Epidemiological Analysis of *Salmonella enterica* subsp. *enterica* Serovar Dublin in German Cattle Herds Using Whole-Genome Sequencing

Silvia García-Soto,[a] Herbert Tomaso,[a] Jörg Linde,[a] Ulrich Methner[a]

[a]Institute of Bacterial Infections and Zoonoses, Friedrich Loeffler Institute, Jena, Germany

ABSTRACT *Salmonella enterica* subsp. *enterica* serovar Dublin is a cattle-adapted serovar that causes enteritis and systemic diseases in animals. In Germany, *S.* Dublin is not detected or is very rarely detected in some federal states but is endemic in certain regions. Information on detailed genetic characteristics of *S.* Dublin is not available. An understanding of the paths and spreading of *S.* Dublin within and between regions and over time is essential to establish effective control strategies. Whole-genome sequencing (WGS) and bioinformatic analysis were used to explore the genetic traits of *S.* Dublin and to determine their epidemiological context. Seventy-four *S.* Dublin strains collected in 2005 to 2018 from 10 federal states were studied. The phylogeny was analyzed using core-genome single-nucleotide polymorphisms (cgSNPs) and core-genome multilocus sequence typing. Genomic clusters at 100 cgSNPs, 40 cgSNPs, and 15 cgSNPs were selected for molecular epidemiology. WGS-based genoserotyping confirmed serotyping. Important specific virulence determinants were detected in all strains, but multidrug resistance in German *S.* Dublin organisms is uncommon. Use of different thresholds for cgSNP analysis enabled a broad view and also a detailed view of the occurrence of *S.* Dublin in Germany. Genomic clusters could be allocated nationwide, to a limited number of federal states, or to special regions only. Results indicate both persistence and spread of *S.* Dublin within and between federal states in short and longer time periods. However, to detect possible routes of infection or persistence of *S.* Dublin indicated by genomic analysis, information on the management of the cattle farms and contacts with corresponding farms are essential.

IMPORTANCE *Salmonella enterica* subsp. *enterica* serovar Dublin is a bovine host-adapted serovar that causes up to 50% of all registered outbreaks of salmonellosis in cattle in Germany. *S.* Dublin is not detected or is only rarely detected in some federal states but has been endemic in certain regions of the country for a long time. Information on genetic traits of the causative strains is essential to determine routes of infection. WGS and bioinformatic analysis should be used to explore the genetic characteristics of *S.* Dublin. Combining the genomic features of *S.* Dublin strains with information on the management of the cattle farms concerned should enable the detection of possible routes of infection or persistence of *S.* Dublin. This approach is regarded as a prerequisite to developing effective intervention strategies.

KEYWORDS whole-genome sequencing, *Salmonella enterica* subsp. *enterica* serovar Dublin, bioinformatics, epidemiology, cattle, Germany

*S*almonella enterica* subsp. *enterica* serovar Dublin is defined as bovine adapted but not bovine restricted and affects calves, young cattle, and adult cattle, causing enteritis and/or systemic disease (1). A wide range of clinical manifestations, from acute clinical disease with high mortality rates and abortions over typhoid-like illness with diarrhea to a chronic carrier state without any clinical signs of infection, may occur (1). Natural infections with the bovine host-adapted serovar *S.* Dublin may occur also in other animal species,

particularly small ruminants like sheep and goats (2). Although it is very rare, *S.* Dublin may also cause human infections; because of its high virulence, it is associated with death and thus is considered a global public health concern (3).

In Germany, outbreaks of salmonellosis in cattle herds officially confirmed by a competent authority are notifiable (4). Between 2010 and 2019, approximately 110 to 130 outbreaks of bovine salmonellosis from different serovars were recorded each year. *Salmonella enterica* subsp. *enterica* serovar Typhimurium was the dominant serovar during this period and caused approximately 40% to 50% of the annually reported cases. *S.* Dublin amounted to 30% to 40% of all registered outbreaks each year. *Salmonella enterica* subsp. *enterica* serovar Enteritidis caused approximately 5% to 10% of all outbreaks. All other serovars are summarized and as a group were the cause of 15% to 20% of all outbreaks of German cattle salmonellosis per year. However, there is no single dominant serovar in this group, and no serovar shows an increasing detection rate over longer periods. Compared with *S.* Typhimurium, the cattle-adapted serovar *S.* Dublin is not detected or is rarely detected in some federal states but is endemic in certain regions (5, 6). In Germany, no detailed information is available on the genetic traits of *S.* Dublin isolates from different regions or over longer periods. However, a better understanding of the paths of infection and spreading of the organism within and between regions and over time might give valuable information to establish effective measures to control the disease. Pulsed-field gel electrophoresis (PFGE) has been shown to provide valuable data to separate *Salmonella* organisms of different serovars (7–9). Whereas a Danish study showed some ability to differentiate between strains of *S.* Dublin (10), results from Sweden have provided unsatisfactory resolution among isolates of this serovar (11). Whole-genome sequencing (WGS) has been used successfully for the characterization of *Salmonella* organisms belonging to this and other serovars (11–14). Core-genome multilocus sequence typing (cgMLST)- and core-genome single-nucleotide polymorphism (cgSNP)-based typing are widely used for *Salmonella* genotyping. While cgMLST uses allele numbers for each gene for typing, single-nucleotide polymorphisms (SNPs) are called and used for typing and for building phylogenetic trees in the second approach. Both methods provide a high capability to discriminate between closely related and nonrelated isolates (15). In addition, WGS data give information regarding the genetic context of the isolates, such as virulence genes and antimicrobial resistance (AMR) determinants.

In view of the lack of information regarding the epidemiology of the host-adapted serovar *S.* Dublin in the German cattle population, this study aimed to characterize *S.* Dublin strains that originated from different federal states in Germany in the period between 2005 and 2018. *S.* Dublin organisms isolated during that period were examined phenotypically at the German National Reference Laboratory (NRL) for Salmonellosis in Cattle. WGS and bioinformatic analysis were used (i) to describe the genetic traits of *S.* Dublin strains, (ii) to analyze the phylogeny using cgSNPs as well as cgMLST, and (iii) to compare the results with recent studies on *S.* Dublin genomes in other European countries.

RESULTS

Serotyping and antimicrobial susceptibility testing of *S.* Dublin strains. The isolates were typed according to the Kauffmann-White scheme and revealed the complete antigenic formula (1,9,12:g,p:−) for *S.* Dublin. All 74 *S.* Dublin strains, apart from strain 197, were not resistant to the antimicrobial substances tested (see Data Set S1 in the supplemental material). The MIC values of the *S.* Dublin organisms were as follows: sulfamethoxazole, <512 μg/ml; trimethoprim, <0.25 to 1 μg/ml; ciprofloxacin, 0.03 to 0.06 μg/ml; tetracycline, 4 to 8 μg/ml; meropenem, 0.03 to 0.06 μg/ml; azithromycin, 8 to 16 μg/ml; nalidixic acid, 8 to 16 μg/ml; cefotaxime, <0.25 μg/ml; chloramphenicol, <8 μg/ml; tigecycline, 1 to 2 μg/ml; ceftazidime, <0.05 μg/ml; colistin, 1 to 2 μg/ml; ampicillin, <1 to 2 μg/ml; gentamicin, 1 to 2 μg/ml. Strain 197 showed resistance to sulfamethoxazole (>512 μg/ml), trimethoprim (>32 μg/ml), tetracycline (>32 μg/ml), nalidixic acid (64 μg/ml), and ampicillin (>32 μg/ml).

General genomic features and *in silico* serotyping. WGS of 74 *S.* Dublin strains resulted in an average of 1,804,798 reads per sample (range, 912,103 to 2,912,074 reads per sample), with an average read length of 230 bp (see Data Set S1), leading to an average read coverage of 84-fold (range, 37- to 139-fold). The assembled genomes

consisted of 29 contigs, on average (range, 24 to 115 contigs). The genome size of the strains was 4,878,835 bp, on average (range, 4,833,405 to 5,034,978 bp).

WGS-based serotyping performed by SISTR and SeqSero correlated with traditional serotyping predicting the formula 9:g,p:− for serovar *S.* Dublin for all strains (see Data Set S1). Classic multilocus sequence typing (MLST) revealed sequence types (STs) ST10 and ST3734, which is a single-locus variant of ST10 in the gene *sucA* (strain 671). The majority of the strains (69/74 strains) were assigned to ST10, and only one sample (strain 713) was assigned to ST3734. In the case of four strains (strains 523, 921, 1720, and 2218), a ST could not be assigned.

AMR genes. All *S.* Dublin strains harbored the chromosomal gene *aac(6′)-Iaa* for aminoglycoside resistance. Only strain 197 from the data set presented a multidrug resistance (MDR) pattern according to ResFinder, NCBI, and AMRFinder databases (see Data Set S2). The MDR pattern consisted of six resistance genes, i.e., *aac(6′)-Iaa* and *aph(6)-Id* for aminoglycoside resistance (streptomycin), *dfrA14* for trimethoprim resistance, *sul2* for sulfonamide resistance, *tet*(A) for tetracycline resistance, and *bla*$_{TEM-1}$ for β-lactamase resistance. No chromosomal point mutations connected to AMR were found among the strains except for the isolates 89 and 1065, which were positive for a point mutation in the gene *acrB* (*acrB*-R717Q) that might cause macrolide (azithromycin) resistance according to AMRFinder.

Virulence determinants. Of the 22 *Salmonella* pathogenicity islands (SPIs) searched, 13 were identified within all 74 *S.* Dublin strains, including SPIs 1 to 6, 9, 11, 13, 14, 16, 17, and 19 (see Data Set S2). Regarding virulence determinants associated with the *S.* Dublin invasome, all strains presented the *spv* operon (*spvRABCD* genes), as well as the putative virulence gene *st313-td* (GenBank accession no. NC_016854), the type VI secretion system (T6SS) (accession no. NC_011205.1), and the lamboid prophage Gifsy-2 (GenBank accession no. NC_010393.1) (56% coverage). Among the other virulence determinants detected, the majority of strains (63/74 strains) revealed a total of 109 virulence genes, 10/74 isolates harbored less than 109 genes, and 1 strain (strain 2303) presented a virulence pattern of 114 genes (see Data Set S2). The 80-kb serovar-specific virulence plasmid pSDV (GenBank accession no. DQ115388.2) was found in all 74 isolates (see Data Set S2).

Phylogenetic analysis and clustering. To establish in-depth phylogeography of *S.* Dublin in Germany, cgSNP calling and cgMLST were performed. The cgSNP alignment consisted of 2,068 bp. The pairwise average cgSNP distance between German strains was 119 SNPs (range, 1 to 592 SNPs). Using cgMLST, the genome sequences of the German *S.* Dublin strains were mapped onto 3,002 predefined cgMLST target loci. For all 74 strains examined, an average of 98.4% of the targets were classified as "good targets" by Ridom SeqSphere (range, 97.4% to 98.6%). The average pairwise distance between strains was 68 alleles (range, 0 to 311 alleles) (see Data Set S3). For a higher resolution of *S.* Dublin phylogeny, we constructed a maximum-likelihood phylogenetic tree based on cgSNPs (Fig. 1) and a minimum spanning tree (MST) based on cgMLST. Both methods indicated (i) the occurrence of highly similar strains in very different regions in Germany and (ii) specific clusters of *S.* Dublin strains in only certain regions.

To study the phylogeography of *S.* Dublin in Germany in more detail, clustering of cgSNP and cgMLST results was performed to identify general genomic clusters (100 cgSNPs), intermediate clusters (40 cgSNPs), and more closely related strains (15 cgSNPs) (Fig. 1). For clustering based on cgMLST results using the MST, the same cutoff values as for cgSNP clustering were used (see Data Set S4). In general, the two methods provided similar results. Clustering based on 15 allelic differences in cgMLST, compared with clustering based on 15 cgSNPs, revealed mainly concordance; in only a few cases, cgMLST-based clusters were larger (see Data Set S4).

Both core-genome methods revealed higher resolution than classic MLST, as indicated by a larger Simpson's index of diversity, which increases with a higher resolution of clustering (see Data Set S5). With a cutoff value of 100 cgSNPs, 70/74 strains clustered into four major clades (clades 1 to 4), and 4 strains were considered outliers (Fig. 1). These four outliers were at least 103 cgSNPs distant from any other sample. Cluster 1, based on 100 cgSNPs, contained strains that were isolated between 2005 and 2018. These organisms originated from all federal states considered in this study except for Mecklenburg-Western
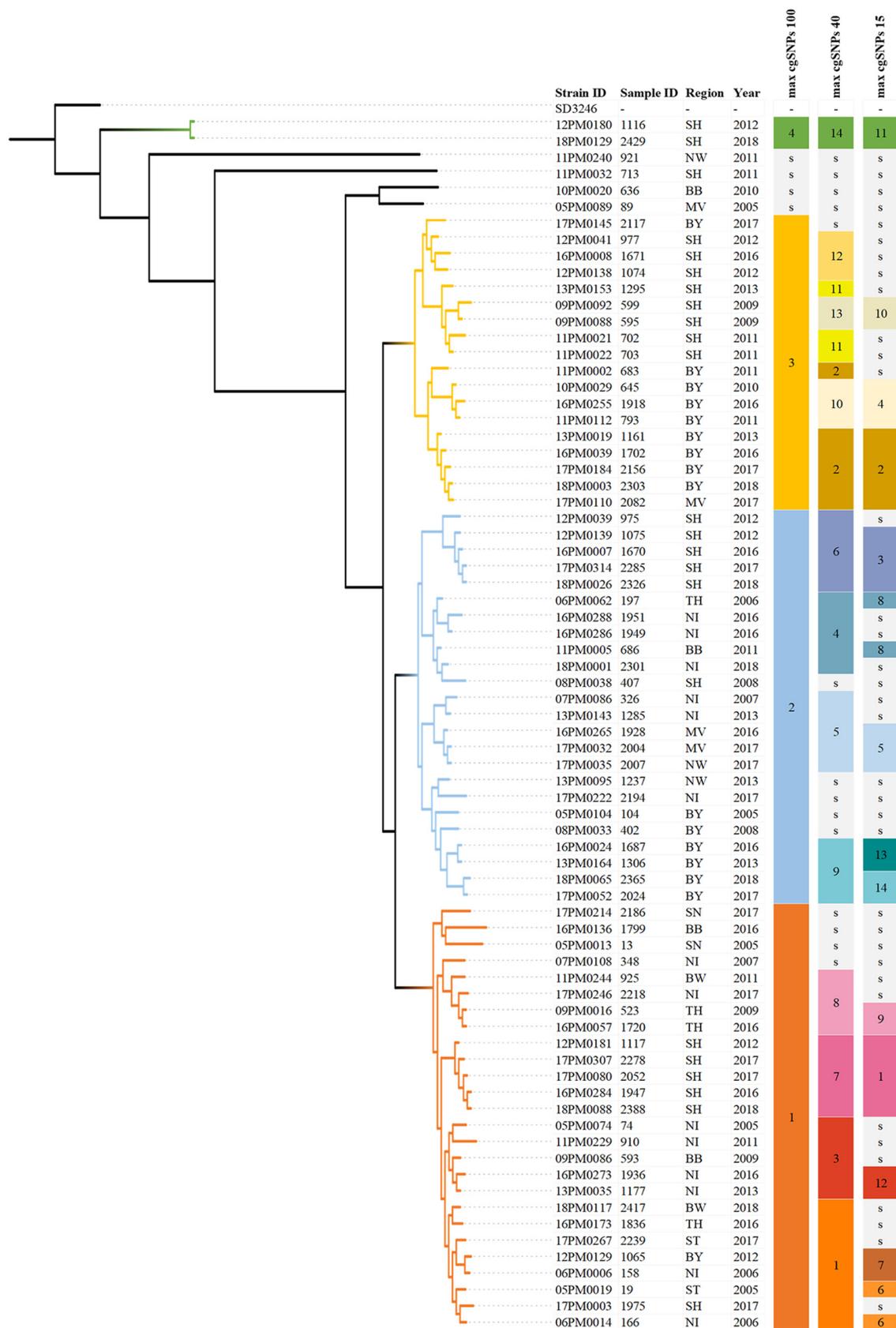
FIG 1 Maximum-likelihood phylogenetic tree based on cgSNP distances for 74 German *Salmonella enterica* subsp. *enterica* serovar Dublin strains. At the right, colored scales show the hierarchical clustering levels using cutoff values of 100 SNPs, 40 SNPs, and 15 SNPs. BB, Brandenburg; BW, Baden-Württemberg; BY, Bavaria; MV, Mecklenburg-Western Pomerania; NI, Lower Saxony; NW, North Rhine-Westphalia; SH, Schleswig-Holstein; SN, Saxony; ST, Saxony-Anhalt; TH, Thuringia.
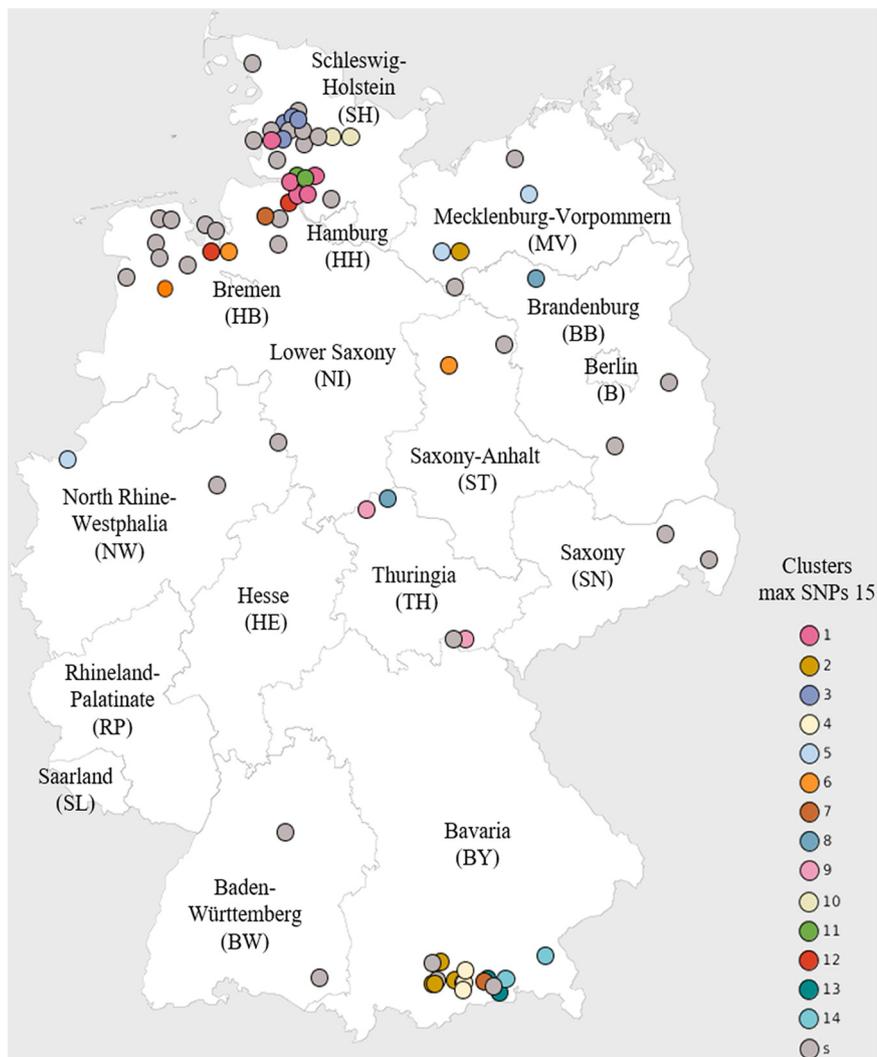
FIG 2 Geographical map of the German federal states, showing the origin of the 74 *Salmonella enterica* subsp. *enterica* serovar Dublin strains used in this study. Each color represents a different cluster according to the hierarchical cgSNP-based clustering with a cutoff value of 15 SNPs.

Pomerania and North Rhine-Westphalia (Fig. 1). Cluster 2 contained 24 samples that also were isolated between 2005 and 2018; they were isolated in Bavaria (6), Schleswig-Holstein (6), Lower Saxony (6), North Rhine-Westphalia (3), Mecklenburg-Western Pomerania (2), and Brandenburg (1) but not in Saxony, Saxony-Anhalt, or Baden-Württemberg. Cluster 3 contained 18 *S.* Dublin strains that were isolated between 2009 and 2018 only in Bavaria (9), Schleswig-Holstein (8), and Mecklenburg-Western Pomerania (1). Cluster 4 contained only 2 identical *S.* Dublin strains, which were isolated in Schleswig-Holstein in 2012 and 2018. Those 2 samples are most distant from any other sample (Fig. 1 and 2). In fact, they are separated from the nearest sample by 254 cgSNPs.

With a maximum distance of 40 cgSNPs, *S.* Dublin strains were grouped into 14 clusters and 14 singletons. Some clusters could be allocated to single federal states (clusters 6, 7, 11, 12, and 13 only in Schleswig-Holstein and clusters 9 and 10 in Bavaria) or were detectable in numerous federal states (clusters 1, 3, 4, 5, and 8) without a more frequent occurrence in certain regions. Except for strain 2082 from Mecklenburg-Western Pomerania, cluster 2 was found only in Bavaria. The 14 singleton strains within the clustering based on 40 cgSNPs originated from very different regions, all together from 7 federal states (Fig. 1).

Using a maximum distance of 15 cgSNPs, clustering of *S.* Dublin strains should give detailed information on their epidemiological links or contacts. Using 15 cgSNPs as the

cutoff value revealed 14 clusters and 36 strains that were considered singletons. Nine clusters (clusters 1, 3, 4, and 9–14) grouped highly similar isolates collected from the same federal state (Fig. 1 and 2). Cluster 1, for example, was detected only in Schleswig-Holstein, over the period from 2012 to 2018. Strains from cluster 4 were found in Bavaria only but also over a period of 6 years. Five clusters (clusters 2 and 5–8) also comprised highly similar isolates; however, strains in these clusters were detected in different federal states. *S.* Dublin strains considered as singletons (36/74 strains) using 15 cgSNPs as the cutoff value occurred both in regions with higher numbers of *S.* Dublin outbreaks (Schleswig-Holstein, Lower Saxony, and Bavaria) and in regions with rarely occurring outbreaks (North Rhine-Westphalia, Saxony-Anhalt, Brandenburg, Thuringia, Saxony, and Baden-Württemberg) (Fig. 1 and 2). Moreover, using a maximum distance of 0 cgSNPs (see Data Set S4), two clusters with two strains each were identified, which were isolated either in different years in Schleswig-Holstein (strains 1947 and 2052) or in different years and different federal states (Mecklenburg-Western Pomerania and North Rhine-Westphalia) (strains 2004 and 2007).

Comparison of *S.* Dublin genomes from Germany, Sweden, and Denmark. The genomes of the 74 German *S.* Dublin strains were compared with raw sequence data from 197 *S.* Dublin isolates from Denmark (13) and 28 strains from Sweden (11). Classic MLST detected ST10 for all isolates from Denmark and Sweden except for 2 isolates from Sweden with unknown ST (see Data Set S6). ST10 was also the most abundant ST (69/74 strains) in the German data set. The cgSNP alignment between the German data set and the data from Denmark and Sweden ($n = 299$) had a length of 1,680 bp and a pairwise average core SNP distance of 107 cgSNPs (range, 0 to 284 cgSNPs), excluding the reference strain. Regarding allele distance, cgMLST analysis revealed an average of 154 alleles (range, 0 to 384 alleles), without considering the reference strain (see Data Set S3). To study the relationship between the German data set and public genomes from Denmark and Sweden, a maximum-likelihood tree and an MST were constructed. Both methods revealed three main clusters, i.e., clusters A, B, and C. Cluster A contained 101 isolates representing the majority (92%) of the German strains but also 6 strains from Denmark and 28 strains from Sweden. The German-derived strains formed their own subcluster within cluster A, in which only 2 isolates (strains 89 and 636) were nearest neighbors to the Danish population and 1 isolate (strain 713) to the Swedish population. In fact, the minimum distance from a German isolate to an isolate not from Germany inside cluster A was 58 cgSNPS when considering strain 89 and strain ERR3491245 from Denmark. Cluster B grouped 42 Danish strains and 2 German isolates (strains 1116 and 2429), which were 58 cgSNPs or 59 cgSNPs away from the closest Danish strains (the clonally related strains ERR3491299 and ERR3491381), respectively. Clade C comprised 76% of the Danish isolates (149/197 strains) and 1 German strain (strain 921) at a distance of at least 141 cgSNPs.

DISCUSSION

In the present study, WGS was used to decipher the genetic traits and phylogeny of the bovine host-adapted serovar *S.* Dublin in Germany. Although the majority of *Salmonella* serovars can be correctly predicted using bioinformatic tools based on WGS data, a certain number of serovars cannot be determined exactly (16, 17). For *S.* Dublin, genoserotyping seems to have no limitations; for all strains analyzed, genoserotyping correlated completely with conventional serotyping. In addition, this study reveals a close agreement between the genotypic prediction of resistance genes and phenotypic antimicrobial susceptibility. Although several clones of different *Salmonella* serovars resistant to fluoroquinolones and third-generation cephalosporins are globally emerging (18), resistance to those compounds has not yet been detected in *S.* Dublin isolates from cattle and calves (19). In this study, 73 of 74 *S.* Dublin strains were not resistant to the antimicrobial substances tested. An MDR pattern consisting of resistance to aminoglycosides, trimethoprim, sulfonamides, tetracyclines, and β-lactamases was detected in only 1 of the *S.* Dublin strains included. Therefore, the occurrence of MDR *S.* Dublin organisms in the German cattle population is uncommon, confirming earlier findings (5, 6) and results from other countries in

Europe (13). In contrast, in the United States, *S.* Dublin has become one of the most important MDR serovars (20, 21).

WGS has been applied to identify the virulence factors that integrate into the *S.* Dublin invasome and that play a role in the pathogenicity of this serovar (22). The *spv* (*Salmonella* plasmid virulence) locus in *S.* Dublin is of special interest because it contributes to enteritis as well as lethal disease in young calves (21, 23). It has been identified in the approximately 80-kb serovar-specific virulence plasmid commonly found in this serovar (18, 24–26). Both the *spv* operon and the serovar-specific virulence plasmid *spv* were detected within the 74 German *S.* Dublin genomes. SPIs are large genetic elements harboring important virulence factors. For example, SPI-1 plays a crucial role in the invasion of host cells and interactions with the host immune system (27). In this study, 13 SPIs were detected within all 74 strains; among them, SPI-1 but also SPI-6 and SPI-9 received special attention, because they encode the T6SS that forms part of the *S.* Dublin invasome (22). The occurrence of the high number of SPIs in all strains analyzed indicates rather low diversity between strains. This is supported by limited sequence variation between *S.* Dublin genomes due to the slow single-gene gain and loss and the low rate of cgSNPs differences (13); thus, reduced diversity of *S.* Dublin genomes over time is suggested (11).

Because the phylogeography of *S.* Dublin in Denmark and Sweden has recently been analyzed using WGS, these data were used to compare the German results with those for neighboring European countries (11, 13, 28). Although there is country-specific phylogeny, the numbers of detected cgSNPs are comparable. Similar to Germany, there are regionally specific clones on one hand and clones spreading all over the country on the other hand (13). Within the analyzed data set, there is no close similarity between German strains and Danish or Swedish isolates. Therefore, exchange of *S.* Dublin strains between Germany and the other two countries has not happened or has happened only occasionally.

In this study, cgSNP typing and cgMLST were used to understand the epidemiological context of *S.* Dublin infections in German cattle herds. Information on genetic characteristics from WGS should be utilized to analyze routes of spreading, transmission, or persistence of *S.* Dublin within and between regions or federal states in Germany over time. However, the use of core genome sequence information for this purpose leads to a number of challenges, especially identification of the required level of relatedness between isolates to characterize a clone (29). WGS-based studies on *S.* Dublin used different cgSNP thresholds to identify closely related strains (11, 13, 28). However, one of the aims of this study was not only identification of clonal organisms within or between single outbreaks but also evaluation of the epidemiological *S.* Dublin situation in Germany over a prolonged period. The special characteristics of *S.* Dublin, the high degree of host adaptation, and the rather low rate at which the serovar mutates at the nucleotide level (30) enable the identification of these epidemiological aspects. Compared to non-host-adapted *Salmonella* serovars occurring in cattle, spreading of bovine host-adapted *S.* Dublin between cattle farms is nearly unanimously associated with the transport or movement of infected animals (30). In the case of other serovars, the first entry of salmonellae into farms may be caused by a large number of biotic and abiotic factors, and thus epidemiological studies on the spreading of those serovars may be much more difficult. Therefore, not only thresholds of 15 cgSNPs but also cutoff values of 40 cgSNPs and 100 cgSNPS were considered to analyze the cattle-derived *S.* Dublin organisms from different regions and different points of time.

In fact, even using 100 cgSNPs two different clusters with strains isolated over the whole study period from 2005 to 2018 from nearly all regions with outbreaks of *S.* Dublin were identified. However, a third cluster contained only strains from two federal states. These findings indicate (i) a rather nationwide occurrence of two general clusters over longer periods and (ii) also a limited spread or persistence of a third cluster with possible exchange of strains between regions in and between federal states. In the case of 40 cgSNPs, clusters became evident showing both (i) the occurrence in one federal state over various periods, indicating a persistence of the infection over time by more related organisms, and (ii) the occurrence in different federal states, indicating

both persistence and probable exchange of the organisms between the federal states. Most interesting in the case of 40 cgSNPs is cluster 1, which includes 8 strains from six federal states; therefore, the question arises of whether any conclusion on spreading of the strains from this cluster is justified. A maximal distance of 15 cgSNPS between *S.* Dublin strains is considered to identify epidemiological links between cattle herds (13, 30). Using this value, several epidemiological aspects not only can be suspected but in fact can be strongly assumed. This refers to the detection of outbreaks in one region in rather short periods (1 year), as well as in very long periods (up to 5 years [e.g., in clusters 2 and 3]). Furthermore, the occurrence of highly related *S.* Dublin strains in very different and distant federal states at different times indicates a possible exchange of the organisms through the movement of animals. These indications by the cgSNP analysis give very valuable data; however, information on the management of the cattle farms and the contacts with associated farms or herds is essential to detect possible routes of infection or persistence of the organism. A number of strains are singletons in the 15-cgSNP clustering, indicating a missing close relationship of these strains to any other isolate in the data set examined. Detailed conclusions regarding routes of transmission, persistence, or contacts between herds for these strains cannot be drawn; however, the information based on 40-cgSNP or 100-cgSNP clustering provides findings on their phylogeography, such as their occurrence in regional clusters.

Moreover, the use of different thresholds for cgSNP analysis in this study enabled a very broad and also detailed view of the occurrence of *S.* Dublin in Germany. The persistence of a rather large cluster in exclusively the same regions in the federal states Schleswig-Holstein and Bavaria over very long time periods clearly indicates a strong persistence and a permanent exchange of the closely related *S.* Dublin organisms within these rather restricted regions. Therefore, control strategies in place seem not to be effective enough to interrupt the existing routes of infection. Apart from the movement of infected animals, attention to prevent the spread of *S.* Dublin in special regions must also be given in the use of common vehicles, the exchange of manure, or the use of shared pastures (31). As also shown in this study, a transfer of persisting *S.* Dublin strains in other, distant regions may occur but obviously not frequently. Final evidence on these transmission routes might be gained by the analysis of both cgSNPs at a cutoff value of ≤15 and data on farming management of the affected cattle herds. Therefore, WGS proved to be an excellent tool for epidemiological studies; however, complex epidemiological analyses require additional information on cattle farming practices. Nonetheless, results from this study also confirm that restriction of movement of *S.* Dublin-infected cattle represents the most effective strategy to prevent regional and nationwide spread of *S.* Dublin.

MATERIALS AND METHODS

Bacterial strains. In this study, we analyzed a data set composed of 74 *S.* Dublin strains covering all federal states in Germany in which this serovar has been detected, over the period between 2005 and 2018. Each strain originated from a single farm with a proven outbreak of salmonellosis in cattle officially confirmed by a competent authority. Because *S.* Dublin is predominantly provable in only a few federal states and not consistently in all German regions, strains from those federal states (e.g., Bavaria, Schleswig-Holstein, and Lower Saxony) contributed in greater numbers to the data set analyzed. Other states (e.g., Thuringia, North Rhine-Westphalia, and Saxony) contributed only a few isolates (Fig. 1 and 2). The annual variation in the number of *S.* Dublin outbreaks in the different regions has also been considered in the selection of the strains examined. The selected *S.* Dublin isolates originated from the collection of bovine host-derived *Salmonella* strains at the NRL for Salmonellosis in Cattle in Germany (see Data Set S7 in the supplemental material). The NRL receives *Salmonella* organisms isolated at cattle farms in Germany with suspected or confirmed outbreaks of salmonellosis from the investigation offices in the federal states for further characterization.

Serotyping and antimicrobial susceptibility testing. All *Salmonella* strains were serotyped using polyvalent and monovalent anti-O and anti-H sera (SIFIN, Germany) according to the Kauffmann-White scheme (32). Antimicrobial susceptibility of the *S.* Dublin strains was assessed by determining the MICs using the broth microdilution method with Sensititre EUVSEC plates (Trek Diagnostic Systems Ltd., East Grinstead, UK). Epidemiological cutoff values according to the European Committee on Antimicrobial Susceptibility Testing (EUCAST) (33) were used. Antimicrobial susceptibilities to sulfamethoxazole, trimethoprim, ciprofloxacin, tetracycline, meropenem, azithromycin, nalidixic acid, cefotaxime, chloramphenicol, tigecycline, ceftazidime, colistin, ampicillin, and gentamicin were examined.

WGS. Genomic DNA of 74 *S.* Dublin strains was prepared using the Qiagen Genomic-tip 20/G kit (Qiagen, Germany) and the genomic DNA buffer set (Qiagen, Germany). The concentration of the DNA

was determined using the Qubit double-stranded DNA (dsDNA) broad-range (BR) assay kit (Invitrogen, USA). Sequencing libraries were prepared using the Nextera XT DNA library preparation kit (Illumina, Inc., USA). Paired-end sequencing of the *S*. Dublin strains was performed with a MiSeq instrument (Illumina) according to the manufacturer's instructions.

Analysis of *S*. Dublin genomes from Sweden and Denmark. To put the findings into context with Germany's neighboring countries, recent studies performed in Denmark (13, 30) and Sweden (11) were selected for analysis. Paired-end Illumina sequencing data (last accessed June 2020) from Sweden (*n* = 28) (see Data Set S7) and Denmark (*n* = 197) (see Data Set S7) were downloaded from the European Nucleotide Archive and the NCBI Sequence Read Archive (SRA).

Data quality control and *in silico* serotyping in bioinformatic analyses. Raw paired-end reads from Illumina sequencing were used as input for the Linux-based bioinformatics pipeline WGSBAC v. 2.0.0 (34) to perform the data analysis, as recently published (14). First, WGSBAC uses FastQC v. 0.11.7 (35) to perform read quality control. The raw coverage, defined as the number of reads by their average read length divided by the genome size of the reference genome, is estimated. WGSBAC performs assembly using Shovill v. 1.0.4 (36) implementation of the SPAdes assembler (37). To check the quality of the assembled genomes, WGSBAC uses QUAST v. 5.0.2 (38). To classify sequences, and thus to identify potential contamination in both reads and assemblies, the pipeline uses Kraken 2 v. 1.1 (39) and the database Kraken2DB. For *in silico* serotyping, WGSBAC relies on SISTR v. 1.0.2 (16) for serovar prediction using the assembled genomes and SeqSero, which works using genome assemblies as well as raw sequence reads (40).

In silico AMR genes and virulence determinants. For detection of AMR genes and virulence determinants, WGSBAC uses the software ABRicate v. 0.8.10 (41) and the databases ResFinder (42), NCBI (43), Virulence Factor Database (VFDB) (44), and PlasmidFinder (45). In addition, WGSBAC uses AMRFinderPlus v. 3.6.10 (46) for the detection of chromosomal point mutations leading to AMR and organism-specific acquired resistance genes. For the identification of SPIs, sequences for 22 SPIs were downloaded from the public data repositories Pathogenicity Island Database (PAIDB) (47) and NCBI (48) and used to create a customized database for ABRicate (https://gitlab.com/FLI_Bioinfo_pub/spis_ibiz_database). A cutoff value of 60% coverage was applied to consider a strain positive for a SPI. In addition to VFDB, customized ABRicate databases were used to detect *S*. Dublin-specific virulence determinants that are part of the invasome, i.e., (i) the *spv* (*Salmonella* plasmid virulence) operon (including *spvRABCD* genes), (ii) the gene *st313-td* for flagellum biosynthesis, and (iii) putative virulence determinants such as the T6SS, the lambdoid Gifsy-2 prophage, and the *fae* cluster (coding for fimbria biosynthesis) (22). BLAST was used to detect the serovar-specific virulence plasmid pSDV (GenBank accession no. DQ115388.2).

Genotyping, phylogenetic analysis, and clustering. For genotyping, WGSBAC uses classic MLST of assembled genomes using the software mlst v. 2.16.1 (49), which incorporates the PubMLST database for the 7-gene *Salmonella* MLST scheme (https://pubmlst.org/salmonella) (50). cgMLST was performed using the external software Ridom SeqSphere+ v. 5.1.0 (51) with default settings, including definition of good cgMLST targets as follows: (i) same length as reference genes plus or minus three triplets, (ii) no ambiguities, and (iii) no frameshifts. The specific core genome scheme (cgMLST v. 2) for *Salmonella enterica* with 3,002 target loci developed by EnteroBase (52) was used within Ridom SeqSphere. To infer phylogeny based on cgMLST results, an MST and a geographical map were depicted and visualized using the external software Ridom SeqSphere+ v. 7.1.0 (https://www.ridom.de/seqsphere). Furthermore, WGSBAC performs mapping-based genotyping using cgSNPs identified by Snippy v. 4.3.6 (53) with standard settings and as reference the complete genome of *S*. Dublin strain 3246 (GenBank accession no. CM001151). For phylogenetic tree construction based on cgSNP analysis, WGSBAC uses the SNP alignment matrix generated by Snippy and reconstructs the tree using RAxML v. 8 (54). The tree was rooted to the reference genome and visualized using the interactive Tree of Life (iTOL) v. 4 web tool (55) (https://itol.embl.de). The pairwise SNP distance between isolates was calculated using the tool snps-dists v 0.63. For hierarchical clustering based on the cgSNP alignment, WGSBAC computes classic hierarchical clustering based on the SNP distance matrix from Snippy using the hierClust function v. 5.1 of R. For clustering based on cgMLST results, cluster calculation based on allele distance was performed by Ridom SeqSphere+, and the same cutoff values as for cgSNPs were applied. As previously applied (13), a cutoff value of 15 cgSNPs was applied to identify closely related strains. In addition, a cutoff value of 100 cgSNPs was chosen to identify more general clusters and 40 cgSNPs for intermediate clusters. Simpson¨s index was calculated using the web-based tool Comparing Partitions (http://www.comparingpartitions.info/index.php?link=Home).

Data availability. Raw sequencing data were deposited in the NCBI SRA under the BioProject accession number PRJNA678856.

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.
SUPPLEMENTAL FILE 1, XLSX file, 0.02 MB.
SUPPLEMENTAL FILE 2, XLSX file, 0.01 MB.
SUPPLEMENTAL FILE 3, XLSX file, 0.9 MB.
SUPPLEMENTAL FILE 4, XLSX file, 0.01 MB.
SUPPLEMENTAL FILE 5, XLSX file, 0.01 MB.
SUPPLEMENTAL FILE 6, XLSX file, 0.02 MB.
SUPPLEMENTAL FILE 7, XLSX file, 0.03 MB.

ACKNOWLEDGMENTS

We sincerely thank A. Hackbart and S. Keiling for their excellent technical assistance.

We are most grateful to all of the colleagues of the Veterinary Investigation Offices in the German federal states who provided *Salmonella* strains.

REFERENCES

1. Uzzau S, Brown DJ, Wallis T, Rubino S, Leori G, Bernard S, Casadesus J, Platt DJ, Olsen JE. 2000. Host adapted serotypes of *Salmonella enterica*. Epidemiol Infect 125:229–255. https://doi.org/10.1017/s0950268899004379.
2. La Ragione R, Metcalfe HJ, Villarreal- Ramos B, Werling D. 2013. *Salmonella* infections in cattle, p 233–262. *In* Barrow PA, Methner U (ed), Salmonella in domestic animals, 2nd ed. CAB International, Wallingford, England.
3. Harvey RR, Friedman CR, Crim SM, Judd M, Barrett KA, Tolar B, Folster JP, Griffin PM, Brown AC. 2017. Epidemiology of *Salmonella enterica* serotype Dublin infections among humans, United States, 1968–2013. Emerg Infect Dis 23:1493–1501. https://doi.org/10.3201/eid2309.170136.
4. Verordnung zum Schutz gegen die Salmonellose der Rinder (Rinder-Salmonellose-Verordnung) vom 14. November 1991 (BGBl. I S. 2118). https://www.gesetze-im-internet.de/rindsalmv/RindSalmV.pdf.
5. Methner U. 2005. Situation of bovine salmonellosis in Germany from 1995–2003 according to the data from the National Animal Disease Reporting System. Berl Munch Tierarztl Wochenschr 118:449–455.
6. Methner U. 2019. Salmonellose der Rinder. *In* Tiergesundheitsjahresberichte 2019. Friedrich-Loeffler-Institute, Greifswald, Germany. www.fli.de/de/publikationen/tiergesundheitsjahresberichte.
7. Methner U, Merbach S, Peters M. 2018. *Salmonella enterica* subspecies *enterica* serovar choleraesuis in a German wild boar population: occurrence and characterisation. Acta Vet Scand 60:65. https://doi.org/10.1186/s13028-018-0422-4.
8. Methner U, Moog U. 2018. Occurrence and characterisation of *Salmonella enterica* subspecies *diarizonae* serovar 61: k: 1, 5, (7) in sheep in the federal state of Thuringia, Germany. BMC Vet Res 14:401. https://doi.org/10.1186/s12917-018-1741-4.
9. Taylor AJ, Lappi V, Wolfgang WJ, Lapierre P, Palumbo MJ, Medus C, Boxrud D. 2015. Characterization of foodborne outbreaks of *Salmonella enterica* serovar Enteritidis with whole-genome sequencing single nucleotide polymorphism-based analysis for surveillance and outbreak detection. J Clin Microbiol 53:3334–3340. https://doi.org/10.1128/JCM.01280-15.
10. Olsen JE, Skov M. 1994. Genomic lineage of *Salmonella enterica* serovar Dublin. Vet Microbiol 40:271–282. https://doi.org/10.1016/0378-1135(94)90116-3.
11. Agren EC, Wahlstrom H, Vesterlund-Carlson C, Lahti E, Melin L, Soderlund R. 2016. Comparison of whole genome sequencing typing results and epidemiological contact information from outbreaks of *Salmonella* Dublin in Swedish cattle herds. Infect Ecol Epidemiol 6:31782. https://doi.org/10.3402/iee.v6.31782.
12. den Bakker HC, Allard MW, Bopp D, Brown EW, Fontana J, Iqbal Z, Kinney A, Limberger R, Musser KA, Shudt M, Strain E, Wiedmann M, Wolfgang WJ. 2014. Rapid whole-genome sequencing for surveillance of *Salmonella enterica* serovar Enteritidis. Emerg Infect Dis 20:1306–1314. https://doi.org/10.3201/eid2008.131399.
13. Kudirkiene E, Sorensen G, Torpdahl M, de Knegt LV, Nielsen LR, Rattenborg E, Ahmed S, Olsen JE. 2020. Epidemiology of *Salmonella enterica* serovar Dublin in cattle and humans in Denmark, 1996 to 2016: a retrospective whole-genome-based study. Appl Environ Microbiol 86:e01894-19. https://doi.org/10.1128/AEM.01894-19.
14. Garcia-Soto S, Abdel-Glil MY, Tomaso H, Linde J, Methner U. 2020. Emergence of multidrug-resistant *Salmonella enterica* subspecies *enterica* serovar Infantis of multilocus sequence type 2283 in German broiler farms. Front Microbiol 11:1741. https://doi.org/10.3389/fmicb.2020.01741.
15. Quainoo S, Coolen JPM, van Hijum S, Huynen MA, Melchers WJG, van Schaik W, Wertheim HFL. 2017. Whole-genome sequencing of bacterial pathogens: the future of nosocomial outbreak analysis. Clin Microbiol Rev 30:1015–1063. https://doi.org/10.1128/CMR.00016-17.
16. Yoshida CE, Kruczkiewicz P, Laing CR, Lingohr EJ, Gannon VP, Nash JH, Taboada EN. 2016. The *Salmonella In Silico* Typing Resource (SISTR): an open Web-accessible tool for rapidly typing and subtyping draft *Salmonella* genome assemblies. PLoS One 11:e0147101. https://doi.org/10.1371/journal.pone.0147101.
17. Zhang S, den Bakker HC, Li S, Chen J, Dinsmore BA, Lane C, Lauer AC, Fields PI, Deng X. 2019. SeqSero2: rapid and improved *Salmonella* serotype determination using whole-genome sequencing data. Appl Environ Microbiol 85:e01746-19. https://doi.org/10.1128/AEM.01746-19.
18. Jajere SM. 2019. A review of *Salmonella enterica* with particular focus on the pathogenicity and virulence factors, host specificity and antimicrobial resistance including multidrug resistance. Vet World 12:504–521. https://doi.org/10.14202/vetworld.2019.504-521.
19. European Food Safety Authority. 2020. The European Union Summary Report on antimicrobial resistance in zoonotic and indicator bacteria from humans, animals and food in 2017/2018. EFSA J 18:6007. https://doi.org/10.2903/j.efsa.2020.6007.
20. Eyler AB, M'Ikanatha NM, Xiaoli L, Dudley EG. 2020. Whole-genome sequencing reveals resistome of highly drug-resistant retail meat and human *Salmonella* Dublin. Zoonoses Public Health 67:251–262. https://doi.org/10.1111/zph.12680.
21. Hsu CH, Li C, Hoffmann M, McDermott P, Abbott J, Ayers S, Tyson GH, Tate H, Yao K, Allard M, Zhao S. 2019. Comparative genomic analysis of virulence, antimicrobial resistance, and plasmid profiles of *Salmonella* Dublin isolated from sick cattle, retail beef, and humans in the United States. Microb Drug Resist 25:1238–1249. https://doi.org/10.1089/mdr.2019.0045.
22. Mohammed M, Le Hello S, Leekitcharoenphon P, Hendriksen R. 2017. The invasome of *Salmonella* Dublin as revealed by whole genome sequencing. BMC Infect Dis 17:544. https://doi.org/10.1186/s12879-017-2628-x.
23. Libby SJ, Adams LG, Ficht TA, Allen C, Whitford HA, Buchmeier NA, Bossie S, Guiney DG. 1997. The *spv* genes on the *Salmonella* Dublin virulence plasmid are required for severe enteritis and systemic infection in the natural host. Infect Immun 65:1786–1792. https://doi.org/10.1128/iai.65.5.1786-1792.1997.
24. Guiney DG, Fang FC, Krause M, Libby S. 1994. Plasmid-mediated virulence genes in non-typhoid *Salmonella* serovars. FEMS Microbiol Lett 124:1–9. https://doi.org/10.1111/j.1574-6968.1994.tb07253.x.
25. Amavisit P, Lightfoot D, Browning GF, Markham PF. 2003. Variation between pathogenic serovars within *Salmonella* pathogenicity islands. J Bacteriol 185:3624–3635. https://doi.org/10.1128/JB.185.12.3624-3635.2003.
26. Chu C, Feng Y, Chien AC, Hu S, Chu CH, Chiu CH. 2008. Evolution of genes on the *Salmonella* virulence plasmid phylogeny revealed from sequencing of the virulence plasmids of *S. enterica* serotype Dublin and comparative analysis. Genomics 92:339–343. https://doi.org/10.1016/j.ygeno.2008.07.010.
27. Lou L, Zhang P, Piao R, Wang Y. 2019. *Salmonella* pathogenicity island 1 (SPI-1) and its complex regulatory network. Front Cell Infect Microbiol 9:270. https://doi.org/10.3389/fcimb.2019.00270.
28. Ung A, Baidjoe AY, Van Cauteren D, Fawal N, Fabre L, Guerrisi C, Danis K, Morand A, Donguy MP, Lucas E, Rossignol L, Lefevre S, Vignaud ML, Cadel-Six S, Lailler R, Jourdan-Da Silva N, Le Hello S. 2019. Disentangling a complex nationwide *Salmonella* Dublin outbreak associated with raw-milk cheese consumption, France, 2015 to 2016. Euro Surveill 24:1700703. https://doi.org/10.2807/1560-7917.ES.2019.24.3.1700703.
29. Tang S, Orsi RH, Luo H, Ge C, Zhang G, Baker RC, Stevenson A, Wiedmann M. 2019. Assessment and comparison of molecular subtyping and characterization methods for *Salmonella*. Front Microbiol 10:1591. https://doi.org/10.3389/fmicb.2019.01591.
30. de Knegt LV, Kudirkiene E, Rattenborg E, Sorensen G, Denwood MJ, Olsen JE, Nielsen LR. 2020. Combining *Salmonella* Dublin genome information and contact-tracing to substantiate a new approach for improved detection of infectious transmission routes in cattle populations. Prev Vet Med 181:104531. https://doi.org/10.1016/j.prevetmed.2018.09.005.
31. Methner U. 2012. Salmonellose der Rinder: Empfehlungen zur Vorgehensweise nach Feststellung eines Ausbruchs. Amtstierärztlicher Dienst Lebensmittelkontrolle 19:1–4.
32. Grimont PAD, Weill F-X. 2007. Antigenic formulae of the Salmonella serovars, 9th ed. Institut Pasteur, Paris, France.
33. EUCAST. 2018. Breakpoint tables for interpretation of MICs and zone diameters, version 8.0. https://www.eucast.org/fileadmin/src/media/PDFs/EUCAST_files/Breakpoint_tables/v_8.0_Breakpoint_Tables.pdf.

34. FLI_Bioinfo. 2020. WGSBAC. Modules for genotyping and characterization of bacterial isolates using whole-genome-sequencing data, v2.0.0. https://gitlab.com/FLI_Bioinfo/WGSBAC.

35. Andrews S. 2018. FastQC: a quality control tool for high throughput sequence data. v0.11.5. https://www.bioinformatics.babraham.ac.uk/projects/fastqc.

36. Seemann T. 2018. Shovill. https://github.com/tseemann/shovill.

37. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 19:455–477. https://doi.org/10.1089/cmb.2012.0021.

38. Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUAST: quality assessment tool for genome assemblies. Bioinformatics 29:1072–1075. https://doi.org/10.1093/bioinformatics/btt086.

39. Wood DE, Lu J, Langmead B. 2019. Improved metagenomic analysis with Kraken 2. Genome Biol 20:257. https://doi.org/10.1186/s13059-019-1891-0.

40. Zhang S, Yin Y, Jones MB, Zhang Z, Deatherage Kaiser BL, Dinsmore BA, Fitzgerald C, Fields PI, Deng X. 2015. *Salmonella* serotype determination utilizing high-throughput genome sequencing data. J Clin Microbiol 53:1685–1692. https://doi.org/10.1128/JCM.00323-15.

41. Seemann T. 2015. Abricate. https://github.com/tseemann/abricate.

42. 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. https://doi.org/10.1093/jac/dks261.

43. Feldgarden M, Brover V, Haft DH, Prasad AB, Slotta DJ, Tolstoy I, Tyson GH, Zhao S, Hsu C-H, McDermott PF, Tadesse DA, Morales C, Simmons M, Tillman G, Wasilenko J, Folster JP, Klimke W. 2019. Using the NCBI AMR-Finder tool to determine antimicrobial resistance genotype-phenotype correlations within a collection of NARMS isolates. bioRxiv. https://doi.org/10.1101/550707.

44. Chen L, Yang J, Yu J, Yao Z, Sun L, Shen Y, Jin Q. 2005. VFDB: a reference database for bacterial virulence factors. Nucleic Acids Res 33:D325–D328. https://doi.org/10.1093/nar/gki008.

45. Carattoli A, Zankari E, Garcia-Fernandez A, Voldby Larsen M, Lund O, Villa L, Moller Aarestrup F, Hasman H. 2014. In silico detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing. Antimicrob Agents Chemother 58:3895–3903. https://doi.org/10.1128/AAC.02412-14.

46. Feldgarden M, Brover V, Haft DH, Prasad AB, Slotta DJ, Tolstoy I, Tyson GH, Zhao S, Hsu CH, McDermott PF, Tadesse DA, Morales C, Simmons M, Tillman G, Wasilenko J, Folster JP, Klimke W. 2019. Validating the AMRFinder tool and Resistance Gene Database by using antimicrobial resistance genotype-phenotype correlations in a collection of isolates. Antimicrob Agents Chemother 63:e00483-19. https://doi.org/10.1128/AAC.00483-19.

47. Yoon SH, Park YK, Kim JF. 2015. PAIDB v2.0: exploration and analysis of pathogenicity and resistance islands. Nucleic Acids Res 43:D624–D630. https://doi.org/10.1093/nar/gku985.

48. Sayers EW, Beck J, Brister JR, Bolton EE, Canese K, Comeau DC, Funk K, Ketter A, Kim S, Kimchi A, Kitts PA, Kuznetsov A, Lathrop S, Lu Z, McGarvey K, Madden TL, Murphy TD, O'Leary N, Phan L, Schneider VA, Thibaud-Nissen F, Trawick BW, Pruitt KD, Ostell J. 2020. Database resources of the National Center for Biotechnology Information. Nucleic Acids Res 48:D9–D16. https://doi.org/10.1093/nar/gkz899.

49. Seemann T. 2014. mlst. https://github.com/tseemann/mlst.

50. Jolley KA, Bray JE, Maiden MCJ. 2018. Open-access bacterial population genomics: BIGSdb software, the PubMLST.org website and their applications. Wellcome Open Res 3:124. https://doi.org/10.12688/wellcomeopenres.14826.1.

51. Junemann S, Sedlazeck FJ, Prior K, Albersmeier A, John U, Kalinowski J, Mellmann A, Goesmann A, von Haeseler A, Stoye J, Harmsen D. 2013. Updating benchtop sequencing performance comparison. Nat Biotechnol 31:294–296. https://doi.org/10.1038/nbt.2522.

52. Alikhan NF, Zhou Z, Sergeant MJ, Achtman M. 2018. A genomic overview of the population structure of *Salmonella*. PLoS Genet 14:e1007261. https://doi.org/10.1371/journal.pgen.1007261.

53. Seemann T. 2014. Snippy. https://github.com/tseemann/snippy.

54. Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30:1312–1313. https://doi.org/10.1093/bioinformatics/btu033.

55. Letunic I, Bork P. 2019. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. Nucleic Acids Res 47:W256–W259. https://doi.org/10.1093/nar/gkz239.