Epidemiol. Infect. (2016), **144**, 2430–2439. © Cambridge University Press 2016
doi:10.1017/S0950268816000492

Whole genome sequencing provides insights into the genetic determinants of invasiveness in *Salmonella* Dublin

M. MOHAMMED[1,2]* AND M. CORMICAN[2]

[1] *Faculty of Science and Technology, Anglia Ruskin University, Cambridge, UK*
[2] *School of Medicine, National University of Ireland Galway, Galway, Ireland*

Received 6 August 2015; Final revision 8 February 2016; Accepted 22 February 2016; first published online 21 March 2016

SUMMARY

Salmonella enterica subsp. *enterica* serovar Dublin (*S.* Dublin) is one of the non-typhoidal *Salmonella* (NTS); however, a relatively high proportion of human infections are associated with invasive disease. We applied whole genome sequencing to representative invasive and non-invasive clinical isolates of *S.* Dublin to determine the genomic variations among them and to investigate the underlying genetic determinants associated with invasiveness in *S.* Dublin. Although no particular genomic variation was found to differentiate in invasive and non-invasive isolates four virulence factors were detected within the genome of all isolates including two different type VI secretion systems (T6SS) encoded on two *Salmonella* pathogenicity islands (SPI), including SPI-6 ($T6SS_{SPI\text{-}6}$) and SPI-19 ($T6SS_{SPI\text{-}19}$), an intact lambdoid prophage (Gifsy-2-like prophage) that contributes significantly to the virulence and pathogenesis of *Salmonella* serotypes in addition to a virulence plasmid. These four virulence factors may all contribute to the potential of *S.* Dublin to cause invasive disease in humans.

Key words: Comparative genomics, invasiveness, *Salmonella* Dublin.

INTRODUCTION

Non-typhoidal *Salmonella* (NTS) is a major cause of foodborne disease in humans worldwide. NTS infection is associated with consumption of contaminated food and is typically characterized by self-limiting gastroenteritis. However, some patients with NTS may suffer from bacteraemia. This is more likely in vulnerable groups (very young, very old, immunocompromised) [1]. Certain NTS serovars including *S.* Choleraesuis and *S.* Dublin are associated with higher frequency of occurrence of invasive disease [2]. The invasiveness of *S.* Dublin may be related in part to expression of the Vi (virulence) antigen [3] which is variably expressed in *S.* Dublin and is also commonly found in human-adapted *S.* Typhi.

S. Dublin (serotype 1,9,12,[Vi]:g,p:-) is closely related to *S.* Enteritidis (serotype 1,9,12:g,m:-); however, they belong to relatively distinct sequence types (STs) based on multilocus sequence typing (MLST) [4]. Furthermore, *S.* Dublin and *S.* Enteritidis differ significantly in their epidemiology. *S.* Dublin is associated with cattle and can be transmitted to humans via contact with infected animals or consumption of contaminated milk and milk products. By contrast, *S.* Enteritidis is associated with poultry and is rarely associated with systemic illness in humans. It has been speculated that extensive genome degradation and pseudogene accumulation within the genome of *S.* Dublin is strongly associated with decreasing the ability of *S.* Dublin to infect a broad range of hosts

* Author for correspondence: Dr M. Mohammed, School of Medicine, National University of Ireland – Galway, Galway, Ireland.
(Email: drmanal20@hotmail.com)

https://doi.org/10.1017/S0950268816000492 Published online by Cambridge University Press

and is involved in bacterial adaptation to certain animal hosts (cattle) [5–7].

Furthermore, *S.* Dublin and *S.* Enteritidis differ in their risk of invasive disease. *S.* Dublin is often associated with invasive disease in humans [8]; however, *S.* Enteritidis is rarely associated with systemic illness in humans. In Ireland, one third of *S.* Dublin clinical isolates received at the National *Salmonella, Shigella* and *Listeria* Reference laboratory were isolated from blood and other sterile body fluids. To test whether there is a genomic signal differentiating invasive (blood and other sterile body fluids) and non-invasive (faecal) isolates of *S.* Dublin we determined the whole genome sequencing (WGS) for a set of representative invasive and non-invasive clinical isolates of *S.* Dublin and performed comparative genomics analysis with the closely related *S.* Enteritidis strain P125109 which belongs to phage type 4 (PT4). We also screened the Irish *S.* Dublin isolates for the previously identified pseudogenes [6] that might be involved in host adaptation.

METHODS

Bacterial strains, genomic DNA extraction and genome sequencing

A set of invasive and non-invasive *S.* Dublin clinical isolates from Ireland (isolated during 2010–2013) were selected for WGS. Isolates included 11 isolates from human blood and other normally sterile body sites in addition to five gastroenteritis isolates for comparison. Bacterial isolates were cultured on nutrient agar media and incubated overnight at 37 °C. DNA was extracted using QIAamp DNA Mini kit (Qiagen, USA) according to the manufacturer's instructions. WGS was performed using an Illumina MiSeq (Illumina Inc., USA) on 100 bp paired-end (PE) libraries.

Both invasive and non-invasive isolates were tested for the presence of the Vi antigen by slide agglutination using *Salmonella* Vi antisera.

Sequence data quality control

The quality of PE Illumina sequence data for each isolate was evaluated using the FastQC toolkit (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Adapter sequences were trimmed and low-quality reads were removed using the ea-utils package (https://code.google.com/p/ea-utils/).

Read mapping, single nucleotide polymorphism (SNP) calling and construction of the phylogenetic tree of the *S.* Dublin sample set

Sequence reads from each isolate were mapped against the reference genome of *S.* Dublin strain CT_02021853 along with its associated plasmid (pCT02021853_74) using Burrows– Wheeler Aligner [9]. Genomic variants including SNPs and insertions and deletions (indels) were identified using Samtools mpileup [10] and filtered with a minimum mapping quality of 60. SNPs were compared against the reference *S.* Dublin strain CT_02021853 and a maximum-likelihood phylogeny of the isolates was constructed using MEGA v. 6 software [11]. Selection of the best-fit model for nucleotide substitution was performed with jModelTest [12].

S. Dublin genome assembly

PE reads were *de novo* assembled using Velvet [13]. The parameters (*k*-mer length, expected coverage, coverage cut-off and insert length) were optimized to obtain the highest N50 value and the best possible assembly. Generated multi-contig draft genomes for each isolate were annotated with the help of the rapid annotation using subsystem technology (RAST) system [14].

Comparative genomic analysis and pseudogene analysis

Comparative genomic analysis with the reference genome of *S.* Enteritidis strain P125109 and the draft genomes of *S.* Dublin Irish clinical isolates was performed using the Artemis Comparison Tool (ACT) (https://www.sanger.ac.uk/resources/software/act/) in order to identify genes and genomic regions that are specific to *S.* Enteritidis and absent from *S.* Dublin in addition to genes and genomic regions that are specific to *S.* Dublin and absent from *S.* Enteritidis.

We also screened *S.* Dublin isolates for the seven inactive pseudogenes identified previously by microarray analysis [6] that might be involved in adaptation of *S.* Dublin to cattle.

RESULTS

Phylogenetic relationship between *S.* Dublin isolates

The phylogenetic relationship between the *S.* Dublin clinical isolates based on SNPs determined by mapping sequences against the genome of reference *S.* Dublin strain CT_02021853 showed that SNPs were found to be randomly distributed around the *S.* Dublin chromosome.

https://doi.org/10.1017/S0950268816000492 Published online by Cambridge University Press

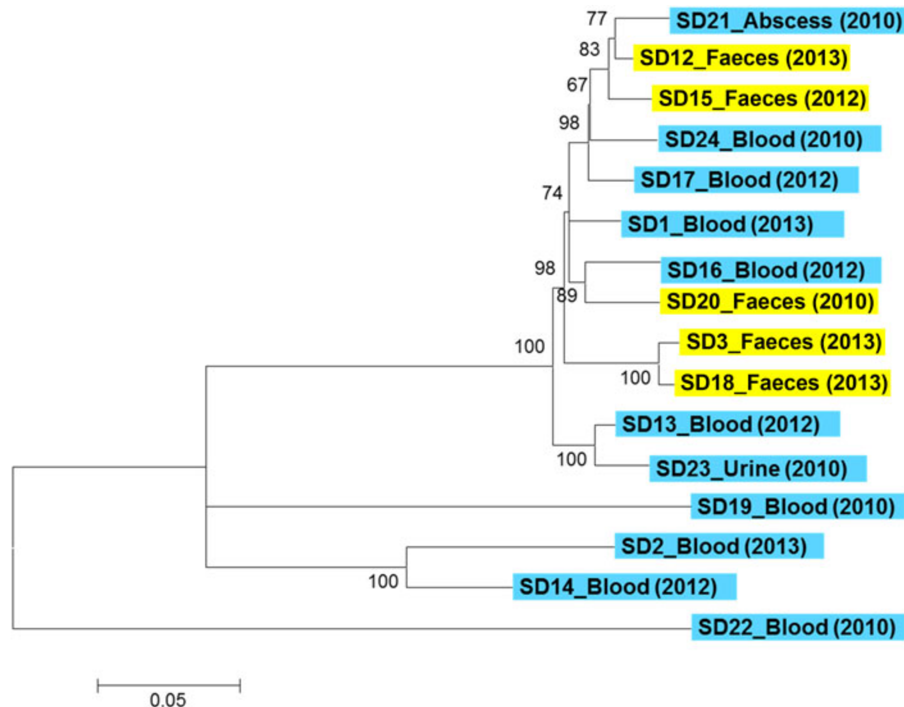

Fig. 1. Phylogenetic tree of *Salmonella* Dublin isolates based on single nucleotide polymorphisms determined from whole genome sequence. Invasive isolates (highlighted in light blue) and faecal isolates (highlighted in yellow) are intermixed. The generalized time-reversible model of nucleotide substitution for each position was used.

Invasive isolates showed higher genetic divergence (5–333 SNPs) compared to the non-invasive isolates (4–30 SNPs); however, invasive and faecal isolates were intermixed (Fig. 1). SNPs within the chromosome of the *S.* Dublin isolates used in this study are provided in Supplementary Table S1.

Specific genes and genomic regions for *S.* Dublin isolates

All *S.* Dublin isolates including invasive and gastroenteritis isolates harbour two different types of type VI secretion systems (T6SS). They are encoded in two *Salmonella* pathogenicity islands (SPI) including SPI-6 ($T6SS_{SPI-6}$, gene range SeD_A0289–SeD_A0326) and SPI-19 ($T6SS_{SPI-19}$, gene range SeD_A1212–SeD_A1243). Interestingly, $T6SS_{SPI-6}$ and $T6SS_{SPI-19}$ share very limited homology with each other.

$T6SS_{SPI-6}$ presents in all *S.* Dublin clinical isolates including invasive and gastroenteritis isolates in addition to other virulent *Salmonella* serovars including *S.* Typhi, *S.* Paratyphi A, *S.* Typhimurium and *S.* Choleraesuis (Table 1, Fig. 2) but it is absent from *S.* Enteritidis PT4.

Two virulence genes within $T6SS_{SPI-6}$; *sciR* (STM0284) and *sciS* (STM0285), encoding the Shiga-like toxin A subunit and putative inner membrane protein, respectively, were present in all *S.* Dublin isolates sequenced in this study.

$T6SS_{SPI-19}$ presents in all *S.* Dublin clinical isolates and other host-restricted serovars including *S.* Gallinarum and *S.* Pullorum (Table 1, Fig. 3). However, SPI-19 is degraded in *S.* Enteritidis PT4.

The draft genome of all *S.* Dublin isolates harbours an intact lambdoid prophage (Gifsy-2-like prophage) as illustrated in Figure 4. Gifsy-2-like prophage is degraded in *S.* Enteritidis PT4. Intact Gifsy-2 is also detected in *S.* Typhimurium and *S.* Newport.

Gifsy-2 prophage encodes several virulence factors including *sodCI* (STM1044), *sseI* (STM1051) and *gtgE* (STM1055) that were found within all *S.* Dublin isolates. Gifsy-2 also encodes *grvA*; a unique avirulence gene in a wild-type situation; however, a non-synonymous mutation detected in *grvA* of *S.* Dublin isolates had changed one amino acid (from polar hydrophilic arginine to non-polar hydrophobic tryptophan) within GrvA protein (Arg53-Trp) that might be associated with loss of its antivirulence function, subsequently increasing bacterial virulence.

Interestingly, the gene encoding the Gifsy-2 prophage attachment and invasion protein (STM1043) was found in all *S.* Dublin Irish strains including

https://doi.org/10.1017/S0950268816000492 Published online by Cambridge University Press

Table 1. *Distribution of T6SS loci in* Salmonella *serotypes*

Serotype	Accession number	$T6SS_{SPI-6}$	$T6SS_{SPI-19}$
Dublin	CP001144	+	+
Enteritidis	AM933172	–	–
Typhimurium*	AE006468	+	–
Typhimurium†	FN424405	+	–
Choleraesuis	CP007639	+	–
Typhi	AL513382	+	–
Paratyphi A	CP000026	+	–
Paratyphi B‡	CP000886	–	–
Paratyphi C	CP000857	+	–
Kentucky	ABAK02000001	+	–
Newport	CP001113	+	–
Heidelberg	CP005995	+	–
Schwarzengrund	CP001127	+	–
Tennessee	CP007505	+	–
Infantis	NZ_CM001274	+	+
Wetervreden	FR775192, FR775207	+	+
Virchow	ABFH02000001, ABFH02000001	–	–
Javiana	CP004027	–	–
Gallinarum	AM933173	–	+
Pullorum	CP006575	–	+
Agona	CP001138	–	+
Montevideo	CP007222	+	–
Anatum	CP007211	+	–
Thompson	CP006717	+	–
Cubana	CP006055	+	–

* *Salmonella* Typhimurium strain LT2 (laboratory adapted strain).
† *Salmonella* Typhimurium strain D23580 (invasive strain).
‡ *Salmonella* Paratyphi B strain SPB7 (d-tartrate-fermenting variant; dT+).

invasive and gastroenteritis isolates; however, it appears to be subject to partial deletion in *S.* Enteritidis PT4.

All *S.* Dublin isolates including invasive and non-invasive isolates harbour an accessory genome that is closely related to the virulence plasmid (pCT02021853_74) of *S.* Dublin strain CT_02021853 with only 14–22 variants difference. This plasmid is exclusive to *S.* Dublin and is not harboured by other *Salmonella* serotypes.

No acquired antimicrobial resistance genes were detected within the plasmid or the chromosome of all clinical isolates sequenced here. However, an isolate from faeces (SD12) and another isolate from a psoas abscess (SD21) were resistant to quinolone antimicrobial agents including nalidixic acid as a result of a single non-synonymous mutation within the housekeeping gene (*gyrA*) that changed the amino acid serine (S) at position 83 to phenylalanine (F) in the DNA gyrase subunit A protein.

The Vi antigen was absent from all *S.* Dublin isolates and the Vi-antigen-encoding genes were absent from the draft genome of *S.* Dublin including invasive isolates. However, the invasion gene (*pagN*) was detected in both invasive and gastroenteritis *S.* Dublin isolates.

Specific genes and genomic regions for *S.* Enteritidis

Forty genes present in *S.* Enteritidis appear to be entirely absent from *S.* Dublin isolates (Table 2). Two genes (SEN2420 and SEN3943) encode a putative exported protein and a conserved hypothetical protein, respectively. The remaining 38 genes were grouped into five genomic regions: SE_Reg1 (SEN0083–SEN0085) encodes two probable secreted proteins and a putative sulfatase, SE_Reg2 (SEN1379–SEN1396) is a part of phage SE14 encoding several phage-related proteins, SE_Reg3 (SEN1432–SEN1436) corresponds to a genomic island named ROD13 and includes genes encoding proteins involved in sugar transport and hexonate uptake, SE_Reg4 (SEN1503–1506) corresponds to another genomic island known as ROD14 encoding regulatory proteins. In addition, another genomic region, SE_Reg5 (SEN1936–SEN1940, SEN1944–SEN1945 and SEN1959), which is absent from *S.* Dublin Irish isolates, encodes several phage-related proteins including terminases and capsid and membrane proteins.

Pseudogene analysis

The draft genomes of the *S.* Dublin clinical isolates were screened for the previously identified seven coding sequences (CDS) that appear to encode active genes in the broad host range serovar *S.* Enteritidis PT4, but are inactive (pseudogenes) in the seven Uruguayan *S.* Dublin clinical isolates [6] as determined by microarray analysis. Interestingly, we found that these seven CDS are also inactive in all of the Irish clinical isolates of *S.* Dublin sequenced in this study and also in the other host-restricted serovars including the poultry-adapted *S.* Gallinarum and the swine-adapted *S.* Cholerasuis (Table 3).

These CDS correspond to a putative transport protein (SEN0042), two hypothetical proteins (SEN0784 and SEN2783), the gene encoding a probable glucarate dehydratase 2 (SEN2806), the gene encoding the outer membrane usher protein LpfC (SEN3461) and the gene

https://doi.org/10.1017/S0950268816000492 Published online by Cambridge University Press

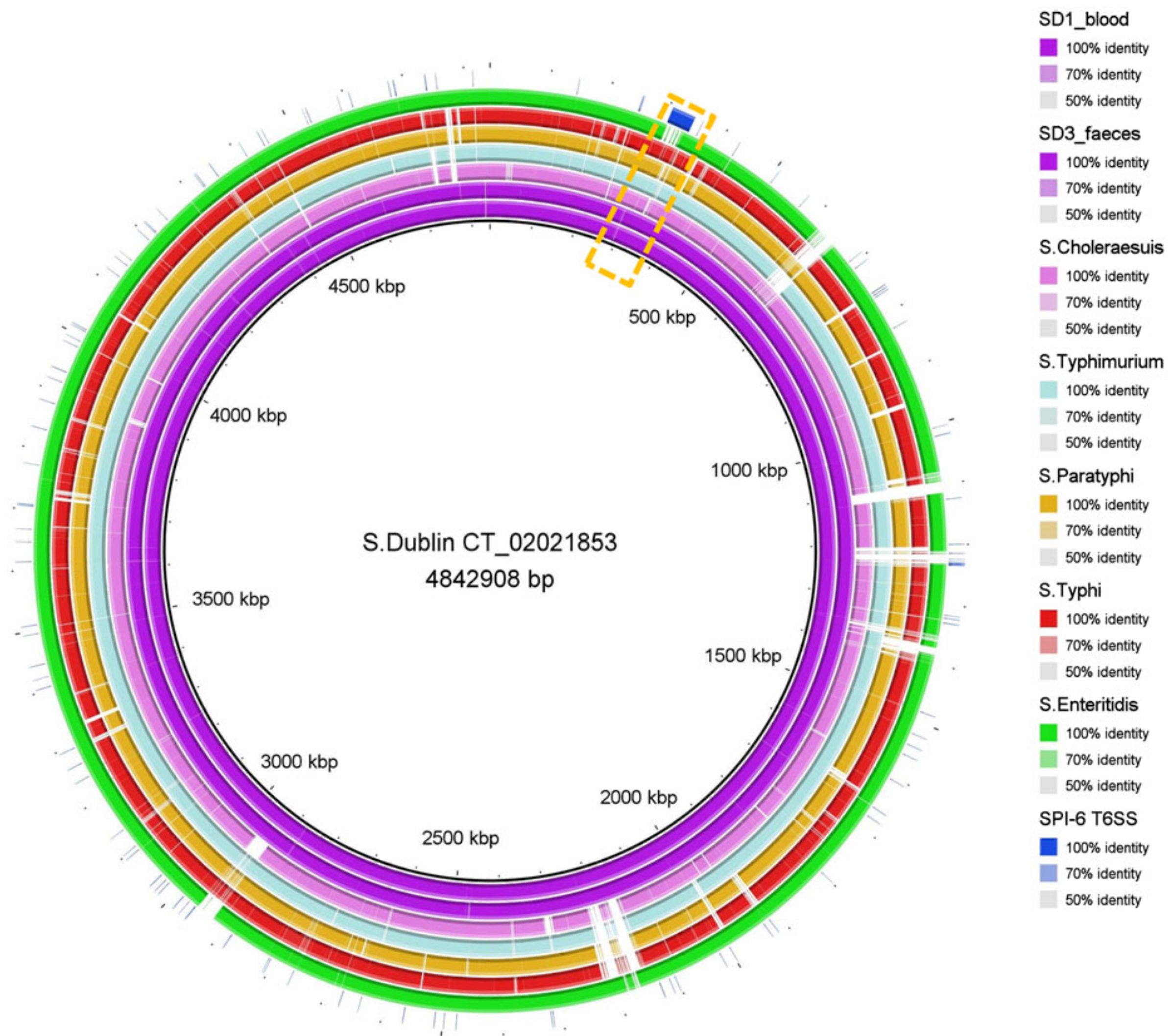

Fig. 2. Complete genome alignment of different *Salmonella* serovars (including the draft genome of representatives of clinical Irish isolates of *S.* Dublin sequenced in this study) generated using BRIG [27]. *S.* Dublin strain CT_02021853 genome is used as a reference. The gene cluster of SPI-6 T6SS (gene range SeD_A0289–SeD_A0326) is also included in the alignment. The genome of clinical *S.* Dublin isolates (including invasive and gastroenteritis isolates) and other invasive serovars including *S.* Choleraesuis, *S.* Typhimurium, *S.* Typhi and *S.* Paratyphi A harbours SPI-6 T6SS while it is absent from *S.* Enteritidis.

encoding probable phosphotransferase system permease (SEN3672). In addition *mglA* (SEN2182) appear to be inactive in all host-adapted typhoidal *Salmonella* serovars including the human-adapted *S.* Typhi and *S.* Paratyphi A (Table 3).

A gene of special interest in clinical microbiology, designated *apeE*, that encodes esterase was found to be inactive (as a result of frameshift mutation) in all *S.* Dublin clinical isolates but is active in other *Salmonella* serovars including Enteritdis, Typhimurium and Typhi. This finding can explain the failure to detect *S.* Dublin on Aes Laboratoire *Salmonella* Agar Plate (ASAP) [15]. Although *apeE* is also absent from some other *Salmonella* serovars including Choleraesuis, Newport, Agona and Heidelberg these serovars harbour a gene closely related to *apeE* (~98% identity) that codes lipase which has the ability to hydrolyse a variety of esters, therefore ASAP chromogenic medium will be suitable for their detection. However, the gene coding lipase was also inactive in all *S.* Dublin clinical isolates.

DISCUSSION

Human infection with *S.* Dublin represents a significant public health problem. In Ireland, an outbreak

https://doi.org/10.1017/S0950268816000492 Published online by Cambridge University Press

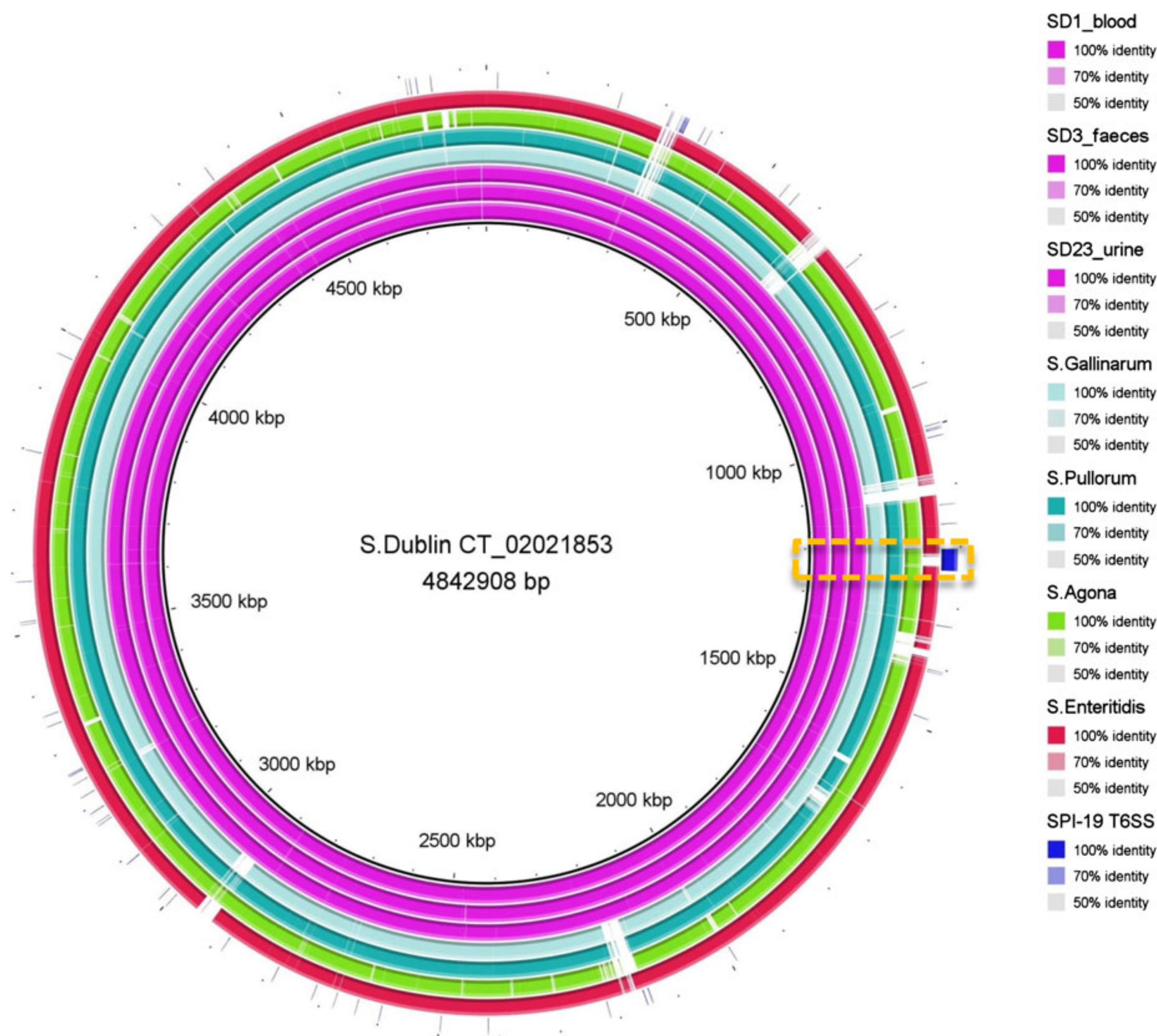

Fig. 3. Complete genome alignment of different *Salmonella* serovars (including the draft genome of representatives of clinical Irish isolates of *S.* Dublin sequenced in this study) generated using BRIG [27]. *S.* Dublin strain CT_02021853 genome is used as a reference. The gene cluster of SPI-19 T6SS (gene range SeD_A1212–SeD_A1243) is also included in the alignment. The genome of *S.* Dublin clinical isolates (including invasive and gastroenteritis isolates), *S.* Agona and the poultry-adapted serovars including *S.* Gallinarum and *S.* Pullorum harbours SPI-19T6SS while it is absent from *S.* Enteritidis.

of *S.* Dublin gastroenteritis confirmed by WGS was reported in October–November 2013 [16]. Although all isolates associated with this specific outbreak were from faecal samples prior to the outbreak, one third of *S.* Dublin clinical isolates received at the National *Salmonella, Shigella* and *Listeria* Reference laboratory were isolated from blood and other sterile body fluids.

Here we applied WGS to a new set of *S.* Dublin clinical isolates including invasive and gastroenteritis isolates and performed comparative genomic analysis in *S.* Dublin clinical isolates and the reference *S.* Enteritidis PT4 to investigate the underlying bacterial genetic determinants that might be associated with virulence in *S.* Dublin.

Two different T6SS encoded on two pathogenicity islands, $T6SS_{SPI\text{-}6}$ and $T6SS_{SPI\text{-}19}$, are present in all *S.* Dublin isolates described in this study but absent from *S.* Enteritidis PT4. Interestingly, $T6SS_{SPI\text{-}6}$ and $T6SS_{SPI\text{-}19}$ share very limited homology with each other.

Experimental studies have showed the crucial role of $T6SS_{SPI\text{-}6}$ in the systemic spread of *S.* Typhimurium [17]. Interestingly, two virulence genes within $T6SS_{SPI\text{-}6}$; *sciR* (STM0284) and *sciS* (STM0285) which encode the Shiga-like toxin A subunit and

https://doi.org/10.1017/S0950268816000492 Published online by Cambridge University Press

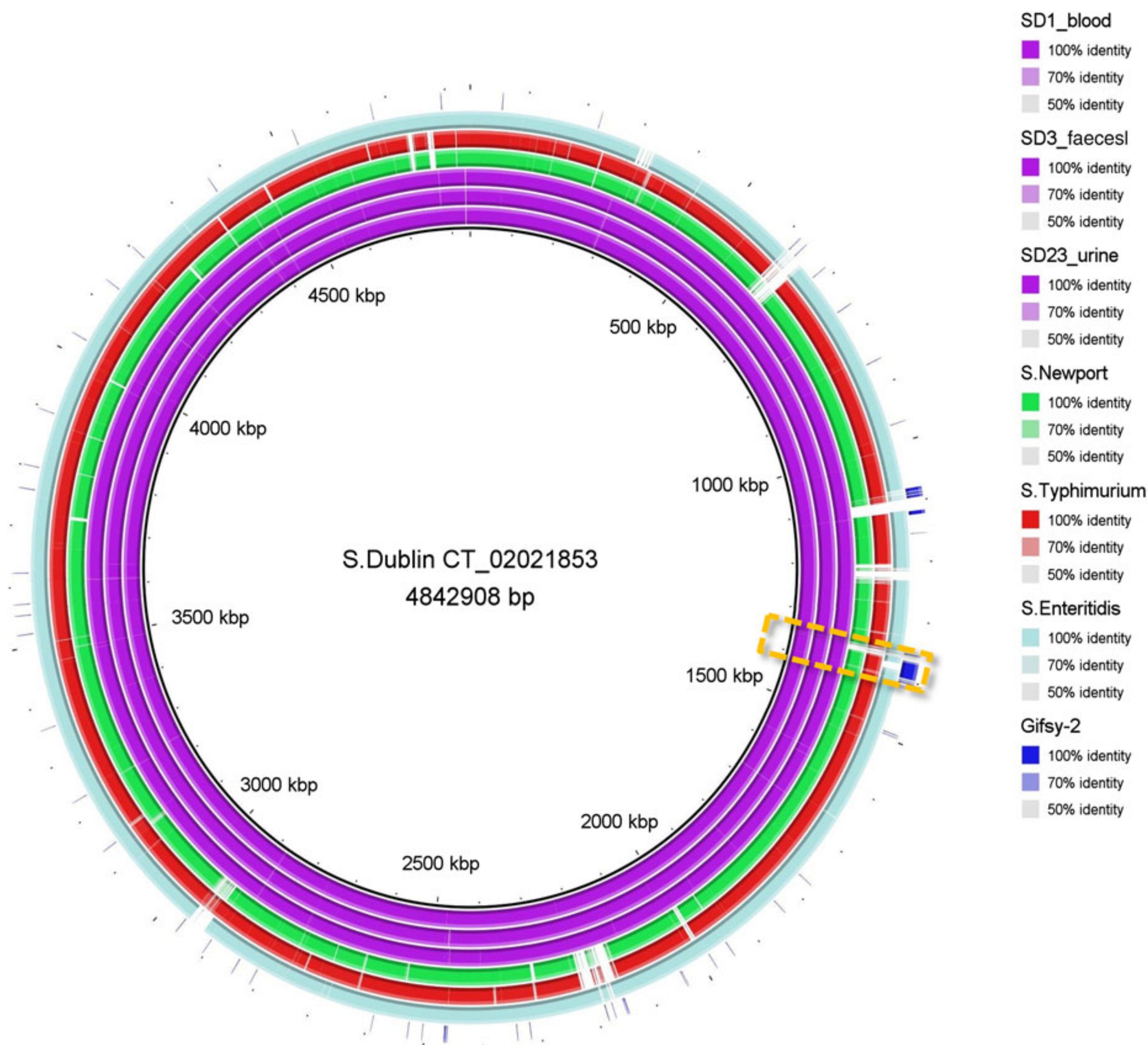

Fig. 4. Complete genome alignment of different *Salmonella* serovars (including the draft genome of representatives of clinical Irish isolates of *S.* Dublin sequenced in this study) generated using BRIG [27]. *S.* Dublin strain CT_02021853 genome is used as a reference. Prophage Gifsy-2 genome is also included in the alignment. The genome of clinical *S.* Dublin isolates (including invasive and gastroenteritis isolates), *S.* Typhimurium and *S.* Newport harbours Gifsy-2 prophage while Gifsy-2 is defective in *S.* Enteritidis PT4.

putative inner membrane protein, respectively, were absent from *S.* Enteritidis PT4 but present in all *S.* Dublin isolates. We also found that $T6SS_{SPI-6}$ is present in other virulent *Salmonella* serovars including Typhi, Paratyphi A, Choleraesuis and Typhimurium suggesting that $T6SS_{SPI-6}$ contributes to the ability of *Salmonella* to cause invasive disease.

$T6SS_{SPI-19}$ was also carried by all *S.* Dublin clinical isolates. Interestingly, $T6SS_{SPI-19}$ is present in other host-restricted serovars including poultry-associated *S.* Gallinarum and *S.* Pullorum suggesting its role in adaptation to specific animals [5, 18]. However, SPI-19 is subject to internal deletion in the broad host range serovar Enteritidis PT4. It has been reported that loss of both T6SSs in *S.* Enteritidis appears to be correlated to adaptation to a wide host range [5].

All *S.* Dublin clinical isolates are lysogenic for a lambdoid prophage (Gifsy-2-like prophage), which is defective in *S.* Enteritidis PT4 [19]. Gifsy-2 contributes significantly to the virulence and pathogenesis of *Salmonella* serotypes [20] as it encodes several virulence factors including *sodCI* (STM1044), *sseI* (STM1051) and *gtgE* (STM1055) [21]. Gifsy-2-like prophage might therefore be associated with the propensity of *S.* Dublin to cause invasive disease.

https://doi.org/10.1017/S0950268816000492 Published online by Cambridge University Press

Table 2. *Single genes and regions (Reg) that are exclusive to* Salmonella *Enteritidis PT4 strains but are absent from* Salmonella *Dublin strains*

Present in Enteritidis PT4 but absent from all Dublin	Function/gene prediction
SE_Reg1 (SEN0083–SEN0085)	Two probable secreted proteins and a putative sulfatase
SE_Reg2 (SEN1379–SEN1396)	Part of phage SE14 encoding several phage-related proteins
SE_Reg3 (SEN1432–SEN1436)	Part of ROD13 genomic island encoding proteins involved in sugar transport and hexonate uptake
SE_Reg4 (SEN1503–1506)	Part of ROD14 encoding regulatory proteins
SE_Reg5 (SEN1936–SEN1940, SEN1944–SEN1945 and SEN1959)	Encode several phage-related proteins including terminases and capsid and membrane proteins
SEN2420	Putative exported protein
SEN3943	Conserved hypothetical protein

Table 3. *List of the seven coding sequences that are predicted to be active genes in* Salmonella *Enteritidis PT4 but inactive (pseudogenes) in* Salmonella *Dublin isolates*

Gene	Prediction	Enteritidis (PT4)	Dublin (Irish isolates)	Gallinarum (287/91)	Cholerasuis (C500)	Typhi (CT18)	Paratyphi A (ATCC 9150)
SEN0042	Putative transport protein	Active	Pseudo	Pseudo	Pseudo	Active	Active
SEN0784	Hypothetical protein	Active	Pseudo	Pseudo	Pseudo	Absent	Absent
SEN2182*	*mglA,* galactoside transport ATP binding protein	Active	Pseudo	Pseudo	Pseudo	Pseudo	Pseudo
SEN2783	Hypothetical protein	Active	Pseudo	Pseudo	Pseudo	Absent	Absent
SEN2806	Probable glucarate dehydratase 2	Active	Pseudo	Pseudo	Pseudo	Active	Partial deletion
SEN3461	Outer membrane usher protein LpfC	Active	Pseudo	Pseudo	Pseudo	Absent	Absent
SEN3672	Probable phosphotransferase system permease	Active	Pseudo	Pseudo	Pseudo	Active	Partial deletion

A unique antivirulence gene, *grvA*, is carried on Gifsy-2. It decreases bacterial pathogenicity in its wild-type situation playing a role in optimizing *Salmonella* virulence [22]. However, the non-synonymous mutation detected in *grvA* of *S.* Dublin isolates had changed one amino acid (polar, hydrophilic arginine to non-polar hydrophobic tryptophan) within the GrvA protein (Arg53-Trp). We postulate that this may be associated with loss of its antivirulence function.

Interestingly, the gene encoding the Gifsy-2 prophage attachment and invasion protein (STM1043) was found in all *S.* Dublin strains; however, it was subjected to partial deletion in *S.* Enteritidis PT4.

Although the massive accumulation of pseudogenes (especially genes involved in metabolic pathways) in *S.* Dublin has been reported to be associated with adaptation to cattle and invasiveness in humans [5], it can also be associated with false-negative results in detection of *S.* Dublin at clinical laboratories when chromogenic medium (ASAP) is used since the *apeE* gene encodes esterase [23], responsible for bacterial ability to hydrolyse ester substrate methyl umbelliferyl caprylate within the chromogenic medium, which was found to be inactive in all *S.* Dublin isolates.

Therefore, *S.* Dublin is not able to produce ApeE esterase during overnight incubation and no expected pink to purple colonies are produced [15]. The unique biochemical features of *S.* Dublin [24] can delay bacterial identification by clinical laboratories where prompt identification is crucial since *S.* Dublin can be associated with invasiveness and systemic illness.

https://doi.org/10.1017/S0950268816000492 Published online by Cambridge University Press

We looked at the genetic differences in invasive and gastroenteritis isolates of *S.* Dublin, but no particular genetic determinants were found to differentiate among isolates. Phylogenetic analysis based on SNPs determined from WGS showed that invasive and faecal isolates are intermixed and SNPs were found to be randomly distributed around the *S.* Dublin chromosome. Furthermore, a plasmid that is closely related to the virulence plasmid (pCT02021853_74) of *S.* Dublin strain CT_02021853 was detected in all *S.* Dublin clinical isolates. This plasmid is limited to *S.* Dublin and has not been detected in other *Salmonella* serotypes. However, this plasmid carries virulence genes that proved to be associated with causing systemic illness [25].

Although the invasiveness of *S.* Dublin has been considered as related to expression of the Vi antigen, the encoding genes [26] were absent from all *S.* Dublin clinical isolates including invasive isolates.

Interestingly, the genome of all *S.* Dublin clinical 'invasive and faecal' isolates sequenced in this study harbours an active invasion gene (*pagN*) which might be associated with the ability to cause invasive illness.

In conclusion, our results reveal several virulence genetic determinants within the genome of *S.* Dublin from Ireland that gives this serovar the potential to cause invasive disease in humans. However, there were no genomic markers that differentiated invasive from gastroenteritis isolates suggesting that host factors are critical in determining the outcome in the individual patient.

Comparison of larger numbers of *Salmonella* isolates from different geographical and temporal spreads may provide detailed insights into the basis of bacterial virulence and invasiveness.

SUPPLEMENTARY MATERIAL

For supplementary material accompanying this paper visit http://dx.doi.org/10.1017/S0950268816000492.

ACKNOWLEDGEMENTS

This research received no specific grant from any funding agency, commercial or not-for-profit sectors.

DECLARATION OF INTEREST

None.

REFERENCES

1. **Bronowski C, *et al.*** Genomic characterisation of invasive non-typhoidal *Salmonella enterica* subspecies *enterica* serovar Bovismorbificans isolates from Malawi. *PLoS Neglected Tropical Diseases* 2013; **7**: e2557.
2. **Jones TF, *et al.*** Salmonellosis outcomes differ substantially by serotype. *Journal of Infectious Diseases* 2008; **198**: 109–114.
3. **Morris C, *et al.*** *Salmonella enterica* serovar Dublin strains which are Vi antigen-positive use type IVB pili for bacterial self-association and human intestinal cell entry. *Microbial Pathogenesis* 2003; **35**: 279–284.
4. **Achtman M, *et al.*** Multilocus sequence typing as a replacement for serotyping in *Salmonella enterica*. *PLoS Pathogens* 2012; **8**: e1002776.
5. **Langridge GC, *et al.*** Patterns of genome evolution that have accompanied host adaptation in *Salmonella*. *Proceedings of the National Academy of Sciences USA* 2015; **112**: 863–868.
6. **Betancor L, *et al.*** Genomic comparison of the closely related *Salmonella enterica* serovars Enteritidis and Dublin. *Open Microbiology Journal* 2012; **6**: 5–13.
7. **Moran NA, Plague GR.** Genomic changes following host restriction in bacteria. *Current Opinion in Genetics and Development* 2004; **14**: 627–633.
8. **Fierer J.** Invasive *Salmonella* Dublin infections associated with drinking raw milk. *Western Journal of Medicine* 1983; **138**: 665–669.
9. **Li H, Durbin R.** Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 2009; **25**: 1754–1760.
10. **Li H, *et al.*** The sequence alignment/map format and SAMtools. *Bioinformatics* 2009; **25**: 2078–2079.
11. **Tamura K, *et al.*** MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Molecular Biology and Evolution* 2013; **30**: 2725–2729.
12. **Posada D.** jModelTest: phylogenetic model averaging. *Molecular Biology and Evolution* 2008; **25**: 1253–1256.
13. **Zerbino DR, Birney E.** Velvet: algorithms for *de novo* short read assembly using de Bruijn graphs. *Genome Research* 2008; **18**: 821–829.
14. **Aziz RK, *et al.*** The RAST Server: rapid annotations using subsystems technology. *BMC Genomics* 2008; **9**: 75.
15. **Gray S, *et al.*** Failure to detect *Salmonella enterica* serovar Dublin on Aes Laboratoire *Salmonella* Agar Plate. *Journal of Clinical Microbiology* 2003; **41**: 4003.
16. **Mohammed M, *et al.*** Whole genome sequencing provides an unambiguous link between *Salmonella* Dublin outbreak strain and a historical isolate. *Epidemiology and Infection* 2016; **144**: 571–581.
17. **Pezoa D, *et al.*** The type VI secretion system encoded in SPI-6 plays a role in gastrointestinal colonization and systemic spread of *Salmonella enterica* serovar Typhimurium in the chicken. *PLoS ONE* 2013; **8**: e63917.
18. **Blondel CJ, *et al.*** The type VI secretion system encoded in *Salmonella* pathogenicity island 19 is required for *Salmonella enterica* serotype Gallinarum survival within infected macrophages. *Infection and Immunity* 2013; **81**: 1207–1220.

https://doi.org/10.1017/S0950268816000492 Published online by Cambridge University Press

19. **Matthews TD, *et al.*** Genomic comparison of the closely-related *Salmonella enterica* serovars Enteritidis, Dublin and Gallinarum. *PLoS ONE* 2015; **10**: e0126883.
20. **Figueroa-Bossi N, Bossi L.** Inducible prophages contribute to *Salmonella* virulence in mice. *Molecular Microbiology* 1999; **33**: 167–176.
21. **Ho TD, *et al.*** Identification of GtgE, a novel virulence factor encoded on the Gifsy-2 bacteriophage of *Salmonella enterica* serovar Typhimurium. *Journal of Bacteriology* 2002; **184**: 5234–5239.
22. **Ho TD, Slauch JM.** Characterization of grvA, an anti-virulence gene on the gifsy-2 phage in *Salmonella enterica* serovar Typhimurium. *Journal of Bacteriology* 2001; **183**: 611–620.
23. **Carinato ME, *et al.*** The *apeE* gene of *Salmonella* Typhimurium encodes an outer membrane esterase not present in *Escherichia coli*. *Journal of Bacteriology* 1998; **180**: 3517–3521.
24. **Fierer J, Fleming W.** Distinctive biochemical features of *Salmonella* Dublin isolated in California. *Journal of Clinical Microbiology* 1983; **17**: 552–554.
25. **Libby SJ, *et al.*** The spv genes on the *Salmonella* Dublin virulence plasmid are required for severe enteritis and systemic infection in the natural host. *Infection and Immunity* 1997; **65**: 1786–1792.
26. **Pulickal AS, *et al.*** Prevalence and genetic analysis of phenotypically Vi- negative *Salmonella* Typhi isolates in children from Kathmandu, Nepal. *Journal of Tropical Pediatrics* 2013; **59**: 317–320.
27. **Alikhan NF, *et al.*** BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. *BMC Genomics* 2011, **12**: 402.

https://doi.org/10.1017/S0950268816000492 Published online by Cambridge University Press