Epidemic multiple drug resistant *Salmonella* Typhimurium causing invasive disease in sub-Saharan Africa have a distinct genotype

Robert A. Kingsley, Chisomo L. Msefula, Nicholas R. Thomson, et al.

Supplemental Material	http://genome.cshlp.org/content/suppl/2009/11/11/gr.091017.109.DC1.html
References	This article cites 49 articles, 20 of which can be accessed free at: http://genome.cshlp.org/content/19/12/2279.full.html#ref-list-1
Open Access	Freely available online through the *Genome Research* Open Access option.
Creative Commons License	This article is distributed exclusively by Cold Spring Harbor Laboratory Press for the first six months after the full-issue publication date (see http://genome.cshlp.org/site/misc/terms.xhtml). After six months, it is available under a Creative Commons License (Attribution-NonCommercial 3.0 Unported License), as described at http://creativecommons.org/licenses/by-nc/3.0/.
Email Alerting Service	Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or **click here.**

Letter

Epidemic multiple drug resistant *Salmonella* Typhimurium causing invasive disease in sub-Saharan Africa have a distinct genotype

Robert A. Kingsley,[1,9,10] Chisomo L. Msefula,[2,9] Nicholas R. Thomson,[1,9] Samuel Kariuki,[1,3] Kathryn E. Holt,[1] Melita A. Gordon,[2] David Harris,[1] Louise Clarke,[1] Sally Whitehead,[1] Vartul Sangal,[4] Kevin Marsh,[5] Mark Achtman,[4,6] Malcolm E. Molyneux,[2] Martin Cormican,[7] Julian Parkhill,[1] Calman A. MacLennan,[2,8] Robert S. Heyderman,[2] and Gordon Dougan[1]

[1]*The Wellcome Trust Sanger Institute, The Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, United Kingdom;* [2]*Malawi-Liverpool-Wellcome Trust Clinical Research Programme, University of Malawi College of Medicine, Blantyre, Malawi;* [3]*Centre for Microbiology Research, Kenya Medical Research Institute, Nairobi, Kenya;* [4]*Department of Molecular Biology, Max-Planck Institute for Infection Biology, D-10117 Berlin, Germany;* [5]*Kenya Medical Research Institute–Wellcome Trust Collaborative Project, Kilifi, Kenya;* [6]*Environmental Research Institute and Department of Microbiology, University College Cork, Cork, Ireland;* [7]*Department of Bacteriology, National University of Ireland, Galway, Ireland;* [8]*Medical Research Council Centre for Immune Regulation, Institute of Biomedical Research, The Medical School, University of Birmingham, Edgbaston, Birmingham B15 2TT, United Kingdom*

Whereas most nontyphoidal *Salmonella* (NTS) are associated with gastroenteritis, there has been a dramatic increase in reports of NTS-associated invasive disease in sub-Saharan Africa. *Salmonella enterica* serovar Typhimurium isolates are responsible for a significant proportion of the reported invasive NTS in this region. Multilocus sequence analysis of invasive *S.* Typhimurium from Malawi and Kenya identified a dominant type, designated ST313, which currently is rarely reported outside of Africa. Whole-genome sequencing of a multiple drug resistant (MDR) ST313 NTS isolate, D23580, identified a distinct prophage repertoire and a composite genetic element encoding MDR genes located on a virulence-associated plasmid. Further, there was evidence of genome degradation, including pseudogene formation and chromosomal deletions, when compared with other *S.* Typhimurium genome sequences. Some of this genome degradation involved genes previously implicated in virulence of *S.* Typhimurium or genes for which the orthologs in *S.* Typhi are either pseudogenes or are absent. Genome analysis of other epidemic ST313 isolates from Malawi and Kenya provided evidence for microevolution and clonal replacement in the field.

[Supplemental material is available online at http://www.genome.org. The sequence data for *S.* Typhimurium strain D23580 chromosome and pSLT-BT have been submitted to EMBL (http://www.ebi.ac.uk/embl/) under accession nos. FN424405 and FN432031, respectively. The sequence data for A130 and 5579 have been submitted to the European Read Archive under accession nos. ERA000075 and ERA000076, respectively, and are available at ftp://ftp.era.ebi.ac.uk/vol1/ERA000/ERA000075/ and ftp://ftp.era.ebi.ac.uk/vol1/ERA000/ERA000076/.]

Bacterial genomes harbor both conserved core genome sequence, shared within a species, and pan-genomic sequences that include variable sequence, specific for clusters of isolates or clonal types (Medini et al. 2005). The core genome encodes many of the genes that define the species, whereas the pan-genome brings diversity in terms of phenotype, including antigenic structure, metabolic profile or, in the case of pathogens, virulence.

In *Salmonella enterica*, ~10% of any individual genome is variable, including many horizontally acquired or mobile elements such as prophages, transposons, plasmids, and insertion elements (McClelland et al. 2001; Parkhill et al. 2001a; Thomson et al. 2008). *S. enterica* is a complex species composed of distinct groups, referred to as serovars, which are defined by antigenic structure and biotype. One of the most common serovars associated with human and animal disease is *S. enterica* serovar Typhimurium (*S.* Typhimurium). This serovar includes a number of variants (pathovars) some of which are promiscuous in terms of their ability to infect different livestock hosts and humans (e.g., DT104), whereas others exhibit a restricted host range (e.g., DT2 and DT99) (Rabsch et al. 2002). In humans, *S.* Typhimurium infection is normally associated with self-limiting gastroenteritis (Zhang et al. 2003), and systemic disease is rare in most parts of the world. However, a highly invasive form of nontyphoidal *Salmonella* (NTS) disease, frequently associated with *S.* Typhimurium isolates, has emerged as a major public health problem in sub-Saharan Africa (Gilks et al. 1990; Vugia et al. 1993; Graham et al. 2000a; Berkley et al. 2005; Kariuki et al. 2006a; Gordon et al. 2008). This disease is characterized by bacteraemia and/or meningitis, with septic arthritis also reported. The clinical presentation is nonspecific, with

fever often the only clinical sign and a history of gastroenteritis present in under half of all cases. Microbiological confirmation is essential for making the diagnosis. Case-fatality rates are around 20%–25% for children and up to 50% in adults. NTS have infected tens of thousands of people, replacing pneumococcus as the most common cause of invasive bacterial disease in some parts of Africa (Vugia et al. 1993; Berkley et al. 2005; Gordon et al. 2008). Invasive NTS infections in adults are almost always associated with HIV-positive individuals (Gordon et al. 2002), but less than one in five cases in infants are HIV-associated, with malaria, anemia, and malnutrition being the most common risk factors (Brent et al. 2006). Importantly, many NTS *S.* Typhimurium isolates causing invasive disease are multiply antibiotic resistant, compromising the clinical treatment of the disease (Gordon et al. 2008).

Although *S.* Typhimurium does not normally cause invasive disease outside of sub-Saharan Africa, other *S. enterica* serovars do, including *S.* Typhi and *S.* Paratyphi A. Infections caused by *S.* Typhi and *S.* Paratyphi A are classically referred to as typhoid or enteric fever. Like typhoid, invasive NTS infection is not always associated with diarrhea (Raffatellu et al. 2008), but overall the diseases are clinically distinct (Graham et al. 2000b; Gordon et al. 2002). Both *S.* Typhi and Paratyphi A are host restricted to humans. In bacteria, host restriction and change of ecological niche (gut to systemic disease) have been linked to the accumulation of large numbers of pseudogenes and an overall reduction in genome size. One or both of these phenomena have been observed previously in other host-restricted pathogens, including *Yersinia pestis* (Parkhill et al. 2001b; Chain et al. 2004) and *Mycobacterium leprae* (Vissa and Brennan 2001).

Previous studies have provided indirect evidence that many invasive *S.* Typhimurium isolates are genotypically related (Kariuki et al. 2006a,b). These studies were performed using subgenomic fingerprinting assays, such as pulsed-field gel electrophoresis (PFGE)- or polymerase chain reaction (PCR)-based methodologies that provide limited insight into phylogenetic and genomic relationships. Although there is evidence that the genome of *S.* Typhimurium is variable (Cooke et al. 2007), this variation has not yet been linked to differences in pathogenicity between field isolates. Here, we address the hypothesis that many *S.* Typhimurium isolates that cause invasive disease in susceptible populations in sub-Saharan Africa are phylogenetically distinct from other *S.* Typhimurium isolates for which whole-genome sequence data are available. We have used a combination of subgenomic assays and whole-genome sequencing to characterize *S.* Typhimurium isolates from invasive disease in two countries and a number of distinct geographical regions where such infections are common. We provide evidence that a distinct phylogenetic lineage of *S.* Typhimurium is a significant cause of invasive NTS disease. We describe genome features, including evidence of genome degradation of genes previously implicated in host pathogen interactions, suggesting that this lineage may represent a distinct pathotype of *S.* Typhimurium.

Results

Novel sequence type (ST) variants of *S.* Typhimurium are associated with invasive NTS in sub-Saharan Africa

S. Typhimurium is a common cause of invasive NTS disease in sub-Saharan Africa, but little is known about the relationship of such strains with those found in other regions of the world. We applied multilocus sequence typing (MLST) (Maiden et al. 1998) analysis to a collection of *S.* Typhimurium isolates from invasive disease in urban and rural areas of the Blantyre region in Malawi, and from both urban Nairobi and rural Kilifi in Kenya (Table 1). These data were compared with ~400 isolates of *S.* Typhimurium for which the ST has been deposited in a global MLST database (http://www.mlst.net/) (Fig. 1). All 31 of the Malawian isolates and 13/20 of the Kenyan isolates were of a single MLST type, ST313. In contrast to the dominance of ST313 in Kenya and Malawi, only

Table 1. Summary of bacterial strains used in this study

Strain	Country of isolation	Year	Source[a]	Sequence type	Phage type[b]	Resistance profile[c]
A018	Malawi	1997	Blood	313	DT56var	AKSuW
A082	Malawi	1997	Blood	313	DT56var	AKSuWT
A130	Malawi	1997	Blood	313	DT56var	AKSuW
A357	Malawi	1998	Blood	313	DT56var	AKSuW
A665	Malawi	1998	Blood	313	DT56var	AKSuW
A680	Malawi	1998	Blood	313	DT56var	AKSuW
A3800	Malawi	1999	Blood	313	DT56var	AKSuW
A4283	Malawi	1999	Blood	313	DT56var	AKSuW
A4447	Malawi	1999	Blood	313	DT56var	AKSuW
C2110	Malawi	2000	CSF	313	Untypable	AKSuW
C2167	Malawi	2000	CSF	313	DT56	AKSuWT
C2364	Malawi	2000	CSF	313	DT56var	AKSuW
A13198	Malawi	2001	Blood	313	DT56var	AKSuW
A13212	Malawi	2001	Blood	313	DT56	AKSuWT
D11578	Malawi	2001	Blood	313	DT56var	AKSuW
A16083	Malawi	2002	Blood	313	DT56var	AKSuWT
D14916	Malawi	2002	Blood	313	DT56var	CAKSSuWT
D15040	Malawi	2002	Blood	313	DT22	AKSuW
A24906	Malawi	2003	Blood	313	DT56var	CAKSSuW
A24910	Malawi	2003	Blood	313	Untypable	CAKSSuW
A24924	Malawi	2003	Blood	313	DT56var	AKSuW
A32751	Malawi	2004	Blood	313	Untypable	AKSuW
A32773	Malawi	2004	Blood	313	DT56var	CAKSSuW
A32796	Malawi	2004	Blood	313	DT56var	CAKSSuW
D23580	Malawi	2004	Blood	313	Untypable	CAKSSuW
A38589	Malawi	2005	Blood	313	DT56var	CAKSSuW
A38596	Malawi	2005	Blood	313	DT56var	CAKSSuW
C13184	Malawi	2005	CSF	313	Untypable	CAKSSuW
A39051	Malawi	2006	Blood	313	DT56var	CAKSSuW
A39129	Malawi	2006	Blood	313	DT56var	CAKSSuW
A39155	Malawi	2006	Blood	313	DT56var	CAKSSuW
5521	Kenya	2003	Blood	313	—	AKSSuW
5575	Kenya	2003	Blood	313	DT56var	ASuW
5576	Kenya	2003	Blood	313	RDNC	ASuW
5577	Kenya	2003	Blood	313	Untypable	ASuW
5578	Kenya	2003	Blood	19	RDNC	K
5579	Kenya	2003	Blood	313	Untypable	CAKSSuW
5580	Kenya	2003	Blood	394	RDNC	ASSuW
5581	Kenya	2003	Blood	313	DT193	ASuW
5582	Kenya	2003	Blood	313	Untypable	AK
5597	Kenya	2003	Blood	313	Untypable	AKSSu
5632	Kenya	2003	Blood	19	DT41	K
5634	Kenya	2003	Blood	19	DT56var	K
5647	Kenya	2003	Blood	19	DT120	CAKSu
5912	Kenya	2004	Soil	313	DT56	ASuW
5964	Kenya	2004	Soil	313	DT56var	CAKSuW
3697	Kenya	NK	Blood	313	RDNC	NK
3762	Kenya	NK	Blood	19	DT193	NK
3831	Kenya	NK	Blood	313	DT193	NK
4191	Kenya	NK	Blood	19	DT193	NK
4246	Kenya	NK	Blood	313	Untypable	NK

[a]Bacterial strains were cultured from blood, soil, or cerebrospinal fluid (CSF).
[b]RDNC, reacts does not conform.
[c]Antibiotic resistance profile. A, 10 μg of ampicillin; C, 30 μg of chloramphenicol; K, 30 μg of kanamycin; S, 10 μg of streptomycin; Su, 200 μg of compound sulphonamide; W, 5 μg of trimethoprim; NK, not known.
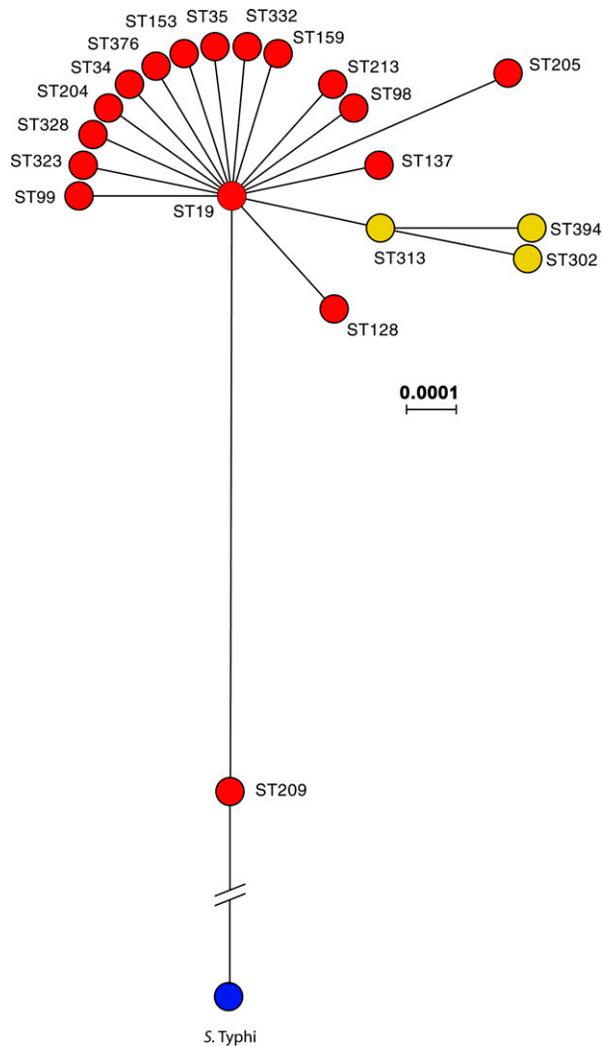
Figure 1. Radial phylogram showing the phylogenetic relationship of *S. enterica* serovar Typhimurium sequence types (STs). The *S.* Typhimurium STs (red and yellow circles) are rooted to *S.* Typhi (blue circle). The length of the connecting line is proportional to the degree of divergence. *S.* Typhimurium STs from various worldwide locations outside sub-Saharan Africa SSA (red) and NTS STs predominantly from SSA (yellow) are indicated.

two other isolates of this ST from outside of sub-Saharan Africa are currently in the *S. enterica* MLST database, which contains over 20 distinct STs of *S.* Typhimurium. Interestingly, both of these other ST313 STs were isolated from cases of invasive NTS disease, one in Scotland and the second in India. One other isolate from Kenya that was assigned to ST394 is highly related to ST313 (Fig. 1). The remaining six isolates from Kenya were of ST19, which is the most common ST for *S.* Typhimurium in the database, with >220 entries. ST19 includes previously sequenced isolates of *S.* Typhimurium associated with classic gastroenteritis, including LT2, SL1344, and DT104 NCTC13348. Most of the ST313 isolates described in this study are of phage types DT56var or untypeable (Table 1). Perhaps significantly, *S.* Typhimurium DT56var is frequently isolated from wild birds, where they cause significant mortality but are rarely isolated elsewhere from humans (Pennycott et al. 2006).

The complete genome sequence of ST313 D23580 from Blantyre, Malawi reveals a novel prophage repertoire and evidence of pseudogene accumulation

As ST313 is distinguishable by MLST from previously sequenced *S.* Typhimurium isolates (Fig. 1), we determined the genome sequence of a representative isolate, D23580, that was isolated in 2004, at the peak of a Blantyre multidrug-resistant (MDR), invasive NTS epidemic (MacLennan et al. 2008). The date of isolation of D23580 corresponded with the emergence of resistance to chloramphenicol, which was the antibiotic treatment of choice until that time (Gordon et al. 2008). As might be expected, whole-genome comparisons between D23580 and LT2, SL1344, and DT104 NCTC13348 revealed extensive synteny and colinearity (Fig. 2). However, in addition to representing a distinct MLST type, the chromosome of D23580 had features that differentiated it from *S.* Typhimurium strains LT2 (McClelland et al. 2001), DT104 NCTC13348, and SL1344, for which genomic sequence data had previously been determined. For example, a comparison of the genome of D23580 with SL1344 and DT104 NCTC13348 identified a number of single nucleotide polymorphisms (SNPs) (Supplemental Fig. 1), insertions, and deletions (Table 2; Supplemental Table 1).

One of the differentiating features was a distinct repertoire of five prophage-like elements, designated BTP1 through BTP5 (Fig. 2; description in Supplemental text). Our analysis of the prophage regions in the D23580 genome revealed that *sseI* that encodes a Type III effector secreted by the *Salmonella* pathogenicity island (SPI)-2 encoded type III secretion system, is a pseudogene. Therefore, we analyzed all the pseudogenes present in D2380 and compared these to the other sequenced *S.* Typhimurium. D23580 contained a total of 77 assignable pseudogenes, of which 23 were specific to this strain (Table 2; Supplemental Table 1). Among the D23580-specific pseudogenes are two, STM_MW29741 and STM_MW10251, which are predicted to encode regulators and may therefore affect expression of multiple genes. A number of predicted transporters are also D23580-specific, including two from the ABC transporter superfamily STM_MW09551 (*ybjZ*), STM_MW30361, and an MFS-family transporter STM_MW15161 (*ydeE*). Eleven of the 23 pseudogenes that are specific for D23580 are also either pseudogenes or completely absent from *S.* Typhi or *S.* Paratyphi A. These represent genes for which the respective orthologous genes or genes in the same biosynthetic pathway in *S.* Typhi or *S.* Paratyphi A are degraded. For example, the *ratB* gene that encodes a secreted protein implicated in long-term persistence in the murine intestine (Kingsley et al. 2003) is inactivated by a frameshift mutation in *S.* Typhi and by a premature stop codon in D23580. Additional orthologous pseudogenes are of currently unknown function, but these may also reflect selection for convergence in the evolution of these pathogens. Twelve D23580-specific pseudogenes that are apparently intact in *S.* Typhi and *S.* Paratyphi A as well as *S.* Typhimurium SL1344, DT2, and LT2, include the *lpxO* gene that encodes a putative dioxygenase required for the synthesis of 2-hydroxymyristate modified lipid A (Gibbons et al. 2000). This modification of the lipid bilayer is controlled by the PhoPQ response regulator (Murata et al. 2007; Gibbons et al. 2008), a central regulator of *Salmonella* virulence genes that contributes to resistance to host defenses including attack by cationic peptides.

In addition to genome degradation due to pseudogene formation, the genome of D23580 is ~15 kb shorter than SL1344, excluding IS and prophage elements, due to four separate deletions impacting a total of 20 genes (Table 2; Supplemental Table 1).
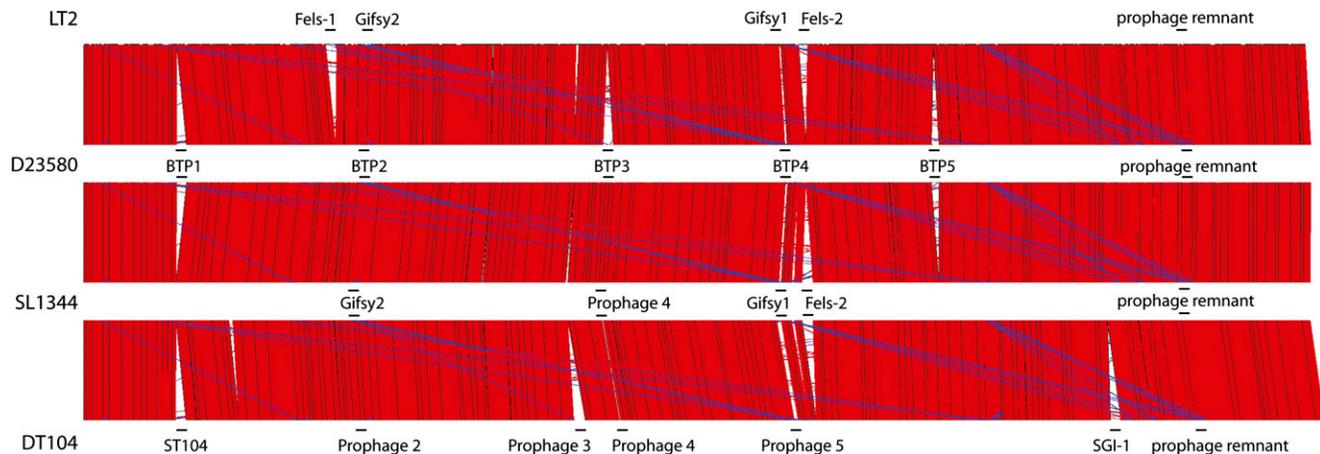
Figure 2. Whole-genome comparison of *S.* Typhimurium LT2, D23580, and DT104 NCTC13348. ACT comparison (http://www.sanger.ac.uk/Software/ACT/) of amino acid matches between the complete six-frame translations (computed using BLASTN) of the whole genome sequences of *S.* Typhimurium LT2 (LT2), D23580, and DT104 NCTC13348. The red bars *between* the DNA lines represent individual TBLASTX matches, with inverted matches colored blue. Large chromosomal features are indicated with a horizontal line labeled with the feature identity. These analyses indicate the overall colinearity and synteny of the chromosome of sequenced *S.* Typhimurium isolates. Large insertions or deletions (mainly prophage elements) are labeled. Smaller indels described in the text are not visible at this resolution.

These deleted regions encode a number of genes that have been implicated in pathogenesis. For example, a deletion of 723bp results in loss of the 5′ end of the *pipD* gene located in SPI-5. The *pipD* gene has homology with dipeptidases in *Lactobacillus* spp. and is required by *S.* Typhimurium for induction of fluid secretion in bovine ileal loops (Wood et al. 1998) and persistence in a murine systemic model (Lawley et al. 2006). A deletion of 3981 bp has removed some or all of STM1548 through STM1553, which are of unknown function. A 9568 bp deletion in a phage remnant resulted in the deletion of at least twelve putative coding sequences (CDSs), including *pagO* and *mig-13*, as well as a number of putative phage-related genes. The *pagO* gene is regulated by the *phoP/phoQ* two-component regulator that also influences the expression of key virulence genes encoded on SPI-1 and SPI-2 (Gunn et al. 1998). The *mig-13* gene encodes a putative membrane protein that was identified on the basis of up-regulation on invasion of cultured macrophages (Valdivia and Falkow 1997). Both *mig-13* and *pagO* have been implicated in systemic persistence in a genome-wide random mutagenesis screen (Lawley et al. 2006). Finally, a 1694 bp deletion affects the *allP* and *allB* genes encoding allantoin permease and allantoinase. Allantoin is used as a sole source of carbon and nitrogen by many enteric bacteria following conversion to glyoxylate by a pathway involving allantoinase (encoded by *allB*), allantoate amidohydrolase (encoded by *allC*), and (S)-ureidoglycolate hydrolase (encoded by *allD*) (Cusa et al. 1999). Interestingly, *allP* and *allB* are also pseudogenes in *S.* Typhi (Parkhill et al. 2001a); *allP* is a pseudogene in *S.* Paratyphi A (Holt et al. 2009); *allS*, a transcriptional activator of the *allD* operon, is a pseudogene in *S.* Gallinarum (Thomson et al. 2008). Perhaps significantly, *S.* Typhi, and *S.* Paratyphi A are adapted to higher primates, and *S.* Gallinarum is restricted to poultry. The loss of allantoin metabolism in human- and poultry-adapted pathogens may be explained by differences in purine metabolism in the tissues of higher primates and avian species. In lower mammals, urate, the breakdown product of purines, is converted to allantoin by the action of uricase, prior to excretion. In contrast, humans, anthropoid apes and new-world monkeys do not express uricase, and therefore allantoin is likely not available in these hosts (Fujiwara and Noguchi 1995). Further-

more, 15 of 20 genes absent from D23580 are also either absent or present as pseudogenes in *S.* Typhi and/or *S.* Paratyphi A (Parkhill et al. 2001a; McClelland et al. 2004). With respect to pseudogenes and deleted genes of D23580, taken together, these data provide evidence that the genome of D23580 has undergone a pattern of degradation distinct from LT2, SL1344, and DT104 NCTC13348.

Draft sequencing of a chloramphenicol sensitive isolate strain from Blantyre (strain A130) and an MDR isolate from Kenya (strain 5579)

The epidemic increase in the incidence of invasive NTS between 1997 and 2007 in Blantyre was associated with the emergence of MDR *S.* Typhimurium (Gordon et al. 2008). In 1997, >90% of isolates were resistant to cotrimoxazole and ampicillin but susceptible to chloramphenicol, the treatment of choice at the time. An increase in the proportion of chloramphenicol resistant (Cm-R) *S.* Typhimurium isolates was first noted in early 2002, and by the end of 2003 Cm-R was predominant, prompting a change to alternative therapies such as ciprofloxacin or ceftriaxone (Gordon et al. 2008).

In order to further investigate the *S.* Typhimurium causing invasive NTS disease during this period we generated draft sequences of *S.* Typhimurium A130 and *S.* Typhimurium 5579 using a 454 Life Sciences (Roche) pyrosequencing machine. *S.* Typhimurium A130 is a chloramphenicol susceptible ST313 isolate from 1997, prior to the emergence of the full MDR phenotype. A total of 680 contigs of >500 bp with an average coverage of 10× were assembled from single-end reads. The ST313 *S.* Typhimurium strain 5579 was isolated from Nairobi, Kenya, in 2003. In this case, a total of 205 contigs of >500 bp with an average coverage of 20× were assembled from paired end reads. Using these data we determined the phylogenetic relationship of these isolates and the gene content. We also determined the presence or absence of pseudogenes, when these were due to deletions or insertions greater than a single base pair, or due to a nonsense SNP resulting in a "STOP" codon. However, it was not possible to determine confidently the presence or absence of pseudogenes from pyrosequencing data when this

Table 2. Genome degradation specific to *S.* Typhimurium strain D23580

STM (ID)	STM_MW (ID)	Gene	Description
157	01631	—	Putative outer membrane protein
522	NP	*allP*	Allantoin transport protein
523	NP	*allB*	Allantoinase
834	08851	*ybiP*	Putative integral membrane protein
942	09551	—	Putative ABC superfamily transport protein
1014	10251	—	Probable regulation protein
1023	10351	—	Hypothetical on gifsy-2 prophage
1051	10631	*sseI*	Type III secretion effector protein (SPI-2)
1092	NP	—	Putative cytoplasmic protein
1093	NP	—	Putative cytoplasmic protein
1094	11041	*pipD*	Similar to dipeptidase A
1228		—	Putative periplasmic protein
1516	15161	*ydeE*	Putative MFS family transport protein
1548	15471	—	Putative ribosyltransferase-isomerase
1549	NP	—	Putative translation initiation inhibitor
1550	NP	—	Putative cytoplasmic protein
1551	NP	—	Putative cytoplasmic protein
1551'	NP	—	Hypothetical protein
1552	NP	—	Putative cytoplasmic protein
1637	16321	—	Putative inner membrane protein
1862	NP	*pagO*	PhoP activated gene
1863	NP	—	Putative inner membrane protein
1864	NP	—	Putative inner membrane protein
1865	NP	—	Putative DNA invertase
1866	NP	—	Hypothetical
1868	NP	*mig-13*	Phage tail assembly protein
1868'	NP	—	Lytic enzyme
1869'	NP	—	Hypothetical protein
1870	NP	—	RecE-like protein
1896	18781	—	Putative cytoplasmic protein
1940	19221	—	Putative cell wall associated hydrolase
2238	22681	—	Putative phage protein
2514	25311	*ratB*	Secreted protein
2589	26091	—	Hypothetical in Gifsy-1 prophage
2680	26941	—	Putative cytoplasmic protein
2932	28951	*ygbE*	Putative inner membrane protein
3012	29741	—	Putative transcriptional regulator
3075	30361	—	Putative ABC-type cobalt transport system
3355	33531	—	Tartrate dehydratase
3624	36131	*yhjU*	Putative inner membrane protein
3745	37341	—	Putative cytoplasmic protein
3768	37571	—	Putative selenocysteine synthase
4196	41451	—	Putative cytoplasmic protein
4286	42371	*lpxO*	Putative dioxygenase

Genome degradation relative to strains LT2, SL1344, and DT104, and comparison of these features with *S.* Typhimurium strains 5579 and A130, and *S.* Typhi Ty2, and *S.* Paratyphi A STM and STM_MW indicate the systematic designation of coding sequences in *S.* Typhimurium LT2 and *S.* Typhimurium D23580, respectively; pseudogenes are indicated by their systematic ID genes that are absent due to probable deletion are indicated NP (not present).

was due to a single nucleotide indel, as this sequencing technology is prone to generating frame-shift errors in homopolymeric tracts.

We identified 1542 SNPs in strains LT2, SL1344, DT104, D23580, A130, and 5579, which were used to determine phylogenetic relationships and calculate relative divergence (Fig. 3). As expected, the ST19 strains LT2 and DT104 NCTC13348, and the ST313 isolates formed distinct clusters. Interestingly, A130 exhibited some divergence from D23580. However, the Malawian isolate D23580 was very closely related to the Kenyan isolate 5579, indicating that these isolates shared a common ancestor relatively recently.

We surveyed the chromosome of *S.* Typhimurium 5579 and A130 to determine the presence or absence of genomic degradation features identified in the genome of D23580 (Supplemental Table 1). The chromosomal gene content of *S.* Typhimurium 5579 was similar to that of D23580, reflecting their recent ancestry. All

of the chromosomal deletions identified in the D23580 chromosome relative to SL1344 and DT104 NCTC13348 were also present in 5579. As in D23580, the *sseI* gene is interrupted by insertion of an IS200 element. All of the assignable pseudogenes formed by nonsense SNPs (STOP codons) were conserved in D23580 and 5579. None of the chromosomal deletions observed in the chromosome of D23580 or 5579 were present in the A130 chromosome, indicating that these regions were all deleted in D23580 since divergence from the A130 lineage (Fig. 3). In contrast, six of eight pseudogenes caused by nonsense SNPs (STOP codons; Table 2) were present in Al30 as well as D23580, indicating that this form of genome degradation occurred predominantly before these lineages diverged and are therefore older than the degradation due to deletions observed in D23580 alone.

The virulence-associated plasmids harbor composite Tn*21*-like transposons encoding multiple antibiotic resistance genes

In common with many other NTS isolates from invasive disease, *S.* Typhimurium D23580 exhibits an MDR phenotype (Table 1). Four putative plasmids were detected in D23580 by gel electrophoresis after plasmid DNA extraction. In agreement, the genomic analyses also identified four plasmids of 117 kbp (pSLT-BT), 84 kbp (pBT1), 2.6 kbp (pBT2), and 1.4 kbp (pBT3). Plasmids pBT1–pBT3 have not previously been described, and appear to be cryptic, in that we could not assign them an obvious phenotype (Supplemental Fig. 2). In contrast, pSLT-BT is closely related to pSLT, a virulence-associated plasmid essential for systemic invasiveness of *S.* Typhimurium in mice (Gulig et al. 1993). Comparison of the sequence of pSLT with pSLT-BT revealed extensive synteny overall, save for a single large insert in pSTL-BT that was absent from pSLT. This insertion resembled a composite Tn*21*-like mobile element (Fig. 4; description in Supplemental text). All the antibiotic resistance expressed by D23580 (CAKSSuW; Table 1) can be explained by genes encoded within this composite element, and no other resistance genes were found in pBT1, pBT2, pBT3, or on the chromosome. Thus, the antibiotic resistance phenotype of the invasive *S.* Typhimurium D23580 can be explained by the possession of multiple-linked antibiotic resistance determinants that are encoded by a plasmid associated with virulence and invasion. The virulence-associated plasmid of *S.* Typhimurium 5579 (pSLT-5579) was identical to pSLT-BT with the exception of deletion of most of the left end of the composite element that had similarity to pU302L. In contrast, pSLT-A130 harbored a distinctly different composite Tn*21*-like element encoding antibiotic resistance genes located at a different site from that on pSLT-BT and pSLT-5579 (Fig. 4; description in Supplemental text). Core genetic features of the Tn*21*-like regions
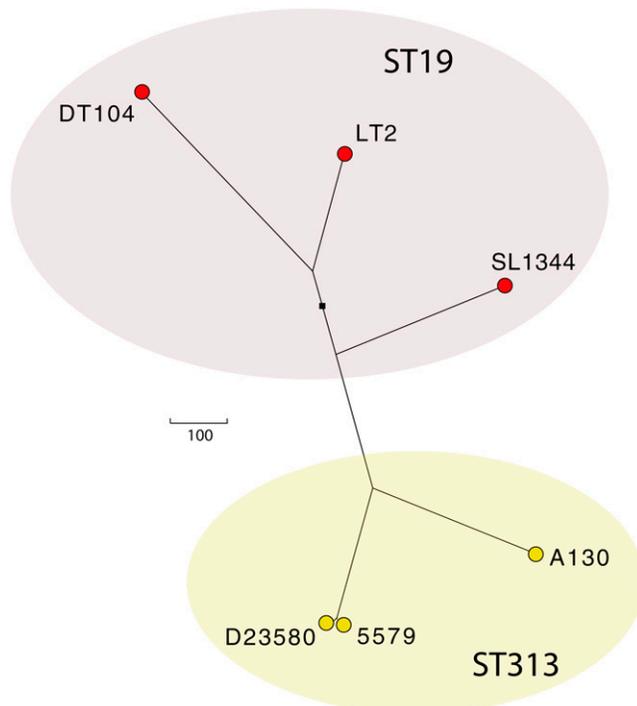
Figure 3. Radial phylogram based on chromosomal SNPs showing the phylogenetic relationship of sequenced *S.* Typhimurium isolates. Branch lengths indicate the number of SNPs, scale as indicated, separating the sequenced isolates indicated by circles. ST19 isolates (red circles) and ST313 isolates (yellow circles) are shown. The black square indicates the ancestral root.

of pSLT-BT and pSLT-A130 share some common features, but the composite elements differ in detail, reflecting a complex evolutionary history. Thus, the emergence of Cm-R in Blantyre may be associated with clonal replacement of different ST313 clones with distinct antibiotic resistance loci.

Evidence for clonal replacement of invasive MDR *S.* Typhimurium during a Blantyre epidemic

We were interested in determining the molecular events that led to the emergence of MDR in Blantyre during the period 1997–2007. We considered two alternative hypotheses to explain the epidemiological data. First, a single clone or clones may have been circulating in the Blantyre region throughout the study period and MDR emerged due to recruitment of a chloramphenicol resistance locus in one or more clones, after which the resistant clones became dominant. If this were the case, we would expect that the chromosomal genotype of isolates prior to the emergence of MDR would be identical to those exhibiting the MDR phenotype. Alternatively, the emergence of MDR may have been the result of clonal replacement by a distinct chromosomal genotype, or multiple genotypes, coupled with a distinct antibiotic resistance locus or loci. If this were the case the chromosomal genotypes would be distinct before and after the emergence of MDR. Prophage elements are hot spots for chromosome evolution in *S.* Typhimurium (Thomson et al. 2004; Cooke et al. 2007) and represent one approach to measuring genetic differences in otherwise highly related isolates. Therefore, oligonucleotide primer pairs were designed based on known prophage insertion sites or internal prophage-

specific sequences present in D23580 and other sequenced *S.* Typhimurium (Supplemental Fig. 3). PCR was then used to amplify genomic DNA template prepared from each of 31 clinical invasive *S.* Typhimurium isolates from Blantyre (Supplemental Table 2). This analysis confirmed that the ST313 isolates are highly conserved at most of the loci harboring the targeted prophage elements but significant genetic differences were detectable within prophages BT2 and BT3. Importantly, PCR amplification of BTP2 and BTP3 associated sequences (primer pairs E and F, Supplemental Table 3) were able to discriminate between Cm-S ST313 isolated early on and Cm-R ST313 isolated later in the Blantyre epidemic (Table 1), suggesting that clonal replacement by a strain with a distinct genotype was responsible for the emergence of MDR. Primer pair E amplifies DNA across the *sseI* CDS within prophage BT2 (Miao et al. 2003), whereas primer pair F amplifies across a highly variable region associated with BT3 that is almost completely absent from D23580 (Supplemental Fig. 2).

In order to investigate the molecular events that led to the emergence of Cm-R we then determined whether a similar composite Tn*21*-like element to that found on the pSLT-BT of D23580 was present in the virulence plasmids of other Malawian isolates. To this end, we used PCR with oligonucleotide primer pairs designed to anneal to the regions flanking the composite element found on pSLT-BT (primer pair P1; Supplemental Fig. 3). Under the experimental conditions employed, the presence of a large insertion at this site would prevent the amplification of a PCR product and the absence of an insertion would result in the amplification of a 1.2-kbp product. Significantly, a 1.2-kbp product was observed for all the Cm-S Malawian isolates, but a similar PCR product was not observed for Cm-R isolates (Supplemental Table 2), suggesting that the composite element locus was empty in the Cm-S isolates. Furthermore, primer pairs designed to amplify fragments internal to the integron-like element (primer pairs 2 through 5), did not amplify a product from the Cm-S isolates, except primer pair 6, which amplified an internal fragment of the *bla*$_{TEM-50}$ gene. However, products were generated with all these primers using Cm-R DNA as a template. These data indicated that while the antibiotic resistance locus of MDR isolates may be related, antibiotic resistance genes encoded by the Malawian strains isolated prior to emergence of chloramphenicol resistance are encoded at a different locus and on a distinct element or elements by isolates with a distinct chromosomal genotype. This was consistent with the presence of a Tn*21*-like element similar to that identified in strain A130 by pyrosequencing.

Discussion

We show here that many *S.* Typhimurium causing invasive NTS in different regions within sub-Saharan Africa belong to a distinct ST, designated ST313. ST313 isolates harbor genomic signatures that differentiate them from *S.* Typhimurium causing gastroenteritis in other regions of the world including a novel repertoire of prophage elements and evidence of genome degradation. Currently, it is not known if ST313 isolates are also associated with gastroenteritis in sub-Saharan Africa, and this awaits further epidemiological investigation. However, limited epidemiological analysis within the Malawian and Kenyan study sites has not identified ST313 *S.* Typhimurium as a significant cause of diarrhea in these specific regions. A framework for a multisite program to address this question has been proposed recently (Clemens 2009). Nonetheless, MLST analysis and comparative whole-genome analysis indicated significant evolutionary distance between ST313 isolates
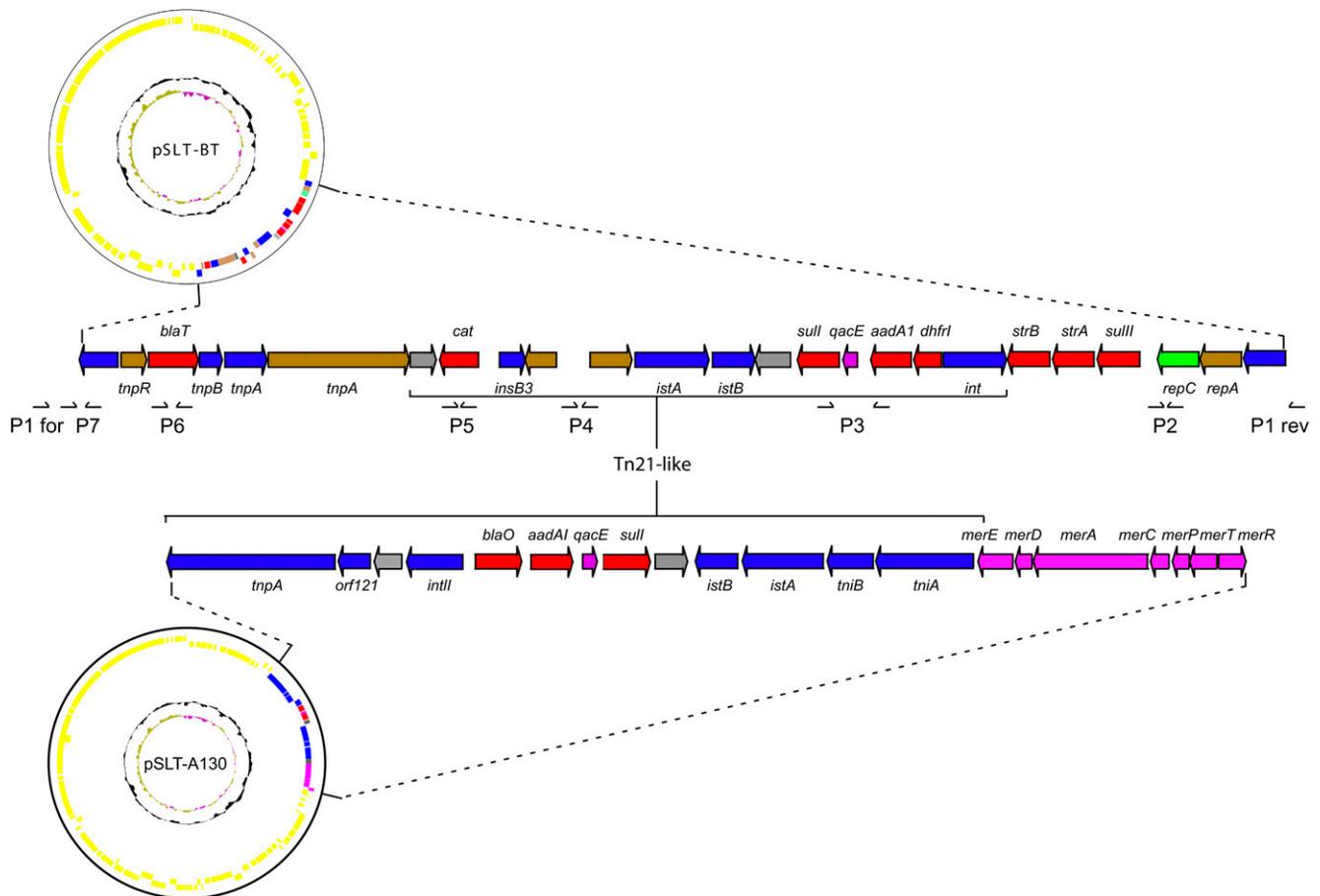
Figure 4. Genes and homologous regions of pSLT-BT and pSKT-A130. A circular map of pSLT-BT shows predicted genes on the forward and reverse strands on the *outer* circle (yellow). Genes with putative antibiotic resistance functions (red), integron- or transposon-related functions (blue), plasmid-associated functions (green), antiseptic and mercury resistance functions (purple), miscellaneous functions (gray), and pseudogenes (brown) are shown. Arrows *below* the pSLT-BT composite element gene map indicate the annealing site of oligonucleotide primers used to probe clinical isolates. Regions of each locus with similarity to Tn*21* are indicated.

and those of ST19 that include common gastroenteritis strains around the world. ST313 isolates rarely have been reported outside of sub-Saharan-Africa, although restricted geographical localization may not be a feature of just this ST. D23580 and other ST313 isolates harbor a unique combination of prophage-like elements and distinct composite mobile elements encoding antibiotic resistance genes. Relatively little is known about how bacterial genomes evolve over time and what impact this variation may have on pathogenicity. The evidence presented here indicates that evolution outside of the core genome of *S.* Typhimurium may be clinically significant. For example, the use of chloramphenicol for the treatment of invasive *S.* Typhimurium infections appears to have contributed to the clonal replacement of a Cm-S ST313 with a Cm-R variant with a novel composite mobile element and distinct genomic signatures.

Analysis of the complete genome sequence of *S.* Typhimurium D23580 suggests that ST313 *S.* Typhimurium may have undergone partial selective genome degradation with some similarities with those host adapted *Salmonella* serovars that cause invasive disease such as *S.* Typhi (Parkhill et al. 2001a; Holt et al. 2008), *S.* Paratyphi A (McClelland et al. 2004), and *S.* Gallinarum (Thomson et al. 2008). D23580 harbors additional pseudogenes and deletions not

present in other *S.* Typhimurium, some of which have been linked to virulence, including *sseI*, encoding a type III-secreted effector protein and ratB, encoding a secreted protein associated with intestinal persistence in the murine model of infection. The *S.* Typhimurium D23580 genome has also lost blocks of genes of unknown function through large deletions, some of which are also absent from *S.* Typhi (Table 3), such as STM1549–1553. Independent mutations in *S.* Typhimurium D23580 and *S.* Typhi have resulted in the convergent lack of allantoin uptake and metabolism. Of 44 novel pseudogenes or deletions in strain D23580 relative to LT2, over half of these (26) are also degraded in *S.* Typhi or *S.* Paratyphi A, in the form of pseudogenes or deletions.

The genome of strain A130 is not degraded to as large an extent as D23580, relative to SL1344 and DT104. None of the deletion events observed in D23580 are present in strain A130, indicating that these events occurred since these two lineages diverged. Although it was not possible to confirm the pseudogene status of most of the A130 orthologs that are pseudogenes in D23580, due to limitations of the pyrosequence methodology, we were able to confirm that six of eight pseudogenes caused by nonsense SNPs were conserved. These data suggested that at least some of the genome degradation, exclusively in the form of

pseudogene formation, occurred before the D23580 and A130 lineages diverged. In contrast, genome degradation due to deletion was restricted to the lineage leading to D23580 and strain 5579. Indeed, that two of the deletions in D23580 are not observed in the chromosome of the closely related strain 5579, indicates that degradation is an ongoing process. The gene content and the structure of the Tn*21*-like locus encoding antibiotic resistance on the virulence plasmid of *S*. Typhimurium strains D23580 and 5579 were very similar. Strain D23580 was isolated from the Blantyre region in 2004 and strain 5579 a year earlier from Nairobi in Kenya. The isolation of these two strains in different regions of sub-Saharan Africa suggests that there may be a connection between the emergence of MDR in these two locations.

An association between the invasion-associated *spv* locus on pSLT (Gulig et al. 1993; Guiney et al. 1995) with antibiotic resistance genes is disconcerting because it suggests that antibiotic usage may be selecting for the maintenance of virulence. A Tn*21*-like element (t-ST4) within the pSLT virulence plasmid of *S*. Typhimurium has been previously described (Villa and Carattoli 2005). t-ST4 was distinct from the Tn*21*-like composite element in pSLT-BT in that t-ST4 was composed of a Tn*21* plus a Tn*1696* type I integron. A distinct antibiotic resistance associated genetic island was reported in *S*. Typhimurium clinical isolates from Asturia, Spain (Herrero et al. 2008). This island is flanked by a toxin/antitoxin system and by genes predicted to encode an iron uptake system. A third antibiotic resistance locus was associated with the virulence plasmid of *S*. Brandenburg, but the complete sequence of this is not yet available (Martinez et al. 2007). Together these observations indicate that the virulence plasmid may act as a general platform for multiple and varied capture of antibiotic, heavy metal, and antiseptic resistance. The composite element detected in at least two sites in the pSLT plasmid is a further example of a high-throughput gene exchange platform facilitating the sampling of genes collected from other bacteria in the environment.

In conclusion, we have exploited genome information to investigate the properties of bacterial pathogens that are currently circulating in nature and that contribute to the disease burden in a clinical setting. The ST313 isolates described in this work are not the only *S. enterica* that can cause invasive disease in humans, but the evidence provided here indicates that they may be becoming a prominent clone that has adapted to this niche and that they are continuing to evolve by microevolution. This study illustrates some of the advantages of a whole-genome approach to such studies, particularly in defining the genetic basis of antibiotic resistance and virulence potential. In future studies we plan to exploit this genomic information to undertake detailed studies on the pathogenic potential of these strains and their epidemiology in the field in terms of transmission routes.

Methods

Bacterial strains and culture conditions

Isolates of *S*. Typhimurium from Blantyre, Malawi, were cultured from venous blood or cerebrospinal fluid of febrile patients at the Queen Elizabeth Central Hospital (QECH) from 1997 to 2006, as described previously (Gordon et al. 2008). Culture and identification of NTS were done using previously described protocols (Kariuki et al. 2006a). For conjugation experiments *Escherichia coli* TOP10 in which the *aph* gene was inserted into the *fhuA* gene was used as the recipient strain. Bacteria were routinely cultured aerobically at 37°C in Luria-Bertani (LB) broth or LB with 1% agar. *S*. Typhimurium strain D23580 was isolated from the blood of

a Malawian 26-mo-old child with malaria and anemia admitted to QECH in January 2004. Antibiotic sensitivity was determined by standard methodology (Supplemental text).

Sequence analysis

Sequence data for MLST and of shotgun libraries for whole-genome sequence determination (strain D23580) was determined using dye terminator chemistry on ABI3730 automated sequencers from PCR products amplified using oligonucleotide primers described previously (Kidgell et al. 2002). Genomic sequences of the *S*. Typhimurium strains A130 and 5579 were determined using a 454 Life Sciences (Roche) GS-FLX sequencer with standard protocols (Margulies et al. 2005). The nucleotide sequences of *S*. Typhimurium strain D23580 chromosome and pSLT-BT have been submitted to EMBL (http://www.ebi.ac.uk/embl/) under accession numbers FN424405 and FN432031, respectively. The raw pyrosequence data for A130 and 5579 have been submitted to the international Short-Read Archive (http://www.ncbi.nlm.nih.gov/Traces/sra/sra.cgi) under accession numbers ERA000075 and ERA000076, respectively. SNPs were identified by comparison with the LT2 finished genome sequence using MUMmer (Kurtz et al. 2004). SNP allele data from whole-genome sequence or MLST were used to fit a phylogenetic tree using PhyML (Guindon and Gascuel 2003) and drawn using a dendroscope. Additional descriptions of molecular biology techniques used are described in the Supplemental text.

References

Berkley JA, Lowe BS, Mwangi I, Williams T, Bauni E, Mwarumba S, Ngetsa C, Slack MP, Njenga S, Hart CA, et al. 2005. Bacteremia among children admitted to a rural hospital in Kenya. *N Engl J Med* **352:** 39–47.

Brent AJ, Oundo JO, Mwangi I, Ochola L, Lowe B, Berkley JA. 2006. *Salmonella* bacteremia in Kenyan children. *Pediatr Infect Dis J* **25:** 230–236.

Chain PS, Carniel E, Larimer FW, Lamerdin J, Stoutland PO, Regala WM, Georgescu AM, Vergez LM, Land ML, Motin VL, et al. 2004. Insights into the evolution of *Yersinia pestis* through whole-genome comparison with *Yersinia pseudotuberculosis*. *Proc Natl Acad Sci* **101:** 13826–13831.

Clemens J. 2009. Meeting on establishment of consortium to study invasive salmonellosis in sub-Saharan Africa. *Emerg Infect Dis* **15:** E2. doi: 10.3201/eid1507.090416.

Cooke FJ, Wain J, Fookes M, Ivens A, Thomson N, Brown DJ, Threlfall EJ, Gunn G, Foster G, Dougan G. 2007. Prophage sequences defining hot spots of genome variation in *Salmonella enterica* serovar Typhimurium can be used to discriminate between field isolates. *J Clin Microbiol* **45:** 2590–2598.

Cusa E, Obradors N, Baldoma L, Badia J, Aguilar J. 1999. Genetic analysis of a chromosomal region containing genes required for assimilation of allantoin nitrogen and linked glyoxylate metabolism in *Escherichia coli*. *J Bacteriol* **181:** 7479–7484.

Fujiwara S, Noguchi T. 1995. Degradation of purines: Only ureidoglycollate lyase out of four allantoin-degrading enzymes is present in mammals. *Biochem J* **312:** 315–318.

Gibbons HS, Lin S, Cotter RJ, Raetz CR. 2000. Oxygen requirement for the biosynthesis of the S-2-hydroxymyristate moiety in *Salmonella typhimurium* lipid A. Function of LpxO, A new Fe^{2+}/α-ketoglutarate-dependent dioxygenase homologue. *J Biol Chem* **275:** 32940–32949.

Gibbons HS, Reynolds CM, Guan Z, Raetz CR. 2008. An inner membrane dioxygenase that generates the 2-hydroxymyristate moiety of *Salmonella* lipid A. *Biochemistry* **47:** 2814–2825.

Gilks CF, Brindle RJ, Otieno LS, Simani PM, Newnham RS, Bhatt SM, Lule GN, Okelo GB, Watkins WM, Waiyaki PG, et al. 1990. Life-threatening bacteraemia in HIV-1 seropositive adults admitted to hospital in Nairobi, Kenya. *Lancet* **336:** 545–549.

Gordon MA, Banda HT, Gondwe M, Gordon SB, Boeree MJ, Walsh AL, Corkill JE, Hart CA, Gilks CF, Molyneux ME. 2002. Non-typhoidal *Salmonella* bacteraemia among HIV-infected Malawian adults: High mortality and frequent recrudescence. *AIDS* **16:** 1633–1641.

Gordon MA, Graham SM, Walsh AL, Phiri LW, Molyneux E, Zijlstra EE, Heyderman RS, Hart CA, Molyneux ME. 2008. Epidemics of invasive *Salmonella enterica* serovar enteritidis and *S. enterica* serovar typhimurium infection associated with multidrug resistance among adults and children in Malawi. *Clin Infect Dis* **46:** 963–969.

Graham SM, Molyneux EM, Walsh AL, Cheesbrough JS, Molyneux ME, Hart CA. 2000a. Nontyphoidal *Salmonella* infections of children in tropical Africa. *Pediatr Infect Dis J* **19:** 1189–1196.

Graham SM, Walsh AL, Molyneux EM, Phiri AJ, Molyneux ME. 2000b. Clinical presentation of non-typhoidal *Salmonella* bacteraemia in Malawian children. *Trans R Soc Trop Med Hyg* **94:** 310–314.

Guindon S, Gascuel O. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol* **52:** 696–704.

Guiney DG, Fang FC, Krause M, Libby S, Buchmeier NA, Fierer J. 1995. Biology and clinical significance of virulence plasmids in *Salmonella* serovars. *Clin Infect Dis* **21:** S145–S151.

Gulig PA, Danbara H, Guiney DG, Lax AJ, Norel F, Rhen M. 1993. Molecular analysis of spv virulence genes of the *Salmonella* virulence plasmids. *Mol Microbiol* **7:** 825–830.

Gunn JS, Belden WJ, Miller SI. 1998. Identification of PhoP-PhoQ activated genes within a duplicated region of the *Salmonella* typhimurium chromosome. *Microb Pathog* **25:** 77–90.

Herrero A, Mendoza MC, Rodicio R, Rodicio MR. 2008. Characterization of pUO-StVR2, a virulence-resistance plasmid evolved from the pSLT virulence plasmid of *Salmonella enterica* serovar typhimurium. *Antimicrob Agents Chemother* **52:** 4514–4517.

Holt KE, Parkhill J, Mazzoni CJ, Roumagnac P, Weill FX, Goodhead I, Rance R, Baker S, Maskell DJ, Wain J, et al. 2008. High-throughput sequencing provides insights into genome variation and evolution in *Salmonella* Typhi. *Nat Genet* **40:** 987–993.

Holt KE, Thomson NR, Wain J, Langridge GC, Hasan R, Bhutta ZA, Quail MA, Norbertczak H, Walker D, Simmonds M, et al. 2009. Pseudogene accumulation in the evolutionary histories of *Salmonella enterica* serovars Paratyphi A and Typhi. *BMC Genomics* **10:** 36. doi: 10.1185-1471-2164-10-36.

Kariuki S, Revathi G, Kariuki N, Kiiru J, Mwituria J, Hart CA. 2006a. Characterisation of community acquired non-typhoidal *Salmonella* from bacteraemia and diarrhoeal infections in children admitted to hospital in Nairobi, Kenya. *BMC Microbiol* **6:** 101. doi: 10.1186-1471-2180-6-101.

Kariuki S, Revathi G, Kariuki N, Kiiru J, Mwituria J, Muyodi J, Githinji JW, Kagendo D, Munyalo A, Hart CA. 2006b. Invasive multidrug-resistant non-typhoidal *Salmonella* infections in Africa: Zoonotic or anthroponotic transmission? *J Med Microbiol* **55:** 585–591.

Kidgell C, Reichard U, Wain J, Linz B, Torpdahl M, Dougan G, Achtman M. 2002. *Salmonella* typhi, the causative agent of typhoid fever, is approximately 50,000 years old. *Infect Genet Evol* **2:** 39–45.

Kingsley RA, Humphries AD, Weening EH, De Zoete MR, Winter S, Papaconstantinopoulou A, Dougan G, Baumler AJ. 2003. Molecular and phenotypic analysis of the CS54 island of *Salmonella enterica* serotype typhimurium: Identification of intestinal colonization and persistence determinants. *Infect Immun* **71:** 629–640.

Kurtz S, Phillippy A, Delcher AL, Smoot M, Shumway M, Antonescu C, Salzberg SL. 2004. Versatile and open software for comparing large genomes. *Genome Biol* **5:** R12. doi: 10.1186/gb-2004-5-2-r12.

Lawley TD, Chan K, Thompson LJ, Kim CC, Govoni GR, Monack DM. 2006. Genome-wide screen for *Salmonella* genes required for long-term systemic infection of the mouse. *PLoS Pathog* **2:** e11. doi: 10.1371/journal.ppat.0020011.

MacLennan CA, Gondwe EN, Msefula CL, Kingsley RA, Thomson NR, White SA, Goodall M, Pickard DJ, Graham SM, Dougan G, et al. 2008. The neglected role of antibody in protection against bacteremia caused by nontyphoidal strains of *Salmonella* in African children. *J Clin Invest* **118:** 1553–1562.

Maiden MC, Bygraves JA, Feil E, Morelli G, Russell JE, Urwin R, Zhang Q, Zhou J, Zurth K, Caugant DA, et al. 1998. Multilocus sequence typing: A portable approach to the identification of clones within populations of pathogenic microorganisms. *Proc Natl Acad Sci* **95:** 3140–3145.

Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, Berka J, Braverman MS, Chen YJ, Chen Z, et al. 2005. Genome sequencing in microfabricated high-density picolitre reactors. *Nature* **437:** 376–380.

Martinez N, Mendoza MC, Rodriguez I, Soto S, Bances M, Rodicio MR. 2007. Detailed structure of integrons and transposons carried by large conjugative plasmids responsible for multidrug resistance in diverse genomic types of *Salmonella enterica* serovar Brandenburg. *J Antimicrob Chemother* **60:** 1227–1234.

McClelland M, Sanderson KE, Spieth J, Clifton SW, Latreille P, Courtney L, Porwollik S, Ali J, Dante M, Du F, et al. 2001. Complete genome sequence of *Salmonella enterica* serovar Typhimurium LT2. *Nature* **413:** 852–856.

McClelland M, Sanderson KE, Clifton SW, Latreille P, Porwollik S, Sabo A, Meyer R, Bieri T, Ozersky P, McLellan M, et al. 2004. Comparison of genome degradation in Paratyphi A and Typhi, human-restricted serovars of *Salmonella enterica* that cause typhoid. *Nat Genet* **36:** 1268–1274.

Medini D, Donati C, Tettelin H, Masignani V, Rappuoli R. 2005. The microbial pan-genome. *Curr Opin Genet Dev* **15:** 589–594.

Miao EA, Brittnacher M, Haraga A, Jeng RL, Welch MD, Miller SI. 2003. *Salmonella* effectors translocated across the vacuolar membrane interact with the actin cytoskeleton. *Mol Microbiol* **48:** 401–415.

Murata T, Tseng W, Guina T, Miller SI, Nikaido H. 2007. PhoPQ-mediated regulation produces a more robust permeability barrier in the outer membrane of *Salmonella enterica* serovar typhimurium. *J Bacteriol* **189:** 7213–7222.

Parkhill J, Dougan G, James KD, Thomson NR, Pickard D, Wain J, Churcher C, Mungall KL, Bentley SD, Holden MT, et al. 2001a. Complete genome sequence of a multiple drug resistant *Salmonella enterica* serovar Typhi CT18. *Nature* **413:** 848–852.

Parkhill J, Wren BW, Thomson NR, Titball RW, Holden MT, Prentice MB, Sebaihia M, James KD, Churcher C, Mungall KL, et al. 2001b. Genome sequence of *Yersinia pestis*, the causative agent of plague. *Nature* **413:** 523–527.

Pennycott TW, Park A, Mather HA. 2006. Isolation of different serovars of *Salmonella enterica* from wild birds in Great Britain between 1995 and 2003. *Vet Rec* **158:** 817–820.

Rabsch W, Andrews HL, Kingsley RA, Prager R, Tschape H, Adams LG, Baumler AJ. 2002. *Salmonella enterica* serotype Typhimurium and its host-adapted variants. *Infect Immun* **70:** 2249–2255.

Raffatellu M, Santos RL, Verhoeven DE, George MD, Wilson RP, Winter SE, Godinez I, Sankaran S, Paixao TA, Gordon MA, et al. 2008. Simian immunodeficiency virus-induced mucosal interleukin-17 deficiency promotes *Salmonella* dissemination from the gut. *Nat Med* **14:** 421–428.

Thomson N, Baker S, Pickard D, Fookes M, Anjum M, Hamlin N, Wain J, House D, Bhutta Z, Chan K, et al. 2004. The role of prophage-like elements in the diversity of *Salmonella enterica* serovars. *J Mol Biol* **339:** 279–300.

Thomson NR, Clayton DJ, Windhorst D, Vernikos G, Davidson S, Churcher C, Quail MA, Stevens M, Jones MA, Watson M, et al. 2008. Comparative genome analysis of *Salmonella* enteritidis PT4 and *Salmonella* Gallinarum 287/91 provides insights into evolutionary and host adaptation pathways. *Genome Res* **18:** 1624–1637.

Valdivia RH, Falkow S. 1997. Fluorescence-based isolation of bacterial genes expressed within host cells. *Science* **277:** 2007–2011.

Villa L, Carattoli A. 2005. Integrons and transposons on the *Salmonella enterica* serovar typhimurium virulence plasmid. *Antimicrob Agents Chemother* **49:** 1194–1197.

Vissa VD, Brennan PJ. 2001. The genome of *Mycobacterium leprae*: A minimal mycobacterial gene set. *Genome Biol* **2:** Reviews1023.1. doi: 10.1186/gb-2001-2-8-reviews1023.

Vugia DJ, Kiehlbauch JA, Yeboue K, N'Gbichi JM, Lacina D, Maran M, Gondo M, Kouadio K, Kadio A, Lucas SB, et al. 1993. Pathogens and predictors of fatal septicemia associated with human immunodeficiency virus infection in Ivory Coast, West Africa. *J Infect Dis* **168:** 564–570.

Wood MW, Jones MA, Watson PR, Hedges S, Wallis TS, Galyov EE. 1998. Identification of a pathogenicity island required for *Salmonella* enteropathogenicity. *Mol Microbiol* **29:** 883–891.

Zhang S, Kingsley RA, Santos RL, Andrews-Polymenis H, Raffatellu M, Figueiredo J, Nunes J, Tsolis RM, Adams LG, Baumler AJ. 2003. Molecular pathogenesis of *Salmonella enterica* serotype typhimurium-induced diarrhea. *Infect Immun* **71:** 1–12.