A SYST-OMICS APPROACH TO FOOD SAFETY:

IDENTIFYING GENETIC DETERMINANTS OF *SALMONELLA* VIRULENCE *IN VIVO*

Alanna Crouse

DEPARTMENT OF HUMAN GENETICS,
MCGILL UNIVERSITY, MONTREAL
June 2018

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Master of Science

© Alanna Crouse, 2018

ABSTRACT

Over the past 50 years, Canadians have increased their consumption of fresh fruits and vegetables. Over this same time period, foodborne related illness caused by contaminated fresh produce has also been on the rise, with *Salmonella* as one of the leading cause of outbreaks. To improve current food safety processes and reduce the burden of salmonellosis, The *Salmonella* Syst-OMICs Consortium has turned to whole genome sequencing. 4500 *Salmonella* strains were isolated and sequenced, 35 of which were tested in 4 models of infection to elucidate the association between sequence and virulence. For our contribution we used susceptible C57BL/6J mice to determine isolate virulence *in vivo*.

Isolates were assigned to "low virulent", "intermediate virulent" and "highly virulent" categories. Correlation analyses comparing isolate behaviour across models revealed a high degree of concordance with nearly all estimates of isolate effects being significantly positive. Comparative genomic analyses identified 4 known and 1 putative virulence associated genes (*sopE*) present in highly virulent isolates and absent in low virulent isolates. *sopE* was confirmed as a virulence associated gene following *in vivo* testing of SL1344Δ*sopE*. The presence or absence of these 5 genes was used to identify a collection of isolates to be tested in mice and *in vitro* using human epithelial cells. 17 isolates showed concordant phenotypes *in vitro* and *in vivo*. WGS of these isolates was analyzed by principal component analysis, revealing that isolates segregate by virulence. 33 isolates of unknown virulence were added to the PCA. Sequence similarity to the 17 isolates of known virulence was used to predict the phenotype of the 33 isolates. 22 out of 33 isolates yielded phenotypes *in vivo* matching their predictions. False negatives were a concern with 4 isolates predicted to be low virulent actually being highly virulent. Application of this system will require further refinement and study into the complexity of host pathogen interactions.

RÉSUMÉ

Au cours des 50 dernières années, la population canadienne a augmenté sa consommation de fruits et légumes frais. Au cours de cette même période, les maladies d'origine alimentaire causées par des produits frais contaminés ont également augmenté. L'infection à salmonelles est devenue une des causes principales d'éclosion de toxi-infection alimentaire. Pour améliorer les processus actuels de sécurité alimentaire et réduire le fardeau de la salmonellose, le consortium de *Salmonella* Syst-OMICs a utilisé une approche de séquençage du génome entier. 4500 souches de *Salmonella* ont été isolées et séquencées, dont 35 ont été caractérisées dans 4 modèles d'infections pour examiner l'association entre la séquence et la virulence. Notre contribution à cette initiative d'envergure a été d'examiner la virulence des souches salmonelles *in vivo*. Pour ce faire, nous avons utilisé un modèle murin bien établi chez les souris C57BL/6J sensibles à l'infection.

Les souches ont été attribuées à des catégories «faiblement virulentes», «virulence intermédiaire» et «de haute virulence». Les analyses de corrélation, comparant le comportement des souches dans tous les modèles d'infection, ont dévoilé un degré élevé de concordance. Les analyses comparatives génomiques ont identifié quatre gènes de virulence connus et un gène présumé virulent (*sopE*) présents chez les souches de haute virulence et absents chez les souches de faible virulence. *SopE* a été examiné *in vivo* en utilisant la souches SL1344deltaSopE, et en outre confirmé comme gène de haute virulence. La présence ou l'absence de ces 5 gènes a été utilisée pour identifier 17 souches qui ont démontré des phénotypes concordants *in vitro* et *in vivo*. Le séquençage du génome entier (*WGS*) de ces souches, suivi par une analyse en composantes principales (ACP), a révélé que ces souches se regroupent suivant leur degré de virulence. Finalement, 33 souches dont la virulence est inconnue ont été ajoutés à l'analyse PCA. L'homologie de leur séquence aux 17 souches de virulence connues a été utilisée pour prévoir leur phénotype de virulence. Nous avons ensuite évalué

ces 33 souches dans le modèle murin. Une majorité des souches (22/33) se sont comportée tel que prédit. Toutefois, 4 souches prédites comme étant de faible virulence, se sont avérées être très virulentes. L'application de ce système nécessitera certainement un raffinement supplémentaire dans l'étude de la complexité de l'interaction entre pathogènes et hôtes.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

CFU	colony forming unit
DC	dendritic cell
delta SL1344	*Salmonella* Typhimurium strain SL1344Δ*invA*Δ*sseB*
DMEM	Dulbecco's modified Eagle medium
DNA	deoxyribonucleic acid
ELISA	enzyme-linked immunosorbent assay
FBS	Fetal bovine serum
GALT	gut-associated lymphoid tissue
HIV	human immunodeficiency virus
IFN	interferon
IL	interleukin
iNTS	invasive non-typhoidal *Salmonella*
I.V.	intravenous
LFA	lateral flow assay
LPS	lipopolysaccharide
M cell	microfold cell
MAP	mitogen-activated protein
MLST	mutlilocus sequence typing
MSMD	Mendelian susceptibility to mycobacterial disease
NADPH	nicotinamide adenine nucleotide phosphate
NASBA	nucleic acid sequence based amplification
NF-κB	nuclear factor κβ
NLR	nod-like receptor
Nramp1	natural resistance-associated macrophage protein 1
NTS	non-typhoidal *Salmonella*
OD	optical density
ORF	open reading frame
PAMP	pathogen-associated molecular patterns
PCA	principal component analysis
PCR	polymerase chain reaction
PFGE	pulse field gel electrophoresis
pg/mL	picograms per milliliter
PMN	polymorphonuclear leukocytes
P.O.	per os
qPCR	quantitative polymerase chain reaction
RCF	relative centrifugal force
ROS	reactive oxygen species
RNS	reactive nitrogen species
RPM	rotations per minute

SCV	*Salmonella* containing vacuole
Sif	*Salmonella* induced filaments
SIGIRR	single Ig IL-1 related receptor
Slc11a1	solute carrier family 11 member 1
SNV	single nucleotide variants
SPI	*Salmonella* pathogenicity island
STAT	signal transducer activator of transcription
T1SS	type one secretory system
T3SS	type three secretory system
TLR	toll-like receptor
TNF	tumor necrosis factor
TSA	tryptic soy agar
TSB	tryptic soy broth
VBNC	viable-but-nonculturable
WGS	whole genome sequencing
WT SL1344	wild-type *Salmonella* Typhimurium strain SL1344

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to begin by acknowledging that McGill University resides on the traditional and unceded land of the Kanien'kehá:ka (Mohawk) Nation. This land has long served as a site of meeting and exchange amongst Indigenous peoples.

I would like to thank my supervisors Dr. Danielle Malo and Dr. Samantha Gruenheid for welcoming me into their labs and for providing guidance, support, and opportunity over the past two years. Thank you for your continuous encouragement and for helping me reach my next step. I am very proud and grateful to have been co-supervised by two amazing female scientists.

Thank you to my supervisory committee: Dr. Silvia Vidal and the late Dr. Hervé Le Moual for their time and insights.

I would like to say a special thank you to Line Larivière and Lei Zhu. Thank you Line for sharing your wisdom, kindness and joy, all of which help make the lab the warm place that it is. To Lei, thank you for the countless hours, mice, strains and patience. Thank you also to Patricia D'Arcy who performed every infection needed for the work presented in this thesis.

Thank you to all the amazing present and past members of the Malo lab: Leila Rached-d'Astous, Kyoko Yuki, Etienne Flammant, Megan Teh, and Jamie Kim. With a special shout out to Megan Eva who has been there for all the trouble shooting, questions and, of course, jokes.

Thank you to my friends, floor members and The Squad for all the TH nights, memes and distractions. Thank you, Caitlin, empowering me and humbling me. I will forever admire your charisma, uniqueness, nerve, and talent. Finally, thank you to my family who has the utmost patience and endless support. Nothing has made me happier than to keep surprising you and making you proud. I promise I'll try to keep it up.

CONTRIBUTIONS OF WORK

I, under the supervision of Dr. Danielle Malo and Dr. Samantha Gruenheid, have performed and analyzed the experiments described in this thesis. The contributions of others are as follows:

THE *SALMONELLA* SYST-OMICS CONSORTIUM:

Dr. Lawrence Goodridge isolated *Salmonella* strains of unknown virulence that were tested *in vivo*. The genomic sequencing and analysis was performed by Dr. Roger Levesque and his team: Jean-Guillaume Emond, Jérémie Hamel, and Brian Boyle. This includes the comparative genomic and principal component analyses performed. Isolates were provided to us by Dr. Roger Levesque and Dr. France Daigle. Genetically modified isolates delta SL1344 and delta *sopE* were a generous gift from Dr. France Daigle. Dr. France Daigle and Maud Kerhoas performed *in vitro Salmonella* phenotyping in both human epithelial cells and human macrophages. Dr. Rafael Garduno, Dr. John Rohde, and Adrian Herod were responsible for *Salmonella* phenotyping in amoeba.

CORRELATION ANALYSIS

Dr. Celia Greenwood and Dr. Catherine Schramm performed contrast analyses for each model and correlation analyses assessing concordance between models of infection.

IN VIVO SALMONELLA PHENOTYPING

Technical assistance was provided by Line Larivière. Maintenance of the mouse colony and technical assistance was provided by Lei Zhu. Mouse infections were performed by Patricia D'Arcy.

INTRODUCTION & LITERATURE REVIEW

SALMONELLA CLASSIFICATIONS

Salmonella are rod-shaped, Gram-negative bacteria belonging to the Enterobactericeae family. Since 1966, when Kauffman began the process of developing a system with which to organize *Salmonella* isolates, *Salmonella* nomenclature has evolved significantly (1). This is due to the immense diversity of *Salmonella* serovars as well as ever evolving technologies which have provided more precise and accurate ways to categorize and delineate isolates. Currently, *Salmonella* is classified into two species: *Salmonella enterica* and *Salmonella bongori*. *S. enterica* is further broken down into six subspecies: enterica, salamae, arizonae, diarizonae, houtenae and indica (Fig. 1). In humans, *Salmonella enterica* subspecies enterica is of greatest concern as this subspecies is found predominantly in mammals and causes approximately 99% of *Salmonella* infections in humans and other warm-blooded animals (2). The other five *Salmonella* subspecies as well as *S. bongori* are found mainly in the environment and in cold-blooded animals (3). Within *S. enterica* subsp. enterica alone, there is over 2600 serovars described to date (4). These serovars can be broken into two groups, Typhoidal *Salmonella* and Non-Typhoidal *Salmonella* (NTS). In the past, serovars within these groups have been classified based on the presence or absence of O (somatic) and H (flagellar) antigens on the cell wall of isolates. Although useful, the ability to assess isolate relatedness at the genetic level has shown that antigenic formulas can be misleading. For example, with the use of DNA-DNA hybridization it was shown that, in contrast to previous classifications, all serovars in subspecies arizonae are actually the same species (5).

Today, gene-based classification systems continue to provide a more discriminatory representation of *Salmonella* divergence. Multilocus sequence typing (MLST) has shown that while some serovar categories are largely supported at the sequence level, as is the case for isolates within *S.* Typhimurium, others serovars mistakenly group isolates that are genetically unrelated. Still other serovars fail to group isolates that have a closely shared evolutionary history (4). With contradictions

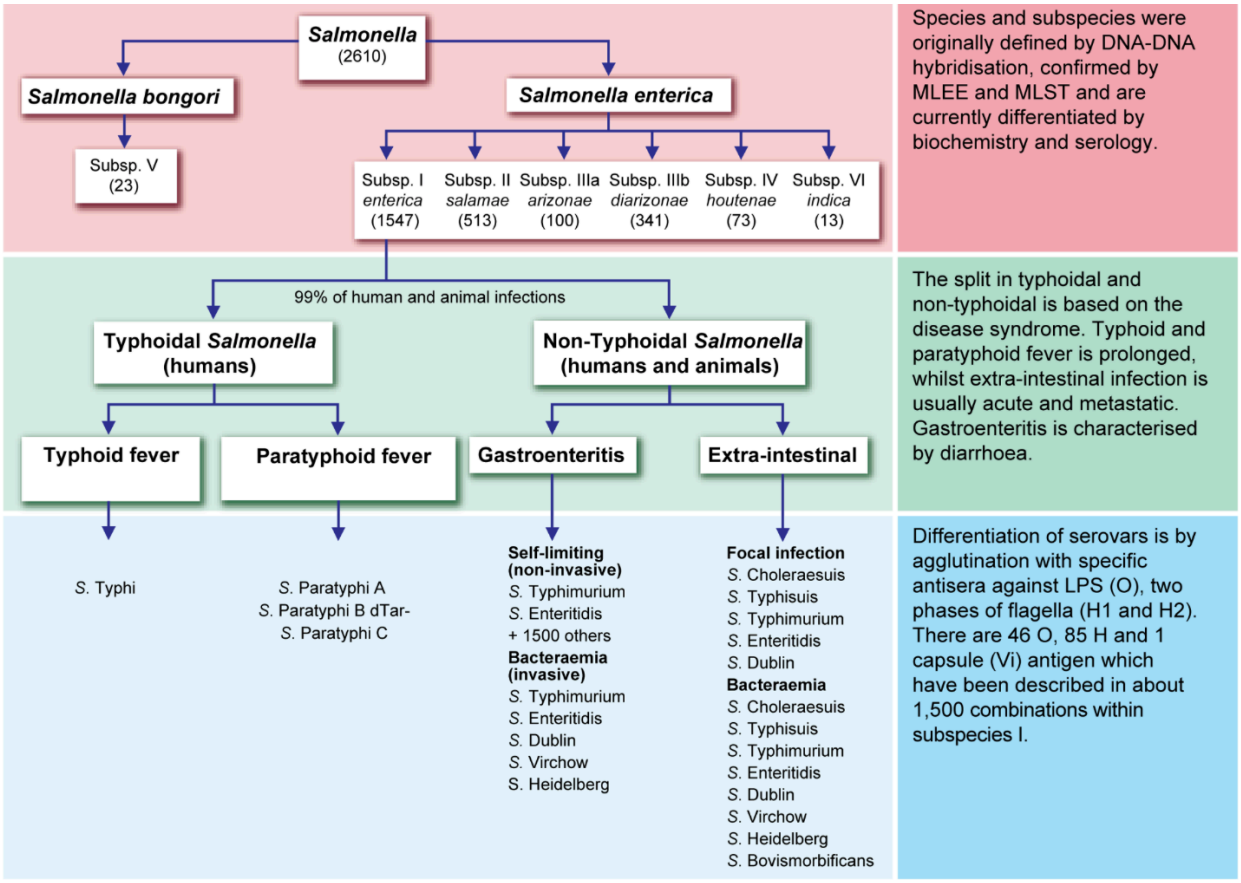

Fig. 1. *Salmonella* diversity broken down by species, subspecies, and clinical manifestations. From Achtman et al., 2012. *Multilocus Sequence Typing as a Replacement for Serotyping in Salmonella enterica*. PLoS Pathogens.

between traditional and modern methods of classifications becoming more prevalent it is likely that *Salmonella* nomenclature will continue to develop into the future.

SALMONELLA EPIDEMIOLOGY AND PATHOGENESIS

The genetic diversity across *Salmonella* coincides with the ability of different isolates to generate different diseases. *Salmonella* infection can lead to primarily three clinical manifestations: Typhoid fever, salmonellosis, and invasive non-Typhoidal *Salmonella* (iNTS) infection. Typhoid, or enteric fever, is a systemic disease that is a major cause of morbidity and mortality in areas of the world with poor sanitation and hygiene. Approximately 22 million cases are diagnosed per year resulting in more than 433,000 deaths worldwide (6, 7) Ingestion of contaminated water is the primary cause of infection. As such, typhoid fever is rare in developed countries with approximately 400 cases diagnosed per year and nearly all associated with international travel (8). Typhoid fever is caused by a select group of typhoidal *Salmonella* serovars that are restricted to human and higher primate hosts. These include *Salmonella enterica* serovar Typhi, and serovar Paratyphi A, B, and C. Symptoms include fever, chills, abdominal pain, loss of appetite and general malaise manifesting 1-3 weeks post-infection. When treated with antibiotics the infection is typically cleared within a few days. When left untreated however, symptoms can last up to 4 weeks with a case fatality rate between 12 and 30% (9). A small proportion (1% to 5%) of those infected become chronic asymptomatic carriers (10).

Salmonellosis, or gastroenteritidis caused by *Salmonella*, is one of the most common and widely distributed food-borne diseases in both Europe and North America (11). Each year, more than 93.8 million people become sick from consuming food that is contaminated with *Salmonella* (12). Salmonellosis in immunocompetent individuals is characterized by a self-limiting gastroenteritis with symptoms of diarrhea, abdominal pain, headache, nausea and vomiting

occurring between 6 and 72 hours post-ingestion of contaminated food. These symptoms generally subside within 4-7 days without treatment. The *Salmonella* serovars most commonly associated with this diarrhoeal disease are *Salmonella enterica serovar* Typhimurium and *serovar* Enteritidis (13).

To cause disease *Salmonella* have to resist the acidic pH of the stomach. After reaching the intestines, *Salmonella* adhere to the apical surface of the epithelial cells via fimbriae adhesins (14). Both typhoidal and NTS invade the intestinal mucosa preferentially through the microfold cells (M cells) overlying the Peyer's patch in the distal ileum (15, 16). Although invading M cells is the primary method for crossing the epithelium *Salmonella* use alternative routes as well including within CD18-expressing phagocytes (17), as well as within dendritic cells (DCs) which open the tight junctions between epithelial cells, reach up into the lumen and engulf the *Salmonella* (18).

After crossing the epithelium *Salmonella* migrate to the lamina propria where they can access the gut-associated lymphoid tissue (GALT) and be taken up by innate immune cells including macrophages, neutrophils and dendritic cells (19, 20). Bacteria then disseminate to systemic sites, either within phagocytic immune cells or directly within the blood, to the mesenteric lymph node, spleen, liver, bone marrow and gallbladder. Multicellular lesions develop at foci of infection through the recruitment of polymorphonuclear (PMN) leukocytes and bone-marrow derived mononuclear cells (21). Typhoidal *Salmonella* replicate and spread to new foci before re-entering circulation and moving back to the intestinal lumen via secretion in the bile, a process which promotes bacterial shedding.

Along with the initial innate immune response, infection with typhoidal *Salmonella* induces a complex adaptive immune response. As with other intracellular bacterial and viral infections, *Salmonella* infection induces a Th1 biased response. Both CD4+ and CD8+ T-cells have been found associated with *Salmonella* infection in patients (22). In addition, Th1 cytokines such as IFN-γ, TNF-α, and IL-10 are increased in humans following vaccination with live attenuated *S*. Typhi (23).

Unlike typhoidal *Salmonella*, NTS do not usually pass beyond the local intestinal lymph nodes in significant number to cause systemic disease in immunocompetent individuals. Following uptake by phagocytes, including macrophages, dendritic cells and neutrophils, NTS induce caspase-1 mediated cell death, initiating the production of IL-1 and IL-18. This, along with the production of chemokines and other pro-inflammatory cytokines including IFN-γ, TNF-α and IL-12, triggers a robust inflammatory response characterized by the recruitment of a large number of neutrophils, increased vascular permeability, mucosal edema, necrosis of the ileal mucosa, fluid loss and diarrhea (24, 25).

iNTS have emerged as an important cause of bacteremia in immunocompromised individuals (26). The clinical presentation of iNTS is characterized by a non-specific febrile illness often without enterocolitis (27). Like NTS, iNTS is caused by non-typhoidal *Salmonella* serovars, with thousands of serovars capable of producing iNTS disease. The manifestation of NTS infection as iNTS depends on host susceptibility, serovar specific propensity for extra-intestinal infection, or a combination of both (28). In contrast to NTS, iNTS resides and replicates intracellularly in the blood and bone-marrow (27). The most-common cause of iNTS disease worldwide is advanced HIV disease. In sub-Saharan Africa, NTS cause an invasive and severe disease with a case fatality rate estimated at 20% in HIV-infected adults (29). Young children with severe malaria and/or malnutrition are also at increased risk of iNTS (30). In high-income country, iNTS bacteremia is rare and occurs in immunocompromised individuals presenting with HIV infection, autoimmune diseases, chronic granulomatous disease or with the syndrome of Mendelian susceptibility to mycobacterial disease (MSMD) (31, 32).

SALMONELLA VIRULENCE FACTORS

Following infection, the interaction between *Salmonella* and an infected host is a complex one that requires a concerted effort from various genes on the side of both host and pathogen. At least 4% of the *Salmonella* genome, which translates to approximately 200 genes, is required for fatal infection in mice (33). These genes must be expressed at the right time and in the right tissues during the course of infection for effective pathogen growth and survival. Genes that enable virulence are encoded on virulence plasmids or throughout the chromosome. The majority of virulence genes are encoded within large gene cassettes known as *Salmonella* pathogenicity islands (SPIs). *Salmonella* as a genus is marked by a high degree of genetic diversity. There are at least 21 SPIs across all *Salmonella* serovars with different serovars containing distinct collections of SPIs in different locations throughout the chromosome (34). As an example, *S.* Typhi contains 601 unique genes and 17 SPIs whereas *S.* Typhimurium contains 479 unique genes and 15 SPIs, with only 13 SPIs common between the two serovars (34).

The two of the most extensively described SPIs are SPI-1 and SPI-2 both of which encode type III secretory systems (T3SSs). T3SSs are hollow, needle-like structures that are rooted in a multi-ring base spanning the inner and outer membranes of the bacterial envelope (35, 36). This structure is used to transfer virulence effector proteins from *Salmonella* into the host cytosol where they can disrupt normal cell function, and favour *Salmonella* uptake and survival. SPIs 1 and 2 each contain genes that encode the structural components of a T3SS, regulatory proteins, secreted effector proteins and chaperones.

SPI-1

Traditionally SPI-1 has been associated with invasion into non-phagocytic cells of the intestinal epithelium. A number of SPI-1 effector proteins including SipA, SipC, InvA, SopB, SopE,

SopE2 and SptP act directly or indirectly to induce actin cytoskeleton rearrangements within host cells via CDC42 and RAC1 signalling (37-43). This process is necessary to cause membrane ruffling and subsequently the induction of *Salmonella* internalization. Once inside host cells a number of SPI-1 effectors help the bacterium transition to an intracellular lifestyle. SptP is involved in shutting down membrane ruffling, while AvrA prevents apoptosis of epithelial cells. (38, 44, 45). Recent findings have shown that the role of SPI-1 genes during host cell invasion is not limited to epithelial cells. Engulfment by phagocytic cells including dendritic cells and B-cells is also modulated by SPI-1. This has been show by mutation of *invC* which results in decreased bacterial uptake into both cell types (46, 47).

SPI-1 has also been found to have important roles outside of cell invasion and uptake. In addition to mediating bacterial internalization SopB, SopE and SopE2 also induce intestinal inflammation. SPI-1 genes disrupt tight junctions between epithelial cells thereby permeating the epithelial barrier (48). These genes also trigger a signalling cascade leading to the production of pro-inflammatory cytokines, which then stimulate the transmigration of PMN cells into the intestinal lumen (49). Inflammation is further promoted by effectors SopA (50) and SipB, the latter of which induces pro-inflammatory IL-1β and IL-18 via caspase-1 activation (51).

SPI-1 is also important for intracellular survival through the development of the *Salmonella* containing vacuole (SCV), a specialized vacuole that holds and protects intracellular bacteria from the intracellular environment of the host cell. SPI-1 encoded SipA coordinates with SPI-2 encoded SifA to modulate the SCV (45). Likewise, SopE, SopB, and SptP also contribute to SCV development with the latter two controlling SCV trafficking and preventing fusion with host lysosomes (52, 53).

SPI-2

SPI-2 has commonly been associated with *Salmonella* proliferation. SPI-2 contains 4 gene groups categorized based on function: *ssa* (secretion system apparatus) genes whose protein products form the structural components for the T3SS, *sse* (secretion system effector) genes, *ssr* (secretion system regulator) gene (54), and *ssc* (secretion system chaperone) genes (55). Proteins SseB, SseC, and SseD act as translocon components which help move virulence effector proteins across the SCV via a T3SS into the host cytosol (56). Although there are a variety of effector proteins associated with the T3SS of SPI-2, not all are encoded within SPI-2. Those encoded within SPI-2 include SpiC, SseF, and SseG each of which promote *Salmonella* replication within macrophages (57-59). SpiC is proposed to have dual functionality as both an effector protein and a translocon components, as it is necessary for the translocation of fellow tranlocon components SseB, SseC, and SseD (60, 61). As an effector protein SpiC is involved in SCV biogenesis and maintenance. Specifically, SpiC inhibits membrane fusion with host endosomes and lysosomes (59). Both SseF and SseG are involved in SCV membrane dynamics and the formation of *Salmonella* induced filaments (Sifs) (62) which are tubular membrane extensions emanating from the SCV membrane necessary for SCV integrity and bacterial replication (63).

SPI-2 associated genes encoded outside of SPI-2 are dependent on those encoded within SPI-2 for their expression and/or translocation. Genes *sifA, sifB, sspH2, sseI,* and *sseJ* are all encoded outside SPI-2 but are dependant of SPI-2 T3SS regulatory proteins SsrA and SsrB for their expression (64) while SopD2, PipB, PipB2, SspH1, SspH2, SlrP, SseI, and SseJ are dependent upon the SPI-2 T3SS for their translocation across the SCV membrane to the extending Sifs (65-68). Many of these proteins, including SifA, SopD2 PipB, and SseJ, work in a coordinated manner to help position the SCV and allow it to grow as the bacteria replicate (58, 67-71).

Another important function of SPI-2 genes is their role in altering host immune responses. By impeding various elements of the immune system *Salmonella* can avoid being killed while creating a favourable environment for replication. One way *Salmonella* achieves this is by interfering with immune signalling pathways including those involving NF-κB and MAP kinase signaling, resulting in diminished downstream cytokine production (72, 73). SPI-2 genes have also been shown to help *Salmonella* avoid the damaging effects of oxidative burst within macrophages (74) and to further impede efficient macrophage functioning by inducing delayed apoptosis (75). Other immune processes altered by SPI-2 genes include altering iNOS localization in macrophages and manipulating phagocyte mobility (76, 77).

SPI-3

SPI-3 shows a high degree of variation between serovars, which has been hypothesized to play a role in determining host specificities (78). So far, SPI-3 has been found to contain 10 open reading frames (ORFs) one of which is the *mgtCB* operon. Despite the variability of SPI-3, this particular operon is present in all *Salmonella* serovars studied and sits in a highly conserved region of SPI-3 (79). The *mgtCB* operon contains genes *mgtB* and *mgtC*. These genes enable intramacrophage replication and are crucial for *Salmonella* survival in low Mg^{2+} conditions (80). Other virulence genes within SPI-3 include *misL* and *marT*. *misL*, which enable *Salmonella* to bind epithelial cells of the intestine, promote cell invasion, and help establish intestinal colonization (81).

SPI-4

As is the case with SPI-1 and SPI-3 encoded genes, SPI-4 is involved in adhesion and invasion into host cells (82). SPI-4 harbours the *siiABCDEF* gene cluster (83). Genes *siiCDF* encode components of a type I secretory system (T1SS) which secrete *siiE*, a large non-fimbrial adhesion

protein used for adhesion to epithelial cells (84). SPI-4 works in cooperation with SPI-1 to mediate a close interaction with microvilli of epithelial cells, leading to ruffling of the apical membrane and the translocation of SPI-1 effectors (85). Following invasion, the activity of SPI-4 contributes to intestinal inflammation (84). The T1SS of SPI-4 has also been suggested to play a role in secreting bacterial toxins. It has been postulated that SPI-4 may induce cytotoxicity by mediating apoptosis in macrophages (86).

CS54 pathogenicity island

Although not designated as an SPI, CS54 is another gene cassette associated with *Salmonella* pathogenicity. CS54 contains 5 known genes, one of which encodes ShdA. ShdA is a surface protein which enables *Salmonella* to bind fibronectin (87). Further study has revealed that ShdA promotes bacterial colonization of the cecum and Peyer's patch (88). Mutation of *shdA* results in a decrease in the duration of fecal shedding indicating that ShdA is important for bacterial persistence and long-term carriages (89). Other genes encoded on CS54 include *sinH, sinI, ratA,* and *ratB*, the latter of which has also been found to promote long-term fecal shedding (88).

Other SPIs

Less is known about the other SPIs that have been found to date. Preliminary work suggests overlapping functions across various SPIs. Like SPI-1, 3 and 4, SPI-5 and 6 have also been found to promote host cell invasion (90, 91). SPI-5 is likely involved in a variety of functions beyond cell invasion given that this cassette encodes a number of SPI-1 and SPI-2 effector proteins including PipB, SopB, PipA, and PipD (90). Mutations in the latter two genes have resulted in similar phenotypes, of decreased fluid secretion and inflammation during gastroenteritis (90).

SPI-11, 12, and 18 all play roles in promoting systemic disease (92-95). Other SPIs have been discovered on the basis of bioinformatic work such as SPI-15, 16, and 17 however their role during infection has yet to be fully understood. Further study is needed to understand the contribution of each SPI to infection, how the presence or absence of SPIs and particular genes translates to varying disease manifestations, and to uncover how proteins across various SPIs interact.

Virulence plasmid

Most *Salmonella* serovars carry a virulence plasmid that aids in the establishment of systemic infection (96). The size and content of the plasmid varies between serovars however the *spv* (*Salmonella* plasmid virulence) operon is highly conserved across serovars (97). The *spv* operon contains five genes designated *spv RABCD,* which are regulated by upstream *spvR* (96). This operon has been shown to enhance virulence in mice, humans, calves, and pigs (98, 99). SpvB and SpvC have been identified as genes of interest in terms of virulence. Both proteins are secreted by the SPI-2 T3SS (73, 100). SpvB is required for proliferation within macrophages and helps trigger apoptosis in macrophages later during infection (99, 101). It has been hypothesized that SpvB may also play a role in preventing oxidase assembly and, in doing so, decrease bacterial death (102). NADPH associates with actin filaments during the organization of oxidases and SpvB is involved in actin depolarization (103). SpvC alters intracellular signalling by inactivating host MAP kinases (104). Whether the two proteins interact or affect common cellular pathways has yet to be demonstrated.

MOUSE MODELS OF SALMONELLA INFECTION

Owing to their cost efficiency, variety in inbred strains, and pathophysiology during infection, mice have been extensively used to study the various clinical manifestations of *Salmonella*

infection. The most widely used mouse model is one of systemic disease caused by *Salmonella* Typhimurium which resembles the clinical syndrome of human iNTS with bacteremia. The pathogenesis of systemic *Salmonella* infection in mice can be divided into four phases (Fig. 2) (105). During the first phase *Salmonella* cross the epithelium and disseminate to the mesenteric lymph node, spleen, liver and bone marrow. This phase is marked by a balance between bacterial growth and bacterial clearing by activated phagocytes, primarily neutrophils and macrophages.

In the second phase, bacteria undergo exponential growth within immune cells while bacteria killing becomes negligible. The interaction between bacterial virulence factors and host defenses during this phase is crucial for determining disease progression and outcome in terms of mortality. In resistant mice *Salmonella* growth is slower and allows for the gradual onset of host defenses as opposed to susceptible mice which are unable to gain control over *Salmonella* replication. Host defenses include the production of antimicrobial peptides, pro-inflammatory cytokines, reactive oxygen species (ROS), and reactive nitrogen species (RNS).

The third phase of pathogenesis involves an innate immune mediated plateau in bacterial growth. Recognition of Pathogen-Associated Molecular Patterns (PAMPs) by Toll-like receptors (TLRs) and Nod-like receptors (NLRs) triggers a pro-inflammatory regulation of intracellular iron levels and coordination of various processes of cell death.

The fourth and final phase relies on an adaptive immune mediated increase in bacterial clearing. B-cells generate antibodies against LPS, flagellin, and other outer membrane proteins. B-cells are also important for initiating T-cell immunity through antigen presenting functions (106, 107). The expansion and differentiation of T-cells is critical during *Salmonella* clearance. Mice with deficient T-cell immunity, such as mice lacking mature CD4+ TCR-alpha beta cells, or deficient in CD28, are highly susceptible to infection with *Salmonella* of attenuated virulence (108, 109). The

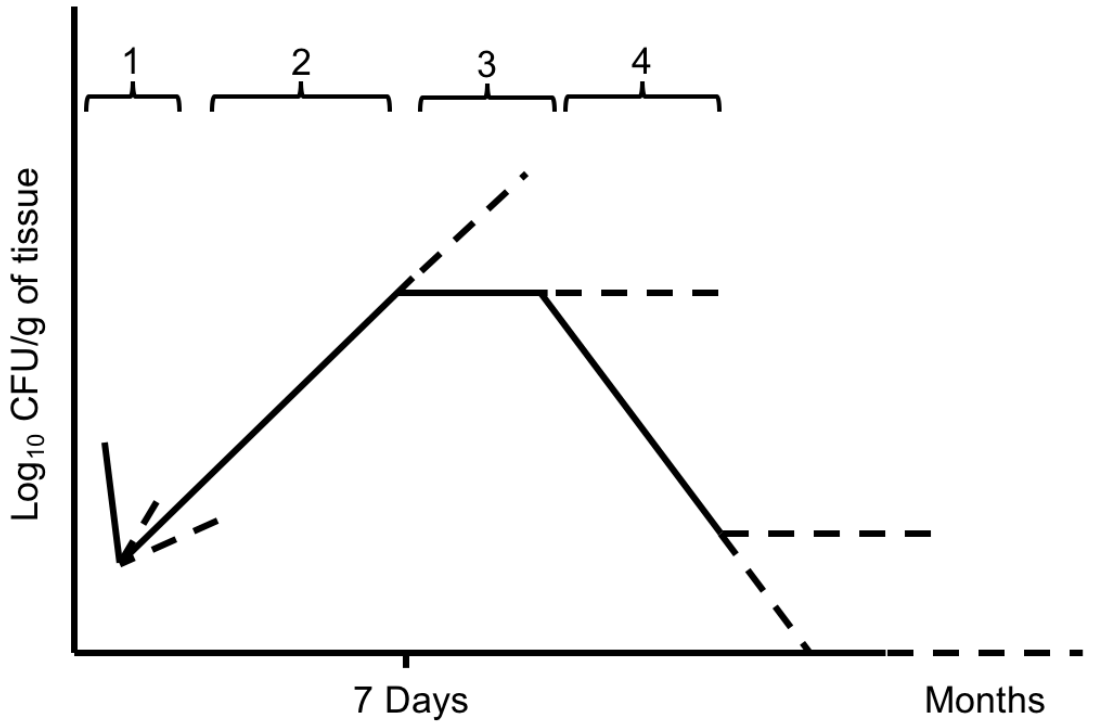

Fig. 2. Four phases of *Salmonella* infection in mice.

During the first phase *Salmonella* crosses the epithelium and disseminates to the mesenteric lymph node, spleen, liver and bone marrow. In the second phase bacteria undergo exponential growth within immune cells while bacteria killing becomes negligible. Bacteria growth rate is dependant on host genetics. The third phase of pathogenesis involves an innate immune mediated plateau in bacterial growth and the fourth phase features an adaptive immune mediated increase in bacterial clearing. Inability to clear bacteria results in the onset of chronic carriage. Adapted from Mastroeni, P. 2002. *Immunity to systemic Salmonella infections*. Curr Mol Med.

Th1 bias observed during infection in mice promotes the production of pro-inflammatory cytokine IFN-γ necessary for *Salmonella* clearance (110).

Throughout the course of infection host resistance to systemic infection depends on the genetic background of the mice. Classical inbred strains of mice can be classified into three distinct categories with respect to their susceptibility to *Salmonella* Typhimurium infection. 129 substrains (129S1, 129S6, 129X1) are extremely resistant to infection compared to A/J or CBA/J mice which present an intermediate susceptibility phenotype. Other strains such as C57BL/6J, BALB/cJ, FVB/J and C3H/HeJ are extremely susceptible to infection as measured by shorter survival time and high bacterial loads in spleen and liver (111).

The primary gene underlying the susceptibility C57BL/6J mice was identified as Solute carrier family 11 member 1 (*Slc11a1* initially known as *Nramp1*) (112). During infection Slc11a1 is recruited to the phagosomal membrane where it contributes to SCV maturation and mediates iron depletion from the phagosome (113). *Slc11a1* has pleiotropic effects and confers resistance to other intracellular pathogens including *Mycobacterium bovis* and *Leishmania donovani*.

In addition to providing a tool with which to identify host resistance genes, the ranging susceptibilities across mouse breeds has allowed for the development of models for multiple clinical manifestations of *Salmonella* infection. Chronic models of infection have been developed to study *Salmonella* pathophysiology later during the course of infection. In these models, mice do not succumb to the infection and carry the bacteria for a prolonged period of time in the reticuloendothelial system and mesenteric lymph nodes. Chronic models use *Salmonella* Typhimurium of attenuated virulence in susceptible mice, or sublethal infection with either *Salmonella* Typhimurium or *Salmonella* Enteritidis in resistant mice (114, 115). The similarities in pathophysiology between mice and humans during chronic infection make mice a valuable model

for the future study of the factors leading to chronic carriage, a phenomenon that remains poorly understood.

Mouse models have also been developed to study intestinal inflammation and enterocolitis. *Salmonella* Typhimurium, although a natural mouse pathogen, does not naturally cause intestinal pathologies in mice. Commensal microbes maintain gut immune homeostasis and compete with *Salmonella* for nutrients and space, preventing intestinal colonization, a phenomenon known as colonization resistance (116, 117). To circumvent this obstacle, the group of WD Hardt have developed a model that uses pre-treatment of mice with a single dose of streptomycin to diminish colonization resistance (118). The antibiotic disrupts the normal intestinal flora allowing *Salmonella* Typhimurium to efficiently colonize the large intestine and trigger a severe acute diffuse inflammation of the cecum (119, 120) thereby permitting the study of salmonellosis-like syndrome in mice.

SALMONELLA AND FRESH PRODUCE IN CANADA

As with other developed countries the clinical manifestation of *Salmonella* most relevant to Canadians is salmonellosis. In Canada, 1 in 8 people are affected by foodborne illness each year with 87,500 cases caused by *Salmonella enterica* (121). Among food products, fresh produce is a major contributor to the over-all burden of foodborne illness. Since the 1970's the consumption of fresh produce has been steadily on the rise (122, 123). Today Canadians are consuming over 10% more vegetables than they did 20 years ago (123). Alongside this rising trend in consumption has been an increase in outbreaks, with outbreaks caused by fresh produce jumping from 1% to 12% of all outbreaks over a 20 year period (124). Non-typhoidal *Salmonella* species are currently ranked as the 2nd and 4th most common cause of foodborne related illness in the U.S. and Canada respectively (11).

The increase in salmonellosis caused by fresh fruits and vegetables presents an overt burden for Canadians in terms of health and quality of life. This trend also causes a less obvious economic burden. Health care costs, loss of productivity associated with employee sick days, food wastage, outbreak tracking investigations and other costs add up to make salmonellosis a multi-million-dollar problem. Reducing this illness alone has the potential to save the fresh produce industry up to $2.8 million and the health care system $36 million. With all costs combined, the Canadian economy could be relieved of a financial burden of $50-$60 million (125).

A SYSTEM-WIDE PROBLEM

In the field

From when fruits and vegetables are grown at the farm to when they are on a consumer's plate there are a number of factors along the way that contribute to the impact of salmonellosis. At the level of the field, the problem begins with the fact that *Salmonella* live freely in the environment (126). In addition to a pathogenic lifestyle within a host, *Salmonella* is also found in the water and soil in which crops are grown, marking the initial point of contamination. Contamination is worsened by proximity to livestock. Increasing demand for fresh produce has translated to larger farms with less distance from farms housing animals. A by-product of this expansion has been increased interactions between crops and animals, and subsequently an increase in cross-contamination.

Beyond the problem of the initial contamination is the detriment that there is currently no method to effectively reduce *Salmonella* presence during plant growth. Traditional methods of treating plants contaminated with *Salmonella* include chlorine solution and hot water treatments. While these methods reduce *Salmonella* numbers on the exterior of plant products, they do not

eliminate *Salmonella* that penetrate the surface of the plant products (127, 128). Furthermore, chemical sanitizers such as chlorine solutions react with organic matter yielding harmful disinfection by-products such as chloroform and bromodichloromethane (129). Therefore, safe and effective alternatives are still needed.

Aside from growing produce domestically, a significant proportion of fruits and vegetables consumed in Canada are imported. In 2016 Canada imported approximately 3 million metric tonnes of fresh fruit (Statistics 130) and 2 million metric tonnes of fresh vegetables (Statistics 131). The U.S and Mexico were the top 2 import sources in both categories with Chile, Guatemala, China and Spain following closely behind. International trade affords Canadians year-round access to a wide variety of produce however; sanitation and health regulations vary between countries. Precautions taken in Canada to reduce *Salmonella* contamination are not necessarily taken in other countries (126). Furthermore, information regarding food safety practices of other countries is not always available making it difficult to estimate the risk of contamination (123).

At the processing plant

At the processing, crops are checked for quality assurance, this includes testing for pathogens. Current methods used to detect *Salmonella* on fresh produce include conventional cultural methods, and rapid methods. Cultural methods are often considered the "gold standard" owing to the consistency of their results however, this is a tedious process. In Canada a 6-step process is performed involving (1) non-selective enrichment to grow up *Salmonella*, (2) enrichment in selective media, (3) plating on selective media, (4) plating of presumptive *Salmonella* isolates on additional selective plates, (5) biochemical screening assays to measure other characteristics of the pathogen, and (6) serological identification using polyvalent or somatic antisera. Performing this

process to completion take 5-7 days, a clear disadvantage when wanting to prevent outbreaks and the number of people impacted.

A second disadvantage of cultural methods is that bacteria can enter a vegetative state known as viable-but-nonculturable (VBNC) in which the bacteria are alive, capable of respiration, metabolic functions, and protein synthesis but are unable to be cultured (132, 133). Cells enter this state following various forms of physiological stress and, likewise, can be resuscitated following external stimuli (133) presenting the opportunity for false-negatives during detection screening.

Rapid methods have now become common practice when testing for foodborne pathogens. Rapid methods include immunoassays and molecular assays. Two immunoassays typically used are lateral flow assays (LFA) and enzyme-linked immunosorbent assays (ELISA). Unlike cultural methods, these tests are quick and relatively easy to perform. However, testing isolates from plant products presents unique complications. Cross-reactivity with resident plant microbes such as *Citrobacter* spp., *Proteus* spp., and *Hafnia* spp. results in false positives and decreases test specificity (134). This contributes to the economic burden of *Salmonella* contamination since once a food product tests positive that product and others from that lot are discarded. Testing isolates from plant products can also lead to false negatives. Interference from the sample matrix can decrease the effectiveness of immunoassays and compromise sensitivity, thus increasing the risk of an outbreak (134).

Alternative to immunoassays are molecular assays such as polymerase chain reaction (PCR) (135, 136), real-time or quantitative polymerase chain reaction (qPCR) (137), DNA hybridization (135-138), and DNA microarray (139). These methods are rapid, automated, and reproducible making them advantageous when testing for foodborne pathogens. However, DNA contamination

from the environment and the complex composition of varying food matrices impact specificity and sensitivity respectively (140, 141).

RNA based assays including reverse transcription PCR (RT-PCR) (142) and nucleic acid sequence based amplification (NASBA) (143) have also been used for pathogen detection. Using RNA enables the estimation of *Salmonella* pathogenicity by assessing the expression of virulence associated genes rather than simply testing for pathogen presence (144). Another advantage of RNA based methods is that, like cultural methods, RNA can be used to assess pathogen viability, which has been previously demonstrated in the context of food safety (145-147). It must be noted that the practicality of this method is suboptimal due to labour intensive protocols (148, 149) and the complexity of cell replication. Bacterial replication relies on various pathways and key proteins (150). Confirming the presence of one critical replication protein does not mean that others are not defective and thus, that replication is impeded. In addition, bacteria in a VBNC state exhibit continued gene expression and mRNA production (151) making long-term viability difficult to determine based on the expression of one or a few genes. Independent of viability, false-positives remain a challenge when detecting pathogens by RNA (149).

Beyond the problems associated with pathogen detection assays, changes to fruit and vegetable processing have also contributed to the rise in *Salmonella* outbreaks. Due to their convenience, pre-cut, ready to eat fruits and vegetables have become common in Canadian households. In some cases, the trade-off with convenience has been food safety. Cutting, slicing, peeling and shredding fruits and vegetables removes their protective outer surfaces, allowing for pathogen contamination and growth after initial sanitizing steps (123). As much as a seven-fold increase in pathogen presence has been found following cutting processes of fresh produce, with

cross-contamination via contamination equipment, such as shredders, knives, etc., playing a major role (152).

At the grocery store

When fresh produce arrives at the grocery store it has a relatively short shelf-life due to the desire for freshness, making outbreaks difficult to prevent once contaminated food has reached this point in the trajectory from field to consumer. By the time a grower is identified as being the source of contamination the harvest has already been completed and the produce in question is no longer available for testing (123). If food products are available for testing, pathogen isolates are identified using one of two sequence based subtyping methods: pulse field gel electrophoresis (PFGE) or multilocus sequence typing (MLST). Subtyping is done to discriminate between isolates of the same species, infer genetic relatedness and ultimately link clinical cases to causative sources.

MLST involves sequencing a variety of housekeeping genes. Compared to traditional serotyping methods, MLST has a high degree of discriminatory power given that genotype is assessed by sequencing (4). MLST is also highly efficient, in part due to the fact that data is inputted into a publicly available dataset that is available for global use (153).

When there are privacy concerns regarding sequencing data and publicly available data set are not an option, PFGE is used. PFGE visualizes genetic variability via distinct banding patterns on a gel. The discriminatory power of PFGE is often insufficient to differentiate isolates, particularly those within the same serovar. In a previous study, (154) showed that among 130 *S*. Newport isolates 58 banding patterns were observed. Eight PFGE patters were associated with 64 isolates, with 41 of the 64 sharing a single pattern. Inadequate discrimination between isolates complicates the process

of linking patients to food sources, thereby prolonging investigations and extending the scope of outbreaks.

PROJECT AIMS: THE SYST-OMICS APPROACH

Given that salmonellosis is a system wide problem, a system-wide solution will likely be the most effective at decreasing the overall burden of salmonellosis. The *Salmonella* Syst-OMICs Consortium is an international collaboration aiming to improve current food safety processes by using whole genome sequencing (WGS) at multiple steps along the path from the farm to the table. Our overarching objective is to combine the phenotypic diversity of *Salmonella* with WGS technology to better understand the association between genotype and phenotype. The major challenge facing this task has been the lack of detail provided by currently used genotyping methods. When using assays like PFGE and PCR of known *Salmonella* genes numerous isolates within a single serovar can appear genetically homogenous yet they show distinct phenotypes during infection (155). As a result, the more detailed data retrieved with WGS presents the opportunity for better detection and subtyping of foodborne pathogens. We further hypothesize that WGS-based characterization of *Salmonella* isolates will enable the identification of isolates posing a public health risk.

WGS in combination with isolate phenotyping is being used to identify genetic commonalities among strains of similar virulence. Beyond improving currently used assays, our goal is to inform the development of new pathogen detection assays that use WGS to assess an isolate's risk of pathogenicity rather than simply testing for contamination. Fresh produce harbouring isolates lacking critical virulence associated genes and thus, having no risk of disease in

humans, could be permitted for sale to consumers, thereby decreasing food wastage and, by extension, the economic burden of salmonellosis.

To begin, *Salmonella* isolates were sequenced using whole genome sequencing. At the outset of this project a data base named SalFos was created to gather isolate identification, source category, source location, sequence, and other details for over 2800 isolates. Isolates were collected and contributed by 20 researchers from The *Salmonella* Syst-OMICs consortium. From this database, a subset of 35 isolates were selected for phenotypic testing. Isolates were selected to maximize genetic diversity, thereby allowing the identification of genetic differences underlying varying virulence phenotypes. Isolates were also selected to gather samples from a range of food sources. Food source categories include human outbreaks, seafood, poultry, nuts and seeds, and fresh produce.

Isolates were tested in 4 models of infection: an *in vitro* model using a human epithelial cell line, an *in vitro* model using a human macrophage cell line, an amoeba model, and a mouse model. This was done to characterize isolate virulence following which comparative genomics was used to identify virulence genes common among highly virulent isolates, thereby creating a link between sequence and phenotype (Fig. 3). For our contribution, we assessed isolate virulence *in vivo* using a mouse model. The objectives for this portion of the project were as follows:

1. **1)** Determine the virulence phenotype of 35 *Salmonella* isolate in vivo
2. **2)** Determine genetic commonalities between highly virulent isolates
3. **3)** Develop and test a method for predicting isolate virulence

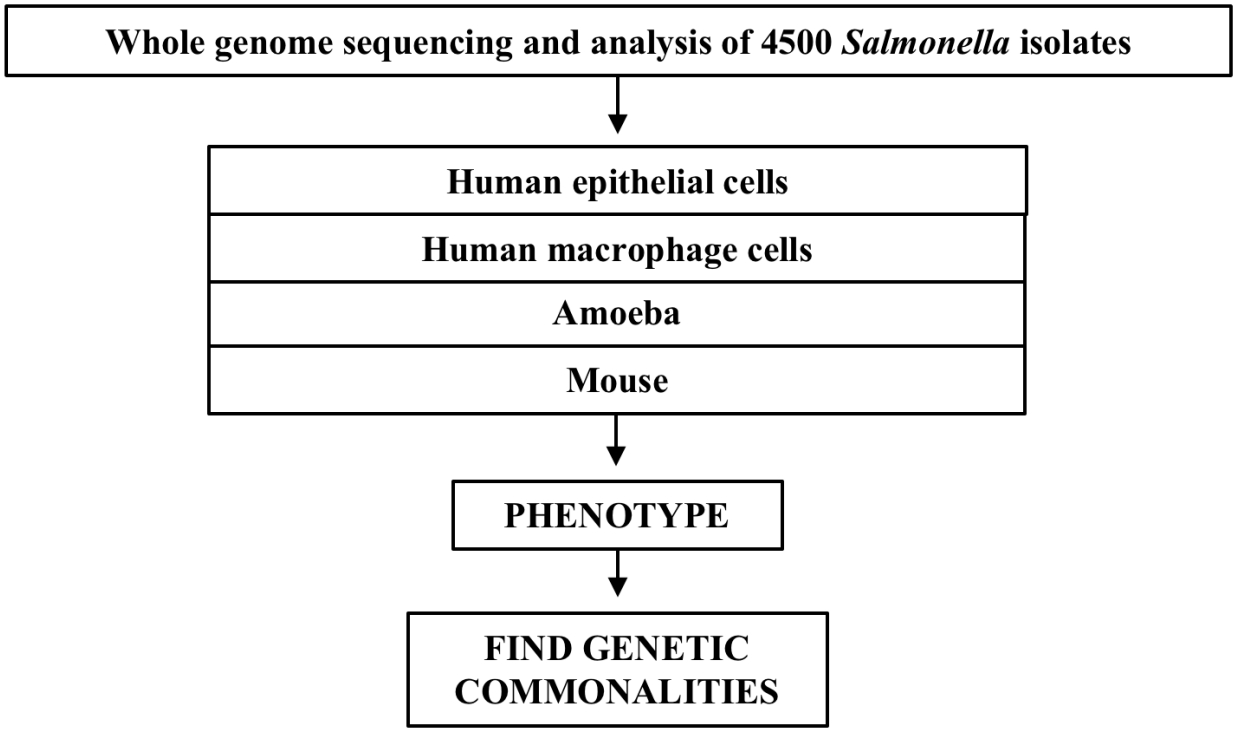

Fig. 3. The *Salmonella* Syst-OMICs phenotyping pipeline.

4500 *Salmonella* isolates were collected and sequenced. A subset of 35 strains was tested in 4 experimental models to determine isolate virulence. These models included human epithelial cells, human macrophages, an amoeba model, and a mouse model. After determining the virulence level of the 35 isolates, comparative genomics was used to identify virulence genes common among virulent strains.

MATERIAL & METHODS

ANIMALS

All animal experiments were performed under guidelines specified by the Canadian Council for Animal Care. The animal-use protocol was approved by the McGill University Animal Care Committee. C57BL/6J mice were obtained from The Jackson Laboratory (Bar Harbor, ME, USA).

SALMONELLA STOCK PREPARATION

To prepare a frozen stock, 1 colony of *Salmonella* was added to 4mL of tryptic soy broth (TSB). Isolates were placed on an oscillating wheel rotating overnight (~16 hours), at 37°C. The following day 1mL of glycerol was added. Stock was frozen at -80°C.

INTRAVENOUS *SALMONELLA* DOSE PREPARATION

1 day prior to infecting, 100µL of frozen *Salmonella* stock was added to 5mL of TSB then placed on an oscillating wheel rotating at 37°C. Isolates were rotated until the mixture reached an optical density (OD) of approximately 1.0 (measured at 600nm). The *Salmonella* was then placed on ice at 4°C for 3-4hours. After this time, 1 in 10 serial dilutions were performed with 10^{-4} and 10^{-5} dilutions being plated on tryptic soy agar (TSA) plates. Plates were incubated overnight at 37°C and tubes containing the *Salmonella* were kept at 4°C overnight. The following day colony forming units (CFUs) were counted to determine the concentration of *Salmonella*. Based on this concentration, *Salmonella* was diluted with 0.9% saline to achieve a concentration of 500 CFUs/µL. Mice were injected with 200µL of 10^5 CFUs via the caudal vein.

PER OS *SALMONELLA* DOSE PREPARATION

1 day prior to infection, 3mL of TSB was inoculated with 1 colony of *Salmonella* then incubated shaking at 220 RPM, 37°C for 17.5 hours. The day of the infection, food and water were removed 4 hours prior to infection. Mice were infected by gavage with 100µL of *Salmonella* Food and water were removed until 2 hours after infection.

IN VIVO TISSUE HARVEST

Mice were euthanized on day 3 post infection after which the spleen, liver, and blood were collected. For per os infected mice, feces were also collected. Bacterial burden in the spleen, liver, and feces was determined by plating serial 1 in 10 dilutions (dilutions done with 0.9% saline) of organ and feces homogenate on TSA plates. Spleen enlargement was measured by spleen index which is calculated by the following equation: $\sqrt{\frac{\text{spleen weight}/10}{\text{body weight}}}$

CYTOKINE ANALYSIS

Blood was collected by cardiac puncture 3 days post-infection with *Salmonella.* Blood was collected into clot activating tubes (Sarstedt). Serum was collected by spinning blood at 2000 RCF for 5 minutes at room temperature. Interleukin-6 levels in the blood serum was determined by enzyme-linked immunosorbent assay (ELISA) using Ready-SET-GO! kits (eBiosciences).

CELL CULTURE

RAW 264.7 macrophages were cultured in Dulbecco's Modified Eagle's Medium (DMEM, Thermo Fisher Scientific) supplemented with 10% fetal bovine serum (FBS, Thermo Fisher Scientific) and 12.5mL of Hepes (Sigma-Aldrich). Cells were kept at 37°C, 5% CO_2.

SALMONELLA PHAGOCYTOSIS ASSAY

Salmonella isolate phagocytosis and survival within murine macrophages was assessed *in vitro* by infection RAW 264.7 macrophages using a gentamicin protection assay (156). *Salmonella* isolates were grown overnight in (1 colony in 10mL of TSB) oscillating at 220 RPM, 37°C. The following day *Salmonella* were opsonized with mouse serum (100µL diluted in 500µL D-PBS). $5x10^5$ cells were plated per well in 24 well flat bottom tissue-culture treated plates (Sarstedt). Following 24 hrs at 37°C, 5% CO_2 cells were infected with *Salmonella* isolates at a multiplicity (MOI) of ~10-20. Infected cells were centrifuged for 5 minutes at 1250 RPM, room temperature then incubated for 40 minutes 37°C, 5% CO_2 after which cells were washed with D-PBS (Wisent Inc.). Media containing 100µg/mL of gentamicin was added for 1 hour at 37°C, 5% CO_2 to kill extracellular bacteria. To measure intercellular bacterial load at timepoint 1 hour post-infection cells were washed twice then lysed by adding D-PBS (Wisent Inc.) with 1% TritonX100. Cell lysate was diluted via 1 in 10 serial dilutions in 0.9% saline and plated on TSA plates. For timepoints longer than 1 hour, following 1 hour incubation with 100µg/mL of gentamicin cells were washed twice incubated with 10µg/mL of gentamicin. Cells were incubated for 2 or 24 hours post infection after which time cell lysates were collected and plated.

STATISTICS

Statistical analyses were performed using GraphPad Prism. Tests used are specified in figure legends.

CORRELATION ANALYSIS

The data was considered as an unbalanced mixed factorial design composed of 258 mice that were infected with 35 isolate of *Salmonella* and two control isolates, the highly virulent control strain was applied to 23 mice and the low virulent control strain to 30 mice. Measurements were made on each mouse in two tissues: liver and spleen (except for 2 mice where measurements were only available in the spleen).

The model we used is as follows:

$$log_{10}(CFU)_{ik} = \mu + S_i + T_k + S_i * T_k + M_i + e_{ik}$$

where S (Strain) and T (Tissue, i.e. Spleen or Liver) are considered as fixed effects and M (Mice) as a random effect. i refers to the mouse and k to the condition. e corresponds to the error term which we assumed to be normally distributed, $e_{ik} \sim N(0, \sigma^2)$. Type III tests of hypothesis were computed to evaluate the significance of the Strain*Condition interaction on the outcome. Contrasts were constructed to evaluate the difference in log10 (CFU) between each Salmonella strain and the low virulent control strain, for each condition (tissue), yielding 70 comparisons. Test results and confidence intervals obtained for each contrast were adjusted for multiple comparisons using Bonferonni's correction.

PCA ANALYSIS

DNA sequencing

Genomic DNA was extracted from overnight LB broth cultures at 37°C using the E-Z 96 Tissue DNA Kit (Omega Bio-tek, Norcross GA, USA). Around 500 ng of genomic DNA was mechanically fragmented during 40 seconds by Covaris M220 (Covaris, Woburn MA, USA) using

the default settings. Libraries were synthesized using NEBNext Ultra II DNA library prep kit for Illumina (New England Biolabs, Ipswich MA, USA) according to manufacturer's instructions and were sequenced by Illumina MiSeq 300 bp paired-end run at the Plateforme d'Analyses Génomiques of the Institut de Biologie Intégrative et des Systèmes (Laval University, Quebec, Canada).

Pan-genome analysis

Pan-genome analysis of 52 *Salmonella* strains, of which 12 were high virulent, 7 were low virulent and 33 strains had an unknown virulence phenotype, was performed using SaturnV (v1.1.0) (https://github.com/ejfresch/saturnV) (Freshi *et al*., in press) with ≥ 90% of identity between proteins and ≥ 85% of alignment covering their length to be considered as ortholog proteins.

Principal component analysis (PCA)

The table of ortholog proteins was converted to binary table using a perl script and imported into R. Ade4 package was used to perform the principal component analysis (PCA) (> data_pca <- dudi.pca(*table.binary.tsv*, scale = F, scannf = FALSE, nf = 5)). The plot was created using the two first principal components and the strains were colored based on their virulence levels (red = high virulent, blue = low virulent and green = unknown virulence phenotype) (> plot(data_pca$li[,1:2], col=color_t_condition, pch =19)).

RESULTS

SELECTING AN *IN VIVO* PROTOCOL FOR TESTING *SALMONELLA* ISOLATES

Treatment groups and Salmonella controls

Given the various ways to model Salmonella infection in mice, we sought to determine the best approach to study the virulence of isolates which had little known about them. The goal was to select a method that provided a high degree of discrimination and sensitivity to tease out phenotypic differences between isolates and to ensure that isolates with pathogenic potential were identified.

Two routes of infection were tested: intravenous (I.V.) injection via the tail vein, and per os (P.O.) by gavage. I.V. injection allows direct administration of Salmonella to the blood stream, allowing bacteria to bypass the harsh environment of the stomach and circumvent the process of adhesion and invasion into intestinal epithelial cells. As a result, I.V. injection offers high sensitivity to Salmonella virulence. P.O. infection, on the other hand, was tested as this model more accurately represents the natural route of infection in both animals and humans.

In addition to comparing different routes of administration, varying doses were also tested. Pathology was assessed following a dose of 10^4 colony forming units (CFUs) or 10^5 CFUs via I.V. injection. One dose of 109 was tested during P.O. infections. Doses were relatively high to better reveal the pathogenic potential of isolates. Combining all doses and routes of infection, a total 3 treatment groups were assessed: I.V. 10^4, I.V. 10^5, and P.O. 10^9.

Susceptible C57BL/6J mice were used for all infections to increase the sensitivity and discrimination of Salmonella screening. Using a more resistant mouse breed may have masked the pathogenic potential of highly virulent isolates and/or failed to show a difference between isolates of varying levels of virulence, given that these mice would be able to overcome infections from isolates with no pathogenic potential as well as those with higher propensity for disease.

Furthermore, using a mouse model with high susceptibility to Salmonella was important for translating this work to humans. As in mice, varying levels of susceptibility are also observed in humans. Isolates labeled as low virulent in mice must also pose little to no risk in humans of varying degrees of susceptibility.

Along with establishing a model of infection, control strains were also tested. S. Typhimurium SL1344 (here referred to as WT SL1344) was used as a positive control for infection. SL1344 is highly effective at invading and surviving within various cell types in vitro (Clark et al. 2011) and is also highly virulent in vivo (Chatfield et al. 1991). An SL1344 derivative harbouring mutations in genes invA and sseB was tested as a low virulent control. This strain was created by our collaborator France Daigle (Université de Montréal) and had yet to be tested in vivo. Salmonella strains with mutations in invA exhibit defective cell entry and decreased virulence in vivo marked by a decreased ability to colonise the Peyer's patch and intestinal wall (Galan and Curtiss 1989). Disruption of invA alone however, does not impair the ability of Salmonella to elicit a systemic infection as has been demonstrated by intraperitoneal infection (Galan and Curtiss 1989). Strains with mutations or deletions in sseB result in decreased replication within macrophages and reduced virulence in vivo (Hensel et al. 1998). We therefore, postulated that with deletions in both invA and sseB, SL1344ΔinvAΔsseB (here referred to as delta SL1344) would display attenuated virulence in vivo compared to wild-type SL1344, and act as a suitable low virulent control.

Mice were euthanized 3 days post-infection at which time tissue and blood were collected (Fig.1.). Given that systemic Salmonella infection in mice is characterized by high bacterial load in the spleen and liver (Richter-Dahlfors et al. 1997; Salcedo et al. 2001), decreased body weight, spleen enlargement (Shea et al. 1999), and increased levels of blood serum pro-inflammatory cytokines (Klimpel et al. 1995), we used these parameters to measure isolate virulence. Spleen index, which acts as a ratio between spleen weight and body weight, was used as a measure of spleen

enlargement. Bacterial burden in the feces was also measured 3 days post P.O. infection to assess bacterial colonization of the intestines.

Virulence of SL1344 and delta SL1344 in 3 treatment groups

In each of the three models of infections delta SL1344 exhibited an attenuation in virulence. Mice infected with delta SL1344 showed a lower bacterial burden in the spleen and the liver compared to those infected with WT SL1344 (Fig. 2.A-B). I.V. infection with 10^5 CFUs produced the largest difference in splenic bacterial burden, with a 5.5 log difference in CFUs between the two strains. This was compared to a 4.2 log difference in I.V. 10^4 infected mice and 2.7 log difference in P.O. infected mice (Fig.2.A). Differences between strains were more moderate in the liver with a 4 log difference in I.V. 10^4 infected mice, a 3.8 log difference in I.V. 10^5 infected mice, and a 3.9 log difference in P.O. infected mice (Fig.2.B). In terms of spleen enlargement, both groups of I.V. treated mice showed a higher spleen index in mice infected with WT SL1344 compared with delta SL1344 while mice infected by P.O. showed no difference (Fig.2.C). All groups showed a decrease in body weight following infection with WT SL1344, with those in the I.V. 10^5 treatment group showing the greatest loss (13.4%) (Fig.2.D). I.V. 10^5 treated mice were also the only ones to show a difference in serum cytokine levels with those infected with WT SL1344 having a mean IL-6 serum level of 2527pg/mL compared to mice infected with delta SL1344 which had a mean level of 38pg/mL (Fig.2.E). Finally, mice infected orally showed no difference in bacterial burden in the feces between WT SL1344 and delta SL1344. (Fig. 2.F).

Based on these findings we selected an I.V. infection model with a dose of 10^5 CFUs to test *Salmonella* isolates of unknown virulence as this group yielded the greatest difference in virulence between the two strains. It was also concluded that WT SL1344 and delta SL1344 would be used in

future experiments as positive and negative controls of infection. These strains would serve as reference strains with which to compare isolates of unknown virulence. These strains provided a basis to show how an isolate of low vs. high virulence would behave in our model of infection.

EVALUATING 35 *SALMONELLA* ISOLATES OF UNKNOWN VIRULENCE

Virulence in vivo

Approximately half (18/35) of all isolates resulted in splenic bacterial burdens significantly lower than that of delta SL1344 (Fig.3.A). 10 of the 35 isolates resulted in bacterial burdens that weren't significantly different from delta SL1344, and 7 showed a significant increase in bacterial burden. Only 2 isolates (S3 and S5), both from the serovar enteritidis, had bacterial burdens similar to highly virulent WT SL1344. A similar distribution was seen in the liver with 19/35 isolates resulting in bacterial burdens lower than delta SL1344, 8/35 that did not differ, and 8/35 resulting in a bacterial burden higher than delta SL1344 (Fig.3.B).

The spleen index, blood serum IL-6 levels and percent body weight change were less variable between isolates with the majority (28/35, 30/35, and 26/35 respectively) showing no difference compared to delta SL1344 (Fig.3.C-E). Strains that did differ in one or more of these 3 parameters typically also differed in terms of bacterial burden. Therefore, if an isolate caused an increased spleen index, increased IL-6 serums levels and/or loss in body weight, it also caused an increase in bacterial burden. Exceptions were S26 and S39 which caused increased levels of IL-6 (mean of 1009pg/mL and 1102pg/mL respectively) but no increase in bacterial burden, and S52 which causes a significant drop in body weight (5.1% loss), with no increase in bacterial burden.

Each of the isolates tested were placed into one of three categories depending on their behaviour *in vivo*. Isolates that caused an increased bacterial burden in the liver and/or spleen relative to delta SL1344 were considered "high virulence" strains. Those whose bacterial burden did not differ from delta SL1344 were considered "intermediate virulence" and all others were considered "low virulence". Therefore, the group of 35 isolates was composed of 9 high virulence isolates, 10 intermediate isolates, and 16 low virulence isolates.

Extended-term experiments

Since our infection model assessed virulence 3 days post-infection, we performed two extended-term experiments to assess how virulence might change with a longer incubation period. It was important to ensure that isolates labelled as low virulent did not appear non-pathogenic because the time over which they were followed was too short. A subset of 5 isolates was selected to test in a 7-day infection model: 2 low virulent isolates (S29 and S37), 2 intermediate isolates (S27 and S30), and 1 highly virulent isolate (S33).

Compared to findings 3 days post-infection, no difference in bacterial burden in the spleen was seen at day 7 for delta SL1344, both low virulent isolates, and highly virulent S33 (Fig.4.A). Opposing results were seen for strains of intermediate virulence. A 0.7 log decrease in bacterial burden was seen for mice infected with S30 while a 0.7 log increase was seen in those infected with S27 (Fig.4.A). In the liver, mice infected with isolates S27, S30, and S33 showed a 0.5 log, 1.5 log, and 1.8 log reduction in the liver at day 7 compared to those assessed at day 3 (Fig.5.B) while delta SL1344, S27, and S37 showed no difference (Fig.4.B). No differences were seen with respect to body weight change or spleen index for any of the five isolates (Fig.4.C-D). Overall, the majority of isolates either showed no change or a reduction in pathology following a longer incubation period.

A survival experiment was also performed to further assess long-term health following infection with S27, S29, S30, and S37. The mice were followed for 30 days during which time none of the mice died or exhibited any symptoms of illness (data not shown).

Virulence in murine macrophages

To further explore isolate behaviour and to bridge the gap between the *in vivo* mouse data and that of other *in vitro* models used in the Syst-OMICs phenotyping pipeline, a subset of the 35 isolates tested *in vivo* was assessed *in vitro* in RAW 264.7 murine macrophages. Two strains from each virulence category (highly virulent S5 and S428, intermediately virulent S27 and S30, and low virulent S29 and S37) along with control strains WT SL1344 and delta SL1344 were tested. Isolates were assessed for their propensity for uptake into, and survival within mouse macrophages. This was done by infecting RAW 264.7 macrophages, then adding gentamicin to kill extracellular bacteria, and plating cell lysate to measuring intracellular levels of *Salmonella*. The intracellular bacterial load was measured at three time points; 1 hour, 2 hours, and 24 hours after infection.

The largest difference between isolates was observed 24 hours post-infection (Fig.5). As was the case *in vivo*, WT SL1344 resulted in a greater bacterial burden than delta SL1344, with 1.0 log difference between the two. Other isolates showed a similar trend in murine macrophages as was seen *in vivo*. Highly virulent S5 produced the highest bacterial burden with 1.3 log more bacteria than delta SL1344, and 0.7 log more than WT SL1344. Other highly virulent isolate S428 lead to a bacterial burden 0.7 log higher than delta SL1344. Intermediately virulent S27 and S30 both lead to bacterial burdens between those of delta SL1344 and WT SL1344. The bacterial burden of S27 was 0.8 log greater than delta SL1344 and 0.2 log lower than WT SL1344. For S29, the bacterial was 0.9 log greater than delta SL1344 and 0.1 log lower than WT SL1344. Finally, low virulent S37 also

behaved similarly *in vitro* as was seen *in vivo* with a bacterial burden 1.3 log lower than that of delta SL1344.

The one isolate that largely deviated from *in vivo* findings was S29. *In vivo*, S29 was low virulent with a bacterial burden significantly lower than delta SL344. In macrophages *in vitro* however, S29 produced a bacterial load 0.9 log higher than delta SL1344 and did not significantly differ from high virulent control WT SL1344. It must be noted that SL1344, as well as S428, have septation defects when grown *in vitro*. Although bacteria from these isolates replicate, they are unable to properly cleave from one another. It has been shown by a group among our collaborators (Gisèle Lapointe, University of Guelph) that this defect reduces the number of colonies observed when plating the cell lysate, thereby resulting in an underestimation of bacterial load. With this is mind, S29 may actualy result in a bacterial burden lower than WT SL1344.

Isolate virulence, phylogeny and source

Using the genome sequence of the 35 isolates, a phylogenetic tree was constructed, displaying the genetic relationship between the isolates (Fig.6.). Within this tree a relationship between isolate sequence and virulence was also observed. Isolates of similar virulence, particularly high virulence isolates, grouped together. 7 out of 9 high virulence isolates clustered together indicating sequence commonalities underlying virulence.

A connection between virulence and isolate source was also observed. Highly virulent isolates had the lowest variety in source categories with isolates coming from either human outbreaks or seafood. 4 out of 5 of samples isolated from human outbreaks were highly virulent in mice with the remaining 1 isolate being of intermediate virulence. Intermediately virulent source categories included: human outbreaks, seafood, or “other” sources. “Other” sources are those other

than: fresh produce, animal, poultry, dairy, nuts and seeds, environmental, seafood, and human outbreaks. In this case "other" included animal feed and chocolate. Low virulent isolates showed the greatest source variety with categories including nuts and seeds, fresh produce, poultry, animal, seafood sources, and "other", with seafood being the largest contributing source. Of note, all isolates from fresh produce were of low virulence.

CORRELATION ANALYSIS ACROSS 4 MODELS OF INFECTION

The 35 isolates tested *in vivo* were also tested in 3 other models of infection: a human epithelial cell line, a human macrophage cell line, and in an amoeba model. To test concordance across models, a correlation analysis was performed. This analysis compared various measures of virulence from each model. In human epithelial cells, three parameters were measured to assess virulence: bacterial adhesion, invasion, and survival within cells. Three parameters were also measured in human macrophages. These included phagocytosis, survival and replication within macrophages. In the amoeba model intracellular bacterial burden was the sole parameter measured. Finally, two parameters from the *in vivo* studies were used, bacterial burden in the spleen and bacterial burden in the liver. All parameters measured bacterial numbers using log CFUs. Rather than using the raw CFU data, the bacterial burden was normalized relative to delta SL1344. The mean CFU value caused by infection with delta SL1344 was set to zero. The difference between the bacterial burden of delta SL1344 and the bacterial burden caused by each of the 35 isolates was plotted showing the contrast between each isolate and delta SL1344 (Fig.7). The contrast data was used to perform correlation analyses in which each of parameters from a given model was compared against every other parameter from other models. Concordance between all combinations of parameters made for 36 comparisons (Fig.8).

Of all comparisons, 86% (31/36) were positive and showed significant correlations. The 5 comparisons that were not significant included: (i) epithelial cell adhesion vs. burden in the mouse liver (r=0.32, CI [-0.02, 0.58]), (ii) survival within epithelial cells vs. burden in the mouse liver (r=0.28, CI [-0.06, 0.56]), (iii) bacterial burden in amoeba vs. invasion into epithelial cells (r=0.31, CI [-0.02, 0.58]), (iv) bacterial burden in amoeba vs. uptake by human macrophages (r=0.14, CI [-0.19, 0.45]), and (v) bacterial adhesion to epithelial cells vs. survival within epithelial cells (r=0.31, CI [-0.02, 0.58]). In contrast, the 5 comparisons with the highest correlation coefficients were (i) bacterial burden in the mouse spleen vs. mouse liver (r=0.93, CI [0.86, 0.96]), (ii) epithelial cells invasion vs. epithelial cell survival (r=0.79, CI [0.63, 0.89]), (iii) uptake into macrophages vs bacterial burden in macrophages 2 hours post-infection (r=0.7, CI [0.49, 0.84]), (iv) bacterial burden in the mouse spleen vs. bacterial burden in amoeba (r=0.67, CI [0.43, 0.82]), and (v) survival within epithelial cells vs. survival within macrophages (r=0.64, CI [0.39, 0.8]).

Specific to the *in vivo* model, the model with the strongest correlation with the mouse data was the amoeba. Comparing the amoeba model with bacterial burden in the mouse spleen yielded a correlation coefficient of r=0.67, CI [0.43, 0.82], and comparing the amoeba with the bacterial burden in the liver resulted in a correlation coefficient of r=0.59, CI [0.32, 0.77]. The model with the lowest level of concordance with the mouse model was the human epithelial cell model. While all correlations of human epithelial cell vs. mouse spleen were significant, 2 out of 3 correlations with the mouse liver were not. These comparisons were with liver vs. adhesion (r=0.32, CI [-0.02, 0.58]) and liver vs. survival (r=0.28, CI [-0.06, 0.56]).

IDENTIFYING VIRULENCE ASSOCIATED GENES

As previously mentioned, the clustering of highly virulent isolates during phylogenetic analysis suggested the presence of genetic commonalities underlying this phenotype. To identify

potential virulence associated genes two comparative analyses were performed each comparing the genomes of 3 high virulence isolates and 2 low virulence isolates. These analyses aimed to find genes present exclusively in highly virulent isolates. A total of 5 genes were identified: *pipB2* and *sopD2*, *shdA*, *sodCl*, and *sopE.*

PipB2, SopD2, ShdA, and SodCI have all been previously identified as virulence associated gene with *Salmonella* isolates carrying mutations in one of each these genes all having been shown to experience attenuated virulence *in vivo* (67, 68, 89, 157). Although SopE has been identified as a virulence associated gene as well, this had yet to be shown *in vivo. In vitro* studies have shown that SopE plays aids in triggering cytoskeleton rearrangement involved in membrane ruffling during *Salmonella* invasion (89, 158-160).

To assess the role of *sopE* as a virulence associated *gene in vivo* we tested a strain of SL1344 carrying a deletion in *sopE* (delta *sopE*) in our mouse model. The virulence of delta *sopE* was compared to high virulent control WT SL1344, low virulent control delta SL1344, as well as an addition strain of SL1344 which carried a deletion only in *invA* (delta *invA*). Compared to both WT SL1344 and delta *invA*, delta *sopE* resulted in a lower bacterial burden in the spleen, however a higher burden than delta SL1344 was observed. Similarly, in the liver delta *sopE* yielded a significantly lower bacterial burden than WT SL1344 and a higher bacterial burden than delta SL1344 (Fig.9.B). In the liver there was no significant difference between delta *sopE* and delta *invA*. In terms of spleen index, a significant difference was only observed between delta *sopE* and delta SL1344 (Fig.9.C). Finally, delta *sopE* lead to a significantly lower drop in body weight compared to WT SL1344 (Fig.9.D). Decreased virulence appeared to be *in vivo* specific as *in vitro* studies showed normal growth and colony formation of delta *sopE.*

PREDICTING SALMONELLA VIRULENCE BASED ON SEQUENCE

17 isolates used to predict virulence

17 *Salmonella* isolates were selected to be tested *in vivo* in mice and *in vitro* in human epithelial cells. These isolates were chosen based on the presence or absence of the 5 virulence associated genes identified from the comparative genomic analyses. Three of these isolates were from the initial 35 isolates tested (S3, S37, and S428) and 14 were additional isolates that were tested.

Following isolate phenotyping it was found that the 17 isolates had similar phenotypes *in vitro* and *in vivo*. *In vitro*, 7 isolates were low virulent and 10 were highly virulent. Similarly to what was done for the initial 35 isolates tested, virulence *in vivo* was measured by bacterial burden in the spleen and liver, spleen index, and changes in body weight (Fig.9. A-D). 16 of the 17 isolates were found to have the same virulence categorization *in vivo* and *in vitro*. S357 was the only isolate that did not display a concordant phenotype. *In vitro*, S357 was low virulent while *in vivo* it produced an intermediate phenotype. S357 resulted in a splenic bacterial burden significantly lower than delta SL1344 however, the bacterial burden in the liver did not differ from delta SL1344 making this an intermediate virulent isolate. Due to the similar behaviour of these 17 isolates overall, they were selected to begin the process of creating a method to predict isolate virulence based on genomics.

The whole genome sequence of the 17 *Salmonella* isolates was analyzed via principal component analysis (PCA) to visualize the relationships between sequence and phenotype (Fig.11). The PCA revealed that isolates appear to cluster with other isolates of similar virulence, indicating a link between sequence and virulence. Based on this finding, it was hypothesized that the position of isolates within the plot, as determined by their sequence, could be used as a predictor of virulence. 33 isolates of unknown virulence were then added to the PCA and their virulence was predicted based on their proximity to isolates of known virulence. 18 isolates clustered with isolates of known

high virulence and were therefore predicted to be highly virulent. The other 15 isolates clustered with known low virulent isolates and were predicted to be low virulent. The next step was to test theses isolates *in vitro* and *in vivo* to assess prediction accuracy.

Testing virulence predictions

Out of 33 isolates whose virulence was predicted, 15 were predicted to be low virulent and 18 were predicted to be highly virulent. Of the 15 low virulent predicted isolates, 11 lead to bacterial burdens significantly lower than delta SL1344 in the spleen, and 4 produced splenic bacterial burden significantly greater than that of delta SL1344 (Fig.13.A). 9 of the isolates that yielded lower bacterial burdens in the spleen also lead to lower burdens in the liver compared to delta SL1344, and 2 did not differ from delta SL1344 liver. Of the 4 isolates causing bacterial burdens higher than delta SL1344 in the spleen, 1 also yielded a bacterial burden higher than delta SL1344 in the liver, 2 lead to burdens lower than delta SL1344, 1 did not differ from delta SL1344 (Fig.13.B). In terms of spleen index 10 isolates lead resulted in a lower spleen index than delta SL1344 while the remaining 5 showed no difference relative to delta SL1344 (Fig.13.C). Finally, all low virulent predicted isolates showed no different in percent body weight change as compared to delta SL1344 (Fig.13.D). In total of the 15 isolates that were predicted to be low virulent, 9 were correctly predicted based on *in vivo* data, 4 were actually highly virulent *in vivo*, and 2 were intermediately virulent.

The remaining 18 isolated were predicted to be highly virulent. Of these 13 resulted in bacterial burdens significantly higher than that caused by delta SL1344, 4 showed no difference, and one lead to a burden lower than delta SL1344 (Fig.13.A). In the liver, only 4 isolates yielded bacterial burdens greater than delta SL1344 while 9 showed no difference, and 5 were lower than delta SL1344 (Fig.13.B). For spleen index, the majority of isolates (14) did not differ from delta

SL1344. 2 isolates were lower than delta SL1344, and 2 isolates were lower (Fig.13.C). Similarly, 16 isolates showed no difference relative to delta SL1344 in terms of percent body weight change, with 2 causing significantly more weight loss than delta SL1344 (Fig.13.D). For highly virulent predicted isolates, 13/18 were correctly predicted, 4 were intermediate and 1 was low virulent.

Overall, 22 out of 33 isolates (67%) showed virulence phenotypes *in vivo* that matched their predictions. 5 isolates were misclassified and 6 caused intermediate phenotypes. The validity, and therefore success rate, of the PCA as a prediction method can be further described in terms of sensitivity and specificity, which allow the accuracy of this method for high and low virulent predictions to be assessed independently. Sensitivity measures the ability to correctly identify positive cases. For our purposes, sensitivity measures the ability of the PCA to correctly predict high virulent isolates. The higher the sensitivity, the lower the number of false negatives. Sensitivity is calculated by the following:

$$\frac{\#\text{ of correct high virulent predictions}}{\#\text{ of correct high virulent predictions} + \#\text{ of false negatives}}$$

Specificity measures the ability to correctly identify negative cases. In the context of *Salmonella* virulence, specificity measures the ability of the PCA to correctly predict low virulent isolates. The higher the specificity, the fewer false positives. Specificity is calculated by the following:

$$\frac{\#\text{ of correct low virulent predictions}}{\#\text{ of correct low virulent predictions} + \#\text{ of false positives}}$$

For these calculations, isolates yielding intermediate phenotypes were excluded as these were neither correctly predicted nor false negatives/positives.

Out of 17 isolates that were highly virulent *in vivo*, 13 were accurately predicted to be highly virulent, yielding a sensitivity of 0.77 or 77%. Conversely, 23% of highly virulent isolates in vivo were mistakenly predicted to be low virulent (false negatives). Of the 10 isolates that were low virulent *in vivo*, 9 were correctly predicted to be low virulent, yielding a specificity of 0.9. Therefore, 90% were correctly predicted while 10% were false positives.

Fig. 1. Infection protocol timeline.

Mice were infected with approximately 10^5 CFUs of *Salmonella* and monitored for 3 days. On the third day post-infection mice were euthanized, and the spleen, liver and blood were collected.

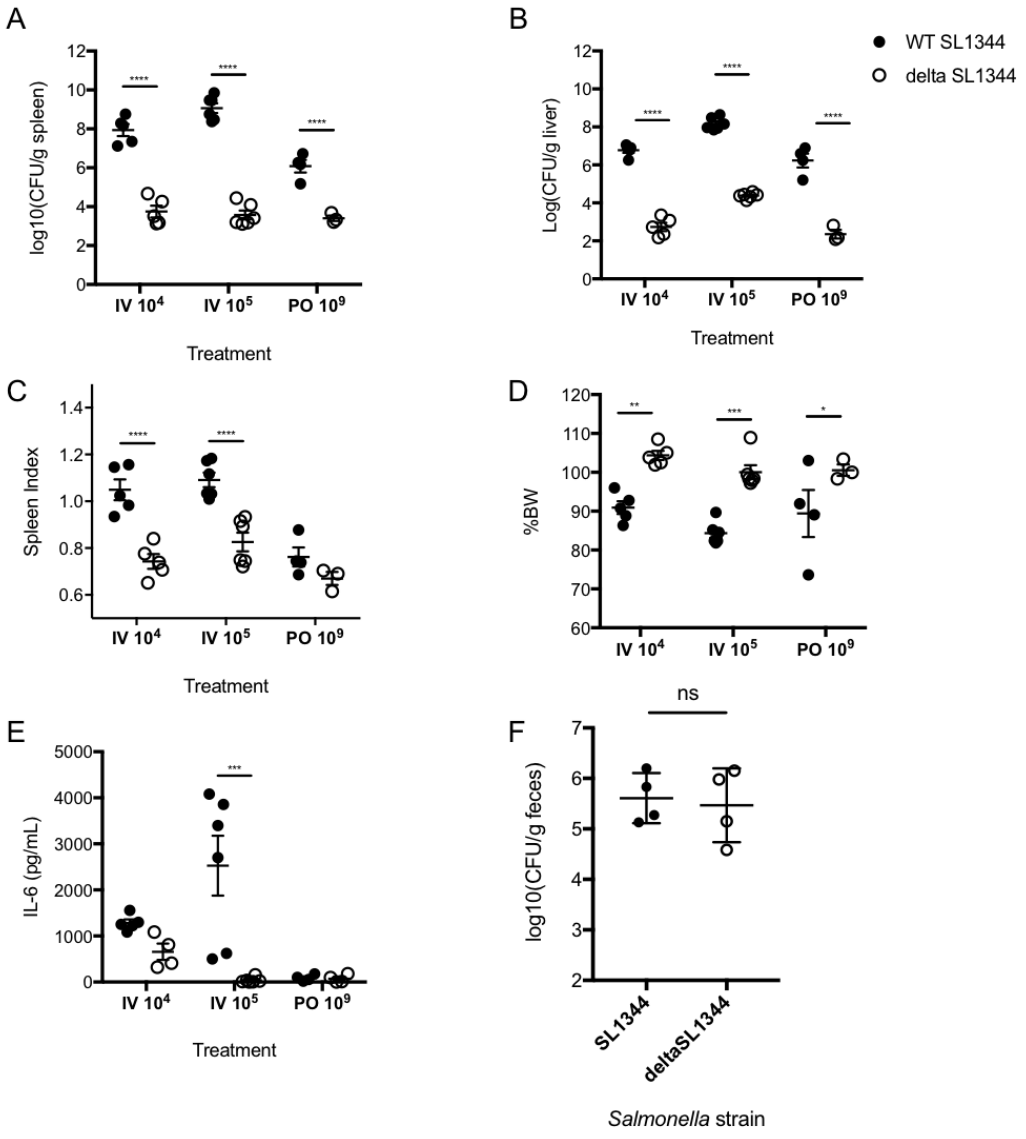

Fig. 2. Selecting an *in vivo Salmonella* testing protocol

Three infection protocols were tested to determine the best protocol for discriminating between *Salmonella* isolates of varying levels of virulence *in vivo*: (i) I.V. infection with 10^4 CFUs (IV 10^4), (ii) I.V. infection with 10^5 CFUs (IV 10^5), and (iii) oral infection via gavage with 10^9 CFUs (PO 10^9). Bacterial burden in the (A) spleen and (B) liver, (C) spleen enlargement, (D) changes in percent body weight, and (E) blood serum IL-6 levels were used as parameters to measure virulence. (F) Bacterial burden in the feces of PO 10^9 infected mice was also measured. Two *Salmonella* isolates were used: highly virulent wild-type SL1344 (WT SL1344), and delta SL1344 which harbours mutations in virulence associated genes *invA* and *sseB*. Statistics were done using a two-way ANOVA with significance levels denoted as: ns= $p > 0.05$, *= $p \leq 0.05$, **= ≤ 0.005, ***= $p \leq 0.0005$, ****= $p < 0.0001$.

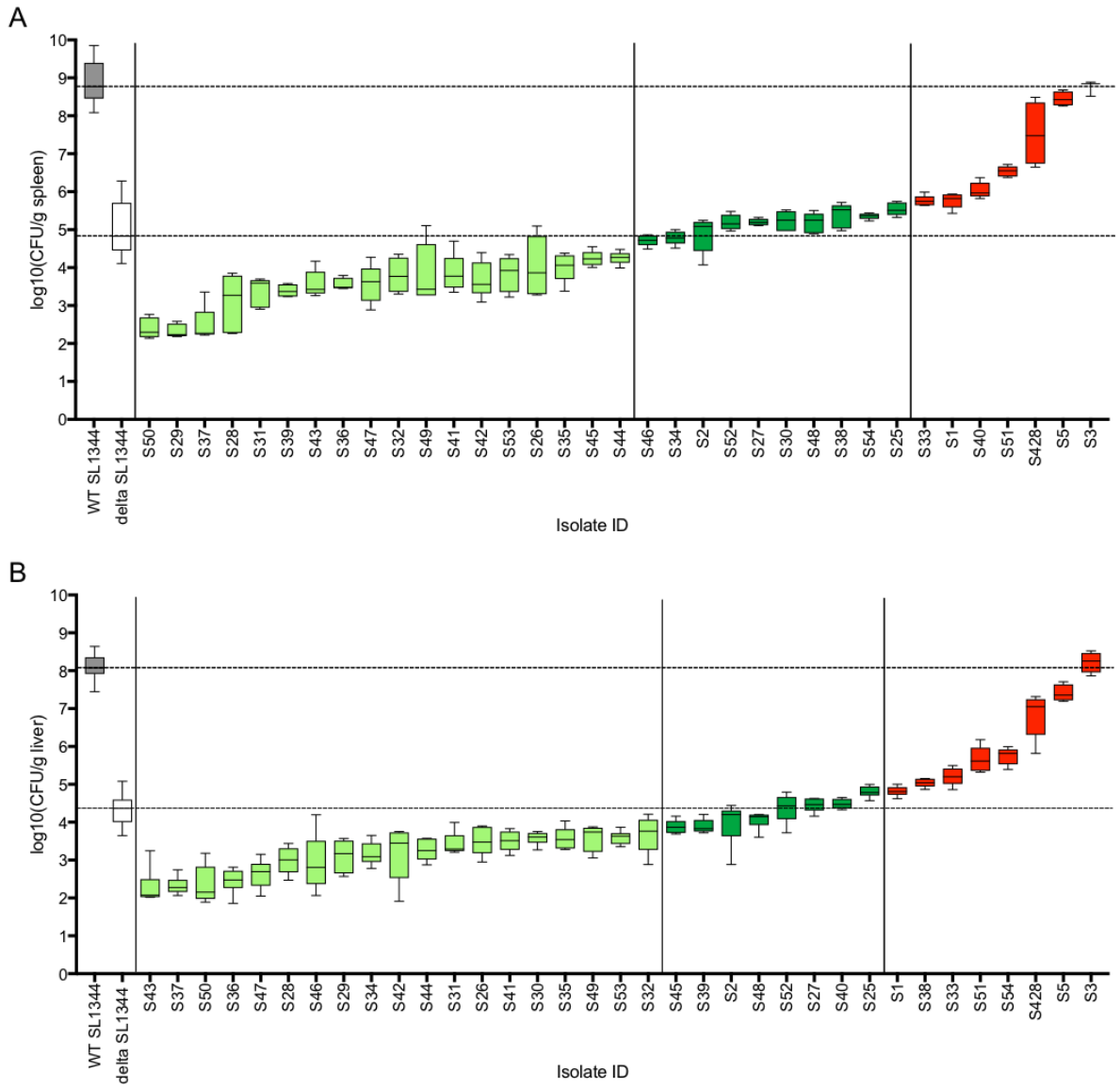

Fig. 3 Virulence of 35 *Salmonella* isolates *in vivo*.

35 strains of unknown virulence were tested *in vivo*. Virulence was assessed based on (A) bacterial burden in the spleen, (B) bacterial burden in the liver, (C) spleen index, (D) change in percent body weight, and (E) blood serum IL-6 levels. Isolates are compared to delta SL1344 using a one-way ANOVA. Isolates causing a bacterial burden significantly lower than that cause by delta SL1344 ($p \leq 0.05$) are light green. Those that are not significantly different ($p > 0.05$) are dark green, and those that caused a bacterial burden higher than delta SL1344 ($p \leq 0.05$) are red. This colour scheme is used for all measures of virulence. Figure is continued on next page.

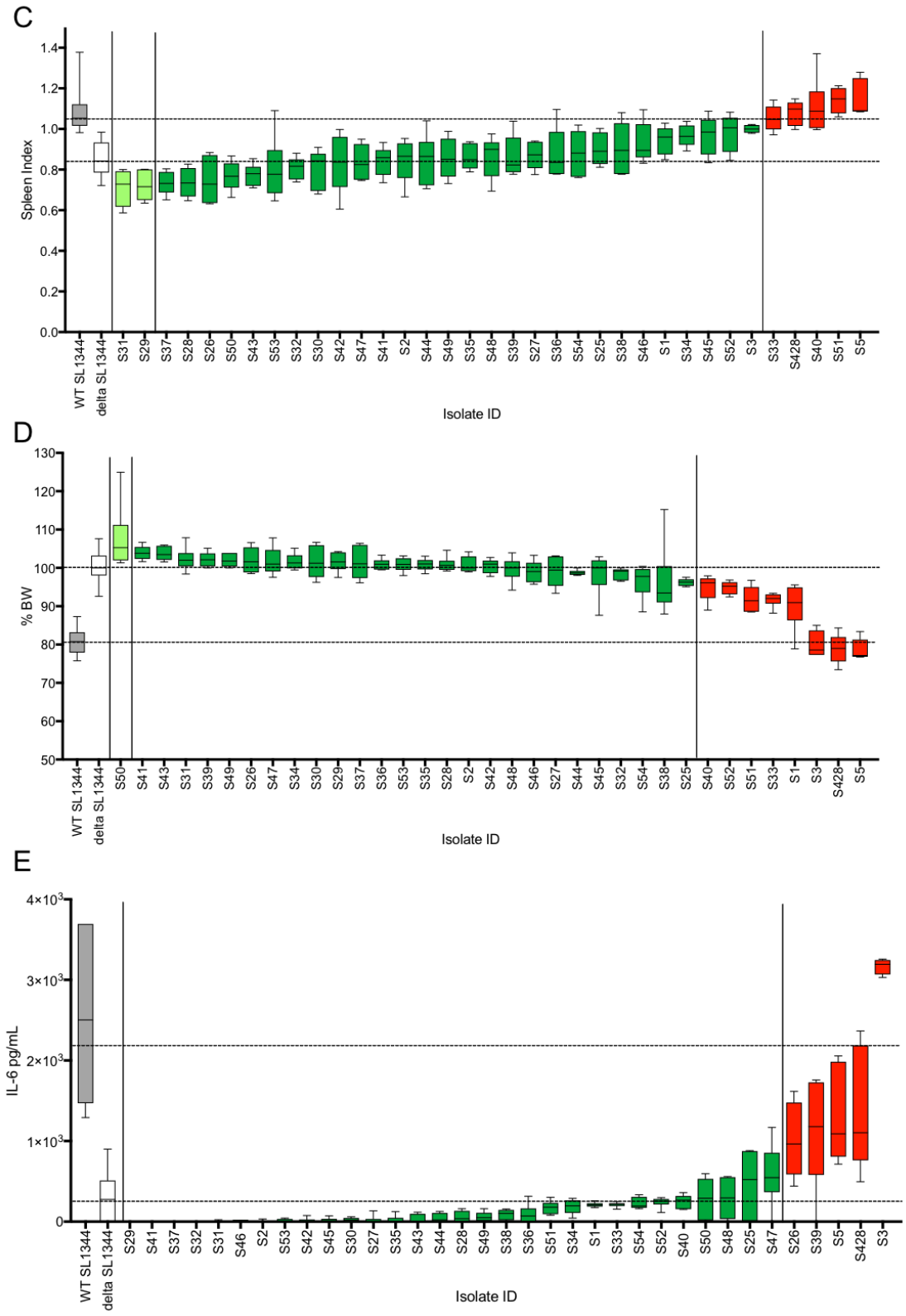

Fig. 3. (Continued) Virulence of 35 *Salmonella* isolates *in vivo*.

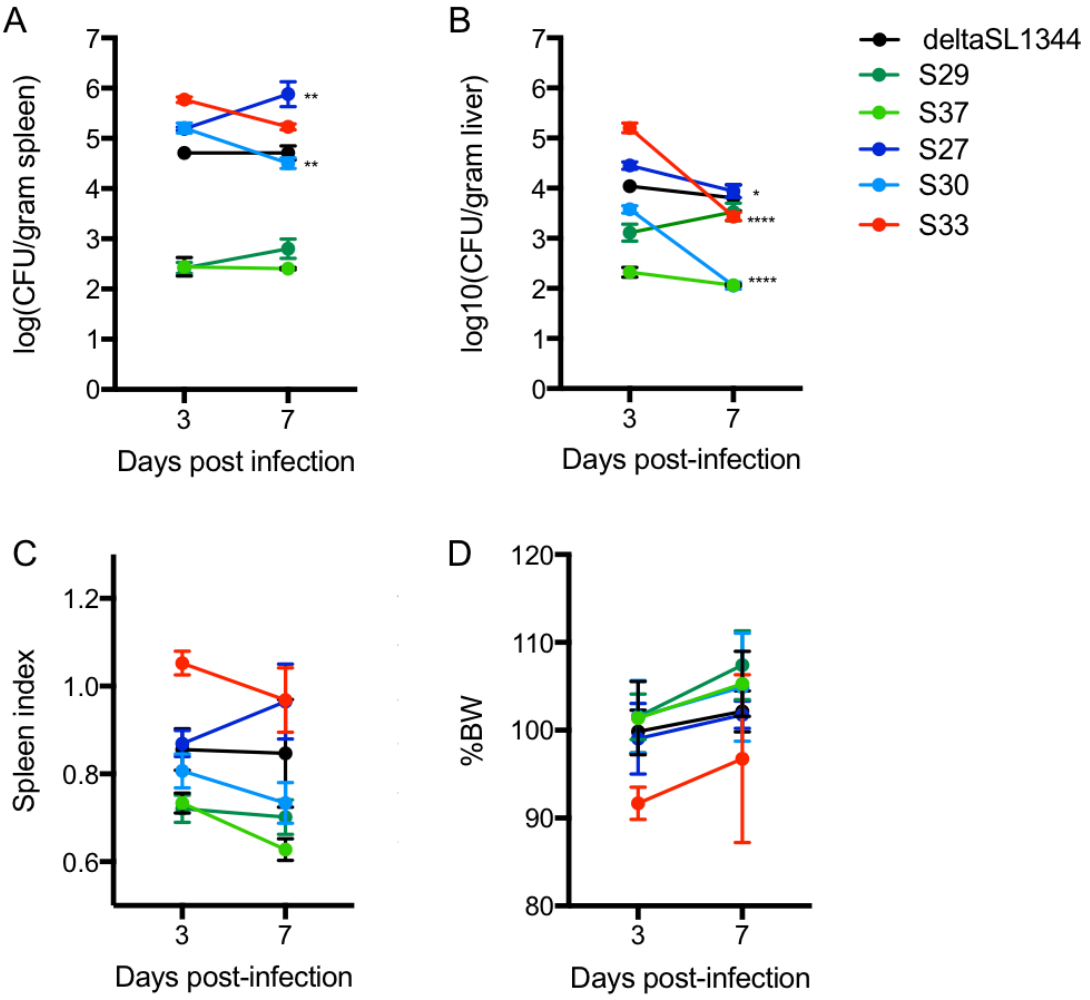

Fig. 4. Virulence at day 3 post-infection vs. day 3 post-infection.

A subset of *Salmonella* isolates was assessed at 3 days post-infection and 7 days post-infection. Isolates in green were low virulent 3 days post infection, isolates in blue were of intermediate virulence, and isolates in red were highly virulent. (A) Bacterial burden in the spleen and (B) liver, (C) spleen index, and (D) percent change in body weight were assessed at both time points. Statistics was done comparing isolate performance at day 3 vs. day 7 using a two-way ANOVA (*= $p \leq 0.05$, **= ≤ 0.005, ***= $p \leq 0.0005$, ****= $p < 0.0001$).

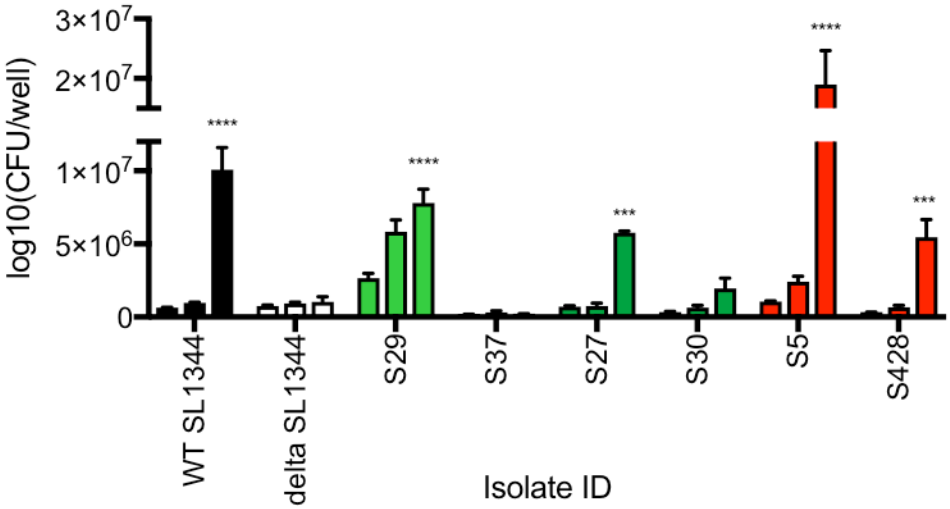

Fig. 5. *Salmonella* isolate virulence *in vitro* using RAW 264.7 macrophages

A subset of 6 *Salmonella* isolates tested *in vivo* were tested *in vitro* using RAW 264.7 macrophages. Intracellular bacterial burden was measure at times 0, 2 and 24 hours post-infection Statistical analysis was done comparing isolates at time 24 hours to low virulence delta SL1344 using a one-way ANOVA, Dunnett's multiple comparisons test (*= $p \leq 0.05$, **= ≤ 0.005, ***= $p \leq 0.0005$, ****= $p < 0.0001$). *In vivo*, S29 and S37 were low virulent (light green), S27 and S30 are of intermediate virulence (dark green), and S5 and S428 are highly virulent (red).

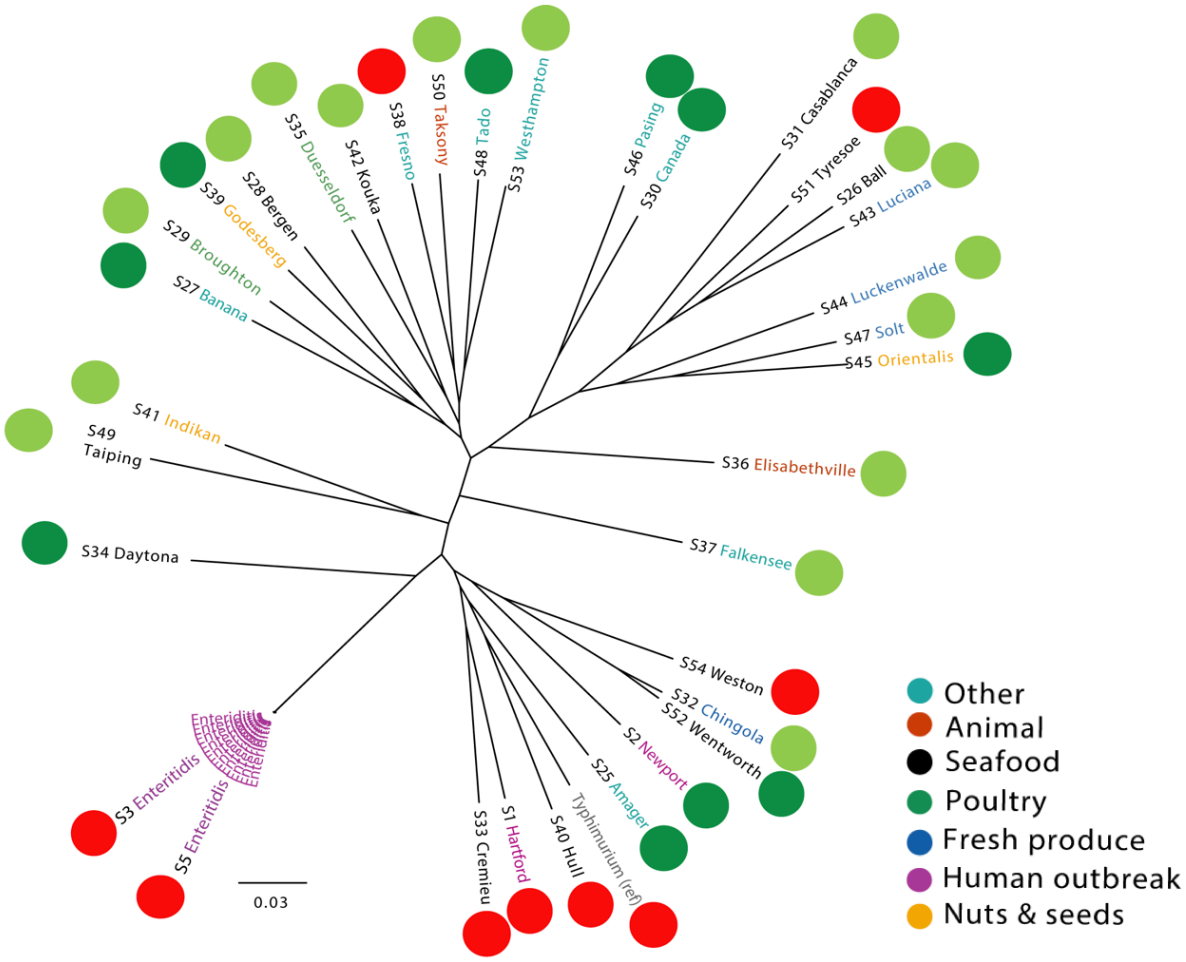

Fig. 6. Phylogenetic relationship between 35 *Salmonella* isolates tested

A phylogenetic tree based on isolate genomic sequence was constructed to visualize the genetic relationships between isolates. Isolates with red dots were highly virulent *in vivo*, those with dark green circles were of intermediate virulence, and those with light green dots low virulent. Serovar font colour indicates isolate source category.

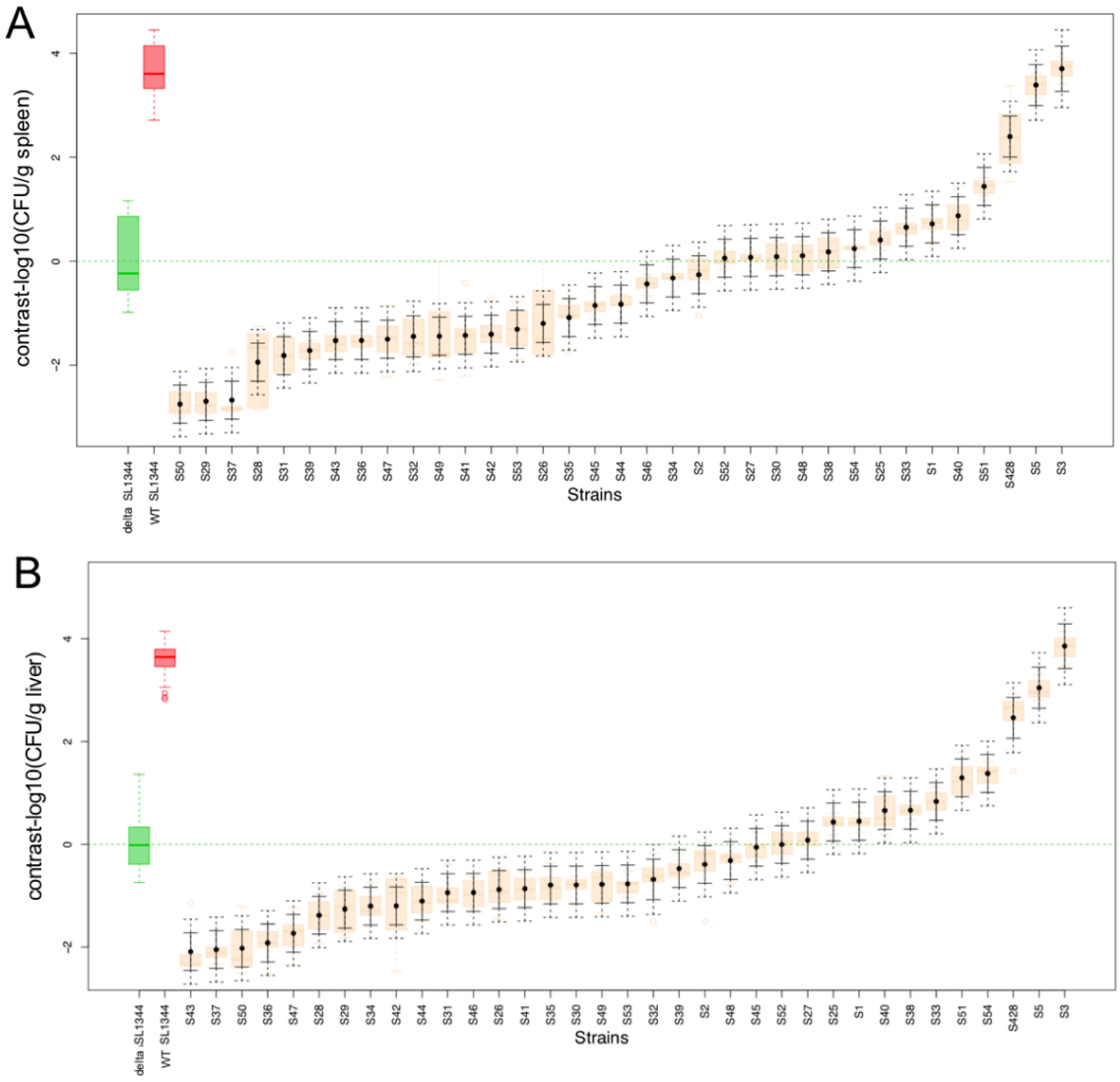

Fig. 7. Contrast in bacterial burden between isolates vs. delta SL1344 in mouse

Contrast between bacterial burden caused by *Salmonella* isolates vs. delta SL1344 (green) shown in the (A) spleen and (B) liver. Highly virulent SL1344 is shown in red. Isolate mean is shown by a black dot. Orange boxes denote the range of values. Solid black bars show the confidence intervals, and dotted black bars show the confidence intervals when adjusted for multiple testing. All isolates whose adjusted confidence intervals do not cross the green line (mean value of delta SL1344) are significant.

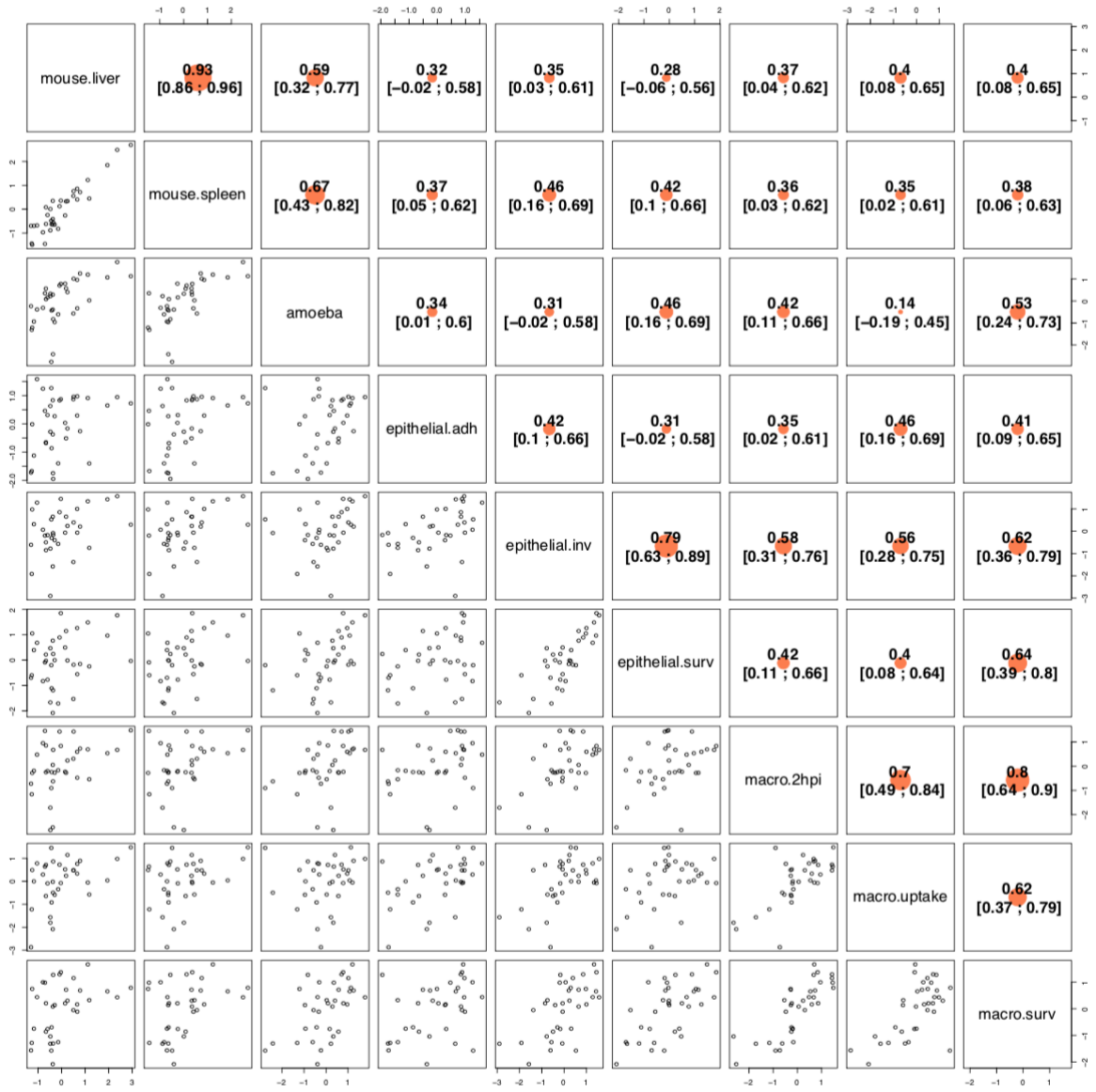

Fig. 8. Correlation of *Salmonella* isolates virulence across 4 models of infection

Concordance between each of the 4 models of infection was assessed by looking at correlation between scaled contrasts from each model. In the lower panel, each point represents a *Salmonella* isolate. Point positions are dictated by the level of contrast between the bacterial burden (CFUs) caused by an isolate and the bacterial burden caused by low virulent control delta SL1344. In the upper panel, the intraclass correlation coefficient is displayed with confidence interval (concordance measure). Correlations are significant for all correlation coefficients whose confidence intervals don't include zero. The size of the orange dot corresponds with the size of the correlation coefficient. adh: adhesion, inv: invasion, surv: survival, macro: macrophage.

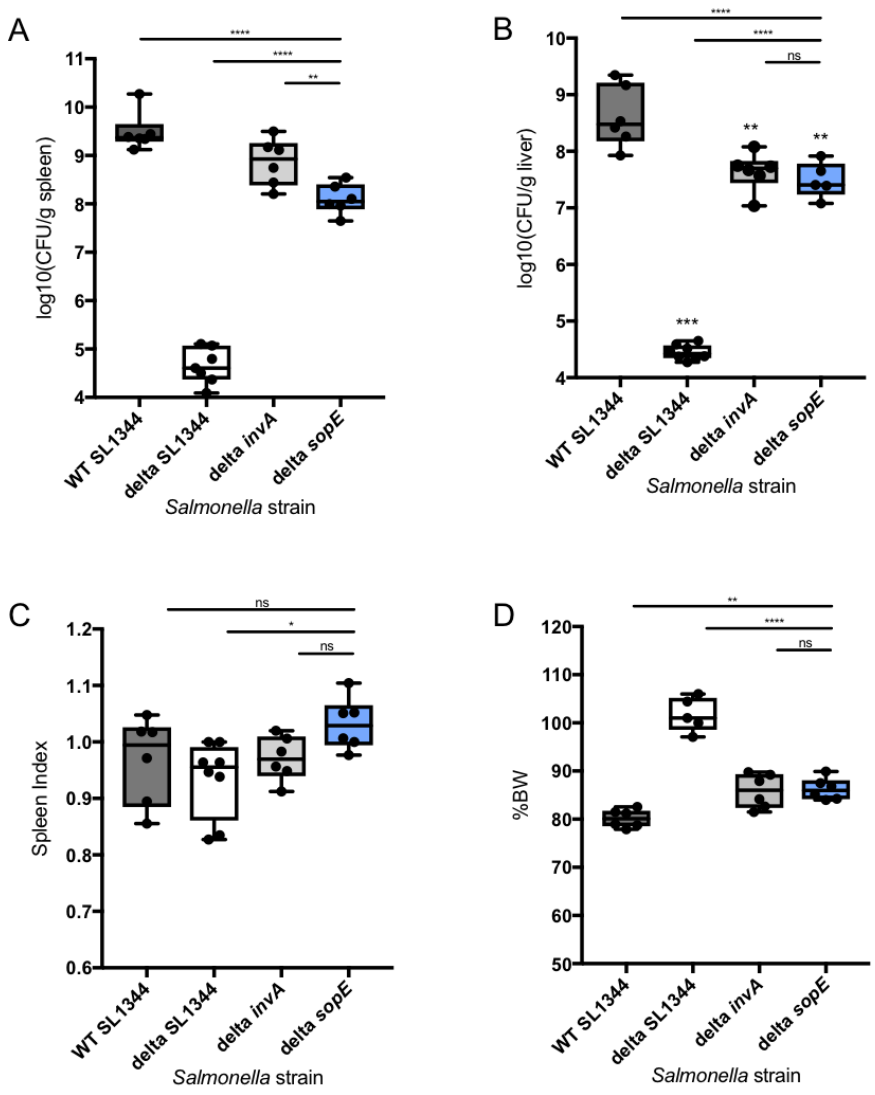

Fig. 9. Testing SopE as a virulence associated gene

To assess the role of SopE as a virulence associated gene the virulence of SL1344Δ*sopE* was compared to that of highly virulent control SL1344, low virulent control delta SL1344 and an addition strain of attenuated virulence SL1344Δ*invA*. Virulence was measured by bacterial burden in the (A) spleen and (B) liver, (C) spleen index, and (D) percent change in body weight. Statistics were done using a one-way ANOVA, Dunnett's multiple comparisons test (ns= $p > 0.05$, *= $p \leq 0.05$, **= ≤ 0.005, ***= $p \leq 0.0005$, ****= $p < 0.0001$).

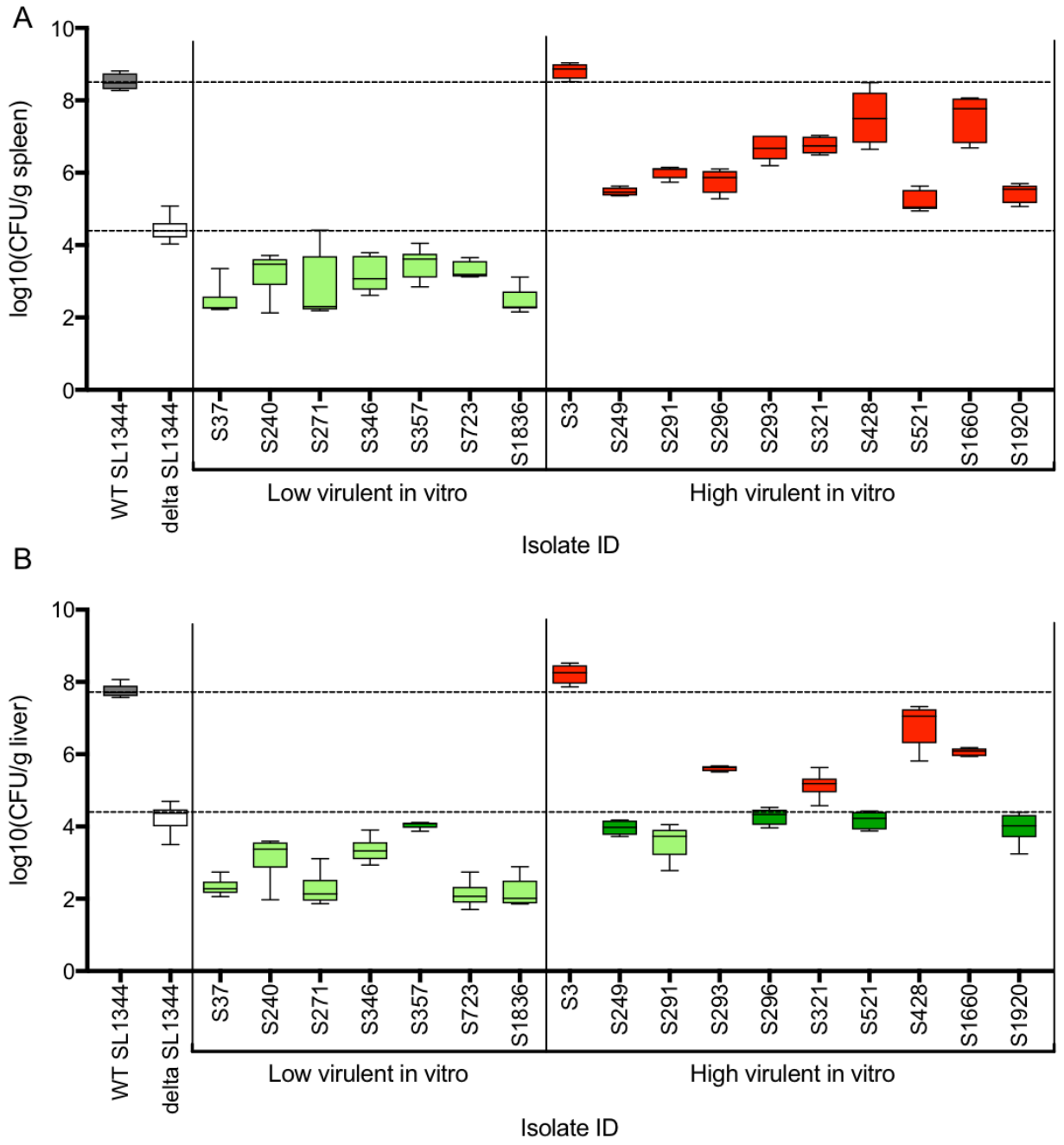

Fig. 10. Virulence of isolates used to create a method of virulence prediction.

In vivo virulence was assessed by (A) spleen and (B) liver, (C) spleen index, and (D) percent change in body weight. Isolates that were low virulent *in vitro* are grouped above "low virulent *in vitro*" and isolates that were highly virulent are grouped above "highly virulent *in vitro*". Isolates are compared to delta SL1344 using a one-way ANOVA, Dunnett's multiple comparisons test. Light green isolates were significantly lower that delta SL1344 ($p \leq 0.05$), Dark green isolates were not significantly different ($p > 0.05$), and red isolates were higher than delta SL1344 ($p \leq 0.05$). Figure is continued on next page.

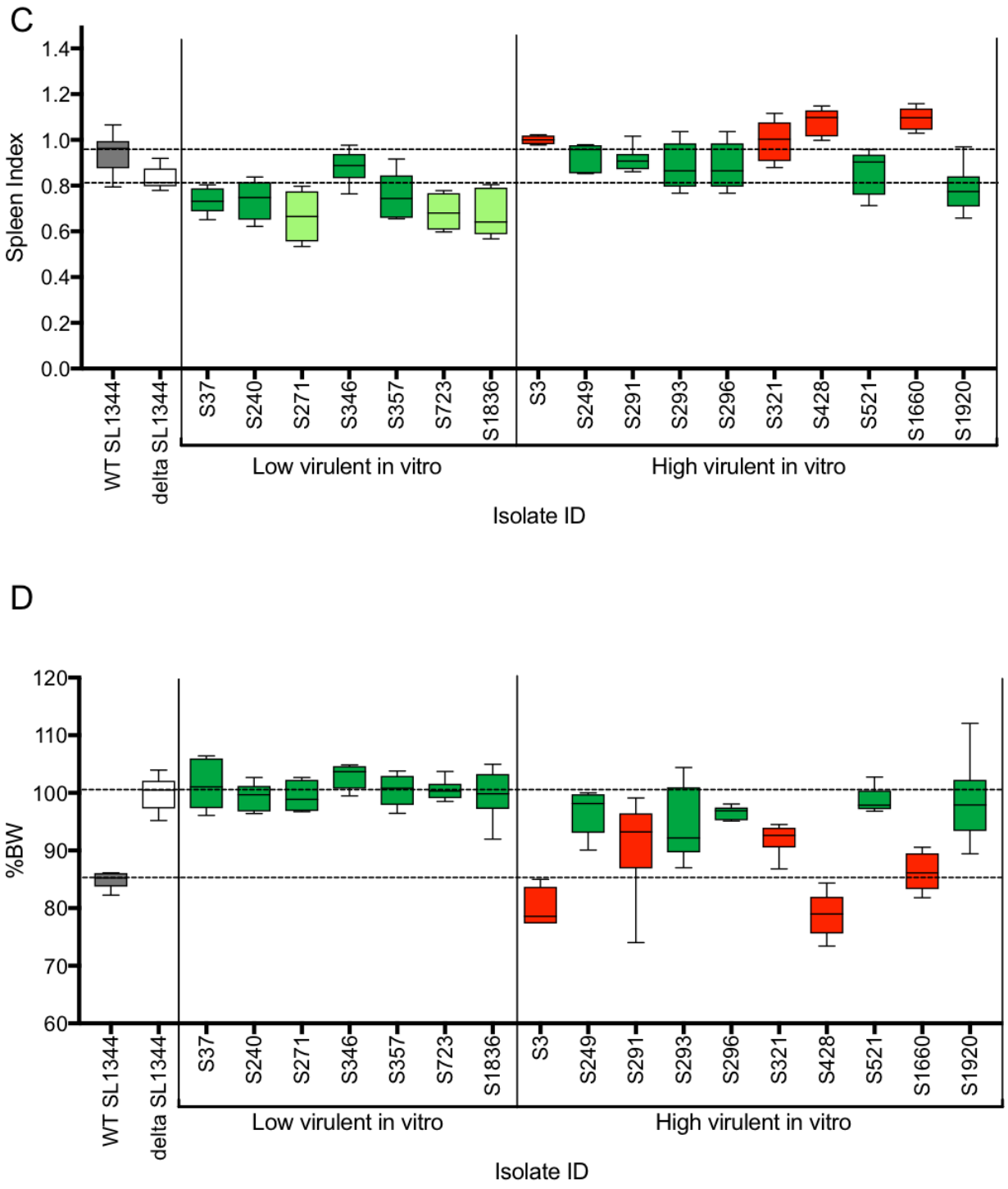

Fig. 10. (Continued) Virulence of isolates used to create a method of virulence prediction.

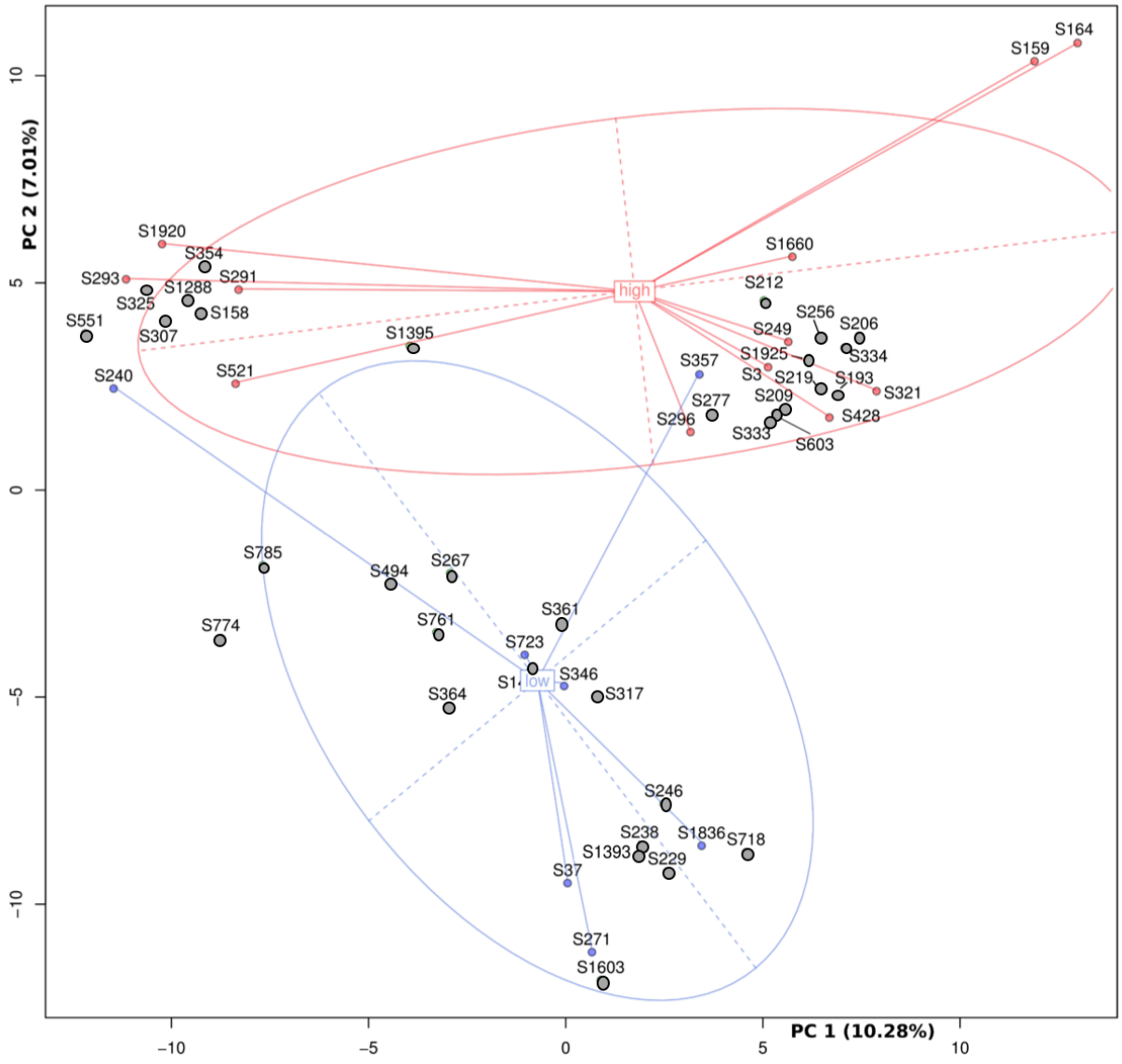

Fig. 11. Principal component analysis based on isolate sequence.

A principal component analysis was done based on the genome sequence of 17 *Salmonella* isolates whose virulence levels were the same both in epithelial cells and in mice. Isolates that were highly virulent are marked with a red dot. Isolates that were low virulent are marked with a blue dot. The red circle denotes a grouping of highly virulent isolates while the blue circle shows a grouping of low virulent isolates. Strains of unknown virulence (marked by grey dots) were added to the PCA.

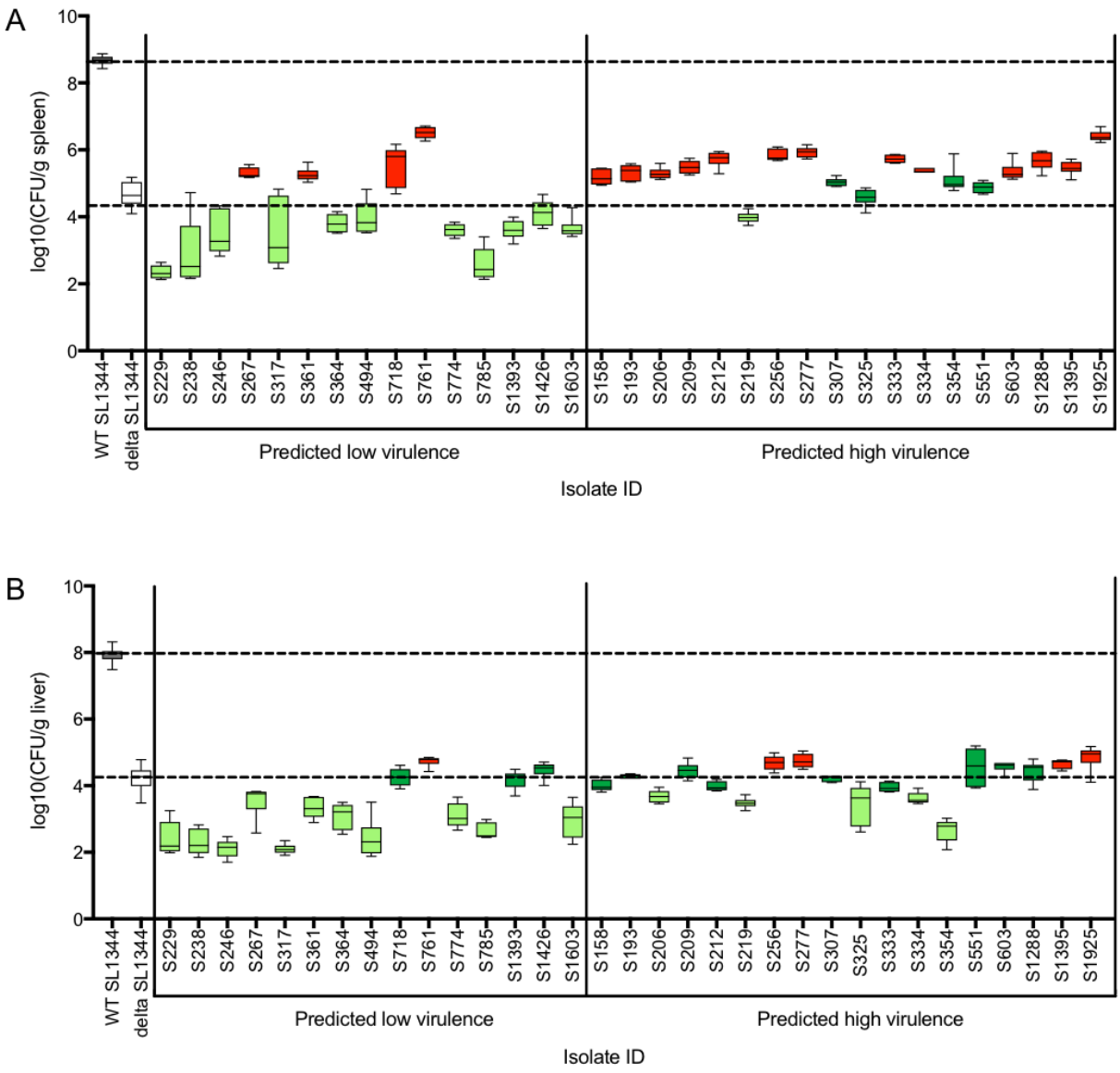

Fig. 12. Isolate virulence *in vivo* compared to sequence-based predictions.

33 *Salmonella* isolates had their virulence levels predicted and were tested *in vivo*. 18 were predicted to be highly virulent and 15 were predicted to be low virulent. These isolates were tested to assess concordance between *in vivo* phenotype and predicted phenotype. Bacterial burden in the (A) spleen and (B) liver, (C) spleen index, and (D) percent change in body weight were assessed. Isolates are compared to delta SL1344 using a one-way ANOVA, Holm-Sidak's multiple comparisons test. Light green isolates were significantly lower that delta SL1344 ($p \leq 0.05$), Dark green isolates were not significantly different ($p > 0.05$), and red isolates were higher than delta SL1344 ($p \leq 0.05$). Figure is continued on next page.

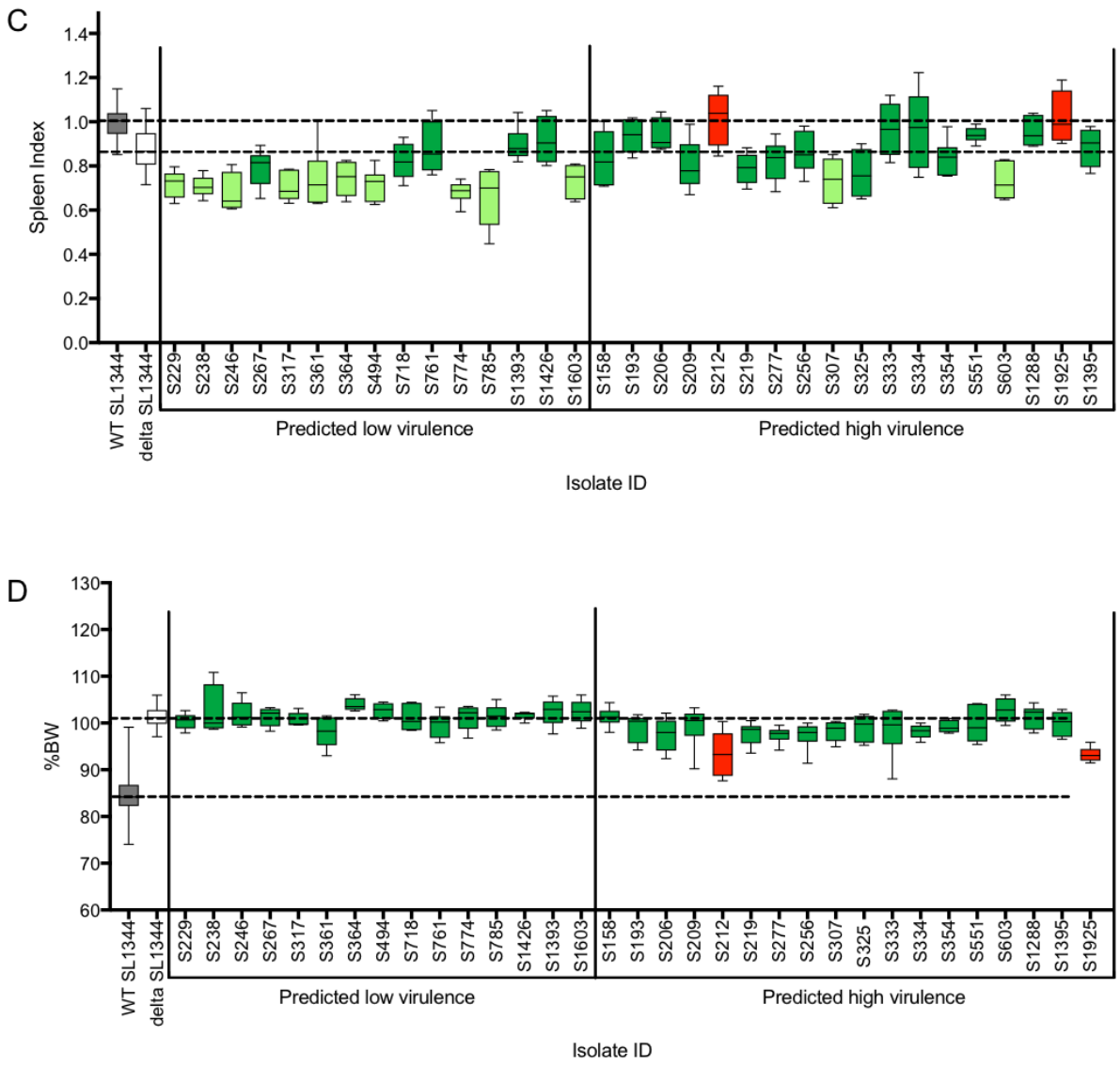

Fig. 12. (Continued) Isolate virulence *in vivo* compared to sequence-based predictions

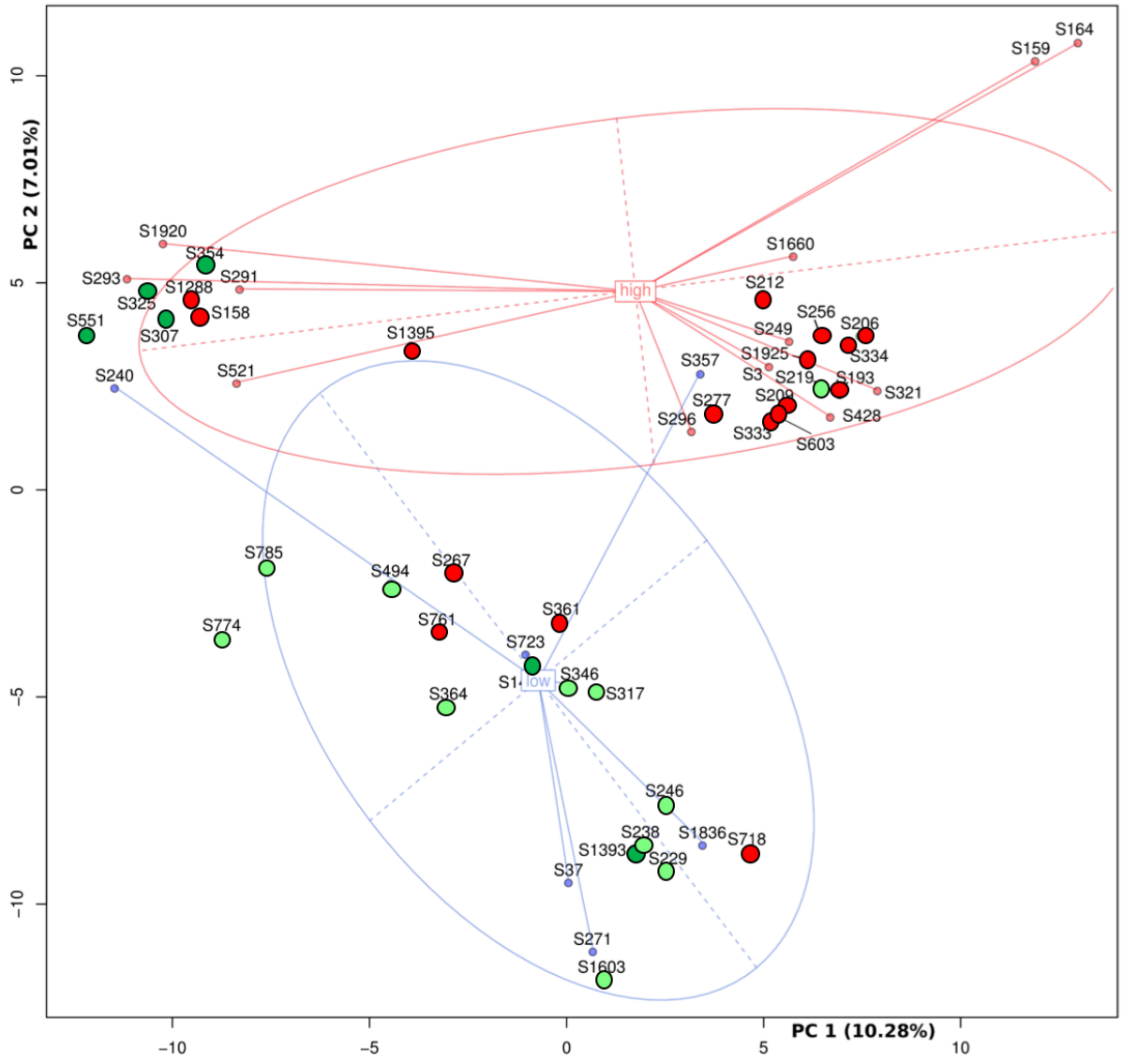

Fig. 13. Principal component analysis with virulence levels for *in vivo* testing.

30 isolates whose virulence level was predicted based on their position within the PCA were tested *in vivo*. Isolates that were found to be low virulent are marked by light green circles, those that were intermediately virulent are marked by dark green circles, and those that were highly virulent are marked by red circles. Strains yet to be tested are marked by grey circles.

Table 1. 35 *Salmonella* isolates of unknown virulence tested *in vivo*.

Isolate ID	Serovar	Source
S1	Hartford	Human
S2	Newport	Human
S3	Enteritidis	Human
S5	Enteritidis	Human
S25	Amager	Other
S26	Ball	Fish/shellfish
S27	Banana	Other
S28	Benger	Fish/shellfish
S29	Broughton	Poultry
S30	Canada	Other
S31	Casablanca	Fish/shellfish
S32	Chingola	Fresh produce
S33	Cremieu	Fish/shellfish
S34	Daytona	Fish/shellfish
S35	Duesseldorf	Poultry
S36	Elisabethville	Animal
S37	Falkensee	Other
S38	Fresno	Other
S39	Godesberg	Nuts/seeds
S40	Hull	Fish/shellfish
S41	Indikan	Nuts/seeds
S42	Kouka	Fish/shellfish
S43	Luciana	Fresh produce
S44	Luckenwalde	Fresh produce
S45	Orientalis	Nuts/seeds
S46	Pasing	Other
S47	Solt	Fresh produce
S48	Tado	Other
S49	Taiping	Fish/shellfish
S50	Taksony	Animal
S51	Tyresoe	Fish/shellfish
S52	Wentworth	Fish/shellfish
S53	Westhampton	Other
S54	Weston	Fish/shellfish
S428	Heidelberg	Human

Table 2. 17 *Salmonella* isolates used to create a predictive PCA.

	Isolate ID	Serovar	Source
Low virulent *in vitro* and *in vivo*	S37	Falkensee	Other
	S240	Montevideo	Nuts/seeds
	S271	Senftenberg	Nuts/seeds
	S346	Liverpool	Fresh produce
	S357	London	Fish/shellfish
	S723	Enderitidis	
	S1836	Kentucky	Nuts/seeds
Highly virulent *in vitro* and *in vivo*	S3	Enteritidis	Human
	S159	Typhimurium	-
	S249	Virchow	Fish/shellfish
	S291	Brendenburg	Fish/shellfish
	S293	Brandenburg	Dairy
	S296	Hvittingfoss	Fish/shellfish
	S321	Blockely	Poultry
	S428	Heidelberg	Human
	S521		Environmental
	S1660	Paratyphi B	Fish/shellfish
	S1920	Montevideo	Nuts/seeds

Table 3. 33 *Salmonella* isolates whose virulence was predicted.

Isolate ID	Serovar	Source
S158	-	-
S193	Thompson	Fresh produce
S206	Muenctten	Fresh produce
S209	Braenderup	Poultry
S212	Stanley	Other
S219	Hadav	Poultry
S229	Derby	Animal
S238	Mbandaka	Other
S246	Kentucky	Other
S256	Bovismobificans	Animal
S267	Kiambu	Poultry
S277	Uganda	Animal
S307	Poona	Fresh produce
S317	Ohio	Fresh produce
S325	Bredeney	Dairy
S333	Berta	Other
S334	Kottbus	Poultry
S354	Gaminara	Other
S361	Adelaide	Animal
S364	Cerro	Poultry
S494	-	Environmental
S551	-	Environmental
S1288	Javiana	Fresh produce
S1426	Berta	Environmental
S1603	Havana	Animal
S1925	MuencHen	Nuts/seeds

DISCUSSION

PHENOTYPIC DIVERSITY WITHIN *SALMONELLA*

The rational for this project was founded on two well established aspects of *Salmonella.* First, that there is a wide range of phenotypic diversity within the *Salmonella* genus, and second, that phenotypic diversity is the result of genetic diversity. The phenotypic variability within *Salmonella* was illustrated in our work by varying levels of virulence observed *in vivo*. The bacterial burdens caused by the 35 initially screened isolates spanned over 6 log CFUs in both the spleen and liver. Furthermore, while some isolates caused weight loss and/or spleen enlargement, others caused no observable signs of disease, even when infected for upwards of a month, embodying the fact that disease manifestation is isolate dependent.

Phenotypic diversity varied similarly across the other models of infection as demonstrated by the positive correlations observed between nearly all measures of virulence, with the amoeba model showing the highest degree of concordance with the *in vivo* data. Given the ability of the amoeba model to replicate isolate virulence seen at the systemic level, this model may be useful in the future for isolate screening that is quick and high-throughput.

In contrast, the human epithelial model was least similar to the mouse model. Multiple isolates were able to adhere to, invade, and/or survive within epithelial cells *in vitro* yet were unable to produce a systemic disease *in vivo*. This makes sense since the response to infection in an entire system is more complex than in one element of that system. Although adhesion, invasion and survival within epithelial cells are critical elements of *Salmonella* infection, having any one or multiple of these capacities is not necessarily sufficient to produce disease within a host.

Based on our findings, studying *Salmonella* infection in epithelial cells alone does not effectively represent the overall pathogenic potential of *Salmonella* isolates *in vivo*. This prevents the findings from this model from being generalizable to risk of disease in humans. Nevertheless, studying isolated components of infection, as is the case for *in vitro* studies, can be invaluable to

understanding specific aspects of *Salmonella* pathogenesis. Studying *Salmonella* infection *in vitro* using epithelial cells has furthered our understanding of *Salmonella* invasion, SCV formation, and modulation of the intestinal environment, among other characteristics of infection (156, 161, 162). By sequencing and testing a collection of isolates in 4 models that assess different aspects of infection, a repertoire of isolate phenotypes and corresponding genotypes has been generated. This repertoire can be used to pinpoint particular elements of pathogenesis that are defective in isolates that are "low virulent" in vivo. Moreover, consulting the genomes of isolates that are impaired during epithelial cells infection can further contribute to our understanding of the mechanisms involved in bacterial invasion, a process which on its own is highly complex.

GENETIC DIVERSITY WITHIN *SALMONELLA*

We sought to gain a better understanding of the genetic diversity underlying *Salmonella* virulence by identifying genes common to highly virulent isolates. WGS and comparative analyses identified 4 known virulence associated genes and 1 putative virulence associated gene, *sopE* which we confirmed following *in vivo* testing. Deletion of *sopE* lead to an attenuation in virulence but not a full abrogation of systemic dissemination. It has been well established that following ingestion, *Salmonella* passes the intestinal barrier primarily by invading non-phagocytic epithelial cells (105). SopE helps facilitate bacterial entry by altering actin filament rearrangement in epithelial cells (160, 163). As such, we would expect that deletion of *sopE* would result in a marked reduction in bacterial burden, particularly following oral infection. The reduced virulence observed using an IV model, which by-passes the process of epithelium invasion, indicates that SopE plays a role beyond known role at the epithelium. Further study is needed to understood the intricate roll of SopE during infection.

One of the purposes of identifying virulence associated genes was to add to the collection of gene targets used to in current molecular detection assays. Other genes used include *invA*, *iagA, spaQ, sipB, sipC,* and *sifA* (164-167). Targeting more genes help validate positive test results by decreasing false-positives caused by cross-reactivity with non-pathogenic microbes, such as *Proteus* spp. and *Citrobacter* spp. which are also present on food products. False negatives can also be decreased by targeting multiple genes as the chances of amplify a portion of the isolate's genome are increased. Having the sequences for 5 additional virulence associated genes may prove useful during the process of designing and testing primer combinations, promoting optimal sensitivity and specificity.

Aside from identifying virulence associated genes, WGS performed on the initial 35 isolates tested was used to construct a phylogenetic tree displaying genetic relatedness. *In vivo* phenotyping of these isolates enabled associations between evolutionary relationship and virulence to be assessed. Determining the degree of genetic similarity and therefore the shared evolutionary history between isolates can be crucial during outbreak investigations. The benefit of WGS during outbreak tracking has already been proven multiple times over the past decade. During a 2010 cholera outbreak in Haiti when PFGE was unable to discriminate between various isolates, sequencing was used to track the outbreak source to a group of UN peacekeepers who had recently been in Nepal during a cholera outbreak (168). During an *E. coli* outbreak in 2011 in Germany, real-time sequencing was used to generate a draft genome sequence in under a day and determine the pathogenic potential of the isolate based on its evolutionary history (169, 170). In the same way WGS was used during these outbreaks, phylogenetic analysis may provide the quick and discriminatory information needed to distinguish and track sources during *Salmonella* outbreaks. By attaining virulence information from isolate sequence in addition to evolutionary history the scope and ramifications of a particular outbreak can be better predicted.

Phylogenetic analyses of the 35 initially tested isolates enabled us to look at the relationship between evolutionary history and virulence. Given that evolutionary history is determined by genetic relatedness, and the fact that sequence underpins phenotype, it was unsurprising, yet reassuring, to see that isolates of similar phenotypes cluster together during phylogenetic analyses. This was particularly true for highly virulent isolates, the majority of which were all located next to one another, denoting a high degree of genetic similarity. What was surprising was the observation that two highly virulent isolates, S38 and S51, were located separately from the other highly virulent isolates and from each other. Furthermore, the isolates most closely related to S38 and S51 were not intermediate isolates but were actually low virulent isolates. Although we did not investigate this finding further here, it would be of interest to study the genetic differences between S38 and S51 and the closely related low virulent isolates surrounding them in the phylogenetic tree.

PREDICTING ISOLATE VIRULENCE

One of the primary objectives of this project was to develop a method for predicting isolate virulence based on genome sequence. The method developed here used the distribution of isolates of known virulence within a PCA plot and the degree of sequence similarity of isolates of unknown virulence to predict phenotype of the latter group. 67% (22/23) of isolates yielded phenotypes *in vivo* that matched their virulence predictions. Nine out of ten low virulent isolates were correctly predicted, and thirteen out of seventeen highly virulent isolates were correctly predicted. The observed 90% specificity indicates a low propensity for false positives. This is advantageous in terms of the economic burden of salmonellosis as this translates to less food wastage.

The sensitivity of this prediction method, on the other hand, poses a major drawback for the practical application of this system. Only 77% of isolates that were highly virulent *in vivo* were predicted to behave as such. An appreciable proportion of false negatives were observed. Although

the sensitivity and specificity were calculated excluding intermediate isolates and therefore, are not fully representative of the entire data set, this does not negate the problem of false negatives. We observed 4 instances where isolates were predicted to be low virulent but were actually highly virulent *in vivo.* False negatives are a major concern for public health, as undetected contamination of food with highly virulent *Salmonella* leads to illness. We would expect to have as few false negatives as possible, ideally having none. Therefore, improvements to the current system need to be made.

CONSIDERATIONS FOR APPLYING A VIRULENCE PREDICTION METHOD

Considerations on the side of the pathogen

Given the false-negative rate of the prediction method, it is not efficient or safe enough for use. If refinement and alterations to the prediction method allowed for consistent and accurate virulence predictions, practical application for pathogen detection and virulence assessment would require continual pathogen isolation, sequencing, and phenotyping. The *Salmonella* genome is not static but rather is are constantly evolving. Acquisition of genetic elements through horizontal transfer and genomic loss through gene degradation contribute largely to serovar specific sequences and phenotypic diversity (34). Sequence delineation over time has resulted in different serovars harbouring distinct collections and genome-wide distribution of SPIs, which helps account for serovar dependant phenotypes (34). Along with broad architectural changes to the genome, more precise changes including deletions, insertions and single point mutations can change an isolate's pathogenic potential, as was demonstrated here by deleting *sopE,* and in other cases of experimentally derived mutations (171). For this reason, continuous isolate phenotyping will be prudent to accurately predict the impact of new mutations, particularly to distinguish between mutations that decrease virulence, confer virulence, and those having neutral effects. As previously

mentioned, of the high-thoughput models used in the phenotyping pipeline the amoeba model yielded results most closely mirroring those found *in vivo*. This may, therefore, be an advantageous model for continued phenotypic surveillance.

Given the impact single nucleotide variants can have on phenotype, the resolution at which isolate genomes are compared must be reconsidered. The gene-based virulence prediction system proposed here is an extension of gene-by-gene subtyping approaches, which themselves are a WGS extension of MLST. Gene-by-gene subtyping involves WGS followed by sequence comparison using the gene as the unit of comparison between isolates (153, 172). This method of subtyping has already been useful in the subtyping other food borne pathogens such as *Campylobacter* (173). While gene-by-gene subtyping approaches compare allelic variations between isolates, as of now, the virulence prediction method developed in this project focuses strictly on gene presence or absence. Since isolate virulence is contingent on allelic variation as well as gene presence, virulence predictions in practice will need to be based off more detailed assessments. Virulence associated sequences rather than the presence of virulence associated genes will need to be examined.

Besides gene-by-gene approaches, sequence-based subtyping methods have also been developed which assess subtle genome-wide variants such as single nucleotide variants (SNVs), and small insertions and deletions (indels), going beyond the unit of the gene to the resolution of a single nucleotide (174). This method requires a high quality reference genome which at times can be challenging however, it has been proven successful at subtyping during outbreak investigations, including outbreaks caused by *Salmonella* (175). Analysis of single nucleotide variants throughout the genome has the added benefit of identifying variants laying outside of genes that may impact gene expression. This method, or a combination of both gene-by-gene and single variant based methods will be needed to more accurately predict isolate virulence.

With all aspects of sequence variation considered, this still does not account for the full complexity of pathogen phenotypes. Epigenetics has been observed in bacteria primarily in the form of DNA methylation. DNA methylase has been associated with DNA replication, DNA repair, and global gene regulation in *E. coli* (176). The study of epigenetics in bacteria appears to be in its infancy with the full impact yet to be determined. Other mechanisms governing bacterial phenotype include small RNAs (177) and interactions with other microbes (178). Therefore, although we've made strides to understand the relationship between sequence and phenotype, the ability to confidently predict *Salmonella* virulence will require further investigation.

Considerations on the side of the host

One consideration with respect to mouse studies is the route of pathogen administration. Here, a systemic model was chosen over an enterocolitis model since this model traditionally requires that mice be pre-treated with streptomycin followed by infection with an isolate harbouring streptomycin resistance. Although pre-treatment with antibiotics and oral inoculation better replicates the natural route of infection and pathology associated with salmonellosis than I.V. injection, the use of antibiotics for our purposes was complicated by varying degrees of antibiotic resistance observed across isolates. Out of 3,377 sequences, 1,003 unique resistomes were identified (179), creating a major obstacle when considering the unique design that would be needed to screen each unique isolate.

More recently, SIGIRR deficient mice have shown exaggerated antimicrobial functions within the gut following infection with enteric bacteria (180). Rather than depleting the infecting pathogen, this activity significantly reduces commensal microbes, abolishing any colonization resistance. SIGIRR mice are highly susceptible to *Salmonella* Typhimurium and develop

enterocolitis upon infection. Compared to wild-type mice, *sigirr* -/- mice show increased colonization of the intestines, severe inflammation and increased mucosal damage (181). Since this phenotype is observed without the use of antibiotics, this may be an advantageous model with which to validate virulence levels found using our systemic model.

As is exemplified by SIGIRR mice, host genetics play a major role in determining the manifestation of infectious disease. We used susceptible C57BL/6J mice with defective Nramp1 because, like mice, humans exhibit a range of resistance phenotypes (182). Isolates labelled as low virulent *in vivo* need to pose no risk to resistant as well as susceptible humans. Further study into the impact of low virulent isolates in susceptible humans should be done using mice that model causes of susceptibility in humans. Immunodeficiencies resulting from defects in the IL-12/STAT4/ IFNγ axis (183), NADPH oxidase activity, CD40LG/CD40 signaling, or the BTK gene have each been associated with *Salmonella* susceptibility in both humans in mice (113, 171), making any of these lines a useful model to examine isolate behaviour in a immunocompromised hosts.

Using C57BL/6J mice we found that the majority of isolates from the "human clinical outbreak" source were also highly virulent in mice, showing the generalizability of isolate behaviour in a systemic mouse model to salmonellosis in humans. However, it remains that various *Salmonella* isolates can produce vastly different phenotypes in mice and humans. *Salmonella* serovars causing Typhoid in humans fail to produce systemic disease in mic while isolates that cause gastroenteritis in humans instead cause systemic disease in mice. Phenotypic concordance between humans and mice can only be determined retroactively. After an outbreak has occurred and the phenotype in humans is known, the isolate responsible can then be tested in mice. Determining the phenotype in humans following testing in mice however is not possible.

Humanized systems may help predict how an isolate will behave in a human host. Nonobese diabetic (NOD)-scid IL2rγnull mice and alymphoid RAG-2−/−/γc−/− mice reconstituted with human leukocytes, known as humanized mice, show increased susceptibility to infection with *S.* Typhi. Hallmarks of systemic Typhoidal-disease including increased bacterial burden in the spleen, liver and bone marrow, as well as significant body weight loss have been observed following infection (184, 185). Given the success with Typhoidal *Salmonella*, humanized mice may also be useful in modeling NTS infection.

Beyond mice, synthetic human models are also possible. As part of the *Salmonella* Syst-OMICs Consortium, a synthetic human gastrointestinal model is being used to assess the activity of a subset of the isolates tested here (Gisèle Lapointe, University of Guelph). Various connecting chambers replicate the physiological conditions of the gut, enabling the estimation of isolates survival and growth following human infection. It will be of interest to compare the phenotypes of isolates in this model with our findings in mice to help build the phenotypic correlation between mice and humans.

CONCLUSIONS & FUTURE DIRECTIONS

In recent years, WGS has proven itself to be an invaluable tool for food regulation and safety in terms of improving *Salmonella* detection, serotyping, and outbreak tracking. Here, we investigated new ways to integrate WGS into food safety practices that relate to salmonellosis. WGS and *Salmonella* phenotyping were used together to further elucidate the role genetics play in manifesting isolate virulence. This process led to the identification of 5 virulence associated genes which genes can be added to the collection of genes currently used in molecular pathogen detection assays. The comparative genomic analyses and PCA virulence assessment method performed can be used further to estimate the scope and ramifications of future outbreaks. The goal of producing a system that can determine isolate virulence based on sequence remains a work in progress, with false negatives being a major concern. Further study into the complex relationship between bacterial virulence and host defenses is needed to develop the current system into one that can be used to accurately predict phenotype during the process of pathogen detection.

To develop a "gene-by-gene" based method for predicting virulence, the repertoire of known virulence associated factors needs to be expanded, therefore more comparative analyses need to be performed. The 5 genes identified in this project were found by performing 2 analyses, each comparing 5 isolates. Future analyses should be performed with a greater number of isolates to identify virulence associated genes that are common to many highly virulent isolates and concordantly, highly relevant to disease. Analyses comparing phylogenetic outliers S38 and S51 should also be done.

Given the insufficient sensitivity of the PCA prediction method, refinements need to be considered. Possible adjustments include using more isolates to form the PCA. Adding more isolates would provide more genetic diversity on which to compare and segregate isolates. In addition, isolates with concordant phenotypes between mice and amoeba should also be used. The current PCA was built using isolates with concordant phenotypes in epithelial cells and in mice. However,

there was relatively low correlation between these two models. The high fidelity between the amoeba model and the mouse model, may provide more accurate predictions for in vivo testing. Of note, S357, which is currently in the PCA, should be removed as this isolate was used as a low virulent isolate but was actually intermediately virulent *in vivo*. Furthermore, more isolates from the "fresh produce" category should be tested.

Finally, in terms of statistics, the correlation analysis must be revisited to correct for multiple testing. Contrast analyses will also be performed comparing the bacterial burden of the 33 predicted isolates to that of low virulent control delta SL1344. This may help may assign current intermediate virulence isolates to either low or highly virulent categories.

REFERENCES

1. Kauffmann F. The Bacteriology of Enterobacteriaceae. Collected Studies of the Author and his Co-Workers. The Bacteriology of Enterobacteriaceae Collected Studies of the Author and his Co-Workers. 1966.
2. Popoff M, Le Minor L. Antigenic formulas of the Salmonella serovars 7th revision: World Health Organization Collaborating Centre for Reference and Research on Salmonella. Pasteur Institute, Paris, France. 1997.
3. Brenner F, Villar R, Angulo F, Tauxe R, Swaminathan B. Salmonella nomenclature. Journal of clinical microbiology. 2000;38(7):2465-7.
4. Achtman M, Wain J, Weill F-X, Nair S, Zhou Z, Sangal V, et al. Multilocus sequence typing as a replacement for serotyping in Salmonella enterica. PLoS pathogens. 2012;8(6):e1002776.
5. Crosa J, Brenner D, Ewing W, Falkow S. Molecular relationships among the Salmonelleae. Journal of bacteriology. 1973;115(1):307-15.
6. Buckle GC, Walker CLF, Black RE. Typhoid fever and paratyphoid fever: Systematic review to estimate global morbidity and mortality for 2010. Journal of global health. 2012;2(1).
7. Crump JA, Luby SP, Mintz ED. The global burden of typhoid fever. Bulletin of the World Health Organization. 2004;82(5):346-53.
8. Jong EC. Enteric Fever: Typhoid and Paratyphoid Fever. Netter's Infectious Diseases: Elsevier; 2012. p. 394-9.
9. Maskell D. Salmonella infections: clinical, immunological and molecular aspects: Cambridge University Press; 2006.
10. Levine MM, Black RE, Lanata C. Precise estimation of the numbers of chronic carriers of Salmonella typhi in Santiago, Chile, an endemic area. The Journal of infectious diseases. 1982;146(6):724-6.
11. Callejón RM, Rodríguez-Naranjo MI, Ubeda C, Hornedo-Ortega R, Garcia-Parrilla MC, Troncoso AM. Reported foodborne outbreaks due to fresh produce in the United States and European Union: trends and causes. Foodborne pathogens and disease. 2015;12(1):32-8.
12. Majowicz SE, Musto J, Scallan E, Angulo FJ, Kirk M, O'Brien SJ, et al. The global burden of nontyphoidal Salmonella gastroenteritis. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America. 2010;50(6):882-9.
13. Keestra-Gounder AM, Tsolis RM, Bäumler AJ. Now you see me, now you don't: the interaction of Salmonella with innate immune receptors. Nature Reviews Microbiology. 2015;13(4):206.
14. Bäumler AJ, Tsolis RM, Heffron F. Contribution of fimbrial operons to attachment to and invasion of epithelial cell lines by Salmonella typhimurium. Infection and immunity. 1996;64(5):1862-5.
15. Clark M, Jepson M, Simmons N, Hirst B. Preferential interaction of Salmonella typhimurium with mouse Peyer's patch M cells. Research in microbiology. 1994;145(7):543-52.
16. Jones BD, Ghori N, Falkow S. Salmonella typhimurium initiates murine infection by penetrating and destroying the specialized epithelial M cells of the Peyer's patches. Journal of Experimental Medicine. 1994;180(1):15-23.
17. Vazquez-Torres A, Jones-Carson J, Bäumler AJ, Falkow S, Valdivia R, Brown W, et al. Extraintestinal dissemination of Salmonella by CD18-expressing phagocytes. Nature. 1999;401(6755):804-8.
18. Rescigno M, Rotta G, Valzasina B, Ricciardi-Castagnoli P. Dendritic cells shuttle microbes across gut epithelial monolayers. Immunobiology. 2001;204(5):572-81.
19. Rescigno M, Urbano M, Valzasina B, Francolini M, Rotta G, Bonasio R, et al. Dendritic cells express tight junction proteins and penetrate gut epithelial monolayers to sample bacteria. Nature immunology. 2001;2(4):361.
20. Vazquez-Torres A, Jones-Carson J, Bäumler AJ, Falkow S, Valdivia R, Brown W, et al. Extraintestinal dissemination of Salmonella by CD18-expressing phagocytes. Nature. 1999;401(6755):804.

21. Mastroeni P, Villarreal-Ramos B, Hormaeche CE. Role of T cells, TNFα and IFNγ in recall of immunity to oral challenge with virulent salmonellae in mice vaccinated with live attenuated aro− salmonella vaccines. Microbial pathogenesis. 1992;13(6):477-91.
22. Sheikh A, Khanam F, Sayeed MA, Rahman T, Pacek M, Hu Y, et al. Interferon-γ and proliferation responses to Salmonella enterica serotype Typhi proteins in patients with S. Typhi bacteremia in Dhaka, Bangladesh. PLoS neglected tropical diseases. 2011;5(6):e1193.
23. Wahid R, Salerno-Gonçalves R, Tacket CO, Levine MM, Sztein MB. Cell-mediated immune responses in humans after immunization with one or two doses of oral live attenuated typhoid vaccine CVD 909. Vaccine. 2007;25(8):1416-25.
24. Mizuno Y, Takada H, Nomura A, JIN CH, Hattori H, Ihara K, et al. Th1 and Th1-inducing cytokines in Salmonella infection. Clinical & Experimental Immunology. 2003;131(1):111-7.
25. Stoycheva M, Murdjeva M. Serum levels of interferon-γ, interleukin-12, tumour necrosis factor-α, and interleukin-10, and bacterial clearance in patients with gastroenteric Salmonella infection. Scandinavian journal of infectious diseases. 2005;37(1):11-4.
26. Dhanoa A, Fatt QK. Non-typhoidal Salmonella bacteraemia: epidemiology, clinical characteristics and its' association with severe immunosuppression. Annals of clinical microbiology and antimicrobials. 2009;8(1):15.
27. Gordon MA, Banda HT, Gondwe M, Gordon SB, Boeree MJ, Walsh AL, et al. Non-typhoidal salmonella bacteraemia among HIV-infected Malawian adults: high mortality and frequent recrudescence. Aids. 2002;16(12):1633-41.
28. Jones TF, Ingram LA, Cieslak PR, Vugia DJ, Tobin-D'Angelo M, Hurd S, et al. Salmonellosis outcomes differ substantially by serotype. The Journal of infectious diseases. 2008;198(1):109-14.
29. Feasey NA, Dougan G, Kingsley RA, Heyderman RS, Gordon MA. Invasive non-typhoidal salmonella disease: an emerging and neglected tropical disease in Africa. The Lancet. 2012;379(9835):2489-99.
30. Brent AJ, Oundo JO, Mwangi I, Ochola L, Lowe B, Berkley JA. Salmonella bacteremia in Kenyan children. The Pediatric infectious disease journal. 2006;25(3):230-6.
31. Gordon MA. Salmonella infections in immunocompromised adults. Journal of Infection. 2008;56(6):413-22.
32. de Beaucoudrey L, Samarina A, Bustamante J, Cobat A, Boisson-Dupuis S, Feinberg J, et al. Revisiting human IL-12Rβ1 deficiency: a survey of 141 patients from 30 countries. Medicine. 2010;89(6):381.
33. Bowe F, Lipps CJ, Tsolis RM, Groisman E, Heffron F, Kusters JG. At least four percent of the Salmonella typhimurium genome is required for fatal infection of mice. Infection and immunity. 1998;66(7):3372-7.
34. Sabbagh SC, Forest CG, Lepage C, Leclerc JM, Daigle F. So similar, yet so different: uncovering distinctive features in the genomes of Salmonella enterica serovars Typhimurium and Typhi. FEMS microbiology letters. 2010;305(1):1-13.
35. Kubori T, Matsushima Y, Nakamura D, Uralil J, Lara-Tejero M, Sukhan A, et al. Supramolecular structure of the Salmonella typhimurium type III protein secretion system. Science. 1998;280(5363):602-5.
36. Kimbrough TG, Miller SI. Assembly of the type III secretion needle complex of Salmonella typhimurium. Microbes and infection. 2002;4(1):75-82.
37. Zhou D, Chen LM, Hernandez L, Shears SB, Galán JE. A Salmonella inositol polyphosphatase acts in conjunction with other bacterial effectors to promote host cell actin cytoskeleton rearrangements and bacterial internalization. Molecular microbiology. 2001;39(2):248-60.
38. Fu Y, Galán JE. A Salmonella protein antagonizes Rac-1 and Cdc42 to mediate host-cell recovery after bacterial invasion. Nature. 1999;401(6750):293.
39. Stender S, Friebel A, Linder S, Rohde M, Mirold S, Hardt WD. Identification of SopE2 from Salmonella typhimurium, a conserved guanine nucleotide exchange factor for Cdc42 of the host cell. Molecular microbiology. 2000;36(6):1206-21.

40. Hayward RD, Koronakis V. Direct nucleation and bundling of actin by the SipC protein of invasive Salmonella. The EMBO journal. 1999;18(18):4926-34.
41. McGhie EJ, Hayward RD, Koronakis V. Cooperation between actin-binding proteins of invasive Salmonella: SipA potentiates SipC nucleation and bundling of actin. The EMBO journal. 2001;20(9):2131-9.
42. Zhou D, Mooseker MS, Galán JE. Role of the S. typhimurium actin-binding protein SipA in bacterial internalization. Science. 1999;283(5410):2092-5.
43. Chen L-M, Hobbie S, Galan JE. Requirement of CDC42 for Salmonella-induced cytoskeletal and nuclear responses. Science. 1996;274(5295):2115-8.
44. Jones RM, Wu H, Wentworth C, Luo L, Collier-Hyams L, Neish AS. Salmonella AvrA coordinates suppression of host immune and apoptotic defenses via JNK pathway blockade. Cell host & microbe. 2008;3(4):233-44.
45. Brawn LC, Hayward RD, Koronakis V. Salmonella SPI1 effector SipA persists after entry and cooperates with a SPI2 effector to regulate phagosome maturation and intracellular replication. Cell host & microbe. 2007;1(1):63-75.
46. Bueno SM, Wozniak A, Leiva ED, Riquelme SA, Carreño LJ, Hardt WD, et al. Salmonella pathogenicity island 1 differentially modulates bacterial entry to dendritic and non-phagocytic cells. Immunology. 2010;130(2):273-87.
47. Rosales-Reyes R, Pérez-López A, Sánchez-Gómez C, Hernández-Mote RR, Castro-Eguiluz D, Ortiz-Navarrete V, et al. Salmonella infects B cells by macropinocytosis and formation of spacious phagosomes but does not induce pyroptosis in favor of its survival. Microbial pathogenesis. 2012;52(6):367-74.
48. Boyle EC, Brown NF, Finlay BB. Salmonella enterica serovar Typhimurium effectors SopB, SopE, SopE2 and SipA disrupt tight junction structure and function. Cellular microbiology. 2006;8(12):1946-57.
49. Hobbie S, Chen LM, Davis R, Galan J. Involvement of mitogen-activated protein kinase pathways in the nuclear responses and cytokine production induced by Salmonella typhimurium in cultured intestinal epithelial cells. The Journal of Immunology. 1997;159(11):5550-9.
50. Zhang Y, Higashide WM, McCormick BA, Chen J, Zhou D. The inflammation-associated Salmonella SopA is a HECT-like E3 ubiquitin ligase. Molecular microbiology. 2006;62(3):786-93.
51. Hersh D, Monack DM, Smith MR, Ghori N, Falkow S, Zychlinsky A. The Salmonella invasin SipB induces macrophage apoptosis by binding to caspase-1. Proceedings of the National Academy of Sciences. 1999;96(5):2396-401.
52. Bakowski MA, Braun V, Lam GY, Yeung T, Do Heo W, Meyer T, et al. The phosphoinositide phosphatase SopB manipulates membrane surface charge and trafficking of the Salmonella-containing vacuole. Cell host & microbe. 2010;7(6):453-62.
53. Mallo GV, Espina M, Smith AC, Terebiznik MR, Alemán A, Finlay BB, et al. SopB promotes phosphatidylinositol 3-phosphate formation on Salmonella vacuoles by recruiting Rab5 and Vps34. The Journal of cell biology. 2008;182(4):741-52.
54. Shea JE, Hensel M, Gleeson C, Holden DW. Identification of a virulence locus encoding a second type III secretion system in Salmonella typhimurium. Proceedings of the National Academy of Sciences. 1996;93(6):2593-7.
55. Hensel M, Shea JE, Raupach B, Monack D, Falkow S, Gleeson C, et al. Functional analysis of ssaJ and the ssaK/U operon, 13 genes encoding components of the type III secretion apparatus of Salmonella pathogenicity island 2. Molecular microbiology. 1997;24(1):155-67.
56. Nikolaus T, Deiwick J, Rappl C, Freeman JA, Schröder W, Miller SI, et al. SseBCD proteins are secreted by the type III secretion system of Salmonella pathogenicity island 2 and function as a translocon. Journal of bacteriology. 2001;183(20):6036-45.
57. Hensel M, Shea JE, Waterman SR, Mundy R, Nikolaus T, Banks G, et al. Genes encoding putative effector proteins of the type III secretion system of Salmonella pathogenicity island 2 are required for bacterial virulence and proliferation in macrophages. Molecular microbiology. 1998;30(1):163-74.

58. Kuhle V, Hensel M. SseF and SseG are translocated effectors of the type III secretion system of Salmonella pathogenicity island 2 that modulate aggregation of endosomal compartments. Cellular microbiology. 2002;4(12):813-24.
59. Uchiya Ki, Barbieri MA, Funato K, Shah AH, Stahl PD, Groisman EA. A Salmonella virulence protein that inhibits cellular trafficking. The EMBO Journal. 1999;18(14):3924-33.
60. Freeman JA, Rappl C, Kuhle V, Hensel M, Miller SI. SpiC is required for translocation of Salmonella pathogenicity island 2 effectors and secretion of translocon proteins SseB and SseC. Journal of bacteriology. 2002;184(18):4971-80.
61. Yu XJ, Ruiz-Albert J, Unsworth KE, Garvis S, Liu M, Holden DW. SpiC is required for secretion of Salmonella pathogenicity island 2 type III secretion system proteins. Cellular microbiology. 2002;4(8):531-40.
62. Guy RL, Gonias LA, Stein MA. Aggregation of host endosomes by Salmonella requires SPI2 translocation of SseFG and involves SpvR and the fms–aroE intragenic region. Molecular microbiology. 2000;37(6):1417-35.
63. Beuzón CR, Méresse S, Unsworth KE, Ruíz-Albert J, Garvis S, Waterman SR, et al. Salmonella maintains the integrity of its intracellular vacuole through the action of SifA. The EMBO journal. 2000;19(13):3235-49.
64. Miao EA, Miller SI. A conserved amino acid sequence directing intracellular type III secretion by Salmonella typhimurium. Proceedings of the National Academy of Sciences. 2000;97(13):7539-44.
65. Sory MP, Cornelis GR. Translocation of a hybrid YopE-adenylate cyclase from Yersinia enterocolitica into HeLa cells. Molecular microbiology. 1994;14(3):583-94.
66. Knodler LA, Celli J, Hardt WD, Vallance BA, Yip C, Finlay BB. Salmonella effectors within a single pathogenicity island are differentially expressed and translocated by separate type III secretion systems. Molecular microbiology. 2002;43(5):1089-103.
67. Brumell JH, Kujat-Choy S, Brown NF, Vallance BA, Knodler LA, Finlay BB. SopD2 is a novel type III secreted effector of Salmonella typhimurium that targets late endocytic compartments upon delivery into host cells. Traffic. 2003;4(1):36-48.
68. Henry T, Couillault C, Rockenfeller P, Boucrot E, Dumont A, Schroeder N, et al. The Salmonella effector protein PipB2 is a linker for kinesin-1. Proceedings of the National Academy of Sciences. 2006;103(36):13497-502.
69. Beuzón CR, Salcedo SP, Holden DW. Growth and killing of a Salmonella enterica serovar Typhimurium sifA mutant strain in the cytosol of different host cell lines. Microbiology. 2002;148(9):2705-15.
70. Ruiz-Albert J, Yu XJ, Beuzón CR, Blakey AN, Galyov EE, Holden DW. Complementary activities of SseJ and SifA regulate dynamics of the Salmonella typhimurium vacuolar membrane. Molecular microbiology. 2002;44(3):645-61.
71. Figueira R, Holden DW. Functions of the Salmonella pathogenicity island 2 (SPI-2) type III secretion system effectors. Microbiology. 2012;158(5):1147-61.
72. Haraga A, Miller SI. A Salmonella enterica serovar Typhimurium translocated leucine-rich repeat effector protein inhibits NF-κB-dependent gene expression. Infection and immunity. 2003;71(7):4052-8.
73. Mazurkiewicz P, Thomas J, Thompson JA, Liu M, Arbibe L, Sansonetti P, et al. SpvC is a Salmonella effector with phosphothreonine lyase activity on host mitogen-activated protein kinases. Molecular microbiology. 2008;67(6):1371-83.
74. Vazquez-Torres A, Xu Y, Jones-Carson J, Holden DW, Lucia SM, Dinauer MC, et al. Salmonella pathogenicity island 2-dependent evasion of the phagocyte NADPH oxidase. Science. 2000;287(5458):1655-8.
75. Van der Velden AW, Lindgren SW, Worley MJ, Heffron F. Salmonella Pathogenicity Island 1-Independent Induction of Apoptosis in Infected Macrophages bySalmonella enterica Serotype Typhimurium. Infection and immunity. 2000;68(10):5702-9.

76. Chakravortty D, Hansen-Wester I, Hensel M. Salmonella pathogenicity island 2 mediates protection of intracellular Salmonella from reactive nitrogen intermediates. Journal of Experimental Medicine. 2002;195(9):1155-66.
77. Worley MJ, Nieman GS, Geddes K, Heffron F. Salmonella typhimurium disseminates within its host by manipulating the motility of infected cells. Proceedings of the National Academy of Sciences. 2006;103(47):17915-20.
78. Retamal P. Modified Intracellular-Associated Phenotypes in a Recombinant Salmonella Typhi Expressing S. Typhimurium SPI-3 Sequences. 2010;5(2).
79. Amavisit P, Lightfoot D, Browning G, Markham P. Variation between pathogenic serovars within Salmonella pathogenicity islands. Journal of bacteriology. 2003;185(12):3624-35.
80. Blanc-Potard AB, Groisman EA. The Salmonella selC locus contains a pathogenicity island mediating intramacrophage survival. The EMBO journal. 1997;16(17):5376-85.
81. Dorsey CW, Laarakker MC, Humphries AD, Weening EH, Bäumler AJ. Salmonella enterica serotype Typhimurium MisL is an intestinal colonization factor that binds fibronectin. Molecular microbiology. 2005;57(1):196-211.
82. Wong K-K, McClelland M, Stillwell LC, Sisk EC, Thurston SJ, Saffer JD. Identification and sequence analysis of a 27-kilobase chromosomal fragment containing a Salmonella pathogenicity island located at 92 minutes on the chromosome map of Salmonella enterica serovar typhimurium LT2. Infection and immunity. 1998;66(7):3365-71.
83. Morgan E, Campbell JD, Rowe SC, Bispham J, Stevens MP, Bowen AJ, et al. Identification of host-specific colonization factors of Salmonella enterica serovar Typhimurium. Molecular microbiology. 2004;54(4):994-1010.
84. Gerlach RG, Jäckel D, Stecher B, Wagner C, Lupas A, Hardt WD, et al. Salmonella Pathogenicity Island 4 encodes a giant non-fimbrial adhesin and the cognate type 1 secretion system. Cellular microbiology. 2007;9(7):1834-50.
85. Gerlach RG, Cláudio N, Rohde M, Jäckel D, Wagner C, Hensel M. Cooperation of Salmonella pathogenicity islands 1 and 4 is required to breach epithelial barriers. Cellular microbiology. 2008;10(11):2364-76.
86. Marcus SL, Brumell JH, Pfeifer CG, Finlay BB. Salmonella pathogenicity islands: big virulence in small packages. Microbes and infection. 2000;2(2):145-56.
87. Kingsley RA, Santos RL, Keestra AM, Adams LG, Bäumler AJ. Salmonella enterica serotype Typhimurium ShdA is an outer membrane fibronectin-binding protein that is expressed in the intestine. Molecular microbiology. 2002;43(4):895-905.
88. Kingsley RA, Humphries AD, Weening EH, De Zoete MR, Winter S, Papaconstantinopoulou A, et al. Molecular and phenotypic analysis of the CS54 island of Salmonella enterica serotype Typhimurium: identification of intestinal colonization and persistence determinants. Infection and immunity. 2003;71(2):629-40.
89. Kingsley RA, van Amsterdam K, Kramer N, Baumler AJ. The shdA gene is restricted to serotypes of Salmonella enterica subspecies I and contributes to efficient and prolonged fecal shedding. Infect Immun. 2000;68(5):2720-7.
90. Wood MW, Jones MA, Watson PR, Hedges S, Wallis TS, Galyov EE. Identification of a pathogenicity island required for Salmonella enteropathogenicity. Molecular microbiology. 1998;29(3):883-91.
91. Folkesson A, Löfdahl S, Normark S. The Salmonella enterica subspecies I specific centisome 7 genomic island encodes novel protein families present in bacteria living in close contact with eukaryotic cells. Research in microbiology. 2002;153(8):537-45.
92. Haneda T, Ishii Y, Danbara H, Okada N. Genome-wide identification of novel genomic islands that contribute to Salmonella virulence in mouse systemic infection. FEMS microbiology letters. 2009;297(2):241-9.
93. Miller SI, Kukral AM, Mekalanos JJ. A two-component regulatory system (phoP phoQ) controls Salmonella typhimurium virulence. Proceedings of the National Academy of Sciences. 1989;86(13):5054-8.

94. Gunn JS, Alpuche-Aranda CM, Loomis WP, Belden WJ, Miller SI. Characterization of the Salmonella typhimurium pagC/pagD chromosomal region. Journal of bacteriology. 1995;177(17):5040-7.
95. Fuentes JA, Villagra N, Castillo-Ruiz M, Mora GC. The Salmonella Typhi hlyE gene plays a role in invasion of cultured epithelial cells and its functional transfer to S. Typhimurium promotes deep organ infection in mice. Research in microbiology. 2008;159(4):279-87.
96. Gulig PA. Virulence plasmids of Salmonella typhimurium and other salmonellae. Microbial pathogenesis. 1990;8(1):3-11.
97. Chu C, Hong S-F, Tsai C, Lin W-S, Liu T-P, Ou JT. Comparative physical and genetic maps of the virulence plasmids of Salmonella enterica serovars Typhimurium, Enteritidis, Choleraesuis, and Dublin. Infection and immunity. 1999;67(5):2611-4.
98. Roudier C, Fierer J, Guiney D. Characterization of translation termination mutations in the spv operon of the Salmonella virulence plasmid pSDL2. Journal of bacteriology. 1992;174(20):6418-23.
99. Libby SJ, Lesnick M, Hasegawa P, Weidenhammer E, Guiney DG. The Salmonella virulence plasmid spv genes are required for cytopathology in human monocyte-derived macrophages. Cellular microbiology. 2000;2(1):49-58.
100. Browne SH, Hasegawa P, Okamoto S, Fierer J, Guiney DG. Identification of Salmonella SPI-2 secretion system components required for SpvB-mediated cytotoxicity in macrophages and virulence in mice. FEMS Immunology & Medical Microbiology. 2008;52(2):194-201.
101. Paesold G, Guiney DG, Eckmann L, Kagnoff MF. Genes in the Salmonella pathogenicity island 2 and the Salmonella virulence plasmid are essential for Salmonella-induced apoptosis in intestinal epithelial cells. Cellular microbiology. 2002;4(11):771-81.
102. Guiney DG, Fierer J. The role of the spv genes in Salmonella pathogenesis. Frontiers in microbiology. 2011;2:129.
103. Lesnick ML, Reiner NE, Fierer J, Guiney DG. The Salmonella spvB virulence gene encodes an enzyme that ADP-ribosylates actin and destabilizes the cytoskeleton of eukaryotic cells. Molecular microbiology. 2001;39(6):1464-70.
104. Li H, Xu H, Zhou Y, Zhang J, Long C, Li S, et al. The phosphothreonine lyase activity of a bacterial type III effector family. Science. 2007;315(5814):1000-3.
105. Mastroeni P. Immunity to systemic Salmonella infections. Current molecular medicine. 2002;2(4):393-406.
106. Nanton MR, Way SS, Shlomchik MJ, McSorley SJ. Cutting edge: B cells are essential for protective immunity against Salmonella independent of antibody secretion. The Journal of Immunology. 2012;189(12):5503-7.
107. Ugrinovic S, Ménager N, Goh N, Mastroeni P. Characterization and Development of T-Cell Immune Responses in B-Cell-Deficient (Igh-6−/−) Mice with Salmonella enterica Serovar TyphimuriumInfection. Infection and immunity. 2003;71(12):6808-19.
108. Mittrücker H-W, Köhler A, Mak TW, Kaufmann SH. Critical role of CD28 in protective immunity against Salmonella typhimurium. The Journal of Immunology. 1999;163(12):6769-76.
109. Hess J, Ladel C, Miko D, Kaufmann S. Salmonella typhimurium aroA-infection in gene-targeted immunodeficient mice: major role of CD4+ TCR-alpha beta cells and IFN-gamma in bacterial clearance independent of intracellular location. The Journal of Immunology. 1996;156(9):3321-6.
110. Pie S, Truffa-Bachi P, Pla M, Nauciel C. Th1 response in Salmonella typhimurium-infected mice with a high or low rate of bacterial clearance. Infection and immunity. 1997;65(11):4509-14.
111. Roy M, Malo D. Genetic regulation of host responses to Salmonella infection in mice. Genes and immunity. 2002;3(7):381.
112. Vidal SM, Malo D, Vogan K, Skamene E, Gros P. Natural resistance to infection with intracellular parasites: isolation of a candidate for Bcg. Cell. 1993;73(3):469-85.
113. Vidal S, Malo D, Marquis J-F, Gros P. Forward genetic dissection of immunity to infection in the mouse. Annu Rev Immunol. 2008;26:81-132.

114. Caron J, Larivière L, Nacache M, Tam M, Stevenson MM, McKerly C, et al. Influence of Slc11a1 on the outcome of Salmonella enterica serovar Enteritidis infection in mice is associated with Th polarization. Infection and immunity. 2006;74(5):2787-802.
115. Monack DM, Bouley DM, Falkow S. Salmonella typhimurium persists within macrophages in the mesenteric lymph nodes of chronically infected Nramp1+/+ mice and can be reactivated by IFNγ neutralization. Journal of Experimental Medicine. 2004;199(2):231-41.
116. Hapfelmeier S, Hardt W-D. A mouse model for S. typhimurium-induced enterocolitis. Trends in microbiology. 2005;13(10):497-503.
117. Bohnhoff M, Drake BL, Miller CP. Effect of streptomycin on susceptibility of intestinal tract to experimental Salmonella infection. Proceedings of the Society for Experimental Biology and Medicine. 1954;86(1):132-7.
118. Wotzka SY, Nguyen BD, Hardt W-D. Salmonella Typhimurium diarrhea reveals basic principles of enteropathogen infection and disease-promoted DNA exchange. Cell host & microbe. 2017;21(4):443-54.
119. Barthel M, Hapfelmeier S, Quintanilla-Martínez L, Kremer M, Rohde M, Hogardt M, et al. Pretreatment of mice with streptomycin provides a Salmonella enterica serovar Typhimurium colitis model that allows analysis of both pathogen and host. Infection and immunity. 2003;71(5):2839-58.
120. Stecher B, Paesold G, Barthel M, Kremer M, Jantsch J, Stallmach T, et al. Chronic Salmonella enterica serovar Typhimurium-induced colitis and cholangitis in streptomycin-pretreated Nramp1+/+ mice. Infection and immunity. 2006;74(9):5047-57.
121. Thomas MK, Murray R, Flockhart L, Pintar K, Pollari F, Fazil A, et al. Estimates of the burden of foodborne illness in Canada for 30 specified pathogens and unspecified agents, circa 2006. Foodborne pathogens and disease. 2013;10(7):639-48.
122. Pollack SL. Consumer demand for fruit and vegetables: the US example. Changing structure of global food consumption and trade. 2001;6:49-54.
123. Kozak G, MacDonald D, Landry L, Farber J. Foodborne outbreaks in Canada linked to produce: 2001 through 2009. Journal of food protection. 2013;76(1):173-83.
124. Sivapalasingam S, Friedman CR, Cohen L, Tauxe RV. Fresh produce: a growing cause of outbreaks of foodborne illness in the United States, 1973 through 1997. Journal of food protection. 2004;67(10):2342-53.
125. Scharff RL, Besser J, Sharp DJ, Jones TF, Peter G-S, Hedberg CW. An economic evaluation of PulseNet: a network for foodborne disease surveillance. American journal of preventive medicine. 2016;50(5):S66-S73.
126. Lynch MF, Tauxe RV, Hedberg CW. The growing burden of foodborne outbreaks due to contaminated fresh produce: risks and opportunities. Epidemiology & Infection. 2009;137(3):307-15.
127. Tyler HL, Triplett EW. Plants as a habitat for beneficial and/or human pathogenic bacteria. Annual review of phytopathology. 2008;46:53-73.
128. Jaquette CB, Beuchat LR, Mahon BE. Efficacy of chlorine and heat treatment in killing Salmonella stanley inoculated onto alfalfa seeds and growth and survival of the pathogen during sprouting and storage. Appl Environ Microbiol. 1996;62(7):2212-5.
129. Reckhow DA, Singer PC, Malcolm RL. Chlorination of humic materials: byproduct formation and chemical interpretations. Environmental science & technology. 1990;24(11):1655-64.
130. Canada S. Statistical Overview of the Canadian Fruit Industry – 2016: Statistic Canada 2017 [Available from: http://www.agr.gc.ca/eng/industry-markets-and-trade/market-information-by-sector/horticulture/horticulture-sector-reports/statistical-overview-of-the-canadian-fruit-industry-2016/?id=1499785006783.
131. Canada S. Statistical Overview of the Canadian Vegetable Industry - 2016: Statistics Canada 2017 [updated 2017-08-23. Available from: http://www.agr.gc.ca/eng/industry-markets-and-trade/market-information-by-sector/horticulture/horticulture-sector-reports/statistical-overview-of-the-canadian-vegetable-industry-2016/?id=1501890793620#a3.3.2.

132. Xu H-S, Roberts N, Singleton F, Attwell R, Grimes D, Colwell R. Survival and viability of nonculturableEscherichia coli andVibrio cholerae in the estuarine and marine environment. Microbial ecology. 1982;8(4):313-23.
133. Pinto D, Santos MA, Chambel L. Thirty years of viable but nonculturable state research: unsolved molecular mechanisms. Critical reviews in microbiology. 2015;41(1):61-76.
134. Goodridge LD, Bisha B. Phage-based biocontrol strategies to reduce foodborne pathogens in foods. Bacteriophage. 2011;1(3):130-7.
135. Cheah Y-K, Salleh NA, Lee L-H, Radu S, Sukardi S, Sim J-H. Comparison of PCR fingerprinting techniques for the discrimination of Salmonella enterica subsp. enterica serovar Weltevreden isolated from indigenous vegetables in Malaysia. World Journal of Microbiology and Biotechnology. 2008;24(3):327.
136. Learn-Han L, Yoke-Kqueen C, Salleh NA, Sukardi S, Jiun-Horng S, Chai-Hoon K, et al. Analysis of Salmonella Agona and Salmonella Weltevreden in Malaysia by PCR fingerprinting and antibiotic resistance profiling. Antonie Van Leeuwenhoek. 2008;94(3):377.
137. Liming SH, Bhagwat AA. Application of a molecular beacon—real-time PCR technology to detect Salmonella species contaminating fruits and vegetables. International journal of food microbiology. 2004;95(2):177-87.
138. Mandal P, Biswas A, Choi K, Pal U. Methods for rapid detection of foodborne pathogens: an overview. American Journal Of Food Technology. 2011;6(2):87-102.
139. Ranjbar R. DNA Microarray for Rapid Detection and Identification of Food and Water Borne Bacteria: From Dry to Wet Lab. 2017;11:330-8.
140. McKillip J, Jaykus L, Drake M. A comparison of methods for the detection of Escherichia coli O157: H7 from artificially-contaminated dairy products using PCR. Journal of applied microbiology. 2000;89(1):49-55.
141. McKillip JL, Drake M. Real-time nucleic acid–based detection methods for pathogenic bacteria in food. Journal of Food Protection. 2004;67(4):823-32.
142. González-Escalona N, Hammack TS, Russell M, Jacobson AP, De Jesús AJ, Brown EW, et al. Detection of live Salmonella sp. cells in produce by a TaqMan-based quantitative reverse transcriptase real-time PCR targeting invA mRNA. Applied and environmental microbiology. 2009;75(11):3714-20.
143. Simpkins S, Chan A, Hays J, Cook N. An RNA transcription-based amplification technique (NASBA) for the detection of viable Salmonella enterica. Letters in applied microbiology. 2000;30(1):75-9.
144. Yaron S, Matthews K. A reverse transcriptase-polymerase chain reaction assay for detection of viable Escherichia coli O157: H7: investigation of specific target genes. Journal of applied microbiology. 2002;92(4):633-40.
145. Sheridan G, Masters C, Shallcross J, Mackey B. Detection of mRNA by Reverse Transcription-PCR as an Indicator of Viability in Escherichia coliCells. Applied and Environmental Microbiology. 1998;64(4):1313-8.
146. McIngvale S, Elhanafi D, Drake M. Optimization of reverse transcriptase PCR to detect viable Shiga-toxin-producing Escherichia coli. Applied and environmental microbiology. 2002;68(2):799-806.
147. Klein PG, Juneja VK. Sensitive detection of viable Listeria monocytogenes by reverse transcription-PCR. Applied and Environmental Microbiology. 1997;63(11):4441-8.
148. Hein I, Flekna G, Krassnig M, Wagner M. Real-time PCR for the detection of Salmonella spp. in food: an alternative approach to a conventional PCR system suggested by the FOOD-PCR project. Journal of microbiological methods. 2006;66(3):538-47.
149. Kao M-C, Durst RA. Detection of Escherichia coli using nucleic acid sequence-based amplification and oligonucleotide probes for 16S ribosomal RNA. Analytical Letters. 2010;43(10-11):1756-69.
150. Murray H. Connecting chromosome replication with cell growth in bacteria. Current opinion in microbiology. 2016;34:13-7.
151. del Mar Lleò M, Pierobon S, Tafi MC, Signoretto C, Canepari P. mRNA detection by reverse transcription-PCR for monitoring viability over time in an Enterococcus faecalis viable but nonculturable population maintained in a laboratory microcosm. Applied and Environmental Microbiology. 2000;66(10):4564-7.

152. Garg N, Churey J, Splittstoesser D. Effect of processing conditions on the microflora of fresh-cut vegetables. Journal of Food Protection. 1990;53(8):701-3.
153. Taboada EN, Graham MR, Carriço JA, Van Domselaar G. Food safety in the age of next generation sequencing, bioinformatics, and open data access. Frontiers in microbiology. 2017;8:909.
154. Fontana J, Stout A, Bolstorff B, Timperi R. Automated Ribotyping and Pulsed-Field Gel Electrophoresis for Rapid Identification of Multidrug-Resistant Salmonella Serotype Newport. Emerging Infectious Diseases. 2003;9(4):496-9.
155. Betancor L, Yim L, Fookes M, Martinez A, Thomson NR, Ivens A, et al. Genomic and phenotypic variation in epidemic-spanning Salmonella enterica serovar Enteritidis isolates. BMC microbiology. 2009;9(1):237.
156. Steele-Mortimer O, Méresse S, Gorvel JP, Toh BH, Finlay BB. Biogenesis of Salmonella typhimurium-containing vacuoles in epithelial cells involves interactions with the early endocytic pathway. Cellular microbiology. 1999;1(1):33-49.
157. Sansone A, Watson PR, Wallis TS, Langford PR, Kroll JS. The role of two periplasmic copper-and zinc-cofactored superoxide dismutases in the virulence of Salmonella choleraesuis. Microbiology. 2002;148(3):719-26.
158. Brooks AB, Humphreys D, Singh V, Davidson AC, Arden SD, Buss F, et al. MYO6 is targeted by Salmonella virulence effectors to trigger PI3-kinase signaling and pathogen invasion into host cells. Proceedings of the National Academy of Sciences. 2017;114(15):3915-20.
159. Hapfelmeier S, Ehrbar K, Stecher B, Barthel M, Kremer M, Hardt W-D. Role of the Salmonella pathogenicity island 1 effector proteins SipA, SopB, SopE, and SopE2 in Salmonella enterica subspecies 1 serovar Typhimurium colitis in streptomycin-pretreated mice. Infection and immunity. 2004;72(2):795-809.
160. Clark L, Perrett CA, Malt L, Harward C, Humphrey S, Jepson KA, et al. Differences in Salmonella enterica serovar Typhimurium strain invasiveness are associated with heterogeneity in SPI-1 gene expression. Microbiology. 2011;157(Pt 7):2072-83.
161. Tahoun A, Mahajan S, Paxton E, Malterer G, Donaldson DS, Wang D, et al. Salmonella transforms follicle-associated epithelial cells into M cells to promote intestinal invasion. Cell host & microbe. 2012;12(5):645-56.
162. Gorvel J-P, Méresse S. Maturation steps of the Salmonella-containing vacuole. Microbes and infection. 2001;3(14-15):1299-303.
163. Galan JE, Ginocchio C, Costeas P. Molecular and functional characterization of the Salmonella invasion gene invA: homology of InvA to members of a new protein family. Journal of bacteriology. 1992;174(13):4338-49.
164. Uyttendaele M, Vanwildemeersch K, Debevere J. Evaluation of real-time PCR vs automated ELISA and a conventional culture method using a semi-solid medium for detection of Salmonella. Letters in applied microbiology. 2003;37(5):386-91.
165. Van Kessel J, Karns J, Perdue M. Using a portable real-time PCR assay to detect Salmonella in raw milk. Journal of food protection. 2003;66(10):1762-7.
166. Ellingson JL, Anderson JL, Carlson SA, Sharma VK. Twelve hour real-time PCR technique for the sensitive and specific detection of Salmonella in raw and ready-to-eat meat products. Molecular and cellular probes. 2004;18(1):51-7.
167. Seo K, Valentin-Bon I, Brackett R, Holt P. Rapid, specific detection of Salmonella Enteritidis in pooled eggs by real-time PCR. Journal of food protection. 2004;67(5):864-9.
168. Eppinger M, Pearson T, Koenig SS, Pearson O, Hicks N, Agrawal S, et al. Genomic epidemiology of the Haitian cholera outbreak: a single introduction followed by rapid, extensive, and continued spread characterized the onset of the epidemic. MBio. 2014;5(6):e01721-14.
169. Chewapreecha C, Marttinen P, Croucher NJ, Salter SJ, Harris SR, Mather AE, et al. Comprehensive identification of single nucleotide polymorphisms associated with beta-lactam resistance within pneumococcal mosaic genes. PLoS genetics. 2014;10(8):e1004547.

170. Rasko DA, Webster DR, Sahl JW, Bashir A, Boisen N, Scheutz F, et al. Origins of the E. coli strain causing an outbreak of hemolytic–uremic syndrome in Germany. New England Journal of Medicine. 2011;365(8):709-17.
171. Gilchrist JJ, MacLennan CA, Hill AV. Genetic susceptibility to invasive Salmonella disease. Nature Reviews Immunology. 2015;15(7):452-63.
172. Maiden MC, Van Rensburg MJJ, Bray JE, Earle SG, Ford SA, Jolley KA, et al. MLST revisited: the gene-by-gene approach to bacterial genomics. Nature Reviews Microbiology. 2013;11(10):728.
173. Sheppard SK, Jolley KA, Maiden MC. A gene-by-gene approach to bacterial population genomics: whole genome MLST of Campylobacter. Genes. 2012;3(2):261-77.
174. Reinert K, Langmead B, Weese D, Evers DJ. Alignment of next-generation sequencing reads. Annual review of genomics and human genetics. 2015;16:133-51.
175. Bekal S, Berry C, Reimer A, Van Domselaar G, Beaudry G, Fournier E, et al. Usefulness of high-quality core genome single-nucleotide variant analysis for subtyping the highly clonal and the most prevalent Salmonella enterica serovar Heidelberg clone in the context of outbreak investigations. Journal of clinical microbiology. 2016;54(2):289-95.
176. Adhikari S, Curtis PD. DNA methyltransferases and epigenetic regulation in bacteria. FEMS microbiology reviews. 2016;40(5):575-91.
177. Storz G, Vogel J, Wassarman KM. Regulation by small RNAs in bacteria: expanding frontiers. Molecular cell. 2011;43(6):880-91.
178. Chanos P, Mygind T. Co-culture-inducible bacteriocin production in lactic acid bacteria. Applied microbiology and biotechnology. 2016;100(10):4297-308.
179. Emond-Rheault J-G, Jeukens J, Freschi L, Kukavica-Ibrulj I, Boyle B, Dupont M-J, et al. A Syst-OMICS approach to ensuring food safety and reducing the economic burden of salmonellosis. Frontiers in microbiology. 2017;8:996.
180. Xiao H, Gulen MF, Qin J, Yao J, Bulek K, Kish D, et al. The Toll–interleukin-1 receptor member SIGIRR regulates colonic epithelial homeostasis, inflammation, and tumorigenesis. Immunity. 2007;26(4):461-75.
181. Sham HP, Yu EYS, Gulen MF, Bhinder G, Stahl M, Chan JM, et al. SIGIRR, a negative regulator of TLR/IL-1R signalling promotes Microbiota dependent resistance to colonization by enteric bacterial pathogens. PLoS pathogens. 2013;9(8):e1003539.
182. Chapman SJ, Hill AV. Human genetic susceptibility to infectious disease. Nature Reviews Genetics. 2012;13(3):175.
183. van de Vosse E, van Dissel JT, Ottenhoff TH. Genetic deficiencies of innate immune signalling in human infectious disease. The Lancet infectious diseases. 2009;9(11):688-98.
184. Mian MF, Pek EA, Chenoweth MJ, Ashkar AA. Humanized mice are susceptible to Salmonella typhi infection. Cellular & molecular immunology. 2011;8(1):83.
185. Libby SJ, Brehm MA, Greiner DL, Shultz LD, McClelland M, Smith KD, et al. Humanized nonobese diabetic-scid IL2rγnull mice are susceptible to lethal Salmonella Typhi infection. Proceedings of the National Academy of Sciences. 2010;107(35):15589-94.

THESIS NON-EXCLUSIVE LICENSE

AUTHOR'S NAME: _____ _____ _____
LAST NAME *FIRST NAME* *INITIALS*

AUTHOR'S DATE OF BIRTH: _____/_____/_____
MONTH / DAY / YEAR

MCGILL EMAIL ADDRESS: alanna.crouse@mail.mcgill.ca **STUDENT NUMBER:** _____

PERMANENT ADDRESS: _____
COUNTRY

MCGILL UNIT: _____

FACULTY: _____

DEGREE SOUGHT: _____

TITLE OF THESIS: A syst-OMICS approach to food safety: Identifying genetic determinants of Salmonella virulence in vivo

I hereby promise that I am author of the thesis above cited.

I confirm that my thesis is my original work, does not infringe any rights of others, and that I have the right to make the grant conferred by this non-exclusive license. I also confirm that if third party copyrighted material was included in my thesis for which, under the terms of the Copyright Act, written permission from the copyright owners is required, I have obtained such permission from the copyright owners and may grant such permission for the full term of copyright protection.

I hereby grant to McGill University a non-exclusive, worldwide, irrevocable, royalty free license, in respect of my thesis, to reproduce, convert, publish, archive, preserve, conserve, communicate and distribute, and loan, in paper form, in microform, electronically by telecommunication or on the internet, and/or any other formats as may be adopted for such use from time to time. I also authorize McGill University to sub-license, sub-contract for any of the acts mentioned.

I hereby grant permission and authorize McGill University to permit access to my thesis and make it available to interested persons in paper or electronic form, through the library, interlibrary and public loan.

I understand that I retain copyright ownership and moral rights in my thesis, and that I may deal with the copyright in my thesis consistent with these rights. I promise to inform any person to whom I may hereafter assign or license my copyright in my thesis of the rights granted by me to **McGill University** and to **Library and Archives Canada.**

I confirm that I have executed a **Non-exclusive license with the Library and Archives Canada** and hereby grant permission to McGill University to submit the abstract and my thesis to the **Library and Archives Canada** in full compliance with these non-exclusive licenses.

The authorization contained in these non-exclusive licenses are to have effect on the date given below (**Effective Date**) unless a deferral of one year from the date has been expressly requested by me, the author, on submitting the thesis, and acknowledged by McGill University **GRADUATE AND POSTDOCTORAL STUDIES OFFICE.**

I agree to indemnify and hold McGill University harmless against any claim and any loss, damage, settlement cost or expense (including legal fees) incurred by McGill University and arising out of, or in connection with, my statements and representations or this license.

SIGNATURE OF AUTHOR: _____ **EFFECTIVE DATE:** _____

THESES NON-EXCLUSIVE LICENSE

Surname:	Given Names:
Full Name of University:	
Faculty, Department, School:	
Degree for which thesis was presented:	Date Degree Awarded:
Thesis Title:	
Supervisor:	
Date of Birth. It is **optional** to supply your date of birth. If you choose to do so please note that the information will be included in the bibliographic record for your thesis.	
E-mail: please provide your e-mail address if you are interested in receiving royalties on sales of your thesis.	
Permanent Address: please provide your permanent address if you are interested in receiving royalties on sales of your thesis.	

In consideration of Library and Archives Canada making my thesis available to interested persons, I,

_____

hereby grant a non-exclusive, for the full term of copyright protection, license to Library and Archives Canada:

(a) to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell my thesis (the title of which is set forth above) worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats;

(b) to authorize, sub-license, sub-contract or procure any of the acts mentioned in paragraph (a).

I undertake to submit my thesis, through my university, to Library and Archives Canada. Any abstract submitted with the thesis will be considered to form part of the thesis.

I represent and promise that my thesis is my original work, does not infringe any rights of others, and that I have the right to make the grant conferred by this non-exclusive license.

If third party copyrighted material was included in my thesis for which, under the terms of the Copyright Act, written permission from the copyright owners is required I have obtained such permission from the copyright owners to do the acts mentioned in paragraph (a) above for the full term of copyright protection

I retain copyright ownership and moral rights in my thesis, and may deal with the copyright in my thesis, in any way consistent with rights granted by me to Library and Archives Canada in this non-exclusive license.

I further promise to inform any person to whom I may hereafter assign or license my copyright in my thesis of the rights granted by me to Library and Archives Canada in this non-exclusive license.

Signature	**Date**