Infection, Genetics and Evolution 10 (2010) 1132–1139

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

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

Association between phylogeny, virulence potential and serovars of *Salmonella enterica*

Eva Litrup [a,b,*], Mia Torpdahl [a], Burkhard Malorny [b], Stephan Huehn [b], Henrik Christensen [c], Eva M. Nielsen [a]

[a] Statens Serum Institut, Artillerivej 5, DK-2300 Copenhagen S, Denmark
[b] Federal Institute for Risk Assessment, Diedersdorfer Weg 1, D-12277 Berlin, Germany
[c] University of Copenhagen, Faculty of Life Sciences, Bülowsvej 17, DK-1870 Frederiksberg C, Denmark

ARTICLE INFO

Article history:
Received 8 February 2010
Received in revised form 16 July 2010
Accepted 16 July 2010
Available online 23 July 2010

Keywords:
Salmonella
Microarray
MLST
Virulence
Pathogenicity
Phylogeny
Typhimurium
Enteritidis
Derby
Java

ABSTRACT

Salmonella enterica subsp. *enterica* is one of the leading causes of zoonotic food-borne disease worldwide. The consequence of these infections is a serious impact on economics of the society in the form of lost productivity and expenses for medical care. The objective of this study was to analyze the difference in genomic content between selected serovars, especially the content of pathogenicity genes and this was done with a DNA microarray. Furthermore, we investigated the phylogenetic relationship between serovars using multilocus sequence typing (MLST). We chose serovars Typhimurium and Enteritidis as they are responsible for 75% of human infections in Europe. Additionally, we included serovars Derby, Dublin, Saintpaul, 4,5,12:i:-, Java and 4,5,12:b:- which are suspected to have different degrees of virulence to humans. MLST analysis clustered strains according to serovar with the exception of Java and Derby. DNA microarray clustered strains according to serovar and serogroup except for serovar 4,5,12:b:-. Differences in content of pathogenicity related genes between serovars with various host preferences and virulence towards humans were not observed. However, our strains from the supposedly less virulent serovar Derby lacked a combination of genes important for virulence. It might be speculated that other serovars can sustain their pathogenicity lacking one or two of these genes, whereas lack of many virulence genes will result in reduced virulence. A partial lack of concordance between MLST and microarray was found and this can be explained by the underlying data. On one hand, microarray data include highly variable regions which are known to be involved in horizontal gene transfer. On the other hand, MLST data is restricted to seven sequences and disregards contribution of horizontally acquired genes when evaluating evolution. The DNA microarray and MLST analysis complement each other giving a clearer image of evolution of these serovars and, furthermore, a visualization of the horizontally acquired genes.

1. Introduction

Salmonella enterica subsp. *enterica* is one of the leading causes of zoonotic food-borne disease worldwide. The consequence of these infections is a serious impact on economics of the society in the form of lost productivity and expenses for medical care. A cost estimated to be several billion dollars annually in the USA alone (Voetsch et al., 2004).

S. enterica subsp. *enterica* consists of over 1500 serovars (Grimont et al., 2007), some of which are pathogenic to humans and animals. A limited segment of these serovars is responsible for causing disease in humans. In Europe, the top two serovars isolated from humans are serovars Typhimurium and Enteritidis (Anon., 2008). In 2006, these two serovars were responsible for over 75% of the human cases of salmonellosis in Europe (Anon., 2008). The *S. enterica* subsp. *enterica* serovars typically cause diarrhea but some serovars like Dublin are more likely to cross the gut-barrier and go into the blood (Jones et al., 2008). Some serovars have preferences for specific hosts (Rabsch et al., 2002). *Salmonella* serovar Typhi is restricted to humans and serovar Gallinarum is adapted to fowl whereas serovar Typhimurium has a broad host-range.

The phylogenetic relationship within the *S. enterica* subsp. *enterica* has been investigated by multilocus enzyme electrophoresis (Selander et al., 1990), and the serovars within *S. enterica* subsp. *enterica* were originally believed to be clonal with a minimum of recombination events between the serovars. Subsequent analysis of multilocus sequence typing (MLST) data has

* Corresponding author at: Statens Serum Institut, Artillerivej 5, 2300 Copenhagen S, Denmark. Tel.: +45 3268 3457; fax: +45 3268 8238.
E-mail address: evl@ssi.dk (E. Litrup).

1567-1348/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.meegid.2010.07.015

shown a high level of recombination within the subspecies of *S. enterica* (Lan et al., 2009; Falush et al., 2006). Within some serovars like *Salmonella* serovar Derby, several sequence types (STs) have been detected. Serovars which originate from more than one common ancestor are said to be polyphyletic and they display several not closely related STs within one serovar. On the contrary, monophyletic serovars originate from a common ancestor and display a few related STs. Other methods have also been employed trying to infer the phylogeny of *S. enterica*.

By sequencing the gene *mutS* from the *Salmonella* reference B collection (SARB) (Boyd et al., 1993) it was suggested that recombination between *S. enterica* strains was extensive (Brown et al., 2003). The SARB collection was established for studying genetic and phenotypic variation in natural populations and the collection consists of strains from 37 different serovars from subspecies *enterica* (Boyd et al., 1993).

Previous studies used microarrays to perform comparative genomic hybridization (Chan et al., 2003; Chang et al., 2008; Porwollik et al., 2004). These investigations were able to cluster strains by serovar using the microarray data. One study used the SARB collection and was able to detect some polyphyletic serovars as well (Porwollik et al., 2004) suggesting that closely related strains of the same serovar are not always genotypically close.

In this study, we investigated a collection of strains isolated from humans whereas previous studies primarily used strains from the SARB collection (Chan et al., 2003; Porwollik et al., 2004; Scaria et al., 2008). We included strains causing different symptoms of disease, from *Salmonella* serovar Typhimurium, 4,5,12:i:- (a monofasic variant of Typhimurium) and *Salmonella* serovar Enteritidis. These serovars are also detected in production animals and were responsible for more than half of the human salmonellosis cases in Denmark in 2007 (Anon., 2007). We also included *Salmonella* serovar Dublin which is primarily detected in cattle and beef. *Salmonella* serovar Dublin caused 1.6% of the human cases in Denmark in 2007 (Anon., 2007). Dublin infections are more invasive in humans as 32.8% of all clinical human samples in Denmark are from blood compared to a few percent for Enteritidis and Typhimurium (source: The Danish National Registry of Gastrointestinal Infections). Furthermore, *Salmonella* serovar Derby was chosen as representative of a *Salmonella* serovar that is typically associated with the pig production (Anon., 2007). *Salmonella* serovar Saintpaul was chosen as a primarily poultry serovar (Beutlich et al., 2010). *Salmonella* serovar Derby and Saintpaul combined are responsible for 1.4% of human infections in Denmark in 2007, so possibly these serovars are less virulent to humans. Finally, we included strains from an European outbreak caused by *Salmonella* serovar Paratyphi B variant Java (from hereon designated Java) (Denny et al., 2007) and strains from a Danish outbreak caused by serovar 4,5,12:b:- which is a monophasic variant of Java.

The objective of this study is to analyze the difference in genomic content between these serovars especially the content of pathogenicity genes and to investigate the phylogenetic relationship between the serovars. We also wanted to assess the concordance between virulence gene content and MLST data. This was done by analyzing strains with a spotted DNA–DNA microarray and comparing these data with MLST data.

2. Materials and methods

2.1. Selection of strains

The selected human strains were isolated from blood (10) or faeces (58) (Table 1). All blood and faecal samples received at Statens Serum Institute in Denmark are screened for double infection with frequently occurring intestinal pathogens such as *Campylobacter*, *Shigella*, *Yersinia* and others. All strains used in this

Table 1
Strain table with serovar, sequence type and source.

Isolate no.	Serovar	Phagetype	ST[a]	Source	Year
0605F40553	Derby		ST39	F[b]	2006
0512H73451	Derby		ST40	F	2005
0611F38817	Derby		ST40	F	2006
0504M62399	Derby		ST678	F	2005
0312F70705	Derby		ST71	F	2003
0504F31802	Derby		ST683	F	2005
0509S47013	Derby		ST682	F	2005
0111H20328	Dublin		ST10	B[c]	2001
0212T49287	Dublin		ST10	B	2002
0301M48513	Dublin		ST10	B	2003
0301H61773	Dublin		ST10	B	2003
0507M75758	Dublin		ST10	F	2005
0502H10416	Dublin		ST10	F	2005
0608T12756	Dublin		ST10	F	2006
2005-60-1160	Dublin		ST10	Pig	2005
2006-60-251	Dublin		ST10	Pig	2006
0308R5237	Enteritidis	3	ST11	B	2003
0402R5577	Enteritidis	4	ST11	B	2004
0508R6682	Enteritidis	21	ST11	B	2005
0501R6203	Enteritidis	8	ST11	F	2005
0507R6514	Enteritidis	21	ST11	F	2005
0509R6904	Enteritidis	4	ST11	F	2005
0506R6431	Enteritidis	3	ST11	F	2005
2005-60-1693	Enteritidis	4	ST11	Cattle	2005
2006-52-197	Enteritidis	RDNC	ST11	Hedgehog	2006
0503R6275	Enteritidis	8	ST310	B	2005
0802T51587	Java		ST88	F	2008
0804W5725	Java		ST127	F	2008
0712S42188	Java		ST149	F	2007
0801H55151	4,5,12:b:-		ST42	F	2008
0608S4784	4,5,12:b:-		ST423	F	2006
0711R8899	4,5,12:b:-		ST679	F	2007
0711T36026	4,5,12:b:-		ST679	F	2007
0803R9185	4,5,12:b:-		ST679	F	2008
0806R9483	4,5,12:b:-		ST679	F	2008
0805W6813	4,5,12:b:-		ST681	F	2008
0605T1143	Saintpaul		ST27	F	2006
0609M17097	Saintpaul		ST27	F	2006
0512H70784	Saintpaul		ST49	F	2005
0604H21578	Saintpaul		ST49	F	2006
0306T23309	Saintpaul		ST50	B	2003
0410W75689	Saintpaul		ST50	B	2004
0208W18237	Saintpaul		ST50	F	2002
0307T28755	Saintpaul		ST50	F	2003
0702W56738	Saintpaul		ST50	F	2007
0610T32185	Saintpaul		ST344	F	2006
2006-60-685	Saintpaul		ST344	pig	2006
0608T15334	Saintpaul		ST680	F	2006
0112F33212	Typhimurium	12	ST19	F	2001
0111H24126	Typhimurium	104	ST19	F	2001
0110R3988	Typhimurium	104a	ST19	F	2001
0112F28702	Typhimurium	104	ST19	F	2001
0111M12249	Typhimurium	RDNC	ST19	F	2001
0110F7002	Typhimurium	120	ST19	F	2001
0205R4381	Typhimurium	12	ST19	F	2002
0202F44678	Typhimurium	12	ST19	F	2002
0211F40143	Typhimurium	RDNC	ST19	F	2002
0207M72344	Typhimurium	RDNC	ST19	F	2002
0210H31581	Typhimurium	104	ST19	F	2002
0208F10996	Typhimurium	193	ST19	F	2002
0210F37188	Typhimurium	3	ST19	F	2002
0210M16322	Typhimurium	170	ST19	F	2002
0201H32554	Typhimurium	10	ST19	F	2002
0506H32341	Typhimurium	12	ST19	F	2005
0509R6852	Typhimurium	104	ST19	F	2005
0511R7026	Typhimurium	104	ST19	F	2005
0209H16582	Typhimurium	120	ST34	F	2002
0207T9764	Typhimurium	12	ST35	F	2002
0110H11581	Typhimurium	10	ST376	F	2001
0704W66075	4,5,12: i:-	U302	ST19	F	2007
0404R5628	4,5,12: i:-	NT	ST34	F	2004
0607H8557	4,5,12: i:-	193	ST34	F	2006
0601T13497	4,5,12: i:-	110	ST34	F	2006

[a] Sequence type.
[b] Isolated from human faeces.
[c] Isolated from human blood.