Antimicrobial resistance genes, virulence markers and prophage sequences in Salmonella enterica serovar Enteritidis isolated in Tunisia using whole genome sequencing

Boutheina Ksibi, Sonia Ktari, Kais Ghedira, Houcemeddine Othman, Sonda Maalej, Basma Mnif, Laetitia Fabre, Faouzia Rhimi, Simon Le Hello, Adnene Hammami

To cite this version:

Boutheina Ksibi, Sonia Ktari, Kais Ghedira, Houcemeddine Othman, Sonda Maalej, et al.. Antimicrobial resistance genes, virulence markers and prophage sequences in Salmonella enterica serovar Enteritidis isolated in Tunisia using whole genome sequencing. Current Research in Microbial Sciences, 2022, 3, pp.100151. 10.1016/j.crmicr.2022.100151. pasteur-04132990

HAL Id: pasteur-04132990
https://pasteur.hal.science/pasteur-04132990
Submitted on 19 Jun 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives| 4.0 International License
Antimicrobial resistance genes, virulence markers and prophage sequences in *Salmonella enterica serovar* Enteritidis isolated in Tunisia using whole genome sequencing

Boutheina Ksibi\(^a, b, \ast\), Sonia Ktari\(^a, b\), Kais Ghedira\(^c\), Houcemeddine Othman\(^d\), Sonda Maalej\(^a, b\), Basma Mnif\(^a, b\), Laetitia Fabre\(^e\), Faouzia Rhimi\(^a, b\), Simon Le Hello\(^e, f\), Adnene Hammami\(^a, b\)

\(^{a}\) University of Sfax, Faculty of Medicine of Sfax-Tunisia, Laboratory of Microbiology, Avenue Majida Bouilla, 3027 Sfax, Tunisia
\(^{b}\) Habib Bourguiba University Hospital Sfax, Tunisia, Research Laboratory Microorganisms and Human Disease – MPH LR03SP03, Tunisia
\(^{c}\) Laboratory of Bioinformatics, Biomathematics and Biostatistics, Institut Pasteur de Tunis, University of Tunis El Manar (UTM), Tunisia
\(^{d}\) Sydney Brenner Institute for Molecular Bioscience, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa
\(^{e}\) Institut Pasteur, Centre National de Référence des Escherichia coli, Shigella and Salmonella, Unité des Bactéries Pathogènes Entériques, 75015 Paris, France
\(^{f}\) Groupe de Recherche sur l’Adaptation Microbienne (GRAM 2.0), Normandie Univ, UNICAEN, UNIROUEN, GRAM 2.0, 14000 Caen, France

**ARTICLE INFO**

Keywords:
Whole genome sequencing
*S. Enteritidis*
Salmonella pathogenicity islands
Antimicrobial resistance
Prophages sequences

**ABSTRACT**

*Salmonella* Enteritidis causes a major public health problem in the world. Whole genome sequencing can give us a lot of information not only about the phylogenetic relatedness of these bacteria but also in antimicrobial resistance and virulence gene predictions. In this study, we analyzed the whole genome data of 45 *S. Enteritidis* isolates recovered in Tunisia from different origins, human, animal, and foodborne samples. Two major lineages (A and B) were detected based on 802 SNPs differences. Among these SNPs, 493 missense SNPs were identified. A total of 349 orthologue genes mutated by one or two missense SNPs were classified in 22 functional groups with the prevalence of carbohydrate transport and metabolism group. A good correlation between genotypic antibiotic resistance profiles and phenotypic analysis were observed. Only resistant isolates carried the respective molecular resistant determinants. The investigation of virulence markers showed the distribution of 11 *Salmonella* pathogenicity islands (SPI) out of 23 previously described. The SPI-1 and SPI-2 genes encoding type III secretion systems were highly conserved in all isolates except one. In addition, the virulence plasmid genes were present in all isolates except two. We showed the presence of two fimbrial operons *sef* and *ste* previously considered to be specific for typhoidal *Salmonella*. Our collection of *S. Enteritidis* reveal a diversity among prophage profiles. SNPs analysis showed that missense mutations identified in *fimbriae* and in SPI-1 and SPI-2 genes were mostly detected in lineage B.

In conclusion, WGS is a powerful application to study functional genomic determinants of *S. Enteritidis* such as antimicrobial resistance genes, virulence markers and prophage sequences. Further studies are needed to predict the impact of the missenses SNPs that can affect the protein functions associated with pathogenicity.

1. Introduction

During the last two decades, whole-genome sequencing (WGS) has become the affordable tool that has the capacity to revolutionize different domains including genetics, microbiology, epidemiology and public health surveillance. The evolution of the current WGS technologies allowed to rapidly increase the realization of bacterial genome sequencing projects (Hu et al., 2021; Punina et al., 2015). Since the entire genome is readily available for analysis, WGS has the capacity to replace traditional methods for characterizations such as serotyping, virulotyping and antimicrobial resistance which can easily be predicted from the genome. This improves the capacity of surveillance systems to quickly provide information on the probable source, to identify the path of disease transmission within a population and to detect the virulence factors of the pathogen. Therefore, WGS-based analyses are becoming the primary subtyping tool of choice for pathogens particularly *Salmonella* species (Deng et al., 2015). To date, there are over 400,000 *Salmonella enterica* genomes published in public databases, that have been

\(^{\ast}\) Corresponding author at: University of Sfax, Faculty of Medicine of Sfax-Tunisia, Laboratory of Microbiology, Avenue Majida Bouilla, 3027 Sfax, Tunisia.

E-mail address: ksibi-taina@outlook.com (B. Ksibi).

https://doi.org/10.1016/j.crmicr.2022.100151

Available online 1 July 2022
2666-5174/© 2022 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
widely exploited to track and investigate outbreaks ("Home - Pathogen Detection - NCBI,"). *Salmonella enterica subsp enterica serovar Enteritidis* (S. Enteritidis) is among the most frequent *Salmonella* serovars isolated worldwide. It is a human and animal pathogen that causes a major public health problem. The genome of Enteritidis serovar has approximately 5 million bases and codes for over 4000 genes, from which more than 200 genes are actively dysfunctional. A global comparison study of S. Enteritidis isolates demonstrates the emergence of closely related isolates during a long period and underline the genetic homogeneity of this serovar. The virulence of *Salmonella enterica* depends on diverse assortment of genes which are required for adhesion, invasion, intra-cellular survival, and replication. These genes are located on various elements of the genome including *Salmonella* pathogenicity islands (SPIs) (Gupta et al., 2019; Zhao et al., 2020). Different strains of S. Enteritidis have mobile genetic elements such as plasmids, prophages and transposon that can carry virulence or antimicrobial resistance genes (Hu et al., 2021).

Recently, WGS has been successfully used in our previous studies to assess the genetic diversity of S. Enteritidis and *Salmonella Typhimurium* serovars (Ksibi et al., 2020; Khari et al., 2020). The use of these technologies has greatly enlarged our view of the genetic diversity of this bacterium. Furthermore, it produces an opportunity to provide more genetic information to study all genetic determinants such as virulence markers, antimicrobial resistance genes, mobile elements, bacterio-phages and to determine genomic changes associated with pathogenicity and antibiotic resistance.

In this study, we sought to investigate the whole genome of 45 S. Enteritidis strains isolated from different sources, human, animals and food in order to predict the antimicrobial resistance, the potential pathogenicity, the prophages sequences and the substitution of an amino acid on the protein sequence.

2. Material and methods

2.1. Bacterial strains

45 S. Enteritidis strains isolated from human, animal, and food samples were included in this study. A collection of 29 clinical S. Enteritidis isolates were recovered in the Laboratory of Microbiology CHU Habib Bourguiba Sfax-Tunisia between 2000 and 2015. In addition, seven S. Enteritidis isolates were collected in 2009, 2014, and 2015 from animal samples. We also included nine S. Enteritidis isolates obtained from two distinct foodborne outbreaks that occurred in 2007 (Table 1).

2.2. Antibiotic susceptibility test

Antibiotic susceptibilities of S. Enteritidis isolates were determined with the disk diffusion method, as recommended by European Committee on Antimicrobial Susceptibility Testing (EUCAST 2013). Antimicrobial susceptibility was performed for 16 antimicrobials, including ampicillin (10 µg), amikacin (30 µg), chloramphenicol (30 µg), azithromycin (30 µg), ciprofloxacin (5 µg), gentamicin (10 µg), streptomycin (10 µg), sulfadime (30 µg), tetracycline (30 µg), tigecycline (15 µg), streptomycin (10 µg), spectinomycin (100 µg), kanamycin (30 µg), netilmicin (10 µg), nalidixic acid (30 µg), pefloxacin (5 µg), and trimethoprim (30 µg). .

2.3. Whole genome sequencing, assembly and phylogeny

Genomic DNA was extracted using a MagNA Pure 96 DNA and Viral NA Small Volume Kit (Roche Life Science). Sequencing libraries were prepared with a Nextera XT DNA sample preparation kit (Illumina, Inc., San Diego, CA) and sequenced on an illumina NextSeq 500 platform (Illumina, San Diego, CA, USA) with 100 to 150-bp paired-end protocol according to the manufacturer’s instructions. A threshold of 30X minimum coverage was applied. Genome sequences were assembled using SPAdes software V.3.6.0 with default settings (Bankevich et al., 2012). Genome assemblies consisted of several contigs ranging from ~3.6 to 4.8 Mbp, with an average GC content of ~52%. The sequencing depth of coverage ranged between 81% to 94%. The genomes were annotated with the NCBI Prokaryote Genomes Annotation Pipeline (PGAP V.4.1) (http://ncbi.nlm.nih.gov/genes/static/Pipeline.html) (Tatusov et al., 2016). We identified a total number of genes ranged between 4313 and 5147, a number of protein coding regions ranging from 4228 to 5075, and a number of pseudogenes ranging from 150 to 665. The phylogenetic relationship based on SNPs was constructed by mapping all the paired-end reads to the S. Enteritidis reference genome (GenBank:AM933172) as previously described (Ksibi et al., 2020).

2.4. In silico analysis of virulence genes and prophage detection

*Salmonella Pathogenicity Islands* (SPIs) were detected using SPIFinder 1.0 database (https://cge.cbs.dtu.dk/services/SPIFinder/). Prophage sequences were detected using PHASTER web server (http://phaster.ca/); only “intact” prophages were considered for this analysis (Zhou et al., 2011).

2.5. In silico identification of resistance genes

Antimicrobial resistance genes were searched using the ResFinder database (https://cge.cbs.dtu.dk/services/ResFinder/) (Zankari et al., 2012) with the following thresholds: minimum length coverage of 90% and nucleotide sequence identity of 96%. PointFinder was used to detect chromosomal structural gene mutations, gyrA, gyrB, parC, and parE genes, which were analyzed for quinolone resistance-determining region (QRDR) mutations (Zankari et al., 2017).

2.6. SNPs annotation and Cog analysis

SNPs annotations were obtained using SnpEff v.4.12 with Ensembl gene annotation database for *Salmonella* (Cingolani et al., 2012). Clusters of Orthologous Groups of proteins (COGs) database was used for the functional annotation of SNPs and determination the most variable COGs for S. Enteritidis (Tatusov, 2000). The amino acid sequences generated from the SnpEff were used as input for functional annotation based on orthologous group.

2.7. Nucleotide sequence accession numbers

Complete genome sequences of these S. Enteritidis isolates are available in GenBank under BioProject no. PRJNA579483 and the GenBank accession numbers listed in Table 1.

3. Results

3.1. Phylogeny and functional SNPs

A phylogenetic analysis based on 806 SNPs revealed the existence of two lineages named A and B with 345 to 470 SNPs difference. Examination of individual lineage revealed that lineage A grouped twenty-six isolates with two clades (C1 and C2) and five subclades (C1–1 to C1–5). Lineage B included eighteen S. Enteritidis subdivided into two clades (C3 to C4) and two subclades (C3–1 and C3–2) placed in lineage B (Fig. 1). Variant annotations showed that 276 were silent SNPs (33.99%), 493 were misssenses SNPs (61.16%) and 37 were nonsense mutations (4.59%). The distribution of 493 missense SNPs among serovar Enteritidis was examined using the COG database. This annotation reflected that the genes are scattered in an assorted range of 22 functional categories throughout the genome. The COG distribution showed that the five most prevalent groups were carbohydrate transport and metabolism
| No. | Tree Label | biosample | Bioproject | genome filename | Isolation date | Isolation type | Isolation source | Resistance patterns | Antimicrobial resistance genes |
|-----|------------|-----------|------------|-----------------|---------------|----------------|------------------|-------------------|-------------------------------|
| 1   | C238       | SAMN13108479 | PRJNA579483 | WICN00000000 | C238.fasta | 2000 | Clinical | Stool | susceptible | gyra, blaTEM |
| 2   | C3532      | SAMN13110851 | PRJNA579483 | WIRC00000000 | C3532.fasta | 2000 | Clinical | unknown | susceptible | |
| 3   | X1500      | SAMN13110876 | PRJNA579483 | WIDC00000000 | X1500.fasta | 2000 | Clinical | unknown | susceptible | |
| 4   | B3015      | SAMN13108482 | PRJNA579483 | WIBL00000000 | B3015.fasta | 2001 | Clinical | Stool | susceptible | |
| 5   | D849       | SAMN13108483 | PRJNA579483 | WICS00000000 | D849.fasta | 2002 | Clinical | Stool | susceptible | |
| 6   | S1078      | SAMN13108484 | PRJNA579483 | WIDB00000000 | S1078.fasta | 2003 | Clinical | Stool | susceptible | |
| 7   | B3544      | SAMN13108480 | PRJNA579483 | WIBM00000000 | B3540.fasta | 2004 | Clinical | Stool | susceptible | |
| 8   | I42        | SAMN13108481 | PRJNA579483 | WICY00000000 | I42.fasta | 2005 | Clinical | Blood | susceptible | |
| 9   | B3790      | SAMN13108485 | PRJNA579483 | WIBN00000000 | B3790.fasta | 2006 | Clinical | Cerebro-Spinal Fluid | susceptible | |
| 10  | C4140      | SAMN13108486 | PRJNA579483 | WICO00000000 | C4140.fasta | 2007 | Clinical | Stool | susceptible | |
| 11  | B1988      | SAMN13108419 | PRJNA579483 | WIK00000000 | B1988.fasta | 2007 | Clinical | Stool | susceptible | |
| 12  | H1104      | SAMN13110852 | PRJNA579483 | WICY00000000 | H1104.fasta | 2007 | Clinical | Stool | susceptible | |
| 13  | B5147      | SAMN13108487 | PRJNA579483 | WICQ00000000 | B5147.fasta | 2008 | Clinical | Blood | susceptible | |
| 14  | G604       | SAMN13110859 | PRJNA579483 | WICU00000000 | G604.fasta | 2008 | Clinical | Stool | NAL | Asp87Asn |
| 15  | C1248      | SAMN13108488 | PRJNA579483 | WICO00000000 | C1248.fasta | 2009 | Clinical | Blood | NAL Pef | Ser83Phe |
| 16  | ED1075     | SAMN13110877 | PRJNA579483 | WICT00000000 | ED1075.fasta | 2009 | Clinical | Blood | NAL Pef | Asp87Asn |
| 17  | H1349      | SAMN13110860 | PRJNA579483 | WICX00000000 | H1349.fasta | 2009 | Clinical | Blood | susceptible | |
| 18  | H175       | SAMN13110878 | PRJNA579483 | WCK00000000 | H175.fasta | 2010 | Clinical | Blood | NAL Pef | Ser83Phe |
| 19  | H444       | SAMN13110884 | PRJNA579483 | WICW00000000 | H444.fasta | 2010 | Clinical | Blood | NAL | p. |
| 20  | C1458      | SAMN13108489 | PRJNA579483 | WICO00000000 | C1458.fasta | 2010 | Clinical | Blood | susceptible | |
| 21  | J316       | SAMN13110853 | PRJNA579483 | WIDA00000000 | J316.fasta | 2010 | Clinical | Blood | AMP NAL | p. |
| 22  | C2644      | SAMN13108490 | PRJNA579483 | WICO00000000 | BK-C2644.fasta | 2012 | Clinical | Blood | susceptible | |
| 23  | B480       | SAMN13108491 | PRJNA579483 | WICA00000000 | BK-B480.fasta | 2013 | Clinical | Stool | NAL Pef | Asp87Asn |
| 24  | NP294      | SAMN13110854 | PRJNA579483 | WICF00000000 | BK-NP294.fasta | 2013 | Clinical | Urine | NAL | p. |
| 25  | W355       | SAMN13110881 | PRJNA579483 | WICW00000000 | BK-W355.fasta | 2013 | Clinical | Blood | NAL Pef | Asp87Thr |
| 26  | B2464      | SAMN13110883 | PRJNA579483 | WICB00000000 | BK-B2464.fasta | 2013 | Clinical | Stool | NAL Pef | Asp87Thr |
| 27  | p573       | SAMN13108492 | PRJNA579483 | WICG00000000 | BK-PL573.fasta | 2014 | Clinical | Stool | susceptible | |
| 28  | F74/14     | SAMN13110858 | PRJNA579483 | WICE00000000 | BK-F74.fasta | 2014 | Clinical | Blood | NAL Pef | Ser83Thr |
| 29  | CR268      | SAMN13108493 | PRJNA579483 | WICD00000000 | BK-CR268.fasta | 2015 | Clinical | Blood | NAL Pef | p. |
| 30  | S2         | SAMN13110855 | PRJNA579483 | WICH00000000 | S2.fasta | 2009 | Animal | Chick | susceptible | |
| 31  | S10        | SAMN13110856 | PRJNA579483 | WICD00000000 | S10.fasta | 2014 | Animal | Poultry | NAL | p. |
| 32  | S12        | SAMN13110879 | PRJNA579483 | WICO00000000 | S12.fasta | 2014 | Animal | Poultry | NAL Pef | |
| 33  | S19        | SAMN13110880 | PRJNA579483 | WICJ00000000 | S19.fasta | 2009 | Animal | Chick | susceptible | |
| 34  | S29        | SAMN13110885 | PRJNA579483 | WICD00000000 | S29.fasta | 2009 | Animal | Chick | susceptible | |
| 35  | 1922       | SAMN13110889 | PRJNA579483 | WIBQ00000000 | 1922.fasta | 2014 | Animal | Poultry | susceptible | |

(continued on next page)
(11.7%), amino acid transport and metabolism [E] (9.46%), transcription [K] (9.17%), cell wall/membrane/envelope biogenesis [M] (6.88%) and energy production and conversion [C] (6.59%) (Fig. 2).

3.2. In vitro and in silico assessments of antimicrobial resistance

Out of 45 isolates, 15 exhibited antimicrobial resistance (Table 1, Fig. 1). Among all these resistant isolates, only four were strictly resistant to nalidixic acid, ten showed resistance to nalidixic acid and intermediate resistance to pefloxacin, and one clinical isolate was resistant to nalidixic acid and ampicillin (Fig. 1). The antimicrobial resistance gene results showed that 15 resistant isolates harbored single nucleotide polymorphism in the gene gyrA: Asp87Thr (n = 6), Ser83Phe (n = 2), Ser83Tyr (n = 1) and Asp87Asn (n = 6) (Fig. 1). No other known quinolone resistance mutations were detected. Only one isolate contained a marker for β-lactamase blaTEMP-1β gene conferring ampicillin resistance. No antimicrobial resistance genes have been identified in the genome of the susceptible isolates. These results agreed with those

Table 1 (continued)

| No. | Tree Label | biosample | Bioproject | genome | filename | Isolation date | Isolation type | Isolation source | Resistance patterns | Antimicrobial resistance genes |
|-----|------------|-----------|------------|--------|----------|----------------|----------------|-------------------|----------------------|--------------------------------|
| BK-1922 | 153 (F) | SAMN13110886 | PRJNA579483 | WIBP00000000 | BK-153. fasta | 2014 | Animal | Poultry | NAL Pef | gyrA: Asp87Asn |
| BK-3174 | 3174 | SAMN13110933 | PRJNA579483 | WIBX00000000 | BK-3174. fasta | 2007 | Foodborne | susceptible |
| BK-3156 | 3156 | SAMN13110887 | PRJNA579483 | WIBT00000000 | BK-3156. fasta | 2007 | Foodborne | susceptible |
| BK-3157 | 3157 | SAMN13110890 | PRJNA579483 | WIBU00000000 | BK-3157. fasta | 2007 | Foodborne | susceptible |
| BK-3220 | 3220 | SAMN13110940 | PRJNA579483 | WIBY00000000 | BK-3220. fasta | 2007 | Foodborne | susceptible |
| BK-3222 | 3222 | SAMN13110941 | PRJNA579483 | WIBZ00000000 | BK-3222. fasta | 2007 | Foodborne | susceptible |
| BK-3155 | 3155 | SAMN13110882 | PRJNA579483 | WIBS00000000 | BK-3155. fasta | 2007 | Foodborne | susceptible |
| BK-3173 | 3173 | SAMN13110895 | PRJNA579483 | WIBW00000000 | BK-3173. fasta | 2007 | Foodborne | susceptible |
| BK-3171 | 3171 | SAMN13110893 | PRJNA579483 | WIBV00000000 | BK-3171. fasta | 2007 | Foodborne | susceptible |
| BK-3152 | 3152 | SAMN13110857 | PRJNA579483 | WIBR00000000 | BK-3152. fasta | 2007 | Foodborne | susceptible |

Fig. 1. Maximum likelihood tree of 45 Salmonella Enteritidis genomes with reference strain AM933172.1. Branch labels represent isolate number_isolation type_isolation date. Heatmaps showing the presence and absence of SPI, prophage and antimicrobial resistance with a white box indicates absence while a black box shows presence. Point mutations in gyrA and SPI profile are indicated by color as shown in the legend.
obtained from the phenotypic analysis.

Phylogenetic tree analysis demonstrated that resistant isolates were separated in two lineages. The six isolates carrying the amino acid change Asp87Thr were grouped in the clade C2–5. The amino acid change from serine to phenylalanine at the site 83 were detected in two isolates grouped in C2–3. At the same site, the singleton isolate showed

Fig. 2. Functional classification of mutated orthologues genes of 45 *Salmonella* Enteritidis genomes with Clusters of Orthologous Groups COGs. Each functional group is indicated by color as shown in the legend.

Fig. 3. Maximum likelihood tree of 45 *Salmonella* Enteritidis genomes with reference strain AM933172.1. Branch labels represent isolate number_Isolation type_isolation date. First heatmap showing the presence and absence of pef and sef genes. Second and third heatmap showing the presence and absence of missense mutations in fimbriae, SPI-1, SPI-2 and phoPQ genes with a white box indicates absence and a black box shows presence.
the variation from serine to tyrosine. Finally, six isolates having the mutation at the site 87 from aspartic acid to asparagine were assembled in lineage B (Fig. 2).

3.3. In silico assessments of virulence genes

To elucidate genomic features of virulence, we used the SPIFinder database. This analysis identified 11 out of 23 previously described SPIs (Fig. 1). All 45 S. Enteritidis carried the same six SPIs, including SPI-1, SPI-2, SPI-3, SPI-9, SPI-13 and CS54. The SPI-14 was detected in all isolates except one. Eight profiles of SPIs, arbitrarily designated as P1 to P8, were identified with two major profiles P3 and P4 (14 isolates for each) which differ at the level of SPI-10. These profiles shared SPI-1_SPI-2_SPI-3_SPI-5_SPI-9_SPI-13_SPI-14_CS54 (Fig. 1).

WGS data was used to screen SPI-1 and SPI-2 genes from 45 S. Enteritidis. SPI-1 and SPI-2 genes were common to all isolates except one clinical strain isolated in 2000 and lacking the spiC and sseA genes. In addition, the phoP and phoQ genes, encoded the phoP/Q system, were detected in all our isolates.

All S. Enteritidis isolates harbored 12 fimbrial operons agf, bcf, fim, lpf, peg, saf, sbh, std, ste, stf, sth and stk. Conversely, sta, stg, stk, stc and stj and tcf genes were absent in all isolates. It is noticeable that fimbrial adherence operon sef was detected in all isolates except five. Moreover, four isolates showed the absence of one or two sef operon genes (Fig. 3).

Finally, the virulence plasmid genes (spvABCDR) involved in intramacrophage survival, plasmid-encoded fimbriae (pefABCD), transfer gene (traAVKELY) and resistance to complement killing (rck) were present in all our isolates except two (Fig. 4). None of these plasmids contained genes associated with antimicrobial resistance. The expression of several genes on the virulence plasmid that contribute to efficient systemic infections was regulated by alternative sigma factor RpoS. This factor across all isolates.

3.4. Mutations of the target virulence genes

We identified 15 synonymous, 20 missense and 7 nonsense mutations among the selected target genes of SPI-1 and SPI-2 genes, fimbrial adherence genes and virulence plasmid genes. The most commonly mutations were changes between G and A (15 times), C and T (15 times) and C and A (9 times); the less frequent changes were those between T and A (2 times) and T and G (1 time). Regarding the mutations in each gene, we identified four missenses SNPs in fimbrial adherence genes (Fig. 3). For one isolate of lineage A, the fimI gene was found to have one missense mutation resulted in an acid amine change of Arginine to cysteine at the site 50. The missense mutations in lpfA and stkA genes, at the site 111 (from threonine to isoleucine) and site 435 (from tyrosine to aspartic acid), respectively, were detected across all the isolates of lineage B (Fig. 3; Table 2).

Among SPI-1 and SPI-2 genes, we found five genes mutated with missense SNPs. The phylogenetic analysis demonstrated that two clinical isolates of lineage A carried a missense mutation in the ttvA gene (from alanine to Aspartate) at the site 950 (Fig. 3; Table 2). For the lineage B isolates, four missense mutations have been identified in invH (Ser146Ala), ttvA (Arg232Cys), ssQ (Thr28Ala) and ssN (His89Gln)
No missense mutations were observed for the other target alternative sigma factor have been identified in genes that code for plasmid transfer missense mutation in (Fig. 4). The 43 carried missense mutation at the site 197 (form valine to alanine) site 364. Another mutation at the site 104 in the gene lineage A carried mutation in the

virulence genes among the isolates studied.

Table 2: Non-synonymous mutations determined in target virulence genes in 45 Salmonella Enteritidis isolates.

| Genes            | Nucleotide position | Nucleotide change | Protein position | Amino acid change | Associated isolates |
|------------------|--------------------|-------------------|------------------|------------------|--------------------|
| chr_genome fimbria | bfpC                | 28,045            | C > A            | 826              | Ala -> Asp         | 17 isolates (13 Human, 1 Animal, 3 Foodborne) |
|                  | fimI                | 584,376           | C > T            | 50               | Arg -> Cys         | 1 Foodborne isolate |
|                  | fliA                | 3,708,664         | G > A            | 111              | Thr -> Ile         | 18 isolates (13 Human, 1 Animal, 4 Foodborne) |
|                  | sth                 | 208,100           | A -> C           | 435              | Tyr -> Asp         | 18 isolates (13 Human, 1 Animal, 4 Foodborne) |
|                  | invH                | 2,929,483         | T > G            | 475              | Ser -> Leu         | 18 isolates (13 Human, 1 Animal, 4 Foodborne) |
|                  | trrA                | 1,760,475         | C > T            | 232              | Arg -> Cys         | 18 isolates (13 Human, 1 Animal, 4 Foodborne) |
|                  | ssaO                | 1,732,159         | T > C            | 28               | Thr -> Ala         | 17 isolates (13 Human, 1 Animal, 3 Foodborne) |
| phoQ-regulated genes | ssaN            | 1,733,278         | A -> C           | 89               | His -> Gln         | 8 isolates (7 Human, 1 Animal) |
|                  | PhoP                | 1,919,806         | A > G            | 156              | Tyr -> Cys         | 1 Human isolate |
|                  | PhoQ                | 1,921,103         | A > C            | 364              | Ser -> Arg         | 1 Human isolate |
|                  | PhoQ                | 1,921,325         | A > G            | 438              | Thr -> Ala         | 18 isolates (13 Human, 1 Animal, 4 Foodborne) |
|                  | PhoP                | 1,919,418         | C > T            | 27               | Gln -> codon       | 8 Human isolates |
| Alternative sigma factor | rpoS            | 2,950,320         | C > T            | 309              | Gly -> Asp         | 1 Human isolate |
|                  | rpoS                | 2,950,824         | C > A            | 141              | Arg -> Leu         | 1 Human isolate |
|                  | rpoS                | 2,950,839         | G > A            | 136              | Pro -> Leu         | 1 Animal isolate |
|                  | rpoS                | 2,950,909         | C > T            | 113              | Gly -> Arg         | 1 Human isolate |
|                  | rpoS                | 2,950,330         | G > A            | 306              | Gln -> codon       | 1 Human isolate |
|                  | rpoS                | 2,950,438         | G > A            | 270              | Gln -> codon       | 1 Human isolate |
|                  | rpoS                | 2,950,803         | C > T            | 148              | Thr -> codon       | 1 Foodborne isolate |
|                  | rpoS                | 2,951,020         | C > A            | 76               | Gln -> codon       | 1 Foodborne isolate |
|                  | rpoS                | 2,951,092         | G > A            | 52               | Gln -> codon       | 1 Human isolate |
| Virulence plasmid genes | pefB              | 17,597            | G > A            | 96               | Val -> Ile         | 1 Human isolate |
|                  | pefC                | 19,667            | T > C            | 343              | Val -> Ala         | All isolates (43 isolates) |
| Transfer gene     | traE                | 35,699            | T > A            | 179              | Am -> Ile          | 24 isolates (16 Human, 3 Animal, 5 Foodborne) |
|                  | traY                | 37,011            | G > A            | 68               | Thr -> Ile         | 1 Human isolate |

For the virulence plasmid genes, we identified missense mutation (Tyr156Cys) in one clinical isolate of the lineage A and one missense mutation (Gln27*) in one clinical isolate of the lineage B. For the phoQ gene, we identified missense mutation (from threonine to alanine) at the site 438 in all isolates of lineage B. In addition, one clinical isolate of the lineage A carried mutation in the phoQ (from Serine to arginine) at the site 364. Another mutation at the site 104 in the gene phoQ, resulting a change from tryptophane to codon stop was detected in one clinical isolate of lineage B (Fig. 3; Table 2).

For the virulence plasmid genes, one clinical isolate of lineage B carried missense mutation at the site 197 (form valine to alanine) (Fig. 4). The 43 S. Enteritidis carried virulence plasmid genes carried a missense mutation in pefC gene (Val343Ala). Other missense mutations have been identified in genes that code for plasmid transfer traE (Asn179Ile) and traY (Thr68Ile). Two mutation types were detected in alternative sigma factor RpoS, including four missense and five nonsense (Fig. 4). No missense mutations were observed for the other target virulence genes among the isolates studied.

3.5. Prophages in S. enteritidis

A total of 21 prophage regions were detected using the PHASTER prophage analysis web server. Five out of these different phages were intact lysogenic phages (Salmon_118,970 sal3; Gifsy_2; Salmon_Fels_1; Gifsy_1 and salmon_re_2010) (Fig. 1). 11 prophage profiles were identified, the most prevalent profiles were Salmon_118,970 sal3 (n = 7); Salmon_RE_2010 (n = 6); Salmon_118,970 sal3-Salmon_Fels_1 (n = 6) and Gifsy_2-Salmon_RE_2010 (n = 5). No intact prophage sequence was detected in six isolates. We identified only one phage sequence in nine isolates. The genomes of the other isolates carried two to three prophage sequences. The most prevalent intact prophages detected in our collection of S. Enteritidis were Salmon_118,970 sal3 and Fels-2 which were present in 20 and 18 genomes, respectively. The prophages Gifsy_2 and Salmon_Fels_1 were defined in 15 isolates, while Gifsy_1 was present in only two isolates. Phylogenetic analysis based on the SNPs of all isolates showed that Salmon_118,970 sal3 and Salmon_RE_2010 were specific for the lineage A and lineage B isolates, respectively while three prophages (Gifsy_2, Salmon_Fels_1 and Gifsy_1) were common among the two lineages (Fig. 1).

4. Discussion

In the present study, a collection of 45 S. Enteritidis isolates were studied through WGS and subsequent in silico to determine the distribution of virulence genes, antimicrobial resistance genes, prophage sequences and the sequence variation of target virulence genes. The use of WGS became the most powerful tool for determining genomic variation especially in homogeneous bacterial genome. Several
studies have demonstrated the homogeneity of S. Enteritidis (Allard et al., 2012; Graham et al., 2018). In the current report, we have identified the existence of two lineages of S. Enteritidis circulating in Tunisia based on 802 SNPs whole-genome. Among these SNPs, 493 missense SNPs were detected. The COG distribution predicted that 349 orthologous genes mutated by one or two missense SNPs were classified into 22 functional groups. The major group was carbohydrate transport and metabolism [G] followed by amino acid transport and metabolism [E] and transcription [K]. Previous studies indicated that the orthologous genes of the majority of Salmonella serovars including Enteritidis are mainly classified in the transport and metabolism of glucid [G], transport and metabolism of amino acids [E], and transcription [K]. Zhou et al. (2018), reported that 1622 missenses SNPs distributed in 928 ORFs were detected in Rissen and Typhi serovar genomes. The distribution of the mutated orthologous genes showed that the majority were classified in the defense mechanisms [V], intracellular passage, secretion and vesicular transport [U] (Zhou et al., 2018). This classification differs from one serovar to another. The most abundant COG functional group was carbohydrate transport and metabolism [G], translation in the Salem et al., 2017; Wasyl et al., 2014). In addition, to single point mutations, a correlation between genotypic antibiotic resistance profiles and epidemiological subtypes of S. Enteritidis. Allard et al., 2012 exhibit the capacity of prophage RE_2010 to separate S. Enteritidis isolates with the same PFGE pattern. Furthermore, Ogunremi et al. (2014) proposed that the prophage composition could differentiate between S. Enteritidis subtypes during foodborne outbreaks (Ogunremi et al., 2014).

Many virulence genes of Salmonella enterica are organized on SPIs. 23 SPIs are identified in genus Salmonella, playing a fundamental role in pathogenesis and host specificity (Zhao et al., 2020). Suez et al. (2013) showed that S. Enteritidis reference strain P125109 contains 14 SPI (SPI-1 to 5, 9, 11, 14, 16, 17, 19 and cs54) (Suez et al., 2013). In our study, the investigation of 45 S. Enteritidis isolates showed that six SPIs, including SPI-1, SPI-2, SPI-3, SPI-9, SPI-13 and CS54 are conserved in all isolates while SPI-11 and SPI-12 were absent. Many studies suggest the universal presence of SPI-1 to SPI 5, SPI-13 and SPI-14, the absence of SPIs 7, 8 and 15 in all non-typhoidal Salmonella isolates, and the mosaic presence of SPIs 6, 10–12 and 16–19 across the serovars (Suez et al., 2013). The characterization of several genes carried by SPI-1 and SPI-2, encoding type III secretion systems which is important for the infection of the host cell and facilitate intracellular survival and replication, were highly conserved in all isolates except one. Previous studies reported that these genes are probably part of the core genes with an essential function for Salmonella serovars and the absence of certain genes could be explained by the possibility of losing the gene during their evolution (Ben Hassena et al., 2021; Suez et al., 2013). In our study, all tested isolates were positive for phoP and phoQ genes which are essential for the regulation of type III secretion systems. PhoP is the transmembrane sensor histidine kinase of the bacterial and phoQ is cytoplasmic regulatory protein (Hu et al., 2021).

Among virulence factors, fimbriae have a major role in pathogenesis and a source of diversity for Salmonella serovars. This factor is one of the most common adhesion systems and are differentially expressed and found in a specific pattern among each serovar (Dufresne and Daigle, 2017). In our analysis, the fimbrial operons agf, bcf, fim, lpf, peg, saf, stb, std, ste, sfc, sth and sti were present in all S. Enteritidis genomes. Suez et al. (2013) showed the same results and indicated that five fimbrial clusters (bcf, agf, stb, sth and sti) were detected to be part of core genome for invasion and systemic disease in humans (Suez et al., 2013). In addition, it has been reported that the five Chaperone-Usher Pathway (CUP) fimbriae sta, stg, ste, tcf and sfe were considered to be specific for typhoidal Salmonella (Dufresne et al., 2018; Forest et al., 2007; Suez et al., 2013). However, our study showed that sfe and ste are not specific for typhoidal Salmonella as we detected the presence of ste operon in all isolates and the sfe operon in 40 isolates. These findings agree with the study of Ben Hassena et al. (2021) that reported the presence and specificity of the sfe operon for the Enteritidis serovar (Ben Hassena et al., 2021).

Virulence-associated plasmids are important genetic elements in Salmonella. They are required for bacterial multiplication in the reticuloendothelial system, which confers to Salmonella the ability to produce systemic disease. In our previous study, the plasmid characterization showed that among 45 isolates 43 carried the replicons incFIB and incFII belonging to FAB type (S1A-B22). This FAB type was previously found to be associated with S. Enteritidis virulence plasmid (Silva et al., 2017; Villa et al., 2010). This largely consistent with our finding that all S. Enteritidis isolates carrying the replicons incF contain pefABC, spvABCD, traAVKELY and rck.

Among the 45 S. Enteritidis isolates investigated in this study, both synonymous and non-synonymous mutations in the selected target virulence genes were detected. In plasmid genes, we note missense mutations in pefC gene (Val343Ala) and in transfer genes as traE (Asn179Ile) and tarY (Thr68Ile). No missense mutations were detected in the spv and rck genes. The spv genes were reported to be highly conserved in Salmonella (Hu et al., 2021). In our study, we observed missense and nonsense variability in the regulated rpoS and phoPQ genes. Hu et al. (2021) observed that cytoplasmic regulatory protein phoP gene has been frequently used for Salmonella detection as it has a conserved sequence
Cingolani, P., Platts, A., Wang, L.L., Coon, M., Nguyen, T., Wang, L., Land, S.J., Lu, X., Ben Hassena, A., Haendiges, J., Zormati, S., Guermazi, S., Gdoura, R., Gonzalez-Aldrich, C., Hartman, H., Feasey, N., Chattaway, M.A., Dekker, D., Al-Emran, H.M., obtained by phenotypic and genotypic analysis were observed. Our.

Tunisia
tion among

Salmonella

lpfA

Q (Thr28Ala),

ssa

(Thr111Ile),

–

Salmonella

disease in Ghana. PLoS Negl. Trop. Dis.

–

597

serovar

M., Litrup, E., 2019. WGS based study of the population structure of

Salmonella enterica

019-6260-6 .

-2

ampicillin resistance in zoonotic

Salmonella

Ø., Woodford, N., 2017. The role of whole genome sequencing in

Salmonella enterica

Ørensen, G., Olsen, J.E., Nielsen, E.M., Litrup, E., 2019. WGS based study of the population structure of

Salmonella enterica

serovar

infantis. BMC Genomics 20, 870. https://doi.org/10.1186/s12864-020-06269-4.

Hu, L., Gao, G., Brown, E.W., Allard, M.W., Ma, L.M., Zhang, G., 2021. Whole genome sequencing and protein structure analyses of target genes for the detection of

Salmonella. Sci. Rep. 11, 20887. https://doi.org/10.1038/s41598-021-02024-7.

Kbbi, B., Kuari, S., Othman, H., Ksibi, B., Ktari, S., Malej, S., Mnif, B., Abbassi, M salah, Fabre, L., Rhimi, F., Le Hello, S., Hammami, A., 2020. Comparison of conventional molecular and whole-genome sequencing methods for subtyping

Salmonella enterica

serovar Enteritidis strains from Tunisia. Eur. J. Clin. Microbiol. Infect. Dis. 40, 5097–5105. https://doi.org/10.1007/s10096-020-04055-8.

Ktari, S., Kbbi, B., Ghedira, K., Fabre, L., Bertrand, S., Malej, S., Mnif, B., Rhimi, F., Le Hello, S., Hammami, A., 2020. Genetic diversity of clinical

Salmonella enterica

serovar Typhimurium in a university hospital of south Tunisia from 2000–2013. Infect. Genet. Evol. 85, 104436. https://doi.org/10.1016/j.meegid.2020.104436.

Mastrorilli, E., Petrin, S., Orsini, M., Longo, A., Cozza, D., Luzzi, I., Ricci, A., Barco, L., Losasso, C., 2020. Comparative genomic analysis reveals high intra-serovar plasticity within

Salmonella Napoli isolated in 2005–2017. BMC Genomics 21, 202. https://doi.org/10.1186/s12864-020-06588-y.

Matsumoto, Y., Izuimya, H., Sekizuka, T., Kuroda, M., Oharishi, M., 2014. Characterization of

β-lactamase-producing

Salmonella enterica

serovar Typhi. Front. Cell. Infect. Microbiol. 8, 26. https://doi.org/10.3389/fcimb.2018.00026.

Mares, M. (Ed.), Current topics in

Salmonella and Salmonellosis. InTech.

Ørensen, G., Olsen, J.E., Nielsen, E.M., Litrup, E., 2019. WGS based study of the population structure of

Salmonella enterica

serovar

infantis. BMC Genomics 20, 870. https://doi.org/10.1186/s12864-020-06269-4.

Ørensen, G., Olsen, J.E., Nielsen, E.M., Litrup, E., 2019. WGS based study of the population structure of

Salmonella enterica

serovar

infantis. BMC Genomics 20, 870. https://doi.org/10.1186/s12864-020-06269-4.

Salmonella

Ørensen, G., Olsen, J.E., Nielsen, E.M., Litrup, E., 2019. WGS based study of the population structure of

Salmonella enterica

serovar

infantis. BMC Genomics 20, 870. https://doi.org/10.1186/s12864-020-06269-4.

Salmonella

Ørensen, G., Olsen, J.E., Nielsen, E.M., Litrup, E., 2019. WGS based study of the population structure of

Salmonella enterica

serovar

infantis. BMC Genomics 20, 870. https://doi.org/10.1186/s12864-020-06269-4.

Ørensen, G., Olsen, J.E., Nielsen, E.M., Litrup, E., 2019. WGS based study of the population structure of

Salmonella enterica

serovar

infantis. BMC Genomics 20, 870. https://doi.org/10.1186/s12864-020-06269-4.
late 1950s: a retrospective, whole-genome sequencing study. Lancet Infect. Dis. 18, 207–214. https://doi.org/10.1016/S1473-3099(17)30765-5.

Villa, I., García-Fernández, A., Fortini, D., Carattoli, A., 2016. Replicon sequence typing of IncF plasmids carrying virulence and resistance determinants. J. Antimicrob. Chemother. 65, 2518–2529. https://doi.org/10.1093/jac/dkg347.

Wahl, A., Battesti, A., Ansaldi, M., 2019. Prophages in Salmonella enterica: a driving force in reshaping the genome and physiology of their bacterial host? Mol. Microbiol. 111, 303–316. https://doi.org/10.1111/mmi.14167.

Wasyl, D., Hoszowski, A., Zającz, M., 2014. Prevalence and characterisation of quinolone resistance mechanisms in Salmonella spp. Vet. Microbiol. 171, 307–314. https://doi.org/10.1016/j.vetmic.2014.01.040.

Zankari, E., Allesøe, R., Joensen, K.G., Cavaco, L.M., Lund, O., Aarestrup, F.M., 2017. PointFinder: a novel web tool for WGS-based detection of antimicrobial resistance associated with chromosomal point mutations in bacterial pathogens. J. Antimicrob. Chemother. 72, 2764–2768. https://doi.org/10.1093/jac/dks217.

Zankari, E., Hasman, H., Cosentino, S., Vestergaard, M., Rasmussen, S., Lund, O., Aarestrup, F.M., Larsen, M.V., 2012. Identification of acquired antimicrobial resistance genes. J. Antimicrob. Chemother. 67, 2640–2644. https://doi.org/10.1093/jac/dks261.

Zhou, Y., Liang, Y., Lynch, K.H., Dennis, J.J., Wishart, D.S., 2011. PHAST: a Fast Phage Search Tool. Nucleic Acids Res. 39, W347–W352. https://doi.org/10.1093/nar/gkr485.