White Paper: Current and Future Development and Use of Molecular Subtyping by USDA-FSIS

USDA-FSIS Subtyping Work Group: Nate Bauer[1], Peter Evans[2], Beth Leopold[3], Jeoffrey Levine[1], and Patricia White[1]

Executive Summary

This White Paper describes current Food Safety and Inspection Service (FSIS) subtyping procedures and provides examples of how the data are used. The document also discusses subtyping methods under development (one to four years to implementation), and ideas for the future (five or more years to potential implementation). It was created to inform FSIS program offices and stakeholders about the availability, uses, strengths, and limitations of current and future methods from the perspective of the laboratorians and epidemiologists, and to provide Agency decision-makers with information necessary to make informed, science-based decisions related to routine monitoring programs and illness investigations, and to evaluate and build FSIS capabilities.

FSIS is the public health agency of the U.S. Department of Agriculture (USDA) responsible for ensuring that the nation's commercial supply of meat, poultry, and processed egg products is safe, wholesome, and correctly labeled and packaged. FSIS conducts microbiological sampling programs to verify that establishments are controlling foodborne hazards. All FSIS microbial testing procedures include cultural isolation and isolate characterization to the species level. In addition, one or more isolates from each sample are subtyped.

FSIS currently uses Pulsed-field Gel Electrophoresis (PFGE) and Multi-Locus Variable Number of Tandem Repeat (VNTR) Analysis (MLVA) to identify outbreaks, trace contamination within an establishment, and investigate potential harborage sites.

[1] Office of Public Health Science (OPHS)
[2] FDA Center for Food Safety and Applied Nutrition (CFSAN)
[3] FSIS Office of Investigation, Enforcement and Audit (OIEA)

Currently, PFGE is considered the best method available for public health laboratories to detect and track outbreaks. MLVA may help to further differentiate subtypes during outbreak investigations. PFGE and MLVA data must be used in conjunction with epidemiological evidence, environmental assessments, trace-back activities, and other information collected during outbreak investigations.

Currently the Outbreak Section of the Eastern Laboratory (OSEL) is responsible for performing all FSIS subtyping analyses. OSEL communicates PFGE pattern information to the national PFGE database maintained by the Centers for Disease Control and Prevention (CDC), where they are compared to patterns from clinical and food isolates provided by local, state, and federal public health and agricultural agencies and by academia. PulseNet is an inter-laboratory network consisting of CDC, state, and local health departments, and food regulatory agencies (FDA and FSIS) which share PFGE pattern data for foodborne pathogens. PulseNet scientists continuously evaluate the pattern databases, and review, oversee quality assurance, and develop new subtyping methods.

Most PFGE pattern-related information communicated within FSIS is related to outbreak investigations, but PFGE pattern information also is communicated to Office of Field Operations (OFO) district offices when there is evidence of recurring subtypes or serotypes of concern in an establishment. The Agency plans to provide this information routinely to federally regulated establishments. FSIS has implemented a Public Health Information System (PHIS), with the capability to detect trends in subtype data.

FSIS has a unique authority to regulate establishments engaged in the production and distribution of meat, poultry, and egg products. Subtyping data can allow FSIS to make scientifically-informed decisions and allocate resources. In general, subtyping information may be used to:

1. Detect and track outbreak clusters or harborage within an establishment;
2. Assess sample-specific risk of illness associated with a product, a process, or an establishment;
3. Provide highly specific data for risk assessments and attribution studies;

4. Define a regulatory adulterant with greater accuracy; and
5. Improve the accuracy and efficiency of current subtyping methods.

FSIS collaborates with USDA Agriculture Research Service (ARS), CDC, and FDA to develop new subtyping and detection assays. FSIS worked with ARS groups to develop PCR-based screening assays for virulence genes of non-O157 STEC and a single nucleotide polymorphism (SNP)-based typing assay for *Listeria monocytogenes (Lm)*. The Agency implemented a molecular serotyping method for *Salmonella* developed by CDC. FSIS participates in meetings designed to bring public health agencies together to discuss the development and implementation of subtyping methods, including next generation methods which may yield useful data from mixed cultures such as enrichment broths. These meetings foster inter-governmental collaboration and coordination.

Section **Page**

Abbreviations

AES	Applied Epidemiology Staff
AFLP	Amplified Fragment Length Polymorphism
APHIS	Animal and Plant Health Inspection Service, USDA
AMS	Agricultural Marketing Service, USDA
ARA	antimicrobial resistance analysis
ARS	Agricultural Research Service, USDA
AST	antimicrobial susceptibility testing
CDC	Centers for Disease Control and Prevention
DNA	deoxyribonucleic acid
EHEC	enterohemorrhagic *E. coli*
ELISA	enzyme-linked immunosorbent assay
ERIC	Enterobacterial Repetitive Intergenic Consensus
FDA-CFSAN	Food and Drug Administration - Center for Food Safety and Applied Nutrition
FDA-CVM	Food and Drug Administration - Center for Veterinary Medicine
FSIS	Food Safety and Inspection Service, USDA
HUS	hemolytic uremic syndrome
ID	identification
Lm	*Listeria monocytogenes*
MALDI-TOF	matrix-assisted laser desorption/ionization time of flight
MDR	multiple-drug resistant
MLG	Microbiology Laboratory Guidebook
MLGT	multi-locus genotyping
MLST	multi-locus sequence typing
MLVA	multi-locus VNTR analysis
NACMCF	National Advisory Committee on Microbiological Criteria for Foods
NARMS	National Antibiotic Resistance Monitoring System
NCBI	National Center for Biotechnology Information
NM	non-motile
NRTE	not ready-to-eat
NVSL	National Veterinary Services Laboratories (APHIS lab in Ames, Iowa)
ODIFP	Office of Data Integration and Food Protection, FSIS
OFO	Office of Field Operations, FSIS
OIEA	Office of Investigation, Enforcement and Audit, FSIS
OPHS	Office of Public Health Science, FSIS
OPPD	Office of Policy and Program Development, FSIS
OSEL	Outbreaks Section of the Eastern Lab, OPHS
PCR	polymerase chain reaction
PFGE	pulsed-field gel electrophoresis
PHIS	Public Health Information System
PHLIS	Public Health Laboratory Information System

PR/HACCP	Pathogen Reduction/Hazard Analysis and Critical Control Point
RTE	ready-to-eat
SE	*Salmonella* Enteritidis
SNP	single nucleotide polymorphism
SPI	*Salmonella* pathogenicity island
ST	sequence type
STEC	Shiga toxin-producing *E. coli*
TMLGT	targeted multilocus genotyping
VNTR	variable number tandem repeat
WGS	whole genome sequencing
WHO	World Health Organization

<back

1 Introduction

1.1 Purpose

This White Paper provides detailed information about the biology and technology associated with subtyping procedures, the interpretation of subtyping data during outbreak investigations, and the potential uses of subtyping data generated by Food Safety and Inspection Service (FSIS) laboratories. FSIS managers use the information to enhance FSIS' food safety mission and to support the adoption of new procedures. The paper combines and updates two FSIS internal documents:

1. **"Proposed Protocol for the Use of Pulsed-Field Gel Electrophoresis (PFGE) Data from FSIS Isolates"** was a technical document created in 2006 which described how PFGE patterns are analyzed, how matches are identified and investigated, and how FSIS investigators interpreted data.
2. **"Analysis of Molecular Subtyping Methods for FSIS Regulatory Testing: The Present and Future of FSIS Regulatory Subtyping"** was created in 2008. It provided information about current subtyping procedures, how the data were used, methods in the immediate pipeline (projected implementation 2009-2013), and proposals for new methods for future work (potential implementation 2014 or later).

In addition, Withee and Dearfield discussed the potential advantages and challenges posed by genomic technologies to FSIS functional activities including outbreak investigation, microbial risk assessment studies and attribution studies (Withee and Dearfield 2007). The National Advisory Committee for the Microbiological Criteria of Food (NACMCF) also have discussed FSIS potential use of subtyping technologies (NACMCF 2010).

1.2 Scope

This White Paper focuses on subtyping procedures[4] for bacterial pathogens of primary concern to FSIS. Pathogens of concern include those specifically mentioned in FSIS

[4] Also referred to as methods and assays in this document.

regulations as well as those subject to testing, which cause foodborne illness from meat, poultry, or egg products. Recent regulatory changes include FSIS Pathogen Reduction-Hazard Analysis and Critical Control Point (PR/HACCP) verification testing for *Campylobacter* and *Campylobacter* performance standards (FSIS 2010) and verification procedures, including sampling and testing of beef trim and other raw ground beef products components for non-O157 STEC (FSIS 2011d). Sections 2 and 5 of this document provide a generic overview of subtyping methods and how subtyping data could be used by a regulatory agency. Sections 3 and 4 focus on the development of subtyping methods and the use of subtyping data to advance FSIS' mission.

1.3 Target Audience

There are three target audiences. The first is internal to FSIS. Staff and managers from specific program areas (Office of Public Health Science (OPHS), Office of Field Operations (OFO), Office of Data Integration and Food Protection ODIFP), Office of Policy and Program Development (OPPD), and Office of Investigation, Enforcement and Audit (OIEA)), as well as the Office of the Undersecretary for Food Safety, may wish to learn more about the biology and technology associated with various subtyping procedures, the interpretation of subtyping data during outbreak investigations, and the potential uses of subtyping data generated by FSIS laboratories. Another audience is FSIS partner federal agencies, including FDA, Centers for Disease Control and Prevention (CDC), and the Agricultural Research Service (ARS), which develop subtyping methods and may be interested in learning more about FSIS' data and technology needs. A third audience is FSIS stakeholders, including industry and consumer groups, who may be interested in learning about FSIS' present and future ability to generate, interpret, and respond to subtype data. This paper strives to present the information in a clear, concise, and non-technical language for all audiences. This version will be edited for submission to a peer-reviewed scientific journal.

<back

2 Background

This section provides a working definition for subtyping, and provides information about subtyping methods (including methods not used currently by FSIS) and procedures to evaluate subtyping method performance. This section also describes how subtype data is used in epidemiological investigations, and how it may be used to classify strains with virulence factors and to determine the familial relationships of bacteria. Lastly, it describes FSIS sampling programs and how FSIS generates, uses, and shares subtype data from its sampling programs.

2.1 Subtyping Definition

A bacterial subtype is a group of organisms with the same attributes (characteristics) within a larger type[5]. Subtyping procedures identify common attributes which assign it to the larger type and different attributes to distinguish it from other subtypes. Typically, the term is used to define groups below the level of bacterial species. Bacteria reproduce asexually when a parent cell divides to produce two daughter cells. The descendants of each daughter cell are considered a lineage, which means they are all descendants from a common ancestor. Mutations are errors in the deoxyribonucleic acid (DNA) sequence that accumulate in each lineage independently, which is referred to as vertical transfer. In bacteria, genetic material, with its associated mutations, also can be exchanged between lineages, a phenomenon referred to as lateral or horizontal transfer. Mutations distinguish individuals within and among lineages. “Subtype” refers to distinguishable entities within and between lineages.

<back

2.2 Subtyping Methods for Bacteria

A large number of procedures, spanning the history of classical and molecular microbiology, have been used for subtyping bacteria. Biochemical, metabolic, and serological attributes are measured using traditional methods, many of which are simple

[5] Subtype is also referred to as a group or clade in this paper.

to perform and very powerful. Growth assays are used to measure bacterial growth in the presence of a variety of compounds that stimulate or inhibit bacterial growth. These assays can be used to develop a comprehensive growth phenotype to identify and characterize bacterial isolates. For example, antimicrobial resistance analysis (ARA), also referred to as antimicrobial susceptibility testing (AST), is a specialized type of growth assay to evaluate resistance to specific compounds, including antibiotics. Another growth assay, based on the failure to rapidly ferment (digest) a specific sugar called sorbitol, is used to distinguish a pathogenic subtype of *E. coli*, O157:H7 from other pathogenic subtypes. Additional typing of *E. coli* subtype O157:H7 is based on its inability to ferment glucoronides. A variety of solid plating media were developed to identify strains that slowly ferment sorbitol and glucoronide. Phenotype microarrays (developed by Biolog) are used to screen bacteria simultaneously for growth in the presence of hundreds of compounds.

Serology is another commonly used method to type bacteria. An extensive set of antisera developed against somatic (O), flagellar (H), and virulence (K, Vi) antigens of *E. coli* and/or *Salmonella* allowed investigators to distinguish thousands of *E. coli* and *Salmonella* serotypes[6]. Serological methods still are used, and latex-coupled antisera are very useful for identifying the major serotypes of concern. However, typing for the lesser known serotypes is both time and labor intensive and maintaining a large panel of typing antisera is costly and not readily available to many labs.

Separation methods also are used to develop new bacterial subtyping methods. For instance, early evolutionary studies of bacteria lineages relied on separation of cellular enzymes using electric fields, a procedure called multilocus enzyme electrophoresis (MLEE). Newer technologies, such as mass spectrometry (developed by Abbott Labs and Sequenom), which can distinguish molecular fragments and estimate nucleic acid sequence, and Raman spectroscopy (developed by River Diagnostics), which is based on the scattering of monochromatic light (e.g., a laser) when it encounters cellular

[6]A serogroup is defined by the O-antigen, and a serotype is defined by a combination of O- and H-antigens

biomass (nucleic acids, carbohydrates, fatty acids, and proteins), are used to identify and subtype bacteria.

The growing importance in molecular biology in the 1970s and the invention of the polymerase chain reaction (PCR) in the 1980s led to the large number of molecular subtyping methods available today. These can be distinguished as fragment-based and sequence-based methods.

Fragment-based methods rely on the separation of PCR-amplified nucleic acid fragments by molecular size[7], such as amplified fragment length polymorphism (AFLP), Enterobacterial Repetitive Intergenic Consensus (ERIC)-PCR, allele-specific PCR, and multi-locus variable number tandem repeat (VNTR) analysis (MLVA). Other methods rely on enzymatic restriction of DNA fragments. These include PFGE, ribotyping, and optical mapping. Optical mapping of bacterial genomes (developed by Opgen) can detect large scale changes like inversions, insertions, deletions, and duplications, which are not detected with assays based on short sequences. Genome Sequencing Scanning (developed by PathoGenetix) is a rapid method based on DNA fragment size and the presence and location of short sequence repeats, which can detect subtypes directly from complex biological samples without the need for culture or primer design.

Sequence-based typing methods are based on the differences in short sequences (multi-locus sequence typing (MLST), SNP, and microarray) or entire genomes. Analysis of sequence data typically focuses on specific informative positions, which are a subset of all differences (Section 2.5). Solid phase or liquid suspension microarrays can simultaneously assay samples for presence or absence or relative expression level (e.g., the quantity of RNA in a cell) for hundreds to millions of genes or sequences, and to detect SNPs. Nucleic acid sequences are determined using a variety of chemical analyses, many of which have been developed in the past few years. Synthetic beads (microspheres) and novel chemistries are used to create nucleic acid libraries consisting

[7] Under the influence of an electric field, nucleic acid molecules, which carry a net negative charge, elongate and migrate through a sieve-like matrix at a rate that is inversely proportional to their molecular size. Larger fragments migrate through the field at a slower pace than smaller fragments.

of beads that contain 1,000,000 or more of an identical DNA fragment. DNA from 100 to 100,000 beads from these libraries can be sequenced or scored for SNP using flow cytometry (e.g., Luminex xMAP) or a series of charge-coupled device camera images (e.g., Roche 454 GS FLX). This process provides extremely robust data for sequence-based studies. Next generation sequencing methods are discussed in section 5.1.

Currently available subtyping methods require a culturally-pure isolate because the presence of multiple subtypes in the sample would create a confusing pattern or signal. Cultural isolation is time- and resource-consuming, and since a small number of available pure colonies are picked for further analysis, this step may preclude subtyping of slow-growing lineages, which can provide misleading data for foodborne pathogen surveillance programs (Singer et al. 2009). Importantly, manufacturers of diagnostic tests for infectious agents are developing formats that do not require cultural isolation therefore making it impossible to subtype strains from positive samples. This trend in diagnostic test design will have an impact on public health surveillance systems for bacterial enteric pathogens (Cronquist et al. 2012). New technologies, including metagenomics (Section 5.1), may lead to the development of methods capable of providing useful subtype data without the need for a cultural isolation step.
<back

2.3 Validating the Performance of Subtyping Methods

Validation is the process of evaluating the performance of methods, including subtyping methods. Guidelines for bacterial subtyping methods evaluation and validation have been published by the European Society for Clinical Microbiology and Infectious Diseases (van Belkum A. et al. 2007). The authors suggested the following performance criteria for evaluating subtyping methods:

1. Stability: the attributes targeted by the subtyping method should not change during lab culture
2. Typeability: the method should assign a subtype to all isolates
3. Discriminatory power: the method should distinguish (assign a different subtype to) isolates from different lineages

4. Epidemiological concordance (relevance): the method should group isolates associated with an outbreak or epidemic while excluding unrelated lineages, such as concurrent, sporadic case-patients
5. Reproducibility: the method should yield the same results when tested on multiple, independent occasions, preferably in different laboratories or by different operators

These criteria should be viewed as guidance, especially when evaluating the performance of a new subtyping method against an established method like PFGE. In addition to these criteria, practical aspects, such as cost of equipment and reagents, the level of technical expertise needed to operate the test, time to result, and compatibility with older data systems, should be considered before investing in a specific subtyping capability. The relative importance of each criterion should be weighed against specific informational needs and public health goals. For example, *Salmonella* serotyping would be intended to place strains into categories that indicate relative risk, whereas PFGE or MLVA would be intended to distinguish outbreak and non-outbreak case-patients. Therefore, PFGE and MLVA would be expected to have maximum discriminatory power and epidemiological concordance, but *Salmonella* serotyping would not. Some high frequency genetic mutations would be expected to affect some subtyping methods and not others. For example, the loss or gain of horizontally-transferrable DNA (*e.g.*, bacteriophage, insertion elements, and plasmids) in humans or animals, or during culture, can result in changes to a PFGE pattern, but may not affect a MLVA or SNP subtype.

Analysts performing subtyping to support epidemiological investigations should demonstrate proficiency with the methods through an externally and routinely administered program. For example, CDC PulseNet administers a proficiency program for PFGE and MLVA. Each analyst at Outbreaks Section of the Eastern Lab (OSEL) completes proficiency sample testing to be certified for the analyses that he or she performs.

Pathogen isolate panels are needed to assist in developing and validating subtyping assays. FSIS shared *L*. *monocytogenes (Lm)* and *E. coli* O157:H7 isolate panels with ARS researchers through an ARS-FSIS material transfer agreement, and FSIS *Salmonella* isolates were shared with Food and Drug Administration - Center for Food Safety and Applied Nutrition (FDA-CFSAN) through an ARS-FDA-CFSAN material transfer agreement. These collections are housed at the ARS facilities in Peoria, IL, and Clay Center, NE, and at a CFSAN facility in College Park, MD. The isolates could be shared with CDC and FDA researchers (and possibly other parties) for the purpose of developing and validating methods.

Organizations using subtyping data to support regulatory decisions should consider the standards for admissibility of scientific evidence. Under the current Federal standard, trial judges are responsible for monitoring the quality and relevance of scientific evidence. Judges should consider (1) testability, (2) peer-review, (3) error rate, and (4) general acceptance of a method or technique before finding that an opinion is more likely than not to be valid and reliable (Wilson et al. 2013).

<back

2.4 The Relationship of Subtypes to Virulence

Virulence is defined as the severity of the disease after infection, as measured by the portion of case-patients developing severe disease or dying after exposure. Subtyping can be used to estimate the potential virulence of a bacterial isolate from a food sample. If the subtype pattern of a bacterial isolate is indistinguishable (or very similar) to the pattern commonly isolated from ill people, it would be reasonable to conclude that the bacterial lineage represented by the pattern can cause disease and may be virulent.

The subtype may contain specific attributes (which may be genes related to the cellular mechanisms responsible for virulence) associated with virulent lineages. For example, *E. coli* isolates containing the Shiga toxin gene may be considered more virulent than those without the gene since the Shiga toxin plays a central role in destroying cells lining the gastrointestinal tract and in the kidneys, resulting in disease and death in

susceptible individuals. In some strains, multiple virulence factor genes are linked together in the genome as “pathogenicity islands.” *Salmonella* serotypes, which are based on the determination of surface molecules (O- and H-antigens), are not believed to be directly responsible for virulence. However, certain *Salmonella* serotypes are isolated more commonly from ill persons than other serotypes. According to CDC Public Health Laboratory Information System (PHLIS), the top 20 *Salmonella* serotypes from human sources accounted for 72.8% of all cases in 2009, while all other serotyped isolates accounted for 15.5% (CDC 2012a). Other DNA sequences, not directly linked to serotype or virulence genes, may similarly be used to identify virulent lineages of bacteria.

In some situations, pathogens with indistinguishable sequences or subtypes have been associated with multiple unrelated outbreaks over months and years, and across multiple jurisdictions and countries. These outbreaks, which are related by a common pathogen lineage, have been called epidemic or pandemic clones (Chen et al. 2007;Kathariou et al. 2006; Lomonaco et al., 2012; Pang et al. 2007;Prager et al. 1999;Ridley and Threlfall 1998) and could represent subtypes containing genes that act together to confer enhanced survival, transmissibility, or virulence. Careful study of these clones may identify individual genes associated with traits that greatly influence risk of illness or outbreaks, referred to as public health risk. Regulatory agencies may wish to perform enhanced surveillance for epidemic or pandemic clones, which may exhibit greater virulence and resistance to antimicrobials, or which may pose a higher risk of contaminating foods or causing illnesses or outbreaks.

<back

2.5 Using Subtyping Data to Determine Familial Relationships among Bacteria

Subtyping data refers to scientific data that is capable of distinguishing groups by directly or indirectly detecting mutations. Subtyping data can provide information about the underlying familial (evolutionary) relationship between clades. A clade is defined as a group of bacteria that have a common ancestor and therefore share mutations. By studying these mutations, we can not only visualize the clades, but also estimate the

familial relationships between clades, analogous to distinguishing first, second and third cousins. In a general sense, clades distinguished by a larger number of mutations are more distantly related. However, the association of mutations and familial relationships is complicated by a number of factors: 1) differences in the rate that mutations are created and incorporated into the population[8], 2) passage of mutations between clades (horizontally-transmitted mutations), and 3) very large mutations, called insertion and deletion mutations (indels), which are created by unpredictable events such as the gain or loss of horizontally-transferred DNA (e.g., bacteriophage, insertion elements, and plasmids).

Phylogenetic analysis is the study of familial relationships based on shared and distinguishable data. Phylogenetic analysis can be performed using any type of attribute, provided it is inherited and affected by mutation. To perform phylogenetic analysis, first the data must be properly aligned so that characters (e.g., sequences or PFGE fragment bands) from one lineage are properly compared to the equivalent, co-descending characters from the other lineages. Next, distance, maximum likelihood, or parsimony methods are applied to the aligned data to produce a similarity matrix which summarizes the occurrences of matching and non-matching characters among the compared isolates. A tree-like diagram (dendrogram) is created to visually represent the familial relationships among these data. A dendrogram is similar to a family tree. However, unlike a family tree, data from the most recent generation is used to estimate familial relationships among the older generations. Each point where the dendrogram splits is called a node and represents a hypothesized common ancestor for the clade. Members of a clade share more mutations than non-members and are thus believed to be more related to each other than to non-members. The cohesiveness of a clade can be tested by repeating the phylogenetic analysis hundreds or thousands of times (referred to as bootstrap analysis). The proportion of analyses leading to the formation of the clade is directly related to clade cohesiveness. The rarity and significance of the data can be evaluated using databases containing appropriate and comparable data

[8] Incorporation of cells with mutations into the population influences whether these mutations are "observed" among isolates characterized by scientists.

from the larger population of related organisms (e.g., the same species or serotype). Analysts would consider the total number of samples that fall within the clade compared with the total number that are excluded from the clade (Wilson et al. 2013).

PFGE and MLVA are not very useful for assigning familial relationships among bacterial clades. PFGE should not be used to determine strain relatedness. PFGE indirectly detects mutations by visualizing restriction enzyme fragments from the genome. Most observable differences in fragment size are caused by indels, many of which are horizontally transmitted (Kudva et al. 2002). Digestion with two or more enzymes enhances the discriminatory power of PFGE by increasing the number of fragments, but does not make the data more amenable to analysis of familial relationships (Zheng et al. 2011). Similarly, MLVA detects indel mutations in specific areas of the genome containing tandem repeats, which are subject to different mutation rates than the rest of the genome and therefore do not accurately represent the overall mutation rate of the genome. For this reason, PFGE and MLVA analyses are used to identify clades with identical or highly similar patterns. Using PFGE and MLVA together enhances the ability to discriminate clades (Broschat et al. 2010). Methods based on direct sequence analysis (MLST, multi-locus genotyping (MLGT), SNP, and whole genome sequencing (WGS)) are more accurate for measuring familial relationships because the mutations are more readily observed and predictable. However, like PFGE and MLVA, the accuracy of DNA or RNA sequence may be compromised by indels and horizontally-transferred mutations. MLST, MLGT, or SNP analysis typically relies on a limited number of mutations culled from all available mutations from the genomes to estimate familial relationships. Therefore, accuracy depends on whether the mutation rates in the chosen genetic regions truly represent the familial relationships among the clades. The same issues could arise when comparing whole genome sequences and therefore programs are designed to find mutations that accurately represent familial relationships between clades.

OSEL and other PulseNet labs use BioNumerics software to create dendrograms from subtype data. For PFGE data, the Unweighted Pair Group Method with Arithmetic Mean

(UPGMA) method is used to create a similarity matrix from the fragment size data. The dendrogram provides an estimate of "similarity," which can range from 0% (all different) to 100% (all the same). PulseNet considers PFGE or MLVA subtypes with 100% similarity to be indistinguishable, and subtypes assigned to the same pattern identification (ID) are referred to as a match. However, they may not be identical because these methods do not measure every potential difference among the matching isolates. Because foodborne illness investigations typically occur over a period of days or weeks, patterns related to a foodborne outbreak should be indistinguishable. In practice, this means that an outbreak clade may not vary by more than one or two PFGE bands, or have only a limited number of variations at MLVA repeat locations. This implies that a limited number of mutations would be likely within an outbreak-responsible clade. However, if the outbreak occurred over a long period of time or if the measured patterns had a high mutation rate, the investigator may consider less similar matches (Barrett et al. 2006)[9]. The degree of acceptable differences during an outbreak should be determined by observation for each organism and subtyping method. The degree of acceptable differences for *Salmonella*, *E. coli* O157:H7 and *Lm* PFGE patterns have been determined through experience over many years and hundreds of outbreak investigations. Expected differences in MLVA patterns are being determined because these methods are relatively new and mutation rates for the tandem repeat loci are different than PFGE patterns. Similarly, next generation methods such as MLST, SNP, and WGS will require similar determinations to be of maximum use during foodborne outbreak investigations (see Section 5.1).

<back

2.6 Laboratory Networks for Sharing Subtype Data: PulseNet and VetNet

In the 1990s, CDC organized PulseNet[10], the first inter-laboratory network for sharing PFGE pattern data for foodborne pathogens. The initial participants included CDC, several state health departments, and the major food regulatory agencies, including FSIS, FDA-CFSAN, and FDA-Center for Veterinary Medicine (CVM). PulseNet now

[9] An example is *Lm* that may persist in a production facility over a period of time that allows for the generation of distinct lineages, any of which may contaminate product leading to illness.
[10] See CDC Web site for more information: http://www.cdc.gov/pulsenet/pathogens.htm

includes all 50 state health departments and a number of local health departments. CDC PulseNet staff maintains databases of patterns submitted by member laboratories, administers a quality assurance program, and develop and validate new protocols and methods. PulseNet networks have been established in Canada, Europe, the Asia Pacific region, and Latin America (Gerner-Smidt et al. 2006;Swaminathan et al. 2006).

PulseNet PFGE procedures and databases have been developed for *E. coli* O157:H7, non-O157 Shiga toxin-producing *E. coli* (STEC) serogroups, *Salmonella* serotypes, *Lm*, *Campylobacter*, and other bacterial pathogens of public health importance. PulseNet member laboratories perform PFGE analyses, process the raw data in BioNumerics, a common software platform that allows for pairwise comparison of patterns generated in different laboratories (Section 3.1.1). CDC staff manages the PulseNet databases and provides identifications for each pattern uploaded by a member lab.

VetNet was developed by ARS to house separate PFGE and antimicrobial resistance (NARMS) databases for *Salmonella* and *Campylobacter* isolated from diagnostic animal specimens, healthy farm animals, carcasses of food-producing animals at slaughter, as well as ground beef, chicken, and turkey products (Jackson et al. 2007). PulseNet and VetNet databases can be searched remotely by CDC, ARS, and FSIS staff. By the end of the 2013 calendar year, OSEL assumed several subtyping duties currently/formerly performed by VetNet, including *Salmonella* and *Campylobacter* PR/HACCP PFGE analysis.

<back

2.7 FSIS' Sampling and Subtyping Programs

FSIS conducts microbiological sampling to assess the effectiveness of industry process controls; provide critical feedback to industry; monitor compliance with performance standards and zero tolerance policies; monitor industry-wide trends; serve as an incentive to reduce the occurrence of pathogens in product; identify the origins of contaminated products that are of public health concern; and capture pathogen PFGE patterns for evaluation in PulseNet. FSIS collects samples at federal establishments,

U.S. ports of entry, in-commerce facilities (e.g., warehouses, distributors, centers, retail firms), and the homes of consumers. The sampling rate at federal establishments is random, volume-adjusted, risk-adjusted, or directed as a follow-up to a positive sample. Sampling of imported products is performance based[11] or as a follow up to a positive sample (FSIS 2011c). FSIS also samples meat, poultry, and egg products associated with outbreak investigations and consumer complaints when product is available. In addition, FSIS receives isolates of *Salmonella*, *Lm*, and *E. coli* O157:H7 referred by state meat and poultry inspection programs and the National School Lunch Program commodity-testing program administered by the Agricultural Marketing Service (AMS). In fiscal year 2010 (Oct 1, 2009 to Sept 30, 2010), FSIS collected 90,678 samples which were tested for either *E. coli* O157:H7, *Lm*, or *Salmonella*[12]. Of these, 1,298 (1.43%) samples were positive[13].

All FSIS testing procedures include cultural isolation and characterization steps, and the isolates are identified to the species level. In addition, one or more isolates from each sample are subtyped.

- All *Lm* isolates are subtyped using two-enzyme PFGE.
- All *E. coli* O157:H7 isolates are subtyped using two-enzyme PFGE and MLVA. Under certain circumstances, *E. coli* O157:H7 are PCR-typed for genes associated with H7 ($fliC_{H7}$) as well as for the presence of Shiga toxin (*Stx*) and intimin (*eae*) genes.
- All *Salmonella* isolates are subtyped using serology, and one- or two-enzyme PFGE (for raw and ready-to-eat (RTE) isolates, respectively). ARA is performed on some *Salmonella* isolates, while some *S.* Typhimurium isolates are subtyped using phage typing[14] or MLVA.

[11] Performance based means that countries with a poor food safety performance record are sampled more frequently than countries with strong food safety performance records.
[12] 17,453 samples (59 positive) were tested for *E. coli* O157:H7, 25,600 (225 positive) samples were tested for *Lm*, and 47,625 samples (1,014 positive) were tested for *Salmonella*
[13] Source provided from the FSIS Data Warehouse on Feb 2, 2011.

[14] Phage-typing is a procedure for characterizng and detecting bacterial strains by their reaction (susceptibility or resistance) to various known strains of phages. Bacteriophages are specific in action.

In recent years, FSIS has added subtyping capabilities, such that MLVA is performed on all *E. coli* O157:H7 isolates and on *S*. Typhimurium by special request during investigations. CDC performs MLVA on *S*. Enteritidis by request from FSIS program staff (requests are typically made through OPHS headquarters). Since July, 2012 OSEL has performed molecular serotyping on *Salmonella* isolates following the CDC method as a replacement for the traditional serotyping method (FSIS 2013a) (Section 3.5).

Several subtyping duties previously performed by VetNet and NARMS-Animal Arm are currently being conducted by OSEL. These include PFGE and ARA laboratory and data analysis for all FSIS *Salmonella* and Campylobacter positive samples obtained through the PR/HACCP sampling program.

<back

2.7.1 Antimicrobial Resistance Analysis (ARA) for *Salmonella*

Through 2013 the NARMS-Animal Arm of ARS performed antimicrobial susceptibility testing on all FSIS *Salmonella* isolates and maintained a database of the resistance information. The FSIS Eastern Laboratory has assumed these duties. The panel of antimicrobials tested is representative of common antimicrobials used in both humans and animals.[15] Testing is done using a semi-automated system (Sensitire®, Trek Diagnostic Systems, Inc., Cleveland, Ohio).

2.8 FSIS' Current and Potential Use of Subtyping Data

Subtype data may be used to:

1. Detect and track outbreak clusters or harborage within an establishment;
2. Assess the sample-specific risk of illnesses or outbreaks associated with a product, process, or establishment ;

Susceptibility to lysis by a particular phage may be the only apparent phenotypic difference between two bacterial strains.

[15] Breakpoints Used for Susceptibility Testing of *Salmonella* 1997-2012. http://www.ars.usda.gov/Main/docs.htm?docid=6750&page=3**).**

3. Provide highly specific data for risk assessments and attribution studies;
4. Define a regulatory adulterant with greater accuracy; or
5. Improve the accuracy and efficiency of current subtyping methods.

Detect outbreak clusters: Isolates sharing the same subtype form a cluster. If the isolates are from case patients, the subtype cluster could indicate an outbreak, especially if one or more of the case patients were related, lived in the same area, ate in the same restaurant, or shared any type of activity, which is referred to as a common exposure. Outbreak clusters may be identified through subtype surveillance before a large number of illnesses are identified, especially important when case-patients are located in different cities or states.

Outbreak investigators often use subtype cluster data to create two groups of people: case-patients who are ill and share the cluster subtype and those who are healthy or who are ill and do not share the cluster subtype. Subtyping data allows investigators to exclude case-patients that do not fit in the cluster. Such sporadic, non-outbreak cases could weaken the association of the illness with an exposure.

Networks such as PulseNet have proven invaluable for identifying widely disseminated clusters with no known evidence of common exposure, actively finding new case-patients during an investigation, and detecting the specific pathogen in environmental samples including foods, surface swabs, water and air. In this way, subtype data can promote timely outbreak investigations and interventions to prevent further illness.

Detect harborage within an establishment: Subtype clusters which include food and other environmental isolates assist food regulatory agencies by providing support for a specific transmission route, or the persistence of a foodborne pathogen within an FSIS regulated establishment. For example, a subtype cluster detected from samples collected over days, weeks or months from samples collected at a single establishment could support the existence of a contaminated product contact surface that is not being properly cleaned and sanitized, or which contains a biofilm that is resistant to the

establishment's cleaning and sanitizing procedures. If the cluster subtype is detected from raw (e.g., uncooked) and processed (e.g., cooked) product samples, the data could indicate the transfer of the pathogen from raw to RTE operation areas, suggesting improper separation of these operations. If the cluster subtype is detected in the finished product and individual ingredients, the data suggest that the ingredient supplier is unable to control the pathogen. Together, these observations, obtained through careful study of subtype data, suggest modifications to operating procedures, such as HACCP plans, prerequisite programs, or Sanitary SOPs

Assess public health risk: Subtyping methods can provide information on the relative probability of a subtype to cause illness, which could be a component of an overall estimate of public health risk. Typically, this is not a rigorously calculated ratio, but rather an assessment of the relative frequency by which the subtype is recovered from outbreak and non-outbreak settings, or finding the subtype associated with multiple independent outbreaks. FSIS currently assesses the relative public health risk of establishment production practices and products based on the use of interventions, production volume, and testing history. FSIS uses *Salmonella* serotype to help determine the *Salmonella* verification testing set sampling frame for raw meat and poultry products and to schedule food safety assessments. FSIS is actively working to include PFGE and antimicrobial susceptibility profiles in its "end of set" letters sent to establishments at the end of a *Salmonella* verification set. However, the Agency has not established a regulatory standard for *Salmonella* subtypes, or formally defined public health risk with respect to pathogen subtypes.

Provide data for risk assessments and attribution studies: Some risk assessments deal with strain (subtype) variation by pooling data from multiple strains, or modeling a worst-case strain, with the intention of providing conservative risk estimates. However, models that ignore potential strain effects could provide a misleading representation of both the public health implications and the confidence that certain risk management strategies would be effective in reducing risk (Coleman et al. 2004). Biological variation among pathogen species of concern to risk assessments and attribution studies include

virulence, dose response, environmental reservoir, growth and inactivation kinetics in foods, and resistance to various food processing procedures. Subtyping methods can distinguish existing biological variation within a pathogen species, either by detecting genes directly controlling such variation or detecting marker sequences that are strongly associated with subpopulations with different properties. When information is available about the degree of between-strain differences, then adjustments can be made to increase the predictive accuracy of risk assessments (Juneja et al. 2003). Subtyping methods support some attribution studies by characterizing the distribution of specific subtypes isolated from different sources (e.g., animals, foods, processing environments) and comparing those subtypes with those isolated from humans. Subtypes that are exclusively, or almost exclusively, associated with a single source are referred to as "indicator subtypes," and are used to assign human illnesses to specific sources (Pires et al. 2009). This approach was taken with *Salmonella* serotypes (Guo et al. 2011;Pires and Hald 2010). Subtyping methods with greater discriminatory power could be more likely to detect indicator subtypes, and therefore improve the accuracy of certain attribution studies.

Define a legal adulterant: The FSIS Federal Meat Inspection Act (FMIA) and Poultry Products Inspection Act (PPIA) define a food as adulterated if it bears or contains a poisonous or deleterious substance that may render a food injurious to health. If the substance is a natural constituent of the food, the food is not considered adulterated if the quantity of the substance does not ordinarily render the food injurious to health, (Title 21 U.S.C Chap. 12, Sec. 601 (m) (1) and 21 U.S.C. Chap. 10, Sec. 453 (g) (1)). FSIS can use data from subtyping assays to interpret the meaning of the acts and establish science-based policies. FSIS may use subtype data to determine that certain subtypes with higher virulence are injurious to health, or to determine that subtypes with lower effective doses are "ordinarily injurious" even when present as a natural constituent of a food. For example, in 1994, FSIS announced its intention to declare raw ground beef containing a single *E. coli* subtype (serotype O157:H7) to be adulterated, and in 2012, expanded this restriction to include additional subtypes of non-O157 STEC. FSIS laboratory methods use traditional serology, biochemical methods (sorbitol

fermentation), and virulence gene detection using real time PCR to define these subtypes from other *E. coli* subtypes. In the future, FSIS may use additional subtyping methods, including molecular methods, to define subtypes of *Campylobacter*, *E. coli*, *Lm*, or *Salmonella* as adulterants. For example, FSIS is considering a petition to declare certain antibiotic resistant serotypes of *Salmonella* to be adulterants (CSPI 2011)].

Improve accuracy and efficiency of current subtyping methods. FSIS labs use PCR-based tests to identify serotypes of *E. coli* O157:H7, non-O157 STEC, and *Salmonella*. These molecular methods reduce FSIS' reliance on antibody reagents, which are expensive to maintain and which are not effective unless the target antigen is produced and located on the surface of the pathogen. With *Salmonella*, molecular serotyping assays reduce time to result from weeks to days, providing actionable data to the field workforce.

<back

3 Current Use of Subtyping Methods by FSIS

This section describes the molecular subtyping methods currently used by FSIS and the ways in which the Agency communicates and interprets the data.

3.1 Pulsed-Field Gel Electrophoresis (PFGE)

PFGE is a DNA-based subtyping method for characterizing bacterial isolates. It is the primary method used by all PulseNet laboratories. Within FSIS, OSEL is responsible for performing PFGE. PFGE is easily applied to different species and generally yields a high amount of pattern diversity that provides good discriminatory power and other measures of subtyping method validity (Section 2.3). PulseNet laboratories use standardized PFGE protocols and computer-assisted pattern normalization techniques, which result in high levels of reproducibility within and between laboratories (Section 2.6).

Isolates with indistinguishable PFGE patterns are more likely to be genetically related, but a PFGE pattern is not as specific as a fingerprint. Relatedness of patterns should be

used as a guide and not a true measure of relatedness. PFGE is one piece of data that should be used in conjunction with epidemiologic evidence or other information collected during investigations (Section 3.1.5).

3.1.1 PFGE Laboratory Method and Pattern Normalization

The PFGE theory and the laboratory method are described in Appendix A and Appendix B respectively. The characteristic of DNA that makes PFGE possible is the restriction site. Restriction sites are DNA sequences that are recognized by one or more restriction enzymes that cut the DNA molecule only at that particular target sequence. Restriction enzymes used for PFGE cut the bacterial genome at multiple sites. Digestion of bacterial DNA with a restriction enzyme cuts all restriction sites, producing DNA fragments of varying sizes. The digestion step is done in a solution called a “restriction digest” that contains genomic DNA, buffer, and the restriction enzyme. The resulting fragment sizes depend on the spacing of the restriction sites along the DNA molecule. The restriction digest containing the fragments from one to 12 isolates are loaded onto an agarose gel along with two to four restriction digests from known strains (referred to as size standards) which contain DNA fragments producing an expected array of fragment sizes. The gel is exposed to an electric field that is constantly changing directions and polarity. Under the electric field, the DNA fragments become arranged in a characteristic pattern, with the large fragments located on the top of the gel and smaller fragments at the bottom. The fragments are stained to produce “bands” on the gel. The PFGE pattern is a characteristic of the subtype. Different restriction enzymes recognize different restriction sites, and therefore produce a different array of fragments when applied to the same isolate.

PFGE data workflow is described in Appendix C. A digital image of the gel is saved to a local computer and analyzed using BioNumerics, a computer program that normalizes and saves PFGE images and transfers images to and from the national PulseNet database at CDC. The normalization step adjusts the image to a standard PulseNet format allowing it to be compared to images from other laboratories. During

normalization, the analyst defines the portion of the gel image area occupied by each isolate (referred to as a lane). The bands are marked in the lanes occupied by the standards, which links them to PulseNet-defined band sizes saved in the BioNumerics program. BioNumerics can perform initial band marking automatically, but this often results in errors, so an OSEL analyst verifies the results for accuracy. The analyst then marks and saves the position of the isolate bands in a file, links names to each lane, and adds additional information (metadata) about the sample as described in Appendix D. Certain metadata are saved by FSIS and not shared with PulseNet or VetNet. Other metadata that reasonably could not be expected to identify a specific establishment are uploaded to PulseNet or VetNet along with the gel image, so that the pattern can be analyzed by CDC-PulseNet database managers and can receive a PulseNet name, which is referred to as pattern confirmation. Processed PFGE images also are available for immediate use within FSIS, while still unconfirmed (unnamed) by CDC-PulseNet. OSEL typically analyzes unconfirmed patterns because there could be some delay in the process of pattern confirmation. To avoid confusion, it is important to distinguish confirmed and unconfirmed patterns in FSIS communications. FSIS's use of confirmed and unconfirmed PFGE patterns and matches is described in Appendix E

To facilitate comparisons of PFGE patterns between laboratories, CDC developed a standardized naming system for designating patterns in PulseNet. Each unique pattern in the database is represented by a 10-character code (e.g., XXXYYY.0000). The first three characters in the code represent the bacterial pathogen, the next three characters denote the enzyme used for DNA restriction, and the last four characters represent the pattern designation. For example, in the pattern designation EXHX01.0026, EXH represents *E. coli* O157:H7, X01 represents restriction endonuclease *Xba*I, and 0026 is the pattern number. Because the pattern numbers are assigned sequentially to unique patterns, no similarity should be implied from the order of pattern numbers. The CDC PulseNet laboratory confirms the pattern ID to ensure consistency. Because pattern confirmation is subjective, pattern names may be changed occasionally by the CDC PulseNet laboratory without informing other PulseNet member laboratories, including

FSIS. Some PulseNet laboratories, and the VetNet laboratory, assign a lab-specific name to their patterns to serve more local needs.

Although PFGE is a powerful tool, there are a variety of considerations involved with its use. Data interpretation is not entirely objective. Analysis is labor-intensive and takes about two or three days after bacterial isolation to complete. Occasionally, isolates are not typed easily and may require multiple attempts to produce a usable pattern. Rigorous standardization of the protocols is necessary to enable inter-laboratory comparison of the PFGE patterns. Even with the aid of image analysis software, the analysis becomes increasingly cumbersome, subjective, and difficult as more profiles are entered into PFGE databases (Hyytia-Trees et al. 2007).

<back

3.1.2 FSIS PFGE Data Analysis and Data Monitoring

FSIS and local, state, and other federal public health agencies contribute PFGE pattern images from human clinical, food, and environmental isolates to the national PulseNet database located at CDC. These entities assist the CDC in monitoring the database for matching patterns. CDC maintains separate databases for *E. coli* O157:H7, non-O157 STEC, *Lm*, *Salmonella*, and other microbial pathogens. The *Salmonella* pattern database is the largest within PulseNet (the number of profiles exceeds 110,000 and contains profiles of approximately 500 serotypes) (Gerner-Schmidt, 2006). Therefore, it is managed slightly differently. Isolates within the *Salmonella* database are classified first by serotype and then pattern ID. Occasionally a PulseNet database manager may confirm a *Salmonella* isolate before the serotype is determined.

Each isolate uploaded by OSEL is digested separately with a primary and secondary enzyme, so two patterns (a primary and secondary pattern) are associated with each isolate. While all PulseNet laboratories submit a primary pattern for each isolate, some may not submit a secondary pattern for all organisms. For *E. coli* O157:H7 and *Lm*, most laboratories immediately submit a primary and secondary pattern. Due to the large number of *Salmonella* isolates, most laboratories submit a primary PFGE pattern for the

top five to 10 serotypes in their state and both a primary and secondary PFGE pattern for isolates associated with outbreak investigations. It is not feasible for most laboratories to PFGE all received *Salmonella* isolates. Within PulseNet, database managers confirm each PFGE image. Isolates with a matching pattern are given the same pattern name. During multistate outbreak investigations, PulseNet may request a secondary pattern from a laboratory if one was not initially submitted. When both patterns are available, isolates are considered to match only if the confirmed pattern names for both the primary and secondary patterns are identical.

In PulseNet, PFGE patterns are compared visually and with the use of BioNumerics software, which is especially helpful for comparing large numbers of patterns. A similarity matrix is used to create a dendrogram to visually depict the relative strengths of all pair-wise pattern comparisons. The dendrogram also provides an estimate of pattern "similarity," which can range from 0% (all different) to 100% (all the same). PulseNet considers PFGE patterns with 100% similarity to be indistinguishable and assigns the same pattern ID. These patterns are considered to match.

A variety of database searches can be performed before and after the pattern ID is confirmed by CDC-PulseNet, including “hotlist search,” “search to server,” “plant comparison,” and “cluster search,” as described below. Database search results and metadata are stored on FSIS servers, which are accessible to OPHS investigators.

Hotlist Searches compare an FSIS pattern (referred to as the query) to patterns recently uploaded to PulseNet (referred to as the Hotlist). OSEL usually will perform a hotlist search before CDC PulseNet confirms the pattern ID. The BioNumerics program is used to determine the relatedness of the query and hotlist patterns, which is visualized as a dendrogram and does not rely on CDC PulseNet pattern confirmation. OSEL determines the PulseNet upload dates in the hotlist and the dates range from 30 to 90 days, with most searches performed against a 60-day hotlist.

Search to Server Searches compare a query pattern to the most closely matching patterns in the entire PulseNet database (i.e., not limited to recent isolates like the hotlist search). If the query matches a confirmed pattern, a combined pattern search is performed to provide a list of the most likely recent matches.

Plant Comparisons compare all patterns derived from the same FSIS-sampled establishment. The BioNumerics program determines the relatedness of the patterns. The output of this type of search is a dendrogram and is accompanied by a spreadsheet that contains all demographic information that OSEL has on the isolates.

Cluster Searches are specific comparisons of an isolate to an active cluster (a group of isolates with a similar pattern that has elicited interest at the CDC). CDC-PulseNet assigns PulseNet cluster codes to these groups of isolates that are of heightened interest due to a number of factors (e.g., increase in frequency above baseline, specificity to a location, epidemiologic link, and recently uploaded food isolate).

While the searches detailed above can be enhanced by PulseNet confirmation of an isolate's pattern, certain searches can be performed *only after* PulseNet confirms the pattern. These searches include combined pattern search, frequency determination, and *Salmonella* pattern and serotype ranking within PulseNet.

Combined Pattern Searches provide a list of isolates that have the same confirmed pattern combination as the query. These searches can be limited to a specific time or include all matching isolates in the PulseNet database.

Frequency Determination provides the frequency of a confirmed primary and/or secondary pattern ID and/or the combination of primary and secondary patterns. Because of the size and complexity of the PulseNet *Salmonella* database, this type of search is difficult and less informative for *Salmonella* isolates, which is why a *Salmonella* pattern and serotype ranking within PulseNet often is requested.

***Salmonella* Serotype and Pattern Ranking within PulseNet** is based on the PulseNet database managers' lists of the 30 most common serotypes and the 5 most common patterns within each *Salmonella* serotype in the PulseNet database. After pattern confirmation, OSEL can report the serotype and pattern rank *Salmonella* if present on the list.

Additional searches may be performed for patterns associated with specific text in the PulseNet database, for example, all *E. coli* O157:H7 patterns submitted by the Ohio PulseNet lab and uploaded to the database in 2010.

Under OSEL's current practice, either a hotlist search or search to server is performed before the pattern is confirmed. A plant comparison always is performed, and a combined pattern search is performed on patterns from all *Lm*, *E. coli* O157:H7, and non-O157 STEC isolates. *Salmonella* isolates are monitored using cluster searches. OPHS investigators and the FSIS liaison to CDC receive routine search results for *Lm*, *E. coli* O157:H7, and non-O157 STEC, or when a *Salmonella* isolate matches a recent PulseNet cluster. If matches to clinical isolates are revealed, investigators in OPHS and other program areas discuss further actions (Section 3.1.4). If a potential match is supported by additional epidemiologic data, OPHS may request that trace forward and/or trace back investigations be initiated to determine whether case-patients were exposed to the product. OPHS investigators may request additional and repeated searches (Section 3.1.4).

Relative pattern frequency is calculated for all primary, secondary, and combined PFGE patterns uploaded to the PulseNet database. Pattern frequency is used to judge the significance of matches. The combined pattern frequency of *Lm* and *E. coli* O157:H7 isolates can be assigned to one of five descriptive terms: common, less common, rare, very rare, and new to the PulseNet database, as described in Appendix F
<back

3.1.3 PFGE on FSIS PR/HACCP *Raw Product Salmonella Isolates*

Prior to 2014 *Salmonella* isolates collected from raw products through the PR/HACCP *Salmonella* Verification Testing Program were not uploaded to the PulseNet database. The ARS Bacterial Epidemiology and Antibiotic Resistance laboratory (co-located with the FSIS Eastern Lab in Athens, GA) typed these isolates by PFGE and uploaded each pattern to the VetNet database. OSEL assumed these duties before the conclusion of the 2013 calendar year; this change has resulted in the inclusion of these isolates into the PulseNet database.

Prior to 2011, VetNet only routinely performed PFGE on *Salmonella* isolates using a single enzyme (*Xba*I). Since 2011, every isolate uploaded to the VetNet database was digested separately with two enzymes, *Xba*I and *Bln*I. The PulseNet and VetNet databases are distinct. However, database managers for each database and OSEL staff have access to both databases, so there is a high level of cooperation and communication. All VetNet isolates are assigned a VetNet pattern ID and the equivalent PulseNet ID, if there is a matching pattern in the PulseNet database, so a list of VetNet isolates matching a PulseNet pattern of interest can be downloaded using a pattern search.

<back

3.1.4 Distinguishing PFGE Patterns—Linking Clinical and Product Isolates

PFGE patterns from strains originating from the same source are more likely to be identical or highly similar compared with strains that do not originate from the same source. If CDC has determined that two isolates share the same PFGE pattern, the isolates are said to ‘match.’ A group of such matches is referred to as a cluster. By continuously monitoring PFGE patterns from foodborne pathogens isolated from humans, it is possible to detect clusters of indistinguishable subtypes that, upon epidemiological investigation may turn out to be outbreaks (illnesses linked to a common source) (Barrett et al. 2006;Hyytia-Trees et al. 2007).

Detection of PFGE pattern matches between clinical and product isolates is typically insufficient to link an illness to a specific product. In many situations, such matches are detected first, followed by corroborating epidemiologic, food history, and trace back data, gathered by FSIS and public health partners. When PFGE pattern clusters containing FSIS-regulated product isolates and clinical isolates are detected, FSIS must decide whether to initiate an investigation to collect additional information necessary to determine if an association exists. This process is described in Appendix C. Clusters that include recent clinical isolate uploads (indicating an ongoing occurrence of illnesses), especially when there is an appropriate temporal (time) and spatial relationship between the product and clinical isolates, are of greatest concern.

Several considerations are necessary in making the decision to investigate clusters involving both FSIS-regulated product and clinical isolates with no known epidemiologic association (Figure 1).

Figure 1. Considerations made by FSIS before Investigating a Cluster

1. CDC designation of a PFGE pattern cluster as an outbreak cluster or knowledge of an ongoing investigation being conducted by local, state, or federal public health officials
    a. Number of clinical isolates
    b. Continued occurrence of illnesses
    c. Severity of disease outcomes
2. Pathogen of concern
3. PFGE pattern frequency
4. Type of FSIS-regulated product (e.g., RTE, not ready-to-eat (NRTE), raw)
    a. Whether product represented by the sample was held or distributed into commerce
    b. Temporal sequence and spatial relationship between the FSIS-regulated product and case-patient isolates
5. Interest or concerns expressed by the FSIS Management Team and public health officials

Pattern frequency is used to help judge the significance of pattern matches. It is important for FSIS staff to consider pattern frequency when deciding whether to commit FSIS resources to an investigation of a pattern match between FSIS regulated products and clinical isolates. Other information should be considered as well, such as evaluating the recent history of the pattern, because a large outbreak can significantly alter frequency statistics in subsequent years. For example, if a specific pattern was observed 100 times during a large outbreak in 2005, but not seen at all in following years, its cumulative frequency still would be high even though the pattern was infrequent in the subsequent years. In addition, a match involving a common pattern should not be disregarded if it has not been seen in the PulseNet database recently, if there is a spike in pattern frequency, or if public health partners are investigating the cluster. Surveillance of patterns for less common serotypes is important to identify the emergence of new pathogens of concern.

After identifying a *Salmonella* illness cluster In PulseNet that may be associated with meat or poultry products, OSEL will provide OPHS investigators with information on FSIS isolates with indistinguishable PFGE patterns. This information may assist the epidemiologists in identifying possible sources of illness in those investigations. Investigation of specific PFGE matches between clinical *Salmonella* isolates and raw product isolates is usually reserved for instances where the isolate pattern is part of a CDC active cluster of human illness with a designated PulseNet cluster code[16] (Section 3.1.2). The Applied Epidemiology Staff (AES) Surveillance Team also follows up on combined pattern searches conducted by OSEL when the PFGE pattern is novel or rare or the serotype is not ranked in the top 5 in humans (as determined annually by the CDC). Investigations of *Salmonella* clusters may also be initiated when the isolates are found to exhibit antibiotic resistance, especially to an important drug class (e.g., cephalosporins or fluroquinolones) used for treating salmonellosis in the United States.

[16] A group of isolates with indistinguishable PFGE pattern that are of heightened interest due to a recent increase in frequency, specificity to a location, epidemiologic link or recent upload of a food isolate are often assigned a distinct PFGE cluster code by CDC/PulseNet staff

Antibiotic resistance may result in treatment failures and increased morbidity and mortality.

When there is limited evidence that FSIS-regulated product is (or is not) the source of a cluster of illnesses, AES monitors the PFGE cluster. Once there is evidence of possible association of an FSIS-regulated product with human illness, a 'watch' is opened and the AES Investigation Team actively collects epidemiologic data from public health partners in order to investigate potential linkage between case-patients and the FSIS-regulated products. According to the FSIS foodborne illness directive 8080.3 (FSIS 2008a), when case-patient food histories indicate the likely involvement of FSIS-regulated product, an official foodborne disease investigation is initiated, an Incident Report (IR) is opened in FSIS's Foodborne Incident Management System (FIMS), and other FSIS program areas, such as OFO and OIEA, become actively involved in the investigation. OPHS continues to collect and analyze epidemiological data, such as information about the exposure of case-patients to specific FSIS-regulated products, and microbiological data. OFO and OIEA conduct trace forward and trace back investigations based upon documented case-patient exposures (e.g., product type/brand, store location, date of purchase) to assist in the identification of the source (e.g. food or other exposure) of illness. During foodborne disease investigations, FSIS requests combined pattern searches or cluster searches to detect newly-uploaded isolates at least weekly, or as needed, until the investigation is closed.

If FSIS-regulated product is implicated and recalled, the PFGE pattern is monitored using hot list or combined pattern searches to assure the adequacy of the recall and to expand the scope of the recall as necessary. Additional clinical isolate uploads to PulseNet with the same PFGE pattern as the recalled product may indicate that the product is linked to those illnesses necessitating further investigation. The AES Surveillance Team, with the assistance of OSEL, also monitors FSIS recalls triggered by positive FSIS product samples without reported human illness. This is referred to as a 'recall watch.' During the recall watch, AES requests follow-up pattern searches for *E. coli* O157:H7 positive isolates approximately every two weeks for at least 60 days

following the recall date. *Lm* follow up searches are requested approximately every two weeks for at least 90 days following the recall date to ensure that illness incubation periods are addressed appropriately. If a new clinical isolate upload to PulseNet matches the cluster pattern, OPHS investigators typically will discuss further actions.

Under most circumstances, cluster investigations involve product isolates where the product yielding the positive isolate with an indistinguishable pattern was produced prior to the clinical isolate collection date. However, this expected temporal sequence may not apply if a portion of common source material (e.g., intact boneless beef) was distributed and used as ingredients in other products (e.g., ground beef) which were subsequently sampled, or if there is evidence of harborage at a producing facility over an extended period of time. To account for long-term harborage *Lm* in production facilities, *Lm* PFGE clusters should be investigated even in situations where the positive product has been held and where clinical isolates were identified prior to the date that the *Lm*-positive product was manufactured. Investigators should assess the trends in PFGE pattern frequency prior to the product isolation date. *Lm* clusters with identified product or environmental isolates should be followed 90 days before and 90 days after the isolation date. Such follow-ups may suggest harborage issues, especially when an outbreak was previously associated with the same PFGE pattern.

<back

3.1.5 Using PFGE in Outbreak Investigations

PFGE is one of several important tools used by FSIS to identify links between FSIS-regulated products and foodborne illnesses and outbreaks. PFGE data may be used in conjunction with microbiological and epidemiologic data, environmental assessments, trace back activities, and other information collected by FSIS and its public health partners during foodborne disease investigations. The success of PFGE cluster investigations relies on timely surveillance, detection of clusters, and the ability to collect adequate information for trace back activities. PFGE analysis is especially valuable in identifying outbreaks in which the illnesses are not clustered geographically, and, as such, would often be considered sporadic cases. PFGE data combined with

epidemiologic data, food consumption history, trace back information, and point of preparation data are used to establish an association between ill individuals and contaminated product.

PFGE patterns generated from FSIS samples and uploaded to the PulseNet or VetNet databases are compared with other patterns, with the goal of detecting illness clusters that may be associated with FSIS-regulated products. When clinical isolates of foodborne pathogens with the same pattern combination are found to increase in frequency within a specific time-frame, referred to as a temporal cluster, an epidemiological investigation can determine whether such a cluster represents an outbreak. If a pattern from an FSIS isolate matches a recently uploaded clinical isolate or recently identified cluster, OPHS collaborates with state and/or local public health partners to determine whether the match represents an outbreak of foodborne illness from a single exposure (e.g., exposure to a specific production lot of a product), often referred to as a point source outbreak.

Epidemiologists use PFGE data to focus their working hypotheses during the course of the outbreak investigation. Investigators expect subtypes from clinical isolates recovered from a single source outbreak to be indistinguishable (high similarity) and subtypes of clinical isolates recovered from different outbreaks should be distinguishable (low similarity). In the absence of epidemiologic information, analysts may conclude that isolates sharing one of the less common patterns are more likely to have come from a common source than isolates sharing a predominant pattern.

The interpretation of PFGE patterns in the context of foodborne disease epidemiology, however, is not simple, and guidelines for interpreting this data are needed. Method reproducibility with a particular organism, the quality of the PFGE gel, the variability (also referred to as the clonality) of the organism being subtyped, and the pattern prevalence in question must be considered. Over-reliance on conclusions from subtyping data in the absence of (or despite) epidemiological data may result in a type I error (false negative conclusion), leading to the premature termination of an

investigation or regulatory action. It also may lead to a type II error (false positive conclusion), leading to an unnecessary expenditure of public health resources (i.e., a "wild goose chase").

Only the indistinguishable pattern should be included in the initial case-patient definition in an outbreak investigation. The shorter the duration of an outbreak, the less time the outbreak strain has to undergo mutations that change the PFGE pattern (Barrett et al. 2006;Tenover et al. 1995). More variability (patterns differing from each other in two to three band positions) may be acceptable if the outbreak has been going on for some time or if person-to-person spread is a prominent feature. If epidemiological information is sufficiently strong, case-patients with isolates exhibiting differing PFGE patterns may be included as part of an outbreak.

OPHS investigators prioritize and may follow up on FSIS-regulated product isolates and clinical isolates with indistinguishable PFGE patterns even if these product isolates are not currently associated with a designated outbreak cluster investigation. Efforts are made to establish any associations between positive FSIS-regulated products and foodborne illness. Priority is given to product isolates collected from recalled products, isolates exhibiting uncommon or new PFGE patterns, and isolates exhibiting increased virulence (e.g., hemolytic uremic syndrome (HUS) in *E. coli* O157:H7 infections). OSEL's VetNet search to determine the frequency of the specific isolate pattern from animal sources and FSIS-regulated products is important. A PFGE pattern may be rare in PulseNet and not a significant cause of human illness, but may be a common pattern in VetNet.

Once clusters are recognized, epidemiological investigations help determine whether a cluster actually represents an outbreak and help identify the clinical isolates resulting from a common source (Besser et al. 2008;Gerner-Smidt et al. 2006). Supporting data, such as case-patient report forms, food history questionnaires, analytical epidemiological study results, and purchase and point of preparation data, are collected from state and local public health departments. When information indicates exposure to

FSIS-regulated product, other FSIS program areas, such as OFO or OIEA, become involved, per FSIS Directive 8080.3 (FSIS 2008a). Epidemiological evidence combined with trace forward and trace back activities (establishment and product information) and laboratory sampling information help determine whether an FSIS-regulated product is associated with foodborne disease.

<back

3.1.6 A Case Study Using PFGE: FSIS Salmonella Multiple-drug Resistant (MDR) Outbreak Investigation

In June, 2009, the Colorado Department of Health and Environment conducted an investigation into a cluster of *Salmonella* Newport MDR infections. The *Xba*I PFGE pattern was not one of the top 10 most common PulseNet patterns of *S*. Newport. Ground beef isolates with patterns matching the case-patient isolates were identified in both PulseNet and VetNet. The ground beef isolates were recovered from an FSIS sample collected from a California processing plant (Establishment A) on May 21 and from a USDA-AMS sample collected from the same plant on June 1, 2009. Twenty of 22 Colorado case-patients reported consumption of ground beef during the week prior to illness onset; 18 of 20 reported purchasing ground beef from the same large supermarket chain (Chain A). The Colorado and Wyoming health departments provided FSIS case-patient ground beef purchase information via Chain A's shopper loyalty cards. FSIS investigators visited 11 Chain A supermarket locations to review ground beef production records. The FSIS review of retail production records in Colorado and Wyoming identified the suppliers of ground beef source materials to Chain A. Establishment A supplied coarse ground beef to Chain A supermarkets that manufactured various ground beef products purchased by case-patients. The first recall of raw ground beef due to *Salmonella* contamination was made possible due to detailed exposure information, adequate retail production records, and laboratory subtyping data, all of which successfully linked ground beef products to illnesses. This investigation illustrates the enhanced surveillance and hypothesis generation during foodborne illness investigations by cross referencing PulseNet and VetNet data.

<back

3.1.7 Using PFGE to Detect Lm Harborage within Establishments

Lm can establish a niche in damp environments and form biofilms in the processing environment, creating harborage sites that are difficult to eliminate with routine sanitation activities. Continued product contamination may occur until an establishment conducts sufficient corrective actions to eliminate the source of harborage or contamination. Inadequate corrective actions may lead to the organism's increased resistance to sanitizers and food preservation strategies, which, in turn, increases the risk of product adulteration. PFGE is an important tool to trace contamination and investigate potential bacterial harborage sites within an establishment and to evaluate the effectiveness of an establishment's corrective actions. OSEL performs a plant comparison on each *Lm* pattern. When the same PFGE pattern is detected over time or at multiple locations within an establishment, harborage is indicated, but not proven, because repeated reintroduction of the same organism also is possible. During a three-year period ending on September 30, 2011, 65 of 387 (17%) sampling events resulted in the detection of a PFGE pattern previously identified in the establishment indicating the potential for *Lm* harborage. In 118 of 387 (30%) sampling events, a PFGE pattern appeared to match a recent clinical isolate illustrating the potential public health risk of these harborage events. This information was communicated within FSIS and was presented to the establishment as evidence for harborage and of its lack of control over *Lm* in its post-lethality production environment[17].

<back

3.1.8 Using PFGE to Refine Attribution Models

Comparing the distribution of different pathogen subtypes in humans, food, animals, and the environment is useful in evaluating the burden of sporadic illness attributable to FSIS-regulated products. FSIS constituents routinely request more data, as well as more accurate attribution data on food/pathogen combinations to implement effective interventions. Pathogen subtyping is a critical component of foodborne disease attribution, by which FSIS should be able to get a better picture of the extent to which

[17] Source: FSIS Quarterly Reports prepared by OPHS and OPPD

FSIS-regulated products contribute to human illness. By combining human surveillance data on laboratory-confirmed infections from CDC's National *Salmonella* Surveillance System with food data from FSIS regulatory testing programs, FSIS and CDC, working with the Foodborne Diseases Active Surveillance Network (FoodNet) adapted the Bayesian attribution model methodology developed by Hald and colleagues (Hald et al. 2004) for food source attribution for *Salmonella* at the point-of-processing (Guo et al. 2011). The principle of the Danish Hald model[18] is to compare the number of human salmonellosis cases caused by different *Salmonella* subtypes with the prevalence of the subtypes isolated from different animal reservoirs or food sources, weighted by the amount of each food source consumed. The U.S. model uses food data collected primarily at processing establishments while the original model from Denmark was based on an integrated surveillance system with data collected from farm to table. In addition, Denmark has a smaller number of *Salmonella* serotypes of public health concern while in the U.S. there are over 100 serotypes of public health concern in food products, though the focus is on the CDC-designated top 20 serotypes (CDC 2012a).

FSIS used the model to estimate the number of human infections resulting from *Salmonella* contamination of raw meat, poultry, and egg products in the United States. The model looks at the attribution by individual *Salmonella* serotypes as well as by all *Salmonella*. Next steps in model refinement may include: 1) incorporating other sources of food data, such as prevalence estimates for *Salmonella* contamination of FDA-regulated commodities; 2) use of PFGE and other molecular subtyping data; and 3) adding other foodborne pathogens.

<back

[18] A method was developed in Denmark (Hald, T. et al, A Bayesian approach to quantify the contribution of animal-food sources to human salmonellosis) that uses information about food contamination, consumption patterns of the population, and human foodborne illnesses to estimate the number of illnesses associated with specific food reservoirs. This method has been adapted in several countries for foodborne illness source attribution of several pathogens.

3.1.9 Using PFGE in FSIS' Salmonella Verification Testing Program

FSIS obtains serotype and PFGE results from *Salmonella*-positive samples as part of the Agency's *Salmonella* Verification Testing Program for raw products. FSIS strives to share *Salmonella* serotype data (provided by OSEL and the Animal and Plant Health Inspection Service (APHIS) National Veterinary Services Laboratory (NVSL)) with establishments within two weeks after a sample has been reported as positive. In 2011, FSIS revised establishment end-of-set letters (EOSL), which are sent to establishments following a pathogen reduction-hazard analysis and critical control point (PR/HACCP) sample set, to include the *Salmonella* serotype for each positive sample. The letters also indicate whether the serotypes detected during the set are on the CDC's top 20 list of most common serotypes from human sources. FSIS schedules *Salmonella* sets according to criteria outlined in a sampling algorithm (FSIS 2011a), which includes whether serotypes of human health concern (defined as serotypes on the CDC's top 20 list for the most recent calendar year) were detected during the establishment's last set. FSIS is actively working to report PFGE pattern and ARA data from PR/HACCP sets to establishments in a separate letter or in a further revision to the EOSL format. This information is intended to assist an establishment to evaluate and improve its *Salmonella* process control[19].

The FSIS *Salmonella* initiative concentrates resources on establishments with higher levels of *Salmonella* and with serotypes of human health concern. FSIS takes follow-up action, which may include scheduling another sample set or assessing the design and execution of the establishment's food safety system. Food safety assessments are conducted at an establishment failing a *Salmonella* verification sample set, at an establishment with repetitive *Salmonella* serotypes of public health concern on its products, and at an establishment with *Salmonella* PFGE patterns associated with foodborne illness in its most recent sample set (FSIS 2008b;FSIS 2009).
<back

[19] FSIS has published a compliance guideline for controlling *Salmonella* and *Campylobacter* in poultry [http://origin-www.fsis.usda.gov/PDF/Compliance_Guideline_Controlling_*Salmonella*_Poultry.pdf].

3.1.10 Campylobacter PFGE

In May 2010, FSIS announced new *Campylobacter* performance standards for young chicken and turkeys at slaughter establishments (FSIS 2010). One year later, FSIS initiated a verification sampling program. ARS-VetNet performed PFGE on all FSIS *Campylobacter* isolates using the primary restriction enzyme, *Sma*I, and by request, the secondary enzyme, *Kpn*I until June, 2013 when OSEL assumed these duties. Current patterns are uploaded to the PulseNet database; the older isolates are included in the VetNet database. VetNet *Campylobacter* PFGE patterns are assigned unique pattern names that are not equivalent to PulseNet pattern names. The VetNet *Campylobacter* database is currently not available online and is not directly accessible by OSEL or CDC-PulseNet.

The PulseNet protocol for preparing *Campylobacter* DNA for PFGE is different from the protocols used for *E. coli*, *Salmonella,* and *Lm*. The *Campylobacter* protocol and the interpretation of *Campylobacter* PFGE data is technically challenging. It is often difficult to culturally isolate *Campylobacter*, which is necessary to avoid getting a mixture of different patterns. In addition, *Campylobacter* patterns are often complex, with numerous bands, which makes it challenging to distinguish clones with different patterns based on size alone.

The PulseNet *Campylobacter* database is much smaller than the *Salmonella, E. coli,* or *Lm* databases, but it is growing. Some submitting laboratories often ‘batch” many months of data before uploading to the PulseNet database. Much of the database is represented by single enzyme digest patterns. While PulseNet does maintain a national *Campylobacter* database, it is not actively used to detect clusters.

Routine PFGE subtyping is of limited value because of the high genetic diversity within the major *Campylobacter* species: C. *jejuni* and *C. coli*. However, when *Campylobacter* is identified as a pathogen in an active outbreak, PFGE may be a useful investigative tool. PulseNet recommends only confirmatory subtyping of the strains when outbreaks

are detected by other means (Gerner-Smidt, 2005). PFGE data generated from FSIS samples may be used for purposes other than outbreak cluster detection. Other groups have used PFGE to determine *Campylobacter* cross-contamination at broiler slaughter (Ellerbroek et al. 2010); to demonstrate a change in genotype during the course of antimicrobial treatment of pigs (Juntunen et al. 2010); to demonstrate the role of poultry products in human campylobacteriosis in Finland (Lyhs et al. 2010); and to characterize samples collected from broilers (Griekspoor et al. 2010;Miller et al. 2010;Wilson et al. 2009). Additional typing methods for *Campylobacter* are discussed in Section 5.3.
<back

3.2. Multiple Locus Variable Number Tandem Repeat Analysis (MLVA)

MLVA takes advantage of tandem repeat sequences that occur in discrete locations of bacterial genomes. Each tandem repeat location contains different numbers of simple sequence units. MLVA is performed as a two-step process.

1. The repeat locations are copied using PCR. Primers specific for areas outside of the repeat locations are used to make these copies.
2. Once copied, the sizes of the PCR-copied fragments are measured by a DNA sequencer or another instrument that can make very accurate estimates of DNA fragment size. The repeat number is determined after accounting for the extra sequence copied through the PCR process and the size of the simple sequence unit.

The process is described in greater detail in Appendix G and Appendix H. A MLVA assay consists of multiple (about eight to 20) assays, each designed to determine the number of units at a specific location. Bacteria from the same source are expected to contain the same number of tandem repeats at a specific location (Example).

Example

The DNA sequence “ACG” in an *E. coli* strain *may* be repeated five times at a specific location in its genome (ACG-ACG-ACG-ACG-ACG). Bacterial isolates from the same source are expected to contain five “ACG” repeats in the same site while isolates from an unrelated source would most likely contain a differing number of ACG repeats at the same locus.

By experimentation, the CDC and other groups identified eight specific locations of the *E. coli* O157:H7 genome that cntains tandem repeats. Each repeat ends to be the same in bacterial isolates generated from the same source and different in bacterial isolates from different sources. The same eight repeat loci are measured for every *E. coli* O157:H7 isolate. Because these loci tend to vary, MLVA assays must be designed or validated for each major lineage, such as serotype or serogroup. In some instances, this involves finding and testing candidate repeat loci for epidemiological concordance, designing and testing primers, and designing algorithms to calculate repeat numbers. A large number of *Salmonella*, *Lm*, and *E. coli* genome sequencing projects are providing raw data that can be used to discover candidate repeat loci. In other instances, an existing MLVA assay can be applied to a related lineage on an experimental basis. For example, the *S.* Typhimurium MLVA assay has been used during recent high-profile *S.* Heidelberg and *S*. Montevideo outbreaks associated with FSIS-regulated foods.

MLVA and PFGE visualize different kinds of mutations. Large indel mutations that would be readily apparent using PFGE may not be detected with MLVA because they occur outside of the tandem repeat loci specified by the assay. Conversely, MLVA detects small mutations at specified loci, which are likely to be invisible to PFGE. During validation of the *E. coli* O157:H7 MLVA protocol by CDC, MLVA was slightly less discriminating than PFGE with two enzymes. However, MLVA provides a useful complement to PFGE because MLVA can subtype further some of the most common PFGE patterns. Epidemiological concordance of MLVA data was found to be better than that of PFGE (Hyytia-Trees et al. 2006). The discriminatory power of MLVA depends on the assay design, including the choice of repeat loci and the level of genetic variation within the lineage to be subtyped. In general, MLVA assays developed by PulseNet have not proven to be more discriminatory than PFGE. However, in some cases, a MLVA assay discriminated strains within a lineage that was indistinguishable by PFGE. MLVA is helpful during outbreaks involving a common PFGE pattern or outbreaks involving organisms that have very limited PFGE pattern diversity. In these situations, MLVA is used to identify case-patients who were not associated with an outbreak.

PulseNet laboratories tend to use MLVA to complement PFGE during certain investigations. The CDC does not expect that MLVA will replace PFGE.

CDC PulseNet has invested significant resources to develop MLVA to a point where it is now useful for discriminating strains within *E. coli* O157:H7 and some *Salmonella* serotypes. Currently, state and federal laboratories submit select isolates to MLVA-certified PulseNet Laboratories for laboratory analysis. CDC personnel perform data analysis using BioNumerics, the same software used to analyze PFGE data. Currently, PulseNet has validated MLVA assays for *E. coli* O157:H7, *S*. Typhimurium and *S*. Enteritidis (Hyytia-Trees et al. 2006;Hyytia-Trees et al. 2010). OSEL analysts are PulseNet-certified for the *E. coli* O157:H7 and *S*. Typhimurium MLVA assays and are planning to acquire certification for the *S*. Enteritidis assay. OSEL performs MLVA on all *E. coli* O157:H7 and on *S.* Typhimurium isolates upon request by FSIS program staff (requests are typically made through OPHS at headquarters).

In FSIS' experience, start-up costs and training for MLVA are more significant than for PFGE. MLVA is substantially easier to perform than PFGE. MLVA data are somewhat easier to interpret than PFGE patterns. The data analysis procedures for MLVA can be automated and therefore require less analyst time. MLVA has a greater reproducibility and comparability between analysts than PFGE. As the PulseNet MLVA database grows, the significance of the differences in MLVA profiles will become more apparent.
<back

3.2.1 Using MLVA in Outbreak Investigations

When the PFGE pattern is common, foodborne disease investigators may use MLVA to separate sporadic case-patients from outbreak-related case-patients to help establish the extent of an outbreak. FSIS consults with CDC and the PulseNet/Enteric Diseases Laboratory Branch laboratorians to determine if MLVA is likely to result in a substantial increase in specificity or discriminatory power to identify a cluster of human cases. CDC and state epidemiologists then may include both MLVA and PFGE patterns in case-patient definitions to guide case-patient ascertainment, refine the case count and

conserve the use of investigative resources. A request is made for MLVA analysis of all isolates in a cluster when MLVA has been included in the epidemiologic case definition to help determine the scope of the outbreak.

When a foodborne disease investigation is in the hypothesis-generating phase and MLVA data are available, CDC recommends that investigators focus on collecting food exposure information from case-patients with an indistinguishable MLVA pattern. If an active epidemiologic investigation is ongoing, PulseNet-MLVA staff may request isolates from all interviewed people to restrict the analysis of exposure information to outbreak-related case-patients.

During outbreak investigations, MLVA may differentiate isolates that are indistinguishable by PFGE but have no common epidemiological source (Noller et al. 2003). Torpdahl (Torpdahl et al. 2006) used MLVA characterization for source identification in a *S.* Typhimurium DT12 outbreak in Denmark. MLVA correlated more closely than PFGE with epidemiological data in this outbreak of a highly clonal strain of *S.* Typhimurium. MLVA also is useful to assess the relatedness of closely related PFGE patterns (e.g., variant PFGE *Bln*I patterns) in temporally related illness clusters.
Highly variable VNTR loci are preferred for short-term cluster detection because of their ability to discriminate between closely related strains. However, this high variability may cause instability in some of these loci during outbreaks. Each VNTR site may have different mutation rates, so variation at some loci would be expected within an outbreak cluster, while variation at other loci would not be expected. This may be determined through experience for each assay. CDC developed rules in this manner to deal with variation in *E. coli* O157:H7 MLVA patterns during an outbreak. From past outbreaks, CDC recognizes that a difference of up to three repeats at a single locus (except at locus 34) or one repeat at two loci (excluding locus 34) can occur during an *E. coli* outbreak.[20] No difference should occur in the number of repeats at VNTR locus 34. In order to develop data interpretation guidelines, thorough mutation studies and an

[20] Dr. Eja Hyytiä-Trees, Personal communication, March 17, 2009

understanding of the factors affecting mutation rates are needed to distinguish VNTR loci that are useful during outbreak investigations (Hyytia-Trees et al. 2007).

Isolate selection for MLVA typing includes: 1) isolates from large multi-state clusters (as prioritized by the CDC PulseNet Laboratory) with a common PFGE pattern and a wide geographical and temporal distribution; 2) isolates from states with multiple PFGE matches; 3) isolates recovered from animal or other food sources with indistinguishable PFGE patterns; 4) PFGE patterns associated with prior large outbreaks (e.g., a more virulent strain); or 5) outbreaks with skewed demographics (e.g., a large number of children, large proportion of one gender).

During foodborne disease investigations involving *E. coli* O157:H7, trace back of ground beef is often complex and may include analyzing data from multiple establishments, evaluating grinding records, and assessing clean-up between production lots of different source material. Sporadic illnesses that might not be associated with a common source could have the potential to confound trace back information. FSIS increasingly uses MLVA to focus trace back activities when investigating PFGE subtype clusters. Trace back should focus on the case-patients who have indistinguishable MLVA patterns. Once the source of the outbreak has been identified, epidemiologic evidence may indicate that close variants to the outbreak primary MLVA pattern should be included in the case definition. Case-patients with similar or indistinguishable MLVA patterns should be included in the case count only if a similar exposure/source of infection has been identified.

<back

3.2.2 A Case Study Using MLVA: FSIS E. coli O157:H7 Outbreak Investigation

Two investigations illustrate the utility of MLVA data.

In August 2009, a Utah cluster of *E. coli* O157:H7 infections was reported and investigated by a multi-agency team. Fourteen primary case-patients, with onset dates ranging from July 11, 2009, through August 27, 2009, were reported with an

indistinguishable PFGE pattern combination common in PulseNet, but relatively uncommon in Utah. Because this pattern was uncommon in Utah, CDC initially did not recommend conducting MLVA. Investigators suspected a common source outbreak in Utah due to the temporality and geographic cluster of the cases. Upon case-patient interviews, it was determined that a high proportion of case-patients reported consuming ground beef (13 or 93%) and attending rodeos (12 or 86%). Because case-patients reported attending four different rodeos, ground beef exposure was an initial hypothesis. FSIS trace back investigations of ground beef did not converge on a single supplier, but rather identified multiple establishments. One establishment was identified as a possible common source supplier accounting for a number of the illnesses, but some infections could not be linked to that establishment's products. Samples of ground beef products were collected from case-patient homes and retail locations, and all samples tested negative. MLVA was used in an attempt to discover if sporadic *E. coli* O157:H7 infections were confounding the trace back investigation of an underlying sub-cluster. CDC performed MLVA on 14 primary cases and two secondary cases. Four different MLVA patterns (A1-A4) were identified —10 isolates were indistinguishable (MLVA pattern A1), five had slight variations at one locus, and one had slight variations at two loci. The case-patients with the outbreak PFGE pattern could be considered part of the same outbreak based upon CDC-PulseNet's interpretation of the variation in *E. coli* O157:H7 MLVA patterns. If the outbreak definition was tightened to only those case-patients with MLVA pattern A1, some case-patients with this MLVA pattern were not linked to the suspect establishment's ground beef products. MLVA did not provide much resolution for the PFGE pattern cluster and did not strengthen the association of the cluster with a particular ground beef source. The Utah Department of Health further investigated the possibility that case-patients were linked through rodeo attendance. Since case-patients did not consume common food items at the rodeos, animal contact was investigated as a possible source of infection. It was determined that the same stock contractor supplied cattle at all four rodeos. An environmental (soil) sample taken from one of the rodeo grounds tested positive for *E. coli* O157:H7 with an isolate that had an indistinguishable PFGE pattern combination to the outbreak strain and with MLVA pattern A1. The evidence supported environmental (cattle) exposure at the

rodeos, not ground beef consumption, as the source of the outbreak. Case-patients whose isolates were the three variant MLVA patterns were possibly sporadic cases with differing exposures; however most of these also reported rodeo attendance. Because of the investigation, the Utah Department of Health alerted the public and rodeo/fair organizers of the risks of *E. coli* O157:H7 exposures at live animal venues and provided guidelines focusing on reducing exposures at such events (Lanier et al. 2011).

In another example, in October 2009, CDC PulseNet staff assigned an outbreak code to a PFGE cluster of *E. coli* O157:H7 isolates due to an increase in uploads. Thirty-one case-patients were reported from eight states with an indistinguishable PFGE pattern combination that is common in PulseNet. Onset dates ranged from August 18, 2009, through October 7, 2009. Among case-patients with exposure information available, 20 of 24 (83%) reported exposure to ground beef. Leftover ground beef samples were collected from case-patient homes and submitted for testing. The leftover samples tested positive for *E. coli* O157:H7 with an indistinguishable PFGE pattern combination from the outbreak strain. CDC performed MLVA on isolates from 31 clinical and 3 ground beef samples and all were indistinguishable. FSIS Investigators, in collaboration with public health partners, performed trace back investigations of ground beef from two grocery store chains, which converged on a single supplying establishment. Because the PFGE pattern combination was common, MLVA provided additional evidence, supported the conclusion, and validated the trace back results. On October 31, 2009, FSIS announced a recall of approximately 545,699 pounds of fresh ground beef products.

In summary, MLVA has the potential to provide powerful insight for ground beef trace back investigations, particularly with common PFGE patterns. However, as with any single point of evidence, it is critical that results be compared with additional data including other laboratory results, epidemiologic information, and results from environmental assessments. MLVA may help to identify epidemiologically linked isolates correctly as being genetically related, to differentiate epidemiologically linked isolates from background sporadic isolates, and to differentiate between outbreaks not

due to a common source. When the PFGE pattern combination is common, MLVA provides additional evidence and validation of product sampling and trace back results.
<back

3.2.3 Using MLVA in Salmonella Enteritidis (SE) Outbreaks

SE is one of the most genetically homogenous serotypes of *Salmonella* and is poorly differentiated by most commonly used subtyping methods (Allard et al. 2013;Pang et al. 2007;Saeed et al. 2006;Zheng et al. 2011). Two PFGE patterns make up nearly 48% of the SE isolates in PulseNet. Therefore, attributing SE to foodborne disease outbreaks is difficult. MLVA provides improved discriminatory power for subtyping. One study showed that, overall, MLVA typing of SE had enhanced resolution, good reproducibility, and good epidemiological concordance (Boxrud et al. 2007).
<back

3.2.4 Caution Needed to Interpret MLVA and PFGE Data

Caution is needed for interpreting a three of more enzyme digests during PFGE analysis, or when adding MLVA to PFGE results. Even when a food isolate and a patient isolate do come from the same source, it is doubtful that they are genetically identical. Therefore, the purpose of genetic subtyping methods ***is not*** to determine if two isolates have genetic differences, but to determine if two isolates are sufficiently similar to each other, and distinguishable from other isolates from non-epidemiologically related sources. The effectiveness of the two-enzyme protocol used in differentiation/categorization of PulseNet isolates has been proven empirically; that is, the hundreds of thousands of isolates in PulseNet have been categorized using one or two of the enzymes. The two-enzyme PulseNet PFGE method is sensitive enough to assign a reliable pattern to hundreds, and for some organisms, thousands of categories (pattern names), while at the same time has detected many clusters representing common source outbreaks. Empirical evidence does not support performing PFGE with additional enzymes not routinely used by PulseNet labs. Using other enzymes may be helpful in some cases, but improper use may be misleading. Additional enzyme digests are therefore performed only when requested by PulseNet.

CDC PulseNet has invested significant resources to develop MLVA to a point where it is now useful for differentiating *E. coli* O157:H7 and some *Salmonella* serotypes. The significance of the information collected through PulseNet is under constant evaluation and review even for organisms that have an effective protocol. Currently, only PulseNet analysts located at the CDC have access to the PulseNet MLVA database, unlike the PFGE databases, which are accessible by all PulseNet laboratories.

<back

3.3 Subtyping Methods for Non-O157 STEC

FSIS and ARS developed screening and isolation procedures for six non-O157 STEC serogroups: O26, O45, O103, O111, O121, and O145. These serogroups accounted for more than 70% of the clinical isolates reported to the CDC during the period from 1983 through 2002 (Brooks et al. 2005). The method was added to the FSIS Microbiology Laboratory Guidebook (MLG) in October 2010. FSIS initiated verification testing of raw beef manufacturing trimmings for these STEC serogroups in June 2012. Food enrichment broths are screened for the presence of the *stx*1 and *stx*2 genes encoding Shiga toxins; the virulence gene *eae* that encodes the intimin protein, which allows close adherence of STEC to human intestinal cells; and the *wzx* or *wbd*l genes associated with the biosynthesis of O-antigens for the six serogroups. If the enrichment broths screen positive for *stx*, *eae*, and one or more of the top six O-antigen associated genes, cultural isolation is initiated and the isolates obtained are confirmed for virulence and serogroup genes using PCR.

The PFGE procedure for non-O157 STEC uses the same standards, enzymes, and equipment as the *E. coli* O157:H7 procedure. Lab analysts who are certified by PulseNet for *E. coli* O157:H7 also are certified for non-O157 STEC, but a separate certification for the analysis of non-O157 STEC gels is required. Data collection procedures also are similar with the following exceptions: OSEL and PulseNet maintain separate databases and separate pattern name format for non-O157 STEC and *E. coli* O157:H7. Currently, data analysis procedures are similar, but this may change as

PulseNet receives more non-O157 STEC isolates. A MLVA test is not available for non-O157 STEC.

<back

3.4 Molecular Confirmation of *E. coli* O157:H7

Routine confirmation of *E. coli* O157:H7 is comprised of three parts: (1) serology for the O-group (H group serology is performed, but is not necessary for confirmation); (2) an enzyme-linked immunosorbent assay (ELISA) method for the Shiga toxin (enterohemorrhagic *E. coli* (EHEC) test kit); and (3) an automated biochemical test (VITEK). As an alternative to the traditional serological method, which uses antisera prepared against the H7 antigen, FSIS laboratories may perform a molecular serotyping method, which is illustrated in Appendix I. OSEL performs PCR-based molecular serotyping of *E. coli* O157:H7 when routine confirmation methods yield indeterminate results (e.g., the isolate is O157 positive by serology and Shiga toxin negative by ELISA). The molecular serotyping PCR test is run on the SmartCycler (an instrument used to perform a real-time PCR) and is designed to simultaneously detect the intimin gene (*eae),* the H7 flagella antigen gene ($fliC_{H7}$), and the Shiga toxin gene. The assay does not differentiate between the major Shiga toxin genes, *stx*1 and *stx*2, and only serves to confirm the presence of at least one of these genes. A positive O157 serology result, combined with a positive PCR test for the flagella antigen gene $fliC_{H7}$ or *stx*, is sufficient to confirm the isolate as *E. coli* O157:H7 subtype (FSIS 2012b). The *eae* result is not used for confirmation of *E. coli* O157:H7 because certain eae types are found in non-O157 STEC strains as well and are therefore not sufficiently specific to *E. coli* O157:H7.

<back

3.5 Molecular Serotyping of *Salmonella*

Traditional serotyping of *Salmonella* uses specific antibodies to type certain cell surface antigens. The method is very specialized, time consuming, and expensive. Laboratories performing traditional serotyping must maintain a large library of antisera to detect all *Salmonella* H- and O- antigens. For this reason, prior to July 2012, FSIS sent all of its

Salmonella isolates to APHIS' NVSL for traditional serotyping resulting in significant delays in reporting serotype results. This makes it difficult to report the results within the Agency and to stakeholders in a useful timeframe. CDC developed a molecular serotyping method for *Salmonella* (Fitzgerald et al. 2007;McQuiston et al. 2011), which has been verified by FSIS. The molecular method is applied to purified isolates and can be divided into two steps.

1. Regions of the *Salmonella* genome responsible for producing the flagellar and somatic antigens (H- and O-antigens encoded by the *fli*B, *fli*C, and *rfb* genes) are amplified using PCR.
2. Specific probes complementary to sequences specific for the fliB, fliC, and rfb genes of major *Salmonella* serotypes are combined with the PCR-amplified DNA. If the probes bind their target sequence, a signal can be detected using the Bioplex platform.

This process is illustrated in Appendix J. The CDC method is quick, straightforward, and economical when compared to traditional serotyping.

In June 2012, FSIS began to use the CDC's molecular serotyping method to identify some *Salmonella* serotypes. As of January 2, 2013, the OSEL laboratory is responsible for the routine identification of twenty-five *Salmonella* serotypes by the molecular method (listed in Appendix K), with all other serotypes confirmed by the NVSL using traditional serotyping methodology (FSIS 2013a).

<back

4 Subtyping Technologies Being Developed or Under Evaluation (Short Term, One to Four Years)

This section describes subtyping methods that FSIS is currently exploring with possible implementation dates in the next one to four years. This includes new subtyping methodologies for *Campylobacter*, non-O157 STEC, *Lm,* and *Salmonella* spp. In addition, FSIS is in the process of implementing a Public Health Information System

(PHIS) that will provide direct access to PFGE pattern information within the Agency and to other federal and state government stakeholders.

4.1 MLVA for Non-O157 STEC, *Lm*, *S.* Newport, and other *Salmonella* serotypes

In general, VNTR regions targeted by MLVA assays are specific to serologically-distinguishable groups of bacteria. Thus, MLVA assays developed for one group are not expected to be valid for others. For example, MLVA assays validated for *E. coli* serotype O157:H7 and *Salmonella* serotypes Typhimurium and Enteritidis are not expected to be applicable to other non-O157 STEC serogroups or to important *Salmonella* serotypes such as Heidelberg, Montevideo, and Newport. Therefore, additional MLVA assays are under development or being validated for use with additional bacterial groups. For non-O157 STEC groups, PFGE has proved to be sufficiently discriminatory, and therefore, development of MLVA for other organisms is currently a higher priority for CDC. However, this may change as more strains from food products are isolated, typed, and added to the PulseNet database. A MLVA protocol for *Lm* is currently undergoing internal validation at CDC's Enteric Diseases Laboratory Branch, and CDC developed a MLVA protocol for *S.* Newport, which has been internally validated and is undergoing external validation. CDC has evaluated additional MLVA assays for *S*. Heidelberg and *S*. Montevideo following large foodborne investigations involving FSIS-regulated products. These assays have been used under an experimental basis and have not been validated for use by PulseNet labs. In 2014 the ultility of whole genome sequence comparison for outbreak surveillance/detection is being pursued by CDC in lieu of new MLVA techniques for foodborne pathogens.

To discover new VNTR regions suitable for MLVA assay development, researchers must first study many isolates from a particular group. Whole genome sequencing initiatives, such as the 100K Genome Project (Section 5.1), will help researchers discover VNTR and understand their variation within a specific group, such as a serotype.

<back

4.2 Public Health Information System (PHIS)

FSIS implemented PHIS, an information technology system that collects, consolidates, and analyzes data. It replaced many of FSIS' other data systems, such as the Performance Based Inspection System (PBIS) and the Automated Import Information System (AIIS). Field and headquarters staff access PHIS through a web-based application. The system has four components: domestic inspection, import activities, export activities, and predictive analytics. The predictive analytics component will allow FSIS staff to perform advanced data analyses to detect temporal and spatial trends. The system will be capable of integrating other data streams generated by ARS, AMS, and CDC. The PHIS will be capable of expanding to accept other data sources and types and new subtyping methods as adapted by FSIS. FSIS is working with CDC PulseNet and ARS to transfer PFGE pattern names, serotype data, and associated metadata from the FSIS Laboratory Information Management System (LIMS), as well as the PulseNet and VetNet databases into the PHIS, which will allow authorized users to access and generate reports on specific isolates and to perform comprehensive studies within the system. Due to the real-time nature of the data transfers, OSEL will still handle time-sensitive needs, such as requests related to ongoing outbreak investigations.

<back

4.3 Targeted Multilocus Genotyping (TMLGT) for *Lm*

ARS, in collaboration with FSIS, developed and validated an SNP-based assay, TMLGT, to facilitate rapid strain characterization and the integration of subtype data into risk-based inspection programs. Where TMLGT has poor discrimination for specific outbreaks, PFGE has strong discrimination. Where PFGE is poorly able to assign *Lm* to lineage, serogroup, and EC clades, TMLGT is able to assign. TMLGT was designed to use Luminex-xMAP technology and the Bioplex platform (also used for molecular serotyping of *Salmonella*, Section 3.5). This technology uses flow cytometry to detect specific nucleic acid sequences. However, the SNP used in the TMLGT methodology could be determined by other means, such as by direct sequencing (Section 5.1). TMLGT was validated in a study that included 906 isolates that had been independently

subtyped using MLGT. Nine hundred three isolates were correctly typed by the method (99.7% typability) (Ward et al. 2010b).

FSIS may consider using more detailed information about *Lm* subtype, which provides information about key differences in public health risk and ecology, to improve existing risk-based sampling programs or inspection policies. In 2005, FSIS implemented a sampling program that incorporates specific data related to the risk of *Lm* contamination (e.g., production volume, outgrowth potential in the product, steps taken to prevent post-lethality contamination, and FSIS sampling history) to create a risk-based sampling framework (FSIS 2007). Through this program, FSIS collects samples from establishments with higher risk more frequently than establishments with lower risk. Along with sampling, FSIS assigns additional inspection to an establishment when there is evidence that it is manufacturing hazardous products, such as when an establishment's products are found to be adulterated with *Lm* or are determined to be associated with an *Lm* outbreak. Data from TMLGT subtyping of FSIS' *Lm* isolates could be used to extend the risk-based sampling framework to include estimates of public health risk (Section 2.8)

<back

5 Subtyping Technologies of the Future (Five or More Years to Implementation)

This section examines examples of how new technologies, including high-throughput sequencing, can be used to provide informative subtype data on bacterial pathogens of concern to FSIS (e.g., STEC, *Lm*, *Salmonella*, and *Campylobacter*). FSIS is not currently evaluating these methods, and Agency implementation may take five or more years. Other U.S. agencies, including CDC, FDA, and ARS, as well as organizations in Canada and the European Union, are developing methods and exploring the use of these data. In recent years, FSIS has participated in meetings designed to bring federal public health agencies together to discuss the development and implementation of subtyping methods. These include the Annual PulseNet meetings organized by CDC, the New Frontiers of Molecular Epidemiology (NFME) meeting organized by FDA and

CDC in 2009 and 2011, the Systems for Food Examination (SAFE) meetings organized by the Department of Defense's Advance Research Projects Agency (DARPA) in 2009, an American Society for Microbiology (ASM) Symposium on the Development and Use of Subtyping Data by Regulatory Agencies organized by FDA and FSIS in 2010 and the Global Microbial Identifier meetings that were initiated in 2011. These meetings provide a forum where FSIS can share ideas and recommendations on subtyping method development with federal partners. The meetings encourage ongoing and two-way communication to ensure that all organizations understand each other's goals and timelines.

5.1 High-Throughput Sequencing of Bacterial Genomes

In 2013, most subtyping methods used by public health and regulatory authorities were based on the interpretation of DNA or RNA fragments or the differential presence or absence of genetic markers associated with virulent strains (Section 2). At present, nucleic acid sequence is used in limited capacity, and includes SNP, MLST, and whole genome sequence comparison.

Instrumentation, technology, and chemistry needed to perform nucleic acid sequencing have progressed enormously in the past decade and further advances are expected in the next five years. High-throughput sequencing usually involves collecting tens of thousands of partially overlapping sequence fragments. By a process called "assembly," the fragments are arranged into their correct order by matching their overlapping sequences, so that large contiguous sequences, and even entire genomes, are reconstructed. In the past, these instruments and capabilities were only available at large facilities. With new, more economical technology, large scale sequencing is beginning to be available to individual laboratories, and large facilities have increased their capabilities dramatically, resulting in an impressive increase in the availability of whole genome sequences deposited into public databases such as the National Center for Biotechnology Information (NCBI). The cost of high-throughput sequencing has decreased in the past decade due to increased availability of instruments and reagents. Next-generation sequencing platforms, such as the 454 GS FLX (Roche Life Sciences),

Miseq and Hiseq (Illumina), PacBio *RS* (Pacific Biosystems), and the Ion Torrent (Life Technologies) sequencers, have become readily available to smaller laboratories including public health labs. These platforms allow multiple genome sequences to be determined in a single run, which will further drive down prices. The availability of genome sequence is driving the need for computational software for assembly and scientists who are trained to perform sequence assembly and annotation (i.e., describing the features of the assembled sequence). Assembly, annotation, and interpretation eventually may be automated. This is being pursued actively by NCBI. Programs eventually may be designed to extract useful information on MLVA, MLST, and SNP subtype from a genome assembly. For FSIS, these trends mean that high-throughput sequence technology and data will be routinely available at a substantially lower cost. CDC, FDA, the Public Health Agency of Canada, New York State Department of Health, and other organizations are beginning to produce and analyze high-throughput and WGS data during the course of an outbreak investigation, and in 2013, CDC initiated a *Lm* real time surveillance project using WGS. WGS data likely will be available to industry as well. The 100K Pathogen Genome project is a consortium initiated by FDA, University of California Davis, and Agilent Technologies, to develop a database of bacterial pathogen genomes accessible through the NCBI. The project seeks to improve the understanding of pathogen genetic variation, and the development of new subtyping assays. FSIS has stated its intention to contribute strains from its regulatory testing program to this project

FSIS will need to consider how to respond to trends in the availability of large scale sequencing capability. For example, should FSIS continue to support the use of these data by contributing strains to the 100K Pathogen Genome Project or other organizations like ARS, CDC, and FDA, and should the Agency acquire the technology to collect these data in real time? Some of the issues raised by the availability of large scale sequencing technology are presented below.

WGS data potentially provides one of the most complete and definitive assessments of subtype. This data is portable (i.e., it can be easily shared between researchers through

the internet, as PFGE and MLVA pattern data currently is shared) and the data are compatible with previous fragment-based, virulence marker, and limited sequenced-based typing schemes like MLVA and SNP analyses. For example, provided the assembly is sufficiently correct, WGS data could be used to find restriction enzyme sites and calculate the size of restriction fragments without the need for PFGE. WGS data can detect large genetic changes, such as inversions, insertions, deletions, and duplications that would not be detected by shorter sequences obtained with MLVA, MLST, SNP, or microarray based typing assays. The data can be used to define new sequences or genes that distinguish and identify hyper-virulent lineages for further research. In summary, WGS provides a wealth of genetic information about a strain, some of which would be relevant during outbreak investigations, but also would drive the development of new subtyping assays to detect lineages or virulent clones of public health concern.

An important issue is data specificity. WGS will detect much greater detail about strains than other subtyping methods, and in many situations will distinguish strains that would be conclusively indistinguishable by PFGE, MLVA, SNP and other methods. Some of these differences are meaningful and distinguish the outbreak lineage from sporadic cases. Other differences may be found both within and outside the outbreak lineage and therefore should be considered inconsequential to the interpretation of the outbreak data. Like any other assay, the proper balance between sensitivity and specificity (i.e., to correctly detect and act to limit the spread of true outbreaks) will need to be determined from practical observation and experience. It is possible that Industry will use WGS data to “exonerate” their products or otherwise limit regulatory action. Interpretative guidelines on the collection, analysis, and interpretation of WGS data are needed to provide standards for this rapidly growing area of investigation.

FDA-CFSAN researchers have used WGS to evaluate *S.* Montevideo from a multistate outbreak of salmonellosis associated with pepper-coated pork salami products that resulted in multiple FSIS and FDA-initiated product recalls. The researchers evaluated complete genomes from patients, boxes of black and crushed red pepper, and an

environmental isolate recovered from a drain at the facility. The density of data provided by the WGS allowed investigators to exclude the involvement of pistachio nuts, while validating the association with multiple production lots of black and crushed red pepper used as an ingredient in the salami products (Lienau et al. 2011). Interestingly, the FDA group concluded that the S. Montevideo strain came from a domestic source, which conflicts with the international origin of the spices (China and India). This could indicate that the spices were cross-contaminated following export to the United States, perhaps at a spice grinding facility. Another group used WGS data (SNP typing assay of 112 isolates) from the same sources to suggest that a majority of the case-patients were sporadic, and not outbreak associated as originally assumed by CDC and FDA (Bakker et al. 2011).

Other outbreak investigations using WGS have been reported, including *Lm* outbreaks in the United States and Canada (Gilmour et al. 2010;Orsi et al. 2008), a large clonal *S*. Enteritidis outbreak associated with eggs (Allard et al. 2013), and STEC outbreaks associated with spinach and sprouts in the U.S. and Europe (Brzuszkiewicz et al. 2011;Kotewicz et al. 2008;Mellmann et al. 2011;Rasko et al. 2011). Researchers will continue to perform WGS from bacteria in strain collections, including FSIS positive product samples, which will help to interpret the degree of variation seen in genomes from outbreak isolates. This information will be extremely useful to define and test whether certain sequences can be used reliably to distinguish unrelated strains and, thus, could be used prospectively for outbreak cluster detection.

The availability of low-cost genome sequence data should benefit public health and regulatory agencies. However, WGS data generally would be more expensive to collect than existing subtyping strategies, require more computer database storage, and will produce replicate datasets that are identical in all but a small number of sequence differences. The per-genome cost could make WGS prohibitive for surveillance samples. Surveillance samples, including isolates from “sporadic” illnesses and routine food and environmental samples may be an important application for subtyping. A great deal of information may be gained from sequencing surveillance samples, such as

detecting virulence genes, characterizing emerging subtypes, or discovering new SNP that could be used to distinguish future outbreak clades. The cost of high-throughput sequencing technologies could be offset by using the data for multiple purposes. Such purposes include: obtaining MLVA, SNP, PFGE subtype from the WGS data, and finding important genes responsible for virulence, persistence, and resistance to antibiotics or sanitizers; developing new assays for detecting future outbreak cluster detection and routine surveillance; and using virulence gene or marker discovery to study mechanisms of pathogenicity for future risk assessments and attribution studies.

Metagenomics is the genetic study of organism communities (Frank and Pace 2008;Nakamura et al. 2008;Nakamura et al. 2011). Molecular methods, including high-throughput sequencing, have been applied to study the relative abundance of bacterial species and subtypes within these communities without the need for cultural isolation. This information could be useful to FSIS. For example, studies could characterize bacterial communities in feces, meat, rinsates, and the processing environment, or studies could characterize the potential for enrichment bias associated with commonly used laboratory methods, such as those initiated by FDA-CFSAN (Pettengill et al. 2012). Although FSIS procedures require a cultured isolate before the Agency takes regulatory action based on the “M1” or “P1” definitions of adulteration (9 CFR 301.1, 381.1, and 590.5), metagenomic data could be used to support other decision-making by the Agency as described in Section 2.8.

<back

5.2 STEC

More than 160 serologically distinct STEC have been identified, but only a subset of these have been definitively implicated with illness (Bettelheim 2007). Thus, the ability to produce Shiga toxin does not by itself render *E. coli* pathogenic; the presence and expression of additional virulence factor genes are required to cause human illness (Wickham et al. 2006). In general, STEC serotype O157:H7 is associated with large foodborne outbreaks and with severe outcomes (including death). Additional STEC serogroups (defined by their O-antigen) more commonly are recovered from clinical

cases in the United States (Brooks et al. 2005). Non-O157 STEC serogroups have been associated with less severe outcomes. However, some infections with non-O157 STEC have led to severe complications, including HUS and death; these strains have been responsible for outbreaks with severity comparable to *E. coli* O157:H7 outbreaks, including a STECO111:NM[21] outbreak associated with an Oklahoma restaurant (Bradley et al. 2012) and a STEC O104:H4 outbreak in Europe associated with sprout consumption (Frank et al. 2011;Mellmann et al. 2011). Seventy to 80% of U.S. clinical non-O157 isolates are serogroups O26, O45, O103, O111, O121, and O145 and form the basis of the current FSIS method for detecting non-O157 STEC (Fratamico et al. 2011;FSIS 2012a;FSIS 2012c). STEC strains not belonging to the top six serogroups or not containing the *eae* (intimin) gene have been isolated from cattle and beef products (Galli et al. 2010;Hussein 2007). This type of strain was responsible for the 2011 European sprout outbreak, but is not identifiable by using the current FSIS method (FSIS 2012a). Other regulatory authorities and industry eventually may rely on more discriminatory methods that seek to distinguish strains with higher or lower pathogenicity, rather than take the approach of FSIS, which is more inclusive. For FSIS, it may become important to identify genes or sequences that are preferentially associated with pathogenic STEC (pSTEC) or enterohemorhagic *E. coli* (EHEC) to facilitate future diagnostic method development.

Adherence of STEC to cells lining the intestinal tract is an important step in the disease process. Strains of *E. coli* O157:H7 and certain non-O157 STEC form attaching and effacing (A/E) lesions, which reshape the human cells, recruit immune cells, and allow Shiga toxin to be transferred to the circulatory and organ systems, resulting in severe outcomes including HUS. Coordinated expression of genes (including but not limited to intimin) found in a specific region in *E. coli* O157:H7 and other STEC strains and referred to as the “locus of enterocyte effacement (LEE) pathogenicity island” are necessary to form an A/E lesion. In LEE-negative STEC strains, other genes substitute for the absence of LEE. Genetic analyses have identified adhesion genes in STEC serogroups O104, O113, and O174 that appear to allow for attachment in the absence

[21] NM indicates the organism is non-motile.

of intimin, and these serotypes have been associated with sporadic and outbreak associated illness in Europe and Australia (Bielaszewska et al. 2011;Paton et al. 1999;Tarr et al. 2008). These additional virulence genes could be used to expand the scope of the pathogenic STEC detected by FSIS tests. Other genes have been associated with pathogenic STEC strains include alpha-hemolysin (*hyl*), serine protease (*esp*), catalase peroxidase (*kat*), immunomodulator (*lif*), subtilase cytotoxin (*sub*), cytolethal distending toxin (*cdt*), and EHEC enterohemolysin (*ehx* or *E-hly*)(FSIS 2012c). In addition, O-islands (OI), which are chromosomal DNA segments first discovered in the whole genome sequence of *E. coli* O157:H7 but absent in the whole genome sequence of a non-pathogenic *E. coli* strain (Perna et al. 2001), could be used to distinguish pathogenic and non-pathogenic strains.

Karmali (Karmali et al. 2003) first proposed the concept of sero-pathotype categories based on reported frequencies of serogroups in human illness of known associations with outbreaks and with severe disease, such as HUS and hemorrhagic colitis. Five sero-pathotype categories were proposed. Category A, which includes *E. coli* O157:H7 and *E. coli* O157:NM, occur at relatively high incidence, are commonly associated with outbreaks, and are associated with severe disease. Category B, which include some but not all non-O157 strains detected by the FSIS method, occur at moderate incidence, are uncommonly involved with outbreaks, and are associated with severe illness. Infections with the category C sero-pathotype occur at low incidence, are rarely involved with outbreaks, but may be associated with severe illness. Category D occur at low incidence, are rarely involved with outbreaks, and do not cause severe illness. Lastly, category E does not cause human illness. A number of virulence genes or other markers could be used to place STEC into these types of categories and therefore distinguish pathogenic STEC (also called EHEC) from non-pathogenic STEC (Bugarel et al. 2011;Coombes et al. 2008a). In particular, specific OI genes, such as the *nle* (non-LEE encoded effector) genes, seem to be associated with the Karmali sero-pathotype categories (Bugarel et al. 2010a;Coombes et al. 2008b;Delannoy et al. 2012;Imamovic et al. 2010;Konczy et al. 2008). A number of methods have been developed to detect pSTEC or EHEC, or to distinguish STEC subtypes including *E. coli* O157:H7.

FSIS is aware of STEC subtyping methods that are under development for commercial applications. Pall has developed PCR assay platform (GeneDisk) for a variety of STEC virulence factors and OI genes. This assay performs multiplex real-time PCR in a plastic tray engraved with reaction chambers preloaded with dried reagents (Beutin et al. 2009). The specific genetic targets can be changed to meet research or regulatory testing needs. Currently, ARS and the French Food Safety Regulatory Agency have evaluated the GeneDisk STEC assay (Beutin et al. 2009;Bugarel et al. 2010b;Fratamico and Bagi 2012) to detect the presence of bacteria that carry *stx* genes in foods. A recently described GeneDisk assay targets genes and probes for 12 O-groups (including the top six strains discussed earlier), 7 H- types, and 15 virulence genes, including *stx* (1 and 2), *eae*, plasmid- and chromosomally-associated genes, and *nle* genes associated with O157, O71, and O122 (Bugarel et al. 2010b). Neogen is developing a mass-spectroscopy-based test, called Neoseek, which could detect more than 70 STEC gene targets, including O-antigen, H-antigen, *stx*, *eae*, virulence factor genes, and OI genes simultaneously from cultures and isolates[22]. Roka Bioscience is developing a commercial assay that claims to distinguish pathogenic and non-pathogenic STEC strains.

The GeneDisk and Neogen assays initially were designed to confirm isolated colonies of STEC, but also have been used on food enrichment broths with good results[23]. FSIS may consider these or other assays for additional screening of enrichment broths as a way to monitor associations and trends in the occurrence of virulence markers. One drawback of using multiplex (many targets) assay to screen food enrichment that contains a large population of mixed bacteria is the inability to determine whether all of the detected markers originated from a single bacterium or from various bacterial cells. From FSIS' perspective, it is critical that the organism is isolated and shown to carry all the relevant virulence and trait markers. To address this problem, immunomagnetic separation (IMS) beads that use various O-specific antibodies can be used to capture

[22] Dr. Eden Hosking, Senior Research Scientist, R&D, Molecular Biology, Neogen Corporation, May 6, 2012
[23] Dr. Pina Fratimico, Personal communication, October 27, 2010

organisms of specific serogroups. PCR then can be performed on the separated fraction containing the serogroup to confirm that they carry trait virulence genes. Although useful, IMS is not a perfect solution as each serogroup is complex and comprised of both known pathogenic and non-pathogenic serotypes, and some non-specific binding will occur. Currently, the only reliable confirmation method is to culturally isolate the organism and confirm the presence of Shiga toxin and a certain combination of virulence factor genes.

<back

5.3 *Campylobacter*

Campylobacter is a leading cause of foodborne illness in the U.S. Infection can lead to clinical outcomes ranging from diarrhea to debilitating arthritis. Most cases of human campylobacteriosis are sporadic (Altekruse et al. 1999). Outbreaks have been associated with ingestion of unpasteurized milk, water from a contaminated municipal source, and undercooked poultry (Blaser 1997).

Campylobacter species are identified and differentiated using phenotypic, biochemical, and PCR methods (FDA 2001;FSIS 2011b). FSIS uses direct microscopy and a commercially-available latex agglutination assay to identify isolates belonging to species *coli, jejuni, or lari* (FSIS 2011b). A method for serotyping *C. jejuni* is based on the detection of heat-stable antigens (Moran and Penner 1999). Genotypic methods include AFLP, MLST, PFGE, ribotyping, SNP profiling, random amplification of polymorphic DNA (RAPD), matrix-assisted laser desorption/ionization time of flight with mass spectroscopy (MALDI-TOF-MS), confocal micro-raman spectroscopy, high-resolution melting (HRM) analysis, PCR–denaturing gradient gel electrophoresis (PCR-DGGE), and PCR–single strand conformation polymorphism (PCR-SSCP), (Ahmed et al. 2012;Eberle and Kiess 2012;Lu et al. 2012).

Although PFGE has been referred to as the “gold standard” for the molecular typing method for *Campylobacter* (Foley et al. 2009a;Nielsen et al. 2000;Nielsen et al. 2010;Sails et al. 2003), a NACMCF panel countered, “…while a number of subtyping

methods have been used with *Campylobacter* species (serotyping, antibiotic resistance, MLST, PFGE, flagellin (*flaA*) sequencing, etc.), none have yet been sufficiently discriminatory to be generally applicable as a gold standard" (NACMCF 2007). Routine PFGE subtyping of sporadic case-patient isolates is of limited value because of the high genetic diversity and non-clonal population structure of *C. jejuni* and *C. coli* (Gerner-Smidt et al. 2006;Wilson et al. 2009).

MLST is used to provide greater discriminatory power. One MLST method (Dingle et al. 2001) includes the sequences of seven housekeeping genes[24]. However, this method is insufficient for differentiating *Campylobacter* strains, especially among *Campylobacter coli*[25]. Another group (Zautner et al. 2012) distinguished six *C. jejuni*-groups by combining MLST with 14 additional markers to distinguish a group with higher virulence for humans (higher prevalence, bloody diarrhea, and hospitalization) from another group that was more prevalent in animal hosts and associated with less severe campylobacteriosis.

Campylobacter subtyping methods are needed to identify and trace the source of foodborne outbreaks, to more fully understand the ecology of this pathogen (e.g., transmission routes and vectors in the food chain, the emergence of novel, including antibiotic resistant, strains); and to understand how strain variation contributes to disease progression and development of debilitating clinical outcomes, including Guillain-Barré (GBS), Miller-Fisher, and Reiter's syndromes (Altekruse et al. 1999). At present, the contribution of *Campylobacter* strain variation to disease manifestation and ecology is not clear (Ahmed et al. 2012), although there are interesting findings. Proteins responsible for colonization, adherence and invasion, induction of host cell death, and one toxin, cytolethal distending toxin, were detected in the genome of *C. jejuni* (Dasti et al. 2010). Whole genome sequence analysis of multiple strains indicated a correlation between *C. jejuni* genomic content, particularly in surface-coding regions,

[24] Seven housekeeping genes: aspA (aspartase A), glnA (glutamine synthetase), gltA (citrate synthase), glyA (serine hydroxymethyltransferase), pgm (phosphoglucomutase), tkt (transketolase), and uncA (ATP synthase α subunit)
[25] Dr. William Miller, personal communication, October 7, 2010

and its capacity for environmental survival (On et al. 2006). In addition, strains with sialic acid modified lipooligosaccharide (LOS) mimic host gangliosides, induces the production of autoantibodies and leading to the development of GBS (Godschalk et al. 2007). Subtypes from FSIS surveillance samples containing genes, or otherwise strongly associated with disease progression could be of greater concern to the agency. Subtype data could be incorporated into risk-based sampling algorithms, or follow up inspection programs.

<back

5.4 *Listeria monocytogenes (Lm)*

Lm strains have substantial genetic diversity which influences traits such as environmental persistence, resistance to sanitizers, as well as human and animal virulence. Using subtyping methods, research groups have discovered subtype clades which correlate with differences in ecology and virulence. These clades include four lineages, four major serogroups, and seven epidemic clones (EC). Of the four *Lm* lineages (called lineage I, II, III, and IV), only lineages I and II are commonly isolated from foods and cases of human listeriosis. Lineage III is commonly isolated from animals, but is not a significant contaminant of foods, nor does it appear to cause illness. The ecology and virulence of lineage IV are poorly understood because it is very rarely isolated. The four major serotype groups (4b, 1/2a, 1/2b, and 1/2c) account for the majority of food contaminants, but three (4b, 1/2a, and 1/2b) are significantly associated with human listeriosis. Serogroup 4b is responsible for many outbreaks and for about 46-49% of sporadic listeriosis cases reported to CDC (CDC 2013). Recently, major serogroup 1/2a has been associated with large outbreaks, including outbreaks linked to FSIS-regulated foods (CDC 2011a;CDC 2011c;Gaul et al. 2012;Gilmour et al. 2010). Lastly, seven EC have each been associated with multiple outbreaks even though they are rare contaminants of food. EC types I, II, and IV belong to lineage I and major serogroup 4b. EC types III, V, and VI belong to lineage II and major serogroups 1/2a and 1/2b. Investigators believe EC possess hyper-virulent properties because these clades have been isolated from multiple outbreaks. Of particular concern to FSIS, EC type II has been associated with two large multistate outbreaks involving FSIS

regulated foods (1998-1999 outbreak linked to contaminated hot dogs, and a 2002 outbreak linked to contaminated turkey deli meats). EC type III is thought to have persisted in an FSIS-regulated establishment over a 12 year period. ECIII serotype 1/2a isolates have caused outbreaks linked to contaminated hot dogs (U.S. 1989) and turkey deli meats (U.S.2000) manufactured in the same food processing plant. Six of the seven EC types have been identified in poultry processing facilities. (Chen et al. 2005;Chen et al. 2007;Kathariou et al. 2006;Lomonaco et al. 2013;Orsi et al. 2008;Verghese et al. 2011).

Lm has a complicated life history owing to its ability to survive in diverse locations (on non-living surfaces, on vegetation and soil, and in animals and humans) and its ability to invade and cause disease in animals and humans. In any given subtype, genes evolve together and influence each other. The outcome may not be predictable by the presence or absence of any single gene but by specific mutations that influence the ecology or pathogenicity within the subtype. Certain genes influence the ability of *Lm* to survive outside and inside of a host animal or human. A regulator of gene expression, *prfA*, plays an important role that allows *Lm* to switch between life outside and inside a host in response to environmental stimuli, including temperature and the availability of certain nutrients (Freitag et al. 2009). Other genes have been studied for their roles in *Lm* ecology and virulence. The internalin (*inl*) genes are responsible for allowing *Lm* to bind and invade animal or human cells, an important step in developing listeriosis. Certain mutations in *inlA* greatly influence virulence as demonstrated by the small number of these mutations in strains from clinical cases and the abundance in strains from food isolates (Jacquet et al. 2004;Tamburro et al. 2010;Van et al. 2010). The importance of *inlA* in *Lm* virulence also has been demonstrated in human cell lines and animal models (Chen et al. 2011;Nightingale et al. 2005;Van Stelten et al. 2011). *InlA* mutations do not render *Lm* completely avirulent, but they are important markers for less virulent subtypes. Notably, *inlA* mutations cannot explain completely why certain serotypes are responsible for most sporadic and outbreak associated listeriosis.

Based on an analysis of subtype diversity among 501 *Lm* isolates from FSIS' RTE sampling program, 34 (6.8%) isolates were from the 4b complex, which includes serotype 4b. Epidemic clones from previously documented outbreaks were identified in 38 (7.6%) isolates, and *InlA* mutations were identified in 243 (48.5%) isolates, which is consistent with previous surveys (Ward et al. 2010a). This survey indicated the range of subtypes with important differences present in FSIS-regulated products for host, ecology, and virulence properties. Information on *Lm* subtypes present in FSIS-regulated foods can be used to prioritize or target additional resources to establishments where subtypes that are more likely to be associated with sporadic or outbreak-associated illnesses have been identified (Section 4.3).

The distribution of *Lm* subtypes in foods is likely to have an important role in risk assessments, including estimates of dose-response and risk associated with specific foods. Older risk assessments sponsored by the World Health Organization (WHO), FDA, and FSIS employed a mouse animal model with a single *Lm* serotype 4b strain to generate the dose response curve and used additional strains to estimate variability (FAO/WHO 2004;FDA/USDA. 2003). The dose response relationship was multiplied by seven orders of magnitude to account for levels of *Lm* likely to be consumed by people during listeriosis outbreaks. Food surveys indicate different relative subtype frequencies in different foods and in clinical patients (Chen et al. 2006;Van et al. 2010;Ward et al. 2010a), possibly due to different ecological and survival characteristics among *Lm* subtypes. Using food survey and epidemiological data, Chen provided subtype-specific dose response estimates for different lineages and ribotypes, which differed by three or four orders of magnitude (Chen et al. 2006). Newer dose-response models have employed pregnant non-human primate, guinea pig, and hamster animal models, along with different *Lm* subtypes with *inl*A mutants. These models, along with improvements in the choice of animal model, disease endpoint, experimental set-up, or substitution of epidemiological data for animal data, may lead to considerable improvements to *Lm* dose-response modeling (Hoelzer et al. 2013).

<back

5.5 *Salmonella*

Traditionally, *Salmonella* diversity has been defined by serotyping and phage typing. *Salmonella* serotypes are distinguished using antisera to four highly variable molecules on the surface of the *Salmonella* bacterium: O (lipopolysaccharide); two H (flagella) variants, H1 and H2; and a virulence antigen, Vi. More than 2,500 unique combinations of O, H1, H2, and Vi types have been discovered (WHO 2007). Phage typing is performed by determining the resistance or sensitivity to infection by an established set of 34 bacterial viruses (phages) (Anderson et al. 1977). There are 209 definitive phage types, referred to as DT (Lan et al. 2009a).

Because *Salmonella* serotypes or phage types are based on a limited number of markers (e.g., O, H, and Vi antigens, as well as phage resistance/sensitivity genes), these methods are imperfect indications of *Salmonella* diversity. Sequence diversity within a serotype or phage type could be due to genetic divergence (i.e., a common ancestor acquired the serotype or phage type) or the independent acquisition of antigens and genes by diverse clones (i.e., distinct clonal groups independently acquire the serotype or phage type). Both of these mechanisms have been demonstrated. Multiple sequence types (STs) have been observed within the same serotype and phage type. For example, genetic studies indicate that the *S.* Typhimurium phage type DT104 is composed of multiple independent clones; although multidrug resistant DT104, which appeared in the United Kingdom in the 1980s and acquired resistance to antimicrobial compounds ampicillin, chloramphenicol, streptomycin, sulfonamides, and tetracycline, likely has a single origin (Cooke et al. 2008;Lan et al. 2009a). Conversely, a single ST may include two closely related serotypes. For example, serotypes *S*. I 4,[5],12:i- and *S*. Typhimurium share a common sequence type ST6 (Alcaine et al. 2006). Newer sequence-based methods (virulence gene profiling, expression profiling, SNP, and WGS) can provide a more specific and scientifically-accurate definition of diversity within *Salmonella*. Measures of genetic diversity within *Salmonella* could be used to detect and track on-farm and illness or outbreak related clusters to support

illness attribution and define pathogenic subtypes with greater precision, as described below.

<back

5.5.1 Use of Subtyping to Track On-Farm and Illness Related Clusters

As described in Section 3.1.1, PFGE, the current gold-standard technique for detection of *Salmonella* clusters, is laborious, and the process of PFGE pattern comparison requires precise standardization and subjective interpretation of band matches. In addition, *Salmonella* PFGE is of limited effectiveness for detecting on-farm or illness clusters with certain serotypes, including Enteritidis and Hadar, which are dominated by a limited number of PFGE pattern subtypes (Pang et al. 2007;Zheng et al. 2011). Sequence-based subtyping methods can increase the discrimination of certain serotypes by distinguishing clonal lineages within each serotype. By improving discrimination, these methods will improve the epidemiological relevance, which is the ability of a method to discriminate outbreak-associated from sporadic case-patients that occur over the same timeframe or geographic location, or to reliably detect clusters associated with a single farm or processing facility In one study, PFGE and MLVA developed for *S.* Enteritidis each distinguished 34 isolates into 13 subtypes; the discriminatory ability, judged by Simpson's and Shannon's Discriminatory Index, was higher for MLVA because the groups were more evenly distributed (Cho et al. 2007). Combinations of methods, such as phage typing and MLVA (Cho et al. 2010), PFGE and MLVA (Broschat et al. 2010), or the use of multiple restriction enzymes (Zheng et al. 2011) may be useful for increasing discrimination of serotypes, especially Enteritidis and Hadar. However, increasing the number of assays to determine subtype may increase the overall probability of an incorrect conclusion due to compounding errors from each assay (Call et al. 2008).

<back

5.5.2 Use of Subtyping to Support Salmonellosis Attribution Estimates

There are considerable differences in the distribution of *Salmonella* serotypes isolated from clinically ill humans and farm animals (Alcaine et al. 2006;Foley et al. 2006;Sarwari

et al. 2001), which indicates non-farm animal origins for some pathogenic *Salmonella*e. Attribution of human *Salmonella* illnesses to general food classes can be accomplished by analyzing data from foodborne disease outbreaks. Such analysis has demonstrated substantial variability in the share of illnesses attributed to foods in the aquatic animal, land animal, and plant classes for six *Salmonella* serotypes (Painter et al. 2013). For example, outbreaks suggest that 80% of serotype Enteritidis illnesses are attributed to foods in the land animal class (predominantly eggs and poultry). Outbreaks suggest that over 75% of serotype Javiana illnesses are attributed to foods in the plant class. There is considerable overlap in serotype occurrence among human and farm animal isolates. The 10 most common *Salmonella* serotypes accounted for a majority (73%) of serotyped isolates from human infections in 10 FoodNet sites in 2009 (CDC 2011b). Nine of these serotypes[26] were among the more commonly identified serotypes in meat and poultry products. As indicated earlier, some serotypes and phage types appear to consist of multiple clones with different origins as determined by genetic sequence analysis, and some serotypes contain more genetic diversity than others. In one analysis performed on publicly-available MLST data, *S.* Newport was the most divergent serotype, followed by *S.* Typhimurium. *SE* had three ST. A majority of isolates belonged to a single sequence type, ST11, which is closely related to *Salmonella* serotypes associated with poultry (Lan et al. 2009b). The other STs were unrelated to the poultry serotypes indicating a possible non-poultry origin. A study of *Salmonella* isolates from clinically ill cattle and humans indicated disparities in ST that were not apparent by traditional serotyping (Alcaine et al. 2006). For example, ST6, which includes *Salmonella* serotypes I 4,[5],12:i:- and Typhimurium, was one of the most common ST among the human and cattle isolates. Serotype 4,[5],12:i:- is considered an emerging serotype that originated from a *S.* Typhimurium ancestor. The clonal population represented by ST6, which *includes S.* Typhimurium, may actually be the emerging subtype that presents a public health threat, at least with respect to contact with cattle and beef. The same study indicated that *S.* Newport isolates consist of two major lineages, type A and type B, which represent cow- and bird-derived lineages,

[26] Enteritidis, Typhimurium, Newport, Montevideo, I 4,[5],12:i:-, Heidelberg, Muenchen, Saintpaul, Oranienburg. Note: Javiana is common in clinically ill humans but uncommon in farm animals.

respectively. Further, MDR *S.* Newport was statistically associated with the cow lineage (Alcaine et al. 2006). These examples indicate that sequence-based subtyping of *Salmonella* strains would provide more specific information that would be useful for illness attribution estimates, as well as defining and tracking lineages that are linked more specifically to serious public health risks.

<back

5.5.3 Use of Subtyping to Differentiate Pathogenic Subtypes of Salmonella

The *Salmonella* genus is very diverse, including two species (*bongori* and *enterica*) and six sub-species within *S. enterica;* 99% of salmonellosis cases are caused by a single subspecies, *Salmonella enterica enterica*, and about 70% of human illnesses are caused by 20 serotypes within this subspecies (Brenner et al. 2000;CDC 2012a;CDC 2012b). Illness outcomes, in terms of invasive disease, hospitalization, and death, differ significantly by serotype (Jones et al. 2008).The intensity of *Salmonella* illness depends on factors related to the host, pathogen, and environment. Under some circumstances, there are no recognizable symptoms. For example, humans are known to chronically carry and shed typhoid and paratyphoid *Salmonella*, sometimes in the complete absence of symptoms. Likewise, many types of non-typhoid *Salmonella* are carried and shed by otherwise healthy animals. In humans, non-typhoidal *Salmonella* infection typically results in gastroenteritis. About 5% of these cases develop bacteremia (blood infection) and focal infection of tissues and organs, resulting in a variety of severe clinical outcomes (Hohmann 2001). Higher rates of bacteremia and focal infection are seen in immunocompromised patients. Although outcomes, including death, depend on underlying comorbidities of bacteremic patients (Hohmann 2001), Dhanoa found higher probability of bacteremia for group D *Salmonella*, including *S*. Enteritidis, among patients at a hospital in Malaysia (Dhanoa and Fatt 2009).

Salmonella virulence factors are encoded by a variety of genetic elements, including plasmids, phages, and *Salmonella* Pathogenicity Islands (SPI). Fourteen SPI have been described. SPI-1 encodes type 3 secretion system 1 (T3SS-1), which causes secretion and translocation of a range of bacterial proteins to the host cell. SPI-2 encodes T3SS-2

that allows intracellular survival and replication (Foley et al. 2009b;Foley and Lynne 2008). Unique combinations of virulence factors may be responsible for certain characteristics associated with specific serotypes, such as the capacity for invasive disease, resistance to antimicrobials, and invasion of hen reproductive organs resulting in internal egg contamination.

It is difficult to determine the exact combination of genes necessary to acquire a given characteristic in *Salmonella*. In some situations, variation may result from differential expression of virulence genes and not the presence or absence of specific genes. For example, a study examining internal egg contamination by *SE* used RAPD, plasmid profiling, phage typing, PCR detection of 30 virulence genes, and relative expression levels for two genes (*agfA* and *fimA*), and was unable to distinguish *SE* strains of egg and non-egg related origin (Botteldoorn et al. 2010), but did find a limited role for one gene (type I fimbrae encoded by the *fim* operon[27]) (Botteldoorn et al. 2010). On the other hand, experimental deletion of an *SE* gene (SPI-2 regulator, coded by the gene ssrA) that controls the expression[28] of other genes drastically reduced SE colonization of the reproductive tract but not the gut (Bohez et al. 2008), indicating a potential role for gene expression as a factor determining whether *SE* strains are capable of invading the internal contents of eggs. Similarly, ARS researchers determined that the cyaA gene, which encodes adenylate cyclase that produces cyclic AMP and pyrophosphate from ATP, is variable in *SE* and separates highly-related PT4 strains into subtypes that either invade or do not invade the internal contents of eggs (Morales et al. 2007).

If many genetic factors are involved, a summary profile may be necessary to distinguish virulent subtypes within a serotype. To date, such a profile has not been identified. A recent study using microarray technology demonstrates this difficulty (Litrup et al. 2010).

---

[27] Fimbrae (plural of fimbria) are proteinaceous structures on the surface of certain bacteria that allow the bacteria to adhere to other bacteria, animal cells, or non-living surfaces.

[28] Bacterial genes are expressed through a process that involves copying the gene to messenger RNA, and in some cases, translating the genetic code on the messenger RNA to create a protein with a structure or activity that contributes to the properties of the bacterium.

In this study, the authors attempted to link the molecular data obtained by microarray analysis of *S.* Typhimurium strains to detailed epidemiological and clinical patient data of the patients previously infected with those strains. The strains were selected from patients with mild infections and from patients with severe infections; clinical data allowed the authors to correct for known underlying diseases and patient age. Strains were analyzed for the presence or absence of 281 genes covering marker groups related to pathogenicity, phages, antimicrobial resistance, fimbriae, mobility, serotype, and metabolism. The authors showed that *S.* Typhimurium strains causing very different symptoms in patients had little genomic variation, and the observed variation did not correlate to the severity of disease. The Litrup study, unlike a previous study (Fierer et al. 1992), also failed to show an association between the presence of the *Salmonella* virulence plasmid, pSLT, and strains causing severe illness, including systemic illness.

In summary, a variety of sequence-based subtyping methods are available or are under development to discriminate and provide information on *Salmonella* isolates that can be used for cluster detection, illness attribution, and defining pathogenic lineages. At this point, a reliable virulence gene screening profile for *Salmonella* is unavailable. A genetic definition of pathogenic *Salmonella* may depend on specific virulence genes or SNP responsible for specific characteristics associated with virulent strains. These characteristics include the capability for systemic infection, reproductive tract colonization, and antibiotic resistance. Serotyping (and phage typing) is currently used as a surrogate to provide information on the potential for an isolate to cause human illness. Although these methods are not tightly linked to virulence factors (Section 2.4), they are a reasonable first cut for differentiating isolates of human health significance (Jones et al. 2008).

<back

6 Conclusions

FSIS' microbiological sampling and testing programs are designed to verify that establishments maintain control of their production processes and to ensure that they adhere to FSIS regulations, policies, and performance standards ((FSIS 2011c). FSIS'

laboratories have implemented subtyping procedures, including PFGE, MLVA, and molecular serotyping of selected STEC serogroups and *Salmonella* serotypes. OSEL is the primary unit within FSIS that performs subtyping of FSIS isolates, and evaluates new methods developed in collaboration with PulseNet (such as MLVA for additional *Salmonella* serotypes) and ARS (such as TMLGT for *Lm* and non-O157 STEC screening assays). FSIS subtyping methods should be consistent with existing regulatory policy, and the Agency has worked closely with PulseNet to share data and to develop and validate methods.

In 2013, PFGE subtyping is applied to all *Lm*, *Salmonella*, STEC, and *Campylobacter* isolates. MLVA methods for *E. coli* O157:H7 and *Salmonella* serotypes supplement, but do not replace, PFGE. New MLVA methods for *Salmonella* serotypes, STEC serogroups, and *Lm* are being developed or validated, but these provide data that are complementary to PFGE and therefore would not be appropriate replacements for PFGE. PCR methods are used to screen samples for declared STEC adulterants. Traditional and molecular serotyping assays are used to determine *Salmonella* sampling frames. In the future, SNP based methods, such as TMLGT, may be employed to improve the Agency's risk-based sampling algorithms.

PFGE and MLVA are designed to detect clusters, and to accurately measure statistical association of illnesses with a variety of exposures allowing public health agencies to investigate and identify exposures of concern, including foodborne exposures. PFGE clusters obtained from FSIS' surveillance samples are used to support evidence for pathogen transmission and persistence in regulated establishments. These observations suggest the need for additional pathogen control through the establishment's existing food safety systems. FSIS collaborates with other Federal, state, and local public health agencies through PulseNet to share PFGE and MLVA data. Since its inception in the late 1990s, the PulseNet network has detected many multi-state or multi-jurisdictional clusters, often associated with the consumption of foods under Federal inspection. Identification of PulseNet clusters lead to outbreak and

trace back investigations and ultimately to the implementation of effective public health actions, including product recalls.

This paper describes FSIS' current procedures for performing PFGE and MLVA and for communicating these data throughout the Agency. FSIS will continue to improve its ability to respond to clusters detected in the PulseNet system by developing guidelines for interpreting subtype data based on its own experience with outbreak and trace back investigations. This White Paper provides current guidance for interpreting PFGE and MLVA data. The lessons learned during future investigations should be used to continually update and refine this guidance.

FSIS' use of subtyping data extends beyond outbreak cluster detection. FSIS also uses a traditional serotyping procedure and a multiplex PCR-based method (Bioplex) to detect *Salmonella* serotypes in samples collected in its PR/HACCP verification program. The Bioplex method was implemented to reduce the time-to-result for certain serotypes (FSIS 2013a). The Agency's risk-based algorithm for determining the monthly PR/HACCP sampling frame considers the number of human health serotype isolates identified in samples from the establishment's most recent set (FSIS 2013b). Human health serotypes are defined as the top 20 serotypes identified most recently among clinical isolates by the PHLIS (CDC 2011b). As FSIS has not established a regulatory performance standard for *Salmonella* serotypes, this information is used only to schedule future sets, not for determining whether the establishment has passed or failed the set.

FSIS has used molecular subtyping methods for detecting adulterants in food samples. In 1994, FSIS defined *E. coli* serotype O157:H7 as an "unusual and urgent food safety problem" and declared that ground beef containing this serotype was adulterated within the meaning of 9 CFR 301.2(1), meaning that the detection of this serotype in ground beef is interpreted as an added poisonous or deleterious substance (FSIS 2012c). In subsequent years, FSIS expanded this definition to include beef cuts (used for preparing ground beef) and other "non-intact" beef products (e.g., mechanically

tenderized beef). In 2011, FSIS expanded the definition to include six additional non-O157 STEC serogroups. The FSIS laboratory method for detecting *E. coli* O157:H7 and non-O157 STEC serogroups includes real time PCR assays. The *E. coli* O157:H7 screen is a commercially available assay for DNA sequences closely associated with the *E. coli* O157:H7 clade, but not with many other *E. coli* groups[29]. The non-O157 STEC screen tests are specific for O-antigen and Shiga toxin genes. These assays allow FSIS to focus on samples most likely to contain adulterants.

FSIS is working with CDC and FDA to generate preliminary foodborne illness source attribution estimates of *Salmonella* serotypes for specific FDA- and FSIS-regulated commodities [30](IFSAC 2012). Subtype data is useful for understanding the source of biological variation that improves risk assessments and attribution studies including variation in virulence, dose response, environmental reservoir, growth and inactivation kinetics in foods, and resistance to various food processing procedures (Coleman et al. 2004).

Over the past decade, there has been explosive growth in the development and availability of sequence based methods for subtyping bacterial pathogens. Researchers with ARS, CDC, FDA, and PulseNet are developing and validating a wide portfolio of subtyping methods for *Salmonella*, *Lm*, STEC, *Campylobacter*, and other pathogens. The 100K Pathogen Genome Project will make 100,000 whole genome sequences for a variety of pathogens, including *Campylobacter*, *Lm*, *Salmonella*, and STEC, available through a publically-accessible database at the NCBI. These data will improve the understanding of variation among pathogen genomes and stimulate the identification of new subtyping tests. FSIS will contribute isolates from its regulatory testing program for sequencing in the 100K Pathogen Genome Project, and looks forward to the benefits that the database could provide federal, state, and local public health agencies.

---

[29] Dupont technical bulletin, Development of the BAX® system multiplex (MP) assay for detecting E. coli O157:H7

[30] IFSAC draft strategic plan 2012
http://www.fsis.usda.gov/wps/wcm/connect/18beaa4a-1105-4a78-91b5-6fc9d7fc5a37/IFSAC_Draft_Strategic_Plan_Attribution.pdf?MOD=AJPERES

Some subtyping methods are ideal for detecting clusters and may, in some situations, be preferable to the gold standard, PFGE. Other methods are optimized for measuring familial relationships or genetic distance between isolates, or they may be useful for distinguishing characteristics important to FSIS. Data from subtyping methods capable of detecting or distinguishing subtype clades with enhanced expression of virulence, historical association with outbreaks, expression of resistance against important antimicrobials, or different ecological characteristics, such as biofilm formation and egg invasion, could be incorporated into FSIS' existing risk-based sampling algorithms. This would allow establishments identified as having a higher risk of producing contaminated products to be sampled more frequently. FSIS may use such data to develop new screening assays specific for adulterants. The Agency will need to calibrate a proper response to these data, keeping in mind prior findings and declarations, especially with respect to existing "zero tolerance" pathogens.

<back

7 Acknowledgements

The members of the FSIS Subtyping Workgroup gratefully thank the following individuals for their timely information and personal communications.

FSIS

William Cray, OPHS Eastern Laboratory

Adrienne Dunham, ODIFP contractor

Eric Ebel, OPHS Risk Assessment and Analytics Staff

Neal Golden, OPHS Risk Assessment and Analytics Staff

Neelam Narang, OPHS Eastern Laboratory

Chuanfa Guo, OPHS Risk Assessment and Analytics Staff

ARS

Jim Bono, Meat Animal Research Center, Clay Center, NE

Mark Englen, Richard B. Russell Research Center, Athens, GA

Jonathan Frye, Richard B. Russell Research Center, Athens, GA

Jean Guard, Richard B. Russell Research Center, Athens, GA

Dayna Harhay, Meat Animal Research Center, Clay Center, NE

Bill Miller, Western Regional Research Center, Albany, CA

Pina Fratamico, Eastern Regional Research Center, Wyndmoor, PA

Todd Ward, National Center for Agricultural Utilization Research, Peoria, IL

CDC

Eja Hyytia-Trees, Foodborne and Diarrheal Diseases Branch, Atlanta, GA

FDA

Eric Brown, Center for Food Safety and Applied Nutrition, College Park, MD

Peter Feng, Center for Food Safety and Applied Nutrition, College Park, MD

Steven Foley, National Center for Toxicological Research, Jefferson, AR

Other

Patrice Arbault, BioAdvantage Consulting, France

Erin Dreyling, Roka Biosciences, Warren, NJ

Eden Hosking, Neogen Corportation, Lansing, MI

Mansour Samadpour, IEH Inc., Lake Forest Park, WA

8 Appendices

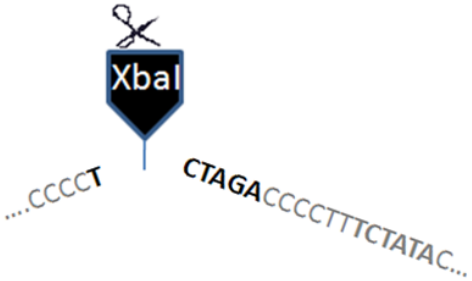
Appendix A: PFGE Theory

Specialized enzymes, called restriction enzymes, are used to cleave bacterial DNA at specific sites to create a collection of well-defined fragments that can be separated by molecular size to create a diagnostic pattern. A variety of restriction enzymes are known to cut DNA. Each enzyme recognizes and cuts DNA at an enzyme-specified sequence. The spacing of these sequences varies within bacteria, creating differences in PFGE patterns. (Note: Bacterial DNA is double stranded, and restriction enzyme cut double stranded DNA. For the sake of simplicity, only a single strand of DNA is shown in the diagrams below.)

.

Figure A1. Site-specific digestion of bacterial DNA using restriction enzymes. The restriction enzyme *Xba*I is used for illustration, but the general idea is the same for all restriction enzymes.

1. The restriction enzyme XbaI recognizes only the DNA sequence TCTAGA .



2. XbaI then cuts at the TCTAGA throughout the bacterial genome.

Figure A2. Between-strain differences in the spacing of restriction sites creates differences in PFGE patterns derived from these strains.

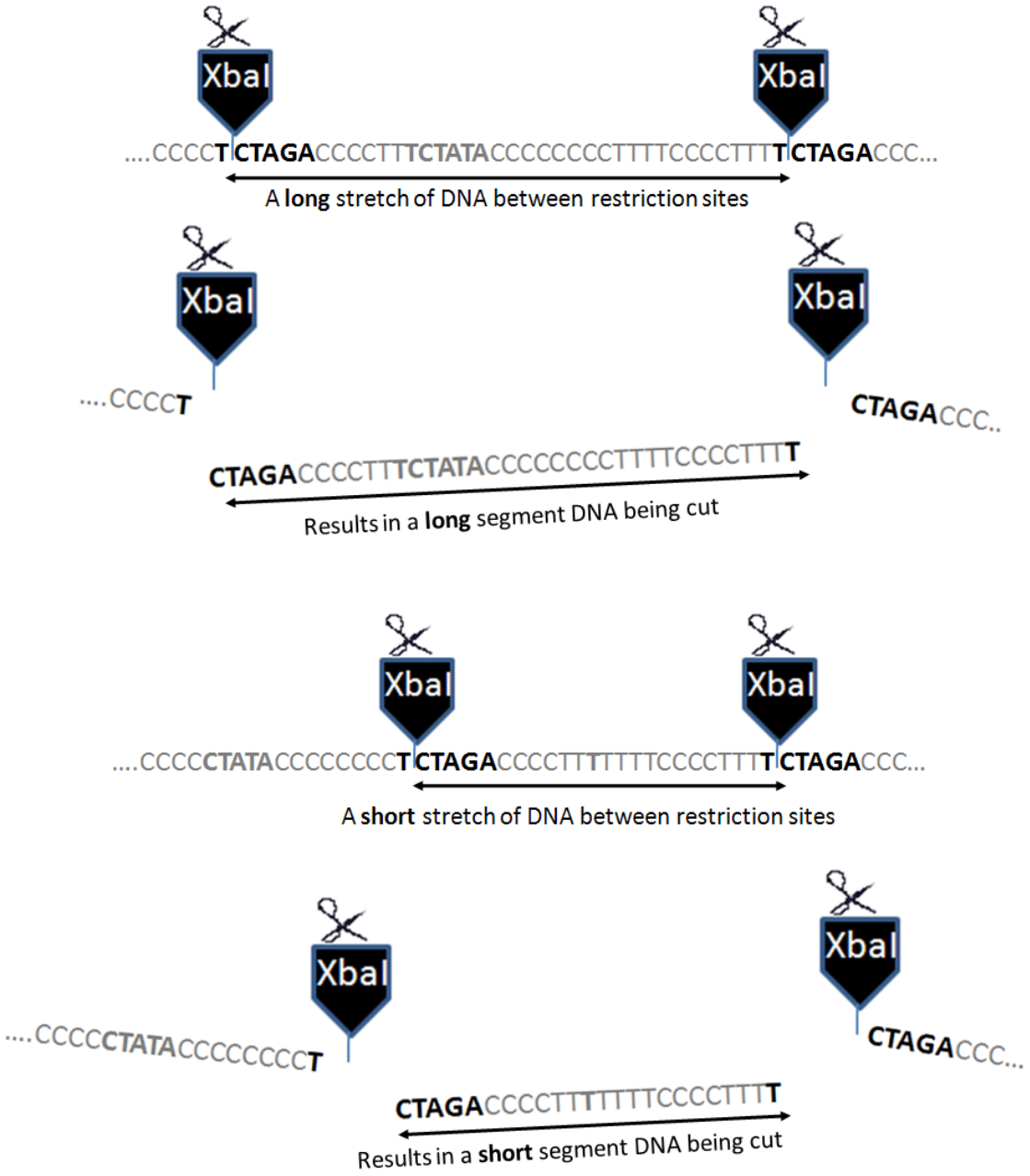


Figure A3. The number of restriction fragments is related to the number of restriction sites. Regions that contain more restriction sites will yield a higher number of DNA segments and more bands.

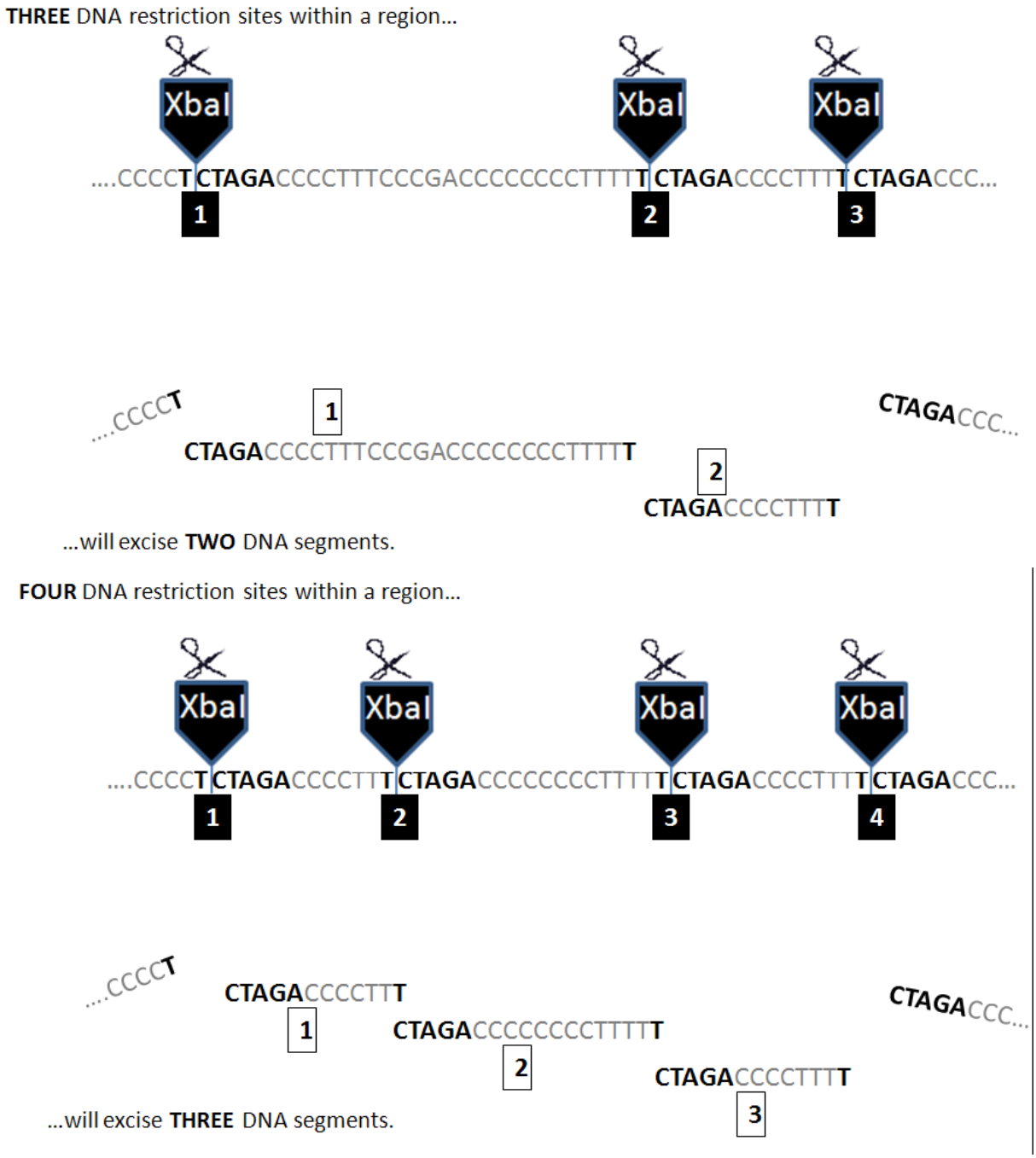


Figure A4. Large DNA fragments derived from restriction digestion of bacterial DNA are visualized by pulsed field gel electrophoresis. A gel image is shown below. Each DNA fragment is visualized as a band on the gel. Smaller fragments travel faster than larger fragments in the electric field, so bands at the top of the gel represent the larger fragments and bands at the bottom of the gel represent the smaller fragments.

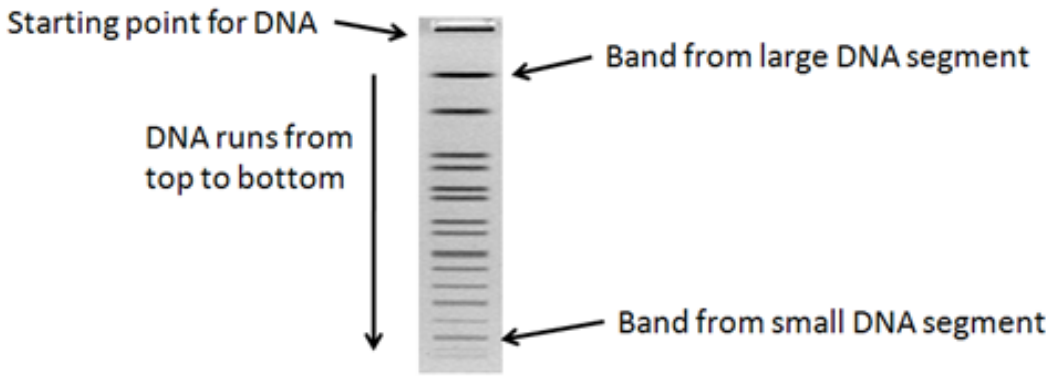
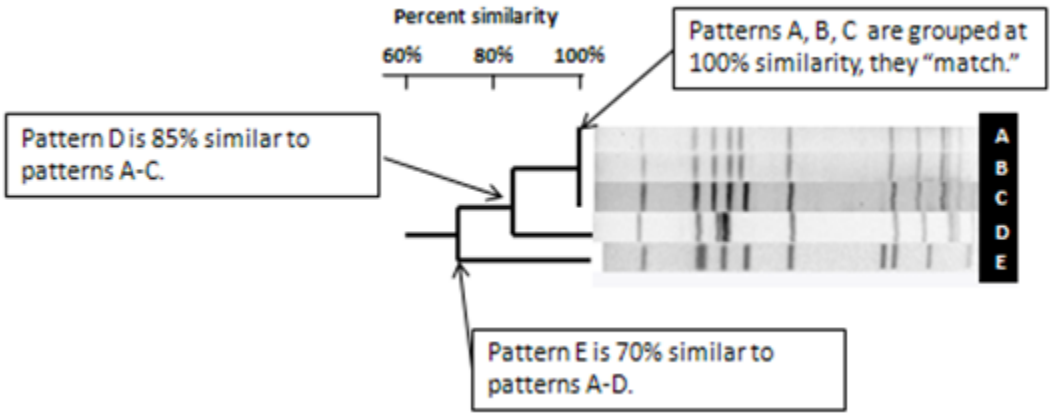


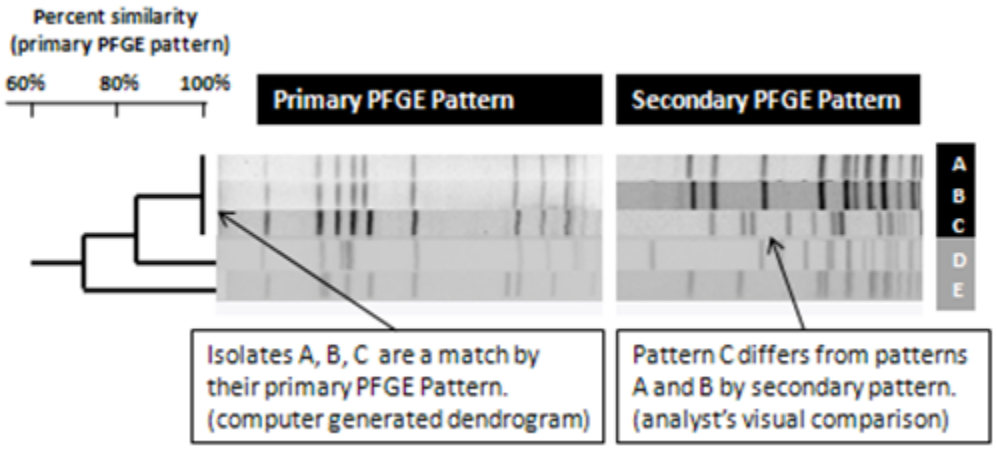
Figure A5: PFGE Pattern Comparison. Five PFGE pattern images labeled A, B, C, D, and E are shown below. The dendrogram on the left side of each figure groups patterns according to their similarity. A percent similarity scale is shown on the top left-hand corner.

<u>One enzyme dendrogram</u>. PFGE patterns A, B, and C are matches and could be considered as an outbreak cluster. Patterns D and E do not match the ABC cluster. D is a closer match than E, but these samples (from case-patients or foods) typically would not be included in the outbreak investigation.



<u>Two enzyme dendrogram.</u> When two or more enzymes are used, there is greater discrimination. The secondary pattern for C does not match A and B. Therefore, C, as well as D and E, would not be included in cluster AB. In this case, the dendrogram was

calculated using the primary, but not the secondary, enzyme pattern. Distinctions among the secondary patterns can be detected visually, or another dendrogram may be calculated.

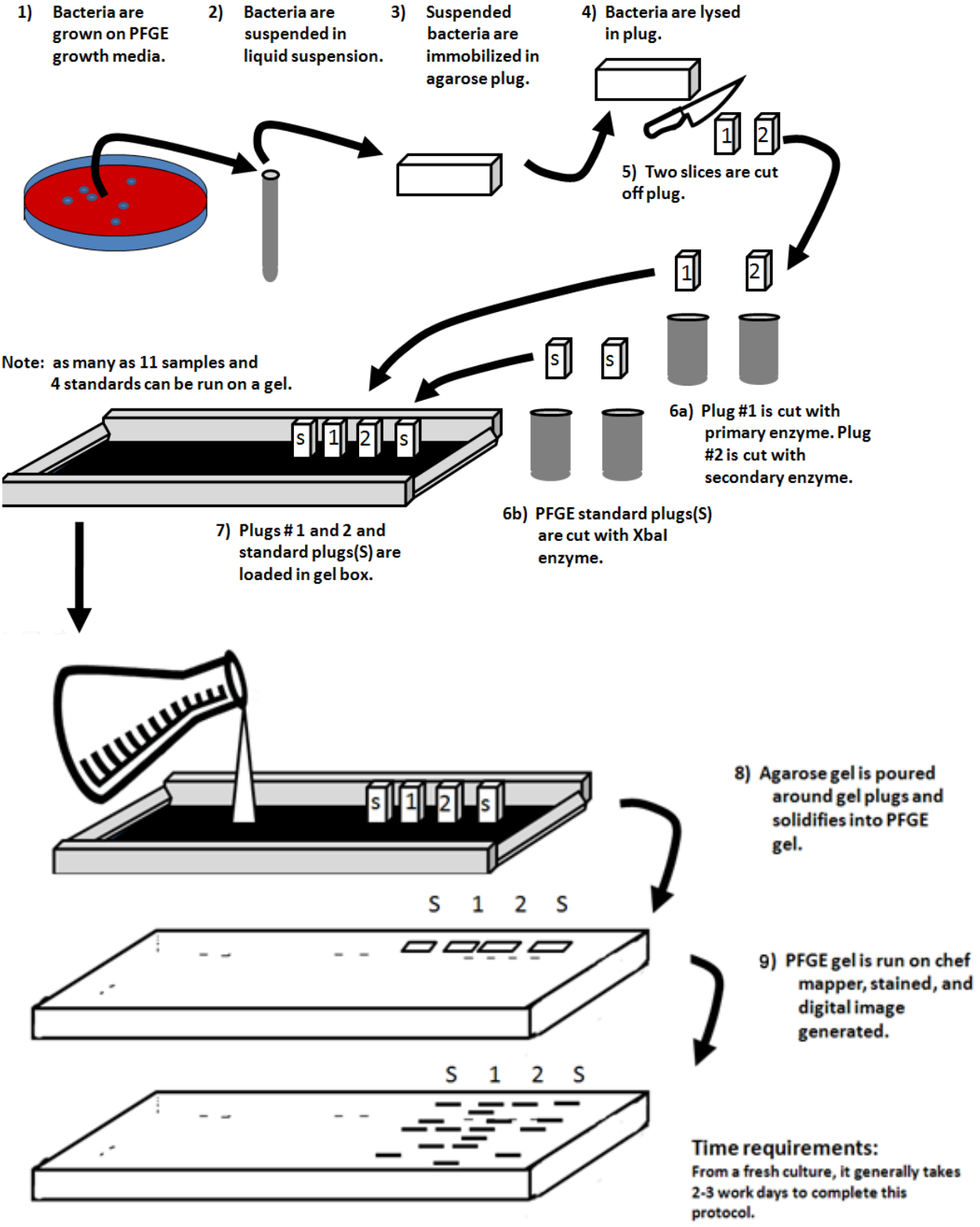


Primary and secondary enzymes. The following primary and secondary enzymes are recommended for use by PulseNet laboratories.

Organism	Primary enzyme	Secondary enzyme
Listeria monocytogenes	*Asc*I	*Apa*I
E. coli O157:H7, non-O157 STEC, and most *Salmonella* serotypes	*Xba*I	*Bln*I
Campylobacter	*Sma*I	*Kpn*I

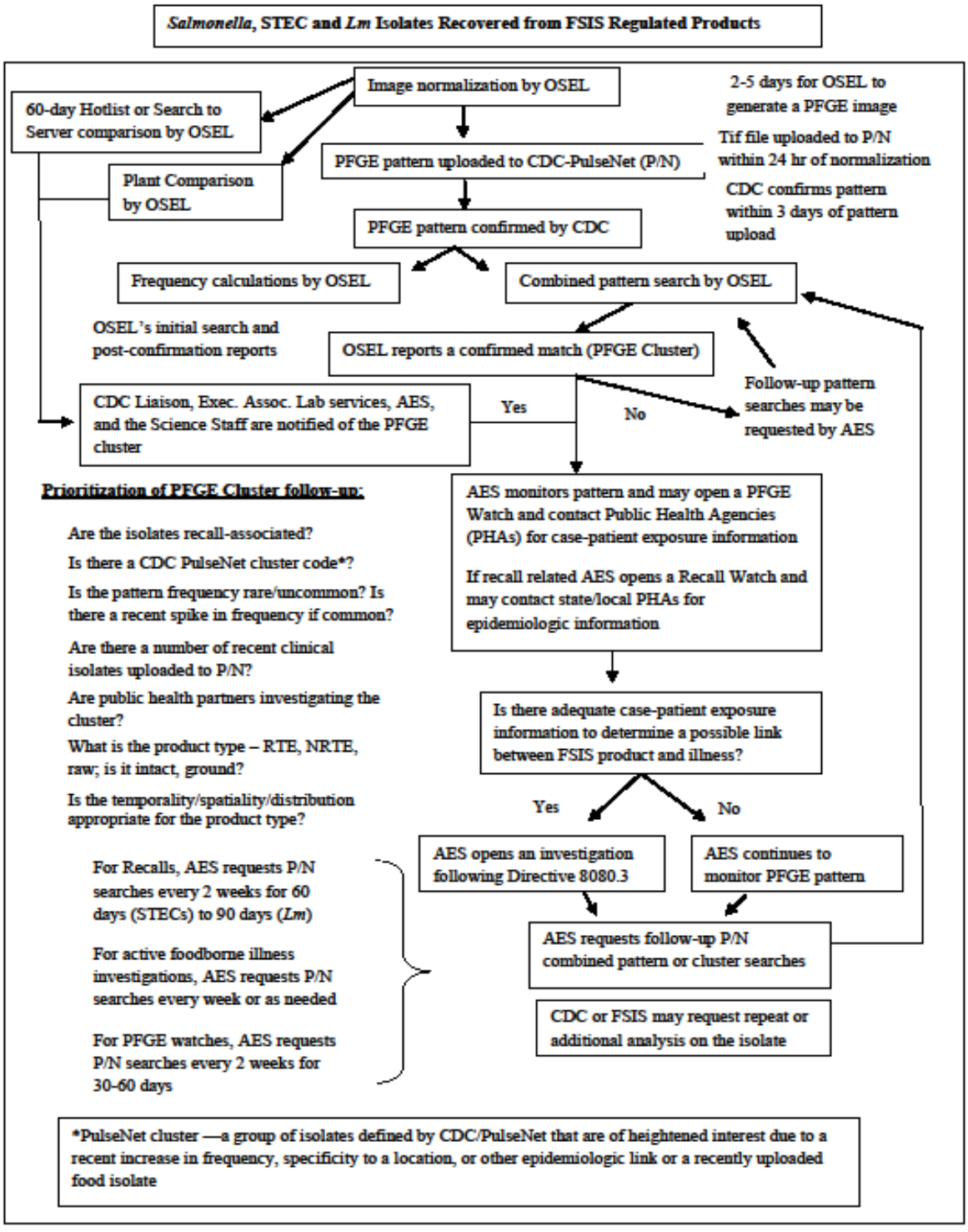
<back

Appendix B: PFGE Laboratory Method: From Isolate to Gel Image



<back

Appendix C: PFGE Data Workflow and Cluster Investigation



<back

Appendix D: PFGE Pattern Record Metadata in OSEL Database, PulseNet, and VetNet

Each PFGE pattern in the OSEL BioNumerics database has certain data electronically linked to it within the database, called metadata. When FSIS patterns are uploaded to the national PulseNet database, only a specific subset of these data are uploaded. Any participating PulseNet laboratory can view information in the PulseNet database.

FSIS metadata available in the OSEL BioNumerics database is listed below. Metadata available to the national CDC-PulseNet database is designated with an asterisk (*).

- OSEL LIMS number (2011-forward)*
- OB (outbreaks) number (discontinued in 2011, was replaced by LIMS number)*
- Species name *
- Serotype *
- Links to routine queries with dates
- Establishment number
- Primary PFGE pattern name (imported from PulseNet) *
- Number of isolates with matching primary PFGE pattern name in PulseNet *
- Secondary PFGE pattern name (imported from PulseNet) *
- Number of isolates with matching Secondary PFGE pattern name in PulseNet *
- Occurrences of isolates with matching primary and secondary PFGE pattern names in PulseNet*
- Outbreak code (imported from PulseNet) *
- Lab MX# (internal laboratory number from Eastern, Midwestern, and Western laboratories)
- Form number
- Establishment state *
- Establishment city
- Establishment country *
- Date sample taken at establishment *
- Date sample received at OSEL *
- Source site (ground beef, workers' gloves, etc. NO BRAND NAMES) *
- Source type (food, environmental, swab) *
- Foreign establishment number
- OSEL gel name
- Lab name *
- Upload date *

Metadata for human clinical isolates that is uploaded to the PulseNet
In addition to the metadata listed above, State and local health departments often provide additional data including patient information linked to the isolates that they upload to the national CDC-PulseNet database. This information is available when OSEL downloads clinical sample PFGE patterns from the national database. This information may be useful during epidemiological investigations.

- Source city
- Patient age
- Patient sex
- Toxin
- Type details (often from notes compiled by epidemiologists)

<back

Appendix E: Confirmed and Unconfirmed PFGE Patterns and Matches

PulseNet database managers confirm PFGE patterns when the pattern is assigned a pattern name (pattern number). Pattern matches may be "confirmed" or "unconfirmed."

Confirmed pattern means that the CDC-PulseNet database manager has analyzed the isolate's PFGE pattern image, assigned it a specific PulseNet pattern name, and considered it a final result. PulseNet database managers rarely reverse confirmed pattern names and FSIS should rely on the names. A confirmed pattern can be referred to as a confirmed pattern name, confirmed isolate, or confirmed PFGE pattern.

Unconfirmed pattern means a pattern that has not been assigned a final name by a CDC-PulseNet database manager. This could indicate that either no pattern name has been assigned, or an automatically assigned unconfirmed pattern has been assigned by the PulseNet computer. PulseNet database managers frequently reverse the temporary unconfirmed pattern names and therefore FSIS should not rely on them.

Unconfirmed match means the CDC-PulseNet database manager has not confirmed a match among isolates. Because PFGE patterns often are not confirmed immediately following upload, OSEL performs its own comparisons. This enables FSIS to pursue epidemiologic leads earlier, alerts OSEL to isolates that should be followed closely, and encourages OSEL to double-check the PulseNet database manager decisions. OSEL's unconfirmed comparisons can be visual (from the subjective judgment of the OSEL PFGE analyst), based on the results of the OSEL BioNumerics program, or a combination of both. If OSEL believes that two unconfirmed isolates match, it uses the description "appears to match." "Appears to match" usually coincides with "confirmed match." Other commonly used terms are "very similar to" and "similar to," which are meant to imply that the isolates may be related, but are not exact matches, and are not expected to have the same confirmed pattern name. Unless noted otherwise, unconfirmed matches are two restriction enzymes; *Salmonella* matches are for one enzyme only (the primary enzyme).

PulseNet database managers use two kinds of computer generated pattern names when referring to unconfirmed matches. The first group is for patterns the computer is able to successfully group with existing pattern names. The second group is for patterns that the computer is unable to type, which likely may result in the creating a new pattern name.

Confirmed match means a match among isolates that have the same confirmed pattern. These isolates usually have not been compared directly to each other, but instead, independently have been given the same pattern name by the CDC-PulseNet database manager. Rarely, OPHS requests that PulseNet staff re-evaluate confirmed patterns. These instances have included high-profile FSIS investigations or situations where FSIS believes that the results of a "pattern confirmation" are questionable. Unless noted otherwise, *E. coli* O157:H7, STEC, and *Lm* matches are for two restriction enzymes and *Salmonella* matches are for one enzyme (the primary enzyme)

<back

Appendix F: Descriptive Terms for PFGE Pattern Frequency

OSEL calculates the frequency for all *E. coli* O157:H7 and *Lm* PFGE patterns. Pattern frequency is calculated as the percentage of isolates in the PulseNet database that have the same primary and secondary PFGE patterns. PulseNet uses some descriptive terms to describe the frequency of individual pattern names. However, these names are somewhat subjective and are applied at the database manager's discretion, which may vary by pathogen. Until PulseNet develops objective terms for describing PFGE pattern frequency, FSIS will use the following descriptive terms for communicating pattern frequency information within the Agency (Tables F1 and F2). These descriptive terms are based on the observed combined pattern frequency for FSIS isolates uploaded to PulseNet, and the size of a typical 60-day Hotlist search (number of isolates submitted to PulseNet in a 60-day period).

For *Lm*, descriptive terms are based on data from 261 pattern combinations calculated during routine PulseNet searches for frequency information for the years 2004 through March 2007. During that time, FSIS uploaded about 20% of the PulseNet *Lm* isolates, so the data is representative of the PulseNet database (Table F1). For *E. coli* O157:H7, descriptive terms were based on routine searches for all isolates uploaded to PulseNet during two 30-day periods (May 2007 and November 2006) (Table F2). Pattern frequency should be considered in context of other information, such as the recent history of the pattern, which can be observed in a Combined Pattern Search.

Table F1: Proposed FSIS Descriptive Terms for *Lm* PFGE Combined Pattern Frequency

Proposed Descriptive Name	Frequency[1]	Number of total hits in PN database[2]	Expected number of PN database hits per search period[3]	Number of pattern combinations found in OSEL data set N=261[4]	Percentage of isolates in OSEL data set N=584[5]
Common	> 0.62%	greater than 40	1 Hit per 60 or less days	12	22%
Less common	≤ 0.62% and > 0.31%	21 to 40	1 Hit per 61 to 120 days	7	4%
Rare	≤ 0.31% and > 0.16%	11 to 20	1 Hit per 121 to 240 days	25	16%
Very rare	≤ 0.16% and ≥2 entries in PN database	2 to 10	1 Hit per 241 or more days to two observations in PN database	116	41%
New to the PN database	1 entry in PN database	1	No prior observations of pattern in PN database	101	17%

[1] Frequency — frequency of pattern combination in PulseNet database
[2] Number of total hits in PulseNet database — total number of isolates with a specific pattern combination in PulseNet database
[3] Expected number of PulseNet database hits per search period — total number of isolates with a specific pattern combination in PulseNet database expected within the defined time period
[4] Number of unique patterns found in OSEL data set (*Lm*) — total number of unique pattern combinations in data set that falls in category specified
[5] Number of isolates found in OSEL data set (*Lm*) — total number of isolates in data set that falls in category specified (includes multiples of same pattern combination)

Table F2: Proposed FSIS Descriptive Terms for *E. coli* O157:H7 PFGE Combined Pattern Frequency

Proposed Descriptive Name	Frequency[6]	Number of total hits in PN database[7]	Expected number of PN database hits per search period[8]	Number of pattern combinations found in 60 day data set N=198[9]	Percentage of isolates in 60 day data set. N=436[10]
Common	> 0.31%	greater than 39	1 Hits per 60 or less days	16	36%
Less common	≤ 0.31% and > 0.16%	20 to 38	1 Hit per 61 to 120 days	13	15%
Rare	≤ 0. 16% and > 0.08%	10 to 19	1 Hit per 121 to 240 days	10	7%
Very rare	≤ 0.08% and ≥2 entries in PN database	2 to 9	1 Hit per 241 or more days to two observations in PN database	57	19%
New to the PN database	1 entry in PN database	1	No prior observations of pattern in PN database	102	23%

[6] Frequency — frequency of pattern combination in PulseNet database
[7] Number of total hits in PulseNet database — total number of isolates with a specific pattern combination in PulseNet database
[8] Expected number of PulseNet database hits per search period — total number of isolates with a specific pattern combination in PulseNet database expected within the defined time period
[9] Number of unique patterns found in 60-day data set (*E. coli* O157:H7) — total number of unique pattern combinations in data set that falls in category specified
[10] Number of isolates found in 60-day data set (*E. coli* O157:H7) — total number of isolates in data set that falls in category specified (includes multiples of same pattern combination)

<back

Appendix G: MLVA Theory and Laboratory Method

The current PulseNet *E. coli* O157:H7 MLVA method used by FSIS analyzes eight DNA loci (specific regions) within the genome. Very similar methods are used for other MLVA assays specific for *Lm*, and *Salmonella* serotypes Enteritidis and Typhimurium.

Figures G1 through G4 show how PCR is used to amplify a specific MLVA locus. Although these figures describe the process for a single locus, the amplification process is the same for all eight loci. Figure G4 shows the end product amplified "forward strand" DNA that is tagged so that it can be detected by a DNA sequencer. The sequencer determines the length of the amplified sections, which effectively determines the number of repeats in the specific locus. By generating repeat data for eight loci, a bacterial isolate is characterized such that it can be compared reliably to other isolates. Isolates that do not have the same repeat number for each of the eight loci are unlikely to have been recently generated from a common "ancestor."

Figure G1. MLVA targets eight DNA loci within the *E. coli* genome. Bacterial DNA is double stranded, which means it is made of two single strands of complementary DNA that are stuck together. One strand is called the "forward strand" and the other is called the "reverse strand." Other MLVA assays target DNA loci in the genomes of *Lm*, and *Salmonella* serotypes Enteritidis and Typhimurium.

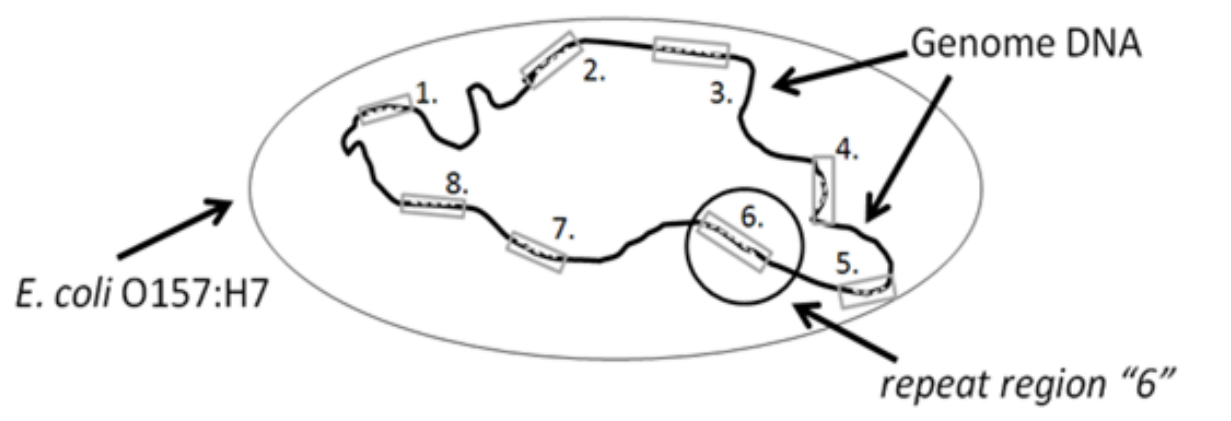


Figure G2. Forward and reverse strands of *E. coli* DNA. During the first step of MLVA, double stranded DNA at the individual loci are copied many times using PCR. In the diagram below, the forward strand of DNA is black with white letters and the reverse strand is white with black letters. The grey area represents the rest of the genome that is not copied by PCR. Forward and reverse strands of DNA are highlighted as "Repeat region 6" in Figure G1. "x" marks additional nucleotides in the forward strand. "y" marks additional nucleotides in the reverse strand.

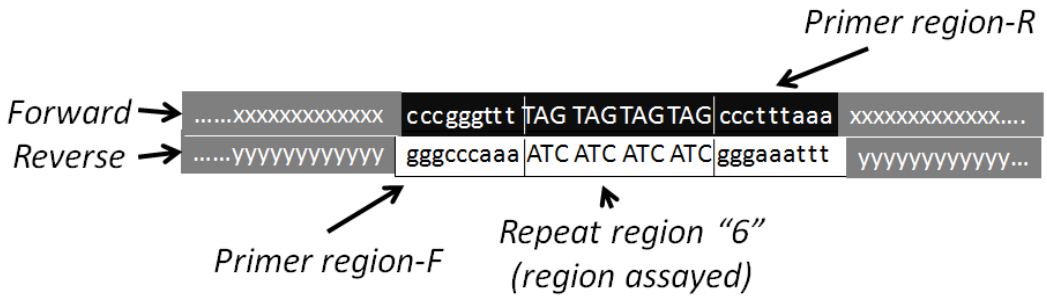


Figure G3. Primers stick to the complementary DNA strand. Heat is used to separate the two DNA strands, and two primers are added to the mixture. The mixture is cooled, which allows the forward primer, which is labeled, to bind to the reverse strand, and the reverse primer, which is unlabeled, to bind to the forward strand. A DNA-producing enzyme, called DNA polymerase, copies the rest of the region of interest using the other strand as a guide. The temperature of the mixture is raised and lowered repeatedly using a thermocycler to produce thousands of copies of the labeled forward strands and unlabeled reverse strands.

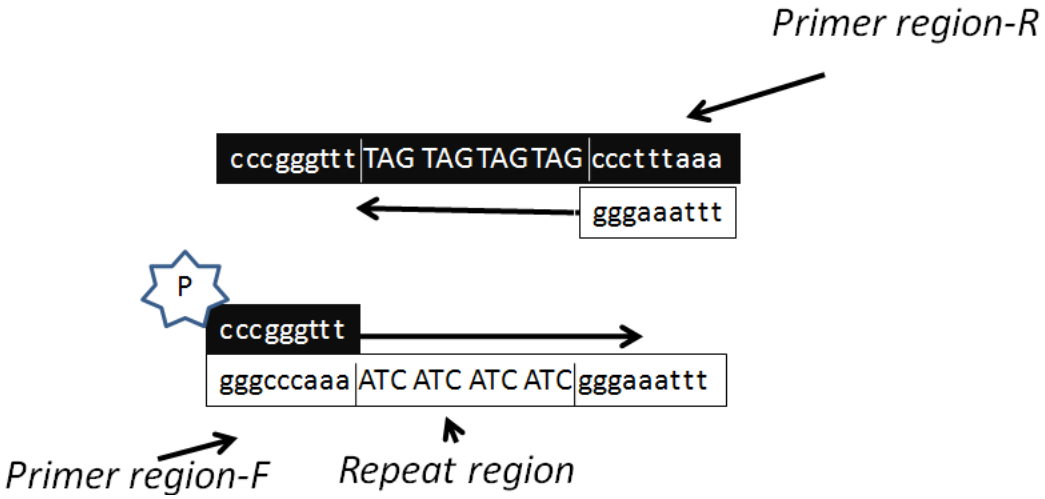
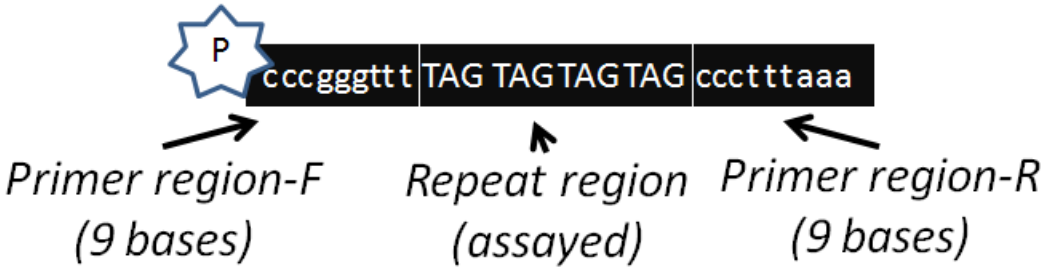
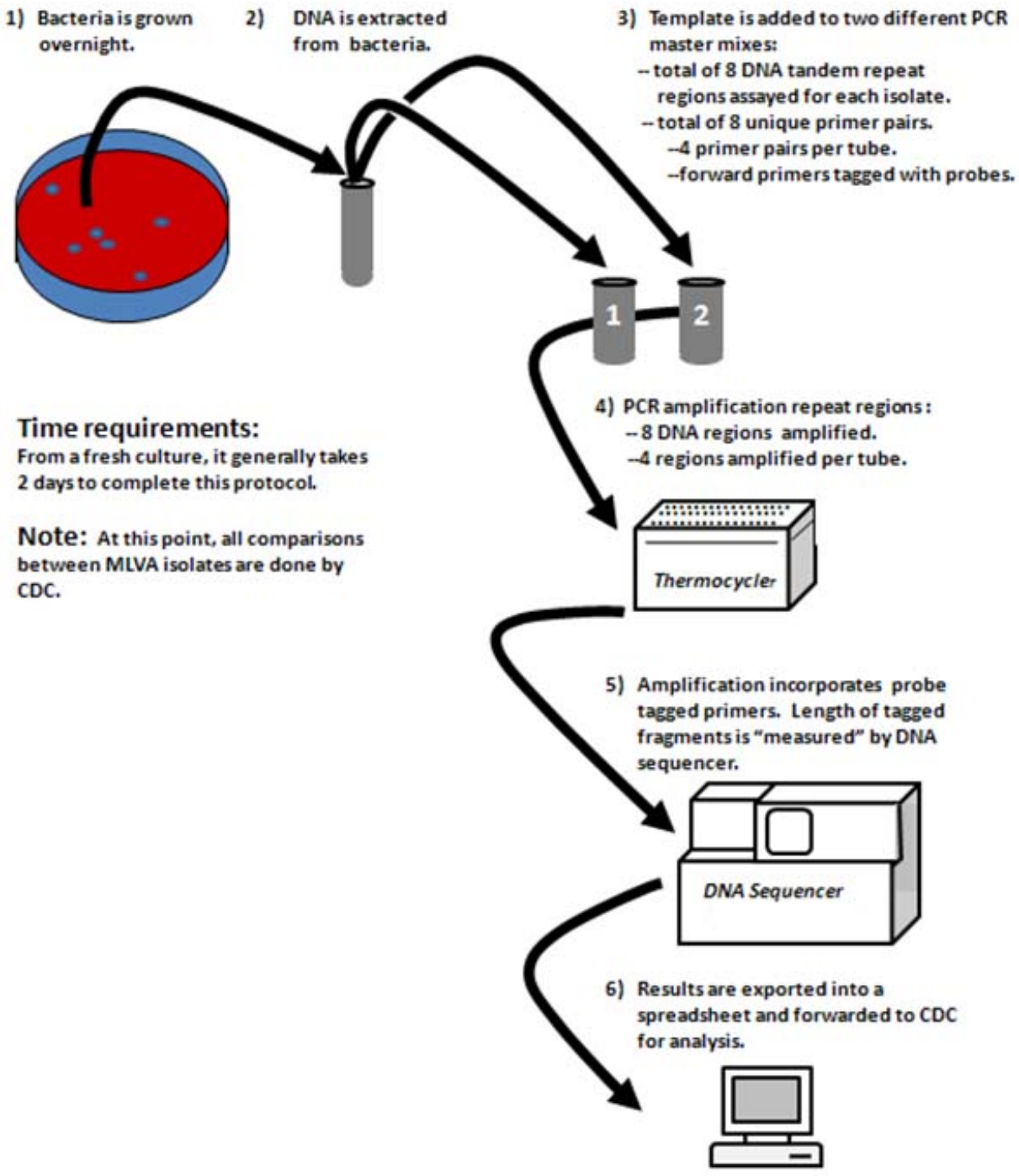


Figure G4. The number of DNA bases in the forward strand is determined. The mixture of amplified DNA is put into a sequencer, which is used to determine the number of bases in the labeled forward strand. The number of bases in the repeat region is calculated by subtracting the number of bases in the forward and reverse primers.



<back

Appendix H: MLVA Protocol: From FSIS Isolate to CDC Analysis



<back

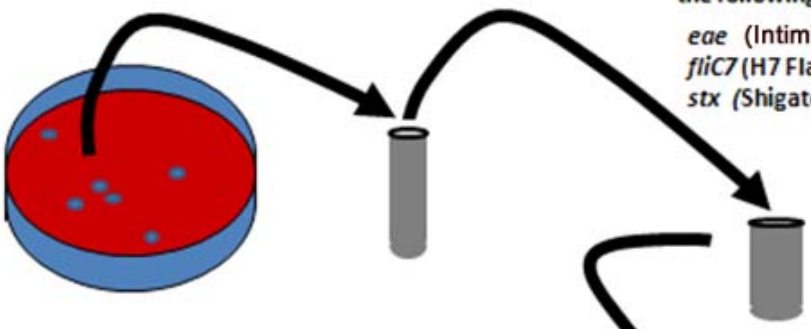
Appendix I: Molecular Subtyping of *E. coli* O157:H7

1) Bacteria is grown overnight.

2) DNA is extracted from bacteria.

3) Template is added to a master mix containing primers and probes for the following:

eae (Intimin gene)
fliC7 (H7 Flagella antigen gene)
stx (Shigatoxin genes *stx1 and/or stx2)*



4) PCR is run on quanitative PCR capabable thermocycler.

This test usually run... *If there is a positive test for O157 in an E.coli isolate, but routine tests for the presence of toxin are negative.*

Interpretation: *a prior positive test for O157 followed by a positive test for either fliC7* (H7 Flagella antigen gene) or either of the *stx genes* (Shigatoxin genes *stx1 and/or stx2)* is a confirmed positive.

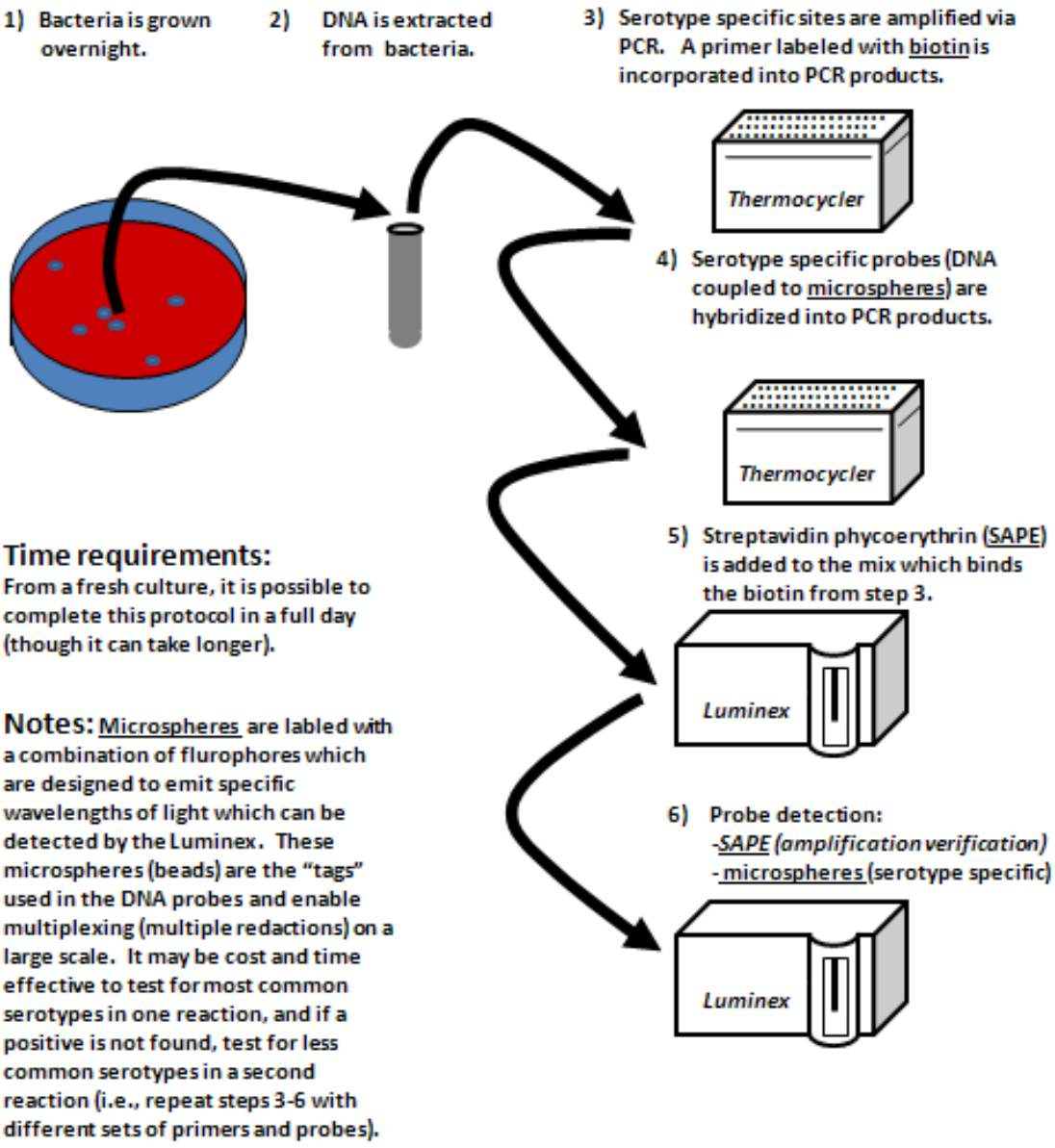
Time requirements:
From a fresh culture, it generally takes 1 day to complete this protocol.

software generates report that confirms the presence or absence of the following:

eae (Intimin gene)
fliC7 (H7 Flagella antigen gene)
stx (Shigatoxin genes *stx1 and/or stx2)*

<back

Appendix J: Molecular Serotyping of *Salmonella*



Time requirements:
From a fresh culture, it is possible to complete this protocol in a full day (though it can take longer).

Notes: Microspheres are labled with a combination of flurophores which are designed to emit specific wavelengths of light which can be detected by the Luminex. These microspheres (beads) are the "tags" used in the DNA probes and enable multiplexing (multiple redactions) on a large scale. It may be cost and time effective to test for most common serotypes in one reaction, and if a positive is not found, test for less common serotypes in a second reaction (i.e., repeat steps 3-6 with different sets of primers and probes).

<back

Appendix K: *Salmonella* Serotypes Reportable by Molecular Detection (Laboratory Guidebook, effective date 1/02/13)

Agona
Albany
Altona
Anatum
Berta
Blockley
Bredeney
Dublin
Enteritidis
Give
Havana
Heidelberg
Infantis
Kentucky
Litchfield
Montevideo
Muenchen
Newport
Ouakam
Panama
Saintpaul
Schwarzengrund
Thompson
Typhimurium*
4,[5],12:i:-**
*"Typhimurium" includes variant 5 ** Molecular result of serotype B:i:-is reported as 4,[5],12:i:-to maintain consistency with existing nomenclature as per CDC recommendation.

<back

9 REFERENCES

Ahmed,M.U., Dunn,L. and Ivanova,E.P. (2012) Evaluation of current molecular approaches for genotyping of Campylobacter jejuni strains. *Foodborne Pathog. Dis.* **9**, 375-385.

Alcaine,S.D., Soyer,Y., Warnick,L.D., Su,W.L., Sukhnanand,S., Richards,J., Fortes,E.D., McDonough,P., Root,T.P., Dumas,N.B., Grohn,Y. and Wiedmann,M. (2006) Multilocus sequence typing supports the hypothesis that cow- and human-associated *Salmonella* isolates represent distinct and overlapping populations. *Applied and environmental microbiology* **72**, 7575-7585.

Allard,M.W., Luo,Y., Strain,E., Pettengill,J., Timme,R., Wang,C., Li,C., Keys,C.E., Zheng,J., Stones,R., Wilson,M.R., Musser,S.M. and Brown,E.W. (2013) On the Evolutionary History, Population Genetics and Diversity among Isolates of *Salmonella* Enteritidis PFGE Pattern JEGX01.0004. *PLoS. One* **8**, e55254.

Altekruse,S.F., Stern,N.J., Fields,P.I. and Swerdlow,D.L. (1999) Campylobacter jejuni--an emerging foodborne pathogen. *Emerg. Infect. Dis.* **5**, 28-35.

Anderson,E.S., Ward,L.R., Saxe,M.J. and de Sa,J.D. (1977) Bacteriophage-typing designations of *Salmonella* typhimurium. *The Journal of hygiene* **78**, 297-300.

Bakker,H.C., Switt,A.I., Cummings,C.A., Hoelzer,K., Degoricija,L., Rodriguez-Rivera,L.D., Wright,E.M., Fang,R., Davis,M., Root,T., Schoonmaker-Bopp,D., Musser,K.A., Villamil,E., Waechter,H., Kornstein,L., Furtado,M.R. and Wiedmann,M. (2011) A whole-genome single nucleotide polymorphism-based approach to trace and identify outbreaks linked to a common *Salmonella* enterica subsp. enterica serovar Montevideo pulsed-field gel electrophoresis type. *Applied and environmental microbiology* **77**, 8648-8655.

Barrett,T.J., Gerner-Smidt,P. and Swaminathan,B. (2006) Interpretation of pulsed-field gel electrophoresis patterns in foodborne disease investigations and surveillance. *Foodborne Pathog. Dis.* **3**, 20-31.

Bettelheim,K.A. (2007) The non-O157 shiga-toxigenic (verocytotoxigenic) Escherichia coli; under-rated pathogens. *Critical reviews in microbiology* **33**, 67-87.

Beutin,L., Jahn,S. and Fach,P. (2009) Evaluation of the 'GeneDisc' real-time PCR system for detection of enterohaemorrhagic Escherichia coli (EHEC) O26, O103, O111, O145 and O157 strains according to their virulence markers and their O- and H-antigen-associated genes. *Journal of applied microbiology* **106**, 1122-1132.

Bielaszewska,M., Mellmann,A., Zhang,W., Kock,R., Fruth,A., Bauwens,A., Peters,G. and Karch,H. (2011) Characterisation of the Escherichia coli strain associated with an outbreak of haemolytic uraemic syndrome in Germany, 2011: a microbiological study. *Lancet Infect. Dis.* **11**, 671-676.

Blaser,M.J. (1997) Epidemiologic and clinical features of Campylobacter jejuni infections. *J Infect. Dis.* **176 Suppl 2**, S103-S105.

Bohez,L., Gantois,I., Ducatelle,R., Pasmans,F., Dewulf,J., Haesebrouck,F. and Van,I.F. (2008) The *Salmonella* Pathogenicity Island 2 regulator ssrA promotes reproductive tract but not intestinal colonization in chickens. *Veterinary microbiology* **126**, 216-224.

Botteldoorn,N., Van,C.E., Goris,J., Werbrouck,H., Piessens,V., Godard,C., Scheldeman,P., Herman,L. and Heyndrickx,M. (2010) Limited genetic diversity and gene expression differences between egg- and non-egg-related *Salmonella* Enteritidis strains. *Zoonoses. Public. Health* **57**, 345-357.

Boxrud,D., Pederson-Gulrud,K., Wotton,J., Medus,C., Lyszkowicz,E., Besser,J. and Bartkus,J.M. (2007) Comparison of multiple-locus variable-number tandem repeat analysis, pulsed-field gel electrophoresis, and phage typing for subtype analysis of *Salmonella* enterica serotype Enteritidis. *Journal of clinical microbiology* **45**, 536-543.

Bradley,K.K., Williams,J.M., Burnsed,L.J., Lytle,M.B., McDermott,M.D., Mody,R.K., Bhattarai,A., Mallonee,S., Piercefield,E.W., McDonald-Hamm,C.K. and Smithee,L.K. (2012) Epidemiology of a large restaurant-associated outbreak of Shiga toxin-producing Escherichia coli O111:NM. *Epidemiology and infection* **140**, 1644-1654.

Brenner,F.W., Villar,R.G., Angulo,F.J., Tauxe,R. and Swaminathan,B. (2000) *Salmonella* nomenclature. *J Clin Microbiol* **38**, 2465-2467.

Brooks,J.T., Sowers,E.G., Wells,J.G., Greene,K.D., Griffin,P.M., Hoekstra,R.M. and Strockbine,N.A. (2005) Non-O157 Shiga toxin-producing Escherichia coli infections in the United States, 1983-2002. *J. Infect. Dis.* **192**, 1422-1429.

Broschat,S.L., Call,D.R., Davis,M.A., Meng,D., Lockwood,S., Ahmed,R. and Besser,T.E. (2010) Improved identification of epidemiologically related strains of *Salmonella* enterica by use of a fusion algorithm based on pulsed-field gel electrophoresis and multiple-locus variable-number tandem-repeat analysis. *Journal of clinical microbiology* **48**, 4072-4082.

Brzuszkiewicz,E., Thurmer,A., Schuldes,J., Leimbach,A., Liesegang,H., Meyer,F.D., Boelter,J., Petersen,H., Gottschalk,G. and Daniel,R. (2011) Genome sequence analyses of two isolates from the recent Escherichia coli outbreak in Germany reveal the emergence of a new pathotype: Entero-Aggregative-Haemorrhagic Escherichia coli (EAHEC). *Archives of microbiology* **193**, 883-891.

Bugarel,M., Beutin,L. and Fach,P. (2010a) Low-density macroarray targeting non-locus of enterocyte effacement effectors (nle genes) and major virulence factors of Shiga toxin-producing Escherichia coli (STEC): a new approach for molecular risk assessment of STEC isolates. *Appl. Environ. Microbiol* **76**, 203-211.

Bugarel,M., Beutin,L. and Fach,P. (2010b) Low-density macroarray targeting non-locus of enterocyte effacement effectors (nle genes) and major virulence factors of Shiga toxin-producing Escherichia coli (STEC): a new approach for molecular risk assessment of STEC isolates. *Applied and environmental microbiology* **76**, 203-211.

Bugarel,M., Martin,A., Fach,P. and Beutin,L. (2011) Virulence gene profiling of enterohemorrhagic (EHEC) and enteropathogenic (EPEC) Escherichia coli strains: a basis for molecular risk assessment of typical and atypical EPEC strains. *BMC microbiology* **11**, 142.

Call,D.R., Orfe,L., Davis,M.A., Lafrentz,S. and Kang,M.S. (2008) Impact of compounding error on strategies for subtyping pathogenic bacteria. *Foodborne Pathog. Dis.* **5**, 505-516.

CDC (2011a) Multistate outbreak of listeriosis associated with Jensen Farms cantaloupe--United States, August-September 2011. *MMWR. Morb. Mortal. Wkly. Rep* **60**, 1357-1358.

CDC (2011b) National *Salmonella* Surveillance Data.

CDC (2011c) Outbreak of invasive listeriosis associated with the consumption of hog head cheese--Louisiana, 2010. *MMWR. Morb. Mortal. Wkly. Rep* **60**, 401-405.

CDC (2012a) National *Salmonella* Surveillance Annual Summary 2009 Table Laboratory-confirmed Laboratory-*Salmonella* isolates from human sources reported to CDC,with the 20 most frequently reported serotypes listed individually, 2009.

CDC (2012b) National *Salmonella* Surveillance Overview**.** *Salmonella* Annual Summary, 2009 [PDF - 478 KB].

CDC (2013) National Case-Based Surveillance and *Listeria*.

Chen,Y., Ross,W.H., Gray,M.J., Wiedmann,M., Whiting,R.C. and Scott,V.N. (2006) Attributing risk to *Listeria monocytogenes* subgroups: dose response in relation to genetic lineages. *J Food Prot* **69**, 335-344.

Chen,Y., Ross,W.H., Whiting,R.C., Van,S.A., Nightingale,K.K., Wiedmann,M. and Scott,V.N. (2011) Variation in *Listeria monocytogenes* dose responses in relation to subtypes encoding a full-length or truncated internalin A. *Appl. Environ. Microbiol* **77**, 1171-1180.

Chen,Y., Zhang,W. and Knabel,S.J. (2005) Multi-virulence-locus sequence typing clarifies epidemiology of recent listeriosis outbreaks in the United States. *Journal of clinical microbiology* **43**, 5291-5294.

Chen,Y., Zhang,W. and Knabel,S.J. (2007) Multi-virulence-locus sequence typing identifies single nucleotide polymorphisms which differentiate epidemic clones and outbreak strains of *Listeria monocytogenes*. *Journal of clinical microbiology* **45**, 835-846.

Cho,S., Boxrud,D.J., Bartkus,J.M., Whittam,T.S. and Saeed,M. (2007) Multiple-locus variable-number tandem repeat analysis of *Salmonella* Enteritidis isolates from human and non-human sources using a single multiplex PCR. *FEMS Microbiol. Lett.* **275**, 16-23.

Cho,S., Whittam,T.S., Boxrud,D.J., Bartkus,J.M., Rankin,S.C., Wilkins,M.J., Somsel,P., Downes,F.P., Musser,K.A., Root,T.P., Warnick,L.D., Wiedmann,M. and Saeed,A.M. (2010) Use of multiple-locus variable number tandem repeat analysis and phage typing for subtyping of *Salmonella* Enteritidis from sporadic human cases in the United States. *Journal of applied microbiology* **108**, 859-867.

Coleman,M.E., Marks,H.M., Golden,N.J. and Latimer,H.K. (2004) Discerning strain effects in microbial dose-response data. *Journal of toxicology and environmental health. Part A* **67**, 667-685.

Cooke,F.J., Brown,D.J., Fookes,M., Pickard,D., Ivens,A., Wain,J., Roberts,M., Kingsley,R.A., Thomson,N.R. and Dougan,G. (2008) Characterization of the genomes of a diverse collection of *Salmonella* enterica serovar Typhimurium definitive phage type 104. *J Bacteriol* **190**, 8155-8162.

Coombes,B.K., Wickham,M.E., Mascarenhas,M., Gruenheid,S., Finlay,B.B. and Karmali,M.A. (2008a) Molecular analysis as an aid to assess the public health risk of non-O157 Shiga toxin-producing Escherichia coli strains. *Appl. Environ. Microbiol.* **74**, 2153-2160.

Coombes,B.K., Wickham,M.E., Mascarenhas,M., Gruenheid,S., Finlay,B.B. and Karmali,M.A. (2008b) Molecular analysis as an aid to assess the public health risk of non-O157 Shiga toxin-producing Escherichia coli strains. *Applied and environmental microbiology* **74**, 2153-2160.

Cronquist,A.B., Mody,R.K., Atkinson,R., Besser,J., Tobin,D.M., Hurd,S., Robinson,T., Nicholson,C. and Mahon,B.E. (2012) Impacts of culture-independent diagnostic practices on public health surveillance for bacterial enteric pathogens. *Clin. Infect Dis.* **54 Suppl 5**, S432-S439.

CSPI (2011) Petition for an Interpretive Rule Declaring Certain Strains of Antibiotic Resistant *Salmonella* in Ground Meat and Poultry to be Adulterants within the Meaning of 21 USC § 601 (m)(1) and (2)(A), and 21 USC § 453 (g)(1) and (g)(2).

Dasti,J.I., Tareen,A.M., Lugert,R., Zautner,A.E. and Gross,U. (2010) Campylobacter jejuni: a brief overview on pathogenicity-associated factors and disease-mediating mechanisms. *Int. J. Med. Microbiol.* **300**, 205-211.

Delannoy,S., Beutin,L. and Fach,P. (2012) Use of clustered regularly interspaced short palindromic repeat sequence polymorphisms for specific detection of enterohemorrhagic Escherichia coli strains of serotypes O26:H11, O45:H2, O103:H2, O111:H8, O121:H19, O145:H28, and O157:H7 by real-time PCR. *J. Clin. Microbiol.* **50**, 4035-4040.

Dhanoa,A. and Fatt,Q.K. (2009) Non-typhoidal *Salmonella* bacteraemia: epidemiology, clinical characteristics and its' association with severe immunosuppression. *Annals of clinical microbiology and antimicrobials* **8**, 15.

Dingle,K.E., Colles,F.M., Wareing,D.R., Ure,R., Fox,A.J., Bolton,F.E., Bootsma,H.J., Willems,R.J., Urwin,R. and Maiden,M.C. (2001) Multilocus sequence typing system for Campylobacter jejuni. *Journal of clinical microbiology* **39**, 14-23.

Eberle,K.N. and Kiess,A.S. (2012) Phenotypic and genotypic methods for typing Campylobacter jejuni and Campylobacter coli in poultry. *Poultry science* **91**, 255-264.

Ellerbroek,L.I., Lienau,J.A. and Klein,G. (2010) Campylobacter spp. in broiler flocks at farm level and the potential for cross-contamination during slaughter. *Zoonoses. Public. Health* **57**, e81-e88.

FAO/WHO (2004) Risk assessment of *Listeria monocytogenes* in ready-to-eat foods.

FDA (2001) Bacteriological Analytical Manual, Chapter 7, *Campylobacter*.

FDA/USDA. (2003) Quantitative Assessment of Relative Risk to Public Health from Foodborne *Listeria monocytogenes* Among Selected Categories of Ready-to-Eat Foods.

Fierer,J., Krause,M., Tauxe,R. and Guiney,D. (1992) *Salmonella* typhimurium bacteremia: association with the virulence plasmid. *J Infect. Dis.* **166**, 639-642.

Fitzgerald,C., Collins,M., Van,D.S., Mikoleit,M., Brown,T. and Fields,P. (2007) Multiplex, bead-based suspension array for molecular determination of common *Salmonella* serogroups. *J Clin Microbiol.* **45**, 3323-3334.

Foley,S.L. and Lynne,A.M. (2008) Food animal-associated *Salmonella* challenges: pathogenicity and antimicrobial resistance. *J Anim Sci* **86**, E173-E187.

Foley,S.L., Lynne,A.M. and Nayak,R. (2009a) Molecular typing methodologies for microbial source tracking and epidemiological investigations of Gram-negative bacterial foodborne pathogens. *Infection, genetics and evolution: journal of molecular epidemiology and evolutionary genetics in infectious diseases* **9**, 430-440.

Foley,S.L., Lynne,A.M. and Nayak,R. (2009b) Molecular typing methodologies for microbial source tracking and epidemiological investigations of Gram-negative bacterial foodborne pathogens. *Infect. Genet Evol.* **9**, 430-440.

Foley,S.L., White,D.G., McDermott,P.F., Walker,R.D., Rhodes,B., Fedorka-Cray,P.J., Simjee,S. and Zhao,S. (2006) Comparison of subtyping methods for differentiating *Salmonella* enterica serovar Typhimurium isolates obtained from food animal sources. *Journal of clinical microbiology* **44**, 3569-3577.

Frank,C., Werber,D., Cramer,J.P., Askar,M., Faber,M., an der,H.M., Bernard,H., Fruth,A., Prager,R., Spode,A., Wadl,M., Zoufaly,A., Jordan,S., Kemper,M.J., Follin,P., Muller,L., King,L.A., Rosner,B., Buchholz,U., Stark,K. and Krause,G. (2011) Epidemic profile of Shiga-toxin-producing Escherichia coli O104:H4 outbreak in Germany. *N. Engl. J. Med.* **365**, 1771-1780.

Frank,D.N. and Pace,N.R. (2008) Gastrointestinal microbiology enters the metagenomics era. *Curr. Opin. Gastroenterol.* **24**, 4-10.

Fratamico,P.M. and Bagi,L.K. (2012) Detection of Shiga toxin-producing Escherichia coli in ground beef using the GeneDisc real-time PCR system. *Frontiers in cellular and infection microbiology* **2**, 152.

Fratamico,P.M., Bagi,L.K., Cray,W.C., Jr., Narang,N., Yan,X., Medina,M. and Liu,Y. (2011) Detection by multiplex real-time polymerase chain reaction assays and isolation of Shiga toxin-producing Escherichia coli serogroups O26, O45, O103, O111, O121, and O145 in ground beef. *Foodborne Pathog. Dis.* **8**, 601-607.

Freitag,N.E., Port,G.C. and Miner,M.D. (2009) *Listeria monocytogenes* - from saprophyte to intracellular pathogen. *Nature reviews. Microbiology* **7**, 623-628.

FSIS (2007) FSIS Risk Assessment for Risk-Based Verification Sampling of *Listeria monocytogenes*.

FSIS (2008a) 8,000 Series: Compliance Evaluation & Enforcement. Directive 8080.3 Foodborne Illness Investigations.

FSIS (2008b) *Salmonella* Verification Sampling Program: Response to Comments and New Agency Policies. pp. 4767-4774.

FSIS (2009) Directive 5100.4. Prioritized Scheduling of Food Safety Assessments.

FSIS (2010) New Performance Standards for *Salmonella* and Campylobacter in Young Chicken and Turkey Slaughter Establishments; New Compliance Guides.

FSIS (2011a) FSIS Scheduling Criteria and Algorithm for the *Salmonella* and *Campylobacter* Verification Sampling Program for Raw Meat and Poultry.

FSIS (2011b) Isolation, Identification, and Enumeration of Campylobacter jejuni/coli/lari from Poultry Rinse and Sponge Samples.

FSIS (2011c) Report on the Food Safety and Inspection Service's Microbiological and Residue Sampling Programs.

FSIS (2011d) Shiga Toxin-Producing Escherichia coli in Certain Raw Beef Products.

FSIS (2012a) Detection and Isolation of non-O157 Shiga Toxin-Producing *Escherichia coli* (STEC) from Meat Products.

FSIS (2012b) MLG 5.06 Detection, Isolation and Identification of *Escherichia coli* O157:H7 from Meat Products.

FSIS (2012c) Risk Profile for Pathogenic **Non-O157 Shiga Toxin-Producing *Escherichia coli*** (non-O157 STEC).

FSIS (2013a) Appendix 1.01, *Salmonella* Serotypes Reportable by Molecular Detection.

FSIS (2013b) FSIS Establishment Eligibility Criteria and Scheduling Algorithm for the *Salmonella* Verification Sampling Program for Raw Meat and Poultry.

Galli,L., Miliwebsky,E., Irino,K., Leotta,G. and Rivas,M. (2010) Virulence profile comparison between LEE-negative Shiga toxin-producing Escherichia coli (STEC) strains isolated from cattle and humans. *Vet. Microbiol.* **143**, 307-313.

Gaul,L.K., Farag,N.H., Shim,T., Kingsley,M.A., Silk,B.J. and Hyytia-Trees,E. (2012) Hospital-Acquired Listeriosis Outbreak Caused by Contaminated Diced Celery--Texas, 2010. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America*.

Gerner-Smidt,P., Hise,K., Kincaid,J., Hunter,S., Rolando,S., Hyytia-Trees,E., Ribot,E.M. and Swaminathan,B. (2006) PulseNet USA: a five-year update. *Foodborne Pathog. Dis.* **3**, 9-19.

Gilmour,M.W., Graham,M., Van,D.G., Tyler,S., Kent,H., Trout-Yakel,K.M., Larios,O., Allen,V., Lee,B. and Nadon,C. (2010) High-throughput genome sequencing of two *Listeria monocytogenes* clinical isolates during a large foodborne outbreak. *BMC Genomics* **11**, 120.

Godschalk,P.C., van,B.A., van den Braak,N., van,N.D., Ang,C.W., Jacobs,B.C., Gilbert,M. and Endtz,H.P. (2007) PCR-restriction fragment length polymorphism analysis of Campylobacter jejuni genes involved in lipooligosaccharide biosynthesis identifies putative molecular markers for Guillain-Barre syndrome. *Journal of clinical microbiology* **45**, 2316-2320.

Griekspoor,P., Engvall,E.O., Olsen,B. and Waldenstrom,J. (2010) Multilocus sequence typing of Campylobacter jejuni from broilers. *Veterinary microbiology* **140**, 180-185.

Guo,C., Hoekstra,R.M., Schroeder,C.M., Pires,S.M., Ong,K.L., Hartnett,E., Naugle,A., Harman,J., Bennett,P., Cieslak,P., Scallan,E., Rose,B., Holt,K.G., Kissler,B., Mbandi,E., Roodsari,R., Angulo,F.J. and Cole,D. (2011) Application of Bayesian techniques to model the burden of human salmonellosis attributable to U.S. food commodities at the point of processing: adaptation of a Danish model. *Foodborne Pathog. Dis.* **8**, 509-516.

Hald,T., Vose,D., Wegener,H.C. and Koupeev,T. (2004) A Bayesian approach to quantify the contribution of animal-food sources to human salmonellosis. *Risk. Anal.* **24**, 255-269.

Hoelzer,K., Chen,Y., Dennis,S., Evans,P., Pouillot,R., Silk,B.J. and Walls,I. (2013) New Data, Strategies, and Insights for *Listeria monocytogenes* Dose-Response Models: Summary of an Interagency Workshop, 2011. *Risk analysis : an official publication of the Society for Risk Analysis*.

Hohmann,E.L. (2001) Nontyphoidal salmonellosis. *Clin Infect. Dis.* **32**, 263-269.

Hussein,H.S. (2007) Prevalence and pathogenicity of Shiga toxin-producing Escherichia coli in beef cattle and their products. *Journal of animal science* **85**, E63-E72.

Hyytia-Trees,E., Lafon,P., Vauterin,P. and Ribot,E.M. (2010) Multilaboratory validation study of standardized multiple-locus variable-number tandem repeat analysis protocol for shiga toxin-producing Escherichia coli O157: a novel approach to normalize fragment size data between capillary electrophoresis platforms. *Foodborne. Pathog. Dis.* **7**, 129-136.

Hyytia-Trees,E., Smole,S.C., Fields,P.A., Swaminathan,B. and Ribot,E.M. (2006) Second generation subtyping: a proposed PulseNet protocol for multiple-locus variable-number tandem repeat analysis of Shiga toxin-producing Escherichia coli O157 (STEC O157). *Foodborne Pathog. Dis.* **3**, 118-131.

Hyytia-Trees,E.K., Cooper,K., Ribot,E.M. and Gerner-Smidt,P. (2007) Recent developments and future prospects in subtyping of foodborne bacterial pathogens. *Future microbiology* **2**, 175-185.

IFSAC (2012) Interagency Food Safety Analytics Collaboration (IFSAC) Strategic Plan for Foodborne Illness Source Attribution.

Imamovic,L., Tozzoli,R., Michelacci,V., Minelli,F., Marziano,M.L., Caprioli,A. and Morabito,S. (2010) OI-57, a genomic island of Escherichia coli O157, is present in other seropathotypes of Shiga toxin-producing *E. coli* associated with severe human disease. *Infect. Immun.* **78**, 4697-4704.

Jackson,C.R., Fedorka-Cray,P.J., Wineland,N., Tankson,J.D., Barrett,J.B., Douris,A., Gresham,C.P., Jackson-Hall,C., McGlinchey,B.M. and Price,M.V. (2007) Introduction to United States Department of Agriculture VetNet: status of *Salmonella* and Campylobacter databases from 2004 through 2005. *Foodborne Pathog. Dis.* **4**, 241-248.

Jacquet,C., Doumith,M., Gordon,J.I., Martin,P.M., Cossart,P. and Lecuit,M. (2004) A molecular marker for evaluating the pathogenic potential of foodborne *Listeria monocytogenes*. *J Infect. Dis.* **189**, 2094-2100.

Jones,T.F., Ingram,L.A., Cieslak,P.R., Vugia,D.J., Tobin-D'Angelo,M., Hurd,S., Medus,C., Cronquist,A. and Angulo,F.J. (2008) Salmonellosis outcomes differ substantially by serotype. *J Infect. Dis.* **198**, 109-114.

Juneja,V.K., Marks,H.M. and Huang,L. (2003) Growth and heat resistance kinetic variation among various isolates of *Salmonella* and its application to risk assessment. *Risk analysis : an official publication of the Society for Risk Analysis* **23**, 199-213.

Juntunen,P., Heiska,H., Olkkola,S., Myllyniemi,A.L. and Hanninen,M.L. (2010) Antimicrobial resistance in Campylobacter coli selected by tylosin treatment at a pig farm. *Veterinary microbiology* **146**, 90-97.

Karmali,M.A., Mascarenhas,M., Shen,S., Ziebell,K., Johnson,S., Reid-Smith,R., Isaac-Renton,J., Clark,C., Rahn,K. and Kaper,J.B. (2003) Association of genomic O island 122 of Escherichia coli EDL 933 with verocytotoxin-producing Escherichia coli seropathotypes that are linked to epidemic and/or serious disease. *Journal of clinical microbiology* **41**, 4930-4940.

Kathariou,S., Graves,L., Buchrieser,C., Glaser,P., Siletzky,R.M. and Swaminathan,B. (2006) Involvement of closely related strains of a new clonal group of *Listeria monocytogenes* in the 1998-99 and 2002 multistate outbreaks of foodborne listeriosis in the United States. *Foodborne Pathog. Dis* **3**, 292-302.

Konczy,P., Ziebell,K., Mascarenhas,M., Choi,A., Michaud,C., Kropinski,A.M., Whittam,T.S., Wickham,M., Finlay,B. and Karmali,M.A. (2008) Genomic O island 122, locus for enterocyte effacement, and the evolution of virulent verocytotoxin-producing Escherichia coli. *Journal of bacteriology* **190**, 5832-5840.

Kotewicz,M.L., Mammel,M.K., Leclerc,J.E. and Cebula,T.A. (2008) Optical mapping and 454 sequencing of Escherichia coli O157 : H7 isolates linked to the US 2006 spinach-associated outbreak. *Microbiology* **154**, 3518-3528.

Kudva,I.T., Evans,P.S., Perna,N.T., Barrett,T.J., Ausubel,F.M., Blattner,F.R. and Calderwood,S.B. (2002) Strains of Escherichia coli O157:H7 differ primarily by insertions or deletions, not single-nucleotide polymorphisms. *J Bacteriol.* **184**, 1873-1879.

Lan,R., Reeves,P.R. and Octavia,S. (2009a) Population structure, origins and evolution of major *Salmonella* enterica clones. *Infect. Genet Evol.* **9**, 996-1005.

Lan,R., Reeves,P.R. and Octavia,S. (2009b) Population structure, origins and evolution of major *Salmonella* enterica clones. *Infect. Genet Evol.* **9**, 996-1005.

Lanier,W.A., Hall,J.M., Herlihy,R.K., Rolfs,R.T., Wagner,J.M., Smith,L.H. and Hyytia-Trees,E.K. (2011) Outbreak of Shiga-toxigenic Escherichia coli O157:H7 infections associated with rodeo attendance, Utah and Idaho, 2009. *Foodborne Pathog. Dis.* **8**, 1131-1133.

Lienau,E.K., Strain,E., Wang,C., Zheng,J., Ottesen,A.R., Keys,C.E., Hammack,T.S., Musser,S.M., Brown,E.W., Allard,M.W., Cao,G., Meng,J. and Stones,R. (2011) Identification of a salmonellosis outbreak by means of molecular sequencing. *N Engl. J Med* **364**, 981-982.

Litrup,E., Torpdahl,M., Malorny,B., Huehn,S., Christensen,H. and Nielsen,E.M. (2010) Association between phylogeny, virulence potential and serovars of *Salmonella* enterica. *Infection, genetics and evolution: journal of molecular epidemiology and evolutionary genetics in infectious diseases* **10**, 1132-1139.

Lomonaco,S., Verghese,B., Gerner-Smidt,P., Tarr,C., Gladney,L., Joseph,L., Katz,L., Turnsek,M., Frace,M., Chen,Y., Brown,E., Meinersmann,R., Berrang,M. and Knabel,S. (2013) Novel Epidemic Clones of *Listeria monocytogenes*, United States, 2011. *Emerg. Infect. Dis.* **19**, 147-150.

Lu,X., Huang,Q., Miller,W.G., Aston,D.E., Xu,J., Xue,F., Zhang,H., Rasco,B.A., Wang,S. and Konkel,M.E. (2012) Comprehensive detection and discrimination of Campylobacter species by use of confocal micro-Raman spectroscopy and multilocus sequence typing. *J Clin Microbiol* **50**, 2932-2946.

Lyhs,U., Katzav,M., Isohanni,P., Heiska,H. and Maijala,R. (2010) The temporal, PFGE and resistance pattern associations suggest that poultry products are only a minor source of human infections in western Finland. *Food Microbiol* **27**, 311-315.

McQuiston,J.R., Waters,R.J., Dinsmore,B.A., Mikoleit,M.L. and Fields,P.I. (2011) Molecular determination of H antigens of *Salmonella* by use of a microsphere-based liquid array. *J Clin Microbiol.* **49**, 565-573.

Mellmann,A., Harmsen,D., Cummings,C.A., Zentz,E.B., Leopold,S.R., Rico,A., Prior,K., Szczepanowski,R., Ji,Y., Zhang,W., McLaughlin,S.F., Henkhaus,J.K., Leopold,B., Bielaszewska,M., Prager,R., Brzoska,P.M., Moore,R.L., Guenther,S., Rothberg,J.M. and Karch,H. (2011) Prospective genomic characterization of the German enterohemorrhagic Escherichia coli O104:H4 outbreak by rapid next generation sequencing technology. *PLoS One* **6**, e22751.

Miller,R.S., Miller,W.G., Behringer,M., Hariharan,H., Matthew,V. and Oyarzabal,O.A. (2010) DNA identification and characterization of Campylobacter jejuni and Campylobacter coli isolated from caecal samples of chickens in Grenada. *J. Appl. Microbiol* **108**, 1041-1049.

Morales,C.A., Musgrove,M., Humphrey,T.J., Cates,C., Gast,R. and Guard-Bouldin,J. (2007) Pathotyping of *Salmonella* enterica by analysis of single-nucleotide polymorphisms in cyaA and flanking 23S ribosomal sequences. *Environmental microbiology* **9**, 1047-1059.

Moran,A.P. and Penner,J.L. (1999) Serotyping of Campylobacter jejuni based on heat-stable antigens: relevance, molecular basis and implications in pathogenesis. *Journal of applied microbiology* **86**, 361-377.

NACMCF (2007) Analytical utility of Campylobacter methodologies. *Journal of food protection* **70**, 241-250.

NACMCF (2010) Response to Questions Posed by the Food Safety and Inspection Service Regarding Determination of the Most Appropriate Technologies for the Food Safety and Inspection Service To Adopt in Performing Routine and Baseline Microbiological Analyses. *Journal of food protection* **73**, 1160-1200.

Nakamura,S., Maeda,N., Miron,I.M., Yoh,M., Izutsu,K., Kataoka,C., Honda,T., Yasunaga,T., Nakaya,T., Kawai,J., Hayashizaki,Y., Horii,T. and Iida,T. (2008) Metagenomic diagnosis of bacterial infections. *Emerg. Infect. Dis.* **14**, 1784-1786.

Nakamura,S., Nakaya,T. and Iida,T. (2011) Metagenomic analysis of bacterial infections by means of high-throughput DNA sequencing. *Experimental biology and medicine (Maywood, N. J.)* **236**, 968-971.

Nielsen,E.M., Engberg,J., Fussing,V., Petersen,L., Brogren,C.H. and On,S.L. (2000) Evaluation of phenotypic and genotypic methods for subtyping Campylobacter jejuni isolates from humans, poultry, and cattle. *J Clin Microbiol* **38**, 3800-3810.

Nielsen,L.N., Sheppard,S.K., McCarthy,N.D., Maiden,M.C., Ingmer,H. and Krogfelt,K.A. (2010) MLST clustering of Campylobacter jejuni isolates from patients with gastroenteritis, reactive arthritis and Guillain-Barre syndrome. *J Appl. Microbiol* **108**, 591-599.

Nightingale,K.K., Windham,K., Martin,K.E., Yeung,M. and Wiedmann,M. (2005) Select *Listeria monocytogenes* subtypes commonly found in foods carry distinct nonsense mutations in inlA, leading to expression of truncated and secreted internalin A, and are associated with a reduced invasion phenotype for human intestinal epithelial cells. *Appl. Environ. Microbiol* **71**, 8764-8772.

Noller,A.C., McEllistrem,M.C., Pacheco,A.G., Boxrud,D.J. and Harrison,L.H. (2003) Multilocus variable-number tandem repeat analysis distinguishes outbreak and sporadic Escherichia coli O157:H7 isolates. *Journal of clinical microbiology* **41**, 5389-5397.

On,S.L., Dorrell,N., Petersen,L., Bang,D.D., Morris,S., Forsythe,S.J. and Wren,B.W. (2006) Numerical analysis of DNA microarray data of Campylobacter jejuni strains correlated with survival, cytolethal distending toxin and haemolysin analyses. *Int. J. Med. Microbiol.* **296**, 353-363.

Orsi,R.H., Borowsky,M.L., Lauer,P., Young,S.K., Nusbaum,C., Galagan,J.E., Birren,B.W., Ivy,R.A., Sun,Q., Graves,L.M., Swaminathan,B. and Wiedmann,M. (2008) Short-term genome evolution of *Listeria monocytogenes* in a non-controlled environment. *BMC Genomics* **9**, 539.

Painter, J., Hoekstra, R., Ayers, T., Tauxe, R., Braden, C. and Angulo, F. (2013) Attribution of foodborne illnesses, hospitalizations, and deaths to food commodities by using outbreak data, United States, 1998–2008.

Pang,J.C., Chiu,T.H., Helmuth,R., Schroeter,A., Guerra,B. and Tsen,H.Y. (2007) A pulsed field gel electrophoresis (PFGE) study that suggests a major world-wide clone of *Salmonella* enterica serovar Enteritidis. *Int. J. Food Microbiol* **116**, 305-312.

Paton,A.W., Woodrow,M.C., Doyle,R.M., Lanser,J.A. and Paton,J.C. (1999) Molecular characterization of a Shiga toxigenic Escherichia coli O113:H21 strain lacking eae responsible for a cluster of cases of hemolytic-uremic syndrome. *Journal of clinical microbiology* **37**, 3357-3361.

Perna,N.T., Plunkett,G., III, Burland,V., Mau,B., Glasner,J.D., Rose,D.J., Mayhew,G.F., Evans,P.S., Gregor,J., Kirkpatrick,H.A., Posfai,G., Hackett,J., Klink,S., Boutin,A., Shao,Y., Miller,L., Grotbeck,E.J., Davis,N.W., Lim,A., Dimalanta,E.T., Potamousis,K.D., Apodaca,J., Anantharaman,T.S., Lin,J., Yen,G., Schwartz,D.C., Welch,R.A. and Blattner,F.R. (2001) Genome sequence of enterohaemorrhagic Escherichia coli O157:H7. *Nature* **409**, 529-533.

Pettengill,J.B., McAvoy,E., White,J.R., Allard,M., Brown,E. and Ottesen,A. (2012) Using metagenomic analyses to estimate the consequences of enrichment bias for pathogen detection. *BMC research notes* **5**, 378.

Pires,S.M., Evers,E.G., van,P.W., Ayers,T., Scallan,E., Angulo,F.J., Havelaar,A. and Hald,T. (2009) Attributing the human disease burden of foodborne infections to specific sources. *Foodborne Pathog. Dis.* **6**, 417-424.

Pires,S.M. and Hald,T. (2010) Assessing the differences in public health impact of *Salmonella* subtypes using a bayesian microbial subtyping approach for source attribution. *Foodborne Pathog. Dis.* **7**, 143-151.

Prager,R., Liesegang,A., Rabsch,W., Gericke,B., Thiel,W., Voigt,W., Helmuth,R., Ward,L. and Tschape,H. (1999) Clonal relationship of *Salmonella* enterica serovar typhimurium phage type DT104 in Germany and Austria. *Zentralblatt fur Bakteriologie : international journal of medical microbiology* **289**, 399-414.

Rasko,D.A., Webster,D.R., Sahl,J.W., Bashir,A., Boisen,N., Scheutz,F., Paxinos,E.E., Sebra,R., Chin,C.S., Iliopoulos,D., Klammer,A., Peluso,P., Lee,L., Kislyuk,A.O., Bullard,J., Kasarskis,A., Wang,S., Eid,J., Rank,D., Redman,J.C., Steyert,S.R., Frimodt-Moller,J., Struve,C., Petersen,A.M., Krogfelt,K.A., Nataro,J.P., Schadt,E.E. and Waldor,M.K. (2011) Origins of the *E. coli* strain causing an outbreak of hemolytic-uremic syndrome in Germany. *N. Engl. J. Med.* **365**, 709-717.

Ridley,A. and Threlfall,E.J. (1998) Molecular epidemiology of antibiotic resistance genes in multiresistant epidemic *Salmonella* typhimurium DT 104. *Microbial drug resistance (Larchmont, N. Y.)* **4**, 113-118.

Saeed,A.M., Walk,S.T., Arshad,M. and Whittam,T.S. (2006) Clonal structure and variation in virulence of *Salmonella* enteritidis isolated from mice, chickens, and humans. *J AOAC. Int.* **89**, 504-511.

Sails,A.D., Swaminathan,B. and Fields,P.I. (2003) Utility of multilocus sequence typing as an epidemiological tool for investigation of outbreaks of gastroenteritis caused by Campylobacter jejuni. *J Clin Microbiol* **41**, 4733-4739.

Sarwari,A.R., Magder,L.S., Levine,P., McNamara,A.M., Knower,S., Armstrong,G.L., Etzel,R., Hollingsworth,J. and Morris,J.G., Jr. (2001) Serotype distribution of *Salmonella* isolates from food animals after slaughter differs from that of isolates found in humans. *J Infect. Dis.* **183**, 1295-1299.

Singer,R.S., Mayer,A.E., Hanson,T.E. and Isaacson,R.E. (2009) Do microbial interactions and cultivation media decrease the accuracy of *Salmonella* surveillance systems and outbreak investigations? *J. Food. Prot.* **72**, 707-713.

Swaminathan,B., Gerner-Smidt,P., Ng,L.K., Lukinmaa,S., Kam,K.M., Rolando,S., Gutierrez,E.P. and Binsztein,N. (2006) Building PulseNet International: an interconnected system of laboratory networks to facilitate timely public health recognition and response to foodborne disease outbreaks and emerging foodborne diseases. *Foodborne Pathog. Dis.* **3**, 36-50.

Tamburro,M., Ripabelli,G., Fanelli,I., Grasso,G.M. and Sammarco,M.L. (2010) Typing of *Listeria monocytogenes* strains isolated in Italy by inlA gene characterization and evaluation of a new cost-effective approach to antisera selection for serotyping. *J Appl. Microbiol* **108**, 1602-1611.

Tarr,C.L., Nelson,A.M., Beutin,L., Olsen,K.E. and Whittam,T.S. (2008) Molecular characterization reveals similar virulence gene content in unrelated clonal groups of Escherichia coli of serogroup O174 (OX3). *J. Bacteriol.* **190**, 1344-1349.

Tenover,F.C., Arbeit,R.D., Goering,R.V., Mickelsen,P.A., Murray,B.E., Persing,D.H. and Swaminathan,B. (1995) Interpreting chromosomal DNA restriction patterns produced by pulsed-field gel electrophoresis: criteria for bacterial strain typing. *Journal of clinical microbiology* **33**, 2233-2239.

Torpdahl,M., Sorensen,G., Ethelberg,S., Sando,G., Gammelgard,K. and Porsbo,L.J. (2006) A regional outbreak of S. Typhimurium in Denmark and identification of the source using MLVA typing. *Euro Surveill.* **11**, 134-136.

van Belkum A., Tassios,P.T., Dijkshoorn,L., Haeggman,S., Cookson,B., Fry,N.K., Fussing,V., Green,J., Feil,E., Gerner-Smidt,P., Brisse,S. and Struelens,M. (2007) Guidelines for the validation and application of typing methods for use in bacterial epidemiology. *Clin. Microbiol. Infect.* **13 Suppl 3**, 1-46.

Van Stelten,A., Simpson,J.M., Chen,Y., Scott,V.N., Whiting,R.C., Ross,W.H. and Nightingale,K.K. (2011) Significant shift in median guinea pig infectious dose shown by an outbreak-associated *Listeria monocytogenes* epidemic clone strain and a strain carrying a premature stop codon mutation in inlA. *Appl. Environ. Microbiol* **77**, 2479-2487.

Van,S.A., Simpson,J.M., Ward,T.J. and Nightingale,K.K. (2010) Revelation by single-nucleotide polymorphism genotyping that mutations leading to a premature stop codon in inlA are common among *Listeria monocytogenes* isolates from ready-to-eat foods but not human listeriosis cases. *Applied and environmental microbiology* **76**, 2783-2790.

Verghese,B., Lok,M., Wen,J., Alessandria,V., Chen,Y., Kathariou,S. and Knabel,S. (2011) comK prophage junction fragments as markers for *Listeria monocytogenes* genotypes unique to individual meat and poultry processing plants and a model for rapid niche-specific adaptation, biofilm formation, and persistence. *Appl. Environ. Microbiol* **77**, 3279-3292.

Ward,T.J., Evans,P., Wiedmann,M., Usgaard,T., Roof,S.E., Stroika,S.G. and Hise,K. (2010a) Molecular and phenotypic characterization of *Listeria monocytogenes* from U.S. Department of Agriculture Food Safety and Inspection Service surveillance of ready-to-eat foods and processing facilities. *J Food Prot* **73**, 861-869.

Ward,T.J., Usgaard,T. and Evans,P. (2010b) A targeted multilocus genotyping assay for lineage, serogroup, and epidemic clone typing of *Listeria monocytogenes*. *Applied and environmental microbiology* **76**, 6680-6684.

WHO (2007) Antigenic Formula or the *Salmonella* Serovars.

Wickham,M.E., Lupp,C., Mascarenhas,M., Vazquez,A., Coombes,B.K., Brown,N.F., Coburn,B.A., Deng,W., Puente,J.L., Karmali,M.A. and Finlay,B.B. (2006) Bacterial genetic determinants of non-O157 STEC outbreaks and hemolytic-uremic syndrome after infection. *J. Infect. Dis.* **194**, 819-827.

Wilson,M.K., Lane,A.B., Law,B.F., Miller,W.G., Joens,L.A., Konkel,M.E. and White,B.A. (2009) Analysis of the pan genome of Campylobacter jejuni isolates recovered from poultry by pulsed-field gel electrophoresis, multilocus sequence typing (MLST), and repetitive sequence polymerase chain reaction (rep-PCR) reveals different discriminatory capabilities. *Microbial ecology* **58**, 843-855.

Wilson,M.R., Allard,M.W. and Brown,E.W. (2013) The forensic analysis of foodborne bacterial pathogens in the age of wholeGÇÉgenome sequencing. *Cladistics*.

Withee,J. and Dearfield,K.L. (2007) Genomics-based food-borne pathogen testing and diagnostics: Possibilities for the US Department of Agriculture's Food Safety and Inspection Service. *Environmental and molecular mutagenesis* **48**, 363-368.

Zautner,A.E., Ohk,C., Tareen,A.M., Lugert,R. and Gross,U. (2012) Epidemiological association of Campylobacter jejuni groups with pathogenicity-associated genetic markers. *BMC microbiology* **12**, 171.

Zheng,J., Keys,C.E., Zhao,S., Ahmed,R., Meng,J. and Brown,E.W. (2011) Simultaneous analysis of multiple enzymes increases accuracy of pulsed-field gel electrophoresis in assigning genetic relationships among homogeneous *Salmonella* strains. *J Clin. Microbiol* **49**, 85-94.