numbers of *Salmonella* Typhimurium isolates and *Salmonella* I 4,[5],12:i:– isolates, respectively, between 2015 and 2016. Also, the numbers of *Salmonella* Infantis isolates increased by 680% between 2018 and 2019 and decreased by 457% between 2019 and 2020.

The overall chi-square test for isolation sources was significant for the NCBI swine ($P < 2e\text{-}16$) data set (*Table S5*), indicating differences in serotype distribution among sources. Our analysis specifically found that *Salmonella* Cerro was highly overrepresented in pet food containing pork ($r_i = 9.31$) (*Fig. 4D*). To a lesser extent, *Salmonella* Schwarzengrund was also overrepresented in pet food containing pork ($r_i = 5.15$). *Salmonella* Agona was highly overrepresented in boneless pork butt ($r_i = 7.59$). In pork chops, *Salmonella* Johannesburg was overrepresented ($r_i = 6.99$), while *Salmonella* I 4,[5],12:i:– was underrepresented ($r_i = -4.31$).

The associations between swine serovars and specific organs and tissue isolates on the basis of the NCBI swine data set are presented in *Fig. S3*. We found that *Salmonella* Enteritidis was highly overrepresented in swine spleens ($r_i = 8.85$) and peritoneum ($r_i = 8.85$). *Salmonella* Infantis was slightly underrepresented in swine intestines ($r_i = -4.13$). *Salmonella* Newport was also overrepresented in swine nasal samples ($r_i = 5.47$) and swine skin ($r_i = 5.47$). In addition, *Salmonella* Anatum was overrepresented in swine lymph nodes ($r_i = 5.06$).

DISCUSSION

Current regulations on *Salmonella* in raw chicken, beef, turkey, and pork products in the United States use performance standards to categorize establishments on the basis of the percentage of raw products positive for *Salmonella* over a rolling window (*42*). This current approach, de facto, implies that all nontyphoidal *Salmonella* serovars are equally likely to cause human disease, even though there are known differences among serovars in the likelihoods to cause invasive disease (*23*) and the ability to cause disease in various hosts (*21, 26, 32*). Although there is increasing consensus that risk-based approaches that target serovars of concern are more likely to have a positive impact on public health than the current approaches used in the United States, ready access to data on *Salmonella* serovar distributions and temporal trends in food animals (e.g., chicken, cattle) and raw meat and poultry is important to facilitate control strategies that target specific serovars. To

better understand available data sources, including potential biases associated with them, we analyzed serovar distribution among data available through FSIS and NCBI. Although our analyses identified differences in the serovar representation among NCBI and FSIS data, suggesting some potential biases in these databases, overall trends are consistent with previous data on *Salmonella* epidemiology, biology, and pathogenesis. It thus appears that despite possible limitations of both FSIS and NCBI databases, analysis of these data sources can help inform risk-based policy options and intervention strategies.

Temporal serovar distribution patterns and overrepresentation data show differences in serovars representation between NCBI and FSIS databases with differences potentially driven by outbreaks, providing initial evidence for possible biases in these databases

Our analyses of temporal patterns of serovars found among different food animals, often were consistent with well-described patterns. For example, among chicken isolates in NCBI and FSIS, *Salmonella* Kentucky was the most frequently identified serovar, consistent with the well-documented role of Kentucky as predominant serovar in chicken in the United States *(8)*. Observed increases in the prevalence for some serovars were likely due to well-documented *Salmonella* outbreaks. For instance, the spike of *Salmonella* Reading turkey isolates in 2018 in the FSIS data set and in 2019 in the NCBI data set likely can be attributed to a large outbreak of multidrug-resistant *Salmonella* Reading infections that were linked to various raw turkey products *(9)*. Similarly, the spike of *Salmonella* Infantis chicken isolates in the NCBI data set in 2019 is likely due to an outbreak of multidrug-resistant *Salmonella* Infantis infections, which was linked to raw chicken products from numerous sources and lasted from 2018 to 2019 *(10)*. In addition, the high numbers of *Salmonella* I 4,[5],12:i:– observed in 2015, followed by a sharp decrease in 2016, may be partially explained by a multistate outbreak of *Salmonella* I 4,[5],12:i:– linked to pork products *(6)*. However, the increased numbers of *Salmonella* Infantis isolates in chickens and *Salmonella* Reading numbers in turkeys may also, more broadly, reflect an emergence of these serovars among chicken and turkey populations *(28, 29)*. The steadily increasing levels of *Salmonella* Infantis in the FSIS chicken data set from 2016 to 2020 also are consistent with emergence of a specific strain representing this serovar in chicken. Although the numbers of all other serovars in the same data set remained relatively constant and showed little change on an annual basis over the 6-year period between 2015 and 2020, *Salmonella* Infantis was the only serovar that steadily increased in frequency in the FSIS data set, as well as the NCBI data set. This illustrates how these databases can help identify serovar prevalence trends that may indicate potential public health concern, particularly if the serovars or subtypes that show an increase have already been characterized as highly virulent and multidrug resistant, as is the case for *Salmonella* Infantis *(24)*. In addition, an increase in prevalence of a given serovar or subtype may also indicate emergence or introduction of a new more virulent or more transmissible subtype, even if no prior data are available for this subtype.

Although our analysis of temporal trends identified instances in which NCBI and FSIS data showed similar trends (e.g., an increase in *Salmonella* Infantis), we also identified instances in which serovar prevalence trends were not consistent between these two databases. For example, while *Salmonella* Montevideo was consistently the most frequently identified serovar among cattle isolates in the FSIS database, serovar *Salmonella* Dublin was typically the most frequently identified serovar among cattle in NCBI (and was always more frequently identified than *Salmonella* Montevideo). As these observations are consistent with the fact that the data sources and acquisition approaches for the FSIS and NCBI databases differ, we performed statistical analyses comparing the serovar distribution, within a given category (e.g., chicken). These analyses showed that serovars were more frequently overrepresented in the NCBI data than the FSIS data. This could be explained by the NCBI data depositions being affected by specific scientific studies or general research interests (e.g., in host-specific serovars, such as *Salmonella* Dublin or Cerro). Overall, the results reveal overrepresentation of different serovars in both the NCBI and FSIS data sets, possibly caused by different biases associated with data collection. Being aware of these caveats will be important, as these data sources are being used in research as well as development of risk assessments and possibly regulatory policies.

Overrepresentation data among different host isolation sources are consistent with previous data on *Salmonella* serovar biology and pathogenesis

Interestingly, our analyses found that certain serovars were overrepresented in specific isolation sources, with a number of instances where findings of overrepresentation were consistent with previous data and studies. In cattle, the overrepresentation of specific serovars in NCBI was consistent with established findings on the biology and pathogenesis of different serovars. For instance, *Salmonella* Dublin was overrepresented among cattle lung isolates, consistent with reports that this serovar frequently tends to cause pneumonia in calves *(35)*. Similarly, *Salmonella* Cerro was found to be overrepresented among cattle feces and hides but underrepresented in organs and beef, consistent with previous research, where WGS was not used for serotyping, which also found that *Salmonella* Cerro was more frequently isolated from cattle feces than other cattle parts *(27)*. These findings are also consistent with numerous studies that found genetic and phenotypic evidence that suggests that *Salmonella* Cerro shows reduced virulence and invasiveness *(13, 25, 37)*. Similarly, the observed overrepresentation

of *Salmonella* Anatum in cattle lymph nodes is consistent with other studies that have reported common isolation of *Salmonella* Anatum from the cattle lymphatic system *(2, 4, 15, 18)*.

Interestingly, the overrepresentation data among the FSIS chicken isolates by source (i.e., ground, raw intact, and young chicken) also provided initial insights that (i) are consistent with previous data and (ii) may provide important insight for future efforts to control targeted serovars in chickens *(17, 19)*. Most strikingly, *Salmonella* Kentucky was overrepresented among young chicken and underrepresented among ground chicken, which not only could be seen as consistent with the reported virulence attenuation of the predominant *Salmonella* Kentucky ST in the United States (i.e., ST152) *(20)* but also could suggest that existing interventions or practices may effectively control transmission of this serovar from live birds to ground chicken. On the other hand, *Salmonella* Infantis was found to be highly overrepresented among ground chicken and turkey and underrepresented in raw intact chicken, which could indicate efficient transmission, with limited reduction, from young chicken to ground chicken, which could indicate *Salmonella* Infantis presence in deeper tissue.

We also find certain overrepresentation patterns that are possibly related to biases in data collection. For example, most serovars that showed evidence for overrepresentation in a given turkey isolation source in NCBI represented serovars that are historically less frequently isolated from turkey *(8)*. For example, *Salmonella* Montevideo was significantly overrepresented in young turkey compared with other serovars, despite its low isolation from turkey sources overall. The overrepresentation of these less common turkey serovars from these isolation sources in the NCBI data set may be an artifact associated with many academic groups submitting data to NCBI; these data may be a result of a specific study or studies focused on one or more selected serovars.

Despite limitations of FSIS and NCBI databases, analysis of them informs risk-based policy options and intervention strategies

As detailed previously, possible limitations of the NCBI and FSIS data sets include potential biases with biases likely more frequent in the NCBI data set. For example, the fact that many data in our NCBI data set analyzed (which excluded FSIS submissions) are from academic sources may lead to bias and associated over- or underrepresentation of specific serovars, as certain serovars may be of higher or lower interest to academia. For instance, *Salmonella* Infantis has recently become of high interest among researchers, which has resulted in numerous studies analyzing the serovar. More broadly, academics might have had an incentive to isolate certain serovars over others, depending on research funding and grants, which may, for example, lead to overrepresentation of antibiotic-resistant *Salmonella*, representing a specific research priority. Second, the NCBI data set may be biased and contain a moderate degree of variability compared with the FSIS data set because samples are collected by various researchers who use different isolation, enrichment, and testing methodologies.

Although the FSIS data set indicates serovars identified from positive samples during routine sampling and isolated through standardized methodologies, there still exists a potential for differences among samples because multiple people are both sampling and handling products. Also, these sampling and isolation techniques may select for certain serovars and bias identification toward those. Finally, the *Salmonella* serotyping protocol used by FSIS does not identify the presence of more than one serovar per sample. This means that the true *Salmonella* serovar diversity in collected samples is not captured and that only the most prevalent serovar (or the serovar that grows best in enrichment media) is identified. It is, however, also likely that in many academic studies only a single isolate per sample is advanced to WGS. A specific limitation of our study relates to the time frame covered by our data sets, which included part of the severe acute respiratory syndrome coronavirus 2 (coronavirus disease 2019) pandemic. This may be why we observed notable decreases in *Salmonella*-positive isolates from the NCBI data set in 2020. It is likely that many academic researchers were unable to collect samples this year. The FSIS data set was likely not affected or at less affected because FSIS continued to collect samples in 2020. Finally, we found that swine-associated *Salmonella* data available from FSIS are more limited than data available for beef, chicken, and turkey.

Despite the limitations detailed in the prior paragraph, our data support that the FSIS and NCBI databases, particularly if used in combination, can be useful to support the implementation of more risk-based control strategies for *Salmonella* in raw meat and poultry. These databases are particularly important as the U.S. Department of Agriculture FSIS may transition into focusing on control of *Salmonella* with the largest public health significance, with a reduced focus on serovars with limited public health impact, even if they are found in high frequency in certain animal sources, such as *Salmonella* Kentucky in chicken. Although NCBI and FSIS databases can allow for timely tracking of changes in serovar prevalences and changes in associations of serovars with different sources, it is important to consider the various biases of FSIS and NCBI data and to also use other foodborne pathogen data sets (e.g., poultry or meat industry data), where possible, to minimize confounding variables and provide a more comprehensive understanding. Future regulatory approaches that will focus on *Salmonella* serovars of public health relevance, as well as risk assessments supporting these approaches, however, will also need to consider other factors that affect risk and that cannot (yet) be easily accessed through these databases, such as

differences in dose-response relationships between subtypes and differences in human illness severity associated with different subtypes. In addition, note that many *Salmonella* serovars are polyphyletic, meaning they can represent two or more genetically distinct groups that can differ in ability to cause human diseases. This has been well documented for *Salmonella* Kentucky, which is polyphyletic with ST152 that is predominant in the United States and linked to a reduced ability to cause human disease, while STs found in other regions (e.g., Europe, Africa) mostly represent a distinct phylogenetic lineage and appear to be fully virulent *(20)*. Hence, future efforts should include analysis of NCBI and FSIS *Salmonella* WGS data, with resolution beyond the serotype levels (e.g., to ST).

REFERENCES

1. Acevedo-Villanueva, K. Y., G. O. Akerele, W. G. Al Hakeem, S. Renu, R. Shanmugasundaram, and R. K. Selvaraj. 2021. A novel approach against *Salmonella*: a review of polymeric nanoparticle vaccines for broilers and layers. *Vaccines* 9:1041.
2. Arthur, T. M., D. M. Brichta-Harhay, J. M. Bosilevac, M. N. Guerini, N. Kalchayanand, J. E. Wells, S. D. Shackelford, T. L. Wheeler, and M. Koohmaraie. 2008. Prevalence and characterization of *Salmonella* in bovine lymph nodes potentially destined for use in ground beef. *J. Food Prot.* 71:1685–1688.
3. Brenner, F. W., R. G. Villar, F. J. Angulo, R. Tauxe, and B. Swaminathan. 2000. *Salmonella* nomenclature. *J. Clin. Microbiol.* 38:2465–2467.
4. Brichta-Harhay, D. M., T. M. Arthur, J. M. Bosilevac, N. Kalchayanand, J. W. Schmidt, R. Wang, S. D. Shackelford, G. H. Loneragan, and T. L. Wheeler. 2012. Microbiological analysis of bovine lymph nodes for the detection of *Salmonella enterica*. *J. Food Prot.* 75:854–858.
5. Brown, E., U. Dessai, S. McGarry, and P. Gerner-Smidt. 2019. Use of whole-genome sequencing for food safety and public health in the United States. *Foodborne Pathog. Dis.* 16:441–450.
6. Centers for Disease Control and Prevention. 2015. Multistate outbreak of multidrug-resistant *Salmonella* I 4,[5],12:i:– and *Salmonella* Infantis infections linked to pork (final update). Available at: https://archive.cdc.gov/#/details?url=https://www.cdc.gov/salmonella/pork-08-15/index.html. Accessed 1 August 2024.
7. Centers for Disease Control and Prevention. 2018. National enteric disease surveillance: Salmonella annual report, 2016. Available at: https://www.cdc.gov/nationalsurveillance/pdfs/2016-Salmonella-report-508.pdf. Accessed 1 August 2024.
8. Centers for Disease Control and Prevention. 2018. National *Salmonella* surveillance 2016 supplement: Animals isolate report. Available at: https://www.cdc.gov/nationalsurveillance/pdfs/salmonella-serotypes-isolated-animals-and-related-sources-508.pdf. Accessed 1 August 2024.
9. Centers for Disease Control and Prevention. 2019. Outbreak of multidrug-resistant *Salmonella* infections linked to raw chicken products. Available at: https://archive.cdc.gov/#/details?url=https://www.cdc.gov/salmonella/infantis-10-18/index.html. Accessed 1 August 2024.
10. Centers for Disease Control and Prevention. 2019. Outbreak of multidrug-resistant *Salmonella* infections linked to raw turkey products. Available at: https://archive.cdc.gov/#/details?url=https://www.cdc.gov/salmonella/reading-07-18/index.html. Accessed 1 August 2024.
11. Centers for Disease Control and Prevention. 2023. FoodNet Fast: Pathogen Surveillance Tool. Available at: https://www.cdc.gov/foodnetfast/. Accessed 1 August 2024.
12. Cohn, A. R., R. A. Cheng, R. H. Orsi, and M. Wiedmann. 2021. Moving past species classifications for risk-based approaches to food safety: *Salmonella* as a case study. *Front. Sustain. Food Syst.* 5:652132.
13. Cohn, A. R., R. H. Orsi, L. M. Carroll, J. Liao, M. Wiedmann, and R. A. Cheng. 2022. *Salmonella* Cerro displays a phylogenetic structure and genomic features consistent with virulence attenuation and adaptation to cattle. *Front. Microbiol.* 13:4725.
14. Dórea, F. C., D. J. Cole, C. Hofacre, K. Zamperini, D. Mathis, M. P. Doyle, M. D. Lee, and J. J. Maurer. 2010. Effect of *Salmonella* vaccination of breeder chickens on contamination of broiler chicken carcasses in integrated poultry operations. *Appl. Environ. Microbiol.* 76:7820–7825.
15. Edrington, T. S., T. M. Arthur, G. H. Loneragan, K. J. Genovese, D. L. Hanson, R. C. Anderson, and D. J. Nisbet. 2020. Evaluation of two commercially-available *Salmonella* vaccines on *Salmonella* in the peripheral lymph nodes of experimentally-infected cattle. *Ther. Adv. Vaccines Immunother.* 8:2515135520957760.
16. Eskin, S. 2020. Vaccines for poultry are crucial for preventing *Salmonella* contamination. Available at: https://www.pewtrusts.org/en/research-and-analysis/articles/2020/09/24/vaccines-for-poultry-are-crucial-for-preventing-salmonella-contamination. Accessed 1 August 2024.
17. Foley, S. L., R. Nayak, I. B. Hanning, T. J. Johnson, J. Han, and S. C. Ricke. 2011. Population dynamics of *Salmonella enterica* serotypes in commercial egg and poultry production. *Appl. Environ. Microbiol.* 77:4273–4279.
18. Gragg, S. E., G. H. Loneragan, K. K. Nightingale, D. M. Brichta-Harhay, H. Ruiz, J. R. Elder, L. G. Garcia, M. F. Miller, A. Echeverry, R. G. Ramirez Porras, and M. M. Brashears. 2013. Substantial within-animal diversity of *Salmonella* isolates from lymph nodes, feces, and hides of cattle at slaughter. *Appl. Environ. Microbiol.* 79:4744–4750.
19. Gutierrez, A., J. De, and K. R. Schneider. 2020. Prevalence, concentration, and antimicrobial resistance profiles of *Salmonella* isolated from Florida poultry litter. *J. Food Prot.* 83:2179–2186.
20. Haley, B. J., S. W. Kim, J. Pettengill, Y. Luo, J. S. Karns, and J. A. S. Van Kessel. 2016. Genomic and evolutionary analysis of two *Salmonella enterica* serovar Kentucky sequence types isolated from bovine and poultry sources in North America. *PLoS One* 11:0161225.
21. Hara-Kudo, Y., and K. Takatori. 2011. Contamination level and ingestion dose of foodborne pathogens associated with infections. *Epidemiol. Infect.* 139:1501–1510.
22. Issenhuth-Jeanjean, S., P. Roggentin, M. Mikoleit, M. Guibourdenche, E. De Pinna, S. Nair, P.I. Fields, and F.-X. Weill. 2014. Supplement 2008–2010 (no. 48) to the White–Kauffmann–Le Minor scheme. *Res. Microbiol.* 165:526–530. https://doi.org/10.1016/j.resmic.2014.07.004.
23. Jones, T. F., L. A. Ingram, P. R. Cieslak, D. J. Vugia, M. Tobin-D'Angelo, S. Hurd, C. Medus, A. Cronquist, and F. J. Angulo. 2008. Salmonellosis outcomes differ substantially by serotype. *J. Infect. Dis.* 198:109–114.
24. Karacan Sever, N., and M. Akan. 2019. Molecular analysis of virulence genes of *Salmonella* Infantis isolated from chickens and turkeys. *Microb. Pathog.* 126:199–204.

25. Kovac, J., K. J. Cummings, L. D. Rodriguez-Rivera, L. M. Carroll, A. Thachil, and M. Wiedmann. 2017. Temporal genomic phylogeny reconstruction indicates a geospatial transmission path of *Salmonella* Cerro in the United States and a clade-specific loss of hydrogen sulfide production. *Front. Microbiol.* 8:737.

26. Langridge, G. C., M. Fookes, T. R. Connor, T. Feltwell, N. Feasey, B. N. Parsons, H. M. B. Seth-Smith, L. Barquist, A. Stedman, T. Humphrey, P. Wigley, S. E Peters, D. J. Maskell, J. Corander, J. A. Chabalgoity, P. Barrow, J. Parkhill, G. Dougan, and N. R. Thomson. 2015. Patterns of genome evolution that have accompanied host adaptation in *Salmonella*. *Proc. Natl. Acad. Sci. USA* 112:863–868.

27. Loneragan, G. H., D. U. Thomson, R. M. McCarthy, H. E. Webb, A. E. Daniels, T. S. Edrington, D. J. Nisbet, S. J. Trojan, S. C. Rankin, and M. M. Brashears. 2012. *Salmonella* diversity and burden in cows on and culled from dairy farms in the Texas High Plains. *Foodborne Pathog. Dis.* 9:549–555.

28. McMillan, E. A., M. D. Weinroth, and J. G. Frye. 2022. Increased prevalence of *Salmonella* Infantis isolated from raw chicken and turkey products in the United States is due to a single clonal lineage carrying the pESI plasmid. *Microorganisms* 10:1748.

29. Miller, E. A., E. Elnekave, C. Flores-Figueroa, A. Johnson, A. Kearney, J. Munoz-Aguayo, K. A. Tagg, L. Tschetter, B. P. Weber, C. A. Nadon, D. Boxrud, R. S. Singer, J. P. Folster, and T. J. Johnson. 2020. Emergence of a novel *Salmonella enterica* serotype Reading clonal group is linked to its expansion in commercial turkey production, resulting in unanticipated human illness in North America. *mSphere* 5:e00056-20.

30. Mukherjee, N., V. G. Nolan, J. R. Dunn, and P. Banerjee. 2019. Sources of human infection by *Salmonella enterica* serotype Javiana: a systematic review. *PLoS One* 14:e0222108.

31. National Library of Medicine. 2024. Pathogen Detection. Available at: http://ncbi.nlm.nih.gov/pathogens. Accessed 1 August 2024.

32. Nuccio, S. P., and A. J. Bäumler. 2014. Comparative analysis of *Salmonella* genomes identifies a metabolic network for escalating growth in the inflamed gut. *mBio* 5:e00929-00914.

33. Office of Disease Prevention and Health Promotion. 2010. 2020 Objectives and topics: Food safety. Available at: https://wayback.archive-it.org/5774/20220414163116/https://www.healthypeople.gov/2020/topics-objectives/topic/food-safety/objectives. Accessed 1 August 2024.

34. Office of Disease Prevention and Health Promotion. 2020. 2030 objectives and topics: Food safety. Available at: https://health.gov/healthypeople/objectives-and-data/browse-objectives/foodborne-illness. Accessed 1 August 2024.

35. Pecoraro, H. L., B. Thompson, and G. E. Duhamel. 2017. Histopathology case definition of naturally acquired *Salmonella enterica* serovar Dublin infection in young Holstein cattle in the northeastern United States. *J. Vet. Diagn. Invest.* 29:860–864.

36. R Core Team. 2020. R: A language and environment for statistical computing. Available at: https://www.R-project.org/. Accessed 1 August 2024.

37. Rodriguez-Rivera, L. D., A. I. Moreno Switt, L. Degoricija, R. Fang, C. A. Cummings, M. R. Furtado, M. Wiedmann, and H. C. Den Bakker. 2014. Genomic characterization of *Salmonella* Cerro ST367, an emerging *Salmonella* subtype in cattle in the United States. *BMC Genomics* 15:427.

38. Scallan, E., R. M. Hoekstra, F. J. Angulo, R. V. Tauxe, M. A. Widdowson, S. L. Roy, J. L. Jones, and P. M. Griffin. 2011. Foodborne illness acquired in the United States—major pathogens. *Emerg. Infect. Dis.* 17:7–15.

39. Siceloff, A. T., D. Waltman, and N. W. Shariat. 2022. Regional *Salmonella* differences in United States broiler production from 2016 to 2020 and the contribution of multiserovar populations to *Salmonella* surveillance. *Appl. Environ. Microbiol.* 88:e0020422.

40. Tack, D. M., L. Ray, P. M. Griffin, P. R. Cieslak, J. Dunn, T. Rissman, R. Jervis, S. Lathrop, A. Muse, M. Duwell, K. Smith, M. Tobin-D'Angelo, D. J. Vugia, J. Zablotsky Kufel, B. J. Wolpert, R. Tauxe, and D. C. Payne. 2020. Preliminary incidence and trends of infections with pathogens transmitted commonly through food—Foodborne Diseases Active Surveillance Network, 10 U.S. sites, 2016–2019. *Morb. Mortal Wkly. Rep.* 69:509–514.

41. U.S. Department of Agriculture, Food Safety and Inspection Service. 2019. Quarterly sampling reports on *Salmonella* fiscal year 2019. Available at: https://www.fsis.usda.gov/sites/default/files/media_file/2021-03/FY2019-Sampling-Summary-Report.pdf. Accessed 1 August 2024.

42. U.S. Department of Agriculture, Food Safety and Inspection Service. 2019. Pathogen reduction—*Salmonella* and *Campylobacter* performance standards verification testing. Available at: https://www.fsis.usda.gov/sites/default/files/media_file/2021-02/24_IM_PR_Sal_Campy.pdf. Accessed 1 August 2024.

43. U.S. Department of Agriculture, Food Safety and Inspection Service. 2020. Quarterly sampling reports on *Salmonella* fiscal year 2020. Available at: https://www.fsis.usda.gov/news-events/publications/test-fy-2020-quarterly-sampling-reports-samonella. Accessed 1 August 2024.

44. U.S. Department of Agriculture, Food Safety and Inspection Service. 2022. Laboratory sampling data. Available at: https://www.fsis.usda.gov/science-data/data-sets-visualizations/laboratory-sampling-data. Accessed 1 August 2024.

45. Zhang, S., H. C. den Bakker, S. Li, J. Chen, B. A. Dinsmore, C. Lane, A. C. Lauer, P I. Fields, and X. Deng. 2019. SeqSero2: Rapid and improved *Salmonella* serotype determination using whole-genome sequencing data. *Appl. Environ. Microbiol.* 85:e01746-19.

SUPPLEMENTAL TABLES AND FIGURES

Supplemental TABLE 1. NCBI search terms for each animal source

Animal	Search terms
Chicken	*Gallus gallus*, chicken, egg, hen
Cattle	*Bos taurus*, cattle, beef, bovine, cow
Turkey	*Meleagris gallopavo*, turkey
Swine	*Sus scrofa*, pig, swine, porcine, pork

Supplemental TABLE 2. Isolation source standardization methodology

Animal source	Term	Action
Chicken	Ground	Combined: finished ground chicken, ground component chicken
Chicken	Unspecified	Combined: finished product chicken, finished chicken, retail chicken, chicken meat, chicken rinse
Chicken	Manure	Deleted
Chicken	Organs and tissues	Combined: liver, tissue pool, organ pool, chicken cecal, chicken cecal tonsils, chicken dirty pool tissue, chicken femur, chicken heart, chicken hock, chicken intestine, chicken joint, chicken spine, chicken spleen, chicken stifle, chicken pericardium, lung, chicken cecum, mixed parts, chicken organs, chicken trachea, chicken cloaca, chicken vertebrae, chicken abdomen, chicken bone
Chicken	Low ash chicken meal	Deleted
Chicken	Egg	Combined: chicken yolk, chicken yolk sac
Chicken	Chicken pet food	Combined: raw chicken dog food
Chicken	Ready-to-eat chicken	Combined: cooked chicken, chicken salad, chicken dish, chicken product breaded nugget, chicken product broccoli and cheese, chicken product broccoli and cheese stuffed
Chicken	Farm environment	Combined: egg environment, chicken litter, chicken paper, chicken house environment
Chicken	Chicken meal	Deleted
Chicken	Chicken feed	Deleted
Chicken	Boneless chicken	Combined: boneless chicken, chicken product boneless meat
Chicken	Carcass	Combined: carcass, raw intact chicken, intact chicken, whole chicken
Turkey	Ground	Combined: ground, ground component, finished ground turkey
Turkey	Unspecified	Combined: unspecified, finished turkey
Turkey	Pet food	Combined: pet food, turkey raw dog food
Turkey	Manure	Deleted
Turkey	Farm environment	Combined: environment, hatchery debris
Turkey	Organs and tissues	Combined: intestine, lung, turkey liver and yolk sac, trachea, nasal swab, liver, turkey bone, caecum, turkey air sac, turkey necks, turkey yolk sac, turkey cloacal, turkey abdomen
Turkey	Turkey patties	Combined: ground turkey patties, turkey patties
Cattle	Ground	Combined: ground, raw ground beef, finished ground beef, ground beef meat
Cattle	Unspecified	Combined: unspecified, finished beef, raw processed beef, raw beef, beef meat, beef enrichment, seasoned beef, calf, raw or partially cooked beef with bone, beef product smoked

Continued on the next page

Supplemental TABLE 2. Isolation source standardization methodology (cont.)

Animal source	Term	Action
Cattle	Bovine booties	Deleted
Cattle	Lymph	Deleted
Cattle	Urine	Deleted
Cattle	Dairy cow manure	Deleted
Cattle	Organs and tissues	Combine: lung, pooled tissue, tissue pool, intestine, lymph node, liver, uterus, placenta, brain, raw beef stomach, ruminate stomachs, spleen, blood, kidney, bone marrow, umbilicus, heart valve, tissue, nasal
Cattle	Fluid	Deleted
Cattle	Bile fluid	Deleted
Cattle	Beef suet	Deleted
Cattle	Intestinal fluid	Deleted
Cattle	Ready-to-eat beef	Combined: roast beef, beef cutlet
Cattle	Feces	Combined: feces, dairy cow fecal
Cattle	Beef patties	Combined: ground beef patty, beef patties, beef hamburger, beef product patties

Supplemental TABLE 3. Top 21 *Salmonella* Serovars from FSIS 2015 to 2020

Serovar	Chicken	Cattle	Turkey	Swine	Total
Adelaide	0	0	0	203	203
Agona	0	24	87	46	157
Anatum	0	137	26	506	669
Cerro	0	92	0	6	98
Derby	0	10	0	295	305
Dublin	0	105	0	0	105
Enteritidis	3,082	5	0	0	3,085
Hadar	30	0	126	0	156
Heidelberg	360	97	19	0	476
I 4,[5],12:i:–	112	12	38	364	526
Infantis	2,294	57	88	417	2,856
Javiana (control)	0	0	0	0	0
Johannesburg	73	4	0	250	327
Kentucky	4,048	38	2	0	4,088
Montevideo	28	344	16	3	391
Muenchen	5	100	69	32	206
Muenster	0	69	29	0	98
Newport	0	59	8	0	67
Reading	0	9	321	7	337
Schwarzengrund	1,016	5	92	7	1,120
Typhimurium	952	67	54	50	1,123

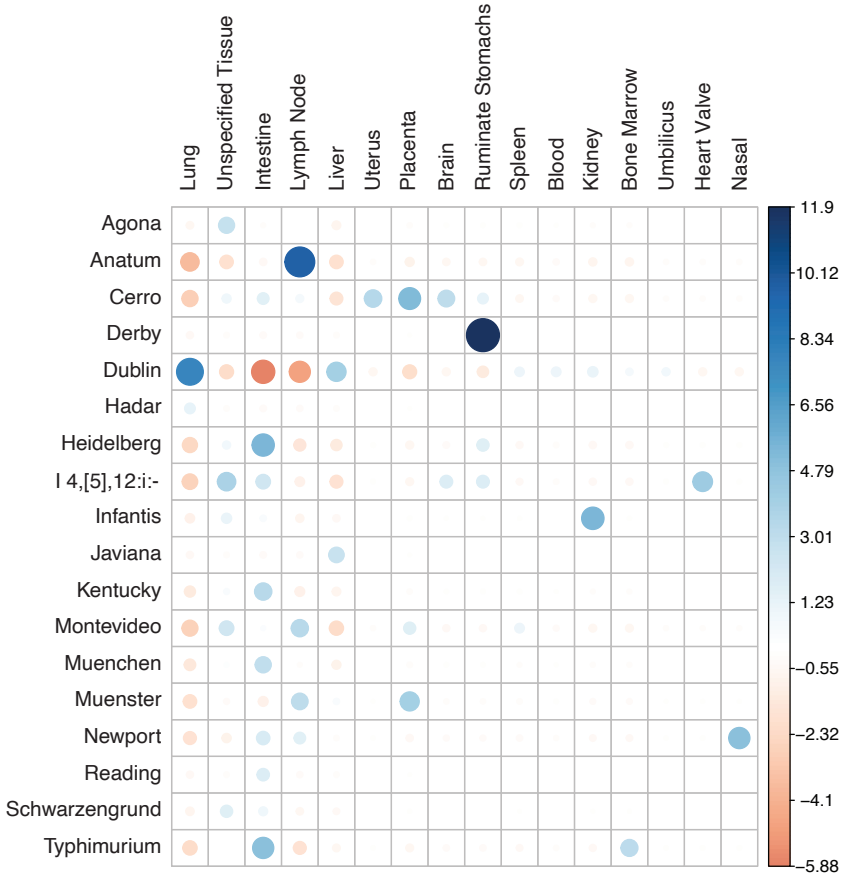
Supplemental TABLE 4. Top 21 *Salmonella* serovars from NCBI 2015 to 2020

Serovar	Chicken	Cattle	Turkey	Swine	Total
Adelaide	1	16	0	370	386
Agona	48	236	285	298	867
Anatum	105	873	135	1,343	2,456
Cerro	36	1,158	2	116	1,312
Derby	74	55	55	1,031	1,215
Dublin	5	1,559	5	7	1,576
Enteritidis	4,517	50	1	41	4,609
Hadar	148	13	17	45	223
Heidelberg	1,327	97	554	158	2,136
I 4,[5],12:i:–	305	187	251	1,101	1,844
Infantis	4,525	264	234	900	5,923
Javiana (control)	6	4	7	9	26
Johannesburg	162	26	3	693	884
Kentucky	7,162	291	85	45	7,583
Montevideo	181	1,635	82	85	1,983
Muenchen	287	408	379	204	1,278
Muenster	20	358	29	58	465
Newport	104	724	71	115	1,014
Reading	32	65	1,049	100	1,246
Schwarzengrund	1,230	62	411	125	1,828
Typhimurium	2,702	624	323	929	4,578

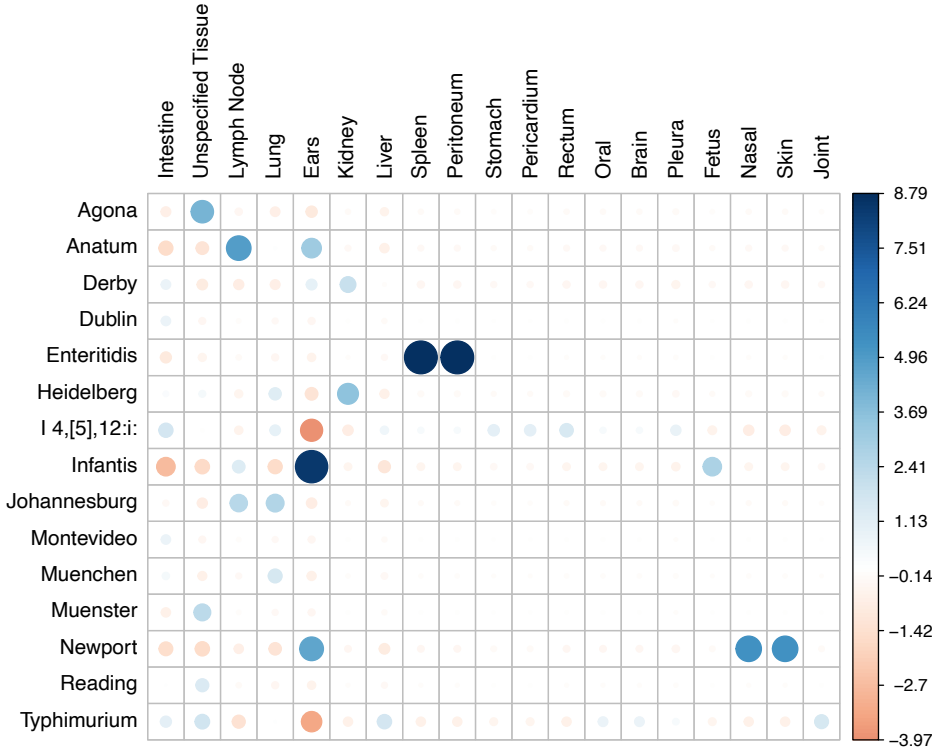
Supplemental TABLE 5. Chi-square results[a]

Animal source	Data source	Data type	χ^2	Degrees of freedom	*P* value
Chicken	FSIS	Isolation source	882.21	32	<2.2e-16
Chicken	NCBI	Isolation source	3,527.00	380	<2.2e-16
Cattle	FSIS	Isolation source	27.23	19	0.10
Cattle	NCBI	Isolation source	1,250.40	220	<2.2e-16
Swine	NCBI	Isolation source	492.5	209	<2.2e-16
Turkey	FSIS	Isolation source	10.67	18	0.91
Turkey	NCBI	Isolation source	1,180.80	144	<2.2e-16
Chicken	NCBI	Organs and tissues	218.21	240	0.84
Cattle	NCBI	Organs and tissues	815.51	255	<2.2e-16
Swine	NCBI	Organs and tissues	511.78	252	<2.2e-16
Turkey	NCBI	Organs and tissues	86.60	81	0.3148

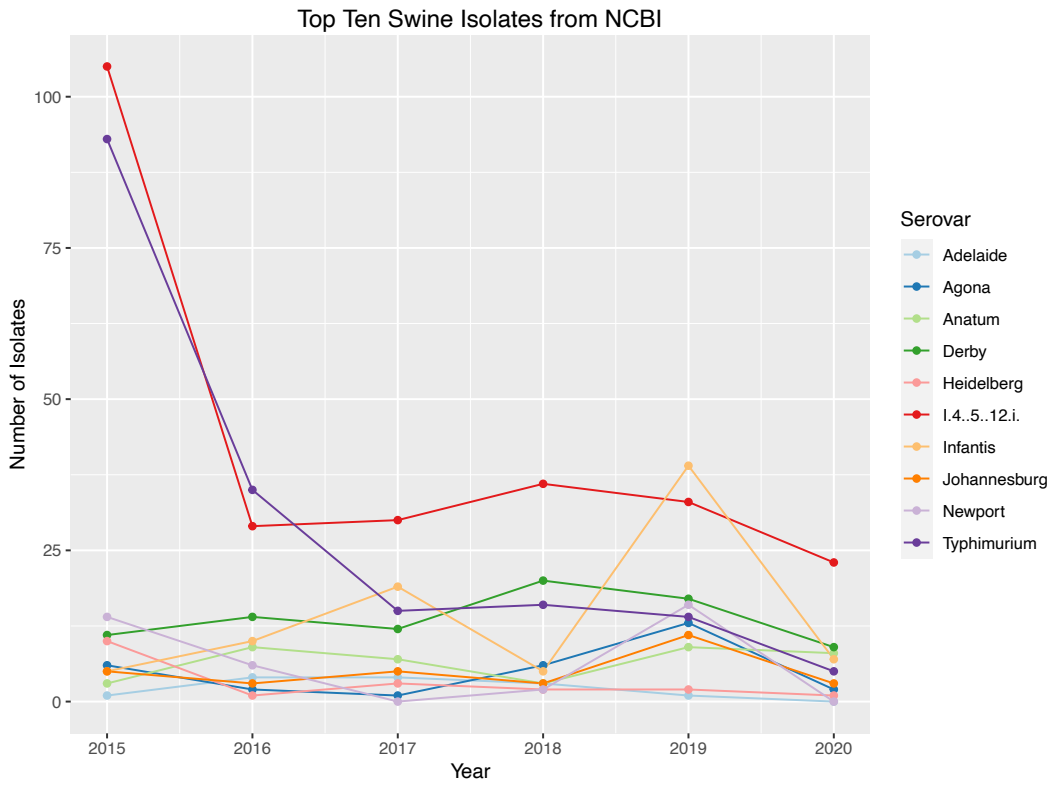
[a]Residual plots were created if $P < 0.05$.



Supplemental Figure 1. NCBI organs and tissues post hoc hypothesis testing with adjusted Pearson residual plot for cattle. Bubble plot of varying degrees of overrepresentation (blue) and underrepresentation (red) in the NCBI data of various serovars in specific cattle organs and tissues.



Supplemental Figure 2. Top 10 *Salmonella* serovars from swine (NCBI) from 2015 to 2020. Line plot of *Salmonella* serovar trends over 6 years in swine from NCBI PD.



Supplemental Figure 3. NCBI organs and tissues post hoc hypothesis testing with adjusted Pearson residual plot for swine. Bubble plot of varying degrees of overrepresentation (blue) and underrepresentation (red) in the NCBI data of various serovars in specific swine organs and tissues.