


# The Typhi colonization factor (Tcf) is encoded by multiple non-typhoidal Salmonella serovars but exhibits a varying expression profile and interchanging contribution to intestinal colonization

Shalhevet Azriel, Alina Goren, Inna Shomer, Gili Aviv, Galia Rahav & Ohad Gal-Mor

To cite this article: Shalhevet Azriel, Alina Goren, Inna Shomer, Gili Aviv, Galia Rahav & Ohad Gal-Mor (2017): The Typhi colonization factor (Tcf) is encoded by multiple non-typhoidal Salmonella serovars but exhibits a varying expression profile and interchanging contribution to intestinal colonization, *Virulence*, DOI: [10.1080/21505594.2017.1380766](https://doi.org/10.1080/21505594.2017.1380766)

To link to this article: <http://dx.doi.org/10.1080/21505594.2017.1380766>

 View supplementary material 

 Accepted author version posted online: 18 Sep 2017.

 Submit your article to this journal 

 Article views: 23

 View related articles 

 View Crossmark data 

**The Typhi colonization factor (Tcf) is encoded by multiple non-typhoidal *Salmonella* serovars but exhibits a varying expression profile and interchanging contribution to intestinal colonization**

Shalhevet Azriel<sup>1</sup>, Alina Goren<sup>1,2,3</sup>, Inna Shomer<sup>1</sup>, Gili Aviv<sup>1,2,3</sup>, Galia Rahav<sup>1,3</sup>, and Ohad Gal-Mor<sup>1,2,3#</sup>

<sup>1</sup>The Infectious Diseases Research Laboratory, Sheba Medical Center, Tel-Hashomer, Israel;

<sup>2</sup>Department of Clinical Microbiology and Immunology, Tel-Aviv University, Tel-Aviv, Israel

<sup>3</sup>Sackler Faculty of Medicine, Tel-Aviv University, Tel-Aviv, Israel

# Address correspondence to Ohad Gal-Mor, Ohad.Gal-Mor@sheba.health.gov.il

ABSTRACT

*Salmonella enterica* serovars Typhi and Paratyphi A are human-restricted pathogens and the leading causative agents of enteric fever. The Typhi colonization factor (Tcf) is a chaperone-usher fimbria, thought to play a role in the host-specificity of typhoidal serovars. Here we show that the *tcf* cluster (*tcfABCD tinR tioA*) is present in at least 25 non-typhoidal *Salmonella* (NTS) serovars and demonstrate its native expression in clinically-important serovars including Schwarzengrund, 9,12:l,v:-, Choleraesuis, Bredeney, Heidelberg, Montevideo, Virchow and Infantis. Although the genetic organization of the *tcf* cluster is well conserved, the N-terminal half of the fimbrial adhesin, TcfD is highly diverse, suggesting different binding properties of distinct *tcfD* variants. Comparison of *tcfA* expression in typhoidal and NTS serovars demonstrated unexpected differences in its expression profiles, with the highest transcription levels in *S. Typhi*, *S. Choleraesuis* and *S. Infantis*. In the latter, *tcf* is induced in rich broth and under microaerobic conditions, characterizing the intestines of warm blooded animals.

Furthermore, Tcf is negatively regulated by the ancestral leucine-responsive transcriptional regulator (Lrp). Using the colitis mouse model, we demonstrate that during mice infection *tcfA* is expressed at higher levels by *S. Infantis* than *S. Schwarzengrund* or *S. Heidelberg*. Moreover, while Tcf is dispensable for *S. Schwarzengrund* and *S. Heidelberg* mouse colonization, Tcf is involved in cecum and colon colonization by *S. Infantis*. Taken together, our results establish that Tcf is broadly encoded by multiple NTS serovars, but presents variable expression profiles and contributes differently to their virulence.

## INTRODUCTION

The bacterial species *Salmonella enterica* (*S. enterica*) is a Gram negative, facultative intracellular and ubiquitous pathogen that can infect a broad range of animals. This highly diverse species contains more than 2600 serovars, which are classified according to somatic (O) and flagellar (H) antigens <sup>1</sup>. *S. enterica* serovars also differ in their adaptation to various hosts (host-specificity) and the disease they cause. For instance, non-typhoidal *Salmonella* (NTS) serovars such as *S. enterica* serovar Typhimurium (*S. Typhimurium*) or *S. Enteritidis* display a broad-host range, capable of infecting many different animal species including reptiles, birds and mammals. In healthy humans, infection with NTS serovars normally develops into a localized self-limiting inflammation of the terminal ileum and colon, known as gastroenteritis. On the other hand, typhoidal serovars including *S. Typhi*, *S. Paratyphi A* or *S. Sendai* can infect only humans and higher primates, manifesting as invasive, systemic life-threatening disease, called typhoid or enteric-fever (reviewed in <sup>2, 3</sup>). Other *Salmonella* serovars or strains, although not fully host-restricted are well adapted to particular animal hosts. *S. Choleraesuis*, *S. Dublin*, *S. Abortusovis* and *S. Typhimurium* phage types DT2 and DT99 are frequently associated with swine, bovine, sheep, and pigeons, respectively <sup>4-6</sup>.

One of the first events in the establishment of a bacterial infection is attachment to host tissues and colonization <sup>7</sup>. Intimate adhesion by bacteria is mediated by surface-exposed proteinaceous hair-like structures (pili) with adhesive properties, known as fimbriae that bind host receptors expressing specific glycoproteins or glycolipids <sup>8,9</sup>. Most of the known fimbriae in Gram-negative bacteria belong to the canonic chaperone-usher biogenesis pathway. This class of

fimbriae requires a periplasmic chaperone and a pore-forming, outer-membrane protein known as the usher for the fimbriae assembly. During this process, the pilus subunits (pilins) are exported into the periplasm by the general secretory (Sec) pathway and bind to an associated specific chaperone, which facilitates pilins folding and prevents premature aggregation. The pilin-chaperone complex is then delivered to the outer-membrane usher that serve as an assembly and secretion apparatus. Secreted pilins are assembled into a long polymeric linear pilus, consisting of numerous copies of the major pilus subunit <sup>10</sup>.

Chaperone-usher fimbriae genes are often arranged within compact gene clusters encoding at least a major structural pilin subunit, a chaperone, and an usher. More complex fimbrial clusters contain accessory genes encoding for additional structural proteins (e.g. minor fimbrial subunits), additional chaperones, or regulators. The minor pilus subunits, which often contain a lectin domain with adhesive properties to a specific host receptor can be translocated to the tip of the fimbrial rod and function as the fimbrial adhesin <sup>11-13</sup>.

The composition of different fimbriae of a pathogen, known as the 'fimbriome', plays a key role in shaping host-tropism. It has been shown that differences in the presence of fimbriae or even allelic variation within a particular fimbria may significantly affect the interaction of *Salmonella* with its host <sup>12, 14-19</sup>.

The Typhi colonization factor (Tcf) belongs to the  $\alpha$  fimbrial clade <sup>12</sup> of the chaperone-usher fimbriae, which was identified in strains of typhoidal serovars including Typhi <sup>20</sup>, Paratyphi A, and Sendai<sup>17</sup>, but is absent from the genome of the ubiquitous broad-host-range serovars Typhimurium and Enteritidis. Moreover, since Tcf displays homology to the CS1 fimbria of human-specific enterotoxigenic *Escherichia coli* <sup>20</sup> and to the cable (Cbl) pili of

*Burkholderia cepacia*, it was suggested that Tcf plays a role in host specificity of typhoidal serovars to humans<sup>17, 20, 21</sup>. In *S. Typhi*, the *tcf* region contains six ORFs encoding the fimbria chaperone (*tcfA*), major fimbrial subunit (*tcfB*), the usher (*tcfC*) and the adhesin (*tcfD*)<sup>21, 22</sup>. Two additional uncharacterized ORFs with the same transcriptional orientation, are designated *tinR* and *tioA*. Interestingly, the *tcf* cluster is variably present in the *Salmonella* pathogenicity island (SPI) 6, which also encodes a type six secretion system (T6SS) and the *Salmonella* atypical fimbria (*saf*)<sup>20</sup>.

Although our group and others have previously shown the presence of the *tcf* operon in a few NTS serovars<sup>17, 23, 24</sup>, *tcf* cluster conservation, its regulatory setup and its role in NTS pathogenicity remained unknown. In this study, we have determined the distribution and conservation of the *tcf* cluster among NTS serovars, characterized its expression profile in typhoidal and NTS serovars, analyzed its regulatory network and demonstrated its variable role in NTS pathogenicity.

## RESULTS

**Distribution of the *tcf* cluster among NTS serovars**

Although the Tcf was initially reported to be a typhoid-specific colonization factor<sup>20</sup>, several studies have reported the presence of the *tcf* cluster in the genome of a few NTS serovars<sup>17, 23-25</sup>. To gain a more systematic view over the distribution of the *tcf* cluster among NTS serovars, we took advantage of the rapidly increasing number of sequenced *Salmonella* genomes and performed a Blastn search, using the DNA sequence *tcfABCD* from *S. Typhi* against the nr database, containing more than 6000 *S. enterica* genomes. The entire *tcfABCD* cluster was identified in 123 genomes from 25 NTS serovars including Crossness, Heidelberg, Infantis, Krefeld, Bredeney, Muenster, Abaetetuba, Djakarta, Choleraesuis, Cerro, Rubislaw, Saintpaul, Montevideo, Senftenberg, Wandsworth, Schwarzengrund, Minnesota, Antsalova, Johannesburg, Panama, Virchow, Koessen, Quebec, Indiana, and 9,12:l,v:-, many of which belong to clade B of *S. enterica* subspecies I<sup>26</sup>. The *tcf* cluster in these NTS genomes was found to be inserted between the *sinR* and *pagN* genes exhibiting a similar genetic organization as in *S. Typhi* and *S. Paratyphi A* (Fig. 1). These results highlight the presence of the *tcf* cluster in multiple NTS serovars and in a much broader distribution than was previously appreciated.

To get further insights into the intraserovar abundance of the *tcf* operon, we examined the presence of *tcfA* in 22 reference strains and clinical isolates of *S. Infantis* from different sources. With the exception of only one isolate (SARB27), this analysis revealed the presence of *tcfA* in 21 out of the 22 examined isolates (Fig S1), suggesting a wide and stable distribution of the *tcf* cluster in strains belong to serovar Infantis.

**Tcf contains interchanging alleles of the fimbria adhesin TcfD**

The Tcf fimbrial subunit mediating the interaction with host cell receptor is the tip adhesin encoded by *tcfD*<sup>21, 22</sup>. Interestingly, while the sequence of TcfA, TcfB, TcfC, TinR and TioA is very similar between the *tcf*-positive *Salmonella* serovars, a significant diversity was found in the N-terminus region of TcfD (positions 1-192), while its C-terminus region (positions 193-359) is highly conserved (Fig. 2 and Table S1). Sequence diversity in this protein is likely to indicate differences in receptor binding between TcfD variants, suggesting that Tcf fimbriae from different serovars may mediate attachment to distinct targets.

#### ***tcf* exhibits varying expression levels in different *Salmonella* serovars**

To study the expression profile of *tcf* in NTS in comparison to typhoidal serovars a reverse transcription PCR approach was applied and the level of *tcfA* transcription was determined in two typhoidal serovars (Typhi and Paratyphi A) and in eight clinically-relevant NTS serovars including Schwarzengrund, Infantis, Virchow, Montevideo, Bredeney, 9,12:l,v:-, Heidelberg and Choleraesuis. *tcfA* expression was compared during growth in rich LB broth, and in M63 minimal medium (pH 7) using semi-quantitative reverse transcription PCR as well as quantitative real-time PCR (qRT-PCR). *Salmonella* RNA without a reverse transcriptase treatment (-RT) was used as a negative control and gDNA from *S. Infantis* was used as a positive control for the PCR reactions. As can be seen in Fig. 3, similar results were demonstrated using both semi-quantitative reverse transcription PCR (Fig. 3A and C) and qRT-PCR (Fig. 3 B and D), demonstrating significant variation in *tcfA* expression both between the typhoidal serovars and between the NTS serovars. In *Salmonella* cultures grown in LB to the late logarithmic phase, *tcfA* expression was significantly higher in *S. Typhi* than in *S. Paratyphi A*, and among the NTS serovars, the highest expression was found in *S. Infantis* (Fig. 3A and B). Growth in M63



minimal media, also exhibited inter-serovars diverse transcription of *tcfA*. Similarly to the expression in LB, higher transcription was found in *S. Typhi* than in *S. Paratyphi A*, and among the NTS serovars, the highest expression was found in serovar Choleraesuis. Serovars Infantis, Virchow and Heidelberg presented comparable levels of *tcfA* expression to the one of *S. Typhi*, but the rest of the examined serovars showed significantly lower levels of expression (Fig. 3C and D). Moreover, while *tcfA* expression in *S. Typhi* and *S. Infantis* was 2.5- and 4.5-fold higher in LB than in M63, respectively, *S. Choleraesuis* presented 3-fold elevated *tcfA* expression in minimal medium relative to the expression in LB (Fig. 3E). These results indicate that despite the high conservation in the cluster organization and genes sequence, *tcfA* presents varying expression profiles among different *S. enterica* serovars and that *tcf* may be induced differently at distinct growth or environmental conditions.

#### ***tcfA* is induced in *S. Infantis* under microaerobiosis**

Due to the high energetic cost involved in fimbriae expression, fimbrial genes are often tightly regulated and assembled under specific environmental stimuli<sup>27-29</sup>. To investigate the effect of growth conditions on *tcf* expression, we next focused on *S. Infantis* that showed the highest *tcfA* expression in LB and applied qRT-PCR to quantitatively measure *tcfA* transcripts in *S. Infantis* 119944 cultures that were grown under five different sets of conditions including: (i) aerobic growth to the mid-logarithmic phase in LB broth, (ii) aerobic growth to the late-logarithmic phase in LB, (iii) aerobic growth to the stationary phase in LB, (iv) microaerobic growth to the stationary phase in LB, and (v) aerobic growth to the late-logarithmic phase in N-minimal medium pH 5.8. Markedly, the highest expression of *tcfA* in *S. Infantis* occurred in cultures grown under microaerobic conditions in LB. Under this set of conditions *tcfA*

transcription was about 90-fold higher than under aerobic growth conditions (Fig. 4). These results indicated that *tcf* is induced in nutrients-rich and microaerobic ambient, a set of conditions found in the intestines of warm-blooded animals.

### **Tcf is negatively regulated by Lrp**

To identify *S. Infantis* regulators controlling *tcf* expression we screened ten null mutant strains, lacking regulators that were previously reported to be involved in oxygen homeostasis (OxyR, SoxR, ArcA, ArcB and FNR)<sup>30</sup>; *Salmonella* pathogenicity (RpoS, PhoP, OmpR and Fur)<sup>31</sup>; and the leucine-responsive regulatory protein (Lrp), which was shown to regulate different types of fimbriae in *E. coli*<sup>32, 33</sup> and *Salmonella*<sup>14, 34, 35</sup>. Using qRT-PCR we compared the *tcfA* expression in the wild-type background and the above ten mutant strains. Interestingly, the most prominent regulatory affect was found for Lrp, where in its absence the expression of *tcfA* increased by four to sixteen-fold compared to the *S. Infantis* wild-type background (Fig. 5). Complementing the expression of *lrp* from a low-copy number plasmid, but not the presence of the empty vector (pWSK29) in the *lrp* null strain, resulted in reduced expression of *tcfA* to similar levels as in the wild-type (Fig. 5B).

Promoter sequence analysis of the intergenic region, upstream to *tcfA* in *S. Infantis* identified six putative Lrp binding sites, all containing the consensus sequence GN(2-3)TTT recognized by Lrp<sup>36</sup> (Fig. 5C), suggesting direct binding of Lrp to the *tcfA* promoter. We concluded from these experiments that Lrp is a key regulator, which negatively controls *tcf* expression.

### **Heterologous expression and imaging of Tcf in surrogate *E. coli* cells**

To further characterize the biology of Tcf, the entire cluster (*pcfABCD tinR tioA*) from *S. Infantis* was cloned under an arabinose inducible promoter (pBAD18) and introduced into a nonfimbriated *E. coli* ORN 172 strain. Cultures that were grown in minimal medium to the mid-logarithmic phase in the presence of arabinose (as inducer) or glucose (as suppressant) were subjected to shearing by a shaft homogenizer. Filtered bacterial cell supernatant, enriched with surface exposed structures was precipitated by trichloroacetic acid (TCA) and proteins precipitate was separated on a SDS-PAGE. A single 13 kDa band that appeared specific to the arabinose-induced cultures (Fig 6A) was isolated from the acrylamide gel and analyzed by LC-MS/MS. This analysis successfully identified the major subunit TcfB (Fig. 6B), which is expected to be exported to the periplasm via the Sec system, following a signal peptide cleavage between positions 24 and 25 (Fig. 6C).

To image the Tcf fimbriae, *E. coli* ORN172 cultures harboring the vector only (pBAD18) and cultures expressing *pcfABCD tinR tioA* were grown in N-minimal medium supplemented with 50 mM arabinose. *E. coli* cultures were negatively stained and imaged by transmission electron microscope (TEM). Short and thin pili that were attached to the cell envelope were imaged in the cultures expressing *pcf*, but not in cultures harboring the empty vector (Fig. 7). Noteworthy, the Tcf pili were very different in their appearance from the flagella, which looked as much longer and thicker hair-like structures (Fig. 7A).

### **Tcf is dispensable for biofilm formation, macrophages uptake and host cell invasion by NTS serovars**

Different chaperone-usher fimbriae have been previously reported to be involved in diverse virulence-associated phenotypes including biofilm formation, macrophages uptake, host

cell adhesion and invasion<sup>37-40</sup>. To determine the role of Tcf in NTS virulence-associated phenotypes, a null deletion of *tcfABCD* was constructed in the genome of clinical isolates of serovars Heidelberg, Schwarzengrund and Infantis and the potential role of Tcf in the above phenotypes was subsequently studied.

The ability of serovars Infantis, Schwarzengrund and Heidelberg to form biofilm was tested under biofilm induction conditions (see Materials and Methods). As a control we used *S. Typhimurium* strain lacking flagellin (*flic fljB* double mutant strain) known to be attenuated in biofilm formation<sup>41, 42</sup>. These results showed that biofilm formation *in-vitro* is independent from the presence of Tcf in these three serovars (Fig. 8A). Similarly, invasion studies using HeLa (Fig. 8B) and Caco-2 (Fig. 8C) cells by serovars Infantis, Schwarzengrund and Heidelberg, grown to the late logarithmic phase, showed that Tcf is dispensable for epithelial cells invasion, while *S. Infantis* and *S. Schwarzengrund* containing a null mutation in *invA* (a structural gene of the type three secretion system one) were significantly impaired in HeLa cells invasion (Fig. 8B). Similar results were obtained when cultures were grown under microaerobic conditions (data not shown). Moreover, expression of *tcfABCD tioA tinR* in *E. coli* ORN172 did not affect its ability to adhere (data not shown) or invade (Fig. 8D) HeLa cells. Likewise, *S. Infantis* lacking *tcfABCD* presented similar adherence to HeLa cells as its wild-type background (Fig. 8E).

Along these lines, uptake experiments conducted with Raw 264.7 macrophage-like cells demonstrated that the absence of Tcf in *S. Infantis* does not affect its uptake even in the  $\Delta lrp$  background, which enhances Tcf expression (Fig. 8F). Collectively, we concluded from these experiments that Tcf is not involved in biofilm formation and host cells entry of serovars Infantis, Schwarzengrund and Heidelberg *in-vitro*.

**Tcf is involved in *S. Infantis* colonization in the colitis mouse model**

To determine the role of Tcf in NTS virulence, the murine colitis model has been applied. Mice that were pretreated with streptomycin were infected orally with equal amount of *S. Heidelberg*, *S. Schwarzengrund* and *S. Infantis* wild-type background and their isogenic *tcf* mutant strains marked with kanamycin and ampicillin, respectively. Four days post infection (p.i.) mice were sacrificed and the ability of the *tcf* mutant to colonize the cecum, colon, and spleen in relation to the wild-type background (competitive index, C.I.) was calculated for each serovar. Competitive infections with two wild-type strains harboring the ampicillin and the kanamycin resistant cassettes in *S. Infantis* (Fig. 9A) and *S. Schwarzengrund* (data not shown) backgrounds showed equal colonization (C.I. =1) in this mouse model, indicating similar fitness of both marked strains. While high bacterial loads were recovered from the gastrointestinal tract organs ( $10^5$ - $10^6$  CFU/ organ), all three serovars were isolated at very low numbers from the mouse spleen (data not shown), indicating the lack of systemic infection by serovars Heidelberg, Schwarzengrund and Infantis following oral infection in this model.

Interestingly, while *tcf* was found unessential for intestinal colonization by *S. Schwarzengrund* (Fig. 9B) or *S. Heidelberg* (Fig. 9C), in *S. Infantis*, a *tcf* mutant was significantly impaired in colonizing the mouse cecum and colon (Fig. 9D). These results show that Tcf contributes differently to the virulence of distinct NTS serovars.

One possible explanation for the differences in the role of Tcf between these serovars is the possibility that Tcf is expressed at different levels among NTS serovars, during mouse infection. To examine this notion, three groups of streptomycin pretreated C57BL/6 mice were infected with  $2-3 \times 10^7$  CFU of *S. Heidelberg*, *S. Schwarzengrund* and *S. Infantis*. One day p.i.

mice were sacrificed and the *tcfA* transcription was determined in the cecum and colon of these infected mice. qRT-PCR analysis showed 2.5-fold higher expression of *S. Infantis tcfA* than its expression in mice infected with *S. Heidelberg* and *S. Schwarzengrund* (Fig. 10). These results are in close agreement with the observed difference of *tcfA* expression *in-vitro* (Fig. 3) and demonstrate that *S. Infantis* exhibits higher *tcfA* expression than *S. Heidelberg* and *S. Schwarzengrund*, both *in-vitro* and during mouse infection. Thus, these results may provide possible explanation for the different contribution of the Tcf fimbria in these serovars to *Salmonella* colonization in the mouse.

## DISCUSSION

*Salmonella* adhesion to host cells is an essential step in colonization, persistence and pathogenicity, which is often facilitated by various fimbrial adhesins. Since *Salmonella* strains harbor multiple and distinct fimbria clusters, the unique composition of specific fimbriae and other adhesion factors has been long suggested to contribute to host tropism<sup>17, 27</sup>. One example is the Tcf that was previously suggested to play a role in the host specificity of typhoidal serovars<sup>17, 20, 21</sup>. Nevertheless, current bioinformatics analysis, exploiting the continuously expanding whole genome sequencing database revealed that the distribution of the *tcf* cluster is actually not limited to the genomes of typhoidal serovars and demonstrated its somewhat broad presence in at least 25 NTS serovars. Therefore, this study provides evidence against the formerly suggested role of Tcf in typhoid host-specificity and point to a more complex part that this fimbria may play in *Salmonella* pathogenicity. The genetic organization of *tcfABCD tinR tioA* was found to be highly conserved in both typhoidal and NTS serovars, inserted between *sinR* and *pagN* within SPI-6. Interestingly, while TcfA, TcfB and TcfC showed high sequence conservation among the

different serovars, TcfD, the adhesin of the Tcf fimbria displayed an extensive sequence diversity in the N-terminal half of the protein. These observations suggest differences in the binding properties of Tcf variants, in a way that might affect host specificity and pathogenesis of each serovar individually. Such diversity between the Tcf variants is expected to provide genetic platform for functional flexibility and further fine tuning host specificity and tissue tropism by NTS serovars. The mechanism(s) responsible for diversification of TcfD only, but not the other fimbrial subunits is not known and presents a fascinating open question. Sequence diversity in the adhesin subunit of other fimbriae was previously shown for variants of the K88 fimbria in enterotoxigenic *E. coli* (ETEC)<sup>43</sup>, FimH<sup>16, 18</sup> and the Klf fimbria, encoded on the *S. Infantis* pESI plasmid<sup>14</sup>. Together, these observations can facilitate further studies that will shed light on the role and mechanisms of fimbrial adhesins diversification and the contribution of such microevolutionary adaptation for the lifestyle, host-range and pathogenesis of *Salmonella* and other related pathogens.

Inducible expression of the entire cluster (*tcfABCD tinR tioA*) in a non-fimbriated *E. coli* strain resulted in the visible expression of thin and short pili that were structurally distinct from the flagella. Moreover, mass spectrometry approach successfully identified TcfB export, indicating that the *tcf* genes cluster is an autonomous functional and structural unit.

Like many other fimbriae, the Tcf of *S. Infantis* is not constitutively expressed and was found to be specifically induced in rich-nutrient medium and at microaerobiosis, characterizing the intestinal conditions of warm-blooded hosts and suggesting a possible role for Tcf during intestinal colonization. These results are in agreement with a previous study that showed

upregulation of *tcfA* transcription in *S. Typhi* grown in high salt LB without agitation (microaerobic conditions) <sup>44</sup>, but are conflicting with a more recent study that reported the induction of a plasmid-born *tcfA::lacZ* fusion in *S. Typhi* grown in minimal medium (M63) and in the absence of the iron regulator Fur <sup>21</sup>. Such discrepancies may be due to differences in the strains used and methodological approach (RT-PCR compared with *lacZ* reporter gene fusion). Induction of fimbria expression under oxygen-limitation was previously reported for several other fimbriae including the MR/P fimbria of uropathogenic *Proteus mirabilis*, type 1 fimbriae of uropathogenic *E. coli* <sup>45</sup> as well as the Klf and IpF pESI encoded fimbriae <sup>14</sup>, suggesting that microaerobiosis is used by different pathogens to sense the environment and regulate fimbriae expression accordingly.

Genetic screen of 10 relevant *S. Infantis* regulators established that *tcf* is negatively regulated by the global regulator Lrp. This regulator is known to function as a transcriptional repressor or activator, controlling the expression of numerous operons in *E. coli* and *Salmonella* <sup>46</sup>. In addition to operons involved in amino acid metabolism, Lrp was previously shown to regulate different fimbrial operons including *pap* (P pilus) and *fan* (K99) <sup>32</sup>, *fim* <sup>35, 47</sup>, *sfa*, *daa* <sup>48</sup>, *klf* <sup>14</sup>, as well as the nonfimbrial adhesin TosA in uropathogenic *E. coli* <sup>33</sup>. The presence of six putative Lrp binding sites in the *tcfA* promoter region strongly suggests that Lrp repression is direct. Similar results have been previously demonstrated for the major type 1 fimbrial subunit gene, *fimA* in *S. Typhimurium* <sup>34</sup>. Transcriptional repression of other genes including the *papBA* operon by Lrp have been shown to occur by cooperative interactions between Lrp and the nucleoid-binding proteins H-NS <sup>48</sup>. The regulation of *tcfA* by an ancestral regulator such as Lrp, demonstrates the assimilation of the horizontally acquired *tcf* cluster into the core regulatory



setup of *Salmonella* and emphasize the regulatory linkage between ancestral metabolic pathways and acquired virulence traits in *Salmonella*.

The role of Tcf in NTS serovars interactions with non-phagocytic host cells was also studied. Adhesion and invasion assays using HeLa and Caco-2 human epithelial cells, failed to exhibit significant difference between the wild-type and an isogenic *tcf* mutant strain of serovars Infantis, Heidelberg and Schwarzengrund. These results are in agreement with previous studies that were unable to demonstrate a role for Tcf in *S. Typhi in-vitro*<sup>21, 44</sup>. Accumulatively, these data suggest that a yet unknown host receptor(s) interacts with Tcf is not expressed on these cell lines or that Tcf is required for a different type of host-pathogen interaction, taking place in the host, but not *in-vitro*.

Consistent with these possibilities, *in-vivo* studies using the streptomycin pretreated mouse model have shown a role for Tcf in *S. Infantis* colonization in the cecum and colon. Nevertheless, Tcf was not found to be involved in colonization of *S. Heidelberg* nor *S. Schwarzengrund* in this infection model. Differences in the contribution of Tcf to mouse colonization may be due to variation of *tcf* expression between these serovars. Indeed, during both *in-vitro* growth and mouse infection, *tcfA* transcription was found to be significantly higher in *S. Infantis* than in serovars Heidelberg and Schwarzengrund. Since both *S. Infantis* and *S. Heidelberg* harbor the same TcfD variant (see Fig. 2), it is unlikely that the different contribution of Tcf to mouse colonization is due to sequences differences in TcfD. Nevertheless, differences in *tcf* expression and the extensive variation demonstrated in the sequence of TcfD in other serovars are expected to affect the role of Tcf in the pathogenicity of distinct *Salmonella* serovars

or strains. To the best of our knowledge, this is the first study demonstrating variation in the sequence of TcfD and an interchanging contribution of Tcf to an *in-vivo* infection.

Besides Tcf, *S. Infantis* (strain 119944) encodes 13 additional Chaperone-usher fimbria (Saf, F17, Lpf, Fim, Stf, Sth, Sti, Sfm, Stc, Stb, Std, Ipf and Klf) and at least six non-fimbrial adhesins (MisL, SiiE, ShdA, BapA, RatB, and CsgA). The presence of multiple adhesin factors and their possibly redundant function, may explain the relatively subtle phenotype seen in the absence of Tcf in *S. Infantis*.

In summary, we show that the Typhi colonization factor is not a unique typhoid virulence determinant, but instead present and expressed in multiple non-typhoidal *Salmonella* serovars. In *S. Infantis*, *tcf* is induced in rich broth, under microaerobiosis and is negatively regulated by Lrp. The *tcf* cluster (*tcfABCD tinR tioA*) is well conserved and exhibits the same genetic organization as in *S. Typhi* and *S. Paratyphi A*. Intriguing sequence variation was found in TcfD, comprising the adhesin of the Tcf fimbria and in the expression profile of *tcfA* in different serovars. In agreement with these findings, we further demonstrated diverse role in virulence for Tcf in three NTS serovars, using the mouse colitis model. Our results suggest that Tcf confers functional flexibility and that Tcf may contribute differently to the host specificity and pathogenesis of distinct *Salmonella* serovars.

## MATERIALS AND METHODS

**Bacterial strains, media and growth conditions.** Bacterial strains utilized in this study are listed in Table S2. Liquid bacterial cultures were routinely maintained in Luria-Bertani (LB) broth (Lennox, BD Difco), N-minimal medium [80 mM MES (for pH 5.8 or 100 mM Tris-HCl (for pH 7), 5 mM KCl, 7.5 mM (NH<sub>4</sub>)SO<sub>4</sub>, 0.5 mM K<sub>2</sub>SO<sub>4</sub>, 337 mM K<sub>2</sub>HPO<sub>4</sub>/KH<sub>2</sub>PO<sub>4</sub>, 24 μM MgCl<sub>2</sub>, 38 mM glycerol, and 0.1% casamino acids] or M63 minimal medium (pH 7) [100 mM KH<sub>2</sub>PO<sub>4</sub>, 15 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 1.7 μM FeSO<sub>4</sub>·7H<sub>2</sub>O, 0.2% glucose, 1 mM MgSO<sub>4</sub>, 21 μg/ml histidine, 40 μg/ml tryptophan, 40 μg/ml cysteine and 0.1% casamino acids] as indicated. *Salmonella* and *E. coli* strains were plated onto LB or xylose lysine deoxycholate (XLD; BD Difco) agar plates. When appropriate, antibiotics were added to the medium as follows: kanamycin (50 μg/ml), ampicillin (100 μg/ml), tetracycline (20 μg/ml) and chloramphenicol (25 μg/ml).

**Bioinformatics.** Blastn search at NCBI against the nr database was used to identify *tcf* homologs. The *tcf* locus (12069 bp long) from *S. Typhi* (strain CT18) was compared using the genome comparison visualizer, Easyfig tool<sup>49</sup> to the corresponding regions in other serovars. Multiple sequence alignment of the TcfD adhesin homologs was performed using CLUSTALW and BOXSHADE 3.2 tools. Identification of the signal peptide in TcfB was done using the SignalP 4.1 Server<sup>50</sup>. Promoter location including the -10 and -35 boxes was predicted by BPPROM

(<http://www.softberry.com/berry.phtml?topic=bprom&group=programs&subgroup=gfindb>).

**Molecular biology and cloning.** All primers used in this study are listed in Table S3. Oligonucleotides were purchased from IDT and PCR was carried out using Phusion Hot Start

Flex DNA Polymerase (New England BioLabs) or with Red Load Taq Master (LAROVA). All *Salmonella* null mutants were constructed using the  $\lambda$ -red-recombination system and a three step PCR method to produce an amplicon containing the antibiotic resistance gene, as described in Serra-Moreno *et al.*<sup>51</sup>. Resistant cassette was then eliminated from the genome by using a helper plasmid encoding the FLP recombinase<sup>52</sup>. *lrp* was amplified using primers 'lrP SacI Fw' and 'lrp XbaI Rv', digested with SacI and XbaI and cloned into pWSK29. The pWSK29::*lrp* was transformed into *S. Infantis lrp* null mutant strain. For expression of *tcf* under arabinose inducible promoter, a PCR fragment containing *tcfABCD*, *tinR* and *tioA* from *S. Infantis* was amplified using the primers 'XbaI Rv tcf tin tio pBAD18' and 'SacI Fw 1ATG tcf pBAD18' and digested with SacI and XbaI. The digested fragment was cloned into pBAD18. Empty pBAD18 or pBAD18::*tcfABCD tinR tioA* were transformed into a fimbriae-less *E. coli* ORN172 strain. The *tcf* cluster (*tcfABCD tinR tioA*) and its regulatory region from *S. Schwarzengrund* was amplified using the primers 'CloneTcf Fw (Schwarzengrund) ' and 'CloneTcf Rev (Schwarzengrund) ' digested with XbaI and SacI and cloned into pWSK29 and transformed into *E. coli* ORN172. *tcfABCD* and its upstream regulatory region from *S. Infantis* was amplified using the primers 'CloneTcf Fw (Infantis) ' and 'CloneTcf Rev (Infantis) ' digested with XbaI and SacI and cloned into pWSK29 and transformed into *S. Infantis tcf* null mutant.

**Reverse transcription PCR.** The cecum and colon of infected mice were isolated immediately freeze in liquid nitrogen and stored at -80°C. Frozen tissues were homogenized in 500  $\mu$ l saline and 1200  $\mu$ l of RNA protect bacterial reagent (Qiagen) were added to homogenates followed by RNA extraction using RNeasy mini kit (Qiagen). RNA was extracted from *Salmonella* cultures grown under different conditions using the Qiagen RNA protect bacterial

reagent and the RNeasy mini kit (Qiagen) according to the manufacturer's instructions, including an on-column DNase I digest. Purified RNA was secondarily treated with an RNase-free DNase I followed by ethanol precipitation. DNase I-treated RNA (1  $\mu$ g from mouse tissues or 200 ng from bacterial cultures) was subjected to cDNA synthesis using the iScript cDNA synthesis kit (Bio-Rad Laboratories). Real-time PCR and data analysis were performed as previously described<sup>53</sup> on a StepOnePlus Real-Time PCR System (Applied Biosystems). The 16S rRNA gene was used as the endogenous normalization controls. Fold-differences in gene expression were calculated as  $2^{-\Delta\Delta C_t}$ .

**Tissue cultures.** Host cell infections were done as previously specified<sup>54</sup>. Briefly, Raw 264.7 macrophage-like cells, the human colonic adenocarcinoma Caco-2 and human epithelial HeLa cells were purchased from the American Type Culture Collection (ATCC) and cultured at 37°C in a humidified atmosphere with 5% CO<sub>2</sub>. Caco-2 cell line was grown in Dulbecco's modified Eagle medium (DMEM)–F-12 medium (Biological Industries) supplemented with 20% fetal bovine serum (FBS) and 2 mM L-glutamine. Raw 264.7 and HeLa cells were maintained in a high-glucose (4.5 g/liter) DMEM (Biological Industries) supplemented with 10% heat-inactivated FBS, 1 mM pyruvate and 2 mM L-glutamine. Eighteen hours prior to bacterial infection, epithelial cells were seeded in a 24-well tissue culture dish at  $5 \times 10^4$  cells/ml and were infected at a multiplicity of infection (MOI) of  $\sim 1:50$  with late logarithmic phase *Salmonella* cultures using the gentamicin protection assay as was previously described<sup>53</sup>. Raw 264.7 cells were seeded at  $2.5 \times 10^5$  cells/ml and were infected at MOI of  $\sim 1:10$  with stationary phase microaerobically grown cultures. Adhesion was determined using cytochalasin D, which inhibits actin-dependent bacterial invasion. Cells were incubated with fresh medium containing 1

µg/ml cytochalasin D 1 h before infection. Bacteria were added and allowed to adhere for 30 min in the presence of 1 µg/ml cytochalasin D. At 30 min (for host cells adhesion) or 2 h (for invasion and uptake) post infection (p.i.), cells were washed three times with phosphate-buffered saline (PBS) and extracted with lysis buffer (containing 0.1% SDS, 1% Triton X-100 in PBS). Serial dilutions of the infected cell lysates were plated onto LB agar plates and incubated at 37°C for bacterial enumeration. *Salmonella* invasion was calculated by the number of intracellular *Salmonella* CFUs divided by the infecting inoculum.

**Biofilm formation.** Overnight cultures grown in LB were diluted 1:100 into fresh LB medium without NaCl (containing 10 g/liter peptone and 5 g/liter yeast extract) and 150 µl was added to cell culture-treated 96-well microplates (Greiner Bio-one). The plates were incubated at 28°C for 96 h. Planktonic cells were discarded, and plastic-adhered cells were fixed for 2 h at 60°C. Fixed bacteria were stained with 150 µl of 0.1% crystal violet for 10 min at room temperature. The plates were washed with phosphate-buffered saline (PBS), and the dye bound to the adherent bacteria was resuspended in 150 µl of 33% acetic acid followed by optical density measurement at 560 nm.

**Competitive index (C.I.) infections.** Eight week old female C57BL/6 mice (Envigo, Israel) were pretreated with streptomycin (20 mg per mouse in 100 µl saline) 24 h prior to infection. Mice were infected with  $10^6$ - $10^7$  CFU of a mixed (1:1) inoculum containing the wild-type *S. Infantis*, *S. Schwarzengrund* or *S. Heidelberg* (harboring pWSK129; Km<sup>R</sup>) and their *tcf* corresponding null mutant strains (harboring pWSK29; Amp<sup>R</sup>). A mixed inoculum of two *S. Infantis* wild-type strains carrying pWSK29 or pWSK129 was used as a control and resulted in a C.I. of 1. At four days post infection the mice were sacrificed and the tested tissues were

homogenized in 700  $\mu$ l of cold PBS. Serial dilutions of the homogenates were plated on XLD agar plates supplemented with ampicillin or kanamycin. CFUs were counted and the competitive index was calculated as  $[\text{mutant/wild-type}]_{\text{output}}/[\text{mutant/wild-type}]_{\text{input}}$ .

**Heterologous expression of the fimbria and mass spectrometry.** A non-fimbriated *E. coli* (ORN172) strain carrying the *tcf* cluster (pBAD18::*tcfABCD tinR tioA*), or the empty vector (pBAD18), which was used as a negative control) were grown aerobically overnight in LB supplemented with ampicillin at 37°C. The next day, the cultures were washed twice with N-minimal medium and diluted 1:50 into N-minimal medium pH 7 containing ampicillin (100  $\mu$ g/ml), L-arabinose (50 mM) or glucose (1M) and grown for 4 h until reaching OD<sub>600</sub> of ~0.5. OD<sub>600</sub>-normalized cultures were centrifuged and resuspended in 2 ml phosphate-buffered saline (PBS). Surface exposed fimbria were separated from the cells by mechanical shearing using a shaft blender (three cycles of 1 min each). Cellular debris were removed by centrifugation (13,000 rpm, 5 min at 4°C) and the supernatant was collected and filtered using a 0.22- $\mu$ m filter (Merck Millipore). The filtered supernatant was then precipitated in 10% Trichloroacetic acid (TCA) for overnight on ice. Precipitated fractions were recovered by centrifugation (13,000 rpm, 45 min at 4°C) and the pellet was washed with 0.8 ml of ice-cold acetone. After acetone was removed, the pellet was air-dried for 10 min at room temperature in a fume hood and resuspended in 20  $\mu$ l of 1 $\times$  SDS-PAGE sample buffer. The boiled samples were separated on 15 % SDS-PAGE followed by Coomassie blue staining. A 13 kDa specific band that was absent from the negative control (pBAD18) was cut from the gel and subjected to mass spectrometry analysis at the Smoler Proteomic Center at the Technion, Haifa, Israel. The samples were digested by trypsin, analyzed by LC-MS/MS on LTQ-Orbitrap (Thermo) and identified by

Discoverer 1.4 software using two algorithms: Sequest (Thermo) and Mascot (Matrix science) against the Tcf subunits sequences and the *E. coli* proteome from the Uniprot database and a decoy database, in order to determine the false discovery rate (FDR). High confidence peptides have passed the 1% FDR threshold.

**Transmission electron microscope (TEM).** A non-fimbriated *E. coli* (ORN172) strain carrying the *tcfABCD tinR* and *tioA* operon (pBAD18::*tcf-tinR tioA*), or the empty vector (pBAD18) were grown aerobically in LB supplemented with ampicillin at 37°C for 5h. Subsequently, the cultures were washed twice with N-minimal medium and diluted 1:100 into N-minimal medium pH 7 containing ampicillin (100 µg/ml) and L-arabinose (50 mM), and grown for overnight. Cells were centrifuged (10,000 rpm, 2 min) and re-suspended in PBS. A drop of 50 µl of bacterial culture was absorbed onto 200-mesh Formvar/carbon-coated copper grids for 3 min and stained with 1% uranyl acetate. Images were taken using FEI F20 Philips-Tecnaï field emission gun TEM operated at an accelerating voltage of 80 kV and a beam current of ~1 nA. Images were acquired with One View CMOS camera (Gatan).

**Statistics.** Statistical analysis was performed using the GraphPad Prism 5 software package (GraphPad Software, Inc.). Analysis of variance (ANOVA) with Dunnett's multiple comparison test was used to determine differences between multiple data sets. A student t-Test against a theoretical mean of 1.0 was used to determine statistical significance of the C.I values. *P*-value smaller than 0.05 was considered statistically significant and was indicated in the figures as follow: \*, *P* <0.05; \*\*, *P* <0.01; \*\*\*, *P* <0.0001; ns, not significant. Error bars show the standard error of the mean.



**Ethics.** Mice experiments were conducted according to the ethical requirements of the Animal Care Committee of the Sheba Medical Center (Approval number 933/14) and in line with the national guidelines.

## FUNDING INFORMATION

This work was supported by grant number 999/14 from the Israel Science Foundation (ISF) and by grant number 3-12435 from Infect-Era /Chief Scientist Ministry of Health awarded to OGM. The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

## ACKNOWLEDGMENTS

We thank to Dr. Tamar Ziv from the Smoler Proteomics Center at the Technion - Israel Institute of Technology, Haifa Israel for her valuable help with the proteomic analysis and for Dr. George Levi from the Wolfson Applied Materials Research Centre at Tel Aviv University for his appreciated help with the TEM.

## REFERENCES

1. Guibourdenche M, Roggentin P, Mikoleit M, Fields PI, Bockemuhl J, Grimont PA, et al. Supplement 2003-2007 (No. 47) to the White-Kauffmann-Le Minor scheme. *Res Microbiol* 2010; 161:26-9.
2. Gal-Mor O, Boyle EC, Grassl GA. Same species, different diseases: how and why typhoidal and non-typhoidal *Salmonella enterica* serovars differ. *Frontiers in microbiology* 2014; 5:391.
3. House D, Bishop A, Parry C, Dougan G, Wain J. Typhoid fever: pathogenesis and disease. *Current opinion in infectious diseases* 2001; 14:573-8.
4. Baumler A, Fang FC. Host specificity of bacterial pathogens. *Cold Spring Harbor perspectives in medicine* 2013; 3.
5. Kingsley RA, Kay S, Connor T, Barquist L, Sait L, Holt KE, et al. Genome and transcriptome adaptation accompanying emergence of the definitive type 2 host-restricted *Salmonella enterica* serovar Typhimurium pathovar. *mBio* 2013; 4:e00565-13.
6. Uzzau S, Brown DJ, Wallis T, Rubino S, Leori G, Bernard S, et al. Host adapted serotypes of *Salmonella enterica*. *Epidemiology and infection* 2000; 125:229-55.
7. Edwards RA, Puente JL. Fimbrial expression in enteric bacteria: a critical step in intestinal pathogenesis. *Trends in microbiology* 1998; 6:282-7.
8. Humphries AD, Townsend SM, Kingsley RA, Nicholson TL, Tsolis RM, Baumler AJ. Role of fimbriae as antigens and intestinal colonization factors of *Salmonella* serovars. *FEMS microbiology letters* 2001; 201:121-5.

9. van der Velden AW, Baumler AJ, Tsolis RM, Heffron F. Multiple fimbrial adhesins are required for full virulence of *Salmonella typhimurium* in mice. *Infection and immunity* 1998; 66:2803-8.
10. Waksman G, Hultgren SJ. Structural biology of the chaperone-usher pathway of pilus biogenesis. *Nature reviews Microbiology* 2009; 7:765-74.
11. Jin LZ, Zhao X. Intestinal receptors for adhesive fimbriae of enterotoxigenic *Escherichia coli* (ETEC) K88 in swine--a review. *Applied microbiology and biotechnology* 2000; 54:311-8.
12. Yue M, Rankin SC, Blanchet RT, Nulton JD, Edwards RA, Schifferli DM. Diversification of the *Salmonella* fimbriae: a model of macro- and microevolution. *PloS one* 2012; 7:e38596.
13. Zav'yalov V, Zavialov A, Zav'yalova G, Korpela T. Adhesive organelles of Gram-negative pathogens assembled with the classical chaperone/usher machinery: structure and function from a clinical standpoint. *FEMS microbiology reviews* 2010; 34:317-78.
14. Aviv G, Elpers L, Mikhlin S, Cohen H, Vitman Zilber S, Grassl GA, et al. The plasmid-encoded Ipf and Klf fimbriae display different expression and varying roles in the virulence of *Salmonella enterica* serovar *Infantis* in mouse vs. avian hosts. *PLoS pathogens* 2017; 13:e1006559.
15. Boddicker JD, Ledebor NA, Jagnow J, Jones BD, Clegg S. Differential binding to and biofilm formation on, HEp-2 cells by *Salmonella enterica* serovar *Typhimurium* is dependent upon allelic variation in the *fimH* gene of the *fim* gene cluster. *Molecular microbiology* 2002; 45:1255-65.

16. Kisiela D, Sapeta A, Kuczkowski M, Stefaniak T, Wieliczko A, Ugorski M. Characterization of FimH adhesins expressed by *Salmonella enterica* serovar Gallinarum biovars Gallinarum and Pullorum: reconstitution of mannose-binding properties by single amino acid substitution. *Infection and immunity* 2005; 73:6187-90.
17. Townsend SM, Kramer NE, Edwards R, Baker S, Hamlin N, Simmonds M, et al. *Salmonella enterica* serovar Typhi possesses a unique repertoire of fimbrial gene sequences. *Infection and immunity* 2001; 69:2894-901.
18. Yue M, Han X, De Masi L, Zhu C, Ma X, Zhang J, et al. Allelic variation contributes to bacterial host specificity. *Nature communications* 2015; 6:8754.
19. Yue M, Schifferli DM. Allelic variation in *Salmonella*: an underappreciated driver of adaptation and virulence. *Frontiers in microbiology* 2014; 4:419.
20. Folkesson A, Advani A, Sukupolvi S, Pfeifer JD, Normark S, Lofdahl S. Multiple insertions of fimbrial operons correlate with the evolution of *Salmonella* serovars responsible for human disease. *Molecular microbiology* 1999; 33:612-22.
21. Leclerc JM, Quevillon EL, Houde Y, Paranjape K, Dozois CM, Daigle F. Regulation and production of Tcf, a cable-like fimbriae from *Salmonella enterica* serovar Typhi. *Microbiology* 2016; 162:777-88.
22. Nuccio SP, Baumler AJ. Evolution of the chaperone/usher assembly pathway: fimbrial classification goes Greek. *Microbiology and molecular biology reviews* : MMBR 2007; 71:551-75.

23. den Bakker HC, Moreno Switt AI, Govoni G, Cummings CA, Ranieri ML, Degoricija L, et al. Genome sequencing reveals diversification of virulence factor content and possible host adaptation in distinct subpopulations of *Salmonella enterica*. *BMC genomics* 2011; 12:425.
24. Jotham Suez SP, Amir Dagan, Alex Marzel, Yosef Ilan Schorr, Prerak T. Desai,, Vered Agmon MM, Galia Rahav, Ohad Gal-Mor. Virulence Gene Profiling and Pathogenicity Characterization of Non-Typhoidal Salmonella Accounted for Invasive Disease in Humans. *PloS one* 2013; 8:e58449. doi:10.1371/journal.pone.0058449.
25. Bronowski C, Winstanley C. Identification and distribution of accessory genome DNA sequences from an invasive African isolate of *Salmonella Heidelberg*. *FEMS microbiology letters* 2009; 298:29-36.
26. Timme RE, Pettengill JB, Allard MW, Strain E, Barrangou R, Wehnes C, et al. Phylogenetic diversity of the enteric pathogen *Salmonella enterica* subsp. *enterica* inferred from genome-wide reference-free SNP characters. *Genome biology and evolution* 2013; 5:2109-23.
27. Clegg S, Wilson J, Johnson J. More than one way to control hair growth: regulatory mechanisms in enterobacteria that affect fimbriae assembled by the chaperone/usher pathway. *Journal of bacteriology* 2011; 193:2081-8.
28. Humphries AD, Raffatellu M, Winter S, Weening EH, Kingsley RA, Droleskey R, et al. The use of flow cytometry to detect expression of subunits encoded by 11 *Salmonella enterica* serotype Typhimurium fimbrial operons. *Molecular microbiology* 2003; 48:1357-76.

29. Low AS, Holden N, Rosser T, Roe AJ, Constantinidou C, Hobman JL, et al. Analysis of fimbrial gene clusters and their expression in enterohaemorrhagic *Escherichia coli* O157:H7. *Environmental microbiology* 2006; 8:1033-47.
30. Bauer CE, Elsen S, Bird TH. Mechanisms for redox control of gene expression. *Annual review of microbiology* 1999; 53:495-523.
31. Clements M, Eriksson S, Tezcan-Merdol D, Hinton JC, Rhen M. Virulence gene regulation in *Salmonella enterica*. *Annals of medicine* 2001; 33:178-85.
32. Braaten BA, Platko JV, van der Woude MW, Simons BH, de Graaf FK, Calvo JM, et al. Leucine-responsive regulatory protein controls the expression of both the pap and fan pili operons in *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America* 1992; 89:4250-4.
33. Engstrom MD, Mobley HL. Regulation of Expression of Uropathogenic *Escherichia coli* Nonfimbrial Adhesin TosA by PapB Homolog TosR in Conjunction with H-NS and Lrp. *Infection and immunity* 2016; 84:811-21.
34. Baek CH, Kang HY, Roland KL, Curtiss R, 3rd. Lrp acts as both a positive and negative regulator for type 1 fimbriae production in *Salmonella enterica* serovar Typhimurium. *PloS one* 2011; 6:e26896.
35. McFarland KA, Lucchini S, Hinton JC, Dorman CJ. The leucine-responsive regulatory protein, Lrp, activates transcription of the fim operon in *Salmonella enterica* serovar typhimurium via the fimZ regulatory gene. *Journal of bacteriology* 2008; 190:602-12.

36. Peterson SN, Dahlquist FW, Reich NO. The role of high affinity non-specific DNA binding by Lrp in transcriptional regulation and DNA organization. *Journal of molecular biology* 2007; 369:1307-17.
37. Berrocal L, Fuentes JA, Trombert AN, Jofre MR, Villagra NA, Valenzuela LM, et al. stg fimbrial operon from *S. Typhi* STH2370 contributes to association and cell disruption of epithelial and macrophage-like cells. *Biological research* 2015; 48:34.
38. Khater F, Balestrino D, Charbonnel N, Dufayard JF, Brisse S, Forestier C. In silico analysis of usher encoding genes in *Klebsiella pneumoniae* and characterization of their role in adhesion and colonization. *PloS one* 2015; 10:e0116215.
39. Schwartz DJ, Kalas V, Pinkner JS, Chen SL, Spaulding CN, Dodson KW, et al. Positively selected FimH residues enhance virulence during urinary tract infection by altering FimH conformation. *Proceedings of the National Academy of Sciences of the United States of America* 2013; 110:15530-7.
40. Ulett GC, Mabbett AN, Fung KC, Webb RI, Schembri MA. The role of F9 fimbriae of uropathogenic *Escherichia coli* in biofilm formation. *Microbiology* 2007; 153:2321-31.
41. Azriel S, Goren A, Rahav G, Gal-Mor O. The Stringent Response Regulator DksA Is Required for *Salmonella enterica* Serovar Typhimurium Growth in Minimal Medium, Motility, Biofilm Formation, and Intestinal Colonization. *Infection and immunity* 2016; 84:375-84.
42. Crawford RW, Reeve KE, Gunn JS. Flagellated but not hyperfimbriated *Salmonella enterica* serovar Typhimurium attaches to and forms biofilms on cholesterol-coated surfaces. *Journal of bacteriology* 2010; 192:2981-90.

43. Zhang W, Zhao M, Ruesch L, Omot A, Francis D. Prevalence of virulence genes in *Escherichia coli* strains recently isolated from young pigs with diarrhea in the US. *Veterinary microbiology* 2007; 123:145-52.
44. Bishop A, House D, Perkins T, Baker S, Kingsley RA, Dougan G. Interaction of *Salmonella enterica* serovar Typhi with cultured epithelial cells: roles of surface structures in adhesion and invasion. *Microbiology* 2008; 154:1914-26.
45. Lane MC, Li X, Pearson MM, Simms AN, Mobley HL. Oxygen-limiting conditions enrich for fimbriate cells of uropathogenic *Proteus mirabilis* and *Escherichia coli*. *Journal of bacteriology* 2009; 191:1382-92.
46. Brinkman AB, Ettema TJ, de Vos WM, van der Oost J. The Lrp family of transcriptional regulators. *Molecular microbiology* 2003; 48:287-94.
47. Blomfield IC, Calie PJ, Eberhardt KJ, McClain MS, Eisenstein BI. Lrp stimulates phase variation of type 1 fimbriation in *Escherichia coli* K-12. *Journal of bacteriology* 1993; 175:27-36.
48. van der Woude MW, Low DA. Leucine-responsive regulatory protein and deoxyadenosine methylase control the phase variation and expression of the *sfa* and *daa* pili operons in *Escherichia coli*. *Molecular microbiology* 1994; 11:605-18.
49. Sullivan MJ, Petty NK, Beatson SA. Easyfig: a genome comparison visualizer. *Bioinformatics* 2011; 27:1009-10.
50. Petersen TN, Brunak S, von Heijne G, Nielsen H. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nature methods* 2011; 8:785-6.



51. Serra-Moreno R, Acosta S, Hernalsteens JP, Jofre J, Muniesa M. Use of the lambda Red recombinase system to produce recombinant prophages carrying antibiotic resistance genes. *BMC molecular biology* 2006; 7:31.
52. Datsenko KA, Wanner BL. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proceedings of the National Academy of Sciences of the United States of America* 2000; 97:6640-5.
53. Elhadad D, Desai P, Rahav G, McClelland M, Gal-Mor O. Flagellin Is Required for Host Cell Invasion and Normal *Salmonella* Pathogenicity Island 1 Expression by *Salmonella enterica* Serovar Paratyphi A. *Infection and immunity* 2015; 83:3355-68.
54. Gal-Mor O, Valdez Y, Finlay BB. The temperature-sensing protein TlpA is repressed by PhoP and dispensable for virulence of *Salmonella enterica* serovar Typhimurium in mice. *Microbes and infection / Institut Pasteur* 2006; 8:2154-62.

## FIGURE LEGENDS

**Figure 1. The genetic organization of the *tcf* cluster is conserved among typhoidal and NTS serovars.** The *sinR* trough *pagN* locus (12069 bp long) from *S. Typhi* (strain CT18) was compared using the Easyfig tool with the corresponding regions in the genomes of *S. Typhimurium* (LT2), *S. Paratyphi A* (AKU12601), *S. Choleraesuis* (SC-B67S)\*, *S. Infantis* (119944), *S. Schwarzengrund* (CVM19633), *S. Montevideo* (507440-20), *S. 9,12:1,v-* (94293), *S. Virchow* (SVQ1), and *S. Heidelberg* (SL476). Pairwise comparison between the related regions is illustrated by the shades of grey that indicate the degree of homology (in %). Gaps in the grey areas point to the lack of sequence similarity. \* This particular *S. Choleraesuis* strain contains a frame shift in *tcfD* sequence leading to a formation of a stop codon in the middle of the gene.

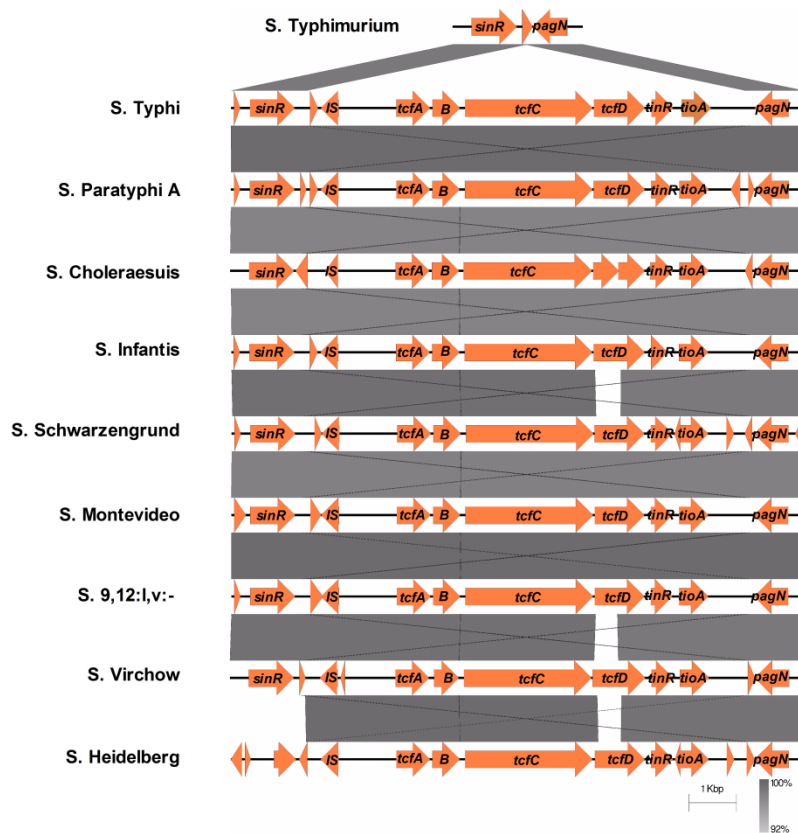


Fig. 1

**Figure 2. The N-terminal half of TcfD is highly diverse.** Multiple sequence alignment of the 359 amino acid sequence of TcfD (the fimbrial adhesin) from *S. Typhi* CT18 (accession number GI:16759298) was compared with the TcfD homologs of serovars Paratyphi A (GI:56128825), Choleraesuis (GI:674188659), Infantis (GI: GI:564962257), Schwarzengrund (GI:194711591), Montevideo (GI:363578184), 9,12:l,v:- (GI619501782), Virchow (GI:554050646), and Heidelberg (GI: GI:381296985). The sequences alignment was performed using CLUSTALW and BOXSHADE 3.2 tools. Identical amino acids are shown in black and similar amino acids are shown in grey.

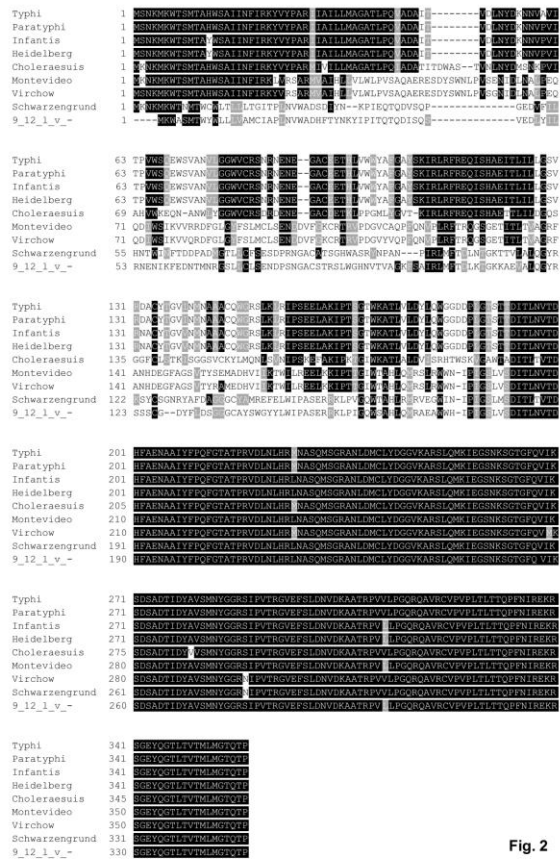


Fig. 2

**Figure 3. *tcf* is differentially expressed in various *Salmonella* serovars.** Reverse transcription PCR was applied to examine the expression of *tcfA* in serovars Typhi, Paratyphi A, Schwarzengrund, Infantis, Virchow, Montevideo, Bredeney, 9,12:l,v:-, Heidelberg and Choleraesuis. Total RNA was extracted from *Salmonella* cultures grown aerobically to the mid-late logarithmic phase (OD<sub>600</sub> 0.5- 1) in LB (**A** and **B**) or in M63 minimal medium (pH 7) (**C** and **D**). Semi-quantitative reverse transcription PCR (**A** and **C**) and real-time reverse transcription PCR (**B** and **D**) were used to determine differences in *tcfA* transcription (normalized to 16S rRNA). 200 ng of RNA were used for cDNA synthesis and equal amounts of cDNA (3 µl) were included as a template for PCR amplification using primers specific to *tcfA* and 16S rRNA. *Salmonella* RNA without a reverse transcriptase stage (-RT) was used as negative control and gDNA from *S. Infantis* was used as a positive control for the PCR reactions. (E) qRT-PCR was applied to determine the fold change in *tcfA* transcription (normalized to 16S rRNA) in *S. Typhi*, *S. Infantis* and *S. Choleraesuis*, while growing aerobically to the mid-late logarithmic phase in M63 minimal medium relative to its expression in LB. One way ANOVA with Dunnett's Multiple Comparison Test were performed to determine statistical significance. ns, not significant; \*\*\*, P<0.001.

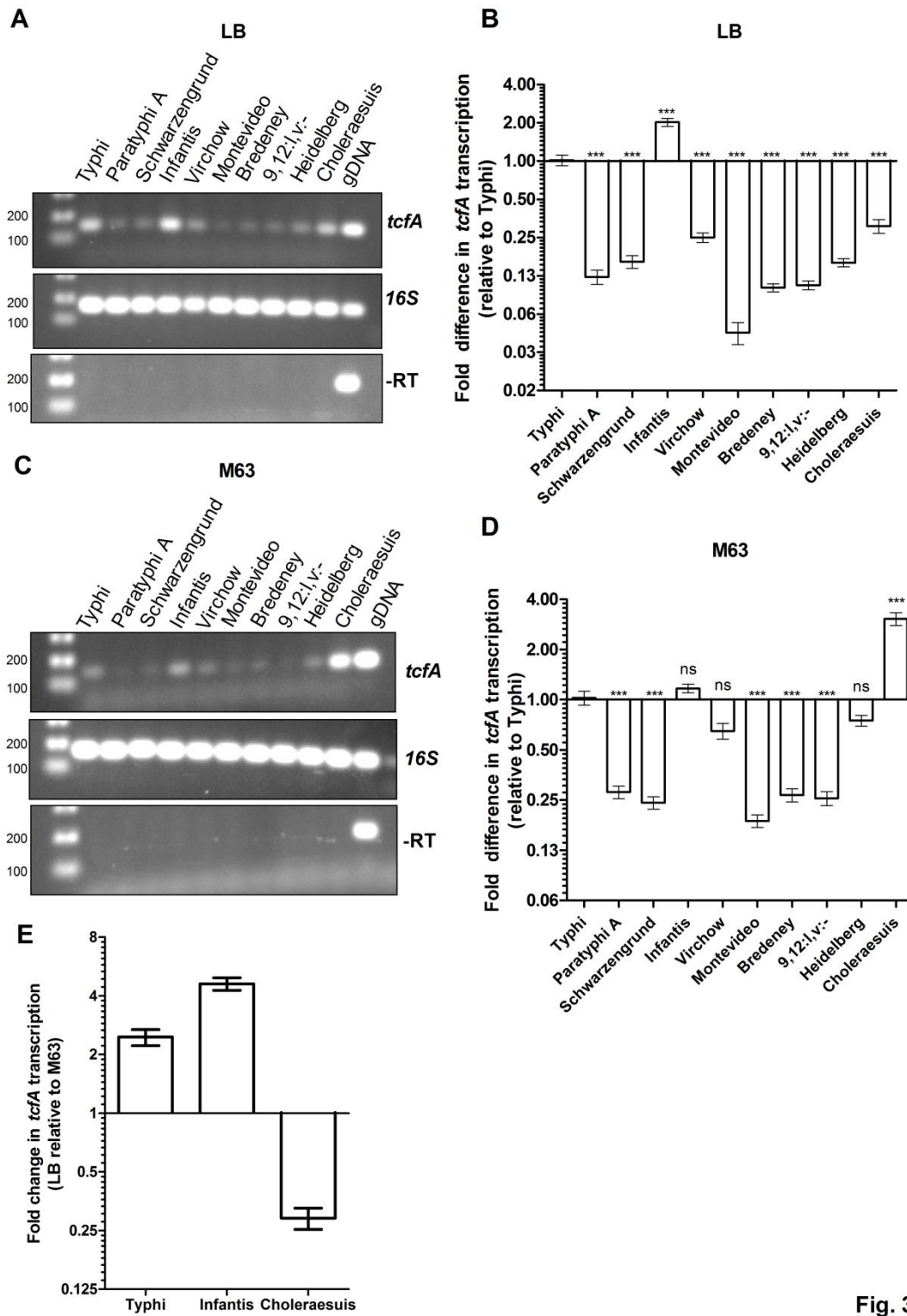


Fig. 3

**Figure 4. *tcf* expression in *S. Infantis* is induced under microaerobic conditions in rich broth.** Total RNA was harvested from *S. Infantis* 119944 grown under different growth conditions and was subjected to a qRT-PCR analysis. Growth conditions that were tested included (1) growth to the late-logarithmic phase (OD<sub>600</sub> ~1) under aerobic conditions in LB broth (late log LB); (2) growth to the mid-logarithmic phase (OD<sub>600</sub> ~0.5) under aerobic conditions in LB broth (mid log LB); (3) growth to the stationary phase (OD<sub>600</sub> ~7) under aerobic conditions in LB broth (stationary LB); (4) growth to the late-logarithmic phase (OD<sub>600</sub> ~1) under aerobic conditions in N-minimal medium pH 5.8 (late log NMM 5.8); and (5) growth to the stationary phase under microaerobic conditions in LB broth (microaerobic LB). The change in *tcfA* transcription (normalized to 16S rRNA) under the different growth conditions relative to *tcfA* expression in the late-logarithmic phase in LB is shown. Indicated values present the mean and SEM of three independent RT-PCR experiments with 2-4 replicates. One way ANOVA with Dunnett's Multiple Comparison Test were performed to determine statistical significance. ns, not significant; \*\*\*, P<0.001.

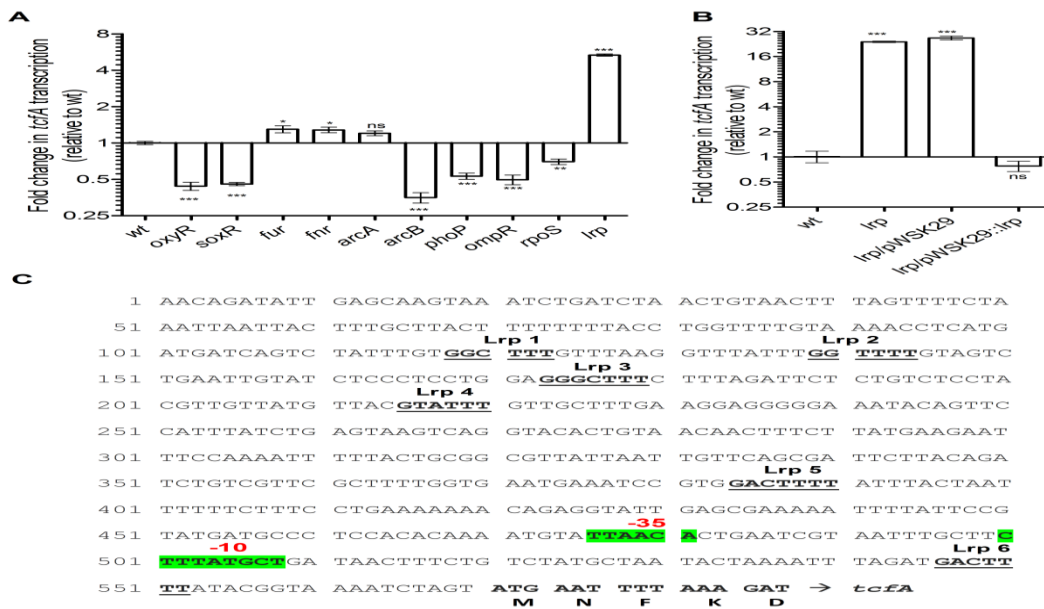


Fig. 5

**Fig. 5** The leucine-responsive regulatory protein (Lrp) is a negative regulator of *tcf*. *S. Infantis* 119944 cultures were grown in LB to the stationary phase under microaerobic conditions at 37°C (*tcf* induction conditions). **(A)** Total RNA was harvested from *S. Infantis* 119944 wild-type (wt) and from ten isogenic regulatory mutant strains (*oxyR*, *soxR*, *fur*, *fnr*, *arcA*, *arcB*, *phoP*, *ompR*, *rpoS* and *lrp*). **(B)** RNA was extracted from *S. Infantis* 119944 wild-type, its isogenic *lrp* null strain, *lrp* harboring pWSK29 (pWSK29) and *lrp* complemented with *lrp* (pWSK29::*lrp*). All RNA samples were reverse transcribed and subjected to qRT-PCR. The change in the transcription of *tcfA* (normalized to 16S rRNA) in all strains relative to *tcfA* expression in the wild-type background is shown. One way Anova with dunnett's test against the wt background was used to determined statistical significance. ns, not significant; \*, $P < 0.05$ ; \*\*, $P < 0.01$ ; \*\*\*, $P < 0.001$ . **(C)** DNA sequence containing the intergenic region upstream from *tcfA* was analyzed *in-silico*. Promoter location including the -10 and -35 boxes was predicted by BPROM highlighted in green and indicated by red text. Six putative Lrp binding sites (Lrp1-6) are numbered, marked in bold and underlined.

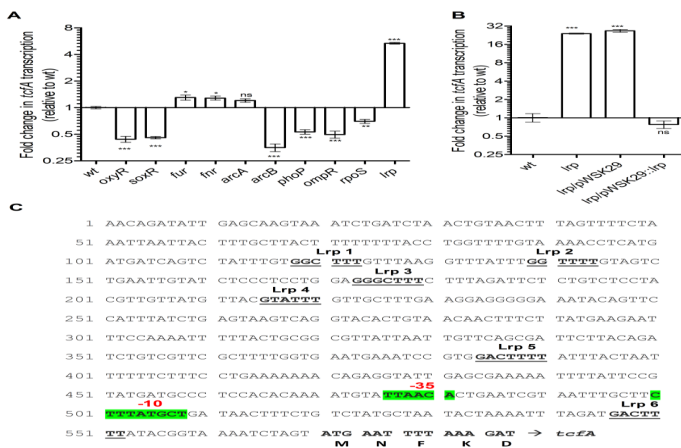


Fig. 5

**Figure 6. Heterologous expression of the *S. Infantis* Tcf in a surrogate bacterial system.** (A) *tcfABCD tinR* and *tioA* from *S. Infantis* were cloned into pBAD18 under the arabinose inducible promoter. pBAD18::*tcfABCD tinR tioA* (pBAD18::*tcf*) was transferred into a fimbriae-less *E. coli* ORN172 strain. *E. coli* ORN172 /pBAD18 (pBAD18) was used as a negative control. Surfaces enriched fractions were extracted, concentrated and separated using 15% SDS-PAGE. (A) Coomassie-stained gel after arabinose induction (lanes 2, 4) or glucose suppression (lane 3) of Tcf (lanes 3, 4) or the empty vector (lane 2) is shown. A specific 13 kDa band that appeared in the arabinose inducible culture was analyzed by LC-MS/MS and identified as the major Tcf subunit, TcfB. (B) Schematic representation of TcfB peptides identified by MS analyses with 42.44% coverage (of the entire protein). The identified peptides are marked in pink and the predicted signal sequence is annotated. (C) The signal peptide of TcfB preceding the Sec cleavage site between positions 24 and 25 (VSA-VQ) as was predicted by the SignalP 4.1 program. Cleavage site prediction is based on three different scores, (C, S and Y) for each position in the TcfB sequence, shown by the red, green and blue lines.

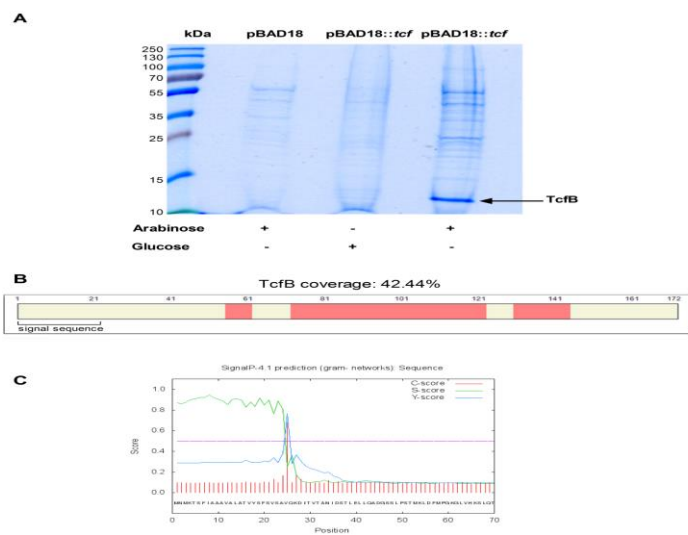


Fig. 6



**Figure 7. Transmission electron microscopy (TEM) images of the Tcf fimbriae.** *E. coli* ORN172 harboring the vector only (pBAD18) (**A**) and *E. coli* ORN172 harboring an inducible Tcf (pBAD18::*tcf-tinR tioA*) (**B**) were grown for 14 h in N-minimal medium (pH 7) supplemented with 50 mM arabinose. Images were taken using FEI F20 Philips-Tecnaï field emission gun transmission electron microscope (TEM) equipped with Gatan One View CMOS camera. Flagella are indicated by arrows and the letter "F" and Tcf fimbriae are shown by arrowheads.

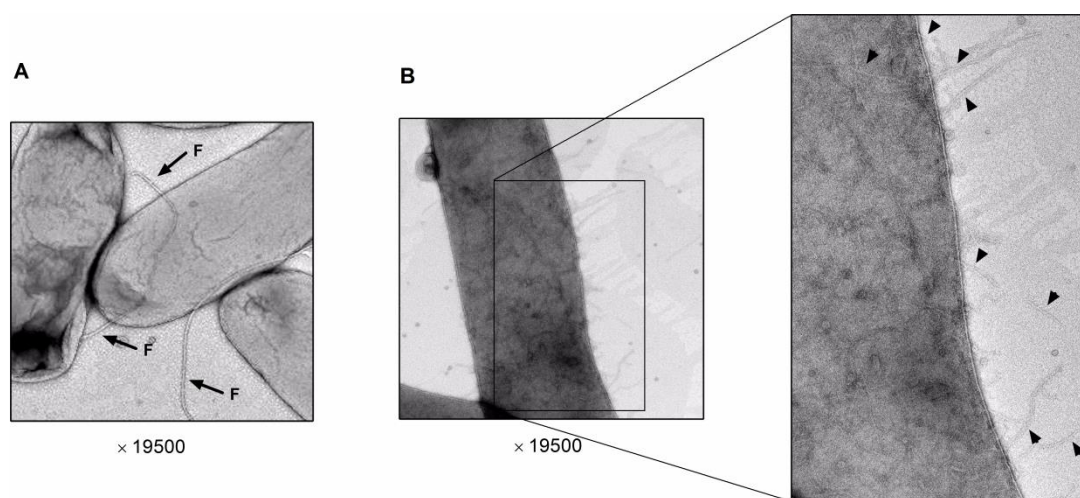
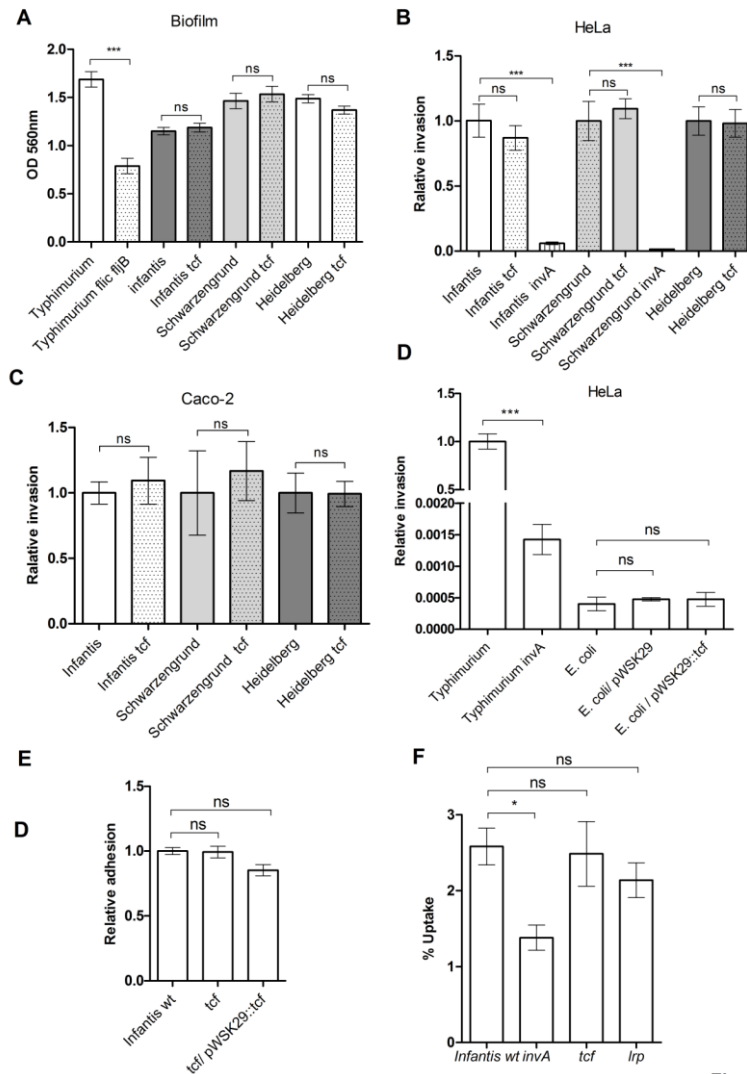


Fig. 7

**Figure 8. Tcf is dispensable for biofilm formation, adhesion to and entry into host cells by NTS serovars.** (A) *S. Infantis*, *S. Schwarzengrund*, and *S. Heidelberg* and their isogenic *tcf* null mutant strains were used to study biofilm formation. As positive and negative controls *S. Typhimurium* SL1344 and its isogenic *fliC fljB* mutant strain, respectively were also included. All strains were grown in LB medium lacking NaCl (biofilm-inducing conditions) at 28°C for 96 h. Biofilm formation was assayed by Crystal Violet staining. The bars represent the mean of six biological repeats with the SEM shown by the error bars. Student t-test was used to determine differences between the wild-type background and its isogenic *tcf* null mutant strain. (B) *Salmonella* strains were grown in LB at 37°C to the late logarithmic phase and used to infect HeLa epithelial cells. Invasion was determined using the gentamicin protection assay and calculated as the percentage of intracellular bacteria (CFU) recovered at 2 h p.i from the total number of inoculum used to infect the cells. *S. Infantis* and *S. Schwarzengrund* strains harboring null deletion in *invA* (required for host cell invasion) were included as a negative control. Invasion of the mutant strains is shown relative to the wild-type backgrounds. (C) *S. Infantis*, *S. Schwarzengrund*, and *S. Heidelberg* and their isogenic *tcf* null mutant strains were used to infect Caco-2 epithelial cells as in (B). (D) *E. coli* ORN172, *E. coli* ORN172 harboring *tcfABCD tinR tioA* from *S. Schwarzengrund* cloned into pWSK29 and *E. coli* ORN172 harboring the empty vector (pWSK29) were used to infect HeLa cells as in (B). *S. Typhimurium* and its *invA* mutant strain were used as a positive and negative controls, respectively. Invasion is shown relative to the *S. Typhimurium* wild-type background invasion. (E) *S. Infantis*, its *tcf* mutant strain and a *tcf* strain harboring the *tcfABCD* genes cloned into pWSK29 were tested for their ability to adhere to HeLa cells in the presence of cytochalasin D (see Materials and Methods). Adhesion is shown

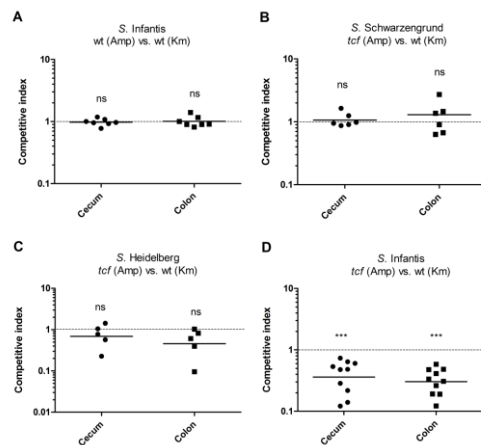
relative to the adhesion of the *S. Infantis* wild-type background. (F) *S. Infantis*, its *invA*, *tcf*, and *lrp* mutant strains were grown to the stationary phase microaerobically and tested for their uptake by RAW264.7 macrophage-like cells using the gentamicin protection assay. *Salmonella* uptake was calculated as the percentage of intracellular bacteria (CFU) recovered at 2 h p.i from the total number of inoculum used to infect the cells. ns, not significant; \*,  $P < 0.05$ ; \*\*,  $P < 0.01$ ; \*\*\*,



$P < 0.001$ .

Fig. 8

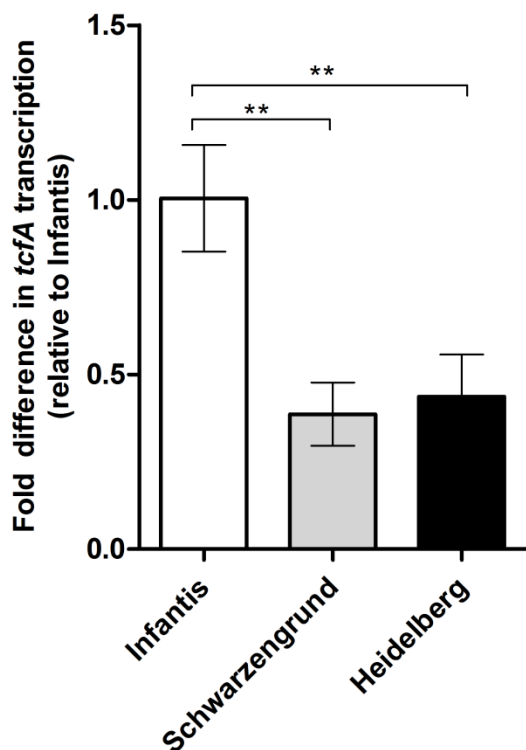
**Figure 9. Tcf fimbria is involved in intestinal colonization of *S. Infantis in-vivo*.** Eight week old female C57BL/6 mice were pretreated with streptomycin and infected by oral gavage with  $1 \times 10^6$ - $1 \times 10^7$  CFU of 1:1 mixture of *S. Infantis* wild-type harboring ampicillin resistant cassette and *S. Infantis* wild-type harboring kanamycin resistant cassette (**A**), and *S. Schwarzengrund* (**B**), *S. Heidelberg* (**C**) and *S. Infantis* (**D**) wild-type background (kanamycin resistant) and their isogenic *tcf* null mutant strains (ampicillin resistant). At day four post infection, mice were sacrificed and organs were harvested and homogenized in sterile PBS. Dilutions of homogenized tissues were plated on selective XLD plates. Competitive index values were determined as  $[\text{CFU } tcf(\text{Amp}^R)/\text{CFU wild-type (Km}^R)]_{\text{output}} / [\text{CFU } tcf/\text{CFU wild-type}]_{\text{input}}$ . Each symbol represents the C.I. result in one animal, while the mean is shown by the horizontal line. Competition experiments between the wild-type and the *tcf* mutants of *S. Heidelberg* and *S. Schwarzengrund* were done once with five or six mice in each experiment and the competition experiment between the wild-type and the *tcf* mutants of *S. Infantis* was done in two independent experiments with five mice in each experiment (10 mice in total). One-sample *t*-test against a theoretical mean of 1.0 (assumes equal fitness of both strains) was used to determined statistical



significance. ns, not significant; \*\*\*,  $P < 0.001$ .

Fig. 9

**Figure 10. *tcfA* is expressed at varying levels by distinct NTS serovars during mouse infection.** Eight week old female C57BL/6 mice were pretreated with streptomycin and infected by oral gavage with  $2-3 \times 10^7$  CFU of *S. Infantis*, *S. Schwarzengrund* and *S. Heidelberg* (four mice per group). One day p.i. the mice were sacrificed and total RNA was extracted from the cecum and colon of these infected mice. qRT-PCR was applied to determine *tcfA* transcription (relative to 16S rRNA). *tcfA* expression is presented relative to its expression in the mice infected with *S. Infantis*. One way ANOVA with Dunnett's Multiple Comparison Test was performed to determine statistical significance. \*\*,  $P < 0.01$ .



**Fig. 10**