Infection, Genetics and Evolution 62 (2018) 109–121

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

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

Research paper

Genomic characterization of endemic *Salmonella enterica* serovar Typhimurium and *Salmonella enterica* serovar I 4,[5],12:i:- isolated in Malaysia

Soo Tein Ngoi, Kien-Pong Yap, Kwai Lin Thong*

Institute of Biological Sciences, University of Malaya, 50603 Kuala Lumpur, Malaysia

ARTICLE INFO

Keywords:
Comparative genomics
Multidrug resistant
Phylogenetic
Sequence type
Virulence
Whole genome sequencing

ABSTRACT

Salmonella enterica serovar Typhimurium (*S.* Typhimurium) and the monophasic variant *Salmonella* I 4,[5],12:i:- are two clinically-important non-typhoidal *Salmonella* serovars worldwide. However, the genomic information of these two organisms, especially the monophasic variant, is still lacking in Malaysia. The objective of the study was to compare the genomic features of a monophasic variant and two endemic *S.* Typhimurium strains isolated from humans. All three strains were subjected to whole genome sequencing followed by comparative genomic and phylogenetic analyses. Extensive genomic deletion in the *fljAB* operon (from STM2757 to *iro*B) is responsible for the monophasic phenotype of STM032/04. The two *S.* Typhimurium genomes (STM001/70 and STM057/05) were essentially identical, despite being isolated 35 years apart. All three strains were of sequence type ST19. Both *S.* Typhimurium genomes shared unique prophage regions not identified in the monophasic STM032/04 genome. Core genome phylogenetic analyses showed that the monophasic STM032/04 was closely-related to the *S.* Typhimurium LT2, forming a distinctive clade separated from the two endemic *S.* Typhimurium strains in Malaysia. The presence of serovar Typhimurium-specific *mdh* gene, conserved Gifsy and Fels-1 prophages, and the close genomic resemblance with *S.* Typhimurium LT2 suggested that the monophasic STM032/04 was originated from an LT2-like *S.* Typhimurium ancestor in Malaysia, following an evolutionary path different from the *S.* Typhimurium strains. In conclusion, the monophasic *Salmonella* I 4,[5],12:i:- and the *S.* Typhimurium strains isolated in Malaysia descended from different phylogenetic lineages. The high genomic resemblance between the two *S.* Typhimurium strains isolated for at least 35 years apart indicated their successful evolutionary lineage. The identification of multiple virulence and antimicrobial resistance determinants in the *Salmonella* I 4,[5],12:i:- and *S.* Typhimurium genomes explained the pathogenic nature of the organisms.

1. Introduction

Salmonella enterica serovar 4,[5],12:i:- (*Salmonella* I 4,[5],12:i:-) is known as the monophasic variant of *Salmonella enterica* serovar Typhimurium (*S.* Typhimurium; antigenic formula 4,[5],12:i:1,2) (Echeita et al., 2001). Since the emergence of this organism in the 1990's, its population had rapidly expanded and became one of the most prevalent non-typhoidal *Salmonella* serovars that cause human infections worldwide (Echeita et al., 1999; Switt et al., 2009). During recent years, *Salmonella* I 4,[5],12:i:- has become the fifth most frequently reported serovar associated with clinical cases in the United States of America (US) (CDC, 2017). The US Centers for Disease Control and Prevention (CDC) have documented an approximately 194% increase in the reported salmonellosis cases that were caused by *Salmonella* I 4,[5],12:i:-, over a ten-year period from 2005 to 2015 (CDC, 2017). In the European Union member states, the monophasic *Salmonella* I 4,[5],12:i:- accounted for approximately 8% of total human salmonellosis cases reported yearly, from 2014 to 2016 and was ranked third among the most reported *Salmonella* serovars in Europe (EFSA, 2017). The emergence of the monophasic *Salmonella* I 4,[5],12:i:- in Asian countries has occurred within the same period as in the European region. The isolation of this organism first reported in Thailand in 1993 (Boonmar et al., 1998). Since then, the occurrence of the monophasic *Salmonella* I 4,[5],12:i:- has been widely documented in Asian countries, and has become one of the most commonly serotyped human *Salmonella* isolates in this region (Bangtrakulnonth et al., 2004; Hendriksen et al., 2011; Yang et al., 2015). The isolation of *Salmonella* I 4,[5],12:i:- from zoonotic sources is mainly associated with pigs and pork products worldwide, in addition to foods from cattle and poultry (Hamilton et al., 2015; Hopkins et al., 2010; Mandilara et al., 2013;

* Corresponding author.
E-mail address: thongkl@um.edu.my (K.L. Thong).

https://doi.org/10.1016/j.meegid.2018.04.027
Received 10 February 2018; Received in revised form 3 April 2018; Accepted 19 April 2018
Available online 21 April 2018

Pornruangwong et al., 2008; Raguenaud et al., 2010; Yang et al., 2015).

Salmonella I 4,[5],12:i:- is genetically similar to the biphasic *S.* Typhimurium, and is thought to be a monophasic variant of *S.* Typhimurium as the two organisms were often indistinguishable by most of the identification methods, including conventional serotyping (Guerra et al., 2000; Herrera-León et al., 2007). Besides, *Salmonella* I 4,[5],12:i:- harbors the *mdh* gene which is specific to serovar Typhimurium (Amavisit et al., 2005; Hopkins et al., 2010). Another supporting evidence of *Salmonella* I 4,[5],12:i:- being a monophasic variant of *S.* Typhimurium is the presence of an *IS*200 fragment in the Typhimurium-specific location within the *fli*B-*fli*A intergenic region (Echeita et al., 2001). Moreover, the monophasic variant is often indistinguishable from *S.* Typhimurium by pulsed-field gel electrophoresis (PFGE), producing highly similar and even identical pulsotypes (Amavisit et al., 2005; Zamperini et al., 2007).

Salmonella I 4,[5],12:i:- strains isolated from the same geographical region often fall into a single genetic lineage, or at least share a very close common ancestry (Guerra et al., 2000; De la Torre et al., 2003). An earlier study showed that the US and Spanish *Salmonella* I 4,[5],12:i:- strains constituted two distinct clones based on the different genomic deletion patterns surrounding the *flj*AB genes (Soyer et al., 2009). Furthermore, a recent study in China has identified unique genomic deletion patterns in the monophasic *Salmonella* I 4,[5],12:i:- isolated from food products in this country (Yang et al., 2015). These evidences showed that independent evolution of multiple successful monophasic clones has occurred after the divergence from *S.* Typhimurium ancestors (Soyer et al., 2009). Besides being genetically clonal, the monophasic variant also showed close genetic relationship with *S.* Typhimurium lineages (Guerra et al., 2000). The majority of the *Salmonella* I 4,[5],12:i:- showed phage patterns similar to *S.* Typhimurium U302 (Amavisit et al., 2005; De la Torre et al., 2003; Echeita et al., 1999; Guerra et al., 2000). However, *Salmonella* I 4,[5],12:i:- belonging to other phage types, mainly DT193 and DT120, had also been identified (Barco et al., 2013; Hauser et al., 2010; Hopkins et al., 2012; Hopkins et al., 2010; Mossong et al., 2007).

Generally, antimicrobial resistance patterns of *Salmonella* I 4,[5],12:i:- range from pan-susceptible to multidrug resistant (MDR); *Salmonella* I 4,[5],12:i:- strains in Europe are often MDR, while strains isolated from the US appear pan-susceptible or only resistant to a few antibiotics (Hopkins et al., 2010; Switt et al., 2009). Similar to European strains, the majority of the monophasic strains isolated in the Asian region were MDR (Huoy et al., 2014; Pornruangwong et al., 2008; Yang et al., 2015). Interestingly, the resistance gene clusters of the *Salmonella* I 4,[5],12:i:- do not share a common ancestor with MDR *S.* Typhimurium, as seen when a new resistance island is present (Hopkins et al., 2010). The new resistance island is found inserted at the *flj*AB operon region of the European *Salmonella* I 4,[5],12:i:-; while the MDR phenotype of *S.* Typhimurium is often attributed to the *Salmonella* genomic island 1 (SGI-1) (Hermans et al., 2006; Hopkins et al., 2010). On the other hand, the virulence and pathogenicity genes repertoire of *Salmonella* I 4,[5],12:i:- are highly similar to that of *S.* Typhimurium, with minor variations which could be attributed to the virulence plasmids (Hauser et al., 2010; Yang et al., 2015).

In Malaysia, *S.* Typhimurium is one of the major pathogens causing human salmonellosis. It has been isolated from different sources, including humans, ready-to-eat food, farm animals, domestic animals, and wild animals (Abatcha et al., 2014; Abatcha et al., 2013; Khoo et al., 2015; Modarressi and Thong, 2010; Ngoi et al., 2013). A previous study on the genotypes of the *S.* Typhimurium strains that had circulated in Malaysia from 1970 to 2009 indicated that a genetically homogeneous population of this organism has persisted in this region (Ngoi et al., 2013). Moreover, the multi-drug resistant phenotype of the endemic *S.* Typhimurium has increased, and hence poses a greater risk to public health in this region (Ngoi et al., 2013). Unlike the better studied *S.* Typhimurium population, the isolation and detailed characterization of the *Salmonella* I 4,[5],12:i:- have not being studied before in Malaysia. Hence, the objective of this study was to describe the genomic features of a monophasic variant in comparison with two selected *S.* Typhimurium strains in Malaysia, in order to understand the phylogenetic relationship and pathogenicity of the two organisms.

2. Materials and methods

2.1. Bacterial strains

A monophasic variant of *S.* Typhimurium (STM032/04) and two *S.* Typhimurium strains (STM001/70 and STM057/05) isolated from clinical sources were selected for genomic comparison (Ngoi et al., 2013). In the previous study, a total of 84 *S.* Typhimurium strains (serotyped according to White-Kauffmann-Le Minor scheme) were subjected to PCR serotyping to confirm the presence of *fli*C H:i (phase 1) and *flj*B H:1,2 (phase 2) flagellin genes (Ngoi et al., 2013). Only one strain (STM032/04) isolated from a patient in 2004 showed absence of *flj*B H:1,2 allele. The selected *S.* Typhimurium strains were isolated 35 years apart (STM001/70 in 1970; STM057/05 in 2005), but showed identical *Xba*I-restricted pulsotype and were single locus variants based on multiple locus variable number tandem repeat analysis (Ngoi et al., 2013). The persistence of this pulsotype over the years and its close genetic proximity with most of the other *S.* Typhimurium strains showed that this genotype may represent the endemic *S.* Typhimurium population in Malaysia. All strains were obtained from the strains collection of Laboratory of Biomedical Science and Molecular Microbiology, Institute of Graduate Studies, University of Malaya. A reference strain of *S.* Typhimurium (ATCC13311) that was available in the laboratory strains collection was used in the PCR detection of *flj*AB operon and the flanking region.

2.2. PCR identification of *fljAB* operon and the flanking region

The monophasic STM032/04 was subjected to PCR identification of the *flj*A, *hin*, *iro*B, *iro*B-*iro*C intergenic regions, STM2757, and STM2758 genes. The gene-specific primers used in this study were previously described by Soyer et al. (2009) and García et al. (2013). PCR amplification was done in a 25 μL monoplex reaction mixture, containing 1 × colourless GoTaq Flexi Buffer, 1.5 mmol/L $MgCl_2$, 200 μM dNTP mix, 0.3 μM of each primer pair (0.45 μM/L for *flj*A and *iro*B-5′ primer pairs), 1 U (1.5 U for *flj*A and *iro*B-5′ genes) of *Taq* DNA polymerase (Promega, Madison, USA), and approximately 100 ng of bacterial genomic DNA. The PCR reaction mixtures were first incubated at 95 °C for 5 min; followed by 25 cycles of 95 °C for 1 min, 55 °C for 45 s (62 °C for STM2757 gene), and 72 °C for 1 min; with a final extension step of 72 °C for 10 min.

2.3. Genomic DNA extraction and whole genome sequencing

The genomic DNA of the bacterial strains was extracted using a DNeasy Blood & Tissue Kit (Qiagen, Hilden, Germany) following manufacturer's instructions. The quality of the extracted genomic DNA was then measured using NanoDrop 2000 UV–Vis Spectrophotometer (Thermo Scientific, Wilmington, USA). Whole genome sequencing was performed using Illumina Miseq platform (GA2x, pipeline version 1.80) by a commercial sequencing vendor.

2.4. Genomic assembly, annotation, and comparative genomic analysis

The quality of the sequence reads was assessed and de novo assembly was performed by using the CLC Genomics Workbench version 5.1 (CLC Bio, Aarhus, Denmark). The assembled contigs were then ordered according to the *S.* Typhimurium LT2 reference genome (GenBank accession no.: NC_003197) using Mauve version 2.4.0 (Darling et al., 2010). For each of the assembled genomes, the subset of contigs that could not be mapped to the chromosome of *S.*