Complete genome for *Salmonella enterica* subsp. *enterica* serotype Derby associated with the pork sector in France.

3.1 Résumé

Afin de procéder à une analyse fine des différences mono nucléiques entre les lignées identifiées et avant de procéder à l'étude de source-attribution (incluant les souches humaines), j'ai produit 2 génomes circularisés pour *S.* Derby. En effet, bien que ce sérovar figure parmi les 5 plus fréquemment isolés chez l'Homme en Europe, aucun génome de référence n'était encore disponible dans les bases de données en libre accès. J'ai choisi une souche isolée en 2014 d'une carcasse de porc et prélevée dans un abattoir en Bretagne, en France, dont le génome appartenait au ST40. Le génome de cette souche faisait partie de la lignée génomique principale, il était caractérisé par la présence de l'îlot de pathogénicité de *Salmonella* 23 (SPI-23) et des gènes conférant à cette souche la résistance à la streptomycine, aux sulphonamides et à la tétracycline (STR-SSS-TET). J'ai d'ailleurs décrit dans le détail dans cette publication la structure du SPI-23 et de l'îlot génomique portant les gènes de résistance aux antibiotiques cité plus haut et un opéron de résistance au mercure.

Même si ça n'a pas été discuté dans cette publication, la résistance aux métaux lourds est une caractéristique à prendre en considération, car elle confère à certains clones de *Salmonella* la capacité de persister dans l'environnement. Le clone majoritaire du variant monophasique de Typhimurium circulant en Europe présente aussi des gènes lui conférant la résistance à certains métaux lourds (Petrovska et al., 2016).

Un autre génome circularisé issu d'une souche des filières volailles a été produit et valorisé dans la publication III.

3.2 Publication

AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Complete Genome Sequence of *Salmonella enterica* subsp. *enterica* Serotype Derby, Associated with the Pork Sector in France

Yann Sévellec,[a,d] Sophie A. Granier,[a,d] Nicolas Randomski,[a,d] Arnaud Felten,[a,d] Simon Le Hello,[b] Carole Feurer,[c] Michel-Yves Mistou,[a,d] Sabrina Cadel-Six[a,d]

[a]Université Paris-Est, Marne-la-Vallée, France

[b]Institut Pasteur, Centre National de Référence des *Salmonella*, Unité des Bactéries Pathogènes Entériques, Paris, France

[c]French Institute for the Pig and Pork Industry (IFIP), Le Rheu, France

[d]ANSES, Laboratory for Food Safety, Maisons-Alfort, France

ABSTRACT In the European Union, *Salmonella enterica* subsp. *enterica* serovar Derby is the most abundant serotype isolated from pork. Recent studies have shown that this serotype is polyphyletic. However, one main genomic lineage, characterized by sequence type 40 (ST40), the presence of the *Salmonella* pathogenicity island 23, and showing resistance to streptomycin, sulphonamides, and tetracycline (STR-SSS-TET), is pork associated. Here, we describe the complete genome sequence of a strain from this lineage isolated in France.

In the European Union, *Salmonella enterica* subsp. *enterica* serovar Derby (*S.* Derby) was the fifth serovar reported from human cases of salmonellosis in 2016 (0.7%; 325/44,462 confirmed cases) (1). European monitoring data linked this serovar predominantly to pigs and pork meat and, to a lesser extent, turkey and cattle (1). Recent studies based on whole-genome sequencing (WGS) have shown that distinct genomic lineages of *S.* Derby exist, associated with either pork or poultry (2, 3). The main genomic lineage associated with pork is characterized by multilocus sequence typing (MLST) profile 40 (sequence type 40 [ST40]), presence of genes mediating resistance to aminoglycosides, sulfonamides, and tetracyclines (3), and presence of *Salmonella* pathogenicity island 23 (SPI-23), which was previously associated with pork enterocyte invasion (2). We present here the complete genome sequence of *S. enterica* subsp. *enterica* serovar Derby strain 2014LSAL02547, which represents this genomic lineage.

Strain 2014LSAL02547 was isolated in 2014 from a pig carcass sampled at a slaughterhouse in Brittany, France, and identified as belonging to *Salmonella* serovar Derby, according to the White-Kauffmann-Le Minor scheme (4). Its genome was sequenced using Illumina HiSeq (i.e., paired-end read sequencing, 2 × 150 bp) and PacBio (i.e., long-read sequencing) technologies. Concerning the Illumina HiSeq sequencing, genomic DNA was isolated from overnight culture at 37°C on a tryptone soy yeast extract agar plate using the Wizard genomic DNA purification kit (Promega, France) according to the manufacturer's instructions for Gram-negative organisms. The DNA concentration was measured with a Qubit fluorometer, and a gel of 0.8% agarose was used to assess the quality of the extraction (and an eventual degradation of the DNA). Library preparation and sequencing were performed by the Institut du Cerveau et de la Moelle épinière (www.icm-institute.org) using NextEra XT technology and a NextSeq 500 sequencer, respectively (Illumina). Concerning the PacBio sequencing, DNA extraction and library preparation were performed by Genoscreen (Lille, France). DNA was extracted using a Gentra Puregene

Received 30 July 2018 **Accepted** 2 September 2018 **Published** 27 September 2018

Citation Sévellec Y, Granier SA, Randomski N, Felten A, Le Hello S, Feurer C, Mistou M-Y, Cadel-Six S. 2018. Complete genome sequence of *Salmonella enterica* subsp. *enterica* serotype Derby, associated with the pork sector in France. Microbiol Resour Announc 7:e01027-18. https://doi.org/10.1128/MRA.01027-18.

Editor Iddo Friedberg, Iowa State University

Address correspondence to Sabrina Cadel-Six, sabrina.cadelsix@anses.fr.

TABLE 1 Structure of the SGI-1C and SPI-23 from *S.* Derby strain 2014LSAL02547[a]

Genomic island and position	CDS	Protein	Length (bp)
SGI-1C			
1	*int*	Integrase	1,158
2	s002	Helix-turn-helix domain protein	327
3	*rep*	Replication protein	954
4	S004	Hypothetical protein	228
5	S005	Hypothetical protein	2,760
6	S006	Hypothetical protein	534
7	S007	Hypothetical protein	612
8	S008	Hypothetical protein	210
9	S009	Hypothetical protein	291
10	S010	Hypothetical protein	255
11	*traG*	Pilus assembly protein TraG	3,405
12	S012	Conjugative relaxosome accessory transposon protein	1,425
13	S013	Hypothetical protein	858
14	S014	Hypothetical protein	411
15	S015	Hypothetical protein	270
16	S016	Hypothetical protein	213
17	S017	Hypothetical protein	243
18	*intI1*	Phage integrase	966
19	prokka_3715	Hypothetical protein	645
20	*aplIR*	Type 2 restriction enzyme AplI	1,092
21	*taqIM*	Modification methylase TaqI	1,830
22	*hin*	DNA-invertase hin	588
23	*intI1*	Integrase Int1	1,014
24	*aadA2*	Streptomycin 3″-adenylyltransferase	780
25	*qacEdelta1*	Quaternary ammonium compound efflux small multidrug resistance transporter QacE delta 1	348
26	*sul1*	Sulfonamide-resistant dihydropteroate synthase	840
27	*ypeA*	Putative acetyltransferase	501
28	*intB*	Transposase/IS protein	786
29	*istA*	Integrase core domain IS*26*	1,515
30	*tniB*	Bacterial TniB protein	861
31	*tnsB*	Transposon Tn*7* transposition protein TnsB	1,680
32	*cph2*	Phytochrome-like protein cph2	708
33	*merE*	Mercury resistance protein MerE	237
34	*mta*	Zinc-responsive transcriptional regulator	363
35	*merA*	Mercuric reductase	1,695
36	*merC*	Mercuric resistance protein MerC	423
37	*merP*	Mercuric transport protein periplasmic component precursor	276
38	*merT*	MerT mercuric transport protein	351
39	*merR*	Mercuric resistance operon regulatory protein	435
40	prokka_3736	Hypothetical protein (putative relaxase)	243
41	*tetR*	Tetracycline repressor protein class A from transposon 1721	678
42	*tetA*	Tetracycline resistance protein, class C	1,200
43	*yedA*	Putative inner membrane transporter YedA	783
44	*tnsB*	Transposon Tn*7* transposition protein TnsB	708
45	prokka_3741	Hypothetical protein	1,035
SPI-23			
1	*intA3*	Integrase A	1,275
2	prokka_01810	Hypothetical protein	252
3	prokka_01811	Abortive infection phage resistance protein	1,113
4	prokka_01812	Bacterial shuffle protein	1,488
5	prokka_01813	Major subunit of bundle-forming pilus precursor	558
6	*bfpA*	Conjugal transfer protein TraD	306
7	*traD*	Hypothetical protein	1,515
8	prokka_01816	Hypothetical protein	735
9	prokka_01817	Hypothetical protein	576
10	prokka_01818	Hypothetical protein	531
11	prokka_01819	Hypothetical protein	342
12	prokka_01820	Hypothetical protein	471
13	prokka_01821	Hypothetical protein	552
14	prokka_01822	Hypothetical protein	180
15	prokka_01823	Hypothetical protein	240
16	prokka_01824	Hypothetical protein	1,251

(Continued on next page)

TABLE 1 (Continued)

Genomic island and position	CDS	Protein	Length (bp)
17	prokka_01825	Hypothetical protein	282
18	prokka_01826	Hypothetical protein	438
19	prokka_01827	Hypothetical protein	558
20	prokka_01828	Hypothetical protein	183
21	prokka_01829	Hypothetical protein	477
22	prokka_01830	Hypothetical protein	2,730
23	prokka_01831	Hypothetical protein	504
24	prokka_01832	Hypothetical protein	93
25	prokka_01833	Hypothetical protein	870
26	prokka_01834	Hypothetical protein	588
27	prokka_01835	Hypothetical protein	984
28	prokka_01836	RNA pyrophosphohydrolase	444
29	prokka_01837	Hypothetical protein	519
30	prokka_01838	Hypothetical protein	261
31	prokka_01839	Hypothetical protein	735
32	prokka_01840	Hypothetical protein	738
33	prokka_01841	hypothetical protein	480
34	hns_2	DNA binding protein H_NS	405
35	prokka_01843	Hypothetical protein	549
36	prokka_01844	Hypothetical protein	1,272
37	prokka_01845	Hypothetical protein	297
38	prokka_01846	Hypothetical protein	291
39	prokka_01847	Hypothetical protein	219
40	prokka_01848	Hypothetical protein	876
41	prokka_01849	Hypothetical protein	525

[a]GenBank accession number CP029486. SGI-1C, *Salmonella* genomic island 1 type C (bases 427735 to 472096); SPI-23, *Salmonella* pathogenicity island 23 (bases 2369809 to 2406412).

kit (Qiagen), the DNA concentration was assessed using a Qubit fluorometer, and the DNA extract's quality was checked by agarose gel electrophoresis. Library preparation was made by DNA fragmentation and ligation of single-molecule real-time (SMRT) adaptors. Prior to sequencing, BluePippin size selection (Sage Science) was set at 15 kb in order to achieve identical sequence overlaps. PacBio sequencing was performed on one SMRT cell. Quality of the Illumina and PacBio reads was examined using FastQC v0.11.5 (5). Prinseq v0.20.4 (6) was used to select Illumina long reads of good quality (no undefined bases; Phred, >30; length, >60 kb). SMRT Analysis v2.3.0 software was used to assemble PacBio reads. In SMRT Analysis, the Hierarchical Genome Assembly Process (HGAP) v3.0 (7) was invoked to correct the subreads (length, >1,000 bases; read score, 0.8), and Celera v8.3 (8) was used for assembly (subread length, >500 bases; deep coverage, >25×). SAMtools v1.5 (9) was used to map the Illumina short paired-end reads against the PacBio assembly to correct potential assembly mistakes and to determine the depth of the final assembly. The final deep coverage obtained was 146×. A unique 4.86-Mb contig was obtained, with a GC content of 51.12%. The genome was annotated using Prokka (10). It included 4,549 coding sequences (CDS) and 88 tRNAs.

Genes mediating antimicrobial resistance phenotype STR-SSS-TET (showing resistance to streptomycin, sulfonamides, and tetracycline) are part of the *Salmonella* genomic island 1 (SGI-1) (3), described in Table 1. This SGI-1 element integrated a cluster of mercury resistance genes (*merA*, *merC*, *merP*, *merT*, and *merR*) located in a Tn*7* transposon. SPI-23 from the 2014LSAL02547 genome was 36,603 bp long and was located between the *DAD50_12070* and *mftA* genes (Table 1).

Data availability. This whole-genome assembly sequence was deposited in the NCBI database under the accession number CP029486. The SRA accession numbers are SRX3643218 (Illumina reads) and SRX4523973 (PacBio reads).

ACKNOWLEDGMENTS

This study was funded by the French Ministry of Agriculture, Food and Forestry, by the *Salmonella* Network from the French Agency for Food, Environmental and Occu-

pational Health and Safety (ANSES) Laboratory for Food Safety and by INAPORC. Y.S. is the recipient of a doctoral fellowship (DGER-ANSES) cofunded by AgroParisTech and the ANSES.

REFERENCES

1. European Food Safety Authority. 2017. The European Union summary report on trends and sources of zoonoses, zoonotic agents and food-borne outbreaks in 2016. EFSA J 15:e05077. https://doi.org/10.2903/j.efsa.2017.5077.
2. Hayward MR, AbuOun M, La Ragione RM, Tchorzewska MA, Cooley WA, Everest DJ, Petrovska L, Jansen VA, Woodward MJ. 2014. SPI-23 of *S.* Derby: role in adherence and invasion of porcine tissues. PLoS One 9:e107857. https://doi.org/10.1371/journal.pone.0107857.
3. Sévellec Y, Vignaud M-L, Granier SA, Lailler R, Feurer C, Le Hello S, Mistou M-Y, Cadel-Six S. 2018. Polyphyletic nature of *Salmonella enterica* serotype Derby and lineage-specific host-association revealed by genome-wide analysis. Front Microbiol 9:891. https://doi.org/10.3389/fmicb.2018.00891.
4. Grimont PAD, Weill FX. 2007. Antigenic formulae of the *Salmonella* serovars, 9th ed. WHO Collaborating Center for Reference and Research on *Salmonella*, Institut Pasteur, Paris, France.
5. Andrews S. 2010. FastQC: a quality control tool for high throughput sequence data. http://www.bioinformatics.babraham.ac.uk/projects/fastqc.
6. Schmieder R, Edwards R. 2011. Quality control and preprocessing of metagenomic datasets. Bioinformatics 27:863–864. https://doi.org/10.1093/bioinformatics/btr026.
7. Chin C-S, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Clum A, Copeland A, Huddleston J, Eichler EE, Turner SW, Korlach J. 2013. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. Nat Methods 10:563–569. https://doi.org/10.1038/nmeth.2474.
8. Myers EW, Sutton GG, Delcher AL, Dew IM, Fasulo DP, Flanigan MJ, Kravitz SA, Mobarry CM, Reinert KH, Remington KA, Anson EL, Bolanos RA, Chou HH, Jordan CM, Halpern AL, Lonardi S, Beasley EM, Brandon RC, Chen L, Dunn PJ, Lai Z, Liang Y, Nusskern DR, Zhan M, Zhang Q, Zheng X, Rubin GM, Adams MD, Venter JC. 2000. A whole-genome assembly of *Drosophila*. Science 287:2196–2204. https://doi.org/10.1126/science.287.5461.2196.
9. Li H. 2011. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. Bioinformatics 27:2987–2993. https://doi.org/10.1093/bioinformatics/btr509.
10. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30:2068–2069. https://doi.org/10.1093/bioinformatics/btu153.

4 Publication III : Phylogenomic analysis of *Salmonella* enterica subsp. enterica serovar Derby circulating in Europe, Asia and the United States and associated with the poultry sector.

4.1 Résumé

Salmonella Derby est considéré à l'origine comme un sérovar adapté au porc. Cependant depuis plusieurs années, on observe en Europe une émergence de *S.* Derby chez les troupeaux d'engraissement de dinde (EFSA, 2015; 2016; 2017; EFSA and ECDC, 2018). Les souches reportées dans cette filière correspondent au profil MLST ST71.

J'ai séquencé au cours de ma thèse à l'aide des technologies Illumina et PacBio un génome complet correspondant à la lignée ST71 de *S*. Derby pour produire un génome de référence. J'ai conduit une analyse génomique sur une sélection de souches internationales (dont les génomes ont été extraits d'Enterobase (Alikhan et al., 2018)) et de souches issues de ma collection Enterobase afin, d'évaluer la diversité génétique et la résistance aux antibiotiques des souches de cette lignée et de comparer les souches humaines avec les souches issues de la volaille (dinde, poulet et caille).

Mes résultats montrent que le ST71 est isolé chez la volaille dans au moins 5 pays européens (France, Allemagne, Angleterre, Italie, Pologne) ainsi qu'en Asie et aux Etats Unis. Cette lignée peut être séparée en 3 clades statistiquement distincts correspondant chacun à une zone géographique : Europe, Asie et USA. Plusieurs souches isolées chez la dinde présentaient un profil de multirésistance aux antibiotiques avec notamment une émergence de la résistance aux quinolones chez les souches italiennes issues de la dinde. Des profils de resistances similaires avaient déjà été isolées en Espagne et au Royaume Uni (EFSA and ECDC, 2018). Un SGI-1A a également été mis en évidence chez les souches asiatiques porteuses de gènes de résistance aux antibiotiques. Une étude des phages portée par les différents clades du ST71 révèle des profils variés dépendant de l'origine géographique des souches.

4.2 Publication

Phylogenomic analysis of Salmonella enterica subsp. enterica serovar Derby circulating in Europe, Asia and the United States and associated with the poultry sector.

Yann Sévellec[1], Arnaud Felten[1], Nicolas Radomski[1], Sophie A. Granier[1], Simon Le Hello[2], Liljana Petrovska[3], Michel-Yves Mistou[1] and Sabrina Cadel-Six[1]*

1. Université PARIS-EST, ANSES, Laboratory for Food Safety, F-94701 Maisons-Alfort, France.

2. Institut Pasteur, Centre National de Référence des *Salmonella*, Unité des Bactéries Pathogènes Entériques, Paris, France.

3. Animal and Plant Health Agency, Addlestone, UK.

Running title: Phylogeny of *S.* Derby from poultry sector and humans.

Corresponding author: Sabrina Cadel-Six, sabrina.cadelsix@anses.fr

ANSES, Laboratory for Food Safety, 14 Rue Pierre et Marie Curie, 94701 Maisons-Alfort

Keywords: *Salmonella* Derby, reference genome ST71, phylogenetic analysis, poultry sector, human infection, antimicrobial resistance, phage analysis.

ABSTRACT

Salmonella Derby seems to be emerging in Europe as a predominant serovar in fattening turkey flocks. This serovar was recorded as predominant in the turkey sector in 2015 in the United Kingdom (UK). In 2017, only two years later, it was recorded in the turkey and broiler sectors in Europe by two other Member States (MSs) (Ireland and Spain). The genomic lineage corresponding to the MLST profile 71 seems to characterise these *S.* Derby isolates. In this study, for the first time, we explored the diversity of 30 *S.* Derby ST71 genomes isolated from poultry and humans in five of Europe's leading producing countries for turkey meat. We also included genomes from the United States and Asia (n=3). The phylogenomic analysis undertaken on the core genome underlined the presence of three clades of which one included all of the European strains. Our results showed that *S.* Derby ST71 notified for the turkey sector is circulating in at least five European MSs: the UK, Germany, Poland, Italy and France. This clone was also isolated in humans in the UK and France, showing its ability to cause gastroenteritis. Another matter of concern is that antimicrobial resistance investigated by genome sequencing showed, in Italian genomes, several mutations involved in resistance to quinolone. Interestingly, phage analysis revealed a variety of profiles that seemed related to geographic origin. These results constitute the baseline for following the routes of transmission within the poultry food trade of this emerging pathogen and identifying appropriate prevention measures of control.

IMPORTANCE. Given that Europe covers 36% of world turkey meat production and that *S.* Derby is known for its ability to infect humans, it is essential to characterize the clones emerging in the European turkey and broiler sectors. The results of this study show that the European *S.* Derby ST71 clones differs from the US and Asian ones by an average of 82 single-nucleotides polymorphism (SNPs) and a standard deviation (SD) of 64 SNPs. Even though the European genomes analysed are closely related with 35 SNPs and a SD of 16 SNPs, certain sub-types can be easily identified through an analysis of the genes involved in antimicrobial resistance (Resfinder) and the phages (PHASTER). The complete closed genome of one European *S.* Derby ST71 clone has been produced, providing a significant contribution to the understanding of this serotype for which no complete genome was available until now.

INTRODUCTION

The European Union (EU) is a major producer of turkey meat worldwide with 80% of production supplied by six Member States (MSs) (Germany, France, Poland, Italy, Spain and the United Kingdom (UK)). Since 2014, the UK has reported a high occurrence of *Salmonella* enterica subsp. enterica serovar Derby (*S.* Derby) in the turkey sector compared to the other MSs (1-3). These *Salmonella* Derby isolates seem to belong to the multilocus sequence typing (MLST) profile 71 (ST71) (4). This profile has been shown to characterise a genomic lineage of *S.* Derby strains isolated in the poultry sector in the UK and France (4, 5). Interestingly, in the EU in 2016, 21.0% of *S.* Derby isolates came from turkeys and 11.3% from broilers, immediately after pigs (64.4%). The most frequent serovar detected in fattening turkey flocks was *S.* Derby (21.8%). Multidrug resistance was reported in 35.7% of the 143 isolates collected by three MSs (the UK, Ireland and Spain). Extremely high levels of resistance (>70%) to sulfamethoxazole and tetracycline were found. Resistance to fluoroquinolone was detected in the UK and Spain. Even though the most common sources of *S.* Derby are pigs and pork meat (6-8), the emergence of a clone in the turkey and broiler sectors is to be expected. This emergence is especially a matter of concern since *S.* Derby ranks among the top five most commonly reported serovars in human cases in the EU and poultry meat is the main source of human contamination (1-3). According to data recorded by the World Health Organization between 2007 and 2015, the burden of pathogens commonly transmitted through the food chain increased worldwide and this global rise was mainly associated with the increased consumption of products of animal origin (9). Among these, white meat (pork and broiler meat) played a major role due to increased consumption in line with population growth and the economic crisis (10). *Salmonella* enterica subsp. enterica is one of the most common and widespread foodborne pathogens. The most frequent food vehicles, reported in Europe for several years and for 2016, were white meat and eggs (3). Serovars Typhimurium and Derby are strongly associated with white meat (2). In Europe in 2015, they ranked first and fifth, accounting for 23.4% and 5.3% of the 14,596 serotyped *Salmonella* isolates from food and animals, respectively. Furthermore, Typhimurium and Derby ranked among the seven serovars most frequently isolated from humans over the last decade in Europe, making them a major public health concern. However, while *S.* Typhimurium in the poultry sector is subject to European regulations, these do not apply for *S.* Derby isolates. Consequently, *S.* Derby clones carrying multiple pathogenicity and antimicrobial resistance traits can circulate within and across European MSs, escaping all measures of control.

This study aimed to explore the genomic diversity of *S.* Derby strains isolated from poultry and humans at the European and international levels. Using phylogenomic and statistical analyses based on single-nucleotide polymorphisms (SNPs), we analysed 30 genomes of *S.* Derby ST71 isolated primarily from the poultry sector (n=19) and from humans (n=11) in various European MSs as well as the United States and Asia. To identify potential genome signatures, we detected and characterised phages within the *Salmonella* Derby genome. Since *S.* Derby is

known to have multidrug-resistant strains (2, 7, 11-14), the identification of acquired antimicrobial resistance genes was also investigated.

RESULTS

The analysed *S.* Derby ST71 genomes were genetically closely related, with an average of 82 SNPs and a standard deviation (SD) of 64 SNPs even though the isolates came from different continents and origins (i.e. humans, animals and food). Nevertheless, the phylogenomic reconstruction showed that three different clades could be identified within this lineage. The pairwise association for each clade is given in Table S1. A Kolmogorov-Smirnov test confirmed statistical differences between the pairwise SNP differences of these three clades. We obtained p-values < 0.01 between the European and American strains as well as between the European and Asian strains. The UK *S.* Derby genomes were genetically related to the strains isolated in the other MSs with an average of 35 SNPs and a standard deviation (SD) of 15 SNPs (Fig. 1).

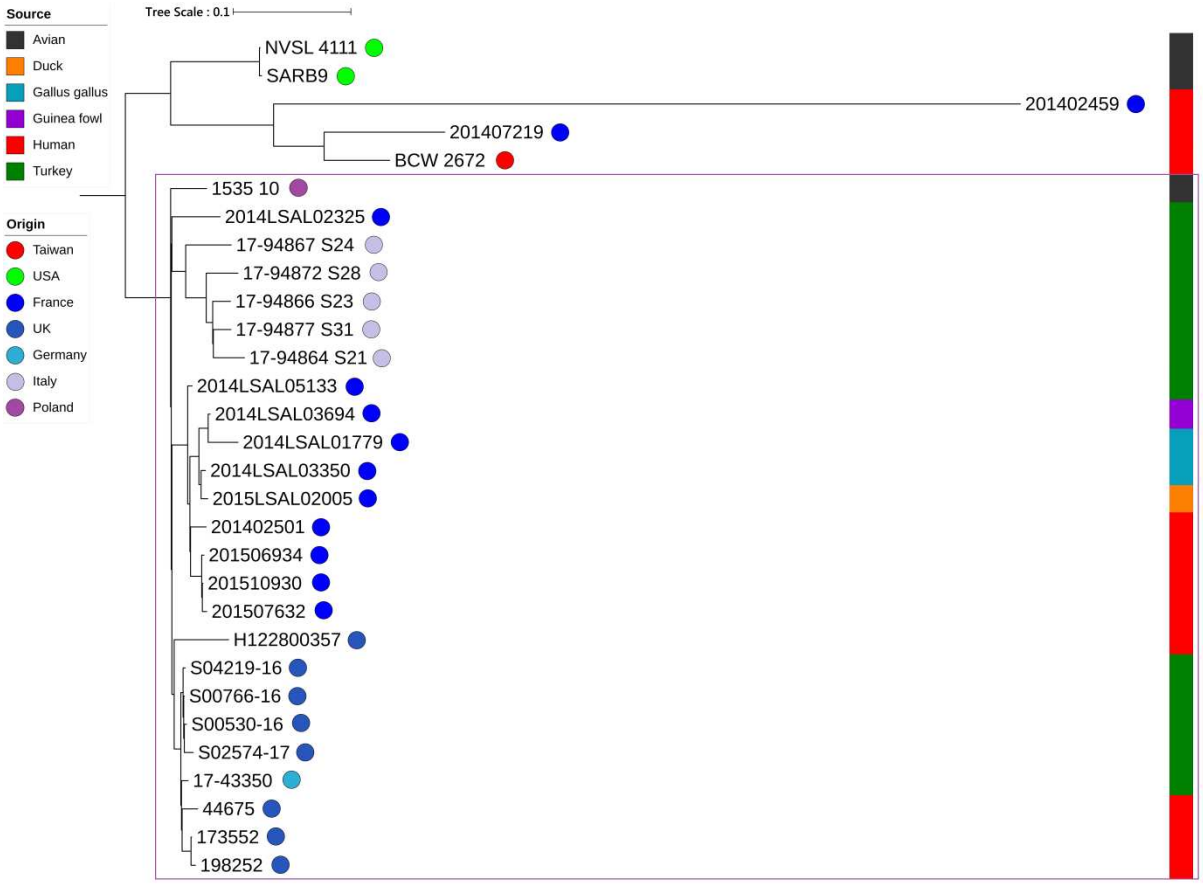

Figure 1. Phylogenetic reconstruction based on SNPs including strains of the ST71 profile and their geographic distributions and origins.

The maximum likelihood criterion and the GTR-gamma model were used. The scale bar indicates the number of substitutions per site. Three clades were identified; the European clade is highlighted by a violet box.

The complete and partial antimicrobial resistance genes identified in the 30 *S.* Derby genomes studied are listed in Table S2. No resistance gene was detected in 60% (18/30) of the 30 studied genomes. Thirty-three percent (10/30) could be classified as multidrug resistant as they harboured more than three different acquired resistance mechanisms, either by mutations or gene acquisitions. An S83Y gyrA mutation, known to mediate resistance to fluoroquinolones, was detected in four poultry-related strains from Italy. Interestingly, the two genomes from the USA displayed no known antimicrobial resistance mechanisms. Conversely, the Asian strain was resistant to aminoglycosides, beta-lactams, sulphonamides, tetracyclines and phenicols and the exact same mechanisms were identified in a French human isolate (Fig. 2).

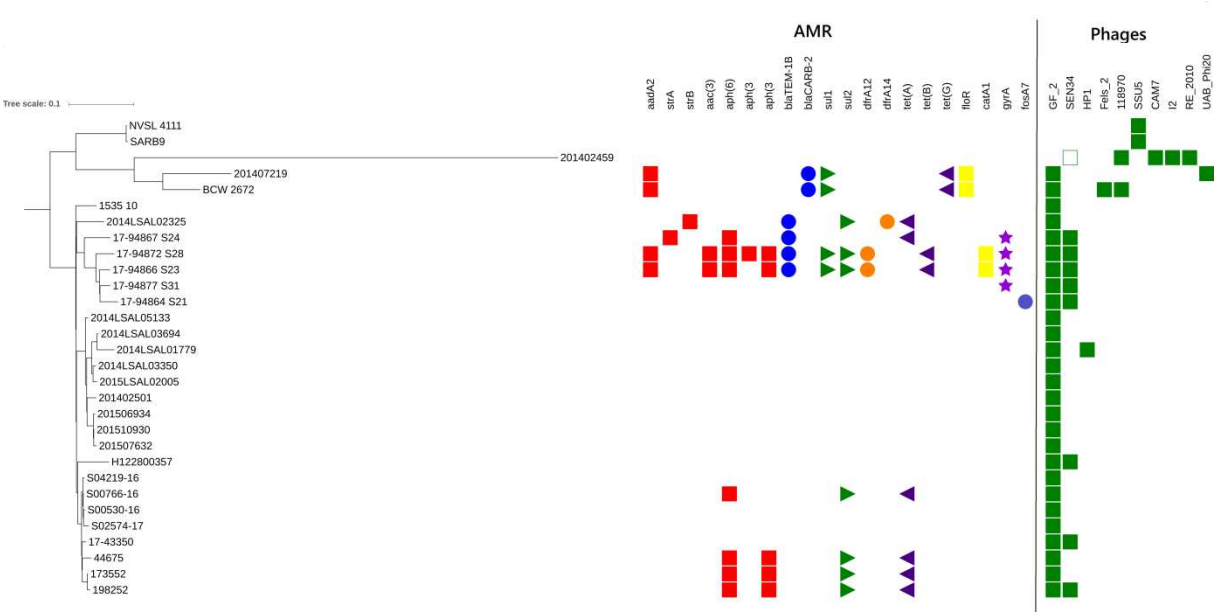

Figure 2. AMR and phage profiles according to the phylogeny.

The red cubes correspond to the resistance gene to aminoglycosides antibiotic, the blue circle to beta-lactam, the green triangle to sulfonamides, the orange circle to trimethoprim, the purple triangle to tetracycline, the yellow cube to phenicols, the purple stars correspond to a gyrA mutation conferring resistance to quinolones and the grey circle to resistance to fosfomycines.

PHASTER software was used to identify intact prophages and their integrase genes within the food-associated and clinical *Salmonella* isolates. A total of 10 intact prophages and 10 different integrase genes (PHASTER was unable to identify any integrase genes in some prophages and identified several in others) were identified and are listed in Table S3 and Fig. 2.

DISCUSSION

The three ST71 Derby clades obtained in this study were correlated with the geographic distribution of the strains (Fig. 1). The first clade contained two strains from the USA, the second had one strain from Asia and two from French patients, one of whom reported travel to Thailand, and the third included all strains from Europe (France, Germany, Poland, Italy and

the UK). The results of this study showed that the ST71 *S.* Derby isolated from turkeys in the UK were genetically related to the strains isolated in the other MSs. Approximately 36% of the world's turkey meat production originates in Europe, with only five countries producing 80% of this: Germany, France, Poland, the UK, Italy and Spain (15). The data on *S.* Derby in turkeys reported by EFSA in 2014 and 2015 for the UK (1, 2) may have been due to the particular attention paid by this country to the incidence of *S.* Derby related to white meat and the tendency of this serovar to become persistent in poultry houses (16). The UK has indeed been the second leading MS for the consumption of white meat after Spain since 2000 (15, 17). Moreover, more than 22 million turkeys are produced for meat each year in the UK and this large-scale activity is dominated by a small number of specialist producers (18). More than half of the farms (>50%) in the UK have a capacity of 50,000 animals (more than 2,400 m^2), compared with only 6.7% of those in France (ITAVI, 2013). In France, the incidence of *S.* Derby in the turkey animal sector is higher than in Gallus gallus in keeping with the European trend. This supports the hypothesis that the "Derby turkey clone" is not confined to the UK but is also circulating within other MSs, suggesting the possibility of an earlier common source. Data from the Anses *Salmonella* Network (jointly with the National Reference Laboratory) show that this serovar is the most frequently isolated from pigs and turkeys in the animal sector in France. *S.* Derby ST71 has been shown to be responsible for human infections in France. Between 2014 and 2015, Institut Pasteur (National Reference Centre for *Salmonella* in France) reported six patients infected with *S.* Derby ST71 representing 2% of the infections due to *S.* Derby. Of these six patients, two were likely to have contracted the infection while travelling abroad (strains 201402459 and 201407219 in Fig. 1) and the remaining four in France. The strains isolated from these four patients were clustered with the French food strains 2015LSAL02005, 2014LSAL03350, 2014LSAL03694, 2014LSAL01779 (13 SNPs with an SD of 8) and 2014LSAL05133 (9 SNPs with an SD of 1). No resistance gene was detected in the French strains responsible for the four French human infections or the five related food strains in the cluster. Nevertheless, the French strain 2014LSAL02325 isolated from turkey meat presented genes involved in resistance to aminoglycosides, sulphonamides and tetracyclines. These genes were also carried by Italian and English strains isolated from patients and turkeys. The Italian strains isolated from turkeys revealed mutations involved in resistance to fluoroquinolones. Most human *Salmonella* infections are self-limiting gastrointestinal illnesses and do not require any antimicrobial treatment. However, in rare cases, infection may spread and lead to severe enteric disease or invasive infection. In such life-threatening cases, rapid and effective antimicrobial treatment is of utmost importance to prevent poor outcomes in patients. Fluoroquinolones are widely recommended to treat severe *Salmonella* infections in adults (21). The emergence of a fluoroquinolone-resistant lineage of the *S.* Derby ST71 European clade in turkey flocks observed in Italy is worrisome. It should be noted that 24% of the *S.* Derby isolated from fattening turkey flocks in Europe in 2016 were resistant to fluoroquinolones. They were identified in Spain and the UK (21).

Interestingly, in the clade associated with Asia, the French human strain 201407219 isolated in 2014 and the strain BCW_2672 isolated from patients in Taiwan in 2000 shared the same

genes conferring resistance to ampicillin, chloramphenicol, streptomycin, sulfamethoxazole and tetracycline. These two genomes showed the presence of a *Salmonella* genomic island (SGI-1) as already described in *S.* Derby by Beutlich et al. (22). Strain 201407219 displayed a partial sequence of the sul1 gene. Remarkably, those two genomic islands were different from the SGI-1C recently identified in two ST40 French strains isolated from pork meat (14).

Additionally, PHASTER identified the presence of a wide variety of prophages in the food-associated isolates compared to the clinical isolates. We were able to identify three profiles within the EU clade, the first characterised by the phage GF-2 (Edward_GF_2_NC_026611) (France, Poland and UK genomes), the second by the phages GF-2 and SEN34 (Salmon_SEN34_NC_028699) (Germany, Italy and UK genomes) and the third by the phages GF-2 and HP1 (Haemop_HP1_NC_001697) (France genomes). Interestingly, with regards to the turkey isolates, we were able to identify only two of the three profiles described above: the profile GF-2 characterising the French and English isolates and the profile GF-2 + SEN34 characterising the German and Italian isolates. The sequence of the GF-2 phage was present in 90% (27/30) of the genomes analysed and in all of the EU genomes. Interestingly, the GF-2 phage, a myovirus lytic bacteriophage, was described for the first time by Yasuike et al. in 2015 (23). It was isolated from a cultured Japanese flounder that succumbed to Edwardsiella tarda, a gram-negative Enterobacteriaceae that is the most serious pathogen in both marine and freshwater fish farms worldwide. Since viruses are indeed the most abundant biological entity in aquatic ecosystems (24-27), it is possible that *Salmonella* Derby may be a natural host of this bacteriophage. Further infectivity tests are needed to confirm this hypothesis.

While the phylogenomic analysis performed on the core genome SNP was able to combine 25 genomes into one ST71 European clade, differences between those genomes were highlighted by the analysis of targeted elements from accessory genomes. Antimicrobial resistance genes and the presence of phages allowed us to distinguish different genomic signatures revealing a diversity of profiles that seemed related to geographic or animal origins. This study demonstrated the presence of an *S.* Derby clone ST71 circulating in Europe, not confined solely to the UK, apparently well adapted to turkey flocks and responsible for human salmonellosis. SNP analysis on the core genome is an undeniably efficient phylogenomic tool for clustering epidemic clones or identifying emerging successful clones. This study also highlighted the power of user-friendly web-based applications such as Resfinder and PHASTER, not requiring any highly qualified bioinformatics expertise, to readily rule out potential sources of contamination during foodborne outbreak investigations. Lastly, SNP analysis of the core genome enabled three *S.* Derby ST71 clades to be identified, one for EU, one for Asian and one for US strains. Analysis of the targeted elements from accessory genomes enabled the identification of potentially useful genomic signatures for monitoring transmission routes for foodborne microbiological hazards along the food chain.

MATERIAL AND METHODS

Whole genome sequencing. To investigate genetic diversity within the *S.* Derby lineage isolated in poultry, the strain 2014LSAL01779 isolated in 2014 from Gallus gallus carcasses in Brittany (France) presenting the MLST profile ST71 was sequenced with Illumina Hiseq (i.e. paired-end read sequencing) and PacBio® (i.e. long read sequencing) technologies and used as reference genome for the downstream analyses. Illumina Hiseq sequencing was performed by the Institut du Cerveau et de la Moelle épinière (ICM) (Pitié-Salpêtrière Hospital, Paris) (www.icm-institute.org) using NextEra XT technology and PacBio® sequencing was performed by Genoscreen (Lille). The quality of the Illumina and PacBio® reads was examined using fastQC V0.11.5 (28). Prinseq V0.20.4 (29) was used to select Illumina long reads of good quality (no undefined bases, phred > 30, length > 60). The PacBio® reads were assembled using SMRT analysis v2.3.0. In SMRT analysis, HGAP V3.0 (30) was invoked to correct the sub-reads (length > 1000 bases, read-Score 0.8) and Celera V8.3 (31) was used for the assembling (sub-reads length > 500 bases, deep coverage > 25x). Samtools V1.5 (32) was used to map the Illumina short paired-end reads against the PacBio assembly to correct potential assembly mistakes and to determine the depth of the final assembly. The final deep coverage obtained was 140x. A unique contig of 4.86Mb was obtained with GC-content of 51.12%. The genome was annotated using Prokka (33); it contained 4499 CDS and 90 tRNA. The reference genome as well as all the other genomes were paired-end sequenced as described above (i.e. Hiseq sequencing) and processed using the iVARCall2 workflow (34) excluding duplications and applying a second alignment around small insertions/deletions.

Accession number. The complete genome sequence of *S.* enterica subsp. enterica serovar Derby strain 2014LSAL01779 included in this work has been submitted to NCBI under accession no. CP026609.

Worldwide *S.* Derby genome selection. Information about the analysed *S.* Derby genomes has been listed in Table S1. Apart from 16 genomes taken from the ANSES (14), Institut Pasteur and HAPA collections, the other full genomes with conventional taxonomic information were collected across public Enterobase databases (http://Enterobase.warwick.ac.uk/species/senterica/search_strains).

Phylogenetic analysis. The phylogenetic analysis was performed with the CSIPhylogeny tool provided by the Center for Genomic Epidemiology (CGE) (https://cge.cbs.dtu.dk/services/CSIPhylogeny/) (35) using the maximum likelihood criterion and GTR-gamma model. The phylogenetic tree was visualised using iTOL (36).

Statistical analyses. The non-normality of the data (i.e. pairwise SNP differences) was checked using the Shapiro test (37) with R from the pairwise matrix generated by the iVARCall2 workflow described above. Equality of variances for pairwise SNP differences was rejected with the Fisher test (38). Distributions of pairwise SNP differences were compared with a Kolmogorov-Smirnov test (KS-test) (39).

Identification of acquired resistance genes. The whole panel of genomes was analysed using the ResFinder 2.1 application (40) on the CGE server. The threshold for reporting a match between a gene in the ResFinder database and the input *S.* Derby genome was set at 90% identity over at least 3/5 of the length of the resistance gene. For strains 201407219 and BCW_2672, SGI-1 coding sequences (NCBI: AF261825.2) were extracted and blasted against the dataset with the BioNumerics BLAST tool. The complete genomic sequence of SGI-1 was investigated using the BioNumerics alignment and sequence visualisation tools.

Detection and characterisation of phages and prophages. Each of the 30 assembled genomes was analysed by PHASTER to identify the presence of prophages and their integrase genes (41). Only prophages identified as “intact” or “questionable” were considered. The identity of all intact prophage sequences detected by PHASTER was confirmed by BLAST (42, 43).

Authors’ contributions

SCS piloted and administered the project. SCS and YS designed and developed the experiments. YS, SG and MLV carried out the experiments and the analyses. SCS, MYM, RL and CF provided acquisitions. YS, SCS, SG, MYM drafted the manuscript. SLH and CF participated in the discussion and reviewed the report.

Acknowledgments

This study was supported by the French Ministry of Agriculture, Food and Forestry and the European Union Horizon 2020 research and innovation program under grant agreement 643476 to the COMPARE project (http://www.compare-europe.eu). We are indebted to the *Salmonella* Network that provided the French food strains analyzed in the study. Yann Sévellec is the recipient of a doctoral fellowship (DGER-ANSES) co-financed by AgroParisTech and the French Agency for Food, Environmental and Occupational Health & Safety (ANSES).

REFERENCES

1. EFSA. 2015. The European Union summary report on trends and sources of zoonoses, zoonotic agents and food-borne outbreaks in 2014. EFSA Journal 13:4329-n/a, 10.2903/j.efsa.2015.4329.

2. EFSA. 2016. The European Union summary report on trends and sources of zoonoses, zoonotic agents and food-borne outbreaks in 2015. EFSA Journal 14:e04634-n/a, 10.2903/j.efsa.2016.4634.

3. EFSA. 2017. The European Union summary report on trends and sources of zoonoses, zoonotic agents and food-borne outbreaks in 2016. EFSA Journal 15:e05077, doi:10.2903/j.efsa.2017.5077.

4. Hayward MR, Petrovska L, Jansen VA, Woodward MJ. 2016. Population structure and associated phenotypes of *Salmonella* enterica serovars Derby and Mbandaka overlap with host range. BMC Microbiol 16:15, 10.1186/s12866-016-0628-4.

5. Hayward MR, AbuOun M, La Ragione RM, Tchorzewska MA, Cooley WA, Everest DJ, Petrovska L, Jansen VA, Woodward MJ. 2014. SPI-23 of *S.* Derby: role in adherence and invasion of porcine tissues. PLoS One 9:e107857, 10.1371/journal.pone.0107857.

6. EFSA. 2009. Analysis of the baseline survey on the prevalence of *Salmonella* in holdings with breeding pigs in the EU, 2008 - Part A: *Salmonella* prevalence estimates. EFSA Journal 7:1377-n/a, 10.2903/j.efsa.2009.1377.

7. Valdezate S, Vidal A, Herrera-Leon S, Pozo J, Rubio P, Usera MA, Carvajal A, Echeita MA. 2005. *Salmonella* Derby clonal spread from pork. Emerg Infect Dis 11:694-8, 10.3201/eid1105.041042.

8. Simon S, Trost E, Bender J, Fuchs S, Malorny B, Rabsch W, Prager R, Tietze E, Flieger A. 2017. Evaluation of WGS based approaches for investigating a food-borne outbreak caused by *Salmonella* enterica serovar Derby in Germany. Food Microbiol doi:https://doi.org/10.1016/j.fm.2017.08.017 https://doi.org/10.1016/j.fm.2017.08.017.

9. WHO. 2015. World Health Day 2015: Food safety, on WHO. http://www.who.int/campaigns/world-health-day/2015/event/en/. Accessed

10. IPCC. 2013. Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change doi:10.1017/CBO9781107415324. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, ISBN 978-1-107-66182-0.

11. Hauser E, Hebner F, Tietze E, Helmuth R, Junker E, Prager R, Schroeter A, Rabsch W, Fruth A, Malorny B. 2011. Diversity of *Salmonella* enterica serovar Derby isolated from pig, pork and humans in Germany. Int J Food Microbiol 151:141-9, 10.1016/j.ijfoodmicro.2011.08.020.

12. Keelara S, Scott HM, Morrow WM, Gebreyes WA, Correa M, Nayak R, Stefanova R, Thakur S. 2013. Longitudinal study of distributions of similar antimicrobial-resistant *Salmonella* serovars in pigs and their environment in two distinct swine production systems. Appl Environ Microbiol 79:5167-78, 10.1128/AEM.01419-13.

13. Kerouanton A, Rose V, Weill FX, Granier SA, Denis M. 2013. Genetic diversity and antimicrobial resistance profiles of *Salmonella* enterica serotype derby isolated from pigs, pork, and humans in France. Foodborne Pathog Dis 10:977-84, 10.1089/fpd.2013.1537.

14. Sévellec Y, Vignaud M-L, Granier SA, Lailler R, Feurer C, Le Hello S, Mistou M-Y, Cadel-Six S. 2018. Polyphyletic Nature of *Salmonella* enterica Serotype Derby and Lineage-Specific

Host-Association Revealed by Genome-Wide Analysis. Front Microbiol 9 10.3389/fmicb.2018.00891.

15. ITAVI. 2013. Structures et organisation des filières volailles de chair en Europe.

16. Mueller-Doblies D, Carrique-Mas JJ, Davies RH. 2014. A study of the dynamics of *Salmonella* infection in turkey breeding, rearing and finishing houses with special reference to elimination, persistence and introduction of *Salmonella*. Avian Pathol 43:146-154, 10.1080/03079457.2014.892569.

17. Malpel GP, Marigeaud M, Marty S. 2014. La filière volaille de chair.,

18. Martin J. 2009. The Commercialisation of British Turkey Production. Rural History 20:209-228, 10.1017/S0956793309990057.

19. DGAl. 2014. bilan 2014 des plans de surveillance et de contrôle. Surveillance sanitaire des denrées animales et végétales en France:69-74,

20. DGAl. 2015. bilan 2015 des plans de surveillance et de contrôle. Surveillance sanitaire des denrées animales et végétales en France:112-125,

21. EFSA, ECDC. 2018. The European Union summary report on antimicrobial resistance in zoonotic and indicator bacteria from humans, animals and food in 2016. EFSA Journal 16:5182, 10.2903/j.efsa.2018.5182.

22. Beutlich J, Jahn S, Malorny B, Hauser E, Huhn S, Schroeter A, Rodicio MR, Appel B, Threlfall J, Mevius D, Helmuth R, Guerra B. 2011. Antimicrobial resistance and virulence determinants in European *Salmonella* genomic island 1-positive *Salmonella* enterica isolates from different origins. Appl Environ Microbiol 77:5655-64, 10.1128/aem.00425-11.

23. Yasuike M, Nishiki I, Iwasaki Y, Nakamura Y, Fujiwara A, Sugaya E, Kawato Y, Nagai S, Kobayashi T, Ototake M, Nakai T. 2015. Full-genome sequence of a novel myovirus, GF-2, infecting Edwardsiella tarda: comparison with other Edwardsiella myoviral genomes. Arch Virol 160:2129-33, 10.1007/s00705-015-2472-5.

24. Suttle C. 1994. The Significance of Viruses to Mortality in Aquatic Microbial Communities, vol 28,

25. Wommack KE, Colwell RR. 2000. Virioplankton: viruses in aquatic ecosystems. Microbiol Mol Biol Rev 64:69-114,

26. Bettarel Y, Amblard C, Sime-Ngando T, Carrias JF, Sargos D, Garabetian F, Lavandier P. 2003. Viral lysis, flagellate grazing potential, and bacterial production in Lake Pavin. Microb Ecol 45:119-27, 10.1007/s00248-002-1016-1.

27. Weinbauer MG. 2004. Ecology of prokaryotic viruses. FEMS Microbiol Rev 28:127-81, 10.1016/j.femsre.2003.08.001.

28. Andrews S. 2014. FastQC A Quality Control tool for High Throughput Sequence Data,

29. Schmieder R, Edwards R. 2011. Quality control and preprocessing of metagenomic datasets. Bioinformatics 27:863-4, 10.1093/bioinformatics/btr026.

30. Chin CS, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Clum A, Copeland A, Huddleston J, Eichler EE, Turner SW, Korlach J. 2013. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. Nat Methods 10:563-9, 10.1038/nmeth.2474.

31. Myers EW, Sutton GG, Delcher AL, Dew IM, Fasulo DP, Flanigan MJ, Kravitz SA, Mobarry CM, Reinert KH, Remington KA, Anson EL, Bolanos RA, Chou HH, Jordan CM, Halpern AL, Lonardi S, Beasley EM, Brandon RC, Chen L, Dunn PJ, Lai Z, Liang Y, Nusskern DR, Zhan M, Zhang Q, Zheng X, Rubin GM, Adams MD, Venter JC. 2000. A whole-genome assembly of Drosophila. Science 287:2196-204,

32. Li H. 2011. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. Bioinformatics 27:2987-93, 10.1093/bioinformatics/btr509.

33. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30:2068-9, 10.1093/bioinformatics/btu153.

34. Felten A, Vila Nova M, Durimel K, Guillier L, Mistou M-Y, Radomski N. 2017. First gene-ontology enrichment analysis based on bacterial coregenome variants: insights into adaptations of *Salmonella* serovars to mammalian- and avian-hosts. BMC Microbiology 17:222, 10.1186/s12866-017-1132-1.

35. Kaas RS, Leekitcharoenphon P, Aarestrup FM, Lund O. 2014. Solving the Problem of Comparing Whole Bacterial Genomes across Different Sequencing Platforms. PLOS ONE 9:e104984, 10.1371/journal.pone.0104984.

36. Letunic I, Bork P. 2016. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. Nucleic Acids Res 44:W242-W245, 10.1093/nar/gkw290.

37. Royston P. 1995. Remark AS R94: A Remark on Algorithm AS 181: The W-test for Normality. Journal of the Royal Statistical Society Series C (Applied Statistics) 44:547-551, 10.2307/2986146.

38. Markowski CA, Markowski EP. 1990. Conditions for the Effectiveness of a Preliminary Test of Variance. The American Statistician 44:322-326, 10.2307/2684360.

39. Huang X, Huang Q, Dun Z, Huang W, Wu S, Liang J, Deng X, Zhang Y. 2016. Nontyphoidal *Salmonella* Infection, Guangdong Province, China, 2012. Emerging Infectious Diseases 22:726-729, 10.3201/eid2204.151372.

40. Zankari E, Hasman H, Cosentino S, Vestergaard M, Rasmussen S, Lund O, Aarestrup FM, Larsen MV. 2012. Identification of acquired antimicrobial resistance genes. J Antimicrob Chemother 67:2640-2644, 10.1093/jac/dks261.

41. Arndt D, Grant JR, Marcu A, Sajed T, Pon A, Liang Y, Wishart DS. 2016. PHASTER: a better, faster version of the PHAST phage search tool. Nucleic Acids Res 44:W16-21, 10.1093/nar/gkw387.

42. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. 2009. BLAST+: architecture and applications. BMC Bioinformatics 10:421, 10.1186/1471-2105-10-421.

43. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J Mol Biol 215:403-10, 10.1016/s0022-2836(05)80360-2.

4.2.1 Supplementary material

Supplementary Table 1. S. Derby genomes information analyzed in this study.

Name	ST profiles	Database	Source Niche	Source Type	Source Details	Collection Year	Continent	Country	Sample ID
17-43350_S13	ST71	Enterobase	Food	Avian	Meleagris; Food, turkey meat	2015	Europe	Germany	SAMEA104379658
17-94867_S24	ST71	Enterobase	Poultry	Avian	Meleagris; Animal, animal	2011	Europe	Italy	SAMEA104379733
17-94864_S21	ST71	Enterobase	Poultry	Avian	Meleagris; Animal, swab	2011	Europe	Italy	SAMEA104379730
17-94877_S31	ST71	Enterobase	Wild animal	Avian	Meleagris; Animal, feces	2013	Europe	Italy	SAMEA104379740
17-94872_S28	ST71	Enterobase	Wild animal	Avian	Meleagris; Animal, feces	2012	Europe	Italy	SAMEA104379737
17-94866_S23	ST71	Enterobase	Wild animal	Avian	Meleagris; Animal, feces	2011	Europe	Italy	SAMEA104379732
1535_10	ST71	Enterobase	Poultry	Avian	NS	2010	Europe	Poland	SAMEA104437530
BCW_2672	ST71	Enterobase	Human	Human	Human; *Homo sapiens*	2000	Asia	Taiwan	SAMN02368552
44675	ST71	Enterobase	Human	Human	Human; *Homo sapiens*	2014	Europe	UK	SAMN03465613
H122800357	ST71	Enterobase	Human	Human	Human; *Homo sapiens*	2012	Europe	UK	SAMN03168996
173552	ST71	Enterobase	Human	Human	human; *Homo sapiens*	2015	Europe	UK	SAMN06680388
198252	ST71	Enterobase	Human	Human	human; *Homo sapiens*	2015	Europe	UK	SAMN06680393
SARB9 (Fidelma Boyd)	ST71	Enterobase	NS	Avian	Bird	1986	North America	United States	NA
NVSL 4111	ST71	Enterobase	Wild animal	Avian	NS	1986	North America	United States	SAMN02367710
2014LSAL02325	ST71	ANSES	Poultry	Avian	Meat from turkey - carcass	2014	Europe	France	SAMN07734901

Name	ST profiles	Database	Source Niche	Source Type	Source Details	Collection Year	Continent	Country	Sample ID
2014LSAL05133	ST71	ANSES	Poultry	Avian	Meat from turkey - carcass	2014	Europe	France	SAMN07734953
2015LSAL02005	ST71	ANSES	Poultry	Avian	Meat from duck - fresh	2015	Europe	France	SAMN07734993
2014LSAL03694	ST71	ANSES	Poultry	Avian	Meat from guinea fowl	2014	Europe	France	SAMN07734940
2014LSAL03350	ST71	ANSES	Poultry	Avian	Meat from broilers - *Gallus gallus* - carcass	2014	Europe	France	SAMN07734914
2014LSAL01779	ST71	ANSES	Poultry	Avian	Meat from broilers - *Gallus gallus* - carcass	2014	Europe	France	SAMN08470240
201402459	ST71	Institut Pasteur	Human	Human	Human; *Homo sapiens*	2014	Europe	France	SAMN09080915
201407219	ST71	Institut Pasteur	Human	Human	Human; *Homo sapiens*	2014	Europe	France	SAMN09080917
S00530-16	ST71	APHA	Poultry	Avian	*Meleagris*	2016	Europe	UK	ERR2230775
S00766-16	ST71	APHA	Poultry	Avian	*Meleagris*	2016	Europe	UK	ERR2230777
S02574-17	ST71	APHA	Poultry	Avian	*Meleagris*	2017	Europe	UK	ERR2230782
S04219-16	ST71	APHA	Poultry	Avian	*Meleagris*	2016	Europe	UK	ERR2230787
201402501	ST71	Institut Pasteur	Human	Human	Human; *Homo sapiens*	2014	Europe	France	SAMN09080916
201506934	ST71	Institut Pasteur	Human	Human	Human; *Homo sapiens*	2015	Europe	France	SAMN09080917
201507632	ST71	Institut Pasteur	Human	Human	Human; *Homo sapiens*	2015	Europe	France	SAMN09080919
201510930	ST71	Institut Pasteur	Human	Human	Human; *Homo sapiens*	2015	Europe	France	SAMN09080920

NS = not specified; NA = not available.

Supplementary table 2. S. Derby antimicrobial resistance genes

Strain	Sector	Source Details	Localization	Aminoglycosides	Beta-lactams	Sulphonamides	Tetracyclines	Trimetroprim	Phenicols	Quinolones	Fosfomycines
17-43350_S13	Poultry	Meleagris; Food, turkey meat	Germany								
17-94864_S21	Poultry	Meleagris; Animal, swab	Italy								fosA7
17-94866_S23	Poultry	Meleagris; Animal, feces	Italy	aac(3)-Iid; aadA2; aph(6)-Id; aph(3'')-Ib	blaTEM-1B	sul1; sul2	tet(B)	dfrA12	catA1	gyrA p.S83Y	
17-94867_S24	Poultry	Meleagris; Animal, Animal	Italy	strA; aph(6)-Id	blaTEM-1B		tet(A)			gyrA p.S83Y	
17-94872_S28	Poultry	Meleagris; Animal, feces	Italy	aac(3)-Iid; aadA2; aph(6)-Id; aph(3')-Ia; aph(3'')-Ib	blaTEM-1B	sul1; sul2	tet(B)	dfrA12	catA1	gyrA p.S83Y	
17-94877_S31	Poultry	Meleagris; Animal, feces	Italy							gyrA p.S83Y	
H122800357	Human	Human; Homo sapiens	UK								

198252	Human	Human; Homo sapiens	UK	aph(3'')-Ib; aph(6)-Id;		sul2	tet(A)				
Strain	**Sector**	**Source Details**	**Localization**	**Aminoglycosides**	**Beta-lactams**	**Sulphonamides**	**Tetracyclines**	**Trimetroprim**	**Phenicols**	**Quinolones**	**Fosfomycines**
44675	Human	Human; Homo sapiens	UK	aph(3'')-Ib; aph(6)-Id;		sul2	tet(A)				
173552	Human	Human; Homo sapiens	UK	aph(3'')-Ib; aph(6)-Id;		sul2	tet(A)				
S00530-16	Poultry	Meleagris	UK								
S00766-16	Poultry	Meleagris	UK	aph(6)-Id		sul2	tet(A)				
S02574-17	Poultry	Meleagris	UK								
S04219-16	Poultry	Meleagris	UK								
1535_10	Avian	NS	Poland								
2014LSAL01779	Poultry	Meat from broilers - Gallus gallus - carcass	France								
2014LSAL02325	Poultry	Meat from turkey - carcass	France	ΔstrA + strB	blaTEM-1B	sul2	tet (A)	drfrA14			
2014LSAL03350	Poultry	Meat from broilers - Gallus gallus - carcass	France								

2014LSAL03694	Poultry	Meat from guinea fowl	France								
Strain	**Sector**	**Source Details**	**Localization**	**Aminoglycosides**	**Beta-lactams**	**Sulphonamides**	**Tetracyclines**	**Trimetroprim**	**Phenicols**	**Quinolones**	**Fosfomycines**
2014LSAL05133	Poultry	Meat from turkey - carcass	France								
2015LSAL02005	Poultry	Meat from duck - fresh	France								
201402501	Human	Human	France								
201506934	Human	Human	France								
201507632	Human	Human	France								
201510930	Human	Human	France								
201407219	Human	Human	France	aadA2	blaCARB-2	sul1	tet(G)		floR		
201402459	Human	Human	France								
BCW_2672	Human	Human; Homo sapiens	Taiwan	aadA2	blaCARB-2	sul1	tet(G)		floR		
NVSL_4111	Avian	NS	USA								
SARB9 (Fidelma Boyd)	Avian	Bird	USA								

NS = not specified. The European clade is underlined by a violet box

Supplementary table 3. S. Derby's phages.

Strain	Sector	Source Details	Localization	Profil name	Edward_GF_2_NC_026611	Haemop_HP1_NC_001697	Salmon_SEN34_NC_028699	Salmon_Fels_2_NC_010463	Salmon_118970_sal3_NC_031940	Salmon_SSU5_NC_018843	Synech_S_CAM7_NC_031927	Entero_I2_2_NC_001332	Salmon_RE_2010_NC_019488	Entero_UAB_Phi20_NC_031019
17-43350_S13	Poultry	Meleagris; Food, turkey meat	Germany	**GF-2 and SEN34**	2	0	2	0	0	0	0	0	0	0
17-94864_S21	Poultry	Meleagris; Animal, swab	Italy	**GF-2 and SEN34**	2	0	2	0	0	0	0	0	0	0
17-94866_S23	Poultry	Meleagris; Animal, feces	Italy	**GF-2 and SEN34**	2	0	2	0	0	0	0	0	0	0
17-94867_S24	Poultry	Meleagris; Animal, Animal	Italy	**GF-2 and SEN34**	2	0	2	0	0	0	0	0	0	0
17-94872_S28	Poultry	Meleagris; Animal, feces	Italy	**GF-2 and SEN34**	2	0	2	0	0	0	0	0	0	0
17-94877_S31	Poultry	Meleagris; Animal, feces	Italy	**GF-2 and SEN34**	2	0	2	0	0	0	0	0	0	0
H122800357	Human	Human; Homo sapiens	UK	**GF-2 and SEN34**	2	0	2	0	0	0	0	0	0	0
198252	Human	Human; Homo sapiens	UK	**GF-2 and SEN34**	2	0	2	0	0	0	0	0	0	0
44675	Human	Human; Homo sapiens	UK	**GF-2 and SEN34**	2	0	0	0	0	0	0	0	0	0

173552	Human	Human; Homo sapiens	UK	**GF-2**	2	0	0	0	0	0	0	0	0	0
S00530-16	Poultry	Meleagris	UK	**GF-2**	2	0	0	0	0	0	0	0	0	0
S00766-16	Poultry	Meleagris	UK	**GF-2**	2	0	0	0	0	0	0	0	0	0
S02574-17	Poultry	Meleagris	UK	**GF-2**	2	0	0	0	0	0	0	0	0	0
S04219-16	Poultry	Meleagris	UK	**GF-2**	2	0	0	0	0	0	0	0	0	0
1535_10	Avian	NS	Poland	**GF-2**	2	0	0	0	0	0	0	0	0	0
Strain	**Sector**	**Source Details**	**Localization**	**Profil name**	Edward_GF_2_NC_026611	Haemop_HP1_NC_001697	Salmon_SEN34_NC_028699	Salmon_Fels_2_NC_010463	Salmon_118970_sal3_NC_031940	Salmon_SSU5_NC_018843	Synech_S_CAM7_NC_031927	Entero_I2_2_NC_001332	Salmon_RE_2010_NC_019488	Entero_UAB_Phi20_NC_031019
2014LSAL01779	Poultry	Meat from broilers - Gallus gallus - carcass	France	**GF-2 and HP1**	2	2	0	0	0	0	0	0	0	0
2014LSAL02325	Poultry	Meat from turkey - carcass	France	**GF-2**	2	0	0	0	0	0	0	0	0	0
2014LSAL03350	Poultry	Meat from broilers - Gallus gallus - carcass	France	**GF-2**	2	0	0	0	0	0	0	0	0	0
2014LSAL03694	Poultry	Meat from guinea fowl	France	**GF-2**	2	0	0	0	0	0	0	0	0	0
2014LSAL05133	Poultry	Meat from turkey - carcass	France	**GF-2**	2	0	0	0	0	0	0	0	0	0

2015LSAL02005	Poultry	Meat from duck - fresh	France	**GF-2**	2	0	0	0	0	0	0	0	0	0
201402501	Human	Human; Homo sapiens	France	**GF-2**	2	0	0	0	0	0	0	0	0	0
201506934	Human	Human; Homo sapiens	France	**GF-2**	2	0	0	0	0	0	0	0	0	0
201507632	Human	Human; Homo sapiens	France	**GF-2**	2	0	0	0	0	0	0	0	0	0
201510930	Human	Human; Homo sapiens	France	**GF-2**	2	0	0	0	0	0	0	0	0	0
201407219	Human	Human; Homo sapiens	France	**GF-2 and UAB**	2	0	0	0	0	0	0	0	0	2
201402459	Human	Human; Homo sapiens	France	**118970-CAM7-I2-RE**	0	0	1	0	2	0	2	2	2	0
BCW_2672	Human	Human; Homo sapiens	Taiwan	**Fels-118970**	2	0	0	2	2	0	0	0	0	0
NVSL_4111	Avian	NS	USA	**SSU5**	0	0	0	0	0	2	0	0	0	0
SARB9 (Fidelma Boyd)	Avian	Bird	USA	**SSU5**	0	0	0	0	0	2	0	0	0	0