strains from the same clonal group from Southeast Asian travelers arriving in Japan. J Clin Microbiol. 1997; 35:3150-55.

32. Nair GB, Ramamurthy T, Bhattacharya SK, Dutta B, Takeda Y, Sack DA. Global dissemination of *Vibrio parahaemolyticus* serotype O3:K6 and its serovariants. Clin Microbiol Rev. 2007; 20:39-48.

33. Matsumoto C, Okuda J, Ishibashi M, Iwanaga M, Garg P, Rammamurthy T, Wong HC, Depaola A, Kim YB, Albert MJ, Nishibuchi M. Pandemic spread of an O3:K6 clone of *Vibrio parahaemolyticus* and emergence of related strains evidenced by arbitrarily primed PCR and *toxRS* sequence analyses. J Clin Microbiol. 2000; 38:578-85.

34. Williams TL, Musser SM, Nordstrom JL, DePaola A, Monday SR. Identification of a protein biomarker unique to the pandemic O3:K6 clone of *Vibrio parahaemolyticus*. J Clin Microbiol. 2004; 42:1657-65.

35. Park KS, Ono T, Rokuda M, Jang MH, Okada K, Iida T, Honda T. Functional characterization of two type III secretion systems of *Vibrio parahaemolyticus*. Infect Immun. 2004; 72:6659-65.

36. Kam KM, Luey CK, Parsons MB, Cooper KL, Nair GB, Alam M, Islam MA, Cheung DT, Chu YW, Ramamurthy T, Pazhani GP, Bhattacharya SK, Watanabe H, Terajima J, Arakawa E, Ratchtrachenchai OA, Huttayananont S, Ribot EM, Gerner-Smidt P, Swaminathan B; *Vibrio parahaemolyticus* PulseNet PFGE Protocol Working Group. Evaluation and validation of a PulseNet standardized pulsed-field gel electrophoresis protocol for subtyping *Vibrio parahaemolyticus:* an international multicenter collaborative study. J Clin Microbiol. 2008 46:2766-2773.

37. Chowdhury NR, Chakraborty S, Ramamurthy T, Nishibuchi M, Yamasaki S, Takeda Y, Nair GB. Molecular evidence of clonal *Vibrio parahaemolyticus* pandemic strains. Emerg Infect Dis. 2000; 6:631-36.

38. González-Escalona N, Martinez-Urtaza J, Romero J, Espejo RT, Jaykus LA, DePaola A. Determination of molecular phylogenetics of *Vibrio parahaemolyticus* strains by multilocus sequence typing. J Bacteriol. 2008; 190:2831-40.

39. Kimura B, Sekine Y, Takahashi H, Tanaka Y, Obata H, Kai A, Morozumi S, Fujii T. Multiple-locus variable-number of tandem-repeats analysis distinguishes *Vibrio parahaemolyticus* pandemic O3:K6 strains. J Microbiol Methods. 2008; 72:313-20.

40. Chowdhury NR, Stine OC, Morris JG, Nair GB. Assessment of evolution of pandemic *Vibrio parahaemolyticus* by multilocus sequence typing. J Clin Microbiol. 2004; 42:1280-82.

Title: Molecular epidemiology of *Shigella sonnei* in Thailand during 2007 to 2008.

Name of researcher; Chaiwat Pulsrikarn
Affiliation; National Institute of Health, Nonthaburi, Thailand

Summary:

Shigella sonnei 200 strains had been confirmed by serological method during 2007 to 2008. That devised to phase I, phase II and phase I&II to 23% (46/200), 41.5% (83/200) and 35.5% (71/200), respectively. In year 2007, we found epidemic of S. sonnei phase I and S. sonnei phase II in May and July, respectively. In year 2008, epidemic of phase of S. sonnei had change to phase I&II in April and July. In this year we had found increasing of S. sonnei phase II in November.

All most of Shigellae were susceptible to all drugs (Amp, Amc, Ctx, Cip and Nal). We found 34.5% (69 from 200 strains) resist to Ampicillin and derivatives as Amoxycillin/clavulinic acid and Cefotaxime.

The phenotypic ESBL had test by double disk method with Cetazidime; Caz and Cetazidime/Clavulanic acid; Caz/Cla and Cefotaxime; Ctx and Cefotaxime/Clavulanic acid; Ctx/Cla found 11.6% (8 from 69 strains) and 18.8 % (13 from 69 strains), respectively.

Purpose:

1. To determine genetic relationship of *Shigella sonnei* in Thailand during 2007-2008.
2. To determine the antimicrobial susceptibility patterns among outbreak and non outbreak of *Shigella* spp. in Thailand isolated during 2007-2008.

Methods:

Bacterial strain:

1) Isolates and identification from patient specimens will be performed followed standard conventional method for detection *Shigella* at collaborative hospital laboratories through Thailand.
2) National Institute of Health receives all isolates suspected to be *Shigella* and test for biochemical utilization follow Ewing(1986) and classification into serotypes with commercial available antiserum, interpretation of the result will be done followed the Kauffmann White Scheme.
3) For Antimicrobial susceptibility of *Shigella* will be tested for phenotypic confirmatory ESBL producing with ceftazidime(30ug), ceftazidime/clavulinic acid(30/10ug) and cefotaxime(30ug), cefotaxime/clavulinic acid(30/10ug)

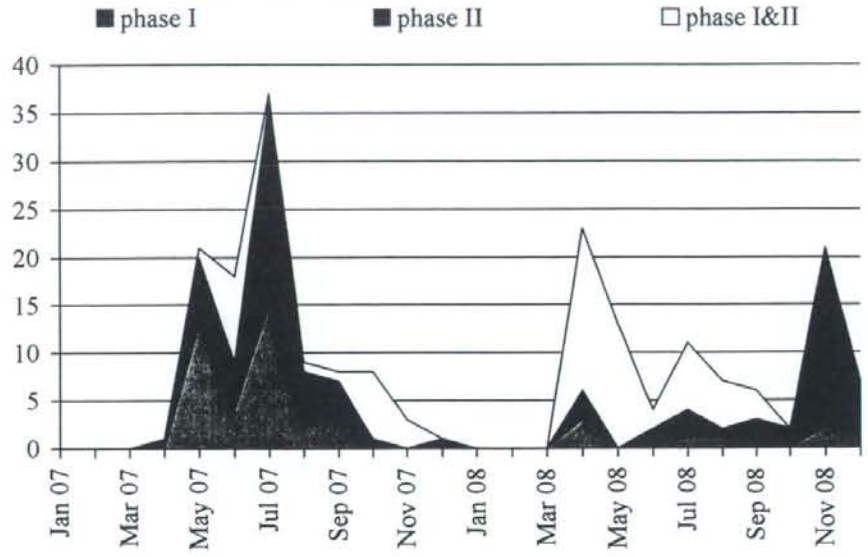
antimicrobial disks followed the procedure as described by National Clinical Laboratory Standard (2002)

4) DNA fingerprint of *Shigella* strains will be performed by Standard Pulsed-field gel electrophoresis protocol for *Escherichia coli* O157:H7.[3]
5) The DNA fingerprint generated by PFGE will analyzed by Bionumeric software according to the criteria proposed by Tenover (1995).

Results:

1. Epidemiological of Shigellosis during 2007 to 2008

Picture 1: Demographic data of *Shigella sonnei* between 2007 to 2008



Shigella sonnei were increase during May, 2007 which was confirmed to S. sonnei phase I 61.9% (13 from 21 strains). And increase again in July by S. sonnei phase II 56.8% (21 from 37 strains). In Year 2008, we found 3 times of Shigellosis first in April by S. sonnei phase I&II 73.9% (17 from 23 strains), the second in July by same phase 63.6% (7 from 11 strains) and found in November by S. sonnei pahse II 90.5% (19 from 21 strains). On this investigate not found Shigellosis during January to February couple year.

2. Antimicrobial susceptibility test

Shigella sonnei 200 strains were test antimicrobial susceptibility with 5 drugs: Ampicillin; Amp, Amoxycillin+Clavulinic acid; Amc, Cefotaxime; Ctx, Ciprofloxacin; Cip and Nalidixic acid; Nal follow NCCLS 2002.

The result show 62.0% of *Shigella sonnei* were susceptible all drugs test. Found 20.0% were resistant to Amp and Amc and 11.5% were resistant to Amp only as show in Table 1. All 69 strains of Shigella that show resistant to Amp had check ESBL test.

Table 1: Antimicrobial Resistant Pattern of *Shigella sonnei*

Antimicrobial Resistant Pattern	**No. of *Shigella sonnei* (%)**			
	phase I (N = 46)	**phase II (N = 83)**	**phase I&II (N = 71)**	**All (N = 200)**
Susceptible	18 (39.1)	49 (59.0)	57 (80.3)	124 (62.0)
AmpAmcCipCtxNal	1 (2.2)	1 (1.2)	-	2 (1.0)
AmpAmcCtxNal	1 (2.2)	1 (1.2)	-	2 (1.0)
AmpAmc	16 (34.8)	16 (19.3)	8 (11.3)	40 (20.0)
AmpNal	1 (2.2)	1 (1.2)	-	2 (1.0)
Amp	6 (13.0)	13 (15.7)	4 (5.6)	23 (11.5)
Nal	3 (6.5)	2 (2.4)	2 (2.8)	7 (3.5)

The result of ESBL initial screening test found suspected 39 strains from 69 strains (56.5%). For ESBL phenotypic confirmed test found 13 and 8 strains with antimicrobial double disk test, respectively. (Picture 1 and Table 2)

Picture 2: ESBL initial screen test* with Cefotaxime (Ctx) and Cetazidime (Caz)

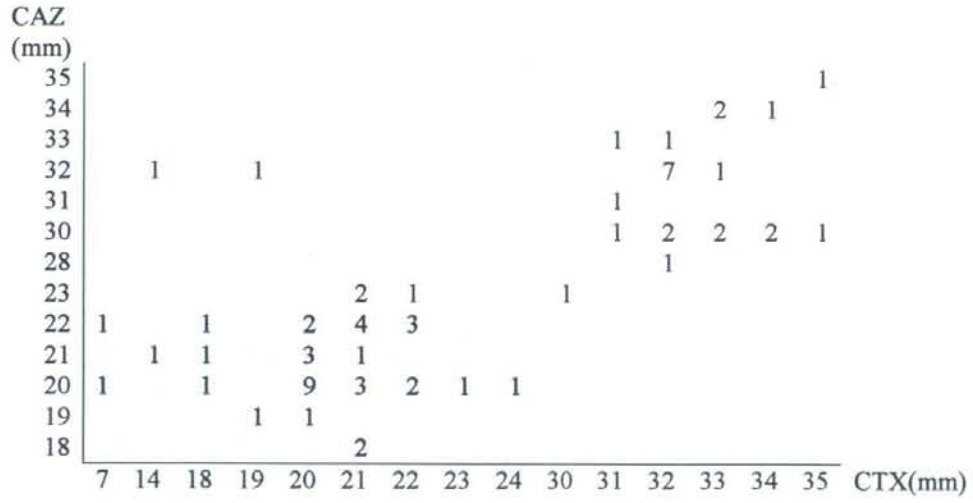


Table 2: ESBL phenotypic confirmatory test with Cefotaxime (Ctx) / Cefotaxime + Clavulanic acid (Ctx+Cla) and Cetazidime (Caz) / Cetazidime + Clavulanic acid (Caz+Cla)

Serotype of *Shigella sonnei*	No. of strains test	Double Disk	
		Ctx / Ctx+Cla	Caz / Caz+Cla
Phase I	25	3	4
Phase II	32	7	2
Phase I&II	12	3	2
All phase	69	13	8

3. PFGE analysis

All 200 strains of Shigellae had digested with XbaI restriction enzyme.

Now, there are in process analysis gel with BioNumberics software.

Discussion:

Preliminary report showed epidemic of shigellosis cause by *Shigella sonnei* various phase. We found increasing of S. sonnei phase II in July 2007 and November 2008. For S. sonnei phase I found in May and July, 2007 and S. sonnei phase I&II found in April and July, 2008.

The phenotypic ESBL of Shigella showed 11.6% and 18.8% when test with Cetazidime and Cefotaxime, respectively.

Reference list:

1. Tenover, F. C., Arbeit, R. D., Goering, R. V., Mickelsen, P. A., Murray, B. E., Persing, D. H. & Swaminathan, B. (1995). Interpreting chromosomal DNA restriction patterns produced by pulsed-field gel electrophoresis: criteria for bacterial strain typing. J Clin Microbiol 33, 2233–2239.

2. NCCLS (2002). Performance Standards for Antimicrobial Susceptibility Testing; 12th Informational Supplement. Document M100-S12. Wayne, PA: National Committee for Clinical Laboratory Standards.

3. Centers for Disease Control and Prevention (2000). Standardized Molecular Subtyping of Foodborne Bacterial Pathogens by Pulsed-Field Gel Electrophoresis: Training Manual. Atlanta: Centers for Disease Control and Prevention.

4. Ewing, W. H. (1986). Edwards and Ewing's Identification of Enterobacteriaceae, 4th ed. New York: Elsevier.

Publication list for this work:

-

Multilocus Variable-Number Tandem Repeat Analysis for Molecular Typing and Phylogenetic Analysis of *Shigella flexneri*

You-Wun Wang[1], Haruo Watanabe[2], Dac Cam Phung[3], Sheng Kai Tung[1], Yeong-Sheng Lee[4], Jun Terajima[2], Shiu-Yun Liang[1], Chien-Shun Chiou, [1,5]*

[1]The Central Region Laboratory, Center for Research and Diagnostics, Centers for Disease Control, Taichung, Taiwan

[2]National Institute for Infectious Diseases, Tokyo, Japan

[3]Division of Enteric Infections, National Institute of Hygiene and Epidemiology, Vietnam

[4]The Sixth Branch Office, Centers for Disease Control, Hualien

[5]Institute of Medicine, Chung Shan Medical University, Taichung, Taiwan

*Corresponding author

Email address:

You-Wun Wang: pleione@cdc.gov.tw

Haruo Watanabe: haruwata@nih.go.jp

Sheng Kai Tung : kbj@cdc.gov.tw

Yeong-Sheng Lee: yslee@cdc.gov.tw

Jun Terajima: terajima@nih.go.jp

Shiu-Yun Liang: sho9004018@yahoo.com.tw

Chien-Shun Chiou: nipmcsc@cdc.gov.tw

ABSTRACT

A total of 36 multilocus variable-number tandem repeat (VNTR) loci were identified by exploring the repeat sequence loci in three *Shigella flexneri*, one *S. dysenteriae* and two *S. sonnei* genomic sequences, and testing with nine isolates of different subserotypes. The 36 VNTR loci were further tested with 242 *S. flexneri* isolates to evaluate the variability in various serotypes. The isolates were also analyzed by pulsed-field gel electrophoresis (PFGE) to compare the discriminatory power and to investigate the usefulness of multilocus VNTR analysis (MLVA) as a tool for phylogenetic analysis of *S. flexneri*. MLVA exhibited higher level of discriminatory power than PFGE and was more powerful in resolving closely related isolates. The phylogenetic groupings using MLVA profiles and PFGE patterns were considerably concordant. The grouping presented two distinct groups for the serogroup 3a/3b and one distinct group for each of the serogroups, 1a/1b/NT, 2a/2b/X/NT, 4a/4b/Y, and 6. The two 3a/3b groups were separated distantly, suggesting that MLVA as well as PFGE was a useful tool in distinguishing *S. flexneri* clonal groups of the same serotype. The present study suggests that MLVA is a powerful tool in discriminating closely related isolates for outbreak investigation and establishing clear phylogenetic relationships among *S. flexneri* strains.

INTRODUCTION

Shigella flexneri, as well as *S. dysenteriae*, *S. boydii*, and *S. sonnei*, are the causative agents of shigellosis, a common acute diarrheal disease in developing countries. The annual number of shigellosis throughout the world was estimated to be 164.7 million, of which 163.2 million were in developing countries with 1.1 million deaths and 1.5 million in industrialized countries (11). *S. flexneri* is the predominant species in the developing countries and ranked as the second in industrialized countries (11, 22). *S. flexneri* comprises eight serotypes, 1, 2, 3, 4, 5, 6, X, and Y, with at least 12 subserotypes, 1a, 1b, 1c, 2a, 2b, 3a, 3b, 4a, 4b, 4c, 5a, and 5b (6, 19), of which 2a is the most prevalent subserotype in the world (11). X and Y serotypes can derive from some subserotypes of serotypes 1, 2, 3, and 4, such as 2b, 4c and 5b for serotype X and 1a, 2a, 4a and 5a for serotype Y, by losing the type factor antigens (3, 6, 19).

A variety of molecular tools have been developed to study genetic relatedness among bacterial strains. In general, molecular markers with low variability can be used to establish phylogenetic relationships among bacterial stains evolved over longer time span and highly variable markers are more useful to resolve closely related strains for the purposes of outbreak investigation and disease surveillance. Among these molecular tools, multilocus sequencing typing (MLST) is a sequence-based method that has been successfully applied to establish phylogenetic structure for some bacterial pathogens, such as Neisseria meningitidis and Streptococcus pneumonia (7). However, to the public health laboratories, MLST is not sufficiently powerful in distinguishing closely related isolates for epidemiological investigation of clusters of infection. Pulsed-field gel electrophoresis (PFGE) is highly discriminatory to most bacterial pathogens and has been taken to be the standard typing method of an international molecular subtyping network—PulseNet, for foodborne disease surveillance (23). Although PFGE has been proven by the PulseNet laboratories to be a powerful tool for the routine subtyping of some foodborne bacterial pathogens to detect clusters of infection, PFGE is, occasionally, not discriminatory enough to distinguish some epidemiologically unrelated *S. sonnei* strains (13). Nevertheless, PFGE is suitable to resolve closely related strains but not an appropriate tool for establishing phylogenetic relationships between bacterial strains that have evolved over a longer time span. Multilocus variable-number tandem repeat (VNTR) analysis (MLVA) has been developed for several bacterial pathogens (14-17, 21); it is based on variation

in the number of tandem repeat sequence units in multiple VNTR loci. VNTR locus can be highly variable or relatively stable with low level of allele diversity. Studies have demonstrated that MLVA based on 4-8 highly variable VNTR loci can exhibit a parallel or higher discriminatory power than PFGE (13, 14, 17) and the method based on combined loci with different variability values can be applied to establish phylogenetic relationships between strains with different evolutionary timescales (4). In the present study, we aim to develop and evaluate a MLVA method for fine typing and phylogenetic analysis of *S. flexneri* strains.

MATERIALS AND METHODS

Bacterial strains. A total of 241 *S. flexneri* isolates were collected in Taiwan between 1995 and 2008, including serotypes and subserotypes 1a (3 isolates), 1b (8), 2a (89), 2b (5), 3a (8), 3b (4), 4a (107), 4b (1), x variant (10), y variant (2), 6 (2) and nontypeable (NT) (2). In addition, one reference strain, *S. flexneri* 2a strain 2457T (ATCC 700930), was purchased from the ATCC company. Among the isolates, 18 were recovered from patients infected during the traveling to Cambodia (4 isolates), China (6), Egypt (1), India (3) and Indonesia (4), and 144 were recovered from eight shigellosis outbreaks. Of the 107 subserotype 4a isolates, 102 were collected from a prolonged shigellosis outbreak occurred in 2001-2007 in a long-stay psychiatric nursing center (12); they should be closely related. In contrast, the serotypes 2a isolates were more divergent even though at least 24 isolates were recovered from five shigellosis outbreaks.

PFGE. The PulseNet PFGE protocol for *S. sonnei* and other enterobacteria was used for PFGE analysis (20) except that 5 U of *Not*I was used instead of *Xba*I for the restriction digestion.

Identification of VNTR loci. Firstly, the three genomes of *S. flexneri* serotype 2a strain 301 (GenBank accession no. AE005674), *S. flexneri* serotype 2a strain 2457T (GenBank accession no. AE014073) and *S. flexneri* serotype 5 strain 8401 (GenBank accession no. CP000266) were explored for potential VNTR loci using VNTRDB computer software developed by Chang et al. (2). As a result, 50 loci were selected for further evaluation. In order to maximize the number of VNTR markers for *S. flexneri*, another search on five genomes of three *Shigella* species, including *S. dysenteriae* strain Sd197 (GenBank accession no. CP000034), *S. flexneri* serotype 2a strains 301and 2457T, *S. sonnei* strains Ss046 (GenBank accession no. CP000038) and 53G (Wellcome Trust Sanger Institute [http://www.sanger.ac.uk]), was performed to explore potential VNTR loci. Consequently, 17 additional loci were selected for further evaluation. In total, 67 loci were examined for the variability with nine *S. flexneri* strains of various subserotypes by PCR amplification. To perform PCR reaction, primers for the 67 loci were designed using the free Primer3 program available on the website at http://frodo.wi.mit.edu. PCR reaction was carried out in a GeneAmp PCR System 9600 (Applied BioSystems). For PCR amplification, crude bacterial DNA was prepared by boiling method as described (13). Each 10-μl PCR mixture contained 1x PCR buffer, 3 mM $MgCl_2$, 0.2 μM of each primer, 200 μM of each deoxyribonucleotide, 1.0 unit of the recombinant SuperNew Taq DNA polymerase (Jier Sheng

Company, Taipei, Taiwan), and 1 μl of DNA template. The PCR reaction was performed with a denaturing step at 94°C for 5 min, followed by 30 cycles of amplification at 94°C for 30 s, 55°C for 45 s, and 72°C for 45 s, and extended at 72°C for 5 min. Loci with varied sizes among the nine strains were considered to be VNTR and the markers were further used to analyze a total of 242 *S. flexneri* isolates.

MLVA. The primer sets for PCR amplification of 36 VNTR loci are listed in Table S1 in the supplemental file 1 (see the supplemental material). The forward primer of each primer set was labeled at its 5' end with an ABI-compatible dye, 6-FAM, NED, VIC or PET by Applied BioSystems (Foster City, CA, USA). Nine multiplex PCR combinations were set for the analysis (Table S1). PCR reaction was performed as described above except that dye-labeled primers were used. The PCR products were analyzed by capillary electrophoresis on an ABI Prism 3130 Genetic Analyzer with GeneScan 500 LIZ Size Standard (cat # 4322682; Applied BioSystems) as described (13).

Data analysis. PFGE images were analyzed using the fingerprint analysis software BioNumerics version 4.5 (Applied Maths; Kortrijk, Belgium). A PFGE genotype was defined as a PFGE pattern with one or more DNA bands different from the others. A dendrogram constructed using the *Not*I-digested PFGE patterns was generated by the UPGMA algorithm with the Dice-predicted similarity value of two patterns. The number of repeat units for each allele was converted from the length of amplicon and was saved as "Character Type" in BioNumerics database and subjected to cluster analysis using the Minimum Spanning Tree algorithm provided in the BioNumerics software. Alleles, which contained not perfect repeat unit(s) due to deletion, insertion or composite different repeat units, were designed by the lengths (in bp) of amplicons. To compare the discriminatory power of PFGE and MLVA with various combinations of VNTR loci, Simpson's index of diversity (*D*) and 95% confidence intervals (CI) were calculated according to the formulas as the described (9, 10). The polymorphism of each locus was represented by Nei's diversity index, calculated as $1-\sum(\text{allelic frequency})^2$.

RESULTS

VNTR loci. In total, 36 VNTR loci were identified after testing nine *S. flexneri* strains of various subserotypes. The characteristics of the 36 VNTR loci are listed in Table S2 in the supplemental file 2 (see the supplemental material). The 36 VNTR markers were analyzed on 242 *S. flexneri* isolates with various serotypes and subserotypes. Of the 36 loci, 27 were detected in all the isolates tested, while 3 were detected only in around 50% of the isolates (Table 1). The low typeability rate for the three loci, SF1, SF6 and SF26, was mainly attributed to the absence of amplicon from the 4a isolates. The serotype groups, 1a/1b/NT, 2a/2b/X/NT, 3a/3b, 4a/4b/Y and 6, were detected with 9, 26, 20, 14 and 3 variable loci, respectively. Three loci were variable among 4 of the 5 serotype groups, 11 loci variable among 3 serotype groups, 8 loci variable among 2 serotype groups, and 11 loci variable in only one serotype group. Three loci (SF20, SF28 and SF35) were invariable to isolates within each of the five serotype groups.

Phylogenetic relationships constructed using PFGE patterns. In total, 92 PFGE-NotI patterns were identified in the 242 isolates. The dendrogram established using the PFGE patterns was shown on Figure 1. Based on serotypes and the genetic relatedness among the isolates, six groups, 4a/Y, 3a/3b, 1a/1b, 2a/2b/X/NT, 6 and 3a, were designated. Serotype 3 (3a/3b) isolates were distributed in two distinct groups. Three isolates (Y, 4b and NT) belonged to singletons. One NT isolate was located within the cluster for the 2a/2b/X isolates. The 4b isolate was distant from the 4a/Y group. A 2b isolate was recovered with X isolates from a shigellosis outbreak; it shared an indistinguishable PFGE pattern with an X isolate.

Phylogenetic relationships constructed using MLVA profiles. In total, 121 MLVA types were identified in the 242 isolates. A phylogenetic tree was constructed using the MLVA profiles by minimum spanning tree algorithm. In this study, nodes (genotypes) differing at seven or fewer loci were considered to be the same group. Based on the criterion, the phylogenetic tree presented six groups (2a/2b/X/NT, 1a/1b/NT, 6, 4a/4b/Y, 3a/3b, and 3a) and 2 singletons for 2b and 3a strains (Figure 2). Thus, the grouping using MLVA profiles was considerably concordant with that using PFGE patterns. The congruence between the two methods was 85.2%. The 2a/2b/X/NT included 90 2a, 4 2b, 10 X and 1 NT isolates that were more divergent. The 2b isolate recovered with X isolates from a common shigellosis outbreak had an identical MLVA profile with one X isolate. Strains in this 2b/X subgroup were close to another 2b strain. The NT

strain differed at only 3 loci from one 2a strain. The 1a/1b/NT group included 3 1a, 8 1b and 1 NT isolates. One 1a strain differed only at one locus from one 1b strain. The 4a/4b/Y group comprised 107 4a isolates, two Y isolates and one 4b isolate. The 4a and Y isolates were closely related. One Y isolate shared a common MLVA type with 10 4a isolates. The isolates of serotype 3 were separated in two groups, 3a/3b and 3a, and a singleton; the two groups had a distance of 14 loci in between. In the 3a/3b group, two isolates with different subserotypes (3a and 3b) differed only at two loci.

Discriminatory power of PFGE and MLVA. The MLVA analysis based on the 36 VNTR markers resulted in 121 MLVA types in the 242 *S. flexneri* isolates, with a discriminatory power index (D) value of 0.9736 (CI, 0.9632-0.9840) (Table 2). Using the PFGE method, the isolates were discriminated into 92 PFGE types, with a D value of 0.9399 (CI, 0.9220-0.9578). The discriminatory power of MLVA was significantly higher than PFGE. For the 90 2a isolates, which were more divergent, the discriminatory power for the MLVA analyses based on four (SF3, SF4, SF6, SF7), eight (SF3, SF4, SF6, SF7, SF8, SF9, SF11, SF27) most variable VNTR loci, and all the 36 loci, was higher than PFGE, though not statistically significantly. For the 103 4a/Y isolates, which were recovered from a common prolonged shigellosis outbreak, MLVA analysis based on the four most variable loci (SF2, SF3, SF17 and SF23) was more discriminatory significantly than PFGE. For the remaining 49 most divergent isolates, both MLVA and PFGE exhibited considerably high discriminatory power.

Subtyping of isolates from outbreaks. A total of 144 isolates recovered from eight shigellosis outbreaks were used to evaluate the usefulness of MLVA in discriminating strains for disease outbreak investigation. Outbreak A, a prolonged outbreak occurring from 2001-2007 in a long-stay psychiatric nursing center, was recovered with isolates of subserotype 4a and serotype Y (Table 3). Isolates of subserotype 2b and serotype X were recovered from outbreak K. Four outbreaks were detected with 2 or more PFGE and MLVA types. The isolates from outbreaks H and I shared an indistinguishable PFGE pattern.

MLVA analysis revealed that four outbreaks had 2 or more MLVA types (Figure 3). The isolates from outbreaks H and I, sharing an indistinguishable PFGE pattern, had different MLVA types. The isolates from the four outbreaks detected with only one MLVA types were recovered in a short time period. The 103 isolates for outbreak A, recovered in a long time period (2001-

2007), were relatively divergent. Outbreaks B and K occurred, respectively, in a military camp and long-stay psychiatric nursing center and lasted only 1-2 weeks, but the isolates had been considerably diversified. The four isolates for outbreak E were recovered in a period of 6 months; the first three isolates had an identical MLVA type but the latest isolate differed at 2 loci from the first three. There was no predominant type for the two outbreaks and a difference of three loci was observed between two genotypes for the outbreak K.

DISCUSSION

VNTR markers usually have a range of degrees of variability or allele diversity that can be used to determine various levels of genetic relatedness between bacterial strains evolved over different evolutionary time spans. In our previous studies, we identified 26 VNTR markers for *S. sonnei* and evaluate the usefulness of the MLVA method as a tool for fine typing and phylogenetic analysis of bacterial strains (4, 13). The studies indicate that MLVA based on four or eight highly variable loci is sufficient to resolve closely related isolates for outbreak investigation and disease surveillance (13); whereas, combined loci with lower variability values are suitable to establish clear phylogenetic relationships between strains over longer evolutionary time (4). Theoretically, the more loci are used, the higher discriminatory power can be achieved and more distinctive phylogenetic relationships among bacterial strains can be established. Unlike *S. sonnei* having only one serotype, *S. flexneri* has eight serotypes with at least 12 subserotypes (6, 19). Since VNTR markers may be serotype-specific, it is necessary to explore more loci at this moment of study.

In order to identify more VNTR markers as possible, we explored TR loci that are potentially variable from the released genomic sequences of *S. flexneri* and other *Shigella* species using the VNTRDB computer program (2). After testing on nine strains with different serotypes, we identified 36 loci that were variable among the serotype groups. The evaluation with 242 *S. flexneri* isolates of various serotypes and subserotypes suggests that variability of some of the 36 loci could be serotype-specific. For instance, 7 loci are invariable to the 2a/2b/X/NT group but are variable to other serotype groups (Table 1). However, this observation could be biased because of the limitation of isolates tested. For the 2a/2b/X/NT and 3a/3b groups, the isolates were relatively divergent and more variable loci were observed. For the 1a/1b/NT, 4a/4b/Y and 6 groups, the number of isolates was too small or most isolates were closely related; only a small portion of the 36 loci was observed to be variable in the groups. It is expected that more variable loci could be found for the serotypes when more divergent strains are analyzed.

Many studies have indicated that MLVA is more powerful than PFGE in discriminating bacterial strains (1, 13, 17); it can complement or replace PFGE as a routine genotyping tool for outbreak investigation and disease surveillance. In the present study we also show that MLVA is more discriminatory than PFGE. The difference of discriminatory power for MLVA and PFGE

is correlated with genetic relatedness among isolates analyzed. The discriminatory power of MLVA based on 36 loci for the 242 isolates is significantly higher than PFGE (CI, 0.9632-0.9840 vs. 0.9220-0.9578) (Table 2). To the panel of divergent 2a isolates, MLVA analysis using the four and eight most variable loci discriminated the 90 isolates into more genotypes than PFGE but the discriminatory power was not significantly greater (Table 2). To the panel of the closely related 4a isolates, in contrast, the discriminatory power for MLVA based on the four most variable loci was significantly higher than that for PFGE. To the panel of the 49 isolates with great divergence, both methods exhibited a high level of resolution. Similarly, as that observed in the study on *S. sonnei* (4), MLVA is more discriminatory than PFGE and is more powerful in discriminating closely related isolates. However, the disadvantage to MLVA is that the highly variable loci could not be common to all the *S. flexneri* serotypes. To find common four or eight loci to replace PFGE for routine typing of all the *S. flexneri* serotypes, the 36 loci need to be evaluated with more isolates for each serotype.

VNTR markers have different degrees of variability; they have been applied to investigate different level of genetic relatedness among bacterial strains of *Yersinia pestis* (18), *Neisseria meningitidis* (14, 21) and *S. sonnei* (4). In the present study, we used the MLVA profiles to establish the phylogenetic relationships among the 242 isolates. When defining a clonal group that includes genotypes differing at seven or fewer loci from the closest one, the grouping based on the MLVA profiles is highly in agreement with that based on the PFGE patterns (Figures 1 and 2). The phylogenetic analysis based on MLVA profiles establishes the genetic relationships among strains with different serotype and subserotypes. The genetic relatedness suggests that the two serotype Y isolates and the 10 serotype X isolates could be derived from serotype 4 and serotype 2 strains, respectively; whereas, the two NT isolates are associated with serotype 1 and serotype 2 groups. The isolates of serotype 3 (subserotypes 3a and 3b) are distributed in two distinct clonal groups, suggesting that MLVA is a useful tool not only in distinguishing strains among different serotypes but also establishing clear clonal patterns for strains with the same serotype. The phylogenetic analyses based on PFGE and MLVA typing also suggest that strains of a serotype with different subserotypes can be very closely-related.

Recently, Gorgé et al. (8) identified 15 VNTR loci by exploring on five genomic sequences of four *Shigella* species for typing of *Shigella* spp. The 15 loci were evaluated on 89 isolates of *Escherichia coli* and *Shigella* species including 19 *S. flexneri* isolates of eight serotypes; 13 of

the 15 loci were variable to the *S. flexneri* isolates. Of the 15 loci, 13 are included in the 36 loci identified in the present study. Although the MLVA based on the 15 loci could effectively discriminate the isolates of the four *Shigella* species and pathogenic *E. coli*, grouping based on the MLVA profiles was not able to clearly separate the four *Shigella* species into distinct groups. Although it may not be able to develop an MLVA method for establishing phylogenetic relationships between different *Shigella* species, it is possible to develop an MLVA method for establishing clear phylogenetic relationships among strains within a *Shigella* species. This has been succeeded on *S. sonnei* (4). In the present study, we identify as many as 36 loci that are variable to *S. flexneri* serotypes. Our study indicates that the MLVA with 36 loci is sufficiently powerful in establishing clear phylogenetic relationships between serotypes and discerning clonal groups within a serotype.

A total of 144 isolates from eight outbreaks were analyzed using MLVA and PFGE to compare the usefulness of both genotyping methods for outbreak investigation. The result of evaluation indicates that MLVA is more powerful than PFGE in resolving closely related isolates. Theoretically, isolates collected from a prolonged outbreak, such as outbreak A, could have many MLVA types, while isolates from an outbreak lasting for a short time period should have little change. However, outbreaks B and K are quite unusual. The outbreaks started and ended in a few days but the isolates were quite divergent. Several PFGE and MLVA genotypes were detected among the isolates for each of the outbreaks and had no predominant type. The strains could be genetically hypervariable; mutants could emerge very quickly in response to host immunity. But, this phenomenon still remains to be explained.

The variability of the 36 loci could be serotype-specific; loci variable to a serotype could be invariable to other serotypes. *S. flexneri* comprises eight serotypes of which X and Y serotypes can derive from some subserotypes of serotypes 1, 2, 3, and 4, such as 2b, 4c and 5b for serotype X and 1a, 2a, 4a and 5a for serotype Y, by losing the type factor antigens (3, 6, 19). The study conducted by Choi et al. (5) by MLST analysis indicates that the *S. flexneri* serotypes 1-5, X and Y are evolutionarily more related than serotype 6. Thus, it could be possible to find a set of common VNTR loci for routine typing of *S. flexneri* serotypes 1-5, X and Y.

In conclusion, we develop a MLVA method comprising 36 VNTR loci for fine typing and phylogenetic analysis of *S. flexneri*. MLVA exhibits a better resolution than PFGE to closely

related isolates so it is a better molecular tool in distinguishing isolates for outbreak investigation. MLVA is also a useful tool for phylogenetic analysis of *S. flexneri*; it separates strains of various serotypes into distinct groups. MLVA separates the strains of serotype 3 into two distinct groups, indicating that it is able to discern clonal groups for a serotype. Thus, the MLVA method is useful tool for resolving closely related isolates and for establishing phylogenetic relationships between *S. flexneri* strains. However, the 36 are not variable to all the serotypes as observed on the isolates tested in the present study. Analysis of the 36 loci with more strains of various serotypes is needed to determine the common loci for serotypes 1-5, X and Y.

ACKNOWLEDGMENTS

This work was supported by a grant (DOH 96-DC-2025) of the Centers for Disease Control, Department of Health, Taiwan, and by a grant (H17-Shinkou-Ippan-019) of Ministry of Health, Labor and Welfare, Japan.

REFERENCES

1. **Boxrud, D., K. Pederson-Gulrud, J. Wotton, C. Medus, E. Lyszkowicz, J. Besser, and J. M. Bartkus.** 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. J Clin Microbiol **45:**536-43.
2. **Chang, C. H., Y. C. Chang, A. Underwood, C. S. Chiou, and C. Y. Kao.** 2007. VNTRDB: a bacterial variable number tandem repeat locus database. Nucleic Acids Res **35:**D416-21.
3. **Chen, J. H., W. B. Hsu, C. S. Chiou, and C. M. Chen.** 2003. Conversion of *Shigella flexneri* serotype 2a to serotype Y in a shigellosis patient due to a single amino acid substitution in the protein product of the bacterial glucosyltransferase gtrII gene. FEMS Microbiol Lett **224:**277-83.
4. **Chiou, C. S., H. Watanabe, Y. W. Wang, W. L. Wang, J. Terajima, K. L. Thong, P. D. Cam, and S. K. Tung.** 2009. Utility of Multilocus Variable-Number Tandem Repeat Analysis as a Molecular Tool for Phylogenetic Analysis of *Shigella sonnei*. J Clin Microbiol **(Accepted)**.
5. **Choi, S. Y., Y. S. Jeon, J. H. Lee, B. Choi, S. H. Moon, L. von Seidlein, J. D. Clemens, G. Dougan, J. Wain, J. Yu, J. C. Lee, S. Y. Seol, B. K. Lee, J. H. Song, M. Song, C. Czerkinsky, J. Chun, and D. W. Kim.** 2007. Multilocus sequence typing analysis of *Shigella flexneri* isolates collected in Asian countries. J Med Microbiol **56:**1460-6.
6. **El-Gendy, A., N. El-Ghorab, E. M. Lane, R. A. Elyazeed, N. I. Carlin, M. M. Mitry, B. A. Kay, S. J. Savarino, and L. F. Peruski, Jr.** 1999. Identification of *Shigella flexneri* subserotype 1c in rural Egypt. J Clin Microbiol **37:**873-4.
7. **Feil, E. J., B. C. Li, D. M. Aanensen, W. P. Hanage, and B. G. Spratt.** 2004. eBURST: inferring patterns of evolutionary descent among clusters of related bacterial genotypes from multilocus sequence typing data. J. Bacteriol. **186:**1518-30.
8. **Gorge, O., S. Lopez, V. Hilaire, O. Lisanti, V. Ramisse, and G. Vergnaud.** 2008. Selection and validation of a multilocus variable-number tandem-repeat analysis panel for typing *Shigella* spp. J Clin Microbiol **46:**1026-36.
9. **Grundmann, H., S. Hori, and G. Tanner.** 2001. Determining confidence intervals when measuring genetic diversity and the discriminatory abilities of typing methods for