Role of a single noncoding nucleotide in the evolution of an epidemic African clade of Salmonella

Disa L. Hammarlöva,b,1, Carsten Krögera,b,c,1, Siàn V. Owenb,1, Rocío Canalsb, Lizeth Lacharme-Lorab, Nicolas Wennera, Anna E. Schagerb, Timothy J. Wellsb, Ian R. Hendersonb, Paul Wigleyb, Karsten Hokampa, Nicholas A. Feaseyba,d,h, Melita A. Gordonb,2, and Jay C. D. Hintonb,2

aInstitute of Integrative Biology, University of Liverpool, L69 7TB Liverpool, United Kingdom; bDepartment of Cell and Molecular Biology, Uppsala University, 751 24 Uppsala, Sweden; cDepartment of Microbiology, School of Genetics and Microbiology, Mayne Institute of Preventive Medicine, Trinity College Dublin, Dublin 2, Ireland; dInstitute of Microbiology and Infection, University of Birmingham, B15 2TJ Birmingham, United Kingdom; eInstitute of Infection and Global Health, University of Liverpool, L69 7ZB Liverpool, United Kingdom; fDepartment of Genetics, School of Genetics and Microbiology, Smurfit Institute of Genetics, Trinity College Dublin, University of Dublin, Dublin 2, Ireland; gDepartment of Clinical Sciences, Liverpool School of Tropical Medicine, L3 5QA Liverpool, United Kingdom; hMalawi-Liverpool-Wellcome Trust Clinical Research Programme, University of Malawi College of Medicine, Blantyre 3, Malawi, Central Africa.

Edited by Roy Curtiss III, University of Florida, Gainesville, FL, and approved February 6, 2018 (received for review August 19, 2017)

Salmonella enterica serovar Typhimurium ST313 is a relatively newly emerged sequence type that is causing a devastating epidemic of bloodstream infections across sub-Saharan Africa. Analysis of hundreds of Salmonella genomes has revealed that ST313 is closely related to the ST19 group of S. Typhimurium that cause gastroenteritis across the world. The core genomes of ST313 and ST19 vary by only ∼1,000 SNPs. We hypothesized that the phenotypic differences that distinguish African Salmonella from ST19 are caused by certain SNPs that directly modulate the transcription of virulence genes. Here we identified 3,597 transcriptional start sites of the ST313 strain D23580, and searched for a gene-expression signature linked to pathogenesis of Salmonella. We identified a SNP in the promoter of the pgtE gene that caused high expression of the PgtE virulence factor in African S. Typhimurium, increased the degradation of the factor B component of human complement, contributed to serum resistance, and modulated virulence in the chicken infection model. We propose that high levels of PgtE expression by African S. Typhimurium ST313 promote bacterial survival and dissemination during human infection. Our finding of a functional role for an extragenic SNP shows that approaches used to deduce the evolution of virulence in bacterial pathogens should include a focus on noncoding regions of the genome.

Salmonella | noncoding genome | transcriptomics | evolution of virulence | host adaptation

The bacterial pathogen Salmonella enterica serovar Typhimurium (S. Typhimurium) is one of the best-understood pathogens and a major cause of gastroenteritis globally. One sequence type of S. Typhimurium, ST313, is the primary cause of invasive nontyphoidal Salmonellosis (INTS) across Africa, resulting in ∼388,000 deaths each year (1). Coinfection with HIV or malaria infection and young age (<5 y of age) are known risk factors for INTS infection (1, 2).

Multidrug resistance has contributed to the expansion of S. Typhimurium ST313. Whole-genome sequence-based phylogenetics revealed clonal replacement of ST313 lineage 1 by lineage 2 in the mid-2000s, accompanied by the acquisition of chloramphenicol resistance (3). The ST313 clade has recently acquired resistance to ceftriaxone, a first-line antibiotic for multidrug-resistant bacterial infections (4). Genomic comparison between the classical gastroenteritis-associated S. Typhimurium ST19 and the African ST313 isolates shows that gene content and synteny are highly conserved, that ST313 has a distinct repertoire of plasmids and prophages, and carries 77 pseudogenes reflecting a degree of genome degradation (5, 6). ST313 and ST19 share >4,000 genes, and their core genomes differ by ∼1,000 SNPs (5). We have reported that 2.7% of the S. Typhimurium isolated from patients in England and Wales are ST313, but lack the characteristic prophages BTP1 and BTP5 that are signatures of African ST313 lineages (7).

Certain virulence-associated phenotypes have been examined in ST313 strains. Compared with the ST19 group of gastroenteritis-associated S. Typhimurium, ST313 is more resistant to complement-mediated killing by human sera (8, 9) and to macrophage-mediated killing (10). ST313 exhibits a stealth phenotype during macrophage infection, consistent with an immune evasion strategy that causes reduced levels of IL-1β cytokine production, apoptosis, and caspase-1-dependent macrophage death (10, 11). Strategies used by ST313 to evade the innate immune system include reduced expression of both FliC and the SPI1 effector protein SopE2 (11). Furthermore, ST313 does not express the SseI effector protein due to pseudogenisation of the sseI gene, allowing ST313 to cause increased systemic infection via dendritic cell-mediated dissemination to extraintestinal sites in the murine infection model (12).

We used a functional genomic approach to search for SNPs responsible for the increased virulence of S. Typhimurium ST313 lineage 2.

Results

The reference strain for S. Typhimurium ST313 lineage 2 is D23580, which was isolated from an HIV+ Malawian child (5).

Significance

Invasive nontyphoidal Salmonella disease is a major and previously neglected tropical disease responsible for an estimated ∼390,000 deaths per year in Africa, largely caused by a variant of Salmonella Typhimurium called ST313. Despite the availability of >100,000 Salmonella genomes, it has proven challenging to associate individual SNPs with pathogenic traits of this dangerous bacterium. Here, we used a transcriptomic strategy to identify a single-nucleotide change in a promoter region responsible for crucial phenotypic differences of African S. Typhimurium. Our findings show that a noncoding nucleotide of the bacterial genome can have a profound effect upon the pathogenesis of infectious disease.
The strain 4/74 was isolated from a calf in the United Kingdom and is a well-characterized representative of S. Typhimurium ST19. Our challenge was to identify which, if any, of the >1,000 SNPs that separate strains D23580 and 4/74 serve to differentiate the strains in terms of gene expression and phenotype. We investigated whether the emergence of the epidemic clade...
of S. Typhimurium ST313 was linked to the altered expression of a core genome-encoded virulence factor. Rather than focusing on a comparison of the core genome, we used comparative transcriptomics to identify transcripts that were both expressed at different levels and associated with a distinct SNP in the promoter region.

This study built upon the primary transcriptome of S. Typhimurium ST19 strain 4/74, which we determined using a combination of RNA-seq and differential RNA-seq (dRNA-seq) under infection-relevant growth conditions (13, 14). By working at the single-nucleotide level, we defined transcriptional start sites (TSS), and cataloged the transcripts expressed in the bacterial cell (15, 16). Here, we used the same approach to define the primary transcriptome and to identify the TSS of D23580, a representative strain of African S. Typhimurium ST313. RNA was isolated from in vitro growth conditions that reflect the extracellular and intracellular stages of infection, namely the early stationary phase (ESP) and the SPI2-inducing condition (InSPI2) (Materials and Methods). To find all relevant TSS, a pooled sample containing RNA from 16 environmental conditions was also analyzed (14) (Materials and Methods). TSS were identified by comparison of mapped sequence reads from each pair of dRNA-seq and RNA-seq samples, as described previously (13, 14, 16). We identified 3,597 TSS for S. Typhimurium strain D23580, revealing the active gene promoters across the genome of an ST313 isolate for the first time. Previously, we reported the locations of 3,838 TSS for the ST19 strain 4/74 (13). Categorization of the TSS into different classes showed that a similar proportion of transcription initiation sites of 4/74 and D23580 were designated as primary (61%) or antisense (11%) (Fig. L4).

We determined the level of conservation of transcriptional organization between D23580 and 4/74 by identifying the TSS shared between the two strains. The locations of the majority of the TSS defined for strain 4/74 were conserved in strain D23580. Specifically, of the 3,838 TSS of strain 4/74, 390 were absent from D23580 and included TSS located in 4/74-specific regions such as prophages sopEΦ and Gifsy-1 (Dataset S1, tab g). We identified 63 D23580-specific TSS, mainly located in the BTP1 and BTP5 prophages of D23580, which are absent from strain 4/74 (5, 6) (Dataset S1, tab d).

To benchmark the transcriptional architecture, we first focused on Salmonella pathogenicity islands SPI1 and SPI2, which are required for key aspects of Salmonella virulence (17). The locations of all TSS within the SPI1 and SPI2 islands were identical in strains D23580 and 4/74 (Fig. 1B and Dataset S1, tab c). In summary, two closely related S. Typhimurium strains that varied by 1,488 SNPs at the core genome level had a high level of conservation at the transcriptional level and shared 90% of promoter regions (Dataset S1, tab a).

To address our hypothesis that the level of expression of certain virulence genes varied between strains D23580 and 4/74 due to changes at the DNA sequence level, we cross-referenced the SNP differences between the two strains with the locations of the TSS (Datasets S1 and S2). We identified 19 TSS that were associated with nucleotide polymorphisms in the −40 to −1 region of the 2,211 primary TSS of D23580 (Dataset S3). We compared the expression level of each TSS between 4/74 and D23580, in two growth conditions, to identify the promoter SNPs responsible for transcriptional changes. No SNPs were identified in the vicinity of the −10 or −35 elements of the fliC or sopE2 gene promoters, indicating that the reported differential expression of FliC and SopE2 between ST313 and ST19 (11) is not caused by alterations to the promoter sequence (Datasets S1–S3).

Identification of a Nucleotide That Modulates Expression of the PgtE Virulence Factor. PgtE is an outer membrane protease that belongs to the Omp15 family of Gram-negative bacteria (18), cleaves and mediates resistance to α-helical antimicrobial peptides, and disrupts the human complement cascade by degrading complement components, such as factor B (19, 20). PgtE does not contribute to intramacrophage replication per se, but stimulates bacterial dissemination during murine infection (21) by facilitating extracellular survival upon release from host cells (22–25). Expression of the pgtE transcript is induced during intramacrophage replication (26, 27), controlled by the SPI2-associated regulators PhoPQ and SnyA (19, 28), and is activated by OmpR/EnvZ and SsrA/B (15).

The promoter and coding regions of the pgtE gene were compared between D23580 and 4/74 at the DNA sequence level, and differed by two SNPs. One SNP was identified in the coding region of pgtE at nucleotide location 2,530,498 in D23580 (2,504,548 in 4/74), generating a synonymous mutation [T54 (ACT) in 4/74 → T54 (ACC) in D23580]. The other SNP was located in the promoter region; the −12 nucleotide (relative to the +1 of the TSS) was C in 4/74 (C107) and T in D23580 (T102) (Fig. 1C). This T nucleotide in the −10 motif is a highly conserved element of highly expressed σ54-dependent promoters (14). We analyzed the functional role of the T102 in the pgtE promoter region experimentally by replacing the T102 nucleotide with C107 by single-nucleotide exchange mutagenesis to generate strain D23580 ΔpgtE C107. Whole-genome sequencing confirmed that the D23580 ΔpgtE C107 strain only contained the intended single-nucleotide difference.

To determine the biological role of the T102 nucleotide, we assayed the level of pgtE transcript in 4/74, D23580, D23580 ΔpgtE C107, and D23580 ΔpgtE strains using qRT-PCR (Fig. 2A). The high level of pgtE expression in D23580 was reduced 10-fold by the introduction of the single C107 nucleotide in the −10 region of the pgtE promoter, P < 0.01 (Fig. 2A). The level of the pgtE transcript expression in 4/74 and the D23580 ΔpgtE C107 SNP mutant was similar.

We hypothesized that the high level of pgtE transcription would increase PgtE protein production in D23580 compared with wild-type strain 4/74, and this was observed (Fig. 2B). A second lineage 2 isolate, D37712, expressed the same high level of PgtE. In contrast, low levels of PgtE were produced by the ST313 lineage 1 isolates A130 and D25248, and ST19 isolates 14028 and LT2 (Fig. 2B). The enhanced production of PgtE by D23580 was reduced to the level of the 4/74 strain by a single-nucleotide change in strain D23580 ΔpgtE C107. To determine whether the T102 SNP would have a reciprocal effect in the S. Typhimurium ST19 genetic background, we introduced the T nucleotide to the −10 promoter motif of pgtE on the chromosome of strain 4/74 (Materials and Methods). The strain 4/74 ΔpgtE D23580 produced increased levels of PgtE protein (Fig. 2B). Taken together, our data show that in D23580, the high-level expression of pgtE at the transcriptional and protein level is driven by the T102 nucleotide in the −10 region of the pgtE promoter.

The T102 SNP Increases Resistance to Human Serum Killing, Modulates Cleavage of Complement Factor B, and Interferes with Complement Deposition. To determine the impact of the increased PgtE activity mediated by the promoter T102 SNP upon extracellular survival, we undertook serum bactericidal assays. Several bacterial factors contribute to the serum-resistance phenotype of Salmonella, including the long heterogenic O-antigen side chains of smooth lipopolysaccharide (LPS), which is the outermost component of the cell envelope of the Gram-negative cell (29–31). Therefore, we assayed resistance to human serum killing of in vitro-grown S. Typhimurium that lacked the LPS biosynthetic α1,3-glucosyltransferase enzyme WaaG. Following treatment with healthy human sera, the level of survival of D23580 ΔwaaG was significantly higher than D23580 ΔwaaG ΔpgtE C107 (P ≤ 0.05) (Fig. 3A). No killing was observed following treatment with heat-inactivated sera that lacked active complement, proving that the observed bactericidal activity was complement-mediated. In
The promoter T\textsuperscript{D23580} SNP increases the resistance of D23580 to serum killing, and the low level of pgtE expression driven by the P\textsuperscript{pgtE}\textsuperscript{4/74} promoter does not. To confirm the phenotypic role of the T\textsuperscript{D23580} SNP in wild-type S. Typhimurium D23580, we determined levels of complement deposition by flow cytometry. Following incubation with human sera, lower levels of complement component 3b (C3b) were found bound to the surface of wild-type D23580 than D23580 P\textsuperscript{pgtE}\textsuperscript{4/74} bacteria (Fig. 3B). Human serum killing of S. Typhimurium is mediated by activation of the classic complement pathway by the binding of specific IgG or IgM antibodies, followed by deposition of membrane attack complexes on the bacterial surface (32–34).

As described above, moving the T\textsuperscript{D23580} SNP into strain 4/74 caused increased PgtE protein production (Fig. 2B). To investigate the function of the T\textsuperscript{D23580} SNP in the S. Typhimurium ST19 genetic background, we examined the role of the T\textsuperscript{D23580} SNP upon complement deposition. The 4/74 wild-type bacteria had higher surface-bound levels of C3b than 4/74 P\textsuperscript{pgtE}\textsuperscript{4/74} D23580 (Fig. 3C), showing that increased expression of PgtE interferes with the complement cascade in both S. Typhimurium ST19 and ST313.

PgtE disrupts the complement cascade by proteolytically cleaving the C3b protein (35). The mechanism of the serum-resistance phenotype was investigated by determining the ability of the S. Typhimurium strains to mediate PgtE-dependent cleavage of complement factor B (Fig. 3D). In agreement with the literature (36), no PgtE activity was detected in strains expressing smooth LPS [Fig. 3D, lanes 1–3 (D23580), lanes 7–8 (4/74)]. Because the results of the serum-resistance assay showed that short (rough) LPS was required to visualize PgtE activity, we repeated experiments in a \textit{ΔwaaG} background (Fig. 3D, lanes 4–6 and 9–10), and discovered that the D23580 \textit{ΔwaaG} mutant showed a high level of complement factor B cleavage. In contrast, 4/74 \textit{ΔwaaG} and the D23580 \textit{ΔwaaG} P\textsuperscript{pgtE}\textsuperscript{4/74} strains showed a low level of complement factor B degradation.

We speculate that after the pathogen exits macrophages, the high level of expression of PgtE in S. Typhimurium ST313 strain D23580 interferes with opsonization and increases resistance to complement-mediated serum killing.

Assessment of PgtE-Mediated Virulence in the Chicken Infection Model. Because S. Typhimurium ST313 has a hyperinvasive phenotype during chicken infection (37), this animal model was used to assess the virulence of the wild-type D23580 and the D23580 P\textsuperscript{pgtE}\textsuperscript{4/74} strains. Following oral infection, the D23580 P\textsuperscript{pgtE}\textsuperscript{4/74} strain and the D23580 \textit{ΔpgtE} strain showed significant attenuation in splenic colonization in comparison with D23580 wild-type (\(P = 0.0035\) and \(P = 0.0379\), respectively), based on two independent repeats of the experiment (Fig. 3E). The data show some bird-to-bird variation between all three tested isolates, which is likely a consequence of the oral route of infection and the use of a commercial outbred chicken line. We found that the D23580 P\textsuperscript{pgtE}\textsuperscript{4/74} strain, the D23580 \textit{ΔpgtE} strain, and the wild-type D23580 strain colonized liver and cecum to similar levels (Fig. S1). Overall, the results showed that
Fig. 3. The pgtE promoter \(D^{23580}\) SNP mediates increased resistance to human serum killing, decreases complement component 3 deposition, enhances cleavage of human complement factor B, and promotes virulence in the chicken infection model. (A) Sensitivity to pooled healthy human serum was assayed in a \(\Delta\text{waaG}\) background (truncated LPS), to observe only the effect of outer membrane proteases. D23580 \(\Delta\text{waaG}\) showed significantly greater serum resistance than D23580 \(\text{P}_{\text{pgtE}} 4/74 \Delta\text{waaG}\) (\(P = 0.048\)). D23580 (B) or 4/74 (C) strains were incubated with pooled human sera and C3 deposition was evaluated by flow cytometry. Data shown is the average of five (B) or three (C) biological replicates, with error bars representing SD. (B) D23580 WT (orange) displayed decreased geometric mean fluorescence intensity (C3 deposition) compared with D23580 \(\text{P}_{\text{pgtE}} 4/74\) (light blue) after 5-min incubation with healthy human sera. (C) 4/74 WT (red) showed an increase of C3 deposition compared with 4/74 \(\text{P}_{\text{pgtE}} D^{23580}\) (dark blue) after 10-min incubation with healthy human sera. (D) PgtE-dependent cleavage of complement factor B detected by Western blotting. Polyclonal antibody against factor B was used. (E) Viable counts of S. Typhimurium D23580-derived strains as log CFU/g of spleen at 3-d postoral infection (10\(^8\) CFU) of 7-d-old Lohmann Brown Layers. Data based on 19 individually sampled birds for each group; combined data for two separately repeated experiments. Each symbol represents the value for an individual chicken and the bars represent the median value for each group.
exchange of the T<sup>D23580</sup> SNP to the C<sup>4/74</sup> genotype resulted in lower median bacterial numbers and a reduced number of animals with splenic infection, equivalent to that seen in the absence of PgtE (10 of 19 for D23580 compared with 2 of 19 for D23580 P<sub>PgtE</sub><sup>4/74</sup> and 4 of 19 for D23580 ΔpgtE). We conclude that the T<sup>D23580</sup> SNP promotes systemic infection of the chicken, and that full virulence of D23580 requires the high level of expression of PgtE that is driven by the T<sup>D23580</sup> nucleotide.

The <i>pgtE</i> Promoter SNP Is Only Carried by African ST313 Lineage 2, and Not Lineage 1. To determine if the <i>pgtE</i> promoter T<sup>D23580</sup> SNP is a characteristic feature of iNTS, the <i>pgtE</i> promoter SNP was analyzed in the context of a phylogeny of 258 genomes of <i>S. Typhimurium</i> ST313 including isolates from Malawi, as well as recently described UK-ST313 genomes from the UK (7). The 228 genomes that carried the T<sup>D23580</sup> SNP formed a monophyletic cluster that included lineage 2, as well as the UK-ST313 strains that share most recent common ancestry with lineage 2 (Fig. 4). The C<sup>4/74</sup> SNP was found to be conserved in all 27 lineage 1 genomes analyzed and certain UK-ST313 genomes which shared more recent common ancestry with lineage 1. This suggests that the T<sup>D23580</sup> SNP first arose in a common ancestor of lineage 2 and a subset of the UK-ST313 (Fig. S2).

The <i>pgtE</i> Promoter Is Highly Conserved in <i>S. enterica</i>. To understand the wider distribution of the <i>pgtE</i> promoter SNP in the <i>Salmonella</i> genus, the conservation of the SNP was assessed in 83 published complete genomes representing the known genomic diversity of <i>Salmonella</i>. The <i>pgtE</i> promoter sequence was not conserved in the three genomes of <i>Salmonella bongori</i> included in the analysis (Dataset S4). Of 80 <i>S. enterica</i> genomes screened, 79 genomes carried the C<sup>4/74</sup> genotype. The T<sup>D23580</sup> SNP was only found in <i>Salmonella</i> Gallinarum strain 287/91, raising the possibility that the SNP has arisen independently in this serovar (Fig. 4). Apart from the –12 SNP present in <i>S. Typhimurium</i> ST313 and <i>S. Gallinarum</i>, the <i>pgtE</i> TSS –35 region was found to be 100% conserved in 75 of 80 <i>S. enterica</i> genomes, with only serovar Agona and subspecies Arizonae showing sequence divergence (Dataset S4).

In summary, the –10 C<sup>4/74</sup> → T<sup>D23580</sup> allele of the D23580 <i>pgtE</i> promoter causes an increase in transcription of the <i>pgtE</i> gene and the production of high levels of PgtE protein in D23580. The increased activity of PgtE in D23580 leads to increased resistance to human serum killing, reduced complement deposition, and enhanced degradation of complement factor B, which is required for activation of the alternative complement pathway. Importantly, the single SNP in the D23580 <i>pgtE</i> promoter drives the ability of

![Fig. 4. Conservation of the <i>pgtE</i> promoter –10 T<sup>D23580</sup> nucleotide across <i>S. enterica</i> subsp. enterica. Maximum-likelihood phylogenetic tree of 336 <i>S. enterica</i> subsp. enterica genomes including 258 ST313 genomes. Presence of C or T nucleotide at the –10 position of the <i>pgtE</i> promoter is indicated by blue and red, respectively. The T nucleotide is found in 100% of ST313 lineage two genomes surveyed, and in a subset of UK-ST313. Across <i>S. enterica</i> subsp. enterica subspecies 1, the C nucleotide is present in 23 serovars, and a T nucleotide was identified in a single genome (<i>S. Gallinarum</i> isolate 287/91).](https://www.pnas.org/cgi/doi/10.1073/pnas.1714718115)
D23580 to cause hyperinvasion in an avian infection model. Outside of ST313, only 1 of 80 complete S. enterica genomes, that of S. Gallinarum, carried the −12 T to C allele. S. Gallinarum is the causal agent of fowl typhoid, suggesting a putative link between the T→C SNP and the ability of S. enterica to cause systemic infection (38). We have shown that the pgtE promoter SNP is a signature of ST313 lineage 2, which clonally replaced ST313 lineage 1 in the early 2000s (3). All isolates of ST313 lineage 2 that we have analyzed carried the same T→C SNP in the −10 promoter motif of pgtE.

Discussion

Previous studies have identified SNP mutations associated with the host tropism of notorious pathogens such as Staphylococcus aureus (39) and Campylobacter jejuni (40). However, these examples involved SNP mutations located within coding genes, and functionally important SNP mutations have rarely been identified in intergenic regions of bacteria. As the expression of a gene is dependent on the −10 and −35 recognition motifs of σ70-dependent promoters (41), a single-nucleotide change can modulate promoter function. Examples include the C → T transition in the −10 promoter motif of the Mycobacterium tuberculosis eis gene, which increases eis expression to generate low-level resistance to Kanamycin (42). Similarly, a promoter SNP that affected expression of Escherichia coli succinate transporter drrA evolved to increase the utilization of citrate as a carbon source in one population for the Long-Term Evolution Experiment (43).

Here, we have identified a single SNP responsible for high levels of expression of the PgtE outer membrane protease that is linked to the virulence of African S. Typhimurium ST313. Our study has implications for bacterial genome-wide association studies, which should clearly include a focus on noncoding regions of the genome. The findings also emphasize the value of identifying all gene promoters in bacterial pathogens, to allow nucleotide differences to be correlated with the process of transcriptional initiation.

It is known that S. Typhimurium bacteria produce high levels of pgtE transcript inside host macrophages (26, 27), and that PgtE protease activity is high in S. Typhimurium bacteria released from infected macrophages (22). The steric hindrance caused by long O-antigen chains makes the modification of LPS unacceptable in one population of the genome. During intramacrophage replication, resulting in shortening of the oligosaccharide chains (22). Accordingly, we used rough S. Typhimurium strains (ΔswaG) that express short O-antigen chains to determine that the pgtE T→C SNP contributes to the resistance of ST313 D23580 to human serum killing.

PgtE is known to interfere with the complement cascade in several ways, including the cleavage of factors B and C3b (35). To demonstrate a role for the pgtE T→C SNP in wild-type (smooth) bacterial strains, we show that after incubation with healthy human serum, the pgtE T→C SNP caused decreased C3b deposition on the Salmonella cell surface in both the 4/74 and D23580 backgrounds. We speculate that the pgtE T→C SNP primes intracellular bacteria for an extracellular lifestyle, and aids the survival of complement-mediated attack by the innate immune system.

The opsonic activity of complement has been shown to be essential for phagocyte-mediated killing of Salmonella in the blood of African people (44), and therefore our data are consistent with the hypothesis that subversion of host innate immunity via interference with the complement cascade contributes to the pathogenesis of invasive nontyphoidal Salmonella in Africa. Because other factors have been reported to enhance the invasiveness of S. Typhimurium ST313 (11, 12, 45), our data suggest that the adaption of African ST313 to a systemic lifestyle is multifactorial. Clearly, further investigation into the interaction between the nucleotide changes responsible for the pseudogenization of eis and the pgtE promoter SNP is required to understand which alteration drove the adaptation to invasive disease, and which could represent “passenger” mutations.

We suggest that the high level of PgtE activity in D23580, together with the inactivated eis effector gene (12) and the acquisition of chloramphenicol resistance (3), has been key to the success of S. Typhimurium ST313 lineage 2. This mechanism has a parallel in the evolution of Yersinia pestis, a pathogen that acquired the Pla Omptin protease, a PgtE homolog, at least 3,000 years ago (46). The Pla protein allowed Y. pestis to transition from a gastrointestinal to an extraintestinal lifestyle and cause plague (47). Here, we propose that the increased PgtE protease activity caused by the T→C SNP in the −10 motif of the pgtE promoter was an important evolutionary event that primed the emergence and clonal expansion of an epidemic S. Typhimurium lineage.

Materials and Methods

Bacterial Strains, Growth Conditions. S. Typhimurium strain 4/74 (accession no. CP002487), a representative strain of nontyphoidal Salmonella sequence type 19, and D23580 (accession no. FN424405), a representative strain of nontyphoidal S. enterica sequence type 313 (ST313) were used in the study. Strain D23580 was isolated from an HIV-1-infected Malawian patient with blood stream infection, and use of this strain has been approved by the Malawian College of Medicine (COMREC ethics no. P.08/14/1614). Other wild-type strains belonging to ST19 and ST313 used in this study are listed in Table S1.

All environmental growth conditions were prepared exactly as previously described (13), with the exception that the “pool” sample was obtained by pooling RNA from 16 environmental conditions: early exponential phase (EEP), mid exponential phase (MEP), late exponential phase (LEP), late stationary phase (LSP), 25 °C, NaCl shock, bile shock, low Fe2+ shock, anaerobic shock, anaerobic growth, oxygen shock, non-SP12, InSI2P, peroxide shock (InSI2P), and nitric oxide shock (InSI2P) (13).

When required, Lennox broth (LB) was supplemented with the following antibiotics: chloramphenicol (25 μg/mL), kanamycin (Km) 50 μg/mL, tetraacycline (Tc) 20 μg/mL, and gentamicin (Gm) 20 μg/mL.

Preparation of cDNA Libraries and Illumina Sequencing. Before RNA-seq, total RNA was extracted using TRIzol, and treated with RNase I, as described previously (14). RNA integrity was inspected visually with the Bioanalyzer (Agilent Technologies). Contaminating DNA was removed using RNase I (Ambion) and RNA samples were not ribo-depleted before cDNA library construction. cDNA library construction, TEX-treatment for dRNA-seq (ESP, InSI2P, and pooled sample), and RNA-seq with the Illumina HiSeq platform was carried out by Vertis Biotechnology, Germany. All protocols were identical to those used previously (13, 14). qPCR assays were performed against the reference genome S. Typhimurium D23580. A pathogen that acquired the Pla Omptin protease, a PgtE homolog, at least 3,000 years ago (46). The Pla protein allowed Y. pestis to transition from a gastrointestinal to an extraintestinal lifestyle and cause plague (47). Here, we propose that the increased PgtE protease activity caused by the T→C SNP in the −10 motif of the pgtE promoter was an important evolutionary event that primed the emergence and clonal expansion of an epidemic S. Typhimurium lineage.

Identification of TSS by a Combination of RNA-Seq and dRNA-Seq. Methods used to assign TSS in D23580 have been described previously (14). Briefly, a TSS was assigned when it was enriched in one of the dRNA-seq libraries (ESP, InSI2P, or Pool) compared with the corresponding RNA-seq library, and was linked to an expressed transcript. This analysis was followed by a second step of validation, in which the transcripts per million approach (50, 51) was used to calculate an expression value for the first 10 nucleotides associated with each TSS, designated the promoter usage value (PUV) (14, 27). A TSS was considered to be expressed when the PUV was ≥10. A TSS was defined as “conserved” between D23580 and 4/74 if the TSS nucleotide sequence was present in both strains, and the PUV value of the TSS was ≥10.

Identification of TSS-Located SNP Mutations Associated with Different Levels of Transcription Expression in D23580 and 4/74. A list of SNP differences between the D23580 and 4/74 reference genomes (accession nos. FN424405 and CP002487) was generated using NUCmer (52) resulting in 1,488 SNPs and small indels (Dataset S2). We identified the SNPs located within the −40 to −1 region of primary TSSs in D23580 (primary = primary + primary plus + primary/intermediate (Datasets S1 and S3). PUV values for each promoter in D23580 and 4/74 (14) were used to analyze the activity of the promoters associated with SNP differences in the −40 to −1 TSS region.
Bacterial Strain Construction Using \( \lambda \) Red Recombination. All of the bacterial strains and plasmids used and constructed in this study are described in Table S1 and the single-stranded DNA (ssDNA) oligonucleotides (primers) in Table S2.

The \( \Delta \)pgTE, \( \Delta \)waaG and pgte-FLAG mutation were constructed in \( S. \) Typhimurium ST319 and ST313 strains using the standard \( \lambda \) red recombination methodology (53). The heat-inducible \( \lambda \) red recombination plasmid pSIMs-tet was used and the induction of the \( \lambda \) red operon was achieved by heat treatment (42 °C, 15 min) of bacterial cultures grown to midexponential phase (OD\(_{600} \) 0.3–0.4 at 30 °C) in LB supplemented with Tc (33 μg/mL). After recombination (33–53), all of the genomic constructs (except the \( \Delta \)waaG::Kan mutation) were transferred into a clean wild-type background by phage transduction, using the P22 HT1051 intZ01 (56), as previously described (6). When required, the antibiotic resistance cassettes were flipped-out using the FLP recombinease expression plasmid pCP20-TRc (57).

The pgTE gene was deleted in \( S. \) Typhimurium using PCR fragments generated with the primers DH95 and DH96 and the template plasmids pKD4 and pKD3. The resulting fragments, carrying, respectively, the Km resistance (kan) or the Cm resistance (cam) cassettes were, respectively, electroporated into DC2580 and 4/74 carrying pSIMs-tet and recombinant \( \Delta \)gtgE::KanCam mutants were selected on Km or Cm LB agar plates. Finally, the mutations were transduced in the corresponding wild-type strain and the resistance cassette was removed, as described above. Similarly, waaG was inactivated using the primers del_waaG_F and del_waaG_R and pKD4 as template.

The FLAG-tagged strains were generated using a forward primer (DH93), which included the sequence encoding the FLAG octa-peptide, and a reverse primer (end pgg) that is identical to that used for construction of mutants using \( \lambda \) red recombination (described above), except that 400 ng of the manufactured primer (DH90, HPCL purified) was used in the transformation reaction. After 2 h of recovery at 30 °C, dilutions of transformation were plated on LB agar (without selection). Clones were restreaked and screened by a stringent PCR with primers DH40 and DH41. Primer DH40 has full complementarity with the sought after mutant, representing one mismatch to the original strain (these two types could be distinguished using a stringent primer set). Primer DH41 has full complementarity with both types of clone. The correct allele of the \( D23580 \) P\(_{gly} \)\(_64/74\) strain was confirmed by whole-genome sequencing using Illumina technology (MicrobesNG, University of Birmingham). Variant-calling analysis confirmed that the DC2580 P\(_{gly} \)\(_64/74\) strain had the intended single-nucleotide difference compared with the wild-type strain.

The single-nucleotide C→T substitution in the pgTE promoter of \( S. \) Typhimurium 4/74 (chromosomal position 2504765) (4/74 P\(_{pgTE} \) DC2580) was carried out by a scarless genome-editing technique based on the pEMG suicide plasmid, as previously described (6, 60). The pEMG derivative pNWA41 that carries the pgTE promoter region with the specific substitution was constructed as follows: the regions flanking the targeted nucleotide were PCR-amplified with the primers pairs NW22/NW123 and NW124/ NW125, using 4/74 genomic DNA as template. The primers NW122 and NW124 were used for T substitution and were complementary to each other over a stretch of 20 nucleotides. The resulting PCR fragments (504 and 505 bp, respectively) were fused by overlap extension PCR and the resulting 898-bp fragment was digested and cloned into pEMG using the BamHI and EcoRI restriction sites. The pNWA41 suicide plasmid was mobilized from E. coli S17-1 8pir into \( S. \) Typhimurium 4/74 by conjugation and trans-conjugants that have integrated pNWA41 by homologous recombination were selected on minimal medium M9 agar supplemented with 0.2% of glucose and Km. The resulting merodiploids were resolved using the pSWW-2 plasmid as previously described (6) and the C→T substitution was confirmed by PCR amplification and sequencing, using the primers NW155 and NW156.

\( qRT-PCR \). Total RNA was extracted from bacteria from midexponential (OD\(_{600} \) 0.3) cultures of bacteria grown in PCN/InSPI2, DNase-treated with Turbo DNA-free kit (Ambion) and the RNA integrity was inspected visually using the Bioanalyzer. Complete DNA-removal was confirmed by a negative PCR with 40 cycles. Four-hundred nanograms of RNA were converted into cDNA using the GoScript Reverse Transcription System (Promega), with random primers according to the manufacturer’s instructions. The Sensifast SYBR HI ROX Kit (Bioline) was used for \( qRT-PCR \) with primer pairs DH54/ DH55 and qPCR_hns_f/qPCR_hns_r. For each \( qRT-PCR \), performed in duplicate, 26.6 ng cDNA was used in total reaction volumes of 20 μL. The amount of pgtE and hns mRNA was calculated using a standard curve based on 10-fold dilutions of genomic DNA (10−0.0001 ng/μL), included in each \( qRT-PCR \) run. The amount of pgte mRNA was normalized to the amount of hns mRNA.

Western Blot Analysis. Salmonella strains carrying the FLAG-tagged version of pgte were grown in PCN/InSPI2 medium to OD\(_{600} \) 0.3. Bacteria were harvested from 10 μL of culture by centrifugation (7,000 × g, 5 min, 4 °C). Cells were washed once with PBS and resuspended in 67.5 μL of the same buffer. Seventy-five microliters of Laemmli buffer 2× [120 mM Tris-HCl pH 6.8, 4% (wt/vol) SDS, 20% (vol/vol) glycerol, Bromophenol blue 0.02% (wt/vol)] and 7.5 μL β-mercaptoethanol were subsequently added to the samples. The extracts were boiled for 10 min, chilled on ice, and cell debris was pelleted by centrifugation (20,000 × g, 5 min, 4 °C). Fifteen microliters of the samples (supernatants) were loaded on a 10% polyacrylamid gel. The gels were separated for 80 min at 150 V in SDS/PAGE running buffer (25 mM Tris, 192 mM glycine, 0.1% SDS). Proteins were transferred onto a methanol-soaked PVDF membrane (Cat. No. 3 010 040; Roche) using a semidry transfer system (#170–3940; Bio-Rad) for 2 h, 125 mA at 4 °C in transfer buffer (25 mM Tris, 192 mM glycine). The membrane was blocked for 15 h at 4 °C in Tris-buffered saline (TBS): 10 mM Tris-HCl pH 7.5, 0.9% NaCl) supplemented with 3% (wt/vol) of dry skimmed milk. Incubation with the antibody (1 h at room temperature) and washing steps (twice for 15 min each) were done in TBS containing 0.1% (vol/vol) Tween 20 and 0.5% (wt/vol) dry skimmed milk. The primary antibodies, monoclonal anti-FLAG M2 antibody (1:3,000 diluted, Sigma-Aldrich F3165) and anti-DnaK mAb 8E2/2 (1:10,000 diluted, ADI-SPA-880; Enzo Life Sciences), were used for the detection of PgE-FLAG and DnaK (loading control), respectively. After incubation with the primary antibodies, the membrane was washed and incubated with the secondary antibody [goat anti-mouse IgG (H+L)-HRP, 1:2,500 diluted, #102–1011; Bio-Rad]. After washes, the membrane was rinsed briefly in TBS before addition of Pierce ECL Western blotting substrate (32109; Thermo Scientific) and the chemiluminescence reaction was measured using the Image Quant LAS 4000 imager (GE Healthcare Life Sciences).

Complement Factor B Cleavage Assay. The complement factor B cleavage assay was carried out as described previously (20), with the following modifications: Bacterial strains were grown to midexponential (OD\(_{600} \) 0.3) cultures of bacteria grown at 37 °C. Overnight culture was pelleted by centrifugation and washed twice in PBS. Two OD\(_{600} \) (in 200 μL PBS containing 33 ng/μL of complement factor B) were incubated at 37 °C in a heat block with agitation (700 rpm) for 2 h and subsequently pelleted by centrifugation. Sixty-six nanograms of complement factor B were separated on an SDS-PAGE, before Western blotting, as described above.

Serum Bactericidal Assays. Serum bactericidal assays were performed four times using a modification of the previously described method (61). Briefly, bacteria were grown to OD\(_{600} \) = 2 in 5 mL of LB at 37 °C and resuspended in PBS to a final concentration of 10⁶ CFU/mL. 10 μL was then mixed with 90 μL of undiluted pooled healthy human sera at 37 °C with shaking (180 rpm), and viable counts determined. Killing was confirmed to be due to the activity of complement by using 56 °C heat-inactivated sera as a control.

Complement Deposition Assays. Detection of C3 complement component deposition was done as previously described (32–34). In brief, overnight cultures of S. Typhimurium were grown to an OD\(_{600} \) of 2. Cells were washed twice with PBS and 5 μL of S. Typhimurium at 2 × 10⁷ CFU/mL was mixed with 45 μL healthy human serum or PBS at room temperature for 5 min (D23580 strains) or 10 min (4/74 strains). Cells were then washed twice with 1 μL of PBS by 5 min centrifugation at 3,300 × g, and incubated with anti-human C3c FITC-conjugated antibody (Thermo Scientific) for 20 min at 4 °C. Samples were washed twice and fixed with 200 μL of formaldehyde-PBS. Cells were acquired on a BD FACSVerse and analyzed with FlowJo v10.4. Three to five replicate assays were conducted for each strain background, and the average geometric mean fluorescence intensity for replicates is shown.
along with the SD. An unpaired, single-tailed t test was used to assess the statistical significance of strain differences.

**Chicken Infection Experiments.** All work was conducted in accordance with United Kingdom legislation governing experimental animals under project license PPL 40/3652 and was approved by the University of Liverpool ethical review process before the award of the license. All birds were checked a minimum of twice daily to ensure their health and welfare. Birds were housed in accommodation meeting United Kingdom legislation requirements. One-day-old Lohmann Brown Layers were obtained from a commercial hatchery, separated into groups on arrival, and given ad libitum access to water and a laboratory-grade vegetable protein-based pellet diet (SDS). Chicks were housed at a temperature of 30 °C. At 7 d of age, chickens were inoculated by oral gavage with 10⁶ CFU of S. Typhimurium strains D23580, D23580 PGTE+, or D23580 ΔpargE. At 3 post infection, 10 birds from each group were killed for postmortem analysis. Samples from spleen, liver, and the caecal contents were removed aseptically from each bird and diluted 1:5 (wt/vol.) in sterile PBS (data from caecal content and liver are shown in Fig. S1). Tissues were then homogenized in a Colworth 80 microstomacher (A.J. Seward & Co. Ltd). Samples were serially diluted and dispensed onto Brilliant green agar (Oxoid) to quantify numbers of *Salmonella*, as described previously (26). – A–Mann–Whitney test was used to determine the statistical significance of strain differences.

**Analysis of the Conservation of the pgtE Promoter SNP.** For analysis of the conservation of the pgtE promoter SNP, 258 genomes of *S. Typhimurium* serovar Typhimurium from Malawi and the United Kingdom (7) were assembled using the AS pipeline (63) and Abacrus (64), unless reference sequence genomes were available (all sequences and accession numbers are given in Dataset S4). Additionally, 83 reference genomes were downloaded from the National Center for Biotechnology Information that represent the known diversity of the *Salmonella* genus. Core genome SNPs were identified using the PanSeq package (65) and a maximum-likelihood phylogenetic tree was constructed from the concatenated SNP alignment using PhyML (66). BLASTn was used to identify the genotype of the entire pgtE TSS – 35 region nucleotides in all genomes (Dataset S4).

**ACKNOWLEDGMENTS.** We thank present and former members of the J.C.D.H. laboratory for helpful discussions, particularly Satheesh Sivasankaran and Aoife Colgan; and Paul Loughnane for his expert technical assistance. This work was supported by funding from a Wellcome Trust Senior Investigator award (to J.C.D.H.) (Grant 106914/Z/15/Z). D.L.H. was supported by the Wenner-Gren Foundation, Sweden. R.C. was supported by an EU Marie Curie International Incoming Fellowship (FP7-PEOPLE-2013-IIF, Project Reference 628450). N.W. was supported by an Early Postdoc Mobility Fellowship from the Swiss National Science Foundation (Project Reference 2PLAP3_158804).
Salmonella as a severe respiratory pathogen.

Hammarlöf et al. | www.pnas.org/cgi/doi/10.1073/pnas.1714718115

Mol Microbiol

Typhi-

serovar Ty-

25:2730

31:587

163:571

64:

Virchow isolates from human

J Bacteriol

80:

Mol Microbiol

1:784

determines composition and length of the O-antigen of the lipopolysac-

Mol Gen Genet

Yersinia pestis

211:

Typhimurium ST313.

Salmonella enterica

Environ Microbiol

Typhimurium

in Eurasia

25:1968

43. Blount ZD, Barrick JE, Davidson CJ, Lenski RE (2012) Genomic analysis of a key in-


407:109

379:109


44. Gondwe EN, et al. (2010) Importance of antibody and complement for oxidative burst


49. Richardson EJ, et al. (2011) Genome sequences of


46. Rasmussen S, et al. (2015) Early divergent strains of


42. Leong SW, et al. (2013) Analysis of the pathogenicity of Salmonella enterica subspecies enterica serotype Typhi in humans and mice.


34. Leong SW, et al. (2013) Analysis of the pathogenicity of Salmonella enterica subspecies enterica serotype Typhi in humans and mice.


