Review

Salmonella source attribution based on microbial subtyping

Lisa Barco a, Federica Barrucci a, John Elmer Dahl Olsen b, Antonia Ricci a,⁎

a OIE, National Reference Laboratory for Salmonella, Istituto Zooprofilattico Sperimentale delle Venezie, Viale dell’Università 10, 35020 Legnaro, Italy
b Department of Veterinary Disease Biology, Faculty of Health and Medical Sciences, University of Copenhagen, Stigboeljsen 4, DK-1870 Frederiksberg C., Denmark

ABSTRACT

Source attribution of cases of food-borne disease represents a valuable tool for identifying and prioritizing effective food-safety interventions. Microbial subtyping is one of the most common methods to infer potential sources of human food-borne infections. So far, Salmonella microbial subtyping source attribution models have been implemented by using serotyping and phage-typing data. Molecular-based methods may prove to be similarly valuable in the future, as already demonstrated for other food-borne pathogens like Campylobacter. This review assesses the state of the art concerning Salmonella source attribution through microbial subtyping approach. It summarizes the available microbial subtyping attribution models and discusses the use of conventional pheno- typic typing methods, as well as of the most commonly applied molecular typing methods in the European Union (EU) laboratories in the context of their potential applicability for Salmonella source attribution studies.

© 2013 Elsevier B.V. All rights reserved.

Contents

1. Introduction ........................................................................................................... 193
2. Salmonella source attribution models based on microbial subtyping ...................... 194
3. Criteria for inclusion of subtyping methods to be used for source attribution .............. 195
4. Phenotypic typing methods for Salmonella subtyping source attribution ................. 195
   4.1. Serotyping ........................................................................................................ 195
   4.2. Phage-typing .................................................................................................. 196
   4.3. Antimicrobial resistance typing ........................................................................ 196
5. Genotypic methods for Salmonella subtyping source attribution ............................. 197
   5.1. Pulsed field gel electrophoresis (PFGE) .......................................................... 197
   5.2. Multiple-locus variable-number tandem repeat analysis (MLVA) ..................... 198
   5.3. Ribotyping ...................................................................................................... 199
   5.4. Plasmid profile analysis .................................................................................... 199
   5.5. Multi-locus sequence typing (MLST) ............................................................... 199
6. Future perspectives in molecular subtyping methods ............................................. 199
7. Conclusions ......................................................................................................... 200
References .................................................................................................................. 200

1. Introduction

Bacteria subtyping is defined as the characterization of bacteria beyond the species and/or subspecies level. Strains of food-borne pathogens responsible for human infections are expected to have indistinguishable or more similar subtypes to those of strains isolated from the originating source than those isolated from unrelated sources (Hyytiä-Trees et al., 2007).

Microbial subtyping is one of the major methodologies to attribute food-borne infectious diseases to their sources (EFSA, 2008). The principle behind the source attribution is the comparison of subtypes of isolates causing human disease with the distribution of these subtypes in their putative sources (e.g. animals, food, environment) (Pires et al., 2009). Microbial subtyping source attribution relies on subtyping laboratory methods to identify overlaps between subtypes
identified from cases of human disease and those from their potential sources.

Human infections caused by subtypes which have been exclusively or almost exclusively isolated from one single source can be attributed to that specific source. On the other hand, when human infections are caused by subtypes isolated in several sources, they can be attributed to those sources proportionally to the reported occurrence of subtypes in the sources (Hald et al., 2004).

The application of this methodology to Salmonella requires i) a collection of isolates from humans and from different sources; non-human isolates should be as representative as possible of the real exposure of humans to the potential vehicles of the pathogen; ii) appropriate subtyping methods that can provide detailed knowledge of Salmonella types in all relevant sources, and the correlation of isolates obtained from humans with those from the possible sources, and iii) a model to infer the sources of human infections based on the subtype distributions obtained within each investigated source.

The application of microbial subtyping approach for source attribution assumes that the distribution of subtypes in the collection of isolates in each source used in the attribution exercise is representative of the true distribution of subtypes in each source (Pires et al., 2009). Hence, it is essential to use a representative sample of isolates from each source investigated obtained through integrated surveillance programs including the major sources (e.g. food animals and food) and humans (EFSA, 2008). For this purpose, a stratified sampling scheme could be used to allocate the adequate sampling fraction in each of the strata considered (sources).

These collections of isolates should be subtyped by using appropriate and harmonized discriminatory typing methods and the profiles obtained should be evaluated through shared interpretative guidelines. Biases related to the lack of standardization of the methods used to subtype the isolates or to establish their clonality should be avoided since it is another essential requirement to guarantee the soundness of the results provided by the models.

Finally, it is equally important to have an in-depth understanding of the epidemiological data that are used as input to the model (EFSA, 2008).

This review aims to assess the state of the art concerning Salmonella source attribution through microbial subtyping approaches. The major attribution models used so far are summarized. Moreover, the usefulness of the most commonly applied typing methods in the European (EU) laboratories are discussed in the context of their potential application in Salmonella source attribution studies.

2. Salmonella source attribution models based on microbial subtyping

Several applications of the microbial subtyping approach for Salmonella source attribution have been described. In 1999, Van Pelt and colleagues developed the so called “Dutch model”, based on the principle of comparing the number of reported human cases, caused by a particular bacterial subtype (Salmonella serovar), with the relative occurrence of that subtype in each source. The number of reported cases per subtype i attributed by the model to source j was estimated by the equation:

\[ \frac{p_i}{\sum_j p_{ij}} \times \chi_i \]

where \( \chi_i \) was the expected number of cases per year of subtype i from source j estimated from \( p_i \), the proportion of Salmonella subtype i in source j among all the sources sampled, \( \sum_j p_{ij} \) the proportion of Salmonella subtype i among all the sources sampled, and \( \chi_i \) the estimated number of human cases of subtype i per year (Van Pelt et al., 1999).

The Dutch model had the advantage of being easy to calculate, but it relied on the false assumptions that all Salmonella subtypes have the same virulence and all sources an equal probability of causing human disease. Conversely, it is widely recognized that different Salmonella subtypes have different probabilities of causing disease in humans and with different levels of severity (EFSA, 2012). Moreover, food types have different abilities to act as vehicles for food-borne disease agents, due to their physical properties, such as pH and water activity, or due to the applied processing and preparation procedures.

In 2004, Hald and colleagues developed a model that, similar to the Dutch model, compared the number of human cases caused by different Salmonella subtypes with their prevalence in different food sources, but this model also incorporated bacterial and food-source dependent factors, using a Bayesian approach. In the Hald model, the number of human cases per subtype i was assumed to follow a Poisson distribution with expected number of cases \( \lambda_i \) (unknown parameter) given by the equations:

\[
\lambda_i = \lambda_i + \epsilon_i - \alpha_i \\
\lambda_j = \sum_i \lambda_{ij} \\
\lambda_y = M_j \cdot p_{ij} \cdot \alpha_i \cdot \alpha_j
\]

where \( \lambda_y \) was the expected number of domestic and sporadic cases, \( \epsilon_i \) the number of travel-related cases, \( \alpha_i \) the number of outbreak-related cases minus one for each outbreak, \( \lambda_i \) the expected number of sporadic and domestic cases per year per Salmonella subtype i from source j, \( p_{ij} \) the prevalence of Salmonella subtype i in source j, \( \alpha_i \) the bacterial-dependent factor for subtype i, \( \alpha_j \) the food-source-dependent factor for source j and \( M_j \) the amount of source j available for consumption per year. Bacterial-dependent factors were represented by survivability during food processing and ability to cause human disease (virulence/pathogenicity), while food-source-dependent factors were represented by likelihood of the investigated sources to cause food-borne infections (e.g. differences in bacterial load, food characteristics influencing growth behavior or preparation procedures). This model, known as “Hald model” or “Danish Salmonella source account model”, also took into consideration the origin of human cases (domestic or travel-related) and whether the cases were sporadic or from an outbreak. For the unknown parameters \( \alpha_i \) and \( \alpha_j \), uninformative uniform distributions were chosen as prior distribution, except for most frequent serotype, and all the subtypes within that serotype, for which \( \alpha_i \) was set to 1. Posterior distributions were calculated through Markov Chain Monte Carlo (MCMC) simulation. This implemented model was used to estimate the contribution to human salmonellosis in Denmark in 1999 due to pork, beef, eggs, broilers, turkeys, ducks, as animal-food sources and included the Salmonella serovars: Enteritidis, Typhimurium, Hadar, Agona, Virchow, Newport, Infantis and Dublin. For Salmonella Enteritidis and Salmonella Typhimurium phage-types were also considered. Moreover, the model estimated the bacterial-dependent factors for each considered serovar and the food-source-dependent factors for each considered source (Hald et al., 2004). In 2007, the Hald model was used to estimate the source of human Salmonella infections in Denmark over the period 2000–2001 by including Salmonella subtypes based on the results of serotyping, antimicrobial resistance profiles and phage-types of S. Enteritidis and S. Typhimurium (Hald et al., 2007).

Wahlström et al. (2011) adapted the Hald model to estimate the number of reported human cases of sporadic Salmonella infections in Sweden during 2004–2006 attributable to each of nine different sources (Wahlström et al., 2011) and similarly, Guo et al. (2011) adapted the Hald model to estimate the burden of food-borne salmonellosis in the United States from 1998 through 2003, attributed to chicken, turkey, egg products, ground beef, intact beef and pork. For the first model serotyping and phage-type data were used, whereas the second model considered only Salmonella serotyping data. Pires and Hald (2010) have further developed the Hald model and presented a three-dimensional version, accommodating data for multiple years in order to attribute human cases of Salmonella in Denmark.
relative to the period 2005–2007. This adaptation of the Hald has been used to attribute human Salmonella infections to the responsible food-animal sources in the EU based on serovar data generated by 24 Member States and considering pigs, broilers, layers and turkeys as animal reservoirs (Pires et al., 2011). Finally, Hald et al. (2012) applied the adaptation of Pires et al. (2011) to assess the relative public health impact deriving from the introduction of a new food safety objective for Salmonella in fattening turkey flocks. The model considered data for 23 Salmonella serovars generated by 25 Member States. The animal-food reservoir studied were again pigs, broilers, layers and turkeys (Hald et al., 2012).

In 2009, the Danish model was modified by Mullner et al. (2009a). The main modifications to the original model included: incorporating uncertainty in the prevalence parameters, modifying prior specification for bacterial-dependent parameters, avoiding food consumption parameters, and changing prior distribution for source-specific parameters. In particular, in order to deal with uncertainty associated with the prevalence estimate, the modified Hald model assumed:

\[ p_{ij} = \pi_j r_{ij} \]

where \( p_{ij} \) was the prevalence of subtype \( i \) in source \( j \), \( \pi_j \) was the prevalence of all subtypes in source \( j \) and \( r_{ij} \) was the relative occurrence of subtype \( i \) among the successfully typed isolates from source \( j \). Unfortunately the full scheme was too complex for the WinBUGS updater, therefore the authors implemented the scheme as a two-stage process. To achieve identifiability Mullner and colleagues introduced a hierarchical structure on the type-dependent parameters assuming \( \log(q_i) \sim N(0,\tau) \) with parameter \( \tau \) Gamma distributed controlling the variation in characteristics between types. In order to avoid food consumption data requirement, the \( M_j \) parameter (amount of source \( j \) available for consumption) was removed and that information was included in the value of the parameters \( q_i \), although this resulted in less meaningful and comparable \( q_i \) estimates. Indeed, a high value of \( q_i \) may reflect a high exposure (i.e. a large market share), but not necessarily a high ability of a food to act as a vehicle for food-borne infection. Moreover, data were divided into observational periods, with the advantage of introducing a temporal dimension into the model. This modified Hald model was applied to several datasets to estimate the contribution of different food-sources to the burden of human salmonellosis and campylobacteriosis in New Zealand (Mullner et al., 2009a).

In 2012 David and colleagues proposed alternative initial parameterizations for the Hald model and studied the effect of the parameters choices on the model robustness, comparing the results with Mullner’s modified Hald model and with a simple attribution model. This analysis was illustrated with the French data on domestic sporadic Salmonella cases registered in 2005. In particular, as reference model, the authors proposed a simple deterministic model for the expected number of human cases due to type \( i \) in source \( j \) as follow:

\[ \lambda_{ij} = \frac{p_{ij} M_j}{\sum_{j} p_{ij} M_j} x_i. \]

The simple attribution model was a revised version of the Dutch model, without source- and type-dependent parameters, but including the parameter \( M_j \) describing the amount of source \( j \) available for consumption. David and colleagues showed that the Hald model was highly sensitive to the parameterization choice, while the Mullner’s modified Hald model provided too large credibility intervals. Finally, in order to improve convergence and avoid biases in the results of the Hald model, they recommended, for those types specific to a unique source, setting \( q_i \) as percentage of human cases due to type \( i \) divided by its prevalence in the source.

An overview of the main features of the studies presented, as well as of other studies based on the same models but applied to different food-borne pathogens is presented in Table 1. As seen from the table, so far, Salmonella subtype-source attribution models have used data generated by phenotypic typing methods, only, while for other bacteria, such as Campylobacter and Listeria monocytogenes, molecular methods have also been applied. The usefulness of molecular subtyping methods for Salmonella attribution models still remains unexplored and it is expected that this approach could find fruitful applications in the near future (EFSA, 2008).

3. Criteria for inclusion of subtyping methods to be used for source attribution

To evaluate the applicability of subtyping methods, several technical parameters, such as discriminatory power, reproducibility, typeability, and stability should be considered (Ramazanzadeh and Mc Nerney, 2007) and they should be balanced against economic considerations (van Belkum et al., 2007).

Many subtyping methods have been developed specifically to characterize very close isolates (e.g. in the framework of investigation of local outbreaks) or, in contrast, specifically to compare quite distant isolates (e.g. long term epidemiological and evolutionary studies). In the first case methods normally investigate fast-evolving markers, whereas in the second case, conserved and stable markers are targeted (Li et al., 2009). In source attribution studies, the level of heterogeneity that should be highlighted is intermediate between those two situations, and consequently no methods appear to be ideal for this purpose. The methods chosen for analysis should be discriminatory enough to enable recognition of correlation between human isolates and their sources but, on the other hand, they should not be too discriminatory, otherwise there is a risk that correlated isolates might not be recognized as related.

Discriminatory power is defined as its ability to distinguish between unrelated strains (Hunter and Gaston, 1988). Some human disease-relevant Salmonella serovars (e.g. S. Enteritidis) and phage-types (e.g. S. Typhimurium DT104) have been described as highly clonal (Cho et al., 2007; Best et al., 2007), hence, for such subtypes, the discriminatory power of subtyping methods used for source attribution needs to be fairly high. Reproducibility is another essential criterion when evaluating a subtyping method. In the framework of a source attribution investigation, it is highly important that the methods are reproducible, as well as not prone to produce false-positive or false-negative results and results produced in one laboratory should be directly comparable to results from other laboratories (definitive methods). To this end the subtyping method should be standardized, so that results can be exchanged between laboratories (Torpdal et al., 2007). Another important criterion in evaluating a typing method is the typeability. This refers to the percentage of distinct strains for which a specific typing profile can be assigned (Hunter and Gaston, 1988). Considering that Salmonella includes more than 2500 different serovars, and that some serovars are considerably different from each other, it is essential for a useful source attribution that the method selected is able to type the vast majority of isolates of interest.

The issue of technical criteria should be balanced against the budget constraints inevitably present when a subtyping method is used on a massive scale as can be reasonably expected for source attribution studies.

Finally, molecular subtyping data must be always interpreted in the proper epidemiological context. When epidemiological studies are conducted the availability of reliable epidemiological information is of paramount importance to draw conclusions and interpret the subtyping data (Barrett et al., 2006).

4. Phenotypic typing methods for Salmonella subtyping source attribution

4.1. Serotyping

As seen in Table 1 serotyping has been the most commonly used subtyping method used for source attribution with Salmonella. Serotyping
takes into account antigenic variability in the outer membrane lipopolysaccharides (O antigens), flagellar proteins (H1 and H2 antigens) and capsular polysaccharide (Vi antigens). The White-Kaufmann-Le Minor scheme (Grimont and Weill, 2007) is the most widely accepted approach used to classify Salmonella serovars. Although more than 2500 different serovars have been described, strains associated with human disease are commonly assigned to relatively few of them. Consequently, different serovars have been described, strains associated with human

approach used to classify

last decade (Malorny et al., 2009; Leader et al., 2009; Barco et al., 2011).

approaches, such as MLST (Achtman et al., 2012), PFGE (Zou et al., 2012),

not easily adapt to a multiplex format. Some molecular subtyping

which can be addressed in a single reaction, since these platforms do

However, PCR typing methods are limited in the number of serovars

PCR protocols, detecting speci

Short Palindromic Repeats (CRISPR) (Fabre et al., 2012), have also

saccharides (O antigens),

takes into account antigenic variability in the outer membrane lipopolysaccharides (O antigens), flagellar proteins (H1 and H2 antigens) and capsular polysaccharide (Vi antigens). The White-Kaufmann-Le Minor scheme (Grimont and Weill, 2007) is the most widely accepted approach used to classify Salmonella serovars. Although more than 2500 different serovars have been described, strains associated with human disease are commonly assigned to relatively few of them. Consequently, serotyping for the most widespread serovars (e.g. S. Typhimurium or S. Enteritidis) is of marginal epidemiological relevance if not supported by further characterization. Conversely, the isolation of strains belonging to rare serovars is more valuable for investigating their possible epidemiological links.

Despite the usefulness of traditional serotyping, this methodology has several drawbacks. It is based on the use of expensive antiserum, the reliability of which is guaranteed only by the expenditure of much quality control effort. Moreover, the procedure is time-consuming, requires well-trained technicians, and, as some isolates are non-typeable, it does not always provide useful results. Nevertheless traditional serotyping remains an important starting point for the characterization of Salmonella isolates within source attribution studies as well as for all other epidemiological investigations. However, this method provides limited information and should be used in conjunction with other subtyping methods.

The limitations of traditional serotyping have sparked interest in searching alternative molecular methods aimed at identifying Salmonella serovars (Wattiau et al., 2011). A number of conventional and real-time PCR protocols, detecting specific serovars, have been published over the last decade (Malorny et al., 2009; Leader et al., 2009; Barco et al., 2011). However, PCR typing methods are limited in the number of serovars which can be addressed in a single reaction, since these platforms do not easily adapt to a multiplex format. Some molecular subtyping approaches, such as MLST (Achtman et al., 2012), PGGE (Zou et al., 2012), ribotyping (Capita et al., 2007) and Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) (Fabre et al., 2012), have also been suggested as methods to classify Salmonella isolates at serovar level. As an alternative Salmonella serovars identification protocols based on microsphere suspension arrays (Fitzgerald et al., 2007; McQuiston et al., 2011) and traditional DNA microarray platforms (Scaria et al., 2008; Tankou-Sandjong et al., 2008) have been developed. For both microarrays (Check&Trace Salmonella, Check-Points BV) and suspension arrays (Luminex xMAP® Salmonella serotyping assay, Luminex Corporation) kits have been recently commercialized, thereby these alternative approaches, based on innovative technologies, are currently available for first line laboratories to obtain serovar and genovar identifications.

4.2. Phage-typing

This is the traditional method for Salmonella subtyping beyond the serovar level (Best et al., 2007). In this phenotypic subtyping approach, Salmonella strains are separated into different phage-types based on their reactivity against a set of serovar specific typing phages. Phase-typing schemes have been developed for some relevant serovars (Olsen et al., 1993). The method is cheap and does not require specific equipment. Unfortunately, only a limited number of reference laboratories, possessing the set of typing phages, can perform the method (Boxrud et al., 2007). Phage-typing requires great experience in the interpretation of results and is not always fully reproducible between laboratories (Ross and Heuzenroeder, 2005). Standardization and quality assurance are crucial to improve the robustness of the method and guarantee comparability among laboratories (Baggesen et al., 2010). Although phage-typing represents a valuable tool in characterizing isolates that exhibit less common phage reaction patterns, the technique is inadequate in the investigation of more common phage-types (Hopkins et al., 2011). Another drawback of the method is the phage conversion (Cho et al., 2008). Different mechanisms, such as loss or acquisition of plasmids, mutation of genes encoding lipopolysaccharide, and expression of temperate phages have been identified as potential causes of phage conversion in Salmonella isolates (Olsen et al., 1993).

Despite these limitations, phage-typing still represents a valuable tool for an initial evaluation of the potential relatedness among isolates, especially in the framework of source attribution studies investigating S. Enteritidis and S. Typhimurium. In particular, when phage-typing is combined with antimicrobial resistance or other subtyping methods, it appears as a valuable tool for epidemiological investigations (Barco et al., 2012) and source tracking, in particular for identifying strains related to specific countries (Hopkins et al., 2010) and/or species (Carrique-Mas et al., 2008).

4.3. Antimicrobial resistance typing

This method determines the profile of resistance of a microbial strain towards a panel of antimicrobial agents. Despite the fact that it has been quite commonly used in the past as a subtyping method to determine the correlation between isolates, during recent years antimicrobial resistance typing has been less frequently used for this specific purpose. Like phage-typing, this method is cheap and does not require specific equipment and reagents. Well-standardized protocols have been developed both for performing the analysis and for interpreting results.

Table 1

Overview of developed models for Salmonella food-borne source attribution studies based on microbial subtyping. Also models initially developed for Salmonella and then adopted for other food-borne pathogens are reported.

<table>
<thead>
<tr>
<th>Reference</th>
<th>Food-borne pathogen</th>
<th>Subtyping method</th>
<th>Country</th>
<th>Period of analysis</th>
<th>Model used</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Hald et al., 2004)</td>
<td>Salmonella spp.</td>
<td>Serotyping, phage-typing</td>
<td>Denmark</td>
<td>1999</td>
<td>Hald model</td>
</tr>
<tr>
<td>(Hald et al., 2007)</td>
<td>Salmonella spp.</td>
<td>Serotyping, antimicrobial resistance profile; phage-typing</td>
<td>Denmark</td>
<td>2000–2001</td>
<td>Hald model</td>
</tr>
<tr>
<td>(Gus et al., 2011)</td>
<td>Salmonella spp.</td>
<td>Serotyping</td>
<td>USA</td>
<td>1998–2003</td>
<td>Hald model</td>
</tr>
<tr>
<td>(Little et al., 2012)</td>
<td>Salmonella spp.</td>
<td>Serotyping</td>
<td>England and Wales</td>
<td>2004–2007</td>
<td>Modified Hald Model</td>
</tr>
<tr>
<td>(Wahlström et al., 2011)</td>
<td>Listeria monocytogenes</td>
<td>Serotyping AFLP subtype</td>
<td>Sweden</td>
<td>2004–2006</td>
<td>Modified Hald Model</td>
</tr>
<tr>
<td>(Pires et al., 2012)</td>
<td>Salmonella spp.</td>
<td>Serotyping</td>
<td>25 EU Member States</td>
<td>2010</td>
<td>Modified Hald model</td>
</tr>
<tr>
<td>(Muliner et al., 2009a)</td>
<td>Campylobacter jejuni</td>
<td>Multi-locus sequence typing</td>
<td>New Zealand</td>
<td>2005–2008</td>
<td>Modified Hald model</td>
</tr>
<tr>
<td>(Muliner et al., 2009b)</td>
<td>Campylobacter jejuni</td>
<td>Serotyping</td>
<td>New Zealand</td>
<td>2002–2004</td>
<td>Modified Hald model</td>
</tr>
<tr>
<td>(David et al., 2012)</td>
<td>Salmonella spp.</td>
<td>Serotyping; antimicrobial resistance profile; phage-typing</td>
<td>France</td>
<td>2005</td>
<td>Hald model, modified Hald model, simple attribution model</td>
</tr>
</tbody>
</table>
Although some resistance determinants are very stable, such as mutations located in chromosomal genes and conferring resistance to fluoroquinolones (Giraud et al., 2003), the resistance factors carried on plasmids, integrons and genomic islands are easily transferable between strains (Miragou et al., 2006). The instability characterizing these latter resistance targets reduces their relevance for epidemiological purposes. The value of resistance patterns as epidemiological markers can be further hampered by the fact that the same phenotypic resistance profiles can be due to the presence of different genetic targets (Le Minor, 1988).

A study conducted to determine the genetic relatedness among Salmonella serovars isolated from domestic animals and humans by using serotyping, PFGE and antimicrobial resistance profiling confirmed the limitation of antimicrobial resistance for tracking potential sources of human infections. Although similar Salmonella PFGE profiles were found in humans and animals, the level of resistance was higher in animals than in humans, since the animals were subjected to higher antimicrobial pressure and the gut environment in animals favors the rapid exchange of plasmids and subsequent development of higher resistance levels (Oloya et al., 2009).

Another main drawback related to antimicrobial resistance typing arises from the lack of agreement between the programs in place in terms of the antibacterial compounds tested, analytical methodologies used, and interpretative criteria followed (Silley et al., 2011). The lack of harmonization is particularly evident between human and veterinary laboratories.

Following such considerations, and despite antimicrobial resistance profiling having been used as a subtyping method for generating data for source attribution models, these data should be used for this purpose after careful assessment (EFSA, 2008), mainly due to the instability of some targets related to resistance during the transfer of Salmonella isolates from sources to human hosts.

5. Genotypic methods for Salmonella subtyping source attribution

In recent years many DNA-based genotyping techniques have been utilized to delineate epidemiological relationships between various Salmonella isolates. These subtyping molecular methods have been classified into three main categories including: DNA binding pattern, DNA sequencing and DNA hybridization-based methods. The first group includes methods discriminating the strains based on differences in the size of the DNA bands (fragments) generated by the amplification of the genomic DNA or by cleavage of DNA using restriction enzymes. The second group includes methods assessing the nucleotide sequences of target regions, whereas the last group mainly refers to microarray methods (Li et al., 2009). The great majority of the methods historically used for Salmonella subtyping are based on a PCR step to generate billions of copies of the genomic fragment of interest. Among the plethora of subtyping methods based on PCR, some of them, such as Random PCR amplification of polymorphic DNA (RAPD), Enterobacterial Repetitive Intergenic Consensus (ERIC) fingerprinting, ISS200-PCR, Amplified fragment length polymorphism (AFLP) and fluorescent AFLP have been traditionally used to subtype Salmonella; some others, such as the Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) have been described only more recently (Fabre et al., 2012).

The present review focuses on a selection of the more widespread molecular methods for Salmonella subtyping used by the European surveillance laboratories and these methods are considered in relation to source attribution studies. In particular, according to a questionnaire survey carried out by the European Food Safety Authority (EFSA) in 2009, the most common molecular methods used by EU laboratories to subtype Salmonella strains isolated from food and animals include Pulsed Field Gel Electrophoresis (PFGE), Multi Locus Variable-Number Tandem Repeat Analysis (MLVA) and ribotyping, performed in 19, 11 and 9 countries respectively. Other less common methods are plasmid profile analysis (3 countries) and Multi Locus Sequence Typing (2 countries) (EFSA, 2009).

5.1. Pulsed field gel electrophoresis (PFGE)

In the past decade PFGE has been a valuable tool for Salmonella subtyping and PFGE continues to be the most widely used molecular subtyping method (Barrett et al., 2006). It is referred as the “gold standard” (Call et al., 2008) and alternative methods for Salmonella subtyping are commonly compared against PFGE. PFGE uses restriction enzymes characterized by infrequent recognition sites, cutting the entire circular DNA of a propagated bacterial isolate and generating large DNA fragments (ranging from 20 to 800 kb), that can be separated in a flat agarose gel by applying alternating electric fields at different angles (Foley et al., 2009).

There is a long tradition of international standardization of PFGE for Salmonella and other food-borne pathogens by PulseNet (Ribot et al., 2006). Protocols defined in the framework of this network are routinely used by surveillance laboratories worldwide (Swaminathan et al., 2006; Soyer et al., 2010), and PFGE patterns are shared in online databases for global comparison (Li et al., 2009). The high degree of standardization, the wide use of the method and the fact that some assays can be used to analyze almost all Salmonella serovars (Broschat et al., 2010) make it attractive to source attribution, however, it is not without drawbacks. The method is complicated, labor-intensive, requires very good technical expertise and it cannot be automated (Fabre et al., 2012). Furthermore, typeability is not excellent for some serovars, such as Salmonella Livingston and Salmonella Cerro because of DNA degradation during PFGE process (Murase et al., 2004). However, this limitation can be solved by supplementing electrophoresis buffer with thiourea (Silbert et al., 2003).

PFGE is recognized as one of the most discriminatory typing techniques and numerous studies have demonstrated its value in successfully tracking infections due to different Salmonella serovars and linking human illness to specific sources (Harbottle et al., 2006; Oloya et al., 2009; Soyer et al., 2010; Hauser et al., 2012). Sometimes, PFGE has been shown too discriminatory and it may show different PFGE patterns for isolates that share a very recent common ancestor (Soyer et al., 2010). This of course will influence conclusion based on source attribution studies. Conversely, sometimes PFGE is not discriminatory enough, since it may classify epidemiologically unrelated isolates into identical PFGE types, specifically when used to type some genetically highly similar strains, such as S. Enteritidis (Boxrud et al., 2007; Cho et al., 2007) or phage types, such as S. Typhimurium DT104 (Lindstedt et al., 2003; Best et al., 2007). In those instances when a single-enzyme (i.e., XbaI) PFGE lacks the discriminatory power to partition strains into epidemiologically meaningful clusters, the combination of profiles generated by using additional restriction enzymes (e.g. BlnI, SpeI) can enhance the value of this method for differentiating highly homogeneous Salmonella strains (Zheng et al., 2011). Therefore the definition of appropriate and fully accepted interpretative criteria to analyze and compare PFGE profiles is a key issue. According to Tenover et al. (1995) PFGE profiles differing in the position of up to three bands should be considered as closely related, since this difference can be generated by a single genetic event, whereas profiles differing in the position of up to six bands should be considered as possibly related, since this is the final result of two genetic events. These criteria were originally set up for nosocomial outbreaks and more restrictive PFGE criteria were then suggested by Barrett et al. (2006) to identify clusters of infections or the index-cases in point-source outbreaks. It is evident that there is no agreement yet on the criteria for evaluating PFGE profiles. Since isolates subtyped for source attribution come from different sources and are geographically and temporally related, but may not represent the direct link between the affected person and the specific source, a certain level of variability between isolates should be accepted when

PFGE profiles are compared. In the framework of source attribution investigations, specific guidelines should be formulated for evaluating PFGE patterns, taking into account the types of isolates studied (in terms of origin as well as of temporal and geographical correlation) and the intrinsic variability of the Salmonella serovars typed.

5.2. Multiple-locus variable-number tandem repeat analysis (MLVA)

MLVA is a PCR genotyping method based on the polymorphic analysis of multiple variable number tandem repeats (VNTR) loci (Li et al., 2009). For routine surveillance of Salmonella, MLVA seems to have notable advantages compared to PFGE. It is cheaper, and the protocol is less labor-intensive and faster than PFGE (Torp dahl et al., 2007). Moreover, MLVA can be completely automated and the generated data are more easily analyzed and shared between laboratories (Torp dahl et al., 2007; Hopkins et al., 2007). In addition, MLVA generally demonstrates a higher discriminatory power than PFGE (Best et al., 2007; Torp dahl et al., 2007). However, MLVA has a notable limitation in relation to source attribution studies: the great variability of Salmonella serovars. Despite some MLVA protocols characterizing several Salmonella serovars having been devised (Table 2), it has been demonstrated that a MLVA scheme effective for one Salmonella serovar is not usually equally discriminative for other Salmonella serovars (Ross et al., 2011). Further, the great majority of protocols have been tested only on a limited number of strains by only a few laboratories. At European level, currently two standardized protocols have been defined for S. Typhimurium (Larsson et al., 2009) and for S. Enteritidis (Hopkins et al., 2011). Moreover, for these serovars, two panels of reference strains with verified fragment sizes, covering the range of the most common alleles for each locus, have been identified to normalize the raw data generated, thereby allowing direct comparison between laboratories irrespective of the platform used for fragment analysis (Larsson et al., 2009).

A main concern in terms of applicability of MLVA in source attribution models is the stability of the genetic targets investigated, since it has been evidenced that the target loci may evolve too quickly to provide a reliable indication about the relationships among potentially related strains (Lindstedt, 2005). For instance, Hopkins et al. (2007, 2011), analyzing strains of S. Typhimurium and S. Enteritidis isolated from outbreaks, demonstrated that despite VNTR being generally stable, small changes in loci occurred. The instability of VNTR loci could hamper the usefulness of MLVA, especially for long-term epidemiological studies (Lindstedt, 2005; Li et al., 2009), and hence the dynamics of MLVA loci over an extended period of time remains an issue that deserves further evaluation (Sintchenko et al., 2012). Despite the drawback mentioned, MLVA seems to be a promising tool for deciphering the potential sources of human infections, especially when it is used in parallel with PFGE as demonstrated by Best et al. (2007). These authors used the “European MLVA S. Typhimurium protocol” to characterize epidemiologically unrelated strains isolated over eight years from animals and humans, and demonstrated that MLVA is an accurate method for tracing strain sources. The same finding emerged when comparing MLVA profiles of S. Enteritidis (Cho et al., 2008, 2010) and of other serovars (van Cuyck et al., 2011) for isolates obtained from humans and different sources. In order to partially overcome the limitation correlated to the high and rapid variability of VNTR and to obtain a more reliable picture of the epidemiological situation investigated, as already mentioned for PFGE, it is also imperative to develop guidelines for interpreting the MLVA profiles. In particular, specific cut-offs, defined as acceptable differences in the numbers of repeats in one or more specific loci, should be defined. Once these specific guidelines and international consensus for the interpretation of VNTR variability are available, MLVA may become an invaluable Salmonella subtyping method, both for outbreak investigations and for long-term surveillance studies such as source attribution models, with the more common serotypes.

Table 2
MLVA schemes developed. Information between brackets corresponds to loci reported in other protocols.

<table>
<thead>
<tr>
<th>Serovar</th>
<th>Loci investigated (alias)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>S. Typhii</td>
<td>Sal02, Sal04, Sal06, Sal10, TR1 (Sal11), Sal15, STTR5 (Sal16), Sal20, TR5 (Sal22), Sal23</td>
<td>(Ramisse et al., 2004)</td>
</tr>
<tr>
<td>S. Typhii; S. Typhimurium; additional serovars</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S. Derby</td>
<td>STTR5 (combined with sequence typing of sopA, sopD and sopD genes)</td>
<td>(van Cuyck et al., 2011)</td>
</tr>
<tr>
<td>S. Enteritidis</td>
<td>SE1; SE2; SE3; SE4; SE5; SE6; SE7; SE8; SE9</td>
<td>(Hauser et al., 2011)</td>
</tr>
<tr>
<td>S. Infantis</td>
<td>STTR2; STTR3; STTR5; STTR6; STTR7; STTR9; STTR10; TR1; TR2; Sal06; Sal23; SE1; SE2</td>
<td>(Boxrud et al., 2007)</td>
</tr>
<tr>
<td>S. Montevideo</td>
<td>M6; M10 (SE5, STTR5, 3184543); M12</td>
<td>(Cho et al., 2007)</td>
</tr>
<tr>
<td>S. Newport</td>
<td>STTR5; STTR6; NewportA; NewportB; NewportM; NewportL</td>
<td>(Malorny et al., 2008)</td>
</tr>
<tr>
<td>S. Paratyphi A</td>
<td>Sty14; Sty16; Sty19; Sty26; Sty37 (Sal02); Sty44 (STTR5)</td>
<td>(Cho et al., 2008);</td>
</tr>
<tr>
<td>S. Typhii</td>
<td>TR1; TR2; TR3</td>
<td>(Beraneik et al., 2009)</td>
</tr>
<tr>
<td>S. Typhii</td>
<td>TR4500; TR6499</td>
<td>(Hopkins et al., 2011)</td>
</tr>
<tr>
<td>S. Gallinarum</td>
<td>SE1; SG1; SG2; SG3; SG4; STTR10pl</td>
<td>(<a href="http://www.pulsenetwork.org">www.pulsenetwork.org</a>) (last visit 20.12.2012)</td>
</tr>
<tr>
<td>S. Gallinarum</td>
<td>SGTR1; SGTR2 (SENTR1; SE10); SGTR3 (SENTR4, SE1); SGTR4 (STTR7, SENTR2)</td>
<td>(Bergamini et al., 2011)</td>
</tr>
<tr>
<td>S. Infantis</td>
<td>STTR2; STTR3; STTR5; STTR6; STTR7; STTR9; STTR10; TR1; TR2; Sal06; Sal23; SE1; SE2</td>
<td>(Kang et al., 2011)</td>
</tr>
<tr>
<td>S. Montevideo</td>
<td>M6; M10 (SE5, STTR5, 3184543); M12</td>
<td>(Rois and Heuzenroeder, 2008)</td>
</tr>
<tr>
<td>S. Newport</td>
<td>STTR5; STTR6; NewportA; NewportB; NewportM; NewportL</td>
<td>(Harada et al., 2011)</td>
</tr>
<tr>
<td>S. Paratyphi A</td>
<td>Sty14; Sty16; Sty19; Sty26; Sty37 (Sal02); Sty44 (STTR5)</td>
<td>(Davis et al., 2009)</td>
</tr>
<tr>
<td>S. Typhii</td>
<td>TR1; TR2; TR3</td>
<td>(Tien et al., 2011)</td>
</tr>
<tr>
<td>S. Typhii</td>
<td>TR4500; TR6499</td>
<td>(Liu et al., 2003)</td>
</tr>
<tr>
<td>S. Typhii</td>
<td>ST2; ST3; STy20 (TR4500); Sty25 (TR6499); Sty37 (Sal02); Sty39 (Sal06); Sty40 (Sal20); STy41 (TR1, Sal11)</td>
<td>(Octavia and Lan, 2009)</td>
</tr>
<tr>
<td>S. Typhii</td>
<td>ST2; ST3; TR22; ST22</td>
<td>(Tien et al., 2012)</td>
</tr>
<tr>
<td>S. Typhii</td>
<td>B194547; 2628542; 2730867 (STTR6)</td>
<td>(Witonski et al., 2006)</td>
</tr>
<tr>
<td>S. Typhimurium; S. Newport</td>
<td>STTR1; STTR2; STTR3; STTR4; STTR5; STTR6; STTR7; STTR8</td>
<td>(Lindstedt et al., 2003)</td>
</tr>
<tr>
<td>S. Typhimurium</td>
<td>STTR2; STTR3; STTR5; STTR6; STTR7; STTR9; STTR10pl</td>
<td>(Lindstedt et al., 2004)</td>
</tr>
<tr>
<td>S. Typhimurium</td>
<td>STTR2; STTR3; STTR5; STTR6; STTR7; STTR9; STTR10; TyR1; TyR2; TyR3</td>
<td>(Ross et al., 2009)</td>
</tr>
<tr>
<td>S. Typhimurium</td>
<td>ST01 (STTR7); ST06 (STTR1); ST16; ST17; ST19 (STTR6); ST20; ST22; ST23; ST25 (STTR5); ST26 (STTR9); ST28; ST30; ST35; ST36; ST38; ST40 (STTR10)</td>
<td>(Chiou et al., 2010)</td>
</tr>
<tr>
<td>S. Typhimurium</td>
<td>STTR5; STTR6; STTR9; STTR10pl; STTR11</td>
<td></td>
</tr>
<tr>
<td>S. Typhimurium</td>
<td>ST2; ST3; ST5; ST6; ST7; ST8; ST10</td>
<td>(Brosch et al., 2010)</td>
</tr>
<tr>
<td>S. Typhimurium</td>
<td>STTR5; STTR6; STTR9; STTR10pl; STTR11</td>
<td></td>
</tr>
<tr>
<td>S. Typhimurium</td>
<td>ST2; ST3; ST5; ST6; ST7; ST8; ST10</td>
<td>(<a href="http://www.pulsenetwork.org">www.pulsenetwork.org</a>) (last visit 20.12.2012)</td>
</tr>
</tbody>
</table>
5.3. Ribotyping

Bacterial rRNA operons are highly conserved, but the flanking DNA regions are more variable. The differences in these variable regions result in different ribopatterns when probed with DNA probes analogous to the conserved domains of rRNA genes (Bauchet et al., 2008). Manual ribotyping is technically demanding and requires experienced technicians. Comparability of results between different laboratories is quite limited and even marginal variations in the protocol can influence the final result (Pavlic and Griffiths, 2009). Therefore, manual ribotyping is inadequate for routine surveillance and cannot be used for source attribution studies where massive subtyping of Salmonella strains is performed. To overcome these limitations, ribotyping was adapted to an automated protocol using a patented technology (RiboPrinter™), thereby making it possible to subtype a large number of isolates (Clark et al., 2003). RiboPrinter is a fully automated system that allows sensitive identification and subtyping of some food-borne pathogens (Pavlic and Griffiths, 2009). In particular, for Salmonella, it was demonstrated that the method is able to differentiate isolates of the same serotype and phage-type (Capita et al., 2007). Since highly specific skills are not required to maintain a consistent protocol, the system can be particularly suitable for food industry laboratories performing source tracking investigations. The results obtained can be readily standardized among different laboratories (Clark et al., 2003). The main limitation of the automated method is the high cost of the equipment needed to perform the analysis. Therefore, its applicability in public health laboratories can be justified only by appropriate cost-benefit analyses (Fontana et al., 2003).

Ribotyping has been demonstrated to be a very discriminatory technique for some Salmonella serovars (Liebana et al., 2001, 2002a, 2002b, 2004; Clark et al., 2003), however some authors have reported less convincing discriminatory power for other serovars (Lindqvist et al., 2004; Adaska et al., 2006), especially compared to PFGE (Erikssson et al., 2005). Taken together, the complexity of the manual method, the high cost of the equipment and the a priori uncertainty with respect to discriminatory power probably means that this method will not be widely used for source attribution studies.

5.4. Plasmid profile analysis

When the bacterial cell divides, copies of the original plasmid are transferred to the daughter cells (Foley et al., 2009). Thus, theoretically isolates of a clonal line carry the same plasmids, however, plasmids can also be horizontally transferred among bacteria, even of different genera and kingdoms though the conjugation process (Carattoli, 2011). It is also recognized that strains with the same chromosomal features may produce different plasmid restriction patterns and that the same plasmid profile may be produced by strains which are different at the chromosomal level (Olsen et al., 1993). Moreover, configurational changes in plasmids (linear versus supercoiled) show different electrophoresis migration patterns, and copies of the same plasmid with different structures could generate separated bands on a gel, confusing the interpretation of results. Conversely, different plasmids with similar molecular weights co-migrate producing indistinguishable bands on a gel (Foley et al., 2009). In addition, in some cases, this method has a low discriminatory power (Foley et al., 2006) and finally, plasmid-free isolates are nontypeable (Adaska et al., 2006). In past decades plasmid profiling has been massively used by laboratories for characterization of Salmonella isolates, partly due to the simplicity of the method and the limited requirement in terms of equipment and reagents. Nowadays, due to the drawbacks reported, the use of the method is limited to specific epidemiological studies and it is always used to support other more discriminative and robust typing methods. Although in some specific cases the plasmid profile resulted the only subtyping method that leads to differentiation between unrelated strains (Lawson et al., 2004), plasmid profiling should be generally considered of limited utility for source attribution studies, since the genomic targets investigated can be inherently unstable.

5.5. Multi-locus sequence typing (MLST)

This method examines sequences of multiple housekeeping genes that are highly conserved, because of the essential functions of the proteins they encode (Li et al., 2009). This method is objective, reproducible and characterized by high typeability. However, since housekeeping genes are under limited selective pressure, the accumulation of mutations is commonly very slow, making MLST an ideal tool for assessing the evolution of organisms, while the usefulness of this method to trace infections and investigate outbreaks is more controversial (Foley et al., 2006). Moreover, since MLST characterization is based on single nucleotide base changes, in order to guarantee the reliability of typing it is essential to get high quality sequencing data (Foley et al., 2009).

Several studies investigating Salmonella isolates have highlighted the main disadvantage of MLST in relation to typing: its low discriminatory power, especially when this method is used to type isolates of the same serovar (Alcaine et al., 2006; Torpdahl et al., 2005; Foley et al., 2006; Harbottle et al., 2006; Hauser et al., 2012). In order to improve the discrimination of traditional MLST, many other protocols investigating sequences in more rapidly changing genes than in housekeeping genes have been developed. For instance, some protocols combined the analysis of sequence of housekeeping genes and temperate phages (Ross and Heuzenroeder, 2005, 2008) or virulence genes (Foley et al., 2006; Tankouo-Sandjong et al., 2007). However, only in very few cases did MLST protocols show a good discriminatory power (Foley et al., 2006; Kotetishvili et al., 2002; Dione et al., 2011).

In general, the low discriminatory power renders MLST unsuitable for source attribution studies, with the possible exception of very long term studies of changes in the importance of sources and types.

6. Future perspectives in molecular subtyping methods

Newer molecular subtyping methods, such as microarrays, single-nucleotide polymorphisms (SNPs) and whole genome sequencing have been recently used to characterize food-borne pathogens (Boxrud, 2010).

Analysis of multiple single-nucleotide polymorphisms (SNPs) takes advantage of nucleotide mutations at specific loci in the bacterial genome to compare isolates of the same species. By simultaneously investigating multiple SNPs, the relatedness of these strains can be elucidated (Foley et al., 2009). Analysis of SNPs has been used to characterize Salmonella isolates (Baker et al., 2008; Octavia and Lan, 2010; Pang et al., 2012). Currently, this approach cannot be considered feasible for source attribution studies since the discriminatory power of the method is strictly related to the SNPs investigated and SNPs polymorphism sometimes does not provide adequate diversity to group related strains into clusters (Boxrud, 2010). However, this methodology may become attractive in future, when more efficient methods for detecting polymorphisms are available (Boxrud, 2010).

Another promising subtyping approach is the whole genome sequencing (WGS). Technologies to perform such investigations have advanced considerably over the last decade and it is reasonable to assume that, in the very near future, WGS will be affordable, rapid and simple to use, becoming a method not restricted to highly specialized laboratories only, but available to a large number of diagnostic laboratories. However, the largest barrier to the implementation of WGS in routine practice is the acquisition of competences to manage, analyze and interpret the colossal amounts of bioinformatic data generated by this technology (Dunner et al., 2012). Hence, the development of user-friendly bioinformatics tools and databases to better cope with these types of data will be a crucial factor determining how WGS will transform diagnostic clinical and public health microbiology (Török and Peacock, 2012). Although there is still room for improvement, in
the not too distant future WGS could replace all other subtyping methods that have been massively used so far for bacteria subtyping. Certainly, this method could be an appealing approach to detect genetic elements and area of variations between strains that may be then used as targets for improving other subtyping methods (Hyytiä-Trees et al., 2007).

Also microarray has to be considered among the innovative subtyping techniques. This technique aims at investigating polymorphic DNA regions using multiple DNA probes spotted on a glass slide, or a combination of single probes bound to uniquely identified bead sets (Boxrud, 2010). Spotted arrays allow to simultaneously query thousands of probes, but the cost of developing an assay is prohibitive for massive application as it would be required for epidemiological studies. For bead arrays the number of probes investigated is more limited, but assays are more flexible, costs are competitive, results are of easier interpretation and the technology is currently available in many public health laboratories (Dunbar, 2006). Both methods have been applied to characterize Salmonella isolates (Tracz et al., 2006; Huehn and Malorny, 2009; Hauser et al., 2011; Fabre et al., 2012; Fang et al., 2012).

These high-throughput subtyping methods may play a significant role in future and it is likely that some of them will ultimately replace current methodologies. For now, they remain limited to a few and very specialized research laboratories. Moreover, they are too cumbersome to be applied in traditional public health laboratories, like the ones usually involved in source attribution studies. Finally, these advanced subtyping methods generate an overload of information that is redundant for epidemiological investigations like source attribution studies.

7. Conclusions

Salmonella source attribution through microarray subtyping requires a collection of isolates representative of the real exposure of humans to the microorganism. When the contribution of different animal sources or reservoirs is compared, it is crucial that strains deriving from all possible sources are correctly represented. The lack of strains from a specific source can lead to the overestimation of another source with a similar distribution of subtypes. Moreover, when more food sources are attributable to the same reservoir (e.g. eggs and poultry meat), the allocation of cases to the different sources within the same reservoir should be complemented by sound epidemiological data.

Non-human isolates are compared to the human ones by using discriminative subtyping methods, in order to identify overlapping profiles between them. Shared veterinary and medical databases incorporating data generated through harmonized protocols would facilitate the achievement of this goal. Moreover, the subtyping methods used for the source attribution studies should be well-established, and their procedures should be described in detail, especially for those cases, such as the serotyping of as S. 4,[5],12:i:-, for which the lack of harmonization in the method used could generate confusion in the final reporting. Furthermore, harmonized criteria should be defined for the selection of the strains to be analyzed according to the subtyping methods chosen. Finally, sources of human infections can be inferred by using a mathematical model which utilizes the subtyping data obtained. One of the most critical issues is how to balance the need for reliable subtyping data, representative of the isolates circulating in the sources of interest, with the feasibility in terms of affordability of laboratory subtyping methods. So far, Salmonella source attribution models have relied on data generated by phenotypic typing methods, such as serotyping and phage-typing. However, since human infections are usually attributed to very few serovars and phage-types, source attribution investigations could gain an advantage by combining data generated by traditional phenotypic and advanced molecular methods. Among the molecular methods currently used by European laboratories, PFGE and MLVA are expected to be increasingly applied for this purpose in the near future. Despite PFGE being a well-established technique, it still remains the most commonly used molecular typing method for Salmonella surveillance. Good knowledge of the variability of pulsotypes for the most common Salmonella serovars and phage-types is available from data collections which have been stored long-term. Although the method is labor-intensive, it is routinely applied in surveillance laboratories worldwide.

MLVA is an innovative and new technique, and it seems a promising alternative to PFGE thanks to its high discriminatory power, reproducibility and the possibility of automation. Nevertheless, the most effective MLVA schemes described so far are serovar-specific and cannot be applied to isolates of different serovars. Moreover, the usefulness of MLVA has been demonstrated for tracing outbreaks, but the significance of this technique deserves further investigations for long-term surveillance, including for source attribution studies. Although the method is becoming more widely used, mainly for subtyping S. Typhimurium isolates, a complete understanding of the variability of the loci investigated is currently lacking. Further information on the mutation rate of VNTR loci would provide indications about the significance of any variations that are detected in order to infer the epidemiological relationship among strains.

Guidelines need to be developed for both PFGE and for MLVA in order to interpret the profiles obtained in isolates from the different sources and identify potential correlations between these profiles and those found in human isolates. An acceptable level of diversity should be defined so that not only isolates producing identical PFGE band patterns or MLVA profiles could be labeled as related, but also similar profiles, presenting an acceptable level of difference, could be clustered and considered as potentially correlated. Since for source attribution models quite heterogeneous isolates are investigated (i.e. collected from different sources, temporal frames and geographical areas), it would be advisable to use highly discriminatory subtyping methods and, when possible, multiple methods in parallel, to get a clear picture of the differences between the strains. On the other hand, profiles obtained should be clustered based on interpretative criteria that allow the correlation of similar profiles, i.e. those presenting a level of diversity that is lower than a specified cut-off value. Such values must be defined according to the scope of the study, the heterogeneity of the serovar/phage-type investigated, and, in case of MLVA, also the variability of the single locus investigated.

Hence, for Salmonella source attribution studies it could be useful to use in parallel discriminative subtyping methods, such as PFGE and MLVA, applying appropriate guidelines to interpret and cluster the profiles obtained. Combining typing methods, that reflect different parts of the genome leads to a higher accuracy in strains discrimination and consequently to a higher confidence in the data generated by the source attribution model used.

Another crucial issue to succeed in such types of projects is represented by the implementation of effective national and international networks combining laboratories that cover the populations of interest and regularly share subtyping data and epidemiological information. The availability of comparable subtyping data obtained from a sufficiently large and representative subset of samples from the sources of interest, supported by consistent epidemiological information and generated through standardized laboratory methodologies is a crucial issue for optimizing these surveillance networks with the specific aims of promoting outbreaks detection and addressing sources identification.

References


Tien, Y.-Y., Wang, Y.-W., Tung, S.-K., Liang, S.-Y., Chiou, C.-S., 2011. Comparison of multilocus variable-number tandem repeat analysis and pulsed-field gel electrophoresis in molecular subtyping of Salmonella enterica serovars Paratyphi A. Diagnostic Microbiology and Infectious Disease 69, 1–6.


