

Polymorphic Amplified Typing Sequences (PATs) Strain Typing System Accurately Discriminates a Set of Temporally and Spatially Disparate *Escherichia coli* O157 Isolates Associated with Human Infection

Indira T. Kudva^{1,*}, Sandra Smole², Robert W. Griffin³, Jeonifer Garren³, Nimisha Kalia², Megan Murray^{3,4}, Manohar John^{3,5,†}, Ralph Timperi^{2,††}, and Stephen B. Calderwood^{3,5,6,*}

¹Food Safety and Enteric Pathogens Research Unit, National Animal Disease Center, Agricultural Research Service, U.S. Department of Agriculture, Ames, IA, 50010

²William A. Hinton State Laboratory Institute, Massachusetts Department of Public Health, Jamaica Plain, MA 02130

³Division of Infectious Diseases, Massachusetts General Hospital, Boston, MA 02114

⁴Department of Epidemiology, Harvard School of Public Health, Boston, MA 02115

⁵Department of Medicine, Harvard Medical School, Boston, MA 02115

⁶Department of Microbiology and Immunobiology Harvard Medical School Boston, MA 02115

[†]Pathovacs Inc., Ames, IA

^{††}Association of Public Health Laboratories, Silver Spring, MA

Abstract: Polymorphic Amplified Typing Sequences (PATs) is a PCR-based *Escherichia coli* O157 (O157) strain typing system. Here, we show that PATs compares excellently with Pulsed-Field Gel Electrophoresis (PFGE) in that both methods cluster geographically diverse O157 isolates similarly. Comparative analysis of the results obtained in this simulated “blind” study attests to the discriminating power and applicability of PATs to epidemiological/nosocomial situations.

Keywords: Bacteria, O157, Strain-typing, PCR, PATs, PFGE.

INTRODUCTION

Escherichia coli O157:H7(O157) isolates are commonly typed by pulsed-field gel electrophoresis (PFGE), following digestion of genomic DNA with the restriction enzyme *Xba*I. The choice of this infrequently cutting restriction enzyme is dictated by the fact that it is an effective cutter that consistently yields a decipherable profile [1-3]. However, inherent limitations of the standard PFGE protocol have spurred several modifications, including: (i) digestion of genomic DNA with additional restriction enzymes to better relate strains; (ii) development of uniform size standards and protocols for universal use; and (iii) institution of protocols to decrease sample turn aroundtime and decrease technical errors. Despite these modifications and use of pattern-recognition computer software, resolution of co-migrating and

undigested DNA bands poses a significant interpretive challenge [2-8].

Polymorphic Amplified Typing Sequences (PATs) was developed based on genetic events underlying the polymorphisms observed at the *Xba*I sites, and another infrequent restriction enzyme site, *Avr*II (*Bln*I), in the O157 genome [9, 10]. Primer pairs targeting 8 polymorphic *Xba*I sites associated with insertion-deletions (Indels) and 7 polymorphic *Avr*II sites associated with Indels or single nucleotide polymorphisms (SNPs) were incorporated, along with those amplifying four virulence genes (*stx*1, *stx*2, *eae*, *hly*A), to develop the PATs O157 strain typing system [9-11]. PATs was shown to accurately discriminate between unrelated isolates and relate closely linked isolates when evaluated against 46 O157 isolates associated with human disease [9-11], as well as 25 bovine O157 isolates [12]. To further confirm the real world applicability of PATs, including epidemiological situations that demand strain typing of unknown isolates, we devised a “blind study” wherein the details of the O157 isolates being typed were withheld from the researcher until after analysis. Such a study, we reasoned, would minimize subjective biases. The results of PATs analyses were compared with those obtained using PFGE for the same isolates.

*Address correspondence to these authors at the Division of Infectious Diseases, Massachusetts General Hospital, Boston, MA 02114; Tel: 617-726-3811; Fax: 617-726-7416; E-mail: scalderswood@partners.org; Food Safety and Enteric Pathogens Research Unit, National Animal Disease Center/ARS/USDA, 1 North, Bldg. 20, 1121, Ames, IA, 50010; Tel: 515-337-7376; Fax: 515-337-7438; E-mail: Indira.Kudva@ars.usda.gov

MATERIALS AND METHODS

Bacterial Strains

A total of 48 O157 isolates that were associated with human disease and isolated from diverse sources/geographic locations were obtained from collections maintained at the William A. Hinton State Laboratory Institute, Jamaica Plain, MA. Details about the strains/isolates were provided only at the end of the study, and can be observed on the two dendograms based on PATS and PFGE profiles (Figs 1 and 2). The isolates were labeled MLVA 001 to 048.

PFGE

Standard PFGE method was used to analyze the 48 O157 isolates as previously described [4, 13]. Briefly, the genomic DNA of each isolate was embedded in separate agarose plugs and digested at 37°C for 2 h with 30 U of *Xba*I (Gibco BRL, Grand Island, N.Y.) per plug. The plugs were loaded onto a 1% agarose–tris buffer gel (SeaKem GoldAgarose; BioWhittaker Molecular Applications, Rockland, Maine), and PFGE was performed with a CHEF Mapper XA apparatus (Bio-Rad Laboratories, Hercules, Calif.). DNA was electrophoresed for 18 h at a constant voltage of 200V (6 V/cm), with a pulse time of 2.2 to 54.2 s, an electric field angle of 120°, and a temperature of 14°C, before being stained with ethidium bromide. PFGE gels were analyzed using Molecular Analyst Fingerprinting Plus software (Bio-Rad). The dendogram for the *Xba*I-based PFGE profiles was constructed using the UPGMA cluster analysis.

PATS

For PATS, a colony lysate of each O157 isolate was subjected to hot start, touchdown PCR with each individual primer pair as described previously [11]. PCR was carried out on the GeneAmp PCR system 2400 thermal cycler (PE Biosystems, Foster City, CA.), using 10 µl of colony lysate, 200 pmoles of each primer, 800 µM dNTPs, 1x diluted Ex Taq™ enzyme buffer and 2.5 units of TaKaRa Ex Taq™ DNA polymerase. An amplification segment of 20 cycles was set where the annealing temperature started at 73°C, to touchdown at 53°C at the end of those cycles. Then, another amplification segment of 10 cycles was set, using the last annealing temperature of 53°C. Each reaction was done in triplicate.

PCR reactions were initially screened for the presence or absence of amplicons. Amplicons, when present, were purified using the QIAquick PCR purification kit (Qiagen, Valencia, Ca.), and digested with the *Avr*II restriction enzyme (New England Biolabs, Beverly, Ma.) to confirm the presence of an *Avr*II restriction site within the amplicon. Undigested and digested DNA fragments were resolved on a 4% agarose gel prepared with a combination of 3% Nusieve GTG agarose (FMC BioProducts, Rockland, Me.) and 1% agarose (Shelton Scientific Inc., Shelton, Ct.) and stained with ethidium bromide. The presence/absence of, and variations in the amplicons was tabulated as before for *Xba*I-based, *Avr*II-based, and combined-PATS; combined-PATS being *Xba*I-, *Avr*II- and virulence genes-based (Tables 1 and 2) [12]. Dendogram for the PATS data was constructed using

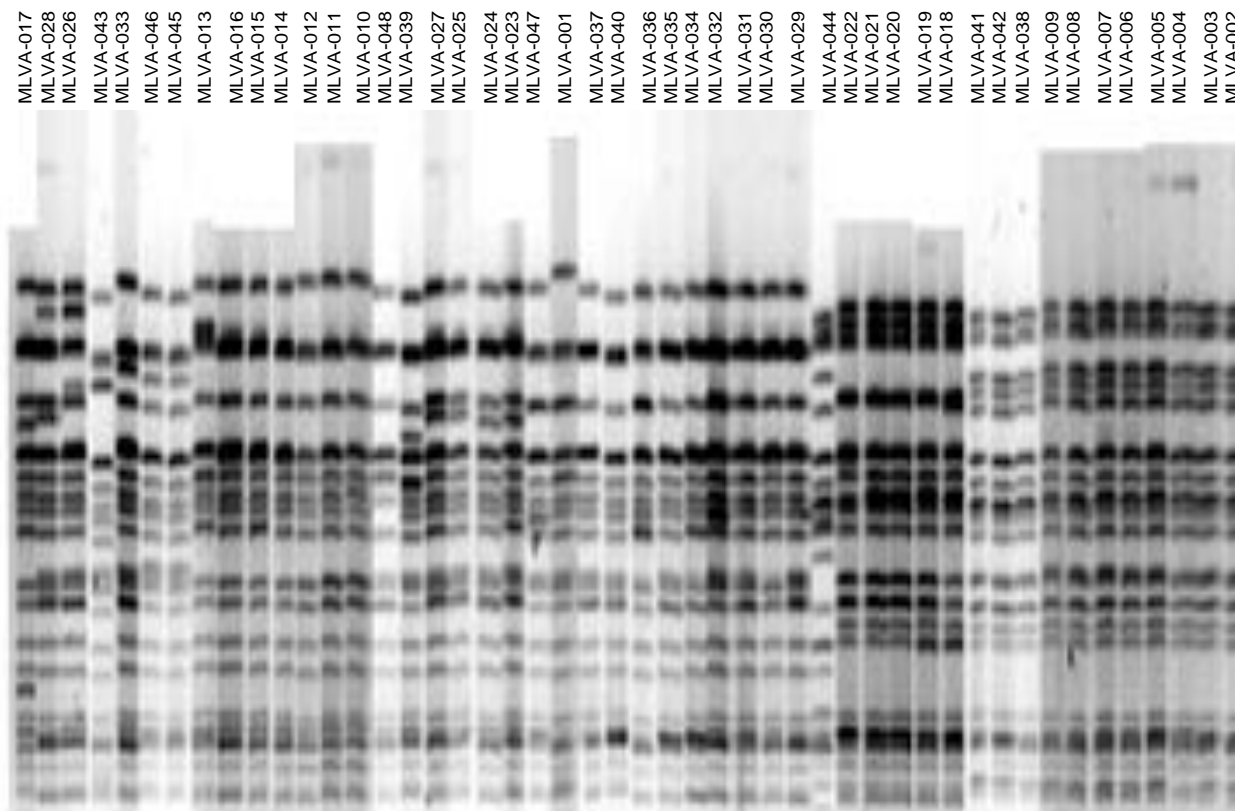


Fig. (1). *Xba*I-based PFGE patterns of the 48 O157 isolates.

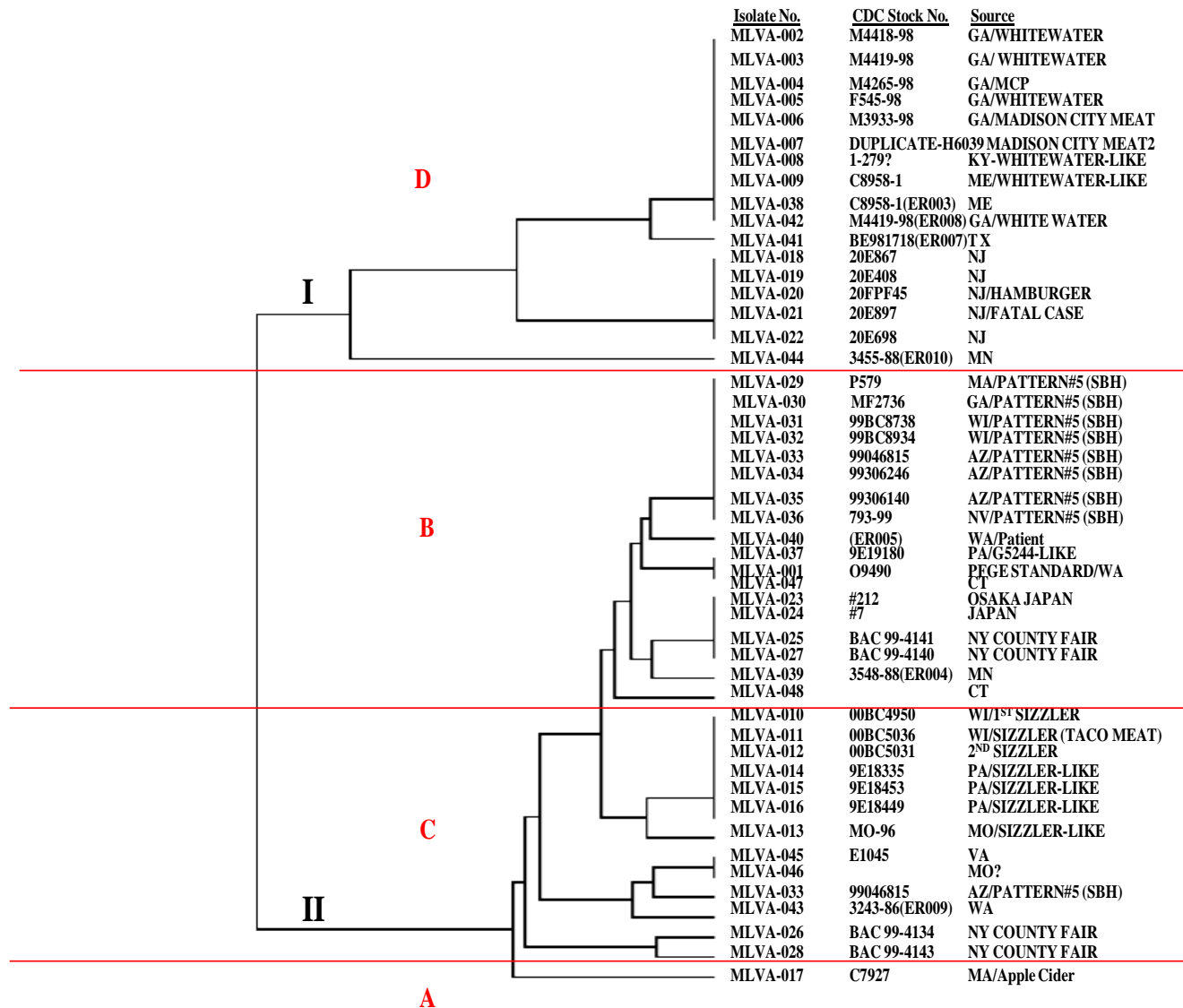


Fig. (2). Analysis of the relatedness between 48 O157 isolates using *XbaI*-based PFGE. Dendrogram was constructed using the UPGMA cluster analysis. PFGE gels were analyzed using Molecular Analyst Fingerprinting Plus software (Bio-Rad). Clades are labeled with roman numerals and groups with alphabets, A-D.

the unweighted pair-group method with arithmetic means (UPGMA) option in phylogenetic analysis using parsimony software (PAUP; Sinauer Associates, Inc., Publishers, Sunderland, Ma.) as described previously [9-11].

RESULTS AND DISCUSSION

Both PATS and PFGE distributed the 48 O157 isolates into two major clades, each including the same set of isolates. Although PFGE generated more clusters in one of the clades, examination of the identities of the isolates showed that PATS had clustered O157 isolates from the same geographical locations better than PFGE. Overall, PATS and PFGE yielded overlapping typing results, which confirmed the ability of PATS to make an unbiased assessment of bacterial isolates while maintaining the discriminatory ability of PFGE.

The PFGE patterns obtained for the 48 O157 isolates can be observed in Fig. (1). The dendrogram based on these PFGE patterns distributed these O157 isolates into two clades containing three and seven clusters respectively (Fig. 2). PFGE clade I was the same as PATS clade II, and PFGE clade II as the same as PATS clade I, with the same O157 isolates in these corresponding clades (Figs, 2 and 3). The additional clusters observed with PFGE within the clade may have been due to its inherent over-discrimination of isolates, which may be attributed to misinterpretation of complex electrophoretic patterns. This often encountered drawback of analysis using PFGE has been largely addressed with specialized instrumentation and sophisticated software to interpret the patterns generated, development of standardized protocols with additional restriction enzymes and standards; yet DNA degradation, incomplete digestion, co-migration of similar sized DNA fragments, as well as non-homologous

Table 1A *Xba*I-based PATS Profiles of the 48 O157 Isolates

PATS type ^a	PCR Amplification and Restriction Digestion Patterns of Amplicons obtained Using 8 PATS Primer Pairs ^b								Isolates ^c
	IK8	IK19	IK25	IK114	IK118	IK123	IKB3	IKB5	
1	2	2	0	2	2	2	0	0	MLVA 002, 003, 004, 005, 006, 007, 008, 009, 038, 041, 042
2	2	2	0	2	2	2	0	2	MLVA 001, 023, 024, 025, 029, 030, 031, 032, 033, 034, 035, 036, 037, 039, 040, 045, 046, 047, 048
3	2	2	0	2	2	2	2	2	MLVA 010, 011, 012, 026, 027, 028, 043
4	2	2	0	2	2	0	0	2	MLVA 013, 014, 015, 016
5	2	2	0	2	2	0	2	0	MLVA 018
6	0	2	0	2	2	0	0	2	MLVA 017
7	2	2	0	2	2	2	2	0	MLVA 019, 020, 021
8	2	0	0	2	2	2	2	0	MLVA 022
9	2	2	0	2	0	2	0	0	MLVA 044

Table 1B. *Avr*II-based PATS Profiles of the 48 O157 Isolates

PATS type	PCR Amplification and Restriction Digestion Patterns of Amplicons obtained Using 7 PATS Primer Pairs ^d							Isolates
	IKNR3	IKNR7	IKNR10	IKNR12	IKNR16	IKNR27	IKNR33	
1	2	1	1	2	1	2	2	MLVA 002, 003, 004, 005, 006, 007, 008, 009, 038, 041, 042
2	2	2	2	2	1	2	2	MLVA 018, 019, 020, 021, 022
3	2	2	2	2	2	3	2	MLVA 017
4	1	2	2	2	1	2	2	MLVA 044
5	2	2	2	2	2	2	2	MLVA 001, 010, 011, 012, 013, 014, 015, 016, 023, 024, 025, 026, 027, 028, 029, 030, 031, 032, 033, 034, 035, 036, 037, 039, 040, 043, 045, 046, 047, 048

^a PATS types are designated arbitrarily with different numbers.

^b Prefixes of each PATS primer pair A/B are indicated. 0, no amplicon; 2, amplicon with one *Xba*I site.

^c *E. coli* O157:H7 isolates from different locations that fell within a given PATS type.

^d Prefixes of each PATS primer pair A/B are indicated. 0, no amplicon; 1, amplicon without *Avr*II site; 2, amplicon with one *Avr*II site; 3, amplicon with an additional *Avr*II site due to a SNP.

DNA migrating as same-size bands, often result in “untypeable” or incorrect profiles [3-8, 14-17]. DNA sequence-based typing methods such as multilocus sequence typing (MLST) and multilocus variable number of tandem repeats analysis (MLVA), are becoming potential alternatives to PFGE, with less subjectivity in the interpretation of results, comparable and transferable data and ease of automation [18, 19]. However, since these techniques still rely on specialized instrumentation and software to interpret data, the PATS technique might be useful for application in any laboratory, particularly

those outside of standardized networks such as PulseNet [2, 8].

PATS grouped the 48 O157 isolates into 9 *Xba*I-PATS types (Table 1A), 5 *Avr*II-based PATS types (Table 1B), and 9 combined-PATS types (Table 2). Both the *Xba*I-based and combined-PATS placed the O157 isolates into same groups which suggests that much of the discrimination was provided primarily by the polymorphic *Xba*I sites with limited discrimination by the polymorphic *Avr*II sites. Of these, PATS type 1, 2, 3, 6, and 9 were the same as PATS type 8, 18, 19,

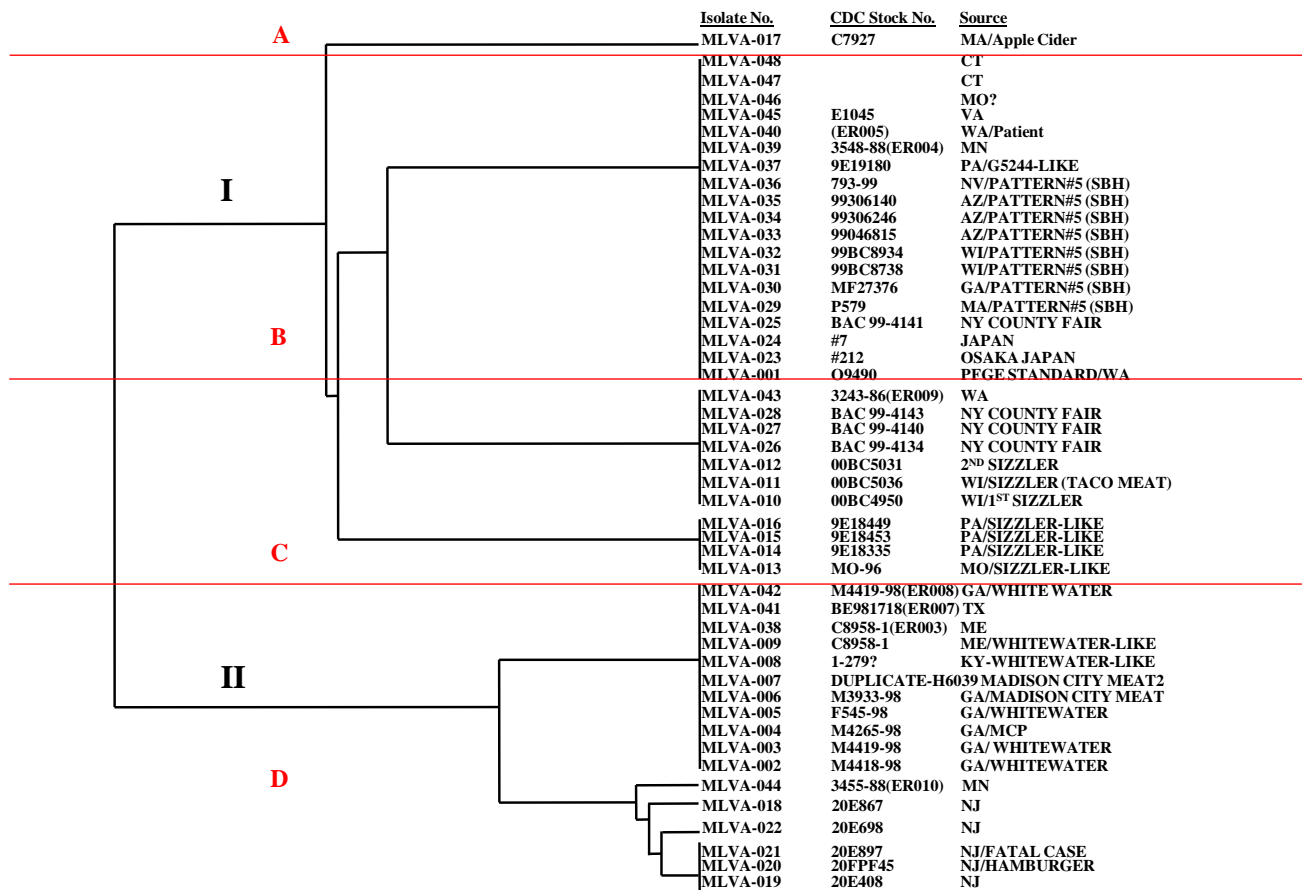


Fig. (3). Analysis of relatedness between 48 O157 isolates using *Xba*I-based PATS. Dendrogram was constructed using the UPGMA option in the phylogenetic analysis using parsimony (PAUP; Sinauer Associates, Inc., Publishers, Sunderland, Ma.). Clades are labeled with roman numerals and groups with alphabets, A-D, corresponding with those seen in the PFGE dendrogram (Fig. 2).

11, and 9 observed in a previous analysis of 46 O157 isolates associated with human disease (Table 2) [11]. Subsequent detailed analysis of the isolates used in this study showed that this repetition of PATS profiles could be associated with the same geographic locations for most of the isolates in these groups [11]. The dendrogram based on the 9 PATS types distributed the 48 O157 isolates into two clades bearing three and four clusters, respectively (Fig. 3). The O157 isolates in each cluster could be linked by geographic location/source once their identities were revealed.

As depicted in Figs. (2 and 3), the 48 O157 isolates could be divided into 2 clades and 4 broad groups (A-D) based on their placement in the dendrogram, and these overlapped between PATS and PFGE. Similar to previous observations, isolates from different outbreaks that shared the same PATS clade/group also shared the same clade/group by PFGE [10,11]. Some of these isolates from different outbreaks were also from different years warranting a need to further investigate their epidemiologic links, however, as these were banked isolates with limited information available this evaluation could not be done. Nevertheless, shared PFGE profiles between O157 isolates from different sources as well as re-emergence of older profiles have been reported in the PulseNet database (www.cdc.gov/ecoli/2012/O157H7-11-12/index.html) making our observation less implausible.

CONCLUSION

As a user-friendly strain typing technique, which is both easy to perform and interpret, PATS favorably complements PFGE and provides comparable results without extensive instrumentation, less subjectivity in the interpretation of results and faster than the currently used 24-hour PFGE protocol. With profiles that are easy to track and transfer between laboratories, which is critical in both epidemiological and nosocomial outbreak situations, this simulated study reaffirmed the ability of PATS to rapidly and accurately type pathogens even in a “field” situation.

ACKNOWLEDGEMENT

This study was supported in part by Grant Cooperative Agreement Number U60-CCU303019-15 from the Centers for Disease Control-Association of Public Health Laboratories to S. B. C. and I. T. K. This work was also supported in part by the Center for Integration of Medicine and Innovative Technology (CIMIT), through U.S. Army Medical Research and Materiel Command Cooperative Agreement Number DAMD17-02-2-0006, award to S. B. C., I. T. K., and M. J.

Table 2. PATS Profiles of 48 O157 Isolates based on Polymorphic XbaI and AvrII Restriction Sites, and Virulence Genes

PATS Type ^a	PCR Amplification and Restriction Digestion Patterns of Amplicons obtained Using 15 PATS – 4 Virulence Gene Primer Pairs ^b																			Isolates ^c		
	Polymorphic XbaI Sites									Polymorphic AvrII sites						Virulence Genes						
	IK8	IK19	IK25	IK114	IK118	IK123	IKB3	IKB5	-	IKNR3	IKNR7	IKNR10	IKNR12	IKNR16	IKNR27	IKNR33	-	stx ₁	stx ₂		eae	hly _A
1 (prev. 8) ^d	2	2	0	2	2	2	0	0		2	1	1	2	1	2	2		0	1	1	1	MLVA 002, 003, 004, 005, 006, 007, 008, 009, 038, 041, 042
2 (prev.18)	2	2	0	2	2	2	0	2		2	2	2	2	2	2	2		1	1	1	1	MLVA 001, 023, 024, 025, 029, 030, 031, 032, 033, 034, 035, 036, 037, 039, 040, 045, 046, 047, 048
3 (prev.19)	2	2	0	2	2	2	2	2		2	2	2	2	2	2	2		1	1	1	1	MLVA 010, 011, 012, 026, 027, 028, 043
4	2	2	0	2	2	0	0	2		2	2	2	2	2	2	2		1	1	1	1	MLVA 013, 014, 015, 016
5	2	2	0	2	2	0	2	0		2	2	2	2	1	2	2		0	1	1	1	MLVA 018
6 (prev.11)	0	2	0	2	2	0	0	2		2	2	2	2	2	3	2		1	1	1	1	MLVA 017
7	2	2	0	2	2	2	2	0		2	2	2	2	1	2	2		0	1	1	1	MLVA 019, 020, 021
8	2	0	0	2	2	2	2	0		2	2	2	2	1	2	2		0	1	1	1	MLVA 022
9 (prev. 9)	2	2	0	2	0	2	0	0		1	2	2	2	1	2	2		0	1	1	1	MLVA 044

^a PATS types are designated arbitrarily with different numbers.

^b Prefixes of each PATS primer pair A/B and virulence gene primer pair F/R are indicated. 0, no amplicon; 1, amplicon without XbaI or AvrII site; 2, amplicon with one XbaI or AvrII site; 3, amplicon with an additional AvrII site due to a SNP

^c Isolates of *E. coli* O157:H7 from various outbreaks that fell within a given PATS type.

^d PATS number of the identical profile observed in a previous study [11].

U.S. Utility Patent Pending for PATS- a bacterial strain typing technology (Filed Nov. 1, 2001).

Mention of trade names or commercial products in this article is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the U.S. Department of Agriculture. USDA is an equal opportunity provider and employer.

REFERENCES

[1] Tenover FC, Arbeit RD, Goering RV, et al. Interpreting chromosomal DNA restriction patterns produced by pulsed-field gel electrophoresis: criteria for bacterial strain typing. J Clin Microbiol 1995; 33: 2233-9.

[2] Olive DM, Bean P. Principles and applications of methods for DNA-based typing of microbial organisms. J Clin Microbiol 1999; 37: 1661-9.

[3] Davis MA, Hancock DD, Besser TE, Call DR. Evaluation of pulsed-field gel electrophoresis as a tool for determining the degree of genetic relatedness between strains of *Escherichia coli* O157:H7. J Clin Microbiol 2003; 41:1843-9.

[4] Birren B, Lai E. Pulsed-Field Gel Electrophoresis - A Practical Guide. Academic Press, Inc., San Diego, Ca.1993.

[5] Harsono KD, Kaspar CW, Luchansky JB. Comparison and genomic sizing of *Escherichia coli* O157:H7 isolates by pulsed-field gel electrophoresis. Appl Environ Microbiol 1993; 59: 3141-4.

[6] Johnson JM, Weagant SD, Jinnemen KC, Bryant JL. Use of pulsed-field gel electrophoresis for epidemiological study of *Escherichia coli* O157:H7 during a food-borne outbreak. Appl Environ Microbiol 1995; 61: 2806-8.

[7] Murase T, Yamai S, Watanabe H. Changes in pulsed-field gel electrophoresis patterns in clinical isolates of enterohemorrhagic *Escherichia coli* O157:H7 associated with loss of Shiga toxin genes. Curr Microbiol 1999; 38: 48-50.

[8] Hunter SB, Vauterin P, Lambert-Fair MA, et al. Establishment of a Universal Size Standard Strain for Use with the PulseNet Standardized Pulsed-Field Gel Electrophoresis Protocols: Converting the National Databases to the New Size Standard. J Clin Microbiol 2005; 43: 1045-50.

[9] Kudva IT, Evans PS, Perna NT, et al. Strains of *Escherichia coli* O157:H7 differ primarily by insertions or deletions, not single-nucleotide polymorphisms. J Bacteriol 2002a; 184: 1873-9.

[10] Kudva IT, Evans PS, Perna NT, et al. Polymorphic amplified typing sequences provide a novel approach to *Escherichia coli* O157:H7 strain typing. J Clin Microbiol 2002b; 40: 1152-1159.

[11] Kudva IT, Griffin RW, Murray M, et al. Insertions, deletions, and single-nucleotide polymorphisms at rare restriction enzyme sites enhance discriminatory power of polymorphic amplified typing sequences, a novel strain typing system for *Escherichia coli* O157:H7. J. Clin. Microbiol. 2004; 42: 2388-2397.

- [12] Kudva IT, Davis MA, Griffin RW, *et al.* Polymorphic Amplified Typing Sequences (PATS) and Pulsed-Field Gel Electrophoresis (PFGE) Yield Comparable Results in the Strain Typing of a Diverse Set of Bovine *Escherichia coli* O157 Isolates. *Int J Microbiol* 2012; Epub: article id 140105; doi:10.1155/2012/140105.
- [13] Bohm H, Karch H. DNA fingerprinting of *Escherichia coli* O157:H7 strains by pulsed field gel electrophoresis. *J Clin Microbiol* 1992; 30: 2169-72.
- [14] Rementería A, Gallego L, Quindos G, Garaizar J. Comparative evaluation of three commercial software packages for analysis of DNA polymorphism patterns. *Clin Microbiol Infect* 2001; 7: 331-6.
- [15] Koort JMK, Lukinmaa S, Rantala M, *et al.* Technical improvement to prevent DNA degradation of enteric pathogens in pulsed-field gel electrophoresis. *J Clin Microbiol* 2002; 40: 3497-8.
- [16] Ribot EM, Fair MA, Gautom R, *et al.* Standardization of pulsed-field gel electrophoresis protocols for the subtyping of *Escherichia coli* O157:H7, *Salmonella*, and *Shigella* for Pulse Net Food Pathog Dis 2006; 3: 59-67.
- [17] Noller AC, McEllistrem MC, Stine OC, *et al.* Multilocus sequence typing reveals a lack of diversity among *Escherichia coli* O157:H7 isolates that are distinct by pulsed-field gel electrophoresis. *J Clin Microbiol* 2003; 41: 675-9.
- [18] Hyytiä-Trees E, Smole SC, Fields PA, Swaminathan B, Ribot EM. Second generation subtyping: a proposed PulseNet protocol for multiple-locus variable-number tandem repeat analysis of Shiga toxin-producing *Escherichia coli* O157 (STEC O157). *Food Pathog Dis* 2006; 3: 118-31.
- [19] Hyytiä-Trees E, Lafon P, Vauterin P, Ribot EM. Multi-laboratory 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. *Food Pathog Dis* 2010; 7: 129-36.

Received: September 27, 2013

Revised: October 02, 2013

Accepted: October 02, 2013

© Kudva *et al.*; Licensee Bentham Open.

This is an open access article licensed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted, non-commercial use, distribution and reproduction in any medium, provided the work is properly cited.