Leonor de Sousa Dâmaso da Silveira

Mestre em Genética Molecular e Biomedicina

Phenotypic and genotypic characterization of *Salmonella* spp. isolates in Portugal

Dissertação para obtenção do Grau de Doutor em
Biologia, Especialidade de Genética Molecular

Orientador: João Paulo dos Santos Gomes, Investigador Auxiliar com Habilitação, INSA
Coorientador: Jorge Cândido Pinheiro da Costa Machado, Coordenador do Departamento de Doenças Infeciosas, INSA
Coorientador: Luís Jaime Gomes Ferreira da Silva Mota, Professor Auxiliar, FCT - UNL

Júri:

Presidente: Prof. Doutor Pedro Viana Batista
Arguentes: Prof. Doutor Mário Nuno Ramos de Almeida Ramirez
Prof. Doutora Paula Cristina Maia Teixeira

Vogais: Prof. Doutora Rita Gonçalves Sobral de Almeida
Prof. Doutor João Paulo dos Santos Gomes

FACULDADE DE CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

Dezembro 2019

Leonor de Sousa Dâmaso da Silveira

Mestre em Genética Molecular e Biomedicina

Phenotypic and Genotypic characterization of *Salmonella* spp. isolates in Portugal

Look up at the stars and not down at your feet. Try to make sense of what you see, and wonder about what makes the universe exist. Be curious. **Stephen Hawking**

Agradecimentos

Primeiro que tudo gostaria de agradecer ao Doutor João Paulo Gomes, a sua orientação e motivação constantes, muitas vezes nos momentos em que mais precisava. O teu apoio foi fundamental para conseguir completar este desafio. A tua confiança em mim e no meu trabalho fizeram-me sentir mais segura e a começar a acreditar mais naquilo que faço. Obrigada!

Ao Doutor Jorge Machado, meu coorientador, Coordenador e Responsável de Laboratório. Agradeço todo o apoio ao longo destes anos.

Ao Doutor Jaime Mota, por ter aceite ser meu coorientador interno, e por todo o apoio e disponibilidade demonstrados ao longo destes anos. Fica a faltar uma futura colaboração!

Aos restantes membros da Comissão de Acompanhamento de Tese, Doutora Mónica Oleastro e Doutora Isabel Couto por todas as sugestões.

Um agradecimento especial às minhas colegas do Laboratório de Infeções Gastrintestinais, Mónica Oleastro e Ângela Pista. Sem a vossa compreensão e apoio constante nada disto teria sido possível.

Aos restantes colegas do Departamento de Doenças Infeciosas, agradeço a motivação no decorrer deste trabalho.

Quero também agradecer aos colegas da Secção de Bioinformática, Alexandra Nunes, Vítor Borges e Miguel Pinto, pelo apoio técnico e não só. Pela companhia nas horas de almoço. Pela discussão de temas técnicos e pela distração nas horas vagas. Obrigada a todos!

Agradeço também à Raquel Rocha pelo seu companheirismo.

Aos colegas da Unidade de Tecnologia e Inovação por todo o apoio na implementação e otimização das técnicas de sequenciação cruciais para desenvolver deste trabalho.

À Faculdade de Ciências e Tecnologia da Universidade Nova de Lisboa, em especial à Professora Isabel Sá Nogueira, coordenadora do Programa Doutoral em Biologia. Toda a minha formação foi feita nesta casa e irei guardar sempre um carinho especial por esses tempos.

A todos meus amigos, por me distraírem quando mais precisava.

Um agradecimento do fundo do coração às minha miúdas, Andrea Santos, Joana Isidro e Rita Macedo. Nem consigo descrever por palavras o quanto vos agradeço. Os vossos conselhos, os vossos mimos, as nossas saídas e sleepovers, ajudaram-me a manter a sanidade mental e a manter-me focada no objetivo. Obrigada! Adoro-vos!

À minha família, de sangue e coração, e que adoro de paixão. Aos meus avós, por estarem sempre presentes, pelo vosso apoio e por serem um excelente exemplo de perseverança e sucesso na vida profissional. À minha avó Arminda, por ter sido tão especial na minha vida e continuar a olhar por mim. Aos meus tios e padrinhos (oficiais e emprestados), pela amizade e por serem os mais fixes de sempre. Adoro-vos a todos! Aos meus sogros e aos meus cunhados, agradeço todo o carinho e apoio. Aos meus miúdos: Catarina, Carolina, João, Francisco e Sofia, espero que lutem sempre por aquilo que querem e que cumpram sempre os vossos objetivos. Nunca deixem de acreditar nas vossas capacidades! Espero ser sempre um bom exemplo para vocês.

Um agradecimento muito especial aos meus pais. O seu apoio incondicional, os exemplos excecionais que são, e a educação que me proporcionaram, fizeram de mim o que sou hoje. E obrigada por me darem sempre guarida quando precisei.

Ao meu irmão, que apesar de longe, é parte do meu coração. Para sempre.

Finalmente, ao Pedro, acima de tudo pela paciência. Eu sei que não foi fácil, especialmente a última fase… obrigada pelo teu apoio, pelo teu carinho, por seres o meu companheiro de viagens e da vida. Por me ajudares a crescer, por seres a minha família, por todos os momentos de partilha. Espero que se sigam muitos mais. Amo-te.

Resumo

Salmonelose permanece uma das causas mais frequentes de infeção gastrintestinal no mundo e representa um enorme impacto económico não só em países em desenvolvimento, mas também em países desenvolvidos. Assim, a vigilância epidemiológica e laboratorial de *Salmonella* é essencial para o controlo da infeção e para a diminuição dos custos associados à doença. Desde 1950 que o Instituto Nacional de Saúde Doutor Ricardo Jorge tem tido um papel relevante na vigilância laboratorial desta doença.

O principal objetivo deste trabalho prendeu-se com a contribuição para o melhor conhecimento de alguns dos serotipos de *Salmonella* que circulam em Portugal, quer mais frequentes, quer serotipos com características vantajosas para a sua perseverança. De forma a cumprir este objetivo, primeiramente foram analisados os dados acumulados ao longo de décadas de vigilância laboratorial e foram determinadas as tendências dos serotipos circulantes no país. Foi também implementada, no Laboratório Nacional de Referência, a sequenciação genómica aplicada à vigilância de *Salmonella enterica*, e foi feita a avaliação da substituição das técnicas clássicas de tipagem. A sequenciação genómica, como método altamente discriminatório permitiu, não só a confirmação de surtos previamente identificados, como também a sinalização de potenciais surtos não sinalizados pelas autoridades de saúde. Neste trabalho, identificámos também vários isolados com múltiplos marcadores de resistência e diversos elementos genéticos móveis que conferem a capacidade de persistência e disseminação a estas bactérias.

Esta dissertação destaca a necessidade de revisão da regulamentação associada a géneros alimentícios e de estabelecimento de novas medidas de controlo de infeção na produção animal, de maneira a evitar a disseminação de marcadores de multirresistência e persistência, que contribuem grandemente para o aumento da severidade das infeções causadas por *Salmonella*, e também da disseminação dessas características altamente transmissíveis. Este trabalho representa uma importante contribuição para o conhecimento dos isolados de *Salmonella enterica* circulantes em Portugal, e também para o reforço da capacitação do Laboratório Nacional de Referência.

Palavras-chave

Salmonella enterica, Multirresistência, Sequenciação de genoma completo, Vigilância, Surto

Abstract

Salmonellosis remains one of the most common causes of gastrointestinal infection in the world, and represents a considerable burden in developing and developed countries. As such, *Salmonella* surveillance is essential to control infection and decrease the economic burden of the disease, and the Portuguese National Institute of Health has been working in this field roughly since 1950.

Considering the importance of the surveillance of *Salmonella enterica*, the main goal of this work was to contribute to a better knowledge of relevant serovars circulating in Portugal, either the most common, or serovars exhibiting important fitness traits that confer ecological advantages over other strains. In order to achieve this objective, first we analysed decades of data and established the trends of *Salmonella enterica* serovars circulating in Portugal. Additionally, whole-genome sequencing (WGS) was implemented at the National Reference Laboratory to the surveillance of *Salmonella enterica*, and was evaluated as a substitute of traditional typing methods, such as serotyping and multiple-locus variable number tandem repeat analysis. WGS proved to be a highly discriminatory method for the detection of suspected outbreaks and even previously undetected clusters. We identified several isolates carrying multiple resistance markers and mobile genetic elements that confer the potential of persistence and spread of these bacteria.

This PhD dissertation highlights the need for the adjustment of the existing regulations for foodstuffs monitoring and the investigation of safe control measures in animal production, as a means to control the spread of resistant and persistent clones that may contribute to the increase of severe infections and to the spread of those markers to other bacteria. This work stands as an important contribution for the insight on *Salmonella enterica* isolates circulating in Portugal, as well as a contribution for the reinforcement of the capacitation of the Portuguese National Reference Laboratory.

Keywords

Salmonella enterica, Multidrug resistance, Whole-genome sequencing, Surveillance, Outbreak

Table of contents

Figure index

Table index

List of abbreviations

% G+C – Percentage of guanine and cytosine

AD – Allele differences

bp – Base pairs

BC – Before Christ

CGE – Centre for Genomic Epidemiology

CLED agar - Cystine lactose electrolyte deficient agar

DNA – Deoxyribonucleic Acid

EC – European Communion

ECDC – European Centre for Disease Prevention and Control

EPIS-FWD – Epidemic Intelligence Information System for food- and waterborne diseases

ENA – European Nucleotide Archive

ESBLs – Extended-spectrum β-lactamases

EU – European Union

EUCAST – European Committee on Antimicrobial Susceptibility Testing

EU/EEA – European Union/European Economic Area

GN broth – Gram-Negative broth

HGT –Horizontal gene transfer

INIAV – Instituto Nacional de Investigação Agrária e Veterinária

iNTS – Invasive non-typhoidal *Salmonella*

INSA – Instituto Nacional de Saúde Doutor Ricardo Jorge

IS – Insertion sequence

LPS – Lipopolysaccharides

MALDI-TOF MS – Matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry

Mb – Megabase

MDR – Multidrug resistant

MLST – Multilocus sequence typing

MLVA – Multiple-locus variable number tandem repeat analysis

MST – Minimum spanning tree

NGS – Next-generation sequencing

NIH – National Institute of Health

NRL – National Reference Laboratory

NTS – Non-typhoidal *Salmonella*

PBP – Penicillin Binding Protein

PCR – Polymerase chain reaction

PFGE – Pulsed-field gel electrophoresis

PHE – Public Health England

PMQR – Plasmid-mediated quinolone resistance genes

PT – Portuguese

QA/QC – Quality Assurance/Quality Control

RFLP – Restriction fragment length polymorphism analysis

rpm – Revolutions per minute

rRNA – Ribosomal RNA

SCV – *Salmonella*-containing vacuoles

SPI – *Salmonella* pathogenicity island

SNP – Single nucleotide polymorphism

ST – Sequence type

T3SS – Type III secretion systems

T4SS – Type IV secretion system

Tn – Transposon

tRNA – Transfer RNA

USA – United States of America

VNTR – Variable number tandem repeat

wgMLST – Whole-genome multi locus sequence typing

WGS – Whole-genome sequencing

WHO – World Health Organization

XLD agar – Xylose Lysine Deoxycholate agar

Notes of the author: thesis organization, format and outline

This PhD dissertation is composed of six chapters, including an Introduction, four research studies, and a final discussion. The main body is based on two studies that were already submitted to peer reviewed international journals, and are presented as individual chapters. In summary, the PhD dissertation includes:

Chapter 1

This chapter consists of an overview of the state of the art regarding the context of this dissertation. It describes the major aspects of *Salmonella spp*. taxonomy, a description of the transition from the classic typing methods to whole-genome sequencing (WGS), insights of *Salmonella* evolution and genome structure, and a summary of all aspects related to human disease, such as diagnostic, epidemiology and treatment. Finally, this chapter presents the general and specific aims of this PhD dissertation.

Chapter 2

This chapter gives insight on the surveillance of *Salmonella spp.* performed at the National Reference Laboratory of Gastrointestinal Infections of the National Institute of Health Doutor Ricardo Jorge, over the last 21 years. It is not presented in the form of a classical manuscript (i.e., with an Abstract, methods section, etc.) since it is not our intent to submit it for publication (in contrast with the subsequent chapters), and because it enrols methodologies that changed over time, and its description would be a cumbersome task and completely irrelevant taking the objective of the chapter into account.

Chapter 3

This chapter, also written in an unconventional manner, focuses on the practical advantages that WGS brought to the National Reference Laboratory in terms of *Salmonella* surveillance. Two studies are approached in this chapter, i) the transition from a traditional molecular typing method, Multiple-locus variable number tandem repeat analysis (MLVA), to WGS and what it entails for surveillance of one of the most common serovars in Portugal, *Salmonella enterica* serovar Enteritidis, ii) and an ongoing study of an outbreak of *Salmonella enterica* serovar Newport and the preliminary findings.

Chapter 4

This chapter is based on a manuscript ("*Silveira L., Pinto M., Isidro J., Pista A., Themudo P., Vieira L., Machado J. and Gomes J.P.. Multidrug resistant Salmonella enterica serovar Rissen clusters detected in Azores archipelago, Portugal*") published in the International Journal of Genomics. The main objective in this chapter was to identify potential epidemiological

clusters of *S. enterica* serovar Rissen linking samples from multiple sources, while gaining insight on the genetic diversity of Portuguese isolates.

Chapter 5

This chapter, also based on a manuscript ("*Leonor Silveira, Alexandra Nunes, Ângela Pista, Joana Isidro, Cristina Belo Correia, Margarida Saraiva, Rita Batista, Isabel Castanheira, Jorge Machado, João Paulo Gomes. Characterization of multidrug-resistant isolates of Salmonella enterica serovars Heidelberg and Minnesota from imported poultry meat in Portugal*", under review in Microbial Drug Resistance), consists in the phenotypic and genotypic characterization of isolates recovered from imported poultry meat.

Chapter 6

Finally, this chapter provides an overview of the main results of this PhD dissertation. Since Chapters 4, and 5 have their own Discussion and Conclusion sections, only the main results are discussed and the main conclusions highlighted, to avoid excessive redundancy. Also, in this section future perspectives are approached.

The chapters were all formatted in the same style, and the references are cited by sequential numbers and listed in a single "References" section. The supplemental tables are also presented in a section at the end of this dissertation.

CHAPTER 1

General introduction

1. General introduction

1.1. The genus *Salmonella* spp.

1.1.1. The discovery of Salmonella

Salmonella is a genus of rod-shaped Gram-negative intracellular facultative bacteria, within the *Enterobacteriaceae* family. Several Historical accounts point to *Salmonella* as the cause of death of many important personalities and civilizations. In 430 BC, the Plague of Athens decimated around 75,000 to 100,000 people during the second year of the Peloponnesian War (against Sparta) and had enormous political, economic and social ramifications, since it was the main reason for Athens losing the war. This outbreak spread from Ethiopia into Egypt, Libya, the Near East and finally the port of Athens. It was also, at that time, the cause of death of the Athenian leader Pericles in 429 BC[1]. A few years later, in 323 BC, Alexander the Great died at 33 years old, most likely due to typhoid fever[2]. In 2001, a mass grave with victims from the Plague was found, and the DNA of *Salmonella enterica* serovar Typhi was detected in three squeletons[3]. Just recently, in 2018, Nature Ecology & Evolution published a study that suggests that *S. enterica* serovar Paratyphi C was introduced in the New World by the European colonizers and was the cause of death of the Aztecs in the 1500s[4]. On November 11th 1861, the beloved Portuguese king D. Pedro V died at the age of 24 years old, from typhoid fever, along with two of his brothers, D. Fernando and D. João[5]. One month later, on December 14th 1861, his cousin, and Queen Victoria's husband, Prince Consort Albert of Saxe-Coburg and Gotha, died from the same illness[6].

The attribution of the discovery of *Salmonella* is still debatable since many were the scientists that contributed to the characterization of typhoid fever and its propagation. In 1873, William Bud, a medical doctor who studied cholera and intestinal fever, stated that typhoid fever was transmitted through faeces, and contaminated hands and water were the probable propagation routes. Bud indicated disinfection as a preventive measure of spread of this disease[7]. Later in 1879, Karl Joseph Eberth described a bacillus in the abdominal lymph nodes and in the spleen of patients with typhoid fever[8]. However, it was only in 1885 that this microorganism was isolated and identified by Theobald Smith as *Bacillus choleraesuis,* the etiologic agent of hog cholera[9,10]. Nevertheless, typhoid fever was still believed to be transmitted through polluted water, milk, putrefying organic matter or even sewer gas[11]. Later, during the 1900s, the work of sanitary engineer George Soper on the typhoid outbreaks caused by Mary Mallon cleared much of the mystery surrounding the transmission of typhoid fever. Mary Mallon was an Irish woman who worked as a cook in several private homes in New York. She was an asymptomatic carrier of *S. enterica* serovar Typhi and was responsible for several outbreaks and the contamination of at least 122

people before she was forcibly admitted in North Brother Island, where she lived until her death in 1938[12].

1.1.2. Taxonomy

The nomenclature of *Salmonella* is somewhat complex and suffered several changes through time. *Bacillus choleraesuis* was initially included in the group *Bacterium* or *Bacillus*, until 1900, Lignières proposed it should be considered a genus and named it *Salmonella*, after Theobald Smith's supervisor Daniel Salmon[9,13]. In 1931, Kauffman proposed one of the first *Salmonella* taxonomic classification methods, the Kauffmann-White Schema, based on the identification of cell wall (O) and flagellar (H) antigens, and according to which serovars were considered different species[14]. At that time, novel *Salmonella* species were given names according to, for example, the disease caused (*Salmonella typhi*), host specificity (*Salmonella abortus-ovis*), and later after the geographical area where the strain was first isolated (*Salmonella london*)[14–16]. By 1940, the number of *Salmonella* species was becoming unmanageable, hence Borman et al proposed the consideration of only three species, determined by serological methods: *S. choleraesuis* for the type species, *Salmonella typhosa* and *Salmonella kauffmannii*, containing all known serovars. This classification was ignored until 1952, when Kauffmann and Edwards proposed the consideration of three species according to their biochemical properties, *Salmonella choleraesuis*, *Salmonella typhosa* and *Salmonella enterica*[17,18]. Subsequent taxonomic classifications divided the serovars in four species or subgenera: *Salmonella kauffmannii*, subgenera I; *Salmonella salamae*, subgenera II; *Salmonella arizonae*, subgenera III; and *Salmonella houtenae*, or subgenera IV[16]. The concept of few species was well accepted, until 1973 DNA hybridization techniques demonstrated that all serovars, except for group V, *Salmonella bongori*, were related at species level[19]. The species name *Salmonella enterica* came up later on, in 1986, since *Salmonella choleraesuis* could easily be confused with the serovar *Salmonella enterica* serovar Choleraesuis[20].

Currently, the genus *Salmonella* comprises two species, *S. bongori* and *S. enterica*, the latter being divided in six subspecies: *enterica* (I), *salamae* (II), *arizonae* (IIIa), *diarizonae* (IIIb), *houtenae* (IV), and *indica* (VI)[15]. According to the White-Kauffman-Le Minor scheme, these subspecies are further divided into serovars[15]. More than 2600 serovars of *S. enterica* have been described to date while only 22 serovars of *S. bongori* are known[15].

1.2. From classical typing to genomic characterization

Typing is by definition the analysis of the phenotype and/or genotype of bacteria in a more specific level, and usually contributes to the detection of outbreaks, definition of transmission patterns, and determination of source of infection[21]. In the following points, a few of the most commonly used methods for typing of *Salmonella* will be described.

1.2.1. Classical typing

The White-Kauffman-Le Minor scheme has been the gold standard method for *Salmonella* serotyping for almost 90 years. It consists, as explained previously, on the identification of O and H antigens, by slide agglutination with specific antisera[15]. The O antigen is part of the lipopolysaccharide that composes the outer membrane of Gram-negative bacteria. There are sixty-seven O groups, 1 to 67, that basically vary in their chemical structure[15,22]. Some bacterial cells can express several O antigens simultaneously[23]. In addition, there are one hundred and fourteen flagellins, that differ in their epitopes[24]. Moreover, most *Salmonella enterica* serovars are diphasic, meaning they have the ability to express two different flagellar filament proteins, FljB and FliC, in a process, controlled by the Hin recombinase, known as flagellar phase variation[25]. All these variations make *Salmonella* serotyping a complex process, since serovars are determined by the combination of O, H1, and H2 antigens. Among several limitations, serotyping is not a discriminatory method when trying to differentiate species, subspecies, or strains, so additional testing is usually required. A typing technique often used to differentiate strains from specific serovars, namely *S. enterica* serovar Enteritidis and *S. enterica* serovar Typhimurium, was phage typing[23]. Phage typing is based on the capacity of certain bacteriophages to selectively infect strains, and consequently, a phage type is assigned upon the variety of phages that are able to infect a specific strain[26]. The complexity of the technique, which calls for active phage stocks, and the subjective interpretation of the results, eventually limited the use of this technique to a few reference laboratories[23,26].

1.2.2. Molecular typing

The development of molecular techniques, based on the genomic differences between bacterial genomes, allowed for the identification and fingerprinting of species, subspecies, and strains in general. Several techniques have been used for *Salmonella* molecular typing, amongst others, Pulsed Field Gel Electrophoresis (PFGE), Restriction Fragment Length Polymorphism analysis (RFLP), Multilocus Sequence Typing (MLST), and Multiple-Locus Variable number tandem repeat Analysis (MLVA)[23,26]. PFGE, used for many years for outbreak detection, is, in a simplistic manner, based on the restriction pattern of

the bacterial genome. The bacterial DNA, internalized in agarose, is isolated and fragmented by restriction enzymes, and then separated by alternating polarities, generating a restriction profile characteristic to each genome[26]. This technique is labour intensive and time-consuming and studies have now shown its discriminatory limitations[27–29]. RFLP is somewhat similar to PFGE, but after the electrophoresis, the fragments are transferred by southern blotting to a membrane with labelled DNA probes that hybridize with repeat motifs[26]. This method, like PFGE, is also currently obsolete. Contrarily, a technique still in use in the great majority of European laboratories is MLVA. MLVA is based on the premise that each strain has a distinct number of tandem repeats in various loci of specific regions of its genome (variable number tandem repeats, VNTR)[30]. Several schemes for many bacteria have been established, including *S. enterica* serovar Enteritidis and *S. enterica* serovar Typhimurium[31,32]. This technique is quite robust, simple, and has high reproducibility, among other advantages, making it an elected tool for worldwide outbreak detection to this day[33]. Finally, MLST is a typing technique based on the variations that naturally occur in the sequence of bacterial housekeeping genes. The most accepted scheme for *Salmonella* is based on 7 housekeeping genes, *aroC*, *dnaN*, *hemD*, *hisD*, *purE*, *sucA*, and *thrA*, and has a high correlation with serotype, making it an alternative to classical serotyping for some laboratories[28]. This molecular typing technique, presents lower resolution than PFGE and MLVA due to the fact that housekeeping genes have a slow mutation rate[35].

1.2.3. Genomic era

In the 1970's the first sequencing technologies started to arise, but the need for toxic and radioactive reagents limited their use, and when Sanger sequencing was developed it quickly prevailed in the following 30 years. Later, the demand for faster, higher throughput and a cheaper solution to determine the human genome sequence, in the context of the Human Genome Project, prompted the development of next-generation sequencing (NGS)[36]. During the 2010's, whole-genome sequencing (WGS) of bacterial pathogens started being used in several public health laboratories following its application in the resolution of two major outbreaks, the cholera outbreak in Haiti and the *Escherichia coli* O104:H4 outbreak[37–39].

1.2.3.1. Illumina sequencing technology

Throughout the years, several sequencing platforms, have been developed by different companies, differing in their method of immobilization of DNA and in the sequencing approach, sequencing by ligation or sequencing by synthesis[40,41]. However,

these second-generation technologies all share the capability for massively parallel analysis, generating an output of millions of short reads in a couple days, at a lower cost than first generation sequencing, and eliminating the need for electrophoresis upon sequencing[37,41]. The Illumina technology, the most popular in the market, and also used in this work, will be described in more detail. Illumina sequencing is based on a sequencing by synthesis approach[37]. It starts with the random fragmentation of DNA and the simultaneous ligation of adapters to both ends of each fragment (tagmentation)[41]. These adaptors allow ligation to a complementary sequence on the flow cell, PCR amplification, binding of sequencing primers, and indexing. After hybridization of forward and reverse primers to the flow cell, bridge amplification begins, in which millions of clusters of clonal sequences are synthetized[40]. Sequencing follows, with the incorporation of fluorescently labelled reversible terminators. These modified nucleotides are blocked at their 3' hydroxyl ends in order to guaranty that incorporation is achieved one nucleotide at a time[41]. During each cycle of incorporation, a laser excites the nucleotide and the signal is detected, followed by cleavage of the fluorescent tag and the restoration of the 3' hydroxyl end, which allows further incorporation of complementary nucleotides[42]. The output of this sequencing platform consists in millions of paired-end reads of 2x150 bp.

1.2.3.2. Application and bioinformatic tools for Salmonella typing

The major applications of NGS in infectious diseases genomic studies, are (i) the characterization of the pathogen, (ii) molecular epidemiology, and (iv) pathogenomics[43]. The importance of *Salmonella* as a foodborne human pathogen has made this bacteria one of the most studied pathogens by WGS, with many bioinformatic tools and pipelines being developed recently[44–46].

Identification of microorganisms by WGS in routine laboratories is still very unlikely given the costs when comparing to current diagnostic methods such as matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry (MALDI-TOF MS). Nonetheless, it can still be a useful tool to identify pathogens that are unidentifiable by traditional methods, unculturable or slow growth microorganisms, and to perform culture-independent metagenomics directly from clinical samples[47]. Specific features of *Salmonella* can be extracted from the high throughput data generated by NGS technologies using freely available online platforms, before or after assembly of the reads, depending on the tool[43,46]. Centre for Genomic Epidemiology (CGE) is an online platform where several tools for *Salmonella* typing are available. Serovar (SeqSero or SISTR) and ST (MLST) determination, identification of resistance markers (ResFinder or CARD), mobile genetic elements (PlasmidFinder, PHASTER), and *Salmonella* pathogenicity islands (SPIFinder) are useful

tools for *Salmonella* fingerprinting that can be used in place of traditional typing methods[48–54].

Comparative genomic analysis based on gene-by-gene approaches (extended MLST schemes), or single nucleotide polymorphisms (SNP), have not only been invaluable tools for infection control, outbreak detection and source attribution, but have also allowed to shed light on the micro and macroevolution processes of several *Salmonella* serovars[29,55–60]. Both approaches can be indexed to core or accessory genome, and easily visualized in software like PHILOViZ or GrapeTree[61,62]. When concerning outbreak detection, core genome MLST (cgMLST) is proving to be the method of choice rather than whole-genome MLST (wgMLST) or SNP typing, since it provides the resolution needed for epidemiological investigations and is slightly less computationally demanding[63,64]. However, when studying evolution of bacterial lineages by accessing the most common recent ancestor and determining transmission events, a combination of core and accessory genomes analysis is more useful[47,65]. Contrary to traditional typing methods, it is difficult to define thresholds of similarity between strains analysed by WGS, essentially due to the natural occurring variability of the genomes over time. It is fundamental to perform large scale studies including epidemiologically verified outbreak isolates and evaluate significant differences case by case[45,66]. An example of such a study is the INNUENDO project, in which a dynamic cgMLST scheme with three levels of differences for *Salmonella enterica* was established, with 0.5% allele differences discriminating outbreak cases from sporadic cases[45]. However, these thresholds should not be considered absolute and static, since the genomic variability in different *Salmonella enterica* serovars is significant.

1.3. *Salmonella* evolution and genome structure

Any bacterial genome is characterized by a common set of gene families, the core genome, and a set of genes subjected to variability, the accessory genome. The core genome together with the accessory genome constitute the pan-genome[66]. *Salmonella* is a close relative of *Escherichia coli*, both species sharing 90% of their core genomes, which suggests the existence of a common ancestor 100 million years ago[67,68]. Bacterial evolutionary processes are strongly correlated with adaptation to new niches and hosts, and it is believed to have occurred in three phases in *Salmonella*[69]. Genome variation in general, results from point mutations (SNP, insertions, and deletions), horizontal gene transfer (HGT), large insertions and deletions, duplications and rearrangements. However, it was mostly HGT that played an important role in *Salmonella* evolution[70]. The acquisition of *Salmonella* pathogenicity island 1 (SPI 1), which encodes several proteins related to invasion of epithelial cells, by plasmid-or phage-mediated HGT, marked the divergence of *E. coli* and

Salmonella[9,69]. The subsequent acquisition of SPI-2, which contributes to survival and replication inside host cells, through a similar mechanism, resulted in the divergence of the two *Salmonella* species, *Salmonella enterica* and *Salmonella bongori*[9,69]. Finally, the acquisition of virulence genes was essential in the divergence of the different *Salmonella* subspecies, evolving primarily from a strictly gastrointestinal pathogen in cold-blooded animals, to then expand to other hosts like warm-blooded animals and also developing the ability to cause systemic infection[69,70]. In some serovars, host restriction was compensated by the ability to increase longevity of transmission, such is the case of *S. enterica* serovar Typhi, as shall be elucidated in the following sections[70]. Other important mark in *Salmonella* evolution was the acquisition of the *fljAB*/*hin* locus, which conferred the bacteria the ability to change flagellar phases, to elude the host immune system[71]. In addition to the SPIs it is also possible to find Insertion sequence (IS) elements, bacteriophages, transposons and plasmids in *Salmonella* genome[71].

Salmonella genome ranges from approximately 4.4 to 5.8 Megabases (Mb)[72]. The core genome evolved from a lower percentage of guanines and cytosines (% G+C) to a G+C rich genome, with approximately 55% G+C content. The acquired virulence islands are clearly identified as regions with higher adenine and thymine content[73]. The pan-genome consists on the entire set of gene families present in all the strains within a clade, therefore it includes the core genome, the genes present in all strains of the clade, and the accessory genome, strain specific genes[74]. *Salmonella* has what is considered to be a closed pan-genome, which means that the average number of gene families specific to each particular strain is lower than the average genome size (Figure 1-1)[75].

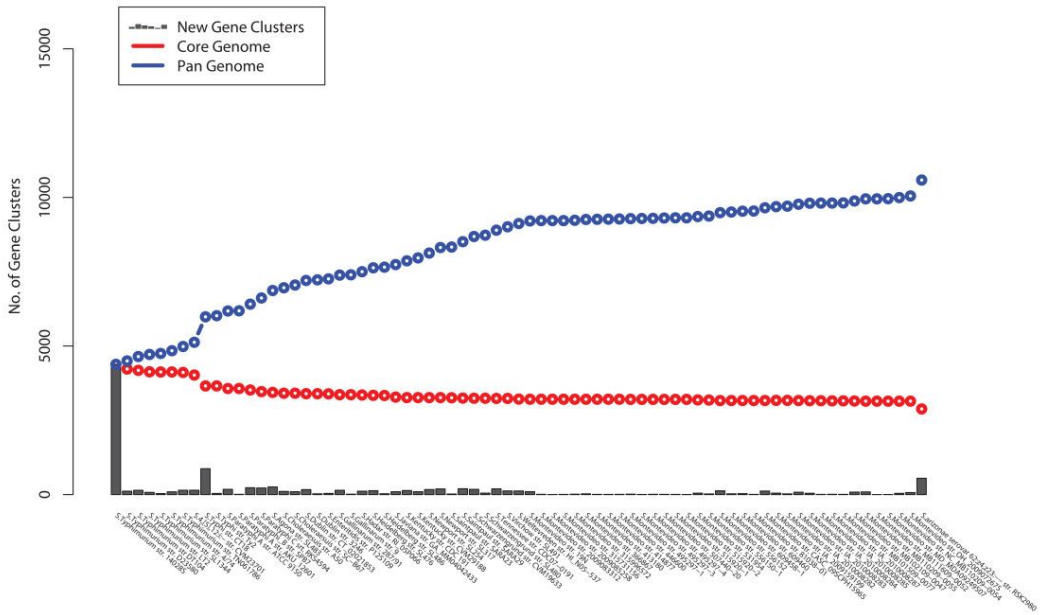


Figure 1-1 – Plot of *Salmonella* pan- and core-genomes, adapted from ([76]).

1.4. Host preference and human disease

Salmonella spp. is a well-adapted pathogen with a wide range of hosts and niches, and depending on the serovar, different pathogenicity. While the main niche of *Salmonella* spp. is the intestinal tract of humans and farm animals, these *Enterobacteriaceae* are widely distributed in nature and have been isolated from a wide variety of animals and their food products[9,77]. *S. bongori*, as well as subspecies *salamae*, *arizonae*, *diarizonae*, *houtenae*, and *indica*, are mainly associated with cold-blooded animals and the environment, while *S. enterica* subsp. *enterica*, that accounts for more than 99.5% of isolated *Salmonella* strains, is usually associated with warm-blooded animals[9]. However, *Salmonella* spp. adapts quite easily to different environments and non-*enterica* subspecies can also be found in warm-blooded animals[9]. *Salmonella* can cause two types of salmonellosis in humans: gastroenteritis, commonly associated with non-typhoidal *Salmonella* (NTS), and enteric fever, caused by *S. enterica* serovar Typhi or *S. enterica* serovar Paratyphi A, B or C[78].

The main vehicle of NTS infection is the ingestion of contaminated food or contaminated water, and by direct contact with animals. NTS is further widespread due to faecal contamination as it is excreted to the environment in the faeces of its hosts[77]. As a result, *Salmonella* can also be found in a variety of fruits and vegetables[10]. Although most animal hosts are asymptomatic, others exhibit signs of infection. For instance, *S. enterica* serovar Gallinarum causes Pullorum disease and *S. enterica* serovar Pullorum fowl typhoid in poultry, *S. enterica* serovar Typhimurium and *S. enterica* serovar Dublin cause salmonellosis in cattle and *S. enterica* serovar Cholerasuis causes septicaemia in pigs[77]. The majority of *Salmonella* infections, caused by NTS, lead to gastroenteritis, presenting self-limiting diarrhoea, associated with fever and abdominal cramping, 12 to 72 hours after infection[10]. However, in immunocompromised individuals, some of those strains may cause systemic infections, frequently lacking diarrheal symptoms, for example, bacteraemia, osteomyelitis and meningitis[78–80].

Typhoid and paratyphoid, serovars *S. enterica* serovar Typhi and *S. enterica* serovar Paratyphi A, B and C, are exclusively human pathogens, meaning that humans are their only reservoir. In endemic countries, typhoid and paratyphoid *Salmonella* are usually transmitted indirectly due to poor hygiene and poor sanitary conditions, which lead to the contamination of the environment either by sewage released in water supplies, or with use of human faeces as fertilizer, or even with untreated piped water[10,81,82]. Other transmission route may be the direct contamination of food and water by a carrier or infected person[82]. The incubation period is often 7 to 14 days and symptoms can sometimes be very similar to other febrile illnesses[82,83]. Enteric fever can be characterized by a wide range of clinical

severity, but the most frequent symptoms include prolonged high fever, nausea, vomiting, and headache. Neurological complications and even death, usually due to intestinal perforation and peritonitis or severe toxic encephalopathy associated with myocarditis and hemodynamic shock, may occur in more severe cases[68,79,80]. Typhoidal *Salmonella* proliferation in the host may endure for extended periods, due to its capacity to invade monocytes and macrophages, and to colonize the gall bladder, with an estimated 5 to 10% of patients experiencing a relapse[79]. Some individuals may even present asymptomatic shedding for several years[10].

1.5. Epidemiology

Salmonellosis is a major public health concern, with huge economic burden in both developed and underdeveloped countries. *Salmonella* is estimated to be responsible for 9% of the 1.9 billion cases of diarrhoea that occur each year worldwide. Nonetheless, 41% of deaths with diarrhoeal causes are attributed to *Salmonella*[84]. In the European Union/European Economic Area (EU/EEA), it remains the second most common cause of gastroenteritis, right after *Campylobacter*[85,86]. NTS is estimated to cause 93.8 million cases of gastroenteritis globally per year, of which 80.3 million are considered foodborne. NTS is also estimated to be responsible for 155 thousand deaths each year[86]. Between 2004 and 2013, a statistically significant decreasing trend of salmonellosis, probably linked to the implementation of control measures in the poultry industry, has been observed in the EU/EEA. However, the countries reporting an increase in the number of cases doubled since 2012, although the trend seems to have stabilized in half of the Member States the past few years. This increase seems to be linked to an increase in the number of cases of *S. enterica* serovar Enteritidis, most specifically to an increase in the number of multi-country outbreaks caused by this serovar (Figure 1-2)[85,87]. In 2016, the most common cause of foodborne outbreaks was *Salmonella*, accounting for 22% of total cases[85]. The highest notification rate continued to be observed in children under 5 years old, and in Portugal it was even 34 times higher than in adults (25-44 years). The hospitalization rate was also very high in Portugal (74–88%), considering the low notification rates (<10 per 100,000)[85].

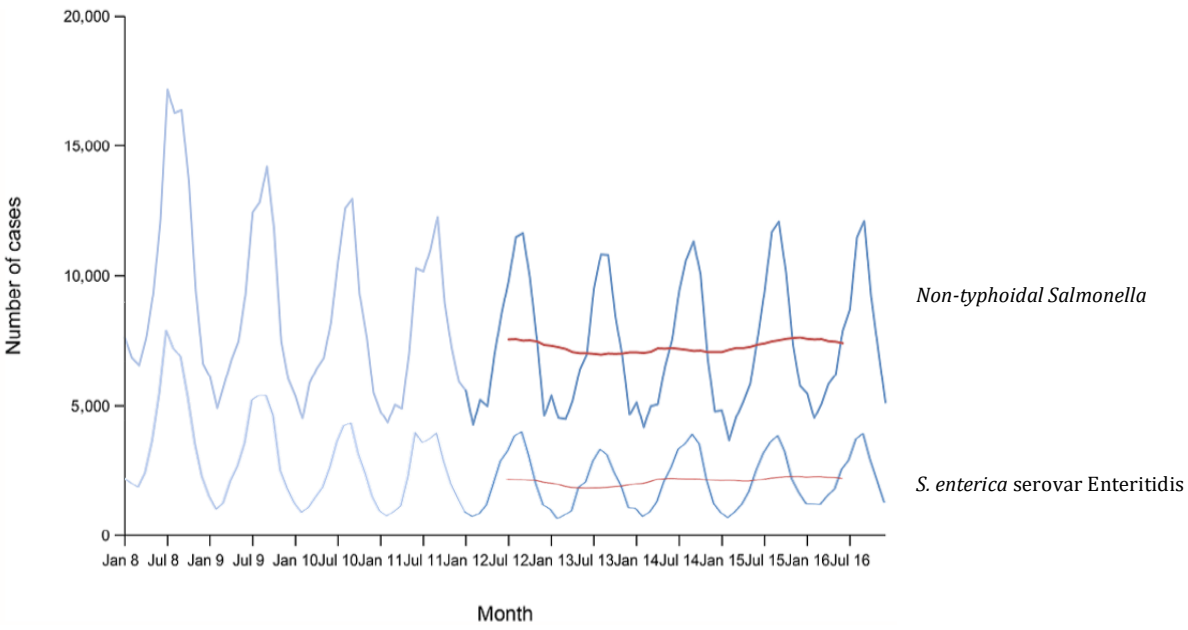


Figure 1-2 - Trends of reported confirmed human cases of non-typhoidal salmonellosis versus trend of reported *Salmonella enterica* serovar Enteritidis cases, in the EU/EEA, 2008–2016. Adapted from ([88]).

While more than 2600 serovars of *S. enterica* have been described to date, *S. enterica* serovar Enteritidis and *S. enterica* serovar Typhimurium remain the most common serovars that cause human disease, accounting for 80% of total cases of salmonelosis[87]. According to the "European Union summary report on zoonoses, zoonotic agents and food-borne outbreaks 2016", the most commonly reported serovars in 2016 were: *S. enterica* serovar Enteritidis, *S. enterica* serovar Typhimurium, monophasic *S. enterica* serovar Typhimurium (*S.* 4,[5],12:i:-), *S. enterica* serovar Infantis, and *S. enterica* serovar Derby[87]. In the European Union (EU), laying hens, broilers and broiler meat remain the main sources of *S. enterica* serovar Enteritidis infections in humans. *S. enterica* serovar Typhimurium is manly associated with cattle, and their meat, and to a lesser extent with pigs, poultry and their meat. Contact with pigs and consumption of pig meat are the most frequent reported sources of monophasic *S. enterica* serovar Typhimurium, while *S. enterica* serovar Infantis is mostly associated with broiler and turkey chains. Finally, *S. enterica* serovar Derby was mostly reported in pigs and pig meat, and to a lesser extent in poultry and cattle (Figure 1-3)[87].

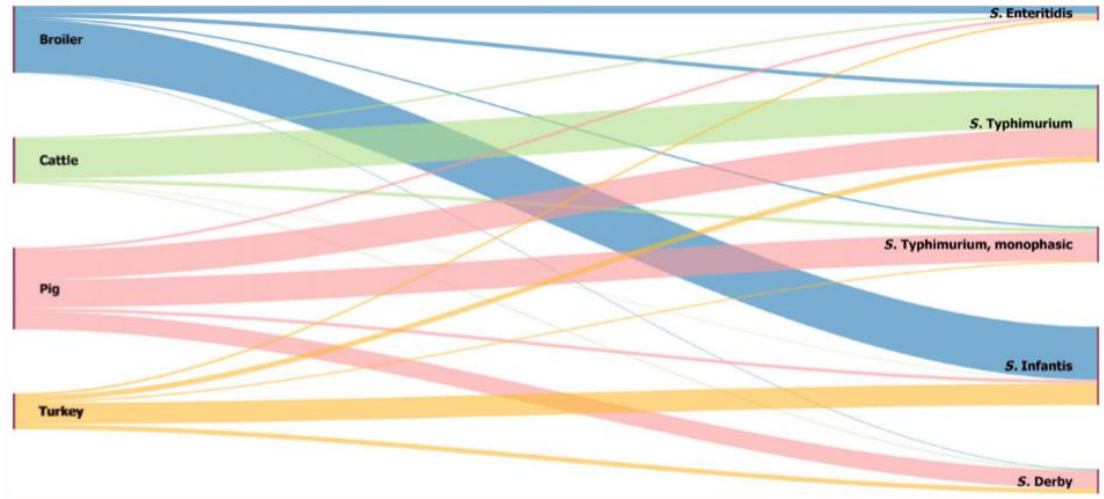


Figure 1-3 - Distribution of the 2016 EU top-five *Salmonella* serovars in human salmonellosis acquired in the Member States, across different food, animal and meat sectors (broiler, cattle, pig and turkey). Adapted from ([88]).

Enteric fever is very frequent in underdeveloped countries, where poor water supply and sanitation is very common. South/South-East Asia, sub-Saharan Africa, and Oceania islands are the most affected areas, presenting the highest mortality rates, especially among children[81]. In 2017, 14.3 million cases of enteric fever occurred globally, with 76.3% of cases caused by *S. enterica* serovar Typhi and 135.9 thousand deaths[81]. Enteric fever is a rare disease in the EU/EEA, mainly acquired while travelling to endemic countries, specially India and Pakistan. In 2016, 22 EU/EEA countries reported 1161 confirmed cases of typhoid/paratyphoid fever, representing an increase comparing to previous years, probably related to higher notification rates and an increase in travels to countries outside the EU[89]. *S. enterica* serovar Typhi was the most frequent cause of enteric fever in EU/EEA, accounting for 60% of cases, followed by *S. enterica* serovar Paratyphi A[89]. When traveling to endemic countries, WHO recommends vaccination with one of the 3 available vaccines[82].

1.6. Pathogenesis

After ingestion, *Salmonella* has to overcome numerous host defence mechanisms, such as gastric acidity, proteases, defensins and aggressins, in order to succeed and cause infection[90,91]. When *Salmonella* enters the small intestine, the bacteria adheres to intestinal epithelial cells, preferably M cells of Peyer's patches (Figure 1-4), and disrupts the epithelial structure causing diarrhoeal symptoms[90]. *Salmonella* then invades enterocytes, via a type III secretion system encoded by SPI-1, T3SS-1, or dendritic cells, in order to reach follicles and mesenteric lymph nodes[78]. Inside the enterocytes, the bacteria resides in *Salmonella*-

containing vacuoles (SCV) where it disrupts the cell's cytoskeleton, alters membrane transport, signal transduction, and cytokine expression, by secreting SPI-2 encoded effector proteins via another type III secretion system, T3SS-2[78,90–92]. This allows *Salmonella* to survive and replicate inside host cells. The SCVs are eventually transported to the basolateral membrane and *Salmonella* is released to the submucosa to either (i) colonize adjacent enterocytes, to repeat the intracellular cycle again, (ii) be phagocytosed by neutrophils, inflammatory monocytes, or dendritic cells, and then spread systemically upon phagocyte migration, (iii) or even to spread directly to the lymph nodes[90–92]. Some serovars, such is the case of *S. enterica* serovar Typhi and *S. enterica* serovar Typhimurium, have yet the ability to form biofilms on the gallbladder stones, prolonging their longevity inside the host[91].

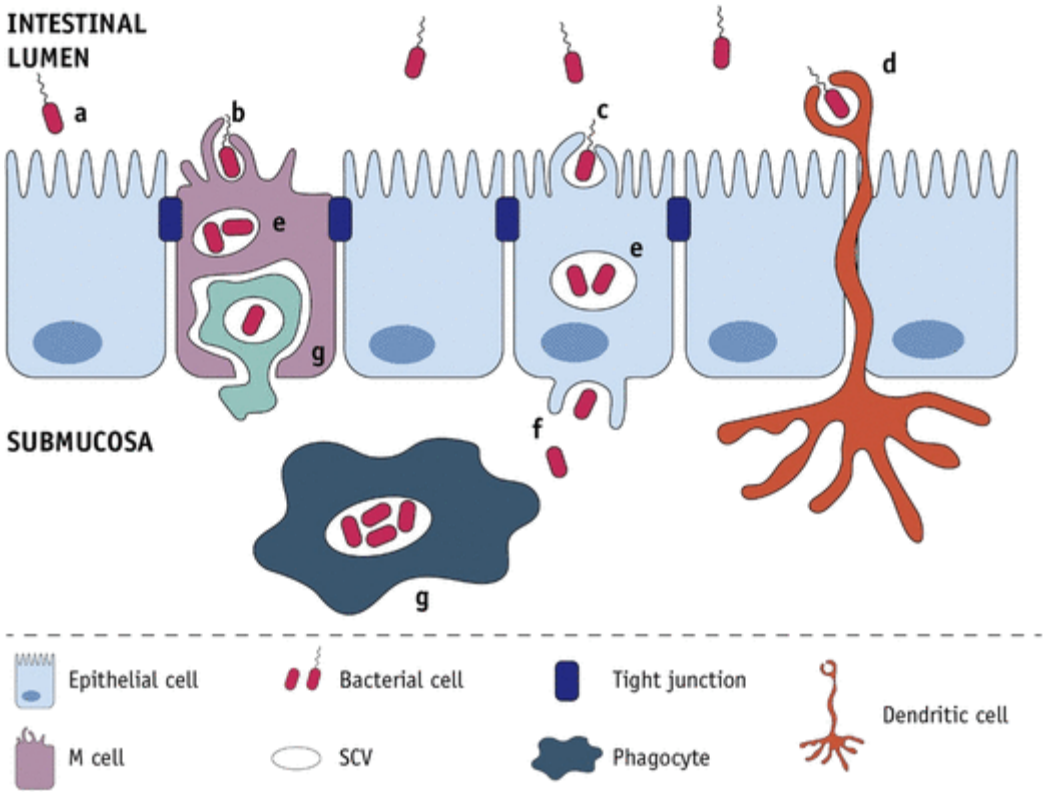


Figure 1-4 – Pathogenesis of *Salmonella*, adapted from ([92]). **a.** *Salmonella* adheres to intestinal epithelium and M cells. **b, c.** Effector proteins cause changes to the cytoskeleton. **d.** Alternatively, *Salmonella* is directly engulfed by dendritic cells from the submucosa. **e.** Inside the enterocytes, *Salmonella* replicates inside *Salmonella*-containing vacuoles. **f.** *Salmonella* is released to the submucosa. **g.** Finally, bacteria is phagocytosed and contained in SCVs.

1.7. Diagnosis, antimicrobial therapy and antimicrobial resistance

1.7.1. Diagnosis

Salmonellosis diagnosis, whether if it is gastrointestinal disease or enteric fever, requires isolation of *Salmonella* from a clinical sample, which may be faeces, blood, bone

marrow, urine, among others. The culture medium used depends on the type of sample collected. Isolation of *Salmonella* in faeces usually requires an enrichment step in Selenite F (*Salmonella* enrichment and inhibition of other bacteria) or GN broth (*Salmonella* and *Shigella* enrichment and inhibition of other bacteria) for 6 to 12h prior to culture in selective media, such as Hektoen agar (selective and differential for *Salmonella* and *Shigella*), SS agar (selective and differential for *Salmonella* and *Shigella*), or XLD agar (selective for *Salmonella* and *Shigella*). Blood cultures require incubation in proper culture bottles followed by subculture in differential media. Urine samples are firstly cultured in CLED agar for inhibition of *Proteus* swarming and colony counting, and also subcultured in differential media. All the obtained cultures are then identified either by their biochemical characteristics using automated systems or more prosaic testing, or by MALDI-TOF MS.

The serological Widal test is still used in several countries as the main diagnostic method for enteric fever given its low cost and straightforward operation. It consists in an agglutination reaction between O and H antigens with the antibodies present in serum of a patient. Given that an individual may present background antibodies due to previous or repeated infections, vaccination, or even for living in an endemic area, this test has low sensitivity and is very inaccurate. Moreover, this test also lacks specificity due to antigens cross reactivity. Rapid serologic tests, represent a slightly improvement from Widal test, but still lack sensitivity and specificity. When culture is not possible or successful, molecular testing by conventional or real-time PCR come as an alternative[93,94].

1.7.2. Antimicrobial therapy

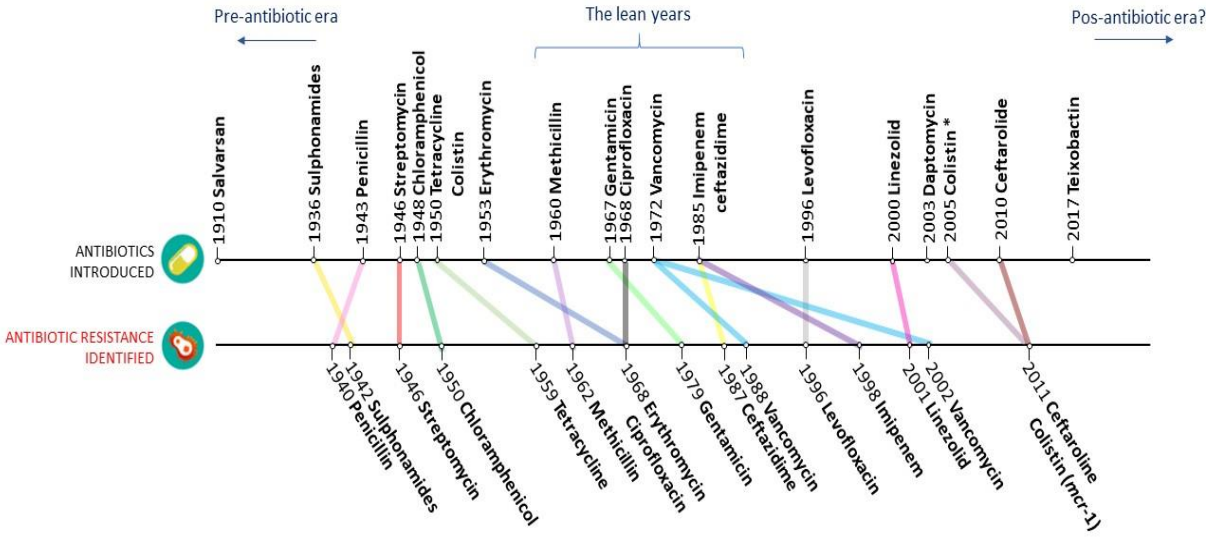
In the majority of cases, salmonellosis caused by NTS does not require hospitalization. Treatment usually involves replacement of fluids and electrolytes to control dehydration. Empirical antibiotic therapy is to be avoided in mild to moderate cases in otherwise healthy individuals, since it gives rise to antimicrobial resistance and may even contribute to prolongation of faecal shedding[10,95]. Nevertheless, antibiotics may be required to cure severe disease, invasive NTS (iNTS) infection and enteric fever, after appropriate testing[96]. Additionally, neonates and young children should always be treated to avoid bacteremia[97]. Treatment of chronic shedding involves long-term oral antibiotic therapy and possible removal of the gallbladder [95].

For several years, the first-line of treatment of enteric fever and NTS severe infections was either ampicillin, chloramphenicol or trimethoprim–sulfamethoxazole[96]. With the emergence of multidrug resistant (MDR) strains, that is, strains presenting resistance phenotypes to at least three different drug classes, the fluoroquinolone

ciprofloxacin and later the third-generation cephalosporin ceftriaxone, have been used as empirical therapy, and azithromycin as the last line of treatment[98]. But eventually, resistance phenotypes to these antibiotics started to emerge and the nephrotoxic colistin was recently re-introduced has the last resort treatment of severe infections caused by MDR pathogens, including *Salmonella*[99–102]. Salmonellosis caused by MDR strains are more invasive, require more hospitalizations and are associated with higher mortality rates[103]. Recently a major concern is threatening the advances of modern medicine, with the identification of plasmids that confer resistance to colistin (*mcr* family), which will probably set us back to a post-antibiotic era, where patients will unfortunately succumb to common infections and minor injuries[104–111].

1.7.3. Antimicrobial resistance

Antibiotics have been one of the major findings of the modern era. The introduction of penicillin for public use in the 1940s revolutionized Medicine, having countless lives been saved since then[112]. However, even at the time of its discovery, Alexander Fleming gave warning of the possibility of occurrence of penicillin resistant strains and in fact, the development and generalized misuse of new antibiotics has always been shortly



accompanied with the arise of resistant strains (Figure 1-5)[113].

Figure 1-5. – Antibiotics timeline depicting the year of introduction and the year when the first resistant strain was identified, adapted from ([114]), ([112]), ([115]), ([102]), https://www.ukri.org/ukri/assets/Image/SiteImages/xorg/amrinfographicfull.jpg, and https://www.cdc.gov/drugresistance/about.html. *The use of colistin was abandoned in the 1980s due to nephrotoxicity effects and was later re-introduced to treat multidrug resistant infections.

The emergence of MDR strains is mainly promoted by the indiscriminate misuse of antibiotics in human and veterinary treatment, and as growth factors in animals for food production[112]. Antibiotic resistance may be intrinsic, when microorganisms exhibit inherent characteristics that confer resistance to certain drugs, or acquired, when resistance genes are acquired from other bacteria (same or different species), or when new mutations in chromosomal genes occur and induce expression of innate resistance mechanisms[116]. Foreign DNA may be acquired through intake of insertion sequences (IS), transposons, plasmids and bacteriophages[117]. In order to avoid an antibiotic apocalypse has preconized by World Health Organization (WHO), we are called for action specifically in respect of harmonization of surveillance in humans and food-producing animals, in a "One Health" approach.

1.7.3.1. β-lactams

β-lactams incorporate three major groups of antibiotics: penicillins, cephalosporins (1st to 4th generation) and carbapenems. They interfere with cell wall formation by inhibiting Penicillin Binding Proteins (PBPs), impeding synthesis of peptydoglican[118]. The most common mechanism of resistance to β-lactams in *Salmonella* is the secretion of β-lactamases, enzymes that hydrolyse the β-lactam ring rendering these antibiotics inactive[118]. These hydrolytic enzymes can either be inherent to some bacteria being carried on chromosomes, or in plasmids with potential to disseminate to other microorganisms (*bla* genes)[119]. There are several classes of β-lactamases that comprise more than one thousand enzymes but the most commonly associated with *Salmonella* resistance are class A β-lactamases, specially the extended-spectrum β-lactamases (ESBLs) gene family TEM[114,118]. Other genes commonly associated with *Salmonella* include class A bla_{PSE-1}, bla_{CARB-2}, bla_{KPC-2}, and bla_{CTX-M}, and class C bla_{ampC}, and bla_{CMY-2}[118]. ESBLs became a major public health concern not only because they have the ability to hydrolyse cephalosporins (cefotaxime, ceftriaxone, ceftazidime, or cefepime) and monobactams (aztreonam), but also because they are widely disseminated worldwide[114,120].

1.7.3.2. Aminoglycosides

Aminoglycosides (streptomycin, neomycin, amikacin, gentamicin, spectinomycin, kanamycin, tobramycin, among others) act by inhibiting protein synthesis by binding to the 16S rRNA of the 30S ribosomal subunit, leading to codon misreading and inhibition of translation[118,121]. The mechanism of resistance to this group of antibiotics in *Salmonella* is usually the modification of the drug although other resistance mechanisms such as reduction of uptake, target alteration and increased efflux exist in other bacteria[118,121].

There are several types of modifying enzymes that confer resistance to aminoglycosides. Aminoglycoside acetyltransferases confer resistance to gentamicin, tobramycin, and kanamycin by modifying their amino groups. They are coded by genes *aac* found in *Salmonella* genomic islands, integrons, and plasmids. Aminoglycoside phosphotransferases, coded by plasmid genes *aph* (*aph*(3') and *aph*(6)-*Id*, also known as *strA* and *strB*, respectively), catalyse ATP-dependent phosphorylation of hydroxyl groups of streptomycin, kanamycin, and neomycin. Finally, nucleotidyltranferases encoded by genes *aad*, some also designated *ant* genes, also target hydroxyl groups by adding AMP from an ATP donor. The *aadA* gene, also known as *ant(3')*, confers streptomycin resistance in *Salmonella*, while *aadB*, or *ant(2')-Ia*, provides resistance to gentamicin and tobramycin[118,121].

1.7.3.3. Quinolones

All quinolones (nalidixic acid, ciprofloxacin, pefloxacin, levofloxacin, etc.) exhibit the same mechanism of action, typically targeting DNA gyrase and DNA topoisomerase IV, disrupting DNA replication and leading to cell death[122,123]. Consequently, the mechanism of resistance conferring resistance to one quinolone, confers resistance to all quinolones, although some mechanisms only decrease quinolone susceptibility and require the presence of additional alterations to confer full resistance[124]. Several mechanisms of resistance to quinolones have been described: (i) mutations in the genes that encode the two topoisomerases, *gyrA*, *gyrB*, *parC* and *parE*, lower the affinity of quinolones to the topoisomerases; (ii) the presence of plasmid-mediated quinolone resistance genes (PMQR), *qnrA*, *qnrB*, *qnrS*, *qnrC*, and *qnrD*, encoding topoisomerase-binding proteins that prevent quinolone binding to the enzymes; (iii) the presence of the *aac(6')lb-cr* gene, encoding a modifying enzyme that decreases quinolone activity; (iv) efflux pumps encoded by *oqxAB* and *qepA*; (v) and finally, the regulation of multidrug efflux pumps AcrAB-TolC, which lower the concentration of quinolones in the cell[122].

1.7.3.4. Macrolides

Macrolides act by binding to the large subunit of 23S rRNA, inhibiting protein synthesis, causing premature dissociation of the tRNA[125]. Macrolide resistance in general is either conferred by (i) ribosomal binding site modification by methylases encoded by *erm* genes; (ii) by antibiotic modifying enzymes esterases (*ereA* and *ereB*) and phosphotranferases (*mphA*, *mphB*, and *mphD*); (iii) by active efflux of the antibiotic (*mefA* and *msrA*); or (iv) by the presence of mutations in the genes *rrl* and *rpl*, encoding the 50S ribosomal proteins L4, and L22 and in 23S rRNA[126].

1.7.3.5. Other antibiotics

Resistance mechanisms to additional clinically relevant antibiotics have been identified in *Salmonella*. Chloramphenicol acts by binding to the 50S ribosomal unit, preventing the formation of new peptide bonds. Resistance to chloramphenicol is conferred by enzymatic inactivation of the antibiotic (*cat1* and *cat2*) or by the presence of efflux pumps (*cmlA* and *floR*)[118].

Tetracyclines also act by inhibiting protein synthesis, by preventing binding of tRNA to the 30S ribosomal subunit. Resistance to tetracyclines in *Salmonella* depends on the presence of efflux pumps encoded by *tet(A)*, *tet(B)*, *tet(C)*, *tet(D)*, *tet(G)*, and *tet(H)*[118].

Sulphonamides and trimethoprim have been used in combination for a few decades. These antimicrobials have a similar mode of action, inhibiting enzymes involved in the synthesis of tetrahydrofolic acid (co-factor in synthesis of amino acids and nucleic acids). The mechanism of resistance to both antibiotics is also similar. Genes *sul1*, *sul2*, and *sul3*, and *dhfr1*, *dfrA1*, and *dhfr12* express enzymes insensitive to sulphonamides and trimethoprim respectively[118].

Finally, and in brief, colistin, also known as polymyxin E, binds in a disruptive way to lipopolysaccharides (LPS) of the outer membrane of Gram-negative bacteria, leading to an increase of permeability of the cell and ultimately to cell death. Colistin resistance is quite complex and involves several mechanisms. For the most part, mechanisms leading to colistin resistance are linked to chromosomal mutations, except for the plasmid mediated *mcr* gene family. In general, in *Salmonella*, they involve mutations that lead to modifications of the outer membrane (*pmrA*/*pmrB*, *phoP*/*phoQ*, *arnBCADTEF*, and *mcr1-5*)[127].

1.8. Heavy metals tolerance

Metals occur naturally in nature and act as important co-factors for several enzymatic processes. In fact, bacterial growth requires the presence of metals, but these compounds become toxic at higher concentrations[128]. In 2006, EU banned the use of antimicrobials in animal feed, which lead to the use of heavy metals such as copper, cobalt, and zinc, as alternative growth promoters or preservatives[129,130]. Additionally, arsenic, copper, mercury, silver and zinc have been used as biocidal agents to treat human and animal infections for several decades. Mercury and silver are also recognized by their disinfectant properties[128]. Meanwhile, heavy metal tolerance has been detected in association with antibiotic resistance genes in several bacteria, including *Salmonella*[130–132]. These genes are usually located in plasmids, transposons, or integrons, therefore, the use of heavy metals has been suggested to promote the spread of antibiotic resistance[130–135]. The

mechanisms that lead to heavy metal tolerance include efflux systems to reduce copper concentration in the cell (*pcoABCDRSE* and *tcrYAZB*), silver (*silCFBAPRSE*) or arsenic (*arsRBC*), detoxification of mercury by enzymatic reduction (*mer* operon), and tellurite resistance system (*terZABCDEF*)[132,136].

1.9. Aims and research plan

The main objective of this PhD dissertation was to contribute to the determination of the genomic structure and antibiotic resistance phenotype of particular *Salmonella enterica* serovars identified in Portugal, through WGS of several isolates with different phenotypes and genotypes. This main objective is divided in several specific aims:

i) to evaluate the main serovars causing salmonellosis in Portugal and to identify serovars of interest for our study (Chapter 1);

ii) to identify clusters of potential epidemiological interest and compare several typing methods with WGS (Chapter 2, 3, and 4);

iii) to study the genomic structure of isolates of the same serovar (Chapter 2, 3, and 4);

iv) to identify resistance markers and mobile genetic elements responsible for the phenotype expressed by the isolates in study (Chapter 2, 3, and 4).

CHAPTER 2

Salmonella spp. surveillance in Portugal

LS performed most of the experimental work from 2014 to 2019 and interpreted all the data analysed for this chapter.

2. *Salmonella spp.* surveillance in Portugal

2.1. Global trends

In 1899, in order to structure and develop a public health system, by offering technical and professional qualification to the sanitary practise, and most specifically, as a means to overcome a large bubonic plague outbreak in Porto, Doctor Ricardo Jorge, a renowned Portuguese doctor and researcher, founded *"Instituto Central de Higiene"* (Central Institute of Health), renamed *"Instituto Nacional de Saúde Doutor Ricardo Jorge"* (National Institute of Health Doctor Ricardo Jorge) (INSA) in 1971 in his honour. Throughout the years, INSA has always been working towards public health improvement in several fields of action, including *Salmonella spp.* surveillance[137].

In this regard, since 1950, *Salmonella* serotyping has been performed at INSA, and the National Reference Laboratory (NRL) of Gastrointestinal Infections, formally created in 1991, has carried on the work of its predecessors to these days. In 1972, Professor Adriana Figueiredo published the results of 21 years of serotyping of *Salmonella spp.* isolated in Portugal. Between 1950 and 1971, the 10 most common serovars in Portugal, in decreasing order, were *S. enterica* serovar Typhi, *S. enterica* serovar Typhimurium, *S. enterica* serovar Enteritidis, *S. enterica* serovar Paratyphi B, *S. enterica* serovar Newport, *S.enterica* serovar Bredeney, *S. enterica* serovar Infantis, *S. enterica* serovar Thompson, *S.enterica* serovar Blockley, and *S. enterica* serovar Anatum. Naturally, typhoid and paratyphoid fever cases decreased significantly through the years, while new serovars appeared and increased in numbers[138]. Figure 2-1 incorporates the data available, from 1950 to July 2019, regarding the number of *Salmonella spp.* isolates serotyped. In total, 21,362 Portuguese isolates were analysed, the most part (approximately 84%) isolated from human clinical samples (Figure 2-2). Between 2003 and 2007 the number of isolates identified peaked, decreasing subsequently until 2011. The introduction of *Salmonella* control measures at the primary production level, established in Regulation (EC) No 2160/2003, and following amendments, are the most likely cause for the *Salmonella* decreasing trend in Europe and also in Portugal. On the contrary, an increase in the number of outbreaks, was a likely cause for the increase in the number of isolates serotyped and in the number of *S. enterica* serovar Enteritidis from 2014 to 2017, which was also observed in most EU/EEA member states[87,139,140]. Interestingly, the trends of the total number of *Salmonella spp.* accompany perfectly the trends of *S. enterica* serovar Enteritidis, since it has been one of the main causes of human infection for several years (Figure 2-3).

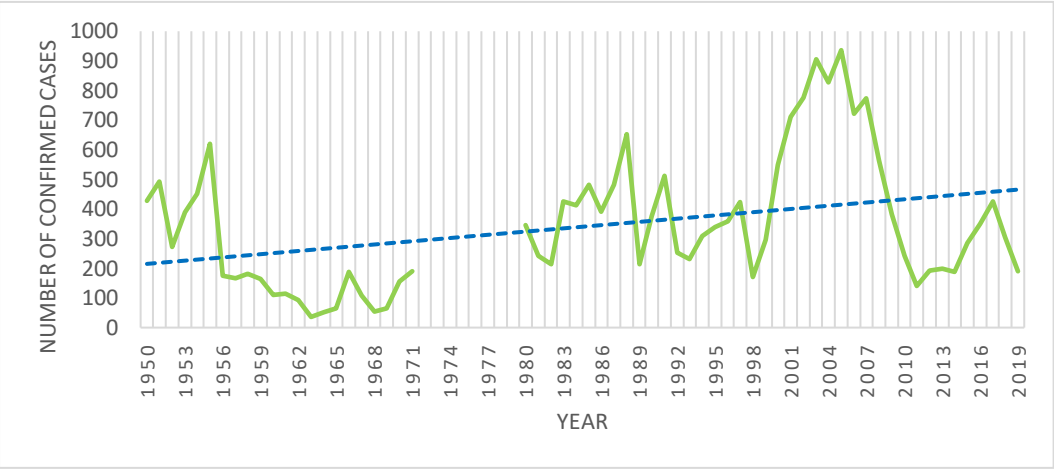


Figure 2-2 – Total of *Salmonella spp.* serovars identified from 1950 to July 2019 in the National Institute of Health Doctor Ricardo Jorge. No data available from 1972 to 1979. The restriction of access to INSA between November 1997 and July 1998, as a result of a contamination hazard, led to a decrease in the number of isolates received at the NRL in both years. Adapted from ([138]).

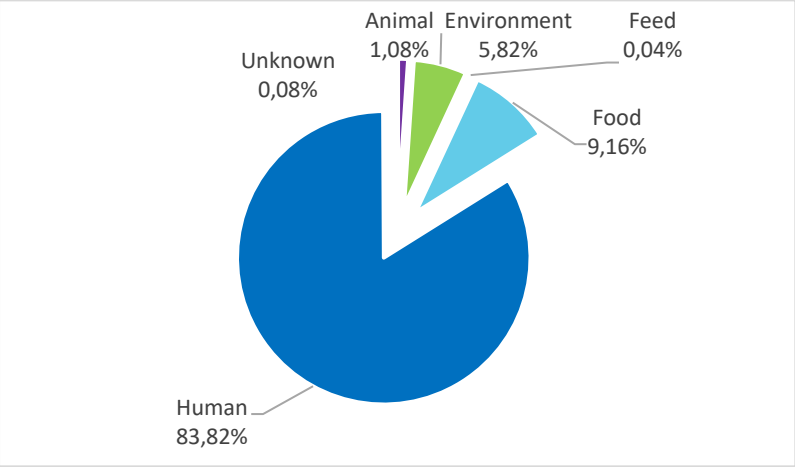


Figure 2-1– Origin of the *Salmonella spp.* isolates received at National Institute of Health Doctor Ricardo Jorge for serotyping, from January 1998 to July 2019.

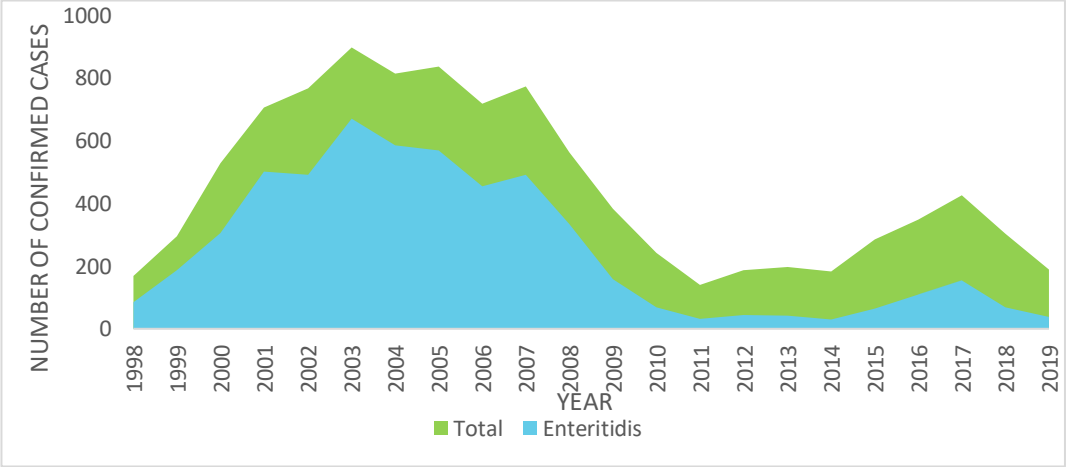


Figure 2-3 – Trends of total *Salmonella spp.* serovars identified versus number of *Salmonella enterica* serovar Enteritidis, from 1950 to July 2019 in the National Institute of Health Doctor Ricardo Jorge.

Regarding the human clinical cases, typhoid and paratyphoid fevers have been compulsorily notifiable diseases in Portugal since 1950, while for salmonellosis, this occurred only after 1987. While comparing the number of notifications of typhoid and paratyphoid fever and the number of cases confirmed by the NRL (Figure 2-4), it can be observed that the discrepancies that once existed gradually diminished. This paralleled the decreasing use of Widal tests in Portuguese hospitals, along with a tighter validation process by the Health Directorate of the cases notified, avoiding an over notification that likely existed in the first years. As far as it is known, the majority of cases of typhoid and paratyphoid fevers in Portugal are acquired while travelling. As for salmonellosis (Figure 2-5), variation between notifications and confirmed cases still remains. Salmonellosis is one of the most commonly notified diseases in Portugal[141–145].

Analysing the data from the isolates received in the NRL, the age-specific confirmed case rate was always highest in the 0-4 age group (in average 24.64 per 100,000 inhabitants) (Figure 2-6), and was 11 to 53 times higher than the 15-24 age group, which was the group with less cases overall. For the most part the male to female ratio was 1:1. Overall, most clinical samples were faeces, followed by blood and urine. It should be taken into account that NTS was responsible for 92.85% of systemic infections. *Salmonella* cases follow a seasonal trend, with most cases occurring during the summer months, mainly July and August, and a smaller peak in December, probably due to overconsumption of egg-based desserts during the holyday season (Figure 2-7). As expected, due to the high population density, most patients reside in Lisboa and Porto, and several districts have few confirmed cases (Figure 2-8). Of note, the detection of 214 outbreaks, between January 1998 and July 2019, with 18.69% being identified in the last 5 years.

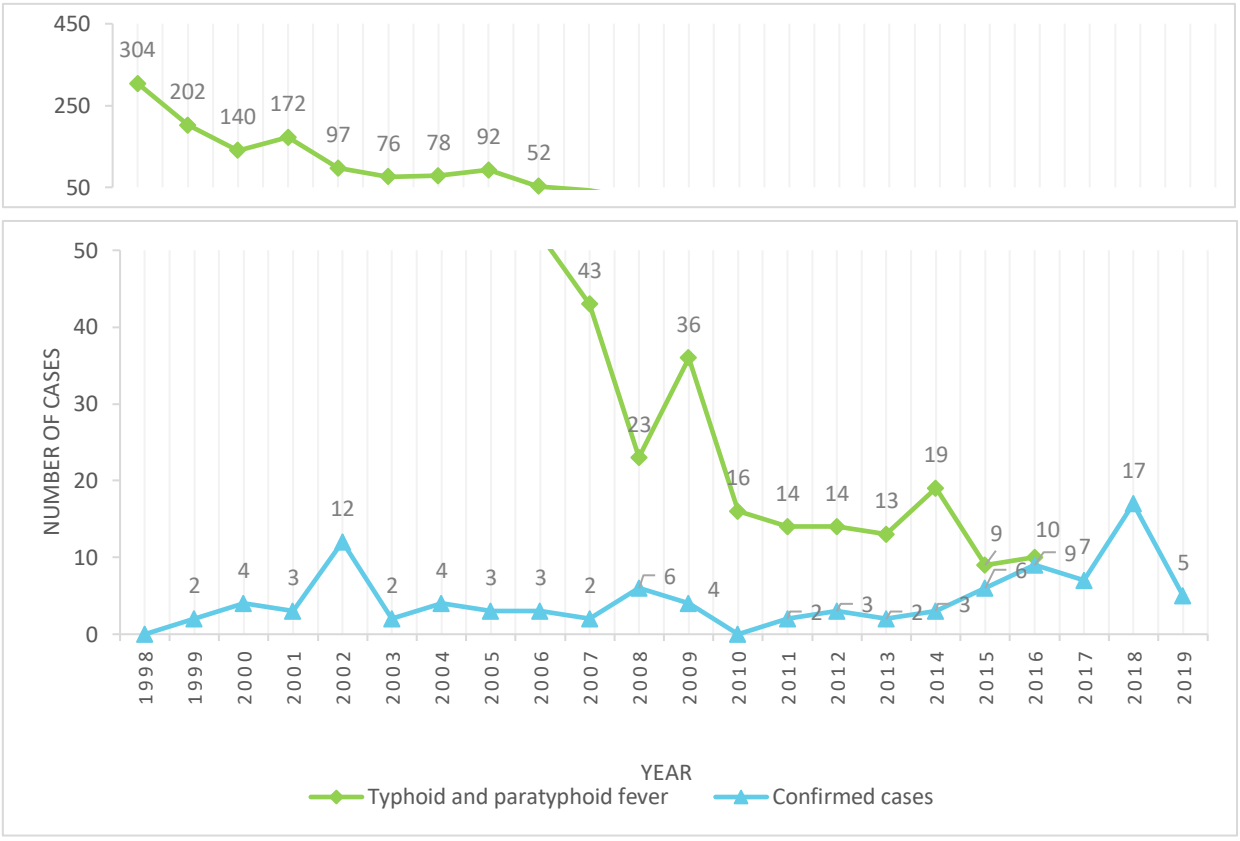


Figure 2-4 – Number of notifications of typhoid/paratyphoid fever from January 1998 to December 2016 and number of laboratory confirmed cases from January 1998 to July 2019, in Portugal. Adapted from ([141–145]). The scale of the y-axis is broken for simplification purposes.

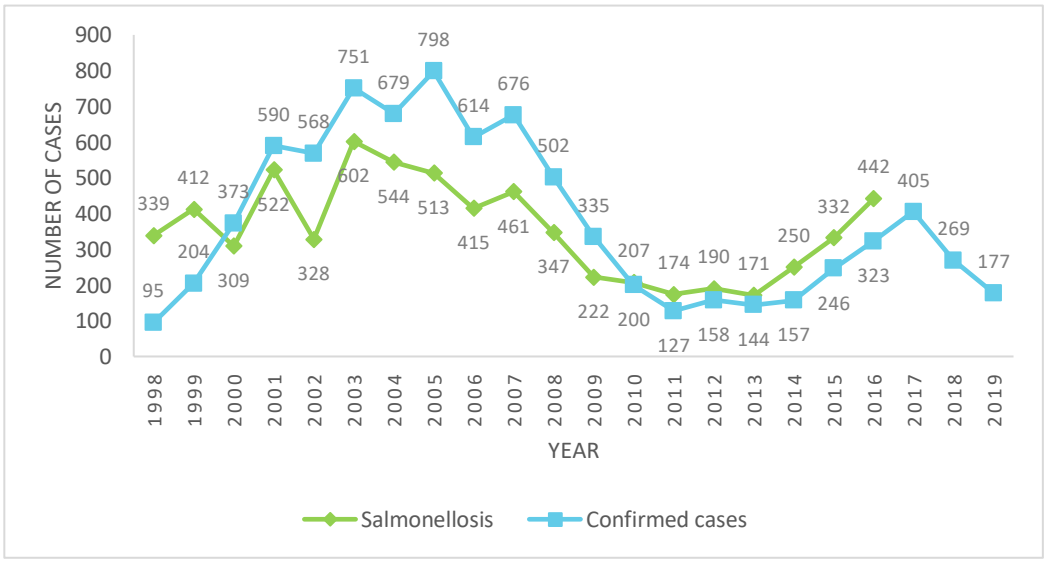


Figure 2-5 – Number of notifications of salmonellosis from January 1998 to December 2016 and number of laboratory confirmed cases from January 1998 to July 2019, in Portugal. Adapted from ([141–145]).

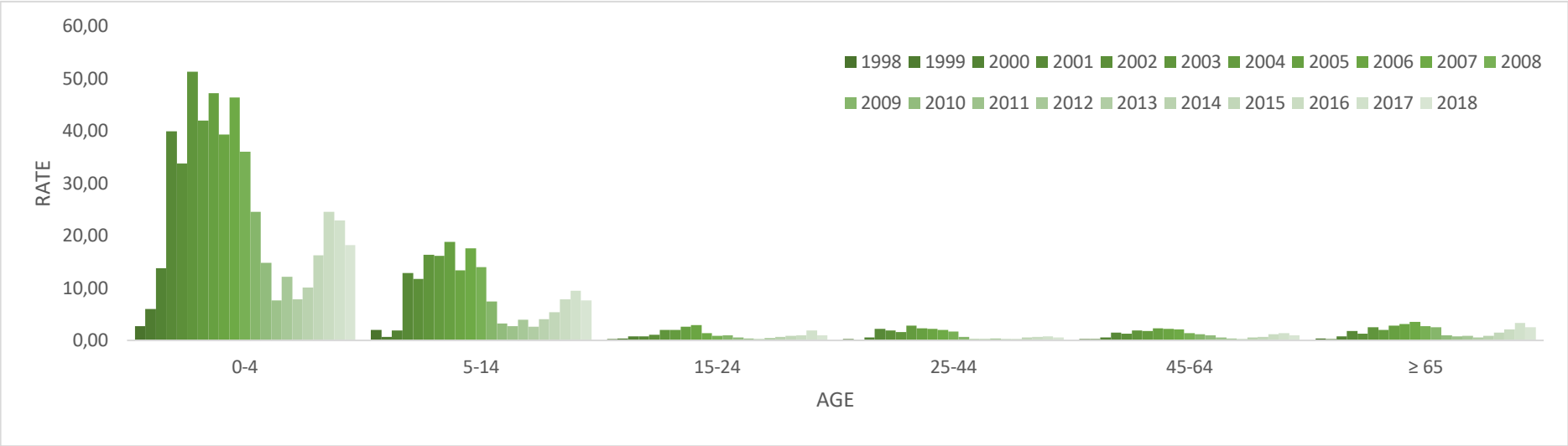


Figure 2-6 – Rates of cases by age group, sent to the Portuguese National Reference Laboratory for serotyping, from January 1998 to December 2018.

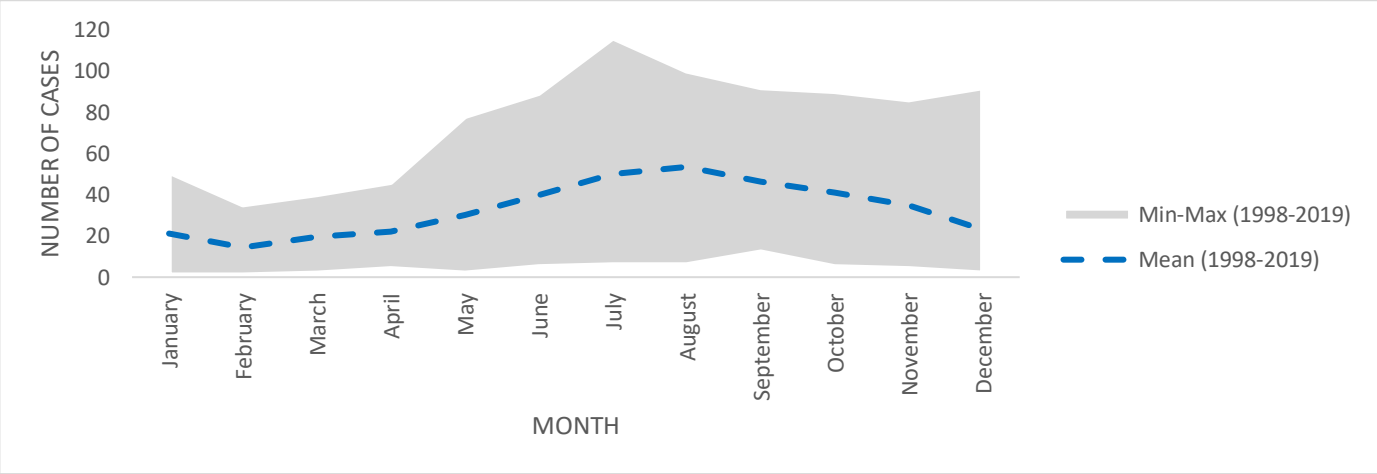


Figure 2-7 – Seasonality of *Salmonella spp.* in Portugal. The graph displays the max, min and mean values obtained for each month in the period between 1998 and 2019.

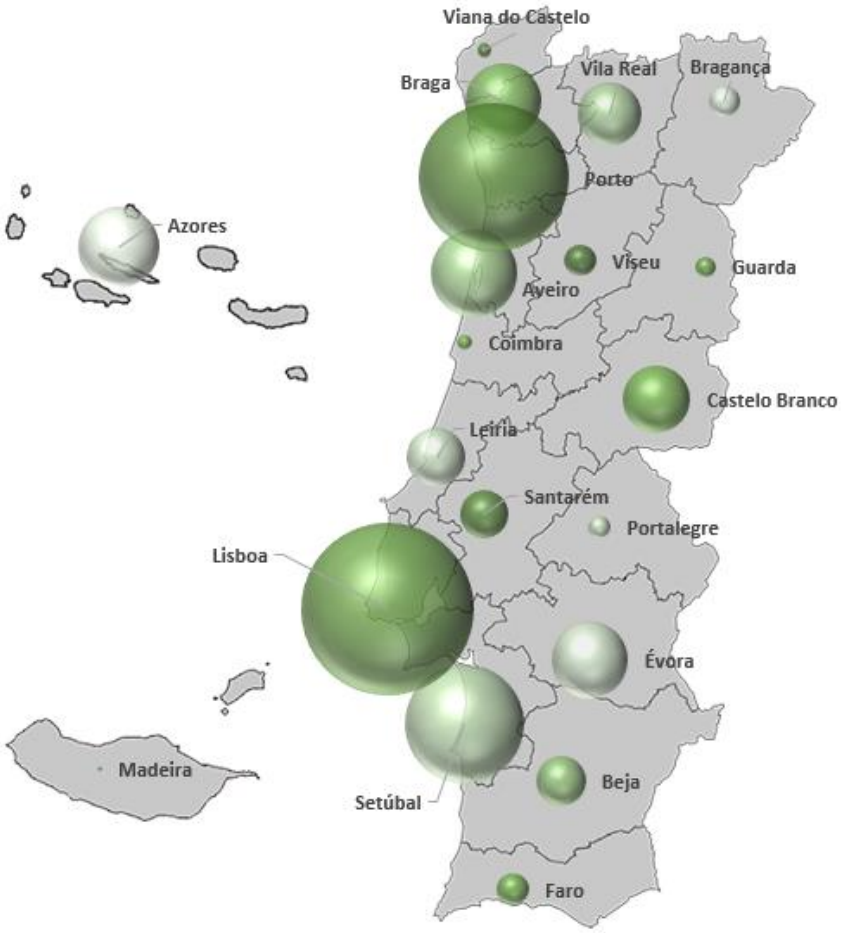


Figure 2-8 – Place of residence of patients with salmonellosis between January 1998 and July 2019.

2.2. Serotyping: trends from January 1998 to July 2019

In Portugal, between January 1998 and July 2019, the most commonly isolated serovars were *S. enterica* serovar Enteritidis, *S. enterica* serovar Typhimurium, and *S. enterica* serovar 1,4,[5],12:i:- (Figure 2-9). As shown in Figure 2-10, *S. enterica* serovar Enteritidis was unequivocally the most common serovar until 2010, whereas this trend is not seen since then. Given that this serovar was mainly found in poultry, the European Commission enforced a framework of legislation targeting *S. enterica* serovar Enteritidis and *S. enterica* serovar Typhimurium in broilers, breeders and turkeys. The turning point was most likely the introduction of Regulation (EC) No 2160/2003 that includes minimum requirements for testing and the elimination of breeding flocks positive for *Salmonella*, and the subsequent introduction of Regulation (EC) No 853/2004, and Regulation (EC) No 854/2004 that includes rules for food safety measures, and rules for the implementation of mandatory official checks, respectively. These regulations contributed for the decrease of cases caused by this serovar, however, in 2017 several *S. enterica* serovar Enteritidis outbreaks, still unexplainable, occurred and this serovar became once again the most prevalent in that year.

In 1986-1987, the first monophasic *S. enterica* serovar Typhimurium, *S. enterica* serovar 1,4,[5],12:i:-, was isolated from poultry carcasses in Portugal[146]. Since then, this serovar has been identified more frequently in Portugal, even surpassing *S. enterica* serovar Enteritidis in 2012, 2014-2016, 2018, and in the first semester of 2019. Regarding other serovars with less representation but also of clinical interest, *S. enterica* serovar Typhi, for example, slightly increased in 2002, 2016 and 2018. Additionally, an increase in the number of isolates of *S. enterica* serovar Rissen in 2015, prompted a study, detailed in Chapter 4 of this dissertation, on the diversity of this serovar in Portugal. In July and August 2019, a sudden increase of *S. enterica* serovar Newport, prompted a still ongoing outbreak investigation, which also involves the Portuguese Health Authorities. In order to establish possible links, epidemiological inquiries are taking place and in WGS of a set of isolates was already carried out (Chapter 3).

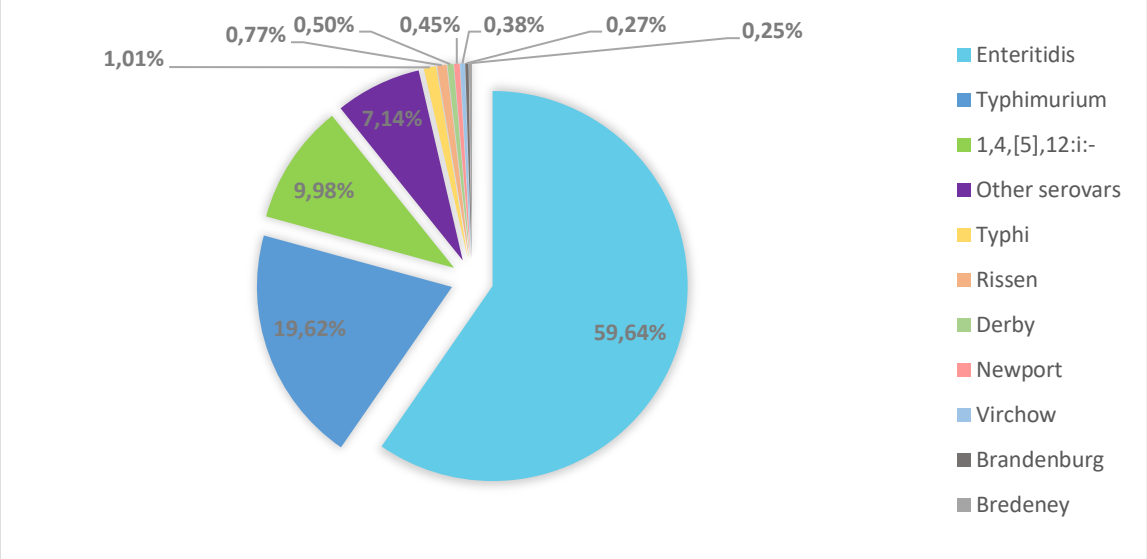


Figure 2-9 – Percentage of *Salmonella* spp. serovars isolated from patients between January 1998 and July 2019, in Portugal.

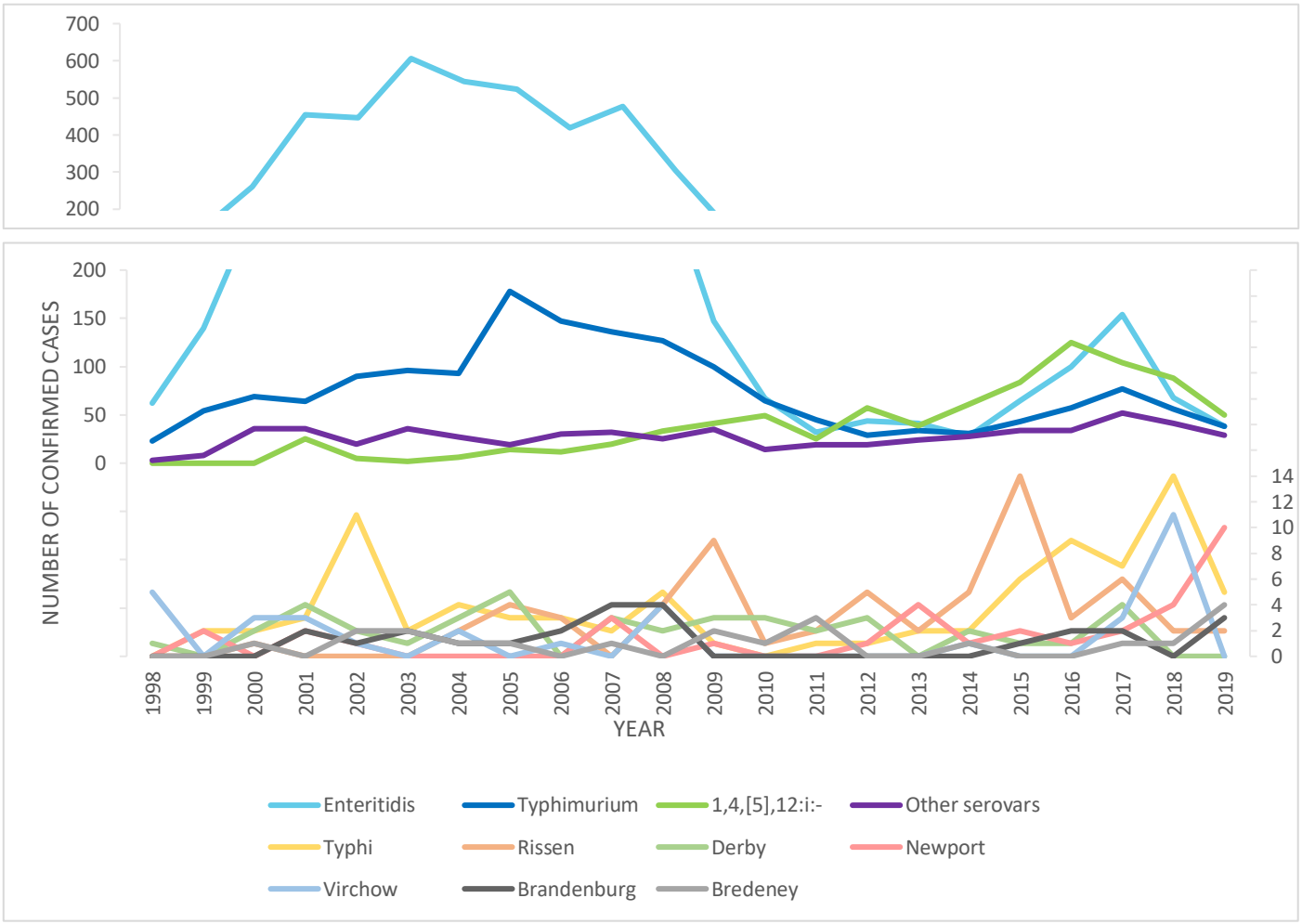


Figure 2-10 – Trends of 10 serovars of interest isolated between January 1998 and July 2019 from Portuguese patients. Due to the major differences in number between the top 3 most common serovars, represented on the top panel, and the next seven, represented on the bottom panel, two different vertical axis had to be included in the graph. Additionally, the y-axis is broken for simplification purposes.

2.3. Multidrug resistance in Portugal

Since 2015, the NRL has been following the European Committee on Antimicrobial Susceptibility Testing (EUCAST) guidelines for antimicrobial surveillance of all *Salmonella* isolates received, on a panel of 17 antimicrobials of 10 groups (penicillins, tetracyclines, cephalosporins, aminoglycosides, chloramphenicol, macrolides, fluoroquinolones, trimethoprim, sulfamethoxazole, and carbapenems). Previously, antimicrobial testing was rarely performed, usually upon request. Of over 1500 susceptibility tests performed, 38.33% of isolates tested were MDR. Most part of the MDR isolates were *S. enterica* serovar 1,4,[5],12:i:-, followed by *S. enterica* serovar Typhimurium (Figure 2-11). Approximately 47% of MDR isolates presented a profile of resistance to penicillins, tetracyclines, and sulfamethoxazole. Overall, resistance to sulfamethoxazole was the most frequent, followed

by penicillins, and tetracycline (Figure 2-12). Fortunately, resistance to carbapenems was not detected and resistance to macrolides remains sporadic.

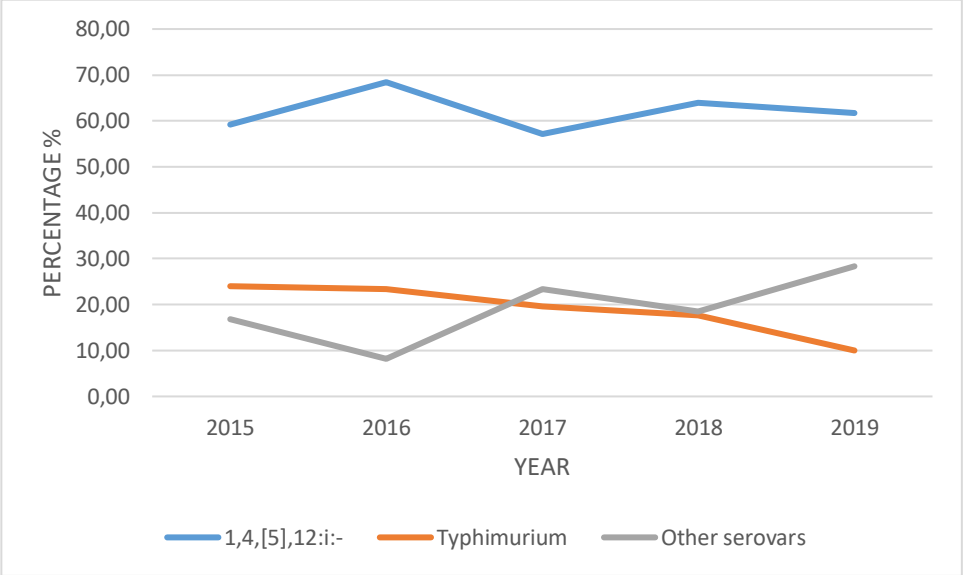


Figure 2-11 – Most relevant multidrug resistant serovars from January 2015 to June 2019.

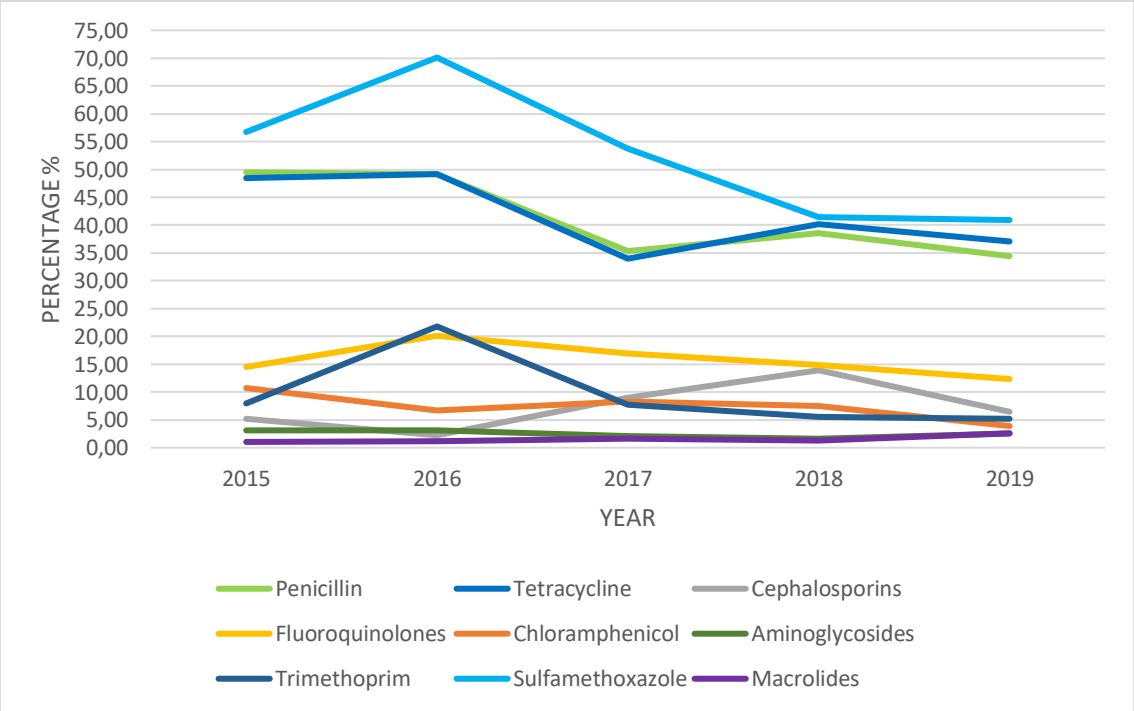


Figure 2-12 – Percentage of isolates resistant to each antibiotic group tested, from January 2015 to June 2019 (isolates from the end of June and July 2019 have not been tested to this date).

CHAPTER 3

The new era of surveillance of *Salmonella*

LS implemented MLVA at the National Reference Laboratory, contributed in the transition of surveillance of *Salmonella* to whole-genome sequencing, performed part of the experimental work and part of the bioinformatics analyses, interpreted data in this chapter, and is preparing the manuscript.

3. The new era of surveillance of *Salmonella*

The first *Salmonella* outbreak study enrolling WGS was published by Baker et al in 2011, and described a SNP based approach to discriminate outbreak related cases and non-related cases in a group of *S. enterica* serovar Montevideo isolates previously characterized by PFGE with a single pattern[147]. Other relevant outbreak studies followed, attesting the discriminatory power of WGS for outbreak resolution[29,55,148–150]. In the Portuguese NRL, the transitioning process from *Salmonella* classical typing methods to WGS began in 2014, in the context of a Master Thesis [151], where 16 isolates of *Salmonella enterica* were sequenced. The results demonstrated high levels of concordance between data previously obtained through classical typing methods, and *in silico* predictions. Moreover, even though the dataset was small, the incomparable high discriminatory power of WGS was demonstrated. Another project that helped prompt the application of WGS in the Portuguese NRL was INNUENDO[45], developed in collaboration with INSA and other European Reference Laboratories. In this context, 22 *S. enterica* serovar Enteritidis isolated in Portugal were part of the large pool provided by the project partners, and were sequenced and subjected to the INNUca pipeline, a platform for bacterial genome assembly, developed in the context of the project, and used throughout the analysis of the isolates in this PhD dissertation. Also, as an output of INNUENDO, a dynamic gene-by-gene scheme was proposed, allowing long-term surveillance and outbreak investigations in simultaneous. This was the approach chosen to analyse the isolates that are the focus of Chapters 3, 4, and 5.

Overall, on behalf of the role of INSA as a reference laboratory for human salmonellosis aiming at transiting from traditional typing methodologies to WGS-based approaches, in the last four years, 315 *Salmonella* spp isolates were sequenced in several contexts described below, excluding some isolates that were purposefully sequenced for the Master thesis already mentioned:

i) 9 isolates with undetermined serovar by the classical serotyping method for confirmation of serovar *in silico*;
ii) 6 isolates with azithromycin resistant phenotype to check if the underlying resistance mechanism was the same in the different serovars;
iii) 29 isolates for confirmation of 12 Urgent Inquiries;
iv) 47 isolates of monophasic *Salmonella enterica* serovar Typhimurium, for a surveillance study which is still ongoing;
v) 20 *Salmonella enterica* serovar Typhi to evaluate the genomic structure of the isolates from Portugal and integrate them with isolates from other

countries, especially endemic countries where this serovar is very frequent, to hopefully identify their origin (ongoing study);

vi) 78 *Salmonella enterica* serovar Enteritidis for the study described in detail in this chapter (not all isolates included for confidentiality reasons);

vii) 9 *Salmonella enterica* serovar Newport of an ongoing outbreak, and that will also be described in this chapter;

viii) 70 *Salmonella enterica* serovar Rissen for the study described in detail in Chapter 4 (not all isolates were included in the study due to quality issues);

ix) And finally, 36 *Salmonella* isolated from imported fresh poultry meat, for the study detailed in Chapter 5.

WGS has already proven to be a valuable tool for the Portuguese NRL, with particular relevance when classical serotyping proves inconclusive. All the isolates subjected to *in silico* serotyping had their serovars confirmed. Additionally, *in silico* serotyping also confirmed the serovars of all the isolates sequenced for other purposes, proving to be a reliable replacement of the labour intensive and expensive classical method. Genome sequencing has also been invaluable to investigate suspected outbreaks allowing the Portuguese NRL to start responding to Urgent Inquiries launched on the Epidemic Intelligence Information System for food- and waterborne diseases (EPIS-FWD), managed by the European Centre for Disease Prevention and Control (ECDC). As described in the points above, four major topics will be discussed in the following chapters. Chapter 4 consists in a study of the genetic diversity of *Salmonella enterica* serovar Rissen isolated in Portugal and its, contextualization among isolates from several countries. Chapter 5 also entails a study of genetic diversity of certain *Salmonella enterica* serovars, in this case serovars Heidelberg and Minnesota, isolated from fresh poultry meat imported to Portugal. In the present chapter, two examples of rather simple situations in which WGS was also applied will be described. First, an example of the transition from a traditional molecular typing technique, MLVA, to WGS for *S. enterica* serovar Enteritidis will follow. Secondly, a practical example, of an ongoing outbreak of *S. enterica* serovar Newport, in which WGS is playing an important role. The decision to present these two subjects in a single chapter with an unconventional format, is mainly due to the methodology being essentially the same when it comes to WGS, and the fact that, contrarily to chapters 4 and 5, both studies are ongoing and are not in a submitted manuscript format. All these topics share most part of the methodology, a WGS-based gene-by-gene analysis. Since chapters 4 and 5 are based on submitted manuscripts, the "Materials and Methods" section may seem repetitive (due to the unavoidable highly specific bioinformatics-based descriptions common to these studies), although minor changes were made to each manuscript to avoid such situation.

Likewise, to avoid excessive repetition, this chapter will include a general "Materials and Methods" section where the methodology common to both topics will be described.

3.1. Materials and Methods

The following WGS methodology was applied to the studies i) Retrospective study of the genetic diversity of *Salmonella enterica* serovar Enteritidis in Portugal; and ii) Ongoing outbreak of *Salmonella enterica* serovar Newport, 2019.

3.1.1.1. Whole Genome Sequencing (WGS)

For each isolate, WGS and bacterial *de novo* assembly were performed as previously described [152]. Total DNA was extracted from fresh cultures on the NucliSens easyMAG platform (Biomerieux), according to the manufacturer's instructions. Quantification and quality assessment of the purified DNA was performed using Qubit Fluorometer with hsDNA Assay Kit (Thermo Fisher Scientific) and agarose gel electrophoresis (0,8%), respectively. High-quality DNA samples were then used to prepare dual-indexed Nextera XT Illumina libraries (Illumina, USA) that were subsequently subjected to cluster generation and paired-end sequencing (2×250bp or 2x150bp) on either a MiSeq or a NextSeq 550 instrument (Illumina, USA) available at the Portuguese NIH, according to the manufacturer's instructions. All genomes were *de novo* assembled using the INNUca v3.1 pipeline (https://github.com/B-UMMI/INNUca), which consists of integrated modules for reads QA/QC, *de novo* assembly and post-assembly optimization steps. After reads' quality analysis using FastQC v0.11.5 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and cleaning with Trimmomatic v0.36 [153], genomes were assembled with SPAdes 3.11 [154] and subsequently improved using Pilon v1.18 [155].

3.1.1.2. Gene-by-gene analysis

Gene-by-gene analysis was performed by taking advantage of a publicly available panel of 8558 loci [156] derived from the EnteroBase *Salmonella* wgMLST schema [63], curated and prepared using the chewBBACA free software (https://github.com/B-UMMI/chewBBACA)[157], downloaded on August 2018 (http://doi.org/10.5281/zenodo.1323684). Allele calling was performed using chewBBACA v2.0.11 (https://github.com/B-UMMI/chewBBACA) with default parameters and a publicly available training file for *S. enterica* (https://github.com/mickaelsilva/prodigal training files). Exact and inferred matches were used to construct an allelic profile matrix, where other allelic classifications (https://github.com/B-UMMI/chewBBACA/wiki) were assumed as "missing" loci.

To evaluate the genetic relationship between strains, minimum spanning trees (MSTs) were constructed taking advantage of goeBURST algorithm[158] implemented in the PHYLOViZ online web-based tool [61], based on 100% shared loci between the analyzed strains (i.e., shared-genome MLST) [159].

3.1.2. Data availability

All raw sequence reads used in the present study were deposited in the European Nucleotide Archive (ENA) under the study accession number PRJEB32515.

3.2. Retrospective study of the genetic diversity of *Salmonella enterica* serovar Enteritidis in Portugal

3.2.1. Introduction

Salmonellosis is one of the main causes of gastroenteritis in the world, causing huge global burden of morbidity and mortality [86]. The genus *Salmonella* consists of two different species, *S. enterica* and *S. bongori*. *S. enterica* comprises six subspecies and more than 2600 different serovars, while *S. bongori* includes only 22 serovars [15]. However, human salmonellosis is mainly caused by *S. enterica* subspecies *enterica*, most frequently by *S. enterica* serovar Enteritidis [88]. In Portugal, this serovar accounted for more than 50% of salmonellosis cases for several years (chapter 2)[160]. The main source of infection by this serovar is the consumption of contaminated poultry and poultry products, mainly eggs. The implementation of control programmes in the poultry industry, involving for example vaccination against *S. enterica* serovar Enteritidis, resulted in the decline of *S. enterica* rates of infection for the past twelve years. Nonetheless, salmonellosis remains the second most common zoonosis in humans in the EU/EEA, with 22.9 cases per 100,000 inhabitants [88]. In fact, salmonellosis remains the most common cause of foodborne outbreaks, and since 2014 an increase in the number of outbreaks caused by *S. enterica* serovar Enteritidis has been observed, accounting for 60.3% of all *Salmonella* outbreaks in the EU in 2015 [161,162].

Since 2016, a Multi Locus Variable-Number Tandem Repeat Analysis (MLVA) scheme has been available for molecular subtyping of *S. enterica* serovar Enteritidis [32]. This typing technique, based on the measurement of the number of repeats in the variable number of tandem repeat regions (VNTRs) of the genome is a useful tool for outbreak detection and a strong contribution for rapid source detection [30]. However, whole-genome sequencing (WGS) has already proven to be even more discriminatory while establishing outbreak links and attributing sources of infection [59,163–165]. In this work, we evaluated the genetic diversity by MLVA of all *S. enterica* serovar Enteritidis isolates recovered from human clinical samples, from January 2012 to July 2019, sent to the NRL of Gastrointestinal

Infections at the Portuguese National Institute of Health (INSA) for serotyping. We then proceeded with WGS of 51 of those isolates to evaluate the clonality of two apparently endemic MLVA types that were the most frequent cause of salmonellosis in Portugal during the last few years. We also included in the WGS analysis, data of 15 isolates recovered in 2001, 2002, and 2003 presenting the same endemic MLVA profiles, that were already sequenced during the INNUENDO project.

3.2.2. Materials and Methods

From January 2012 to July 2019, 528 *S. enterica* serovar Enteritidis isolates were serotyped using the White-Kauffman-Le Minor classification scheme [15]. The isolates were sent to INSA for serotyping by several Portuguese hospitals and laboratories, covering several regions of the country. MLVA was performed using the 5-locus MLVA method published by Hopkins et al, 2011[166]. Isolates were grown overnight in non-selective medium at 37ºC. An aliquot of each isolate was prepared in 100 µl of sterile water and using a sterile 1 µl loop to pick the colonies. The microtubes were placed in a water bath at 100ºC for 10 minutes, followed by cooling in ice for 1 minute and centrifugation for 10 minutes at 10,000 rpm. The supernatant was used in the amplification reaction. PCR products were diluted 1:10 and subjected to capillary electrophoresis [32,166]. In order to evaluate the genetic diversity of the two most frequent MLVA profiles in Portugal, 51 isolates recovered in several years were selected for WGS (methods previously described). Fifteen additional isolates, recovered in 2001 and 2002, previously sequenced and assembled following the same methodology described above, in the context of the INNUENDO project, were also included in the subsequent analysis (Table 3-1). The ResFinder 3.1 web server[49] (https://cge.cbs.dtu.dk/services/ResFinder/) was used to identify acquired antimicrobial resistance genes and/or chromosomal mutations, using the default threshold (Table 3-1). PlasmidFinder 2.0[50] was used to detect and characterize the plasmids present in these isolates (https://cge.cbs.dtu.dk/services/PlasmidFinder/), also using the default threshold (Table 3-1).

Table 3-1 – Sixty-six *Salmonella enterica* serovar Enteritidis sequenced isolates with MLVA profiles 3-10-5-4-1 and 3-11-5-4-1. Since genome assembly of isolate PT50 failed, it was excluded from the subsequent analysis.

ID	Year	MLVA profile	Sample type	Region	Age	Gender	Genotype	Plasmids
PT1	2012	3-10-5-4-1	Faeces	Setúbal	21	M	aac(6')-Iaa, gyrA p.D87Y	IncFIB(S), IncFII(S)
PT2	2012	3-10-5-4-1	Faeces	Setúbal	7	F	aac(6')-Iaa, gyrA p.D87Y	IncFIB(S), IncFII(S), Col440I
PT3	2012	3-10-5-4-1	Faeces	Lisboa	7	M	aac(6')-Iaa, gyrA p.S83F	IncFIB(S), IncFII(S)
PT4	2012	3-10-5-4-1	Faeces	Lisboa	5	F	aac(6')-Iaa, gyrA p.D87Y	IncFIB(S), IncFII(S)
PT5	2012	3-10-5-4-1	Faeces	Faro	Unknown	M	aac(6')-Iaa, gyrA p.D87Y	IncFIB(S), IncFII(S)
PT6	2013	3-10-5-4-1	Faeces	Porto	1	F	aac(6')-Iaa, gyrA p.S83F	IncFIB(S), IncFII(S)
PT7	2013	3-10-5-4-1	Faeces	Braga	4	M	aac(6')-Iaa, gyrA p.S83F	IncFIB(S), IncFII(S)
PT8	2013	3-10-5-4-1	Faeces	Beja	6	M	aac(6')-Iaa, blaTEM-1B	IncFIB(S), IncFII(S), IncX1
PT9	2013	3-10-5-4-1	Faeces	Lisboa	3	F	aac(6')-Iaa, gyrA p.D87Y	IncFIB(S), IncFII(S)
PT10	2013	3-10-5-4-1	Faeces	Lisboa	10	M	aac(6')-Iaa, gyrA p.D87Y	IncFIB(S), IncFII(S)
PT11	2013	3-10-5-4-1	Unknown	Ponta Delgada	13	M	aac(6')-Iaa, gyrA p.D87Y	IncFIB(S), IncFII(S)
PT12	2013	3-10-5-4-1	Faeces	Beja	6	F	aac(6')-Iaa, gyrA p.D87Y	IncFIB(S), IncFII(S)
PT13	2014	3-10-5-4-1	Faeces	Setúbal	3	M	aac(6')-Iaa, gyrA p.S83F	IncFIB(S), IncFII(S)
PT14	2014	3-10-5-4-1	Faeces	Beja	6	M	aac(6')-Iaa, gyrA p.D87Y	IncFIB(S), IncFII(S)
PT15	2014	3-10-5-4-1	Faeces	Braga	1	F	aac(6')-Iaa, gyrA p.D87Y	IncFIB(S), IncFII(S)
PT16	2014	3-10-5-4-1	Faeces	Lisboa	25	M	aac(6')-Iaa, gyrA p.D87Y	IncFIB(S), IncFII(S)
PT17	2014	3-10-5-4-1	Faeces	Porto	6	F	aac(6')-Iaa, blaTEM-1B	IncFIB(S), IncFII(S), IncX1

ID	Year	MLVA profile	Sample type	Region	Age	Gender	Genotype	Plasmids
PT18	2015	3-10-5-4-1	Unknown	Unknown	Unknown	M	aac(6')-Iaa, gyrA p.D87Y	IncFIB(S), IncFII(S), IncI1
PT19	2015	3-10-5-4-1	Faeces	Lisboa	4	M	aac(6')-Iaa	IncFIB(S), IncFII(S)
PT20	2015	3-10-5-4-1	Blood	Porto	67	M	blaTEM-1B	IncFIB(S), IncFII(S), IncX1
PT21	2015	3-10-5-4-1	Faeces	Castelo Branco	3	M	aac(6')-Iaa	IncFIB(S), IncFII(S)
PT22	2016	3-10-5-4-1	Faeces	Porto	57	F	aac(6')-Iaa, gyrA p.D87Y	IncFIB(S), IncFII(S)
PT23	2016	3-10-5-4-1	Faeces	Porto	4	F	aac(6')-Iaa, gyrA p.D87Y	IncFIB(S), IncFII(S)
PT24	2016	3-10-5-4-1	Faeces	Lisboa	9	F	aac(6')-Iaa, gyrA p.D87Y	IncFIB(S), IncFII(S)
PT25	2016	3-10-5-4-1	Faeces	Porto	5	F	aac(6')-Iaa	IncFIB(S), IncFII(S)
PT26	2016	3-10-5-4-1	Faeces	Bragança	8	F	aac(6')-Iaa	IncFIB(S), IncFII(S)
PT27	2016	3-10-5-4-1	Faeces	Unknown	3	M	aac(6')-Iaa	IncFIB(S), IncFII(S)
PT28	2016	3-10-5-4-1	Faeces	Unknown	3	F	aac(6')-Iaa	IncFIB(S), IncFII(S)
PT29	2017	3-10-5-4-1	Faeces	Unknown	11	F	aac(6')-Iaa, gyrA p.S83F	IncFIB(S), IncFII(S)
PT30	2017	3-10-5-4-1	Faeces	Porto	5	M	aac(6')-Iaa, gyrA p.D87Y	IncFIB(S), IncFII(S)
PT31	2017	3-10-5-4-1	Faeces	Unknown	Unknown	F	aac(6')-Iaa, gyrA p.D87Y	IncFIB(S), IncFII(S)
PT32	2017	3-10-5-4-1	Faeces	Ponta Delgada	Unknown	F	aac(6')-Iaa	IncFIB(S), IncFII(S)
PT33	2017	3-10-5-4-1	Faeces	Ponta Delgada	Unknown	M	aac(6')-Iaa	IncFIB(S), IncFII(S)
PT34	2017	3-10-5-4-1	Faeces	Vila Real	7	F	aac(6')-Iaa, gyrA p.D87Y	IncFIB(S), IncFII(S)
PT35	2017	3-10-5-4-1	Faeces	Ponta Delgada	Unknown	M	aac(6')-Iaa, gyrA p.D87Y	IncFIB(S), IncFII(S)

ID	Year	MLVA profile	Sample type	Region	Age	Gender	Genotype	Plasmids
PT36	2018	3-10-5-4-1	Faeces	Lisboa	7	M	aac(6')-Iaa, gyrA p.D87Y	IncFIB(S), IncFII(S)
PT37	2018	3-10-5-4-1	Faeces	Porto	7	M	aac(6')-Iaa, gyrA p.D87Y	IncFIB(S), IncFII(S)
PT38	2018	3-10-5-4-1	Faeces	Bragança	8	M	aac(6')-Iaa, gyrA p.D87Y	IncFIB(S), IncFII(S)
PT39	2018	3-10-5-4-1	Faeces	Unknown	5	M	aac(6')-Iaa, gyrA p.D87Y	IncFIB(S), IncFII(S)
PT40	2001	3-10-5-4-1	Faeces	Lisboa	2	F	aac(6')-Iaa	IncFIB(S), IncFII(S)
PT41	2001	3-10-5-4-1	Faeces	Lisboa	Unknown	M	aac(6')-Iaa	IncFIB(S), IncFII(S)
PT42	2001	3-10-5-4-1	Faeces	Lisboa	Unknown	Unknown	aac(6')-Iaa	IncFIB(S), IncFII(S)
PT43	2001	3-10-5-4-1	Faeces	Lisboa	0	M	aac(6')-Iaa	IncFIB(S), IncFII(S)
PT44	2002	3-10-5-4-1	Blood	Lisboa	35	M	blaTEM-1B, sul1, sul2, aac(6')-Iaa, aadA1, aph(3'')-Ib, aph(6)-Id, dfrA1	IncFIB(S), IncFII(S), IncI1, IncQ1
PT45	2002	3-10-5-4-1	Faeces	Setúbal	11	M	aac(6')-Iaa	IncFIB(S), IncFII(S)
PT46	2003	3-10-5-4-1	Faeces	Lisboa	6	M	aac(6')-Iaa	IncFIB(S), IncFII(S)
PT47	2001	3-10-5-4-1	Faeces	Leiria	3	F	aac(6')-Iaa	IncFIB(S), IncFII(S)
PT48	2002	3-10-5-4-1	Faeces	Lisboa	35	M	aac(6')-Iaa, gyrA p.D87Y	IncFIB(S), IncFII(S)
PT49	2012	3-11-5-4-1	Faeces	Setúbal	6	F	aac(6')-Iaa, gyrA p.D87Y	IncFIB(S), IncFII(S)
PT50	2015	3-11-5-4-1	Faeces	Setúbal	5	F	ND	ND
PT51	2016	3-11-5-4-1	Faeces	Unknown	8	M	aac(6')-Iaa, gyrA p.D87Y	IncFIB(S), IncFII(S)
PT52	2016	3-11-5-4-1	Faeces	Lisboa	8	M	aac(6')-Iaa, gyrA p.D87Y	IncFIB(S), IncFII(S)
PT53	2016	3-11-5-4-1	Faeces	Lisboa	3	M	aac(6')-Iaa, gyrA p.D87Y	IncFIB(S), IncFII(S)

ID	Year	MLVA profile	Sample type	Region	Age	Gender	Genotype	Plasmids
PT54	2016	3-11-5-4-1	Faeces	Lisboa	5	M	aac(6')-Iaa, gyrA p.D87Y	IncFIB(S), IncFII(S)
PT55	2016	3-11-5-4-1	Faeces	Lisboa	7	F	aac(6')-Iaa, gyrA p.D87Y	IncFIB(S), IncFII(S)
PT56	2016	3-11-5-4-1	Faeces	Unknown	Unknown	Unknown	aac(6')-Iaa, gyrA p.D87Y	IncFIB(S), IncFII(S)
PT57	2016	3-11-5-4-1	Faeces	Unknown	Unknown	Unknown	aac(6')-Iaa, gyrA p.D87Y	IncFIB(S), IncFII(S)
PT58	2016	3-11-5-4-1	Faeces	Unknown	Unknown	Unknown	aac(6')-Iaa, gyrA p.D87Y	IncFIB(S), IncFII(S)
PT59	2016	3-11-5-4-1	Faeces	Unknown	Unknown	Unknown	aac(6')-Iaa, gyrA p.D87Y	IncFIB(S), IncFII(S)
PT60	2018	3-11-5-4-1	Faeces	Setúbal	60	F	aac(6')-Iaa, gyrA p.D87Y	IncFIB(S), IncFII(S)
PT61	2001	3-11-5-4-1	Faeces	Porto	26	M	aac(6')-Iaa	IncFIB(S), IncFII(S), Col156
PT62	2002	3-11-5-4-1	Faeces	Setúbal	70	F	aac(6')-Iaa	IncFIB(S), IncFII(S)
PT63	2002	3-11-5-4-1	Faeces	Setúbal	10	M	aac(6')-Iaa	IncFIB(S), IncFII(S)
PT64	2002	3-11-5-4-1	Faeces	Lisboa	5	F	aac(6')-Iaa	IncFIB(S), IncFII(S)
PT65	2003	3-11-5-4-1	Faeces	Lisboa	6	F	aac(6')-Iaa, gyrA p.D87Y	IncFIB(S), IncFII(S)
PT66	2002	3-11-5-4-1	Faeces	Porto	Unknown	F	aac(6')-Iaa, gyrA p.D87Y	IncFIB(S), IncFII(S)

3.2.3. Results

For several years the number of *S. enterica* serovar Enteritidis had been decreasing in Portugal (chapter 2), until in 2015 the trend inverted reaching a new peak in 2017 (Figure 3-1). Among the 528 isolates analyzed, MLVA revealed a total of 56 different profiles (Figure 3-2). Between January 2012 and July 2019, the three most common profiles were 3-10-5-4-1, 3-11-5-4-1, and 2-10-7-3-2, corresponding to almost 46% of all cases (Figure 3-3). Remarkably, profiles 3-10-5-4-1, 3-11-5-4-1, and 3-9-5-4-1, seem to be endemic in Portugal, consistently appearing throughout the entire period of study (Figure 3-4). Moreover, profiles 3-10-5-4-1 and 3-11-5-4-1, increased quite significantly in 2013 and 2016 respectively, probably related with undetected national outbreaks. Interestingly, the most frequent profiles in 2017 were 2-10-7-3-2, 2-9-7-3-2, and 2-11-7-3-2, profiles previously infrequent or undetected. No relevant relation between a specific profile and demographic data was found.

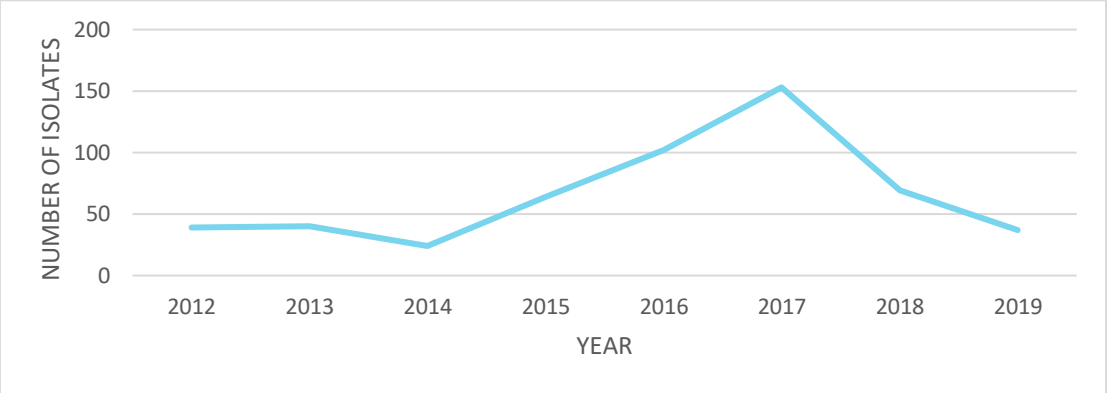


Figure 3-1 – Trend of *Salmonella enterica* serovar Enteritidis received at the National Institute of Health in Portugal.

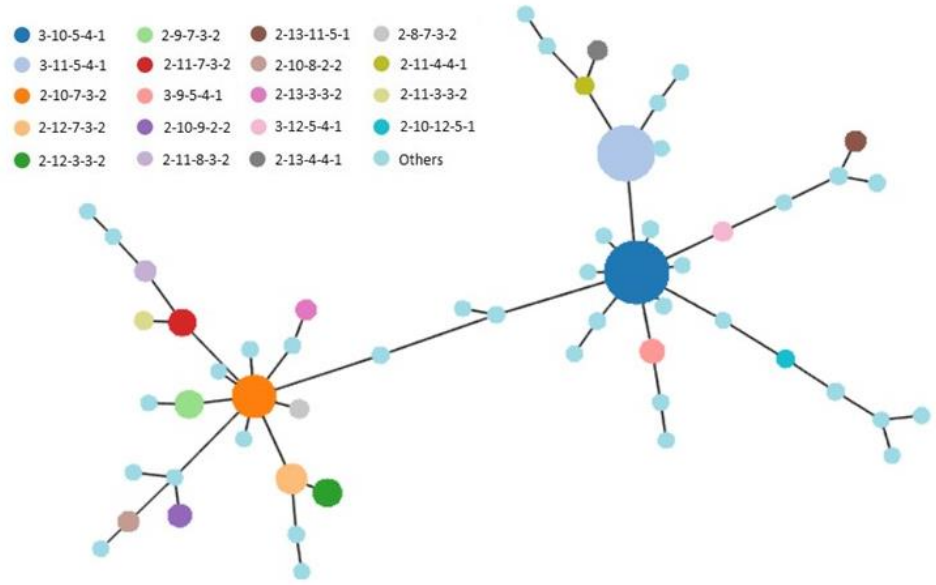


Figure 3-2 – Minimum spanning tree of the 56 different MLVA profiles detected among the 528 isolates analysed (2012-2019). For simplification purposes, rare profiles are in the same colour.

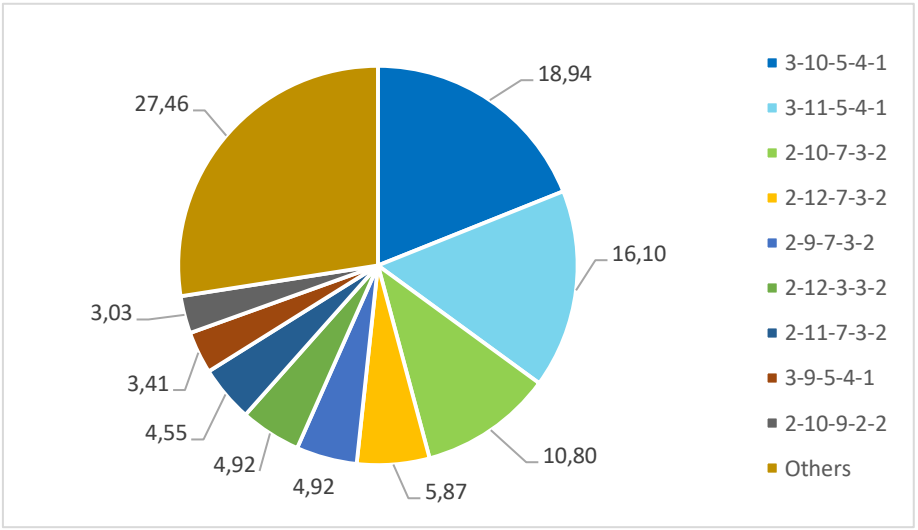


Figure 3-3 – Percentage of the most common MLVA profiles in Portugal, between 2012 and July 2019.

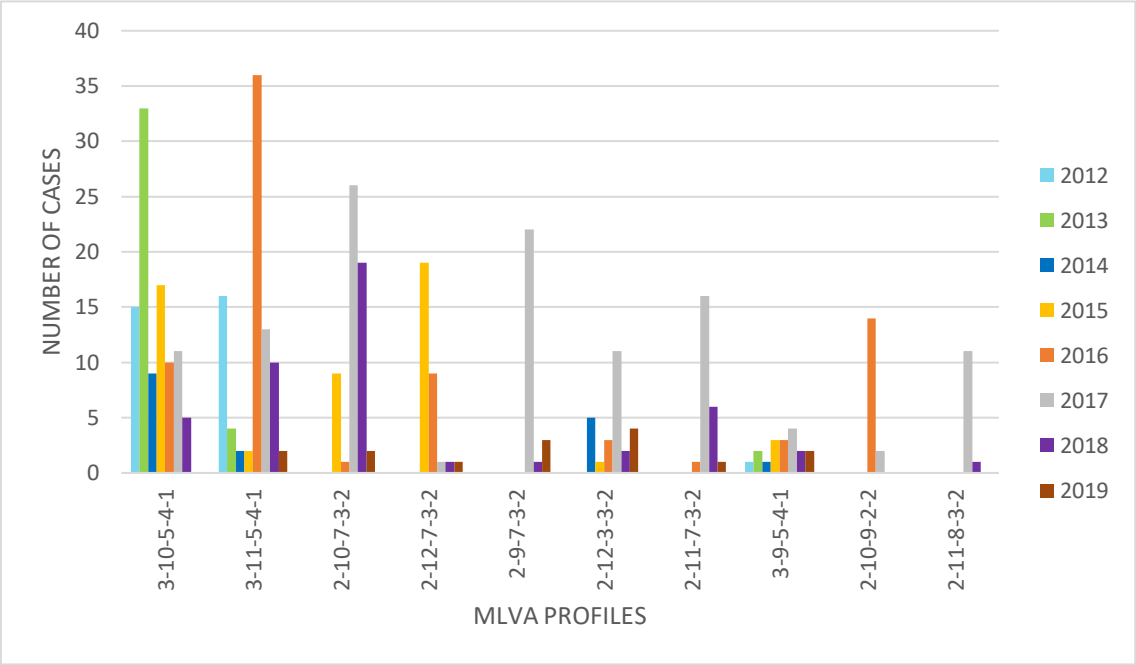


Figure 3-4 – Number of cases of the 10 most frequent profiles per year.

In silico MLST analysis of the 65 sequenced isolates revealed that all enrolled isolates belonged to ST11. In order to assess their genomic diversity and phylogenetic relationships, using a gene-by-gene approach, a MST based on 3103 shared loci between all 65 isolates was generated. The MST revealed an overall genetic relatedness between the two MLVA profiles (Figure 5-5). In fact, some isolates of profile 3-11-5-4-1 were closer to isolates of profile 3-10-5-4-1 than to isolates of the same profile. Interestingly, profile 3-11-5-4-1 reveals a considerably higher degree of diversity than profile 3-10-5-4-1. This profile included two isolates (PT66 and PT62) very distant from the others, both recovered in 2002. When comparing the other isolates with PT66, we found 111 exclusive loci that will require further analysis. Additionally, nine isolates were closely related, confirming a suspected outbreak that occurred in 2016. In parallel, two additional outbreaks involving isolates of profile 3-10-5-4-1, seem to have occurred.

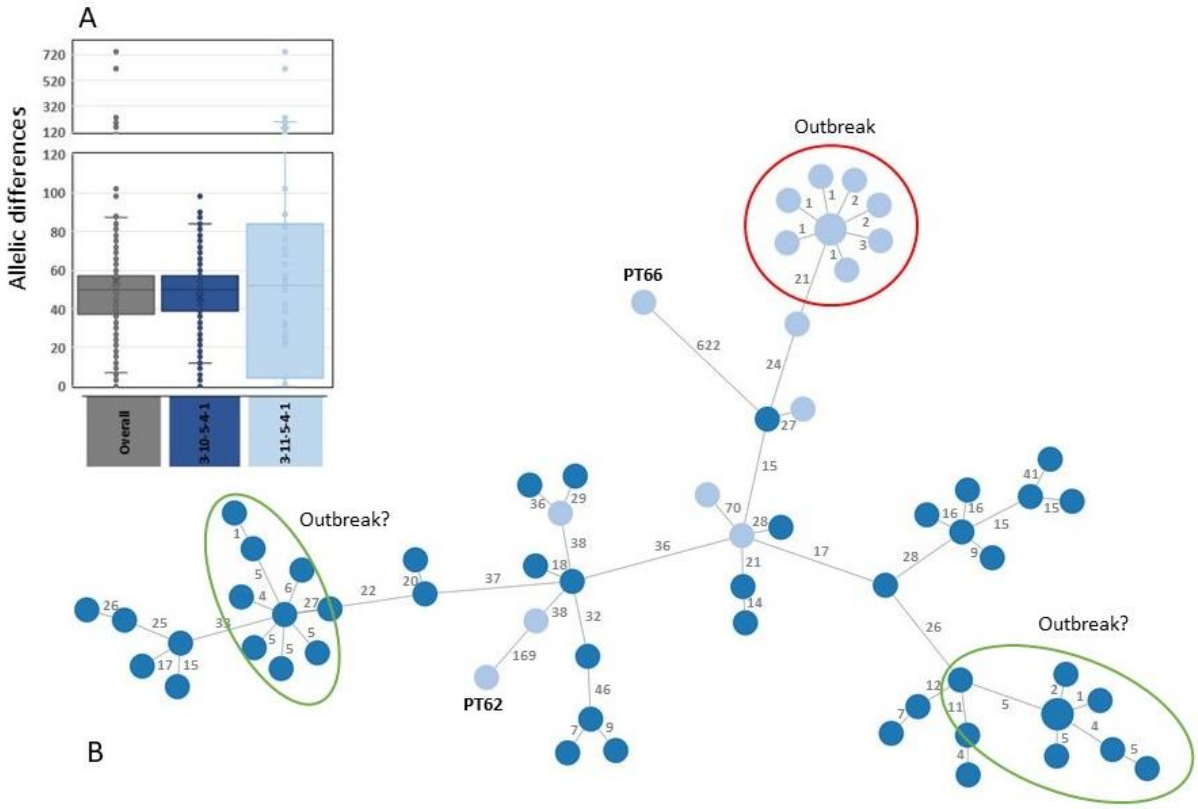


Figure 3-5 – Global phylogeny of *S. enterica* serovar Enteritidis isolates with MLVA profiles 3-10-5-4-1 and 3-11-5-4-1, based on a dynamic gene-by-gene approach using a wgMLST schema with 8558 loci. **A** Plot depicting the distribution of the number of allelic differences between all 65 isolates (PT50 assembly failed) and between isolates of the same MLVA profile. **B** The minimum spanning tree enrolling 65 isolates retrieved in Portugal, was constructed based on the allelic diversity of 3103 shared loci, using the goeBURST algorithm implemented in the PHYLOViZ online platform. Filled circles (nodes) represent unique allelic profiles, and are colored according to the MLVA profile. The allele differences between isolates are represented by a number next to each line. Circled in red, a confirmed outbreak that occurred in 2016, and in green two potential clusters not previously detected.

3.2.4. Discussion

S. enterica serovar Enteritidis is one of the most common serovars in Portugal. It is also the serovar most frequently associated with outbreaks in the EU[85,162]. MLVA usefulness to the confirmation of outbreaks of *S. enterica* serovar Typhimurium and *S. enterica* serovar Enteritidis is well established[164,167]. In this work, we characterized 528 *S. enterica* serovar Enteritidis isolates recovered from January 2012 to July 2019, in Portugal. We identified 56 different MLVA profiles, although 46% of all cases belonged to only three profiles, 3-10-5-4-1, 3-11-5-4-1, and 2-10-7-3-2, that constantly appear throughout the period of study This fact points to the possibility that these three profiles are endemic in Portugal. Also, the huge

variability of profiles among sporadic cases may reveal a great multitude of sources, maybe including importation of isolates, but additional studies would be required to support this hypothesis.

Given the representativity of MLVA profiles 3-10-5-4-1 or 3-11-5-4-1 in Portugal, we performed WGS on several isolates to access their genomic diversity and conclude about the potential bias associated with the use of the MLVA for molecular epidemiology studies. A gene-by-gene analysis of 65 isolates revealed, as expected, that multiple isolates within and between these two MLVA profiles are very closely related considering the low number of allelic differences between them. Even so, WGS undoubtedly has higher discriminatory power than MLVA, as 63 nodes were detected among the 65 isolates, although these were clustered in only two distinct MLVA profiles. This is easily explained by the fact that while MLVA typing evaluates the number of tandem repeats present in five alleles, the WGS schema, upon which all the analysis of this dissertation is based, enrols the allelic diversity among >3000 shared loci (from a total of 8558 loci in the wgMLST schema). Interestingly, profile 3-11-5-4-1 presents higher genomic diversity than profile 3-10-5-4-1, and some isolates of profile 3-11-5-4-1 are closer to isolates of profile 3-10-5-4-1, although the later observation is not very surprising as these two profiles differ by only one repetition in one of the five loci analysed.

In 2016, MLVA proved to be an efficient and timely tool to clarify a suspected outbreak, involving 11 isolates, nine belonging to profile 3-11-5-4-1, and two to profile 2-10-9-2-2. WGS confirmed the extremely high genetic relatedness of the nine isolates sharing the same MLVA profile (Figure 3-5). In retrospective, WGS also revealed other potential clusters that were not signalized by the Public Health Authorities. That may be the case of the two clusters signalled in Figure 3-5, each comprising eight isolates, that differ in no more than five or six alleles within the 3103 shared loci. Both these clusters correspond to cases in children under 8 years old, from several regions of the country. The resistance genotypes and plasmids present in the isolates of these clusters are identical in both cases (Table 3-1). Eventually, surveillance of *S. enterica* serovar Enteritidis through WGS, in real time, will now enable the NRL to inform the Public Health Authorities of possible outbreaks that require investigation.

In summary, *S. enterica* serovar Enteritidis, once thought to be a highly clonal population, appears to have a higher heterogenicity, revealed by WGS. This methodology already presents higher discriminatory power than molecular methods once considered gold standard, even allowing the detection of previously undetectable outbreaks.

3.3. Ongoing outbreak of *Salmonella enterica* serovar Newport, 2019

3.3.1. Introduction

Although *Salmonella enterica* serovar Newport is rarely reported in Portugal (chapter 2), it was the 5th most reported serovar causing salmonellosis in 2016 in the EU. Nevertheless, more than one third of the cases were associated with travel outside the EU[87]. In the United States of America (USA), this serovar was the third most common cause of salmonellosis in 2013[168]. Furthermore, several outbreaks of *S. enterica* serovar Newport have been detected through the years, and the type of food implicated, is quite diverse. Susceptible isolates are usually associated with produce, and resistant isolates with products of animal origin, such as, ground beef, unpasteurized cheese, and poultry and their products[168]. In 2019, a sudden increase of the number of cases of *S. enterica* serovar Newport identified in the Portuguese NRL, gave rise to the suspicion of an outbreak. Therefore, nine isolates from different regions of the country and recovered from different sample types (faeces, blood, urine and ascitic fluid) were selected for WGS to confirm the suspected outbreak.

3.3.2. Epidemiological information

In July of 2019 (Figure 3-6; week 27 to week 30), 15 *S. enterica* serovar Newport were detected in the NRL. In total, until week 34 (August), 30 cases were identified. This fact prompted the suspicion of the existence of an outbreak since only 44 cases had been identified, for the past 10 years (Figure 3-7).

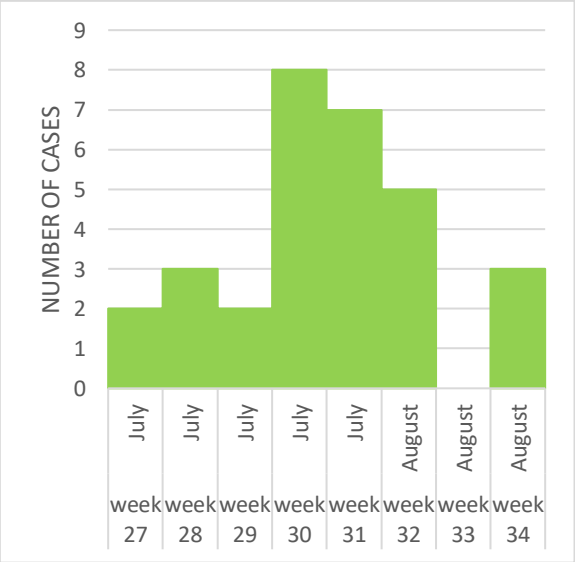


Figure 3-6 – Distribution of suspected outbreak cases by month, July to August 2019 (n=30).

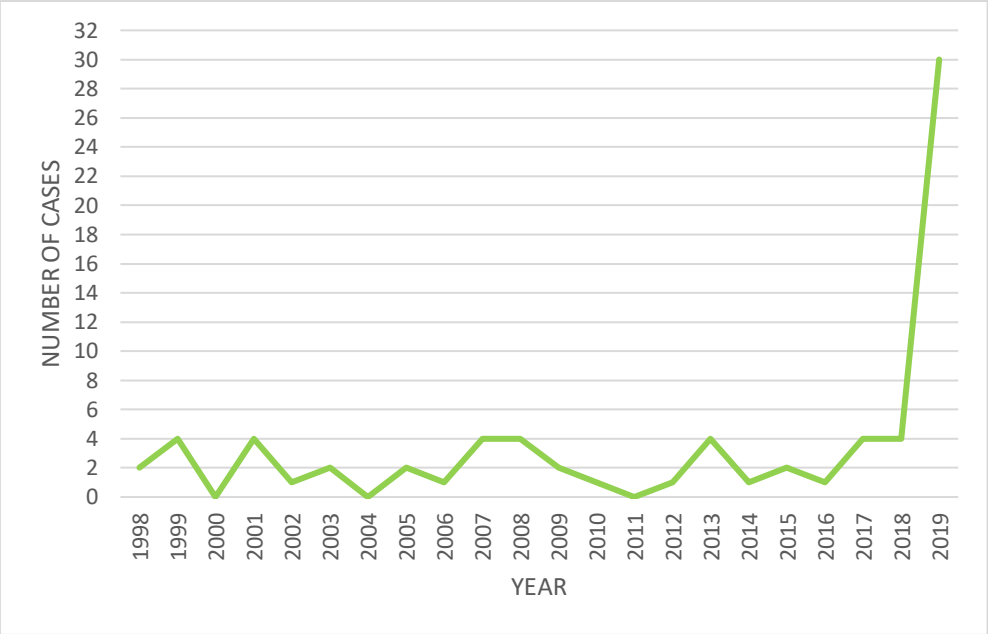


Figure 3-7 – Trend of *Salmonella enterica* serovar Newport identified in the Portuguese National Reference Laboratory, 1998-2019.

Patients ages ranged from 0 to 91 years old with most of the cases affecting patients over 65 years old. Most part of the cases were from Lisboa, but there were also cases from other regions of the country (Figure 3-8). All isolates were pan-susceptible (susceptible to all antimicrobials tested). The isolates selected for WGS were from Lisboa (n=5), Leiria (n=1), Funchal (n=1), Évora (n=1) e Porto (n=1).

Figure 3-8 – Cases' distribution in the Portuguese continent, and Madeira and Açores archipelagos.

3.3.3. WGS-based analysis

In silico MLST revealed that one of the nine isolates selected for WGS analysis belonged to ST2368, while the other eight isolates belonged to ST118. The gene-by-gene analysis revealed ST118 isolates shared 3728 loci. In fact, 4 isolates shared the same genomic background and only presented 1 to 8 Allele Differences (AD) from the other isolates. The isolate with a different ST presented 425 AD from the closest isolate (Figure 3-9).

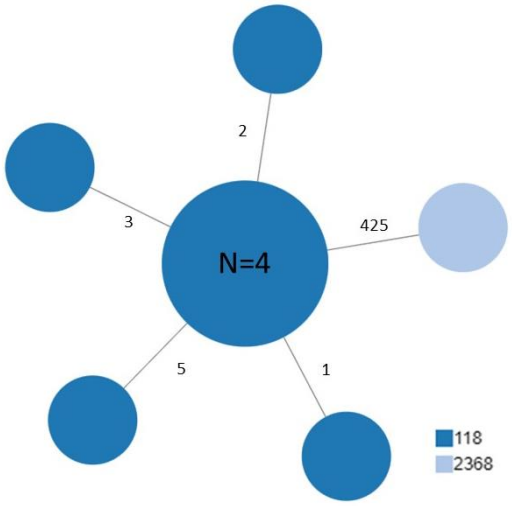


Figure 3-9 – Allelic diversity analysis of *Salmonella enterica* serovar Newport ST118 and ST2368 based on a gene-by-gene approach.

3.3.4. Conclusions

Given the epidemiological data available, and the WGS data obtained, there is a strong possibility that all the cases of *S. enterica* serovar Newport ST118 that occurred in the 8 weeks of study are closely related. ST118 belongs to lineage Newport-III that mostly encloses pan-susceptible isolates[169]. In fact, all *S. enterica* serovar Newport isolates recovered in the period of the study were pan-susceptible. Infection caused by susceptible isolates is usually less severe, and tends to affect young children, the elderly, and immunocompromised patients. In this case, the most affected age group was ≥65 years old. This information was shared with Portuguese Public Health Authorities that will conduct a proper investigation of each case. Although this work is still ongoing, as other isolates will soon be sequenced, it constitutes a very practical example of what WGS brought to the NRL, namely the capacity to establish genetic links that allow the Epidemiologists to conduct targeted inquires when suspecting of outbreaks, and propose measures to identify the infectious sources and stop the transmission chains.

CHAPTER 4

Multidrug resistant *Salmonella enterica* serovar Rissen clusters detected in Azores archipelago, Portugal

Manuscript published in the International Journal of Genomics

Silveira L., Pinto M., Isidro J., Pista A., Themudo P., Vieira L., Machado J. and Gomes J.P., "Multidrug resistant *Salmonella enterica* serovar Rissen clusters detected in Azores archipelago, Portugal", International Journal of Genomics, vol. 2019, Article ID 1860275, 9 pages, 2019. https://doi.org/10.1155/2019/1860275.

LS contributed to the design of the study, performed most of the experimental work and part of the bioinformatics analyses, interpreted data and wrote the manuscript.

4. Multidrug resistant *Salmonella enterica* serovar Rissen clusters detected in Azores archipelago, Portugal

4.1. Abstract

Gastrointestinal infections caused by Nontyphoidal *Salmonella* (NTS) remain one of the main causes of foodborne illness worldwide. Within the multiple existing *Salmonella enterica* serovars, the serovar Rissen is rarely reported, particularly as a cause of human salmonellosis. Between 2015 and 2017, the Portuguese National Reference Laboratory of Gastrointestinal Infections observed an increase in the number of clinical cases caused by multidrug resistant (MDR) *S. enterica* serovar Rissen, particularly from the Azores archipelago. In the present study, we analysed by whole genome sequencing (WGS) all clinical, animal, food and environmental isolates received up to 2017 in the Portuguese Reference Laboratories. As such, through a gene-by-gene analysis, we aimed to identify potential epidemiological clusters linking clinical and samples from multiple sources, while gaining insight into the genetic diversity of *S. enterica* serovar Rissen. We also investigated the genetic basis driving the observed multidrug resistance. By integrating 60 novel genomes with all publicly available serovar Rissen genomes, we observed a low degree of genetic diversity within this serovar. Nevertheless, the majority of Portuguese isolates showed high degree of genetic relatedness and a potential link to pork production. An in-depth analysis of these isolates revealed the existence of two major clusters from the Azores archipelago composed of MDR isolates, most of which were resistant to at least five antimicrobials. Considering the well-known spread of MDR between gastrointestinal bacteria, the identification of MDR circulating clones should constitute an alert to public health authorities. Finally, this study constitutes the starting point for the implementation of the "One Health" approach for *Salmonella* surveillance in Portugal.

4.2. Introduction

Nontyphoidal *Salmonella* (NTS) are zoonotic pathogens that remain one of the main causes of gastrointestinal infection and one of the most important causes of foodborne illness around the world. Annually, an estimated 93.8 million cases of gastroenteritis are caused by NTS worldwide, of which 80.3 million are considered foodborne [86]. Salmonellosis is also estimated to be responsible for 155,000 deaths each year [86]. In 2015, over 95,000 cases of salmonellosis were reported in the European Union (EU) [162]. Although more than 2600 *Salmonella enterica* serovars have been identified to date, most of the cases in developed countries are caused by *S. enterica* serovar Enteritidis or *S. enterica* serovar Typhimurium, accounting for 63% of all reported cases in the EU in 2012 [170,171]. In contrast, *S. enterica* serovar Rissen is rarely reported as a cause of human salmonellosis in Europe, but is frequently reported in the United States of America and particularly in Asia [172–174]. As a matter of fact, between 2014 and 2016, this serovar was not even among the 20 most frequently reported serovars responsible for human salmonellosis in the EU/EEA [87]. In Portugal, only 31 cases of human salmonellosis caused by *S. enterica* serovar Rissen were identified is a 12 year period (2000-2012) [175]. However, this is one of the most commonly reported serovars in pigs and pork, in several European and Asian countries [176–183]. This serovar has also been isolated less frequently from other sources, namely beef, chicken and seafood [180,181,184,185]. In Portugal, it has been identified in several studies, not only in pig and pork, but also in beef, chicken and wild animals [177,186–188].

Salmonella serotyping has been the gold standard for *Salmonella* surveillance for years, allowing monitoring of shifts in prevalence of certain serovars in specific regions, which are strong indicatives of existing clusters [54,163]. Until recently, *Salmonella* outbreak investigations have been conducted using different molecular typing methods, such as phage typing, MLVA or PFGE [164,189,190]. With the development of next generation sequencing technologies, those classical typing methods are being used to a lesser extent and genomic approaches based on single nucleotide polymorphisms and gene-by-gene analysis are progressing as frontline tools for high-resolution isolate characterization and outbreak detection [64,191,192].

Between 2015 and 2017, an increase in the number of *S. enterica* serovar Rissen isolated from clinical samples, especially multidrug resistant (MDR) isolates from the Azores archipelago, was observed. We used whole-genome sequencing (WGS) to analyse all clinical isolates received up to 2017 at the National Reference Laboratory (NRL) of Gastrointestinal Infections at the Portuguese National Institute of Health (INSA), in order to gain insight into the genetic diversity of *S. enterica* serovar Rissen Portuguese (PT) isolates

and eventually identify suspected outbreaks. All animal, food and environmental *S. enterica* serovar Rissen isolates received at the NRL from the National Institute of Agrarian and Veterinary Research (INIAV), between 2014 and 2017, were also included in this work to investigate potential sources of infection.

4.3. Materials and Methods

4.3.1. Bacterial isolate typing and antimicrobial susceptibility testing

S. enterica isolates included in the present study were obtained from the INSA and INIAV culture collections. The isolates were serotyped by the slide agglutination method, according to the Kauffman-White-Le Minor scheme [15]. In total, 60 *S. enterica* serovar Rissen isolates, collected from 2014 to 2017 in Portugal, were selected for WGS (Supplementary Table 4-1). Twenty-two were isolated from human clinical samples, 14 from animals, mostly pigs (N=9) but also bovine (N=4), and chicken (N=1), 22 from food products of animal origin, and 2 from environmental samples.

Antimicrobial Susceptibility Testing was performed by disc diffusion, following the European Committee on Antimicrobial Susceptibility Testing (EUCAST) [193] recommendations, on a panel of 17 antimicrobials: ampicillin, amoxicillin-clavulanic acid, cefoxitin, cefotaxime, ceftazidime, ceftriaxone, cefepime, meropenem, pefloxacin, nalidixic acid, gentamicin, azithromycin, tetracycline, tigecycline, chloramphenicol, sulfamethoxazole, and trimethoprim. Results were interpreted using current EUCAST breakpoints and Epidemiological cut-off values [193–197].

4.3.2. Whole Genome Sequencing and genome characterization

DNA was extracted from each PT isolate using the NucliSens easyMAG platform (bioMérieux, France) for total nucleic acid extraction according to the manufacturer's instructions. DNA was then subjected to the NexteraXT library preparation protocol (Illumina, USA) prior to paired-end sequencing (2x250 bp or 2x150 bp) on either a MiSeq or a NextSeq 550 instrument (Illumina, USA) according to the manufacturer's instructions (detailed in Supplementary Table 4-1).

All genome sequences were assembled using the INNUca v3.1 pipeline (https://github.com/B-UMMI/INNUca) an integrative bioinformatics pipeline for read quality analysis and *de novo* genome assembly. Read quality analysis and improvement is performed respectively using FastQC v0.11.5 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and Trimmomatic v0.36 [153] (with sample-specific read trimming criteria determined automatically based on FasQC

report). Genomes are assembled with SPAdes v3.10 (Bankevich et al., 2012) and subsequently polished using Pilon v1.18[155], with QA/QC statistics (such as depth of coverage and number of contigs) being monitored and reported throughout the analysis. *In silico* MLST prediction is performed using the *mlst* v2.4 software (https://github.com/tseemann/mlst). The full characterization of isolates, including specimen type and source, sampling date, sequence type (ST), final genome assembly sizes and depth of coverage values are reported in Supplementary Table 4-1.

For all isolates, the serovar was predicted *in silico* using SeqSero software [48]. As a means to potentiate strain discrimination, assemblies were also analysed using PHASTER [52] as to determine the presence of phages. The ResFinder 3.1 web server [49] (https://cge.cbs.dtu.dk/services/ResFinder/) was used to identify acquired antimicrobial resistance genes and/or chromosomal mutations, using a threshold of 80% identity. The predicted results from both platforms were then compared with antimicrobial susceptibility testing results. After genome annotation using Prokka v1.13 [198], metal tolerance was accessed by inspecting the presence of several genes from different metal export systems, such as the copper tolerance genes *pcoABCDRSE*, silver tolerance genes *silCFBAPRSE*, arsenite transmembrane pump genes *arsABCR*, mercury tolerance genes *merACDE*, and tellurite resistance gene *tehAB* [132].

4.3.3. Gene-by-gene analysis

Gene-by-gene analysis was performed by taking advantage of a publicly available panel of 8558 loci [156] derived from the EnteroBase *Salmonella* wgMLST schema [63], curated and prepared using chewBBACA [157], downloaded on August 2018 (http://doi.org/10.5281/zenodo.1323684). Allele calling was performed on all genomes using chewBBACA v2.0.11 [157] with default parameters and a publicly available training file for *S. enterica* (https://github.com/mickaelsilva/prodigal_training_files). Exact and inferred matches were used to construct an allelic profile matrix, where other allelic classifications (see https://github.com/B-UMMI/chewBBACA/wiki) were assumed as "missing" loci. Minimum spanning trees (MSTs) were constructed using the goeBURST algorithm [158] implemented in the PHYLOViZ online web-based tool [61], based on 100% shared loci between all isolates (i.e., shared-genome MLST) [159].

For comparative purposes, all *S. enterica* genomes from serovar Rissen identified in the EnteroBase database were downloaded (on November 2018) from the European Nucleotide Archive (ENA) and were assembled as described above using the INNUca pipeline. After post-assembly inspection and confirmation of serovar using SeqSero, a total

of 270 genomes from strains isolated worldwide, described in Supplementary Table 4-2, were used to construct an initial MST enrolling all genomes (i.e. 60 PT plus 270 retrieved from ENA) to integrate all these novel PT genomes within the known *S. enterica* serovar Rissen diversity. Additionally, in order to perform WGS-based epidemiological cluster analysis a second MST was constructed enrolling only the 60 novel PT genomes. To increase the resolution power for cluster analysis of the PT isolates for both initial MSTs, we took advantage of PHYLOViZ online 2.0 Beta version (http://online2.phyloviz.net/). This platform allows maximization of the shared genome in a dynamic manner, i.e., for each sub-set of isolates under comparison, the maximum number of shared loci (at 100%) between them is automatically used for sub-tree construction. All allelic distance thresholds used during cluster investigation were expressed as percentages of allele differences (AD) (i.e., the number of observed allelic differences divided by the total number of shared loci under comparison). Thus, to explore isolates sub-sets, a conservative step-by-step approach was performed by applying three allelic distance cut-offs of 1, 0.5 and 0.25% to both initial MSTs, based on previously described data for cluster investigation in gene-by-gene based surveillance [45].

4.4. Results

4.4.1. Antimicrobial susceptibility and heavy metal tolerance

All antimicrobial resistance phenotype and genotype data are presented in Supplementary Table 4-1. Although none of the 60 PT isolates are resistant to either meropenem, cefoxitin, cefotaxime, ceftazidime, ceftriaxone, cefepime or tigecycline, most are resistant to at least one of the remaining antimicrobials tested. Moreover, resistance to more than one antimicrobial was verified in 88.3% of the isolates and 83.3% are MDR. Only one isolate (PT11) is fully susceptible to the antimicrobials tested (1.7%). Sulfamethoxazole resistance is the most common (83.3%), followed by tetracycline (81.7%), trimethoprim (80.0%), ampicillin (73.3%), chloramphenicol (53.3%), and azithromycin (50.0%) resistance. Of note, two distinct food-associated isolates exhibit resistance to quinolones, with PT60 being resistant to both pefloxacin and nalidixic acid while PT44 only to nalidixic acid. Additionally, only one isolate (PT03) reveals intermediate susceptibility to gentamicin (1.7%). None of the isolates presents the genes that confer resistance to colistin (i.e., the *mcr* genes).

Metal resistance-associated genes for copper (*pcoABCDRSE*), arsenic (*arsABCR*) and tellurite (*tehAB*) were observed in all PT isolates analysed (Supplementary Table 4-1). Thirteen isolates (21.7%) presented the mercury resistance-associated genes *merACDE*, which was always collocated with the ampicillin and sulphonamide resistance genes *bla-*

TEM-1B and *sul1*, respectively. All these isolates also presented *cmlA1*, conferring resistance to chloramphenicol, and *dfrA1*, conferring resistance to trimethoprim. Fifty-three isolates (88.3%) also present the complete silver tolerance cassette *silCFBAPRSE*, which was located contiguously with the *pco* gene cluster.

4.4.2. Global genetic diversity of S. enterica serovar Rissen

All novel PT isolates were firstly integrated with all publicly available *S. enterica* serovar Rissen genomes (N=270), using a gene-by-gene approach, in order to assess their genomic diversity and phylogenetic relationships within the worldwide circulating population. *In silico* MLST analysis revealed that all enrolled isolates belonged to ST469. The initial MST (Figure 4-1 A), based on 2305 shared loci between all 330 isolates, reveals low genetic diversity between all isolates, with an overall mean pairwise AD of 35 ± 9, and that most PT isolates from the present study are closely related. While an initial conservative threshold of 1% (i.e., an AD of 24) still maintains all PT isolates phylogenetically linked, when applying a cut-off of 0.5% to the MST (due to the overall low genetic diversity observed), 10 out of the 60 isolates showed up as unlinked (with two pairs and six single isolates segregating independently) (Figure 4-1 A), potentially indicating that they are epidemiologically unrelated. In order to further analyse the cluster containing most PT isolates (at a 0.5% threshold), a sub-MST of this cluster was generated (Figure 4-1 B) which increased the number of shared loci to 3162 and an overall mean pairwise AD of 29 ± 10 was observed. This sub-set of 97 isolates comprises, not only most PT isolates, but also isolates from the United States of America, the United Kingdom, Spain, Denmark and Vietnam. Applying a cut-off of 0.5% to this sub-set, corresponding to an AD of 16, two main clusters containing PT isolates remain and one isolate segregates independently (PT10). Nevertheless, when a more restrict cut-off is applied (0.25%; 8 AD), more consistent with outbreak clustering investigation [45], all the PT isolates separate from strains of other countries (with the exception of an isolate from the United Kingdom, ENA accession # SAMN09298461) and two main clusters containing most of the PT isolates are observed, suggesting two main circulating clones.

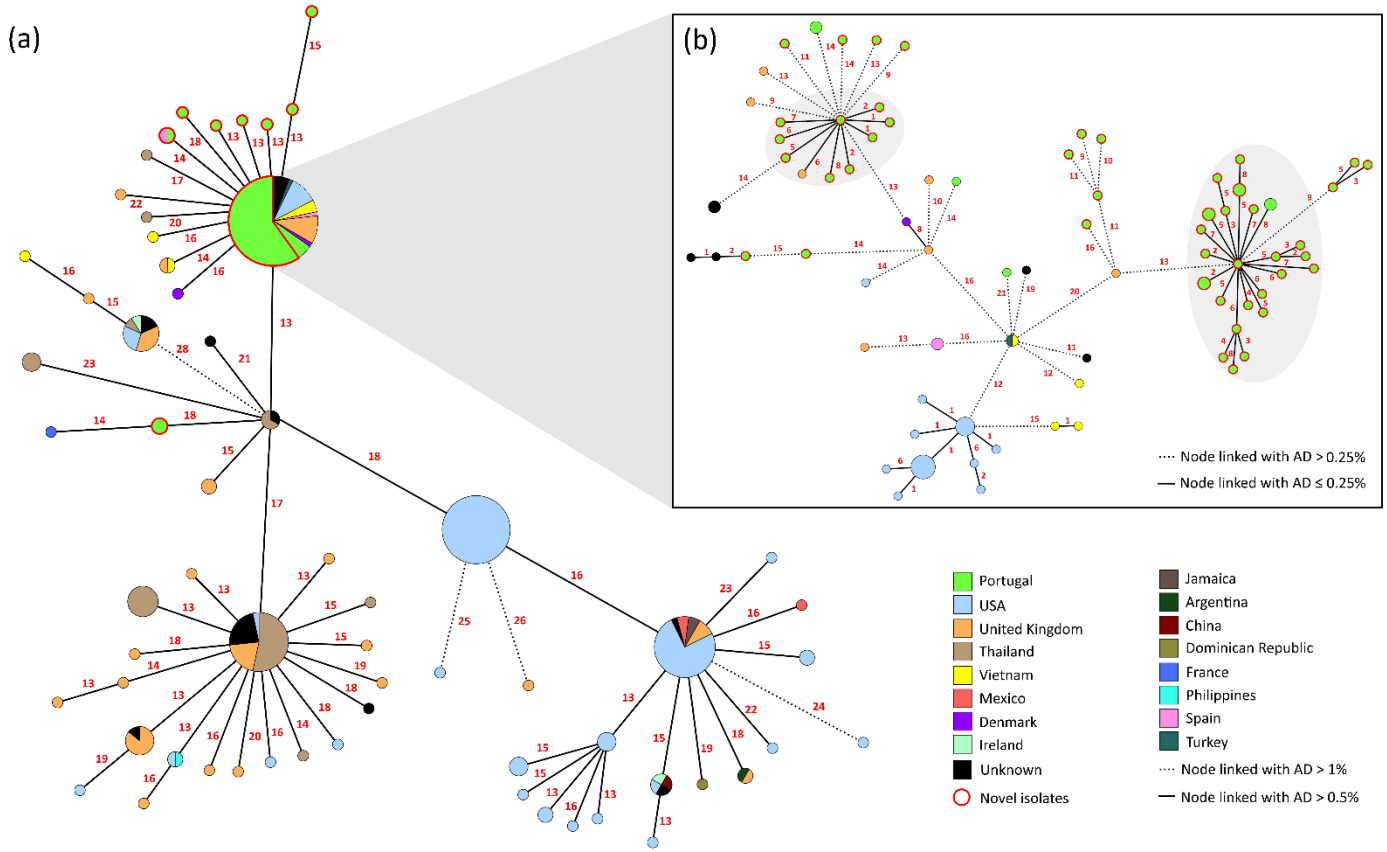


Figure 4-1 – Phylogenetic analysis of *S. enterica* serovar Rissen, based on a gene-by-gene approach using a wgMLST schema with 8558 loci. (a) – Minimum spanning tree (MST) enrolling 270 publicly available genomes and the 60 novel "Portuguese" (PT), constructed based on the allelic diversity of 2305 shared loci. The numbers in red on the connecting lines represent the AD between isolates. Nodes linked with allelic distances (AD) equal to or below 0.5% (i.e., 12 AD) have been collapsed for visualization purposes. Node sizes are proportional to the number of isolates they represent. Nodes are coloured according to the country of origin. (b) – Sub-MST constructed based on the maximum number of shared loci (3162 loci) between the sub-set of isolates linked at an allelic distance of 0.5% and containing most PT isolates. Two major clusters containing mostly PT isolates linked with AD ≤ 0.25% are highlighted in grey.

4.4.3. WGS-based epidemiological analysis of the PT isolates

We then proceeded with the same gene-by-gene approach, strictly for the 60 PT *S. enterica* serovar Rissen isolates, to assess their potential epidemiological relatedness (Figure 4-2). The initial MST reveals that the isolates share 3465 loci, with a mean pairwise AD of 35 ± 17 (ranging from 0 up to 47). As such, we observed that, while the number of shared loci between the PT isolates was increased by more than 1100 loci, in comparison with the initial analysis, the overall genetic diversity is still low. As a means to exclude potential epidemiologically unrelated cases of *S. enterica* serovar Rissen within this set of isolates, an initial conservative threshold of 1% (i.e., 35 AD) reveals that at least 3 isolates (PT60, PT40 and PT17) segregate independently. As two of these isolates are linked (cluster E), we performed a sub-MST analysis that sustains their close genetic relatedness as the increase in the number of shared loci to 3801 only increments their AD by 3. Nevertheless, an epidemiological link between these isolates could not be traced.

In order to generate potential clusters to be subjected to the dynamic MST analysis, we then applied a lower threshold of 0.5% (i.e., 17 AD) which reveals four additional clusters and even more potential isolated cases (PT01, PT10, PT11, PT12, PT16, PT18, PT19, PT50, and PT56). One of these clusters (Cluster D) contains two food isolates (Figure 4-2 C) retrieved from chicken meat in 2016 (Figure 4-2 B), one from the Lisbon Metropolitan area (PT45) and the other from Spain (PT44) (Figure 4-2 A). After increasing the number of shared loci under analysis to 3828, the sub-MST shows that these isolates are distinguishable by only 7 AD. This suggests the existence of either a *S. enterica* serovar Rissen strain already circulating within the Iberian Peninsula, though more isolates are required for confirmation, or a discrete phenomenon as only two cases were detected. Nevertheless, these isolates present the resistance genes *sul*1 (Supplementary Table 4-1), *dfrA*12 (Supplementary Table 4-1), and *aadA*2 (Supplementary Table 4-1), as reported by a previous study on this area [183]. We also observed a more heterogeneous cluster (Cluster C) where the epidemiological linkage between the four isolates is unclear, due to the differences in geographical location (Figure 4-2 A), isolation date (Figure 4-2 B), source (Figure 4-2 C) and antibiotic resistance profile (Figure 4-2 D). Still, the sub-MST for this cluster (which is based on 3805 loci) suggests that these isolates are genetically related, as they present a mean pairwise AD of 14 ± 3. Regarding Cluster B, sub-MST analysis now enrolling 3686 shared loci shows that isolates are still linked at the 0.5% threshold, with a mean pairwise AD of 14 ± 6. Although this cluster is comprised by isolates from animal, food and clinical samples (Figure 4-2 C), it is hard to suggest a direct transmission link from these sources to human, with the clinical cases all detected prior to 2016, contrarily to all but one

non-human sample (PT24) (Figure 4-2 B). However, all isolates from this cluster are MDR (Figure 4-2 D). In addition, 12 out of the 13 isolates from this cluster possess the mercury tolerance genes *merACDE*, in association with the chloramphenicol resistance gene, *cmlA1*, and trimethoprim resistance gene *dfrA1* which further distinguishes this cluster from all others where these genes are absent. The only other isolate possessing these genes in the entire dataset is PT50, which is very closely related to this cluster at an AD of 18, suggesting its genetic close relatedness but lacking epidemiological relationship. Moreover, the absence of the silver tolerance-associated genes (*silCFBAPRSE*) was only observed in isolates from this cluster (7 out of the 13, including the five isolates from pork skewers). Of note, the five 2016 isolates from pork skewers with an undisclosed origin (PT29, PT30, PT31, PT32, PT33) are very likely meat products from the same pig holding facility, as within 3831 shared loci they only exhibit up to 6 AD between them and share the same resistance profile. These isolates share the same year of isolation and resistance profile, to both antibiotics and heavy metals, with a pork isolate from the Azores archipelago (PT35) with a maximum AD of 8, all indications of the existence of a possible cluster in Azores.

Finally, the largest cluster (Cluster A) is mostly comprised by isolates from the Azores archipelago (n =21) but also includes two isolates from Lisbon Metropolitan area, two from Centre region and one from North region (Figure 4-2 A). Analysis of this cluster reveals that the isolates share a maximum of 3639 loci with a mean pairwise AD of 12 ± 4, with all isolates still linked after sub-MST construction. This cluster presents distinct sources (Figure 4-2 C), with the majority of isolates (14 out of 29) originating from pigs (Supplementary Table 4-1) or being human clinical cases (11 out of 29), suggesting that this epidemiological clone may originate from pig holding facilities. Of note, we observed that a clinical isolate (PT20) and a food isolate (PT13), collected two months apart in Azores, are only distinguishable by 1 AD in 3866 loci, strongly indicating an epidemiological link between them. Moreover, with the exception of PT48 and PT49, all isolates from this cluster are MDR, presenting 4 to 7 resistance determinants (Figure 4-2 D). The two non-MDR isolates are likely epidemiologically linked (1 AD over 3866 shared loci between the two) and present the same resistance profile, being resistant to ampicillin, amoxicillin-clavulanic acid and chloramphenicol. Most isolates are resistant to azithromycin, with the exception of PT02, PT48, and PT49. In addition, four sets of isolates present the same resistance profiles between them (Supplementary Table 4-1): i) PT26, PT28, and PT46; ii) PT37, PT51, and PT55; iii) PT25, PT38, PT39 and PT52; iv) PT06, PT13, PT14 PT15, PT20, PT34 PT41, PT43, PT57 and PT58. Phage presence analysis for all 60 isolates also reveals that all isolates from Cluster A (except PT42) and Cluster C possess an Entero P88 Phage-like structure, which further distinguishes them from isolates of Cluster B. In fact, only PT09 from Cluster B

contains a similar phage but its nucleotide sequence is different from the ones detected in Cluster A and C. In summary, the described data points to the simultaneous existence of at least two multidrug resistant epidemic *S. enterica* serovar Rissen clones circulating in the Azores archipelago at least since 2014, which were already introduced throughout the Portuguese continent.

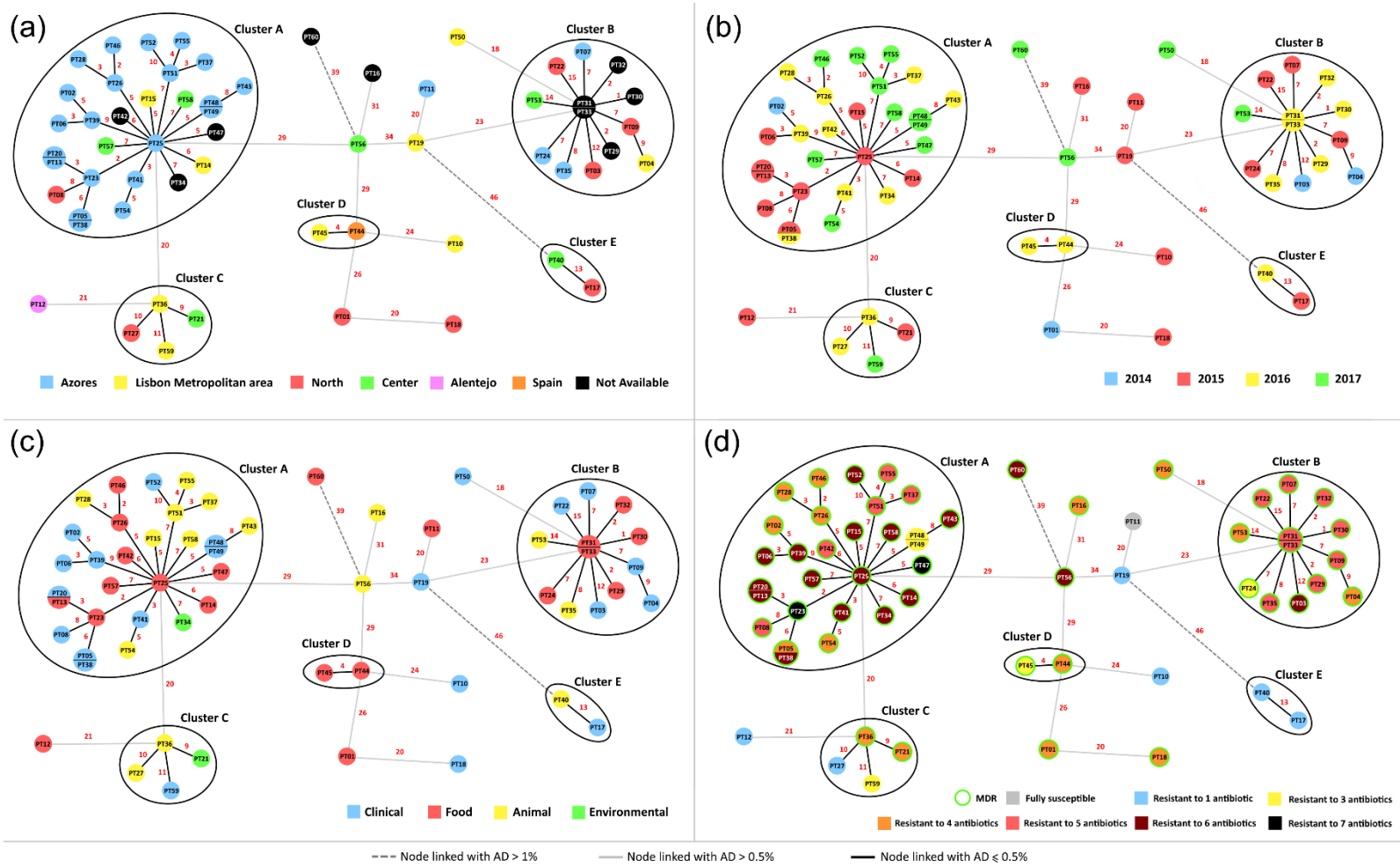


Figure 4-2 – Allelic diversity analysis of the "Portuguese" (PT) S. enterica serovar Rissen, using a wgMLST-based gene-by-gene approach. Minimum spanning tree enrols a total of 3465 shared loci. Nodes are coloured by (a) region of isolation, (b) year of isolation, (c) sample type and (d) antibiotic resistance profile. The numbers in red on the connecting lines represent the allelic distance between isolates. MDR – Multidrug resistant.

4.5. Discussion

WGS is quickly supplanting traditional procedures for *Salmonella* surveillance and outbreak detection in Reference Laboratories. In this regard, food- and water-borne outbreaks are detected either when a common source is determined through epidemiological inquiries, followed by the characterization of all the isolates, or when a group of similar isolates is identified, followed by the common source by epidemiological investigation [191]. The current study aimed for the identification of *S. enterica* serovar Rissen genetic clusters circulating in Portugal, and the detection of potential sources of infection, as a follow-up of an unusual increment in the number of isolates that arrived at the NRLs since 2015.

Even though *S. enterica* serovar Rissen is rarely reported worldwide as a cause of human salmonellosis, it has previously been identified in Portugal associated with pig, pork, beef, chicken and wild animals [177,186–188,199], which was also observed in this work. Using a dynamic shared-genome based approach, by progressively maximizing the number of shared loci between isolates, we detected five potential clusters of closely related clinical, animal, food, and environmental *S. enterica* serovar Rissen ST469 isolates [45], with the two largest clusters containing all the isolates from the Azores archipelago (Cluster A and Cluster B) (Figure 4-2 A). This approach revealed a high degree of similarity among the *S. enterica* serovar Rissen population, contrary to what was previously described through PFGE [178]. In fact, among the 330 studied isolates, we found a mean genetic distance of about 35 AD (with a maximum AD of 81) within the shared 2305 loci. Apart from a few isolates that segregate independently, a great number of the PT isolates formed very closely related clusters. Increasing the resolution of the initial shared wgMLST approach by increasing the number of loci analysed, reinforced the relatedness of the Portuguese clusters, most specifically the clusters containing MDR isolates from the Azores archipelago (Cluster A and Cluster B). Even though this genomic approach seems to be highly discriminatory, there is no universal cut-off defined for identification of outbreaks, therefore epidemiological investigation should be carried out to facilitate the interpretation of WGS data. Given the high degree of genetic similarity within this serovar revealed in this study, several isolates that seem very closely related may in fact be epidemiologically unlinked. Nonetheless, the genomic analysis together with the scarce epidemiological information points to the existence of two non-related MDR *S. enterica* serovar Rissen clones circulating in the Azores archipelago for the past years. Additionally, the identification of clinical isolates as well as isolates from animals and food in the Portuguese continent that show a perfect clustering with the isolates from Azores strongly suggests the spread of the circulating clones

throughout the Portuguese territory, with a putative origin in Azores. The fact that the Azores archipelago is composed by nine small islands with livestock as one of the major economic resources reinforces this possibility. Another detected cluster containing a PT isolate and a Spanish isolate seems to concern an already described successful clone circulating in the Iberian Peninsula, as a result of intensive trade of live pigs and pork between Portugal and Spain [183].

Increased antimicrobial resistance in pig-associated *S. enterica* serovars has become a reality for the past decades, including the successful clone *S. enterica* serovar Rissen ST469 [131,183,200]. MDR bacteria emerge as a direct consequence of selective pressure derived from overall antibiotic misuse. The use of antibiotics in food-producing animals has been associated with the emergence of certain MDR clones [104]. Additionally, the acquisition of novel properties, such as antibiotic resistance and metal tolerance may occur by horizontal gene transfer between different bacteria and even between bacterial species [201]. In the present study, 88.3% of the isolates were resistant to more than one antimicrobial and 83.3% were MDR (Supplementary Table 4-1 and Figure 4-2 D). A high level of resistance to several antibiotics was observed, although resistance to carbapenems, cephalosporins and colistin was not detected. Moreover, 50% of the isolates, mainly isolates from Cluster A, were resistant to azithromycin, which is widely used for the treatment of invasive *Salmonella* infections. According to the genomic analysis of these isolates, azithromycin resistance is likely mediated by the macrolide inactivation gene *mph*A, while *bla*TEM-1B_1 seems to be responsible for ampicillin resistance. Also, *tet*(A) appears in all the tetracycline resistant isolates of this serovar, confirming that *tet*(A) is most likely the gene responsible for tetracycline resistance in *S. enterica* serovar Rissen [183]. We also accessed the presence of metal tolerance genes in these isolates as heavy metals are widely used in biocidal products, feed additives and soil fertilizers, leading to bioaccumulation processes [131]. The acquisition of copper and silver resistance genes by *Salmonella* seems to be related to the emergence and widespread of certain MDR clones, as a means of survival in metal contaminated environments [131]. In this study, 53 PT isolates presented the *pco*ABCDRSE and *sil*CFBAPRSE cassettes in the same genetic structure, suggestive of tolerance to those heavy metals. This fact reinforces the association of *pco* and *sil* cassettes with successful MDR clones in pig production, most specifically the ST469 *S. enterica* serovar Rissen clone, as described previously [109,131].

4.6. Conclusions

In summary, we identified at least two MDR *S. enterica* serovar Rissen clones in the Azores archipelago, already circulating in the continent. The presence of MDR isolates with zoonotic potential in food-producing animals is a growing public health concern, having not only a severe burden to human health, but also great economic impact. Patients infected by MDR bacteria have an increased risk of developing severe infections with high mortality and morbidity rates, and represent an increased healthcare cost [103]. International trade of food-producing animals and their products contributes greatly to the global spread of MDR *Salmonella* clones, which calls for continuous monitoring, especially in pig-production. Although WGS has great potential in supporting epidemiological investigations, the availability of epidemiological data is critical for timely and efficient source detection and outbreak control. This WGS-based *S. enterica* serovar Rissen surveillance study in Portugal results from the collaboration between the Portuguese *Salmonella* NRLs of human and animal health. Hopefully, this stands as the starting point for the implementation of the "One Health" approach for *Salmonella* surveillance in Portugal.

Acknowledgments: The authors would like to acknowledge all the laboratories that sent the isolates to the NRLs for typing. This work was partially funded by the GenomePT project [POCI-01-0145-FEDER-022184], supported by COMPETE 2020 – Operational Programme for Competitiveness and Internationalisation (POCI), Lisboa Portugal Regional Operational Programme (Lisboa2020), Algarve Portugal Regional Operational Programme (CRESC Algarve2020), under the PORTUGAL 2020 Partnership Agreement, through the European Regional Development Fund (ERDF), and by Fundação para a Ciência e a Tecnologia (FCT).

CHAPTER 5

Characterization of multidrug-resistant isolates of *Salmonella enterica* serovars Heidelberg and Minnesota from imported poultry meat in Portugal

Manuscript in revision in Microbial Drug Resistance

Silveira L, Nunes A, Pista Â, Isidro J, Correia C B, Saraiva M, Batista R, Castanheira I, Machado J, Gomes J P. Characterization of multidrug-resistant isolates of *Salmonella enterica* serovars Heidelberg and Minnesota from imported poultry meat in Portugal.

LS contributed to the design of the study, performed most of the experimental work and part of the bioinformatics analyses, interpreted data and wrote the manuscript.

5. Characterization of multidrug-resistant isolates of *Salmonella enterica* serovars Heidelberg and Minnesota from imported poultry meat in Portugal

5.1. Abstract

Multidrug resistant bacteria are a major public health problem, mainly linked to antimicrobials misuse in human and veterinary medicine, as well as in animal production. *Salmonella enterica* serovar Heidelberg and *Salmonella enterica* serovar Minnesota are a frequent cause of foodborne infection in North and South America, but are rarely reported in Europe. Their frequent multidrug resistance (MDR) character make them potential spreaders of antibiotic resistance genes. In this work, we determined the antibiotic resistance profile and performed whole-genome sequencing (WGS) on a set of *S. enterica* serovar Heidelberg and *S. enterica* serovar Minnesota isolates recovered from samples of fresh poultry meat, collected in 2012-2019 within the Portuguese Official Inspection Plan for imported foodstuffs, and integrated these results with data already reported in the literature. Several isolates of both serovars showed extremely high genetic relatedness either with isolates from raw poultry meat preparations imported from South America to the Netherlands or to isolates from samples from the broiler production chain in Brazil, an important exporter country. This similarity was extended to their general MDR profile, including the presence of the extended-spectrum cephalosporins-resistant CMY-2. The MDR character was also common to the vast majority (94.4%) of isolates from both serovars, where, several of them also carried the plasmid IncX1 containing a Type IV Secretion System, a well-known virulence mechanism. These results somehow mirror the scenario observed in the Netherlands, showing the introduction, through fresh imported poultry meat in compliance with European legislation, of MDR *S. enterica* serovar Heidelberg and serovar Minnesota isolates in Europe with the potential spread of their diverse resistance markers. The present study suggests the hygiene criteria for foodstuffs monitoring are in need of revision, with surveillance of the resistome, before foodstuffs are placed on the market, being an important contribute to avoid further dissemination of resistance markers.

5.2. Introduction

One of the main problems that 21st century medicine is facing concerns the rise of multidrug-resistant (MDR) bacteria (bacteria resistant to more than three different classes of antimicrobials)[104,202]. The use of antibiotics in humans, animals and plants has led to the widespread of mobile genetic elements carrying multiple antimicrobial resistance traits [112]. As such, the food production industry and its globalization play an important role in the transmission, evolution and dissemination of clinically relevant MDR pathogens worldwide[203]. *Salmonella* remains the second most common cause of gastrointestinal disease in the EU[162], and MDR *Salmonella* have become a major public health concern over the years, causing invasive infections that frequently require hospitalization and present high mortality rates[204].

In the food chain, poultry, especially chicken and turkey, are regarded as significant contributors for the salmonellosis burden, since their intestinal tracts are very often colonized with *Salmonella*[205]. Additionally, the transmission of *Salmonella* in flocks may occur from different routes, either from the infected breeding flocks (vertical transmission), or through previously infected flocks, contaminated feed or water (horizontal transmission). Consequently, the elimination of *Salmonella* from a facility may prove a difficult task [206]. Several countries, including United States of America (USA) and Brazil, and the European Union (EU) have implemented, in the past decade, *Salmonella* control programmes in poultry flocks, aimed to reduce the presence of *Salmonella* spp., especially *Salmonella enterica* serovar Enteritidis and *Salmonella enterica* serovar Typhimurium, that cause most part of human infections[205,207]. Although resulting in a significant decrease of *S. enterica* serovar Enteritidis infection, these measures yielded a shift in the prevalence of *Salmonella* serovars in poultry and also in human salmonellosis, with an increase in the dissemination of successful clones of less frequent serovars, carrying several advantageous adaptation features accompanied with MDR mobile elements [205]. However, between 2013 and 2017, the proportion of *S. enterica* serovar Enteritidis human cases increased, mostly due to importation of eggs from one of the Member States[208].

S. enterica serovar Heidelberg and *S. enterica* serovar Minnesota are rarely reported in European countries, not only in human clinical cases but also in animals and foodstuffs[87,162]. However, *S. enterica* serovar Heidelberg is a frequent cause of gastroenteritis in North and South America, and is frequently isolated in poultry meat[209–215]. *S. enterica* serovar Minnesota has also been frequently isolated in poultry production in Brazil[207,216]. In this study, the Portuguese National Institute of Health (INSA) analysed samples of fresh poultry meat, collected within the Portuguese Official Inspection Plan for

imported foodstuffs coordinated by the General Directorate of Food and Veterinary (DGAV). In order to access the potential need for a more precise survey of these serovars, we determined the antibiotic resistance profile and performed whole-genome sequencing (WGS) on a set of these isolates, and put these data in frame with data reported in the literature regarding the isolation of these *Salmonella* serovars.

5.3. Materials and methods

5.3.1. Sample dataset characterization

S. enterica isolates were obtained from samples of imported fresh poultry meat, collected between 2012 and 2019. In total 163 samples, corresponding to 815 units of fresh poultry meat (mostly chicken gizzards) were collected by the General Directorate of Food and Veterinary, as part of official border control, and analysed in INSA. Forty-eight samples, corresponding to 58 units, tested positive for *Salmonella* spp.. *Salmonella* isolates were then serotyped by the slide agglutination method, according to the White-Kauffman-Le Minor scheme[15] and antimicrobial susceptibility testing was performed by disc diffusion, following the European Committee on Antimicrobial Susceptibility Testing (EUCAST)[193] recommendations, on a panel of 18 antimicrobials: ampicillin, amoxicillin-clavulanic acid, azithromycin, cefepime, cefotaxime, cefoxitin, ceftazidime, ceftriaxone, chloramphenicol, erythromycin, gentamicin, meropenem, nalidixic acid, pefloxacin, sulfamethoxazole, tetracycline, tigecycline, and trimethoprim. Results were interpreted using current EUCAST breakpoints and Epidemiological cut-off values[193,195–197,217–221].

5.3.2. Whole Genome Sequencing and genome characterization

Total DNA was extracted from fresh cultures on the NucliSens easyMAG platform (bioMérieux, France), according to the manufacturer's instructions. Thirty-six isolates were selected for WGS, in order to have at least one representative from each batch of imported fresh poultry meat. WGS and bacterial *de novo* assembly were performed as previously described[152]. Briefly, for WGS, high-quality DNA samples quantified using Qubit (ThermoFisher, USA) were subjected to dual-indexed Nextera XT Illumina library preparation (Illumina, USA), prior to cluster generation and paired-end sequencing (2×250bp or 2x150bp) on either a MiSeq or a NextSeq 550 instrument (Illumina, USA) available at INSA, according to the manufacturer's instructions.

5.3.3. Genome assembly and annotation

All genomes were *de novo* assembled using the INNUca v3.1 pipeline (https://github.com/B-UMMI/INNUca), an integrative bioinformatics pipeline that consists of

several integrated modules for reads QA/QC, *de novo* assembly and post-assembly optimization steps. Briefly, after reads' quality analysis using FastQC v0.11.5 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and cleaning with Trimmomatic v0.36 [153], genomes were assembled with SPAdes 3.10 [154] and subsequently improved using Pilon v1.18 [155]. Draft genome sequences were annotated with RAST server (http://rast.theseed.org/FIG/rast.cgi) [222–224].

5.3.4. Strains' genomic characterization

For each strain, *in silico* Multi Locus Sequence Type (MLST) prediction was performed using the *mlst* v2.4 software (https://github.com/tseemann/mlst), while serovar was predicted with SeqSero software [48], using both raw and trimmed reads as well as the assembled genomes. The ResFinder 3.1 web server [49] (https://cge.cbs.dtu.dk/services/ResFinder/) was used to identify acquired antimicrobial resistance genes and/or chromosomal mutations, using a threshold of 80% identity. The predicted results were then compared with antimicrobial susceptibility testing results. PlasmidFinder 2.0 and plasmid MLST [50] were used to detect and characterize the plasmids present in these isolates (https://cge.cbs.dtu.dk/services/PlasmidFinder/; https://cge.cbs.dtu.dk/services/pMLST/), with a threshold for %ID ≥80% and a minimum % coverage of 60%. As a means to potentiate isolate discrimination, assemblies were also analysed using PHASTER [52] to determine the presence of phages. For simplification purposes, only hits with intact phages were considered for further analysis. Finally, SPIFinder 1.0 help to identify *Salmonella* pathogenicity islands (SPIs) in the sequenced isolates (https://cge.cbs.dtu.dk/services/SPIFinder/).

5.3.5. Gene-by-gene analysis

For comparative purposes regarding genomic relatedness, along with the 36 isolates imported to Portugal, 88 isolates identified in Brazil (82 serovar Heidelberg and 6 serovar Minnesota) [203] and 133 isolates of serovar Heidelberg identified in the Netherlands, that had also been found in poultry meat imported from Brazil [225], were included in the analysis (Supplementary Table 5-1). The genomes of this set of 221 isolates were also assembled as described above. Gene-by-gene analysis was performed by taking advantage of a publicly available panel of 8558 loci [156] derived from the EnteroBase *Salmonella* wgMLST schema [63], curated and prepared using the chewBBACA free software (https://github.com/B-UMMI/chewBBACA) [157], downloaded on August 2018 (http://doi.org/10.5281/zenodo.1323684). For all genomes, allele calling was performed using chewBBACA v2.0.11 [157] with default parameters and a publicly available training file for *S.*

enterica (https://github.com/mickaelsilva/prodigal_training_files). Exact and inferred matches were used to construct an allelic profile matrix, where other allelic classifications (see https://github.com/B-UMMI/chewBBACA/wiki) were assumed as "missing" loci.

To evaluate the genetic relationship between strains, minimum spanning trees (MSTs) were constructed taking advantage of goeBURST algorithm [158] implemented in the PHYLOViZ online web-based tool [61], based on 100% shared loci between all strains (i.e., shared-genome MLST) [159]. In order to increase the resolution power for cluster analysis, the PHYLOViZ online 2.0 Beta version (http://online2.phyloviz.net/) was used, as it allows maximizing the shared genome in a dynamic manner, i.e., for each sub-set of strains under comparison, the maximum number of shared loci between them is automatically used for sub-tree construction.

5.4. Results

5.4.1. Characterization of the isolates

In total, we identified 37 *S. enterica* serovar Heidelberg isolates, 19 *S. enterica* serovar Minnesota, and 1 *S. enterica* 4,[5],12:i:-. The *in silico* serotyping of raw reads of the 36 selected isolates (23 serovar Heidelberg and 13 serovar Minnesota) confirmed the classical serotyping results for the sequenced isolates (Table 5-1).

All the isolates tested for antimicrobial resistance were susceptible to meropenem, chloramphenicol, and cefepime, however 94.4% were multidrug resistant (MDR) (Table 5-1). Sulfamethoxazole and tetracycline resistance were the most common (94.0%), followed by nalidixic acid and ampicillin (83.3%), and pefloxacin resistance (80.5%). Resistance to 2nd generation cephalosporin cefoxitin (77.7%), to 3rd generation cephalosporins cefotaxime (80.5%), ceftazidime (77.7%), and ceftriaxone (72.2%), were also frequently observed. Of note, resistance to tigecycline, trimethoprim, gentamicin and azithromycin were also found in a few isolates. For the most part, these phenotypic results were confirmed by the *in silico* predictions of ResFinder. Sulphonamide resistance was mainly caused by the presence of *sul2*, although four *S. enterica* serovar Heidelberg isolates additionally presented *sul1* (PT15, PT23, PT30, PT34). Tetracyclines resistance was mediated by *tet(A)* (Table 5-1), similarly to what we previously found in *S. enterica* serovar Rissen (Chapter 4). Also, fluoroquinolone resistance was caused by the presence of mutations in *gyrA* (S83F) and *parC* (T57S) in *S. enterica* serovar Heidelberg, while in *S. enterica* serovar Minnesota it was mediated by the presence of *qnrB* genes (Table 5-1). Generally, β-lactams resistance was caused by the presence of the β-lactamase CMY-2 (77.7%), but other β-lactamase (*bla*) genes, including $bla_{TEM\text{-}1B}$, $bla_{TEM\text{-}116}$, $bla_{CTX\text{-}M\text{-}8}$, and

bla$_{CTX-M-55}$ were also detected (PT1, PT8, PT15, PT31, PT33) (Table 5-1). The gene *mphB*, encoding the macrolide 2'-phosphotransferase II was detected in an erythromycin resistant *S. enterica* serovar Heidelberg isolate (PT15), while no known mutations associated with macrolides resistance or any common genes other than *macA* and *macB* were found in the azithromycin resistant *S. enterica* serovar Heidelberg PT29 isolate (Table 5-1). Gentamycin resistance was mediated either by aminoglycoside acetyltransferases (AAC(3)-IV and AAC(3)-Via) or by an aminoglycoside nucleotidyltransferase (ANT(2")-Ia). The gene *dfrA25* was also detected in the only trimethoprim resistant isolate (PT30). Additionally, 17 *S. enterica* serovar Heidelberg isolates presented the gene *fosA7*, conferring resistance to fosfomycin. Finally, none of the isolates sequenced presented the *mcr* genes that confer resistance to colistin.

Regarding the presence of plasmids, the 36 sequenced isolates harboured at least one, with 16 different plasmids being detected in total (Table 5-1). Overall, we detected five different Col plasmids (Col440I, Col8282, ColpVC, Col156, and ColRNAI), and 11 plasmids of seven different incompatibility groups (IncA/C2, IncFIB(pHCM2), IncFIB(AP001918), IncFII, IncFII(pHN7A8), IncFII(29), IncX1, IncI1, IncHI2, IncHI2A, and IncQ1). IncA/C2 (ST2) was present in 35 isolates and frequently carried *tet(A)*, *sul2*, *bla*$_{CMY-2}$, and the mercury resistance (*mer*) operon. In five *S. enterica* serovar Heidelberg isolates, the *bla*$_{CMY-2}$ gene seemed to be chromosomally inserted. Additionally, IncX1 was present in 18 *S. enterica* serovar Heidelberg isolates and harboured the Type IV secretion system (T4SS). The *mer* operon was absent only in non-MDR isolate PT13. The *qnr* genes in *S. enterica* serovar Minnesota were all located in Col440I.

PHASTER analysis identified 15 different phages (Supplementary Table 5-2) among both serovars, six in *S. enterica* serovar Heidelberg isolates (Gifsy-2, Sfi19, UAB Phi20, SPN1S, phiV10, SPN1S, and Gifsy-1) and 11 in *S. enterica* serovar Minnesota (phiV10, Gifsy-1, Fels-2, Fels-1, SEN34, SfV, vB SemP Emek, SfII, ENT47670, P88, and SSU5). All *S. enterica* serovar Heidelberg isolates harboured lambdoid prophage Gifsy-2, which was absent in all *S. enterica* serovar Minnesota isolates. SPI analysis revealed that all isolates harbour SPI13, SPI14, and Centisome 63 pathogenicity island (C63PI) (Table 5-1). Also present in several isolates were SPI1 (only in *S. enterica* serovar Heidelberg), SPI2, SPI3, SPI4, and SPI5.

.

Table 5-1 – Serovar, antibiotic resistance phenotype and genotype, plasmids and pathogenicity islands of the *Salmonella enterica* serovar Heidelberg and *Salmonella enterica* serovar Minnesota isolates enrolled in the present study. ST – sequence type. Amp - ampicillin, Azm - azithromycin, Ctx - cefotaxime, Fox - cefoxitin, Caz - ceftazidime, Cro - ceftriaxone, Ery - erythromycin, Gmn - gentamicin, Nal - nalidixic acid, Pef - pefloxacin, Smx - sulfamethoxazole, Tet - tetracycline, Tgc - tigecycline, and Tmp – trimethoprim.

Isolate ID	Year	Serovar	Batch	ST	Phenotype	Genotype	Plasmids	Pathogenicity islands
PT1	2012	*S.*Minnesota	1	548	Amp, Fox, Caz, Ctx, Tet, Smx	*blaTEM-116*, *blaCMY-2*, *blaCMY-6*, *tet(A)*, *sul2*, *qnrB*, *parC* p.T57S	IncA/C2, Col440I	SPI2, SPI3, SPI13, SPI14, C63PI
PT2	2012	*S.*Minnesota	2	548	Amp, Ctx, Tet, Smx	*blaCMY-2*, *tet(A)*, *sul2*, *parC* p.T57S	IncA/C2, ColpVC	SPI2, SPI3, SPI4, SPI13, SPI14, C63PI
PT3	2013	*S.*Minnesota	3	548	Amp, Fox, Caz, Ctx, Cro, Tet, Smx, Pef, Nal	*blaCMY-2*, *tet(A)*, *sul2*, *qnrB*, *parC* p.T57S	IncA/C2, Col440I, ColRNAI	SPI2, SPI13, SPI14, C63PI
PT4	2013	*S.*Minnesota	3	ND	Amp, Fox, Caz, Ctx, Cro, Tet, Smx	*blaCMY-2*, *tet(A)*, *sul2*, *parC* p.T57S	IncA/C2, IncFIB(AP001918), IncFII, IncFII(29), ColRNAI	SPI13, SPI14, C63PI
PT5	2013	*S.*Minnesota	4	ND	Amp, Fox, Caz, Ctx, Cro, Tet, Smx	*blaCMY-2*, *tet(A)*, *sul2*, *parC* p.T57S	IncA/C2	SPI2, SPI4, SPI13, SPI14, C63PI
PT6	2013	*S.*Heidelberg	5	15	Amp, Fox, Caz, Ctx, Cro, Tet, Smx, Pef, Nal	*blaCMY-2*, *tet(A)*, *sul2*, *parC* p.T57S, *gyrA* p.S83F, *fosA7*	IncA/C2, IncI1, IncX1, ColpVC	SPI2, SPI3, SPI4, SPI13, SPI14, C63PI
PT7	2013	*S.*Minnesota	6	548	Amp, Fox, Caz, Ctx, Cro, Tet, Smx	*blaCMY-2*, *tet(A)*, *sul2*, *qnrB*, *parC* p.T57S	IncA/C2, Col440I	SPI2, SPI3, SPI4, SPI13, SPI14, C63PI
PT8	2013	*S.*Heidelberg	7	ND	Amp, Fox, Caz, Ctx, Cro, Tet, Smx, Pef, Nal	*blaCMY-61*, *tet(A)*, *sul2*, *parC* p.T57S, *gyrA* p.S83F	IncA/C2, IncI1, IncX1, ColpVC	SPI1, SPI3, SPI5, SPI13, SPI14, C63PI
PT9	2013	*S.*Heidelberg	8	15	Amp, Fox, Caz, Ctx, Tet, Smx, Pef, Nal	*blaCMY-2*, *tet(A)*, *sul2*, *parC* p.T57S, *gyrA* p.S83F	IncA/C2, ColpVC	SPI1, SPI3, SPI5, SPI13, SPI14, C63PI
PT10	2013	*S.*Heidelberg	9	15	Tet, Smx, Pef, Nal	*tet(A)*, *sul2*, *parC* p.T57S, *gyrA* p.S83F, *fosA7*	IncA/C2, IncX1, ColpVC	SPI2, SPI3, SPI4, SPI13, SPI14, C63PI
PT11	2014	*S.*Heidelberg	10	15	Amp, Fox, Caz, Ctx, Cro, Tet, Smx, Pef, Nal	*blaCMY-2*, *tet(A)*, *sul2*, *parC* p.T57S, *gyrA* p.S83F, *fosA7*	IncA/C2, IncX1, ColpVC	SPI1, SPI3, SPI5, SPI13, SPI14, C63PI
PT12	2014	*S.*Minnesota	11	548	Amp, Fox, Caz, Ctx, Cro, Tet, Smx	*blaCMY-2*, *tet(A)*, *sul2*, *parC* p.T57S	IncA/C2, Col156	SPI2, SPI13, SPI14, C63PI
PT13	2014	*S.*Heidelberg	12	15	Pef, Nal	*parC* p.T57S, *gyrA* p.S83F, *fosA7*	IncX1, ColpVC	SPI2, SPI3, SPI4, SPI13, SPI14, C63PI
PT14	2014	*S.*Heidelberg	13	15	Pef, Nal	*blaCMY-2*, *tet(A)*, *sul2*, *parC* p.T57S, *gyrA* p.S83F	IncA/C2, IncI1, ColpVC, ColRNAI	SPI5, SPI13, SPI14, C63PI
PT15	2015	*S.*Heidelberg	14	15	Amp, Tet, Smx, Gmn, Pef, Nal, Ery	*blaTEM-1B*, *tet(A)*, *sul1*, *sul2*, *mphB*, *aac(3)-VIa*, *parC* p.T57S, *gyrA* p.S83F	IncA/C2, IncHI2, IncI1, IncQ1, ColRNAI	SPI13, SPI14, C63PI
PT16	2015	*S.*Heidelberg	15	15	Amp, Fox, Caz, Ctx, Cro, Tet, Smx, Nal	*blaCMY-2*, *tet(A)*, *sul2*, *parC* p.T57S, *gyrA* p.S83F, *fosA7*	IncA/C2, IncX1, ColpVC	SPI3, SPI4, SPI5, SPI13, SPI14, C63PI
PT17	2017	*S.*Heidelberg	16	15	Amp, Fox, Caz, Ctx, Cro, Tet, Smx, Pef, Nal	*blaCMY-2*, *tet(A)*, *sul2*, *parC* p.T57S, *gyrA* p.S83F, *fosA7*	IncA/C2, IncX1, ColpVC, ColRNAI	SPI3, SPI5, SPI13, SPI14, C63PI
PT18	2017	*S.*Heidelberg	17	15	Tet, Smx, Pef, Nal	*tet(A)*, *sul2*, *parC* p.T57S, *gyrA* p.S83F	IncA/C2, IncI1, ColpVC, ColRNAI	SPI3, SPI5, SPI13, SPI14, C63PI
PT19	2017	*S.*Heidelberg	18	15	Amp, Fox, Caz, Ctx, Cro, Tet, Smx, Pef, Nal	*blaCMY-2*, *tet(A)*, *sul2*, *parC* p.T57S, *gyrA* p.S83F, *fosA7*	IncA/C2, IncX1, ColpVC	SPI1, SPI2, SPI3, SPI5, SPI13, SPI14, C63PI

Isolate ID	Year	Serovar	Batch	ST	Phenotype	Genotype	Plasmids	Pathogenicity islands
PT20	2017	*S.*Heidelberg	19	15	Amp, Fox, Caz, Ctx, Cro, Tet, Smx, Pef, Nal	*blaCMY-2*, *tet(A)*, *sul2*, *parC* p.T57S, *fosA7*	IncA/C2, IncX1, ColpVC	SPI1, SPI2, SPI3, SPI4 SPI5, SPI13, SPI14, C63PI
PT21	2018	*S.*Minnesota	20	ND	Amp, Fox, Caz, Ctx, Cro, Tet, Smx, Pef, Nal	*blaCMY-2*, *tet(A)*, *sul2*, *qnrB*, *parC* p.T57S	IncA/C2, Col440I	SPI2, SPI3, SPI4, SPI13, SPI14, C63PI
PT22	2018	*S.*Heidelberg	21	15	Amp, Fox, Caz, Ctx, Cro, Tet, Smx, Pef, Nal	*blaCMY-2*, *tet(A)*, *sul2*, *parC* p.T57S, *gyrA* p.S83F, *fosA7*	IncA/C2, IncI1, IncX1	SPI1, SPI2, SPI3, SPI5, SPI13, SPI14, C63PI
PT23	2018	*S.*Heidelberg	22	15	Amp, Fox, Caz, Ctx, Cro, Tet, Smx, Gmn, Pef, Nal	*blaCMY-2*, *tet(A)*, *sul1*, *sul2*, *aac(3)-VIa*, *parC* p.T57S, *gyrA* p.S83F, *fosA7*	IncA/C2, IncHI2A, IncHI2, IncX1, ColRNAI	SPI2, SPI3, SPI5, SPI13, SPI14, C63PI
PT24	2018	*S.*Minnesota	23	548	Amp, Fox, Caz, Ctx, Cro, Tet, Smx, Pef, Nal	*blaCMY-2*, *tet(A)*, *sul2*, *qnrB*, *parC* p.T57S	IncA/C2, Col440I	SPI2, SPI3, SPI13, SPI14, C63PI
PT25	2018	*S.*Heidelberg	23	15	Amp, Fox, Caz, Ctx, Cro, Tet, Smx, Pef, Nal	*blaCMY-2*, *tet(A)*, *sul2*, *parC* p.T57S, *gyrA* p.S83F, *fosA7*	IncA/C2, ColpVC	SPI5, SPI13, SPI14, C63PI
PT26	2018	*S.*Heidelberg	24	15	Amp, Fox, Caz, Ctx, Cro, Tet, Smx, Pef, Nal	*blaCMY-2*, *tet(A)*, *sul2*, *parC* p.T57S, *gyrA* p.S83F, *fosA7*	IncA/C2, IncQ1, IncX1	SPI2, SPI3, SPI5, SPI13, SPI14, C63PI
PT27	2018	*S.*Heidelberg	24	15	Amp, Fox, Caz, Ctx, Cro, Tet, Smx, Pef, Nal	*blaCMY-2*, *tet(A)*, *sul2*, *parC* p.T57S, *gyrA* p.S83F, *fosA7*	IncA/C2, IncQ1, IncX1	SPI2, SPI3, SPI5, SPI13, SPI14, C63PI
PT28	2018	*S.*Minnesota	25	548	Amp, Fox, Caz, Ctx, Cro, Tet, Smx, Pef, Nal	*blaCMY-2*, *tet(A)*, *sul2*, *qnrB*, *parC* p.T57S	IncA/C2, Col440I	SPI2, SPI3, SPI5, SPI13, SPI14, C63PI
PT29	2019	*S.*Heidelberg	26	15	Amp, Fox, Caz, Ctx, Cro, Tet, Tgc, Smx, Azm, Pef, Nal	*blaCMY-2*, *tet(A)*, *sul2*, *parC* p.T57S, *gyrA* p.S83F, *macA*, *macB*	IncA/C2, IncX1, ColpVC	SPI2, SPI5, SPI13, SPI14, C63PI
PT30	2019	*S.*Heidelberg	27	15	Amp, Fox, Caz, Ctx, Cro, Tet, Tmp, Smx, Gmn, Pef, Nal	*blaCMY-2*, *tet(A)*, *dfrA25*, *sul1*, *sul2*, *aac(3)-IV*, *parC* p.T57S, *gyrA* p.S83F, *fosA7*	IncA/C2, IncX1, ColRNAI	SPI2, SPI3, SPI5, SPI13, SPI14, C63PI
PT31	2019	*S.*Minnesota	28	548	Amp, Fox, Caz, Ctx, Cro, Tet, Smx, Pef, Nal	*blaCTX-M-8*, *tet(A)*, *sul2*, *qnrB*, *parC* p.T57S	IncA/C2, Col440I	SPI2, SPI3, SPI4, SPI13, SPI14, C63PI
PT32	2019	*S.*Minnesota	29	548	Amp, Fox, Caz, Ctx, Cro, Tet, Tgc, Smx, Pef, Nal	*blaCMY-2*, *tet(A)*, *sul2*, *qnrB*, *parC* p.T57S	IncA/C2, IncFIB(pHCM2), Col440I, Col8282	SPI2, SPI3, SPI13, SPI14, C63PI
PT33	2019	*S.*Minnesota	29	548	Amp, Fox, Caz, Ctx, Cro, Tet, Tgc, Smx, Pef, Nal	*blaTEM-1B*, *blaCTX-M-8*, *blaCTX-M-55*, *blaCMY-2*, *tet(A)*, *sul2*, *qnrB*, *parC* p.T57S	IncA/C2, IncFII(pHN7A8), Col440I	SPI2, SPI3, SPI13, SPI14, C63PI
PT34	2019	*S.*Heidelberg	29	15	Tet, Tgc, Smx, Gmn, Pef, Nal	*tet(A)*, *sul1*, *sul2*, *ant(2")-Ia*, *parC* p.T57S, gyrA p.S83F, *fosA7*	IncA/C2, IncX1, Col440I, ColpVC	SPI5, SPI13, SPI14, C63PI
PT35	2019	*S.*Heidelberg	30	15	Tet, Tgc, Smx, Gmn, Pef, Nal	*tet(A)*, *sul2*, *parC* p.T57S, gyrA p.S83F, *fosA7*	IncA/C2, IncX1, ColpVC	SPI5, SPI13, SPI14, C63PI
PT36	2019	*S.*Heidelberg	30	15	Amp, Fox, Caz, Ctx, Cro, Tet, Smx, Pef, Nal	*blaCMY-2*, *tet(A)*, *sul2*, *parC* p.T57S, *gyrA* p.S83F, *fosA7*	IncA/C2, IncI1, IncX1, ColpVC	SPI1, SPI2, SPI3, SPI5, SPI13, SPI14, C63PI

5.4.2. WGS-based genomic relatedness analysis

Although in four isolates the *in silico* MLST analysis was not possible to determine because one of the alleles was not found, all *S. enterica* serovar Heidelberg isolates obtained from fresh poultry meat imported from third countries to Portugal, belong to ST15 and all *S. enterica* serovar Minnesota belong to ST548. The initial MST (Figure 5-1A), based on 2843 shared loci revealed two genetically distant clusters (2705 AD), each corresponding to a different serovar. Most part of the isolates obtained in the samples imported to Portugal clustered together with different sets of isolates from Brazil and/or Netherlands. Regarding only serovar Minnesota, the isolates share a total of 3458 loci with an overall mean pairwise distance of 85.2±80.6 allelic differences (AD) (Figure 5-2). Taking the isolates studied at INSA into consideration, they only present a mean pairwise distance of 26.3±7.1 AD. However, two *S. enterica* serovar Minnesota Brazilian isolates seem to be closer to the isolates of our study, presenting a mean pairwise distance of 25.7±6.6 AD, and sharing the same antibiotic resistance phenotype and genotype of some of the isolates. SRR7130551, is resistant to cephalosporins, penicillins, tetracyclines, and fluoroquinolones[203], presenting the same resistance markers of isolates PT24 and PT33, including plasmids IncA/C2 and Col440I (Table 5-1). However, this isolate contains an additional pathogenicity island, SPI4, absent in both PT isolates. SRR7130561, presents the same resistance profile and plasmids of PT31, but not SPI4 that is present in this strain.

.

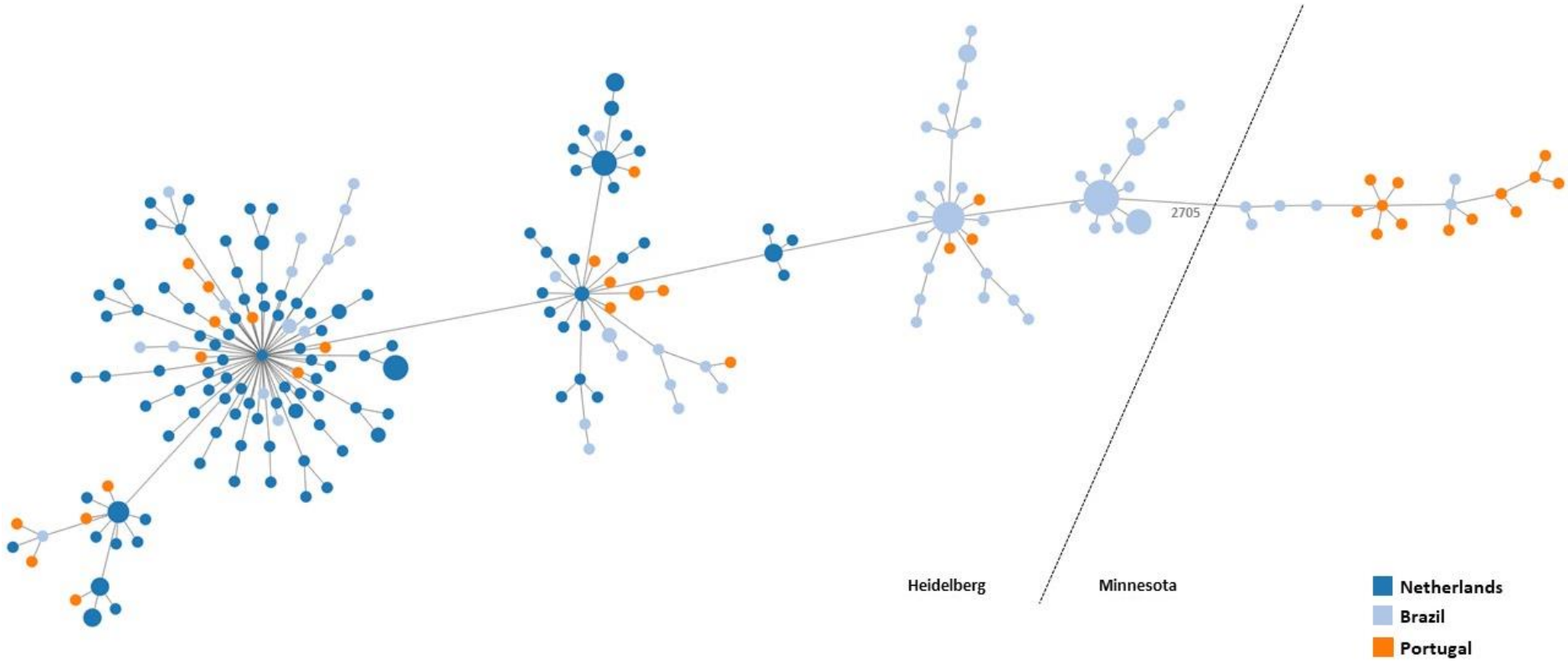


Figure 5-1 – Global phylogeny of *Salmonella enterica* serovars Heidelberg and Minnesota isolates, based on a dynamic gene-by-gene approach using a wgMLST schema with 8558 loci. The minimum spanning tree enrolling 133 Brazil-imported isolates identified in Netherlands (van der Berg RR et al, 2019), 88 isolates identified in Brazil (Monte DF et al, 2019) and 36 imported isolates identified in Portugal, was constructed based on the allelic diversity of 2843 shared loci, using the goeBURST algorithm implemented in the PHYLOViZ online platform. Filled circles (nodes) represent unique allelic profiles, and are colored according to isolates' isolation country. For simplification purposes, only the allele difference between both serovars is shown. The dashed black line segregates S. enterica serovars Heidelberg and Minnesota isolates.

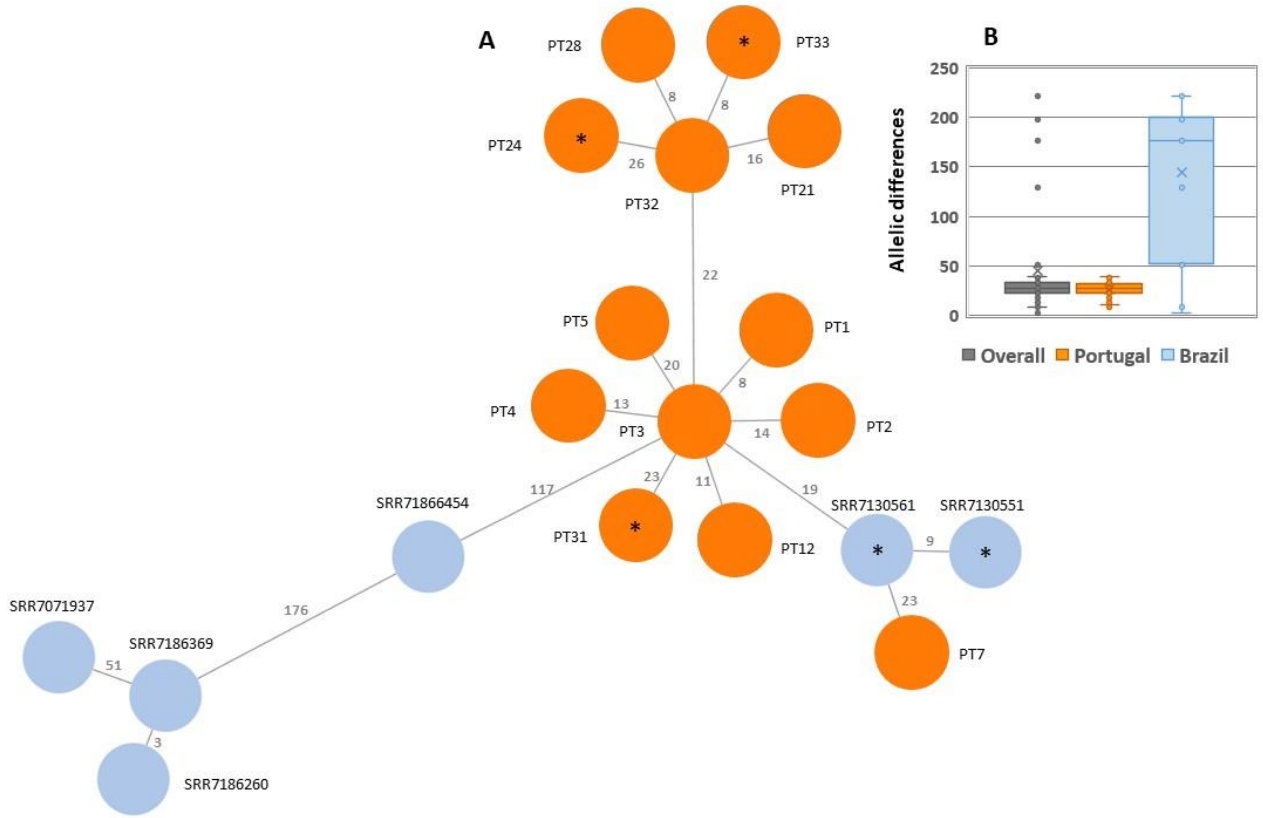


Figure 5-2 – Phylogenic relationships based on a dynamic gene-by-gene approach and allelic differences between *Salmonella enterica* serovar Minnesota isolates. **A** The minimum spanning tree was constructed using the goeBURST algorithm implemented in the PHYLOViZ online platform, and is based on the allelic diversity found among the 3458 genes shared by all *S. enterica* serovar Minnesota isolates. Unique allelic profiles are shown as filled circles coloured according to isolates' isolation country. The numbers in grey on the connecting lines represent the allele differences between isolates. Marked with an asterisk are the two isolates from Brazil, SRR7130551 and SRR7130561, which display an AD to some PT isolates lower than the mean AD among all PT isolates (26.3). **B** Plot depicting the allelic differences between the overall population and between isolates of each country.

Nevertheless, we also found cases for which an apparent discrete allelic distance is not concordant with the differential presence of other genetic features. For instance, although *S. enterica* serovar Minnesota PT3 and PT4 were isolated from units of the same batch and reveal 13 AD, PT3 presents SPI2, absent in PT4, and Col440I harbouring *qnrB*, while PT4 presents several IncF plasmids and phages Fels 1 and Fels 2. Similarly, isolates PT32, PT33, and PT34, from another batch, also presented several differences. Firstly, this batch contained isolates of serovar Minnesota (PT32 and PT33) and serovar Heidelberg (PT34). Secondly, PT33, that presents a mean pairwise distance of 8 AD with PT32, besides harbouring *blaCMY-2* like PT32, harbours additional *bla* genes, *blaTEM1*, *blaCTX-M-8*, and *blaCTX-M55* (Table 5-1). These isolates also have different phage content (Supplementary Table 5-2).

Analysis of *S. enterica* serovar Heidelberg isolates revealed they share 3041 loci, with a mean pairwise distance of 18.2±9.5 AD between all isolates (Figure 5-3). This serovar revealed a low level of genetic variability, as previously described[226,227]. Interestingly, isolate PT10, Brazilian isolate SRR7130373[203], and isolates from cluster VI[225] from the Netherlands are closely related, sharing a mean pairwise distance of 2.7±2.1. PT11 presents a mean pairwise distance of 5 AD from Batch BI[225], original from Santa Catarina state in Brazil. The same resistance genotype and plasmids were detected in these isolates. Although PT13 only shares 5 AD with isolates from batch AI from Netherlands[225], we identified a high level of genetic diversity among these isolates, mainly regarding their antimicrobial resistance profiles.

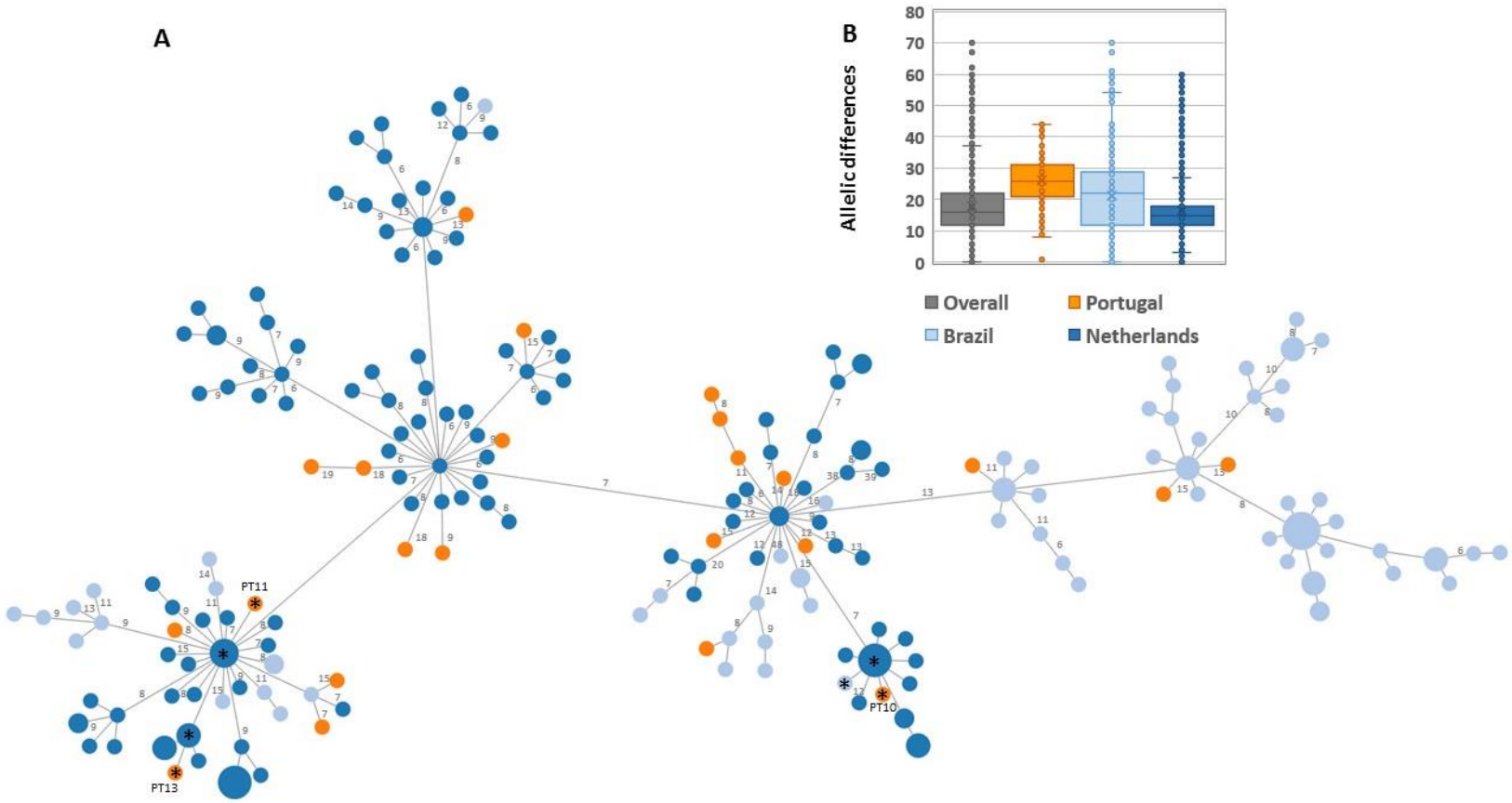


Figure 5-3 – Phylogenic relationships based on a dynamic gene-by-gene approach and allelic differences of *Salmonella enterica* serovar Heidelberg isolates. **A** The minimum spanning tree enrolls a total of 3041 shared loci and was constructed using the goeBURST algorithm implemented in the PHYLOViZ online platform. Unique allelic profiles are shown as filled circles colored according to isolates' isolation country. The numbers in grey on the connecting lines represent the allele differences (AD) between isolates. For simplification purposes, AD <5 are not shown. Marked with an asterisk are isolates from Brazil, and Netherlands, which are closest to the, also marked with an asterisk, isolates PT10, PT11, and PT13 imported to Portugal. **B** Plot depicting the allelic differences between the overall population and between isolates of each country.

5.5. Discussion

One of the most known and appreciated typical dishes in Portugal are stewed chicken gizzards, to an extent that national production is not sufficient to cover the market's requirements. As such, part of this fresh poultry meat consumed in Portugal is imported from third countries. Portugal, being one of the entry points to the EU market, has Border Inspection Posts in its airports and ports, where samples of imported foods of animal origin are inspected by the competent authority DGAV. Regarding *Salmonella* in fresh poultry meat, the specific requirements were revised in the light of the serovars of *Salmonella* causing the higher percentage of human salmonellosis, namely *S. enterica* serovar Enteritidis, *S. enterica* serovar Typhimurium, and *S. enterica* serovar 4,[5],12:i:-. The Commission Regulation (EC) No 2073/2005[228], and subsequent amendments[229], lay down the food hygiene criterion for *Salmonella* in poultry carcasses of broilers and turkeys, after chilling in slaughterhouses, where 50 units must be tested (n=50) and 7 units may test positive for *Salmonella* (c=7). Additionally, testing against the food safety criterion must be carried out in five units (n=5) and all five units must test negative for *Salmonella* (c=0). For compliance with the regulation, for *S. enterica* serovar Typhimurium, *S. enterica* serovar 1,4,[5],12:i:-, and *S. enterica* serovar Enteritidis the result must be "*Not detected*", in 25 g of fresh poultry meat. This criterion for products placed on the market during their shelf-life aims to control the two most common serovars with public health significance in the EU, without detrimental economic consequences due to the application of the legislation. However, this means that if any other serovar of *Salmonella* is detected, the foodstuff is proper for human consumption and in compliance with the food safety microbiological criteria.

In this study, we isolated, between 2012 and 2019, from 163 samples (each one comprising 5 units) of fresh poultry meat imported from third countries, 37 *S. enterica* serovar Heidelberg isolates, 19 *S. enterica* serovar Minnesota, and 1 *S. enterica* 4,[5],12:i:-. Moreover, WGS of a set of these isolates involving all batches revealed a high genetic diversity within serovar Heidelberg and serovar Minnesota, suggesting multiple sources. In fact, when comparing the isolates detected in fresh poultry meat imported to Portugal with isolates characterized in Brazil[203], and in the Netherlands, we found several genetically related isolates also sharing phenotypic and genotypic features that suggest the existence of common geographical contaminated sources, translated in several entry points in the EU and in the dissemination of diverse MDR clones.

Although *S. enterica* serovar Heidelberg and *S. enterica* serovar Minnesota are infrequently reported in Europe, the dissemination of these two serovars in the American

continent has been frequently reported[203,213–216,230–232]. While the detection of these serovars in fresh poultry meat placed on the market, is in compliance with the Commission Regulation (EC) Nº 2073/2005, the presence of antibiotic resistance markers may constitute a health threat due to the frequent and well-known horizontal contamination. This is especially a concern when those markers confer resistance to antimicrobials used to treat severe infections, such is the case of cephalosporins and fluoroquinolones. Considering that *Salmonella* can spread to humans through contaminated foods, to prevent or reduce this risk, whenever *Salmonella* is identified but the requirements established in the European legislation for *Salmonella* in fresh poultry meat are fulfilled, "(…) *the batches of those products placed on the market must be clearly labelled by the manufacturer in order to inform the consumer of the need of thorough cooking prior to consumption*". Additionally, the principles contained in the World Health Organization (WHO) Manual "Five Keys to Safer Food" must be followed by consumers and retailers[233]. *S. enterica* serovar Heidelberg and *S. enterica* serovar Minnesota have been found to present extended-spectrum cephalosporin-resistance associated with AmpC β-lactamases CMY[207,209,211,212,216,226,230,231,234]. In our study, antimicrobial resistance testing of several isolates revealed a high percentage of multidrug resistance (94%) which is most likely due to misuse of antimicrobials in poultry production[216]. A high frequency of isolates presenting resistance to sulfamethoxazole, tetracycline, β-lactams and fluoroquinolones was detected. Fluoroquinolones, sulphonamides and tetracyclines are usually administered in poultry industry as therapeutic agents for bacterial infections and growth promoters, which might explain the high frequency of isolates with resistance to these groups of antibiotics[216,235]. WGS of the isolates identified in INSA further revealed the generalized presence of bla_{CMY-2}, but also, less frequently, the presence of other widely disseminated *bla* genes. In this work, bla_{CMY-2} was associated with IncA/C2 other than Inc1, as previously reported[209,230]. In a few cases, bla_{CMY-2} was apparently inserted in the chromosome. Typically, IncA/C2 also carried *tet(A)*, *sul2*, and the mercury tolerance operon. This plasmid is widely distributed across the American continent and is known to be responsible for the dissemination of several important resistance markers[230]. Additionally, IncX1 carrying T4SS was identified in 18 *S. enterica* serovar Heidelberg isolates. This secretion system is known to play not only an important part in pathogenesis, namely in the ability to invade and persist in macrophages and intestinal epithelial cells, but also in horizontal gene transfer, since it can transfer nucleoprotein complexes besides effector proteins[227,236,237]. Remarkably, these plasmids also carried toxin-antitoxin systems previously described to confer plasmid stability[227]. Other Inc plasmids were also detected as well as Col plasmids. We also detected the gene

fosA7, previously identified in *S. enterica* serovar Heidelberg isolated from broiler chickens in Canada, in 17 isolates of this serovar[238].

Although the Portuguese National Reference Laboratory for Gastrointestinal Infections has not identified to date any *S. enterica* serovar Heidelberg nor *S. enterica* serovar Minnesota with these resistance profiles, associated with human infection, the presence of these MDR isolates in fresh poultry meat for human consumption is still a risk, especially if in the environment of food production, distribution, and also in private houses, good hygiene practices concerning foods and surfaces are not followed. In the United States and Canada, *S. enterica* serovar Heidelberg is more frequently associated with invasive human infections, such as myocarditis and septicaemia, than other Non-typhoidal *Salmonella*, and is the second serovar mostly associated with deaths in the US[226,227]. *S. enterica* serovar Heidelberg and *S. enterica* serovar Minnesota seem to be highly successful and persistent clones, surely due to their enhanced capacity to intake so many resistance markers and virulence traits. Moreover, the isolates identified in this study are resistant to antimicrobials typically used as treatment of severe *Salmonella* infections. The further spread of these traits to other bacteria, and even to commensal microbiota (prior to cooking the contaminated foodstuffs), represents an added risk that should be taken into account.

In summary, we detected two distinct serovars of MDR *Salmonella* in fresh poultry meat imported from third countries and entering the EU through Portugal. *S. enterica* serovar Heidelberg seems to be a highly clonal population, while *S. enterica* serovar Minnesota isolates present higher genomic variability. Nonetheless, both serovars present several fitness traits that enhance their capacity to spread and persist alike other successful *Salmonella* epidemic clones. In an era where the global spread of MDR bacteria is one of the major public health concerns, the criteria for foodstuffs monitoring, especially when concerning imported goods, are quite possibly in need of reviewing. The surveillance of the antibiotic resistance of *Salmonella* serovars isolated from fresh poultry meat, would be an important contribute for evaluating and understanding the dissemination of antibiotic resistance from animals for food production, to food, and then humans, before foodstuffs of this food category are placed on the market.

CHAPTER 6

Final overview, concluding remarks and future directions

6. Final overview, concluding remarks and future perspectives

The main objective of this PhD dissertation was to contribute to the determination of the genomic structure of *Salmonella enterica* serovars identified in Portugal, mainly through whole-genome sequencing. The importance of this objective stems from the fact that salmonellosis remains the second most common cause of gastrointestinal disease in the European Union, and as such has a high economic impact and a high burden of disease worldwide. The use of whole-genome sequencing in *Salmonella* surveillance has gained steady ground during the course of this PhD, and so, the need for the Portuguese National Reference Laboratory for *Salmonella* to implement this methodology was inevitable. An important collaborative work, INNUENDO, allowed the implementation of well-established pipelines for bacterial genome assembly used throughout the analysis of the isolates in this dissertation.

The role of the NRL implies, not only the implementation of new molecular approaches for *Salmonella* surveillance, but also the analysis of the data obtained and its release both to the scientific community and health authorities. Decades of work in *Salmonella* surveillance resulted in the collection of data that was in dire need of a retrospective analysis to illustrate the dynamics of the most relevant serovars in Portugal. In this scope, in chapter 2, we analysed the existing data for the past 21 years, from classical typing to antimicrobial susceptibility testing, maintaining the focus on the data and not on the methodology, which suffered a few changes over time. We found that, the trend of the total number of detected *Salmonella enterica* has, over the years, been accompanied by the trend of *Salmonella enterica* serovar Enteritidis, a serovar that used to be responsible for most part of salmonellosis cases in Portugal. This serovar, mainly found in poultry and poultry products, suffered a significant decrease when control measures at the primary production level were introduced. The decrease of *Salmonella enterica* serovar Enteritidis, allowed for other serovars, less frequent at that time, to occupy new niches and infect new hosts. Consequently, a new serovar, also isolated in poultry, *Salmonella enterica* serovar 1,4,[5],12:i:-, emerged and became the most frequent serovar in recent years in parallel with the serovar Enteritidis.

A better understanding of *Salmonella enterica* in general, and certain serovars in specific is crucial so that the control measures introduced result in the reduction of the number of cases and not just an inversion of a trend. In chapter 3, we investigated the genomic diversity of the most frequent serovar in Portugal (in the period of study), *Salmonella enterica* serovar Enteritidis. Also, due to a suspected outbreak of a rare serovar, *Salmonella enterica* serovar Newport was also included in this study as it relies on an

ongoing investigation with the participation of the NRL. A retrospective MLVA typing analysis of *Salmonella enterica* serovar Enteritidis revealed that almost half of the isolates belonged to three endemic MLVA profiles while the other half was mainly composed by several sporadic MLVA types. While MLVA allowed the confirmation of a suspected outbreak signalled by the Portuguese Health Authorities in 2016, it hardly gave any information without any epidemiological context. Whole-genome sequencing, however, has higher resolution than MLVA, and allowed, not only the confirmation of the 2016 outbreak, but also allowed the detection of two undetected outbreaks. WGS discriminatory power is unquestionable, as 63 nodes were detected among the 65 isolates, although these were clustered in only two distinct MLVA profiles. Whole-genome sequencing also allowed us to confirm that the sudden increase in 2019 of the number of cases of salmonellosis caused by *Salmonella enterica* serovar Newport was quite possibly an outbreak as this serovar was rarely isolated in the previous years. Although other isolates will be sequenced in the following weeks in order to have a more complete scenario, our preliminary data confirms the existence of an outbreak due to the high genetic relatedness of most of the isolates sequenced so far. On the epidemiological point of view, the recent data collection signifies that the investigation of this outbreak by the competent authorities is still ongoing. In summary, whole-genome sequencing proved to be a highly discriminatory tool for outbreak detection, which will surely lead to signalling of even more clusters of interest allowing proper investigations to take part.

Other serovars, non-related with poultry, such as *Salmonella enterica* serovar Rissen also emerged recently, mainly in Azores. Chapter 4 is focused on the study of the genetic diversity of this serovar in Portugal precisely to investigate the putative existence of outbreaks and try to decipher transmission links. Another hypothesis underlying the increase of this serovar is that these strains might carry genetic determinants conferring fitness advantages, which could also be investigated through the genetic analysis. Whole-genome sequencing revealed a high degree of similarity among the *Salmonella enterica* serovar Rissen population, enrolling 60 Portuguese isolates and 270 isolates from different countries. We establish several genetic links among isolates from different sources, including human, animal, and food, in the Portuguese continent that show a perfect clustering with the isolates from Azores. Although the apparent relatedness between several isolates, given the high degree of genetic similarity within this serovar, an epidemiological link could be inexistent. Nonetheless, we identified two multidrug-resistant clones in the Azores archipelago, already spread to the continent. The fact that the majority of the isolates are multidrug-resistant, including antimicrobials of great clinical importance such as macrolides and cephalosporins, is worrisome. An additional concern was the

identification of metal tolerance genes, believed to have a significant role in emergence and widespread of certain multidrug-resistant clones, as a means of survival in metal contaminated environments. Similarly, chapter 5 consisted in the study of *Salmonella enterica* serovar Heidelberg and *Salmonella enterica* serovar Minnesota isolates identified in Portugal. Given the origin of the isolates (imported poultry meat from third countries) and the recent studies on the genetic diversity of isolates from poultry meat in Brazil and imported to the Netherlands, we aimed for the phenotypic and genotypic integration of the isolates recovered in Portugal. Our major findings suggest that, the isolates imported to Portugal present high genetic relatedness with isolates imported to the Netherlands and from isolated identified in Brazil. Interestingly, while *Salmonella enterica* serovar Heidelberg population seems to be highly clonal, *Salmonella enterica* serovar Minnesota isolates present great genetic variability. Nonetheless, both serovars, like *Salmonella enterica* serovar Rissen, present fitness traits that enhance their capacity to spread and persist in the environment. Although these serovars are not frequent human pathogens, their antibiotic resistance markers can be potentially transferred to typical pathogens that share the same biological niche (i.e., food and environment), thus increasing the probability of severe human infections. In highlight, there is an urgent need for the revision of the regulations for foodstuffs monitoring and for the implementation of safe control measures in animal production, in order to avoid the spread of successful clones carrying genetic markers that can either cause severe infections, or further spread those markers to other relevant gastrointestinal bacteria.

In conclusion, this PhD dissertation constitutes an important contribution to gain insight on the characteristics of relevant *Salmonella enterica* isolates circulating in Portugal, as well as a contribution for the reinforcement of the capacitation of the Portuguese National Reference Laboratory. Not only was it possible to frame *Salmonella* surveillance in the past decades, but also to implement whole-genome sequencing as a tool for the detection of several outbreaks, allowing a more precise and real-time communication with the Health Authorities towards a more effective control of salmonellosis.

Future perspectives

As future perspectives we will carry on the surveillance of *Salmonella enterica* through whole-genome sequencing, continuing the study of the serovars included in this dissertation, as well as other relevant serovars. In particular, we will soon focus on *Salmonella enterica* serovar 1,4,[5],12:i:-, since we have a large collection of isolates, including the first isolates to be identified in the world, back in 1987. We already started sequencing these isolates and will soon start studying the genomic evolution of this

particular serovar, through several decades, as well as resistance markers (this serovar is usually multidrug resistant), and other fitness traits of interest.

7. References

1. Littman, R. J. The Plague of Athens: Epidemiology and Paleopathology. *Mt. Sinai J. Med. A J. Transl. Pers. Med.* **76**, 456–467 (2009).

2. Cunha, B. A. The death of Alexander the Great: malaria or typhoid fever? *Infect. Dis. Clin. North Am.* **18**, 53–63 (2004).

3. Papagrigorakis, M. J., Yapijakis, C., Synodinos, P. N. & Baziotopoulou-Valavani, E. DNA examination of ancient dental pulp incriminates typhoid fever as a probable cause of the Plague of Athens. *Int. J. Infect. Dis.* **10**, 206–214 (2006).

4. Vågene, Å. J. *et al.* Salmonella enterica genomes from victims of a major sixteenth-century epidemic in Mexico. *Nat. Ecol. Evol.* **2**, 520–528 (2018).

5. Zúquete, A. *Nobreza de Portugal* | *Os Reis* | *Volume 8*. (Alêtheia Editores, 2018).

6. Richard Adler & Mara, E. *Typhoid Fever: A History*. (McFarland, 2016).

7. Budd, W. Typhoid Fever: its nature, mode of spreading, and prevention. in *Typhoid Fever* 610–612 (The American Journal of Public Health, 1918).

8. Eberth, K. J. Die Organismen in den Organen bei Typhus abdominalis. *Arch. für Pathol. Anat. und Physiol. und für Klin. Med.* **81**, 58–74 (1880).

9. Lamas, A. *et al.* A comprehensive review of non-enterica subspecies of Salmonella enterica. *Microbiol. Res.* **206**, 60–73 (2018).

10. Chlebicz, A. & Śliżewska, K. Campylobacteriosis, Salmonellosis, Yersiniosis, and Listeriosis as Zoonotic Foodborne Diseases: A Review. *Int. J. Environ. Res. Public Health* **15**, 863 (2018).

11. Soper, G. A. The curious career of Typhoid Mary. *Bull N Y Acad Med* **15**, 698–712 (1939).

12. Marineli, F., Tsoucalas, G., Karamanou, M. & Androutsos, G. Mary Mallon (1869-1938) and the history of typhoid fever. *Ann. Gastroenterol.* **26**, 132–134 (2013).

13. Schultz, M. Theobald Smith. *Emerg. Infect. Dis.* **14**, 1940–1942 (2008).

14. Salmonella Subcommittee. The Genus Salmonella Lignières, 1900. *J. Hyg. (Lond).* **34**, 333–50 (1934).

15. Grimont, P. A. D. & Weill, F.-X. *Antigenic formulae of the Salmonella serovars*. (WHO

Collaborating Centre for Reference and Research on Salmonella, 2007).

16. Le Minor, L., Rohde, R. & Taylor, J. Nomenclature des Salmonella. *Ann. l ' Inst. Pasteur* **119**, 206–210 (1970).

17. Borman, E. K., Stuart, C. A. & Wheeler, K. M. Taxonomy of the Family Enterobacteriaceae. *J. Bacteriol.* **48**, 351–367 (1944).

18. Kauffmann, F. & Edwards, P. R. Classification and nomenclature of Enterobacteriaceae. *Int. J. Syst. Evol. Microbiol.* **2**, 2–8 (1954).

19. Crosa, J. H., Brenner, D. J., Ewing, W. H. & Falkow, S. Molecular Relationships Among the Salmonelleae. **115**, 307–315 (1973).

20. Brenner, F. W., Villar, R. G., Angulo, F. J., Tauxe, R. & Swaminathan, B. Salmonella Nomenclature. *J. Clin. Microbiol.* **38**, 2465–2467 (2000).

21. Whole Genome Epidemiological Typing of Salmonella. (2014).

22. Hong, Y., Cunneen, M. M. & Reeves, P. R. The Wzx translocases for Salmonella enterica O-antigen processing have unexpected serotype specificity. *Mol. Microbiol.* **84**, 620–630 (2012).

23. Wattiau, P., Boland, C. & Bertrand, S. Methodologies for Salmonella enterica subsp. enterica Subtyping: Gold Standards and Alternatives. *Appl. Environ. Microbiol.* **77**, 7877–7885 (2011).

24. McQuiston, J. R., Waters, R. J., Dinsmore, B. A., Mikoleit, M. L. & Fields, P. I. Molecular Determination of H Antigens of Salmonella by Use of a Microsphere-Based Liquid Array. *J. Clin. Microbiol.* **49**, 565–573 (2011).

25. Bonifield, H. R. & Hughes, K. T. Flagellar Phase Variation in Salmonella enterica Is Mediated by a Posttranscriptional Control Mechanism. **185**, 3567–3574 (2003).

26. Steve Yan, S. *et al.* An overview of Salmonella typing. *Clin. Appl. Immunol. Rev.* **4**, 189–204 (2004).

27. Deng, X. *et al.* Comparative Analysis of Subtyping Methods against a Whole-Genome-Sequencing Standard for Salmonella enterica Serotype Enteritidis. *J. Clin. Microbiol.* **53**, 212–218 (2015).

28. den Bakker, H. C. *et al.* Rapid Whole-Genome Sequencing for Surveillance of Salmonella enterica Serovar Enteritidis. *Emerg. Infect. Dis.* **20**, 1306–1314 (2014).

29. Leekitcharoenphon, P., Nielsen, E. M., Kaas, R. S., Lund, O. & Aarestrup, F. M. Evaluation of Whole Genome Sequencing for Outbreak Detection of Salmonella enterica. *PLoS One* **9**, e87991 (2014).

30. Bertrand, S. *et al.* Multi Locus Variable-Number Tandem Repeat (MLVA) Typing Tools Improved the Surveillance of Salmonella Enteritidis: A 6 Years Retrospective Study. *PLoS One* **10**, e0117950 (2015).

31. Nielsen, E. M. Laboratory standard operating procedure for MLVA of Salmonella enterica serotype Typhimurium. *Eur. Cent. Dis. Prev. Control.* **1**, (2011).

32. ECDC. *Laboratory standard operating procedure for multiple-locus variable-number tandem repeat analysis of Salmonella enterica serotype Enteritidis*. (2016). doi:10.2900/973540

33. Pijnacker, R. *et al.* An international outbreak of Salmonella enterica serotype Enteritidis linked to eggs from Poland: a microbiological and epidemiological study. *Lancet Infect. Dis.* **3099**, 1–9 (2019).

34. Achtman, M. *et al.* Multilocus sequence typing as a replacement for serotyping in Salmonella enterica. *PLoS Pathog.* **8**, (2012).

35. Ruppitsch, W. Molecular typing of bacteria for epidemiological surveillance and outbreak investigation / Molekulare Typisierung von Bakterien für die epidemiologische Überwachung und Ausbruchsabklärung. *Die Bodenkultur J. L. Manag. Food Environ.* **67**, 199–224 (2016).

36. van Dijk, E. L., Auger, H., Jaszczyszyn, Y. & Thermes, C. Ten years of next-generation sequencing technology. *Trends Genet.* **30**, 418–426 (2014).

37. Besser, J., Carleton, H. A., Gerner-Smidt, P., Lindsey, R. L. & Trees, E. Next-generation sequencing technologies and their application to the study and control of bacterial infections. *Clin. Microbiol. Infect.* **24**, 335–341 (2018).

38. King, L. A. *et al.* Outbreak of Shiga Toxin – Producing Escherichia coli O104:H4 Associated With Organic Fenugreek Sprouts, France, June 2011. *Clin. Infect. Dis.* **54**, 1588–94 (2012).

39. Barzilay, E. J. *et al.* Cholera Surveillance during the Haiti Epidemic — The First 2 Years. *N. Engl. J. Med.* **368**, 599–609 (2013).

40. Cao, Y., Fanning, S., Proos, S., Jordan, K. & Srikumar, S. A Review on the Applications

of Next Generation Sequencing Technologies as Applied to Food-Related Microbiome Studies. *Front. Microbiol.* **8**, 1–16 (2017).

41. Kchouk, M., Gibrat, J. F. & Elloumi, M. Generations of Sequencing Technologies: From First to Next Generation. *Biol. Med.* **09**, 1–8 (2017).

42. Ambardar, S., Gupta, R., Trakroo, D., Lal, R. & Vakhlu, J. High Throughput Sequencing: An Overview of Sequencing Chemistry. *Indian J. Microbiol.* **56**, 394–404 (2016).

43. Lynch, T., Petkau, A., Knox, N., Graham, M. & Van Domselaar, G. A Primer on Infectious Disease Bacterial Genomics. *Clin. Microbiol. Rev.* **29**, 881–913 (2016).

44. Rantsiou, K. *et al.* Next generation microbiological risk assessment: opportunities of whole genome sequencing (WGS) for foodborne pathogen surveillance, source tracking and risk assessment. *Int. J. Food Microbiol.* **287**, 3–9 (2018).

45. Llarena, A.-K. *et al.* INNUENDO: A cross-sectoral platform for the integration of genomics in the surveillance of food-borne pathogens. *EFSA Support. Publ.* **15**, (2018).

46. Carriço, J. A., Rossi, M., Moran-Gilad, J., Van Domselaar, G. & Ramirez, M. A primer on microbial bioinformatics for nonbioinformaticians. *Clin. Microbiol. Infect.* **24**, 342–349 (2018).

47. Kwong, J. C., Mccallum, N., Sintchenko, V. & Howden, B. P. Whole genome sequencing in clinical and public health microbiology. *Pathology* **47**, 199–210 (2015).

48. Zhang, S. *et al.* Salmonella Serotype Determination Utilizing High-Throughput Genome Sequencing Data. *J. Clin. Microbiol.* **53**, 1685–1692 (2015).

49. Zankari, E. *et al.* Identification of acquired antimicrobial resistance genes. *J. Antimicrob. Chemother.* **67**, 2640–2644 (2012).

50. Carattoli, A. *et al.* In Silico Detection and Typing of Plasmids using PlasmidFinder and Plasmid Multilocus Sequence Typing. *Antimicrob. Agents Chemother.* **58**, 3895–3903 (2014).

51. Larsen, M. V *et al.* Multilocus Sequence Typing of Total-Genome-Sequenced Bacteria. *J. Clin. Microbiol.* **50**, 1355–1361 (2012).

52. Arndt, D. *et al.* PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Res.* **44**, W16–W21 (2016).

53. Jia, B. *et al.* CARD 2017: expansion and model-centric curation of the comprehensive

antibiotic resistance database. *Nucleic Acids Res.* **45**, D566–D573 (2017).

54. Yoshida, C. E. *et al.* 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 (2016).

55. Dallman, T. *et al.* Phylogenetic structure of European Salmonella Enteritidis outbreak correlates with national and international egg distribution network. *Microb. Genomics* **2**, 1–8 (2016).

56. Portmann, A. C. *et al.* A validation approach of an end-to-end whole genome sequencing workflow for source tracking of Listeria monocytogenes and Salmonella enterica. *Front. Microbiol.* (2018). doi:10.3389/fmicb.2018.00446

57. Petrovska, L. *et al.* Microevolution of monophasic Salmonella typhimurium during epidemic, United Kingdom, 2005–2010. *Emerg. Infect. Dis.* (2016). doi:10.3201/eid2204.150531

58. Hammarlöf, D. L. *et al.* Role of a single noncoding nucleotide in the evolution of an epidemic African clade of Salmonella. *Proc. Natl. Acad. Sci.* **115**, E2614–E2623 (2018).

59. Graham, R. M. A., Hiley, L., Rathnayake, I. U. & Jennison, A. V. Comparative genomics identifies distinct lineages of S. Enteritidis from Queensland, Australia. *PLoS One* **13**, e0191042 (2018).

60. Langridge, G. C. *et al.* Patterns of genome evolution that have accompanied host adaptation in Salmonella. *Proc. Natl. Acad. Sci.* **112**, 863–868 (2015).

61. Ribeiro-Gonçalves, B., Francisco, A. P., Vaz, C., Ramirez, M. & Carriço, J. A. PHYLOViZ Online: web-based tool for visualization, phylogenetic inference, analysis and sharing of minimum spanning trees. *Nucleic Acids Res.* **44**, W246–W251 (2016).

62. Zhou, Z. *et al.* GrapeTree: visualization of core genomic relationships among 100,000 bacterial pathogens. *Genome Res.* **28**, 1395–1404 (2018).

63. Alikhan, N.-F., Zhou, Z., Sergeant, M. J. & Achtman, M. A genomic overview of the population structure of Salmonella. *PLOS Genet.* **14**, e1007261 (2018).

64. Nadon, C. *et al.* Pulsenet international: Vision for the implementation of whole genome sequencing (WGS) for global foodborne disease surveillance. *Eurosurveillance* (2017). doi:10.2807/1560-7917.ES.2017.22.23.30544

65. McNally, A. *et al.* Combined Analysis of Variation in Core, Accessory and Regulatory Genome Regions Provides a Super-Resolution View into the Evolution of Bacterial Populations. *PLOS Genet.* **12**, e1006280 (2016).

66. Schürch, A. C., Arredondo-Alonso, S., Willems, R. J. L. & Goering, R. V. Whole genome sequencing options for bacterial strain typing and epidemiologic analysis based on single nucleotide polymorphism versus gene-by-gene–based approaches. *Clin. Microbiol. Infect.* **24**, 350–354 (2018).

67. Doolittle, R. F., Feng, D., Tsang, S., Cho, G. & Little, E. Determining Divergence Times of the Major Kingdoms. **271**, (1996).

68. de Jong, H. K., Parry, C. M., van der Poll, T. & Wiersinga, W. J. Host–Pathogen Interaction in Invasive Salmonellosis. *PLoS Pathog.* **8**, e1002933 (2012).

69. Bäumler, A. J., Tsolis, R. M., Ficht, T. A. & Adams, L. G. Evolution of host adaptation in Salmonella enterica. *Infect. Immun.* **66**, 4579–4587 (1998).

70. Tanner, J. R. & Kingsley, R. A. Evolution of Salmonella within Hosts. *Trends Microbiol.* **26**, 986–998 (2018).

71. Porwollik, S. & McClelland, M. Lateral gene transfer in Salmonella. *Microbes Infect.* **5**, 977–989 (2003).

72. NCBI. Salmonella enterica. Available at: https://www.ncbi.nlm.nih.gov/genome/genomes/152.

73. Desai, P. *et al.* Evolutionary Genomics of Salmonella enterica Subspecies. *MBio* **4**, (2013).

74. David Alcaraz, L. Pan-genomics: Unmasking the gene diversity hidden in the bacteria species. *PeerJ Prepr.* | 1–7 (2014). doi:10.7287/peerj.preprints.113v2

75. Gordienko, E. N., Kazanov, M. D. & Gelfand, M. S. Evolution of Pan-Genomes of Escherichia coli, Shigella spp., and Salmonella enterica. *J. Bacteriol.* **195**, 2786–2792 (2013).

76. Leekitcharoenphon, P., Lukjancenko, O., Friis, C., Aarestrup, F. M. & Ussery, D. W. Genomic variation in Salmonella enterica core genes for epidemiological typing. *BMC Genomics* **13**, 88 (2012).

77. Andino, A. & Hanning, I. Salmonella enterica: Survival, Colonization, and Virulence Differences among Serovars. *Sci. World J.* **2015**, 1–16 (2015).

78. Hurley, D., McCusker, M. P., Fanning, S. & Martins, M. Salmonellaâ€“Host Interactions â€“ Modulation of the Host Innate Immune System. *Front. Immunol.* **5**, (2014).

79. MacKenzie, K. D., Palmer, M. B., Köster, W. L. & White, A. P. Examining the Link between Biofilm Formation and the Ability of Pathogenic Salmonella Strains to Colonize Multiple Host Species. *Front. Vet. Sci.* **4**, (2017).

80. Dekker, J. P. & Frank, K. M. Salmonella, Shigella, and Yersinia. *Clin. Lab. Med.* **35**, 225–246 (2015).

81. Stanaway, J. D. *et al.* The global burden of typhoid and paratyphoid fevers: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet Infect. Dis.* **19**, 369–381 (2019).

82. World Health Organization. *Weekly epidemiological record - Typhoid vaccines: WHO position paper*. (2018).

83. Crump, J. A. Progress in Typhoid Fever Epidemiology. *Clin. Infect. Dis.* **68**, S4–S9 (2019).

84. Besser, J. M. Salmonella epidemiology: A whirlwind of change. *Food Microbiol.* **71**, 55–59 (2018).

85. ECDC (European Centre for Disease Prevention and Control). *Annual Epidemiological Report for 2016 Salmonellosis*. (2019).

86. Majowicz, S. E. *et al.* The Global Burden of Nontyphoidal Salmonella Gastroenteritis. *Clin. Infect. Dis.* **50**, 882–889 (2010).

87. EFSA&ECDC. *The European Union summary report on trends and sources of zoonoses, zoonotic agents and food-borne outbreaks in 2016. EFSA Journal* **15**, (2017).

88. European Food Safety Authority. *The European Union summary report on trends and sources of zoonoses, zoonotic agents and food-borne outbreaks in 2016*. *EFSA Journal* **15**, (2017).

89. ECDC. *Annual Epidemiological Report for 2016 - Typhoid and paratyphoid fevers*. (2018).

90. Fabrega, A. & Vila, J. Salmonella enterica Serovar Typhimurium Skills To Succeed in the Host: Virulence and Regulation. *Clin. Microbiol. Rev.* **26**, 308–341 (2013).

91. Yue, M. & Schifferli, D. M. Allelic variation in Salmonella: an underappreciated driver of adaptation and virulence. *Front. Microbiol.* **4**, (2014).

92. dos Santos, A. M. P., Ferrari, R. G. & Conte-Junior, C. A. Virulence Factors in Salmonella Typhimurium: The Sagacity of a Bacterium. *Curr. Microbiol.* **76**, 762–773 (2019).

93. Crump, J. A., Sjölund-Karlsson, M., Gordon, M. A. & Parry, C. M. Epidemiology, Clinical Presentation, Laboratory Diagnosis, Antimicrobial Resistance, and Antimicrobial Management of Invasive Salmonella Infections. *Clin. Microbiol. Rev.* **28**, 901–937 (2015).

94. Deksissa, T. & Gebremedhin, E. Z. A cross-sectional study of enteric fever among febrile patients at Ambo hospital: prevalence, risk factors, comparison of Widal test and stool culture and antimicrobials susceptibility pattern of isolates. *BMC Infect. Dis.* **19**, 288 (2019).

95. Hohmann, E. L. Nontyphoidal Salmonellosis. *Clin. Infect. Dis.* **32**, 263–269 (2001).

96. Eng, S.-K. *et al.* Salmonella : A review on pathogenesis, epidemiology and antibiotic resistance. *Front. Life Sci.* **8**, 284–293 (2015).

97. Guarino, A. *et al.* European Society for Pediatric Gastroenterology, Hepatology, and Nutrition/European Society for Pediatric Infectious Diseases Evidence-Based Guidelines for the Management of Acute Gastroenteritis in Children in Europe. *J. Pediatr. Gastroenterol. Nutr.* **59**, 132–152 (2014).

98. Cuypers, W. L. *et al.* Fluoroquinolone resistance in Salmonella: insights by whole-genome sequencing. *Microb. Genomics* **4**, (2018).

99. Eaves, D. J. *et al.* Prevalence of Mutations within the Quinolone Resistance-Determining Region of gyrA, gyrB, parC, and parE and Association with Antibiotic Resistance in Quinolone-Resistant Salmonella enterica. *Antimicrob. Agents Chemother.* **48**, 4012–4015 (2004).

100. García-Fernández, A., Fortini, D., Veldman, K., Mevius, D. & Carattoli, A. Characterization of plasmids harbouring qnrS1, qnrB2 and qnrB19 genes in Salmonella. *J. Antimicrob. Chemother.* **63**, 274–281 (2009).

101. Yanagi, D. *et al.* Emergence of fluoroquinolone-resistant strains of Salmonella enterica in Surabaya, Indonesia. *Diagn. Microbiol. Infect. Dis.* **64**, 422–426 (2009).

102. Falagas, M. E., Kasiakou, S. K. & Saravolatz, L. D. Colistin: The Revival of Polymyxins for the Management of Multidrug-Resistant Gram-Negative Bacterial Infections. *Clin. Infect. Dis.* **40**, 1333–1341 (2005).

103. Parisi, A. *et al.* Health Outcomes from Multidrug-Resistant Salmonella Infections in High-Income Countries: A Systematic Review and Meta-Analysis. *Foodborne Pathog. Dis.* **15**, 428–436 (2018).

104. WHO. *Antimicrobial resistance — Global Report on Surveillance*. (2014).

105. Liu, Y.-Y. *et al.* Emergence of plasmid-mediated colistin resistance mechanism MCR-1 in animals and human beings in China: a microbiological and molecular biological study. *Lancet Infect. Dis.* **16**, 161–168 (2016).

106. Quesada, A. *et al.* Detection of plasmid mediated colistin resistance (MCR-1) in Escherichia coli and Salmonella enterica isolated from poultry and swine in Spain. *Res. Vet. Sci.* **105**, 134–135 (2016).

107. Veldman, K. *et al.* Location of colistin resistance gene mcr-1 in Enterobacteriaceae from livestock and meat: Table 1. *J. Antimicrob. Chemother.* **71**, 2340–2342 (2016).

108. Figueiredo, R. *et al.* Detection of an mcr-1 -encoding plasmid mediating colistin resistance in Salmonella enterica from retail meat in Portugal: Table 1. *J. Antimicrob. Chemother.* **71**, 2338–2340 (2016).

109. Campos, J., Cristino, L., Peixe, L. & Antunes, P. MCR-1 in multidrug-resistant and copper-tolerant clinically relevant Salmonella 1,4,[5],12:i:- and S . Rissen clones in Portugal, 2011 to 2015. *Eurosurveillance* **21**, 30270 (2016).

110. Doumith, M. *et al.* Detection of the plasmid-mediated mcr-1 gene conferring colistin resistance in human and food isolates of Salmonella enterica and Escherichia coli in England and Wales. *J. Antimicrob. Chemother.* **71**, 2300–2305 (2016).

111. Nang, S. C., Li, J. & Velkov, T. The rise and spread of mcr plasmid-mediated polymyxin resistance. *Crit. Rev. Microbiol.* **0**, 31 (2019).

112. Ventola, C. L. The Antibiotic Resistance Crisis Part 1 : Causes and Threats. *Pharm. Ther.* **40**, 277–283 (2015).

113. Alexander Fleming. Nobel Lecture, December 11, 1945. (1945).

114. Davies, J. & Davies, D. Origins and Evolution of Antibiotic Resistance. *Microbiol. Mol. Biol. Rev.* **74**, 417–433 (2010).

115. Lewis, K. Platforms for antibiotic discovery. *Nat. Rev. Drug Discov.* **12**, 371–387 (2013).

116. Exner, M. *et al.* Antibiotic resistance: What is so special about multidrug-resistant

Gram-negative bacteria? *GMS Hyg Infect Control* **12**, 1–24 (2017).

117. Porwollik, S. & McClelland, M. Lateral gene transfer in Salmonella. *Microbes and Infection* (2003). doi:10.1016/S1286-4579(03)00186-2

118. Alcaine, S. D., Warnick, L. D. & Wiedmann, M. Antimicrobial Resistance in Nontyphoidal Salmonella. *J. Food Prot.* **70**, 780–790 (2007).

119. Dhillon, R. H.-P. & Clark, J. ESBLs: A Clear and Present Danger? *Crit. Care Res. Pract.* **2012**, 1–11 (2012).

120. Cantón, R., González-Alba, J. M. & Galán, J. C. CTX-M Enzymes: Origin and Diffusion. *Front. Microbiol.* **3**, (2012).

121. Krause, K. M., Serio, A. W., Kane, T. R. & Connolly, L. E. Aminoglycosides: An Overview. *Cold Spring Harb. Perspect. Med.* **6**, a027029 (2016).

122. Cuypers, W. L. *et al.* Fluoroquinolone resistance in Salmonella: insights by whole-genome sequencing. *Microb. Genomics* **4**, (2018).

123. Hopkins, K. L., Davies, R. H. & Threlfall, E. J. Mechanisms of quinolone resistance in Escherichia coli and Salmonella: Recent developments. *Int. J. Antimicrob. Agents* **25**, 358–373 (2005).

124. Hopkins, K. L., Davies, R. H. & Threlfall, E. J. Mechanisms of quinolone resistance in Escherichia coli and Salmonella: Recent developments. *Int. J. Antimicrob. Agents* **25**, 358–373 (2005).

125. Vester, B. & Douthwaite, S. Macrolide Resistance Conferred by Base Substitutions in 23S rRNA. *Antimicrob. Agents Chemother.* **45**, 1–12 (2001).

126. Nair, S. *et al.* WGS for surveillance of antimicrobial resistance: A pilot study to detect the prevalence and mechanism of resistance to azithromycin in a UK population of non-typhoidal Salmonella. *J. Antimicrob. Chemother.* **71**, 3400–3408 (2016).

127. Aghapour, Z. *et al.* Molecular mechanisms related to colistin resistance in Enterobacteriaceae. *Infect. Drug Resist.* **12**, 965–975 (2019).

128. Argudín, M. A., Hoefer, A. & Butaye, P. Heavy metal resistance in bacteria from animals. *Res. Vet. Sci.* **122**, 132–147 (2019).

129. EC (1831/2003). Regulation (EC) No 1831/2003 of the European Parliament and of the Council of 22 September 2003 on additives for use in animal nutrition. *Off. J. Eur. Union* **4**, L 268/29-42 (2003).

130. Figueiredo, R. *et al.* Multidrug-Resistant Salmonella enterica Isolated from Food Animal and Foodstuff May Also Be Less Susceptible to Heavy Metals. *Foodborne Pathog. Dis.* **16**, 166–172 (2019).

131. Mourão, J. *et al.* Tolerance to multiple metal stressors in emerging non-typhoidal MDR Salmonella serotypes: A relevant role for copper in anaerobic conditions. *J. Antimicrob. Chemother.* **71**, 2147–2157 (2016).

132. Mourão, J., Novais, C., Machado, J., Peixe, L. & Antunes, P. Metal tolerance in emerging clinically relevant multidrug-resistant Salmonella enterica serotype 4,[5],12:i:- clones circulating in Europe. *Int. J. Antimicrob. Agents* **45**, 610–6 (2015).

133. Romero, J. L., Grande Burgos, M. J., Pérez-Pulido, R., Gálvez, A. & Lucas, R. Resistance to Antibiotics, Biocides, Preservatives and Metals in Bacteria Isolated from Seafoods: Co-Selection of Strains Resistant or Tolerant to Different Classes of Compounds. *Front. Microbiol.* **8**, 1–16 (2017).

134. Pal, C., Bengtsson-Palme, J., Kristiansson, E. & Larsson, D. G. J. Co-occurrence of resistance genes to antibiotics, biocides and metals reveals novel insights into their co-selection potential. *BMC Genomics* **16**, 1–14 (2015).

135. Kaur, J. & Jain, S. K. Role of antigens and virulence factors of Salmonella enterica serovar Typhi in its pathogenesis. *Microbiol. Res.* **167**, 199–210 (2012).

136. Nascimento, A. M. A. & Chartone-Souza, E. Operon mer: Bacterial resistance to mercury and potential for bioremediation of contaminated environments. *Genet. Mol. Res.* **2**, 92–101 (2003).

137. Instituto Nacional de Saúde Doutor Ricardo Jorge. (2017). Available at: http://www.insa.pt/. (Accessed: 21st August 2019)

138. Figueiredo, A. Serotipos de Salmonella identificados no Instituto Superior de Higiene Dr. Ricardo Jorge de 1950 a 1971. *Arq. do Inst. Nac. Saúde* **1**, 187–193 (1972).

139. ECDC. *Annual epidemiological report 2014 food and waterborne diseaes and zoonoses*. (2014). doi:10.17113/ft

140. Comissão Europeia. Regulamento (CE) N.º 2160/2003 DO PARLAMENTO EUROPEU E DO CONSELHO de 17 de Novembro de 2003 relativo ao controlo de salmonelas e outros agentes zoonóticos específicos de origem alimentar. *J. Of. da União Eur.* **325**, 1–25 (2003).

141. Direção-Geral da Saúde. Doenças de declaração obrigatória 2010-2013. **II**, (2015).

142. Ministério da Saúde & Direção-Geral da Saúde. Doenças de Declaração Obrigatória 2013-2016. **I**, 1–173 (2017).

143. Direção-Geral da Saúde. Doenças de Declaração Obrigatória - 1996-2000. **II**, (2013).

144. Direção-Geral da Saúde. Doenças de Declaração Obrigatória - 2000-2004. **II**, (2005).

145. Direção-Geral da Saúde. Doenças de Declaração Obrigatória - 2004-2008. (2010).

146. Machado, J. & Bernardo, F. Prevalence of Salmonella in chicken carcasses in Portugal. *J. Appl. Bacteriol.* **69**, 477–480 (1990).

147. den Bakker, H. C. *et al.* A Whole-Genome Single Nucleotide Polymorphism-Based Approach To Trace and Identify Outbreaks Linked to a Common Salmonella enterica subsp. enterica Serovar Montevideo Pulsed-Field Gel Electrophoresis Type. *Appl. Environ. Microbiol.* **77**, 8648–8655 (2011).

148. Taylor, A. J. *et al.* 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 (2015).

149. Ashton, P. M. *et al.* Whole Genome Sequencing for the Retrospective Investigation of an Outbreak of Salmonella Typhimurium DT 8. *PLoS Curr.* **7**, (2015).

150. Angelo, K. M. *et al.* Outbreak of salmonella newport infections linked to cucumbers — United States, 2014. *Morb. Mortal. Wkly. Rep.* **64**, 144–147 (2015).

151. Ferreira, M. P. R. Transition from the traditional typing methods to the application of whole genome sequencing (WGS) for surveillance of food- and waterborne diseases. (Universidade de Lisboa - Faculdade de Farmácia, 2016).

152. Macedo, R. *et al.* Evaluation of a gene-by-gene approach for prospective whole-genome sequencing-based surveillance of multidrug resistant Mycobacterium tuberculosis. *Tuberculosis* **115**, 81–88 (2019).

153. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).

154. Bankevich, A. *et al.* SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. *J. Comput. Biol.* **19**, 455–477 (2012).

155. Walker, B. J. *et al.* Pilon: An Integrated Tool for Comprehensive Microbial Variant Detection and Genome Assembly Improvement. *PLoS One* **9**, e112963 (2014).

156. Rossi, M. *et al.* INNUENDO whole genome and core genome MLST schemas and datasets for Salmonella enterica. (2018). doi:10.5281/ZENODO.1323684

157. Silva, M. *et al.* chewBBACA: A complete suite for gene-by-gene schema creation and strain identification. *Microb. Genomics* **4**, 1–7 (2018).

158. Francisco, A. P. *et al.* PHYLOViZ: phylogenetic inference and data visualization for sequence based typing methods. *BMC Bioinformatics* **13**, 87 (2012).

159. Francisco, A. P., Bugalho, M., Ramirez, M. & Carriço, J. A. Global optimal eBURST analysis of multilocus typing data using a graphic matroid approach. *BMC Bioinformatics* **10**, 152 (2009).

160. Soares, A. R. & Machado, J. Molecular epidemiology of Salmonella enteritidis phage type 1b and 6a isolates in Portugal. *Epidemiol. Infect.* **131**, 607–612 (2003).

161. ECDC & EFSA. *Multi-country outbreak of Salmonella Enteritidis phage type 8, MLVA profile 2-9-7-3-2 and 2-9-6-3-2 infections*. (2017).

162. ECDC. *Annual epidemiological report for 2015 - Salmonellosis*. (2018).

163. Ashton, P. M. *et al.* Identification of Salmonella for public health surveillance using whole genome sequencing. *PeerJ* **4**, e1752 (2016).

164. Inns, T. *et al.* A multi-country Salmonella Enteritidis phage type 14b outbreak associated with eggs from a German producer: 'near real-time' application of whole genome sequencing and food chain investigations, United Kingdom, May to September 2014. *Eurosurveillance* **20**, (2015).

165. Inns, T. *et al.* Prospective use of whole genome sequencing (WGS) detected a multi-country outbreak of Salmonella Enteritidis. *Epidemiol. Infect.* **145**, 289–298 (2017).

166. Hopkins, K. L., Peters, T. M., de Pinna, E. & Wain, J. Standardisation of multilocus variable-number tandemrepeat analysis (MLVA) for subtyping of Salmonella enterica serovar Enteritidis. *Eurosurveillance* **16**, pii=19942 (2011).

167. Freidl, G. *et al.* Tracing Back the Source of an Outbreak of Salmonella Typhimurium; National Outbreak Linked to the Consumption of Raw and Undercooked Beef Products, the Netherlands, October to December 2015. *PLoS Curr.* 1–14 (2018). doi:10.1371/currents.outbreaks.1c667d62b51eb9840f5f7eb617e56bc1

168. Crim, S. M. *et al.* Salmonella enterica Serotype Newport Infections in the United States, 2004–2013: Increased Incidence Investigated Through Four Surveillance Systems. *Foodborne Pathog. Dis.* **15**, 612–620 (2018).

169. Sangal, V. *et al.* Evolution and Population Structure of Salmonella enterica Serovar Newport. *J. Bacteriol.* **192**, 6465–6476 (2010).

170. Hendriksen, R. S. *et al.* Global Monitoring of Salmonella Serovar Distribution from the World Health Organization Global Foodborne Infections Network Country Data Bank: Results of Quality Assured Laboratories from 2001 to 2007. *Foodborne Pathog. Dis.* **8**, 887–900 (2011).

171. ECDC. *Annual epidemiological report 2014 - food- and waterborne diseases and zoonoses*. (2014).

172. Bonardi, S. Salmonella in the pork production chain and its impact on human health in the European Union. *Epidemiol. Infect.* **145**, 1513–1526 (2017).

173. Bangtrakulnonth, A. *et al.* Salmonella Serovars from Humans and Other Sources in Thailand. *Emerg. Infect. Dis.* **10**, 131–136 (2004).

174. Prathan, R., Bitrus, A. A., Sinwat, N., Angkititrakul, S. & Chuanchuen, R. Phylogenetic characterization of Salmonella enterica from pig production and humans in Thailand and Laos border provinces. *Vet. World* **12**, 79–84 (2019).

175. Silveira, L., Marques, A. & Machado, J. Infeções por Salmonella enterica no período entre 2000-2012. *Bol. Epidemiológico Obs.* 14–16 (2013).

176. EFSA&ECDC. *The European Union summary report on antimicrobial resistance in zoonotic and indicator bacteria from humans, animals and food in 2015*. *EFSA Journal* (2017). doi:10.2903/j.efsa.2017.4694

177. Clemente, L. *et al.* Occurrence of extended-spectrum β-lactamases among isolates of Salmonella enterica subsp. enterica from food-producing animals and food products, in Portugal. *Int. J. Food Microbiol.* **167**, 221–228 (2013).

178. Hendriksen, R. S. *et al.* Antimicrobial Resistance and Molecular Epidemiology of Salmonella Rissen from Animals, Food Products, and Patients in Thailand and Denmark. *Foodborne Pathog. Dis.* **5**, 605–619 (2008).

179. Boschi, T. *et al.* Cluster of Cases of Salmonella enterica Serotype Rissen Infection in a General Hospital, Italy, 2007. *Zoonoses Public Health* **57**, 518–522 (2010).

180. Nguyen, D. T. A. *et al.* Prevalence, antibiotic resistance, and extended-spectrum and AmpC β-lactamase productivity of Salmonella isolates from raw meat and seafood samples in Ho Chi Minh City, Vietnam. *Int. J. Food Microbiol.* (2016). doi:10.1016/j.ijfoodmicro.2016.07.017

181. Zhang, L. *et al.* Highly Prevalent Multidrug-Resistant Salmonella From Chicken and Pork Meat at Retail Markets in Guangdong, China. *Front. Microbiol.* **9**, (2018).

182. Nadimpalli, M. *et al.* CTX-M-55-type ESBL-producing Salmonella enterica are emerging among retail meats in Phnom Penh, Cambodia. *J. Antimicrob. Chemother.* (2018). doi:10.1016/0038-1098(79)91043-3

183. García-Fierro, R., Montero, I., Bances, M., González-Hevia, M. Á. & Rodicio, M. R. Antimicrobial Drug Resistance and Molecular Typing of Salmonella enterica Serovar Rissen from Different Sources. *Microb. Drug Resist.* **22**, 211–217 (2016).

184. Im, M. C. *et al.* Prevalence and characteristics of Salmonella spp. isolated from commercial layer farms in Korea. *Poult. Sci.* **94**, 1691–1698 (2015).

185. Kumar, R., Surendran, P. K. & Thampuran, N. Distribution and genotypic characterization of Salmonella serovars isolated from tropical seafood of Cochin, India. *J. Appl. Microbiol.* **106**, 515–524 (2009).

186. Vieira-Pinto, M. *et al.* Salmonella sp. in Game (Sus scrofa and Oryctolagus cuniculus). *Foodborne Pathog. Dis.* **8**, 739–740 (2011).

187. Gomes-Neves, E. *et al.* Salmonella cross-contamination in swine abattoirs in Portugal: Carcasses, meat and meat handlers. *Int. J. Food Microbiol.* **157**, 82–87 (2012).

188. Gomes-Neves, E. *et al.* Clinically relevant multidrug resistant Salmonella enterica in swine and meat handlers at the abattoir. *Vet. Microbiol.* **168**, 229–233 (2014).

189. Wuyts, V. *et al.* MLVA as a Tool for Public Health Surveillance of Human Salmonella Typhimurium: Prospective Study in Belgium and Evaluation of MLVA Loci Stability. *PLoS One* **8**, e84055 (2013).

190. Barco, L. *et al.* Molecular Characterization of Salmonella enterica Serovar 4,[5],12:i:-DT193 ASSuT Strains from Two Outbreaks in Italy. *Foodborne Pathog. Dis.* **11**, 138–144 (2014).

191. Pearce, M. E. *et al.* Comparative analysis of core genome MLST and SNP typing within a European Salmonella serovar Enteritidis outbreak. *Int. J. Food Microbiol.* (2018).

doi:10.1016/j.ijfoodmicro.2018.02.023

192. Palma, F. *et al.* Genome-wide identification of geographical segregated genetic markers in Salmonella enterica serovar Typhimurium variant 4,[5],12:i:-. *Sci. Rep.* (2018). doi:10.1038/s41598-018-33266-5

193. EUCAST. Disk Diffusion Method for Antimicrobial Susceptibility Testing Antimicrobial susceptibility testing EUCAST disk diffusion method. (2017).

194. EUCAST. *The European Committee on Antimicrobial Susceptibility Testing. Breakpoint tables for interpretation of MICs and zone diameters European Committee on Antimicrobial Susceptibility Testing Breakpoint tables for interpretation of MICs and zone diameters. Ver*. (2015).

195. EUCAST. *The European Committee on Antimicrobial Susceptibility Testing. Breakpoint tables for interpretation of MICs and zone diameters. Version 6.0, 2016*. (2016).

196. EUCAST. *The European Committee on Antimicrobial Susceptibility Testing. Breakpoint tables for interpretation of MICs and zone diameters. Version 7.0, 2017*. (2017).

197. EUCAST. *The European Committee on Antimicrobial Susceptibility Testing. Breakpoint tables for interpretation of MICs and zone diameters. Version 7.1, 2017*. (2017).

198. Seemann, T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* **30**, 2068–2069 (2014).

199. Vieira Pinto, M., Tenreiro, R. & Martins, C. Unveiling contamination sources and dissemination routes of Salmonella sp. in pigs at a Portuguese slaughterhouse through macrorestriction profiling by pulsed-field gel electrophoresis. *Int. J. Food Microbiol.* **110**, 77–84 (2006).

200. Antunes, P., Mourao, J., Pestana, N. & Peixe, L. Leakage of emerging clinically relevant multidrug-resistant Salmonella clones from pig farms. *J. Antimicrob. Chemother.* **66**, 2028–2032 (2011).

201. Partridge, S. R., Kwong, S. M., Firth, N. & Jensen, S. O. Mobile Genetic Elements Associated with Antimicrobial Resistance. *Clin. Microbiol. Rev.* **31**, 1–61 (2018).

202. Magiorakos, A.-P. *et al.* Multidrug-resistant, extensively drug-resistant and pandrug-resistant bacteria: an international expert proposal for interim standard definitions for acquired resistance. *Clin. Microbiol. Infect.* **18**, 268–281 (2012).

203. Monte, D. F. *et al.* Genomic Features of High-Priority Salmonella enterica Serovars

Circulating in the Food Production Chain, Brazil, 2000–2016. *Sci. Rep.* **9**, 11058 (2019).

204. Parisi, A. *et al.* Increasing incidence of invasive nontyphoidal Salmonella infections in Queensland, Australia, 2007-2016. *PLoS Negl. Trop. Dis.* **13**, e0007187 (2019).

205. Antunes, P., Mourão, J., Campos, J. & Peixe, L. Salmonellosis: The role of poultry meat. *Clin. Microbiol. Infect.* **22**, 110–121 (2016).

206. Voss-Rech, D. *et al.* Longitudinal study reveals persistent environmental Salmonella Heidelberg in Brazilian broiler farms. *Vet. Microbiol.* **233**, 118–123 (2019).

207. Voss-Rech, D. *et al.* A temporal study of Salmonella enterica serotypes from broiler farms in Brazil. *Poult. Sci.* **94**, 433–441 (2015).

208. EFSA&ECDC. *Multi-country outbreak of Salmonella Enteritidis infections linked to Polish eggs*. (2017).

209. Edirmanasinghe, R. *et al.* A Whole-Genome Sequencing Approach To Study Cefoxitin-Resistant Salmonella enterica Serovar Heidelberg Isolates from Various Sources. *Antimicrob. Agents Chemother.* **61**, (2017).

210. Santin, E., Hayashi, R. M. & Wammes, J. C. Phenotypic and Genotypic Features of a Salmonella Heidelberg Strain Isolated in Broilers in Brazil and Their Possible Association to Antibiotics and Short-Chain Organic Acids Resistance and Susceptibility. **4**, 1–13 (2017).

211. Antony, L. *et al.* Genome divergence and increased virulence of outbreak associated Salmonella enterica subspecies enterica serovar Heidelberg. *Gut Pathog.* **10**, 53 (2018).

212. Castellanos, L. R. *et al.* Genomic Characterization of Extended-Spectrum Cephalosporin-Resistant Salmonella enterica in the Colombian Poultry Chain. *Front. Microbiol.* **9**, 1–11 (2018).

213. Bearson, B. L., Bearson, S. M. D., Looft, T., Cai, G. & Shippy, D. C. Characterization of a Multidrug-Resistant Salmonella enterica Serovar Heidelberg Outbreak Strain in Commercial Turkeys: Colonization, Transmission, and Host Transcriptional Response. *Front. Vet. Sci.* **4**, 1–7 (2017).

214. Amand, J. A. S., Otto, S. J. G., Cassis, R. & Christianson, C. B. A. Antimicrobial resistance of Salmonella enterica serovar Heidelberg isolated from poultry in Alberta

Antimicrobial resistance of Salmonella enterica serovar Heidelberg isolated from poultry in Alberta. *Avian Pathol.* **46**, 379–386 (2013).

215. Vincent, C. *et al.* Comparison of advanced whole genome sequence-based methods to distinguish strains of Salmonella enterica serovar Heidelberg involved in foodborne outbreaks in Québec. *Food Microbiol.* **73**, 99–110 (2018).

216. Mendonça, E. P. *et al.* Spread of the serotypes and antimicrobial resistance in strains of Salmonella spp. isolated from broiler. *Brazilian J. Microbiol.* **50**, 515–522 (2019).

217. EUCAST. *The European Committee on Antimicrobial Susceptibility Testing. Breakpoint tables for interpretation of MICs and zone diameters. Version 4.0, 2014*. (2014).

218. EUCAST. *The European Committee on Antimicrobial Susceptibility Testing. Breakpoint tables for interpretation of MICs and zone diameters. Version 5.0, 2015*. (2015).

219. EUCAST. *The European Committee on Antimicrobial Susceptibility Testing. Breakpoint tables for interpretation of MICs and zone diameters. Version 8.1, 2018*. (2018).

220. EUCAST. *The European Committee on Antimicrobial Susceptibility Testing. Breakpoint tables for interpretation of MICs and zone diameters. Version 9.0, 2019*. (2019).

221. EUCAST. *The European Committee on Antimicrobial Susceptibility Testing. Breakpoint tables for interpretation of MICs and zone diameters. Version 8.0, 2018*. (2018).

222. Aziz, R. K. *et al.* The RAST Server: Rapid Annotations using Subsystems Technology. *BMC Genomics* **9**, 75 (2008).

223. Brettin, T. *et al.* RASTtk: A modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. *Sci. Rep.* **5**, 8365 (2015).

224. Overbeek, R. *et al.* The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic Acids Res.* **42**, 206–214 (2014).

225. Berg, R. R. Van Den *et al.* Characterization and whole genome sequencing of closely related multidrug- resistant Salmonella enterica serovar Heidelberg isolates from imported poultry meat in the Netherlands. 1–20 (2019). doi:10.6084/m9.figshare.8796875.Multidrug-resistant

226. Bekal, S. *et al.* Usefulness of High-Quality Core Genome Single-Nucleotide Variant Analysis for Subtyping the Highly Clonal and the Most Prevalent Salmonella enterica Serovar Heidelberg Clone in the Context of Outbreak Investigations. *J. Clin. Microbiol.*

54, 289–295 (2016).

227. Hoffmann, M. *et al.* Comparative genomic analysis and virulence differences in closely related salmonella enterica serotype heidelberg isolates from humans, retailmeats, and animals. *Genome Biol. Evol.* (2014). doi:10.1093/gbe/evu079

228. European Commission. Commission regulation (EC No 2073/2005 of 15 November 2005 on microbial criteria for foodstuffs. *Off. J. Eur. Union* **L338**, 1–26 (2005).

229. European Commission. COMMISSION REGULATION (EC) No 2073/2005 of 15 November 2005 on microbiological criteria for foodstuffs. *Eur. Union Off. J.* **32**, (2019).

230. Tiba-casas, M. R. & Camargo, C. H. Emergence of CMY-2-Producing Salmonella Heidelberg Associated with IncI1 Plasmids Isolated from Poultry in Brazil. **00**, 1–6 (2018).

231. Giuriatti, J. *et al.* Salmonella Heidelberg: Genetic profile of its antimicrobial resistance related to extended spectrum β-lactamases (ESBLs). *Microb. Pathog.* **109**, 195–199 (2017).

232. Antony, L. *et al.* Genome divergence and increased virulence of outbreak associated Salmonella enterica subspecies enterica serovar Heidelberg. *Gut Pathog.* **10**, 1–9 (2018).

233. WHO. *Five keys for safer food*. (2006).

234. Santin, E. *et al.* Phenotypic and Genotypic Features of a Salmonella Heidelberg Strain Isolated in Broilers in Brazil and Their Possible Association to Antibiotics and Short-Chain Organic Acids Resistance and Susceptibility. *Front. Vet. Sci.* **4**, 1–13 (2017).

235. Gouvêa, R., Santos, F. dos, Aquino, M. de & Pereira VL de, A. Fluoroquinolones in industrial poultry production, bacterial resistance and food residues:a review. *Rev. Bras. Ciência Avícola* **17**, 1–10 (2015).

236. Gokulan, K. *et al.* Impact of Plasmids, Including Those EncodingVirB4/D4 Type IV Secretion Systems, on Salmonella enterica serovar Heidelberg Virulence in Macrophages and Epithelial Cells. *PLoS One* **8**, e77866 (2013).

237. Juhas, M. Type IV secretion systems and genomic islands-mediated horizontal gene transfer in Pseudomonas and Haemophilus. *Microbiol. Res.* **170**, 10–17 (2015).

238. Rehman, M. A., Yin, X., Persaud-Lachhman, M. G. & Diarra, M. S. First Detection of a

Fosfomycin Resistance Gene, fosA7, in Salmonella enterica Serovar Heidelberg Isolated from Broiler Chickens. *Antimicrob. Agents Chemother.* **61**, 1–6 (2017).

8. Supplementary material

Supplementary Table 4-1 – Metadata, genome assembly statistics, antibiotic resistance phenotype and genotype and heavy metal tolerance genotype of the Salmonella enterica serovar Rissen isolates enrolled in the present study. R -Resistant; S - Susceptible; (+) - Present; (-) - Absent; MDR - Multidrug resistant; NUT - Nomenclature of Territorial Units for Statistics; Amc – amoxicillin/clavulanic acid; Amp – ampicillin; Azm – azithromycin; Chl – chloramphenicol; Fep – cefepime; Ctx – cefotaxime; Fox – cefoxitin; Caz – ceftazidime; Cro – ceftriaxone; Gmn – gentamicin; Mem – meropenem; Nal - nalidixic acid; Pef – pefloxacin; Smx – sulfamethoxazole; Tet – tetracycline; Tgc – tigecycline; and Tmp – trimethoprim.

Isolate ID	Lab. ID	Location	Region (NUTII)	Collection Date	Sample	Source	Product	Mean depth of coverage	Genome size (bp)	# of contigs	ST	Read size (bp)	Sequencing Aparatus (Illumina)	ENA Accession #	ENA ID
PT01	Se_157-14	Braga	North	July 2014	Food	Pig	Fresh sausages	41.14	4907173	108	469	250	MiSeq	ERS3404912	PT_SE0001
PT02	Se_197-14	Azores	Azores	August 2014	Clinical	Human	Urine	64.84	4905194	93	469	150	NextSeq	ERS3404913	PT_SE0002
PT03	Se_290-14	Aveiro	North	October 2014	Clinical	Human	Not Available	63.47	4914066	104	469	250	MiSeq	ERS3404914	PT_SE0003
PT04	Se_297-14	Lisbon	Lisbon Metropolitan area	November 2014	Clinical	Human	Faeces	90.53	4919725	112	469	150	NextSeq	ERS3404915	PT_SE0004
PT05	Se_4-15	Azores	Azores	January 2015	Clinical	Human	Blood	39.5	4937155	112	469	250	MiSeq	ERS3404916	PT_SE0005
PT06	Se_72-15	Azores	Azores	February 2015	Clinical	Human	Faeces	55.52	4938746	76	469	250	MiSeq	ERS3404917	PT_SE0006
PT07	Se_90-15	Azores	Azores	March 2015	Clinical	Human	Faeces	53.44	4944290	103	469	250	MiSeq	ERS3404918	PT_SE0007
PT08	Se_123-15	Oporto	North	April 2015	Clinical	Human	Faeces	90.93	4892381	107	469	150	NextSeq	ERS3404919	PT_SE0008
PT09	Se_128-15	Oporto	North	April 2015	Clinical	Human	Faeces	80.99	4963973	112	469	150	NextSeq	ERS3404920	PT_SE0009
PT10	Se_171-15	Lisbon	Lisbon Metropolitan area	June 2015	Clinical	Human	Faeces	58.84	4863424	161	469	150	NextSeq	ERS3404921	PT_SE0010
PT11	Se_PIGA-2	Azores	Azores	June 2015	Food	Bovine	Frozen minced meat	53.46	4867520	74	469	250	MiSeq	ERS3404922	PT_SE0011
PT12	Se_PIGA-4	Not Available	Alentejo	June 2015	Food	Not Available	Blood chorizo	56.62	4975525	103	469	250	MiSeq	ERS3404923	PT_SE0012
PT13	Se_PIGA-7	Azores	Azores	June 2015	Food	Pig	Meat chorizo	59.02	4944337	84	469	250	MiSeq	ERS3404924	PT_SE0013
PT14	Se_PIGA-8	Not Available	Lisbon Metropolitan area	June 2015	Food	Pig	Meat	64.89	4945078	102	469	250	MiSeq	ERS3404925	PT_SE0014
PT15	Se_PIGA-9	Not Available	Lisbon Metropolitan area	June 2015	Animal	Pig	Piglet	75.94	4946168	79	469	250	MiSeq	ERS3404926	PT_SE0015
PT16	Se_PVRAM-7-15	Not Available	Not Available	June 2015	Animal	Pig	Piglet carcass	60.58	4995231	144	469	150	NextSeq	ERS3404927	PT_SE0016
PT17	Se_227-15	Aveiro	North	July 2015	Clinical	Human	Not Available	52.81	4843182	131	469	150	NextSeq	ERS3404928	PT_SE0017
PT18	Se_254-15	Oporto	North	August 2015	Clinical	Human	Faeces	48.15	5012266	89	469	250	MiSeq	ERS3404929	PT_SE0018
PT19	Se_266-15	Setúbal	Lisbon Metropolitan area	August 2015	Clinical	Human	Faeces	63.42	4886200	80	469	250	MiSeq	ERS3404930	PT_SE0019
PT20	Se_270-15	Azores	Azores	August 2015	Clinical	Human	Faeces	42.3	4945266	97	469	250	MiSeq	ERS3404931	PT_SE0020
PT21	Se_305-15	Coimbra	Centre	August 2015	Environmental	NA	Water	125.42	5066848	125	469	250	MiSeq	ERS3404932	PT_SE0021
PT22	Se_479-15	Oporto	North	November 2015	Clinical	Human	Faeces	58.32	4909000	109	469	250	MiSeq	ERS3404933	PT_SE0022
PT23	Se_PIGA-2-16	Azores	Azores	November 2015	Food	Pig	Hamburger	80.69	4940239	84	469	250	MiSeq	ERS3404934	PT_SE0023
PT24	Se_PIGA-3-16	Azores	Azores	November 2015	Food	Pig	Raw Chorizo	48.04	4873930	94	469	250	MiSeq	ERS3404935	PT_SE0024
PT25	Se_PIGA-4-16	Azores	Azores	November 2015	Food	Pig	Raw meat	58.24	4940367	98	469	250	MiSeq	ERS3404936	PT_SE0025

Isolate ID	Lab. ID	Location	Region (NUTII)	Collection Date	Sample	Source	Product	Mean depth of coverage	Genome size (bp)	# of contigs	ST	Read size (bp)	Sequencing Aparatus (Illumina)	ENA Accession #	ENA ID
PT26	Se_PIGA-5-16	Azores	Azores	January 2016	Food	Pig	Chorizo	52.54	4938840	85	469	250	MiSeq	ERS3404937	PT_SE0026
PT27	Se_R33-16	Oporto	North	March 2016	Animal	Bovine	Carcass	60.86	4924248	96	469	250	MiSeq	ERS3404938	PT_SE0027
PT28	Se_R39-16	Azores	Azores	March 2016	Animal	Pig	Carcass	55.48	4938922	93	469	250	MiSeq	ERS3404939	PT_SE0028
PT29	Se_R70-16	Not Available	Not Available	May 2016	Food	Pig	Pork skewers	56.54	4879946	112	469	150	NextSeq	ERS3404940	PT_SE0029
PT30	Se_R71-16	Not Available	Not Available	May 2016	Food	Pig	Pork skewers	138.65	4880717	91	469	150	NextSeq	ERS3404941	PT_SE0030
PT31	Se_R72-16	Not Available	Not Available	May 2016	Food	Pig	Pork skewers	99.05	4889369	102	469	150	NextSeq	ERS3404942	PT_SE0031
PT32	Se_R73-16	Not Available	Not Available	May 2016	Food	Pig	Pork skewers	68.58	4876907	110	469	150	NextSeq	ERS3404943	PT_SE0032
PT33	Se_R74-16	Not Available	Not Available	May 2016	Food	Pig	Pork skewers	221.84	4881425	87	469	150	NextSeq	ERS3404944	PT_SE0033
PT34	Se_R92-16	Not Available	Not Available	May 2016	Environmental	NA	Mud	75.75	4917595	110	469	150	NextSeq	ERS3404945	PT_SE0034
PT35	Se_R110-16	Azores	Azores	June 2016	Animal	Pig	Carcass	56.04	4919561	123	469	150	NextSeq	ERS3404946	PT_SE0035
PT36	Se_R134-16	Setúbal	Lisbon Metropolitan area	August 2016	Animal	Bivalve	Bivalve mollusc	186.4	4910852	85	469	150	NextSeq	ERS3404947	PT_SE0036
PT37	Se_R151-16	Azores	Azores	August 2016	Animal	Pig	Carcass	64.27	4953004	147	469	150	NextSeq	ERS3404948	PT_SE0037
PT38	Se_S176	Azores	Azores	August 2016	Clinical	Human	Faeces	71.69	4937185	83	469	250	MiSeq	ERS3404949	PT_SE0038
PT39	Se_S208	Azores	Azores	August 2016	Clinical	Human	Urine	57.31	4941276	117	469	250	MiSeq	ERS3404950	PT_SE0039
PT40	Se_PVRAM-46-16	Viseu	Centre	October 2016	Animal	Chicken	Carcass	68.29	4854659	105	469	150	NextSeq	ERS3404951	PT_SE0040
PT41	Se_S308	Azores	Azores	October 2016	Clinical	Human	Exudate	86.48	4979091	118	469	250	MiSeq	ERS3404952	PT_SE0041
PT42	Se_R196-16	Not Available	Not Available	November 2016	Food	Pig	Raw sausage	291.87	4877328	89	469	150	NextSeq	ERS3404953	PT_SE0042
PT43	Se_R199-16	Azores	Azores	November 2016	Animal	Bovine	Carcass	39.49	4947413	203	469	150	NextSeq	ERS3404954	PT_SE0043
PT44	Se_R200-16	Spain	Spain	November 2016	Food	Chicken	Meat	69.09	4888699	110	469	150	NextSeq	ERS3404955	PT_SE0044
PT45	Se_R217-16	Lisbon	Lisbon Metropolitan area	December 2016	Food	Chicken	Raw meat	63.82	4887212	120	469	150	NextSeq	ERS3404956	PT_SE0045
PT46	Se_PIGA-7-16	Azores	Azores	January 2016	Food	Pig	Chorizo	77.72	4921583	108	469	150	NextSeq	ERS3404957	PT_SE0046
PT47	Se_R65-17	Not Available	Not Available	March 2017	Food	Turkey	Turkey skewers	83.93	4930828	79	469	150	NextSeq	ERS3404958	PT_SE0047
PT48	Se_S398	Azores	Azores	March 2017	Clinical	Human	Bronchial aspirate	62.66	4920188	116	469	250	MiSeq	ERS3404959	PT_SE0048
PT49	Se_S402	Azores	Azores	March 2017	Clinical	Human	Urine	46.6	4917517	100	469	250	MiSeq	ERS3404960	PT_SE0049
PT50	Se_S420	Lisbon	Lisbon Metropolitan area	March 2017	Clinical	Human	Urine	39.09	4903866	123	469	250	MiSeq	ERS3404961	PT_SE0050
PT51	Se_R74-17	Azores	Azores	April 2017	Animal	Bovine	Carcass	77.31	4963946	117	469	150	NextSeq	ERS3404962	PT_SE0051
PT52	Se_S473	Azores	Azores	May 2017	Clinical	Human	Faeces	46.96	4939749	93	469	250	MiSeq	ERS3404963	PT_SE0052
PT53	Se_PVRAM-2-17	Leiria	Centre	June 2017	Animal	Bovine	Carcass	59.02	4925286	115	469	150	NextSeq	ERS3404964	PT_SE0053
PT54	Se_PVRAM-3-17	Azores	Azores	July 2017	Animal	Pig	Carcass	46.47	4936820	229	469	150	NextSeq	ERS3404965	PT_SE0054
PT55	Se_PVRAM-10-17	Azores	Azores	August 2017	Animal	Pig	Carcass	127.85	4961638	130	469	150	NextSeq	ERS3404966	PT_SE0055
PT56	Se_PVRAM-17-17	Leiria	Centre	August 2017	Animal	Pig	Carcass	49.42	4913428	123	469	150	NextSeq	ERS3404967	PT_SE0056
PT57	Se_PIGA-8-17	Not Available	Centre	September 2017	Food	Pig	Chorizo	49.41	4909916	176	469	150	NextSeq	ERS3404968	PT_SE0057
PT58	Se_PVRAM-19-17	Leiria	Centre	September 2017	Animal	Pig	Carcass	63.03	4917601	123	469	150	NextSeq	ERS3404969	PT_SE0058
PT59	Se_S674	Lisbon	Lisbon Metropolitan area	September 2017	Clinical	Human	Pus	44.34	4924957	182	469	150	NextSeq	ERS3404970	PT_SE0059
PT60	Se_R258-17	Not Available	Not Available	October 2017	Food	Bovine/pork	Gourmet Hamburguer	65.2	5090350	122	469	150	NextSeq	ERS3404971	PT_SE0060

Isolate ID	Resistance Profile	Penicillins			Tetracyclines			Chloramphenicol				Trimethoprim			Sulphonamides			Macrolides		Aminoglycosides						Fluoroquinolones						Carbapenems	Cephalosporins					Heavy metal tolerance/resistance genes				
		Amp	Amc	*bla*TEM-1B_1	Tet	Tgc	*tet(A)*	Chl	*catA2*	*cmlA1*	*floR*	Tmp	*dfrA12*	*dfrA1*	Smx	*sul1*	*sul3*	Azm	*mphA*	Gmn	*aac(3)ib*	*aadA2*	*aadA1*	*aac(6')-Iaa*	*aadA24*	Pef	Nal	*qnrS1*	*qnrB14*	*qnrB19*	*parC* (T57S)	Mem	Fox	Ctx	Caz	Cro	Fep	*merACDE*	*arsABCR*	*pcoABCDESR*	*silABCEFRS*	*tehAB*
PT01	MDR	R	S	+	R	S	+	S	-	-	-	R	+	-	R	+	-	S	-	S	-	+	+	+	-	S	S	-	-	-	+	S	S	S	S	S	S	-	+	+	+	+
PT02	MDR	S	S	-	R	S	+	R	+	-	-	R	+	-	R	+	-	S	-	S	-	+	-	+	-	S	S	-	-	-	+	S	S	S	S	S	S	-	+	+	+	+
PT03	MDR	R	R	+	R	S	+	R	-	+	-	R	-	+	R	+	+	S	-	I	+	+	+	+	-	S	S	-	-	-	+	S	S	S	S	S	S	+	+	+	+	+
PT04	MDR	R	S	+	S	S	-	R	-	+	-	R	-	+	R	+	+	S	-	S	-	+	+	+	-	S	S	-	-	-	+	S	S	S	S	S	S	+	+	+	+	+
PT05	MDR	S	S	+	R	S	+	S	-	-	-	R	+	-	R	+	-	R	+	S	-	-	-	+	+	S	S	-	-	-	+	S	S	S	S	S	S	-	+	+	+	+
PT06	MDR	R	S	+	R	S	+	R	+	-	-	R	+	-	R	+	-	R	+	S	-	+	+	+	-	S	S	-	-	-	+	S	S	S	S	S	S	-	+	+	+	+
PT07	MDR	R	S	+	R	S	+	R	-	+	-	R	-	+	R	+	+	S	-	S	-	+	+	+	-	S	S	-	-	-	+	S	S	S	S	S	S	+	+	+	-	+
PT08	MDR	R	S	+	R	S	+	S	-	-	-	R	+	-	R	+	-	R	+	S	-	+	+	+	-	S	S	-	-	-	+	S	S	S	S	S	S	-	+	+	+	+
PT09	MDR	R	S	+	R	S	+	R	-	+	-	R	-	+	R	+	+	S	-	S	-	+	+	+	-	S	S	-	-	-	+	S	S	S	S	S	S	+	+	+	+	+
PT10	non-MDR	R	S	+	S	S	-	S	-	-	-	S	-	-	S	-	-	S	-	S	-	-	-	+	-	S	S	-	-	-	+	S	S	S	S	S	S	-	+	+	+	+
PT11	non-MDR	S	S	-	S	S	-	S	-	-	-	S	-	-	S	-	-	S	-	S	-	-	-	+	-	S	S	-	-	-	+	S	S	S	S	S	S	-	+	+	+	+
PT12	non-MDR	S	S	+	S	S	+	S	-	-	-	S	+	-	R	+	-	S	-	S	-	+	+	+	-	S	S	-	-	-	+	S	S	S	S	S	S	-	+	+	+	+
PT13	MDR	R	S	+	R	S	+	R	+	-	-	R	+	-	R	+	-	R	+	S	-	-	-	+	+	S	S	-	-	-	+	S	S	S	S	S	S	-	+	+	+	+
PT14	MDR	R	S	+	R	S	+	R	+	-	-	R	+	-	R	+	-	R	+	S	-	+	+	+	-	S	S	-	-	-	+	S	S	S	S	S	S	-	+	+	+	+
PT15	MDR	R	S	+	R	S	+	R	+	-	-	R	+	-	R	+	-	R	+	S	-	+	+	+	-	S	S	-	-	-	+	S	S	S	S	S	S	-	+	+	+	+
PT16	MDR	S	S	-	R	S	+	S	-	-	-	R	+	-	R	+	-	R	+	S	-	+	-	+	-	S	S	-	-	-	+	S	S	S	S	S	S	-	+	+	+	+
PT17	non-MDR	S	S	-	R	S	+	S	-	-	-	S	-	-	S	-	-	S	-	S	-	-	-	+	-	S	S	-	-	-	+	S	S	S	S	S	S	-	+	+	+	+
PT18	MDR	R	S	+	R	S	+	S	-	-	-	R	+	-	R	+	-	S	-	S	-	+	+	+	-	S	S	-	-	-	+	S	S	S	S	S	S	-	+	+	+	+
PT19	non-MDR	S	S	-	R	S	+	S	-	-	-	S	-	-	S	-	-	S	-	S	-	-	-	+	-	S	S	-	-	-	+	S	S	S	S	S	S	-	+	+	+	+
PT20	MDR	R	S	+	R	S	+	R	+	-	-	R	+	-	R	+	-	R	+	S	-	-	-	+	+	S	S	-	-	-	+	S	S	S	S	S	S	-	+	+	+	+
PT21	MDR	R	S	+	R	S	+	S	-	-	-	R	+	-	R	+	-	S	-	S	-	+	+	+	-	S	S	-	-	-	+	S	S	S	S	S	S	-	+	+	+	+
PT22	MDR	R	S	+	R	S	+	R	-	+	-	R	-	+	R	+	+	S	-	S	-	+	+	+	-	S	S	-	-	-	+	S	S	S	S	S	S	+	+	+	+	+
PT23	MDR	R	R	+	R	S	+	R	+	-	-	R	+	-	R	+	+	R	+	S	-	-	-	+	+	S	S	-	-	-	+	S	S	S	S	S	S	-	+	+	+	+
PT24	MDR	S	S	-	R	S	+	R	-	+	-	S	-	-	R	-	+	S	-	S	-	+	+	+	-	S	S	-	-	-	+	S	S	S	S	S	S	-	+	+	+	+
PT25	MDR	R	R	+	R	S	+	S	-	-	-	R	+	-	R	+	-	R	+	S	-	-	-	+	+	S	S	-	-	-	+	S	S	S	S	S	S	-	+	+	+	+
PT26	MDR	S	S	-	R	S	+	S	-	-	-	R	+	-	R	+	-	R	+	S	-	+	+	+	-	S	S	-	-	-	+	S	S	S	S	S	S	-	+	+	+	+
PT27	non-MDR	R	S	-	S	S	-	S	-	-	-	S	-	-	S	-	-	S	-	S	-	-	-	+	-	S	S	-	-	-	+	S	S	S	S	S	S	-	+	+	+	+
PT28	MDR	S	S	-	R	S	+	S	-	-	-	R	+	-	R	+	-	R	+	S	-	+	+	+	-	S	S	-	-	-	+	S	S	S	S	S	S	-	+	+	+	+
PT29	MDR	R	S	+	R	S	+	R	-	+	-	R	-	+	R	+	+	S	-	S	-	+	+	+	-	S	S	-	-	-	+	S	S	S	S	S	S	+	+	+	-	+
PT30	MDR	R	S	+	R	S	+	R	-	+	-	R	-	+	R	+	+	S	-	S	-	+	+	+	-	S	S	-	-	-	+	S	S	S	S	S	S	+	+	+	-	+
PT31	MDR	R	S	+	R	S	+	R	-	+	-	R	-	+	R	+	+	S	-	S	-	+	+	+	-	S	S	-	-	-	+	S	S	S	S	S	S	+	+	+	-	+

Isolate ID	Resistance Profile	Penicillins			Tetracyclines			Chloramphenicol				Trimethoprim			Sulphonamides			Macrolides		Aminoglycosides						Fluoroquinolones						Carbapenems	Cephalosporins					Heavy metal tolerance/resistance genes				
		Amp	Amc	*bla*TEM-1B_1	Tet	Tgc	*tet(A)*	Chl	*catA2*	*cmlA1*	*floR*	Tmp	*dfrA12*	*dfrA1*	Smx	*sul1*	*sul3*	Azm	*mphA*	Gmn	*aac(3)ib*	*aadA2*	*aadA1*	*aac(6')-Iaa*	*aadA24*	Pef	Nal	*qnrS1*	*qnrB14*	*qnrB19*	*parC* (T57S)	Mem	Fox	Ctx	Caz	Cro	Fep	*merACDE*	*arsABCR*	*pcoABCDESR*	*silABCEFRS*	*tehAB*
PT32	MDR	R	S	+	R	S	+	R	-	+	-	R	-	+	R	+	+	S	-	S	-	+	+	+	-	S	S	-	-	-	+	S	S	S	S	S	S	+	+	+	-	+
PT33	MDR	R	S	+	R	S	+	R	-	+	-	R	-	+	R	+	+	S	-	S	-	+	+	+	-	S	S	-	-	-	+	S	S	S	S	S	S	+	+	+	-	+
PT34	MDR	R	S	+	R	S	+	R	+	-	-	R	+	-	R	+	-	R	+	S	-	+	+	+	-	S	S	-	-	-	+	S	S	S	S	S	S	-	+	+	+	+
PT35	MDR	R	S	+	R	S	+	R	-	+	-	R	-	+	R	+	+	S	-	S	-	+	+	+	-	S	S	-	-	-	+	S	S	S	S	S	S	+	+	+	-	+
PT36	MDR	R	S	+	R	S	+	S	-	-	-	R	+	-	R	+	-	S	-	S	-	+	+	+	-	S	S	-	-	-	+	S	S	S	S	S	S	-	+	+	+	+
PT37	MDR	R	S	+	R	S	+	S	-	-	-	R	+	-	R	+	-	R	+	S	-	-	-	+	+	S	S	-	-	-	+	S	S	S	S	S	S	-	+	+	+	+
PT38	MDR	R	R	+	R	S	+	S	-	-	-	R	+	-	R	+	-	R	+	S	-	-	-	+	+	S	S	-	-	-	+	S	S	S	S	S	S	-	+	+	+	+
PT39	MDR	R	R	+	R	S	+	S	+	-	-	R	+	-	R	+	-	R	+	S	-	+	+	+	-	S	S	-	-	-	+	S	S	S	S	S	S	-	+	+	+	+
PT40	non-MDR	S	S	-	R	S	+	S	-	-	-	S	-	-	S	-	-	S	-	S	-	-	-	+	-	S	S	-	-	-	+	S	S	S	S	S	S	-	+	+	+	+
PT41	MDR	R	S	+	R	S	+	R	+	-	-	R	+	-	R	+	-	R	+	S	-	-	-	+	+	S	S	-	-	-	+	S	S	S	S	S	S	-	+	+	+	+
PT42	MDR	S	S	-	R	S	+	R	+	-	-	R	+	-	R	+	-	R	+	S	-	+	-	+	-	S	S	-	-	-	+	S	S	S	S	S	S	-	+	+	+	+
PT43	MDR	R	S	+	R	S	+	R	+	-	-	R	+	-	R	+	-	R	+	S	-	-	-	+	+	S	S	-	-	-	+	S	S	S	S	S	S	-	+	+	+	+
PT44	MDR	S	S	-	S	S	-	S	-	-	-	R	+	-	R	+	-	R	+	S	-	+	-	+	-	S	R	-	+	+	+	S	S	S	S	S	S	-	+	+	+	+
PT45	MDR	S	S	-	S	S	-	S	-	-	-	R	+	-	R	+	-	R	+	S	-	+	-	+	-	S	S	-	-	+	+	S	S	S	S	S	S	-	+	+	+	+
PT46	MDR	S	S	-	R	S	+	S	-	-	-	R	+	-	R	+	-	R	+	S	-	+	+	+	-	S	S	-	-	-	+	S	S	S	S	S	S	-	+	+	+	+
PT47	MDR	R	R	+	R	S	+	R	+	-	-	R	+	-	R	+	-	R	+	S	-	+	+	+	-	S	S	-	-	-	+	S	S	S	S	S	S	-	+	+	+	+
PT48	non-MDR	R	R	+	S	S	-	R	+	-	-	S	-	-	S	-	-	S	-	S	-	-	+	+	-	S	S	-	-	-	+	S	S	S	S	S	S	-	+	+	+	+
PT49	non-MDR	R	R	+	S	S	-	R	+	-	-	S	-	-	S	-	-	S	-	S	-	-	+	+	-	S	S	-	-	-	+	S	S	S	S	S	S	-	+	+	+	+
PT50	MDR	R	S	+	S	S	-	R	-	+	-	R	-	+	R	+	+	S	-	S	-	+	+	+	-	S	S	-	-	-	+	S	S	S	S	S	S	+	+	+	+	+
PT51	MDR	R	S	+	R	S	+	S	-	-	-	R	+	-	R	+	-	R	+	S	-	-	-	+	+	S	S	-	-	-	+	S	S	S	S	S	S	-	+	+	+	+
PT52	MDR	R	R	+	R	S	+	S	-	-	-	R	+	-	R	+	-	R	+	S	-	-	-	+	+	S	S	-	-	-	+	S	S	S	S	S	S	-	+	+	+	+
PT53	MDR	R	S	+	S	S	-	R	-	+	-	R	-	+	R	+	+	S	-	S	-	+	+	+	-	S	S	-	-	-	+	S	S	S	S	S	S	+	+	+	+	+
PT54	MDR	S	S	-	R	S	+	S	-	-	-	R	+	-	R	+	-	R	+	S	-	-	+	+	-	S	S	-	-	-	+	S	S	S	S	S	S	-	+	+	+	+
PT55	MDR	R	S	+	R	S	+	S	-	-	-	R	+	-	R	+	-	R	+	S	-	-	-	+	+	S	S	-	-	-	+	S	S	S	S	S	S	-	+	+	+	+
PT56	MDR	R	S	+	R	S	+	R	+	-	-	R	+	-	R	+	-	R	+	S	-	+	+	+	-	S	S	-	-	-	+	S	S	S	S	S	S	-	+	+	+	+
PT57	MDR	R	S	+	R	S	+	R	+	-	-	R	+	-	R	+	-	R	+	S	-	-	+	+	-	S	S	-	-	-	+	S	S	S	S	S	S	-	+	+	+	+
PT58	MDR	R	S	+	R	S	+	R	+	-	-	R	+	-	R	+	-	R	+	S	-	+	+	+	-	S	S	-	-	-	+	S	S	S	S	S	S	-	+	+	+	+
PT59	non-MDR	R	R	+	R	S	+	S	-	-	-	S	-	-	S	-	-	S	-	S	-	-	+	+	-	S	S	-	-	-	+	S	S	S	S	S	S	-	+	+	+	+
PT60	MDR	R	R	+	R	S	+	R	-	-	+	S	-	-	S	-	-	S	-	S	-	-	-	+	-	R	R	+	-	-	+	S	S	S	S	S	S	-	+	+	+	+

Supplementary Table 4-2 - List of publicly available *Salmonella enterica* serovar Rissen genomes used in the present study. All read datasets were downloaded from the European Nucleotide Archive.

Strain ID	Sample source	Year	Location	ENA Accession #
BCW_2764	Human	2002	Denmark	SAMN02368641
CRJJGF_00114	Food	2004	United States	SAMN02908621
BCW_2766	Unknown	2004	Thailand	SAMN02368643
BCW_2765	Livestock	2004	Denmark	SAMN02368642
FDA00001049	Environment	2006	Turkey	SAMN02844434
FDA00001046	Environment	2006	Vietnam	SAMN02844431
FDA00004496	Animal feed	2008	United States	SAMN02918550
IEH_NGS_SAL_00172	Environment	2009	United States	SAMN02265259
FDA00002917	Environment	2009	United States	SAMN02846301
FDA00002912	Environment	2009	United States	SAMN02846296
FDA00002855	Food	2009	United States	SAMN02846239
FDA00002853	Environment	2009	United States	SAMN02846237
FDA00002914	Environment	2009	United States	SAMN02846298
FDA00002915	Environment	2009	United States	SAMN02846299
FDA00002913	Environment	2009	United States	SAMN02846297
FDA00002854	Food	2009	United States	SAMN02846238
FDA00002916	Environment	2009	United States	SAMN02846300
FDA00002850	Environment	2009	United States	SAMN02846234
FDA00002849	Environment	2009	United States	SAMN02846233
FDA00002481	Environment	2009	United States	SAMN02845865
FDA00002848	Environment	2009	United States	SAMN02846232
FDA00002852	Environment	2009	United States	SAMN02846236
FDA00002851	Environment	2009	United States	SAMN02846235
FDA00002939	Food	2009	United States	SAMN02846323
FDA00002788	Food	2009	United States	SAMN02846172
FDA00002938	Food	2009	United States	SAMN02846322
CFSAN031935	Environment	2009	United States	SAMN03577468
CFSAN031930	Environment	2009	United States	SAMN03577463
CFSAN031937	Environment	2009	United States	SAMN03577470
32309.31	Environment	2009	United States	SAMN10253329
FDA00003115	Environment	2010	United States	SAMN02846535
FDA00003545	Aquatic	2010	Vietnam	SAMN02846965
FDA00003544	Aquatic	2010	Vietnam	SAMN02846964
FNW19G96	Aquatic	2010	Vietnam	SAMN02344903
FDA00004662	Food	2010	United States	SAMN02918711
FDA00001453	Environment	2010	Dominican Republic	SAMN02844838
NC_S803	Livestock	2010	United States	SAMN07469595
NC_S812	Livestock	2010	United States	SAMN07469592
NC_S810	Livestock	2010	United States	SAMN07469594
NC_S811	Livestock	2010	United States	SAMN07469593
NC_S801	Livestock	2010	United States	SAMN07469597
NC_S802	Livestock	2010	United States	SAMN07469596
FAR0094	Food	2011	China	SAMN02345345
FSE0050	Food	2011	Vietnam	SAMN02345332
FSW0031	Environment	2011	Philippines	SAMN02345262
CFSAN030098	Food	2012	United States	SAMN03464584

H124020495	Food	2012	United Kingdom	SAMN03168707
H125180332	Food	2012	United Kingdom	SAMN03168603
FAR0091	Food	2012	Unknown	SAMN02345586
FAR0099	Environment	2012	Mexico	SAMN02345539
PNUSAS002751	Unknown	2012	United States	SAMN06198443
CFSAN031334	Livestock	2013	Thailand	SAMN03576868
CFSAN031330	Livestock	2013	Thailand	SAMN03576864
CFSAN031335	Livestock	2013	Thailand	SAMN03576869
CFSAN031332	Livestock	2013	Thailand	SAMN03576866
CFSAN031331	Livestock	2013	Thailand	SAMN03576865
CFSAN031321	Livestock	2013	Thailand	SAMN03576855
CFSAN031320	Livestock	2013	Thailand	SAMN03576854
CFSAN031313	Livestock	2013	Thailand	SAMN03576847
CFSAN031314	Livestock	2013	Thailand	SAMN03576848
CFSAN031318	Livestock	2013	Thailand	SAMN03576852
CFSAN031312	Livestock	2013	Thailand	SAMN03576846
CFSAN031308	Livestock	2013	Thailand	SAMN03576842
CFSAN031315	Livestock	2013	Thailand	SAMN03576849
CFSAN031324	Livestock	2013	Thailand	SAMN03576858
CFSAN031309	Livestock	2013	Thailand	SAMN03576843
CFSAN031327	Livestock	2013	Thailand	SAMN03576861
CFSAN031325	Livestock	2013	Thailand	SAMN03576859
CFSAN031322	Livestock	2013	Thailand	SAMN03576856
CFSAN031316	Livestock	2013	Thailand	SAMN03576850
CFSAN031323	Livestock	2013	Thailand	SAMN03576857
CFSAN031311	Livestock	2013	Thailand	SAMN03576845
CFSAN031310	Livestock	2013	Thailand	SAMN03576844
CFSAN031319	Livestock	2013	Thailand	SAMN03576853
CFSAN031307	Livestock	2013	Thailand	SAMN03576841
CFSAN031306	Livestock	2013	Thailand	SAMN03576840
FNE0189	Food	2013	Thailand	SAMN02698353
CFSAN045275	Livestock	2013	Thailand	SAMN04431411
CFSAN045273	Livestock	2013	Thailand	SAMN04431409
CFSAN045268	Livestock	2013	Thailand	SAMN04431404
CFSAN045265	Livestock	2013	Thailand	SAMN04431401
CFSAN045263	Livestock	2013	Thailand	SAMN04431399
CFSAN045340	Livestock	2013	Thailand	SAMN04431476
CFSAN045327	Livestock	2013	Thailand	SAMN04431463
CFSAN045326	Livestock	2013	Thailand	SAMN04431462
CFSAN045324	Livestock	2013	Thailand	SAMN04431460
OH-1302304	Unknown	2013	United States	SAMN05721584
CVM N44364F	Livestock	2013	United States	SAMN06287508
CFSAN084068	Environment	2013	United States	SAMN09756362
S280	Unknown	2014	Portugal	SAMEA3476860
73107	Human	2014	United Kingdom	SAMN03479841
5965	Human	2014	United Kingdom	SAMN03479663
40781	Food	2014	United Kingdom	SAMN03478634
68669	ND	2014	United Kingdom	SAMN03477972
31626	Human	2014	United Kingdom	SAMN03476772
5839	Human	2014	United Kingdom	SAMN03476472

13420	Human	2014	United Kingdom	SAMN03473985
CFSAN045347	Livestock	2014	Thailand	SAMN04431483
CFSAN045343	Livestock	2014	Thailand	SAMN04431479
CVM N57959F	Unknown	2014	United States	SAMN04577502
CVM N57219F	Livestock	2014	United States	SAMN04577270
NC_NCF3D0-L5	Environment	2014	United States	SAMN06322129
FDA00008934	Food	2015	Mexico	SAMN03495909
FDA00008935	Food	2015	Mexico	SAMN03495910
91210	Food	2015	United Kingdom	SAMN03480405
91281	Human	2015	United Kingdom	SAMN03480200
91246	Human	2015	United Kingdom	SAMN03480194
91209	Food	2015	United Kingdom	SAMN03479124
91211	Food	2015	United Kingdom	SAMN03478922
91208	Food	2015	United Kingdom	SAMN03478809
91212	Food	2015	United Kingdom	SAMN03476328
FDA00009504	Aquatic animal	2015	Vietnam	SAMN04148250
FSIS1502596	Livestock	2015	United States	SAMN04331735
FSIS1503283	Livestock	2015	United States	SAMN04331725
ADRDL-15-8159	Poultry	2015	United States	SAMN04240670
FSIS1605466	Livestock	2015	United States	SAMN04530405
CVM N57971	Poultry	2015	United States	SAMN04576648
182090	Human	2015	United Kingdom	SAMN04600580
185971	Human	2015	United Kingdom	SAMN04600398
129524	Human	2015	United Kingdom	SAMN04600236
CVM N57978	Poultry	2015	United States	SAMN05771731
MS150107	Human	2015	Ireland	SAMEA81466918
NC_NCF6D0-A13	Environment	2015	United States	SAMN06322119
NC_NCF6D0-A18	Environment	2015	United States	SAMN06322118
NC_NCF6D0-A4	Environment	2015	United States	SAMN06322126
NC_NCF6D0-L9	Environment	2015	United States	SAMN06322127
NC_NCF6D0-L8	Environment	2015	United States	SAMN06322128
NC_NCF6D0-A20	Environment	2015	United States	SAMN06322116
151916	Livestock	2015	Spain	SAMEA104142963
PAT-15-27861SA	Livestock	2015	Portugal	SAMEA104142939
PAT-15-19702SA	Livestock	2015	Portugal	SAMEA104142937
15Q003557	Livestock	2015	France	SAMEA104142871
ADRDL-15-2378	Livestock	2015	United States	SAMN07351348
CFSAN069222	Environment	2015	United States	SAMN07714159
2187_Se_151916	Livestock	2015	Spain	SAMEA104354224
2176_Se_PAT_15_27861SA	Livestock	2015	Portugal	SAMEA104354222
2174_Se_PAT_15_19702SA	Livestock	2015	Portugal	SAMEA104354221
34927	Livestock	2015	United States	SAMN07420494
170460	Human	2015	United Kingdom	SAMN09423139
FDA00009757	Food	2015	Jamaica	SAMN04385807
FDA00009756	Food	2015	Jamaica	SAMN04385806
191884	Human	2015	United Kingdom	SAMN09484447
195784	Human	2015	United Kingdom	SAMN10140147
FDA00010268	Food	2016	Unresolved	SAMN05232953
FDA00010267	Environment	2016	Unresolved	SAMN05232952
PNUSAS004417	Unknown	2016	United States	SAMN05877103

ADRDL-16-9998	Livestock	2016	United States	SAMN05781501
FSIS1608386	Livestock	2016	United States	SAMN06048830
CFSAN058818	Environment	2016	United States	SAMN06175263
CFSAN058816	Environment	2016	United States	SAMN06175307
FSIS1609588	Livestock	2016	United States	SAMN06229944
FDA00011195	Animal feed	2016	United States	SAMN06214692
FDA00011193	Animal feed	2016	United States	SAMN06214694
271097	Human	2016	United Kingdom	SAMN06278620
CFSAN059804	Environment	2016	United States	SAMN06270186
ADRDL-902	Livestock	2016	United States	SAMN06330644
OH-16-23867-16	Poultry	2016	United States	SAMN07138205
OH-16-26883-1	Poultry	2016	United States	SAMN07184770
CFSAN071463	Environment	2016	United States	SAMN08017022
CVM N16S193	Poultry	2016	United States	SAMN08114066
CVM N16S274	Poultry	2016	United States	SAMN08114143
CFSAN075243	Feed	2016	United States	SAMN08395358
CFSAN071961	Environment	2016	United States	SAMN08057881
CFSAN080464	Feed	2016	United States	SAMN09071683
227073	Human	2016	United Kingdom	SAMN09403226
216348	Human	2016	United Kingdom	SAMN09403232
217006	Human	2016	United Kingdom	SAMN09610602
212747	Food	2016	United Kingdom	SAMN09634347
FSIS1710520	Poultry	2017	United States	SAMN06459939
FSIS1710555	Livestock	2017	United States	SAMN06459285
FSIS1700028	Livestock	2017	United States	SAMN06701838
FDA00011518	Food	2017	Mexico	SAMN06689639
FSIS1700448	Livestock	2017	United States	SAMN06882344
FSIS1700956	Livestock	2017	United States	SAMN06899446
FSIS1701130	Livestock	2017	United States	SAMN07141738
367310	Human	2017	United Kingdom	SAMN07155624
360374	Human	2017	United Kingdom	SAMN07155217
367265	Human	2017	United Kingdom	SAMN07180765
333422	Food	2017	United Kingdom	SAMN07180676
333421	Food	2017	United Kingdom	SAMN07180660
362108	Human	2017	United Kingdom	SAMN07180311
363492	Human	2017	United Kingdom	SAMN07180216
FSIS1701469	Livestock	2017	United States	SAMN07237724
FSIS1702121	Livestock	2017	United States	SAMN07260890
FSIS1703115	Livestock	2017	United States	SAMN07424747
FSIS1703367	Livestock	2017	United States	SAMN07501501
FSIS1703368	Livestock	2017	United States	SAMN07501502
FDA00012093	Animal feed	2017	Argentina	SAMN07510071
PNUSAS021809	Unknown	2017	United States	SAMN07561485
FSIS21720373	Livestock	2017	United States	SAMN07819022
FSIS11704588	Livestock	2017	United States	SAMN07835489
FDA00012295	Aquatic animal	2017	Vietnam	SAMN07981379
FSIS11705536	Livestock	2017	United States	SAMN08114352
CFSAN071971	Environment	2017	United States	SAMN08057833
CFSAN071970	Environment	2017	United States	SAMN08057822
FSIS1701253	Livestock	2017	United States	SAMN09098834

412138	Human	2017	United Kingdom	SAMN09298461
CFSAN081790	Feed	2017	United States	SAMN09262024
456191	Human	2017	United Kingdom	SAMN09388861
371818	Human	2017	United Kingdom	SAMN09423133
388672	Human	2017	United Kingdom	SAMN09431486
445933	Human	2017	United Kingdom	SAMN09445526
378690	Food	2017	United Kingdom	SAMN09643791
CVM N17S1456	Livestock	2017	United States	SAMN09771059
413029	Human	2017	United Kingdom	SAMN09433805
PNUSAS032562	Environment	2018	United States	SAMN08383811
FSIS31800061	Livestock	2018	United States	SAMN08432325
ADRDL-1843	Livestock	2018	United States	SAMN08579906
FSIS11808313	Livestock	2018	United States	SAMN08767328
FSIS21821362	Poultry	2018	United States	SAMN08848643
FSIS11808806	Livestock	2018	United States	SAMN08886742
FSIS11809218	Livestock	2018	United States	SAMN08967083
529250	Human	2018	United Kingdom	SAMN09076426
FSIS21821653	Poultry	2018	United States	SAMN09225336
520599	Human	2018	United Kingdom	SAMN09388937
FSIS11810650	Livestock	2018	United States	SAMN09444133
533995	Human	2018	United Kingdom	SAMN09522091
sam	Human	2018	Ireland	SAMEA4730812
550127	Human	2018	United Kingdom	SAMN09607167
503417	Human	2018	United Kingdom	SAMN09634155
526142	Human	2018	United Kingdom	SAMN09651916
FSIS11811376	Livestock	2018	United States	SAMN09633410
570850	Human	2018	United Kingdom	SAMN09683619
FSIS11810870	Poultry	2018	United States	SAMN09533147
FSIS31800980	Poultry	2018	United States	SAMN09994491
FSIS11813906	Livestock	2018	United States	SAMN10064108
FSIS11808310	Livestock	2018	United States	SAMN08767326
06-1346	Unknown	Unknown	Unknown	SAMEA1484347
PNUSAS001935	Unknown	Unknown	United States	SAMN04893774
VNB1121-sc-2280660	Unknown	Unknown	Unknown	SAMEA3447772
2452-sc-2280574	Unknown	Unknown	Unknown	SAMEA3447686
74_H_097-sc-2280751	Unknown	Unknown	Unknown	SAMEA3447861
72_H_175-sc-2280732	Unknown	Unknown	Unknown	SAMEA3447843
71_H_195-sc-2280689	Unknown	Unknown	Unknown	SAMEA3447801
71_H_131-sc-2280683	Unknown	Unknown	Unknown	SAMEA3447795
VNSC2361-sc-2280610	Unknown	Unknown	Unknown	SAMEA3447722
VNB1504-sc-2280704	Unknown	Unknown	Unknown	SAMEA3447816
01-0479	Unknown	Unknown	Unknown	SAMN03264955
00-0084	Unknown	Unknown	Unknown	SAMN03264954
2012K-0157	Unknown	Unknown	Unknown	SAMN03264953
BCW_2864	Unknown	Unknown	Unknown	SAMN02368730
PNUSAS002711	Unknown	Unknown	United States	SAMN05603672
PNUSAS004670	Unknown	Unknown	United States	SAMN05919885
S02948-14	Unknown	Unknown	Unknown	SAMEA4064117
S02707-14	Unknown	Unknown	Unknown	SAMEA4064114
PNUSAS006373	Unknown	Unknown	United States	SAMN06213966

SAMEA4412641	Unknown	Unknown	Unknown	SAMEA4412641
PNUSAS007242	Unknown	Unknown	United States	SAMN06220355
PNUSAS007916	Unknown	Unknown	United States	SAMN06310818
PNUSAS009904	Unknown	Unknown	United States	SAMN06625047
PNUSAS014086	Unknown	Unknown	United States	SAMN07161029
PNUSAS023317	Unknown	Unknown	United States	SAMN07806985
PNUSAS025277	Unknown	Unknown	United States	SAMN07949738
PNUSAS027824	Unknown	Unknown	United States	SAMN08016445
PNUSAS030901	Unknown	Unknown	United States	SAMN08240237
PNUSAS032846	Unknown	Unknown	United States	SAMN08432357
PNUSAS029879	Unknown	Unknown	United States	SAMN08470840
PNUSAS035109	Unknown	Unknown	United States	SAMN08616364
PNUSAS036743	Unknown	Unknown	United States	SAMN08823502
PNUSAS037800	Unknown	Unknown	United States	SAMN08865617
PNUSAS037136	Unknown	Unknown	United States	SAMN08865508
PNUSAS037130	Unknown	Unknown	United States	SAMN08865514
PNUSAS037097	Unknown	Unknown	United States	SAMN08922760
PNUSAS042247	Unknown	Unknown	United States	SAMN09381922
PNUSAS042313	Unknown	Unknown	United States	SAMN09501102
PNUSAS043909	Unknown	Unknown	United States	SAMN09534158
PNUSAS044558	Unknown	Unknown	United States	SAMN09636247
PNUSAS045202	Unknown	Unknown	United States	SAMN09650791
14ARS_VSM0382	Unknown	Unknown	Unknown	SAMEA104162236
14ARS_VSM0381	Unknown	Unknown	Unknown	SAMEA104162235

Supplementary Table 5-1 – List of publicly available *Salmonella enterica* serovar Heidelberg and *Salmonella enterica* serovar Minnesota genomes used in the present study.

Id	Species	Serotype	Origin_country	Isolation
SRR6881714	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7064459	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7064627	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7064868	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7064945	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7064960	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7064964	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7065001	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7065005	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7071937	*S. enterica*	Minnesota	Brazil	Brazil
SRR7071939	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7071959	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7071960	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7072121	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7072128	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7072141	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7072143	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7072145	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7072176	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7072177	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7072178	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7072198	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7072200	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7072202	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7072208	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7072217	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7073432	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7079287	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7079314	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7130373	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7130375	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7130551	*S. enterica*	Minnesota	Brazil	Brazil
SRR7130561	*S. enterica*	Minnesota	Brazil	Brazil
SRR7130567	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7186260	*S. enterica*	Minnesota	Brazil	Brazil
SRR7186369	*S. enterica*	Minnesota	Brazil	Brazil
SRR7186454	*S. enterica*	Minnesota	Brazil	Brazil
SRR7188845	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7189240	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7189510	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7190735	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7192261	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7192445	*S. enterica*	Heidelberg	Brazil	Brazil

SRR7192446	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7219557	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7221031	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7221243	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7221348	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7221431	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7221474	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7221475	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7221476	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7221528	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7221529	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7221614	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7221645	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7221647	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7221657	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7229075	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7229077	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7230389	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7230673	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7230743	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7230883	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7232306	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7232379	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7232380	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7232613	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7232983	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7232984	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7232990	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7233008	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7233581	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7233615	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7233616	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7241938	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7241940	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7242751	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7242913	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7244087	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7244211	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7249331	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7249775	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7249857	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7250105	*S. enterica*	Heidelberg	Brazil	Brazil
SRR7250125	*S. enterica*	Heidelberg	Brazil	Brazil
SRR8663248	*S. enterica*	Heidelberg	Venezuela	Brazil
SRR7232219	*S. enterica*	Heidelberg	Brazil	Brazil
ERR3420975	*S. enterica*	Heidelberg	Brazil	Netherlands

ERR3420976	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3420977	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3420978	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3420979	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3420980	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3420981	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3420982	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3420983	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3420984	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3420985	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3420986	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3420987	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3420988	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3420989	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3420990	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3420991	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3420992	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3420993	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3420994	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3420995	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3420996	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3420997	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3420998	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3420999	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421000	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421001	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421002	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421003	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421004	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421005	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421006	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421007	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421008	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421009	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421010	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421011	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421012	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421013	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421014	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421017	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421018	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421019	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421020	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421021	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421022	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421023	*S. enterica*	Heidelberg	Brazil	Netherlands

ERR3421024	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421025	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421026	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421027	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421028	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421029	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421030	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421031	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421032	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421033	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421034	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421035	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421036	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421037	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421038	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421039	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421040	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421041	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421042	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421043	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421044	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421045	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421046	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421047	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421048	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421049	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421050	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421051	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421052	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421053	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421054	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421055	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421058	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421059	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421060	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421061	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421062	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421063	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421064	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421065	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421067	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421068	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421069	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421070	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421071	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421072	*S. enterica*	Heidelberg	Brazil	Netherlands

ERR3421073	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421074	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421075	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421076	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421077	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421078	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421079	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421080	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421081	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421082	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421083	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421084	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421085	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421086	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421087	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421088	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421089	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421090	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421091	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421092	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421093	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421094	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421095	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421096	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421097	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421098	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421099	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421100	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421101	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421102	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421103	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421104	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421105	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421106	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421107	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421108	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421109	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421110	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421111	*S. enterica*	Heidelberg	Brazil	Netherlands
ERR3421112	*S. enterica*	Heidelberg	Brazil	Netherlands

Supplementary table 5-2 - Phage presence in the 36 sequenced isolates of *Salmonella enterica* serovar Heidelberg and *Salmonella enterica* serovar Minnesota

Isolate ID	Region Length	Score	# Total Proteins	Most Common Phage (proteins)	GC %
PT1	33.3Kb	100	18	Fels 1 (5)	49.39%
	40Kb	150	49	Fels 2 (35)	52.92%
PT2	92.5Kb	150	112	phiV10 (33)	52.13%
	33.3Kb	100	18	Fels 1 (5)	49.39%
	9.8Kb	100	16	Gifsy 1 (1)	51.41%
	33.8Kb	120	45	SEN34 (15)	49.68%
	31.9Kb	150	46	SfV (38)	51.75%
PT3	9.8Kb	100	16	Gifsy 1 (1)	51.41%
PT4	40Kb	150	48	Fels 2 (36)	52.92%
	33.3Kb	100	17	Fels 1 (5)	49.39%
	9.8Kb	100	18	Gifsy 1 (1)	51.41%
PT5	44.1Kb	120	57	phiV10 (33)	51.99%
PT6	44.9Kb	150	62	Gifsy 2 (34)	51.03%
PT7	44.1Kb	120	57	phiV10 (33)	51.99%
	33.3Kb	100	18	Fels 1 (5)	49.39%
PT8	44.9Kb	150	62	Gifsy 2 (34)	51.03%
PT9	37.9Kb	150	49	Gifsy 2 (31)	51.55%
PT10	44.9Kb	150	62	Gifsy 2 (34)	51.03%
PT11	44.9Kb	150	62	Gifsy 2 (34)	51.03%
PT12	45Kb	140	56	phiV10 (32)	51.82%
	42.9Kb	150	57	SfV (39)	51.97%
	30.8Kb	150	37	Fels 2 (28)	53.38%
	9.8Kb	100	18	Gifsy 1 (1)	51.41%
PT13	44.9Kb	150	62	Gifsy 2 (34)	51.03%
PT14	34.8Kb	150	47	Gifsy 2 (30)	51.95%
	9.8Kb	100	18	Gifsy 1 (1)	51.41%
PT15	34Kb	150	45	Gifsy 2 (30)	52.26%
PT16	42.3Kb	132	52	SPN1S (46)	50.53%
	45.4Kb	150	57	Gifsy 2 (33)	51.25%
PT17	44.9Kb	150	60	Gifsy 2 (34)	51.03%
PT18	44.9Kb	150	60	Gifsy 2 (34)	51.03%
PT19	44.9Kb	150	60	Gifsy 2 (34)	51.03%
PT20	44.9Kb	150	60	Gifsy 2 (34)	51.03%
PT21	33.3Kb	100	18	Fels 1 (5)	49.39%
	44.3Kb	150	66	SfII (37)	51.31%
	44.8Kb	140	56	phiV10 (32)	51.89%
PT22	58.3Kb	132	52	SPN1S (46)	50.79%
	53Kb	150	56	Gifsy 2 (33)	51.13%

PT23	44.9Kb	150	60	Gifsy 2 (34)	51.03%
	52.1Kb	121	45	SPN1S (42)	52.67%
	9.8Kb	100	16	Gifsy 1 (1)	51.41%
PT24	33.3Kb	100	18	Fels 1 (5)	49.39%
	44.1Kb	120	57	phiV10 (33)	51.99%
	9.8Kb	100	16	Gifsy 1 (1)	51.41%
PT25	39Kb	150	52	Gifsy 2 (31)	51.32%
	51.3Kb	115	44	SPN1S (40)	52.94%
PT26	58.3Kb	132	52	SPN1S (46)	50.79%
	44.9Kb	150	60	Gifsy 2 (34)	51.03%
PT27	42.3Kb	132	52	SPN1S (46)	50.53%
	37.8Kb	100	58	UAB Phi20 (34)	47.52%
	44.9Kb	150	60	Gifsy 2 (34)	51.03%
PT28	42.2Kb	110	58	vB SemP Emek (15)	46.73%
	33.3Kb	100	18	Fels 1 (5)	49.39%
	44.8Kb	140	56	phiV10 (32)	51.89%
PT29	45.4Kb	150	57	Gifsy 2 (33)	51.25%
	9.8Kb	100	16	Gifsy 1 (1)	51.41%
PT30	44.9Kb	150	60	Gifsy 2 (34)	51.03%
	43.7Kb	92	57	phiV10 (33)	51.01%
	9.8Kb	100	16	Sfi19 (1)	51.41%
PT31	44.8Kb	140	56	phiV10 (32)	51.89%
	41.3Kb	120	62	ENT47670 (14)	52.30%
	35.6Kb	150	51	SfII (38)	51.71%
	39.9Kb	150	35	Fels 2 (28)	52.18%
	33.3Kb	100	17	Fels 1 (5)	49.39%
	9.8Kb	100	17	Gifsy 1 (1)	51.41%
PT32	44.8Kb	140	56	phiV10 (32)	51.89%
	9.8Kb	100	17	Gifsy 1 (1)	51.41%
	33.3Kb	100	18	Fels 1 (5)	49.39%
	36.4Kb	150	48	P88 (35)	52.87%
	24Kb	100	22	SSU5 (21)	53.63%
PT33	33.3Kb	100	17	Fels 1 (5)	49.39%
	39.9Kb	150	35	Fels 2 (27)	52.18%
	44.8Kb	140	56	phiV10 (32)	51.89%
	9.8Kb	100	17	Gifsy 1 (1)	51.41%
PT34	42.8Kb	150	55	Gifsy 2 (33)	51.54%
PT35	31Kb	150	38	Gifsy 2 (29)	52.44%
	9.8Kb	100	16	Sfi19 (1)	51.41%
PT36	31.5Kb	100	35	SPN1S (35)	49.98%
	34.8Kb	150	47	Gifsy 2 (30)	51.95%