Dynamic time warping assessment and sensitive high resolution melting analysis for subtyping *Salmonella* isolates from the Northern Thailand

Nuttachat Wisittipanit [Equal first author, 1], **Chaiwat Pulsrikarn** [2], **Saranya Wutthiosot** [3], **Sitthisak Pinmongkhonkul** [4], **Kritchai Pooncharean** [Corresp. Equal first author, 5]

[1] Department of Material Engineering, School of Science, Mae Fah Luang University, Chiang Rai, Thailand

[2] Department of Medical Sciences, WHO National Salmonella and Shigella Center, National Institute of Health, Ministry of Public Health, Bangkok, Bangkok, Thailand

[3] Veterinary Research and Development Center (Upper Northern Region), Lampang, Thailand

[4] Division of Biology, School of Science, University of Phayao, Phayao, Thailand

[5] Division of Biochemistry, School of Medical Sciences, University of Phayao, Phayao, □□□□□□□□□□□□□, Thailand

Corresponding Author: Kritchai Pooncharean
Email address: kritchai.po@up.ac.th

Background: Nontyphoidal *Salmonella* spp. transmitted through various routes are a major concern of food poisoning due to the consumption of contaminated food.

Objective: To establish a molecular-based protocol for simple and rapid subtyping of Salmonella isolates from various sources.

Materials and methods: Sensitive High-Resolution Melting-curve analysis (S-HRMa) and Dynamic Time Warping assessment (DTW) were applied for serotyping forty Salmonella spp. isolates from various origins and locations in seven provinces in the north of Thailand; the results were compared to those from conventional serotyping and ERIC- PCR.

Results: HRM serotyping of forty *Salmonella* spp. initially produced fourteen melting-curves with two predominant clusters: C1 (n=18) and C2 (n=9). Applying S-HRMa and serogroups generated twenty-five sensitive clusters. Conventional serotyping revealed that cluster C1 and C2 comprised of six different Salmonella serotypes with S. Weltevradent (n=14) as the predominant one. The S-HRMa also suggested the possible subtyping in some serotypes. In addition, DTW was performed to cluster those forty *Salmonella* spp. into twenty-eight clusters, assigned into different four clades corresponding to S-HRMa. The two clustering methods indicated that the S. Weltevreden was the predominant subtype (DTW4-S1, n=6). Three ERIC clusters at 92% similarity index also corresponded to the results of those two clustering methods. With important and related epidemiological data, S. Derby and S. Monophasic were suggested to be related to the slaughterhouse and swine. In this study, the ERIC cluster 10 comprising two Salmonella isolates of S. Weltevraden suggested the transmission route was likely to be farm-to-farm in the same province.

Conclusions: The DTW assessment and S-HRMa effectively increased the discrimatory power of clustering to the same level as that of ERIC - PCR and were a simple and rapid protocol to perform *Salmonella* subtyping for the epidemiological research.

Dynamic time warping assessment and Sensitive High Resolution Melting analysis for subtyping Salmonella isolates from the Northern Thailand

Nuttachat Wisitipanit[1], Chaiwat Pulsrikarn[2], Saranya Wutthiosot[3], Sitthisak Pinmongkhonkul[4] Kritchai Pooncharean[5]

[1] Department of Material Engineering, School of Science, Mae Fah Luang University, 333 Moo.1, Ta Sud, Muang, Chiang Rai 57100, Thailand
[2] Department of Medical Sciences, WHO National Salmonella and Shigella Center, National Institute of Health, Ministry of Public Health, Tiwanond Road, Amphur Muang, Nonthaburi, 11000, Thailand
[3] Veterinary Research and Development Center (Upper Northern Region) Lampang Hang Chat, Lampang, Thailand
[4] Division of Biology, School of Science, University of Phayao, 19 Moo 2, Tambon Maeka, Amphur Muang, Phayao, 56000, Thailand.
[5] Division of Biochemistry, School of Medical Sciences, University of Phayao, 19 Moo 2, Tambon Maeka, Amphur Muang, Phayao, 56000, Thailand

Corresponding Author:
Kritchai Pooncharean[5]
Division of Biochemistry, School of Medical Sciences, University of Phayao, 19 Moo 2, Tambon Maeka, Amphur Muang, Phayao, 56000, Thailand
Email address: kof_of@hotmail.com

Abstract

Background:

Nontyphoidal *Salmonella spp*. transmitted through various routes are a major concern of food poisoning due to the consumption of contaminated food.

Objective:

To establish a molecular-based protocol for simple and rapid subtyping of *Salmonella* isolates from various sources.

Materials and methods:

Sensitive High-Resolution Melting-curve analysis (S-HRMa) and Dynamic Time Warping assessment (DTW) were applied for serotyping forty *Salmonella* spp. isolates from various origins and locations in seven provinces in the north of Thailand; the results were compared to those from conventional serotyping and ERIC- PCR.

Results:

HRM serotyping of 40 *Salmonella* spp. initially produced 14 melting-curves with 2 predominant clusters: C1 (n=18) and C2 (n=9). Applying S-HRMa and serogroups generated 25 sensitive clusters. Conventional serotyping revealed that cluster C1 and C2 comprised of 6 different *Salmonella* serotypes with *S.* Weltevradent (n=14) as the predominant one. The S-HRMa also suggested the possible subtyping in some serotypes. In addition, DTW was performed to cluster those 40 *Salmonella* spp. into 28 clusters, assigned into different 4 clades corresponding to S-HRMa. The two clustering methods indicated that the S. Weltevreden was the predominant subtype (DTW4-S1, n=6). Three ERIC clusters at 92% similarity index also corresponded to the results of those two clustering methods. With important and related epidemiological data, *S.* Derby and *S.* Monophasic were suggested to be related to the slaughterhouse and swine. In this study, the ERIC cluster 10 comprising 2 *Salmonella* isolates of *S.* Weltevraden suggested the transmission route was likely to be farm-to-farm in the same province.

Conclusions:

The DTW assessment and S-HRMa effectively increased the discriminatory power of clustering to the same level as that of ERIC - PCR and were a simple and rapid protocol to perform *Salmonella* subtyping for the epidemiological research.

Introduction

Nontyphoidal *Salmonella* (NTS) is the major pathogen causing gastroenteritis to victims mostly young children (Hohmann, 2001). Commonly, the victims were usually infected by their exposure to *Salmonella*-contaminated food or animals together with non-hygienic behaviour. Various Nontyphoidal *Salmonella* spp. were unique due to different cellular appearances of somatic antigens (O antigens) combined with flagella antigens (H antigens) (Grimont and Weill, 2007). To initially determine *Salmonella* serogroups, only unique O antigens were diagnosed, obtaining more than 50 distinct groups. For the complete serotyping, the determination of phase 1 and 2 H antigens was executed to derive *Salmonella* serotype according to the complete standard *Salmonella* typing system (Kauffmann-White scheme).

The incidence of salmonellosis was documented through several epidemiological research in Thailand (Hendriksen et al., 2009; Pulsrikarn et al., 2013). The epidemiological survey showed that the prevalence of *Salmonella* serotypes significantly varied according to different sampling locations, animal hosts and sources of samples, causing regional health problems (Jackson et al., 2013). *Salmonella* contamination possibly occurred at any responsible site through the entire food chain (Forshell and Wierup, 2006). Thus, the *Salmonella* surveillance and monitoring systems were essentially established elsewhere by independent laboratories and funded agencies from government to provide enough informative data to track the transmission route of *Salmonella* isolates causing health problems (Herikstad, Motarjemi and Tauxe, 2002). Some *Salmonella* serotypes were specifically linked to their preferred environmental niche such as *Salmonella* Typhimurium and *Rissen* with swine (Arguello et al., 2012) and *S. Kentucky* with poultry (Crespo et al., 2016). On the other hand, some serotypes can adapt to various environmental niche such as *Salmonella Monophasic* which is classified as the most ubiquitous in several zoonotic reservoirs, responsible for the majority of human and animal infections. (Davies et al., 2018).

The conventional typing of *Salmonella* serotypes is usually performed as the standard protocol providing sufficient data to several epidemiological researches. The assay is based on culture method using different selective media to identify *Salmonella* spp. and then the serotyping is done by serological agglutination based on a different combination of O and H antigens. However, these methods need qualified personnel to effectively complete the complicated protocols. The molecular typing for *Salmonella* serovar identification is based on the sequence polymorphism of *rfb* locus and flagellar alleles as gene targets (Cardona-Castro et al., 2009). Other molecular modifications such as High-resolution melting-curve analysis, coupled with the multiplex PCR can be used for detection of polymorphisms of 16S rDNA, *fljB*, *gyrB* and *ycfQ* (Athamanolap et al., 2014; Zeinzinger et al., 2012).

Together with the above mentioned techniques for serotyping, High-resolution melting (HRM) analysis and dynamic time warping (DTW) are rapid molecular techniques used in species discrimination. High-resolution melting (HRM) analysis was established as the effective method using single-nucleotide polymorphisms for *Salmonella* typing. HRM data was effectively assessed through the analysis of curve differences generated by subtraction of a reference curve from the unknown curves after normalization with the temperature shifting. Visual assessment of the curve differences was performed to discriminate many prevalent *Salmonella* serotypes despite its discrepancy in some serotypes. Previously, the sequentially ordered data of a melt curve was analyzed with dynamic time warping (DTW) to produce DTW distances used to correctly cluster 51 strains of 18 fungal species. In addition, the DTW analysis was applied to rapidly and correctly identify 243 clinical fungal isolates (Lu et al., 2017).

The objective of this study was to use the rapid molecular techniques sufficient to suggest the genetic relatedness of *Salmonella* isolates with various epidemiological data. The DTW assessment and modified HRM analysis were first introduced to perform the post analysis of the HRM serotyping to characterize 40 *Salmonella* isolates collected from seven provinces in the north of Thailand during February 2018 to September 2019.

Materials & Methods

The sample collection and Salmonella isolation and identification

Samples in this study were randomly collected from various types of specimens, animals and sources in seven provinces in the Northern Thailand during February 2018 to September 2019 under the approval by the IACUC of University of Phayao (project number 62-02-04-001). This study had a total of 718 samples. The types of specimens were randomly selected within six categories; organs, intestinal content, stool, cecal content, carcass and meat; the animals consisted of chicken, goat, swine, cow, and rat; and the sample sources were randomly selected within three categories: house, farm, slaughter house. The specimen collections were performed at the collection sites under sterile technique i.e. all samples were collected in sterile plastic bag, kept in sample holder filled with ice during transportation and carefully kept at 4 °C until further isolation and identification process at the Veterinary Research and Development Center (Upper Northern Region), Lampang, Thailand.

The samples were then transferred to buffered peptone water (BPW; Oxoid, Hampshire, UK) with overnight incubation at 37 °C, and later transferred to both TT broth and RVS broth (Oxoid, Hampshire, UK) with overnight incubation at 37 C° and 42 C° respectively. Both overnight TT and RVS were separately plated on XLD agar and incubated overnight at 37 C°. Black centre dot colonies, referred to suspected Salmonella colonies, were picked to perform 2 biochemical tests; triple sugar iron (TSI) slant, and motility indole lysine agar (MIL) (Biomedia, Nonthaburi, Thailand). The determination of serotypes was performed on the positive colonies by

biochemical and serum agglutination tests at the WHO national Salmonella and Shigella Center, the National Institute of Health, Ministry of Public Health, Nonthaburi Province, Thailand.

Determination of Salmonella serotypes by HRM serotyping

Salmonella serotyping using the multiplex PCR coupled with HRM analysis was initially performed with the genomic extracted DNA from *Salmonella* isolates, as previously described by McNerney et al (McNerney et al., 2017). Briefly, 1 ml of overnight culture was centrifuged, and the pellet was washed twice with 400 µl of TE buffer (10 mM Tris HCl, pH 8.0, 1 mM EDTA) and resuspended in 400 µl of TE buffer. The resuspended solution was heated at 80°C for 20 minutes. After cooling at room temperature, 50 µL lysozyme (10 mg/mL) was added to the solution with occasionally shaking at 37°C for one hour and then 75 µL of 10% SDS/proteinase K solution was added with vigorous vertex and incubated at 65 °C for10 minutes. A 100 µL of 5 M NaCl and 100 µL of prewarmed (65°C) CTAB/NaCl solution were added, causing white precipitate; then the solution was further incubated for 10 minutes at 65°C. A 750 µl of chloroform/isoamyl alcohol (24:1) was added and then centrifuged for 5 minutes at 13,000 rpm at 4°C. The upper aqueous solution was collected to a fresh microcentrifuge tube whereby ethanol precipitation was performed. Finally, the pellet was resuspended in a tube with 50 µL water and the DNA solution was kept at -20°C until used.

Multiplex PCR coupled with HRM analysis using a combination of primers to amplify *flj*B (170 bps), *gyr*B (171 bps) and *ycf*Q (241 bps) (Table 1) were conducted in a BIORAD CFX96™ Real-Time System (Bio-Rad, Hercules, CA, USA). First, multiplex PCR mixture was prepared, containing 1 µL of DNA, 0.1 µM of *gyr*B, 0.075 µM of *flj*B and *ycf*Q primer pairs as well as 2 µL of HOT FIREPol EvaGreen: no ROX Mix (Solis Biodye, Tartu, Estonia) adjusted to 10 µL with water. Thermocycling conditions were as follows: 95°C for 15 minutes, followed by 45 cycles of 95°C for 10 seconds, 60°C for 10 seconds and 72°C for 20 seconds. To perform the HRM analysis, PCR mixtures were heated at 95°C for 1 minute and cooled to 40°C for 1 minute and then continuously heated from 70°C to 95°C, rising at 0.2°C/second, with 25 acquisitions per degree Celsius. HRM profile was generated using the Precision Melt Analysis software version 1.2 with a sensitivity setting at 0.30, a temperature shift at threshold 5, a pre-melt normalization range from 80.4 °C to 81.1°C, and a post-melt normalization range from 89.2 °C to 89.6 °C. After normalizing and temperature shifting of the melting curves, difference plots were generated using *S. Barille* as reference.

The sensitive High Resolution Melting analysis (S-HRMa) for Salmonella subtyping.

This analysis was performed by a modified protocol for clustering several closely related HRM melting curves using the Precision Melt Analysis software version 1.2. All closely related HRM

melting curves using *S. Barille* as reference were considered in a single panel. In order to increase variation in the graphs, a distinct HRM melting curve from all members in the cluster was selected as the new reference. Then, the new HRM melting curves were created which produced some peaks with high variation of peaks' patterns. The melting-curve analysis specified the temperature range covering each peak. The difference of peaks' patterns was observed visually and each distinct pattern was assigned the alphabet A to L. The new HRM melting curves, named the sensitive clusters, were assigned to the unique combination of all peaks' patterns in this analysis.

The hierarchical clustering of the normalized melting curves from HRM analysis using DTW assessment

Normalized melt curves generated from HRM were used to construct a dendrogram of hierarchical clustering, using DTW as a distance measure (Lu et al., 2017); all the dendrogram construction steps were performed in Python (Rossum, 1995). First, a smooth spline approximation was determined from each normalized melt curve using cubic splines of the splrep function in scipy module; then, a rate curve was calculated from the negative first derivative of the resulting spline using the splev function. The obtained curve was z-normalized using the zscore function. Then, these z-normalized curves were used to calculate DTW distances where only the region between 80 and 94 degrees Celsius was accounted for. The hierarchical clustering based on neighbor-joining method, using a distance matrix of the DTW distances, was performed by the linkage function. Finally, the dendrogram presenting the clustering was created by the dendrogram function.

*Molecular analysis of bla*TEM *and flo*R

Amplifications of two antibiotic determinants (*bla*TEM and *flo*R) were performed by Realtime multiplex PCR using the primers (IDT, Singapore) listed in Table 1. The reaction mixture (10 µL) contained 1 µL of DNA, primer set at concentration listed in Table 1 and 2 µL of HOT FIREPol Blend Master Mix Plus 10 mM $MgCl_2$ (Solis Biodye). In multiplex PCR 1 and 2, thermocycling was as follows: 95°C for 15 minutes; 40 cycles of 95°C for 15 seconds; 60°C for 45 seconds and 72°C for 1 minute; and a final step at 72°C for 7 minutes. The HRM analysis was performed by heating the PCR mixture at 95°C for 1 minute, cooled to 40°C for 1 minute and then continuously heated from 60°C to 95°C, rising at 0.2°C/s, with 25 acquisitions per degree Celsius. A melt curve plot was created between the negative derivative of fluorescence versus temperature. The melting temperature (T_m) values for each PCR fragments was observed as a distinct peak at 81.9°C and 83.2°C from *bla*TEM and *flo*R, respectively.

The genomic DNA fingerprinting assays using ERIC-PCR

The DNA extraction of *Salmonella* isolates followed the protocol of McNerney et al (Mcnerney et al., 2017). To perform PCR reaction, the primers of repetitive element fingerprinting-based PCR or ERIC-PCR were listed in Table 1. The 20 µL reaction mixture contained 0.2 µL of DNA, primer concentration (25 uM) and 2 µL of HOT FIREPol Blend Master Mix Plus 10 mM $MgCl_2$ (Solis Biodye). The conditions were as follows; 95°C for 15 minutes; 40 cycles of 95°C for 60 seconds, 54°C for 2 minutes and 72°C for 4 minutes; and a final step at 72°C for 10 minutes. Amplicons were separated using 4% agarose gel electrophoresis with 1X TBE at constant DC voltage of 60V for 3 hours at room temperature. DNA bands were stained with RedSafe dye (INiRON, Washington, United States) and visualized under Molecular Imager (Gel DOC™ XR+, Bio-RAD). The image of the gels were exported as JPEG images at 300 dpi resolution in which the Image Lab™ software was used for further analysis.

The analysis of DNA fingerprint patterns and phylogeny tree construction

Analysis of the genetic fingerprint patterns and construction of phylogeny tree was performed by a temporary BioNumeric evaluation license from Applied Maths, using curve based algorithm (pearson correlation) to create similarity scale and unweight pair group method with arithmetic mean (UPGMA) for cluster analysis.

The determination of Discrimatory index from different clustering protocols

The Discrimatory Power (D) was assessed to obtain the average probability that a clustering method will assign a different type from two unrelated strains randomly sampled in the given population of *Salmonella* isolates. The discrimatory power (D) of a clustering method was evaluated using Simpson's diversity index described by Hunter and Gaston (Hunter and Gaston, 1998) according to the equation:

$$D = 1 - \frac{1}{N(N-1)} \sum_{j=1}^{S} x_j(x_j - 1) \qquad (1)$$

where D is the index of discrimatory power, N the number of unrelated strains tested, S the number of different types, and x_j the number of strains belonging to the j^{th} type. D value is in a range of 0 (identical types) to 1 suggesting that the typing method was capable of distinguishing each member of a population from all other members of that population.

Results

The HRM serotyping of the 40 Salmonella isolates from various sources indicated 12 unique clusters together with two groups of closely related clusters C1 and C2

In the preliminary project to develop the HRM serotyping protocol, all 40 *Salmonella* isolates in this study were randomly selected from 201 *Salmonella*-contaminated samples which were initially collected from the total number of 718 samples from various sources, origin and provinces in northern Thailand during February 2018 to September 2019 as illustrated in Figure 1.

Rapid *Salmonella* serotyping was performed by the real-time triplex PCR generating two principal T_m at 87 °C (*gyr*B and *ycf*Q) and additional T_m at 83 °C (*flj*B). To generate HRM melt curves for *Salmonella* typing, the HRM was performed upon normalized and temperature-shifted view with pre-melt range 80.4 – 81.1°C and post-melt range 89.2 – 89.6°C (Figure 2). To cluster HRM melt curves, auto-clustering was correctly assigned by the software-equipped machine to most isolates with high confidence (> 97.0%); nevertheless, manual clustering was performed to yield six *Salmonella* isolates in cluster 2 which generated percent confidence ranging from 50.1% to 95.1% due to variation in HRM melt curves (Table 2). In this analysis, twelve unique HRM melt curves, easily differentiated by visual differentiation, were assigned as H3 to H14 while Cluster 1 (n=16) and 2 (n=9) were predominated in this study.

S-HRMa effectively differentiated closely related HRM melt curves of Cluster 1 and 2 based on the curve analysis of high variation-based regions.

Sixty-two percent of all forty isolates in this study, classified as Cluster 1 and 2, were further analyzed by the S-HRMa (Figure 3). The assumption of this analysis was that different regions of closely related HRM melt curves still contain unique and corresponding curve-variation which was further analyzed specifically in their own region using a suitable reference capable of promoting such variation. Containing the most closely related HRM melt curves (n=16), Cluster 1 showed less curve–variation across the temperature range from 77.0°C to 83.6°C except for *Salmonella* isolate L35 and L40 although various references were performed (data not shown) (Figure 3A). However, cluster 2 initially showed high curve–variation at two specific regions which were at 79.6°C – 83.0°C region and 85.0°C – 87.3°C region using *S.* Barille as the reference (Figure 3B). To increase such curve-variation, in-cluster reference was selected using *Salmonella* isolate L16 as the new reference. Two further analyses assigned as low temp region (79.6°C – 83.0°C) and high temperature region (85.0°C – 87.3°C) were conducted separately using L16 as reference (Figure 3C). The results showed nine novel HRM melt curves (A-G) using S-HRMa. The

low and high temperature region generated three HRM melt curves (A-C) and six HRM melt curves (G-L) respectively suggesting that two different regions apparently contain unique and specific information enough for further and functional clustering of *Salmonella* isolates. The S-HRMa was performed to further assign the HRM melt curves in cluster 1 and 2 to 3 and 10 unique patterns including the reference (L16), respectively.

S-HRMa and serogroups practically provide both salmonella typing and subtyping compared with conventional serotyping.

The rapid HRM serotyping initially classified 40 *Salmonella* isolates into 12 clusters with 0.7949 discrimatory index. The Salmonella serogroups and the S-HRMa were effectively performed to increase the discrimatory index to 0.9603 comparable to the conventional serotyping and the ERIC clusters (91% similarity) (Table 3). The informative serogroup of *Salmonella* isolates in Cluster 1 was practically employed to adequately classify the identical 18 HRM melt curves in the cluster 1 to *S.* Weltevraden (n=14, Group E), *S.* Agona (n=2, Group B) and *S.* Corvallis (n=2 Group C). For cluster 2, only *S.* Braenderup (n=2) was correctly classified to the distinct serogroup C while both *S.* Monophasic and *S.* Derby were the same serogroup B. Moreover, the informative serogroup and the S-HRMa effectively assigned thirteen novel clusters or the sensitive clusters. The initial 12 clusters plus the novel 13 clusters increased the discrimatory index of the rapid HRM serotyping of all 40 *Salmonella* isolates to 0.9603, more than that of the conventional serotyping (0.891). The subtyping capability of the HRM serotyping was evidently observed with four serotypes: *S.* Kentucky, *S.* Corvallis, *S.* Stanlay and *S.* Enteritidis as indicated, with two different HRM clusters assigned to each serotype (Figure 4A). After employing serogroups and performing S-HRMa to HRM serotyping, more than one sensitive cluster were assigned to some *Salmonella* isolates in the same serotypes such as *S.* Weltevraden, *S.* Monophasic *S.* Derby and *S.* Braenderup (Table 3). As a result, the capability of *Salmonella* subtyping was observed in this study.

The hierarchical clustering assessment through DTW assessment effectively provided the rapid protocol for constructing the phylogenetic tree with four clades and corresponded to S-HRMa.

The DTW analysis was performed to build a dendrogram, reflecting hierarchical clustering and phylogenic relatedness of 40 *Salmonella* isolates (Figure 5). Applying the distance level of clustering at 0.0003, 28 DTW clusters were generated with the high discrimatory index at 0.9679 compared to 0.9603 obtained from the S-HRMa (25 sensitive clusters) and 0.7949 from HRM Serotyping (14 HRM clusters). In addition, the DTW clustering significantly corresponded to the HRM serotyping with the majority of cluster 1 and cluster 2 observed in a separated clade 1 and 2, respectively. Compatible with the S-HRMa, DTW clustering located two distinct *S.*

Weltevreden of sensitive cluster 2 and 3 (L35 and L40) at the clade 3 compared to all nine *S.* Weltevreden of sensitive cluster 1, located exclusively at the clade 1. For these two methods, *Salmonella* subtyping was observed in cluster 2 more than that in cluster 1. In addition, four *S.* Monophasic isolates were effectively subtyped by the DTW clustering but those of S-HRMa showed clonally related property. For two *S.* Derby, only S-HRMa indicated different subtypes. In general, clade 3 and 4 contained several *Salmonella* serotypes generating visually differentiated melt curves except two *Salmonella* isolates from two *S.* Stanley (L30, L22) and *S.* Typhimurium (L17, L28) which were classified by these two clustering methods. However, the clonally related property was assumed due only to the genetic similarity of only three genes: *flj*B (170 bp), *gyr*B (171 bp) and *ycf*Q (241 bp), targeted for the *Salmonella* clustering.

The molecular subtyping of salmonella serotypes indicated by the combined results of the clustering assessment through DTW and S-HRMa

The HRM melt curves were generated to basically indicate information of amplified PCR fragments through the sequence polymorphism; thus, the different patterns of HRM melt curves were applied to *Salmonella* clustering by visual observation (HRM clustering), modified visual observation (S-HRMa) and algorithm-based methods with DTW assessment. HRM serotyping initially classified *Salmonella* spp. to fourteen clusters, and later both the S-HRMa and DTW clustering generated twenty-five sensitive clusters and twenty-eight DTW clusters, respectively. Moreover, the combination of sensitive clusters and DTW clusters were applied to indicate the serotypes and subtypes of forty *Salmonella* isolates. The plots of negative first derivative of normalized HRM melt curves, indicating different *Salmonella* serotypes, are shown in Figure 6. Twelve *Salmonella* isolates of *S.* Weltevreden in Cluster 1 were further characterized to five different sub clusters or subtypes after performing the S-HRMa and DTW clustering. The predominant *Salmonella* isolates of *S.* Weltevreden subtype DTW4-S1 (n=6) was observed in the following DTW clusters: DTW1-S1 (n=2), DTW5-S1 (n=1), DTW16-S2 (n=1) and DTW17-S3 (n=1). In addition, two subtypes of *S.* Covallis were observed in cluster 1 and H11 of HRM cluster (DTW 21-S22) while two *S.* Agona isolates (DTW 5-S4) were observed to be clonally related. The same analysis also suggested all nine *Salmonella* isolates in cluster 2 or *S.* Monophasic (n=4), *S.* Braederup (n=2) and *S.*Derby (n=3) were of different *Salmonella* subtypes. Furthermore, the HRM cluster H3 to H14 were mostly correlated to different *Salmonella* serotypes with the exception of *S.* Stanley (H4 and H5), *S.* Enteritidis (H6 and H7) and *S.* Kentucky (H10 and H13). However, two subtypes of *S.* Typhimurium were diagnosed within the same H3 cluster as DTW26-S15 (n=2) and DTW27-S15 (n=1). In this study, the S-HRMa and DTW assessment were considered to be the effective methods for rapid subtyping of *Salmonella* spp.

The S-HRMa and DTW were observed to produce the clustering results corresponding to the ERIC-PCR clusters.

All 40 Salmonella isolates were phylogenetically analyzed by performing ERIC-PCR to create DNA fingerprint and its corresponded phylogeny tree with all related clustering and epidemiological dataset (Figure 7). The ERIC-PCR gave exceptionally high discriminatory index of 0.9962 at 91% similarity. Significantly, 27 *Salmonella* isolates in clusters C1 and C2 were observed to be clustered together in the phylogeny tree suggesting the phylogenetic related properties of those *Salmonella* isolates in the two clusters especially in cluster 2 of which nine *Salmonella* isolates of three serotypes arranged in close proximity. Evidently, the phylogeny tree from ERIC-PCR revealed three clusters (E10, E18, E21) to be clonally related at 91%. In this experiment, the S-HRMa and the DTW analysis were together applied to cluster *Salmonella* isolates with the combined clusters of the sensitive and the DTW clusters. The combined clusters of DTW4 and S1 effectively suggest the genetic relatedness property of E10 cluster. However, S6 or DTW7 sufficiently indicate the genetic relatedness property of E18 and E21, respectively. The results revealed the possibility of applying those two clustering methods to perform *Salmonella* subtyping to the level of 91% similarity from the ERIC-PCR. Further analysis of antibiotic determinants, *bla*TEM and *flo*R suggested only E18, possibly possessing clonally related properties due to acquiring the same antibiotic determinants; on the contrary, the *Salmonella* isolates in both E10 and E21 showed different antibiotic determinants and were classified as different *Salmonella* subtypes. In conclusion, the S-HRMa and DTW analysis could be applied to indicate different subtypes of *Salmonella* isolates at the high level of 91% genetic similarity, comparable to that performed by ERIC-PCR.

S-HRMa and DTW assessment suggested the distribution of the clonally related Salmonella isolates in Northern Thailand.

All related epidemiological data of all 40 *Salmonella* isolates including isolated locations, sources, types of animals and samples of *Salmonella* isolates were provided for each *Salmonella* strains in their corresponded serotypes in Table 4. Of all 40 *Salmonella* isolates in this study, *S.* Weltevraden (n=12) was the most prevalent *Salmonella* serotype found in nearly all provinces in Northern Thailand except Chaing Rai and showed that the most associated epidemiological data was farm, chicken and stool sample, containing the group of eight *Salmonella* serotypes as follows: *S.* Albany, *S.* Corvallis, *S.* Enteritidis, *S.* Kentucky, *S.*agona, *S.* Braenderup, *S.* Weltevraden and *S.* Stanley. Distinctively, *S.* Derby (n=3) and *S.* Monophasic (n=4) were mostly associated with the slaughterhouse and swine while S. Typhimurium (n=3) associated with goats and house. *Salmonella* isolates showing high similarity (>91%), effectively identified by the S-HRM analysis

and the DTW clustering, were indicated in three *Salmonella* serotypes as *S.* Monophasic, *S.* Derby and *S.* Weltevraden. From the associated epidemiological data, the transmission route of *S.* Weltevraden was likely to be farm-to-farm in the same province while that of *S.* Derby and *S.* Monophasic was predicted to be house-to-slaughterhouse and vice versa in different provinces.

Discussion

The HRM serotyping was well established to rapidly identify the frequently encountered *Salmonella* serotypes associated with hospitalized patients and minced pork in the Northern Thailand (Pooncharoen et al., 2019). The capability of Salmonella subtyping in this study was observed with *S.* Stanley, *S.* Enteritis and *S.* Kentucky, which exhibited two visually different HRM patterns to each serotype. To explain, the underlying principle of this multiplex HRM serotyping was the DNA polymorphism of three Salmonella gene targets; *flj*B, *gyr*B, *ycf*Q, in each Salmonella serotypes or subtypes (Zeinzinger et al., 2012). In this study, the sensitive high - resolution melting - curve analysis (S-HRMa) was effectively established to differentiate two clusters containing the majority of Salmonella isolates exhibiting closely related HRM patterns. The S-HRMa was performed by using internal reference which effectively promoted unified HRM patterns for further clustering. With the informative serogroup, the established protocol effectively improved the discriminatory index from 0.7949 to 0.9679 with the good correlation to the standard Salmonella subtyping by ERIC - PCR based phylogeny as observed in several publications (Johnson et al., 2001). In comparison to the S-HRMa, the raw data set of normalized melting curves from HRM analysis was performed to construct the hierarchical clustering by using Dynamic Time Warping assessment (DTW) as shown to be the rapid and robust technique capable of identifying 243 clinical fungal isolates (Lu et al., 2017). The results of the DTW clustering was indicated with high discriminatory index 0.9679 together with the construction of the dendrogram compatible with the ERIC - PCR based phylogeny. Furthermore, the combination of the S-HRMa and the DTW clustering successfully subtyped eleven *S.* Weltevraden, four *S. Monophasic* and three *S.* Derby isolates to five, four and three subtypes respectively. The evidence of *S.* Monophasic subtypes in this study was well correlated to the continually evolving property of this virulent serotype during clonal expansion in many countries (Petrovska et al., 2016) (Izumiya et al., 2018). In this study, the combination of S-HRMa and DTW clustering was introduced to rapidly subtype *Salmonella* isolates with sufficient efficacy to the level of Rep based PCR molecular typing compared with other standard protocols such as PFGE or MLST which are costly and complicated protocols.

The most prevalent S. Weltevreden (n=12) isolates was observed in nearly all provinces in the Northern Thailand with the highly associated pattern to chickens and their line of production. *S.* Weltevreden was regarded as the prevalent serotype found in various contaminated food as well as ~~the~~ humans in Thailand (Bangtrakulnonth et al., 2004) (Lertworapreecha, Sutthimusik & Tontikapong, 2013). After further analysis, *S.* Weltevreden (n=12) isolates were rapidly subtyped by the combination of two clustering methods to five different subtypes with the predominant subtype (DTW4-S1, n=6) significantly correlated with

the stool samples of chickens from farms. The prediction of transmission route by this *Salmonella* subtyping and related dataset was likely to be the spread of *S.* Weltevraden between chicken farms in the same province possibly due to natural carriers (Skov et al., 2004). Thus, the combination of two protocols were performed after the HRM serotyping to identify *Salmonella* subtypes and the different transmission route of *Salmonella* contamination as previously achieved by the subtyping of *S.* Weltevreden (n=22) isolates by ERIC - PCR fingerprints (Kumar, Surendran & Thampuran, 2009). Additionally, *S.* Monophasic was found exclusively in two populated provinces such as Chiang mai and Lampang in the Northern Thailand and observed to be correlated to swine of which its line of industry is regarded as the main reservoir of human infection in Thailand (Padungtod & Kaneene, 2006). Additionally, two *S.* Monophasic isolates exhibiting 91.8 similarity of ERIC pattern also suggested their transmission route to possibly be house-to-slaughterhouse and vice versa in different provinces. An interesting observation was *S.* Typhimurium association with goats and houses raising concern of Salmonella contamination in goat - associated products (Duffy et al., 2009).

Conclusions

In this study, we described the effective protocols that can be performed after HRM serotyping in order to further subtype *Salmonella* spp. with high discrimatory index comparable to Rep PCR fingerprints. The S-HRMa, informative serogroups and the DTW were firstly introduced to the epidemiological research of *Salmonella* spp. as the simple and rapid molecular typing with the high discrimatory index comparable to ERIC - PCR fingerprints together with our future prospect of this protocol to rapidly serotype Salmonella samples directly from infected patients for effective treatment and epidemiological concern.

Acknowledgements

The authors are also deeply indebted to Asst. Prof. Dr. Suphak Mahatthontanahak for providing both equipment and research areas at the central lab (CLUP) and especially to Professor Emeritus Torpong Sanguansermsri at Thalassemia unit for inspiring the High - resolution melting - curve analysis facility to this study. Finally, the first author of this study wishes to sincerely pay his respects to Prof. Prapon Wilairat, who always supports both knowledge and the passion for mastering knowledge.

Conflict of Interest

The authors confirm that there are no conflicts of interest associated with this publication and there has been no significant financial support for this work that could have influenced its outcome.

References

Arguello H, Carvajal A, Collazos JA, García-Feliz C, Rubio P. 2012. Prevalence and serovars of Salmonella enterica on pig carcasses, slaughtered pigs and the environment of four Spanish slaughterhouses. Food Research International 45:905–912. DOI: 10.1016/j.foodres.2011.04.017.

Athamanolap P, Parekh V, Fraley SI, Agarwal V, Shin DJ, Jacobs MA, Wang TH, Yang S. 2014. Trainable high resolution melt curve machine learning classifier for large-scale reliable genotyping of sequence variants. PLoS ONE 9. DOI: 10.1371/journal.pone.0109094.

Bangtrakulnonth A, Pornreongwong S, Pulsrikarn C, Sawanpanyalert P, Hendriksen RS, Lo Fo Wong DM a, Aarestrup FM. 2004. Salmonella Serovars from Humans and Other Sources in Thailand, 1993-2002. Emerging Infectious Diseases 10:131–136. DOI: 10.3201/eid1001.02-0781.

Cardona-Castro N, Sánchez-Jiménez M, Lavalett L, Múñoz N, Moreno J. 2009. Development and evaluation of a multiplex polymerase chain reaction assay to identify Salmonella serogroups and serotypes. Diagnostic Microbiology and Infectious Disease 65:327–330. DOI: 10.1016/j.diagmicrobio.2009.07.003.

Crespo R, Sischo WC, Guard J, Shah DH, Paul NC. 2016. Population dynamics and antimicrobial resistance of the most prevalent poultry-associated Salmonella serotypes. Poultry Science 96:687–702. DOI: 10.3382/ps/pew342.

Davies RH, Gosling RJ, Rabie A, Wales AD, Mueller-Doblies D, Nunez-Garcia J, Kell N, Martelli F. 2018. Observations on the distribution and persistence of monophasic Salmonella Typhimurium on infected pig and cattle farms. Veterinary Microbiology 227:90–96. DOI: 10.1016/j.vetmic.2018.10.032.

Duffy L, Barlow R, Fegan N, Vanderlinde P. 2009. Prevalence and serotypes of Salmonella associated with goats at two Australian abattoirs. Letters in Applied Microbiology 48:193–197. DOI: 10.1111/j.1472-765X.2008.02501.x.

Forshell LP, Wierup M. 2006. <Plym Forshell, L. 2006.pdf>. Rev. sci. tech. Off. int. Epiz 25:541–554.

Grimont P, Weill F-X. 2007. Antigenic formulae of the Salmonella servovars. France.

Rossum GV. 1995. Python tutorial, Technical Report CS-R9526, Centrum voor Wiskunde en Informatica (CWI). Amsterdam.

Hauser E, Tietze E, Helmuth R, Junker E, Blank K, Prager R, Rabsch W, Appel B, Fruth A, Malorny B. 2010. Pork contaminated with Salmonella enterica serovar 4,[5],12:i:-, an

emerging health risk for humans. Applied and Environmental Microbiology 76:4601–4610. DOI: 10.1128/AEM.02991-09.

Hendriksen RS, Bangtrakulnonth A, Pulsrikarn C, Pornruangwong S, Noppornphan G, Emborg H, Aarestrup FM. 2009. Risk Factors and Epidemiology of the Ten Most Common. 6.

Herikstad H, Motarjemi Y, Tauxe R V. 2002. Salmonella surveillance: A global survey of public health serotyping. Epidemiology and Infection 129:1–8. DOI: 10.1017/S0950268802006842.

Hohmann EL. 2001. Nontyphoidal Salmonellosis. Clin Infect Dis 32:263–9.

Hunter PR, Gaston MA. 1988. Numerical Index of the Discriminatory Ability of Typing Systems : an Application of Simpson ' s Index of Diversity. Journal of Clinical Microbiology 26:2465–2466.

Izumiya H, Kuroda M, Tamamura Y, Sekizuka T, Hinenoya A, Tanaka K, Kusumoto M, Arai N, Iwata T, Barco L, Yamasaki S, Watanabe A, Uchida I, Akiba M. 2018. Phylogenetic Characterization of Salmonella enterica Serovar Typhimurium and Its Monophasic Variant Isolated from Food Animals in Japan Revealed Replacement of Major Epidemic Clones in the Last 4 Decades . Journal of Clinical Microbiology 56:1–14. DOI: 10.1128/jcm.01758-17.

Jackson BR, Griffin PM, Cole D, Walsh KA, Chai SJ. 2013. Salmonella enterica Serotypes and Food Commodities ,. Emerg Infect Dis 19:1239–1244. DOI: 10.3201/eid1908.121511.

Johnson JR, Clabots C, Azar M, Boxrud DJ, Besser JM, Thurn JR. 2001. Molecular analysis of a hospital cafeteria-associated salmonellosis outbreak using modified repetitive element PCR fingerprinting. Journal of Clinical Microbiology 39:3452–3460. DOI: 10.1128/JCM.39.10.3452-3460.2001.

Kumar R, Surendran PK, Thampuran N. 2009. Distribution and genotypic characterization of Salmonella serovars isolated from tropical seafood of Cochin, India. Journal of Applied Microbiology 106:515–524. DOI: 10.1111/j.1365-2672.2008.04020.x.

Lertworapreecha M, Sutthimusik S, Tontikapong K. 2013. Antimicrobial resistance in Salmonella enterica isolated from pork, chicken, and vegetables in southern Thailand. Jundishapur Journal of Microbiology 6:36–41. DOI: 10.5812/jjm.4321.

Lu S, Mirchevska G, Phatak SS, Li D, Luka J, Calderone A, Fonzi WA. 2017. Dynamic time warping assessment of high- resolution melt curves provides a robust metric for fungal identification. PLoS ONE 12:1–21. DOI: 10.1371/journal.pone.0173320.

McNerney R, Clark TG, Campino S, Rodrigues C, Dolinger D, Smith L, Cabibbe AM, Dheda K, Schito M. 2017. Removing the bottleneck in whole genome sequencing of Mycobacterium tuberculosis for rapid drug resistance analysis: a call to action. International Journal of Infectious Diseases 56:130–135. DOI: 10.1016/j.ijid.2016.11.422.

Padungtod P, Kaneene JB. 2006. Salmonella in food animals and humans in northern Thailand. International Journal of Food Microbiology 108:346–354. DOI: https://doi.org/10.1016/j.ijfoodmicro.2005.11.020.

Petrovska L, Mather AE, Abuoun M, Branchu P, Harris SR, Connor T, Hopkins KL, Underwood A, Lettini AA, Page A, Bagnall M, Wain J, Parkhill J, Dougan G, Davies R, Kingsley RA. 2016. Microevolution of monophasic Salmonella typhimurium during epidemic, United Kingdom, 2005–2010. Emerging Infectious Diseases 22:617–624. DOI:

10.3201/eid2204.150531.

Pooncharoen K, Narong Nuanmuang, Prommuang P, Sriisan S. 2019. High-resolution melting-curve analysis for serotyping of Salmonella spp. group B isolated from minced pork in the Northern part of Thailand. Journal of Associated Medical Sciences 52:62–71. DOI: 10.14456/jams.2018.3.

Pulsrikarn C, Pornreongwong S, Tribuddharat C, Meethai C, Srifuengfung S. 2013. Serogroup and Serovar Distribution of Salmonella in Siriraj Hospital. Siriraj Med J 65:s34–s37.

Singh P, Mustapha A. 2014. Development of a real-time PCR melt curve assay for simultaneous detection of virulent and antibiotic resistant Salmonella. Food Microbiology 44:6–14. DOI: 10.1016/j.fm.2014.04.014.

Skov MN, Spencer AG, Hald B, Petersen L, Nauerby B, Carstensen B, Madsen M. 2004. The Role of Litter Beetles as Potential Reservoir for Salmonella enterica and Thermophilic Campylobacter spp. Between Broiler Flocks. Avian Diseases 48:9–18. DOI: 10.1637/5698.

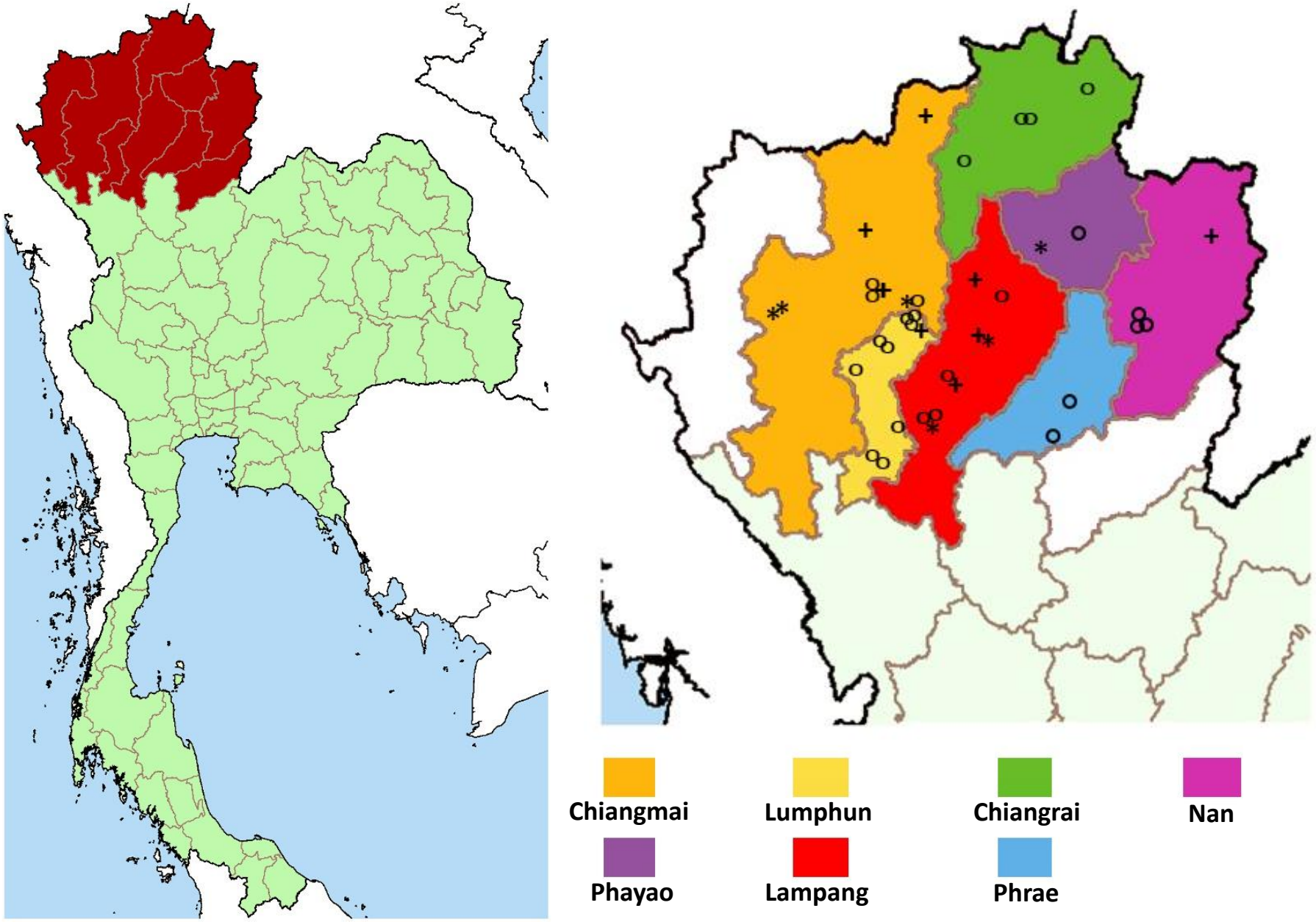
Versalovic J, Koeuth T, Lupski JR. 1991. Distribution of repetitive DNA sequences in eubacteria and application to fingerprinting of bacterial genomes. Nucleic Acids Research 19:6823–6831.

Zeinzinger J, Pietzka AT, Stöger A, Kornschober C, Kunert R, Allerberger F, Mach R, Ruppitsch W. 2012. One-step triplex high-resolution melting analysis for rapid identification and simultaneous subtyping of frequently isolated Salmonella serovars. Applied and Environmental Microbiology 78:3352–3360. DOI: 10.1128/AEM.07668-11.

Figure 1(on next page)

Geographic data of different sampling locations of 40 *Salmonella* isolates during February 2018 to September 2019 from the northern part of Thailand.

(A) National Thai map is illustrated in green color and the dark red color represents Northern Thailand, the focused area of sampling locations. (B) Seven provinces of Northern Thailand as the different sites of sampling locations of 40 *Salmonella* isolates with different labeled colors to each province. The sampling sites of 40 *Salmonella* isolates were shown as different symbols in the map



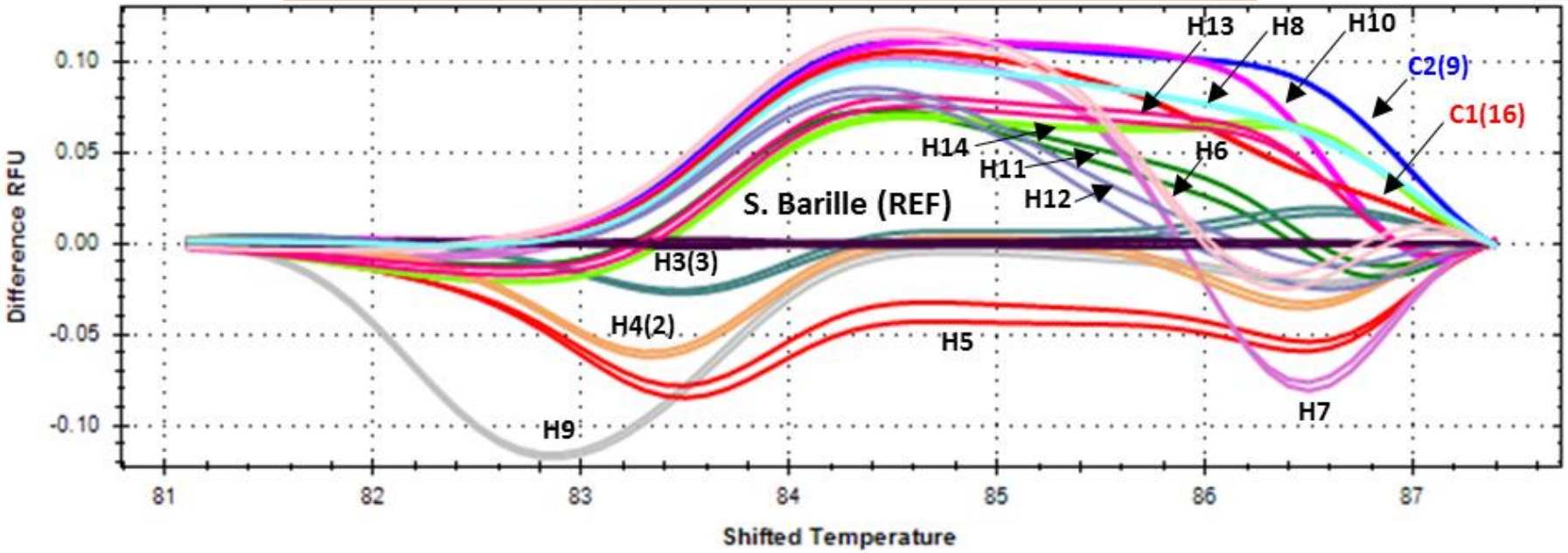
PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27664v1 | CC BY 4.0 Open Access | rec: 19 Apr 2019, publ: 19 Apr 2019

Figure 2(on next page)

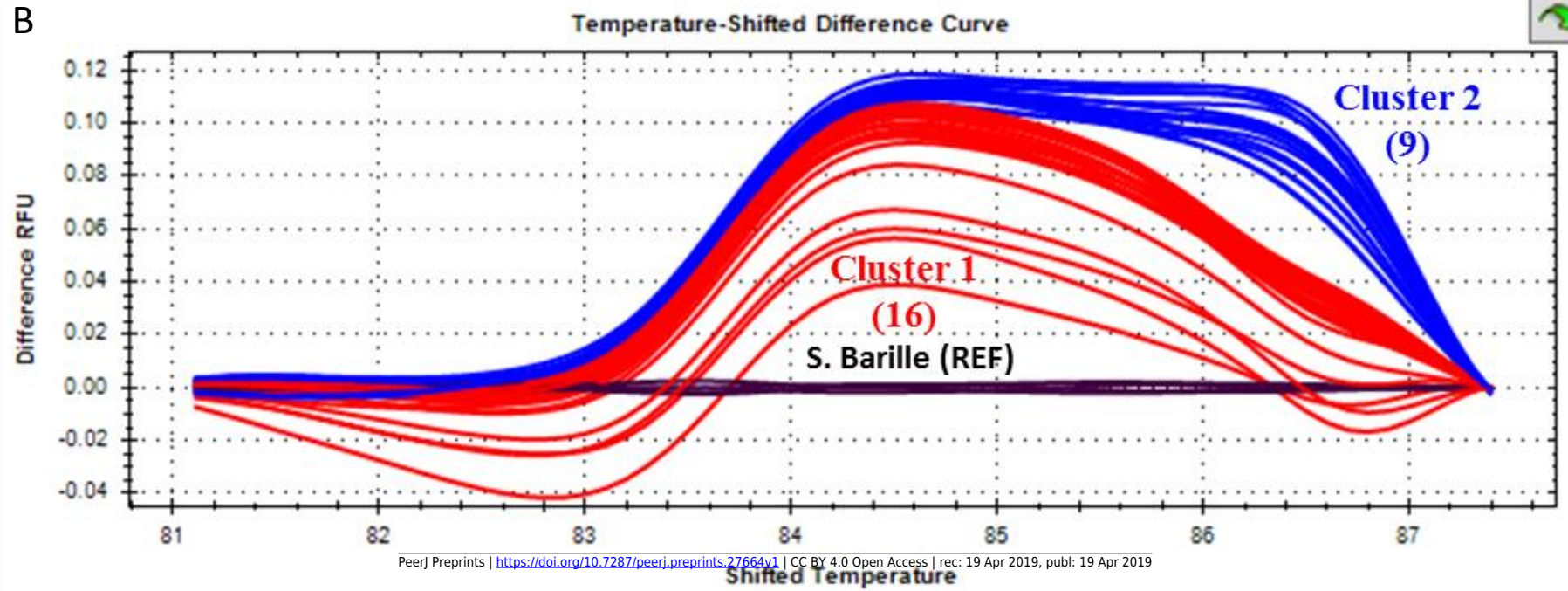
Sixteen HRM patterns assigned to 40 *Salmonella* isolates from different specimens, animals, sources and provinces from the northern part of Thailand.

(A) Sixteen visually different HRM curves with their duplication using *S*. Barille as reference was illustrated in different colors and labeled with the alphabet “C” or “H” followed by a specified number with the number in parenthesis suggesting the number of *Salmonella* isolates exhibiting the patterns. (B) The closely related similar clusters, assigned as Cluster 1 (blue color) and Cluster 2 (red color), were the dominant HRM patterns in this study.

A



B

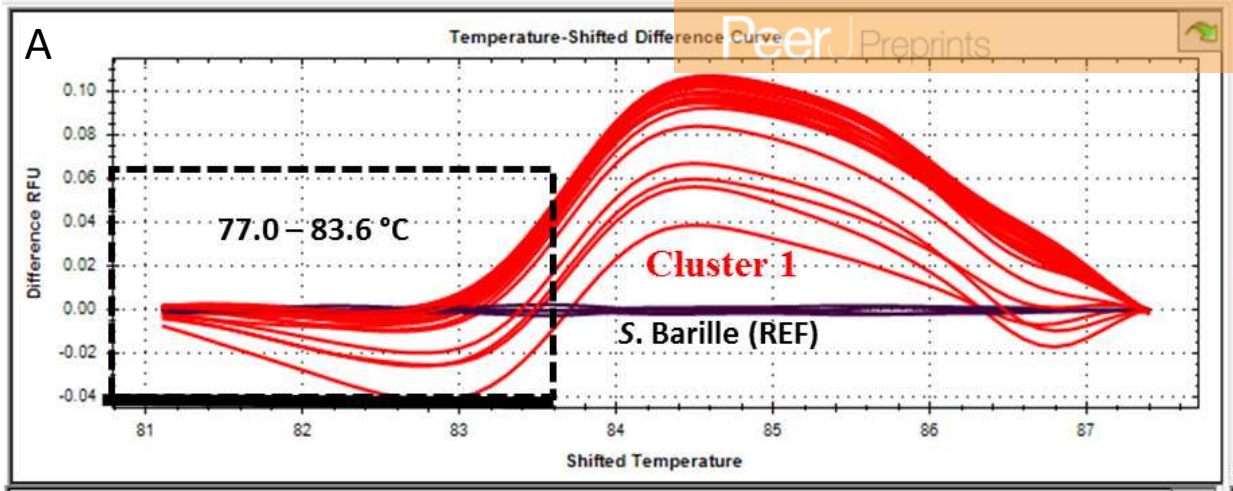


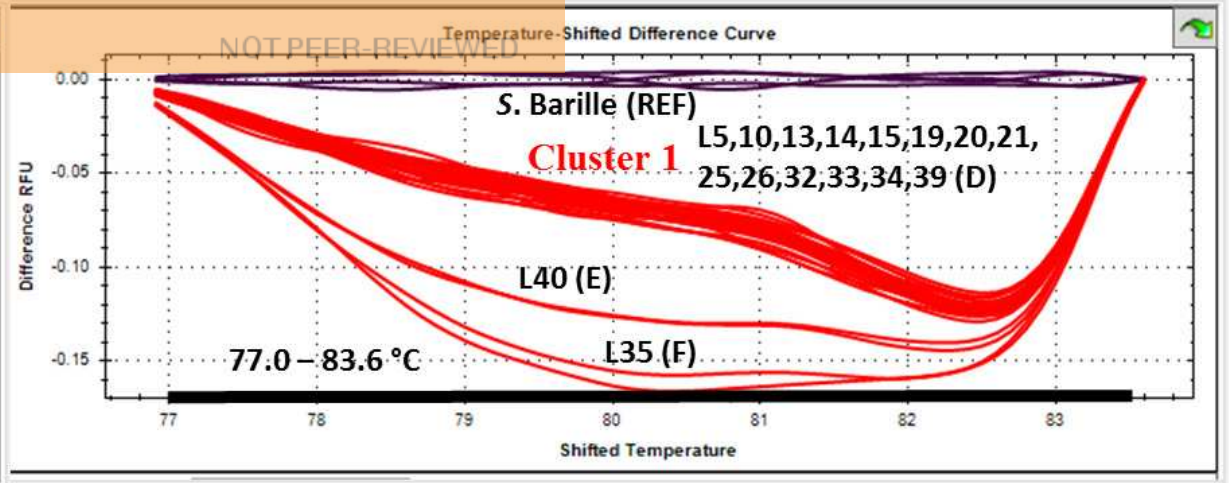
PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27664v1 | CC BY 4.0 Open Access | rec: 19 Apr 2019, publ: 19 Apr 2019

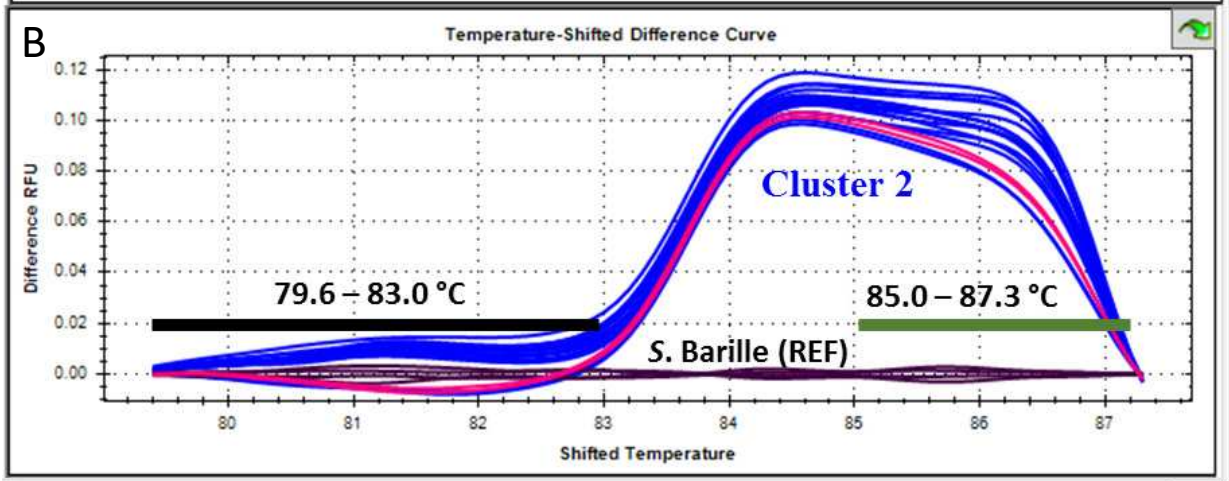
Figure 3(on next page)

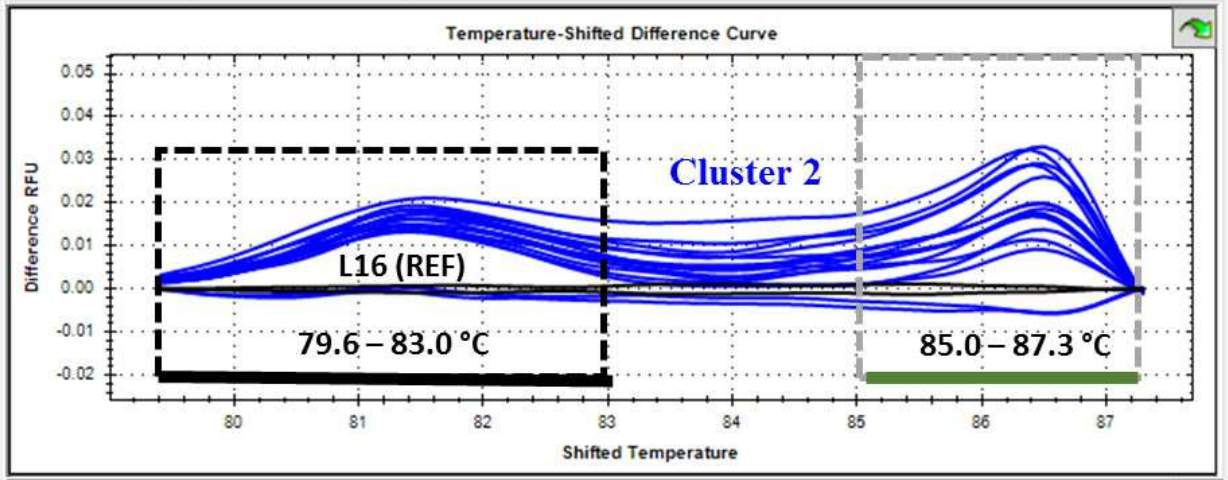
The Sensitive High Resolution Melting analysis (S-HRMa) differentiated cluster 1 and cluster 2 to various curve patterns patterns (A – L) used for further clustering.

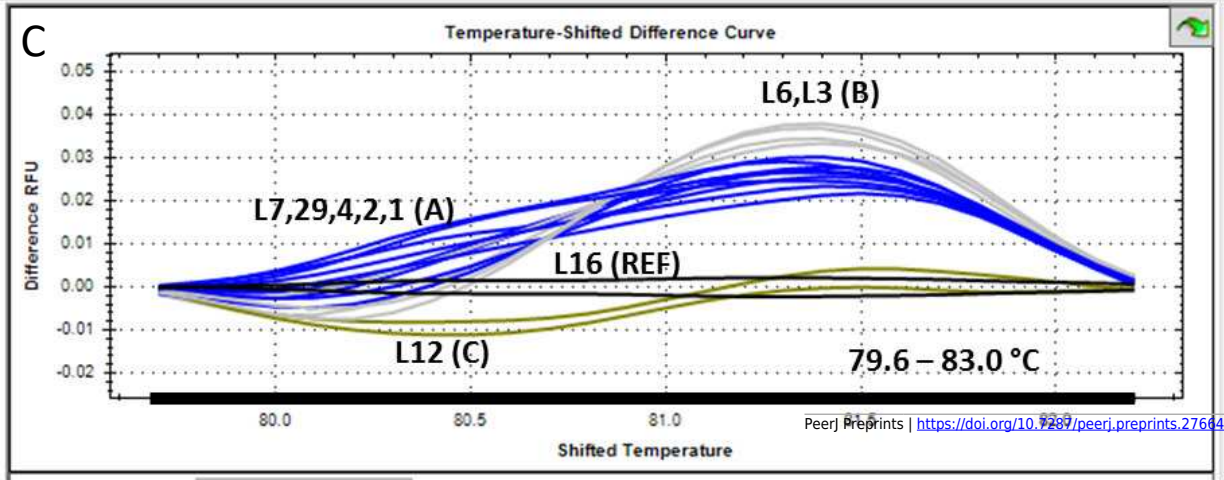
(A) Cluster 1 was differentiated by performing HRM analysis with one temperature range (77.0°C - 83.6°C) using *S*. Barille as reference. (B) Cluster 2 was differentiated by performing HRM analysis ranging from 79.6°C to 87.47°C using *S.* Barille (Left) L16 (Right) as reference. (C) Three (A to C) and six different colored HRM melt curves (G to L) were derived by performing HRM analysis with two temperature range (79.6°C – 83.0°C) and (85.0°C – 87.3°C) respectively with L16 as reference.











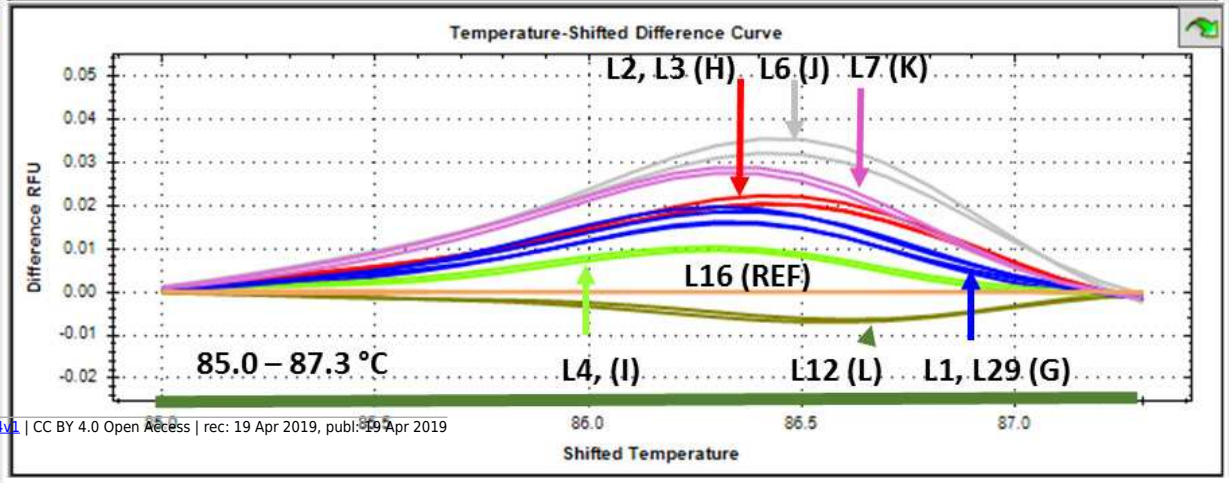
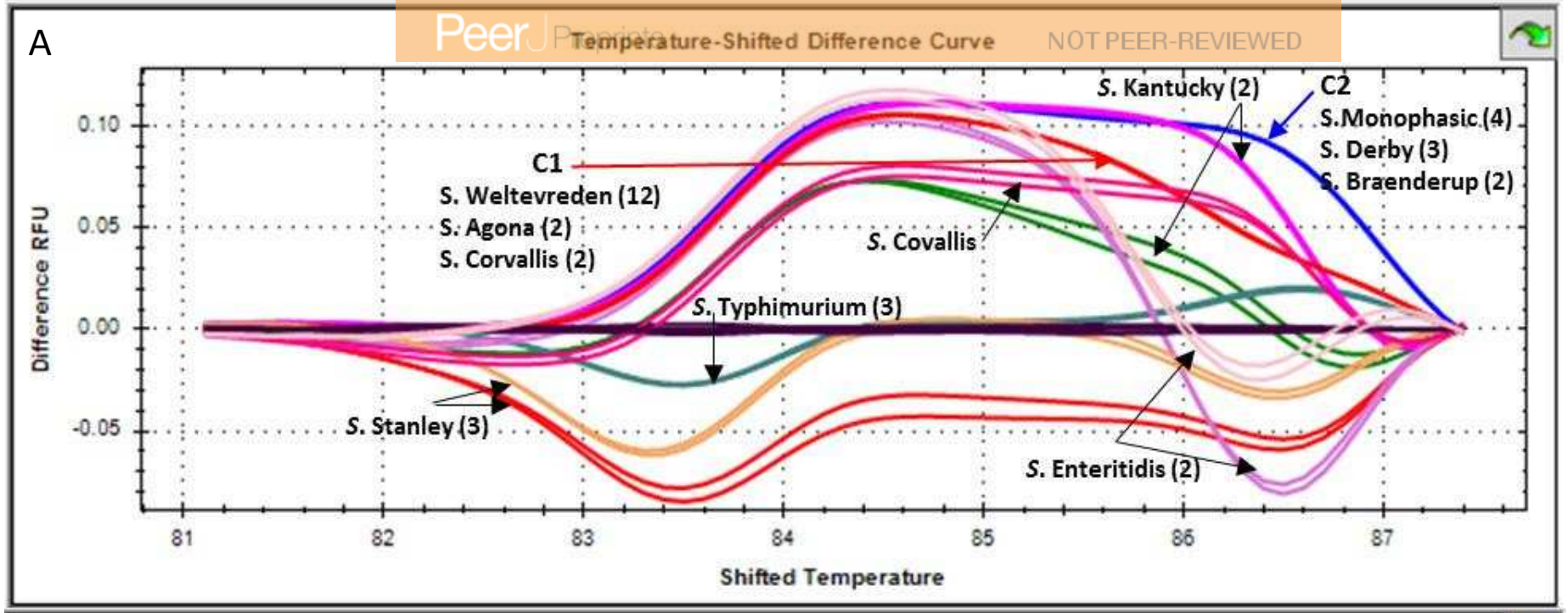


Figure 4(on next page)

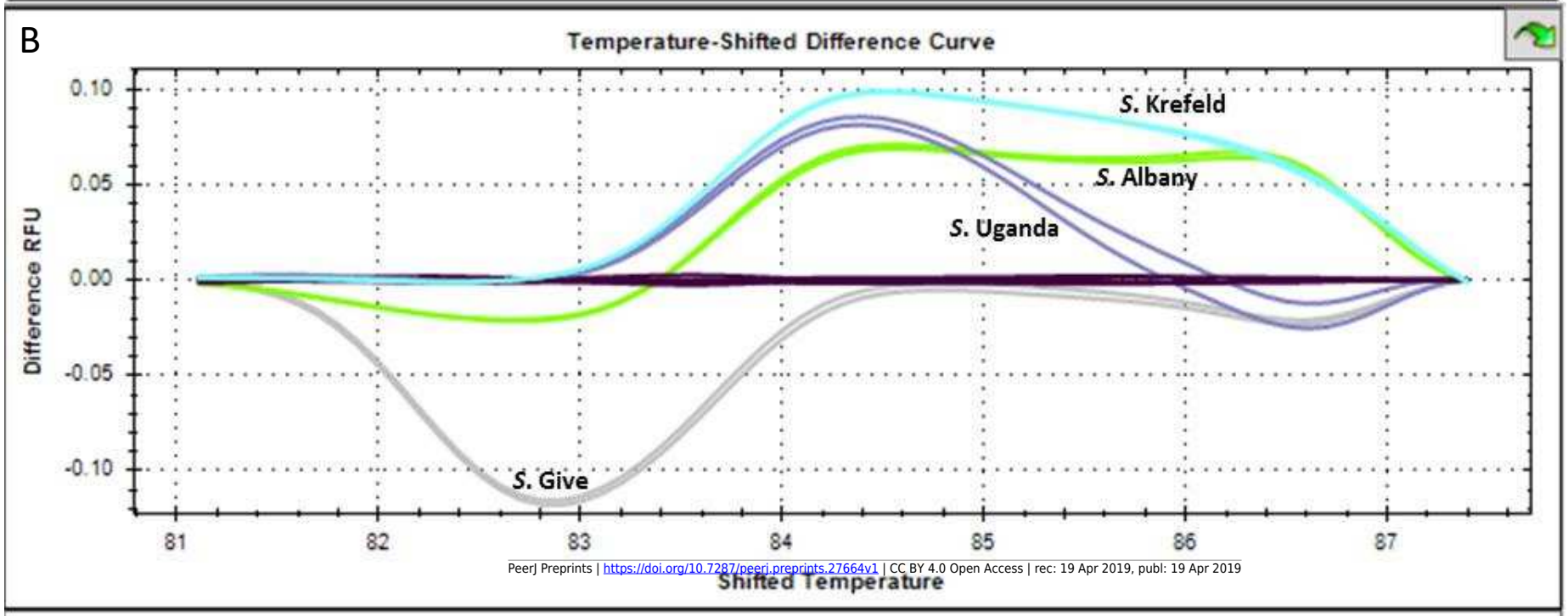
The corresponded *Salmonella* serotypes to sixteen HRM melt patterns assigned to the all 40 Salmonella from the northern part of Thailand.

(A) Eleven *Salmonella* serotypes assigned to ten different HRM curves, corresponded to the majority of all *Salmonella* isolates(n=36, 90 percent). The associated information with each HRM patterns was the cluster name, corresponded *Salmonella* serotypes and the number of *Salmonella* isolates in parenthesis. (B) Four *Salmonella* serotypes assigned to four different HRM curves, corresponded to the minority of *Salmonella* isolates (n=4, 10 percent)

A



B

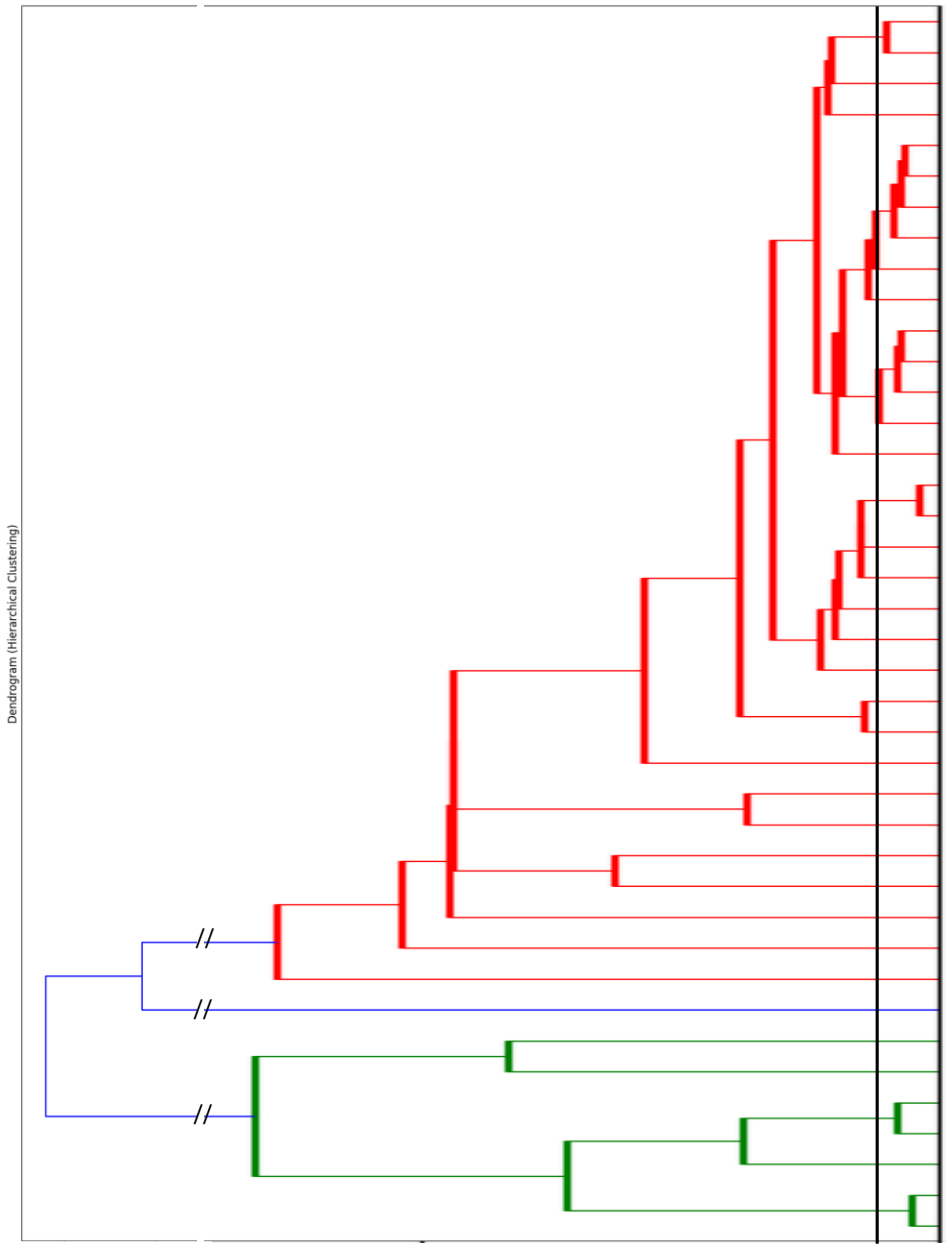


PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27664v1 | CC BY 4.0 Open Access | rec: 19 Apr 2019, publ: 19 Apr 2019

Figure 5(on next page)

The hierarchical phylogeny tree of 40 *Salmonella* isolates at the Northern Thailand during February 2018 to September 2019.

This phylogeny was created by performing the Dynamic Time Warping assessment (DTW) of the normalized melting curves from High - resolution melting - curve analysis using neighbor-joining method for the tree construction. The right side of the phylogeny tree was the illustration of four clades (black rectangular boxes) assigned at the phylogenic distance 0.0003. The associated clustering data from the HRM clusters, sensitive clusters, DTW clusters and corresponded serotypes were provided in the table. The red and blue color displayed in the table indicated the associated clustering data of Cluster 1 and Cluster 2 respectively. The Discrimatory power and number of types for each clustering was in the bottom of the table in gray color.

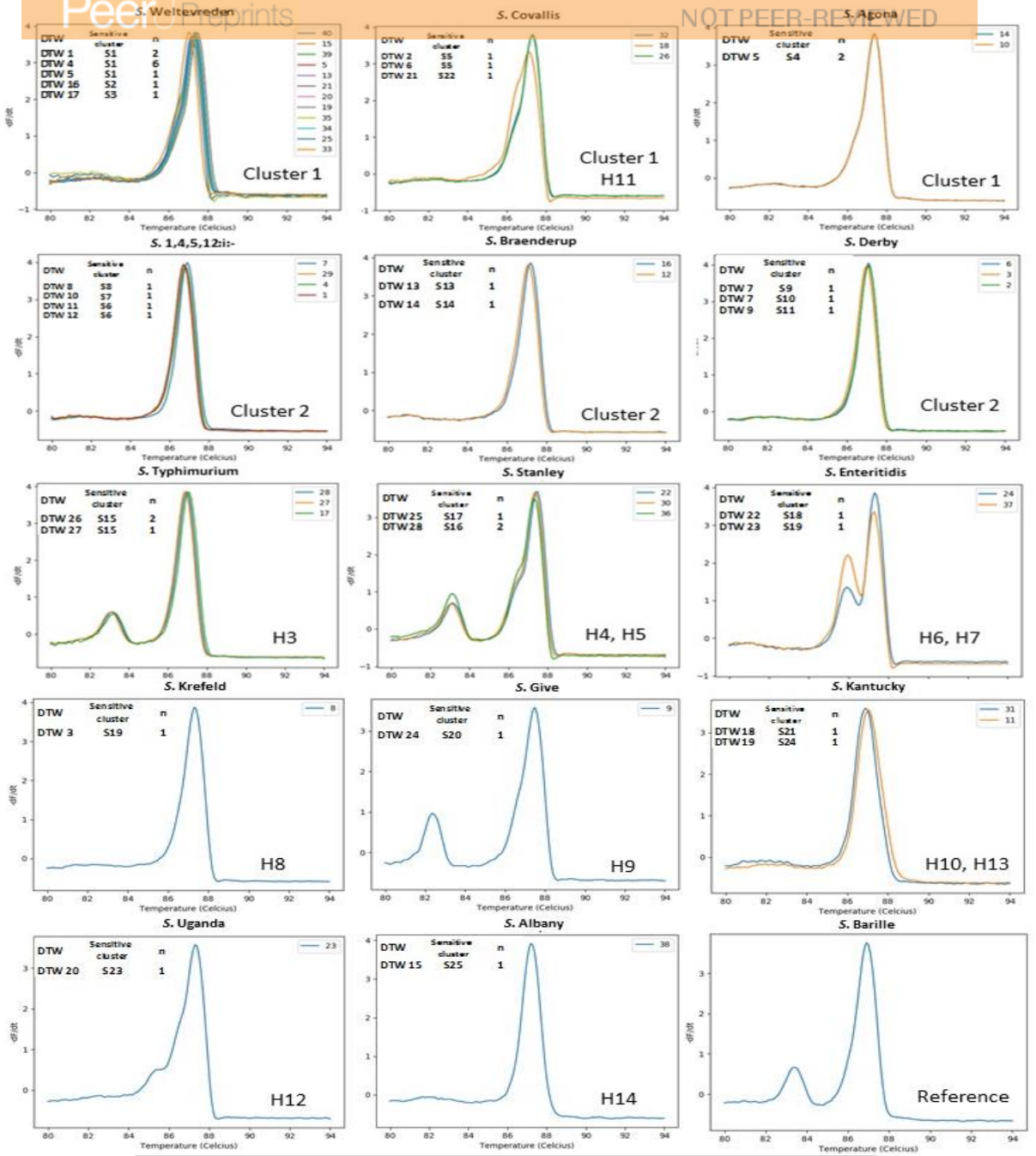


Strain number	HRM clusters 79-89	Sensitive clusters	DTW clusters	Conventional Serotypes
L13	Cluster 1	S1	DTW 1	*S*. Weltevreden
L5				
L26		S5	DTW 2	*S*. Corvallis
L8	H8	S19	DTW 3	S.Krefeld
L19	Cluster 1	S1	DTW 4	*S*. Weltevreden
L15				
L21				
L39				
L34				
L33				
L20			DTW 5	
L14		S4		*S*. Agona
L25		S1		*S*. Weltevreden
L10		S4		*S*. Agona
L32		S5	DTW 6	*S*. Corvallis
L2	Cluster 2	S9	DTW 7	*S*.Derby
L3		S10		
L7		S8	DTW 8	*S*. I.4,5,12:i:-
L6		S11	DTW 9	*S*. Derby
L4		S7	DTW 10	*S*. I.4,5,12:i:-
L1		S6	DTW 11	
L29			DTW 12	
L12		S13	DTW 13	*S*.Braenderup
L16		S14	DTW 14	
L38	H14	S25	DTW 15	*S*. Albany
L35	Cluster 1	S2	DTW 16	*S*. Weltevreden
L40		S3	DTW 17	
L11	H10	S21	DTW 18	*S*. Kantucky
L31	H13	S24	DTW 19	
L23	H12	P23	DTW 20	*S*. Uganda
L18	H11	S22	DTW 21	*S*. Corvallis
L24	H6	P18	DTW 22	*S*. Enteritidis
L37	H7	S19	DTW 23	
L9	H9	S20	DTW 24	*S*. Give
L36	H5	S17	DTW 25	*S*. Stanley
L17	H3	S15	DTW 26	*S*. Typhimurium
L28	H3			
L27	H3		DTW 27	
L30	H4	S16	DTW 28	*S*. Stanley
L22	H4			
	14	25	28	Number of types
	0.7949	0.9603	0.9679	Discrimatory power

Figure 6(on next page)

The *Salmonella* serotypes illustrated as the negative first derivative of normalized HRM melt graphs.

The HRM melting curves were normalized and negative first derivative (-dF/dt) was performed between the temperature 80°C to 94°C to the normalized curves. Inside each rectangular boxes. Each graph corresponded to each *Salmonella* isolates with the same serotypes were displayed with the overlying fashion. The HRM clusters and the reference were indicated at the bottom right as well as the top left, the subtyping signatures displayed as the DTW (The DTW cluster), the Sensitive cluster (S) and "n" as the number of Salmonella isolates. In addition, the top right, each number represents the strain number with their corresponded color of graphs. Fifteen *Salmonella* Serotypes including reference were displays in this figure.

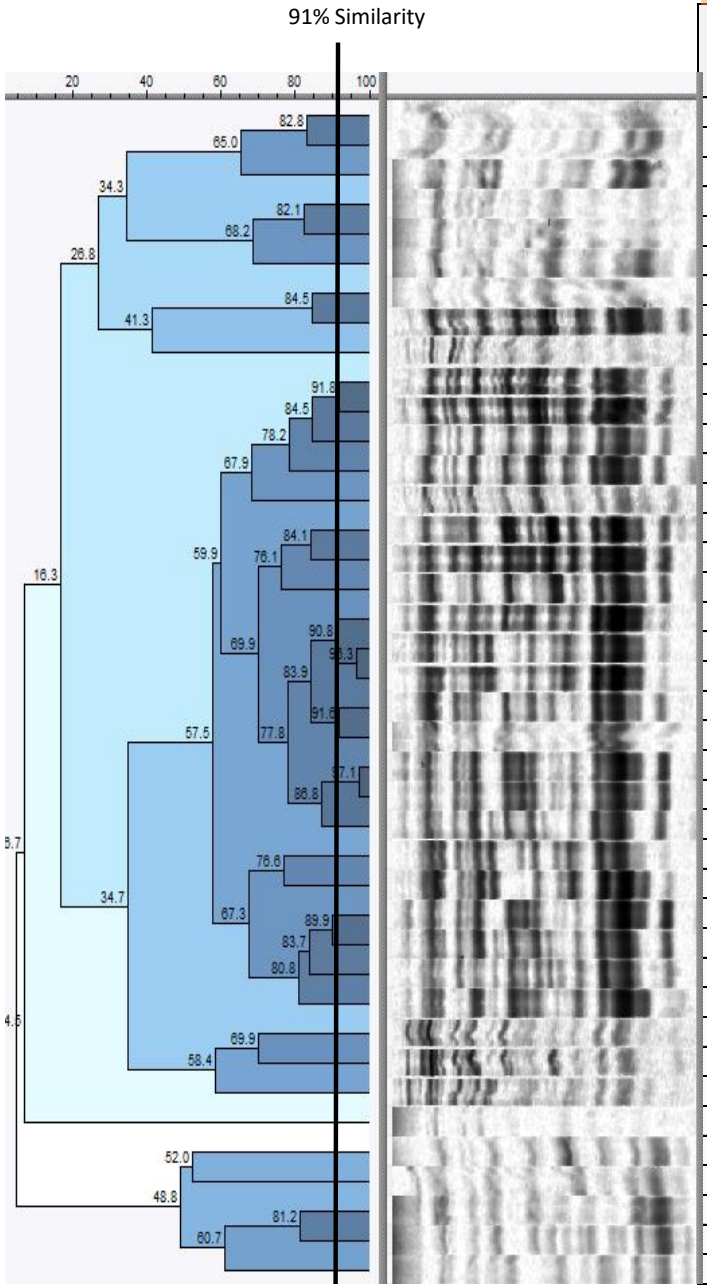


PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27664v1 | CC BY 4.0 Open Access | rec: 19 Apr 2019, publ: 19 Apr 2019

Figure 7 (on next page)

The phylogeny tree construction based on ERIC- PCR fingerprint patterns of 40 *Salmonella* isolates from different specimens, animals, sources and provinces from the Northern part of Thailand during February 2018 to September 2019.

Phylogeny tree constructed using curve based algorithm as pearson correlation and UPGMA for clustering fingerprints according to percent similarity was represented with the different level of percent similarity from low similarity (light blue color) to high similarity (more intense blue color). The dataset included the four clustering results from HRM clusters, sensitive clusters from S-HRM analysis, DTW based clustering and ERIC patterns (91% similarity) as well as general information of the *salmonella* isolates such as serotype, specimens, animals, source and locations of collection.



Strain number	HRM clusters 79-89	Sensitive clusters	DTW clusters	ERIC -PCR Patterns (91.8%)	Antibiotic resistant genotype *bla* TEM	Antibiotic resistant genotype *flo* R	Conventional Serotypes	Specimens	Animals	Sources	Province	Month	Year
L27	H3	S14	DTW27	E1	P	P	*S*. Typhimurium	Organs	Chickens	House	Chiang rai	June	2061
L28			DTW26	E2		N		Intestinal contents	Goat			October	
L17				E3		P					Lampang	May	
L20	H1	S1	DTW4	E4		N	*S*. Weltevreden	Stool sample	Chicken	Farm	Chiang mai	June	
L21				E5		P					Lampang	May	
L19				E6							Chiang mai	June	
L15				E7		N		cecal contents	Swine	Slaughterhouse	Phayao		2060
L39				E8		P		Stool sample	Chicken	Farm	Lampang	July	2061
L40		S3	DTW17	E9		N		Intestinal contents	Cow	House	Lamphun	June	
L33		S1	DTW4	E10	P	P		Stool sample	Chicken	Farm		July	
L34					N	N							
L5		S1	DTW1	E11	P	N		Cecal contents	Swine	Slaughterhouse*	Nan	May	
L13				E12		P		Stool sample	Chicken	Farm	Phrae	June	2060
L35		S2	DTW5	E13		N					Lamphun	July	2061
L31	H13	S24	DTW19	E14			*S*. Kantucky						
L32	H1	S5	DTW6	E15		P	*S*. Corvallis						
L11	H10	S21	DTW18	E16			*S*. Kantucky				Chiang mai	June	2060
L30	H4	S15	DTW28	E17		N	*S*. Stanley				Lamphun	July	2061
L29	H2	S6	DTW12	E18	P	N	*S*. I.4,5,12:i:-	Intestinal contents	Chicken	House	Lampang	August	
L1			DTW11					Cecal contents	Swine	Slaughterhouse	Chiang mai	Feburary	2060
L12		S13	DTW13	E19	P	N	*S*. Braenderup	Stool sample	Chicken	Farm	Phrae	June	
L16		S14	DTW14	E20		P					Chiang mai		2061
L2		S9	DTW7	E21	P	P	*S*. Derby	Cecal contents	Swine	Slaughterhouse*	Nan	May	
L3		S10			P	N		Intestinal contents		House	Chiang rai		
L14	H1	S4	DTW5	E22	P	N	*S*. Agona	Stool sample	Chicken	Farm	Lampang		
L4	H2	S7	DTW10	E23			*S*. I.4,5,12:i:-	Cecal contents	Swine	Slaughterhouse	Chiang mai	Feburary	2060
L7		S8	DTW8	E24							Lampang	June	
L6		S11	DTW9	E25		P	*S*. Derby				Nan	May	2061
L10	H1	S4	DTW5	E26		N	*S*. Agona		Chicken		Chiang mai	June	
L8	H8	S19	DTW3	E27			*S*. Krefeld	Intestinal contents	Swine	House	Lampang	December	2060
L9	H9	S20	DTW24	E28		P	*S*. Give	Cecal contents		Slaughterhouse		Feburary	
L36	H5	S16	DTW25	E29		N	*S*. Stanley	Stool sample	Chicken	Farm	Lamphun	July	2061
L38	H14	S25	DTW15	E30		P	*S*. Albany						
L37	H7	S18	DTW23	E31			*S*. Enteritidis						
L18	H11	S22	DTW21	E32	N		*S*. Corvallis				Phayao		
L25	H1	S1	DTW5	E33	P	N	*S*. Weltevreden	Organs		House	Chiang mai	August	
L26		S5	DTW2	E34			*S*. Corvallis	Carcess				April	
L22	H4	P15	DTW28	E35	N		*S*. Stanley	Stool sample		Farm	Lampang	May	
L23	H12	P23	DTW20	E36		P	*S*. Uganda	Organs	Rat	House	Nan		
L24	H6	P17	DTW22	E37	P	N	*S*. Enteritidis	Meat	Chicken		Chiang rai	August	

Table 1(on next page)

Primers used in this study

Primer	Genes	Sequence (5' ⟶ 3')	Size of PCR-product (bps)	Primer Concentration (pmol/ul)	Reference
HRM Multiplex *fljB,, gyrB* and *ycfQ* genes (HRM-rt PCR)					
fljB_f	*fljB*	GTGAAAGATACAGCAGTAACAACG	170	0.1	(Zeinzinger et al., 2012)
fljB_r		CAAAGTACTTGTTATTATCTGCG		0.1	
gyrB_f	*gyrB*	AAACGCCGATCCACCCGA	171	0.075	(Zeinzinger et al., 2012)
gyrB_r		TCATCGCCGCACGGAAG		0.075	
ycfQ_f	*ycfQ*	GCCTACTCTCTATGCGGAATTCAC	241	0.075	(Zeinzinger et al., 2012)
ycfQ_r		GATATCGCGCGAGGAGGCG		0.075	
Multiplex 1 *bla*TEM and *flo*R					
blaTEM_f	*bla*TEM	CAGCGGTAAGATCCTTGAGA	323	0.15	(Singh & Mustapha, 2014)
blaTEM_r		TTACATGATCCCCCATGTTG		0.15	
Chl F	*flo*R	GGCAGGCGATATTCATTACT	197	0.12	(Singh & Mustapha, 2014)
Chl R		CGAGAAGAAGACGAAGAAGG		0.12	
Molecular typing					
ERIC_f	ERIC-PCR	ATGTAAGCTCCTGGGGATTCAC		25	(Versalovic, Koeuth & Lupski, 1991)
ERIC_r		AAGTAAGTGACTGGGGTGAGCG		25	

[b]Y=T or C; R=A or G; S=G or C; D=A or G or T

1

2

3

4

Table 2(on next page)

HRM serotyping of 40 *Salmonella* isolates from different specimens, animals, sources and provinces during February 2018 to September 2019.

The HRM clusters derived from automatically and manually clustering using the Precision Melt Analysis software V 1.2.

Isolate no.	Isolate name	HRM Serotyping			
		Tm Peak	Clustering	HRM clusters	Percent confidence
1	L5	1(87.5)	Auto	Cluster 1	99.5
2	L13	1(87.3)	Auto		97.5
3	L15	1(87.5)	Auto		99.5
4	L19	1(87.5)	Auto		99.6
5	L20	1(87.3)	Auto		99.4
6	L21	1(87.4)	Auto		99.0
7	L25	1(87.3)	Auto		98.5
8	L33	1(87.4)	Auto		98.3
9	L34	1(87.3)	Auto		99.2
10	L39	1(87.4)	Auto		99.1
11	L35	1(87.3)	Auto		98.0
12	L40	1(87.3)	Manual		92.0
13	L14	1(87.4)	Auto		99.0
14	L10	1(87.4)	Auto		98.3
15	L26	1(87.3)	Auto		98.5
16	L32	1(87.3)	Auto		97.8
17	L1	1(87.2)	Manual	Cluster 2	77.4
18	L29	1(86.9)	Manual		71.6
19	L4	1(87.0)	Manual		54.1
20	L7	1(87.0)	Auto		94.6
21	L2	1(87.3)	Auto		82.0
22	L3	1(87.2)	Auto		90.2
23	L6	1(87.3)	Manual		95.1
24	L12	1(87.2)	Manual		75.0
25	L16	1(87.3)	Manual		51.3
26	L17	2(87.1,83.2)	Auto	H3	98.1
27	L27	2(86.9,83.0)	Auto		99.8
28	L28	2(87.0,83.1)	Auto		99.4
29	L22	2(87.4,83.2)	Auto	H4	98.4
30	L30	2(87.3,83.1)	Auto		98.8
31	L36	1(87.3)	Auto	H5	99.6
32	L24	2(87.5,86.1)	Auto	H6	99.3
33	L37	2(87.3,86.0)	Auto	H7	99.8
34	L8	1(87.4)	Auto	H8	98.1
35	L9	2(87.6,82.5)	Auto	H9	98.5
36	L11	1(87.0)	Manual	H10	97.5
37	L18	1(87.2)	Auto	H11	97.0
38	L23	1(87.4)	Auto	H12	98.7
39	L31	1(86.9)	Auto	H13	89.3
40	L38	1(87.2)	Auto	H14	97.0
Reference		1(87.4)	Auto	REF	99.6

1

Table 3(on next page)

S-HRMa of 40 *Salmonella* isolates from different specimens, animals, sources and provinces during February 2018 to September 2019.

The Sensitive clusters was further assigned to all *Salmonella* isolates especially those in cluster 1 and 2 using serogroups and S-HRMa. The molecular analyzation of two antibiotic resistant genes, their conventional serotypes, and ERIC clusters (91%) were conducted and compared in this study.

Clustering of *Salmonella* isolates (n=40)									
Strain number	**HRM clusters**	**Serogroup**	**S-HRM analysis**		**Sensitive clusters**	**Antibiotic resistent genes**		**Conventional Serotypes**	**ERIC clusters (91%)**
	79-89 °C		**79 - 83 °C**	**85 - 88 °C**		***bla*TEM**	***flo*R**		
L5	**Cluster 1**	1E	ND	D	S1	P	N	S. Weltevreden	E1
L13							P	S. Weltevreden	E2
L15							N	S. Weltevreden	E3
L20								S. Weltevreden	E4
L21							P	S. Weltevreden	E5
L33								S. Weltevreden	E6
L34						N	N	S. Weltevreden	
L39						P	P	S. Weltevreden	E7
L19								S. Weltevreden	E8
L25							N	S. Weltevreden	E9
L35				E	S2	N		S. Weltevreden	E10
L40				F	S3	P		S. Weltevreden	E11
L10		1B		D	S4			S. Agona	E12
L14								S. Agona	E13
L26		1C			S5			S. Corvallis	E14
L32							P	S. Corvallis	E15
L1	**Cluster 2**	2B	A	G	S6		N	S .I.4,5,12:i:-	E16
L29								S .I.4,5,12:i:-	
L4				I	S7			S .I.4,5,12:i:-	E17
L7				K	S8			S .I.4,5,12:i:-	E18
L2				H	S9		P	S. Derby	E19
L3			B		S10		N	S. Derby	
L6				J	S11		P	S. Derby	E20
L12		2C	C	L	S12		N	S. Braenderup	E21

L16			REF	REF	S13		P	S. Braenderup	E22
L17	H3	3B	ND		S14			S. Typhimurium	E23
L27								S. Typhimurium	E24
L28							N	S. Typhimurium	E25
L22	H4	4B			S15			S. Stanley	E26
L30								S. Stanley	E27
L36	H5	5B			S16	N		S. Stanley	E28
L24	H6	6D			S17	P		S. Enteritidis	E29
L37	H7	7D			S18		P	S. Enteritidis	E30
L8	H8	8E			S19		N	S. Krefeld	E31
L9	H9	9E			S20		P	S. Give	E32
L11	H10	10C			S21			S. Kantucky	E33
L18	H11	11C			S22	N	N	S. Corvallis	E34
L23	H12	12E			S23	P	P	S. Uganda	E35
L31	H13	13C			S24		N	S. Kantucky	E36
L38	H14	14E			S25		P	S. Albany	E37
Discriminatory power	**0.7949**	**0.8795**	ND		**0.9603**	**ND**	**ND**	**0.891**	**0.9962**
Number of types	**14**	**17**			**25**			**14**	**37**

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27664v1 | CC BY 4.0 Open Access | rec: 19 Apr 2019, publ: 19 Apr 2019

Table 4(on next page)

The associated epidemiological data of 40 *Salmonella* isolates derived from DTW and S-HRMa

The associated data of *Salmonella* isolates was classified to four categories and subcategories as provinces/types displayed in the first two columns on the left and the next fourteen column displayed the corresponded *Salmonella* serotypes in this study. The total number of *Salmonella* isolates displaying each associated data was indicated in the outermost, left column while the total number of *Salmonella* isolates in each serotype was provided at the row labeled in orange color. Two *Salmonella* isolates showing high similarity (>91%) or in the same ERIC clusters were in the yellow column.

Cartigories	Provinces/Types	Salmonella serotypes																	
		Albany	Corvallis	Derby		Enteritidis	Kantucky	Agona	Braenderup	Give	I.4,5,12:i:-		Krefeld	Typhimurium	Uganda	Weltevreden		Stanley	Total
Provinces	Chiang mai		L26				L11	L10	L16		**L1**	L4					L19,20,25		**9**
	Chiang rai			**L3**		L24								L27, 28					**4**
	Lampang							L14		L9	**L29**	L7	L8	L17			L21,39	L22	**9**
	Lamphun	L38	L32			L37	L31									**L33,L34**	L35,40	L36,30	**10**
	Nan			**L2**	L6										L23		L5		**4**
	Phayao		L18														L15		**2**
	Phrae								L12								L13		**2**
	Total	**1**	**3**	**2**	**1**	**2**	**2**	**2**	**2**	**1**	**2**	**2**	**1**	**3**	**1**	**2**	**10**	**3**	**40**
Sources	Slaughter house			**L2**	L6			L10		L9	**L1**	L4 L7					L5,15		**9**
	Farm	L38	L32,18			L37	L11,31	L14	L12,16							**L33,L34**	L13,19,20,21,39,35	L22,36,30	**20**
	House		L26	**L3**		L24					**L29**		L8	L17,27,28	L23		L25,40		**11**
Animals	Chicken	L38	L32,18,26			L24,37	L11,31	L10,14	L12,16	L9	**L29**			L27		**L33,L34**	L13,19,20,21,39,25,35	L22,36,30	**27**
	Swine			**L3**	**L2**	L6					**L1**	L4 L7	L8				L5,15		**9**
	Goat													L27, 28					**2**
	Rat														L23				**1**
	Cow																L40		**1**
Samples	Stool sample	L38	L32,18			L37	L11,31	L10,14	L12,16							**L33,L34**	L13,19,20,21,39,35	L22,36,30	**21**
	Intestinal contents			**L3**							**L29**		L8	L17, 28			L40		**6**
	Cecal contents			**L2**	L6					L9	**L1**	L4 L7					L5,15		**8**
	Carcess		L26																**1**
	Organs													L27	L23		L25		**3**
	Meat					L24													**1**

1