METHODS AND RESOURCES

Adding function to the genome of African *Salmonella* Typhimurium ST313 strain D23580

Rocío Canals[1], Disa L. Hammarlöf[1¤a], Carsten Kröger[1¤b], Siân V. Owen[1¤c], Wai Yee Fong[1], Lizeth Lacharme-Lora[1], Xiaojun Zhu[1], Nicolas Wenner[1], Sarah E. Carden[2], Jared Honeycutt[2], Denise M. Monack[2], Robert A. Kingsley[3], Philip Brownridge[1], Roy R. Chaudhuri[4], Will P. M. Rowe[1¤d], Alexander V. Predeus[1], Karsten Hokamp[5], Melita A. Gordon[6,7], Jay C. D. Hinton[1]*

1 Institute of Integrative Biology, University of Liverpool, Liverpool, United Kingdom, 2 Department of Microbiology and Immunology, Stanford University School of Medicine, Stanford, California, United States of America, 3 Quadram Institute Bioscience, Norwich Research Park, Norwich, United Kingdom, 4 Department of Molecular Biology and Biotechnology, University of Sheffield, Sheffield, United Kingdom, 5 Department of Genetics, School of Genetics and Microbiology, Smurfit Institute of Genetics, Trinity College Dublin, Ireland, 6 Institute of Infection and Global Health, University of Liverpool, Liverpool, United Kingdom, 7 Malawi-Liverpool-Wellcome Trust Clinical Research Programme, University of Malawi College of Medicine, Malawi, Central Africa

¤a Current address: Department of Cell and Molecular Biology, Uppsala University, Uppsala, Sweden
¤b Current address: Department of Microbiology, School of Genetics and Microbiology, Moyne Institute of Preventive Medicine, Trinity College Dublin, Ireland
¤c Current address: Department of Biomedical Informatics, Harvard Medical School, Boston, Massachusetts, United States of America
¤d Current address: Scientific Computing Department, STFC Daresbury Laboratory, Warrington, United Kingdom
* jay.hinton@liverpool.ac.uk

🔓 OPEN ACCESS

Abstract

Salmonella Typhimurium sequence type (ST) 313 causes invasive nontyphoidal *Salmonella* (iNTS) disease in sub-Saharan Africa, targeting susceptible HIV⁺, malarial, or malnourished individuals. An in-depth genomic comparison between the ST313 isolate D23580 and the well-characterized ST19 isolate 4/74 that causes gastroenteritis across the globe revealed extensive synteny. To understand how the 856 nucleotide variations generated phenotypic differences, we devised a large-scale experimental approach that involved the global gene expression analysis of strains D23580 and 4/74 grown in 16 infection-relevant growth conditions. Comparison of transcriptional patterns identified virulence and metabolic genes that were differentially expressed between D23580 versus 4/74, many of which were validated by proteomics. We also uncovered the *S*. Typhimurium D23580 and 4/74 genes that showed expression differences during infection of murine macrophages. Our comparative transcriptomic data are presented in a new enhanced version of the *Salmonella* expression compendium, SalComD23580: http://bioinf.gen.tcd.ie/cgi-bin/salcom_v2.pl. We discovered that the ablation of melibiose utilization was caused by three independent SNP mutations in D23580 that are shared across ST313 lineage 2, suggesting that the ability to catabolize this carbon source has been negatively selected during ST313 evolution. The data revealed a novel, to our knowledge, plasmid maintenance system involving a plasmid-encoded CysS cysteinyl-tRNA synthetase, highlighting

Abbreviations: bp, base pairs; CDS, coding sequence; Cm, chloramphenicol; CPM, counts per million; EEP, early exponential phase; ESP, early stationary phase; FC, fold-change; FDR, false discovery rate; Gm, gentamicin; InSPI2, SPI-2-inducing; iNTS, invasive nontyphoidal *Salmonella*; Km, kanamycin; LB, Lennox broth; LC-MS/MS, liquid chromatography-tandem mass spectrometry; LEP, late exponential phase; LPS, lipopolysaccharide; LSP, late stationary phase; MEP, middle exponential phase; MNP, multinucleotide polymorphism; Nal, nalidixic acid; ncRNA, noncoding RNA; NonSPI2, SPI-2-noninducing; OD, optical density; PCN, phosphate carbon nitrogen; PNK, polynucleotide kinase; RNA-seq, RNA sequencing; SCV, *Salmonella*-containing vacuole; SNP, single-nucleotide polymorphism; SPI, *Salmonella* pathogenicity island; sRNA, small RNA; ST, sequence type; Tc, tetracycline; TIS, transposon-insertion sequencing; TPM, transcripts per million; tRNA, transfer RNA; TSS, transcriptional start site; UTR, untranslated region; WT, wild type.

the power of large-scale comparative multicondition analyses to pinpoint key phenotypic differences between bacterial pathovariants.

Author summary

Invasive nontyphoidal *Salmonella* (iNTS) is associated with a major and largely unreported tropical disease that is responsible for hundreds of thousands of deaths per year in Africa. The main causative agent is a pathovariant of *Salmonella* Typhimurium called ST313, which is closely related to the well-characterized ST19 sequence type of *Salmonella* that causes gastroenteritis globally. ST313 and ST19 vary by just 856 core genome single-nucleotide polymorphisms (SNPs). To understand how genetic changes generate phenotypic and mechanistic differences between African and global *Salmonella*, we used functional transcriptomic and proteomic approaches. By investigating the transcriptome of African and global *S.* Typhimurium in 17 growth conditions, we discovered that 677 genes and small RNAs were differentially expressed between strains D23580 (ST313) and 4/74 (ST19). A parallel proteomic approach linked gene expression differences to alterations at the protein level. We also identified differentially expressed genes during the actual infection of murine macrophages. Our data revealed the genetic basis of the loss of a carbon source utilization in African *Salmonella* and the discovery of a new mechanism for maintaining a plasmid in a bacterial population involving a plasmid-encoded essential bacterial gene.

Introduction

S. enterica serovar Typhimurium (*S.* Typhimurium) infects a wide range of animal hosts and generally causes self-limiting gastroenteritis in humans. Variants of this serovar, belonging to sequence type (ST) 313, are associated with invasive nontyphoidal *Salmonella* (iNTS) disease in susceptible HIV[+], malaria-infected, or malnourished individuals in sub-Saharan Africa [1]. iNTS causes around 681,000 deaths per year worldwide, killing 388,000 people in Africa alone [2]. The multidrug resistance of ST313 isolates complicates patient treatment and accounts for the high case fatality rate (20.6%) of iNTS disease [3]. Two ST313 lineages have been associated with iNTS, and the clonal replacement of lineage 1 by lineage 2 is hypothesized to have been driven by the gain of chloramphenicol (Cm) resistance by lineage 2 [4]. Genetically distinct ST313 isolates that do not belong to lineages 1 and 2 have been described in the United Kingdom [5] and in Brazil [6].

The globally distributed *S.* Typhimurium ST19 causes gastroenteritis in humans and invasive disease in mice. Following oral ingestion, these bacteria colonize the gut and stimulate inflammation by a *Salmonella* pathogenicity island (SPI)-1-mediated process. Subsequently, ST19 can survive and proliferate in a "*Salmonella*-containing vacuole" (SCV) within epithelial cells or macrophages that involves the SPI-2 type three secretion system responsible for systemic disease in mammalian hosts [7]. Host restriction of other *Salmonella* pathovariants has been associated with genome degradation caused by pseudogene formation [8–11]. This process involves the loss or inactivation of virulence genes required for colonization of the mammalian gut while the ability to thrive inside macrophages is maintained.

Phenotypic differences between ST313 and ST19 have been summarized previously [12], and new studies have since been published. S1 Table lists 20 phenotypic features that

differentiate ST313 from ST19 isolates at the level of metabolism, motility, and stress resistance [13–24]. In terms of infection biology, reports of the relative ability of ST313 and ST19 isolates to invade epithelial cells and macrophages have yielded conflicting results (S1 Table) [6,13,15,17,25–27]. It is clear that ST313 infection of macrophages stimulates lower levels of cytotoxicity and inflammasome response than ST19 infections [13,25]. Following treatment with human serum, more complement was required for antibody-mediated bactericidal killing of ST19 than for ST313 isolates [14]. Animal infection experiments have demonstrated that ST313 isolates can infect nonhuman hosts, including mice, cows, chickens, and macaques [15–18,28,29]. Taken together, these findings confirm that ST313 is a distinct pathovariant of *S.* Typhimurium [30]. However, the molecular mechanisms responsible for the phenotypic signature of the ST313 pathovariant remain to be understood and require a bespoke experimental approach.

D23580 is the ST313 lineage 2 reference strain, a typical representative Malawian strain isolated from an HIV-negative child in 2004 [19]. We previously defined the transcriptional start sites (TSSs) of this strain and identified a SNP in the promoter of the *pgtE* gene specific to ST313 lineage 2 that modulated virulence [20]. To investigate whether the ability of ST313 and ST19 of *S.* Typhimurium to cause different types of human disease was a genetic characteristic of the two types of bacteria, we identified all genomic differences between D23580 and 4/74. We then generated a comprehensive dataset for studying the mechanisms of infection-relevant differences between ST313 and ST19 listed in S1 Table. We hypothesized that transcriptional differences between the two strains would account for specific phenotypic differences, and we present a multicondition transcriptomic comparison of the ST313 strain, D23580, with the ST19 strain, 4/74 (S1 Fig).

Results

Resequencing and reannotation of D23580, the *S.* Typhimurium ST313 reference strain

S. Typhimurium D23580 was the first ST313 isolate to be genome sequenced [19]. At that time, the presence of one D23580-specific plasmid, pBT1, was reported. To facilitate a robust transcriptomic analysis of D23580, we resequenced the strain using a combination of long-read PacBio and short-read Illumina technologies. Following a hybrid assembly approach (Materials and Methods), three contigs were identified: the 4,879,402 base pair (bp) chromosome, the 117,046 bp pSLT-BT plasmid, and the 84,543 bp pBT1 plasmid (accession: PRJEB28511). Comparison with the published D23580 genome (accession: FN424405) [19] identified just three nucleotide differences in the chromosome. Specifically, an extra nucleotide at the 304,327 position (1 bp downstream of Asp-transfer RNA [tRNA]), at the 857,583 position (1 bp upstream of Lys-tRNA), and one nucleotide change at position 75,492 (T-to-C; intergenic region) were identified. The sequence of the pSLT-BT plasmid had a single-nucleotide deletion difference at position 473 in an intergenic region. The sequence of the pBT1 plasmid has not been reported previously, and a primer-walking approach was used to sequence the two remaining small plasmids carried by D23580 (Materials and Methods), pBT2 and pBT3 (2,556 bp and 1,975 bp, respectively) (accession: PRJEB28511).

To maximize the functional insights to be gained from a transcriptomic analysis, a well-annotated genome is required. The published annotation for D23580 dates back to 2009 [19] and lacked certain essential bacterial genes such as the genes encoding the two outer membrane proteins LppA and LppB [31]. Accordingly, we searched for important nonannotated bacterial genes and used D23580 transcriptomic data (described below) to cross-reference the locations of transcripts with the location of coding genes (S1 Text). This analysis allowed us to

update the published annotation of D23580 by adding 86 new coding genes and 287 small RNAs (sRNAs) and correcting the start or end locations of 13 coding genes (S2 Table). The resequenced and reannotated *S*. Typhimurium D23580 genome is subsequently referred to as D23580_liv (accession: PRJEB28511).

The *S*. Typhimurium D23580 and 4/74 genomes are 95% identical

Previously, the D23580 genome had been compared with the attenuated laboratory *S*. Typhimurium LT2 strain [19,32,33]. To assess the similarities and differences between the ST313 strain D23580 and a virulent ST19 isolate, a detailed comparative genomic analysis was performed against the ST19 strain 4/74 (S1 Text). 4/74 is a prototrophic *S*. Typhimurium ST19 strain that is highly virulent in four animal models [34] and is the parent of the widely used SL1344 auxotrophic strain [35]. D23580 and 4/74 share 92% and 95% of coding genes and sRNAs, respectively (S2 Table). Genetic differences included 788 SNPs, three multinucleotide polymorphisms (MNPs) [36], and 65 indels, as well as 77 D23580-specific pseudogenes that have been listed elsewhere [19]. Analysis of the SNPs, using the 4/74 annotation as a reference, showed that 379 were nonsynonymous, 255 were synonymous, six were located in sRNAs, nine generated stop codons in coding genes, and seven lost stop codons in D23580 (S3 Table). The final 132 SNPs were in intergenic regions. Fig 1 compares the chromosome and pSLT plasmid organization of strains 4/74 and D23580 and shows the distribution of the indels and three SNP classes that differentiate the two strains. We recently reported the presence of 3,597 TSSs in D23580 and discussed the key differences in comparison with the 3,838 TSSs identified in 4/74 [20,37]. Seventeen of the SNPs and indels were located ≤40 nucleotides upstream of one of the D23580 TSSs [20], raising the possibility of a direct influence upon the level of transcription.

Regarding prophage complement, SopEφ [38] was absent from D23580 and present in strain 4/74 [19,24]. As we established earlier, D23580 carries two ST313-specific prophages, BTP1 and BTP5 [5,19,24]. In terms of plasmids, the genome of 4/74 includes pSLT$^{4/74}$, pCol1B9$^{4/74}$, and pRSF1010$^{4/74}$ [35]. In contrast, D23580 carries a distinct plasmid complement, namely, pSLT-BT, pBT1, pBT2, and pBT3 [19]. The pSLT-BT plasmid of D23580 carries a Tn*21*-based insertion element that encodes resistance to five antibiotics [19].

The D23580 and 4/74 strains carry 4,396 orthologous coding genes (S1 Text). Ten of the orthologs were encoded by the D23580-specific prophages BTP1 and BTP5 or by the 4/74-specific pRSF1010$^{4/74}$ plasmid and so were excluded from further analysis. A total of 279 orthologous sRNAs were found in both strains (S2 Table). The sRNA-associated differences included three 4/74-specific sRNAs (STnc3640, STnc1400, and STnc3800), and the duplication of IsrB-1 in D23580. Eight new sRNAs were found in the BTP1 prophage region of D23580, and the existence of four was confirmed by northern blot (S2 Fig).

We identified 93 D23580-specific chromosomal genes that were encoded within prophage regions and absent from 4/74 (S2 Table): specifically, 59 BTP1 genes, 27 BTP5 genes, one Gifsy-2 gene, and six Gifsy-1 genes. We found 89 chromosomal genes that were 4/74-specific and were absent from D23580 (S2 Table). Most were associated with the SopEφ prophage region (68 genes) or located in the Def2 remnant phage (13 genes) or in three separate non-phage–associated regions in D23580: *allB* (associated with allantoin utilization), the SPI-5 genes *orfX* and *SL1344_1032*, and an approximately 4-kb deletion that included genes *SL1344_1478* to *SL1344_1482*.

A total of 4,675 orthologous coding genes and noncoding sRNAs were shared by strains D23580 and 4/74. The sRNA IsrB-1 was removed from the list of orthologs because it was duplicated in D23580. To search for a distinct transcriptional signature of D23580, the
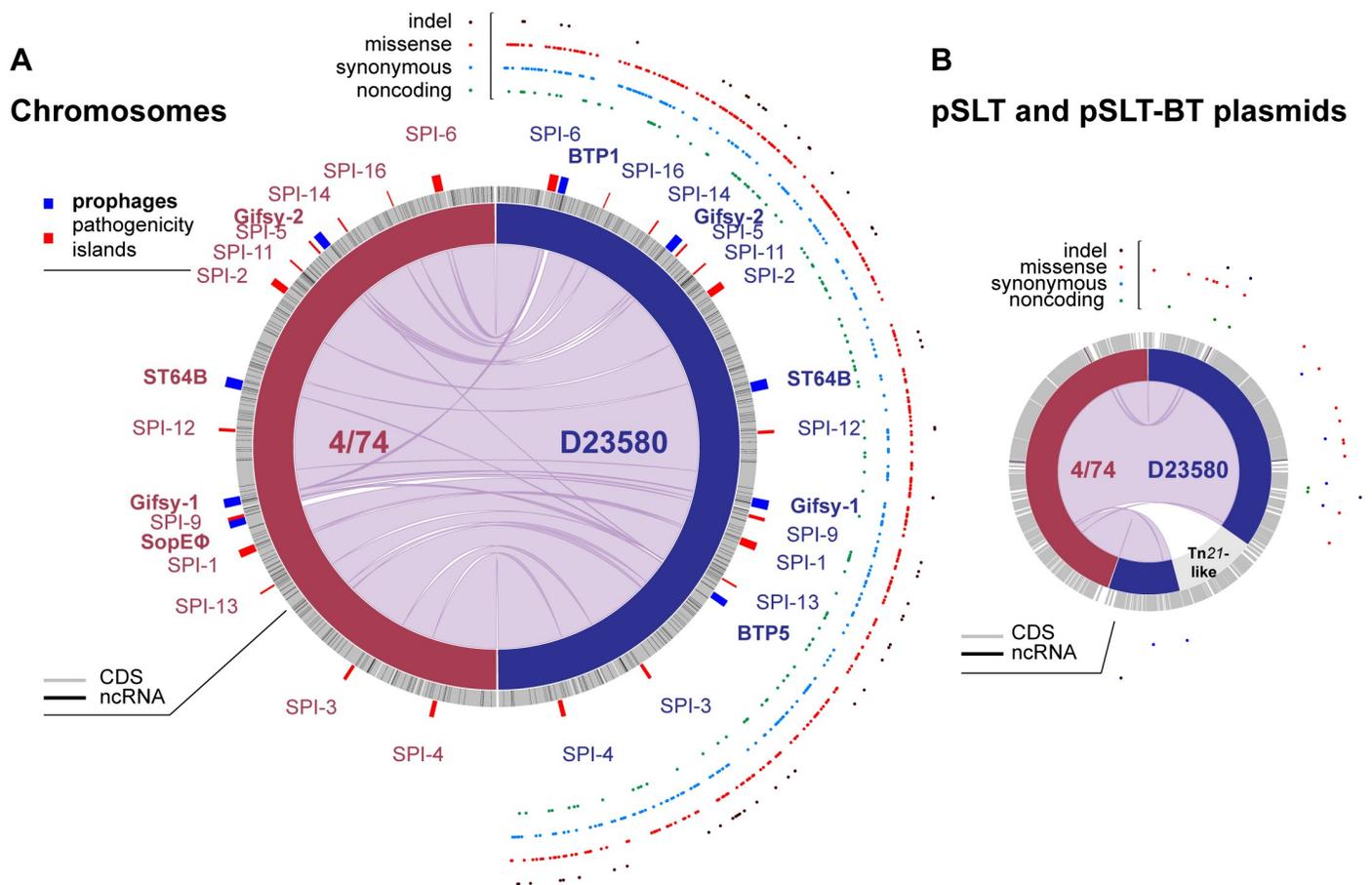
Fig 1. Comparative genomic analysis between S. Typhimurium 4/74 and D23580. Plots were obtained using the Circa software (http://omgenomics.com/circa/). (A) 4/74 and D23580 chromosomes; (B) 4/74 pSLT and D23580 pSLT-BT plasmids. In both panels, 4/74 data are represented on the left and D23580 data on the right. The four functional types of variants between D23580 and 4/74 are shown on the right-hand side of each panel (S3 Table). CDS and ncRNAs for *S.* Typhimurium D23580 are detailed in S2 Table and have already been reported for 4/74 in Kröger and colleagues [35]. CDS, coding sequence; ncRNA, noncoding RNA; SPI, *Salmonella* pathogenicity island; ST, sequence type.

https://doi.org/10.1371/journal.pbio.3000059.g001

expression levels of the 4,674 orthologs was compared between D23580 and 4/74 using a transcriptomic approach.

Comparison of transcriptional response to infection-relevant stress between S. Typhimurium ST313 D23580 and ST19 strain 4/74

To discover the similarities and differences in the transcriptome of strains D23580 and 4/74, we first used our established experimental strategy: the transcriptome of D23580 was determined using RNA isolated from 16 infection-relevant in vitro growth conditions [37] and during intra-macrophage infection [39,40]. To allow direct comparison of the D23580 transcriptomic data with previously published 4/74 data, experiments were performed exactly as Kröger and colleagues [37] and Srikumar and colleagues (Materials and Methods) [40].

The RNA-sequencing (RNA-seq)-derived reads were mapped to the D23580_liv chromosome and the pSLT-BT, pBT1, pBT2, and pBT3 plasmid sequences (Materials and Methods). Numbers of mapped sequence reads and other RNA-seq-derived statistical information are

detailed in S4 Table. The level of expression of individual genes and sRNAs was calculated as transcripts per million (TPM) [41,42] for the chromosome and the pSLT-BT and pBT1 plasmids (S5 Table). To achieve a complete transcriptomic comparison, we first reanalyzed our published 4/74 transcriptomic data [37,40] to add all transcripts expressed by the three plasmids pSLT[4/74], pCol1B9[4/74], and pRSF1010[4/74] (Materials and Methods, S4 Table).

Initial analysis focused on the expression characteristics of the strains D23580 and 4/74 in 17 distinct environmental conditions. The number of genes and sRNAs expressed in at least one condition for strain D23580 was 4,365 (85%) out of 5,110. 745 genes and sRNAs (15%) were not expressed in any of the 17 conditions. For strain 4/74, the number of genes and sRNAs that were expressed in at least one condition was 4,306 (86%) out of 5,026, consistent with our earlier findings [37] (S3 Fig). 3,958 of the 4,674 orthologous coding genes and sRNAs shared by strains D23580 and 4/74 were expressed in at least one growth condition in both strains.

A small minority (117) of orthologous genes were expressed in at least one condition in strain 4/74 but not in any of the conditions in D23580, with most showing low levels of expression (close to the threshold TPM = 10) (S5 Table). In contrast, we identified 82 orthologous coding genes and sRNAs that were expressed in at least one of the 17 growth conditions for D23580 but not expressed in 4/74 (S5 Table).

To compare the expression profiles of D23580 and 4/74, we made 17 individual pairwise comparisons between the 17 growth conditions with the two strains (Materials and Methods, S3 Fig). The data confirmed that *S.* Typhimurium reacts to particular infection-relevant stresses with a series of defined transcriptional programs that we detailed previously [37]. By comparing the transcriptomic response of two pathovariants of *S.* Typhimurium, the conservation of the transcriptional response is apparent (S3 Fig).

A complementary analytical approach was used to identify the transcriptional differences that relate to the distinct phenotypes of the ST313 and ST19 pathovariants (S1 Table). Overall, 1,031 of the orthologous coding genes and sRNAs were differentially expressed ($\geq$3 fold-change) between strains D23580 and 4/74 in at least one growth condition (Fig 2A, S5 Table). Transcriptional differences are highlighted in S3 Fig.

The terms "D23580-up-regulated" and "D23580-down-regulated" refer to genes that show a higher or lower level of expression in D23580 compared to 4/74. Three coding genes were D23580-up-regulated, and six genes were D23580-down-regulated in almost all growth conditions (Fig 2B and 2C). The up-regulated genes included *pgtE*, a gene that is highly expressed in D23580, responsible for resistance to human serum killing and linked to virulence [20]. The other two up-regulated genes were *nlp*, encoding a *ner*-like regulatory protein, and the *STM2475* (*SL1344_2438*) gene, which encodes a hypothetical protein. Two sRNAs, STnc3750 and STnc2050, were D23580-up-regulated. STnc3750 overlaps, and is transcribed in the same direction, with the last 32 nucleotides of the 3′ end of *pgtE* and is up-regulated in strain 4/74 in the intra-macrophage environment [40]. The function of these two Hfq-associated sRNAs remains unknown.

Three of the genes that were D23580-down-regulated in most conditions (*pSLT043-5*) were located downstream of the Tn*21*-like element in the pSLT-BT plasmid (S4 Fig). Because the Tn*21*-like multidrug resistance island was inserted between the *mig-5* promoter region and the *pSLT043-5* genes, we hypothesize that the differential expression reflects transcriptional termination mediated by the Tn*21* cassette. Two other D23580-down-regulated genes were located in the Gifsy-1 prophage region, *dinI-gfoA*. The presence of a SNP in the P$_{dinI-gfoA}$ promoter of D23580 is responsible for the lack of viability of the Gifsy-1 phage in D23580 [24]. The final gene that was D23580-down-regulated in most growth conditions was the *cysS* chromosomal gene, which encodes a cysteinyl-tRNA synthetase. Aminoacyl-tRNA synthetases are generally
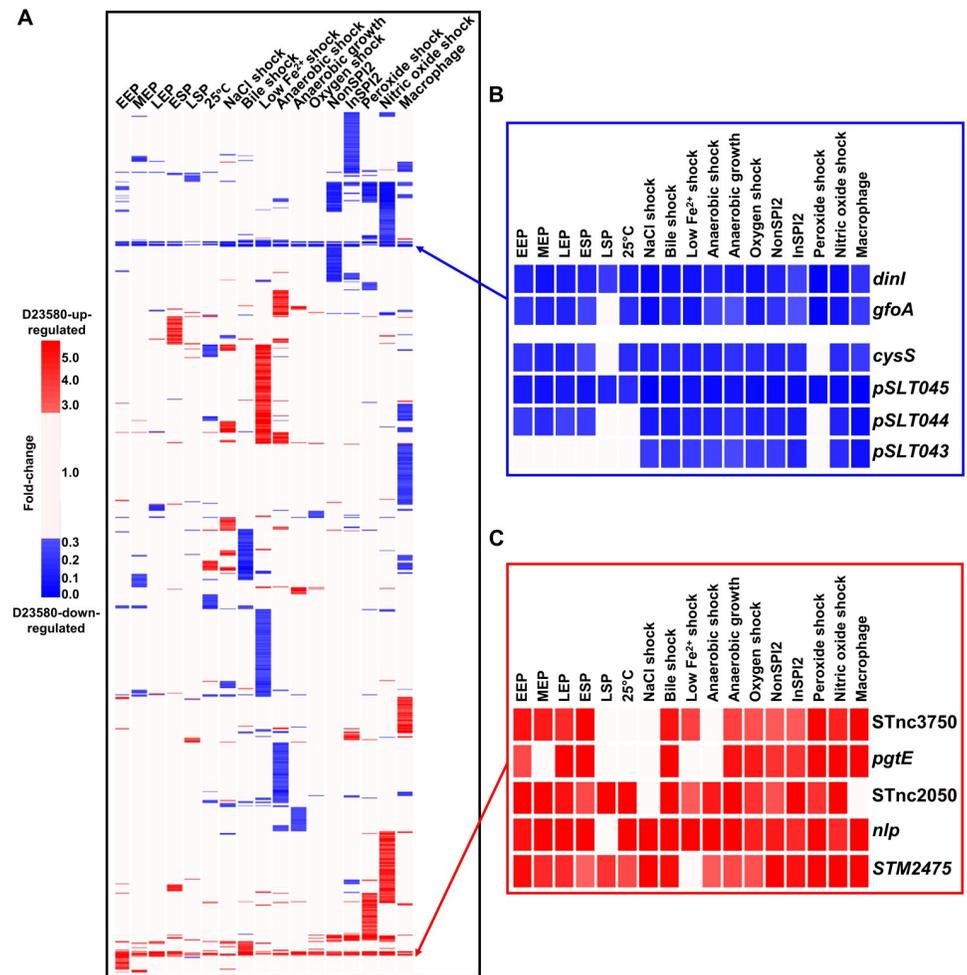
Fig 2. Interstrain transcriptomic comparison of *S*. Typhimurium D23580 versus 4/74. Expression of orthologous coding genes and sRNAs was compared between strains D23580 and 4/74 (reanalyzed data from Kröger and colleagues [37] and Srikumar and colleagues [40]) during growth in 17 infection-relevant conditions. The TPM value for each coding gene and sRNA in each condition in D23580 was divided by the TPM value for the same gene/sRNA and condition in 4/74. Heat maps were obtained using the GeneSpring GX7.3 software (Agilent, Santa Clara, CA, USA). Cluster analysis was performed using data with ≥3 fold-change. (A) Heat map of the 1,031 coding genes and sRNAs that showed significant difference (≥3 fold-change) between the two strains in at least one condition (S5 Table). The ≥3 fold-changes are shown in red (D23580-up-regulated) or blue (D23580-down-regulated). (B) Heat map representing D23580-down-regulated genes observed in all or most growth conditions. (C) Heat map of the D23580-up-regulated coding genes and sRNAs observed in most growth conditions. EEP, early exponential phase; ESP, early stationary phase; InSPI2, SPI-2-inducing; LEP, late exponential phase; LSP, late stationary phase; MEP, middle exponential phase; NonSPI2, SPI-2-noninducing; SPI, *Salmonella* pathogenicity island; sRNA, small RNA; TPM, transcripts per million.

essential genes, required for cell growth and survival [43]. The unexpected low level of *cysS* expression in D23580 in several growth conditions (TPM values ranging from 5 to 18 excluding the late stationary phase and shock conditions) was investigated further (see below).

Intriguing patterns of differential expression were observed between strains D23580 and 4/74 in particular growth conditions for certain functional groups of *Salmonella* genes. For example, the flagellar regulon and associated genes showed a characteristic pattern of expression in the phosphate carbon nitrogen (PCN)-related minimal media and inside macrophages

(S5 Fig). To allow us to make statistically significant findings, a larger-scale experiment was designed.

Identification of the transcriptional signature of *S.* Typhimurium ST313 D23580

To generate a robust transcriptional signature of D23580, we focused on the five environmental conditions with particular relevance to *Salmonella* virulence, namely, early stationary phase (ESP), anaerobic growth, SPI-2–noninducing (NonSPI2) and SPI-2–inducing (InSPI2) conditions, and intra-macrophage. The ESP and anaerobic growth conditions stimulate expression of the SPI-1 virulence system, and SPI-2 expression is induced by the InSPI2 and macrophage conditions [37,40]. RNA was isolated from three biological replicates of both D23580 and 4/74 grown in the four in vitro environmental conditions. For the two strains, three biological replicates were generated in parallel in a new set of experiments for this study. Additionally, RNA was extracted from additional biological replicates of intra-macrophage *S.* Typhimurium following infection of murine RAW264.7 macrophages for both D23580 (two additional replicates) and 4/74 (one additional replicate). Following RNA-seq, the sequence reads were mapped to the D23580 and 4/74 genomes using our bespoke software pipeline (Materials and Methods). The RNA-seq mapping statistics are detailed in S4 Table. To ensure that biologically meaningful gene expression differences were reported, we used very conservative cutoffs to define differential expression (Materials and Methods).

Following RNA-seq analysis of the three biological replicates of D23580 and 4/74 in five growth conditions, differential expression analysis of orthologous genes and sRNAs was performed with a rigorous statistical approach (Materials and Methods, S6 Table). We identified 677 genes and sRNAs that showed ≥2 fold-change (false discovery rate [FDR] ≤0.001) in at least one growth condition (Fig 3A). Between 6% (anaerobic growth) and 2% (InSPI2 condition) of orthologous genes and sRNAs were differentially expressed between the two strains (Fig 3B).

The ability to swim in semisolid agar is a key phenotypic difference between D23580 and 4/74 [13,16]. We confirmed that D23580 was less motile than 4/74 (S5 Fig) but did not observe significant differences in motility gene expression in complex media between the strains at the transcriptional level (S5 Fig). One nucleotide deletion and 11 SNP differences were found in the flagellar regulon between the two strains: one mutation in the promoter region of *mcpA*; three synonymous mutations in *flgK*, *cheA*, and *fliP*; four nonsynonymous mutations in *flhA*, *flhB*, *fliB*, and *mcpC*; and three mutations in the 5′ untranslated regions (UTRs) of *motA*, *flhD*, and *mcpA*. FlhA and FlhB are transmembrane proteins that are essential for flagellar protein export [44]. The SNP in *flhB* was conserved in all ST313 strains tested but also in the ST19 strains LT2 and 14028. The D23580 *flhA* SNP was specific to ST313 lineage 2.

To investigate the function of the 4/74 *flhA* SNP, the mutation was introduced to the chromosome of D23580 by single-nucleotide engineering (D23580 *flhA*$^{4/74}$). Motility of the D23580 *flhA*$^{4/74}$ mutant was significantly increased compared to the D23580 wild-type (WT) strain (S5 Fig). We originally hypothesized that the *flhA* SNP was related to the reported decreased inflammasome activation in macrophages, which is thought to contribute to the stealth phenotype of *S.* Typhimurium ST313 that involves evasion of the host immune system during infection [25]. However, no significant differences in cell death due to inflammasome activation were found between WT D23580 and the D23580 *flhA*$^{4/74}$ mutant (S5 Fig).

The transcriptomic data did offer an explanation for the reduced motility of D23580 on minimal media. In the NonSPI2 condition, all flagellar genes were D23580-down-regulated, with the exception of the master regulators *flhDC* (S5 Fig). In contrast, in the InSPI2 and intra-
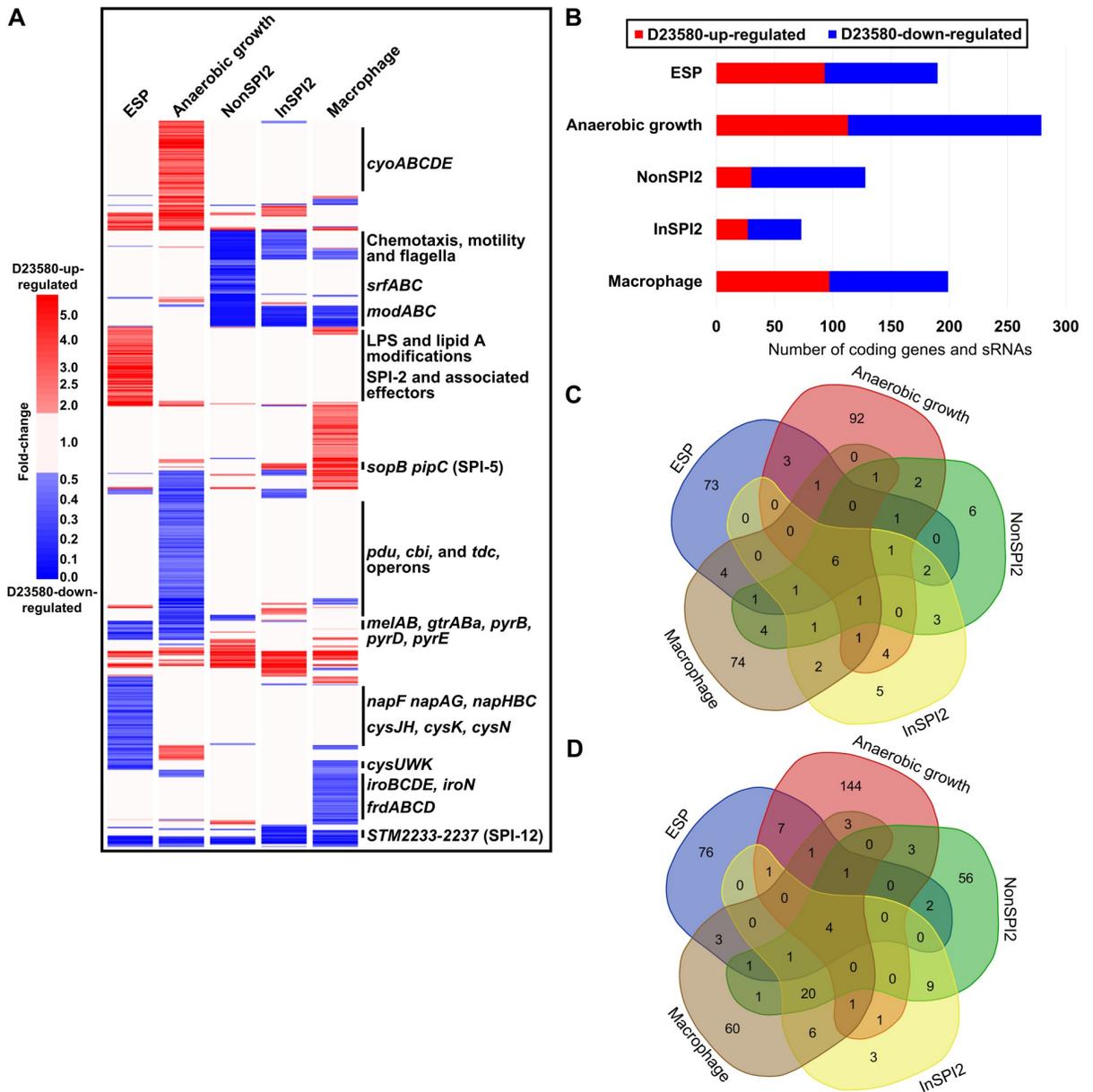
Fig 3. Transcriptional signature of *S.* Typhimurium D23580. Differential expression analysis of orthologous coding genes and sRNAs between strains D23580 and 4/74 during growth in five infection-relevant conditions. (A) Heat map highlighting biological relevant clusters. The CPM values of three biological replicates for each coding gene and sRNA in each condition in D23580 were compared to the CPM values for the same gene/sRNA and condition in 4/74. The heat map was obtained using GeneSpring GX7.3 (Agilent). Cluster analysis was performed with CPM values of the 677 coding genes and sRNAs that showed ≥2 fold-change and ≤0.001 of FDR in the differential expression analysis generated using Degust in at least one condition. Only fold-changes ≥2 with an FDR ≤0.001 are represented in red (D23580-up-regulated) or blue (D23580-down-regulated) (S6 Table). (B) Number of coding genes and sRNAs differentially expressed in each of the five growth conditions, based on Degust results (≥2 fold-change, ≤0.001 FDR). (C) Venn diagram of the D23580-up-regulated genes in the five growth conditions (http://bioinformatics.psb.ugent.be/webtools/Venn/) (S6 Table). (D) Venn diagram of the D23580-down-regulated genes in the five growth conditions (S6 Table). CPM, counts per million; ESP, early stationary phase; FDR, false discovery rate; InSPI2, SPI-2-inducing; LPS, lipopolysaccharide; NonSPI2, SPI-2-noninducing; SPI, *Salmonella* pathogenicity island; sRNA, small RNA.

https://doi.org/10.1371/journal.pbio.3000059.g003

macrophage condition, only the flagellar class 2 genes (such as *flgA*) were significantly down-regulated. RflP (YdiV) is a post-transcriptional negative regulator of the flagellar master

transcriptional activator complex FlhD$_4$C$_2$ [45–47]. We speculate that the down-regulation of the flagellar regulon in NonSPI2 could be due to a significant up-regulation (3.5 fold-change) of *rflP* in this low-nutrient environmental condition. This differential expression was not seen in the InSPI2 growth condition, which only differs from NonSPI2 by a lower pH (5.8 versus 7.4) and a reduced level of phosphate [37].

We identified six genes and sRNAs that were D23580-up-regulated in all five growth conditions, specifically *pgtE*, *nlp*, *ydiM*, *STM2475* (*SL1344_2438*), the ST64B prophage-encoded *SL1344_1966*, and the sRNA STnc3750 (Fig 3C). Just four genes were D23580-down-regulated in all conditions, namely, *pSLT043-5* (*SLP1_0062–4*) and *cysS* (Fig 3D). These findings confirmed that biologically significant information can be extracted from the initial 17-condition experiment (Fig 2B and 2C) because similar genes were up-/down-regulated across the multiple conditions of the replicated experiment.

The transcriptomic data were interrogated to identify virulence-associated genes that were differentially expressed between D23580 and 4/74. Coding genes and sRNAs located within SPI-1, SPI-2, SPI-5, SPI-12, and SPI-16 showed differential expression between D23580 and 4/74 in at least one growth condition (S6 Fig). The SPI-5-encoded *sopB* gene (encoding a SPI-1 effector protein) and its associated chaperone gene (*pipC*) were significantly D23580-up-regulated in the InSPI2 and intra-macrophage conditions. In contrast, the SPI-12-associated genes *STM2233-7* (*SL1344_2209–13*) were D23580-down-regulated in the same two growth conditions. Most SPI-2 genes were significantly D23580-up-regulated in the ESP condition, raising the possibility that the noninduced level of expression of SPI2 is higher in D23580 than 4/74.

The most highly differentially expressed genes in ESP (≥4 fold-change, FDR ≤0.001) (Fig 4) included the D23580-up-regulated genes required for itaconate degradation (*ripC*), *myo*-inositol utilization (*reiD*), and proline uptake (*putA*). D23580-down-regulated genes in the same growth condition included those involved in uptake of uracil and cytosine (*uraA* and *codB*), melibiose utilization (*melAB*), carbamoyl-phosphate metabolism and pyrimidine biosynthesis (*carAB* and *pyrEIB*), nitrate reductase (*napDF*), and sulfate metabolism (*cysPU* and *sbp*).

We also identified genes that were differentially expressed between D23580 and 4/74 during infection of RAW264.7 macrophages (Fig 4). The 16 genes that were most highly D23580-up-regulated (≥4 fold-change, FDR ≤0.001) included a β-glucosidase, *STM3775* (*SL1344_3740*); genes involved in cysteine metabolism, *cdsH*; oxidation of L-lactate, *STM1621* (*SL1344_1551*); and the transcriptional regulator *rcsA*. Genes that were D23580-down-regulated during infection of macrophages were involved in the uptake of sialic acid (*nanM*) and maltose or maltodextrin (*malEFK*) or the secretion and import of siderophores (*iroC* and *iroD*).

Malaria is one of the risk factors frequently associated to iNTS disease in sub-Saharan Africa, especially in young children. The recent work of Lokken and colleagues [48] reported an increase in intracellular iron availability in liver mononuclear cells caused by the infection of the malaria parasite. In addition, systemic growth of a *S*. Typhimurium mutant that was unable to acquire exogenous iron was significantly increased in a malaria coinfection mouse model [48]. These observations raise the possibility that the D23580-down-regulation of iron acquisition genes reflects adaptation to the host-associated niche.

Six of the 677 genes and sRNAs that were differentially expressed between D23580 and 4/74 were among the 17 genes that had a SNP/indel in the promoter region [20]: *pgtE*, *yohM*, *mcpA*, *yrbL*, *nanM*, and *STM2475*. The promoter region of *STM2475* was further studied (S7 Fig) because *STM2475* was D23580-up-regulated in all five growth conditions tested. Three TSSs controlled expression of this gene in 4/74: one primary, one secondary, and one internal [37]. The same number of TSSs was identified in D23580 [20]. Further investigation of the region revealed the absence of the 4/74 internal TSS in D23580. The 4/74 secondary TSS became the
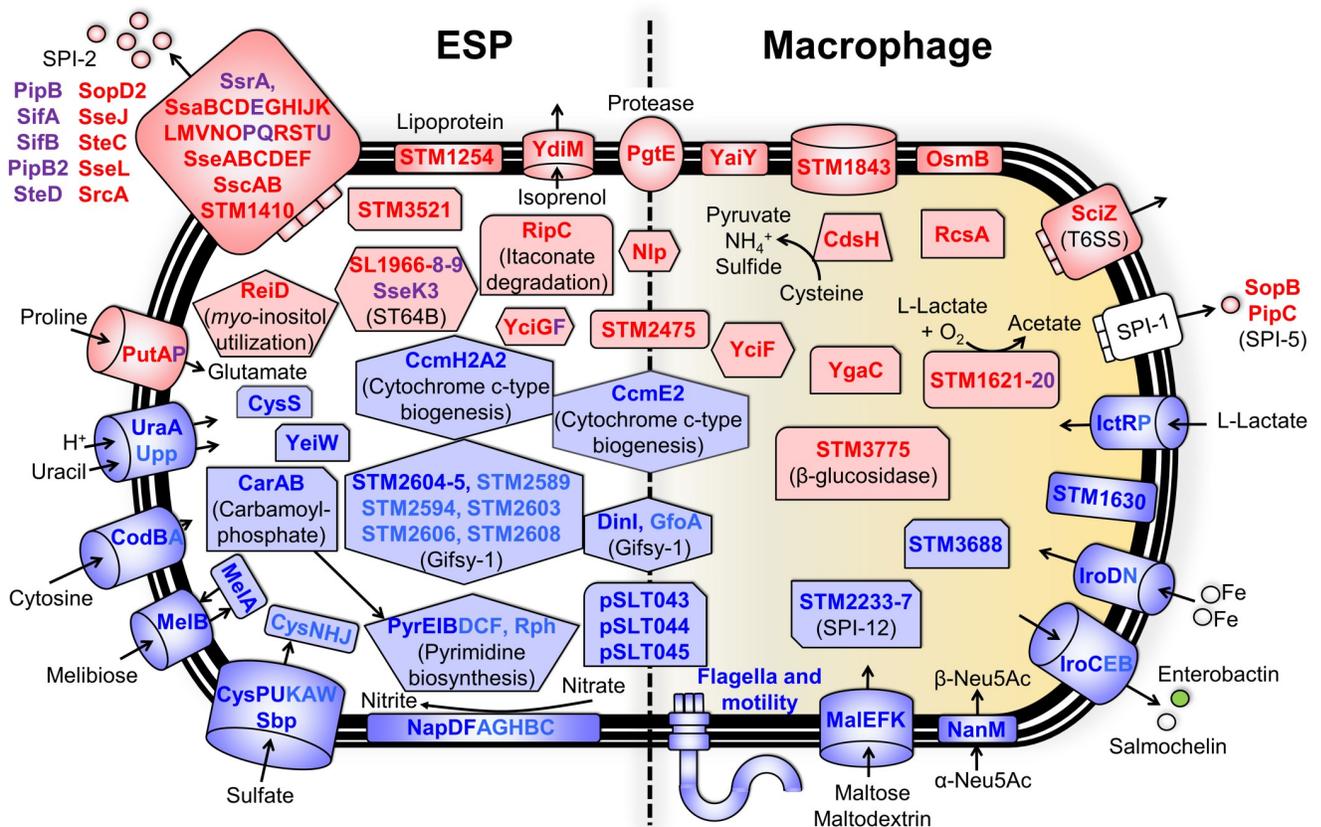
Fig 4. Differentially expressed genes between *S.* Typhimurium 4/74 and D23580 in ESP and the intra-macrophage environment. Differential gene expression of D23580 versus 4/74 in the ESP (on the left) and the macrophage (on the right) conditions. Colors refer to fold-changes of D23580 versus 4/74 from differential expression analysis using Degust; red = D23580-up-regulated, blue = D23580-down-regulated. The figure includes genes that are differentially expressed ≥4 fold-change with ≤0.001 FDR (red and blue font color). Purple and light blue font colors represent up-regulated or down-regulated genes, respectively, that are related to the previous functional groups of genes but have a fold-change ≤4 and ≥2 (≤0.001 FDR). ESP, early stationary phase; FDR, false discovery rate; SPI, *Salmonella* pathogenicity island.

https://doi.org/10.1371/journal.pbio.3000059.g004

primary TSS in D23580, likely because of the presence of an "A" insertion in the −10 element of the *STM2475* promoter region. The 4/74 primary TSS corresponded to the secondary TSS in D23580. D23580 carried a third *STM2475* TSS within the upstream gene *ypfG*. We speculate that the "A" insertion in the −10 element in D23580 could explain the differential expression between strains.

The key features of the transcriptional signature of D23580 included the differential expression of the flagellar and associated genes, genes involved in aerobic and anaerobic metabolism, and iron-uptake genes. Specifically, the aerobic respiratory pathway *cyoABCDE* was D23580-up-regulated in the anaerobic growth condition, and anaerobic-associated-pathway *pdu*, *cbi*, and *tdc* operons were D23580-down-regulated. Importantly, genes associated with the acquisition of iron through production and uptake of siderophores were D23580-down-regulated in the intra-macrophage environment. In summary, the transcriptional signature of D23580 suggests that the biology of ST313 lineage 2 differs from ST19 under anaerobic conditions in vitro and during infection of murine macrophages.

The challenge of data reproducibility in experimental science is widely acknowledged [49,50]. To assess the robustness of our experiments, the RNA-seq-derived expression profiles

that we generated from five replicated conditions were compared with five relevant individual conditions. There was a high level of correlation between the individual versus replicated data-sets (correlation coefficients between 0.88 and 0.97) (S8 Fig). However, different levels of expression were seen between the individual and the replicated ESP growth conditions of D23580 for a small minority of genes. The main variations in terms of functional gene groups involved cysteine metabolism, carbamoyl-phosphate and pyrimidine biosynthesis, and nitrate metabolism. Variation in expression of the *ripCBA-lgl* operon was also observed during anaerobic growth. We speculate that these alterations in gene expression reflect experimental variations such as the use of different batches of media.

1.5% of proteins were differentially expressed between D23580 and 4/74

RNA-seq-based transcriptomic analysis does not reflect the translational and post-translational levels of regulation [51]. To identify proteins that differentiated strains D23580 and 4/74, we used a proteomic strategy that involved a liquid chromatography-tandem mass spectrometry (LC-MS/MS) platform and analyzed proteins from D23580 and 4/74 bacteria grown in the ESP condition (S7 Table). A label-free quantification approach identified 66 differentially expressed orthologous proteins (≥2 unique peptides, ≥2 fold-change, *p*-value <0.05) (Fig 5), including 54 D23580-up-regulated proteins and 12 D23580-down-regulated proteins. The most highly D23580-up-regulated protein was PgtE, corroborating our previous study [20]. Up-regulated proteins included those required for carbamoyl-phosphate and pyrimidine biosynthesis (CarAB and PyrIB), some SPI-1 proteins and associated effectors (PrgH, SipAB, InvG, SlrP, SopB, SopE2, SopA, and SopD), RipAC (itaconate degradation), and Lgl (methylglyoxal detoxification).

To identify genes that were differentially expressed at both the transcriptional and translational levels, the quantitative proteomic data were integrated with the transcriptomic data. Eight D23580-up-regulated proteins (YciF, SopA, PgtE, STM2475, RipC, RibB, Nlp, and STM3775) were significantly up-regulated in the transcriptomic data (≥2 fold-change, FDR ≤0.001). The promoter regions of two of the related genes, *pgtE* and *STM2475*, carried a D23580-specific SNP (S3 Table). Four differentially expressed proteins (pSLT043, CysS, YgaD, and MelA) were D23580-down-regulated at the transcriptomic level (≥2 fold-change,

Fig 5. Differentially expressed proteins between *S.* Typhimurium D23580 and 4/74. Representation of significant D23580-up-regulated proteins (red dots) and D23580-down-regulated proteins (blue dots) in the ESP growth condition by Log$_2$ fold-change and the chromosome location in D23580 (≥2 unique peptides, ≥2 fold-change, *p*-value <0.05). Gray dots refer to proteins that showed nonsignificant differences. ESP, early stationary phase.

https://doi.org/10.1371/journal.pbio.3000059.g005

FDR ≤0.001) ([Fig 6](#)). Overall, 12 genes were differentially expressed at both the transcriptional and protein levels.

The evolution of *S.* Typhimurium ST313 involved the SNP-based inactivation of melibiose utilization genes

The melibiose utilization system consists of three genes: *melR*, which encodes an AraC-family transcriptional regulator; *melA*, encoding the alpha-galactosidase enzyme; and *melB*. MelB is
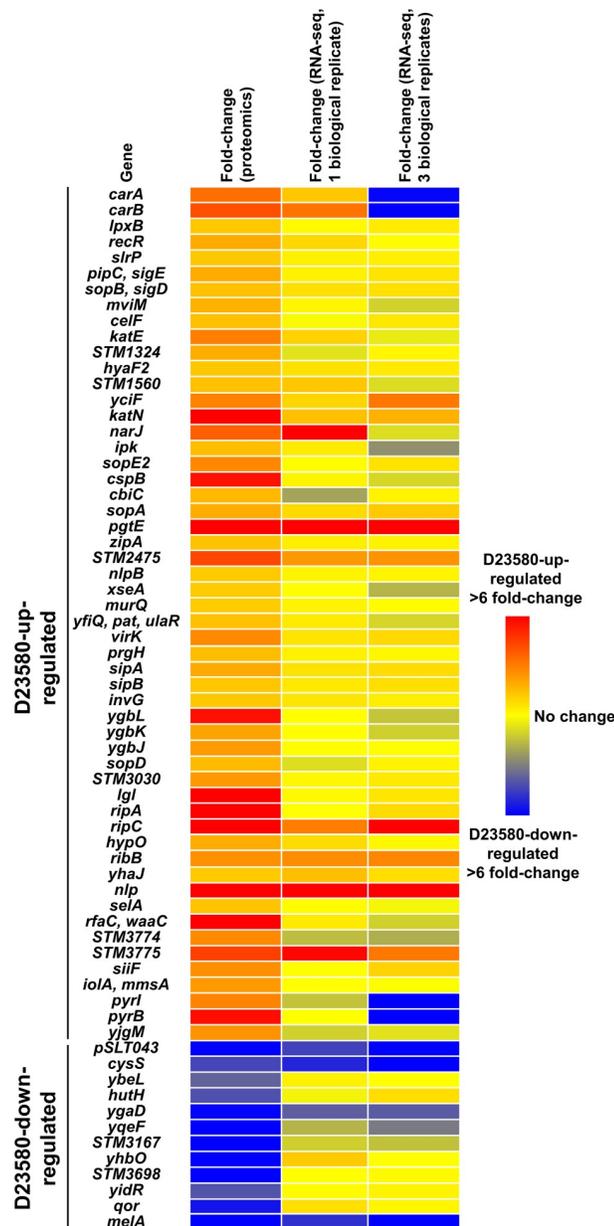

Fig 6. Heat map of the 66 proteins that are differentially expressed between *S.* Typhimurium 4/74 and D23580. Data represent levels of gene expression at the proteomic level in the ESP growth condition and in two independent ESP RNA-seq datasets (one biological replicate versus three biological replicates). ESP, early stationary phase; RNA-seq, RNA sequencing.

https://doi.org/10.1371/journal.pbio.3000059.g006

responsible for the active transport of melibiose across the bacterial cell membrane. We found that the *melAB* genes were D23580-down-regulated at the transcriptomic level (Fig 7A). The differential expression of *melA* was confirmed at the proteomic level (Fig 5).

In strain D23580, the melibiose utilization genes contain three nonsynonymous SNPs (4/74 → D23580). Two are present in *melB* (Pro → Ser at the 398 AA and Ile → Val at the 466 AA) and one in *melR* (Phe → Leu) (Fig 7B). The three SNPs were analyzed in the context of a phylogeny of 258 genomes of *S.* Typhimurium ST313 that included isolates from Malawi, as well as more distantly related ST313 genomes from the UK [5] (S8 Table). All three SNPs were found to be monophyletic, allowing us to infer the temporal order in which they arose and representing an accumulation of SNPs in melibiose utilization genes over evolutionary time. The first SNP, *melB* I466V, was present in all 258 ST313 strains tested and therefore arose first. The second SNP, in *melR*, was present in all ST313 lineage 2 and UK-ST313 genomes, suggesting that it appeared prior to the divergence of these phylogenetic groups [5]. The final SNP, *melB* P398S, was present in all ST313 lineage 2 and a subset of UK-ST313 genomes, consistent with this being last of the three mutations to arise (Fig 7C). ST313 strains can therefore be classified into groups of strains containing one, two, or three SNPs in melibiose utilization genes.

It has been reported that D23580 did not ferment melibiose, whereas a ST313 lineage 1 isolate (A130), *S.* Typhimurium SL1344, and *S.* Typhi Ty2 were able to utilize melibiose as a sole carbon source [18]. MelB catalyzes the symport of melibiose with Na^+, Li^+, or H^+ [53]. We confirmed that ST19 strains and strains belonging to the ST313 lineage 1 were positive for alpha-galactosidase activity. In contrast, isolates representing the ST313 lineage 2 and a subset of UK-ST313 strains were unable to utilize melibiose.

To determine the biological role of the SNPs in the *melB* and *melR* genes, we employed a genetic approach. Single-nucleotide engineering was used to generate isogenic strains that reflect all three melibiose gene SNP states for determination of the role of the SNP differences between ST313 lineage 2 and ST19 in the alpha-galactosidase (MelA)-mediated phenotypic defect (Fig 7D). Melibiose utilization in D23580 was rescued by nucleotide exchange of the three SNP mutations (D23580 *melB*⁺ *melR*⁺) (Fig 7E, S1 Data). D23580 recovered its ability to grow with melibiose as the sole carbon source after exchanging only the *melR* SNP with 4/74 (D23580 *melR*⁺). In contrast, D23580 did not grow in the same medium when the exchange only involved the two *melB* SNPs (D23580 *melB*⁺). 4/74 lost its ability to utilize melibiose as sole carbon source when we introduced the D23580 *melR* SNP (4/74 *melR*⁻ and 4/74 *melB*⁻ *melR*⁻). However, an exchange of the two nucleotides in *melB* did not eliminate the ability of 4/74 to grow in minimal medium with melibiose (4/74 *melB*⁻). These data correlated with the alpha-galactosidase activity of the mutants, although a slight difference was observed between strains D23580 *melR*⁺ (light green) and D23580 *melB*⁺ *melR*⁺ (green) and between strains 4/74 (green) and 4/74 *melB*⁻ (light green) (Fig 7F), suggesting an altered efficiency of melibiose utilization between the two strains. To completely restore alpha-galactosidase activity in D23580, the reversion of the nonsynonymous SNPs in both the *melR* and *melB* genes was required. Our data suggest that the *melR* SNP is critical for the loss of function of the melibiose utilization system.

In a chicken infection model, the *melA* transcript of *S.* Typhimurium strain F98 is more highly expressed in the caecum than during in vitro growth [54]. In a chronic infection model, accumulation of melibiose was observed in the murine gut after infection with *S.* Typhimurium strain 14028s [55]. More recently, it has been reported that some gut bacteria are able to extracellularly hydrolyze raffinose into melibiose and fructose, causing the accumulation of melibiose [56]. We speculate that the ability to metabolize melibiose could provide a fitness advantage to *S.* Typhimurium ST19 during gut colonization and that the loss of the melibiose catabolic pathway in *S.* Typhimurium ST313 lineage 2 could reflect niche adaptation. The
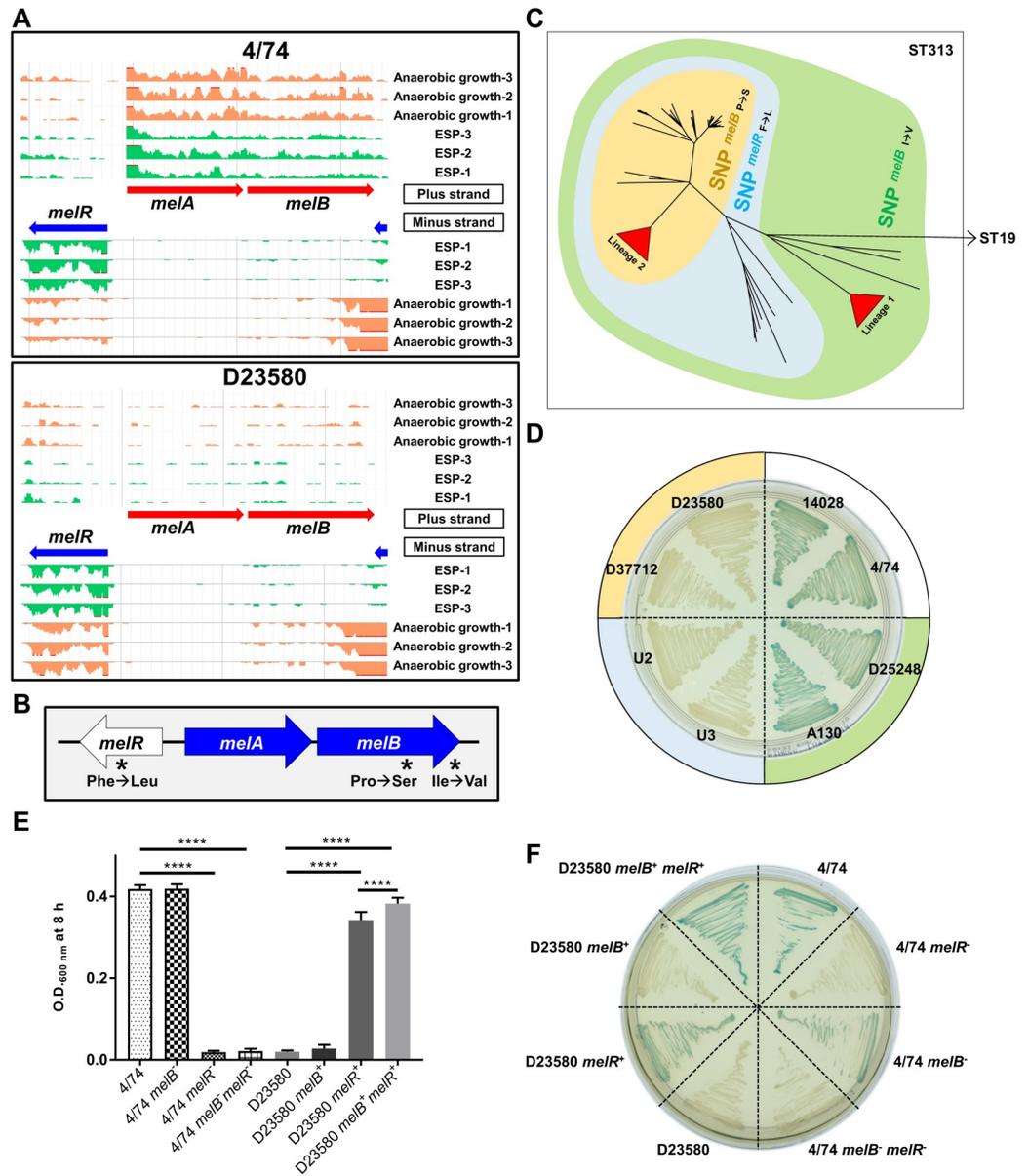
Fig 7. Melibiose phenotype differentiates the *S.* Typhimurium ST19 and ST313 strains. (A) Visualization of RNA-seq data with three biological replicates in the ESP and anaerobic growth conditions using JBrowse [52] for the melibiose utilization operon. The scale of the mapped reads was 1 to 100. (B) Presence of three nonsynonymous SNPs in the melibiose utilization genes (4/74 → D23580). (C) Accumulation of SNPs in the melibiose utilization genes during the evolution of ST313 in the context of a whole-genome-core SNP phylogeny. Isolate names, ST313 lineage, and genotype for the three SNPs are included in S8 Table. (D) Alpha-galactosidase activity of representatives of ST19 and ST313 strains on Pinnacle *Salmonella* ABC medium (Lab M, Heywood, UK); green = positive, colorless = negative. The colors of the external circle correlate with the colors represented in the tree in (C). (E) Bacterial growth in minimal M9 medium supplemented with 0.4% melibiose (S1 Data). Statistical analysis was performed using one-way ANOVA and Tukey's multiple comparison test. Bars represent the mean of seven biological replicates and standard deviation. Significant differences (****) indicate p-value <0.0001. (F) Alpha-galactosidase activity of 4/74 and D23580 WT strains and corresponding mutants. The ability to use melibiose is rescued in D23580 by exchange of the three SNP mutations. ESP, early stationary phase; RNA-seq, RNA sequencing; SNP, single-nucleotide polymorphism; ST, sequence type; WT, wild type.

https://doi.org/10.1371/journal.pbio.3000059.g007

inactivation of melibiose catabolism by SNPs that are conserved throughout the ST313 lineage 2 is consistent with a functional role in ST313 virulence, and we are currently examining this possibility.

A plasmid-encoded cysteinyl-tRNA synthetase is required for growth in D23580

The dramatic down-regulation of the chromosomal *cysS* gene at both the transcriptomic (Fig 8A) and proteomic levels (Fig 5) was studied experimentally. The coding and noncoding regulatory regions of the chromosomal *cysS* were identical at the DNA level in strains D23580 and 4/74. The chromosomal *cysS* gene encodes a cysteinyl-tRNA synthetase, which is essential for cell growth in *S*. Typhimurium and other bacteria [43,57,58]. To investigate *cysS* gene function, we consulted a transposon-insertion sequencing (TIS) dataset for *S*. Typhimurium D23580 (S1 Data). Genes that show the absence or low numbers of transposon-insertion sites are considered to be "required" for bacterial growth in a particular condition [58,59]. The data suggested that a functional chromosomal *cysS* was not required for growth in rich medium (Fig 8B). We searched for *S*. Typhimurium D23580 genes that encoded a cysteinyl-tRNA synthetase and identified the pBT1-encoded gene, *pBT1-0241* ($cysS^{pBT1}$), which the TIS data suggested to be "required" for growth in rich medium (Fig 8B).

To investigate cysteinyl-tRNA synthetase function in D23580, individual knock-out mutants were constructed in the chromosomal *cysS* gene ($cysS^{chr}$) and the $cysS^{pBT1}$ gene. These genes were 89% identical at the amino acid level and 79% at the nucleotide level. The $cysS^{pBT1}$ mutant was whole-genome sequenced to confirm the absence of secondary unintended mutations. The pBT1 plasmid was also cured from D23580. We determined the relative fitness of the two *cysS* mutants and the pBT1-cured strain. The WT D23580 and D23580 Δ$cysS^{chr}$ and D23580 ΔpBT1 mutants grew at similar rates in Lennox broth (LB), while the D23580 Δ$cysS^{pBT1}$ mutant showed an extended lag phase (Fig 9A, S1 Data). The D23580 Δ$cysS^{pBT1}$ mutant showed a more dramatic growth defect in minimal medium with glucose as the sole carbon source (Fig 9B, S1 Data).

To determine whether the presence of the pBT1 plasmid was linked to the decrease in $cysS^{chr}$ expression, RNA from two biological replicates was isolated from the pBT1-cured strain in the ESP growth condition. Differential expression analysis between this mutant and the WT D23580 strain showed a significant increase in expression of $cysS^{chr}$, with TPM levels close to those seen in 4/74 (Fig 9C, S5 Table, S1 Data). These results suggested the pBT1 plasmid is responsible for the down-regulation of $cysS^{chr}$ expression in D23580.

The conservation of pBT1 was studied among 233 ST313 strains and compared to the presence of the pSLT-BT plasmid, which was found in all lineage 2 isolates (Fig 9D, S8 Table). Approximately 37% of ST313 lineage 2 isolates carried the pBT1 plasmid. The pBT1 plasmid has rarely been seen previously but did show significant sequence similarity to five plasmids found in *Salmonella* strains isolated from reptiles and elsewhere (98% to 99% nucleotide identity over 92% to 97% of the plasmid sequence; accessions JQ418537, JQ418539, CP022141, CP022036, and CP022136, S1 Text).

Examples of essential bacterial genes located on plasmids are rare, and this phenomenon has been previously explored [62]. We conclude that the essentiality of the $cysS^{pBT1}$ gene provides a novel strategy, to our knowledge, for plasmid maintenance in a bacterial population.

The SalComD23580 community data resource

To allow scientists to gain new biological insights from analysis of this rich transcriptomic dataset, we have made it available as an online resource for the visualization of similarities and
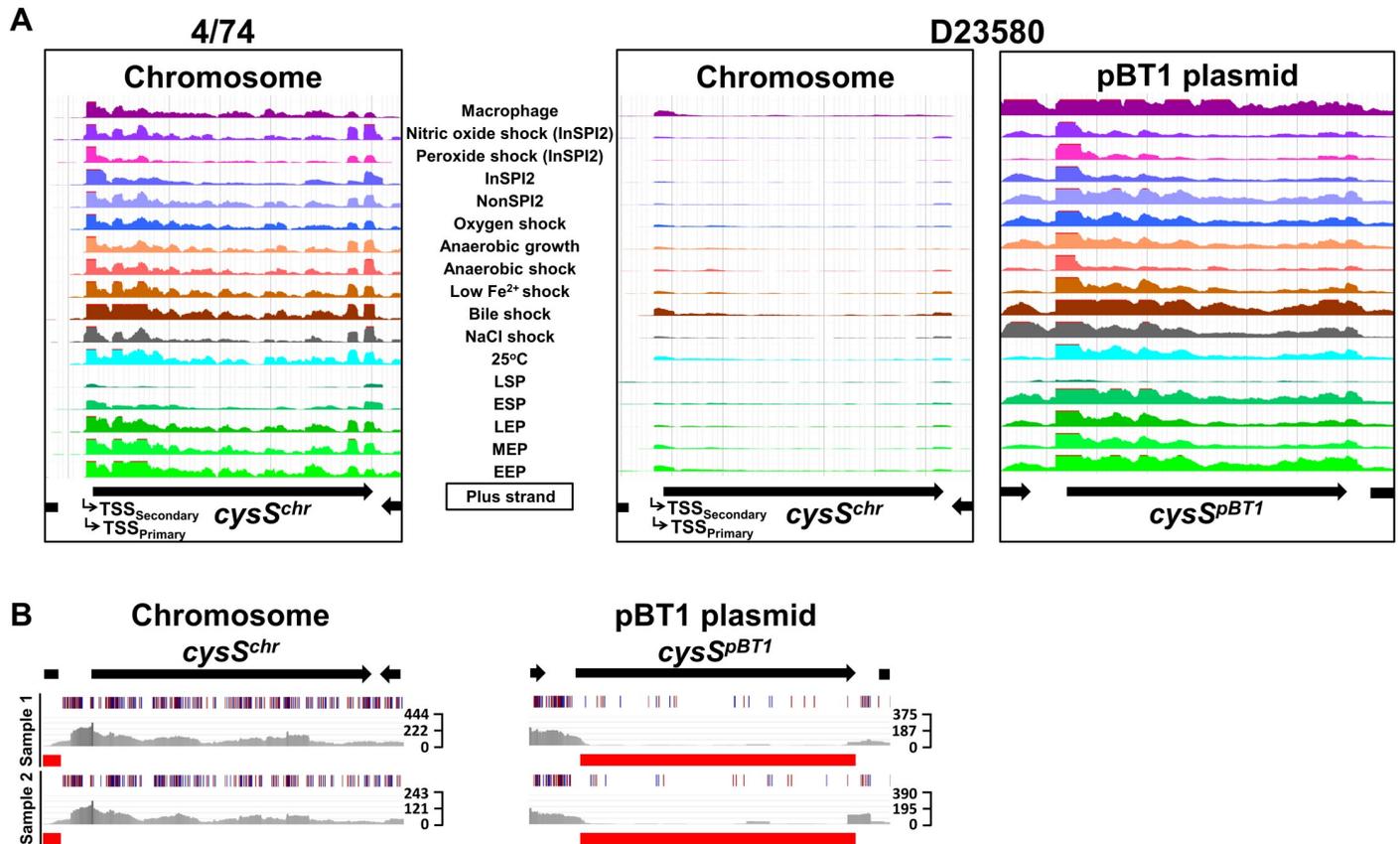
Fig 8. The pBT1 plasmid encodes the functional cysteinyl-tRNA synthetase in S. Typhimurium D23580. (A) RNA-seq data for *cysS* in 4/74 and chromosomal and pBT1-plasmid-encoded *cysS* in D23580 from the online JBrowse resources provided in this study. The scale of the mapped reads was 1 to 500. (B) Transposon library results for the *cysS^chr* and *cysS^pBT1* genes in D23580. Figures were obtained using the Dalliance genome viewer [60]. Black arrows at the top represent genes. Each sample is represented by three tracks. The first track contains blue and red lines that correspond to transposon-insertion sites; red = orientation of the transposon is the same as the direction of the gene, blue = opposite direction. The second track shows raw data for the Illumina sequencing reads. The third track highlights in red those genes that were considered "required" for growth in that condition based on an insertion index. The insertion index was calculated for each gene as explained in [59,61], and genes with insertion index values <0.05 were considered as "required" for growth in the Lennox rich medium. The scale on the right represents sequence read coverage. EEP, early exponential phase; ESP, early stationary phase; InSPI2, SPI-2-inducing; LEP, late exponential phase; LSP, late stationary phase; MEP, middle exponential phase; NonSPI2, SPI-2-noninducing; RNA-seq, RNA sequencing; SPI, *Salmonella* pathogenicity island; tRNA, transfer RNA; TSS, transcriptional start site.

https://doi.org/10.1371/journal.pbio.3000059.g008

differences in gene expression between ST313 (D23580) and ST19 (4/74), using an intuitive heat map-based approach (http://bioinf.gen.tcd.ie/cgi-bin/salcom_v2.pl). To examine the transcriptional data in a genomic context, we generated two strain-specific online browsers that can be accessed from the previous link, one for D23580 and one for 4/74. The value of this type of online resource for the intuitive interrogation of transcriptomic data has been outlined recently [63].

Perspective

To investigate the functional genomics of *S*. Typhimurium ST313, we first resequenced and reannotated the genome of the D23580 isolate. Our comparative genomic analysis of two *S*. Typhimurium ST313 and ST19 isolates confirmed the findings of Kingsley and colleagues [19], identifying 856 SNPs and indels, many instances of genome degradation, and the presence of specific prophages and plasmids. To discover the genetic differences that impact upon
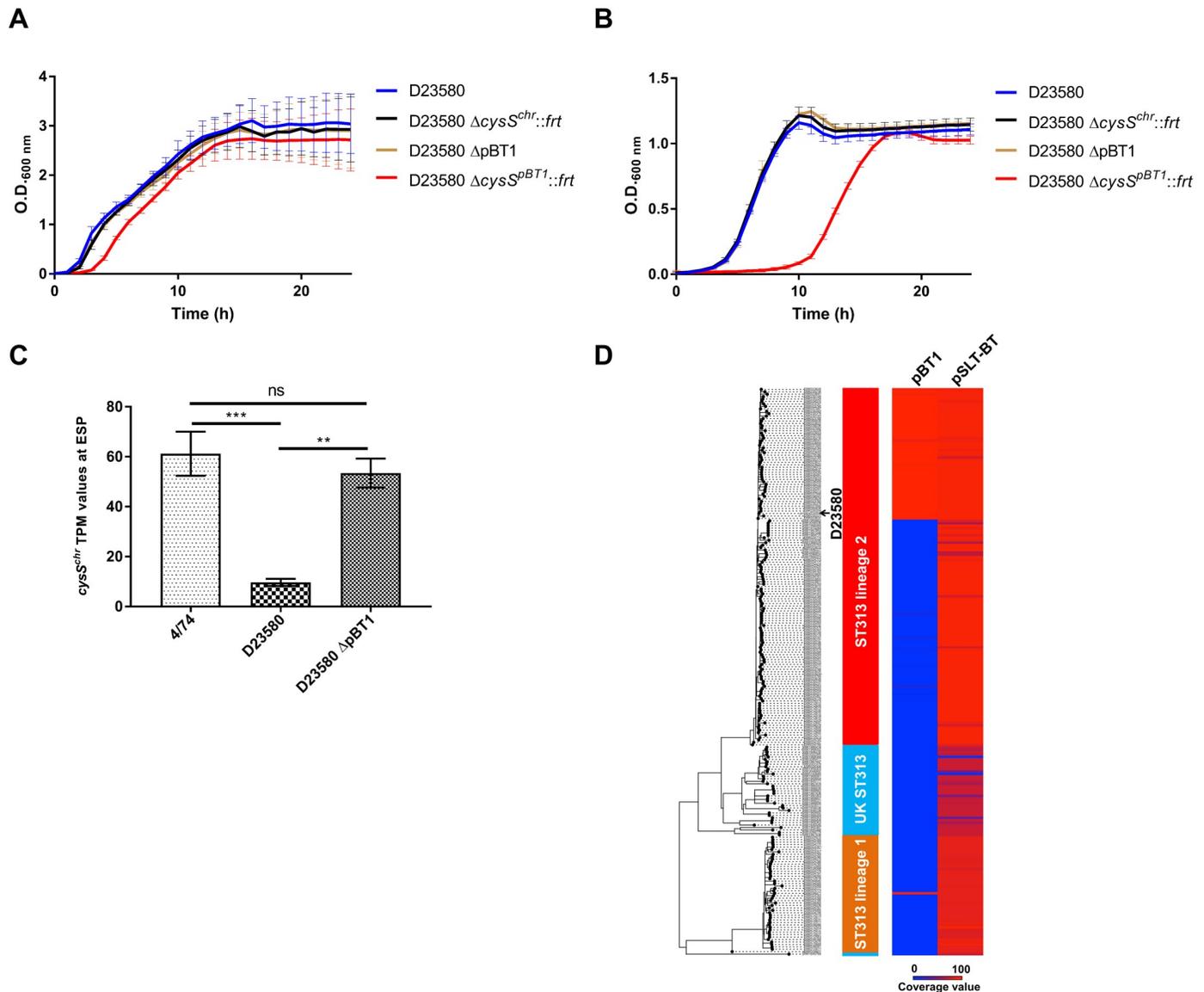
Fig 9. The pBT1-encoded *cysS* is required for optimal growth of *S.* Typhimurium D23580. (A) Growth curves of D23580 WT, D23580 Δ*cysS*chr::*frt*, D23580 pBT1-cured strain, and D23580 Δ*cysS*pBT1::*frt* strains in LB medium, *n* = 8 (standard deviations are represented) (S1 Data). (B) Growth curves in minimal M9 medium supplemented with 0.4% glucose, *n* = 5 (standard deviations are represented) (S1 Data). (C) Comparison of *cysS*chr expression levels (TPM values) of 4/74 (*n* = 3), D23580 (*n* = 3), and the D23580 pBT1-cured strain (*n* = 2) in the ESP growth condition (S1 Data). Bars represent mean values and standard deviations. Significant differences (***) indicate *p*-value <0.001. (D) The pBT1 plasmid is present in a subset of ST313 isolates of lineage 2 and in one isolate from lineage 1. Isolate names, ST313 lineage, and coverage value for pBT1 and pSLT-BT are included in S8 Table. ESP, early stationary phase; LB, Lennox broth; ns, not significant; OD, optical density; ST, sequence type; TPM, transcripts per million; WT, wild type.

https://doi.org/10.1371/journal.pbio.3000059.g009

the biology of *S.* Typhimurium ST313, we used a functional transcriptomic approach to show that the two *S.* Typhimurium pathovariants shared many responses to environmental stress.

By investigating global gene expression in multiple infection-relevant growth conditions, we discovered that 677 genes and sRNAs were differentially expressed between strains D23580 and 4/74. A parallel proteomic approach confirmed that many of the gene expression differences led to alterations at the protein level. The differential expression of 199 genes and sRNAs within macrophages allowed us to predict functions of African *S.* Typhimurium ST313 that

are modified during infection. The comparative gene expression data were used to predict key phenotypic differences between the pathovariants, which are summarized in S1 Table. The power of our experimental approach is highlighted by our discovery of the molecular basis of the melibiose utilization defect of D23580 and a novel, to our knowledge, bacterial plasmid maintenance system that relied upon a plasmid-encoded essential gene.

In the future, similar functional transcriptomic approaches could shed light on the factors responsible for the phenotypic differences that distinguish the pathovariants of many bacterial pathogens.

Materials and methods

Bacterial strains

The clinical isolate *S. enterica* serovar Typhimurium D23580 was obtained from the Malawi-Liverpool-Wellcome Trust Clinical Research Programme, Blantyre, Malawi [19]. This strain, isolated from the blood of an HIV⁻ child from Malawi, is used as a representative of the *Salmonella* ST313 after approval by the Malawian College of Medicine (COMREC ethics no. P.08/14/1614). *S.* Typhimurium 4/74 was originally isolated from the bowel of a calf with salmonellosis [64] and is used as a representative strain of *Salmonella* ST19. Other *Salmonella* strains referenced in this study are listed in S9 Table.

Growth conditions

All strains were routinely grown in LB containing 10 g/L tryptone, 5 g/L yeast extract, and 5 g/L NaCl. Liquid bacterial cultures were incubated at 37°C, 220 rpm for 16 h. Agar plates were prepared with 1.5% Bacto Agar (BD Difco, Franklin Lakes, NJ, USA). To test the ability to grow with melibiose as the sole carbon source, strains were grown in M9 minimal medium with 0.4% of melibiose. M9 minimal medium consisted of 1× M9 Minimal Salts (Sigma Aldrich, St. Louis, MO, USA), 2 mM $MgSO_4$, and 0.1 mM $CaCl_2$. Glucose was added at a final concentration of 0.4% to M9 minimal medium to study growth behavior of the *cysS* mutants. Media were supplemented with antibiotics when required: kanamycin (Km) 50 µg/mL, gentamicin (Gm) 20 µg/mL, tetracycline (Tc) 20 µg/mL, nalidixic acid (Nal) 50 µg/mL, and Cm 20 µg/mL.

Details for growing bacteria in the 16 in vitro infection-relevant conditions and inside murine RAW264.7 macrophages (ATCC TIB-71) have been published previously [37,40] and are summarized in S10 Table.

Resequencing of *S.* Typhimurium D23580 genome

For PacBio sequencing, *S.* Typhimurium D23580 was grown for 16 h in Lennox medium at 37°C, 220 rpm. DNA was extracted using the Bioline mini kit for DNA purification (Bioline, London, UK). Genomic quality was assessed by electrophoresis in a 0.5% agarose gel at 30–35 V for 17–18 h. A 10 kb library was prepared for DNA sequencing using three SMRT cells on a PacBio RSII (P5/C3 chemistry) at the Centre for Genomic Research, University of Liverpool, UK. Illumina sequencing of *S.* Typhimurium D23580 was performed by MicrobesNG, University of Birmingham, UK.

All the SNP and indel differences found between the chromosome and pSLT-BT sequences of the D23580 strain used in this study (accession: PRJEB28511) and the published D23580 (accession: FN424405 and FN432031) were confirmed by PCR with external primers and subsequent Sanger sequencing.

Draft sequences of the pBT2 and pBT3 plasmids were provided by Robert A. Kingsley [19] and were used to design oligonucleotides for primer-walking sequencing (all primer sequences are listed in S11 Table; Eurofins Genomics, Luxembourg, Luxembourg). Plasmid DNA from *S.* Typhimurium D23580 was isolated using the ISOLATE II Plasmid Mini Kit (Bioline). For pBT2, the following oligonucleotides were used: Fw-pBT2-1 and Rv-pBT2-1, Fw-pBT2-2 and Rv-pBT2-2; and for pBT3, the following oligonucleotides were used: Fw-pBT3-3 and Rv-pBT3-3, Fw-pBT3-1 and Rv-pBT3-4, Fw-pBT3-4 and Rv-pBT3-2.

The resulting genome sequence was designated D23580_liv (accession: PRJEB28511).

Assembly of the *S.* Typhimurium D23580 complete genome

HGAP3 [65] was used for PacBio read assembly of the D23580 chromosome and for the large plasmids pSLT-BT and pBT1. A hybrid assembly approach, Unicycler v0.4.5 [66], was used to combine the long reads from PacBio and the short reads from the Illumina platform in order to assemble small plasmids (not covered by PacBio due to size selection in library preparation) and to improve the large plasmid assemblies.

RNA isolation, cDNA library preparation, and Illumina sequencing

Total RNA from *S.* Typhimurium D23580 grown in 16 in vitro infection-relevant conditions (EEP, MEP, LEP, ESP, LSP, 25˚C, NaCl shock, bile shock, low Fe^{2+} shock, anaerobic shock, anaerobic growth, oxygen shock, NonSPI2, InSPI2, peroxide shock, and nitric oxide shock) and murine RAW264.7 macrophages was isolated using TRIzol and treated with DNase I, as described previously [37,40]. For a more robust comparative transcriptomic analysis, a second round of RNA-seq experiments involved RNA isolation from three new D23580 biological replicates grown in ESP, anaerobic growth, NonSPI2, and InSPI2, and two more intra-macrophage samples. In addition, RNA purifications of three *S.* Typhimurium 4/74 biological replicates grown in ESP, anaerobic growth, NonSPI2, and InSPI2, and one more biological replicate from the intra-macrophage environment were performed for this study.

For RNA-seq, cDNA libraries were prepared and sequenced by Vertis Biotechnologie AG (Freising, Germany). Briefly, RNA samples were fragmented with ultrasound (4 pulses of 30 sec at 4˚C), treated with Antarctic phosphatase, and rephosphorylated with polynucleotide kinase (PNK). RNA fragments were poly(A)-tailed, and an RNA adapter was ligated to the 5′-phosphate of the RNA. First-strand cDNA synthesis was carried out using an oligo(dT)-adapter primer and M-MLV reverse transcriptase. cDNA was subsequently amplified by PCR to 20–30 ng/μL and purified using the Agencourt AMPure XP kit (Beckman Coulter Genomics, Chaska, MN, USA). cDNA samples were pooled in equimolar amounts, size selected to 150–500 bp, and sequenced on an Illumina HiSeq 2000 system (single-end 100 bp reads). Minor changes were applied to different RNA-seq runs. For the third macrophage biological replicate of D23580, cDNA was PCR amplified to 10–20 ng/μL and size selected to 200–500 bp, and samples were sequenced on an Illumina HiSeq 2500 platform (1×100 bp). For RNA samples of D23580 and 4/74 grown in the four in vitro growth conditions with three biological replicates, the third macrophage replicate of 4/74, and the D23580 pBT1-cured strain, cDNA was PCR amplified to 10–20 ng/μL and size selected to 200–500 bp, and cDNA libraries were single-read sequenced on an Illumina NextSeq 500 system using 75-bp read length.

Read processing and alignment

The quality of each RNA-seq library was assessed using FastQC v0.11.5 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and then processed with Trimmomatic v0.36 [67] to remove Illumina TruSeq adapter sequences, leading and trailing bases with a Phred

quality score below 20, and trim reads with an average base quality score of 20 over a 4-bp sliding window. All reads less than 40 nucleotides in length after trimming were discarded from further analysis.

The remaining reads of each library were aligned to the corresponding genomes using Bowtie2 v2.2.9 [68], and alignments were filtered with Samtools v1.3.1 [69] using a MAPQ cutoff of 15. For *S.* Typhimurium D23580, reads were aligned to the sequences of the chromosome and the pSLT-BT, pBT1, pBT2, and pBT3 plasmids (accession: PRJEB28511). For *S.* Typhimurium 4/74, reads were aligned to the sequences of the published 4/74 chromosome and the plasmids pSLT^{SL1344}, pCol1B9^{SL1344}, and pRSF1010^{SL1344} (accession: CP002487, HE654724, HE654725, and HE654726, respectively). The RNA-seq mapping statistics are detailed in S4 Table. Reads were assigned to genomic features using featureCounts v1.5.1 [70].

The complete RNA-seq pipeline used for this study is described in https://github.com/will-rowe/rnaseq.

Two strain-specific browsers were generated for the visualization of the transcriptional data in a genomic context online (http://bioinf.gen.tcd.ie/cgi-bin/salcom_v2.pl). The different tracks in each JBrowse [52] were normalized using a published approach [71].

Quantifying differences in expression with only one biological replicate

Expression levels of strain D23580 were calculated as TPM values [41,42] that were generated for coding genes and noncoding sRNAs in the chromosome and pSLT-BT and pBT1 plasmids using the reannotated D23580_liv genome (S2 Table). For strain 4/74, TPM values were recalculated from our published RNA-seq data [37,40] for coding genes and noncoding sRNAs in the chromosome and the three plasmids pSLT$^{4/74}$, pCol1B9$^{4/74}$, and pRSF1010$^{4/74}$ [35]. Based on those values and following previously described Materials and Methods, the expression cutoff was set as TPM > 10 for genes and sRNAs [37].

For comparative analysis between the two *S.* Typhimurium strains D23580 and 4/74, TPM values were obtained for the 4,675 orthologous genes and noncoding sRNAs. These values were used to calculate fold-changes between strains. TPM values ≤10 (representing nonexpressed genes or sRNAs) were set to 10 before calculation of fold-changes. Because of the availability of only one biological replicate per growth condition, a conservative cutoff of ≥3 fold-change was used as a differential expression threshold between strains.

Differential gene expression analysis with three biological replicates

Raw read counts from the 4,674 orthologous coding genes and noncoding sRNAs for the three replicates of the five conditions (ESP, anaerobic growth, NonSPI2, InSPI2, and macrophage) for 4/74 and D23580 were uploaded into Degust (S2 Data) (http://degust.erc.monash.edu/). To identify statistically significant gene expression differences between the two bacterial strains, data were analyzed using the Voom/Limma approach [72,73] with an FDR of ≤0.001 and Log$_2$ fold-change of ≥1. Pairwise comparisons were generated between the two strains for each specific condition. To remove genes with low counts across all samples, thresholds of ≥10 read counts and ≥1 count per million (CPM) in at least the three biological replicates of one sample were used [72,74].

For differential expression analysis of the D23580 ΔpBT1 strain grown in ESP, two RNA-seq biological replicates were compared with the three biological replicates of D23580 WT.

Sample processing for proteomics

An LC-MS/MS (Q Exactive Orbitrap; Thermo Fisher Scientific, Waltham, MA, USA) 4-h reversed-phase C18 gradient was used to generate proteomic data from six biological replicates

of each strain, 4/74 and D23580, grown in the ESP condition in LB. The pellet from bacterial cultures was resuspended in 50 mM phosphate buffer (pH 8), sonicated (10 sec on, 50 sec off, for 10 cycles at 30% amplitude), and supernatants were analyzed after centrifugation at 16,000 × *g* for 20 min. Subsequent experimental procedures were performed at the Centre for Proteome Research at the University of Liverpool, UK. In brief, 100 μg of protein were digested (RapiGest, in-solution trypsin digestion), and 1 μg of digested protein was run on an LC-MS/MS platform.

Analysis of proteomic data

A database was generated merging the amino acid sequences of the annotated genes in 4/74 [35] and our reannotated D23580 to allow homologous proteins as well as strain-specific proteins to be identified. The merged database was clustered using the program Cd-hit and an identity threshold of 95% [75]. Clusters with a single protein, representing strain-specific proteins, were included in the database with their accession ID. Clusters with more than one protein represented orthologs, and only peptides common to all proteins of the cluster were included in the database. Common peptides allowed label-free comparison of proteins that had a low level of sequence variation.

Raw data obtained from the LC-MS/MS platform (data available from the ProteomeXchange Consortium via the PRIDE database [76]) were loaded into the Progenesis QI software (Nonlinear Dynamics, Newcastle upon Tyne, UK) for label-free quantification analysis. Differential expression analysis between the two strains, 4/74 and D23580, is shown in S7 Table. From those results (2,013 proteins), multihit proteins (peptides assigned to more than one protein in the same strain) were removed, leaving a total of 2,004 proteins. Cutoffs of ≥2 unique peptides per identified protein (1,632 proteins), ≥2 fold-change expression, and *p*-value <0.05 between strains (121 proteins) were used. Among the 121 proteins, 25 were 4/74-specific, 30 were D23580-specific, and 66 were encoded by orthologous genes between strains.

Alpha-galactosidase activity

To assess alpha-galactosidase activity, strains were grown on Pinnacle *Salmonella* ABC (chromogenic *Salmonella* medium, Lab M). Bacteria that are able to produce alpha-galactosidase in the absence of beta-galactosidase appear as green colonies on this medium because of the hydrolysis of X-alpha-Gal. This enzymatic activity was correlated to the ability to grow in M9 minimal medium with melibiose as the sole carbon source.

Construction of scarless single-nucleotide substitution mutants

Two strategies were used for single-nucleotide replacement as previously described [20]. For D23580 *flhA*^4/74, a single-strand DNA oligonucleotide recombination approach was used [77]. Briefly, the flhA-474SNP oligonucleotide containing the SNP in 4/74 ("C") was used to replace the SNP in D23580 ("T"). The methodology followed the same strategy used for λ Red recombination explained below. After electroporation of the ssDNA oligonucleotide into D23580 carrying the pSIM5-*tet* plasmid, screening for D23580 recombinants was performed using a PCR with a stringent annealing temperature and primers Fw-flhA and Rv-flhA-474SNP. The reverse primer contained the 4/74 SNP in *flhA*. The SNP mutation in D23580 *flhA*^4/74 was confirmed by Illumina whole-genome sequencing (MicrobesNG, University of Birmingham). Variant-calling bioinformatic analysis confirmed the presence of the intended mutation and the absence of any secondary mutations.

The second strategy for constructing scarless SNP mutants followed a previously described approach based on the pEMG suicide plasmid [24,78]. Oligonucleotides melR-EcoRI-F and melR-BamHI-R were used to PCR amplify, in 4/74 and D23580, an *melR* region containing the SNP described between strains. Additionally, primers melB-EcoRI-F and melB-BamHI-R were used for amplification, in 4/74 and D23580, of a *melB* region containing the two SNPs described in this gene. PCR products were cloned into the pEMG suicide plasmid and transformed into *Escherichia coli* S17-1 λ*pir*. The resulting recombinant plasmids were conjugated into 4/74 or D23580, depending on the strain that was used for the PCR amplification. For *S*. Typhimurium 4/74, transconjugants were selected on M9 minimal medium with 0.2% glucose and Km. For *S*. Typhimurium D23580, transconjugants were selected on LB Cm Km plates. As described previously [24], transconjugants were transformed with the pSW-2 plasmid to promote the loss of the integrated pEMG by a second homologous recombination. The single-nucleotide substitutions were confirmed by PCR amplification with external primers and sequencing. Mutants D23580 *melR*+ and D23580 *melR*+*melB*+ were confirmed by Illumina whole-genome sequencing (MicrobesNG, University of Birmingham). Variant-calling bioinformatic analysis confirmed the intended mutations and the absence of any secondary mutations in D23580 *melR*+*melB*+. The D23580 *melR*+ mutant had a secondary spontaneous synonymous mutation at the chromosomal location 436,081 in *STMMW_04211* (GCC → GCA).

Construction of the Δ*cysS*^{chr} and Δ*cysS*^{pBT1} mutants in *S*. Typhimurium D23580 by λ Red recombineering

The D23580 mutants in *cysS*^{chr} (*STMMW_06051*) and *cysS*^{pBT1} (*pBT1-0241*) were constructed using the λ Red recombination strategy [79]. The Km resistance cassette (*aph*) of pKD4 was amplified by PCR using the primer pairs NW_206/NW_207 and NW_210/NW_211, respectively. The resulting PCR fragments were electroporated into D23580 carrying the recombineering plasmid pSIM5-*tet* following the previously described methodology [20,80]. The Δ*cysS*^{chr}::*aph* mutation was transduced into WT D23580 using the high-frequency-transducing bacteriophage P22 HT 105/1 *int-201* [81] as previously described [24]. The D23580 Δ*cysS*^{pBT1}:: *aph* mutant was whole-genome sequenced using the Illumina technology (MicrobesNG, University of Birmingham). Variant-calling bioinformatic analysis confirmed the intended mutation and the absence of secondary nonintended mutations with the exception of a six-nucleotide insertion in a noncoding region at the chromosomal position 2,755,248 (A → AGCAAGG). The Km resistance cassettes of the two recombinant strains, Δ*cysS*^{chr}::*aph* and D23580 Δ*cysS*^{pBT1}::*aph*, were flipped out using the FLP recombinase expression plasmid pCP20-TcR [22].

Construction of the *S*. Typhimurium D23580 pBT1-cured strain

The pBT1 plasmid was cured from D23580 using published methodology [82]. First, the *pBT1-0211* gene of pBT1, encoding a putative RelE/StbE replicon stabilization toxin, was replaced by an I-SceI-*aph* module by λ Red recombination. The I-SceI-*aph* module was amplified from pKD4-I-SceI [24] using primers NW_163 and NW_164, and the resulting PCR fragment was electroporated into D23580 carrying pSIM5-*tet*. The resulting Δ*pBT1-0211*::I-SceI-*aph* mutants were selected on LB Km plates, and the mutation was transduced into WT D23580 as described above. D23580 Δ*pBT1-0211*::I-SceI-*aph* was subsequently transformed with the I-SceI meganuclease-producing plasmid pSW-2 [78], and transformants were selected on LB Gm agar plates supplemented with 1 mM *m*-toluate, which induces high expression of the I-SceI nuclease from pSW-2. The absence of pBT1 was confirmed by whole-genome sequencing of the D23580 Δ*pBT1* strain (MicrobesNG, University of Birmingham).

Growth curves in M9 melibiose, LB, and M9 glucose media

Overnight bacterial cultures were washed twice with PBS and resuspended in the specific growth medium at an optical density $(OD)_{600\ nm}$ of 0.01. Growth curves of strains grown in LB and M9 minimal medium supplemented with melibiose or glucose were based on OD at 600 nm measurements every hour of samples growing in a 96-well plate. Microplates were incubated at 37°C on an orbital shaker set at 500 rpm in a FLUOstar Omega (BMG Labtech) plate reader. Only the values of the $OD_{600\ nm}$ at 8 h were plotted for strains grown in M9 melibiose medium.

Analysis of SNP conservation in the melibiose utilization operon

The conservation of the two SNPs in *melB* and one SNP in *melR* that distinguished the *S.* Typhimurium strains D23580 and 4/74 was analyzed in the genomes of 258 *S.* Typhimurium ST313 isolates from Malawi and the United Kingdom. The A5 assembly pipeline [83] and ABACAS [84] were used when a reference-quality genome was not available. The PanSeq package allowed the identification of core genome SNPs [85], and the concatenated SNP alignment served to obtain a maximum-likelihood phylogenetic tree using PhylML [86]. BLASTn was used to identify the genotype of the melibiose SNPs shown in Fig 7C in all genomes (S8 Table).

Conservation of pBT1 and pSLT-BT plasmids among ST313 isolates

For phylogenetic analysis of ST313 isolates, all available FASTQ data were downloaded from the ENA using FASTQ dump v2.8.2 (accessions in S8 Table, ENA access date: 01.02.2017). Data quality was assessed using FastQC v0.11.5 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and then processed with Trimmomatic v0.36 [67] to any adapter sequences, leading and trailing bases with a Phred quality score below 20, and trim reads with an average base quality score of 20 over a 4-bp sliding window. All reads less than 40 nucleotides in length after trimming were discarded from further analysis.

A multiple-sequence alignment was generated by mapping isolate FASTQ data to the ST313 D23580 reference genome (pSLT-BT and pBT1 plasmids) (accession: PRJEB28511) using Bowtie2 v2.2.9 [68]. Alignments were filtered (MAPQ cutoff 15) and then deduplicated, sorted, and variant called with Samtools v1.3.1 [69]. For each alignment, recombination was masked using Gubbins v2.2.0 [87] and the variable sites were used to construct a maximum-likelihood tree using RAxML [88]. Phylogenetic trees were visualized using Figtree (http://tree.bio.ed.ac.uk/software/figtree/) and Dendroscope [89]. Coverage information was extracted from the alignment files using bedtools v2.26.0 [90] and visualized using R. Results are shown in Fig 9D (S8 Table).

Statistical analysis for phenotypic studies

One-way ANOVA and Tukey's multiple comparison test were performed using GraphPad Prism 6.0 (GraphPad Software Inc., La Jolla, CA, USA).

Supporting information

S1 Table. Phenotypic features that distinguish *S.* Typhimurium ST313 lineage 2 from ST19 isolates from the literature. ST, sequence type.
(PDF)

S2 Table. Complete *S*. Typhimurium D23580 updated annotation, including 4/74 orthologs.
(XLSX)

S3 Table. SNPs, MNPs, and indels between *S*. Typhimurium 4/74 and D23580. MNP, multinucleotide polymorphism; SNP, single-nucleotide polymorphism.
(XLSX)

S4 Table. RNA-seq sequence reads for *S*. Typhimurium 4/74 and D23580. RNA-seq, RNA sequencing.
(XLSX)

S5 Table. TPM values for *S*. Typhimurium 4/74 and D23580 from the two RNA-seq datasets. RNA-seq, RNA sequencing; TPM, transcripts per million.
(XLSX)

S6 Table. RNA-seq results for *S*. Typhimurium 4/74 and D23580 from Degust. RNA-seq, RNA sequencing.
(XLSX)

S7 Table. Proteomic data of *S*. Typhimurium 4/74 and D23580 grown in ESP. Complete output table from Progenesis analysis of the *S*. Typhimurium 4/74 and D23580 proteomic data. ESP, early stationary phase.
(XLSX)

S8 Table. Conservation of the SNPs in the melibiose operon and the pBT1 and pSLT-BT plasmids among *S*. Typhimurium ST313 isolates. SNP, single-nucleotide polymorphism; ST, sequence type.
(XLSX)

S9 Table. Bacterial strains and plasmids.
(PDF)

S10 Table. Infection-relevant growth conditions used for the RNA-seq experiments in this study. RNA-seq, RNA sequencing.
(PDF)

S11 Table. Oligonucleotides used in this study.
(PDF)

S1 Text. Supporting materials and methods.
(PDF)

S1 Fig. Schematic representation of the RNA-seq-based comparative transcriptomic approach. RNA-seq, RNA sequencing.
(TIF)

S2 Fig. Novel *S*. Typhimurium D23580 noncoding sRNAs. Northern blots confirming the existence of novel sRNAs annotated in the BTP1 prophage region (S1 Text). For every individual sRNA, a northern blot and mapped reads in the same conditions are shown. The arrowheads indicate the most prominent bands. 5S rRNA was used as a loading control. Estimated length of the sRNAs is in brackets and was based on RNA-seq data and sequence analysis. TSSs of the individual sRNAs are indicated at the bottom. RNA-seq, RNA sequencing; sRNA, small RNA; TSS, transcriptional start site; 5S rRNA, ribosomal 5S RNA.
(TIF)

S3 Fig. Transcriptional response to infection-relevant stress of *S*. Typhimurium 4/74 and D23580. (A) Percentage of expressed genes (TPM >10) for each individual strain (S5 Table). (B) Number of coding genes and sRNAs differentially expressed (fold-change ≥3) for each of the 17 infection-relevant conditions (S5 Table). (C) Heat map of the cluster analysis of all orthologous coding genes and sRNAs between the two strains obtained using GeneSpring GX7.3 (Agilent). The TPM for each coding gene and sRNA in each condition in D23580 was divided by the TPM value for the same gene/sRNA and condition in 4/74. TPM values ≤10 (representing nonexpressed genes or sRNAs) were set to 10 before calculating fold-changes (S5 Table). (D) Bubble chart for 4/74 representing up-regulated coding genes and sRNAs versus down-regulated. The following comparisons based on TPM values were obtained for each specific condition: MEP, LEP, ESP, and LSP were compared to EEP; NaCl shock, bile shock, low Fe^{2+} shock, and anaerobic shock were compared to MEP; oxygen shock was compared to anaerobic growth; peroxide shock and nitric oxide shocks were compared to InSPI2; InSPI2 was compared to NonSPI2; and macrophage was compared to ESP. (E) Bubble chart for D23580. EEP, early exponential phase; ESP, early stationary phase; InSPI2, SPI-2-inducing; LEP, late exponential phase; LSP, late stationary phase; MEP, middle exponential phase; NonSPI2, SPI-2-noninducing; SPI, *Salmonella* pathogenicity island; sRNA, small RNA; TPM, transcripts per million.
(TIF)

S4 Fig. The Tn*21*-like antibiotic resistance cassette is inserted in the *mig-5* operon, preventing expression of *rlgAb*, *rlgAa*, and *pSLT043*. Visualization of the RNA-seq data in the 17 infection-relevant conditions from the online JBrowse resources provided in this study. Red arrows represent genes that showed up-regulation in 4/74 versus D23580, and blue arrows represent D23580-down-regulated genes. Scale of the mapped reads was 1 to 500. The insertion of the Tn*21*-like element is indicated by dotted lines. RNA-seq, RNA sequencing.
(TIF)

S5 Fig. Differences in expression of the flagellar regulon between *S*. Typhimurium 4/74 and D23580. (A) Heat map of the flagellar regulon genes representing the relative expression of D23580 versus 4/74 in five growth conditions. TPM values were obtained from the RNA-seq dataset with only one biological replicate. (B) Swimming motility assay of 4/74, D23580, and D23580 *flhA*$^{4/74}$ (S1 Text, S1 Data). Bars represent the mean of 12 independent replicates and standard deviation. Statistical analysis was determined by one-way ANOVA and Tukey's multiple comparison test. Significant differences indicate ****, *p*-value <0.0001; and **, *p*-value <0.01. (C) LDH cytotoxicity assay using C57BL/6 BMDM (S1 Text, S1 Data). Bars represent the mean of six independent replicates and standard deviation. Groups were compared using one-way ANOVA and Tukey's multiple comparisons test, significant differences indicate ****, *p*-value <0.0001; and **, *p*-value <0.01. (D) Heat map of the flagellar regulon using the RNA-seq data with three biological replicates. Results represent the fold-change (D23580 versus 4/74) and FDR values obtained from Degust. BMDM, bone marrow-derived macrophages; FC, fold-change; FDR, false discovery rate; LDH, lactate dehydrogenase; RNA-seq, RNA sequencing; TPM, transcripts per million.
(TIF)

S6 Fig. SPI-associated genes differentially expressed between *S*. Typhimurium 4/74 and D23580. Heat map of the SPI genes and sRNAs that show ≥2 fold-change and ≤0.001 FDR (D23580 versus 4/74), obtained using GeneSpring GX7.3 (Agilent). The CPM values of three biological replicates for each coding gene and sRNA in each condition in D23580 were compared to the CPM values for the same gene/sRNA and condition in 4/74 (S6 Table). Only fold-

changes ≥2 with an FDR ≤0.001 are represented in red (D23580-up-regulated) or blue (D23580-down-regulated). CPM, counts per million; FDR, false discovery rate; SPI, *Salmonella* pathogenicity island; sRNA, small RNA.
(TIF)

S7 Fig. One indel distinguishes the promoter region of *STM2475* between *S*. Typhimurium 4/74 and D23580. (A) RNA-seq data of 4/74 and D23580 in the 17 in vitro growth conditions for the *STM2475* gene. Previously identified TSSs are indicated in both strains [20,37]. Scale of the mapped reads was 1 to 500. (B) *STM2475* promoter regions of the secondary TSS in 4/74 and the corresponding primary TSS in D23580 highlight an "A" insertion in the −10 element of D23580. RNA-seq, RNA sequencing; TSS, transcriptional start site.
(TIF)

S8 Fig. Reproducibility of transcriptomic experiments. Correlation coefficient plots of Log$_2$[TPM values] for five infection-relevant conditions. The "different sequencing runs" plots compare the two RNA-seq datasets (one biological replicate versus three biological replicates). The "same sequencing run" plots compare two samples of the RNA-seq dataset with three biological replicates. (A) 4/74, (B) D23580. RNA-seq, RNA sequencing; TPM, transcripts per million.
(TIF)

S1 Data. Raw data of phenotypic assays.
(XLSX)

S2 Data. Input file for Degust. Data include counts for the three biological replicates in five growth conditions (ESP, anaerobic growth, NonSPI2, InSPI2, macrophage) in 4/74 and D23580, and for the two biological replicates of D23580 ΔpBT1 in ESP. ESP, early stationary phase; InSPI2, SPI-2-inducing; NonSPI2, SPI-2-noninducing; SPI, *Salmonella* pathogenicity island.
(CSV)

Author Contributions

Conceptualization: Rocío Canals, Jay C. D. Hinton.

Data curation: Rocío Canals, Will P. M. Rowe.

Formal analysis: Rocío Canals, Siân V. Owen, Philip Brownridge, Roy R. Chaudhuri, Will P. M. Rowe, Alexander V. Predeus.

References

1. Feasey NA, Dougan G, Kingsley RA, Heyderman RS, Gordon MA. Invasive non-typhoidal salmonella disease: an emerging and neglected tropical disease in Africa. Lancet. 2012; 379: 2489–2499. https://doi.org/10.1016/S0140-6736(11)61752-2 PMID: 22587967

2. Ao TT, Feasey NA, Gordon MA, Keddy KH, Angulo FJ, Crump JA. Global Burden of Invasive Nonty-phoidal Salmonella Disease, 20101. Emerg Infect Dis. 2015; 21: 941–949. https://doi.org/10.3201/eid2106.140999 PMID: 25860298

3. Uche IV, MacLennan CA, Saul A. A Systematic Review of the Incidence, Risk Factors and Case Fatality Rates of Invasive Nontyphoidal Salmonella (iNTS) Disease in Africa (1966 to 2014). PLoS Negl Trop Dis. 2017; 11: e0005118. https://doi.org/10.1371/journal.pntd.0005118 PMID: 28056035

4. Okoro CK, Kingsley RA, Connor TR, Harris SR, Parry CM, Al-Mashhadani MN, et al. Intra-continental spread of human invasive Salmonella Typhimurium pathovariants in sub-Saharan Africa. Nat Genet. 2012; 44: 1215–1221. https://doi.org/10.1038/ng.2423 PMID: 23023330

5. Ashton PM, Owen SV, Kaindama L, Rowe WPM, Lane CR, Larkin L, et al. Public health surveillance in the UK revolutionises our understanding of the invasive Salmonella Typhimurium epidemic in Africa. Genome Med. 2017; 9(1):92. https://doi.org/10.1186/s13073-017-0480-7 PMID: 29084588

6. Almeida F, Seribelli AA, da Silva P, Medeiros MIC, dos Prazeres Rodrigues D, Moreira CG, et al. Multilocus sequence typing of Salmonella Typhimurium reveals the presence of the highly invasive ST313 in Brazil. Infect Genet Evol. 2017; 51: 41–44. https://doi.org/10.1016/j.meegid.2017.03.009 PMID: 28288927

7. Fields PI, Swanson RV, Haidaris CG, Heffron F. Mutants of Salmonella typhimurium that cannot survive within the macrophage are avirulent. Proc Natl Acad Sci. 1986; 83: 5189–5193. https://doi.org/10.1073/pnas.83.14.5189 PMID: 3523484

8. Langridge GC, Fookes M, Connor TR, Feltwell T, Feasey N, Parsons BN, et al. Patterns of genome evolution that have accompanied host adaptation in Salmonella. Proc Natl Acad Sci. 2015; 112: 863–868. https://doi.org/10.1073/pnas.1416707112 PMID: 25535353

9. Tanner JR, Kingsley RA. Evolution of Salmonella within Hosts. Trends Microbiol. 2018; 26(12): 986–998. https://doi.org/10.1016/j.tim.2018.06.001 PMID: 29954653

10. Wheeler NE, Gardner PP, Barquist L. Machine learning identifies signatures of host adaptation in the bacterial pathogen Salmonella enterica. PLoS Genet. 2018; 14: e1007333. https://doi.org/10.1371/journal.pgen.1007333 PMID: 29738521

11. Nuccio S-P, Bäumler AJ. Comparative Analysis of Salmonella Genomes Identifies a Metabolic Network for Escalating Growth in the Inflamed Gut. mBio. 2014; 5: e00929–14. https://doi.org/10.1128/mBio.00929-14 PMID: 24643865

12. Lokken KL, Walker GT, Tsolis RM. Disseminated infections with antibiotic-resistant non-typhoidal Salmonella strains: contributions of host and pathogen factors. Pathog Dis. 2016; 74: ftw103. https://doi.org/10.1093/femspd/ftw103 PMID: 27765795

13. Ramachandran G, Perkins DJ, Schmidlein PJ, Tulapurkar ME, Tennant SM. Invasive Salmonella Typhimurium ST313 with Naturally Attenuated Flagellin Elicits Reduced Inflammation and Replicates within Macrophages. PLoS Negl Trop Dis. 2015; 9: e3394. https://doi.org/10.1371/journal.pntd.0003394 PMID: 25569606

14. Goh YS, MacLennan CA. Invasive African nontyphoidal Salmonella requires high levels of complement for cell-free antibody-dependent killing. J Immunol Methods. 2013; 387: 121–129. https://doi.org/10.1016/j.jim.2012.10.005 PMID: 23085530

15. Okoro CK, Barquist L, Connor TR, Harris SR, Clare S, Stevens MP, et al. Signatures of Adaptation in Human Invasive Salmonella Typhimurium ST313 Populations from Sub-Saharan Africa. PLoS Negl Trop Dis. 2015; 9(6): e0003848. https://doi.org/10.1371/journal.pntd.0003848 PMID: 26076129

16. Ramachandran G, Panda A, Higginson EE, Ateh E, Lipsky MM, Sen S, et al. Virulence of invasive Salmonella Typhimurium ST313 in animal models of infection. PLoS Negl Trop Dis. 2017; 11: e0005697. https://doi.org/10.1371/journal.pntd.0005697 PMID: 28783750

17. Singletary LA, Karlinsey JE, Libby SJ, Mooney JP, Lokken KL, Tsolis RM, et al. Loss of Multicellular Behavior in Epidemic African Nontyphoidal Salmonella enterica Serovar Typhimurium ST313 Strain D23580. mBio. 2016; 7(2): e02265. https://doi.org/10.1128/mBio.02265-15 PMID: 26933058

18. Yang J, Barrila J, Roland KL, Kilbourne J, Ott CM, Forsyth RJ, et al. Characterization of the Invasive, Multidrug Resistant Non-typhoidal Salmonella Strain D23580 in a Murine Model of Infection. PLoS Negl Trop Dis. 2015; 9: e0003839. https://doi.org/10.1371/journal.pntd.0003839 PMID: 26091096

19. Kingsley RA, Msefula CL, Thomson NR, Kariuki S, Holt KE, Gordon MA, et al. Epidemic multiple drug resistant Salmonella Typhimurium causing invasive disease in sub-Saharan Africa have a distinct genotype. Genome Res. 2009; 19: 2279–2287. https://doi.org/10.1101/gr.091017.109 PMID: 19901036

20. Hammarlöf DL, Kröger C, Owen SV, Canals R, Lacharme-Lora L, Wenner N, et al. Role of a single non-coding nucleotide in the evolution of an epidemic African clade of Salmonella. Proc Natl Acad Sci U S A. 2018; 115: E2614–E2623. https://doi.org/10.1073/pnas.1714718115 PMID: 29487214

21. Ramachandran G, Aheto K, Shirtliff ME, Tennant SM. Poor biofilm-forming ability and long-term survival of invasive Salmonella Typhimurium ST313. Pathog Dis. 2016; 74(5): ftw049. https://doi.org/10.1093/femspd/ftw049 PMID: 27222487

22. Kintz E, Davies MR, Hammarlöf DL, Canals R, Hinton JCD, van der Woude MW. A BTP1 prophage gene present in invasive non-typhoidal Salmonella determines composition and length of the O-antigen of the lipopolysaccharide. Mol Microbiol. 2015; 96: 263–275. https://doi.org/10.1111/mmi.12933 PMID: 25586744

23. Micoli F, Ravenscroft N, Cescutti P, Stefanetti G, Londero S, Rondini S, et al. Structural analysis of O-polysaccharide chains extracted from different Salmonella Typhimurium strains. Carbohydr Res. 2014; 385: 1–8. https://doi.org/10.1016/j.carres.2013.12.003 PMID: 24384528

24. Owen SV, Wenner N, Canals R, Makumi A, Hammarlöf DL, Gordon MA, et al. Characterization of the Prophage Repertoire of African Salmonella Typhimurium ST313 Reveals High Levels of Spontaneous Induction of Novel Phage BTP1. Front Microbiol. 2017; 8: 235. https://doi.org/10.3389/fmicb.2017.00235 PMID: 28280485

25. Carden S, Okoro C, Dougan G, Monack D. Non-typhoidal Salmonella Typhimurium ST313 isolates that cause bacteremia in humans stimulate less inflammasome activation than ST19 isolates associated with gastroenteritis. Pathog Dis. 2015; 73(4): ftu023. https://doi.org/10.1093/femspd/ftu023 PMID: 25808600

26. Barrila J, Yang J, Crabbé A, Sarker SF, Liu Y, Ott CM, et al. Three-dimensional organotypic co-culture model of intestinal epithelial cells and macrophages to study Salmonella enterica colonization patterns. NPJ Microgravity. 2017; 3. https://doi.org/10.1038/s41526-017-0011-2 PMID: 28649632

27. Herrero-Fresno A, Wallrodt I, Leekitcharoenphon P, Olsen JE, Aarestrup FM, Hendriksen RS. The Role of the st313-td Gene in Virulence of Salmonella Typhimurium ST313. PLoS ONE. 2014; 9: e84566. https://doi.org/10.1371/journal.pone.0084566 PMID: 24404174

28. Carden SE, Walker GT, Honeycutt J, Lugo K, Pham T, Jacobson A, et al. Pseudogenization of the Secreted Effector Gene sseI Confers Rapid Systemic Dissemination of S. Typhimurium ST313 within Migratory Dendritic Cells. Cell Host Microbe. 2017; 21: 182–194. https://doi.org/10.1016/j.chom.2017.01.009 PMID: 28182950

29. Parsons BN, Humphrey S, Salisbury AM, Mikoleit J, Hinton JCD, Gordon MA, et al. Invasive Non-Typhoidal Salmonella Typhimurium ST313 Are Not Host-Restricted and Have an Invasive Phenotype in Experimentally Infected Chickens. PLoS Negl Trop Dis. 2013; 7: e2487. https://doi.org/10.1371/journal.pntd.0002487 PMID: 24130915

30. Branchu P, Bawn M, Kingsley RA. Genome Variation and Molecular Epidemiology of Salmonella enterica Serovar Typhimurium Pathovariants. Infect Immun. 2018; 86: e00079–18. https://doi.org/10.1128/IAI.00079-18 PMID: 29784861

31. Lee S-J, Liang L, Juarez S, Nanton MR, Gondwe EN, Msefula CL, et al. Identification of a common immune signature in murine and human systemic Salmonellosis. Proc Natl Acad Sci. 2012; 109: 4998–5003. https://doi.org/10.1073/pnas.1111413109 PMID: 22331879

32. McClelland M, Sanderson KE, Spieth J, Clifton SW, Latreille P, Courtney L, et al. Complete genome sequence of *Salmonella enterica* serovar Typhimurium LT2. Nature. 2001; 413: 852–856. https://doi.org/10.1038/35101614 PMID: 11677609

33. Wilmes-Riesenberg MR, Foster JW, Curtiss R. An altered rpoS allele contributes to the avirulence of Salmonella typhimurium LT2. Infect Immun. 1997; 65: 203–210. PMID: 8975913

34. Richardson EJ, Limaye B, Inamdar H, Datta A, Manjari KS, Pullinger GD, et al. Genome Sequences of Salmonella enterica Serovar Typhimurium, Choleraesuis, Dublin, and Gallinarum Strains of Well-Defined Virulence in Food-Producing Animals. J Bacteriol. 2011; 193: 3162–3163. https://doi.org/10.1128/JB.00394-11 PMID: 21478351

35. Kröger C, Dillon SC, Cameron ADS, Papenfort K, Sivasankaran SK, Hokamp K, et al. The transcriptional landscape and small RNAs of Salmonella enterica serovar Typhimurium. Proc Natl Acad Sci. 2012; 109: E1277–E1286. https://doi.org/10.1073/pnas.1201061109 PMID: 22538806

36. Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. Fly (Austin). 2012; 6: 80–92. https://doi.org/10.4161/fly.19695 PMID: 22728672

37. Kröger C, Colgan A, Srikumar S, Händler K, Sivasankaran SK, Hammarlöf DL, et al. An Infection-Relevant Transcriptomic Compendium for Salmonella enterica Serovar Typhimurium. Cell Host Microbe. 2013; 14: 683–695. https://doi.org/10.1016/j.chom.2013.11.010 PMID: 24331466

38. Hardt WD, Urlaub H, Galán JE. A substrate of the centisome 63 type III protein secretion system of Salmonella typhimurium is encoded by a cryptic bacteriophage. Proc Natl Acad Sci U S A. 1998; 95: 2574–2579. PMID: 9482928

39. Eriksson S, Lucchini S, Thompson A, Rhen M, Hinton JCD. Unravelling the biology of macrophage infection by gene expression profiling of intracellular Salmonella enterica. Mol Microbiol. 47: 103–118. https://doi.org/10.1046/j.1365-2958.2003.03313.x PMID: 12492857

40. Srikumar S, Kröger C, Hébrard M, Colgan A, Owen SV, Sivasankaran SK, et al. RNA-seq Brings New Insights to the Intra-Macrophage Transcriptome of Salmonella Typhimurium. PLoS Pathog. 2015; 11: e1005262. https://doi.org/10.1371/journal.ppat.1005262 PMID: 26561851

41. Wagner GP, Kin K, Lynch VJ. Measurement of mRNA abundance using RNA-seq data: RPKM measure is inconsistent among samples. Theory Biosci. 2012; 131: 281–285. https://doi.org/10.1007/s12064-012-0162-3 PMID: 22872506

42. Wagner GP, Kin K, Lynch VJ. A model based criterion for gene expression calls using RNA-seq data. Theory Biosci. 2013; 132: 159–164. https://doi.org/10.1007/s12064-013-0178-3 PMID: 23615947

43. Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, et al. Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. Mol Syst Biol. 2006; 2: 2006.0008. https://doi.org/10.1038/msb4100050 PMID: 16738554

44. Minamino T. Hierarchical protein export mechanism of the bacterial flagellar type III protein export apparatus. FEMS Microbiol Lett. 2018; 365: fny117. https://doi.org/10.1093/femsle/fny117 PMID: 29850796

45. Spöring I, Felgner S, Preuße M, Eckweiler D, Rohde M, Häussler S, et al. Regulation of Flagellum Biosynthesis in Response to Cell Envelope Stress in Salmonella enterica Serovar Typhimurium. mBio. 2018; 9: e00736–17. https://doi.org/10.1128/mBio.00736-17 PMID: 29717015

46. Takaya A, Erhardt M, Karata K, Winterberg K, Yamamoto T, Hughes KT. YdiV: a dual function protein that targets FlhDC for ClpXP-dependent degradation by promoting release of DNA-bound FlhDC complex. Mol Microbiol. 2012; 83: 1268–1284. https://doi.org/10.1111/j.1365-2958.2012.08007.x PMID: 22380597

47. Wada T, Morizane T, Abo T, Tominaga A, Inoue-Tanaka K, Kutsukake K. EAL domain protein YdiV acts as an anti-FlhD4C2 factor responsible for nutritional control of the flagellar regulon in Salmonella enterica Serovar Typhimurium. J Bacteriol. 2011; 193: 1600–1611. https://doi.org/10.1128/JB.01494-10 PMID: 21278297

48. Lokken KL, Stull-Lane AR, Poels K, Tsolis RM. Malaria parasite-mediated alteration of macrophage function and increased iron availability predispose to disseminated non-typhoidal Salmonella infection. Infect Immun. 2018: IAI.00301–18. https://doi.org/10.1128/IAI.00301-18 PMID: 29986892

49. Casadevall A, Fang FC. Reproducible science. Infect Immun. 2010; 78: 4972–4975. https://doi.org/10.1128/IAI.00908-10 PMID: 20876290

50. Leek JT, Scharpf RB, Bravo HC, Simcha D, Langmead B, Johnson WE, et al. Tackling the widespread and critical impact of batch effects in high-throughput data. Nat Rev Genet. 2010; 11: 733–739. https://doi.org/10.1038/nrg2825 PMID: 20838408

51. Creecy JP, Conway T. Quantitative bacterial transcriptomics with RNA-seq. Curr Opin Microbiol. 2015;0: 133–140. https://doi.org/10.1016/j.mib.2014.11.011 PMID: 25483350

52. Westesson O, Skinner M, Holmes I. Visualizing next-generation sequencing data with JBrowse. Brief Bioinform. 2013; 14: 172–177. https://doi.org/10.1093/bib/bbr078 PMID: 22411711

53. Ethayathulla AS, Yousef MS, Amin A, Leblanc G, Kaback HR, Guan L. Structure-based mechanism for Na+/melibiose symport by MelB. Nat Commun. 2014; 5: 3009. https://doi.org/10.1038/ncomms4009 PMID: 24389923

54. Harvey PC, Watson M, Hulme S, Jones MA, Lovell M, Berchieri A, et al. Salmonella enterica Serovar Typhimurium Colonizing the Lumen of the Chicken Intestine Grows Slowly and Upregulates a Unique Set of Virulence and Metabolism Genes. Infect Immun. 2011; 79: 4105–4121. https://doi.org/10.1128/IAI.01390-10 PMID: 21768276

55. Kaiser BLD, Li J, Sanford JA, Kim Y-M, Kronewitter SR, Jones MB, et al. A Multi-Omic View of Host-Pathogen-Commensal Interplay in Salmonella-Mediated Intestinal Infection. PLoS ONE. 2013; 8: e67155. https://doi.org/10.1371/journal.pone.0067155 PMID: 23840608

56. Mao B, Tang H, Gu J, Li D, Cui S, Zhao J, et al. In vitro fermentation of raffinose by the human gut bacteria. Food Funct. 2018; 9: 5824–5831. https://doi.org/10.1039/c8fo01687a PMID: 30357216

57. Barquist L, Langridge GC, Turner DJ, Phan M-D, Turner AK, Bateman A, et al. A comparison of dense transposon insertion libraries in the Salmonella serovars Typhi and Typhimurium. Nucleic Acids Res. 2013; 41: 4549–4564. https://doi.org/10.1093/nar/gkt148 PMID: 23470992

58. Canals R, Xia X-Q, Fronick C, Clifton SW, Ahmer BM, Andrews-Polymenis HL, et al. High-throughput comparison of gene fitness among related bacteria. BMC Genomics. 2012; 13: 212. https://doi.org/10.1186/1471-2164-13-212 PMID: 22646920

59. Langridge GC, Phan M-D, Turner DJ, Perkins TT, Parts L, Haase J, et al. Simultaneous assay of every Salmonella Typhi gene using one million transposon mutants. Genome Res. 2009; 19: 2308–2316. https://doi.org/10.1101/gr.097097.109 PMID: 19826075

60. Down TA, Piipari M, Hubbard TJP. Dalliance: interactive genome viewing on the web. Bioinformatics. 2011; 27: 889–890. https://doi.org/10.1093/bioinformatics/btr020 PMID: 21252075

61. Barquist L, Mayho M, Cummins C, Cain AK, Boinett CJ, Page AJ, et al. The TraDIS toolkit: sequencing and analysis for dense transposon mutant libraries. Bioinformatics. 2016; 32: 1109–1111. https://doi.org/10.1093/bioinformatics/btw022 PMID: 26794317

62. Tazzyman SJ, Bonhoeffer S. Why There Are No Essential Genes on Plasmids. Mol Biol Evol. 2015; 32: 3079–3088. https://doi.org/10.1093/molbev/msu293 PMID: 25540453

63. Perez-Sepulveda BM, Hinton JCD. Functional Transcriptomics for Bacterial Gene Detectives. Microbiol Spectr. 2018; 6. https://doi.org/10.1128/microbiolspec.RWR-0033-2018 PMID: 30215343

64. Rankin JD, Taylor RJ. The estimation of doses of Salmonella typhimurium suitable for the experimental production of disease in calves. Vet Rec. 1966; 78: 706–707. PMID: 5336163

65. Chin C-S, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, et al. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. Nat Methods. 2013; 10: 563–569. https://doi.org/10.1038/nmeth.2474 PMID: 23644548

66. Wick RR, Judd LM, Gorrie CL, Holt KE. Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. PLoS Comput Biol. 2017; 13: e1005595. https://doi.org/10.1371/journal.pcbi.1005595 PMID: 28594827

67. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014; 30: 2114–2120. https://doi.org/10.1093/bioinformatics/btu170 PMID: 24695404

68. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012; 9: 357–359. https://doi.org/10.1038/nmeth.1923 PMID: 22388286

69. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009; 25: 2078–2079. https://doi.org/10.1093/bioinformatics/btp352 PMID: 19505943

70. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics. 2014; 30: 923–930. https://doi.org/10.1093/bioinformatics/btt656 PMID: 24227677

71. Conway T, Creecy JP, Maddox SM, Grissom JE, Conkle TL, Shadid TM, et al. Unprecedented High-Resolution View of Bacterial Operon Architecture Revealed by RNA Sequencing. mBio. 2014; 5: e01442–14. https://doi.org/10.1128/mBio.01442-14 PMID: 25006232

72. Law CW, Chen Y, Shi W, Smyth GK. voom: precision weights unlock linear model analysis tools for RNA-seq read counts. Genome Biol. 2014; 15: R29. https://doi.org/10.1186/gb-2014-15-2-r29 PMID: 24485249

73. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015; 43: e47–e47. https://doi.org/10.1093/nar/gkv007 PMID: 25605792

74. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 2010; 26: 139–140. https://doi.org/10.1093/bioinformatics/btp616 PMID: 19910308

75. Li W, Godzik A. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. Bioinformatics. 2006; 22: 1658–1659. https://doi.org/10.1093/bioinformatics/btl158 PMID: 16731699

76. Vizcaíno JA, Csordas A, del-Toro N, Dianes JA, Griss J, Lavidas I, et al. 2016 update of the PRIDE database and its related tools. Nucleic Acids Res. 2016; 44: D447–D456. https://doi.org/10.1093/nar/gkv1145 PMID: 26527722

77. Sawitzke JA, Costantino N, Li X-T, Thomason LC, Bubunenko M, Court C, et al. Probing cellular processes with oligo-mediated recombination and using the knowledge gained to optimize recombineering. J Mol Biol. 2011; 407: 45–59. https://doi.org/10.1016/j.jmb.2011.01.030 PMID: 21256136

78. Martínez-García E, Lorenzo V de. Engineering multiple genomic deletions in Gram-negative bacteria: analysis of the multi-resistant antibiotic profile of Pseudomonas putida KT2440. Environ Microbiol. 2011; 13: 2702–2716. https://doi.org/10.1111/j.1462-2920.2011.02538.x PMID: 21883790

79. Datsenko KA, Wanner BL. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. Proc Natl Acad Sci U S A. 2000; 97: 6640–6645. https://doi.org/10.1073/pnas.120163297 PMID: 10829079

80. Koskiniemi S, Pränting M, Gullberg E, Näsvall J, Andersson DI. Activation of cryptic aminoglycoside resistance in Salmonella enterica. Mol Microbiol. 2011; 80: 1464–1478. https://doi.org/10.1111/j.1365-2958.2011.07657.x PMID: 21507083

81. Schmieger H. Phage P22-mutants with increased or decreased transduction abilities. Mol Gen Genet MGG. 1972; 119: 75–88. https://doi.org/10.1007/BF00270447 PMID: 4564719

82. de Moraes MH, Teplitski M. Fast and efficient three-step target-specific curing of a virulence plasmid in Salmonella enterica. AMB Express. 2015; 5(1): 139. https://doi.org/10.1186/s13568-015-0139-y PMID: 26272479

83. Coil D, Jospin G, Darling AE. A5-miseq: an updated pipeline to assemble microbial genomes from Illumina MiSeq data. Bioinformatics. 2015; 31: 587–589. https://doi.org/10.1093/bioinformatics/btu661 PMID: 25338718

84. Assefa S, Keane TM, Otto TD, Newbold C, Berriman M. ABACAS: algorithm-based automatic contiguation of assembled sequences. Bioinformatics. 2009; 25: 1968–1969. https://doi.org/10.1093/bioinformatics/btp347 PMID: 19497936

85. Laing C, Buchanan C, Taboada EN, Zhang Y, Kropinski A, Villegas A, et al. Pan-genome sequence analysis using Panseq: an online tool for the rapid analysis of core and accessory genomic regions. BMC Bioinformatics. 2010; 11: 461. https://doi.org/10.1186/1471-2105-11-461 PMID: 20843356

86. Guindon S, Delsuc F, Dufayard J-F, Gascuel O. Estimating Maximum Likelihood Phylogenies with PhyML. Bioinformatics for DNA Sequence Analysis. Humana Press; 2009. pp. 113–137. https://doi.org/10.1007/978-1-59745-251-9_6

87. Croucher NJ, Page AJ, Connor TR, Delaney AJ, Keane JA, Bentley SD, et al. Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. Nucleic Acids Res. 2015; 43: e15–e15. https://doi.org/10.1093/nar/gku1196 PMID: 25414349

88. Stamatakis A. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics. 2006; 22: 2688–2690. https://doi.org/10.1093/bioinformatics/btl446 PMID: 16928733

89. Huson DH, Scornavacca C. Dendroscope 3: An Interactive Tool for Rooted Phylogenetic Trees and Networks. Syst Biol. 2012; 61: 1061–1067. https://doi.org/10.1093/sysbio/sys062 PMID: 22780991

90. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics. 2010; 26: 841–842. https://doi.org/10.1093/bioinformatics/btq033 PMID: 20110278