Comparative Genome Analysis of Hungarian and Global Strains of *Salmonella* Infantis

Tibor Nagy[1†], Ama Szmolka[2†], Tímea Wilk[1], János Kiss[1], Mónika Szabó[1], Judit Pászti[3], Béla Nagy[2] and Ferenc Olasz[1]*

[1] National Agricultural Research and Innovation Centre, Agricultural Biotechnology Institute, Gödöllő, Hungary, [2] Centre for Agricultural Research, Institute for Veterinary Medical Research, Budapest, Hungary, [3] National Center for Epidemiology, Budapest, Hungary

Introduction: The emergence and spread of new strains of zoonotic bacteria, such as multidrug resistant (MDR) *Salmonella* Infantis, represent a growing health risk for humans in and outside Europe due to foodborne infections of poultry meat origin.

Objectives: In order to understand genome relations of *S*. Infantis strains from Hungary and from different geographic regions, we performed a comprehensive genome analysis of nine Hungarian and 67 globally selected strains of *S*. Infantis and 26 *Salmonella* strains representing 13 non-Infantis serovars.

Results: Analyses of whole-, and accessory genomes, showed that almost all *S*. Infantis strains were separated from the non-Infantis serovars. *S*. Infantis strains from Hungary formed subclusters based on their time of isolation. In whole genome sequence analysis, the Swiss strains of *S*. Infantis were closely related to each other and clustered together with subclusters of strains from Hungary, Japan, Italy, United States, and Israel. The accessory genome analysis revealed that the Swiss strains were distinct from most of the strains investigated, including the Hungarian ones. Analysis of the cloud genes offered the most detailed insight into the genetic distance and relationship of *S*. Infantis strains confirming that the Swiss and Hungarian strains belonged to different lineages. As expected, core genome analysis provided the least discriminatory power for analysis of *S*. Infantis. Genomic sequences of nine strains from Brazil, Israel, Mexico, Nigeria, and Senegal (deposited as *S*. Infantis) proved to be outliers from the *S*. Infantis clade. They were predicted to be *Salmonella* Rissen, *Salmonella* Ouakarm, *Salmonella* Kentucky, *Salmonella* Thompson, and *Salmonella* enterica subsp. *diarizonae*.

Conclusion: Accessory genome of *S*. Infantis showed the highest diversity suggesting a faster evolution than that of the whole genomes contributing to the emergence of multiple genetic variants of *S*. Infantis worldwide. Accordingly, in spite of the comprehensive analysis of several genomic characteristics, no epidemiologic links between these *S*. Infantis strains from different countries could be established. It is also concluded that several strains originally designated as *S*. Infantis need in databanks reclassification.

Keywords: *Salmonella* Infantis, whole-, core-, and accessory genomes, antibiotic resistance genotypes, genomic-, and pathogenicity islands, outlier *S*. Infantis isolates

INTRODUCTION

For several decades, non-typhoidal salmonellosis has become the first or second most frequent foodborne zoonosis worldwide with *Salmonella* Enteritidis as the leading serovar, followed by serovars Typhimurium (especially of monophasic *Salmonella* Typhimurium) and *Salmonella* Infantis, which have increased during the last decade in humans. The primary sources of *S.* Enteritidis infections are layers and layed eggs while those of *S.* Infantis infections are broiler chicken (European Food Safety Authority [EFSA] and European Centre for Disease Prevention [ECDC], 2017a,b, 2019). In spite of the implementation of *Salmonella* monitoring and control programs in the EU and in contrast to the earlier declining trend of human salmonellosis (O'Brien, 2013), there have been some unfavorable changes on this trend since 2016 in Europe. Especially worrying is that the prevalence of *S.* Infantis infections has almost doubled among humans in the last few years. About 90% of such human cases were of broiler origin. Furthermore, recent isolates of *S.* Infantis show increasing prevalence and diversity of antibiotic resistance (European Food Safety Authority [EFSA] and European Centre for Disease Prevention [ECDC], 2017a,b, 2019). As a consequence, we are witnessing an increased public awareness and increasing health concerns due to *S.* Infantis infection of broilers in Europe.

Salmonella Infantis appears to be an emerging serovar worldwide as well. It has been reported as the most frequently isolated serovar from fresh poultry meat and broiler flocks as well as from human in Japan (Asai et al., 2006; Shahada et al., 2006; Murakami et al., 2007), Belgium, France (Cloeckaert et al., 2007), and Hungary (Nógrády et al., 2007, 2008). *S.* Infantis has also emerged in Israel and became the leading serovar in humans and broilers (Gal-Mor et al., 2010). Similar trend has been reported in Italy (Dionisi et al., 2011). The first reports about the emergence and clonal spread of multidrug resistant (MDR) *S.* Infantis in human populations and in broilers in several European countries came from Nógrády et al. (2007, 2008). The emerging "clone B" strains, as determined by pulsed field gel electrophoresis (PFGE), were predominantly resistant for nalidixic acid, streptomycin, sulfonamides, and tetracycline (NSSuT) and carried a conjugative ~270 kb MDR plasmid pSI54/04 containing a class 1 integron with *aad*A1 gene cassette and the *tet*(A) gene (Nógrády et al., 2007, 2008, 2012; Olasz et al., 2015; Szmolka et al., 2018). Sequence analysis and PCR-typing of plasmid pSI54/04 revealed a high sequence similarity to the megaplasmid pESI of human *S.* Infantis from Israel (Aviv et al., 2014). However, both *in vivo* and *in vitro* evidence for the pathogenic significance of this plasmid in chicks is missing (Szmolka et al., 2018). In order to obtain underlying genomic data for subsequent analysis of the relations of *S.* Infantis clones in Hungary, we have performed whole genome sequencing (WGS) projects (Olasz et al., 2015; Wilk et al., 2016, 2017).

Bioinformatic analyses of sequence data from our projects and those available in different databases provided a promising approach to obtain new information and in-depth insights into *Salmonella* genomics. In case of *S.* Infantis, the amount of relevant data available in public databases and genomic studies increased significantly in the past few years (Yokoyama et al., 2014; Franco et al., 2015; Hindermann et al., 2017; Acar et al., 2019). However, the type, genome status, and quality of these datasets are usually diverse depending on the intention of the depositors. Sequence data deposited in the NCBI GenBank include reads or contigs, scaffolds, or even whole genomes and contain relevant information on the completeness of the sequence, which was required to select suitable data set for our analyses.

So far genomic studies on *S.* Infantis have primarily dealt with sequences from national collections and studied genomic relations between strains of the respective countries, i.e., Japan, Italy, and Switzerland (Yokoyama et al., 2014; Franco et al., 2015; Hindermann et al., 2017; Acar et al., 2019). In order to understand genomic relations of *S.* Infantis strains from different geographic origin, here we performed a comprehensive genome analysis of nine Hungarian and further 67 globally selected *S.* Infantis and 26 other *Salmonella* strains representing 13 non-Infantis serovars. Comparisons of the whole-, core-, accessory genomes allowed us to obtain a more comprehensive view about possible genomic relations of *S.* Infantis from different countries and geographic regions.

MATERIALS AND METHODS

Selection and Origin of *Salmonella* Genome Sequences

For the analyses, whole genome sequence of 76 *S.* Infantis strains was selected. The isolates were selected on the basis of already published papers. The concept of the choice of the isolates was as follows: (i) Isolates from a single country were preferred if published in a single publication (see **Supplementary Table S1**). (ii) Each country was represented by a maximum 10 strains of poultry (food) or human (clinical) origin if possible. (iii) Further criterion for the selection was the time of publications (between 2014 and 2017 with the exception of Senegal SARB27, which was published earlier; Boyd et al., 1993), representing recent isolates. (iv) Beyond that, the choice of the sequences of the isolates was random.

Among the 76 *S.* Infantis strains, nine strains derived from Hungary and the other 67 strains were selected from diverse European and non-European countries. Hungarian strains covered two periods of isolation, and were designated here as pre-emerging (1980–1994) and emerging (2000–2016) strains. The latter group representing the emergence of MDR *S.* Infantis in Hungary (Nógrády et al., 2007; Olasz et al., 2015; Wilk et al., 2016, 2017). The annotated genome of the earliest sequenced *S.* Infantis strain UK-1973 1326/28 was included as reference.

In addition to *S.* Infantis, genome sequences of 26 *Salmonella* strains representing 13 most common *Salmonella* serovars in animals and humans based on the paper of Zou et al. (2013) were also included into the analyzes. Each of these serovars was represented by two strains, except Paratyphi, where three strains were included for the biovars A, B, and C (**Supplementary Table S1**). The sequence of *Salmonella bongori* strain NCTC12419 was used as a non-*enterica* control

(Fookes et al., 2011). In this way, a total of 102 *Salmonella* genomes were downloaded from the NCBI database (Agarwala et al., 2016). Raw sequences deposited in SRA database (Leinonen et al., 2011) were assembled into contigs using the SPAdes v.3.11.1 program (Bankevich et al., 2012) with default settings. Contigs were annotated using Prokka v1.14.1 with default settings (Seemann, 2014).

Bioinformatics and Phylogenetic Analysis

Core- and accessory genome analyses were performed with Roary (Page et al., 2015). The input files were produced by Prokka and the core gene alignment method was set to MAFFT. IslandViewer4 (Bertelli et al., 2017) was used to identify genomic islands (GIs) in the reference strain UK-1973 1326/28. *Salmonella* pathogenicity islands (SPIs) were also defined according to this reference strain.

Blast + (Camacho et al., 2009) was used for identification of the GI, SPI genes, as well as of core-, accessory-, and flagellin genes. The identification criteria for GIs were the following: the region should be covered 90% and the similarity should be larger than 90%. ProgressiveMauve with default settings was used to perform multiple alignments with default settings (Darling et al., 2010). Dendroscope 3.6.3 (Huson et al., 2007) was used to draw the final trees from Mauve's guided tree with rectangular phylogram option. The whole genome analysis—including the custom Python scripts is available at this link: https://github.com/TravisCG/articles/tree/master/infantis2020/. The serovar-marker flagellin gene *flj*B of *S.* Infantis was identified according to Kardos et al. (2007). The flagellin genes *flj*A, *flj*B, *fli*A, *fli*B, *fli*C, *fli*D, *fli*S, and *hin* were identified based on sequences of the reference strain (Liu et al., 2017). For *in silico* prediction of the serovar and clonal identity of the strains, the web-based softwares SeqSero and MLST were used (Larsen et al., 2012; Zhang et al., 2015). The acquired antibiotic resistance genes of *S.* Infantis were identified by using the ResFinder v3.2 (Zankari et al., 2012). Selected thresholds for the prediction of resistance genes were: identity 90%; minimum length 80%. The assembled contigs were used as input sequence for the *in silico* prediction programs.

RESULTS

Analysis of Whole Genome Sequences

In order to obtain detailed information on relationships of the 102 selected *Salmonella* strains (**Supplementary Table S1**), whole genome sequences were compared, including the two *S. enterica* subsp. *arizonae* and the *S. bongori* strains as internal subspecies- and species controls. The genomic reconstruction showed that the nine Hungarian *S.* Infantis strains grouped in three subclusters according to their time of isolation (**Figure 1**). One cluster contained two subclusters representing six recent (emerging) strains (2004–2016), while the other included three pre-emerging strains isolated between 1980 and 1994. This observation was in good agreement with our results obtained by PFGE (**Supplementary Figure S1**). In general, some geographic clustering of *S.* Infantis strains was recognized. The genomes of

nine Swiss strains from 2010 to 2013 proved to be closely related to each other and formed a separate subcluster and clustered together with subclusters of strains from Hungary, Japan, Italy, United States, and Israel.

Salmonella Infantis strains formed a large well-defined clade clearly separated from the non-Infantis serovars in the whole genome analysis, with the exception of nine strains. Out of these, the strains Mexico-2008-4 and -5 formed a divergent outlier cluster, which was a sister group of *S. enterica* subsp. *arizonae*. Six other outlier *S.* Infantis, Sal280[2012], Sal147[2011], Senegal SARB27, Brasil-2013 LMP-ST01, and Israel-2014 FDA00004302 grouped together with *Salmonella* Agona, while Mexico-2008-CFSAN047352 was most closely related to *Salmonella* ParatyphiC_RKS4594. The strain Nigeria-2009 BCW_2699 showed moderate similarity with the *S.* Infantis clade. The non-Infantis strains clustered separately in pairs according to their serovars (**Figure 1**).

Analysis of the Core- and Accessory Genomes

Genomes were analyzed according to the genomic units as determined by Inglin et al. (2018). Core genes, representing genes present in all selected genomes, are generally responsible for fundamental functions such as metabolism, basic regulation of cell cycle, maintenance of the genome, or cell envelope. Softcore genes are present in 95–99% of the strains. The accessory genome includes genes that are present occasionally in several strains. These are often responsible for strain specific traits and can be divided into the classes of shell genes that are present in a remarkable proportion of the strains (15–95%) and cloud genes that are unique or specific to a smaller fraction of strains (<15%) (**Table 1**).

The pan-genome of the 102 *Salmonella* strains contained 18,847 genes, while 2030 genes were identified as core genes (**Table 1** and **Supplementary Table S2**). In the 76 *S.* Infantis strains investigated, the number of core genes was a bit higher (2316) and the number of cloud genes was lower (8175). In **Table 1**, the number and distribution of the core, soft core, shell, and cloud genes in *S.* Infantis strains and their percentual overlap from different countries (Hungary, Switzerland, United States, Japan, and Italy) were also compared. The number of core genes in the studied countries is very similar (3814–4139) and no major geographical differences were observed. The number of soft core genes is low (0–238), while the number of shell and cloud genes is more variable (shell genes: 438–1014; cloud genes: 684–1219).

Core genome-based phylogenetic reconstruction of these 102 *Salmonella* strains showed that a large set of *S.* Infantis strains clustered separately from the non-Infantis strains without indication of *S.* Infantis subclusters (**Figure 2**). The nine outliers of the Infantis clade appeared in different branches of the core-genome tree together with the non-Infantis strains. The separate core genome analysis of *S.* Infantis strains did not show significant diversity of these strains either, except for the outlier strains as pointed out above (**Supplementary Figure S2**).

Cloud gene tree of the 102 *Salmonella* genomes reflected the separation of *S.* Infantis and non-Infantis strains in a way that
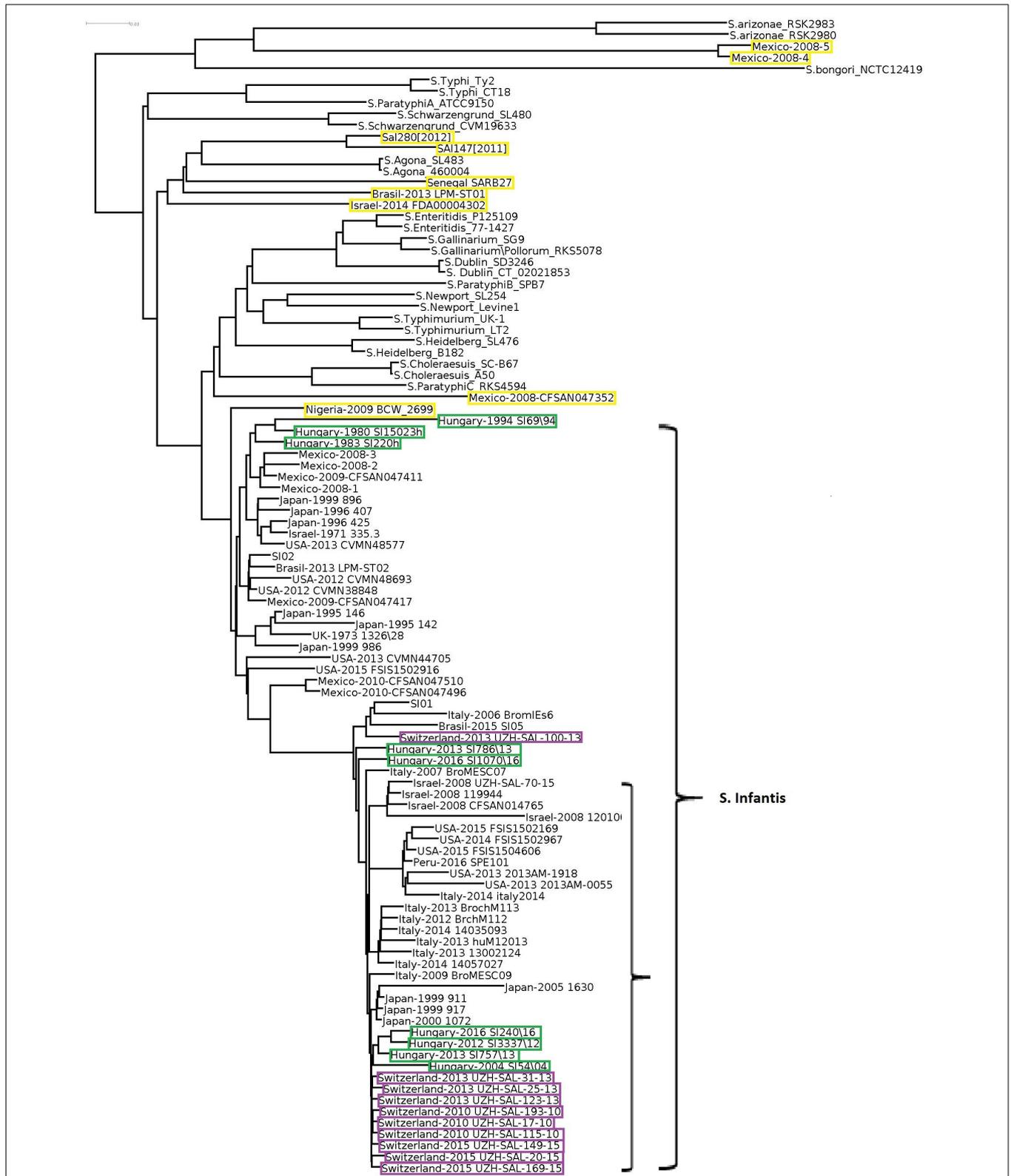
FIGURE 1 | Whole genome tree of all selected *Salmonella* strains including nine Hungarian and 67 global *S.* Infantis strains and 26 strains of 13 non-Infantis *Salmonella* serovars. Green boxes show the pre-emerging and the recent Hungarian isolates; purple boxes indicate Swiss strains, while yellow boxes indicate the *Salmonella* strains originally deposited as *S.* Infantis but outlying from the Infantis cluster. The main *S.* Infantis subclade is indicated. The bar represents average similarity between the genomes corresponding to genetic distance. These signs were consequently applied to all figures. *S. bongori* was used as outgroup control.

TABLE 1 | Distribution of core and accessory genes according to the tested set of Salmonella strains.

Group	Core genes (99–100%)	%	Soft core genes (95–99%)	%	Shell genes (15–95%)	%	Cloud genes (<15%)	%	Whole genome (100%)	%
All *Salmonella* (*n* = 102)	2030	10,8	969	5,1	1924	10,2	13924	73,9	18847	100
All *S.* Infantis (*n* = 76)	2316	18,0	1268	9,8	1132	8,8	8175	63,4	12891	100
S. Infantis Hungary (*n* = 9)	4123	78,8	0	0	711	13,6	395	7,6	5229	100
Hungary + Swiss (*n* = 19)	4053	75,0	0	0	670	12,4	684	12,7	5407	100
Hungary + United States (*n* = 19)	3968	65,2	0	0	898	14,8	1220	20,0	6086	100
Hungary + Japan (*n* = 19)	3863	65,5	0	0	1014	17,2	1020	17,3	5897	100
Hungary + Italy (*n* = 19)	4102	72,3	0	0	648	11,4	920	16,2	5670	100
Swiss + Italy (*n* = 20)	4139	77,6	82	1,538	438	8,2	672	12,6	5331	100
Swiss + Japan (*n* = 20)	3869	69,1	223	3,984	715	12,8	790	14,1	5597	100
Swiss + United States (*n* = 20)	3975	68,8	127	2,198	681	11,8	996	17,2	5779	100
Italy + Japan (*n* = 20)	3887	66,7	228	3,911	727	12,5	988	16,9	5830	100
Italy + United States (*n* = 20)	4004	67,2	115	1,93	668	11,2	1171	19,7	5958	100
Japan + United States (*n* = 20)	3814	61,7	238	3,849	912	14,8	1219	19,7	6183	100

S. Infantis strains were divided into two large clades by a third clade of non-Infantis serovars (including four outlier strains) in the central position (**Figure 3**). The Hungarian and Swiss strains formed separate well identifiable clusters in both clades of S. Infantis. Besides, the Hungarian strains consistently exhibited the time-related separation (pre-emerging and recent isolates) as it was observed in WGS analysis (**Figure 1**). Furthermore, one of the Swiss strains (Switzerland 2010 UZH-SAL-193-10) branched further away from the main clade together with one Brazilian strain of *S.* Infantis. Besides, the outliers from Mexico (Mexico-2008-4 and Mexico-2008-5) were positioned also near to this cluster together with *S. bongori* and *S. arizonae* (**Figure 3**).

When cloud genes of the *S.* Infantis strains were analyzed, their distribution appeared much more diverse (**Figure 4**). The Hungarian strains consistently exhibited the time-related separation observed previously. One large cluster of six Hungarian emerging strains also included one Swiss strain (Switzerland-215-UZH-SAL-2015) representing the only Swiss strain with genetic relatedness to Hungarian strains. The Swiss isolates did not form a distinct cluster but they were separated into four subclusters grouped together with some strains from Italy, Israel, and Japan, showing no close genomic relation to either the recent or the pre-emergent Hungarian strains in this cloud gene-based analysis (**Figure 4**). The outlier Infantis strains Senegal SARB27, Mexico-2008-4 and -5, Brasil-2013 LPM-ST01, Sal280[2012] and SAl147[2011], Israel-2014 FDA00004302, Mexico-2008-CFSAN047352, and Nigeria-2009 BCW_2699 were located separately. The congruent results of the whole-, core-, and cloud-genome-based analyses imply that these outlier strains do not belong to the serovar Infantis (**Figures 3, 4**).

Antibiotic Resistance Genotype of S. Infantis Strains

Out of the 76 *S.* Infantis strains tested, 45 were predicted as multiresistant on the basis of the co-existence of at least three resistance genes (**Supplementary Table S3**).

According to this, the association between genes *tet*(A), *sul1*, and *aadA1* was found in 64.4% of the MDR strains. In some strains from the United States and Italy, a high abundance of antibiotic resistance genes was detected including ESBL genes $bla_{CTX-M-65}$ or $bla_{CTX-M-1}$. Multiresistance genotypes were more characteristic to the recent strains, isolated between 2000 and 2016, while most of the old isolates (years 1971–1999) did not carry acquired antibiotic resistance genes. Examining the geographical distribution of resistance genes, we found that no strain isolated from Mexico exhibited antibiotic resistance, unlike strains isolated in other countries.

Analysis of Flagellin Genes of *Salmonella* Serovar Infantis

The presence of *flj*A, *flj*B, *fli*A, *fli*B, *fli*C, *fli*D, *fli*S, and *hin* flagellin- or flagella related genes was tested in all 102 selected *Salmonella* strains. The *fli*A,B,D,S genes were detected in almost all serovars represented here (**Supplementary Table S4**). Most strains of *S.* Infantis were characterized by the co-occurrence of all the above flagellar genes. Exceptions were the strains Hungary-2013 SI757/13, United States-2014 FSIS1502967, 7 strains from Japan, Mexico-2008-1, and Brasil-2013 LPM-ST02 which showed the absence of one of the genes. Three of the outlier strains, Israel-2014 FDA00004302, Senegal SARB27, and Nigeria-2009_BCW_2699 harbored all the eight flagellar genes; however, they were consistently separated from the Infantis clusters in all phylogenetic analyses (**Figures 1–3**). The other six outlier *Salmonella* isolates were deficient in several flagellar genes that further strengthened the need for validation of the serovar on the basis of genomic sequences as well.

Results on the *in silico* prediction of the antigenic profiles (O:H1:H2) and of the MLST profile of the above nine outlier *Salmonella* strains are presented in **Table 2**. According to this, the Senegal SARB27 (ST79) was the only strain that was predicted as *S.* Infantis, but it was assigned to
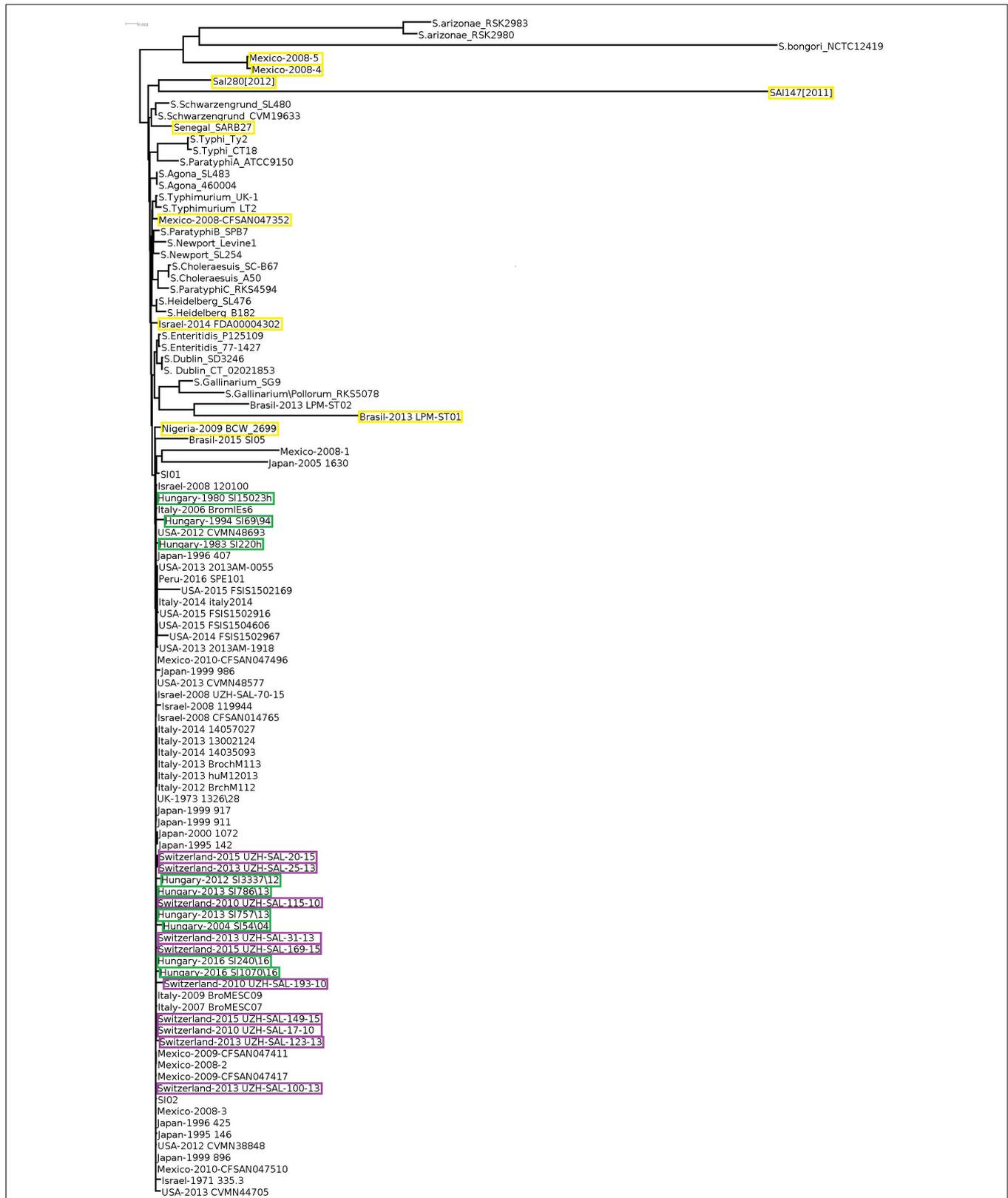
FIGURE 2 | Core genome tree of all *Salmonella* strains investigated. Yellow boxes indicate the *Salmonella* strains outlying from the *S.* Infantis cluster. Only the strains from Hungary, Switzerland, and the outlier *S.* Infantis isolates are highlighted by color boxes.

FIGURE 3 | Cloud-gene-based tree of all *Salmonella* serovars. The outlier *S.* Infantis isolates are indicated by yellow boxes, while strains from Hungary and Switzerland are highlighted by green and purple boxes, respectively.
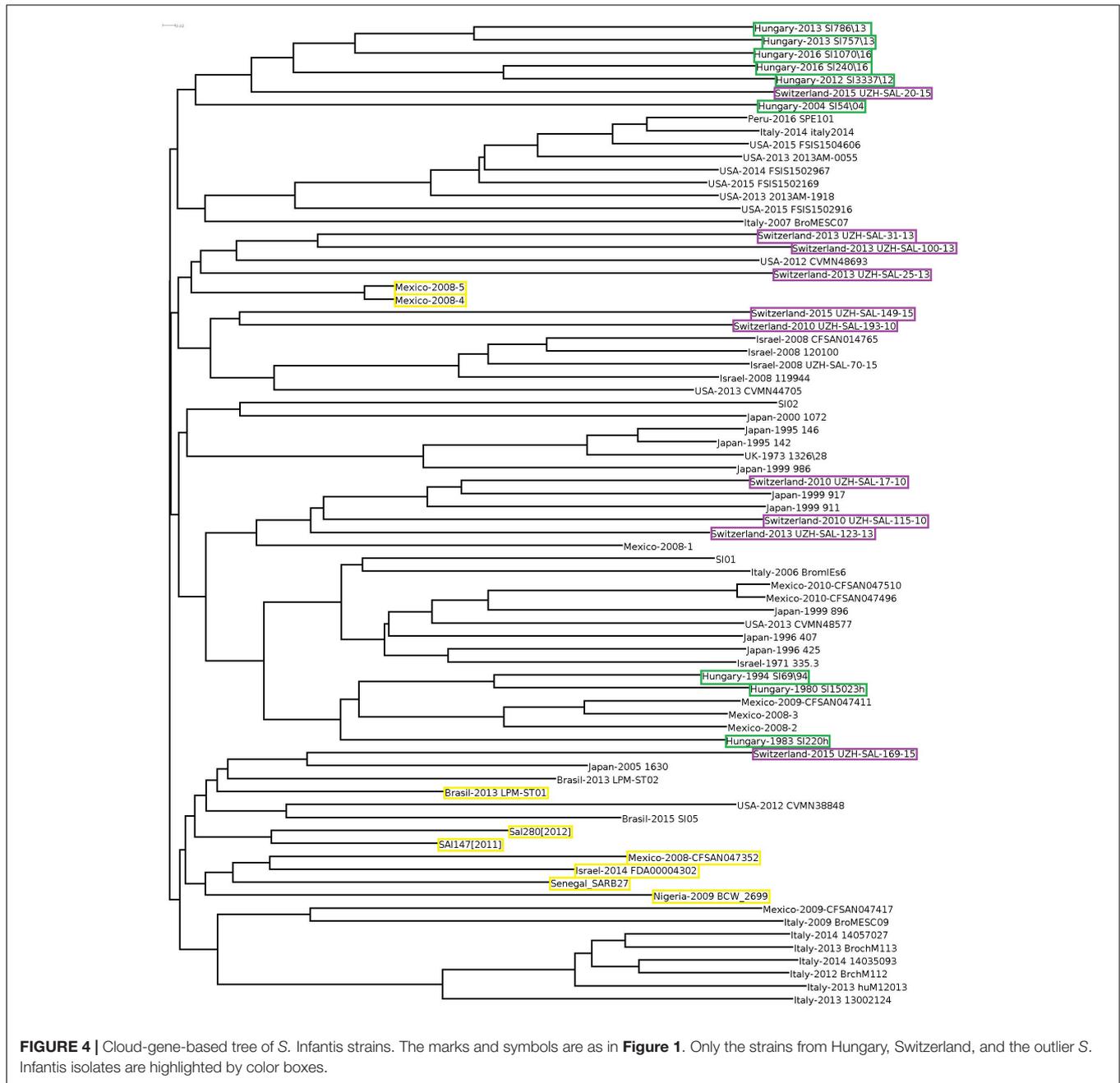
FIGURE 4 | Cloud-gene-based tree of *S.* Infantis strains. The marks and symbols are as in **Figure 1**. Only the strains from Hungary, Switzerland, and the outlier *S.* Infantis isolates are highlighted by color boxes.

ST79 instead of ST32 characteristic for *S.* Infantis. Two strains, Sal147[2011] and Sal280[2012], were identified as *S.* Rissen (7:f,g:-) both belonging to ST469. Further tree strains deposited as *S.* Infantis were predicted as Kentucky, Thompson, and Ouakarm assigned to ST198, ST26, and ST1610, respectively. The Mexican isolate Mexico-2008-4 was predicted to be *S. enterica* subsp. *diarizonae* with the antigenic profile of 60:r:e,n,x,z$_{15}$ and ST63. For two outlier strains (Mexico-2008-5 and Nigeria-2009 BCW_2699), there was not possible to detect the serovar on the basis of genomic sequences (**Table 2**). All other *S.* Infantis strains proved to be ST32 including those that were regarded

as Infantis-like in **Table 2**. The only exception was the serovar Gege (ST36).

Identification and Analysis of Genomic and Pathogenicity Islands

In order to reveal the prevalence and diversity of GIs of *Salmonella*, the first annotated complete *S.* Infantis genome UK-1973 1326/28 (Olasz et al., 2015) was used as a reference to determine their number, location, length and coding capacity (**Supplementary Table S5**—under tab "1326_28 island summary" and "1326_28 island annotation," respectively). In

TABLE 2 | Serovars and sequence types (STs) of the outlier *Salmonella* Infantis isolates on the basis of *in silico* prediction.

Indoor reference	WGS	Core	Cloud	Number of Infantis related flagellin genes	Serovar	Antigenic profile (O:H1:H2)	MLST
						Outliers according to the genome analyses	
UK-1973 1326/28	−	−	−	8	Infantis	7:r:1,5	ST32
Senegal SARB27	+	+	+	8	Infantis	7:r:1,5	ST79
SAl147[2011]	+	+	+	2	Rissen	7:f,g:−	ST469
Sal280[2012]	+	+	+	3	Rissen	7:f,g:−	ST469
Israel-2014 FDA00004302	+	+	+	8	Kentucky	8:i:z6	ST198
Mexico-2008-CFSAN047352	+	+	+	6	Thompson	7:k:1,5	ST26
Brasil-2013 LPM-ST01	+	+	+	3	Ouakarm	9,46:z29:e,n,x	ST1610
Mexico-2008-4	+	+	+	6	subsp. *diarizonae (IIIb)*	60:r:e,n,x,z15	ST63
Nigeria-2009 BCW_2699	+	+	+	8	N/A	7:-:−	ST603
Mexico-2008-5	+	+	+	7	N/A	60:-:−	ST63
Brasil-2013 LPM-ST02	−	+	+	7	Infantis	7:r:1,5	ST32
Japan-2005 1630	−	+	−	7	N/A	7:-:1,5	ST32
Mexico-2008-1	−	+	−	7	Gege	30:r:1,5	ST36
Brasil-2015 SI05	−	+	−	8	N/A	7:-:−	ST32
Switzerland-2015 UZH-SAL-169-15	−	−	+	8	Infantis	7:r:1,5	ST32
USA-2012 CVMN38848	−	−	+	8	Infantis	7:r:1,5	ST32

The reference genome of *S*. Infantis is indicated in gray.

The maximum number of flagellin or flagellin-related genes is eight (data from **Supplementary Table S4**).

Outliers (nine strains) are above the black line (except the reference strain) while Infantis-like isolates below the black line will be discussed consequently as Infantis.

WGS: whole genome analysis (data from **Figure 1**);

core: core gene analysis (data from **Figure 2**);

cloud: cloud genome analysis (data from **Figure 4**);

N/A: serovar is not available by *in silico* analysis (SeqSero).

the genome of the reference strain, 52 segments with coding capacity for 633 annotated genes were detected as parts of putative GIs whose length ranged between 4082 and 47,931 bp (**Supplementary Table S5**). The genes *avr*A and *sop*E2 characteristic for SPI-1 were found on the predicted islands gi.20 and gi.23, respectively, and the SPI2-specific *spi*C and *ssa*Q gene was located on island gi.19 and gi.10, respectively, indicating that the prediction method applied was able to identify the well-known SPIs, i.e., SPI-1 and SPI-2.

Using the sequences of the 52 predicted islands (GIs and SPIs), their prevalence was tested in all *S*. Infantis strains (**Supplementary Table S6**). The presence of a particular island varied between 5 and 76 occurrences with the average of 70.9 (93.2%) among the 76 genomes. It also should be mentioned that the outlier strains Brasil-2013_LPM-ST01, Senegal_SARB27, SAl147[2011], Sal280[2012] Mexico-2008-4, and -5, harbored less than 83% of the 52 predicted islands (**Supplementary Table S6**), indicating their divergent evolutionary lineages.

DISCUSSION

Over the past decades, *S*. Infantis has become more prevalent among broiler flocks, representing an increasing health hazard for humans in Europe, Israel, Japan, and United States. Accordingly, *S*. Infantis strains of humans and poultry have been studied worldwide (Shahada et al., 2006, 2010; Cloeckaert et al., 2007; Nógrády et al., 2007, 2008, 2012; Gal-Mor et al., 2010; Dionisi et al., 2011; Tate et al., 2017) and a growing portion of strains proved to be MDR. So far, only few genomic analyses have been reported on different *Salmonella* serovars (Zou et al., 2013) and *S*. Infantis strains (Yokoyama et al., 2014), with special regard to MDR *S*. Infantis of the respective countries (Franco et al., 2015; Hindermann et al., 2017; Szmolka et al., 2018; Acar et al., 2019). Here we provide a comprehensive analysis about the possible relations between *S*. Infantis strains from different countries and between *S*. Infantis and other non-Infantis serovars. For this, we applied whole-, core-, and accessory genome analyses.

Whole genome comparisons showed that a tendency for geographic origin was reflected in clustering of *S*. Infantis isolates from countries represented by 9–10 isolates from Switzerland, Hungary, Japan, Italy, and United States (**Figure 1** and **Table 2**). This may indicate the differences in contamination and re-contamination of production units from their environments together with the differing local practices of the antibiotic usage which could influence the selection and persistence of various clones and subclones of *S*. Infantis.

Our data on antimicrobial resistance confirm the recent global emergence of resistance of S. Infantis. From the presented data cannot be concluded that the spread of a single clone could be responsible for the wide spread of antibiotic resistance, but it is more likely to be due to multiple simultaneously

appearing clones with similar resistance patterns (Gal-Mor et al., 2010; Franco et al., 2015; Tate et al., 2017; Acar et al., 2019; Bogomazova et al., 2019).

In this study, the nine Hungarian isolates showed a time-related separation as well: the pre-emerging and recent isolates formed two distinct clusters. Furthermore, nine out of 10 genomes of recent Swiss isolates also formed a single cluster in WGS analyses (**Figure 1**). Our results (**Figures 1**, **2**, **4**) demonstrated that the core genome analysis as performed by Hindermann et al. (2017) is not discriminative enough to determine whether or not the *S.* Infantis strains from Switzerland are really clustered together with, and could be epidemiologically related to strains from Hungary, Japan, Italy, United States, and Israel. Our results also indicated that the analysis of core genomes provides less detailed information about the different genetic lineages as discussed by Pornsukarom et al. (2018). In this respect, the phylogenetic reconstructions based on cloud genes (**Figures 3**, **4**) remarkably differ also from that based on the whole genomes (**Figure 1**). According to our cloud genome analysis, the recent Hungarian isolates represent a cluster being separated not only from the Swiss isolates but also from most of the European and global isolates (**Figures 3**, **4**). Therefore, the core genome-based suggestion of Hindermann et al. (2017) that the Hungarian MDR clone of *S.* Infantis would be documented in Swiss food and human strains and it would have "spread within and outside Europe" seems to be unsubstantiated. Our present data and those from the literature (Gal-Mor et al., 2010; Franco et al., 2015; Tate et al., 2017; Acar et al., 2019) indicate that some *S.* Infantis clones may emerge and persist in different geographic areas almost simultaneously, suggesting that vertical dissemination of certain clones within the poultry sector should also be considered.

All of our genome analyses suggested that nine of the 76 strains (Sal147[2011], Sal280[2012], Israel-2014 FDA00004302, Mexico-2008-CFSAN047352, Brasil-2013 LPM-ST01, Mexico-2008-4, Mexico-2008-5, Nigeria-2009 BCW_2699, and Senegal SARB27) previously identified as *S.* Infantis are distant from the main Infantis clade and could represent other serovars, except the strain Senegal SARB27. For instance, the isolates (Mexico-2008-4 and Mexico-2008-5) are much closer to strains of *S. enterica* subsp. *arizonae* and even to *S. bongori* than to *S.* Infantis. Indeed, one of these Mexican isolates proved to be *S. enterica* subsp. *diarizonae* in *in silico* antigenic profile prediction. Senegal SARB27 was typed by *in silico* MLST typing as ST79 that is more likely characteristic for some exotic strains of *S.* Infantis[1]. These discrepancies between the results of traditional and whole genome sequence-based typing methods point to the need of exploiting the available *in silico* tools for strain typing in order to avoid misidentification.

Our analysis is in accordance to the data of Liu et al. (2017) that the PCR systems based on the eight selected flagellin and flagellin related gene sequences can be applicable for the detection of Infantis serovar; however, some exceptions are possible. Accordingly, in case of 76 *S.* Infantis genomes, this analysis mainly corresponds to the results of "conventional"

---

[1]http://enterobase.warwick.ac.uk/species/index/senterica

serovar classification. However, it should be mentioned that according to our MLST analysis, the nine outlier isolates need to be reclassified.

Analyzing "non-Infantis" and "Infantis only" genomes separately confirmed that the distinguishing resolution was generally better in the Infantis only system as compared to the combined analysis on "non-Infantis" plus "Infantis." As the *S.* Infantis genomes were missing from the analyses of Zou et al. (2013), we also wanted to get some idea about possible differentiation of *S.* Infantis from the main *Salmonella* serovars. Here we did not only confirm the results of Zou et al. (2013) about genetic distances between the tested *Salmonella* serovars and *S.* Infantis, but also confirmed our findings on genomic distances between certain lineages of *S.* Infantis.

CONCLUSION

The comprehensive genome analysis based on the whole-, core-, and accessory genomes could better elucidate genetic relatedness of local and global as well as of pre-emerging and emerging strains of *S.* Infantis. Among these methods the core genome analysis has the least discriminatory power, while analysis of cloud genes offered the most detailed insight into the genetic distances and relationships of *S.* Infantis strains. Our analyses demonstrated that, contrary to the claim of Hindermann et al. (2017), the Swiss strains are not directly related to the Hungarian ones. Accordingly, in spite of the comprehensive analysis of several genomic characteristics, no direct epidemiologic links between *S.* Infantis strains from different countries could be established. Our results also suggest the need for genome-based re-classification of some *Salmonella* strains for which the sequences were originally deposited in the databases as *S.* Infantis.

DATA AVAILABILITY STATEMENT

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

AUTHOR CONTRIBUTIONS

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2020.00539/full#supplementary-material

REFERENCES

Acar, S., Bulut, E., Stasiewicz, M. J., and Soyer, Y. (2019). Genome analysis of antimicrobial resistance, virulence, and plasmid presence in Turkish *Salmonella* serovar Infantis isolates. *Int. J. Food. Microbiol.* 307:108275. doi: 10.1016/j.ijfoodmicro.2019.108275

Agarwala, R., Barrett, T., Beck, J., Benson, D. A., Bollin, C., Bolton, E., et al. (2016). Database resources of the national center for biotechnology information. *Nucleic Acids Res.* 44, D7–D19. doi: 10.1093/nar/gkv1290

Asai, T., Itagaki, M. Y., Shiroki, Y., Yamada, M., Tokoro, M., Kojima, A., et al. (2006). Antimicrobial resistance types and genes in *Salmonella enterica* infantis isolates from retail raw chicken meat and broiler chickens on farms. *J. Food Prot.* 69, 214–216. doi: 10.4315/0362-028x-69.1.214

Aviv, G., Tsyba, K., Steck, N., Salmon-Divon, M., Cornelius, A., Rahav, G., et al. (2014). A unique megaplasmid contributes to stress tolerance and pathogenicity of an emergent *Salmonella enterica* serovar Infantis strain. *Environ. Microbiol.* 16, 977–994. doi: 10.1111/1462-2920.12351

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., et al. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* 19, 455–477. doi: 10.1089/cmb.2012.0021

Bertelli, C., Laird, M. R., Williams, K. P., Lau, B. Y., Hoad, G., Winsor, G. L., et al. (2017). IslandViewer 4: expanded prediction of genomic islands for larger-scale datasets. *Nucleic Acids Res.* 45, W30–W35. doi: 10.1093/nar/gkx343

Bogomazova, A. N., Gordeeva, V. D., Krylova, E. V., Soltynskaya, I. V., Davydova, E. E., Ivanova, O. E., et al. (2019). Mega-plasmid found worldwide confers multiple antimicrobial resistance in *Salmonella* Infantis of broiler origin in Russia. *Int. J. Food Microbiol.* \ 26:108497. doi: 10.1016/j.ijfoodmicro.2019.108497

Boyd, E. F., Wang, F.-S., Beltran, P., Plock, S. A., Nelson, K., and Selander, R. K. (1993). Salmonella reference collection B (SARB): strains of 37 serovars of subspecies I. *J. Gen. Microbiol.* 139, 1125–1132. doi: 10.1099/00221287-139-6-1125

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., et al. (2009). BLAST+: architecture and applications. *BMC Bioinformatics* 10:421. doi: 10.1186/1471-2105-10-421

Cloeckaert, A., Praud, K., Doublet, B., Bertini, A., Carattoli, A., Butaye, P., et al. (2007). Dissemination of an extended-spectrum-beta-lactamase bla(TEM-52) gene-carrying IncI1 plasmid in various *Salmonella enterica* serovars isolated from poultry and humans in Belgium and France between 2001 and 2005. *Antimicrob. Agents Chemother.* 51, 1872–1875. doi: 10.1128/aac.01514-06

Darling, A. E., Mau, B., and Perna, N. T. (2010). progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. *PLoS One* 5:e11147. doi: 10.1371/journal.pone.0011147

Dionisi, A. M., Lucarelli, C., Benedetti, I., Owczarek, S., and Luzzi, I. (2011). Molecular characterisation of multidrug-resistant *Salmonella enterica* serotype Infantis from humans, animals and the environment in Italy. *Int. J. Antimicrob. Agents* 38, 384–389. doi: 10.1016/j.ijantimicag.2011.07.001

European Food Safety Authority [EFSA] and European Centre for Disease Prevention [ECDC] (2017a). The European Union summary report on antimicrobial resistance in zoonotic and indicator bacteria from humans, animals and food in 2015. *EFSA J.* 15:e04694. doi: 10.2903/j.efsa.2017.4694

European Food Safety Authority [EFSA] and European Centre for Disease Prevention [ECDC] (2017b). The European Union summary report on trends and sources of zoonoses, zoonotic agents and food−borne outbreaks in 2016. . *EFSA J.* 15:e05077. doi: 10.2903/j.efsa.2017.5077

European Food Safety Authority [EFSA] and European Centre for Disease Prevention [ECDC] (2019). The European Union summary report on antimicrobial resistance in zoonotic and indicator bacteria from humans, animals and food in 2017. *EFSA J.* 17:5598 doi: 10.2903/j.efsa.2019.5598

Fookes, M., Schroeder, G. N., Langridge, G. C., Blondel, C. J., Mammina, C., Connor, T. R., et al. (2011). Salmonella bongori provides insights into the evolution of the Salmonellae. *PLoS Pathog.* 7:e1002191. doi: 10.1371/journal.ppat.1002191

Franco, A., Leekitcharoenphon, P., Feltrin, F., Alba, P., Cordaro, G., Iurescia, M., et al. (2015). Emergence of a clonal lineage of multidrug-resistant ESBL-Producing *Salmonella* Infantis Transmitted from broilers and broiler meat to humans in Italy between 2011 and 2014. *PLoS One* 10:e0144802. doi: 10.1371/journal.pone.0144802

Gal-Mor, O., Valinsky, L., Weinberger, M., Guy, S., Jaffe, J., Schorr, Y. I., et al. (2010). Multidrug-resistant *Salmonella enterica* serovar Infantis. *Israel. Emerg. Infect. Dis.* 16, 1754–1757. doi: 10.3201/eid1611.100100

Hindermann, D., Gopinath, G., Chase, H., Negrete, F., Althaus, D., Zurfluh, K., et al. (2017). *Salmonella enterica* serovar infantis from food and human infections, Switzerland, 2010-2015: poultry-related multidrug resistant clones and an emerging ESBL producing clonal lineage. *Front. Microbiol.* 8:1322. doi: 10.3389/fmicb.2017.01322

Huson, D. H., Richter, D. C., Rausch, C., Dezulian, T., Franz, M., and Rupp, R. (2007). Dendroscope: an interactive viewer for large phylogenetic trees. *BMC Bioinformatics* 8:460. doi: 10.1186/1471-2105-8-460

Inglin, R. C., Meile, L., and Stevens, M. J. A. (2018). Clustering of Pan- and core-genome of *lactobacillus* provides novel evolutionary insights for differentiation. *BMC Genomics* 19:284. doi: 10.1186/s12864-018-4601-5

Kardos, G., Farkas, T., Antal, M., Nógrády, N., and Kiss, I. (2007). Novel PCR assay for identification of *Salmonella enterica* serovar Infantis. *Lett. Appl. Microbiol.* 45, 421–425. doi: 10.1111/j.1472-765X.2007.02220.x

Larsen, M. V., Cosentino, S., Rasmussen, S., Friis, C., Hasman, H., Marvig, R. L., et al. (2012). Multilocus sequence typing of total-genome-sequenced bacteria. *J. Clin. Microbiol.* 50, 1355–1361. doi: 10.1128/jcm.06094-11

Leinonen, R., Sugawara, H., Shumway, M., and International Nucleotide Sequence Database (2011). The sequence read archive. *Nucleic Acids Res.* 39, D19–D21. doi: 10.1093/nar/gkq1019

Liu, Y., Zhang, D.-F., Zhou, X., Xu, L., Zhang, L., and Shi, X. (2017). Comprehensive analysis reveals two distinct evolution patterns of *Salmonella* flagellin gene clusters. *Front. Microbiol.* 8:2604. doi: 10.3389/fmicb.2017.02604

Murakami, K., Ishihara, T., Horikawa, K., and Oda, T. (2007). Features of *Salmonella* serovars among food handlers in Kyushu. *Japan. New Microbiol.* 30, 155–159.

Nógrády, N., Kardos, G., Bistyák, A., Turcsányi, I., Mészáros, J., Galantai, Z., et al. (2008). Prevalence and characterization of *Salmonella* infantis isolates originating from different points of the broiler chicken-human food chain in Hungary. *Int. J. Food Microbiol.* 127, 162–167. doi: 10.1016/j.ijfoodmicro.2008.07.005

Nógrády, N., Király, M., Davies, R., and Nagy, B. (2012). Multidrug resistant clones of *Salmonella* Infantis of broiler origin in Europe. *Int. J. Food Microbiol.* 157, 108–112. doi: 10.1016/j.ijfoodmicro.2012.04.007

Nógrády, N., Tóth, Á, Kostyák, Á, Pászti, J., and Nagy, B. (2007). Emergence of multidrug-resistant clones of *Salmonella* Infantis in broiler chickens and humans in hungary. *J. Antimicrob. Chemother.* 60, 645–648. doi: 10.1093/jac/dkm249

O'Brien, S. J. (2013). The "decline and fall" of nontyphoidal *Salmonella* in the United Kingdom. *Clin. Infect. Dis.* 56, 705–710. doi: 10.1093/cid/cis967

Olasz, F., Nagy, T., Szabó, M., Kiss, J., Szmolka, A., Barta, E., et al. (2015). Genome Sequences of Three *Salmonella enterica* subsp. enterica serovar infantis strains from healthy broiler chicks in hungary and in the United Kingdom. *Genome Announc* 3:e01468-14. doi: 10.1128/genomeA.01468-14

Page, A. J., Cummins, C. A., Hunt, M., Wong, V. K., Reuter, S., Holden, M. T. G., et al. (2015). Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics* 31, 3691–3693. doi: 10.1093/bioinformatics/btv421

Pornsukarom, S., van Vliet, A. H. M., and Thakur, S. (2018). Whole genome sequencing analysis of multiple *Salmonella* serovars provides insights into phylogenetic relatedness, antimicrobial resistance, and virulence markers across humans, food animals and agriculture environmental sources. *BMC Genomics* 19:801. doi: 10.1186/s12864-018-5137-4

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

Shahada, F., Chuma, T., Dahshan, H., Akiba, M., Sueyoshi, M., and Okamoto, K. (2010). Detection and characterization of extended-spectrum beta-lactamase (TEM-52)-producing *Salmonella* serotype Infantis from broilers in Japan. *Foodborne Pathog. Dis.* 7, 515–521. doi: 10.1089/fpd.2009.0454

Shahada, F., Chuma, T., Tobata, T., Okamoto, K., Sueyoshi, M., and Takase, K. (2006). Molecular epidemiology of antimicrobial resistance among *Salmonella enterica* serovar Infantis from poultry in Kagoshima, Japan. *Int. J. Antimicrob. Agents* 28, 302–307. doi: 10.1016/j.ijantimicag.2006.07.003

Szmolka, A., Szabó, M., Kiss, J., Pászti, J., Adrián, E., Olasz, F., et al. (2018). Molecular epidemiology of the endemic multiresistance plasmid pSI54/04 of *Salmonella* Infantis in broiler and human population in Hungary. *Food Microbiol.* 71, 25–31. doi: 10.1016/j.fm.2017.03.011

Tate, H., Folster, J. P., Hsu, C. H., Chen, J., Hoffmann, M., and Li, C. (2017). Comparative analysis of extended-spectrum-β-Lactamase CTX-M-65-producing *salmonella enterica* serovar infantis isolates from humans, food animals, and retail chickens in the United States. *Antimicrob. Agents Chemother.* 61, e488–e417. doi: 10.1128/AAC.00488-17

Wilk, T., Szabó, M., Szmolka, A., Kiss, J., Barta, E., Nagy, T., et al. (2016). Genome Sequences of Multidrug-Resistant *Salmonella enterica* subsp. enterica Serovar Infantis Strains from Broiler Chicks in Hungary. *Genome Announc* 4:e01400-16. doi: 10.1128/genomeA.01400-16

Wilk, T., Szabó, M., Szmolka, A., Kiss, J., Olasz, F., and Nagy, B. (2017). Genome Sequences of *Salmonella enterica* subsp. enterica serovar infantis strains from hungary representing two peak incidence periods in three decades. *Genome Announc* 5:e01735-16. doi: 10.1128/genomeA.01735-16

Yokoyama, E., Murakami, K., Shiwa, Y., Ishige, T., Ando, N., Kikuchi, T., et al. (2014). Phylogenetic and population genetic analysis of *Salmonella enterica* subsp enterica serovar Infantis strains isolated in Japan using whole genome sequence data. *Infect. Genet. Evol.* 27, 62–68. doi: 10.1016/j.meegid.2014.06.012

Zankari, E., Hasman, H., Cosentino, S., Vestergaard, M., Rasmussen, S., Lund, O., et al. (2012). Identification of acquired antimicrobial resistance genes. *J Antimicrob Chemother.* 67, 2640–2644. doi: 10.1093/jac/dks261

Zhang, S. K., Yin, Y. L., Jones, M. B., Zhang, Z. Z., Kaiser, B. L. D., Dinsmore, B. A., et al. (2015). *Salmonella* serotype determination utilizing high-throughput genome sequencing data. *J. Clin. Microbiol.* 53, 1685–1692. doi: 10.1128/jcm.00323-15

Zou, Q. H., Li, R. Q., Wang, Y. J., and Liu, S. L. (2013). Identification of genes to differentiate closely related *Salmonella* lineages. *PLoS One* 8:e55988. doi: 10.1371/journal.pone.0055988