Ministry for Primary Industries
Manatū Ahu Matua

Source attribution for salmonellosis using microbial subtyping

MPI Technical Paper No: 2016/44

Prepared for the Ministry for Primary Industries by
Professor Nigel French, Dr Eve Pleydell and Dr Jonathan Marshall (Massey University), Dr Philip Carter (ESR), Dr Craig Thornley and Dr Donald Campbell (MPI)

ISBN No: 978-1-77665-329-4 (online)
ISSN No: 2253-3923 (online)

December 2013

Disclaimer

While every effort has been made to ensure the information in this publication is accurate, the Ministry for Primary Industries does not accept any responsibility or liability for error of fact, omission, interpretation or opinion that may be present, nor for the consequences of any decisions based on this information.

Requests for further copies should be directed to:

Publications Logistics Officer
Ministry for Primary Industries
PO Box 2526
WELLINGTON 6140

Email: brand@mpi.govt.nz
Telephone: 0800 00 83 33
Facsimile: 04-894 0300

This publication is also available on the Ministry for Primary Industries website at http://www.mpi.govt.nz/news-and-resources/publications/

Scientific Interpretative Summary

This SIS is prepared by MPI risk assessors to provide context to the following report for MPI risk managers and external readers

Source attribution for human salmonellosis using microbial subtyping

Background

Non-typhoidal *Salmonella* are a significant cause of foodborne illness in New Zealand and of concern to the Ministry for Primary Industries (MPI) as an occasional contaminant of food and food products. Source attribution is the process of evaluating patterns and trends of potential reservoirs for human illness e.g. types of animals, and is helpful for prioritising further food safety research and targeting interventions. A 12-month nationwide source attribution study of human salmonellosis was carried out so as to better inform future risk reduction measures.

For this study, a microbial subtyping approach was applied and the degree of similarity in *Salmonella* subtypes from human and animal specimens was used to estimate the relative importance of animal reservoirs as origins for bacteria infecting humans.

Source reservoir sampling was limited to animals and their immediate environment. Human cases were restricted to those likely to have acquired infection within New Zealand. The study focus was on *Salmonella enterica* subsp.*enterica.*

Methods

This study brought together the *Salmonella* typing expertise of the Enteric Reference Laboratory (ERL) at ESR, with the source attribution modelling skills of the Molecular Epidemiology and Public Health Laboratory (mEpiLab) at Massey University. A total of 939 non-typhoidal *Salmonella* isolates from separate samples were collected over a 12 month period.. Isolates were obtained from the ESR Enteric Reference Laboratory collection of human diagnostic and non-human isolates (veterinary diagnostic specimens and environments associated with food animals) and from proactive sampling of faecal material in animal handling areas (lairages) at abattoirs. Faecal samples were also collected from wild birds in urban habitats. Isolates included in the study were: 499 from human cases, 198 from cattle, 135 from sheep and sheep lairages, 76 from poultry and poultry environments, 21 from wild birds and 10 from pig lairages. Isolates were phenotyped using serotype and phage type, and genotyped using pulsed-field gel electrophoresis (PFGE) and multilocus variable-number tandem repeat analysis (MLVA). The data were evaluated using hierarchical cluster analysis, rarefaction curves, proportional similarity indices and multidimensional scaling analysis, prior to being fitted within modified Hald source attribution models.

Results and Discussion

Attribution models fitted to the serotype and PFGE typing data indicated that, between May 2011 and April 2012, 60% (95% credible interval [CI] 30-86%) of human salmonellosis cases were associated with *Salmonella* subtypes that are also present in cattle. Sixteen percent (95% CI 1-44%) and 15% (95% CI 0-42%) of human salmonellosis cases were associated with *Salmonella* subtypes that are also present in poultry and wild birds respectively. Fewer than 10% of cases were associated with *Salmonella* subtypes from other sources (e.g. sheep and pigs). *Salmonella* was rarely isolated from faecal samples collected from healthy pigs and sheep in abattoir lairages.

Salmonella Typhimurium was the most frequently occurring serotype, accounting for almost half of the isolates typed from human cases and over two-thirds of isolates from cattle and wild birds. Of the human cases of *Salmonella* Typhimurium infection, 59% (95% CI 29-83%) and 63% (95% CI 37-85%) were associated with *S.* Typhimurium subtypes that are also present in cattle when estimated using phage typing and MLVA typing, respectively. Forty-nine percent (95% CI 20-76%) and 24% (95% CI 2-52%) of human cases with non-Typhimurium salmonellosis were associated with *Salmonella* subtypes that are also present in cattle and wild birds, respectively.

In respect of the cattle findings, it is possible that the range of cattle subtypes in the study may have been biased towards those associated with salmonellosis illness in cattle, particularly as *Salmonella* detections in cattle were common during the study period. Nonetheless, the range of cattle *Salmonella* subtypes were similar to those previously observed.

This is the first time in New Zealand that such a large dataset of non-typhoidal *Salmonella* isolates had been assembled from human and animals. It should be noted that source was the only attribute available to describe the isolates so no information is available on pathways. Estimates may have been affected by the large number of serotypes from human cases that were not recovered from the animal reservoirs.

MASSEY UNIVERSITY
COLLEGE OF SCIENCES
TE WĀHANGA PŪTAIAO

Final report: Project 11779 / 15084
Source attribution for salmonellosis using microbial subtyping

13th December 2013

A report prepared for the

Ministry for Primary Industries

by

Professor Nigel French
Dr Eve Pleydell
Dr Philip Carter
Dr Jonathan Marshall

[m]EpiLab, Hopkirk Research Institute
IVABS, College of Sciences
Massey University
Palmerston North
New Zealand
Email: E.J.Pleydell@massey.ac.nz
Phone: +64 (06) 356 9099 extn 81193

Enteric Reference Laboratory, ESR
34 Kenepuru Drive
Porirua
New
Email: Philip.Carter@esr.cri.nz
Phone: +64 (04) 914 0789
Fax: +64 (06) 350 5714

Contents

1. Executive summary

This study brought together the *Salmonella* typing expertise of the Enteric Reference Laboratory (ERL) at ESR, with the source attribution modelling skills of the Molecular Epidemiology and Public Health Laboratory (mEpiLab) at Massey University. For the first time in New Zealand, a comprehensive dataset of non-typhoidal *Salmonella* isolates was assembled from human clinical cases and potential animal reservoirs. Nine hundred and fifty six isolates of *Salmonella enterica* were collected and typed from both human clinical cases and potential animal reservoirs over a 12 month period. The typing methods utilised included phenotyping (serotype and phagetype) and genotyping (Pulsed Field Gel Electrophoresis (PFGE) and Multilocus variable-number tandem repeat analysis (MLVA)). The data were explored using hierarchical cluster analysis, rarefaction curves, proportional similarity indices and multidimensional scaling analysis, prior to being fitted within modified Hald source attribution models. This use of phenotypic and genotypic typing techniques, combined with molecular epidemiological analyses and source attribution modelling, revealed several patterns and trends.

Attribution models fitted to both the serotype and PFGE typing data of the 939 isolates indicated that, between May 2011 and April 2012, cattle were a major source of domestically-acquired human salmonellosis in New Zealand and, in particular, were the predominant source of *Salmonella* Typhimurium infections. Although the estimates were associated with wide confidence intervals, cattle were estimated to be associated with 60% (30-86%) of all salmonellosis cases using the most discriminatory genotyping method (PFGE data). Poultry and wild birds were assigned as the source of 16% (1-44%) and 15 (0-42%) of salmonellosis cases respectively, and other sources (e.g. sheep and pigs) were estimated to account for fewer than 10% of cases. *Salmonella* Typhimurium was the most frequently occurring serotype in New Zealand, accounting for almost half of the isolates that were typed from human clinical cases and over two thirds of the isolates from cattle and wild birds.

Cattle were attributed as the source of 59% (29-83%) of these *Salmonella* Typhimurium cases using phage typing and 63% (37-85%) using MLVA. Similarly, cattle were identified as the most common source of infection with non-Typhimurium serotypes; they were estimated to account for 49% (20-76%) of cases, followed by wild birds at 24% (2-52%). As for all the attribution estimates in this report, it is very important to note the wide and overlapping confidence (credible) intervals.

The sampling frame of 939 isolates was from a single year's submission to ERL from which 499 human isolates were randomly selected and subtyped. A notable feature of this collection of isolates is that there are a large number of serotypes causing human disease in New Zealand that we have not recovered from the animal reservoirs that were studied. This implies that there are important sources of disease that have not been considered here, and/or that the isolates collected from the reservoirs are not representative of the total number of strains present within those reservoirs.

In addition to the human isolates, a target of 150 typed isolates was set for each of four potential livestock reservoirs: cattle, sheep, pigs and poultry. The livestock isolates available included all the National Microbiological Database (NMD) isolates over the 12 month period (or a random sample if there were more than 150). For food-animal sources that were unlikely to provide 150 isolates from ERL submissions, such as pigs and sheep, a feasibility study of abattoir lairage sampling was carried out at five slaughter plants located in North and South Island (two sheep plants and three pig plants). A second pilot study was also conducted to examine the carriage of *Salmonella* in faeces from wild birds located in urban centres.

The collection of isolates from sheep was dominated by the two serotypes that most commonly cause clinical disease in this species; namely, *S.* Hindmarsh and *S.* Brandenburg. These serotypes were minority subtypes within the human isolates in this study. This, coupled with a lack of positive faecal samples collected from healthy sheep in abattoir lairages, implies that sheep do not present a high risk of *Salmonella* infection to the population of New Zealand at this current time.

Salmonella was rarely isolated from faecal samples collected from healthy pigs at abattoir lairages, and very few *Salmonella* were sent directly to the ERL, resulting in a very low number of isolates typed from this livestock species. It is unknown whether this reflects a true low prevalence of *Salmonella* within the New Zealand pigs.

The *Salmonella* isolates collected from wild birds showed strong similarities to strains causing human disease, thus highlighting the potential for some species of wild birds to act as vectors for disease spread.

At the current time, no single typing system of appropriate discrimination is available for all serotypes. Therefore, source attribution studies of *Salmonella* need to incorporate a variety of typing systems and gauge the reproducibility of the estimates from multiple models fitted to different data.

2. Introduction

Non-typhoidal salmonellosis is recognised as a significant cause of food-borne human illness in New Zealand. To this end, the Ministry of Primary Industries (MPI) in 2007 developed three strategic goals to reduce the public health burden of infection, including the aim of achieving a 30% reduction in

reported annual incidence of food-borne salmonellosis after five years. To inform the MPI *Salmonella* Strategy a 12-month, nationwide source attribution study was established.

Source attribution is the process of identifying which sources or reservoirs are responsible for human illness. Determining the proportion of cases that originate from each food source and the principal pathways that lead to contamination of the food supply are important for prioritising interventions and developing control strategies to reduce the burden of disease on the human population. Source attribution can be achieved using a variety of methods, which include microbial typing, epidemiological data, intervention strategies and expert opinion.

This study has taken a targeted, risk-based approach to source attribution, which is likely to be more cost-effective. Efficiency was further enhanced by utilising information that was already gathered by on-going surveillance activities.

Isolates from a single year's submission to ERL formed the sampling frame from which 499 human isolates were randomly selected and subtyped using a variety of genotyping methods. In tandem, a target of 150 typed isolates was set for each of four potential livestock reservoirs: cattle, sheep, pigs and poultry. The livestock isolates available included all the NMD isolates over the 12 month period (or a random sample if there are more than 150). For food-animal sources that were unlikely to provide 150 isolates from ERL submissions, such as pigs and sheep, a feasibility study of abattoir lairage sampling was carried out at five slaughter plants located in North and South Island (two sheep plants and three pig plants). A second pilot study was also conducted to examine the carriage of *Salmonella* in faeces from wild birds located in urban centres.

3. Methods

3.1. Sampling strategy – ERL isolates

The period of isolate collection ran from 1st May 2011 to 30th April 2012. Isolates of *Salmonella* Typhi were not included in the study, and non-typhoidal isolates obtained from people who were recorded as having travelled overseas during the incubation period for salmonellosis were also excluded.

In the first instance, every other eligible isolate received by ERL was selected for inclusion and subtyping. At regular intervals the numbers of isolates that had been selected for subtyping were assessed and the species of origin and geographical locations of source. If the numbers of subtyped isolates had not reached the anticipated numbers for that period, then a stratified random selection of other isolates was included.

All cattle isolates included in the study were obtained from the ERL archive over the relevant time period. Only 6.6% of these were from NMD testing, and all of these were isolated from bobby calves. The remainder were of unknown provenance, but most likely from samples submitted via veterinary diagnostic laboratories as part of veterinary investigations of animal health problems (outbreaks or sporadic cases of suspect salmonellosis).

The poultry samples submitted to ERL are not always accompanied by specific information as to their source and reason for sampling. Looking at the data available for the samples of poultry origin that were submitted during the study period, it was decided that only the samples identified as faeces, cloacal swabs, and environmental would be included in the typing dataset. It was assumed

that the environmental samples were taken from within the poultry sheds and would thereby represent faecal contaminants in that environment. The uncertainty as to the source and reason for sampling of the poultry isolates means that the representativeness of this dataset is uncertain. Note that all samples identified as poultry were considered to be from chicken (*Gallus gallus domesticus*); specimens and strains identified as "duck" were taken from wildfowl, as described in 3.3.

3.2. Supplementary livestock sampling – ^{m}EpiLab, Massey University

Sufficient numbers of isolates from cattle were obtained from the ERL. However, in order to increase the number of *Salmonella* collected from pigs and sheep, proactive supplementary sampling was conducted by the molecular epidemiology laboratory team at Massey University. After consultation with Dr Rob Davies at the Animal Health and Veterinary Laboratories Agency, UK, sampling methods for collecting faecal samples from animals at an abattoir were developed. These consisted of collecting freshly voided faeces in the lairage. These samples were collected from the floor whilst a mob was in a pen, or immediately after a mob had vacated a pen. Care was taken to ensure that only fresh faecal pats that had not been trampled were selected, and the samples were not collected from areas of the faecal pat that were in contact with the floor itself. Up to three well-spaced, samples were taken from a single group of animals in a pen.

The samples were collected using two methods, and the choice of method depended upon the size and consistency of the faecal pat being sampled. If the faeces were large and well-formed, then pinches of faeces were taken with a gloved hand and placed into a sterile pot, and gloves were changed between samples. Smaller or looser faeces were sampled using wand swabs that were placed into charcoal transport media.

Five abattoirs were recruited onto the study: two sheep plants and three pig plants. The participating processing plants were situated on North and South Island. Due to a lower than anticipated throughput for the selected South Island pig plant in 2011, a second South Island plant was included in the 2012 round of sampling.

In addition to the samples collected during the abattoir visits, mEpiLab researchers visiting sheep farms (n=9) for other projects also collected freshly voided faecal samples. The details of the 633 samples that were collected are shown in Table 1.

Table 1. Details of the 633 faecal samples collected for attempted *Salmonella* isolation from pigs and sheep in New Zealand between 30th June 2011 and 6th March 2012.

Region of farm of origin	Pig		Sheep		Total no. Samples
	No. batches[1]	No. samples	No. batches	No. samples	
North island					
Hawke's Bay	3	52	4	38	90
Manawatu-Wanganui	6	51	10	61	112
Taranaki	9	61	0	0	61
Subtotal	18	164	14	99	263
South Island					
Canterbury	12	134	14	104	238
Marlborough	0	0	2	16	16
Otago	0	0	3	18	18
Southland	2	6	2	12	18
Subtotal	14	140	21	150	290
Unknown	1	3	13	77	80
Grand Total	33	307	68	326	633

[1] A batch of animals is defined as a group arriving together at the abattoir and originating from the same farm.

At the laboratory, up to four samples originating from the same mob of animals were pooled into 20ml buffered peptone water (BPW) prior to processing. After incubation overnight, 100ul of BPW was transferred into Rappaport Vassiliadis Soy peptone enrichment broth (RVS) and incubated for another 24 hours. The incubated selective broth was then plated onto MacConkey agar, prior to plating onto two further selective agars: XLD and BGM. Potential *Salmonella* growing on the agar plates were inoculated into two differential media TSI and LIA, and were then tested against poly-O and poly-H antisera. Presumptive *Salmonella* were frozen in glycerol, and were also inoculated onto Dorset egg slopes for transport to ERL for serotyping and genotyping.

3.3. Pilot study of urban wild birds – mEpiLab, Massey University

During 2011, initial work sampling wild birds was undertaken in Palmerston North, Whanganui and Dunedin. Following this, over the summer of 2011-2012, wild bird populations were sought and sampled twice in each of the three main urban centres, Auckland, Wellington, and Christchurch. The local councils for these cities were contacted regarding sites in the cities where birds such as gulls, starlings, sparrows and ducks were known to roost or gather. Contact was also made with Dr Yolanda van Heezik, a zoologist at Otago University with expertise in urban wildlife and biodiversity. Published prevalence of *Salmonella* shedding in other countries by higher risk wild bird species range between 2.5-15.9% (Snoeyenbos, Morin et al. 1967; De Sousa, Berchieri et al. 2010; Kitadai, Ninomiya et al. 2010; Carlson, Franklin et al. 2011).

The aim of this study was to try to collect 20-25 isolates per visit, which entailed collecting a large number of samples per visit. Therefore, pooling of faecal samples prior to laboratory processing was necessary. The details of the 817 samples collected from wild birds are shown in Table 2.

Table 2. Summary of 817 fresh faecal samples collected from wild birds in urban habitats.

Species	City						Total
	Auckland	Wellington	Christ-church	Palmerston North	Dunedin	Whanganui	
Gulls							
Red-billed	9	-	1	-	8	-	18
Black-billed		-	37	-		-	37
Black-backed	1	-		-	1	-	2
Unspecified	45	37	34	-		-	116
Subtotal	55	37	72	0	9	0	173
Waterfowl							
Duck	41	37	48	22	9	-	157
Pukeko	10	-	-	-	-	-	10
Goose	4	-	-	-	1	-	5
Black swan	4	-	-	-	-	-	4
Unspecified		-	-	1	-	-	1
Subtotal	59	37	48	23	10	0	177
Passerines							
Sparrow	55	54	55	16	21	1	202
Starling	65	26	42	3	-	2	138
Unspecified	-	-	-	11	-	-	11
Subtotal	120	80	97	30	21	3	351
Pigeon	44	42	29	0	1	0	116
Grand Total	278	196	246	53	41	3	817

A validation study was undertaken in the laboratory to test whether the methodology chosen was able to detect low numbers of *Salmonella*. In summary, a *Salmonella* isolate originating from an avian host was grown in the laboratory and serial dilutions were made to provide four test dilutions of approximately 400, 80, 16 and 3 cfu/ml, respectively. The actual concentrations of these dilutions were checked using spiral plate counts. For each dilution, 100ul of the *Salmonella* dilution was transferred into 20ml of BPW along with a swab of waterfowl faeces. The BPW was inoculated overnight and 100ul of the incubated broth was added to a selective broth, incubated for another 24 hours and then plated onto selective agar. Three different selective broths (RVS, selenite and tetrathionate) were tested on each of three different selective agars (bile green, XLD and CHROMagar). An uninoculated broth of RVS was used as a negative control. The culture results for this study are shown in Table 3.

Table 3. An assessment of the growth of *Salmonella* of avian origin using three types of selective broth and three types of selective agar.

Approximate concentration of *Salmonella* (cfu/ml)	RVS			Selenite			Tetrathionate		
	BG	XLD	CHROM	BG	XLD	CHROM	BG	XLD	CHROM
3	-	-	-	-	-	3	-	-	-
16	++	++	++	+	-	+	++	++	++
80	++	+ (+ others)	+	1	1	3	+	++	++
400	++	++	+	+	-	+	++	++	++
Negative control	-	-	-	+	-	-	-	-	-

Key: - = no growth; 1-3 = numbers of colonies present; + = medium growth; ++ = profuse growth.

Following this validaton study, RVS broth was chosen with subsequent plating onto MacConkey, bile green agar and XLD.

3.4.Characterisation of isolates

3.4.1. Serotyping

All isolates are being serotyped by the *Salmonella* Reference Laboratory at NCBID using the Kauffman-White classification scheme.

3.4.2. Pulsed Field Gel Electrophoresis (PFGE)

All samples for PFGE were prepared and the gels were run following the protocols outlined in the CDC guidelines for the application of PFGE for *Salmonella*. A single enzyme, *Xba*I, is used to digest genomic DNA from each isolate and the gels are analysed using BioNumerics.

3.4.3. Multilocus variable-number tandem repeat analysis (MLVA)

MLVA was performed according to Lindstedt et al. (Lindstedt et al 2006) on the 424 S. Typhimurium isolates.

3.4.4. Phagetyping

Phage types were identified using the scheme described by Anderson (Anderson et al 1977).

3.5. Descriptive analyses

3.5.1. Rarefaction curves of "type" richness

Rarefaction is a technique that was developed by ecologists to assess species richness in a given habitat (Heck, van Belle et al. 1975). In the context of this work, a rarefaction curve is a plot of the number of "types" identified (either phenotypes or genotypes) as a function of the number of isolates that were typed. If the curve flattens towards a plateau it is likely that the majority of

different types have been sampled for that particular source species; if the curve remains steep it is likely that more types would be identified if more isolates were typed.

3.5.2. Descriptive cluster analysis using BioNumerics

The PFGE gel pictures were imported into BioNumerics to allow for cluster analysis of similar banding patterns. To ascertain clusters of isolates, Dice's similarity coefficient was used to produce a distance matrix with an optimisation setting of 1% and a tolerance setting of 1.5%. The unweighted pair group method with arithmetic means (UPGMA) algorithm was used to produce a dendrogram and a cluster was defined as a group that contained five or more isolates.

BioNumerics was also used to identify unique "pulse types". A pulse type was defined as a single isolate or group of isolates that were identified as sharing a unique PFGE banding pattern. To obtain pulse type designations, a distance matrix was constructed using the number of band differences with an optimisation setting of 2% and a tolerance setting of 2%. The complete linkage algorithm was used to construct a dendrogram that grouped isolates of 100% similarity together. Each group was designated a unique pulse type identity.

3.5.3. Multidimensional scaling analysis of *S.* Typhimurium pulse types

The relationships between sub-populations within the *Salmonell*a Typhimurium dataset were further explored using multidimensional scaling (MDS) and permutational multivariate analysis of variance (PERMANOVA) (McArdle and Anderson, 2001). These techniques use a pairwise similarity matrix generated from the PFGE profiles to:

1) Represent the relatedness of isolates in a 2-dimensional space
2) Test specific hypotheses concerning the contribution of covariates to the variation in pairwise distances and calculate components of variation.

3.6.Source attribution mdoels

3.6.1. Proportional similarity index matrices

The proportional similarity index (PSI) is a measure of the overlap between the frequency distributions of "types" obtained from two source groups (Rosef, Kapperud et al. 1985). A diagonal matrix of pairwise PSI values from different groups can be produced and visualised as a neighbour-net network allowing for the simultaneous assessment of the relative dissimilarities in frequency distributions of all groups (Huson and Bryant 2006).

3.6.2. Modified Hald source attribution models

The modified Hald source attribution model developed by Muellner et al. was fitted to each set of typing data individually (Mullner, Jones et al. 2009).

4. Results

4.1. Supplementary livestock sampling

Presumptive *Salmonella* was grown from five of the 307 (1.6%) pooled faecal samples that were collected from pig abattoir lairages. Four of these positive samples were collected from two consecutive visits to the same pig processing plant, and the positive animals were identified as being from the same farm on both occasions. The details of these samples are shown in Table 2 together with the serotyping results of the isolates obtained from them.

Table 4. Details of the five pooled faecal samples from pigs that were culture-positive for *Salmonella.*

Collection date	No. of positive pooled samples	Serotypes
7/11/2011	2	*Salmonella* Derby
12/12/2011	2	*Salmonella* Derby
06/03/2012	1	*Salmonella* Brandenburg

Although 326 pooled faecal samples were also collected from sheep lairages during this study, no presumptive *Salmonella* were cultured from any of the pooled ovine samples.

4.2. Pilot study of urban wild birds

Using these methods, presumptive *Salmonella* was cultured from eight pooled samples (Table 5). Serotyping has been undertaken on isolates grown from three pooled avian samples collected in Christchurch. All three pools contained a mix of two serotypes: *Salmonella enterica* subsp. *enterica* (I) ser. 4,12 : - : 2 and *Salmonella* Typhimurium. A mix of two serotypes (Typhimurium and Enteriditis) was also found in a pool of samples collected from gulls at an Auckland refuse station; whilst, *Salmonella* Oranienburg was isolated from gulls at a Wellington refuse station. Full serotyping results of the wild bird isolates are shown in Table 5.

Table 5. Details of the eight pooled faecal samples from wild birds that were culture-positive for *Salmonella.*

Collection site	No. positive pooled samples	Collection date	Species	Serotypes identified
Christchurch				
Site 1: Bush	1	28/10/2011	Starlings	*Salmonella e.* subsp. *enterica* (I) ser. 4,12 : - : 2 & *Salmonella* Typhimurium
Site 2: Lake	1	29/10/2011	Sparrows	*Salmonella e.* subsp. *enterica* (I) ser. 4,12 : - : 2 & *Salmonella* Typhimurium
Site 3: Wastewater treatment plant	1	30/10/2011	Sparrows	*Salmonella e.* subsp. *enterica* (I) ser. 4,12 : - : 2 & *Salmonella* Typhimurium
Wellington				
Site 4: Refuse station	1	6/01/2012	Gulls	*Salmonella* Oranienburg
Site 5: Seafront	1	7/01/2012	Ducks	*Salmonella* Enteriditis
Auckland				
Site 6: Park	1	11/01/2012	Starlings	*Salmonella* Typhimurium
Site 7: Refuse station	2	11/01/2012	Gulls	*Salmonella* Enteriditis & *Salmonella* Typhimurium

4.3. General description of the library of isolates

A total of 960 *Salmonella* isolates were collected, digested and typed using PFGE. Of these 939 were identified as originating from separate samples and these were taken forward for further analyses. This collection of 939 *Salmonella* isolates contained 499 isolates from human cases of disease and 440 isolates from potential animal reservoirs (referred to as source groups). Table 6 shows the numbers of isolates obtained from each source group, along with the numbers of serotypes and *Xba*I pulse types identified within each source group. An "*Xba*I pulse type" is defined as one or more isolates that share a unique PFGE banding pattern after digestion with the *Xba*I restriction enzyme.

Table 6. The numbers of 956 *Salmonella* isolates originating from each of six source groups and the numbers of different serotypes and unique PFGE banding types (*Xba*I pulse types) identified within each group. Note the total includes all sources so the values in the last row are not merely the sum of the other rows.

Source group	Number of isolates (%)	Number of serotypes		Number of pulse types	
		Total	Major[1]	Total	Major[1]
Human	499 (52%)	57	16	269	19
Cattle	198 (21%)	22	8	88	14
Sheep	135 (14%)	5	3	39	5
Pigs	10 (1%)	4	1	7	0
Poultry	76 (8%)	19	5	41	5
Wild birds - urban	21 (2%)	6	1	11	1
Total	939 (100%)	73	25	366	47

[1]A major serotype or pulse type is defined as one which is represented by 4 or more isolates within that source group.

4.4. Attributing cases of human salmonellosis using serotype data

4.4.1. Descriptive analysis

In total, 73 serotypes were identified within the collection of 939 isolates. The 12 most frequently occurring isolates are shown in Table 7. The most common serotype was *Salmonella* Typhimurium, which accounted for 44% of isolates. A further 31% of isolates were identified as one of four serotypes: *S.* Hindmarsh, *S.* Brandenburg, *S.* Enteriditis and *S.* Infantis, each individually accounting for 7-9% of the total collection. Sixty one of the serotypes identified each accounted for less than 1% of the total collection.

Table 7. The 12 most frequently identified serotypes within a collection of 939 *Salmonella enterica* isolates that were obtained from clinical cases of human disease (n=499) and potential animal reservoirs (n=423) between 1 May 2011 and 30 April 2012 in New Zealand.

Serotype	Number of isolates (% of total)	Number of major[1] *Xba*I pulse types	Number of phage types present within the major[1] pulse types
Typhimurium	420 (44%)	29	30
Hindmarsh	82 (9%)	4	-
Brandenburg	76 (8%)	2	-
Infantis	65 (7%)	6	-
Enteritidis	63 (7%)	2	5
Agona	20 (2%)	1	-
Saintpaul	16 (2%)	0	-
Emek	12 (1%)	0	-
Stanley	12 (1%)	0	-
Mbandaka	9 (1%)	1	-
Virchow	9 (1%)	0	-
Subsp. I ser. 4,5,12 : i : -	9 (1%)	0	-
Other (61 serotypes)	156 (16%)	0	-

[1]A major pulse type is defined as a unique *Xba*I PFGE banding pattern that was represented by 4 or more isolates within that serotype.

The serotypes present within a source group differed between the sources, as shown in Table 8 *Salmonella* Typhimurium was the most commonly identified serotype from three of the six source groups: humans (49% of isolates from this source), cattle (67%) and wild birds (57%). However, *S.* Typhimurium only accounted for 7% of the isolates sourced from sheep; within this source group *S.* Hindmarsh accounted for 59% of isolates and *S.* Brandenburg a further 31%. In contrast, only 1 isolate from a single, human, clinical case was identified as *S.* Hindmarsh, whilst 4% of human isolates were *S.* Brandenburg.

Within the collection of isolates of poultry origin, the most commonly occurring serotype was *S.* Infantis (28% of isolates). This serotype accounted for 6% of human isolates and 2% of cattle isolates.

Of the small number of isolates recovered from porcine samples, the four *S.* Derby isolates originated from the same farm; and, therefore, it is not possible to interpret the distribution of serotypes across this source group.

Table 8. The serotypes and numbers (N) of *Salmonella enterica* isolates derived from six source groups. Colour formatting has been used to highlight the distribution of isolates across the serotypes within each source group.

Human (N=499)		Cattle (N=198)		Sheep (N=135)		Pigs (N=10)		Poultry (N=76)		Wild Birds (N=21)	
Serotype	N	Serotype	N	Serotype	N	Serotype	N	Serotype	N	Serotype	N
Typhimurium	244	Typhimurium	133	Hindmarsh	80	Derby	4	Infantis	21	Typhimurium	12
Enteritidis	56	Brandenburg	12	Brandenburg	42	Agona	2	Typhimurium	19	Enteritidis	3
Infantis	30	Emek	12	Typhimurium	10	Brandenburg	2	Agona	10	Subsp. I ser. 4,12 : - : 1,2	3
Brandenburg	19	Lexington var. 15+	5	Rissen	2	Typhimurium	2	Mbandaka	4	Oranienburg	1
Saintpaul	13	Enteritidis	4	Subsp. I ser. 8 : - : - (non-motile)	1			Thompson	4	Mississippi	1
Stanley	12	Infantis	4					Anatum var.	2	Saintpaul	1
Subsp. I ser. 4,5,12 : i : -	8	Ruiru	4					Derby	2		
Virchow	8	Senftenberg	4					Lexington var. 15+	2		
Mississippi	7	Mbandaka	3					Senftenberg	2		
Agona	6	Agona	2					Others [n=10]	10		
Oranienburg	6	Kentucky	2								
Weltevreden	6	Lexington	2								
Newport	5	Saintpaul	2								
Oslo	5	Others [n=9]	9								
Others [n=43]	74										

KEY
% of isolates from that source
60-69%
50-59%
40-49%
30-39%
20-29%
10-19%
1-9%

4.4.2. Serotype richness

Using rarefaction curves to visualise the relative richness of serotypes (i.e. the number of different serotypes) within the source groups highlighted the far greater serotype richness of isolates obtained from human cases compared to those sourced from cattle and sheep (Figure 1A). The curve for the ovine isolates almost reached a plateau, which implied that even if more isolates from this species had been serotyped it was unlikely that there would be many further serotypes identified. In contrast, the curves for both the human and bovine isolates remained steep, implying that a greater sampling effort in both species was likely to identify even more unique serotypes.

The serotype richness of the poultry isolates appeared to fall between that of humans and cattle, although the 95% confidence intervals overlapped with both sources, and the steep trajectory of the curve implied that a greater number of isolates were required from this source in order to sample a meaningful range of the serotypes that were truly present within it.

A second rarefaction curve was plotted, in order to compare serotype richness for humans versus that of the combined animal source groups (cattle, sheep, pigs, poultry and wild birds) (Figure 1B). These two groups were much closer in terms of the numbers of isolates that were serotyped, but the lack of overlap between the 95% CI envelopes of the two groups implied that the serotype richness of human clinical isolates was indeed significantly greater than that of isolates from potential animal reservoirs.

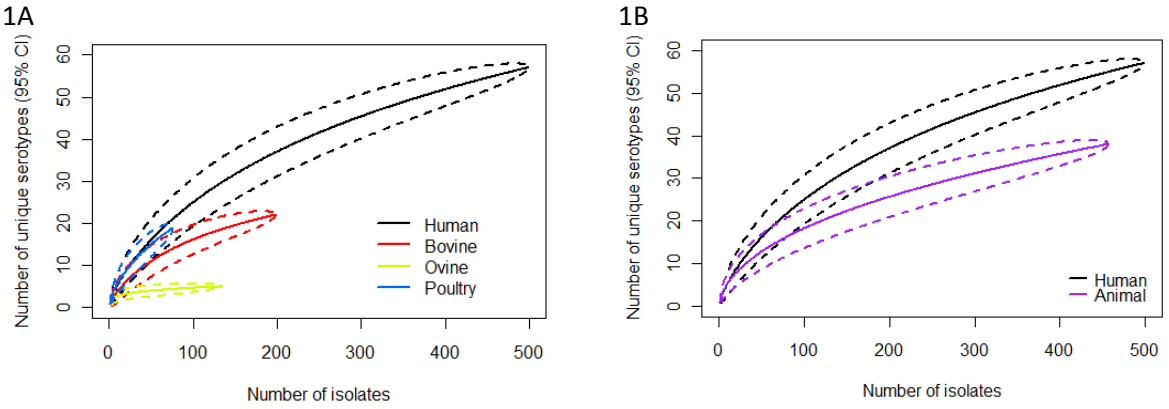

Figure 1. Rarefaction curves showing the serotype richness of libraries of *Salmonella enterica* isolates obtained from different source groups. 1A) Shows the relative serotype richness of isolates from human cases compared to each of the three potential reservoir species for which over 70 isolates were serotyped [cattle, sheep, poultry]. 1B) Shows the serotype richness of human isolates compared to the combined richness of all the isolates collected from non-human sources (cattle, sheep, pigs, poultry and wild birds).

4.4.3. Proportional similarity indices

Figure 2A is a diagonal similarity matrix showing the pairwise proportional similarities between the serotype frequency distributions of each of the six source groups. Comparing the human serotypes to those identified within the potential animal reservoirs showed that the greatest similarities were to cattle (0.62) and urban wild birds (0.66), and the least similarity was to sheep (0.12). Correspondingly, the cattle and wild bird serotypes were relatively similar to each other, whilst the sheep serotypes were distinct from every other group. The split network in Figure 2B also clearly shows this closer alignment of serotypes from humans, cattle and wild birds. The low number of isolates harvested from porcine samples impedes interpretation of the results shown here.

2A

	Human	Bovine	Ovine	Porcine	Poultry	Wild Bird	n
Human	1	0.55-0.66	0.07-0.17	0.04-0.51	0.27-0.47	0.50-0.67	499
Bovine	0.62	1	0.08-0.19	0.06-0.52	0.24-0.45	0.37-0.72	198
Ovine	0.12	0.14	1	0.06-0.41	0.04-0.14	0.03-0.12	135
Porcine	0.25	0.27	0.27	1	0.13-0.47	0-0.48	10
Poultry	0.38	0.36	0.09	0.37	1	0.16-0.36	76
Wild Bird	0.66	0.60	0.07	0.20	0.25	1	21

2B

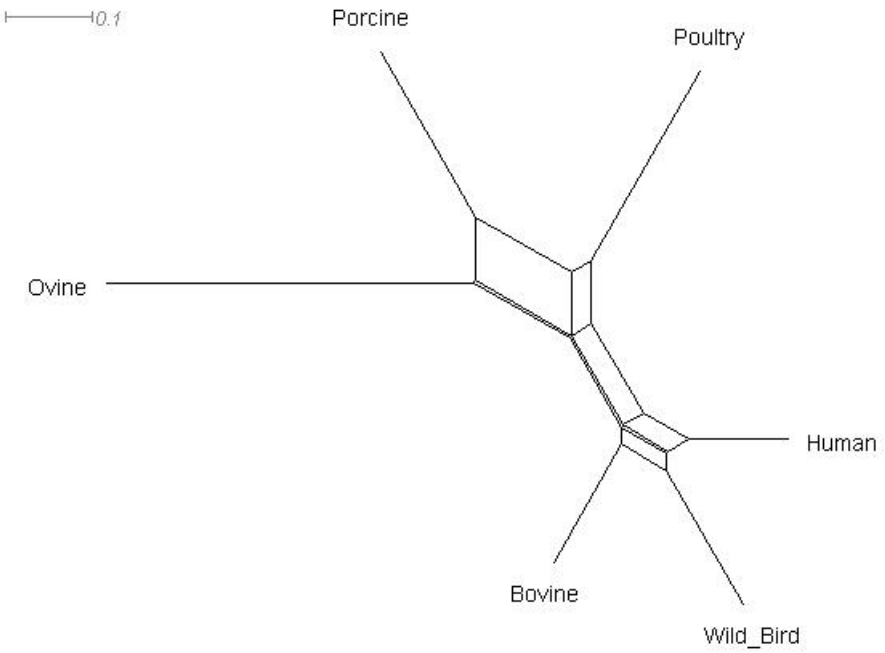

Figure 2. Examination of the similarities between the serotypes of *Salmonella enterica* originating from six source groups 2A) Diagonal matrix showing pairwise proportional similarity indices (PSI) for the serotypes of 939 *Salmonella enterica* isolates originating from 6 source groups (below diagonal), the corresponding 95% CIs (above diagonal), and the number of isolates that were serotyped within each source group (n). The PSI values have been shaded using a grey colour gradient that goes from light grey for those pairs showing the least similarity to dark grey for those showing the greatest similarity. 2B) Splits network of the PSI dissimilarity matrix showing the relative distances between the serotype frequency distributions of each of the source groups.

4.4.4. Modified Hald source attribution modelling using serotypes

Figure 3 shows the estimates of the probabilities that a human case of salmonellosis was attributable to each of the potential animal reservoirs based on the posterior distributions from a Hald model that was fitted to the serotype data. Using this serotype analysis, cattle represented the greatest probability of infection at a median estimate of 38% of human cases; the estimates from poultry and wild birds were 19% and 28% of cases, respectively; and the lowest risks of infection were associated with sheep and pigs each being responsible for an estimated 7-8% of cases each. The 95% credible intervals were extremely wide for each source group, showing that there was a high degree of uncertainty around these estimates.

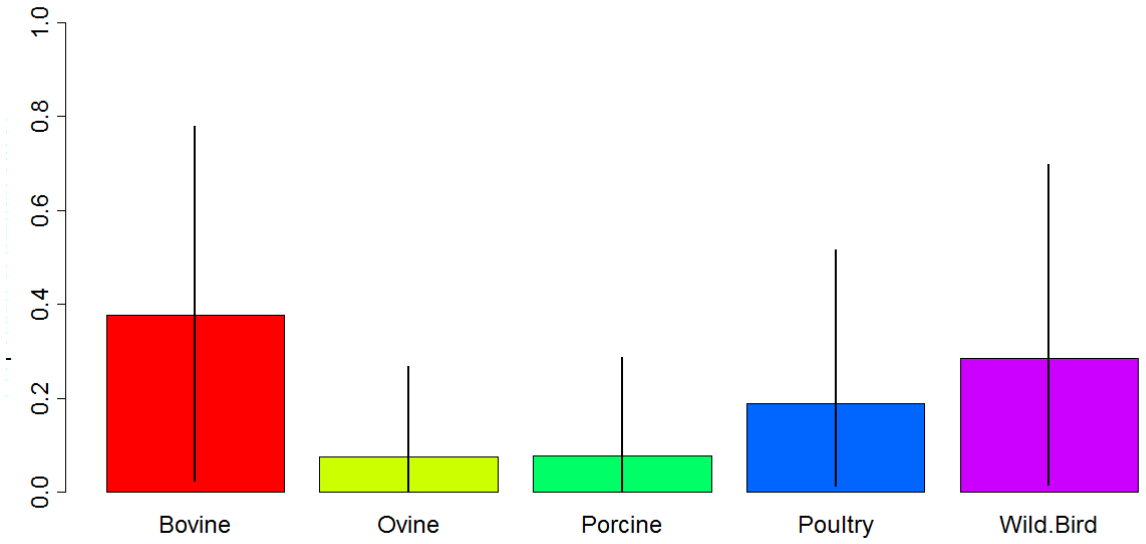

Figure 3. The estimated proportion of human cases of salmonellosis attributable to each of 5 potential animal reservoirs derived from a modified Hald source attribution model fitted to serotype data for 939 *Salmonella enterica* isolates from human clinical cases and potential animal reservoirs. The heights of each bar represent the median values of the posterior distributions from the model; the vertical lines represent the 95% credible intervals.

4.5. Attributing cases of human salmonellosis using pulse type data

4.5.1. Hierarchical clustering of PFGE banding patterns

Isolates from human cases

Using the Dice similarity coefficient to ascertain clusters of similar *Xba*1 PFGE banding patterns, 10 clusters consisting of 5 or more isolates were identified among the PFGE profiles of the human *S.* Typhimurium isolates. The 10 clusters comprised 115 of the 244 clinical *S.* Typhimurium isolates. Phage typing data from the human isolates indicated that the majority of these clusters were associated with one predominant phage type (Figure 4).

One major cluster consisting of 20 isolates was identified among the 56 human *S.* Enteritidis isolates. This cluster was also of a single phage type, phage type 11. Comparison of the total collection of 65 *S.* Enteritidis isolates from all sources identified two major clusters that both contained isolates from humans and one animal reservoir source (Figure 5). At the top of the figure the red-marked cluster shows isolates from four human cases and two faecal samples from gulls clustering together; the larger, green-marked cluster contains 25 human and three bovine isolates; the blue-marked cluster shows three human case isolates that did not cluster with NZ poultry isolates but were indistinguishable from a collection of overseas poultry *S.* Enteritidis (not shown).

One cluster was identified among the 30 *S.* Infantis human isolates. The cluster consisted of 8 isolates obtained from across New Zealand (Figure 6).

The 19 human *S.* Brandenberg isolates gave 1 major cluster of 13 isolates and six other distinguishable patterns.

Isolates from cattle

The 198 bovine isolates consisted of 22 serotypes, the most dominant being *S.* Typhimurium (133 isolates). Using the Dice coefficient, 10 clusters of five or more isolates were identified among the *S.* Typhimurium isolates. These clusters appeared to be a mix of two or more different phage types with one phage type predominating. A comparison of the human and bovine *S.* Typhimurium isolates showed that isolates with the same phage type had indistinguishable or very similar patterns. Interestingly no bovine phage type DT160 isolates were among the random selection of isolates that were typed using PFGE, indicating that this phage type was rarely isolated from bovine sources.

Isolates from sheep

The 80 *S.* Hindmarsh isolates obtained from sheep formed two major clusters; and the single human isolate of this serotype grouped with one of these clusters. Similarly, the majority of the 19 human *S.* Brandenburg isolates and five of the 12 bovine *S.* Brandenburg isolates were indistinguishable from the ovine *S.* Brandenburg (Figure 7). The 10 *S.* Typhimurium isolates gave eight different patterns all but one of which were seen among the human *S.* Typhimurium isolates.

Isolates from poultry and wild birds

Twenty-three serotypes were identified among the 109 poultry and avian isolates. One major cluster of 20 isolates was identified consisting of four from poultry, two from starlings, one from a wild duck and 13 from birds of unspecified species. Two serotype variants were identified in the cluster, *S.* Typhimurium and *S.* Subsp. I ser. 4,12 : - : 1,2. This cluster was indistinguishable from the large cluster of human *S.* Typhimurium RDNC-May06 isolates shown in Figure 4.

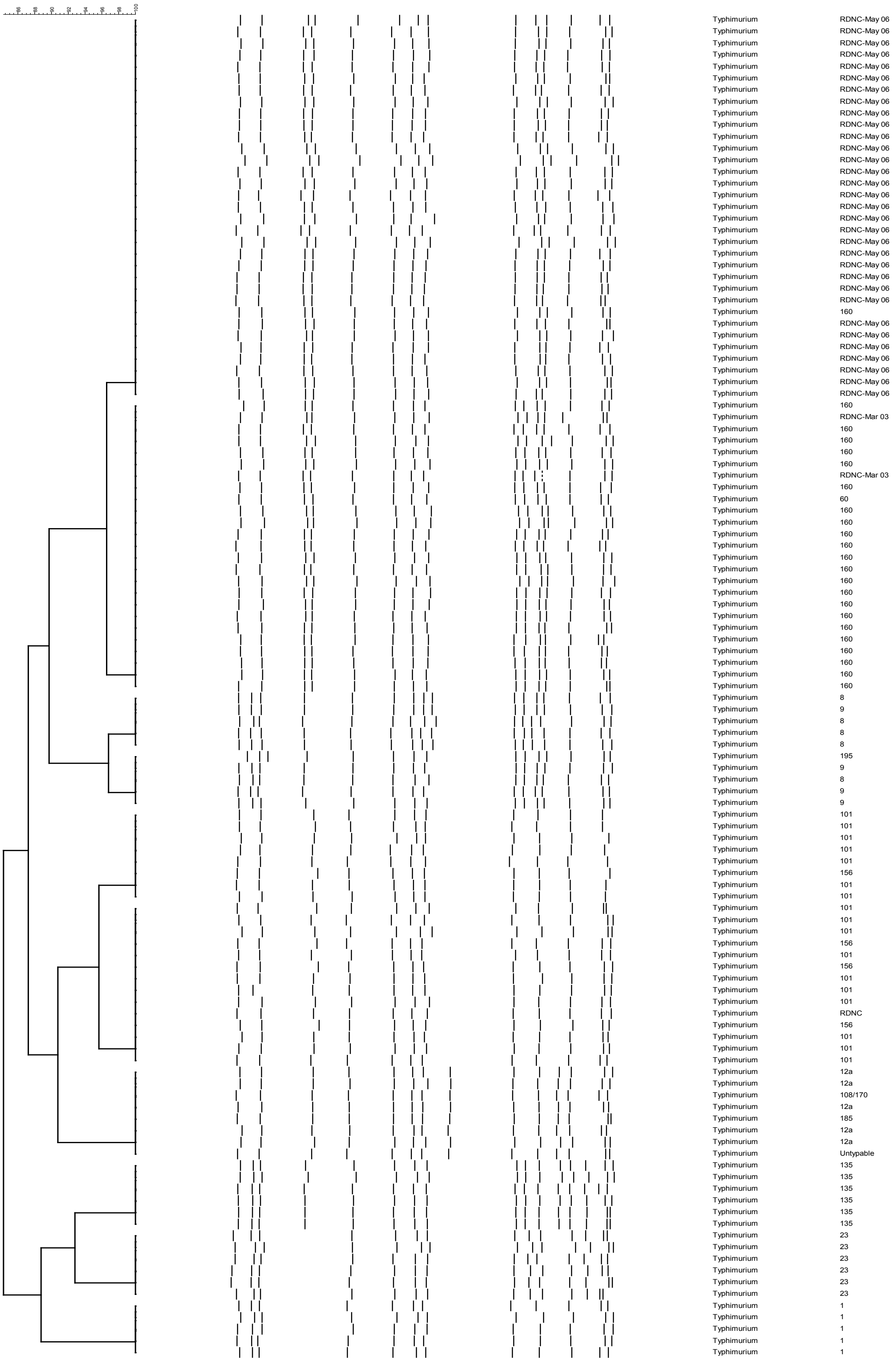

Figure 4. The 10 major clusters of PFGE banding patterns found within the *Salmonella* Typhimurium isolates that originated from human cases of disease. The phage types of the isolates within these 10 clusters are shown in the final column.

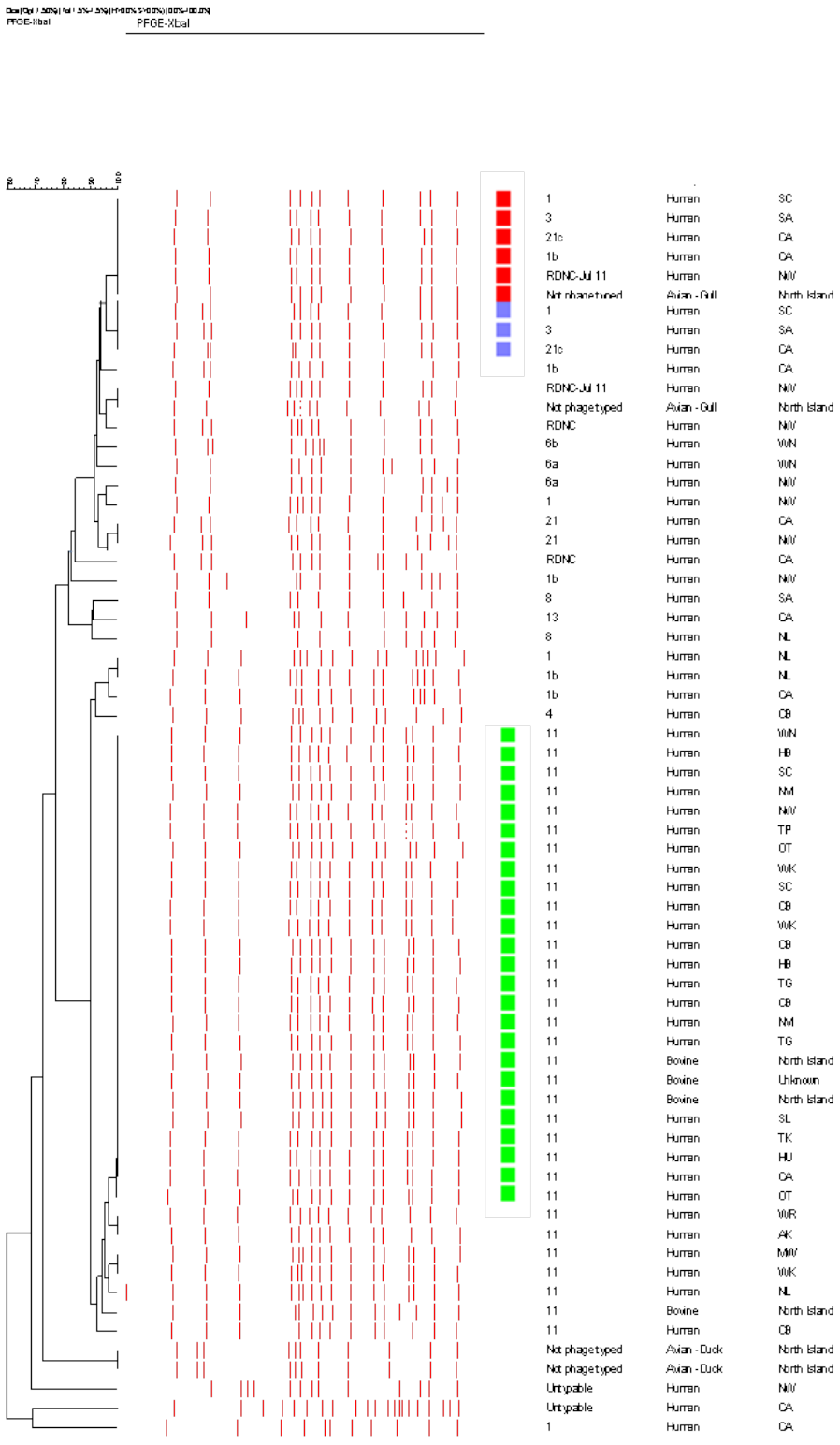

Figure 5. Cluster analysis of the PFGE banding patterns obtained from 65 *Salmonella* Enteriditis from human clinical cases (n=57) and potential animal reservoirs. The 3 major clusters are highlighted using coloured squares. The columns from left to right show the phage type, source of isolate, and area of country. Note the Avian-Duck samples were from wildfowl.

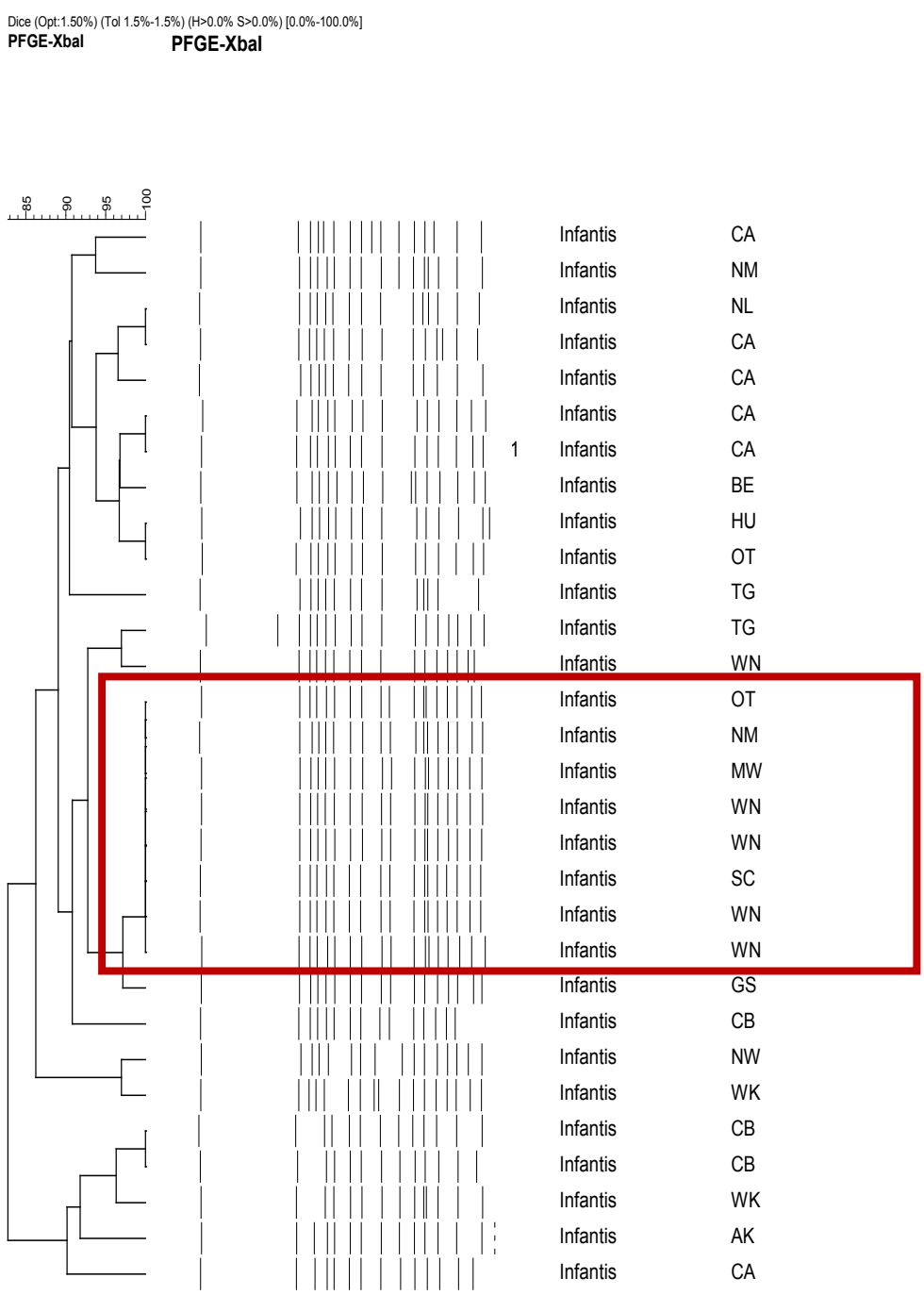

Figure 6. Cluster analysis of the PFGE banding patterns obtained from 30 *S.* Infantis isolates sourced from human clinical cases. The columns show from left to right: ERL id key, serotype and geographical location of case. The red box highlights the only major cluster identified within this dataset.

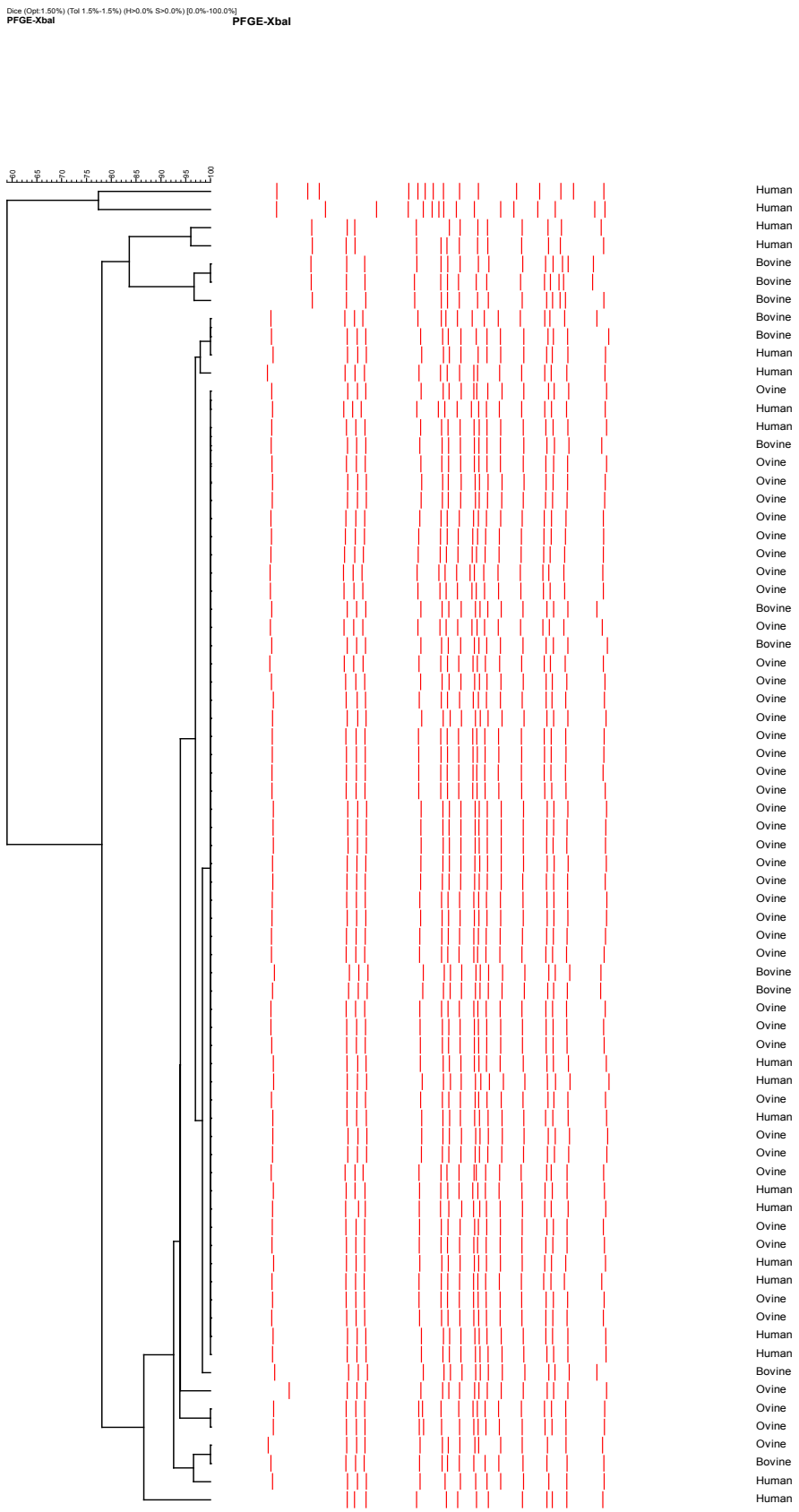

Figure 7. Cluster analysis of the PFGE banding patterns obtained from 70 *Salmonella* Brandenburg from human clinical cases (n=56), sheep (n=42) and cattle (n=12).

4.5.2. *Xba*1 PFGE pulse type richness

Using the number of band differences to construct a distance matrix and the complete linkage algorithm to draw the dendrograms, 367 unique *Xba*1 PFGE pulse types were identified amongst the PFGE banding data (Table 6). Forty nine of these pulse types were represented by four or more isolates and these were termed "major" pulse types.

Figure 8 shows the rarefaction curves of the pulse types from the four source groups that contained over 70 isolates. In line with the results of the serotype analysis, the human clinical isolates showed a significantly greater number of pulse types than the other sources. However, in contrast to the serotype curves, the pulse type curves are all of a steep trajectory – including the ovine isolates that do not appear to be reaching a plateau. This is demonstrating the greater discrimination of the PFGE typing scheme compared to serotyping.

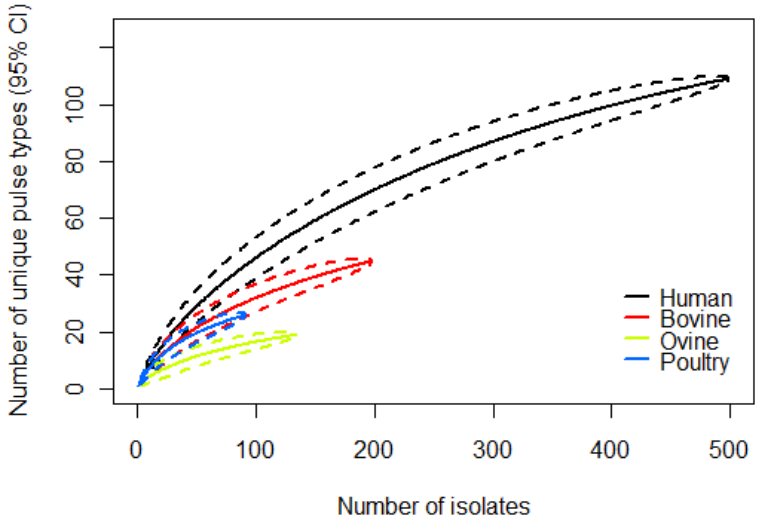

Figure 8. Rarefaction curves showing the *Xba*1 PFGE pulse type richness of libraries of *Salmonella enterica* isolates obtained from different source groups.

4.5.3. Proportional similarity of pulse type frequency distributions

Due to the increased discriminatory power, the pairwise proportional similarity indices (PSI) for the pulse type data (Figure 9A) are correspondingly lower than those seen for the serotype data (Figure 2). The most similar animal pulse types to the human clinical cases are those from cattle (PSI = 0.39); then wild bird and poultry (0.26 and 0.24, respectively); then sheep (0.18). At this level of discrimination the similarity between cattle and wild birds is also lower (PSI = 0.07).

The split network of the dissimilarity matrix is very star-like (Figure 9B), also indicating a high degree of discrimination between isolates from different source groups using this method.

All *Salmonella:* PFGE pulse types

9A

	Human	Bovine	Ovine	Porcine	Poultry	Wild Bird	n
Human	1	0.31-0.42	0.11-0.23	0-0.07	0.15-0.30	0.15-0.29	499
Bovine	0.39	1	0.11-0.23	0-0.19	0.15-0.27	0.01-0.18	198
Ovine	0.18	0.18	1	0-0.04	0.04-0.15	0.01-0.10	135
Porcine	0.06	0.12	0.02	1	0-0.11	0-0	10
Poultry	0.24	0.24	0.11	0.07	1	0.01-0.13	76
Wild Bird	0.26	0.07	0.06	0	0.07	1	21

9B

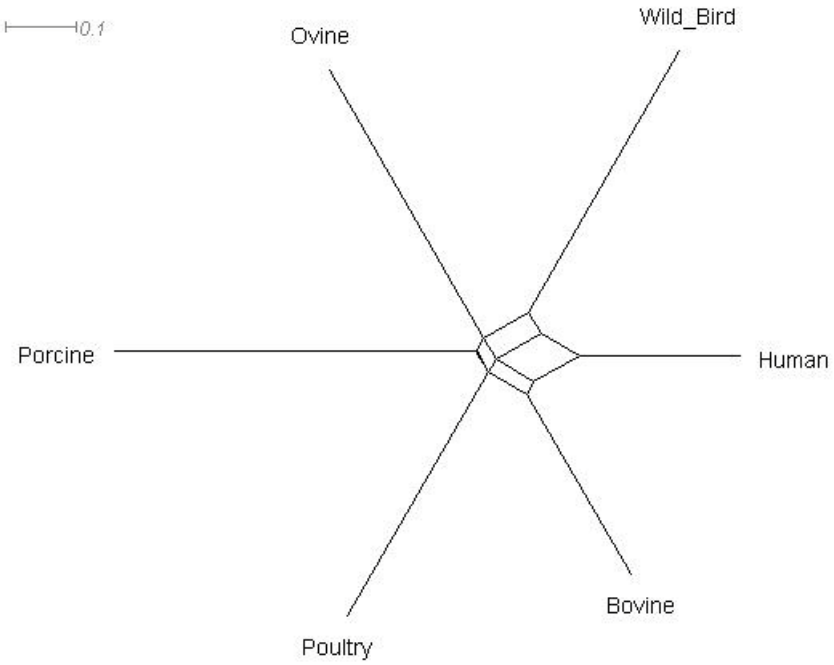

Figure 9. Examination of the similarities between the *Xba*1 PFGE pulse types of *Salmonella enterica* originating from 6 source groups. 2A) Diagonal matrix showing pairwise proportional similarity indices (PSI) for the pulse types of 939 *Salmonella enterica* isolates originating from 6 source groups (below diagonal), the corresponding 95% CIs (above diagonal), and the number of isolates that were serotyped within each source group (n). The PSI values have been shaded using a grey colour gradient that goes from light grey for those pairs showing the least similarity to dark grey for those showing the greatest similarity. 2B) Splits network of the PSI dissimilarity matrix showing the relative distances between the pulse type frequency distributions of each of the source groups.

4.5.4. Modified Hald source attribution modelling using PFGE pulse types

Figure 10 shows the estimates of the probabilities that a human case of salmonellosis was attributable to each of the potential animal reservoirs based on the posterior distributions from a Hald model that was fitted to the PFGE pulse type data. Compared to the serotype model, the pulse type model estimated that lower proportions of human cases were likely to originate from poultry and wild bird reservoirs with median values of 16% and 15% respectively. Conversely, the pulse type model estimated that a greater proportion of human cases were attributable to cattle (60%) compared to the serotype model (38%). The 95% credible intervals were generally narrower for this model compared to the serotype model, which is likely to be due to the greater discrimination giving rise to lower uncertainty.

All *Salmonella*: PFGE pulse types

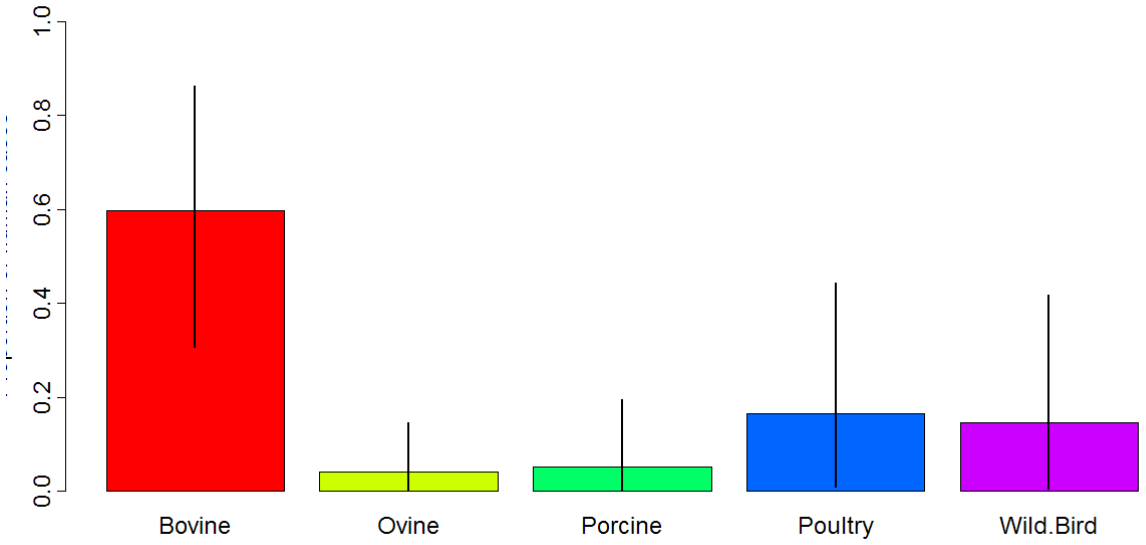

Figure 10. The estimated proportion of human cases of salmonellosis attributable to each of 5 potential animal reservoirs derived from a modified Hald source attribution model fitted to *Xba1* PFGE pulse type data for 939 *Salmonella enterica* isolates from human clinical cases and potential animal reservoirs. The heights of each bar represent the median values of the posterior distributions from the model; the vertical lines represent the 95% credible intervals.

4.6. Attributing cases of human salmonellosis caused by *S.* Typhimurium

Forty nine percent (n=244) of the human isolates were identified as *S.* Typhimurium, along with 67% (n=133) of the bovine isolates and 17% (n=43) of the isolates from other animal sources (Table 8). Due to the prominence of this serotype, two additional typing techniques were available and utilised: phage typing and multilocus variable-number tandem repeat analysis (MLVA). The sections that follow describe the in-depth analyses that were possible for this single serotype.

4.6.1. Phage typing of *S.* Typhimurium isolates

Descriptive analysis

Phage typing results were available for 418 *S.* Typhimurium isolates. Table 9 shows that 8 phage types accounted for 4% or more of this subset of *S.* Typhimurium isolates. The three most frequently occurring phage types were RDNC-May 06, DT 101 and DT 1. Table 9 also shows that both MLVA profiling and PFGE typing enabled the discrimination of genetically distinct strains within a phage type.

Table 9. The 13 most frequently identified phage types within a collection of 418 *Salmonella* Typhimurium isolates that were obtained from clinical cases of human disease (n=244) and potential animal reservoirs (n=174) between 1 May 2011 and 30 April 2012 in New Zealand.

Phage type	Number of isolates (%)	No. major[1] MLVA profiles within phage type (total[2])	No. major pulse[1] types within phage type (total[2])
RDNC-May 06	58 (14%)	3 (8)	4 (9)
101	54 (13%)	4 (15)	3 (12)
1	41 (10%)	2 (12)	5 (14)
160	33 (8%)	2 (17)	2 (6)
12a	30 (7%)	3 (8)	3 (12)
156	30 (7%)	4 (10)	2 (11)
8	20 (5%)	2 (9)	2 (10)
9	15 (4%)	0 (13)	1 (5)
23	14 (3%)	1 (4)	2 (5)
135	14 (3%)	1 (9)	1 (7)
42	14 (3%)	1 (6)	1 (8)
RDNC	11 (3%)	0 (10)	0 (9)
191	11 (3%)	1 (2)	2 (4)
Other (n=28)	73 (17%)	4 (55)	1 (64)

[1] A major MLVA profile or pulse type was one that was represented by 4 or more isolates within that phage type group. [2] The total number of MLVA profiles or pulse types present within that phage type group.

The phage types that were most frequently identified within each of the source groups are shown in Table 10. The RDNC-May 06 phage type was the most frequently encountered type within the human isolates. This phage type was also frequently present within the avian source groups, both poultry and wild bird. Conversely, within the cattle, sheep and pig isolates there were only two representatives of this phage type, and they both were contained within the cattle collection.

The most common poultry phage type, however, was DT 191, but only one human isolate of this phage type was identified. In contrast, the top four phage types within the cattle isolates were all also within the top six phage types from human cases. Only 10 *S.* Typhimurium isolates were recovered from ovine samples, and this set of 10 contained seven different phage types, all of which were also present within the human clinical case isolates.

Five of the wild bird isolates had not been phage typed, but three of these displayed identical MLVA profiles and clustered within the same unique pulse type groups as other RDNC-May 06 isolates from this source group; therefore, these 3 have been labelled as “presumptive” RDNC-May 06 within Table 10.

Phage type richness

The rarefaction curves for the two largest groups of isolates (human and bovine) are shown in Figure 11. This figure demonstrates that the human and bovine *S.* Typhimurium isolates are equally rich in phage types with complete overlap of the 95% confidence intervals.

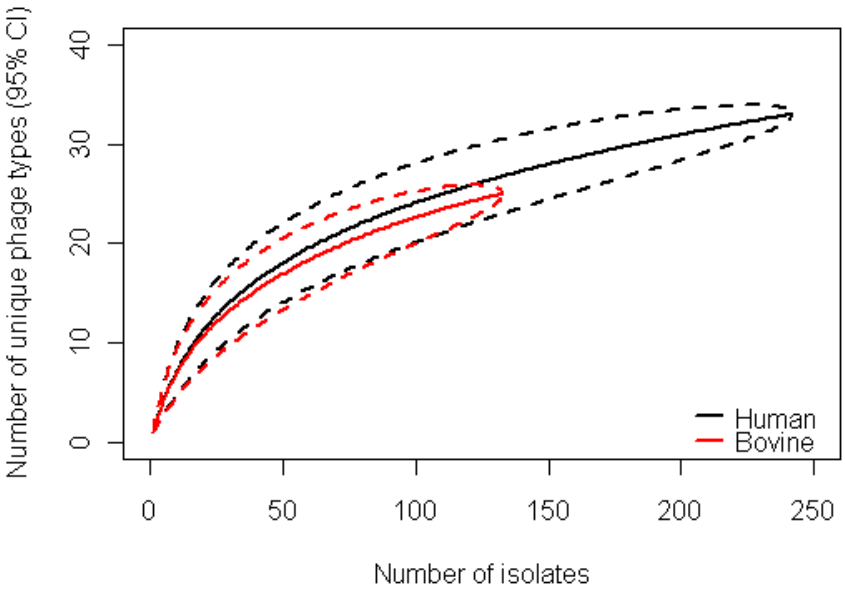

Figure 11. Rarefaction curves showing the relative richness of phage types within two libraries of *S.* Typhimurium isolates, one sourced from human clinical cases and the other from cattle.

Table 10. The phage types and numbers (N) of *Salmonella enterica* Typhimurium isolates derived from six source groups. Colour formatting has been used to highlight the distribution of isolates across the phage types within each source group. In the wild bird column, the 3 pRDNC-May 06 isolates were not phage typed, but had identical pulse types and MLVA profiles to other *S.* Typhimurium RDNC-May 06 from that source group; therefore, the "p" represents "presumptive".

Human (N=244)		Bovine (N=133)		Ovine (N=10)		Porcine (n=2)		Poultry (N=19)		Wild Bird (N=12)	
Phage type	N	Phage type	N	Phage type	N	Phage type	N	Phage type	N	Phage type	N
RDNC-May 06	42	101	26	9	3	195	1	191	10	RDNC-May 06	7
160	33	12a	16	8	2	12a	1	RDNC-May 06	4	pRDNC-May 06	3
101	27	1	16	185	1			199	1	Untyped	2
1	23	156	15	101	1			155	1		
156	15	8	11	74	1			135	1		
135	12	42	7	12a	1			89	1		
12a	12	23	6	1	1			1	1		
23	8	9	5								
42	7	108/170	4								
9	7	RDNC	4								
8	7	Other (15)	23								
RDNC	7										
Other (22)	44										

KEY
% of isolates from that source
50-59%
40-49%
30-39%
20-29%
10-19%
1-9%

Proportional similarity indices for phage types

The pairwise PSI matrix in Figure 12A shows a 56% similarity between *S.* Typhimurium phage types from human cases and those from cattle. There is also a strong similarity between bovine and ovine phage types (46%), with much lower similarity between ruminant and avian isolates. The split network plot (Figure 12B) clearly shows this ruminant versus avian split, with the human Typhimurium isolates closer to the ruminant grouping than the avian.

Salmonella Typhimurium: phage types

12A

	Human	Bovine	Ovine	Porcine	Poultry	Wild Bird	n
Human	1	0.45-0.59	0.08-0.34	0.01-0.10	0.11-0.34	0.13-0.22	244
Bovine	0.56	1	0.16-0.54	0-0.17	0.01-0.16	0-0.04	133
Ovine	0.32	0.46	1	0-0.3	0-0.11	0-0	10
Porcine	0.07	0.12	0.10	1	0-0	0-0	2
Poultry	0.29	0.08	0.05	0	1	0.05-0.42	19
Wild Bird	0.17	0.02	0	0	0.21	1	12

12B

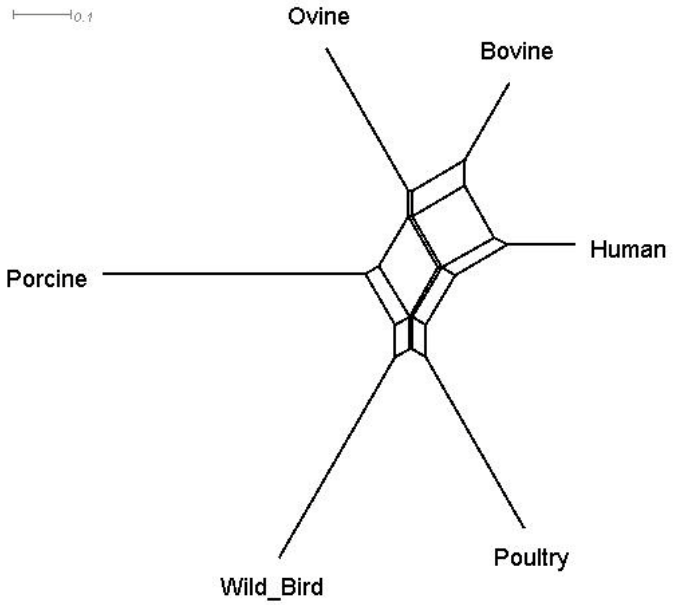

Figure 12. Examination of the similarities between the phage types of *Salmonella enterica* Typhimurium originating from six source groups. 12A) Diagonal matrix showing pairwise proportional similarity indices (PSI) for the phage types of 420 *Salmonella* Typhimurium isolates originating from 6 source groups (below diagonal), the corresponding 95% CIs (above diagonal), and the number of isolates that were phage typed within each source group (n). The PSI values have been shaded using a grey colour gradient that goes from light grey for those pairs showing the least similarity to dark grey for those showing the greatest similarity. 12B) Splits network of the PSI dissimilarity matrix showing the relative distances between the phage type frequency distributions of each of the source groups.

Modified Hald source attribution modelling of phage types

Fitting the Hald model to the *S.* Typhimurium phage type data also showed a strong association between bovine phage types and those from human cases (Figure 13). This model produced an estimate of 59% of human cases of salmonellosis due to *S.* Typhimurium could be attributable to bovine sources (95% Credible Interval = 29%-83%). The other sources were predicted to be of low importance to Typhimurium salmonellosis (<5% of cases), due in part to the low number of isolates recovered from those sources, with the exception of urban wild birds for which an estimate of 20% of human cases was returned (albeit with very a wide 95% Credible Interval of 1% - 44%).

Salmonella Typhimurium: phage types

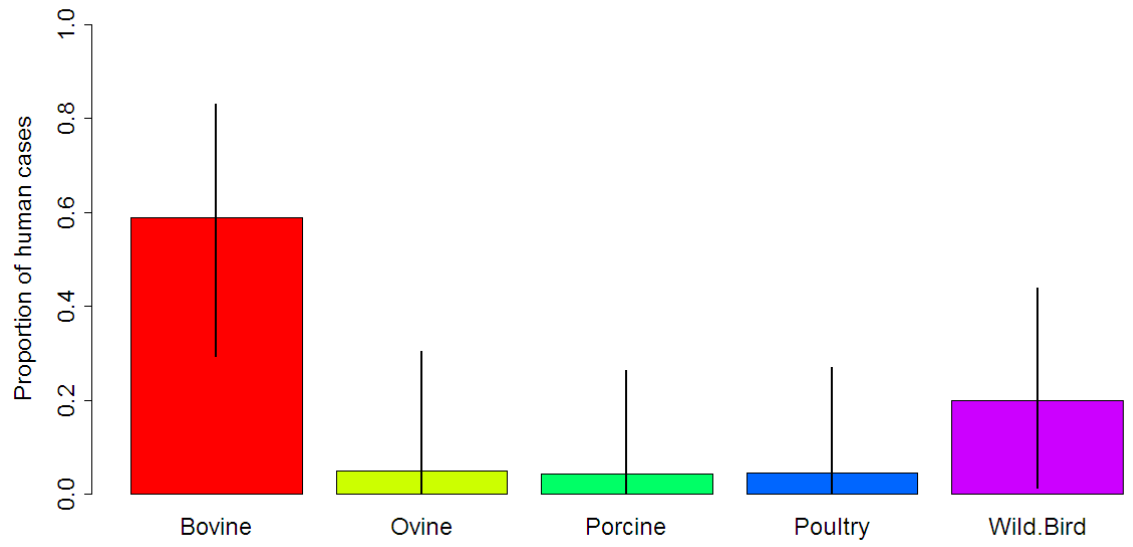

Figure 13. The estimated proportion of human cases of *Salmonella* Typhimurium attributable to each of 5 potential animal reservoirs derived from a modified Hald source attribution model fitted to phage type data for 420 *S.* Typhimurium isolates from human clinical cases and potential animal reservoirs. The heights of each bar represent the median values of the posterior distributions from the model; the vertical lines represent the 95% credible intervals.

4.6.2. MLVA typing of *Salmonella* Typhimurium isolates

All the isolates gave a multilocus variable-number tandem repeat (MLVA) profile, although a percentage of the isolates produced no amplicon at one or two of the loci. Repeat PCR of these isolates also failed to produce a product suggesting that these loci are truly absent in these isolates. Similar results have been observed in other MLVA studies (Lindstedt et al. 2004; Dyet et al 2010). A minimum spanning tree showed that some individual MLVA types clustered according to the phage type (Figure 14) but many of the MLVA patterns were associated with a range of different phage types.

MLVA profile richness

Compared to the phage type rarefaction plots (Figure 11), the MLVA typing showed greater discrimination between human and bovine *S.* Typhimurium isolates with an indication that the human isolates showed a wider range of MLVA profiles than the bovine isolates, although there is some overlap between the 95% CI envelopes of the two groups (Figure 15).

Proportional similarity indices for MLVA profiles

The pattern of pairwise PSI values was similar to that seen for the phage typing data of the same isolates, but the PSI values were lower for the MLVA profiles (Figure 16), which again indicates a greater degree of differentiation of this typing technique compared to phage typing.

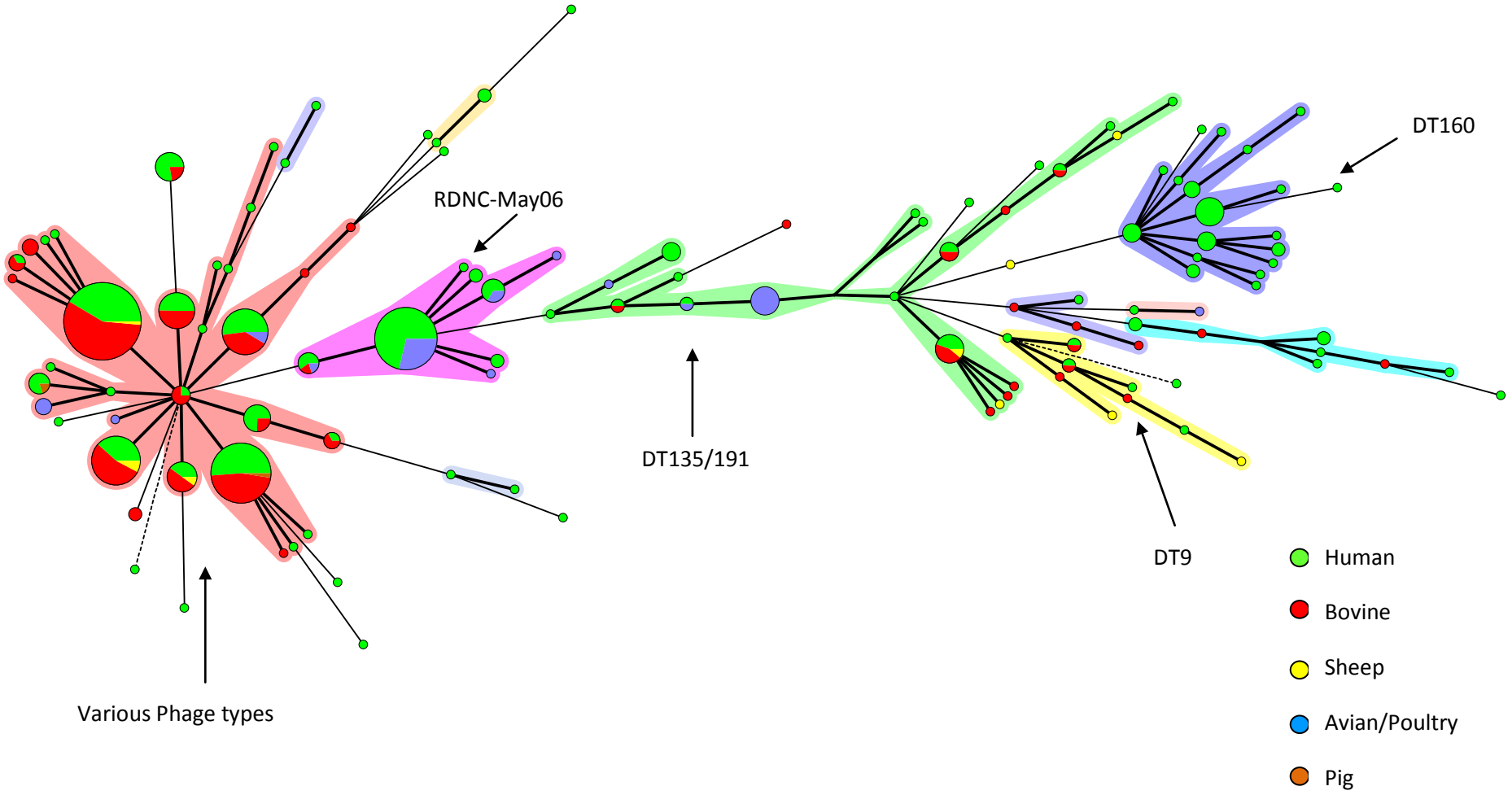

Figure 14. Using a minimum spanning tree to explore the relationships between the MLVA profiles and phage types obtained for *Salmonella* Typhimurium isolates originating from 240 cases of human disease together with 176 isolates derived from various animal reservoir species. Each circle represents a unique MLVA profile; the size of the circle indicates the relative abundance of that profile; the colours within the circles indicate the proportion of isolates with that profile that originated from the different source groups; the thickness of the lines between the circles indicates the number of loci differences between profiles (thick bands indicating single locus variants); the phage types of the isolates are represented by the colour blocks around clusters of MLVA profiles.

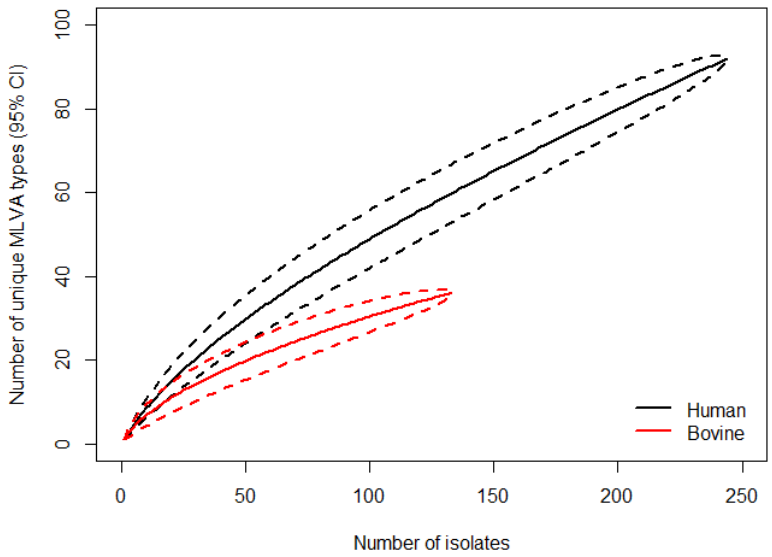

Figure 15. Rarefaction curves showing the relative richness of MLVA profiles within two libraries of *S.* Typhimurium isolates, one sourced from human clinical cases and the other from cattle.

Salmonella Typhimurium: MLVA

16A

	Human	Bovine	Ovine	Porcine	Poultry	Wild Bird	n
Human	1	0.33-0.46	0.03-0.21	0.01-0.13	0.06-0.22	0.09-0.18	244
Bovine	0.42	1	0.08-0.42	0-0.19	0-0.11	0-0	133
Ovine	0.17	0.27	1	0-0	0-0	0-0	10
Porcine	0.10	0.14	0	1	0-0	0-0	2
Poultry	0.19	0.08	0	0	1	0-0.32	19
Wild Bird	0.14	0	0	0	0.16	1	12

16B

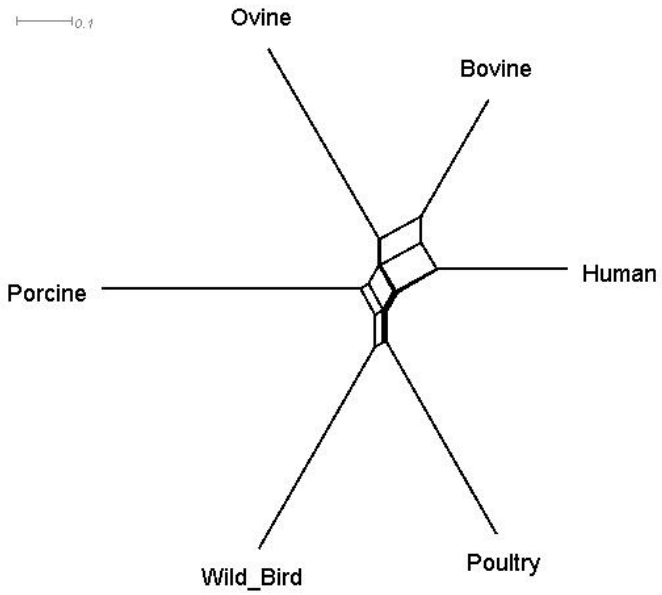

Figure 16. Examination of the similarities between the MLVA profile of *Salmonella enterica* Typhimurium originating from six source groups. 16A) Diagonal matrix showing pairwise proportional similarity indices (PSI) for the MLVA types of 420 *Salmonella* Typhimurium isolates originating from 6 source groups (below diagonal), the corresponding 95% CIs (above diagonal), and the number of isolates that were MLVA typed within each source group (n). The PSI values have been shaded using a grey colour gradient that goes from light grey for those pairs showing the least similarity to dark grey for those showing the greatest similarity. 16B) Splits network of the PSI dissimilarity matrix showing the relative distances between the MLVA profile frequency distributions of each of the source groups.

Modified Hald source attribution modelling of MLVA profiles

The estimates obtained from the Hald model that was fitted to the MLVA profiles of the .*S*. Typhimurium isolates (Figure 17) were in broad agreement with the results of the model fitted to the phage types of the same isolates (Figure 13). The estimated proportion of human cases of *S.* Typhimurium that were attributed to the cattle reservoir was 63%, which was slightly higher than the phage type estimate, and at 37%-85% the 95% Credible Interval for the MLVA model was slightly narrower than the phage model. The estimated contribution of wild birds was also similar between the two models, with the MLVA model giving an estimate of 18% (95% CrI 12%-40%).

Salmonella Typhimurium: MLVA profiles

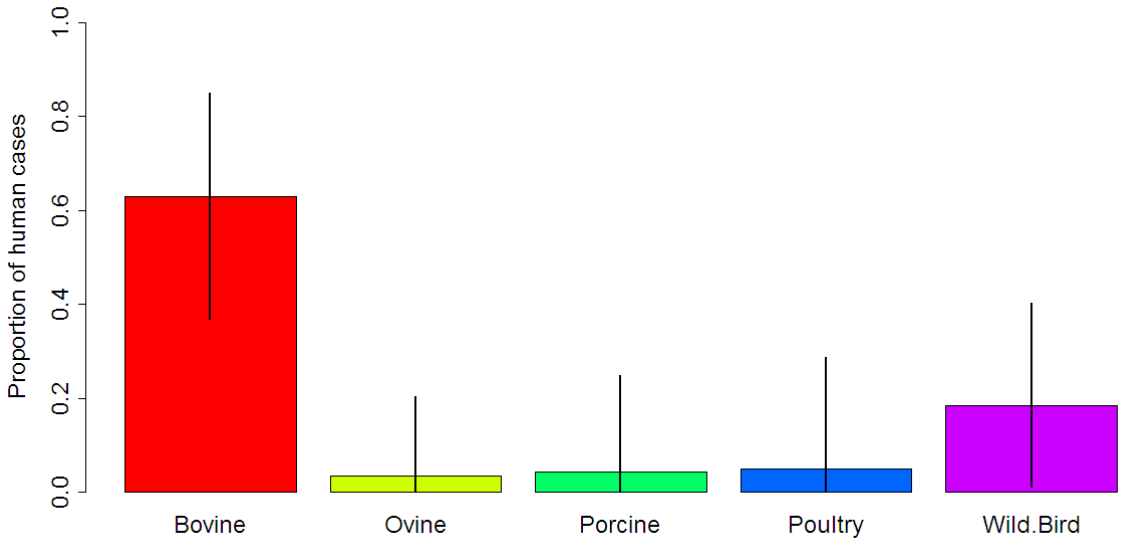

Figure 17. The estimated proportion of human cases of *Salmonella* Typhimurium attributable to each of 5 potential animal reservoirs derived from a modified Hald source attribution model fitted to the MLVA profiles of 420 *S.* Typhimurium isolates from human clinical cases and potential animal reservoirs. The heights of each bar represent the median values of the posterior distributions from the model; the vertical lines represent the 95% credible intervals.

4.7.PFGE typing results for the *S.* Typhimurium subset

4.7.1. Multidimensional scaling analysis of *Salmonella* Typhimurium pulse types

Multidimensional scaling plots were used to explore the *Salmonella* Typhimurium dataset. Specifically, the relationships between *Xba*1 pulse types and phage types were explored; as well as the relationships between pulse type and host species of origin.

Exploring pulse types and phage types

Figure 18 is a multidimensional scaling (MDS) plot of the PFGE pulse types of the total collection of *S.* Typhimurium isolates. The plot is a 2D representation of the multidimensional relative distances between the different PFGE banding patterns, such that isolates with similar banding patterns will cluster together and those with highly dissimilar banding patterns will be located in very different areas of the plot. The MDS plot in Figure 18 shows definite areas where pulse types are clumping together, as well as outlying single isolates that are more distinct from the rest of the isolates.

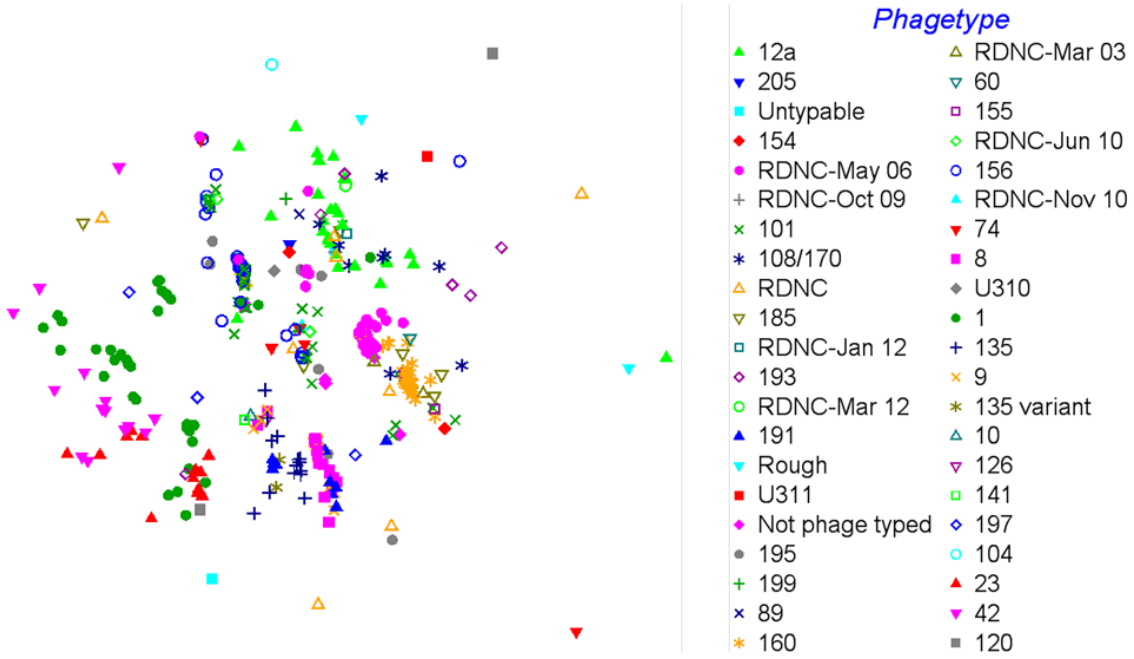

Figure 18. Multidimensional scaling plot showing the similarities of the PFGE pulse types of 420 *Salmonella* Typhimurium isolates. The individual isolates have been coloured according to their phage type, thus allowing for a visual assessment of the relationships between pulse types and phage types.

In Figure 18 and the following plots, the phage type of each isolate has been superimposed onto its pulse type position, thus allowing for a visual assessment of the relationships between pulse type and phage type. These plots show that some of the clusters of pulse types also represent clusters of phage types, indicating there is a degree of congruity between the two typing systems. However, there are some phage types that are present in multiple pulse type clusters and in some single isolates that are not grouping closely with other isolates. This indicates that some phages are able to infect multiple genetic sub-strains of *S.* Typhimurium. Plots 19 – 21 explore some of the pulse type clusters in greater detail.

Two densely-packed clusters occur very close together on the right-hand side of the plot. Figure 19 shows a magnified look at these two clusters, revealing that one is predominantly composed of isolates of phage type DT 160 (dark yellow stars) and the other is predominantly phage type RDNC-May 06. This implies that these phage types are closely genetically related.

Figure 20 shows a magnified plot of a very tight central cluster that aligns with two phage types: DT 156 (blue circles) and DT 101 (green crosses).

The diffuse cluster of pulse types on the left-hand side of the plot is magnified in Figure 21. Much of the left-hand side of the plot is composed of 3 commonly occurring phage types DT 1 (green circles), DT 23 (red triangles) and DT 42 (magenta triangles), with a lot of overlap between all 3.

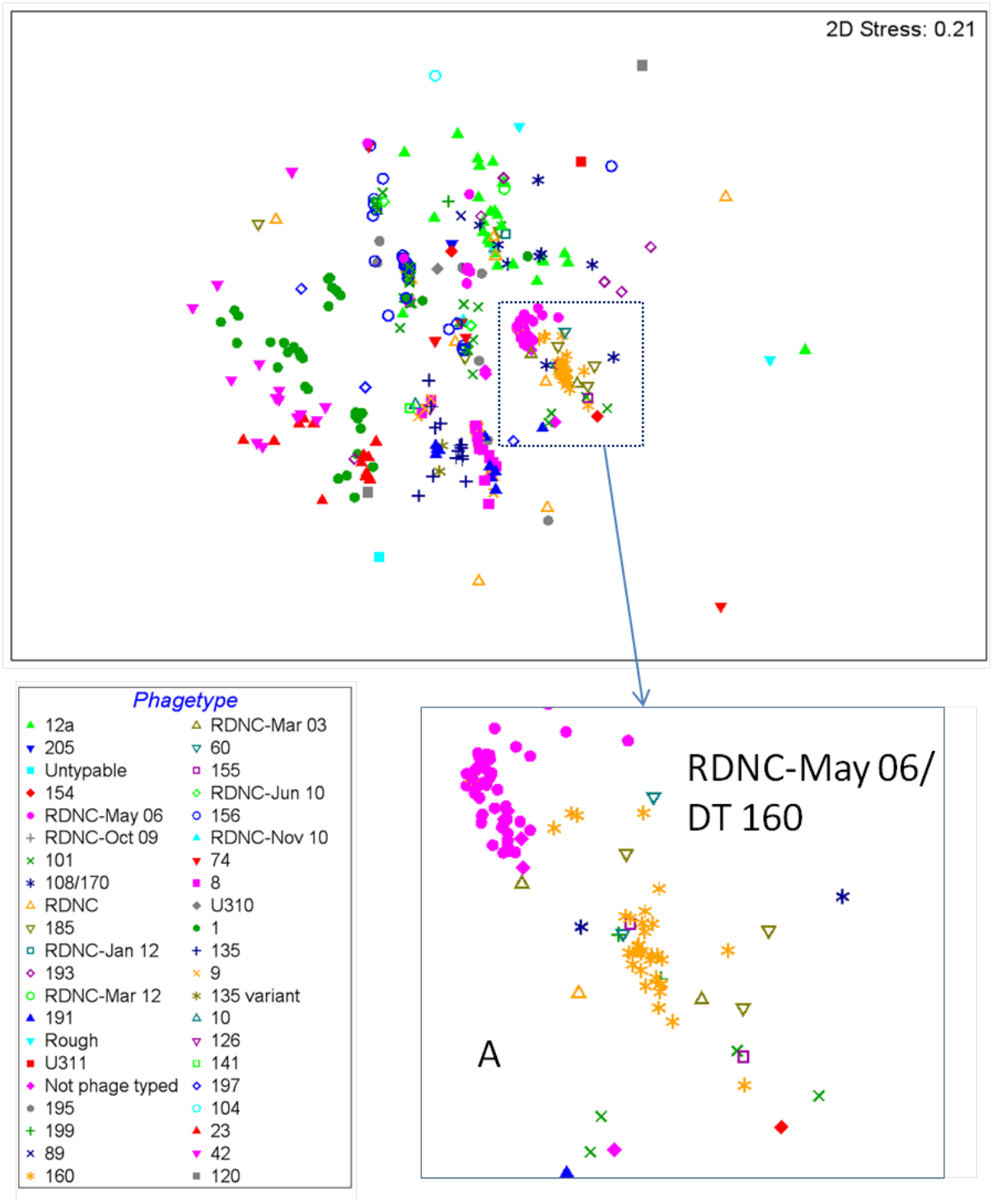

Figure 19. Multidimensional scaling plot showing the similarities of the PFGE pulse types of 420 *Salmonella* Typhimurium isolates (as shown in Figure 18) with a magnified look at two neighbouring clusters on the right-hand side of the plot. The individual isolates have been identified according to phage type.

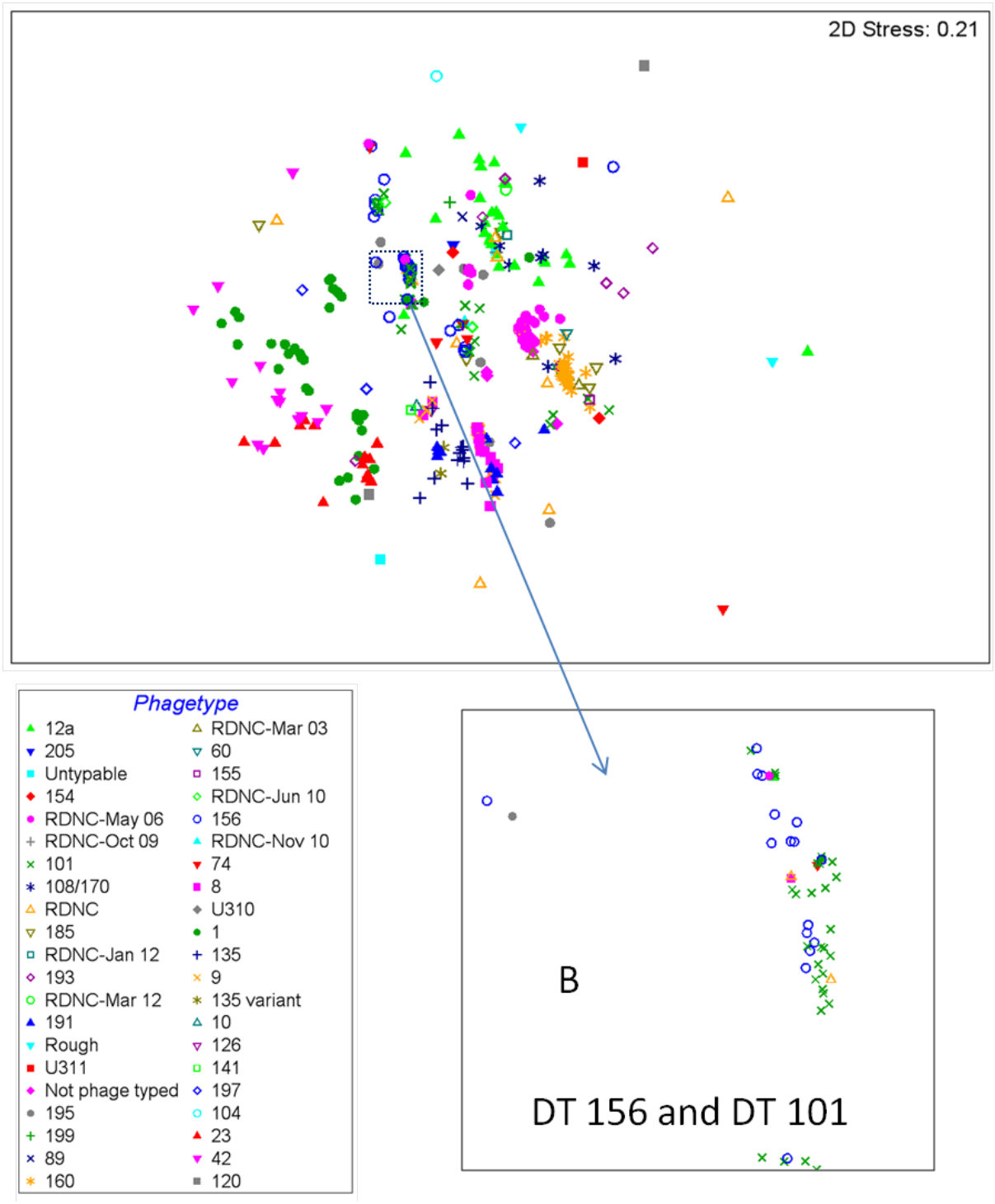

Figure 20. Multidimensional scaling plot showing the similarities of the PFGE pulse types of 420 *Salmonella* Typhimurium isolates (as shown in Figure 18) with a magnified look at a tight cluster in the centre of the plot. The individual isolates have been identified according to phage type.

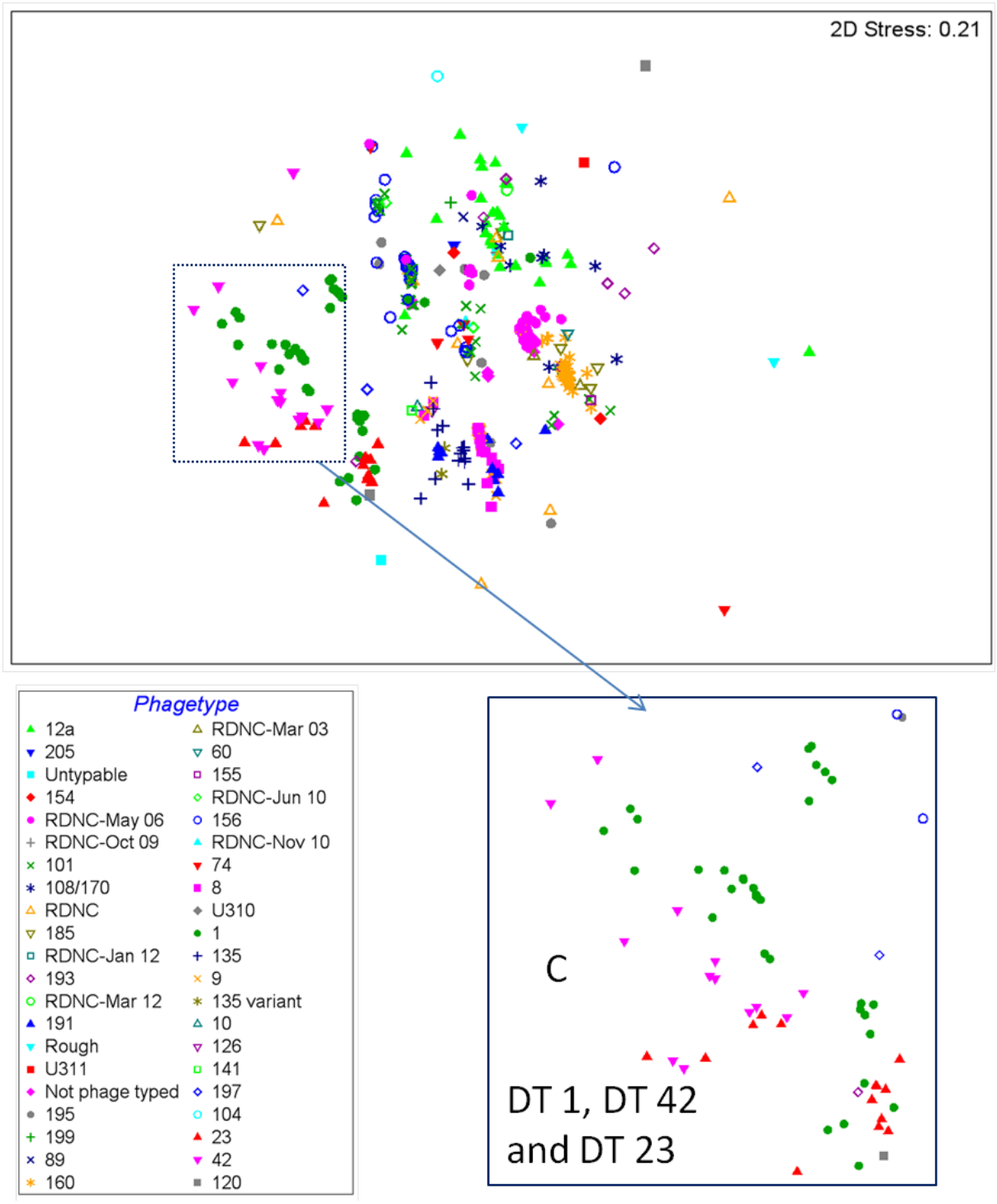

Figure 21. Multidimensional scaling plot showing the similarities of the PFGE pulse types of 420 *Salmonella* Typhimurium isolates (as shown in Figure 18) with a magnified look at a diffuse cluster on the left-hand side of the plot. The individual isolates have been identified according to phage type.

Exploring pulse types, phage types and source groups of isolates

If the individual isolates on the same MDS plot are coloured according to their source group (i.e. human clinical cases, bovine, ovine, porcine, poultry, wild bird) then preliminary assessments can be made of relationships between source groups, pulse types and phage types.

Figure 22 shows the close-up detail of Figure 19 but in this figure the isolates have been identified according to their source species of origin. Human isolates predominate in both of these clusters, but the cluster of RDNC-May 06 isolates also contains a number of isolates from wild birds (grey and green crosses, and blue stars) and poultry (red diamonds). In fact, of the 12 *S.* Typhimurium isolates from wild birds, 10 were identified as RDNC-May 06 and 2 were untypable by phage typing (1 of these had an identical MVLA profile and pulse type to other RDNC-May 06 isolates from wild birds; and the other was more genetically distinct).

In contrast, Figure 23 shows the two areas highlighted in Figures 20 and 21 that were related to phage types DT 156 and DT 101, and DT 1, DT 23 and DT 42. In this figure it is evident that the isolates present in both these clusters originated predominantly from human and bovine sources.

Figure 24 continues this exploration of associations between pulse type, phage type and host species of origin. Each of these six bubble plots is a multidimensional scaling plot of pulse types, but the individual plots only show the isolates obtained from a single source group. The number of bubbles in each subplot represents the number of different pulse types obtained from that source, whilst the sizes of the bubbles denote the relative abundance of each pulse type. Common phage types that have been found to be associated with particular pulse types have been highlighted using arrows. This plot very readily shows the abundance of pulse types within the set of human clinical isolates, but the major sources of some of those larger bubbles on the human plot can be clearly seen on the plots of the reservoir species. The association between RDNC-May 06 and DT 160 and avian species is clearly visible; whilst the predominance of DT 101 and DT 156 amongst the cattle isolates is notable with two pulse type bubbles associated with both of these phage types. Another collection of bubbles in the human plot is associated with phage types DT 8 and DT 9, which are clearly visible on both the cattle and sheep plots. These bubbles cluster closely with those containing by DT 191, which is notable within the poultry plot.

4.7.2. Attributing human cases of *S.* Typhimurium infections using pulse type data

S. Typhimurium *pulse type richness*

The rarefaction curves for human and bovine *S.* Typhimurium pulse types are shown in Figure 25. This plot suggests that there is a greater richness of pulse types within the human set of isolates, but there is much overlap of the 95% confidence envelopes between the two sources.

Proportional similarity indices for S. Typhimurium *pulse types*

Figure 26 demonstrates that PFGE is a useful tool for discriminating between genetic subsets of *S.* Typhimurium. The pairwise PSI matrix shows considerable overlap between the pulse type frequency distributions for the human and bovine isolates (PSI 0.53), but there is also considerable overlap with the poultry and wild bird isolates (both PSI 0.43). The cattle and sheep pulse types are also similar (PSI 0.53).

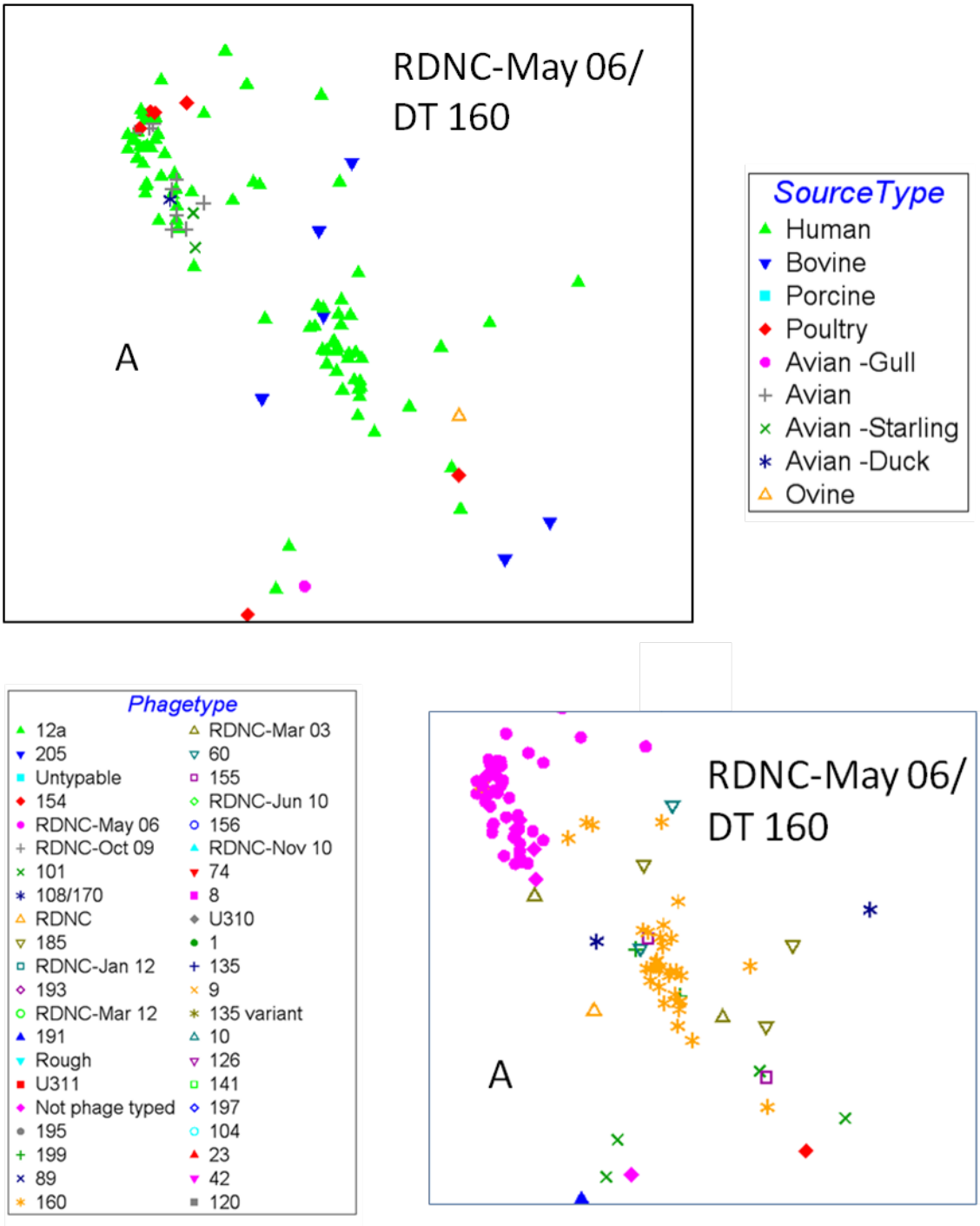

Figure 22. Two views of the same area of a multidimensional scaling plot showing the similarities of the PFGE pulse types of 420 *Salmonella* Typhimurium isolates (as shown in Figure 18). In the upper plot, the isolates are labelled according to their source species of origin. In the lower plot, the isolates are labelled according to their phage type.

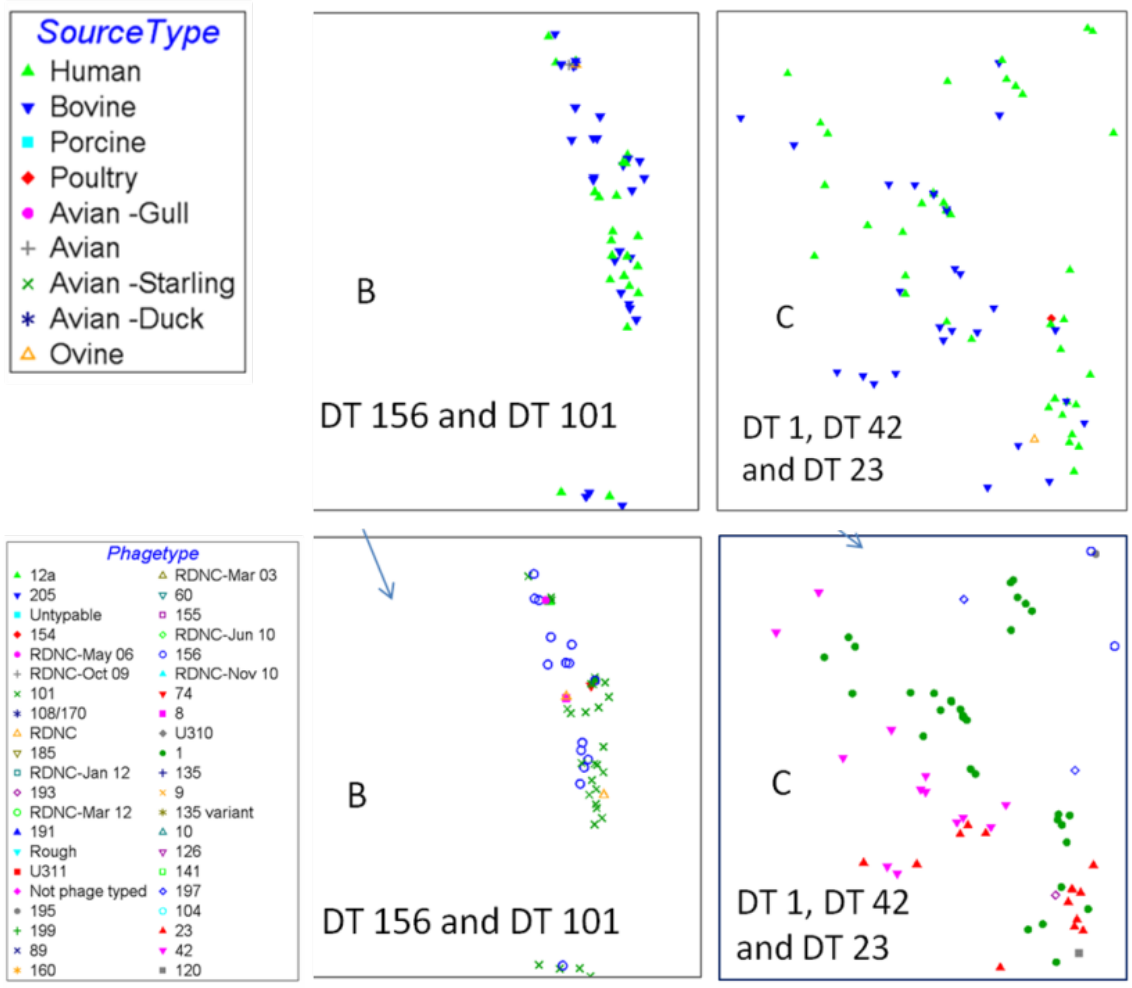

Figure 23. Two views of the same areas of a multidimensional scaling plot showing the similarities of the PFGE pulse types of 420 *Salmonella* Typhimurium isolates (as shown in Figure 18). In the upper plots, the isolates are labelled according to their source species of origin. In the lower plots, the isolates are labelled according to their phage type.

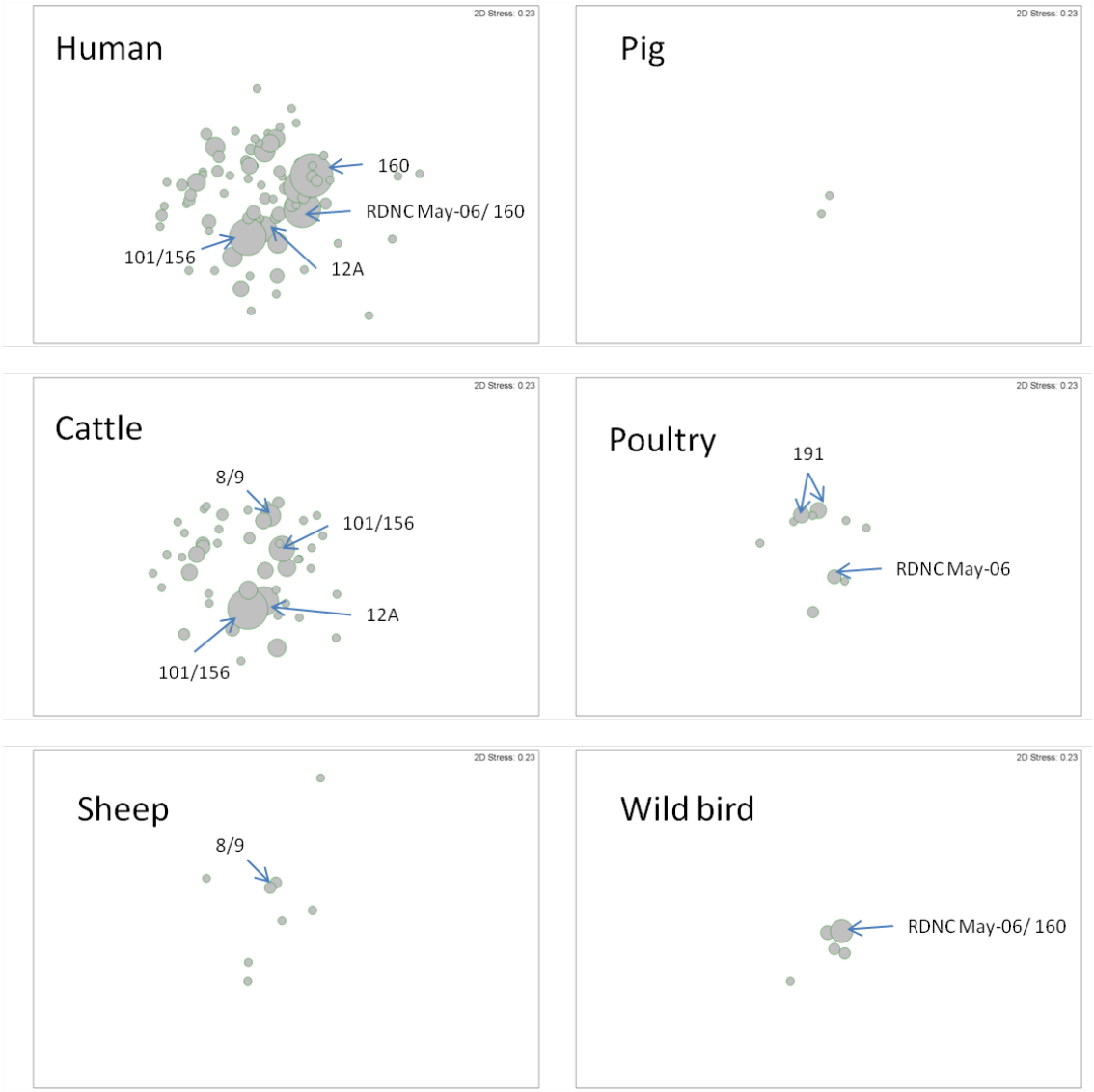

Figure 24. Multi-dimensional scaling bubble plots showing 2-D representations of the *XbaI* PFGE similarity matrix (based on the DICE coefficient). Each bubble represents a pulse type; and the location of each bubble is determined by the similarity matrix generated from the PFGE banding patterns. The bubble sizes are scaled according to the number of isolates from each host. Common pulse types are highlighted and labeled according to the predominant phage type associated with that banding pattern.

Proportional similarity indices for S. Typhimurium *pulse types (cont)*

Simultaneously comparing all pairwise PSI using a split network (Figure 26B), places the bovine and ovine isolates closely together on the left side of the plot, whilst poultry and wild bird are together at the top of the plot and the human isolates are in close apposition to those from wild birds.

Proportional similarity indices for non-Typhimurium serotypes

As a contrast, Figure 28 shows the PSI matrix and corresponding split network for all the serotypes present within the dataset except *S.* Typhimurium. This network is very star like (indicating over discrimination) and the PSI values are lower across the matrix, nonetheless there is a suggestion that human, bovine and poultry pulse types show some similarities with PSI values ranging from 0.19 to 0.25.

Modified Hald source attribution modelling of S. Typhimurium pulse types

In agreement with the PSI results in Figure 26, the Hald model of *S.* Typhimurium pulse types (Figure 27) estimates that the greatest burden of human disease with this serotype comes from cattle (33%, 7%-58%) and wild birds (43%, 6%-69%).

Modified Hald source attribution modelling of non-Typhimurium Salmonella *pulse types*

For this set of isolates, the Hald model estimated that cattle are the most important source of non-Typhimurium infections with an estimated 49% (95% CrI 20%-76%) of human cases attributable to this source (Figure 29). The proportions of non-Typhimurium cases attributable to wild birds and poultry were estimated to be 24% (2%-52%) and to poultry 14% (1%-38%).

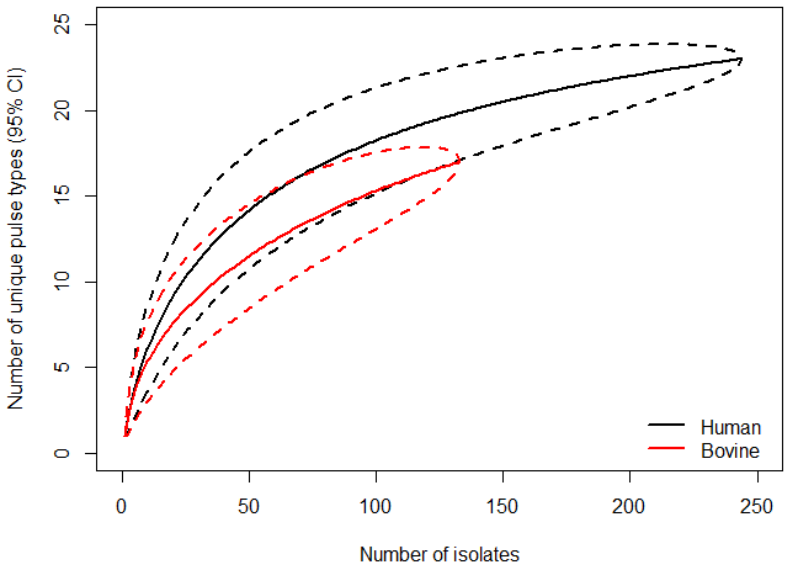

Figure 25. Rarefaction curves showing the relative richness of *Xba*1 PFGE pulse types within two libraries of *S.* Typhimurium isolates, one sourced from human clinical cases and the other from cattle.

Salmonella Typhimurium: pulse types

26A

	Human	Bovine	Ovine	Porcine	Poultry	Wild Bird	n
Human	1	0.43-0.58	0.16-0.36	0.01-0.13	0.23-0.54	0.30-0.53	244
Bovine	0.53	1	0.21-0.62	0.02-0.29	0.13-0.33	0.01-0.29	133
Ovine	0.32	0.53	1	0-0	0.10-0.60	0-0.25	10
Porcine	0.10	0.23	0	1	0-0.26	0-0	2
Poultry	0.43	0.25	0.45	0.11	1	0.10-0.47	19
Wild Bird	0.43	0.11	0.08	0	0.26	1	12

26B

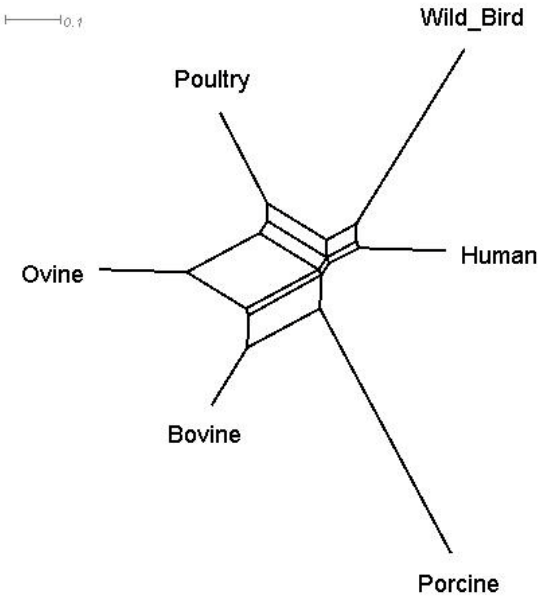

Figure 26. Examination of the similarities between the *Xba*1 pulse types of *Salmonella enterica* Typhimurium originating from six source groups. 26A) Diagonal matrix showing pairwise proportional similarity indices (PSI) for the pulse types of 420 *Salmonella* Typhimurium isolates originating from 6 source groups (below diagonal), the corresponding 95% CIs (above diagonal), and the number of isolates that were pulse typed within each source group (n). The PSI values have been shaded using a grey colour gradient that goes from light grey for those pairs showing the least similarity to dark grey for those showing the greatest similarity. 26B) Split network of the PSI dissimilarity matrix showing the relative distances between the pulse type frequency distributions of each of the source groups.

Salmonella Typhimurium: pulse types

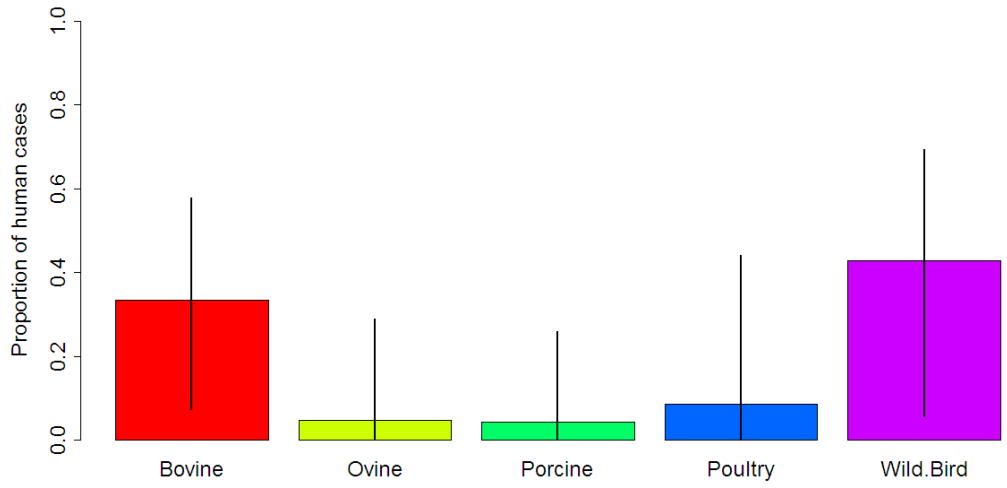

Figure 27. The estimated proportion of human cases of *Salmonella* Typhimurium attributable to each of 5 potential animal reservoirs derived from a modified Hald source attribution model fitted to the *Xba*1 pulse types of 420 *S.* Typhimurium isolates from human clinical cases and potential animal reservoirs. The heights of each bar represent the median values of the posterior distributions from the model; the vertical lines represent the 95% credible intervals.

All *Salmonella* serotypes except Typhimurium: pulse types

28A

	Human	Bovine	Ovine	Porcine	Poultry	Wild Bird	n
Human	1	0.15-0.30	0.05-0.12	0-0.03	0.10-0.22	0-0.10	255
Bovine	0.25	1	0.08-0.25	0-0	0.09-0.28	0-0.05	65
Ovine	0.09	0.15	1	0-0.02	0-0.07	0-0.06	125
Porcine	0.02	0.00	0.01	1	0-0.07	0-0	8
Poultry	0.19	0.22	0.03	0.04	1	0-0	57
Wild Bird	0.07	0.02	0.03	0.00	0.00	1	9

28B

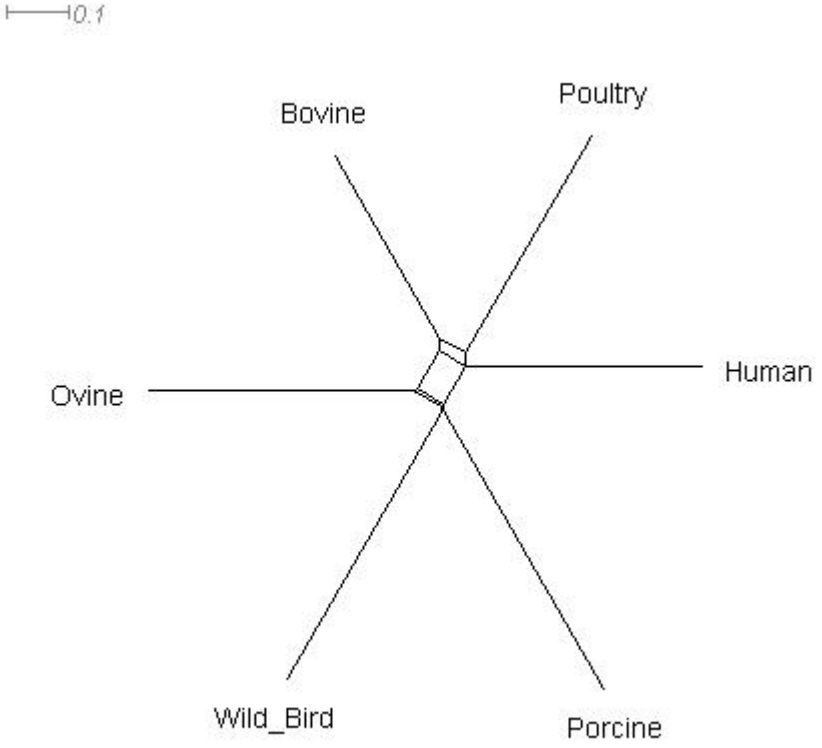

Figure 28. Examination of the similarities between the *Xba*1 pulse types of non-Typhimurium *Salmonella enterica* serotypes originating from six source groups. 27A) Diagonal matrix showing pairwise proportional similarity indices (PSI) for the pulse types of 519 *Salmonella* isolates originating from 6 source groups (below diagonal), the corresponding 95% CIs (above diagonal), and the number of isolates that were pulse typed within each source group (n). The PSI values have been shaded using a grey colour gradient that goes from light grey for those pairs showing the least similarity to dark grey for those showing the greatest similarity. 27B) Split network of the PSI dissimilarity matrix showing the relative distances between the pulse type frequency distributions of each of the source groups.

All *Salmonella* serotypes except Typhimurium: pulse types

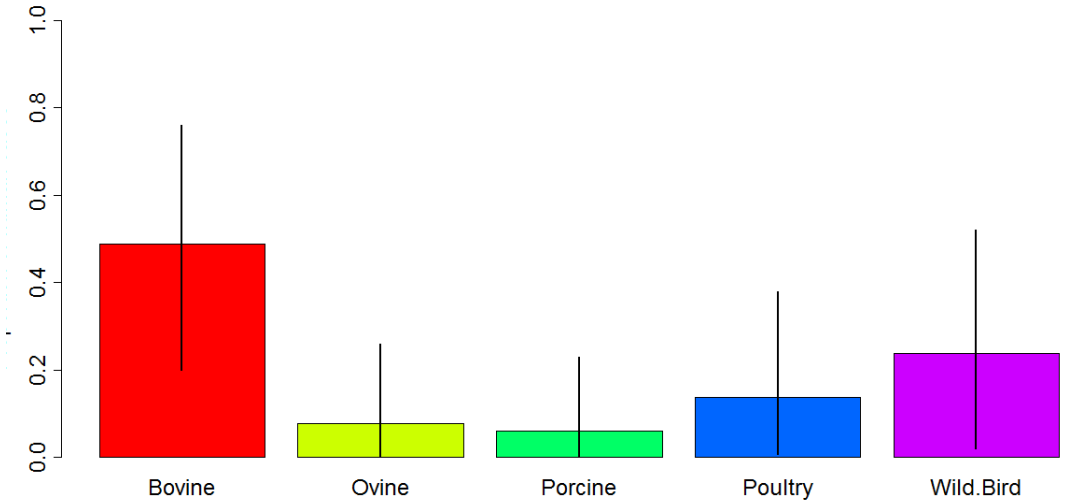

Figure 29. The estimated proportion of human cases with salmonellosis caused by serotypes other than *S.* Typhimurium attributable to each of 5 potential animal reservoirs derived from a modified Hald source attribution model fitted to the *Xba*1 pulse types of 536 *Salmonella* isolates from human clinical cases and potential animal reservoirs. The heights of each bar represent the median values of the posterior distributions from the model; the vertical lines represent the 95% credible intervals.

4.8. Summaries of Hald model estimates

4.8.1. Entire dataset

Two typing techniques were applied to all isolates: Serotyping and *Xba*I PFGE typing. The estimates of the proportion of human cases attributable to each potential animal reservoir that were derived from the modified Hald models fitted to both these complete datasets are shown in Table 11. The results of the two models were in rough agreement that the most common source of human disease was cattle, with a reasonable proportion of human cases also attributable to wild birds and poultry. The model fitted to the more discriminatory PFGE data increased the proportion of cases attributable to cattle and decreased the proportions attributable to avian sources, and these estimates were associated with narrower CrI.

Table 11. Summary of the estimated proportions of human cases of salmonellosis in New Zealand in 2011-2012 that were attributable to each of 5 potential animal reservoirs. Two kinds of typing data were fitted in separate modified Hald models: serotypes and PFGE banding data (pulse types).

Source group	Serotype		Pulse type	
	Estimate	95% Cr. I	Estimate	95% Cr. I
Bovine	38%	2-78%	60%	30-86%
Ovine	7%	0-27%	4%	0-14%
Porcine	8%	0-29%	5%	0-20%
Poultry	19%	1-52%	16%	1-44%
Wild Bird	28%	2-70%	15%	0-42%

4.8.2. Subset of *Salmonella* Typhimurium

Due to the availability of further typing techniques for the *Salmonella* Typhimurium isolates, three modified Hald models were fitted to the typing data for this substantial subset of the total isolates.

Table 12 shows the results of these 3 models. The estimates derived from phage typing data (a phenotypic test) and MLVA profile data (a genotypic test) were in close agreement with tighter credible intervals around the MLVA estimates. These two models estimate that 59-63% of the human cases of *S.* Typhimurium infections were attributable to cattle sources, with wild birds accounting for a further 18-20% of cases. The estimates derived from the PFGE pulse type data showed a different balance between those two sources, with 43% of human infections attributable to wild birds and 33% to cattle.

Table 12. Summary of the estimated proportions of human cases of salmonellosis caused by *S.* Typhimurium in New Zealand in 2011-2012 that were attributable to each of 5 potential animal reservoirs. Three kinds of typing data were fitted in separate modified Hald models: phage types, MLVA profiles and PFGE banding data (pulse types).

Source group	Phage type		MLVA profile		Pulse type	
	Estimate	95% Cr. I	Estimate	95% Cr. I	Estimate	95% Cr. I
Bovine	59%	29-83%	63%	37-85%	33%	7-58%
Ovine	5%	0-30%	3%	0-20%	5%	0-29%
Porcine	4%	0-26%	4%	0-25%	4%	0-26%
Poultry	4%	0-27%	5%	0-29%	9%	0-44%
Wild Bird	20%	1-44%	18%	1-40%	43%	6-69%

4.8.3. Subset of *Salmonella* other than Typhimurium

The final Hald model was fitted to the PFGE pulse types of the non-Typhimurium serotypes, and the results of this model are shown in Table 13. These data suggest that cattle are a prominent source of the non-Typhimurium infections, with estimates of 49% of cases attributable to cattle.

Table 13. Summary of the estimated proportions of human cases of salmonellosis (caused by serotypes other than *S.* Typhimurium) in New Zealand in 2011-2012 that were attributable to each of 5 potential animal reservoirs. The modified Hald model was fitted to PFGE banding data (pulse types).

Source group	Pulse type	
	Estimate	95% Cr. I
Bovine	49%	20-76%
Ovine	8%	0-26%
Porcine	6%	0-23%
Poultry	14%	1-38%
Wild Bird	24%	2-52%

5. Discussion

This is the largest molecular epidemiological study of non-typhoidal *Salmonella* that has been conducted in New Zealand. It was made possible by combining the *Salmonella* typing capabilities of the Enteric Reference Laboratory (ERL) at ESR, with the molecular modelling expertise of the Molecular Epidemiology and Public Health Laboratory (mEpiLab) at Massey University. The study is also a functional "One Health" study whereby *Salmonella* isolates from human cases were collected prospectively, and isolates from potential animal reservoirs were amassed during the same time period. This has resulted in the construction of a more representative library of isolates for source attribution modelling than has been available previously.

This work has confirmed that *Salmonella* Typhimurium is currently the most frequently occurring serotype in New Zealand, accounting for almost half of the isolates that were typed from human clinical cases and over two thirds of the isolates from cattle and wild birds. Hald models fitted to both the serotype and PFGE typing data of the total 939 isolates indicated that, between May 2011 and April 2012, cattle were a major source of domestically-acquired human salmonellosis in New Zealand and, in particular, the predominant source of *Salmonella* Typhimurium infections. Importantly this study is only concerned with New Zealand acquired disease, and does not attempt to identify reservoirs or pathways for the approximately 25% of cases that reported having been abroad during the incubation period. Further, due to the design of the sampling plan it was not possible to consider salmonellosis that may have been acquired from imported foods.

The observation that cattle were the most important reservoir is consistent with the increasing number of reports over recent years of outbreaks of salmonellosis in dairy herds (Marchant 2012). It has been hypothesised that the trend towards feeding increased quantities of concentrated feed to dairy cows may be playing a part in this emergence of clinical salmonellosis in adult animals (Teague 2011). Whatever the causal and contributing factors are behind the upsurge in salmonellosis within dairy herds in the country, this study strongly suggests that cattle are a primary source of salmonellosis for the human population in New Zealand. Therefore, measures to control salmonellosis in dairy herds would be expected to have measurable, beneficial effects for public health. This is of particular relevance to the current debate around the increasing consumption of unpasteurised milk by non-farming families because "raw" milk is an obvious pathway for the transmission of *Salmonella* from a clinically or sub-clinically infected dairy herd to a person. Other pathways of concern would be the contamination of retail meat with *Salmonella*, and also the direct contact route for farmers and veterinarians, who are in-contact with sick cattle and/or otherwise healthy animals that are shedding *Salmonella* in their faeces, and meat-plant workers. The young children of this group of occupationally exposed adults would also be at risk of contracting infection through contact with contaminated hands, clothing or equipment.

This study has not identified pigs as an important source of salmonellosis in New Zealand. This is in contrast to the findings of Mullner, Jones et al. (2009) who, on the basis of a modified Hald model fitted to serotype and phage type data, suggested that pork may be a significant source of human disease. Muellner et al. used a variety of datasets that had been collected for different purposes, and they report that the data from pork "were very sparse and in parts unrepresentative". They were also only able to look at a limited number of serotypes (7) and *S.* Typhimurium phagetypes (5) with all other serotypes and phage types being grouped together in "Other/Unknown" categories.

Unfortunately, there is limited data on *Salmonella* carriage from pigs. This is partly because pigs are not routinely included within the NMD. However, during 2008-2009 when pork carcasses were swabbed within the NMD programme, *Salmonella* isolations were rare. Furthermore, within this study, despite substantial sampling efforts being made at abattoir lairages, *Salmonella* was only detected in five pooled samples (1.6% of a total of 307 samples), with four of these samples being collected during two consecutive visits to a single abattoir and which were subsequently identified as originating from animals from the same farm. Compliance managers at the abattoirs also indicted that *Salmonella* was a rare finding within their internal quality control sampling regimes (personal communications). The number of isolates originating from pigs remains sparse, therefore it is hard to interpret the modelling results for this sector and whether the findings reflect a true low

prevalence of *Salmonella* in New Zealand raised pigs. Farm-based epidemiological studies of *Salmonella* could be a useful approach to investigate whether, in contrast to reports from other countries (McDowell, Porter et al. 2007; Fosse, Seegers et al. 2009), New Zealand pig farms are indeed relatively *Salmonella* free.

Although, no *Salmonella* was recovered from 326 pooled faecal samples collected from sheep on farms or at lairages during this study, 135 *Salmonella* isolates were submitted to ERL during the 12-month sampling period. This collection of ovine isolates showed a substantially different distribution of serotypes than the other sources with a predominance of *S.* Hindmarsh (59%) and *S.* Brandenburg (31%). Both of these serotypes can cause clinical disease in sheep, and both have been linked to outbreaks of disease in some flocks (Clark et al 1999; Baker, Thornley et al. 2007). In contrast, *S.* Hindmarsh and *S.* Brandenburg accounted for a very low percentage of the human case isolates in this study (0.2% and 4%, respectively). Therefore, the modelling work suggested that sheep were currently an uncommon source of human infection in New Zealand. The validity of this result relies upon the representativeness of the collection of isolates that arrived at ERL; however, given the extensive nature of sheep production in New Zealand and the fact that, despite intensive sampling of healthy animals on farms and at abattoirs, no positive pooled faecal samples were identified it is possible that this distribution of serotypes does reflect the true picture of *Salmonella* carriage by sheep in this country.

There is an interesting disparity between the pairwise PSI value for the *S.* Typhimurium phage types from humans and sheep (0.32), and the Hald model estimate of the proportion of human cases that may be attributable to sheep sources (0.04). This is likely to be due to the facts that a low number of *S.* Typhimurium were recovered from sheep (10), but all of these were identified as phage types that were also present within the human dataset. Therefore, the "overlap" in phage distributions could appear to be high resulting in a higher PSI value, but the low numbers of isolates from sheep suggests that sheep are not a common source of *S.* Typhimurium for people and this is taken into account within the Hald model. The Hald model of the MLVA profiles of the 10 sheep isolates returned a very similar estimate (5%) to the model fitted to the phage type data, despite the two typing systems targeting very different characteristics of the bacteria. This reproducibility of model results using different typing data increases the certainty that the model results are reliable.

The two avian reservoirs that were investigated, poultry and wild birds, were also highlighted as potentially substantial sources of infection for people. The models fitted to the entire serotype and PFGE pulse type datasets estimated that these two source groups could each be responsible for a 15-28% of human cases. However, when the data was split in two (*S.* Typhimurium isolates and isolates of non-Typhimurium *Salmonella* serotypes), a divergence in epidemiology between these two sources was highlighted. It was revealed that the *S.* Typhimurium isolates from wild birds, whilst low in numbers, were very similar (both phenotypically and genotypically) to those from human cases. This resulted in the Hald models for phage type data and MLVA profile data both returning a median estimate of about one fifth of human *S.* Typhimurium cases being attributable to wild bird sources.

Salmonella Enteriditis was the second most frequently occurring serotype within the collection of isolates from human cases where it accounted for disease in 11% of the case patients. Looking at the reservoir species, this serotype was rare amongst the collections of isolates from ruminants and pigs,

and absent from poultry, but contributed 14% of the wild bird isolates. Similarly, *S.* Infantis was the third most common serotype amongst human cases, and amongst the reservoir groups this serotype was almost exclusively isolated from poultry. However, it should be noted that very few poultry samples have been positive in National Microbiological Database (NMD) the national surveillance system of the microbial quality of ruminant and poultry carcasses at abattoirs, and retail surveys (summarised in Wilson and Baker 2009, Chrystal et al 2008, Wong et al 2009) and similar low prevalences have been observed on the surface of eggs, and zero prevalence in egg contents (Wilson, 2007).

One of the current limitations of this study was the uncertain origin of the poultry samples submitted to ERL. Only the samples identified as faeces, cloacal swabs, and environmental were included in the typing dataset, and those from poultry-feed were excluded. A major assumption was therefore that the environmental samples represented faecal contaminant from poultry in the shed environment. Further it was not known whether the samples were from sporadic sampling or incident investigations. This uncertainty of the precise source and epidemiology of the poultry isolates needs to be taken into consideration when interpreting the typing data and source attribution estimates.

During the supplementary sampling efforts, four of the eight positive pooled faecal samples that were collected from urban wild birds were collected around waste water plants and refuse stations. Two of the other four positive samples originated from recreational areas in Christchurch in October 2011; at that time there were signs along the waterways in the city warning people of potential sewage contamination due to earthquake related damage to the waste treatment infrastructure in the city. No further positive samples were identified during subsequent sampling trips to the city. These observations concur with previous publications that have noted the recovery of enteric pathogens such as *Salmonella* from birds present around waste water treatment plants or refuse sites (Fenlon 1981). From this standpoint, it could be hypothesised that wild birds are acting as vectors of infection rather than a primary source of infection. Thus they are encountering *Salmonella* of human origin in certain high-risk environments and can then take infection to other sites. If those other sites are kerbside café tables, favoured picnic spots in recreational areas, public barbeque facilities, or even (as has been reported) taking a late night bath in the green-lipped mussel outlet within supermarkets, then even the potentially transient carriage of enteric pathogens by free-living wild bird species could pose important public health risks. Other wild bird-to-person transmission pathways could include the contamination of drinking water collected from roof catchments, as well as indirect pathways such as the contamination of cattle feed leading to infection of cattle. This study focussed on urban wild birds, it would be interesting to conduct studies of wild birds in areas experiencing *Salmonella* outbreaks on dairy farms to investigate whether birds such as gulls, starlings and sparrows are an important part of the transmission cycle of *Salmonella* between farms and between farms and residents in the surrounding area.

Interestingly, all the *S.* Typhimurium isolates from wild birds that had been phage typed were identified as RDNC-May 06. Furthermore, several untyped isolates clustered within the same pulse type groups as these RDNC-May 06 isolates and showed identical MLVA profiles. The hierarchical cluster and multidimensional scaling analyses of the PFGE data show that RDNC-May 06 is highly similar to DT 160, the phage type that was associated with die-offs in sparrow and other passerine populations.

Fitting the source attribution models to the PFGE pulse type data had mixed results. Several of the PSI analyses of pulse type data produced split networks with a very star-like structure, indicating that the isolates within each source group were quite different to those within other source groups. This reflects the greater degree of discrimination of this typing system, which is known to be a very useful tool for outbreak investigations but to have some limitations for longer term epidemiological studies. Nevertheless, comparing the Hald models fitted to the serotype data and the PFGE banding data of the whole collection of isolates showed that both told a similar story, but the 95% credible intervals were narrower for the PFGE model estimates, implying that the greater discrimination of PFGE compared to serotype was lowering the uncertainty within the model. However, the ability to use phage typing and MLVA profiling for the *S.* Typhimurium dataset found that the results of Hald models fitted to these 2 types of data were very comparable, whereas the model fitted to the PFGE pulse type data lowered the estimate for the contribution of cattle and increased that of the contribution of wild birds. Thus the availability of alternative genotyping methods for serotypes in addition to Typhimurium, could result in modifications to the current Hald model estimates.

A notable feature of the collection of isolates from human cases is that there are a large number of serotypes causing human disease in New Zealand that we have not recovered from the animal reservoirs that were studied. This implies that there are important sources of disease that have not been considered here, and/or that the isolates collected from the reservoirs are not representative of the total number of strains present within those reservoirs. Other studies have recognised the importance of imported food products as potential sources of human salmonellosis and as potential sources of animal infection (summarised by Wilson and Baker 2009) and these were not considered in the present study. All of these possibilities could affect the reliability of the model estimates, which could therefore be improved by more extensive testing of a wider range of potential sources. Farm-based epidemiological studies could also be a useful tool to assess the representativeness of the collections of isolates from the livestock reservoirs. It is also important to consider potential biases that may result from the surveillance pyramid, and the fact that only a fraction of human salmonellosis cases are reported and isolates available for typing and consideration in attribution studies (recent estimates of the surveillance multiplier was between 6.7 and 40 in a study of European Union countries, Haagsma et al 2012).

6. Conclusions

A comprehensive dataset of non-typhoidal *Salmonella* isolates has been assembled from human clinical cases and potential animal reservoirs. The use of phenotypic and genotypic typing techniques combined with molecular epidemiological analyses and source attribution modelling has revealed several trends. Cattle are highlighted as the most important source of human infections, particularly infections with *S.* Typhimurium – the serotype causing the greatest proportion of human infections at the current time. Cattle may also be the most important source of infection with serotypes other than Typhimurium. The *Salmonella* collected from wild birds showed strong similarities to strains causing human disease, highlighting the potential for some species of wild birds to act as vectors for disease spread. *Salmonella* was rarely isolated from faecal samples collected from healthy pigs and sheep in abattoir lairages. Source attribution studies of *Salmonella* need to incorporate a variety of typing system as at this current time there is no single typing system available of appropriate discrimination for all strains of *Salmonella*.

7. References

Baker, M. G., C. N. Thornley, et al. (2007). "A recurring salmonellosis epidemic in New Zealand linked to contact with sheep." Epidemiology and Infection **135**(1): 76-83.

Carlson, J. C., A. B. Franklin, et al. (2011). "The role of starlings in the spread of *Salmonella* within concentrated animal feeding operations." Journal of Applied Ecology **48**(2): 479-486.

Chrystal, N. D., S. J. Hargraves, A. C. Boa, and C. J. Ironside. (2008). Counts of Campylobacter spp. and prevalence of Salmonella associated with New Zealand broiler carcasses. Journal of Food Protection. **71**:2526-2532.

Clark, G., Fenwick, S., Boxall, N., Swanney, S., & Nicol, C. (1999). Salmonella Brandenburg abortions in sheep, pathogenesis and pathology. In Proceedings of the 29th Seminar of the Society of Sheep and Beef Cattle Veterinarians of the New Zealand Veterinary Association 12-22.

De Sousa, E., A. J. Berchieri, et al. (2010). "Prevalence of *Salmonella* spp. Antibodies to toxoplasma gondii, and Newcastle disease virus in feral pigeons (Columba livia) in the city of Jaboticabal, Brazil." Journal of Zoo and Wildlife Medicine **41**(4): 603-607.

Fenlon, D. R. (1981). "Seagulls (Larus spp.) as vectors of salmonellae: an investigation into the range of serotypes and numbers of salmonellae in gull faeces." J Hyg (Lond) **86**(2): 195-202.

Fosse, J., H. Seegers, et al. (2009). "Prevalence and Risk Factors for Bacterial Food-Borne Zoonotic Hazards in Slaughter Pigs: A Review." Zoonoses and Public Health **56**(8): 429-454.

Haagsma, J. A., et al. "Community incidence of pathogen-specific gastroenteritis: reconstructing the surveillance pyramid for seven pathogens in seven European Union member states." Epidemiology and Infection 1.1 (2012): 1-15. doi:10.1017/S0950268812002166

Heck, K. L., G. van Belle, et al. (1975). "Explicit Calculation of the Rarefaction Diversity Measurement and the Determination of Sufficient Sample Size." Ecology **56**(6): 1459-1461.

Huson, D. H. and D. Bryant (2006). "Application of phylogenetic networks in evolutionary studies." Molecular Biology and Evolution **23**(2): 254-267.

Kitadai, N., N. Ninomiya, et al. (2010). "*Salmonella* isolated from the feces of migrating cranes at the Izumi Plain (2002-2008): Serotype, antibiotic sensitivity and PFGE type." Journal of Veterinary Medical Science **72**(7): 939-942.

Lindset

Marchant, R. (2012). "Salmonellosis in dairy herds." VetScript(March 2012): 15-16.

McDowell, S. W. J., R. Porter, et al. (2007). "Salmonella in slaughter pigs in Northern Ireland: Prevalence and use of statistical modelling to investigate sample and abattoir effects." International Journal of Food Microbiology **118**(2): 116-125.

Mullner, P., G. Jones, et al. (2009). "Source Attribution of Food-Borne Zoonoses in New Zealand: A Modified Hald Model." Risk Analysis **29**(7): 970-984.

Rosef, O., G. Kapperud, et al. (1985). "Serotyping of Campylobacter jejuni, Campylobacter coli, and Campylobacter laridis from domestic and wild animals." Applied and Environmental Microbiology **49**(6): 1507-1510.

Snoeyenbos, G. H., E. W. Morin, et al. (1967). "Naturally Occurring Salmonella in "Blackbirds" and Gulls." Avian Diseases **11**(4): 642-646.

Teague, W. (2011). *Salmlonella* Typhimurium outbreaks in dairy herds. Proceedings of the Dairy Cattle Veterinarians of the New Zealand Veterinary Association. VetLearn, NZVA, New Zealand.

Wilson, N., and M. Baker. 2009. A systematic review of the aetiology of salmonellosis in New Zealand.

Wilson, M. 2007. Survey of retail eggs for *Salmonella*. Client report (FW0779) prepared for the New Zealand Food Safety Authority by ESR.

Wong, T. L., L. Hollis, A. Cornelius, and J. Bennett**.** 2005. *Salmonella* in uncooked retail meats in New Zealand. Client report FW05104. Prepared for the New Zealand Food Safety Authority. Institute of Environmental Science and Research Limited, Christchurch, New Zealand.

Acknowledgements

We would like to thank the staff at ERL and mEpiLab for their valuable contributions to the microbiology and molecular biology contained within this report. In particular we would like to thank Sebastien Hudault for sterling work collecting samples from lairages and urban wild birds, Dr Anne Midwinter for laboratory isolation and characterisation from said samples, Sarah Moore for help with lab work and data management, Dr Yolanda Van Heezik, Dept Zoology, U of Otago, for helpful discussions around gaining samples from wild birds in urban habitats, Hugo Strydom and Muriel Dufour Enteric Reference Laboratory ESER and Beth Robson and Brent Gilpin from the Water Programme at ESR.