International Journal of Food Microbiology 254 (2017) 18–24

Contents lists available at ScienceDirect

International Journal of Food Microbiology

journal homepage: www.elsevier.com/locate/ijfoodmicro

Determination of regional relationships among *Salmonella* spp. isolated from retail pork circulating in the Chiang Mai municipality area using a WGS data approach

Prapas Patchanee, Thanaporn Eiamsam-ang, Juntakarn Vanaseang, Phacharaporn Boonkhot, Pakpoom Tadee*

Integrative research center for Veterinary Preventive Medicine, Faculty of Veterinary Medicine, Chiang Mai University, Chiang Mai, Thailand, 50100

ARTICLE INFO

Keywords:
Salmonella spp.
Virulence genes
Pork
Multilocus sequence typing
Whole genome sequencing data

ABSTRACT

Salmonella is recognized as a significant zoonotic foodborne pathogen, and pork products are involved in one-fifth of infections. Whole genome sequencing data of *Salmonella* isolated from retail's pork circulating in the Chiang Mai Municipality area between April 2013 and September 2014, were used to focus on genetic diversity and proven in pig-human transmission based on Multilocus Sequence Typing (MLST). Additionally, WGS data were used to investigate virulence genes, to assess the hazard or pathogenic potential transferred into the food production chain. In this study, all 32 *Salmonella* strains were classified into 11 Sequence Types (STs). ST469 accounted for the majority (41%). The sequence types of two other strains, 6% of the total, could not be identified. All tested strains carried at least 15 virulence genes. The most frequent gene profile was "*sfm-fim-sop-inv.-org-sip-spa-sif-fli-flg-hil-spr-ssa-sse-pag-bss*" (47%). *Salmonella* circulating in the study area demonstrated competence in biofilm production, host cell adhesion, host cell invasion, and host cell survival. Based on the phenotypic and genotypic findings, as well as pathogen source, it appears possible that a common supply chain or common infection source might be presented in the retail pork system in the study area. In addition, an epidemiological comparison of the *Salmonella* genotypes from the current study with those from other areas such as People's Republic of China (PR China) and the Lao People's Democratic Republic (Lao PDR) was generated by Minimum spanning tree (MST). Identical strains originating from humans, animals and food were found. The findings indicate that contamination can be occured at all levels including pre-harvest, the farm-slaughterhouse-retail chain and consumers over different geographical areas. Acquiring information about infection sources and transmission routes will hopefully motivate all sectors to enforce strict sanitation controls at all production stages including the consumer level.

1. Introduction

Salmonella is recognized as important zoonotic pathogen which causes acute food-borne disease in humans (Hendriksen et al., 2008; Campioni et al., 2012; Van Hoek et al., 2012). Approximately 80 million cases are reported globally each year (Rostagno & Callaway, 2012). In Thailand, the annual surveillance summary for 2015 by the Bureau of Epidemiology indicated *Salmonella* is the most frequently detected pathogen involving for hospitalized patients from 77 provinces (Bureau of Epidemiology, 2015). Products of farm animal origin such as contaminated eggs, adulterated dairy products, as well as raw or under-cooked meats, are considered as the most important source of the disease. Pork products have been reported to be concerned in one-fifth of cases in some areas (Mürmann et al., 2009). In Chiang Mai, Thailand, several studies of *Salmonella* prevalence in pork samples have been conducted. Padungtod & Kaneene (2006) and Sanguankiat et al. (2010) reported the prevalence of *Salmonella* isolated from retail pork to be 29% and 35%, respectively. More recently, a related study found a prevalence of 42% (Patchanee et al., 2016). These findings suggested that *Salmonella* risk has been continued to be an important public health issue in the area for a decade or more, and cannot be eradicated certainly. In the pig production chain, the "farm-slaughterhouse-retail" transmission route of *Salmonella* contamination has been documented using the DNA-fingerprinting molecular technique Pulse Field Gel Electrophoresis (PFGE) (Tadee et al., 2015; Patchanee et al., 2016). Although the regional epidemiological knowledge is expanding, the available data are as yet insufficient to reach conclusions regarding disease transmission patterns related to different hosts, genetic micro-

* Corresponding author.
E-mail address: td.pakpoom@gmail.com (P. Tadee).

http://dx.doi.org/10.1016/j.ijfoodmicro.2017.05.006
Received 1 February 2017; Received in revised form 8 May 2017; Accepted 9 May 2017
Available online 10 May 2017

evolution, or virulence of pathogens transmitted in food production.

Pathogen identification is an essential part of epidemiological surveillance and outbreak investigation. Molecular techniques are important tools which can be used to erase doubt in areas that have not yet been clearly determined by previous studies (Maslow et al., 1993; Foxman, 2012). Multilocus Sequence Typing (MLST) is a technique that utilizes the allelic variations in nucleotide sequences of numerous housekeeping gene loci to distinguish pathogen strains (Mirajkar et al., 2014). This characterization method overcomes some of the limitations associated with gel-based techniques as PFGE, results from different laboratories cannot easily be compared. MLST has been demonstrated to be very useful in global epidemiological studies of numerous infectious agents (Urwin & Maiden, 2003; Harbottle et al., 2006; Shi et al., 2015).

Salmonella pathogenicity has been associated with the presence of virulence genes (Ochman et al., 2000; Pallen et al., 2007). Those genes are enhanced for host cell invasion, survival ability and proliferation capability of the pathogen (Arunava et al., 2012). Severity of infection in human salmonellosis is determined by each of the virulence gene's pathogenic potential.

The polymerase chain reaction (PCR) technique is the most common method for the gene detection (Gao et al., 2016). Both MLST and virulence gene recognition are PCR reaction based methods. However, they require a large number of specific primers which are desired in various conditions of DNA amplification and provide a lower throughput compared with other techniques such as whole genome sequencing (WGS). WGS is a new molecular technique that might break other molecular methods (Enright & Spratt, 1999; Pérez-Losada et al., 2013). “Typing” and “exploring inside” with clear and precise results are also delivered by WGS (McClelland et al., 2001; Chiu et al., 2005). Accordingly, whole genome sequencing data was applied in this study.

This study attempted to obtain whole genome sequencing data for *Salmonella* isolated from retail pork circulating in the Chiang Mai Municipality area between April 2013 and September 2014 (from the study by Patchanee et al., 2016) with a focus on genetic diversity and demonstrating pig-human transmission based on a Multilocus Sequence Typing (MLST) database. Additionally, the virulence gene investigation was conducted from WGS data to assess the pathogenic potential transferred by the food production chain and to provide advice regarding implementation of disease control measures for the study area.

2. Materials and methods

2.1. Salmonella strains

A total of 32 *Salmonella* strains identifying in the research by Patchanee et al. (2016) were tested in this study. Those strains were isolated from retail pork circulating in the Chiang Mai Municipality area during the period April 2013 through September 2014. Separating each serotype found the following: *Salmonella* Rissen ($n = 11$), *Salmonella* group I 4, 5, 12: I: - ($n = 3$), *Salmonella* Anatum ($n = 3$), *Salmonella* Kedougou ($n = 3$), *Salmonella* Krefeld ($n = 2$), *Salmonella* Weltevreden ($n = 2$), *Salmonella* Give ($n = 2$), *Salmonella* Corvallis ($n = 2$), *Salmonella* Newport ($n = 1$), *Salmonella* Lexington ($n = 1$), *Salmonella* Agona ($n = 1$) and *Salmonella* Yoruba ($n = 1$). All strains were examined using whole genome sequencing (WGS) at Shepperd Laboratory, Swansea University, United Kingdom.

2.2. Whole genome sequencing

DNA of all 32 *Salmonella* strains tested were extracted using QIAamp DNA mini kit (Qiagen, Crawley, UK). Afterward, the DNA were sequenced as the short reads by an Illumina MiSeq sequencer (Illumina, Cambridge UK). *Salmonella* Typhimurium str. LT2 chromosome [complete genome accession number: NC_003197] was used as the reference strain of DNA assembly. Finally, high coverage short reads were assembled via SPAdes software following Bankevich et al. (2012).

2.3. Multilocus sequence typing (MLST) based on whole genome sequencing data

Whole genome sequencing data of 32 *Salmonella* strains digested in 7 loci of housekeeping genes, including *aroC* (chorismate synthase), *dnaN* (DNA polymerase III subunit beta), *hemD* (uroporphyrinogenIII cosynthase), *hisD* (histidinol dehydrogenase), *purE* (phosphoribosy laminomidazole), *sucA* (alpha-ketoglutarate dehydrogenase) and *thrA* (aspartokinase I/homoserine dehydrogenase) were submitted for allelic number determination to the Center for Genomic Epidemiology (https://cge.cbs.dtu.dk/services/MLST/). Finally, sets of 7 allelic numbers from each *Salmonella* strain were submitted to the *Salmonella enterica* MLST Database, Warwick Medical School, The University of Warwick (http://mlst.warwick.ac.uk/mlst/dbs/Senterica) to obtain the Sequence Type (ST).

2.4. Virulence genes investigation based on whole genome sequencing data

Whole genome sequencing data of all strains were brought into virulence genes exploring the Rapid annotation serveration server: RAST (http://rast.nmpdr.org/). Twenty-one virulence genes were investigated in this study, including adhesion effectors (*sef*, *pef*, *sfm* and *fim*), invasion effectors (*sop*, *inv.*, *org*, *sip*, *spa*, *sif*, *fli*, *flg*, *hil* and *spr*), host cell survival effectors (*ssa*, *sse*, *prg* and *pag*), virulence plasmid gene (*spv*), enterotoxin-producing gene (*stn*) and biofilm regulator (*bss*).

2.5. Statistical analysis

The frequencies of each ST, virulence genes identified and virulence gene profiles of tested strains were obtained using Epi Info™7.

2.6. Discriminatory index

Simpson's diversity index was used to evaluate the discriminatory power of serotyping, MLST and virulence profiles using an online tool at “Comparing partitions: Diversity and partition congruence coefficients calculation” (http://comparingpartitions.info). This process uses the following formula (Hunter & Gaston, 1988):

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

where D is the Simpson's diversity index, N represents the overall sample population, S denotes the total number of each type and n_j is the number of strains fitting into each type. This index approximates the possibility that any two consecutively sampled strains from a sample population will be placed into a different group. The index computes values in a range of “0” to “1”, representing “definitely not diversity” and “infinite diversity”, respectively.

2.7. Minimum spanning tree (MST) analysis

MST analysis was used in this study. The process is based on evaluation of MLST characteristics for each locus by Bionumerics® software version 7.2 from “advanced cluster analysis for categorical data” command (Applied Maths, Sint-Martens Latem, Belgium). The relatedness among STs in terms of geographical area can be determined by this analysis. The STs which were closely related in loci characteristics are displayed close together in the chart. All of the 32 *Salmonella* strains acquired during the current study and an additional 528 *Salmonella* strains originating from surrounding areas during the period 2000 through 2016 were submitted on MLST databases and analyzed. Distribution of the 528 strains included those derived from Thailand ($n = 40$), P.R. China ($n = 477$) and Lao PDR ($n = 11$). Splitting in each