UNIVERSIDAD SAN FRANCISCO DE QUITO USFQ

Colegio de Posgrados

Exploring prophages in *Salmonella enterica*: an *in-silico* approach

Mecanismo de Titulación: Tesis en torno a una hipótesis o problema de investigación y su contrastación

Francisco Antonio Quelal Madrid

PhD. Sonia Zapata

Directora de Trabajo de Titulación

Trabajo de titulación de posgrado presentado como requisito para la obtención del título de Máster en Microbiología

Quito, 22 de diciembre de 2023

UNIVERSIDAD SAN FRANCISCO DE QUITO USFQ
COLEGIO DE POSGRADOS

HOJA DE APROBACIÓN DE TRABAJO DE TITULACIÓN

Exploring prophages in *Salmonella enterica*: an *in-silico* approach

Francisco Antonio Quelal Madrid

Nombre del Director del Programa:	Patricio Rojas-Silva
Título académico:	M.D., Ph.D.
Director del programa de:	Maestría en Microbiología
Nombre del Decano del colegio Académico:	Carlos Valle
Título académico:	Ph.D.
Decano del Colegio:	COCIBA
Nombre del Decano del Colegio de Posgrados:	Hugo Burgos
Título académico:	Ph.D.

Quito, 22 de diciembre de 2023

© DERECHOS DE AUTOR

Nombre del estudiante:	Francisco Antonio Quelal Madrid
Código de estudiante:	00326901
C.I.:	1722906276
Lugar y fecha:	Quito, 22 de diciembre de 2023

ACLARACIÓN PARA PUBLICACIÓN

Nota: El presente trabajo, en su totalidad o cualquiera de sus partes, no debe ser considerado como una publicación, incluso a pesar de estar disponible sin restricciones a través de un repositorio institucional. Esta declaración se alinea con las prácticas y recomendaciones presentadas por el Committee on Publication Ethics COPE descritas por Barbour et al. (2017) Discussion document on best practice for issues around theses publishing, disponible en http://bit.ly/COPETheses.

UNPUBLISHED DOCUMENT

Note: The following graduation project is available through Universidad San Francisco de Quito USFQ institutional repository. Nonetheless, this project – in whole or in part – should not be considered a publication. This statement follows the recommendations presented by the Committee on Publication Ethics COPE described by Barbour et al. (2017) Discussion document on best practice for issues around theses publishing available on http://bit.ly/COPETheses.

DEDICATORIA

A mis padres y hermanos, por su amor incondicional y cuidados de siempre: todo lo bueno que tengo ahora empezó con ustedes. A mis amigos por sus enseñanzas y compañía a lo largo del tiempo. A mis maestros y mentores por transmitir sus enseñanzas y confiar en que nuestras enseñanzas van a trascender. A todos por mostrarme su amor y cariño, ya que este logro no es solo mío sino también de los que han caminado junto a mí todo este tiempo.

AGRADECIMIENTOS

A la Universidad San Francisco de Quito, al Instituto de Microbiología, por abrirme las puertas y hacerme sentir como en casa, y por todo el apoyo académico y logístico que me han brindado.

A Sonia Zapata, por su guía profesional y experta durante este proceso, y su apoyo incondicional que me inspira a seguir por este camino de científico.

A Carolina Armijos y Lorena Mejía, por sus comentarios y apoyo en la investigación.

A todo el equipo del Laboratorio de Parasitología, por su amistad y hacerme sentir siempre bienvenido.

Al equipo de Bioinformática del Instituto de Microbiología, por su guía y enseñanza en muchos aspectos que me permitieron realizar este estudio.

A mis amigos y compañeros de Maestría, especialmente Gaby, Santi y Pame, por permitirme aprender de ustedes y compartir lindos momentos.

A todos mis profesores y profesoras de la Maestría de Microbiología, por enseñarme muchas cosas útiles que no solo están a nivel académico, sino también permiten mi crecimiento personal.

A Bianca, por ser una gran amiga que me enseña y me apoya en los proyectos que me he propuesto, y a pesar de la distancia siempre me ha demostrado su amor y apoyo.

A mi familia, mamá, papá y hermanos, por su infinito amor que me permite seguir y alcanzar mis sueños y metas.

RESUMEN

Los profagos en *S. enterica* constituyen hasta un 30% del genoma accesorio, y pueden influir en la adaptación al medio ambiente y la susceptibilidad a otros bacteriófagos.

En el presente estudio se identificaron profagos en 142 genomas de *S. enterica* provenientes de granjas avícolas de los cuales 137 correspondieron al serovar Infantis, 3 a Typhimurium y 3 Enteritidis. Para esto se preparó una base de datos de 160 genomas de profagos de *S. enterica, E. coli* y otros profagos reportados en genomas de *Salmonella* pero que provenían de hospedadores no entéricos.

Luego del análisis bioinformático se identificaron cuatro profagos con >95% de identidad y >60% de integridad de los cuales, *Peduovirus pro483* fue el profago más prevalente, presente en 132 genomas de *S.* Infantis. *Enterobacteria ST104* y *Gifsy*-2, estuvieron presentes en un genoma de S. Typhimurium, y *Peduovirus fiAA91ss* presente solo en un genoma de *S.* Infantis.

Por otro lado, se identificaron genes de virulencia tales como *gtgA* y *sodC*, y de exclusión de superinfección como *sieB* y *gp17*, en las secuencias correspondientes a *Enterobacteria ST104* y *Gifsy-2*.

El gen *terL*, común entre los profagos encontrados y otras secuencias relacionadas permitió esclarecer la relación de *Enterobacteria ST104* con otros *Lederbergvirus*, *Peduovirus pro483* y *fiAA91ss* con otros Peduovirus, y *Gifsy-2* con otros fagos lamboides.

A nuestro conocimiento este es el primer estudio que describe la presencia de profagos en cepas de *Salmonella enterica* aisladas en granjas avícolas ecuatorianas

Palabras clave: *S. enterica*, Infantis, profagos, *in-silico*, Ecuador.

ABSTRACT

Prophages in *S. enterica* constitute up to 30% of the accessory genome, potentially influencing adaptation to the environment and susceptibility to other bacteriophages.

In this study, prophage sequences were identified in 142 *S. enterica* genomes from poultry farms, with 137 corresponding to serovar Infantis, and 3 each to Typhimurium and Enteritidis. A comprehensive database containing 160 prophage genomes from *S. enterica, E. coli,* and other *Salmonella* prophages originating from non-enteric hosts was employed for bioinformatics analysis.

Four prophages, exhibiting >95% identity and >60% integrity, were identified. *Peduovirus pro483* emerged as the most prevalent, present in 132 S. Infantis genomes. *Enterobacteria ST104* and *Gifsy-2* were detected in a *S.* Typhimurium genome, while *Peduovirus fiAA91ss* was exclusive to a *S.* Infantis genome.

Notably, virulence genes such as *gtgA* and *sodC*, along with superinfection exclusion genes like sieB and gp17, were identified in the sequences corresponding to *Enterobacteria ST104* and *Gifsy-2*.

The *terL* gene, shared among the identified prophages and related sequences, provided insights into the relationships, elucidating the connection of *Enterobacteria ST104* with other *Lederbergviruses, Peduovirus pro483* and *fiAA91ss* with other *Peduoviruses,* and *Gifsy-2* with other lamboid phages.

To our knowledge, this marks the first study describing the presence of prophages in Salmonella enterica strains isolated from Ecuadorian poultry farms.

Keywords: *S. enterica,* Infantis, prophages, *in-silico*, Ecuador

CONTENT INDEX

TABLE INDEX

FIGURE INDEX

PART 1: GENERAL INTRODUCTION

Salmonella General Aspects

Salmonella is a Gram-negative bacterium from the Enterobacteriaceae family. This genus comprises two species: *S. enterica* and *S. bongori*, being the first the most studied due to its pathogenicity (Hiyoshi et al., 2018). More than 2500 serovars have been identified, showcasing the remarkable diversity that enables *Salmonella enterica* to adapt to specific environments and hosts (Boyle et al., 2007; Coburn et al., 2007; Gymoese et al., 2019). Typically, *Salmonella* serovars fall into two primary groups: typhoidal and non-typhoidal. Typhoidal serovars exhibit a distinctive trait of colonizing and causing diseases exclusively in humans (Popa & Papa, 2021).

Non-typhoidal serovars represent the majority and are commonly associated with outbreaks and contamination of water sources through farming and agricultural practices (Guerrero et al., 2021). Among these serovars, some exhibit host specificity, with notable examples being *S.* Dublin in cattle and *S.* Cholerasuis in pigs (Hiyoshi et al., 2018). In humans, susceptibility to infection linkage on factors such as the individual's immune system and age, with early childhood, the elderly, and immunocompromised individuals being particularly vulnerable (Ao et al., 2015). Symptoms can range from mild, self-limiting gastroenteritis to more severe cases involving the invasion of epithelial cells in the host's intestinal tract through intracellular infection (Boyle et al., 2007).

The most severe cases of non-typhoidal salmonellosis are related to the virulence factors that confer different levels of adaptation to the host environment (Coburn et al., 2007). Six main virulence factors include secretion systems, fimbriae, flagella and flagellin, ion acquisition genes, superoxide dismutase genes, and virulence plasmids (Ibarra & Steele-Mortimer, 2009). Some virulence factors are characteristic to specific serovars and are also used for their

differentiation. However, these factors can also be acquired from mobile genetic elements, particularly plasmids and bacteriophages (Cui et al., 2021; Figueroa-Bossi & Bossi, 1999; Jajere, 2019). These virulence factors are crucial in the epidemiology of certain *Salmonella* serovars due to their implication in the emergence and success of some strains associated to specific outbreaks (Boyle et al., 2007).

On the other hand, antimicrobial resistance has become another noteworthy characteristic in *Salmonella*, exhibiting a notable increase over the past few decades. Most cases of multidrug resistant (MDR) *Salmonella* affect developing countries primarily due to inadequate health systems, deficient prevention policies, and substandard food quality control measures (Jajere, 2019; WHO, 2023). MDR *Salmonella* is also a consequence of the intensive use of antimicrobials within the healthcare system and the food industry (Jajere, 2019; Van Boeckel et al., 2015; Vinueza Burgos, 2017). Currently, *S.* Typhi and several non-typhoidal *Salmonella* have been classified as serious threats by the Center for Disease Control and Prevention in the United States (CDC, 2022). Although epidemiologic surveillance of antimicrobial resistance is a common practice enforced by the World Health Organization (WHO), some local governments have failed to implement surveillance, complicating the identification of emerging resistance patterns (WHO, 2017).

The WHO (2015) has identified *Salmonella enterica* as one of the 31 pathogens responsible for causing intestinal or systemic diseases, ranking as the third leading cause of death among food-related illnesses. Particularly in Latin America, approximately 77 million people are victims of foodborne infections, with 9,000 fatalities resulting from contaminated food. Non-typhoid *Salmonella* alone contributes for 95% of these cases, alongside other foodborne pathogens such as *Campylobacter*, *Escherichia coli*, and Norovirus (WHO, 2015). This alarming situation underscores the pressing need to concentrate efforts on enhancing quality control measures and

implementing effective policies within agricultural systems, food safety practices, and public policies in the region (Guerrero et al., 2021).

Salmonella enterica serovar Infantis.

S. Enteritidis and *S.* Typhimurium are recognized as the main serovars responsible for infections in both humans and broilers. Contaminated products linked to these serovars, particularly raw chicken meat and eggs, are widely acknowledged as the primary sources of human infections (Guerrero et al., 2021). Since 2014, there has been a growing prevalence of *S.* Infantis in poultry farms, potentially attributed to the displacement of S. Enteritidis and S. Typhimurium through quality control and vaccination policies (Montoro-Dasi et al., 2023). This trend raises concerns about potential risks to human health.

Although the prevalence of certain serovars may vary geographically, *S.* Infantis has exhibited a clear predominance on a global scale (Moulana & Asgharpour, 2022). Since 2020, *S.* Infantis has been recognized as the fourth among the 10 most prevalent non-typhoidal serovars associated with human infections worldwide (Montoro-Dasi et al., 2023). This recognition underscores the importance of ongoing monitoring and research efforts to understand and mitigate the evolving dynamics of *S.* Infantis and its impact on public health.

In the European Union, the most prevalent serovars associated with human infections are *S.* Enteritidis, *S.* Typhimurium and its monophasic variant, *S.* Infantis and *S.* Derby, respectively, accounting for 54.6%, 11.4%, 8.8%, 2.0% and 0.93% of cases according to the European Centre for Disease Prevention and Control (ECDC) (2022). All serovars have been associated with different environments in the food and meat production industry (Ferrari et al., 2019).

In Latin America, *S.* Infantis has also been reported as an emerging serovar among poultry farms with different prevalence depending on the country (Vinueza-Burgos, 2017). Reports from Ecuador and Peru indicate that S. Infantis is the predominant serovar in broilers and

poultry farms, constituting a significant proportion ranging from approximately 83% to 98% of the isolates associated with this serovar (Vallejos-Sánchez et al., 2019; Vinueza-Burgos et al., 2019). This high prevalence may be attributed to the close interaction between poultry industries of Peru and Ecuador (Vinueza-Burgos et al., 2019). In other countries of the region, *S.* Infantis was reported as an underrepresented serovar in broiler, chicken meat and poultry farms (Lapierre et al., 2020; Medeiros et al., 2011). This highlights the need for more research focused on *S.* Infantis to understand the epidemiology of this emerging serovar in the region.

The emergence of this serovar could be attributed to vaccination and control programs aimed at *S.* Enteritidis and *S.* Typhimurium (Montoro-Dasi et al., 2023). The decline or near-eradication of these previous serovars creates an ecological niche that facilitates the colonization and success of S. Infantis (ECDC, 2022; Montoro-Dasi et al., 2023). Additionally, *S.* Infantis strains have been reported to present different genetic features that include antimicrobial resistance, acquisition of mobile genetic elements and virulence factors which enhanced its pathogenicity and fitness (Montoro-Dasi et al., 2023).

The global dominance of *S.* Infantis in poultry farms is thought to be facilitated by a megaplasmid called pESI that imparts resistance and virulence features, including the adaptation to oxidative stress, mercury tolerance, fimbriae and a diverse number of resistance genes which explains the MDR patterns (Hall et al., 2022; Moulana & Asgharpour, 2022). It appears that pESI-like megaplasmids in *S.* Infantis have been vertically acquired, but other enteric commensals have also reported the presence of this type of megaplasmids (McMillan et al., 2020; Shah et al., 2017).

On the other hand, the presence of prophage sequences within *S.* Infantis could explain part of the success of this serovar in the poultry industry (Cohen et al., 2020; D'Alessandro et al., 2018; Trofeit et al., 2023). Prophages sequences normally carry genes associated with resistance

genes and virulence factors (Żbikowska et al., 2020) as well as genes associated with host cellular invasion, intracellular survival and biofilm formation (Wahl et al., 2019). These prophage sequences could have been acquired by interacting with other poultry-associated serovars, such as *S.* Gallinarum (Kipper et al., 2022). Nonetheless, information about prophages in *Salmonella* remains scarce and not deeply studied.

Prophages of *Salmonella enterica*

Up to 30% of the accessory genome of *Salmonella enterica* is comprised of prophages and these could play a crucial role in the genetic diversity of this microorganism (Wahl et al., 2019). Their influence has been extensively studied because of the expression of toxins, virulence factors and genes associated with metabolic intake affecting bacterial fitness and physiology (Kraushaar et al., 2017). However, prophages can also affect the interaction of *S. enterica* with its environment through the expression of specific genes that modulate signaling to other microorganisms or phages (Mottawea et al., 2018; Wahl et al., 2019).

While certain prophage species infecting *Salmonella* genomes have been reported in other bacterial hosts like *Escherichia coli, Yersinia, Klebsiella and Haemophilus*, induction assays have demonstrated that most prophages in *Salmonella* have narrowed their host range to specific serotypes and strains (Gao et al., 2020). The most common prophages found in *Salmonella* are *Lederbergvirus*-like phages (formerly p22-like phages), from which *Lederbergvirus* p22 and *Lederbergvirus* ST64T features as the most studied prophages found in *S.* Typhimurium conferring O-antigen modifications, phage resistance, and other characteristics to the host (Gao et al., 2020; Switt et al., 2015; Wahl et al., 2019). Other common phages found in *Salmonella* are lamboid phages *Gifsy-1, Gifsy-2* and *Gifsy-3*, conferring virulent characteristics and promoting transduction or interaction with other genomic elements through recombination (Wahl et al., 2019). Most prophages have been described within *S.*

Typhimurium genomes due to its importance as a clinical and environmental prevalent serovar, but studies focused on other serovars are crucial to understand the adaptation of *Salmonella* to specific circumstances (Fong et al., 2022; Gao et al., 2020; Trofeit et al., 2023). A study by Cohen et al. (2021) has established that the predominant prophages in *S.* Intantis are *Bcepmuvirus bcepmu* (as a conserved sequence in different *S.* Infantis lineages), *Peduovirus pro483* (emergent and diverse among *S.* Infantis lineages) and *Gifsy-1* (predominant in *Salmonella* along with other lamboids and *Gifsy*-like phages).

Prophages remain in a dormant state in the *Salmonella* genome, and they tend to be inherited vertically. Under specific stressful conditions, prophages can excise from the bacterial chromosome or integrate into it (Wahl et al., 2019). The excision or integration of prophages is typically triggered by factors such as DNA damage, the availability of hosts, or even the influence of other phages. These events can significantly impact the fitness of the host (Mottawea et al., 2018).

S. enterica genomes also present defective prophage sequences that have lost the ability to form infectious particles but retain certain features (Bobay et al., 2014; Henrot & Petit, 2022). The interaction with other mobile genetic elements, host defensive systems, recombination events and a high mutation rate are the main reasons prophages can lose function (Bobay et al., 2014). However, some defective prophages can be reactivated by the influence of other infecting phages that provide the correct signaling or specific functions. *S. enterica* can contain defective phages that can no longer be induced or reactivated and are fixed, similar to pathogenicity and genomic islands (Kaur & Jain, 2012).

Prophage virulence genes are the most studied for their effect on *S. enterica* survival and adaptability to the environment (Wahl et al., 2019). The most common functions provided by prophages while lysogenized in their host are the expression of toxins, secretion systems and

antimicrobial resistance (Bobay et al., 2014; Canchaya et al., 2004; Wahl et al., 2019). A common virulence gene is *sopE, a*ssociated with cellular invasion. This gene was originally identified in *SopEPhi* prophage in *S. enterica* SL1344 and has been found in many other prophage species modulating the adaptation of *S. enterica* to new environments (Bachmann et al., 2014; Wahl et al., 2019). The impact of *sopE* harboring genomes has been reported to be essential in the emergence of certain serovars like *S.* Infantis (Petrovska et al., 2016; Rusconi et al., 2016).

The contribution of prophages in *S. enterica* is important especially for their implications on phage therapy success (Wahl et al., 2019). Some prophage genes can modulate the infection of other phages through the expression of genes that modify surface receptors used by other phages to recognize their target (Owen et al., 2020). An example of these modifications is found in *p22*-like phages, which carries the *gtrABC* operon for O-antigen recognition, but when lysogenized modifies the host O-antigen composition by glycolysation (Wahl et al., 2019). Another known function is the superinfection immunity or exclusion found in many *Salmonella* phages. The most notable examples are the *SieA* and *SieB* found in *p22*-like phages, blocking the C2 receptors (Cenens et al., 2016).

Some prophage genes have been reported to have different functions depending on the *Salmonella* serovar in which they are integrated. For example, *Lederbergvirus BTP1* has been reported to confer immunity to *p22*-like phages when is integrated in *S.* Typhimurium through expression of *bstA* gene (Owen et al., 2021), but when infecting *S. enterica* ST313 it was found to act as a virulence factor affecting the macrophage uptake during intracellular infection (Herrero-Fresno et al., 2017). These differences in function highlight the need to study the effects of prophages in different bacterial hosts. Since *S. enterica* is characterized for its ubiquity in a diversity of environments and hosts, the interaction of prophages and their

influence on the species behavior, the study of prophages is important to understand the biology of this bacterium.

PART 2: SCIENTIFIC PAPER

Introduction

Prophages are specific forms of bacteriophages that are integrated into the bacterial genome as lysogens and remain latent until certain environmental or host physiological conditions trigger their excision (Dion et al., 2020). Phageomes are estimated to comprise approximately 80% of the virome in some environments, particularly gut microbiomes (Dion et al., 2020; Henrot & Petit, 2022; Wendling, 2023), and can represent >20% of their host's genome (Dion et al., 2020; Nishijima et al., 2022; Wendling, 2023). While residing latent in a genome, prophages can influence several aspects of their host's physiology, such as virulence, metabolic intake, antimicrobial resistance, and phage infections, making them an interesting focus, especially for gastrointestinal tract pathogens like *Salmonella* (Nayfach et al., 2021; Tisza & Buck, 2021; Wahl et al., 2019).

In *S. enterica*, prophages constitute up to 30% of the accessory genome and are recognized as one of the main contributors to the species' diversity (Wahl et al., 2019). Their influence has been traditionally studied through the expression of virulence genes and other characteristics that influence the adaptation of *S. enterica* to different environments, but recent studies have focused on the influence of prophages in the infection dynamics, through the expression of genes that confers resistance to certain groups of phages (Dion et al., 2020; Henrot & Petit, 2022; Wahl et al., 2019). A clear example is observed in many *p22*-like phages which harbor superinfection exclusion genes that prevent the infection of multiple phages by inducing abortive infection or surface antigen modification (Owen et al., 2021).

Understanding the influence of prophages in *S. enterica* is crucial for comprehending the emergence of certain strains or serovars. This influence operates through the expression of

virulence genes, promoting the success of these strains in specific environments (Cohen et al., 2020; Gymoese et al., 2019; Moura de Sousa et al., 2021; Weissman et al., 2018).

The main objective of this study was to detect *S. enterica* prophages within 142 S. *enterica* genomes from poultry farms, chicken carcasses and clinical casesto assess their functionality, presence of virulence genes and influence in the bacterial fitness. This investigation aims to unravel the relationships between the identified prophages and other related bacteriophages associated with *S. enterica*, providing comprehensive insights into the intricate microbial ecosystems.

Methodology

Genome Assembly

The genome Assembly Protocol utilized 151 Illumina NextSeq paired-end sequences obtained from a previous study (Mejía et al., 2020) (available under bioproject PRJEB37560). Details regarding sequence names and serovars are provided in Supplementary Table 1. Trimmomatic (v0.39) (Bolger et al., 2014) was employed to trim ambiguous nucleotides from paired-end sequences, ensuring a quality threshold >20 and a minimum length of 100 bp. Quality control was performed using MultiQC (v1.16) (Ewels et al., 2016). SPAdes (v3.15.4) (Prjibelski et al., 2020) facilitated the assembly of filtered sequences, employing a cut-off value of 25 and a careful pipeline flag. Only assembled scaffolds were utilized to ensure the identification of large prophage genomes. Assembly quality was assessed with Abyss (v2.3.5) (Jackman et al., 2017), utilizing the abyss-fac flag, and BUSCO (v5.5.0) (Manni et al., 2021). To guarantee less fragmented identified prophage genomes, and to obtain more comprehensive and informative results, we selected genomes meeting specific criteria: assemblies with <150 scaffolds, N50 >150 kb, L50 <15, genome completeness >85%, and a file size >4.1 MB.

Prophage database and genome identification

A database of *Salmonella* prophage genomes was created for *in silico* identification using BLAST Command Line Applications (v2.13.0+) (Altschul et al., 1990; Morgulis et al., 2008) with the *megablast* task, with a percent identity >95%, and an E-value <E-100, to ensure high similarity and statistical significance in our BLAST results. The database consisted of 160 reference prophage genomes retrieved from prior studies (Gao et al., 2020; Switt et al., 2015) that established that these genomes were exclusive or related to *S. enterica* genomes to avoid other prophages with a wide range of hosts (Supplementary Table 2). The completeness and correct species assignment of the references were assessed through manual comparison with the NCBI Nucleotide and RefSeq databases (Sayers et al., 2022) and the latest information from the International Committee on Taxonomy of Viruses (ICTV) (Turner et al., 2021; Walker et al., 2020). Only matches with lengths >6 kB were selected because *Inovirus M13* was the smallest prophage genome in our database with 6408 bp.

Prophage genome annotation and gene functionality

Each identified prophage genome species was concatenated and realigned with its respective reference from our custom database (Supplementary Table 2) using MAFFT (v7.520) (Katoh et al., 2019). Previous studies (Ha & Denver, 2018; Hatfull et al., 2010; Pope et al., 2011) used alignments representing 45%-58% of the reference genome with >90% of identity to evaluate gene function and completeness level. In our case we used prophage genomes whose alignment represented >60% of the reference genome and a percent identity >95% to ensure a higher proportion of the prophage genomes recovered prior to evaluate their function through gene annotation.

Gene and ORF calling were performed using Prodigal (v2.6.3) (Hyatt et al., 2010) with the *–g 11* flag to specify the bacterial-plastid genetic code. Gene sequence identification was

performed using BAKTA (v1.8.1) (Schwengers et al., 2021). Annotated genomes were visualized and curated using Geneious Prime (v2023.0.2) (Kearse et al., 2012) to evaluate nonsense stop codons and incorrectly identified ORFs. Hypothetical gene sequences were aligned against the UniProt (Coudert et al., 2023) database using the BLAST online service. Curated annotated genomes were compared with taxonomically closely related references according to the ICTV (Turner et al., 2023; Walker et al., 2020) taxonomy to identify shared gene features with the same protein coding product.

Prophage genome analysis

Genomes of *Salmonella* prophages were mapped to their respective reference genomes from our database using Bowtie2 (v1.3.1) (Langmead & Salzberg, 2012) to identify the location of the genes found. Subsequently, the annotated genomes were aligned using progressiveMauve (v2.4.0) (Darling et al., 2010) to identify conserved regions or locally colinear blocks when comparing the *Salmonella* prophages with their corresponding reference genomes. Genome comparisons between the identified prophages and the reference genomes were generated using graphical representations in Geneious Prime (v2023.0.2) (Kearse et al., 2012).

Salmonella prophage species estimation

Further validation of *Salmonella* prophage identification involved a manual comparison to identify shared genes among the detected prophage genomes, their corresponding references in our database, and other taxonomically related phages from the ICTV (Turner et al., 2023; Walker et al., 2020) database (Supplementary Table 2). te (Kumar et al., 2018). Subsequently, a maximum likelihood tree was constructed with IQ-TREE2 (v2.2.5) (Nguyen et al., 2015) using the previously generated alignment, employing a fast bootstrap of 1000 iterations. The resulting tree was visualized with FigTree (v1.4.4) (http://tree.bio.ed.ac.uk/software/figtree/).

Additionally, using the same parameters, a second maximum likelihood tree was constructed with complete genome sequences.

Results

Salmonella genomes and prophage identification

After trimming the 151 paired-end reads to remove duplicated or ambiguous sequences, an average of 1.24 million reads (ranging from 0 to 14.1 million reads) remained. Subsequently, these reads were assembled into genomes, of which only 142 genomes met the criteria of <150 scaffolds, N50 >150 kB, L50 <15, genome completeness >85%, and file size >4.1 MB, necessary for unfragmented prophage genome identification (Supplementary Table 1; Supplementary Figure 1).

The 142 assembled genomes exhibited an average of 152.63 scaffolds, with scaffolds averaging 153,849 bp for N50 and 6.62 bp for L50. To ensure no data loss, the genome size and file size of the assembled genomes maintained an average of 4.71 Mbp and 4.84 MB, respectively (Supplementary Figure 1). Of there genomes 136 were identified as *S. Infantis,* 3 were *S.* Enteritidis and 3 were *S.* Typhimurium (Supplementary Table 3).

Thirteen distinct *Salmonella* prophage-associated sequences were identified with an identity >95% (Supplementary table 3). Notably, only one genome from *S.* Infantis (0.70%) did not harbor any *Salmonella* prophages. Of the remaining genomes, 11 out of 142 (7.75%) contained 1 prophage, 120 out of 142 (84.51%) harbored 2 prophages, 9 out of 142 (6.34%) hosted 3 prophages, and only one (0.70%) featured 4 prophages.

Peduovirus pro483 was the most prevalent, followed by *Stockinghallvirus FSL SP-004.* Other prophage matches found constituted less than 6% of the analyzed genomes. *Punavirus SJ46, Phage Gifsy-1, Lambdavirus lambda* and *Peduovirus fiAA91ss* were uniquely identified in their respective samples.

Phage Gifsy-2 was identified in one *S.* Infantis, two *S.* Typhimurium, and two *S.* Enteritidis genomes. *Salmonella phage ST64B* was present in three *S.* Enteritidis and two *S.* Typhimurium genomes. Only one *S.* Typhimurium genome harbored Enterobacteria phage ST104 with 95.02% query coverage (Table 1).

Table 1. Number of *Salmonella serovars* harboring each of the prophages detected

Prophage Species	Detected genome size	Reference genome size	Query coverage	Number of prophage harboring genomes (%)	Serovar
Peduovirus pro483	12887 bp	29237 bp	44.07 %	49 (34.51%)	Infantis
	12914 bp		44.17 %	1 (0.70%)	Enteritidis
	12945 bp		44.27 %	5 (3.52%)	Infantis
	22183 bp*		75.87 %	63 (44.37%)	Infantis
	22230 bp*		76.03%*	14 (9.86%)	Infantis
Stockinghallvirus FSL SP-004	6055 bp	29742 bp	20.36 %	120 (84.51%)	Infantis
Peduovirus Wphi	6853 bp	32684 bp	20.97 %	8 (5.63%)	Infantis
Shigella phage SfII	6087 bp	41475 bp	14.68 %	1 (0.70%)	Infantis
	10031 bp		24.18 %	2 (1.41%)	Infantis
Enterobacteria phage mEp460	13532 bp	44510 bp	30.40 %	1 (0.70%)	Infantis
	14114 bp		31.71 %	1 (0.70%)	Infantis
Salmonella phage 118970-sal3	13543 bp	77375 bp	17.50 %	1 (0.70%)	Infantis
	25721 bp		33.24 %	1 (0.70%)	Infantis
Peduovirus fiAA91ss	22122 bp	33628 bp	65.78%*	1 (0.70%)**	Infantis
Lambdavirus lambda	16347 bp	48502 bp	33.70 %	1 (0.70%)	Infantis
Phage Gifsy-1	22005 bp	48491 bp	45.38 %	1 (0.70%)**	Infantis
Phage Gifsy-2	6091 bp	45840 bp	13.29 %	1 (0.70%)	Typhimurium
	6091 bp		13.29 %	2 (1.41%)	Enteritidis
	6429 bp		20.40 %	1 (0.70%)	Infantis
	31512 bp*		68.71 %	1 (0.70%)	Typhimurium
Enterobacteria phage ST104	38150 bp*	41391 bp	95.02 %	1 (0.70%)**	Typhimurium
Salmonella phage ST64B	15325 bp	40149 bp	38.17 %	2 (1.41%)	Typhimurium
	15325 bp		38.17 %	3 (2.11%)	Enteritidis
Punavirus SJ46	10688 bp	103445 bp	10.33 %	1 (0.70%)**	Infantis

(*)Selected prophages with query coverage >60%
(**) Only sample presenting the current Salmonella prophage

We selectively retained matches with an identity >95% and query coverages >60% to determine gene function through genome annotation. Only 4 prophage genomes met these criteria:

Peduovirus pro483 and *Peduovirus fiAA91ss* in *S.* Infantis samples, and both *Enterobacteria phage ST104* and *Phage Gifsy-2* in the same *S.* Typhimurium sample (Table 1).

Prophage genome analysis

Finally, four *Salmonella* prophage genomes selected for genome annotation were mapped and aligned to their respective references to identify their gene location and conserved regions.

Only *Enterobacteria phage ST104* showed a similar number of features compared to its reference genome. Matches for *Peduovirus pro483* and *fiAA91ss* revealed a clear loss of features associated with structural proteins, infection, and excision proteins in contrast to their reference genomes. Regarding *Phage Gifsy-2*, genes encoding phage structural proteins, transcription, and excision proteins were absent from the identified sequence (Supplementary Figure 3).

The prophage *FQ-Enterobacteria ST104* showed the presence of the gene *lar*, encoding a restriction alleviation protein, which was absent in its reference genome. Our recovered sequence also lacked the following genes: *gtrA, gtrB,* and *gtrC,* encoding O-antigen conversion proteins; *xis*, for an excisionase enzyme; *ninY* and *ninZ,* for phage nuclease enzymes; *gp64*, for a holin enzyme; and *gp13*, for a DNA transfer protein (Figure 1A).

Sequences of *FQ-Peduovirus pro483* and *Peduovirus pro483* revealed clear differences with their reference sequences. *FQ-Peduovirus pro483* lacked a block of 12 genes, including *int* (phage integrase), *ogr* (a late control protein), *cox* (excisionase), *dnaB* (recombinase), *gpB* (replication initiation), *dksA* (zinc-finger-like protein), *gpA* (nicking at origin protein), and 4 hypothetical proteins (Figure 1B). Notably, the gene *tfaE*, a tail fiber assembly protein, was found to be inverted in our sequence. Furthermore, in contrast to the absence of the 12 genes upstream, a conserved block of 3 genes was identified downstream, in which *ogr* was duplicated in the reference genome (Figure 1B).

In *FQ-Peduovirus fiAA91ss*, a block of 13 downstream genes was absent (Figure 1C). These genes included *int* (phage integrase), *gpC* (a transcriptional regulator), *istA* and *istB* (genes for a transposable element), *dnaB* (recombinase), *gpB* (for replication initiation), *gpA* (nicking at origin protein), *cdtvA*, *cdtvB*, and *cdtvC* (encoding cytholethal toxins), and 2 hypothetical proteins. Additionally, 2 inverted blocks were identified: *gpG* (a tail fiber assembly protein) and a block consisting of *fI, fII*, and *e*, associated with phage tail proteins. Notably, 2 genes were found to be unique to *FQ-Peduovirus fiAA91ss: tfaE* (a tail fiber assembly protein) and pinE *(*a DNA invertase) (Figure 1C)

FQ-Phage Gifsy-2 lacked 21 genes compared with its reference (Figure 1D). Upstream was evident the absence of a block consisting of *int*, for a phage integrase, *xis*, for excisionase, and *recT*, a recombinase, and downstream a block of 3 genes for *gpN*, *gpT*, for tail fiber assembly, *sseI*, a virulence determinant, was missing. Important features present for both were also found: *sodC*, for superoxide dismutase, and *gtgA*, for a type III secretion system.

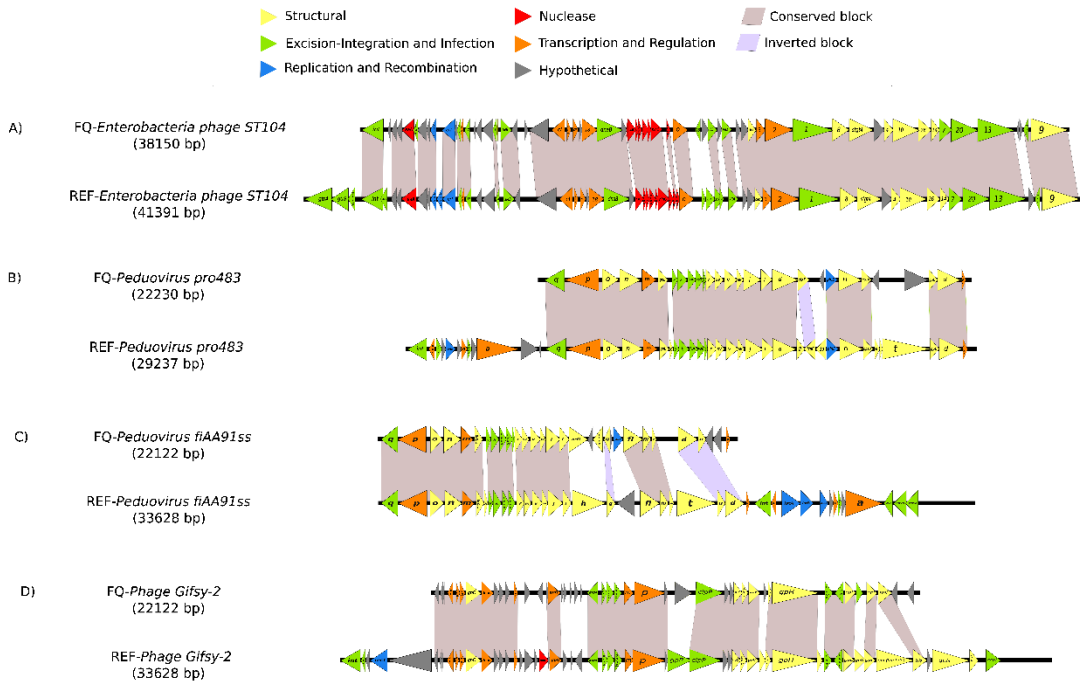


Figure 1. Genome alignments between four most complete *Salmonella* prophages identified and their respective reference genomes in our custom database. Gene functions were color-coded as indicated in the figure legend. Conserved genetic blocks were also highlighted.

Ultimately, *FQ-Enterobacteria phage ST104* and *FQ-Phage Gifsy-2* both presented 3 genes with potential influence on host physiology and phage infection dynamics: *ral* (restriction alleviation protein), *sieB*, and *gp17* (superinfection exclusion), *lomR* (an opacity protein for surface antigens), and *sodC* (superoxide dismutase). In contrast, both *Peduovirus pro483* and *fiAA91ss* only presented *pinE*, which has a potential influence on phage infection as it encodes a DNA invertase (Table 2).

Table 2. List of genes identified with potential influence on phage infection dynamics and host physiology.

Prophage identified	Gene name	Function
FQ-Enterobacteria phage ST104	*ral*	Restriction alleviation protein
	sieB	Superinfection exclusion protein
	gp17	Superinfection exclusion protein
FQ-Peduovirus pro483	*pinE*	DNA invertase
FQ-Peduovirus fiAA91ss	*pinE*	DNA invertase
FQ-Phage Gifsy-2	*lomR*	Opacity protein for surface antigens
	gtgA	O-antigen conversion protein
	sodC	Superoxide dismutase

Salmonella prophages positioning and relationships

A maximum likelihood tree was constructed based on the alignment of the *terL* gene, which is the only gene shared among all analyzed prophage sequences (Figure 2).

Upon comparing the prophage sequences using the *terL* gene, it was evident that all *Peduovirus*-like sequences formed a cohesive cluster (Figure 2, Clade A). Notably, FQ-*Peduovirus pro483* and FQ-Peduovirus fiAA91ss exhibited strong similarity, with a robust bootstrap value, despite their corresponding references being positioned in different branches within the *Peduovirus* clade.

FQ-*Phage Gifsy-2* clustered with its reference and other lamboid prophages like *Gifsy-1*, *Enterobacteria phage Fels-1*, and *Salmonella phage Fels-2*. Intriguingly, *Lederbergvirus HK620* and *Lederbergvirus sf6* also grouped within this cluster (Figure 2, Clade B).

All *Lederbergvirus*-like sequences formed a distinct clade, with FQ-*Enterobacteria ST104* and its corresponding reference placed in the same branch, as expected (Figure 2, Clade C).

Since the *terL* gene was the sole shared gene among the analyzed prophage sequences, another maximum likelihood tree was constructed, this time utilizing complete genomes

(Supplementary Figure 2). Here, a different topology and relationships emerged in contrast to Figure 2. The *Peduovirus*-like clade included all *Peduoviruses*, except *Peduovirus pro147*, with FQ-*Peduovirus pro483* and FQ-*Peduovirus fiAA91ss* more closely related to REF-*Peduovirus pro483* (Supplementary Figure 2, Clade A).

FQ-*Phage Gifsy-2* was associated with *Gifsy-1*, forming a smaller clade, and its corresponding reference showed a closer relationship with *Peduovirus* rather than other lamboids, as observed previously (Supplementary Figure 2, Clade B).

Finally, *Lederbergvirus*-like sequences were found in a larger clade, encompassing prophages previously grouped with *Gifsy-2* in Figure 2, while maintaining FQ-*Enterobacteria ST104 associated* with its reference and other *Lederbergviruses* (Supplementary Figure 2, Clade C).

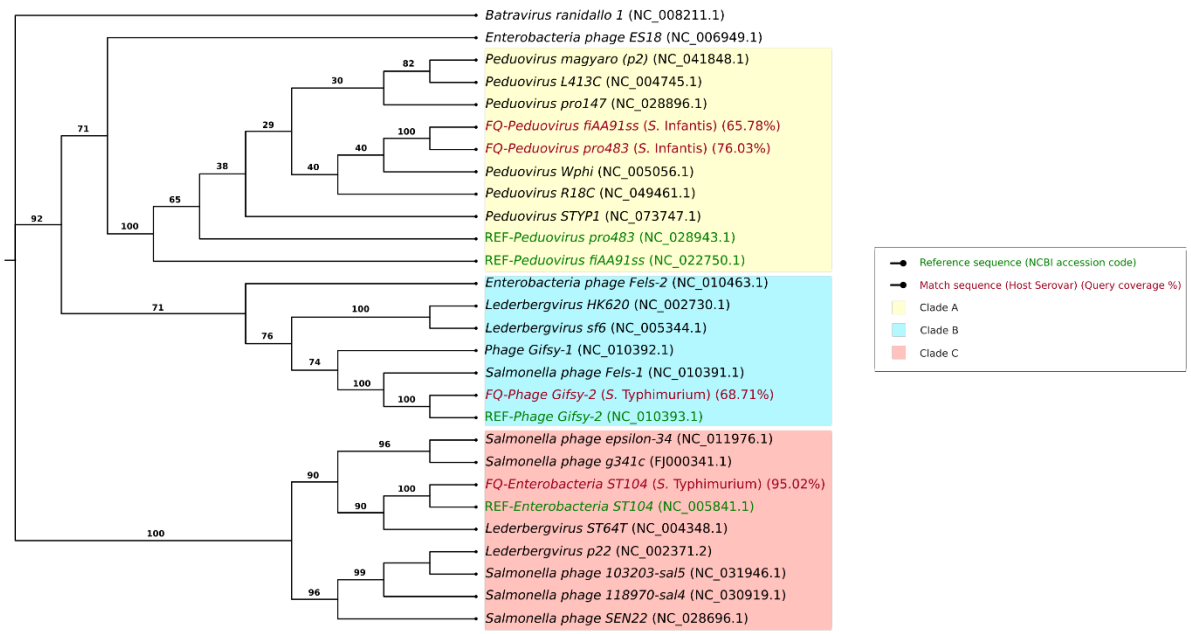


Figure 2. Phylogenetic relationships using the terminase large subunit gene (*terL*). Reference sequences are written in green and found sequences in red. 3 clusters were observed: Cluster A (yellow) with *Peduovirus*-like sequences, ClusterB (cyan) with *Gifsy*-like sequences, and Cluster C with *Lederbergvirus*-like sequences. Maximum likelihood tree constructed with IQ-TREE2 with a fast bootstrap of 1000 iterations using a general time reversible model with gamma distribution (GTR+G) predicted with MEGAX.

Discussion

Prophages in *S. enterica* are significant contributors to the species' genetic diversity, playing a crucial role in the adaptability of *S. enterica* to diverse environments and influence susceptibility to other infecting phages (Hu et al., 2021). Despite their ubiquity in *S. enterica* genomes comprising up to 30% of the accessory genome, studies on prophages remain scarce (Wahl et al., 2019).

In this study, we identified four prophages that best met our selection criteria: *Enterobacteria phage ST104, Peduovirus pro483, Peduovirus fiAA91ss* and *Phage Gifsy-2*. Although additional matches were excluded, there is a possibility of their association with these prophage

sequences. This is supported by observed mosaicism in phages, where specific genes or genetic blocks exhibit conservation, despite not sharing the same evolutionary history due to horizontal gene transfer (Jonge et al., 2019; Yu et al., 2017). Furthermore, our observations revealed that the identified prophages exhibited genome fragmentation or incompleteness, particularly in the cases of *Peduovirus pro483* and *Gifsy-2* matches. It is noteworthy that the phenomenon of prophage genome fragmentation or the presence of defective prophages is a common occurrence in bacterial genomes (Wahl et al., 2019). This reflects the rapid degradation of prophage sequences upon integration into the bacterial chromosome. Prophages play a crucial role in host survival, balancing the risk of bacterial lysis with the potential for adaptation through the expression of prophage genes. The degradation of prophages can be viewed because of selective pressure to stabilize defective or incomplete prophages (Bobay et al., 2014). Although the mechanisms governing the selection of prophages in bacterial genomes remain incompletely understood. The most described stabilization mechanisms involve recombination with other infecting phages or mobile genetic elements that share insertion sites or homologous genes (Bobay et al., 2014; Cohen et al., 2020; Gymoese et al., 2019). Additionally, host defense strategies target key phage genes, thereby promoting the selection of prophage sequences that enhance host fitness (Moura de Sousa et al., 2021; Weissman et al., 2018). Since our objective was to recover the most complete and functional prophage sequences, we narrowed our genome analysis to only the prophage matches mentioned above to ensure that functionality was not lost and to evaluate their influence on bacterial genomes and host physiology.

Prophages play a significant role in *S. enterica* genomes, with earlier studies linking specific serovars to prophages and environments (D'Alessandro et al., 2018; Switt et al., 2015). This is congruent with our results where *Peduovirus pro483* emerged as the most prevalent prophage sequence suggesting this prophage adapted to *S.* Infantis in the poultry farm environment, since *Peduovirus pro483* was originally isolated from an avian pathogenic *Escherichia coli*, and

present in *S.* Typhimurium isolated from poultry farms (Gymoese et al., 2019; Petrovska et al., 2016). Additionally, *Peduvovirus pro483* was commonly reported carrying the gene *sopE*, facilitating bacterial host epithelial cell invasion as in other *Peduovirus P2*-like prophages found in emerging clones from poultry environments (Gymoese et al., 2019; Pretovska et al., 2016; Rusconi et al., 2016). Although our match did not present the *sopE* gene, its absence could be attributed to recombination events, suggesting that *Peduovirus pro483* might be fixed in *S.* Infantis as evidence of its adaptation to poultry farms (Moura de Sousa et al., 2021; Weissman et al., 2018), as observed in the loss of functions and the presence of *pinE* (DNA invertase) in our recovered sequence.

Our study found only one *S.* Infantis genome harboring *Peduovirus fiAA91ss*. Our sequence matching also harbored the *pinE* gene, indicating a relationship to *Peduovirus P2* and other *Peduovirus* like *Peduovirus pro483*. In contrast to the reference genome used in our database, our match did not harbored cytolethal toxins and transposon insertion sequences, reinforcing the idea that certain phage functions are lost immediately after integration in the host genome due to recombination and influence of other mobile genetic elements (Bobay et al., 2014; Moura de Sousa et al., 2021; Weissman et al., 2018).

Although our *S. enterica* genomes did not exhibit a prevalence for *Phage Gifsy-2*, all sequences related to this prophage were identified across all serovars in our genomes. While lamboid phages such as *Phage Gifsy-1, Phage Gifsy-2,* and *Phage Gifsy-3,* have been documented to facilitate adaptability to the environment and horizontal gene transfer between *S. enterica* serovars through other phages, resulting in common conserved genetic blocks in Typhimurium and Enteritidis serovars (Braetz et al., 2023; Kurasz et al., 2023; Svahn et al., 2023), reports of *Phage Gifsy-2* in *S.* Infantis are limited. Interestingly a report by Calarga et al. (2021) presented *Phage Gifsy-2* in *S.* Infantis genomes along with *Peduovirus L413-C*, a *Yersinia pestis* phage also related to *Peduovirus fiAA91ss* (Qi et al., 2022), reinforcing the idea of *Gifsy*-like phages

in *Salmonella* enhancing adaptation to certain environments through interaction with other phages.

Enterobacteria phage ST104 present in *S.* Typhimurium is congruent to previous reports that associated this prophage with *S.* Typhimurium DT104 phage type strain (Parker et al., 2021). While this prophage is reported to be exclusive of *S.* Typhimurium DT104 and other related strains in cattle and ground beef production (Parker et al., 2021), our match stands out as one of few *Enterobacteria phage ST104* found in *S.* Typhimurium from poultry environments. The acquisition of this prophage by *S.* Typhimurium from poultry farms may be a result of the intensive use of antimicrobials in poultry production as prophylactics and growth promotors (Van Boeckel et al., 2015; Vinueza Burgos, 2017) since *Enterobacteria phage ST104* has been reported promoting antimicrobial resistance gene transduction (Emmanuel, 2021; Parker et al., 2021).

Apart from the four prophages previously described, there were 9 prophage sequences that were discarded from further analysis due to no fulfilling the selection criteria (>60% complete) (Table 1). We want to highlight the presence of *Stockinghallvirus FSL SP-004* and *Peduovirus Wphi* among these 9 prophage sequences because they represented the second and third most prevalent prophage sequences found, respectively. Previous reports of *Stockinghallvirus FSL SP-004* (previously known as *Salmonella phage FSL SP-004)* have found this prophage associated to *S.* Newport in dairy farms (Moreno Switt, 2013; Moreno Switt et al., 2013; Silva et al., 2016). This differs from our findings in which we reported *Stockinghallvirus FSL SP-004* only in *S.* Infantis, indicating that this prophage could be circulating in more than one serovar and could be associated to other environmental settings. As for *Peduovirus WPhi,* this prophage has been reported only in human microbiota and isolated from *E. coli* O145 in cattle feces (Farahmandzad et al., 2022; Shridhar et al., 2019), being a similar case as Stockinghallvirus FSL SP-004. Further studies focused on the identification of these prophages

could clarify the origin and influence of these prophages in *S.* Infantis genomes. The fragmentation and incompleteness of these prophage sequences could indicate the degradation and later stabilization of these sequences in the host genome as previously described (Bobay et al., 2014; Weissman et al., 2018).

Genes influencing *S.* enterica physiology through virulence factors and immunity to phage superinfection have also been found in *S.* Typhimurium. These genes were commonly carried by *Enterobacteria phage ST104* and *Phage Gifsy-2* when integrated in the same genome, which is congruent to our findings (Emmanuel, 2021; Parker et al., 2021). While avian-related *S. Typhimurium* is known to also harbor the *sopE* gene found in various emerging clones with *Enterobacteria phage ST104* (Kirkwood et al., 2021), our *S.* Typhimurium did not present this virulence gene in the sequence matching corresponding to this prophage. Both prophages featured superinfection exclusion genes like *gp17* and *sieB*, known to protect from superinfection of *p22*-like and lamboid phages (Berngruber et al., 2010; Folimonova, 2012; Monteiro et al., 2019) and *sodC*, a virulence factor which protects *S. enterica* from oxidative stress from macrophages during intracellular infection (Ammendola et al., 2005; Figueroa-Bossi & Bossi, 1999; Golubeva & Slauch, 2006). This could explain the relationship observed of *Enterobacteria phage ST104* with *Lederbergvirus p22* and *Phage Gifsy-2* with *Phage Gifsy-1* and *Salmonella phage Fels-1* (Wahl et al., 2019; Sattar et al., 2023).

Studies focusing on *S.* Infantis with a similar approach to ours have reported 4 to 8 prophages per genome (Cohen et al., 2020; Gymoese et al., 2019), while for other serovars, the median number was 5 prophages per genome (Mottawea et al., 2018). In comparison, our study detected a lower number of prophages in our genomes. This difference may be attributed to the contiguity of the genomes used, as ours were not contiguous enough, despite adapting our assembly protocol to ensure less fragmented prophage genomes and more comprehensive results. This circumstance represented a limitation of our study which can be solved by using

more contiguous genomes in future studies. This could yield similar results to the findings by Cohen et al. (2020) and Gymoese et al. (2019) where complete gap-free genome sequences were obtained using both short and long read sequencing pipelines for prophage identification.

Although we only identified the *terL* gene as the shared gene among our four prophages and other related phages, it proved sufficient to construct a maximum likelihood tree and validate the accuracy of our prophage identification. The *terL* gene is recognized for its utility in phage group classification due to its conservation, which arises from its crucial role in correct phage DNA length packaging (Wangchuk et al., 2021). While other phage genes encoding essential proteins, such as phage integrases and capsid proteins, are also valuable for distinguishing between phage groups (Dion et al., 2020; Switt et al., 2015), the prevalence of recombination events and high mutation rates currently hinders the establishment of clear phylogenetic relationships (Turner et al., 2021). Viruses, in general, lack a distinct common ancestor. To further affirm the precision of our prophage identification protocol, we constructed another maximum likelihood tree using complete genomes. The observed relationships were different, especially for *Lederbergviruses*, which encompassed a larger cluster, the reduction of the *Gifsy*-like cluster and the dillucidation of *Peduovirus pro483* and *fiAA91ss* relationships. This could be due to the larger component of the genome included which established a more robust classification evidenced by higher bootstrap values. Similar approaches have been employed using Hidden Markov Models (HMM) to classify prophages using the similarity between certain proteins or sequences (Gauthier et al., 2022; Mohammed et al., 2023). We did not use this approach because our objective was to obtain an estimation of the identification method.

Finally, we want to acknowledge that the study of prophages in *Salmonella* is crucial due to their impact on the host's adaptation to the environment and susceptibility to other infecting phages (Howard-Varona et al., 2017; Hu et al., 2021). Previous studies have isolated novel phages through different approaches, with prophage induction being the most common (Shin et

al., 2014; Segall et al., 2019; Lakshminarasimhan, 2022). However, this approach is not suitable for our detected prophages since they lack essential genes for prophage excision and the following bacterial lysis and our *S. enterica* genomes exhibited resistance to ciprofloxacin (Mejía et al, 2020), an antimicrobial agent frequently used for prophage induction due to its effect on bacterial replication (Devos et al., 2017; Segall et al., 2019). Despite these limitations, there are other approaches focused on the ability of some prophages to influence the infection of specific lytic phages by protecting their host from superinfection when lysogenized (Das & Jha, 2020; Owen et al., 2020, 2021), as also observed in this study in *S.* Typhimurium harboring *Enterobacteria phage ST104* and *Phage Gifsy-2*.

To the best of our knowledge, this study marks the first attempt to utilize an *in-silico* approach, examining the impact of prophages on phage infection dynamics and their implications for *S.* Infantis emergence in Ecuador. This is particularly relevant for MDR *S.* Infantis, given its increasing prevalence worldwide, especially in association with poultry farms and the presence of the pESI-like megaplasmids which enhance its adaptation and success in these environments (Donado-Godoy et al., 2015; Mejía et al., 2020; Vinueza-Burgos et al., 2019; Voss-Rech et al., 2015). In the country, previous studies focused on prophages have primarily explored antimicrobial resistance (Burnett et al., 2021) and served as genotyping tools (Corrales-Martínez et al., 2023). Data about prophages circulating in *Salmonella* strains of interest are crucial for the development of prevention policies in the epidemiologic surveillance and studies regarding the biology of *Salmonella* in different environments.

Conclusions

Our study was centered on the identification of prophages in *S.* Infantis, particularly those associated with poultry farms, and exhibiting resistance to multiple antimicrobials. Despite challenges in genome contiguity affecting prophage identification, we identified 4 prophages,

with *Peduovirus pro483* being the most prevalent. While some prophages showed evidence of adaptation, loss of functions, or possible recombination, their influence on host physiology remains unclear, having detected this characteristic only in 2 prophages within a *S.* Typhimurium genome. Our findings underscore the need for more contiguous genomes, experimental validation, and focused research on prophages, especially in the context of the rising prevalence of multidrug-resistant *S.* Infantis, emphasizing their potential impact on phage therapy efficacy.

REFERENCES

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. Journal of Molecular Biology, 215(3), 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2

Ammendola, S., Ajello, M., Pasquali, P., Kroll, J. S., Langford, P. R., Rotilio, G., Valenti, P., & Battistoni, A. (2005). Differential contribution of sodC1 and sodC2 to intracellular survival and pathogenicity of Salmonella enterica serovar Choleraesuis. Microbes and Infection, 7(4), 698–707. https://doi.org/10.1016/j.micinf.2005.01.005

Bachmann, N. L., Petty, N. K., Ben Zakour, N. L., Szubert, J. M., Savill, J., & Beatson, S. A. (2014). Genome analysis and CRISPR typing of Salmonella enterica serovar Virchow. BMC Genomics, 15(1), 389. https://doi.org/10.1186/1471-2164-15-389

Berngruber, T. W., Weissing, F. J., & Gandon, S. (2010). Inhibition of Superinfection and the Evolution of Viral Latency. Journal of Virology, 84(19), 10200–10208. https://doi.org/10.1128/jvi.00865-10

Bobay, L.-M., Touchon, M., & Rocha, E. P. C. (2014). Pervasive domestication of defective prophages by bacteria. Proceedings of the National Academy of Sciences, 111(33), 12127–12132. https://doi.org/10.1073/pnas.1405336111

Boyle, E. C., Bishop, J. L., Grassl, G. A., & Finlay, B. B. (2007). Salmonella: From Pathogenesis to Therapeutics. In Journal of Bacteriology (Vol. 189, Issue 5, pp. 1489–1495).

Braetz, S., Schwerk, P., Figueroa-Bossi, N., Tedin, K., & Fulde, M. (2023). Prophage Gifsy-1 Induction in Salmonella enterica Serovar Typhimurium Reduces Persister Cell Formation after Ciprofloxacin Exposure. Microbiology Spectrum, 11(4), e01874-23. https://doi.org/10.1128/spectrum.01874-23

Burnett, E., Ishida, M., de Janon, S., Naushad, S., Duceppe, M.-O., Gao, R., Jardim, A., Chen, J. C., Tagg, K. A., Ogunremi, D., & Vinueza-Burgos, C. (2021). Whole-Genome Sequencing Reveals the Presence of the blaCTX-M-65 Gene in Extended-Spectrum β-Lactamase-Producing and Multi-Drug-Resistant Clones of Salmonella Serovar Infantis Isolated from Broiler Chicken Environments in the Galapagos Islands. Antibiotics, 10(3), Article 3. https://doi.org/10.3390/antibiotics10030267

Calarga, A. P., Gontijo, M. T. P., de Almeida, L. G. P., de Vasconcelos, A. T. R., Nascimento, L. C., de Moraes Barbosa, T. M. C., de Carvalho Perri, T. M., Dos Santos, S. R., Tiba-Casas, M. R., & Marques, E. G. L. (2022). Antimicrobial resistance and genetic background of non-typhoidal Salmonella enterica strains isolated from human infections in São Paulo, Brazil (2000–2019). Brazilian Journal of Microbiology, 53(3), 1249–1262.

CDC. (2022, July 15). The biggest antibiotic-resistant threats in the U.S. Centers for Disease Control and Prevention. https://www.cdc.gov/drugresistance/biggest-threats.html

Cenens, W., Makumi, A., Govers, S. K., Lavigne, R., & Aertsen, A. (2016). Viral Transmission Dynamics at Single-Cell Resolution Reveal Transiently Immune Subpopulations Caused by a Carrier State Association. PLOS Genetics, 11(12), 1–19. https://doi.org/10.1371/journal.pgen.1005770

Coburn, B., Grassl, G. A., & Finlay, B. B. (2007). Salmonella, the host and disease: A brief review. Immunology & Cell Biology, 85(2), 112–118. https://doi.org/10.1038/sj.icb.7100007

Cohen, E., Rahav, G., & Gal-Mor, O. (2020). Genome Sequence of an Emerging Salmonella enterica Serovar Infantis and Genomic Comparison with Other S. Infantis Strains. Genome Biology and Evolution, 12(3), 223–228. https://doi.org/10.1093/gbe/evaa048

Corrales-Martínez, J., Jaramillo, K., Satán, C., Tadesse, D., Villavicencio, F., Sánchez-Gavilanes, L., Rivadeneira-Martínez, B., Balcázar, J. L., & Calero-Cáceres, W. (2023).

Genomic Characterization of a Who Critical Priority Isolate Enterobacter Kobei Harboring Oxa-10, Kpc-2, and Ctx-M-12 Recovered from a Water Irrigation Channel in Ecuador (SSRN Scholarly Paper 4610213). https://doi.org/10.2139/ssrn.4610213

Coudert, E., Gehant, S., de Castro, E., Pozzato, M., Baratin, D., Neto, T., Sigrist, C. J. A., Redaschi, N., Bridge, A., & The UniProt Consortium. (2023). Annotation of biologically relevant ligands in UniProtKB using ChEBI. Bioinformatics, 39(1), btac793. https://doi.org/10.1093/bioinformatics/btac793

Cui, L., Wang, X., Zhao, Y., Peng, Z., Gao, P., Cao, Z., Feng, J., Zhang, F., Guo, K., Wu, M., Chen, H., & Dai, M. (2021). Virulence Comparison of Salmonella enterica Subsp. Enterica Isolates from Chicken and Whole Genome Analysis of the High Virulent Strain S. Enteritidis 211. Microorganisms, 9(11), Article 11. https://doi.org/10.3390/microorganisms9112239

D'Alessandro, B., Perez Escanda, V., Balestrazzi, L., Iriarte, A., Pickard, D., Yim, L., Chabalgoity, J. A., & Betancor, L. (2018). A novel prophage identified in strains from Salmonella enterica serovar Enteritidis is a phylogenetic signature of the lineage ST-1974. Microbial Genomics, 4(3), e000161.

Darling, A. E., Mau, B., & Perna, N. T. (2010). progressiveMauve: Multiple Genome Alignment with Gene Gain, Loss and Rearrangement. PLoS ONE, 5(6), e11147. https://doi.org/10.1371/journal.pone.0011147

Das, J., & Jha, G. (2020). Bacteriophages: Co-evolving with bacteria to provide useful traits. Polymorphism, 4, 72–82.

Devos, S., Van Putte, W., Vitse, J., Van Driessche, G., Stremersch, S., Van Den Broek, W., Raemdonck, K., Braeckmans, K., Stahlberg, H., Kudryashev, M., Savvides, S. N., & Devreese, B. (2017). Membrane vesicle secretion and prophage induction in multidrug-resistant

Stenotrophomonas maltophilia in response to ciprofloxacin stress. Environmental Microbiology, 19(10), 3930–3937. https://doi.org/10.1111/1462-2920.13793

Dion, M. B., Oechslin, F., & Moineau, S. (2020). Phage diversity, genomics and phylogeny. Nature Reviews Microbiology, 18(3), Article 3. https://doi.org/10.1038/s41579-019-0311-5

Donado-Godoy, P., Byrne, B. A., León, Castellanos, R., Vanegas, C., Coral, A., Arevalo, A., Clavijo, V., Vargas, M., Romero Zuñiga, J. J., Tafur, M., Pérez-Gutierrez, E., & Smith, W. A. (2015). Prevalence, Resistance Patterns, and Risk Factors for Antimicrobial Resistance in Bacteria from Retail Chicken Meat in Colombia. Journal of Food Protection, 78(4), 751–759. https://doi.org/10.4315/0362-028X.JFP-14-349

Edgar, R. C. (2004). MUSCLE: Multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Research, 32(5), 1792–1797. https://doi.org/10.1093/nar/gkh340

Emmanuel, I. E. (2021). Challenges of Phage Therapy as a Strategic Tool for the Control of Salmonella Kentucky and Repertoire of Antibiotic Resistance Genes in Africa. In Bacteriophages in Therapeutics. IntechOpen. https://doi.org/10.5772/intechopen.95329

European Centre for Disease Prevention and Control. (2022). The European Union One Health 2021 Zoonoses Report. EFSA Journal, 20(12), e07666. https://doi.org/10.2903/j.efsa.2022.7666

Farahmandzad, F., Lanjanian, H., Arefian, E., & Kavousi, K. (2022). Bacteriophages in Gut Microbiome Interactions in Patients with Inflammatory Bowel Disease: Differences by Sex and The Disease Type [Preprint]. In Review. https://doi.org/10.21203/rs.3.rs-1240997/v1

Ferrari, R. G., Rosario, D. K. A., Cunha-Neto, A., Mano, S. B., Figueiredo, E. E. S., & Conte-Junior, C. A. (2019). Worldwide Epidemiology of Salmonella Serovars in Animal-Based

Foods: A Meta-analysis. In Applied and Environmental Microbiology (Vol. 85, Issue 14, pp. e00591-19).

Figueroa-Bossi, N., & Bossi, L. (1999). Inducible prophages contribute to Salmonella virulence in mice. Molecular Microbiology, 33(1), 167–176. https://doi.org/10.1046/j.1365-2958.1999.01461.x

Folimonova, S. Y. (2012). Superinfection Exclusion Is an Active Virus-Controlled Function That Requires a Specific Viral Protein. Journal of Virology, 86(10), 5554–5561. https://doi.org/10.1128/jvi.00310-12

Fong, K., Lu, Y. T., Brenner, T., Falardeau, J., & Wang, S. (2022). Prophage Diversity Across Salmonella and Verotoxin-Producing Escherichia coli in Agricultural Niches of British Columbia, Canada. Frontiers in Microbiology, 13. https://www.frontiersin.org/articles/10.3389/fmicb.2022.853703

Gao, R., Naushad, S., Moineau, S., Levesque, R., Goodridge, L., & Ogunremi, D. (2020). Comparative genomic analysis of 142 bacteriophages infecting Salmonella enterica subsp. Enterica. BMC Genomics, 21(1), 374. https://doi.org/10.1186/s12864-020-6765-z

Gauthier, C. H., Abad, L., Venbakkam, A. K., Malnak, J., Russell, D. A., & Hatfull, G. F. (2022). DEPhT: A novel approach for efficient prophage discovery and precise extraction. Nucleic Acids Research, 50(13), e75. https://doi.org/10.1093/nar/gkac273

Golubeva, Y. A., & Slauch, J. M. (2006). Salmonella enterica Serovar Typhimurium Periplasmic Superoxide Dismutase SodCI Is a Member of the PhoPQ Regulon and Is Induced in Macrophages. Journal of Bacteriology, 188(22), 7853–7861. https://doi.org/10.1128/JB.00706-06

Guerrero, T., Erazo, E., Bayas-Rea, R. D. L. Á., & Zapata Mena, S. (2021). Systematic Review: Nontyphoidal Salmonella in Food from Latin America. Foodborne Pathogens and Disease, 19. https://doi.org/10.1089/fpd.2020.2925

Gymoese, P., Kiil, K., Torpdahl, M., Østerlund, M. T., Sørensen, G., Olsen, J. E., Nielsen, E. M., & Litrup, E. (2019). WGS based study of the population structure of Salmonella enterica serovar Infantis. BMC Genomics, 20(1), 870. https://doi.org/10.1186/s12864-019-6260-6

Ha, A. D., & Denver, D. R. (2018). Comparative Genomic Analysis of 130 Bacteriophages Infecting Bacteria in the Genus Pseudomonas. Frontiers in Microbiology, 9, 1456. https://doi.org/10.3389/fmicb.2018.01456

Hall, J. P. J., Botelho, J., Cazares, A., & Baltrus, D. A. (2022). What makes a megaplasmid? Philosophical Transactions of the Royal Society B: Biological Sciences, 377(1842), 20200472. https://doi.org/10.1098/rstb.2020.0472

Hatfull, G. F., Jacobs-Sera, D., Lawrence, J. G., Pope, W. H., Russell, D. A., Ko, C.-C., Weber, R. J., Patel, M. C., Germane, K. L., Edgar, R. H., Hoyte, N. N., Bowman, C. A., Tantoco, A. T., Paladin, E. C., Myers, M. S., Smith, A. L., Grace, M. S., Pham, T. T., O'Brien, M. B., … Hendrix, R. W. (2010). Comparative Genomic Analysis of 60 Mycobacteriophage Genomes: Genome Clustering, Gene Acquisition, and Gene Size. Journal of Molecular Biology, 397(1), 119–143. https://doi.org/10.1016/j.jmb.2010.01.011

Henrot, C., & Petit, M.-A. (2022). Signals triggering prophage induction in the gut microbiota. Molecular Microbiology, 118(5), 494–502. https://doi.org/10.1111/mmi.14983

Herrero-Fresno, A., Espinel, I. C., Spiegelhauer, M. R., Guerra, P. R., Andersen, K. W., & Olsen, J. E. (2017). The Homolog of the Gene bstA of the BTP1 Phage from Salmonella enterica Serovar Typhimurium ST313 Is an Antivirulence Gene in Salmonella enterica Serovar Dublin. Infection and Immunity, 86(1), e00784-17. https://doi.org/10.1128/IAI.00784-17

Hiyoshi, H., Tiffany, C. R., Bronner, D. N., & Bäumler, A. J. (2018). Typhoidal Salmonella serovars: Ecological opportunity and the evolution of a new pathovar. FEMS Microbiology Reviews, 42(4), 527–541. https://doi.org/10.1093/femsre/fuy024

Howard-Varona, C., Hargreaves, K. R., Abedon, S. T., & Sullivan, M. B. (2017). Lysogeny in nature: Mechanisms, impact and ecology of temperate phages. The ISME Journal, 11(7), Article 7. https://doi.org/10.1038/ismej.2017.16

Hu, J., Ye, H., Wang, S., Wang, J., & Han, D. (2021). Prophage activation in the intestine: Insights into functions and possible applications. Frontiers in Microbiology, 12, 785634.

Hyatt, D., Chen, G.-L., LoCascio, P. F., Land, M. L., Larimer, F. W., & Hauser, L. J. (2010). Prodigal: Prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics, 11(1), 119. https://doi.org/10.1186/1471-2105-11-119

Ibarra, J. A., & Steele-Mortimer, O. (2009). Salmonella – the ultimate insider. Salmonella virulence factors that modulate intracellular survival. Cellular Microbiology, 11(11), 1579–1586. https://doi.org/10.1111/j.1462-5822.2009.01368.x

Jackman, S. D., Vandervalk, B. P., Mohamadi, H., Chu, J., Yeo, S., Hammond, S. A., Jahesh, G., Khan, H., Coombe, L., Warren, R. L., & Birol, I. (2017). ABySS 2.0: Resource-efficient assembly of large genomes using a Bloom filter. Genome Research, 27(5), 768–777. https://doi.org/10.1101/gr.214346.116

Jajere, S. M. (2019). A review of Salmonella enterica with particular focus on the pathogenicity and virulence factors, host specificity and antimicrobial resistance including multidrug resistance. Veterinary World, 12(4), 504–521. https://doi.org/10.14202/vetworld.2019.504-521

Jonge, P. A. de, Nobrega, F. L., Brouns, S. J. J., & Dutilh, B. E. (2019). Molecular and Evolutionary Determinants of Bacteriophage Host Range. Trends in Microbiology, 27(1), 51–63. https://doi.org/10.1016/j.tim.2018.08.006

Katoh, K., Rozewicki, J., & Yamada, K. D. (2019). MAFFT online service: Multiple sequence alignment, interactive sequence choice and visualization. Briefings in Bioinformatics, 20(4), 1160–1166. https://doi.org/10.1093/bib/bbx108

Kaur, J., & Jain, S. K. (2012). Role of antigens and virulence factors of Salmonella enterica serovar Typhi in its pathogenesis. Microbiological Research, 167(4), 199–210. https://doi.org/10.1016/j.micres.2011.08.001

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P., & Drummond, A. (2012). Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics (Oxford, England), 28(12), 1647–1649. https://doi.org/10.1093/bioinformatics/bts199

Kipper, D., Mascitti, A. K., De Carli, S., Carneiro, A. M., Streck, A. F., Fonseca, A. S. K., Ikuta, N., & Lunge, V. R. (2022). Emergence, Dissemination and Antimicrobial Resistance of the Main Poultry-Associated Salmonella Serovars in Brazil. Veterinary Sciences, 9(8), 405.

Kirkwood, M., Vohra, P., Bawn, M., Thilliez, G., Pye, H., Tanner, J., Chintoan-Uta, C., Branchu, P., Petrovska, L., Dallman, T., Hall, N., Stevens, M. P., & Kingsley, R. A. (2021). Ecological niche adaptation of Salmonella Typhimurium U288 is associated with altered pathogenicity and reduced zoonotic potential. Communications Biology, 4(1), Article 1. https://doi.org/10.1038/s42003-021-02013-4

Kraushaar, B., Hammerl, J. A., Kienöl, M., Heinig, M. L., Sperling, N., Dinh Thanh, M., Reetz, J., Jäckel, C., Fetsch, A., & Hertwig, S. (2017). Acquisition of virulence factors in livestock-

associated MRSA: Lysogenic conversion of CC398 strains by virulence gene-containing phages. Scientific Reports, 7(1), Article 1. https://doi.org/10.1038/s41598-017-02175-4

Kumar, S., Stecher, G., Li, M., Knyaz, C., & Tamura, K. (2018). MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. Molecular Biology and Evolution, 35(6), 1547–1549. https://doi.org/10.1093/molbev/msy096

Kurasz, J. E., Crawford, M. C., Porwollik, S., Gregory, O., Tadlock, K. R., Balding, E. C., Weinert, E. E., McClelland, M., & Karls, A. C. (2023). Strain-Specific Gifsy-1 Prophage Genes Are Determinants for Expression of the RNA Repair Operon during the SOS Response in Salmonella enterica Serovar Typhimurium. Journal of Bacteriology, 205(1), e00262-22. https://doi.org/10.1128/jb.00262-22

Lakshminarasimhan, A. (2022). Prophage induction therapy: Activation of the lytic phase in prophages for the elimination of pathogenic bacteria. Medical Hypotheses, 169, 110980. https://doi.org/10.1016/j.mehy.2022.110980

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. Nature Methods, 9(4), 357–359. https://doi.org/10.1038/nmeth.1923

Lapierre, L., Cornejo, J., Zavala, S., Galarce, N., Sánchez, F., Benavides, M. B., Guzmán, M., & Sáenz, L. (2020). Phenotypic and Genotypic Characterization of Virulence Factors and Susceptibility to Antibiotics in Salmonella Infantis Strains Isolated from Chicken Meat: First Findings in Chile. Animals : An Open Access Journal from MDPI, 10(6), 1049. https://doi.org/10.3390/ani10061049

Manni, M., Berkeley, M. R., Seppey, M., Simão, F. A., & Zdobnov, E. M. (2021). BUSCO Update: Novel and Streamlined Workflows along with Broader and Deeper Phylogenetic Coverage for Scoring of Eukaryotic, Prokaryotic, and Viral Genomes. Molecular Biology and Evolution, 38(10), 4647–4654. https://doi.org/10.1093/molbev/msab199

McMillan, E. A., Jackson, C. R., & Frye, J. G. (2020). Transferable Plasmids of Salmonella enterica Associated With Antibiotic Resistance Genes. Frontiers in Microbiology, 11. https://www.frontiersin.org/articles/10.3389/fmicb.2020.562181

Medeiros, M. A. N., Oliveira, D. C. N. D., Rodrigues, D. D. P., & Freitas, D. R. C. D. (2011). Prevalence and antimicrobial resistance of Salmonella in chicken carcasses at retail in 15 Brazilian cities. Revista Panamericana de Salud Pública, 30(6), 555–560. https://doi.org/10.1590/S1020-49892011001200010

Mejía, L., Medina, J. L., Bayas, R., Salazar, C. S., Villavicencio, F., Zapata, S., Matheu, J., Wagenaar, J. A., González-Candelas, F., & Vinueza-Burgos, C. (2020). Genomic Epidemiology of Salmonella Infantis in Ecuador: From Poultry Farms to Human Infections. Frontiers in Veterinary Science, 7. https://www.frontiersin.org/articles/10.3389/fvets.2020.547891

Mohammed, M., Casjens, S. R., Millard, A. D., Harrison, C., Gannon, L., & Chattaway, M. A. (2023). Genomic analysis of Anderson typing phages of Salmonella Typhimrium: Towards understanding the basis of bacteria-phage interaction. Scientific Reports, 13(1), Article 1. https://doi.org/10.1038/s41598-023-37307-6

Monteiro, R., Pires, D. P., Costa, A. R., & Azeredo, J. (2019). Phage Therapy: Going Temperate? Trends in Microbiology, 27(4), 368–378. https://doi.org/10.1016/j.tim.2018.10.008

Montoro-Dasi, L., Lorenzo-Rebenaque, L., Marco-Fuertes, A., Vega, S., & Marin, C. (2023). Holistic Strategies to Control Salmonella Infantis: An Emerging Challenge in the European Broiler Sector. Microorganisms, 11(7), Article 7. https://doi.org/10.3390/microorganisms11071765

Moreno Switt, A. (2013). Genomic Characterization Of Salmonella Free-Living Phages, Plasmids And Chromosomally Inserted Mobile Elements. https://hdl.handle.net/1813/33807

Moreno Switt, A. I., Orsi, R. H., den Bakker, H. C., Vongkamjan, K., Altier, C., & Wiedmann, M. (2013). Genomic characterization provides new insight into Salmonella phage diversity. BMC Genomics, 14(1), 481. https://doi.org/10.1186/1471-2164-14-481

Morgulis, A., Coulouris, G., Raytselis, Y., Madden, T. L., Agarwala, R., & Schäffer, A. A. (2008). Database indexing for production MegaBLAST searches. Bioinformatics (Oxford, England), 24(16), 1757–1764. https://doi.org/10.1093/bioinformatics/btn322

Mottawea, W., Duceppe, M.-O., Dupras, A. A., Usongo, V., Jeukens, J., Freschi, L., Emond-Rheault, J.-G., Hamel, J., Kukavica-Ibrulj, I., Boyle, B., Gill, A., Burnett, E., Franz, E., Arya, G., Weadge, J. T., Gruenheid, S., Wiedmann, M., Huang, H., Daigle, F., … Ogunremi, D. (2018). Salmonella enterica Prophage Sequence Profiles Reflect Genome Diversity and Can Be Used for High Discrimination Subtyping. Frontiers in Microbiology, 9, 836. https://doi.org/10.3389/fmicb.2018.00836

Moulana, Z., & Asgharpour, F. (2022). Prevalence and Antimicrobial Resistance of Salmonella Enterica Serovar Infantis Isolates from Poultry: A review. Poultry Science Journal, 10(1), 13–26. https://doi.org/10.22069/psj.2022.19783.1759

Moura de Sousa, J. A., Pfeifer, E., Touchon, M., & Rocha, E. P. C. (2021). Causes and Consequences of Bacteriophage Diversification via Genetic Exchanges across Lifestyles and Bacterial Taxa. Molecular Biology and Evolution, 38(6), 2497–2512. https://doi.org/10.1093/molbev/msab044

Nayfach, S., Páez-Espino, D., Call, L., Low, S. J., Sberro, H., Ivanova, N. N., Proal, A. D., Fischbach, M. A., Bhatt, A. S., Hugenholtz, P., & Kyrpides, N. C. (2021). Metagenomic compendium of 189,680 DNA viruses from the human gut microbiome. Nature Microbiology, 6(7), Article 7. https://doi.org/10.1038/s41564-021-00928-6

Nishijima, S., Nagata, N., Kiguchi, Y., Kojima, Y., Miyoshi-Akiyama, T., Kimura, M., Ohsugi, M., Ueki, K., Oka, S., Mizokami, M., Itoi, T., Kawai, T., Uemura, N., & Hattori, M. (2022). Extensive gut virome variation and its associations with host and environmental factors in a population-level cohort. Nature Communications, 13(1), Article 1. https://doi.org/10.1038/s41467-022-32832-w

Owen, S. V., Canals, R., Wenner, N., Hammarlöf, D. L., Kröger, C., & Hinton, J. C. D. (2020). A window into lysogeny: Revealing temperate phage biology with transcriptomics. Microbial Genomics, 6(2), e000330. https://doi.org/10.1099/mgen.0.000330

Owen, S. V., Wenner, N., Dulberger, C. L., Rodwell, E. V., Bowers-Barnard, A., Quinones-Olvera, N., Rigden, D. J., Rubin, E. J., Garner, E. C., Baym, M., & Hinton, J. C. D. (2021). Prophages encode phage-defense systems with cognate self-immunity. Cell Host & Microbe, 29(11), 1620-1633.e8. https://doi.org/10.1016/j.chom.2021.09.002

Parker, C. T., Huynh, S., Alexander, A., Oliver, A. S., & Cooper, K. K. (2021). Genomic Characterization of Salmonella typhimurium DT104 Strains Associated with Cattle and Beef Products. Pathogens, 10(5), Article 5. https://doi.org/10.3390/pathogens10050529

Petrovska, L., Mather, A. E., AbuOun, M., Branchu, P., Harris, S. R., Connor, T., Hopkins, K. L., Underwood, A., Lettini, A. A., & Page, A. (2016). Microevolution of monophasic Salmonella typhimurium during epidemic, United Kingdom, 2005–2010. Emerging Infectious Diseases, 22(4), 617.

Pires, D. P., Costa, A. R., Pinto, G., Meneses, L., & Azeredo, J. (2020). Current challenges and future opportunities of phage therapy. FEMS Microbiology Reviews, 44(6), 684–700. https://doi.org/10.1093/femsre/fuaa017

Popa, G. L., & Papa, M. I. (2021). Salmonella spp. Infection—A continuous threat worldwide. Germs, 11(1), 88–96. https://doi.org/10.18683/germs.2021.1244

Pope, W. H., Jacobs-Sera, D., Russell, D. A., Peebles, C. L., Al-Atrache, Z., Alcoser, T. A., Alexander, L. M., Alfano, M. B., Alford, S. T., Amy, N. E., Anderson, M. D., Anderson, A. G., Ang, A. A. S., Ares, M., Barber, A. J., Barker, L. P., Barrett, J. M., Barshop, W. D., Bauerle, C. M., … Hatfull, G. F. (2011). Expanding the Diversity of Mycobacteriophages: Insights into Genome Architecture and Evolution. PLoS ONE, 6(1), e16329. https://doi.org/10.1371/journal.pone.0016329

Prjibelski, A., Antipov, D., Meleshko, D., Lapidus, A., & Korobeynikov, A. (2020). Using SPAdes De Novo Assembler. Current Protocols in Bioinformatics, 70(1), e102. https://doi.org/10.1002/cpbi.102

Qi, Z., Meng, B., Wei, X., Li, X., Peng, H., Li, Y., Feng, Q., Huang, Y., Zhang, Q., Xu, X., Zhao, H., Yang, X., Wang, C., & Zhao, X. (2022). Identification and characterization of P2-like bacteriophages of Yersinia pestis. Virus Research, 322, 198934. https://doi.org/10.1016/j.virusres.2022.198934

Rusconi, B., Sanjar, F., Koenig, S. S. K., Mammel, M. K., Tarr, P. I., & Eppinger, M. (2016). Whole Genome Sequencing for Genomics-Guided Investigations of Escherichia coli O157:H7 Outbreaks. Frontiers in Microbiology, 7. https://www.frontiersin.org/articles/10.3389/fmicb.2016.00985

Sattar, S., Carbone, V., Basharat, Z., Bostan, N., Tariq, A., Javed, S., Ullah, I., Khanum, S., Yaqoob, A., Ahmed, I., Abbas, M. A., Shah, S. T. A., & Altermann, E. (2023). Isolation and

Molecular Characterization of a Novel Podo-virus BIS08P22 having a Unique Terminase Enzyme Representing a new Genus in Class Caudoviricetes (2023100094). Preprints. https://doi.org/10.20944/preprints202310.0094.v1

Sayers, E. W., Bolton, E. E., Brister, J. R., Canese, K., Chan, J., Comeau, D. C., Connor, R., Funk, K., Kelly, C., Kim, S., Madej, T., Marchler-Bauer, A., Lanczycki, C., Lathrop, S., Lu, Z., Thibaud-Nissen, F., Murphy, T., Phan, L., Skripchenko, Y., … Sherry, S. T. (2022). Database resources of the national center for biotechnology information. Nucleic Acids Research, 50(D1), D20–D26. https://doi.org/10.1093/nar/gkab1112

Schwengers, O., Jelonek, L., Dieckmann, M. A., Beyvers, S., Blom, J., & Goesmann, A. (2021). Bakta: Rapid and standardized annotation of bacterial genomes via alignment-free sequence identification. Microbial Genomics, 7(11), 000685. https://doi.org/10.1099/mgen.0.000685

Segall, A. M., Roach, D. R., & Strathdee, S. A. (2019). Stronger together? Perspectives on phage-antibiotic synergy in clinical applications of phage therapy. Current Opinion in Microbiology, 51, 46–50. https://doi.org/10.1016/j.mib.2019.03.005

Shah, D. H., Paul, N. C., Sischo, W. C., Crespo, R., & Guard, J. (2017). Population dynamics and antimicrobial resistance of the most prevalent poultry-associated Salmonella serotypes. Poultry Science, 96(3), 687–702. https://doi.org/10.3382/ps/pew342

Shin, H., Lee, J.-H., Yoon, H., Kang, D.-H., & Ryu, S. (2014). Genomic Investigation of Lysogen Formation and Host Lysis Systems of the Salmonella Temperate Bacteriophage SPN9CC. Applied and Environmental Microbiology, 80(1), 374–384. https://doi.org/10.1128/AEM.02279-13

Shridhar, P. B., Worley, J. N., Gao, X., Yang, X., Noll, L. W., Shi, X., Bai, J., Meng, J., & Nagaraja, T. G. (2019). Analysis of virulence potential of Escherichia coli O145 isolated from

cattle feces and hide samples based on whole genome sequencing. PLOS ONE, 14(11), e0225057. https://doi.org/10.1371/journal.pone.0225057

Silva, C., Calva, E., Puente, J. L., Zaidi, M. B., & Vinuesa, P. (2016). Complete Genome Sequence of Salmonella enterica Serovar Typhimurium Strain SO2 (Sequence Type 302) Isolated from an Asymptomatic Child in Mexico. Genome Announcements, 4(2), 10.1128/genomea.00253-16. https://doi.org/10.1128/genomea.00253-16

Svahn, A. J., Suster, C. J. E., Chang, S. L., Rockett, R. J., Sim, E. M., Cliff, O. M., Wang, Q., Arnott, A., Ramsperger, M., Sorrell, T. C., Sintchenko, V., & Prokopenko, M. (2023). Pangenome Analysis of a Salmonella Enteritidis Population Links a Major Outbreak to a Gifsy-1-Like Prophage Containing Anti-Inflammatory Gene gogB. Microbiology Spectrum, 11(2), e02791-22. https://doi.org/10.1128/spectrum.02791-22

Switt, A. I. M., Sulakvelidze, A., Wiedmann, M., Kropinski, A. M., Wishart, D. S., Poppe, C., & Liang, Y. (2015). Salmonella Phages and Prophages: Genomics, Taxonomy, and Applied Aspects. In H. Schatten & A. Eisenstark (Eds.), Salmonella (Vol. 1225, pp. 237–287). Springer New York. https://doi.org/10.1007/978-1-4939-1625-2_15

Tisza, M. J., & Buck, C. B. (2021). A catalog of tens of thousands of viruses from human metagenomes reveals hidden associations with chronic diseases. Proceedings of the National Academy of Sciences, 118(23), e2023202118. https://doi.org/10.1073/pnas.2023202118

Trofeit, L., Sattler, E., Künz, J., & Hilbert, F. (2023). Salmonella Prophages, Their Propagation, Host Specificity and Antimicrobial Resistance Gene Transduction. Antibiotics, 12(3), Article 3. https://doi.org/10.3390/antibiotics12030595

Turner, D., Kropinski, A. M., & Adriaenssens, E. M. (2021). A Roadmap for Genome-Based Phage Taxonomy. Viruses, 13(3), 506. https://doi.org/10.3390/v13030506

Turner, D., Shkoporov, A. N., Lood, C., Millard, A. D., Dutilh, B. E., Alfenas-Zerbini, P., van Zyl, L. J., Aziz, R. K., Oksanen, H. M., Poranen, M. M., Kropinski, A. M., Barylski, J., Brister, J. R., Chanisvili, N., Edwards, R. A., Enault, F., Gillis, A., Knezevic, P., Krupovic, M., … Adriaenssens, E. M. (2023). Abolishment of morphology-based taxa and change to binomial species names: 2022 taxonomy update of the ICTV bacterial viruses subcommittee. Archives of Virology, 168(2), 74. https://doi.org/10.1007/s00705-022-05694-2

Vallejos-Sánchez, K., Tataje-Lavanda, L., Villanueva-Pérez, D., Bendezú, J., Montalván, Á., Zimic-Peralta, M., Fernández-Sánchez, M., & Fernández-Díaz, M. (2019). Whole-Genome Sequencing of a Salmonella enterica subsp. Enterica Serovar Infantis Strain Isolated from Broiler Chicken in Peru. Microbiology Resource Announcements, 8(43), 10.1128/mra.00826-19. https://doi.org/10.1128/mra.00826-19

Van Boeckel, T. P., Brower, C., Gilbert, M., Grenfell, B. T., Levin, S. A., Robinson, T. P., Teillant, A., & Laxminarayan, R. (2015). Global trends in antimicrobial use in food animals. Proceedings of the National Academy of Sciences, 112(18), 5649–5654.

Vinueza-Burgos, C. V. (2017). Salmonella and Campylobacter in broilers at slaughter age: A possible source for carcasses contamination in Ecuador. Ghent University.

Vinueza-Burgos, C., Baquero, M., Medina, J., & De Zutter, L. (2019). Occurrence, genotypes and antimicrobial susceptibility of Salmonella collected from the broiler production chain within an integrated poultry company. International Journal of Food Microbiology, 299, 1–7. https://doi.org/10.1016/j.ijfoodmicro.2019.03.014

Voss-Rech, D., Vaz, C. S. L., Alves, L., Coldebella, A., Leão, J. A., Rodrigues, D. P., & Back, A. (2015). A temporal study of Salmonella enterica serotypes from broiler farms in Brazil. Poultry Science, 94(3), 433–441. https://doi.org/10.3382/ps/peu081

Wahl, A., Battesti, A., & Ansaldi, M. (2019). Prophages in Salmonella enterica: A driving force in reshaping the genome and physiology of their bacterial host? Molecular Microbiology, 111(2), 303–316. https://doi.org/10.1111/mmi.14167

Walker, P. J., Siddell, S. G., Lefkowitz, E. J., Mushegian, A. R., Adriaenssens, E. M., Dempsey, D. M., Dutilh, B. E., Harrach, B., Harrison, R. L., Hendrickson, R. C., Junglen, S., Knowles, N. J., Kropinski, A. M., Krupovic, M., Kuhn, J. H., Nibert, M., Orton, R. J., Rubino, L., Sabanadzovic, S., … Davison, A. J. (2020). Changes to virus taxonomy and the Statutes ratified by the International Committee on Taxonomy of Viruses (2020). Archives of Virology, 165(11), 2737–2748. https://doi.org/10.1007/s00705-020-04752-x

Wangchuk, J., Chatterjee, A., Patil, S., Madugula, S. K., & Kondabagil, K. (2021). The coevolution of large and small terminases of bacteriophages is a result of purifying selection leading to phenotypic stabilization. Virology, 564, 13–25. https://doi.org/10.1016/j.virol.2021.09.004

Weissman, J. L., Holmes, R., Barrangou, R., Moineau, S., Fagan, W. F., Levin, B., & Johnson, P. L. F. (2018). Immune loss as a driver of coexistence during host-phage coevolution. The ISME Journal, 12(2), Article 2. https://doi.org/10.1038/ismej.2017.194

Wendling, C. C. (2023). Prophage mediated control of higher order interactions—Insights from multi-level approaches. Current Opinion in Systems Biology, 35, 100469. https://doi.org/10.1016/j.coisb.2023.100469

WHO. (2017). Prioritization of pathogens to guide discovery, research and development of new antibiotics for drug-resistant bacterial infections, including tuberculosis (p. 87). World Health Organization. https://www.who.int/publications/i/item/WHO-EMP-IAU-2017.12

WHO. (2015, December 3). WHO's first ever global estimates of foodborne diseases find children under 5 account for almost one third of deaths. WHO's First Ever Global Estimates of

Foodborne Diseases Find Children under 5 Account for Almost One Third of Deaths. https://www.who.int/news/item/03-12-2015-who-s-first-ever-global-estimates-of-foodborne-diseases-find-children-under-5-account-for-almost-one-third-of-deaths

WHO. (2023, November 21). Antimicrobial resistance. https://www.who.int/news-room/fact-sheets/detail/antimicrobial-resistance

Yu, P., Mathieu, J., Yang, Y., & Alvarez, P. J. J. (2017). Suppression of Enteric Bacteria by Bacteriophages: Importance of Phage Polyvalence in the Presence of Soil Bacteria. Environmental Science & Technology, 51(9), 5270–5278. https://doi.org/10.1021/acs.est.7b00529

Żbikowska, K., Michalczuk, M., & Dolka, B. (2020). The use of bacteriophages in the poultry industry. Animals, 10(5), 872.

SUPPLEMENTARY MATERIAL INDEX

Supplementary Table 1. Details for all Salmonella enterica genomes assembled, their quality control statistics and indicators

Sample name	NCBI Code	Serovar	File Size before trimming (MB)		File Size after trimming (MB)		Assembly summary: abyss-fac										BUSCO	Assembly: SPAdes
			Forward	Reverse	Forward	Reverse	n	n:500	L50	min	N75	N50	N25	E-size	max	sum	Genome Complete (%)	Assembled File Size (MB)
Se_Q_001	CSG217	*S.* Infantis	103	105	73	74	157	62	5	515	138387	266844	470431	451299	1195276	4982149	98,4	5,1
Se_Q_002	CSG232	*S.* Typhimurium	92	93	65	66	69	28	4	545	229450	478958	647238	499866	978044	4698564	98,4	4,8
Se_Q_003	CSG252	*S.* Typhimurium	69	70	49	49	80	33	5	545	171440	406336	478270	427904	977752	4697863	98,4	4,8
Se_Q_004	CSG177	*S.* Infantis	58	60	41	42	178	79	8	523	149779	259203	340779	233524	376736	4925717	98,4	5,1
Se_Q_005	CSG175	*S.* Infantis	59	60	42	43	151	63	8	501	111884	194191	385887	228439	525933	4706473	98,4	4,8
Se_Q_006	CSG132	*S.* Infantis	63	64	44	45	157	72	9	503	111845	181160	226718	253770	719507	4914638	98,4	5
Se_Q_007	U113s	*S.* Typhimurium	57	58	40	41	178	89	10	537	83215	178330	266195	179917	333658	4955233	99,2	5,1
Se_Q_008	U114s	*S.* Enteritidis	56	57	39	40	79	39	6	542	153089	291342	375907	302264	647270	4632449	98,4	4,8
Se_Q_009	U120s	*S.* Infantis	116	118	81	82	176	77	8	515	89905	150188	474554	335336	1029493	5083069	100,0	5,2
Se_Q_010	U121s	*S.* Infantis	54	55	38	39	153	78	9	523	103540	183601	323905	222476	554236	4871727	98,4	5
Se_Q_011	U123s	*S.* Infantis	49	50	35	36	137	78	11	523	58371	129590	183859	123842	225091	3686054	90,3	3,8
Se_Q_012	U127s	*S.* Infantis	55	56	38	39	143	72	8	617	105437	192292	333656	237973	617933	4782068	100,0	4,9
Se_Q_013	U128s	*S.* Infantis	78	80	54	55	163	66	6	515	115018	201642	371030	417143	1195275	4962512	98,4	5,1
Se_Q_014	U129s	*S.* Infantis	94	96	68	69	156	72	8	515	112205	165469	438969	262897	604074	5082850	100,0	5,2
Se_Q_015	U219s	*S.* Infantis	76	78	55	55	167	75	9	515	111806	180235	324280	228602	590272	5077223	100,0	5,2
Se_Q_016	U634s	*S.* Infantis	49	50	35	36	129	63	10	515	105545	183598	237218	191325	492353	4597871	96,8	4,7
Se_Q_017	U638s	*S.* Infantis	73	74	52	53	159	71	6	502	114987	183947	525590	429418	1195182	4973292	100,0	5,1
Se_Q_018	U639s	*S.* Infantis	74	75	53	53	152	71	8	515	114952	183598	370826	250527	589934	4955057	98,4	5,1
Se_Q_019	U652s	*S.* Infantis	73	74	52	53	148	67	6	515	103469	201608	370969	411990	1195009	5001213	100,0	5,1
Se_Q_020			**D**	**D**	**D**	**D**	**D**	**D**	**D**	**D**	**D**	**D**	**D**	**D**	**D**	**D**	**D**	0
Se_Q_021	U664s	*S.* Infantis	97	98	68	69	147	132	14	542	56723	113808	187395	123342	313697	4887325	100,0	5
Se_Q_022	U666s	*S.* Infantis	62	63	44	44	166	75	8	503	111806	194184	324230	256584	719600	4949368	100,0	5,1
Se_Q_023	U669s	*S.* Infantis	64	65	46	46	152	68	6	503	103482	194191	525799	377336	1028952	4943934	100,0	5,1
Se_Q_024	U672s	*S.* Infantis	72	73	51	52	165	71	7	552	94393	183600	525862	278981	719785	4972344	100,0	5,1
Se_Q_025	U676s	*S.* Infantis	59	60	42	43	173	69	8	515	105321	180609	525731	288308	824615	4919924	100,0	5
Se_Q_026	U679s	*S.* Infantis	106	108	75	76	149	67	5	515	105545	183599	525922	439834	1192981	4941018	99,2	5,1
Se_Q_027	U682s	*S.* Infantis	88	89	62	62	156	66	6	515	124631	183687	469482	370959	1029473	4966517	100,0	5,1
Se_Q_028	U684s	*S.* Infantis	72	73	51	51	160	73	8	501	93967	183687	438783	252912	589935	4938821	99,2	5,1
Se_Q_029	U689s	*S.* Infantis	73	74	52	52	155	64	7	552	111884	183947	525921	362993	1029044	4963870	99,2	5,1
Se_Q_030	U692s	*S.* Infantis	79	81	55	56	162	67	6	512	105437	194416	525918	380846	1044953	4962358	98,4	5,1
Se_Q_031	U706s	*S.* Infantis	65	66	46	46	152	76	9	515	96311	181513	324225	208948	521940	4898829	100,0	5

Se_Q_032	U707s	*S.* Infantis	77	79	55	56	170	75	9	515	93966	180521	368507	214594	522039	4956781	100,0	5,1
Se_Q_033	U708s	*S.* Infantis	73	74	51	52	149	60	7	501	111884	194278	438876	273065	590544	4969235	100,0	5,1
Se_Q_034	U711s	*S.* Infantis	55	56	39	39	150	77	9	515	105321	183859	268385	206387	471316	4936538	100,0	5,1
Se_Q_035	U712s	*S.* Infantis	96	97	68	69	153	68	6	515	96310	201656	371360	414389	1195275	4971725	98,4	5,1
Se_Q_036	U715s	*S.* Infantis	141	144	99	101	165	63	5	515	127845	203919	469779	444181	1195182	4975527	99,2	5,1
Se_Q_037	U718s	*S.* Infantis	24	24	16	16	29	4	2	1137	1601	1750	3653	2488	3653	8141	0,0	15,0 kB
Se_Q_038	U719s	*S.* Infantis	198	200	134	136	158	65	6	515	105545	183688	444040	419315	1179805	4953029	98,4	5,1
Se_Q_039	U723s	*S.* Infantis	120	123	78	80	169	70	6	515	105321	183687	470132	371174	1029473	4966517	100,0	5,1
Se_Q_040			**D**	**D**	**D**	**D**	**D**	**D**	**D**	**D**	**D**	**D**	**D**	**D**	**D**	**D**	**D**	0
Se_Q_041	U728s	*S.* Infantis	179	182	122	124	168	66	5	507	105545	203918	525853	448781	1188867	4972216	100,0	5,1
Se_Q_042	U729s	*S.* Infantis	105	107	74	75	163	67	6	515	105321	183600	525862	381778	1029474	4965114	99,2	5,1
Se_Q_043	U758s	*S.* Infantis	92	94	64	65	161	71	8	515	103520	183600	370969	254967	604576	4947710	99,2	5,1
Se_Q_044	U759s	*S.* Infantis	97	99	67	68	155	63	6	515	148183	201638	443813	432023	1195275	4973905	99,2	5,1
Se_Q_045	U762s	*S.* Infantis	115	117	70	71	156	63	6	515	111885	203916	470180	423071	1195182	4967830	99,2	5,1
Se_Q_046	U769s	*S.* Infantis	94	95	66	67	137	64	6	515	105438	201620	444101	426266	1195275	4974672	99,2	5,1
Se_Q_047	U773s	*S.* Infantis	89	90	63	64	156	67	8	515	93967	183599	371597	253888	649759	4970622	98,4	5,1
Se_Q_048	U801s	*S.* Infantis	161	164	110	111	152	61	5	515	114988	203915	525921	451228	1187838	4956900	98,4	5,1
Se_Q_049	U804s	*S.* Infantis	142	144	99	101	171	67	5	504	105545	203915	470256	446519	1195089	4942619	98,4	5,1
Se_Q_050	U806s	*S.* Infantis	101	102	71	72	161	62	6	504	148183	201623	444039	435349	1200177	4970943	100,0	5,1
Se_Q_051	U814s	*S.* Infantis	192	195	131	133	147	65	5	515	105429	203915	525858	446007	1195273	4973872	98,4	5,1
Se_Q_052	U815s	*S.* Infantis	95	96	66	67	177	65	6	515	93904	183597	470263	370584	1044861	4925562	99,2	5,1
Se_Q_053	U818s	*S.* Infantis	114	116	81	82	168	66	6	515	89905	201627	444040	417117	1195271	4974881	99,2	5,1
Se_Q_054	U855s	*S.* Infantis	104	106	49	49	146	65	7	515	105321	183687	525901	361229	1029084	4931517	99,2	5,1
Se_Q_055	U857s	*S.* Infantis	101	103	71	72	159	66	6	515	105545	201626	371299	417759	1195089	4971699	100,0	5,1
Se_Q_056	U860s	*S.* Infantis	75	76	53	54	137	63	7	515	124631	203916	371597	284253	676384	4916076	98,4	5
Se_Q_057	U863s	*S.* Infantis	1,2 GB	1,2 GB	820	829	153	60	5	515	127741	203920	525828	454696	1195275	4971725	100,0	5,1
Se_Q_058	U865s	*S.* Infantis	148	151	105	106	166	69	6	515	92521	183687	470134	395552	1099117	4973961	99,2	5,1
Se_Q_059	U866s	*S.* Infantis	75	76	53	54	156	70	6	502	103519	201621	444101	423020	1195097	4976315	98,4	5,1
Se_Q_060	U869s	*S.* Infantis	146	148	14	14	27	5	2	543	869	2112	3543	2400	3543	7684	0,0	14,6 kB
Se_Q_061	U871s	*S.* Infantis	71	72	51	51	183	76	8	502	103515	183859	371030	311801	942896	4959522	100,0	5,1
Se_Q_062	U873s	*S.* Infantis	112	114	78	80	163	67	6	515	127731	201644	444101	420547	1191517	4976344	97,6	5,1
Se_Q_063	U875s	*S.* Enteritidis	147	150	104	106	76	30	3	618	185781	406400	1455556	672327	1455556	4773271	98,4	4,9
Se_Q_064	U878s	*S.* Infantis	82	84	58	58	169	69	7	515	96311	181222	525954	358858	1029772	4977338	99,2	5,1
Se_Q_065	U880s	*S.* Infantis	104	106	75	76	160	63	5	515	105545	203916	525868	451211	1195275	5016594	100,0	5,1
Se_Q_066	U883s	*S.* Infantis	80	82	57	58	156	63	5	515	114968	203919	525791	444421	1194908	4971129	99,2	5,1
Se_Q_067	U885s	*S.* Infantis	78	80	55	56	141	68	7	515	105545	201635	370969	328926	960342	4961154	100,0	5,1
Se_Q_068	U888s	*S.* Infantis	204	207	145	147	167	67	6	515	92521	183600	470196	418552	1145098	4973494	98,4	5,1

Se_Q_069	U889s	*S.* Infantis	59	60	42	42	149	63	10	515	114982	183687	226203	202617	510646	4723271	98,4	4,9
Se_Q_070	U890s	*S.* Infantis	79	81	56	57	170	69	8	515	113153	183687	371360	300266	915601	4972832	100,0	5,1
Se_Q_071	U895s	*S.* Infantis	61	62	43	44	142	66	7	515	124631	203915	525920	298640	743739	4928924	100,0	5
Se_Q_072	U897s	*S.* Infantis	56	57	40	40	147	65	8	515	111884	194406	438876	239230	510645	4869524	100,0	5
Se_Q_073	U899s	*S.* Infantis	73	74	15	15	28	5	2	561	869	2130	2574	1868	2574	6751	0,0	13,4 kB
Se_Q_074	U969s	*S.* Infantis	70	71	49	49	160	68	8	515	103470	183687	277902	270230	807038	4927197	99,2	5
Se_Q_075	U972s	*S.* Infantis	88	90	63	64	159	66	7	504	92095	183600	371658	345249	1029138	4972348	98,4	5,1
Se_Q_076	U973s	*S.* Infantis	66	67	33	33	112	52	7	523	89700	183599	224894	164281	298879	3144061	87,1	3,3
Se_Q_077	U976s	*S.* Infantis	43	44	31	31	111	42	6	509	102704	183598	224481	177563	324332	2634265	75,8	2,8
Se_Q_078	U979s	*S.* Infantis	73	74	52	53	148	61	6	502	127843	203917	525795	372931	1028771	4830648	98,4	5
Se_Q_079	U980s	*S.* Infantis	72	73	51	51	145	71	8	515	114902	183687	525908	271306	719606	4968981	99,2	5,1
Se_Q_080	U983s	*S.* Infantis	87	88	62	63	158	68	7	503	115007	183860	438969	256380	604576	4950623	99,2	5,1
Se_Q_081	U986s	*S.* Infantis	89	90	63	64	166	69	7	515	145663	194156	444101	282854	590383	4971259	100,0	5,1
Se_Q_082	U998s	*S.* Infantis	58	59	41	41	169	74	7	515	103500	180946	470149	319161	897909	4930452	99,2	5,1
Se_Q_083	U1000s	*S.* Infantis	151	154	107	109	170	65	5	515	96310	203916	470191	434277	1177679	4958205	98,4	5,1
Se_Q_084	U1003s	*S.* Infantis	65	66	46	47	169	68	6	515	145663	194285	525919	349482	949950	4962491	99,2	5,1
Se_Q_085	U1006s	*S.* Infantis	60	61	43	43	170	78	9	503	114985	180829	324316	215536	510820	4928362	99,2	5,1
Se_Q_086	U1010s	*S.* Infantis	217	222	147	150	157	63	5	515	127966	203919	525767	453608	1195274	4970923	100,0	5,1
Se_Q_087	U1019s	*S.* Infantis	175	180	121	123	152	60	5	515	127835	203831	470164	440745	1195276	5003523	100,0	5,1
Se_Q_088	U1051s	*S.* Infantis	142	145	96	98	157	58	5	515	105429	203919	525920	452315	1193237	4976698	98,4	5,1
Se_Q_089	U1052s	*S.* Infantis	147	149	100	101	153	55	5	515	114978	218061	525850	459568	1195276	4910636	98,4	5
Se_Q_090	U1056s	*S.* Infantis	138	141	96	98	155	64	5	515	105545	203916	470195	438765	1195275	4974139	100,0	5,1
Se_Q_091	U1057s	*S.* Infantis	1,3 GB	1,3 GB	444	452	157	57	5	515	127741	203922	525792	455204	1195276	4958329	98,4	5,1
Se_Q_092	U1059s	*S.* Infantis	186	190	129	131	156	63	6	515	115806	183687	469954	428704	1195269	4979617	98,4	5,1
Se_Q_093	U1066s	*S.* Infantis	235	239	157	160	175	67	7	515	115806	183687	328299	411906	1195275	4978508	98,4	5,1
Se_Q_094	U1068s	*S.* Infantis	139	141	91	92	185	75	8	515	75953	194484	371300	236424	605857	4969036	99,2	5,1
Se_Q_095	U1071s	*S.* Infantis	246	252	170	173	160	66	6	515	105653	181631	470589	439083	1195276	4973584	98,4	5,1
Se_Q_096	U1092s	*S.* Infantis	281	287	74	76	173	76	7	515	92032	183599	371360	362990	1096372	4945769	98,4	5,1
Se_Q_097	U1095s	*S.* Infantis	301	308	208	212	168	61	6	504	105653	217634	444101	426654	1195274	4942761	98,4	5,1
Se_Q_098	U1111s	*S.* Infantis	169	172	119	121	171	62	5	515	105545	217656	470238	454390	1200266	4983958	100,0	5,1
Se_Q_099	U1114s	*S.* Infantis	206	210	146	148	145	59	4	515	116776	332783	525930	463678	1195276	4970958	98,4	5,1
Se_Q_100	U1119s	*S.* Infantis	95	97	65	66	161	66	7	515	105429	183687	333629	409911	1195079	4958535	99,2	5,1
Se_Q_101	U1121s	*S.* Infantis	153	155	108	110	174	64	5	515	96310	203919	470590	441553	1195275	4975149	99,2	5,1
Se_Q_102	U1125s	*S.* Infantis	173	176	124	126	173	65	6	515	105429	183647	470176	431332	1195275	4971014	99,2	5,1
Se_Q_103	U1128s	*S.* Infantis	134	137	95	96	133	49	5	515	127734	217930	525798	462431	1195275	4918610	98,4	5
Se_Q_104	U1132s	*S.* Infantis	135	137	95	96	182	64	6	515	92095	183687	525921	427536	1195272	4964167	98,4	5,1
Se_Q_105	U1133s	*S.* Infantis	147	150	103	105	166	64	5	515	94331	203916	525876	439913	1200267	5017389	100,0	5,1

Se_Q_106	U1135s	*S.* Infantis	1010	1,1 GB	480	487	142	59	5	515	145663	203922	525921	458314	1195275	4972749	99,2	5,1
Se_Q_107	U1143s	*S.* Infantis	91	92	64	65	143	63	5	515	124634	203920	526047	451960	1190461	4973014	99,2	5,1
Se_Q_108	U1145s	*S.* Infantis	86	87	61	62	160	65	6	502	105429	183687	525971	371372	1029604	5001339	100,0	5,1
Se_Q_109	U1175s	*S.* Infantis	66	67	47	47	150	72	9	502	115034	201639	280583	207648	438873	4953178	98,4	5,1
Se_Q_110	U1178s	*S.* Infantis	151	154	105	107	152	70	6	515	105437	203888	525922	327898	824584	4973125	98,4	5,1
Se_Q_111	U1181s	*S.* Infantis	86	89	60	61	133	57	7	515	148183	201627	370742	357697	1028877	4916214	98,4	5
Se_Q_112	U1187s	*S.* Infantis	72	74	50	51	174	72	8	515	103538	180947	525800	328857	951266	5012740	99,2	5,1
Se_Q_113	U1192s	*S.* Infantis	73	74	51	52	156	68	7	515	93967	183687	438876	266279	590274	4955045	98,4	5,1
Se_Q_114	U1193s	*S. Enteritidis*	88	90	63	63	73	26	4	618	400868	478531	694896	528125	978047	4698575	98,4	4,8
Se_Q_115	U1196s	*S.* Infantis	134	136	93	94	168	64	6	515	115007	183687	470195	433446	1195276	4971350	98,4	5,1
Se_Q_116	U1401	*S. Infantis*	67	69	48	48	163	70	8	515	105429	183687	269563	260207	753860	4969119	100,0	5,1
Se_Q_117	U1405s	*S.* Infantis	89	92	62	63	148	63	6	502	148183	201620	444101	430796	1195276	4973634	98,4	5,1
Se_Q_118	U1407s	*S.* Infantis	141	143	98	99	172	75	6	504	105429	201632	444101	422133	1195182	4981059	100,0	5,1
Se_Q_119	U1410s	*S.* Infantis	56	58	39	40	138	56	8	515	137723	194184	324237	227470	443813	4662115	99,2	4,8
Se_Q_120	U1411s	*S.* Infantis	66	67	46	47	151	65	8	586	138005	194184	383560	229155	444101	4831418	98,4	5
Se_Q_121	U1412s	*S. Infantis*	59	61	42	43	333	133	16	508	59252	103944	171465	112579	224615	4710491	100,0	4,9
Se_Q_122	U1418s	*S.* Infantis	98	100	68	69	170	78	6	515	105545	201629	371360	414892	1195276	4988717	98,4	5,1
Se_Q_123	U1424s	*S.* Infantis	79	81	55	56	172	75	8	515	93976	181340	324484	321795	1004693	4943578	100,0	5,1
Se_Q_124	U1431s	*S.* Infantis	59	60	41	42	150	71	10	515	114985	201626	239404	198987	439942	4880964	100,0	5
Se_Q_125	U1432s	*S.* Infantis	57	58	40	40	150	73	9	586	105321	194285	232171	206695	511727	4858079	98,4	5
Se_Q_126	U1436s	*S.* Infantis	57	58	40	41	127	67	8	515	105321	194175	444101	240905	521906	4872597	100,0	5
Se_Q_127	U1437s	*S.* Infantis	85	87	59	60	144	64	6	515	145663	194191	525799	380438	1029076	4968250	98,4	5,1
Se_Q_128	U1447s	*S.* Infantis	85	87	60	61	152	58	7	515	138104	194184	438969	274360	590383	4912998	98,4	5
Se_Q_129	U1459s	*S.* Infantis	92	94	64	65	180	80	8	515	103482	180526	324266	334060	1045248	4979507	100,0	5,1
Se_Q_130	U1467s	*S. Infantis*	53	54	37	38	219	93	9	524	47492	122678	209274	155807	437028	3334242	82,3	3,5
Se_Q_131	U1470s	*S.* Infantis	58	60	40	41	147	72	8	523	105429	194184	278475	253896	653954	4911304	100,0	5
Se_Q_132	U1473s	*S.* Infantis	47	48	33	33	134	56	8	552	87822	160941	225182	199285	505523	3804469	91,9	4
Se_Q_133	U1483s	*S.* Infantis	98	100	68	70	176	75	7	515	124227	201649	438969	290447	756072	4975275	98,4	5,1
Se_Q_134	U1489s	*S.* Infantis	33	34	23	23	69	10	1	544	1862	15179	15179	9533	15179	25781	0,0	39,5 kB
Se_Q_135	U1493s	*S.* Infantis	71	72	49	50	168	76	11	515	82933	149868	270578	211586	669891	4996844	99,2	5,1
Se_Q_136	U1495s	*S.* Infantis	84	85	58	59	179	73	8	515	90033	166870	324680	326000	1029381	4971687	98,4	5,1
Se_Q_137			59	60	42	42	67	31	4	564	225385	341534	1266045	541359	1266045	4989379	99,2	5,1
Se_Q_138	U1505s	*S.* Infantis	88	90	63	64	163	76	7	515	107976	183687	336225	340933	1029044	4946285	100,0	5,1
Se_Q_139	U1506s	*S.* Infantis	61	63	42	43	147	68	8	515	145663	192205	324412	302782	915788	4906403	99,2	5
Se_Q_140	U1508s	*S.* Infantis	145	148	102	104	138	60	5	515	105545	203917	525920	452305	1195275	4975692	99,2	5,1
Se_Q_141	U1519s	*S.* Infantis	244	248	171	173	297	109	8	509	105653	204075	375192	234190	490266	4996052	99,2	5,1
Se_Q_142			**D**	**D**	**D**	**D**	**D**	**D**	**D**	**D**	**D**	**D**	**D**	**D**	**D**	**D**	**D**	0

Se_Q_143	U1671s	*S.* Infantis	137	139	96	98	160	67	5	515	145663	203920	525920	454465	1195089	4974094	100,0	5,1
Se_Q_144	U1672s	*S.* Infantis	123	125	87	88	178	69	5	515	105545	203916	525896	441228	1195182	4979276	100,0	5,1
Se_Q_145	U1673s	*S.* Infantis	80	81	56	57	166	67	7	502	114905	183947	443813	361724	1044874	5018874	100,0	5,1
Se_Q_146	U1678s	*S.* Infantis	89	91	64	64	161	62	7	502	127731	183687	525737	294937	719599	4962469	98,4	5,1
Se_Q_147	U1680s	*S.* Infantis	89	90	61	62	148	53	6	515	124631	194757	525863	389544	1029137	4915668	99,2	5
Se_Q_148	U1682s	*S.* Infantis	**D**	**D**	**D**	**D**	**D**	**D**	**D**	**D**	**D**	**D**	**D**	**D**	**D**	**D**	**D**	0
Se_Q_149	U1683s	*S.* Infantis	120	122	85	86	152	62	6	515	145663	201624	444101	429267	1195100	4965818	99,2	5,1
Se_Q_150	U1686s	*S.* Infantis	141	143	100	101	163	62	4	515	152221	398405	525799	480844	1195275	5024068	99,2	5,1
Se_Q_151	U1688s	*S.* Infantis	125	127	88	89	187	67	6	515	105545	183600	525879	379183	1027435	4973991	99,2	5,1

Samples discarded prior to the study due to low quality data or no data obtained after all the assely protocol in the current study (D); Yellow colored cells indicate human clinical isolates.

Supplementary Table 2. Details of the 160 prophage genomes used for the custom database using information from Switt et al. (2015) and Gao et al. (2020), and curated using ICTV taxonomy release database

NCBI Accession number	Phage name	Genome Size (bp)	Virus Type	Host
NC_006552.1	*Hollowayvirus F116*	65195	dsDNA	*Pseudomonas aeruginosa*
NC_005357.1	*Rauchvirus BPP1*	42493	dsDNA	*Bordetella bronchiseptica*
NC_005887.1	*Ryyoungvirus bcepC6B*	42415	dsDNA	*Burkholderia cepacia*
NC_015266.1	*Kayeltresvirus KL3*	40555	dsDNA	*Burkholderia ambifaria*
NC_025115.1	*Arsyunavirus RSY1*	40002	dsDNA	*Ralstonia solanacearum*
NC_015273.1	*Kisquattuordecimvirus KS14*	32317	dsDNA	*Burkholderia cenocepacia*
NC_009237.1	*Bcepmuvirus E255*	37446	dsDNA	*Burkholderia thailandensis*
NC_005882.1	*Bcepmuvirus bcepMu*	36748	dsDNA	*Burkholderia cenocepacia*
NC_005178.1	*Casadabanvirus D3112*	37611	dsDNA	*Pseudomonas aeruginosa*
NC_008717.1	*Casadabanvirus DMS3*	36415	dsDNA	*Pseudomonas aeruginosa*
NC_011976.1	*Salmonella phage epsilon34*	43016	dsDNA	*Salmonella enterica*
NC_030919.1	*Salmonella phage 118970_sal4*	42418	dsDNA	*Salmonella enterica*
NC_031019.1	*Enterobacteria phage UAB_Phi20*	41809	dsDNA	*Salmonella enterica*
NC_005841.1	*Enterobacteria phage ST104*	41391	dsDNA	*Salmonella enterica*
NC_028696.2	*Salmonella phage SEN22*	41338	dsDNA	*Salmonella enterica*
NC_014900.1	*Salmonella phage ST160*	40986	dsDNA	*Salmonella enterica*
NC_013059.1	*Salmonella phage g341c*	40975	dsDNA	*Salmonella enterica*
NC_004348.1	*Lederbergvirus ST64T*	40679	dsDNA	*Salmonella enterica*
NC_031946.1	*Salmonella Phage 103203_sal5*	40443	dsDNA	*Salmonella enterica*
NC_017985.1	*Salmonella phage SPN9CC*	40128	dsDNA	*Salmonella enterica*
NC_018275.1	*Salmonella phage vB_SemP_Emek*	39783	dsDNA	*Salmonella enterica*
NC_019501.1	*Enterobacteria phage IME10*	39646	dsDNA	*Escherichia coli*
NC_005344.1	*Lederbergvirus Sf6*	39043	dsDNA	*Shigella flexneri*
NC_027398.1	*Enterobacteria phage Sf101*	38742	dsDNA	*Escherichia coli*
NC_002730.1	*Lederbergvirus HK620*	38297	dsDNA	*Escherichia coli*
NC_019445.1	*Escherichia phage TL-2011b*	44784	dsDNA	*Escherichia coli*
NC_031077.1	*Enterobacter phage Tyrion*	41760	dsDNA	*Salmonella enterica*
NC_004775.2	*Uetakevirus epsilon15*	39672	dsDNA	*Salmonella enterica*
NC_016761.1	*Uetakevirus SPN1S*	38684	dsDNA	*Salmonella enterica*
NC_028656.1	*Oslovirus PA2*	65955	dsDNA	*Escherichia coli*
NC_010237.1	*Traversvirus min27*	63395	dsDNA	*Escherichia coli*
NC_028685.1	*Oslovirus VASD*	62851	dsDNA	*Escherichia coli*
NC_025434.1	*Diegovirus POCJ13*	62699	dsDNA	*Shigella sonnei*
NC_000924.1	*Traversvirus tv933W*	61670	dsDNA	*Escherichia coli*
NC_000902.1	*Traversvirus II*	60942	dsDNA	*Escherichia coli*
NC_018846.1	*Oslovirus ov191*	60894	dsDNA	*Escherichia coli*
NC_029120.1	*Diegovirus dv7502Stx*	60875	dsDNA	*Shigella sonnei*
NC_008464.1	*Traversvirus tv86*	60238	dsDNA	*Escherichia coli*
NC_004813.1	*Marienburgvirus BP4795*	57930	dsDNA	*Escherichia coli*
NC_011356.1	*Pankowvirus YYZ2008*	54896	dsDNA	*Escherichia coli*
NC_011357.1	*Pankowvirus pv1717*	62147	dsDNA	*Escherichia coli*
NC_018279.1	*Salmonella phage vB_SosS_Oslo*	49116	dsDNA	*Salmonella enterica*

NC_006949.1	*Enterobacteria phage ES18*	46900	dsDNA	*Salmonella enterica*
NC_019721.1	*Enterobacterial phage mEp390*	40029	dsDNA	*Escherichia coli*
NC_019705.1	*Shamshuipovirus mEpX2*	38759	dsDNA	*Escherichia coli*
NC_016160.1	*Saikungvirus HK75*	36661	dsDNA	*Escherichia coli*
NC_019709.1	*Cuauhtlivirus mEpX1*	41567	dsDNA	*Escherichia coli*
NC_019719.1	*Saikungvirus HK633*	41528	dsDNA	*Escherichia coli*
NC_019714.1	*Kwaitsingvirus HK446*	39026	dsDNA	*Escherichia coli*
NC_019708.1	*Nochtlivirus mEp235*	37595	dsDNA	*Escherichia coli*
NC_002166.1	*Shamshuipovirus HK022*	40751	dsDNA	*Escherichia coli*
NC_002167.1	*Byrnievirus HK97*	39732	dsDNA	*Escherichia coli*
NC_019768.1	*Wanchaivirus HK106*	41468	dsDNA	*Escherichia coli*
NC_021190.1	*Enterobacteria phage phi80*	46150	dsDNA	*Escherichia coli*
NC_019717.1	*Enterobacteria phage HK225*	45366	dsDNA	*Escherichia coli*
NC_019704.1	*Enterobacteria phage mEp237*	44375	dsDNA	*Escherichia coli*
NC_019706.1	*Aguilavirus mEp043*	42780	dsDNA	*Escherichia coli*
NC_031940.1	*Salmonella phage 118970_sal3*	77375	dsDNA	*Salmonella enterica*
NC_003356.1	*Enterobacteria phage phiP27*	42575	dsDNA	*Escherichia coli*
NC_021857.1	*Shigella phage SfII*	41475	dsDNA	*Shigella flexneri*
NC_004313.1	*Salmonella phage ST64B*	40149	dsDNA	*Salmonella enterica*
NC_022749.1	*Shigella phage SfIV*	39758	dsDNA	*Shigella flexneri*
NC_003444.1	*Enterobacteria phage SfV*	37074	dsDNA	*Shigella flexneri*
NC_001895.1	*Peduovirus magyaro (p2)*	33593	dsDNA	*Escherichia coli*
NC_004745.1	*Peduovirus L413-C*	30728	dsDNA	*Yersinia pestis*
NC_005340.1	*Eganvirus PsP3*	30636	dsDNA	*Salmonella enterica*
NC_001317.1	*Eganvirus ev186*	30624	dsDNA	*Salmonella enterica*
NC_005056.1	*Peduovirus WPhi*	32684	dsDNA	*Escherichia coli*
NC_022750.1	*Peduovirus fiAA91-ss*	33628	dsDNA	*Escherichia coli*
NC_028701.2	*Senquatrovirus SEN4*	33509	dsDNA	*Salmonella enterica*
NC_021774.1	*Stockinghallvirus FSL SP004*	29742	dsDNA	*Salmonella enterica*
NC_029003.2	*Eganvirus SEN1*	29733	dsDNA	*Salmonella enterica*
NC_028943.1	*Peduovirus pro483*	29237	dsDNA	*Escherichia coli*
NC_019488.1	*Felsduovirus RE2010*	34117	dsDNA	*Salmonella enterica*
NC_010463.1	*Enterobacteria phage Fels-2*	33693	dsDNA	*Salmonella enterica*
NC_026014.1	*Xuanwuvirus P88*	35814	dsDNA	*Escherichia coli*
NC_019932.1	*Entnonagintavirus ENT90*	29564	dsDNA	*Erwinia amylovora*
NC_010393.1	*Phage Gifsy-2*	45840	unclassified	*Salmonella enterica*
NC_010392.1	*Phage Gifsy-1*	48491	unclassified	*Salmonella enterica*
NC_001416.1	*Lambdavirus lambda*	48502	dsDNA	*Escherichia coli*
NC_020845.1	*Eurybiavirus MED4213*	180977	dsDNA	*Prochlorococcus marinus*
NC_023693.1	*Justusliebigvirus phi92*	148612	dsDNA	*Escherichia coli*
NC_009904.1	*Kochikohdavirus EF24C*	142072	dsDNA	*Enterococcus faecalis*
NC_001697.1	*Hpunavirus HP1*	32355	dsDNA	*Haemophilus influenzae*
NC_003315.1	*Hpunavirus HP2*	31508	dsDNA	*Haemophilus influenzae*
NC_005856.1	*Punavirus P1*	94800	dsDNA	*Escherichia coli*
NC_031129.1	*Punavirus SJ46*	103445	dsDNA	*Escherichia coli*
NC_010495.1	*Macdonaldcampvirus ViIIE1*	45051	dsDNA	*Salmonella enterica*
NC_031924.1	*Shuimuvirus IME207*	47564	dsDNA	*Salmonella enterica*
NC_005284.1	*Stanholtvirus sv1026b*	54865	dsDNA	*Burkholderia pseudomallei*

NC_024365.1	*Pseudomonas phage phiPSA1*	51090	dsDNA	*Pseudomonas syringae*
NC_031091.1	*Pseudomonas phage MD8*	43277	dsDNA	*Pseudomonas aeruginosa*
NC_005859.1	*Tequintavirus T5*	121750	dsDNA	*Escherichia coli*
NC_028748.2	*Waukeshavirus BMBtp3*	51366	dsDNA	*Bacillus thuringiensis*
NC_028841.1	*Lilyvirus lily*	44952	dsDNA	*Paenibacillus larvae*
NC_019401.1	*Mimasvirus GAP32*	358663	dsDNA	*Cronobacter sakazakii*
NC_009821.1	*Krischvirus georgiaone*	164270	dsDNA	*Escherichia coli*
NC_020079.1	*Escherichia virus phAPEC8*	147737	dsDNA	*Escherichia coli*
NC_004827.1	*Bacteriophage Aaphi23*	43033	dsDNA	*Actinobacillus actinomycetemcomitans*
NC_019934.1	*Cronobacter phage ENT39118*	39012	dsDNA	*Cronobacter sakazakii*
NC_013594.1	*Muvirus mu*	37235	dsDNA	*Escherichia coli*
NC_019455.1	*Haemophilus phage SuMu*	37151	dsDNA	*Haemophilus parasuis*
NC_028898.1	*Baylorvirus bv1127AP1*	35764	dsDNA	*Mannheimia haemolytica*
NC_028896.1	*Peduovirus pro147*	32675	dsDNA	*Escherichia coli*
NC_003313.1	*Longwoodvirus K139*	33106	dsDNA	*Vibrio cholerae*
NC_019522.1	*Pectobacterium phage ZF40*	48454	dsDNA	*Pectobacterium carotovorum*
NC_019927.1	*Cronobacter phage ENT47670*	47611	dsDNA	*Cronobacter sakazakii*
NC_015295.1	*Erwinia phage phiEt88*	47279	dsDNA	*Escherichia coli*
NC_025458.1	*Shewanella sp. phage 1_41*	43510	dsDNA	*no hit found*
NC_026611.1	*Gofduovirus GF2*	43129	dsDNA	*Escherichia coli*
NC_027995.1	*Jilinvirus CVM10*	41666	dsDNA	*Escherichia coli*
NC_028699.1	*Brunovirus SEN34*	40740	dsDNA	*Salmonella enterica*
NC_027339.1	*Enterobacteria phage SfI*	38389	dsDNA	*Shigella flexneri*
NC_024369.2	*Vibrio phage X29*	41569	dsDNA	*Vibrio cholerae*
NC_019514.1	*Waedenswilvirus S6*	74669	dsDNA	*Erwinia amylovora*
NC_025445.1	*Bonnellvirus J865*	40981	dsDNA	*Pantoea agglomerans*
NC_025443.1	*Nonanavirus nv9NA*	52869	dsDNA	*Salmonella enterica*
NC_011551.1	*Sendosyvirus APSE2*	39867	dsDNA	*Candidatus Hamiltonella defensa*
NC_000935.1	*Sendosyvirus APSE1*	36524	dsDNA	*Candidatus Hamiltonella defensa*
NC_009514.1	*Phage cdtI DNA*	47021	dsDNA	*Escherichia coli*
NC_031264.1	*Brucella phage BiPBO1*	46877	dsDNA	*Brucella abortus*
NC_001901.1	*Ravinvirus N15*	46375	dsDNA	*Escherichia coli*
NC_005069.1	*Yersinia phage PY54*	46339	dsDNA	*Yersinia enterocolitica*
NC_019716.1	*Enterobacteria phage mEp460*	44510	dsDNA	*Escherichia coli*
NC_018843.1	*Salmonella phage SSU5*	103299	dsDNA	*Salmonella enterica*
NC_029028.1	*Nonagvirus JenP1*	60754	dsDNA	*Escherichia coli*
NC_028776.1	*Seuratvirus cajan*	59670	dsDNA	*Escherichia coli*
NC_019545.1	*Salmonella phage SPN3UB*	47355	dsDNA	*Salmonella enterica*
NC_005857.1	*Klebsiella phage phiKO2*	51601	dsDNA	*Klebsiella oxytoca*
NC_016158.1	*Escherichia phage HK639*	49576	dsDNA	*Escherichia coli*
NC_009552.2	*Geobacillus virus E2*	40863	dsDNA	*Bacillus anthracis*
NC_018454.1	*Cronobacter phage phiES15*	39974	dsDNA	*Cronobacter sakazakii*
NC_015296.1	*Kuttervirus ViI*	157061	dsDNA	*Salmonella enterica*
NC_001609.1	*Enterobacteria phage P4*	11624	dsDNA	*Escherichia coli*
NC_023575.1	*Pseudomonas phage vB_PaeP_Tr60_Ab31*	45550	dsDNA	*Pseudomonas aeruginosa*
NC_020850.1	*Vibrio phage VBM1 genomic sequence*	38374	dsDNA	*Vibrio parahaemolyticus*
NC_010391.1	*Salmonella phage Fels-1*	42723	unclassified	*Salmonella enterica*

NC_001954.1	*Infulavirus If1*	8454	ssDNA	*Escherichia coli*
NC_006294.1	*Capistrivirus KSF1*	7107	ssDNA	*Vibrio cholerae*
NC_001332.1	*Enterobacteria phage I2-2*	6744	ssDNA	*Escherichia coli*
NC_025824.1	*Inovirus M13*	6408	ssDNA	*Escherichia coli*
NC_002371.2	*Lederbergvirus p22*	41724	dsDNA	*Salmonella enterica*
NC_049461.1	*Peduovirus R18C*	31834	dsDNA	*Escherichia coli*
NC_073747.1	*Peduovirus STYP1*	28946	ssDNA	*Salmonella enterica*
NC_020414.2	*Zindervirus UAB78*	43984	ssDNA	*Salmonella enterica*
NC_011802.1	*Lederbergvirus SE1Spa*	41941	ssDNA	*Salmonella enterica*
NC_016073.1	*Kuttervirus SFP10*	157950	Circular DNA	*Salmonella enterica*
NC_019910.1	*Agtrevirus SKML39*	159624	ssDNA	*Salmonella enterica*
NC_005282.1	*Felixounavirus felixO1*	86155	ssDNA	*Salmonella enterica*
NC_016071.1	*Seunavirus PVPSE1*	145964	ssDNA	*Salmonella enterica*
NC_020416.1	*Gelderlandvirus s16*	160221	ssDNA	*Salmonella enterica*
NC_015269.1	*Tequintavirus SPC35*	118351	ssDNA	*Salmonella enterica*
NC_010583.1	*Epseptimavirus EPS7*	111382	ssDNA	*Salmonella enterica*
NC_021777.1	*Jerseyvirus jersey*	43447	ssDNA	*Salmonella enterica*
NC_016763.1	*Jerseyvirus SE2*	43221	ssDNA	*Salmonella enterica*
NC_009232.2	*Jerseyvirus SETP3*	42572	ssDNA	*Salmonella enterica*
NC_006940.2	*Jerseyvirus SS3e*	40793	ssDNA	*Salmonella enterica*
NC_019417.1	*Chivirus SPN19*	59203	Circular DNA	*Salmonella enterica*
NC_004831.2	*Zindervirus SP6*	43769	ssDNA	*Salmonella enterica*
NC_010807.1	*Teetrevirus SGJL2*	38815	ssDNA	*Salmonella enterica*

Supplementary Table 3. Details for all *Salmonella enterica* prophages identified with their respective detected size, prophage reference ID, identity percentage, reference genome size and completeness level

Sample name	NCBI Code	Number of phages detected	Phage Size (pB)	Phage ID	% ID	Reference phage size (pB)	Phage Completeness
Se_Q_001	CSG217	2	22183	*Peduovirus pro483*	97,54	29237	75,87 %
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_002	CSG232	2	15325	*Salmonella phage ST64B*	98,016	40149	38,17%
			6091	*Phage Gifsy-2*	98,375	45840	13,29%
Se_Q_003	CSG252	2	15325	*Salmonella phage ST64B*	98,016	40149	38,17%
			6091	*Phage Gifsy-2*	98,375	45840	20,40%
Se_Q_004	CSG177	4	13543	*Salmonella phage 118970 sal3*	99,801	77375	17,50%
			6087	*Shigella phage SfII*	97,601	41475	14,68%
			22005	*Phage Gifsy-1*	99,973	48491	45,38 %
			6429	*Phage Gifsy-2*	99,684	45840	13,29%
Se_Q_005	CSG175	2	12887	*Peduovirus pro483*	97,54	29237	44,07%
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_006	CSG132	2	22183	*Peduovirus pro483*	97,54	29237	75,87 %
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_007	U113s	3	38150	*Enterobacteria phage ST104*	99,971	41391	95,02 %
			25721	*Salmonella phage 118970 sal3*	99,642	77375	33,24%
			31512	*Phage Gifsy-2*	99,743	45840	68,71 %
Se_Q_008	U114s	1	15325	*Salmonella phage ST64B*	98,016	40149	38,17%
Se_Q_009	U120s	2	22183	*Peduovirus pro483*	97,54	29237	75,87 %
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_010	U121s	2	12887	*Peduovirus pro483*	97,54	29237	44,07%
			6853	*Peduovirus Wphi*	97,344	32684	20,97%
Se_Q_011	U123s	1	22122	*Peduovirus fiAA91ss*	95,257	33628	65,78%
Se_Q_012	U127s	2	22183	*Peduovirus pro483*	97,54	29237	75,87 %
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_013	U128s	2	22183	*Peduovirus pro483*	97,54	29237	75,87 %
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_014	U129s	2	22183	*Peduovirus pro483*	97,54	29237	75,87 %
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_015	U219s	2	22183	*Peduovirus pro483*	97,54	29237	75,87 %
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_016	U634s	2	22183	*Peduovirus pro483*	97,54	29237	75,87 %
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%

Se_Q_017	U638s	2	22183	*Peduovirus pro483*	97,54	29237	75,87 %
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_018	U639s	2	22183	*Peduovirus pro483*	97,54	29237	75,87 %
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_019	U652s	2	22230	*Peduovirus pro483*	97,54	29237	76,03 %
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_021	U664s	3	16347	*Lambdavirus lambda*	96,525	48502	33,70%
			14114	*Enterobacteria phage mEp460*	98,838	44510	31,71%
			10031	*Shigella phage SfII*	96,999	41475	24,18%
Se_Q_022	U666s	2	22230	*Peduovirus pro483*	97,54	29237	76,03 %
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_023	U669s	2	12887	*Peduovirus pro483*	97,532	29237	44,07%
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_024	U672s	2	12887	*Peduovirus pro483*	97,532	29237	44,07%
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_025	U676s	2	22183	*Peduovirus pro483*	97,54	29237	75,87 %
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_026	U679s	2	22183	*Peduovirus pro483*	97,54	29237	75,87 %
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_027	U682s	2	22183	*Peduovirus pro483*	97,54	29237	75,87 %
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_028	U684s	2	12887	*Peduovirus pro483*	97,532	29237	44,07%
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_029	U689s	2	22230	*Peduovirus pro483*	97,532	29237	76,03 %
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_030	U692s	2	22183	*Peduovirus pro483*	97,532	29237	75,87 %
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_031	U706s	2	22183	*Peduovirus pro483*	97,54	29237	75,87 %
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_032	U707s	2	12887	*Peduovirus pro483*	97,54	29237	44,07%
			6853	*Peduovirus Wphi*	97,344	32684	20,97%
Se_Q_033	U708s	2	22183	*Peduovirus pro483*	97,54	29237	75,87 %
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_034	U711s	2	22183	*Peduovirus pro483*	97,54	29237	75,87 %
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_035	U712s	2	12945	*Peduovirus pro483*	97,458	29237	44,27%
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%

Se_Q_036	U715s	2	22183	*Peduovirus pro483*	97,54	29237	75,87 %
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_037	U718s	0					
Se_Q_038	U719s	1	12887	*Peduovirus pro483*	97,532	29237	44,07%
Se_Q_039	U723s	2	22183	*Peduovirus pro483*	97,54	29237	75,87 %
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_040							
Se_Q_041	U728s	1	12887	*Peduovirus pro483*	97,532	29237	44,07%
Se_Q_042	U729s	1	12887	*Peduovirus pro483*	97,532	29237	44,07%
Se_Q_043	U758s	2	22183	*Peduovirus pro483*	97,54	29237	75,87 %
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_044	U759s	2	22183	*Peduovirus pro483*	97,54	29237	75,87 %
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_045	U762s	2	22183	*Peduovirus pro483*	97,54	29237	75,87 %
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_046	U769s	2	22183	*Peduovirus pro483*	97,54	29237	75,87 %
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_047	U773s	1	12887	*Peduovirus pro483*	97,532	29237	44,07%
Se_Q_048	U801s	2	22183	*Peduovirus pro483*	97,54	29237	75,87 %
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_049	U804s	1	12887	*Peduovirus pro483*	97,532	29237	44,07%
Se_Q_050	U806s	2	22183	*Peduovirus pro483*	97,54	29237	75,87 %
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_051	U814s	1	12887	*Peduovirus pro483*	97,532	29237	44,07%
Se_Q_052	U815s	3	12887	*Peduovirus pro483*	97,532	29237	44,07%
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
			6853	*Peduovirus Wphi*	97,344	32684	20,97%
Se_Q_053	U818s	3	12887	*Peduovirus pro483*	97,532	29237	44,07%
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
			6853	*Peduovirus Wphi*	97,344	32684	20,97%
Se_Q_054	U855s	3	12887	*Peduovirus pro483*	97,532	29237	44,07%
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
			6853	*Peduovirus Wphi*	97,344	32684	20,97%
Se_Q_055	U857s	3	12887	*Peduovirus pro483*	97,54	29237	44,07%
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
			6853	*Peduovirus Wphi*	97,344	32684	20,97%
Se_Q_056	U860s	2	22183	*Peduovirus pro483*	97,54	29237	75,87 %

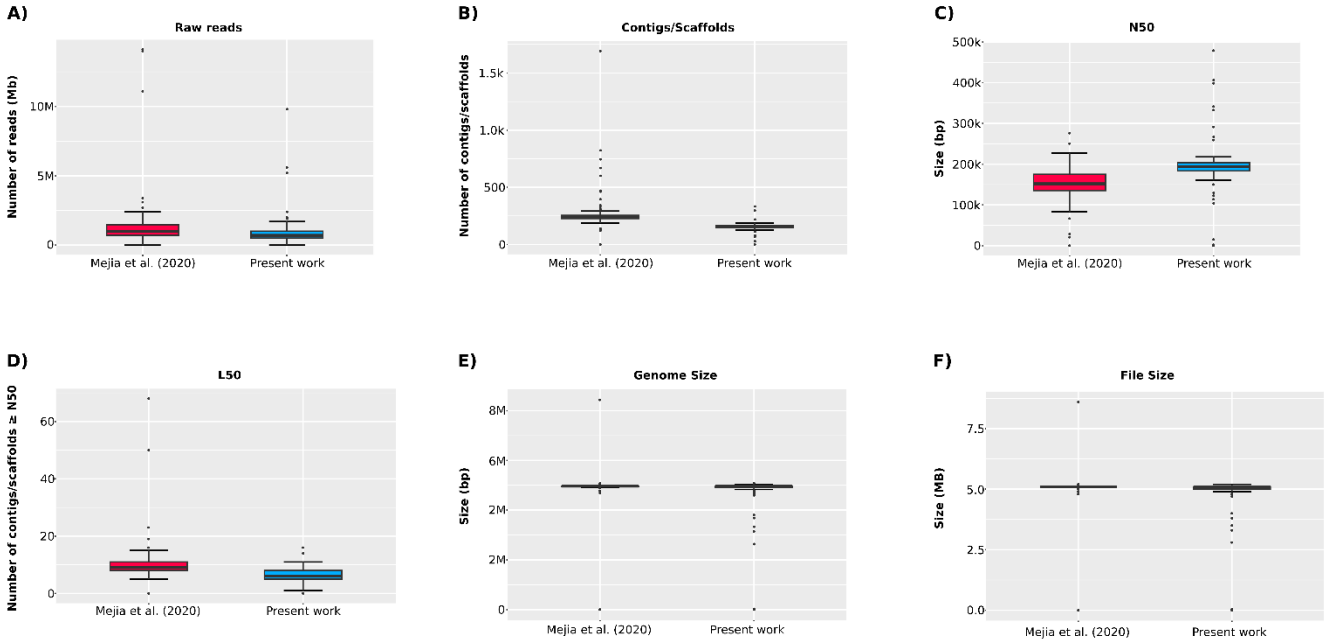
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_057	U863s	2	12945	*Peduovirus pro483*	97,458	29237	44,27%
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_058	U865s	2	12887	*Peduovirus pro483*	97,54	29237	44,07%
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_059	U866s	2	22230	*Peduovirus pro483*	95,182	29237	76,03 %
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_060	U869s	0					
Se_Q_061	U871s	2	22183	*Peduovirus pro483*	97,54	29237	75,87 %
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_062	U873s	1	22183	*Peduovirus pro483*	97,54	29237	75,87 %
Se_Q_063	U875s	3	15325	*Salmonella phage ST64B*	98,16	40149	38,17%
			6091	*Phage Gifsy-2*	98,375	45840	13,29%
			12914	*Peduovirus pro483*	96,856	29237	44,17%
Se_Q_064	U878s	2	12887	*Peduovirus pro483*	97,54	29237	44,07%
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_065	U880s	2	12887	*Peduovirus pro483*	97,54	29237	44,07%
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_066	U883s	2	22183	*Peduovirus pro483*	97,54	29237	75,87 %
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_067	U885s	2	12887	*Peduovirus pro483*	97,54	29237	44,07%
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_068	U888s	2	12887	*Peduovirus pro483*	97,54	29237	44,07%
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_069	U889s	2	22183	*Peduovirus pro483*	97,54	29237	75,87 %
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_070	U890s	2	22183	*Peduovirus pro483*	97,54	29237	75,87 %
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_071	U895s	2	12887	*Peduovirus pro483*	97,54	29237	44,07%
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_072	U897s	2	22183	*Peduovirus pro483*	97,54	29237	75,87 %
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_073	U899s	0					
Se_Q_074	U969s	2	22183	*Peduovirus pro483*	97,54	29237	75,87 %
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_075	U972s	2	12945	*Peduovirus pro483*	97,458	29237	44,27%
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%

Se_Q_076	U973s	1	12887	*Peduovirus pro483*	97,54	29237	44,07%
Se_Q_077	U976s	2	22183	*Peduovirus pro483*	97,54	29237	75,87 %
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_078	U979s	2	22183	*Peduovirus pro483*	97,54	29237	75,87 %
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_079	U980s	2	22183	*Peduovirus pro483*	97,54	29237	75,87 %
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_080	U983s	2	22183	*Peduovirus pro483*	97,54	29237	75,87 %
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_081	U986s	2	22230	*Peduovirus pro483*	95,182	29237	76,03 %
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_082	U998s	2	22183	*Peduovirus pro483*	97,54	29237	75,87 %
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_083	U1000s	2	12887	*Peduovirus pro483*	97,54	29237	44,07%
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_084	U1003s	2	22230	*Peduovirus pro483*	95,182	29237	76,03 %
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_085	U1006s	2	22183	*Peduovirus pro483*	97,54	29237	75,87 %
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_086	U1010s	2	12887	*Peduovirus pro483*	97,54	29237	44,07%
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_087	U1019s	2	22183	*Peduovirus pro483*	97,54	29237	75,87 %
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_088	U1051s	2	12887	*Peduovirus pro483*	97,54	29237	44,07%
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_089	U1052s	2	12887	*Peduovirus pro483*	97,54	29237	44,07%
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_090	U1056s	2	12945	*Peduovirus pro483*	97,458	29237	44,27%
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_091	U1057s	2	12887	*Peduovirus pro483*	97,54	29237	44,07%
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_092	U1059s	2	12887	*Peduovirus pro483*	97,54	29237	44,07%
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_093	U1066s	2	12887	*Peduovirus pro483*	97,54	29237	44,07%
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_094	U1068s	2	22230	*Peduovirus pro483*	95,182	29237	76,03 %
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%

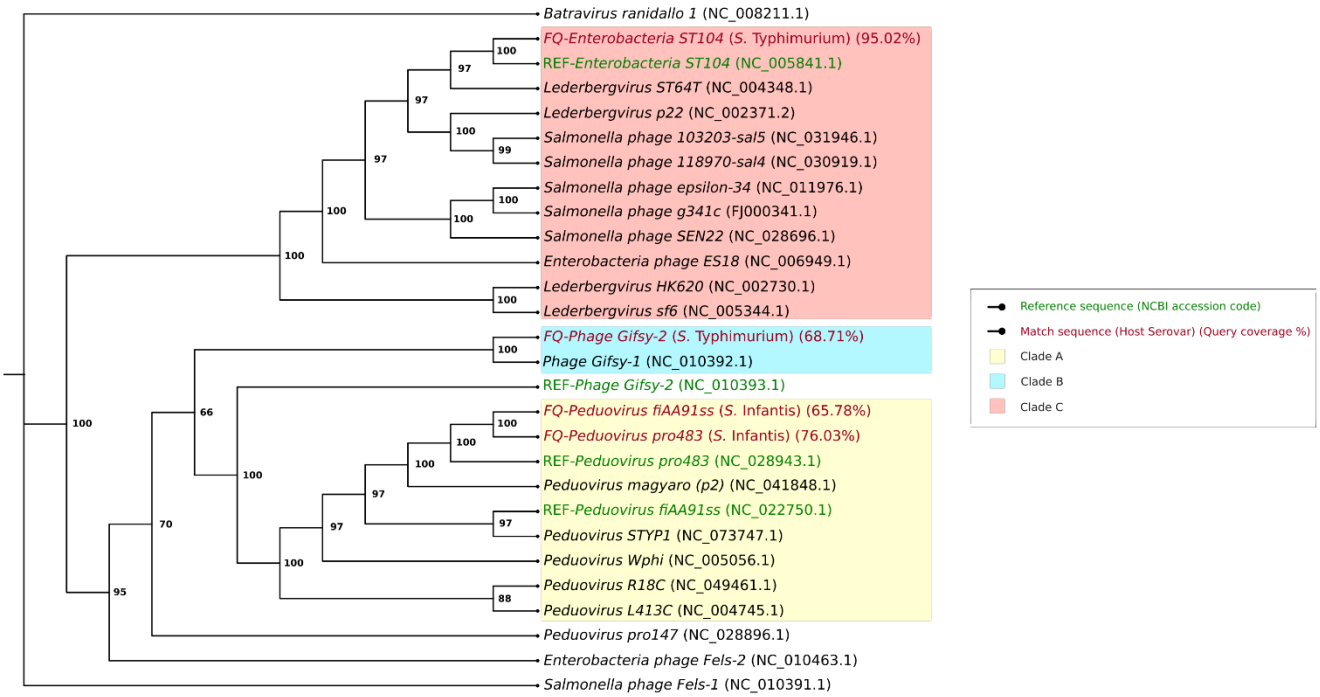
Se_Q_095	U1071s	2	22183	*Peduovirus pro483*	97,54	29237	75,87 %
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_096	U1092s	2	12887	*Peduovirus pro483*	97,54	29237	44,07%
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_097	U1095s	2	12887	*Peduovirus pro483*	97,54	29237	44,07%
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_098	U1111s	2	12887	*Peduovirus pro483*	97,54	29237	44,07%
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_099	U1114s	2	12945	*Peduovirus pro483*	97,458	29237	44,27%
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_100	U1119s	2	22183	*Peduovirus pro483*	97,54	29237	75,87 %
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_101	U1121s	2	12887	*Peduovirus pro483*	97,54	29237	44,07%
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_102	U1125s	2	12887	*Peduovirus pro483*	97,54	29237	44,07%
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_103	U1128s	2	12887	*Peduovirus pro483*	97,54	29237	44,07%
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_104	U1132s	2	12887	*Peduovirus pro483*	97,54	29237	44,07%
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_105	U1133s	2	12887	*Peduovirus pro483*	97,54	29237	44,07%
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_106	U1135s	2	22183	*Peduovirus pro483*	97,54	29237	75,87 %
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_107	U1143s	2	22230	*Peduovirus pro483*	95,182	29237	76,03 %
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_108	U1145s	2	22230	*Peduovirus pro483*	95,182	29237	76,03 %
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_109	U1175s	2	22183	*Peduovirus pro483*	97,54	29237	75,87 %
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_110	U1178s	2	12887	*Peduovirus pro483*	97,54	29237	44,07%
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_111	U1181s	2	22183	*Peduovirus pro483*	97,54	29237	75,87 %
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_112	U1187s	2	22183	*Peduovirus pro483*	97,54	29237	75,87 %
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_113	U1192s	2	12887	*Peduovirus pro483*	97,54	29237	44,07%

			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_114	U1193s	2	15325	*Salmonella phage ST64B*	98,016	40149	38,17%
			6091	*Phage Gifsy-2*	98,375	45840	13,29%
Se_Q_115	U1196s	2	22183	*Peduovirus pro483*	97,54	29237	75,87 %
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_116	U1401	2	22183	*Peduovirus pro483*	97,54	29237	75,87 %
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_117	U1405s	2	22183	*Peduovirus pro483*	97,54		
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_118	U1407s	2	12887	*Peduovirus pro483*	97,532	29237	44,07%
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_119	U1410s	2	22183	*Peduovirus pro483*	97,54	29237	75,87 %
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_120	U1411s	2	22230	*Peduovirus pro483*	95,182	29237	76,03 %
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_121	U1412s	2	13532	*Enterobacteria phage mEp460*	98,928	44510	30,40%
			10031	*Shigella phage SfII*	96,999	41475	24,18%
Se_Q_122	U1418s	2	12887	*Peduovirus pro483*	97,54	29237	44,07%
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_123	U1424s	2	22183	*Peduovirus pro483*	97,54		
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_124	U1431s	2	22230	*Peduovirus pro483*	97,54	29237	76,03 %
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_125	U1432s	2	22183	*Peduovirus pro483*	97,54	29237	75,87 %
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_126	U1436s	2	12887	*Peduovirus pro483*	97,532	29237	44,07%
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_127	U1437s	2	22183	*Peduovirus pro483*	97,54	29237	75,87 %
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_128	U1447s	2	22183	*Peduovirus pro483*	97,54	29237	75,87 %
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_129	U1459s	2	22230	*Peduovirus pro483*	95,182	29237	76,03 %
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_130	U1467s	1	10688	*Punavirus SJ46*	98,765	103445	10,33%
Se_Q_131	U1470s	2	22183	*Peduovirus pro483*	97,54	29237	75,87 %
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_132	U1473s	2	22230	*Peduovirus pro483*	97,54	29237	76,03 %

			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_133	U1483s	2	22230	*Peduovirus pro483*	97,54	29237	76,03 %
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_134	U1489s	0					
Se_Q_135	U1493s	2	22183	*Peduovirus pro483*	97,54	29237	75,87 %
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_136	U1495s	2	22183	*Peduovirus pro483*	97,54	29237	75,87 %
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_137		0					
Se_Q_138	U1505s	2	12887	*Peduovirus pro483*	97,54	29237	44,07%
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_139	U1506s	2	22183	*Peduovirus pro483*	97,54	29237	75,87 %
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_140	U1508s	3	12887	*Peduovirus pro483*	97,532	29237	44,07%
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
			6853	*Peduovirus Wphi*	97,344	32684	20,97%
Se_Q_141	U1519s	3	12887	*Peduovirus pro483*	97,532	29237	44,07%
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
			6853	*Peduovirus Wphi*	97,344	32684	20,97%
Se_Q_142							
Se_Q_143	U1671s	2	22183	*Peduovirus pro483*	97,54	29237	75,87 %
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_144	U1672s	2	12887	*Peduovirus pro483*	97,54	29237	44,07%
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_145	U1673s	2	22183	*Peduovirus pro483*	97,54	29237	75,87 %
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_146	U1678s	2	22183	*Peduovirus pro483*	97,54	29237	75,87 %
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_147	U1680s	2	12887	*Peduovirus pro483*	97,54	29237	44,07%
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_148	U1682s	0					
Se_Q_149	U1683s	2	22183	*Peduovirus pro483*	97,54	29237	75,87 %
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_150	U1686s	2	22183	*Peduovirus pro483*	97,54	29237	75,87 %
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%
Se_Q_151	U1688s	2	12887	*Peduovirus pro483*	97,54	29237	44,07%
			6055	*Stockinghallvirus FSL SP004*	95,574	29742	20,36%



Supplementary Figure 1. Box plots illustrating the difference between the prior study by Mejía et al. (2020) (depicted in red) and the current work (depicted in blue), focusing on various assembly and quality control statistics. A) Differences in the number of reads obtained after quality control for raw reads. B) Differences in the number of contigs/scaffolds obtained. C) Differences in N50 measured in fragment size. D) Differences in L50 measured in number of contigs/scaffolds ≥N50. E) Differences in genome size obtained. F) Differences in file size obtained.



Supplementary Figure 2. Phylogenetic relationships using the complete prophage genome. Reference sequences are written in green and found sequences in red. 3 clusters were observed: Cluster A (yellow) with *Peduovirus*-like sequences, ClusterB (cyan) with *Gifsy*-like sequences, and Cluster C with *Lederbergvirus*-like sequences. Maximum likelihood tree constructed with IQ-TREE2 with a fast bootstrap of 1000 iterations using a general time reversible model with gamma distribution (GTR+G) predicted with MEGAX.