Infection, Genetics and Evolution 26 (2014) 146–152

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

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

Salmonella Typhi *shdA*: Pseudogene or allelic variant?

I.M. Urrutia [a,1], J.A. Fuentes [a,1], L.M. Valenzuela [a], A.P. Ortega [a], A.A. Hidalgo [a], G.C. Mora [b,*]

[a] Facultad de Ciencias Biológicas, Universidad Andres Bello, República 217, Santiago de Chile, Chile
[b] Facultad de Medicina, Universidad Andres Bello, República 330, Santiago de Chile, Chile

ARTICLE INFO

Article history:
Received 29 November 2013
Received in revised form 9 May 2014
Accepted 12 May 2014
Available online 21 May 2014

Keywords:
ShdA
Pseudogene
Host-specificity
Fibronectin

ABSTRACT

ShdA from *Salmonella* Typhimurium ($ShdA_{STm}$) is a large outer membrane protein that specifically recognizes and binds to fibronectin. $ShdA_{STm}$ is involved in the colonization of the cecum and the Peyer's patches of terminal ileum in mice. On the other hand, *shdA* gene from *Salmonella* Typhi ($shdA_{STy}$) has been considered a pseudogene (i.e. a nonfunctional sequence of genomic DNA) due to the presence of deletions and mutations that gave rise to premature stop codons. In this work we show that, despite the deletions and mutations, $shdA_{STy}$ is fully functional. *S.* Typhi Δ*shdA* mutants presented an impaired adherence and invasion of HEp-2 pre-treated with TGF-β1, an inducer of fibronectin production. Moreover, *shdA* from *S.* Typhi and *S.* Typhimurium seem to be equivalent since $shdA_{STm}$ restored the adherence and invasion of *S.* Typhi Δ*shdA* mutant to wild type levels. In addition, anti-FLAG mAbs interfered with the adherence and invasion of the *S.* Typhi *shdA*-3xFLAG strain. Finally, $shdA_{STy}$ encodes a detectable protein when heterologously expressed in *Escherichia coli* DH5α.

The data presented here show that $shdA_{STy}$ is not a pseudogene, but a different functional allele compared with $shdA_{STm}$.

1. Introduction

Salmonella enterica subspecies *enterica* includes serovars that commonly cause infections in warm-blooded animals (Baumler, 1997; Boyd et al., 1993; Groisman and Ochman, 1997; Shelobolina et al., 2004). Genome sequences of closely related *S. enterica* serovars share more than 90% identity at the nucleotide level. Despite their remarkable similarity, most serovars differ in their host specificity and disease manifestations (McClelland et al., 2001; Parkhill et al., 2001). Some *S. enterica* serovars, such as *S. enterica* serovar Typhimurium (*S.* Typhimurium) are considered "generalists" because they infect a broad range of hosts. Other serovars are host-restricted, such as *S. enterica* serovar Typhi (*S.* Typhi), a human-restricted pathogen that causes typhoid fever (Barrow and Duchet-Suchaux, 1997; Coburn et al., 2007; Collins, 1974; Parkhill et al., 2001; Parry et al., 2002; Soyer et al., 2009).

The evolution of a broad host pathogen, such as *S.* Typhimurium, to a host-restricted pathogen, such as *S.* Typhi, might have occurred by acquisition of new genes through horizontal transfer, loss of genetic information by deletions or pseudogene formation, or by a combination of these mechanisms (Andersson and Andersson, 1999; Hacker and Carniel, 2001; Moran and Plague, 2004). Pseudogenes are defined as sequences homologous to functional genes that do not encode a functional product. Pseudogenes arise by point mutations and/or deletions/rearrangements that produce nonsense or frameshift mutations in the coding sequence, resulting in a truncated version of the gene (Dagan et al., 2006). It is understandable that mutations would readily accumulate in nonessential genes acquired by horizontal gene transfer, but until recently it was not obvious that such loss of function mutations play an important role in the evolution of bacterial pathogens. Comparative genomics studies, that compare host-restricted pathogens with their host-generalist relatives, indicate that accumulation of pseudogenes is a hallmark of host-restricted pathogenic bacteria (Andersson and Andersson, 2001; McClelland et al., 2004; Parkhill et al., 2003). When genomes were compared across the bacterial domain, pathogens have a higher number of pseudogenes than non-pathogen bacteria (Liu et al., 2004).

In *S. enterica*, CS54 (a pathogenicity island corresponding to SPI-24) harbors the *shdA* gene, annotated as a pseudogene in *S.*

Abbreviations: $shdA_{STm}$, *S.* Typhimurium *shdA* gene; $shdA_{STy}$, *S.* Typhi *shdA* gene.

* Corresponding author. Address: Facultad de Medicina, Universidad Andres Bello, República 330, Santiago de Chile 8370146, Chile. Tel.: +56 02 2 661 82 25.

E-mail addresses: i.urrutia@uandresbello.edu (I.M. Urrutia), jfuentes@unab.cl (J.A. Fuentes), lu.valenzuela.v@uandresbello.edu (L.M. Valenzuela), all.ortega@uandresbello.edu (A.P. Ortega), alejandro.a.hidalgo@gmail.com (A.A. Hidalgo), gmora@unab.cl (G.C. Mora).

[1] Both authors contributed equally to this manuscript.

http://dx.doi.org/10.1016/j.meegid.2014.05.013
1567-1348/

Typhi due to massive deletions and non-sense mutations; therefore, some authors suggested that *shdA* might not be relevant for human restricted *S.* Typhi (Betancor et al., 2012; Deng et al., 2003b; Kingsley et al., 2003). In contrast, *shdA* is fully functional in *S.* Typhimurium since Δ*shdA* mutants exhibited a reduced colonization of Peyer's patches of terminal ileum, cecum, mesenteric lymph nodes and spleen, and a reduced fecal shedding in mice (Kingsley et al., 2003). *S.* Typhimurium ShdA ($ShdA_{STm}$) has been characterized as a large outer membrane protein belonging to the autotransporter family characterized by a passenger domain consisting of two regions: an N-terminal non repeat region and a repeat region constituted by two types of imperfect direct amino acid repeats (named A and B). The A region is repeated three times (A1 to A3), while the B region is repeated 9 times (B1 to B9) (Kingsley et al., 2004) (Fig. 1C). Moreover, ShdA specifically recognizes and binds to fibronectin, a glycoprotein abundantly produced by epithelial intestinal cells after inflammatory damage induced by bacterial colonization (Kingsley et al., 2004, 2002).

In this work, we found that despite the large deletions and the premature stop codons, *shdA* from *S.* Typhi ($shdA_{STy}$) is fully functional and participated in the adherence and invasion to a fibronectin-producing epithelial cell line. Our results show that $shdA_{STy}$ is not a pseudogene, but a different functional allele compared with $shdA_{STm}$. This finding underlines the need for experimental research in order to unequivocally identify a pseudogene.

2. Materials and methods

2.1. Bacterial strains, media and culture conditions

Strain *S.* Typhi STH2370 was obtained from the Infectious Diseases Hospital Lucio Córdova, Chile. *S.* Typhimurium 14028s was obtained from the Instituto de Salud Pública (ISP), Chile. The strains were grown routinely in liquid culture using Luria Bertani (LB) medium (Bacto peptone, 10 g/L; Bacto yeast extract, 5 g/L; NaCl, 5 g/L) at 37 °C, with aeration, or anaerobically by adding an overlay of 500 μl of sterile mineral oil as a barrier to oxygen prior to cell assays with cultured human cells HEp-2. When required, medium was supplemented with kanamycin (Kan; 50 mg/ml), chloramphenicol (Cam; 20 mg/ml), or ampicillin (Amp; 50 mg/ml). Media were solidified by adding agar (15 g/L).

2.2. Bioinformatic analyses

Comparative sequence analyses were made with the *shdA* sequences available in http://www.ncbi.nlm.nih.gov/ (*S.* Typhi strains CT18, Ty2, Ty21a, and P-stx-12; and *S.* Typhimurium strains 14028s, 798, D23580, LT2, SL1344, ST4/74, T000240, U288, UK-1). *S.* Typhi STH2370 *shdA* gene was sequenced at the Pontificia Universidad Católica, Chile. Sequences were analyzed using BLAST

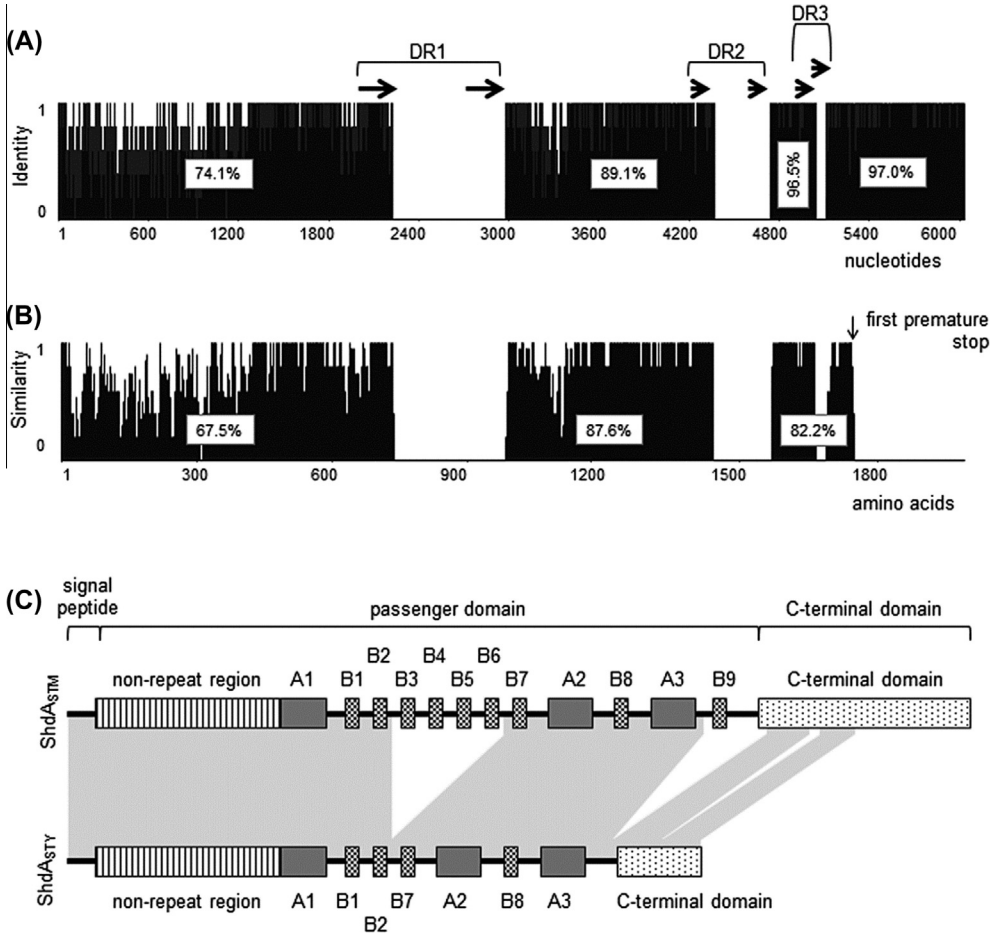

Fig. 1. (A) Nucleotide alignment between *shdA* from *S.* Typhi STH2370 ($shdA_{STy}$) and *shdA* from *S.* Typhimurium 14028s ($shdA_{STm}$). $shdA_{STy}$ presents three large deletions, each of one flanked by different imperfect direct repeats (DR1, DR2, DR3). The percentages correspond to the identity between different gene segments. (B) Predicted protein alignment between $ShdA_{STy}$ and $ShdA_{STm}$. The figure depicts the first premature stop codon. The percentages correspond to the similarity between different protein segments. (C) Regions and domains found in ShdA. $ShdA_{STm}$ presents a passenger domain that can be divided into two regions: an N-terminal non repeat region and a repeat region constituted by two types of imperfect direct amino acid repeats (named A and B). The A region is repeated three times (A1 to A3), while the B region is repeated 9 times (B1 to B9). The $ShdA_{STm}$ domains and regions were taken from Kingsley et al. (2004). The $ShdA_{STy}$ domains and regions were deduced from the comparison with $ShdA_{STm}$.