



Horizontal Transmission of Stress Resistance Genes Shape the Ecology of Beta- and Gamma-Proteobacteria

Shady Mansour Kamal¹, David J. Simpson², Zhiying Wang², Michael Gänzle^{2*} and Ute Römling^{1*}

¹ Department of Microbiology, Tumor and Cell Biology, Karolinska Institutet, Solna, Sweden, ² Department of Agricultural, Food and Nutritional Science, University of Alberta, Edmonton, AB, Canada

OPEN ACCESS

Edited by:

Kira Makarova,
National Center for Biotechnology
Information (NLM), United States

Reviewed by:

Alan J. Wolfe,
Loyola University Chicago,
United States
Torsten Waldmühling,
Darmstadt University of Technology,
Germany

*Correspondence:

Ute Römling
ute.romling@ki.se
Michael Gänzle
mgaenzle@ualberta.ca

Specialty section:

This article was submitted to
Evolutionary and Genomic
Microbiology,
a section of the journal
Frontiers in Microbiology

Received: 21 April 2021

Accepted: 07 June 2021

Published: 06 July 2021

Citation:

Kamal SM, Simpson DJ, Wang Z,
Gänzle M and Römling U (2021)
Horizontal Transmission of Stress
Resistance Genes Shape the Ecology
of Beta- and Gamma-Proteobacteria.
Front. Microbiol. 12:696522.
doi: 10.3389/fmicb.2021.696522

The transmissible locus of stress tolerance (tLST) is found mainly in beta- and gamma-Proteobacteria and confers tolerance to elevated temperature, pressure, and chlorine. This genomic island, previously referred to as transmissible locus of protein quality control or locus of heat resistance likely originates from an environmental bacterium thriving in extreme habitats, but has been widely transmitted by lateral gene transfer. Although highly conserved, the gene content on the island is subject to evolution and gene products such as small heat shock proteins are present in several functionally distinct sequence variants. A number of these genes are xenologs of core genome genes with the gene products to widen the substrate spectrum and to be highly (complementary) expressed thus their functionality to become dominant over core genome genes. In this review, we will present current knowledge of the function of core tLST genes and discuss current knowledge on selection and counter-selection processes that favor maintenance of the tLST island, with frequent acquisition of gene products involved in cyclic di-GMP signaling, in different habitats from the environment to animals and plants, processed animal and plant products, man-made environments, and subsequently humans.

Keywords: *Escherichia coli*, *Pseudomonas aeruginosa*, *lebsiella pneumoniae*, *Cronobacter sakazakii*, protease, disaggregase, small heat shock protein, heat tolerance

INTRODUCTION

Bacteria are found in most extreme habitats as these organisms possess an almost unrestricted potential to adapt to altering and adverse environmental conditions including survival of a temporary rise to lethal conditions and occupation of novel ecological niches. Gradual adaptation is mediated by mutations of the core genome, while a rapid and quantum-leap adaptation beyond the functional plasticity of the available genetic repertoire is conferred by mobile genetic elements, plasmids, and transposons, in combination with a vast repertoire of genome engineering tools and repetitive DNA sequences; and the acquisition of novel genes (Shintani, 2017). The horizontally transferred physiological characteristics that are commonly paid attention to include resistance against antimicrobial agents and heavy metals, virulence properties, and widening of catabolic capabilities and resistance (Lan et al., 2001; Elliott et al., 2002; Herold et al., 2004; Lin et al., 2015). Theoretically, and perhaps even practically, there is no restriction to which type of genetic elements are to be horizontally transferred upon exposure to a certain selective pressure; however, properties and transfer of mobile genetic elements that confer resistance to environmental stress

(Berendsen et al., 2016) are not as well understood when compared to mobile genetic elements that enhance virulence or mediate antimicrobial resistance.

A mobile genomic island conferring heat resistance was independently identified in *Escherichia coli* and *Pseudomonas aeruginosa* and termed locus of heat resistance (LHR) and transmissible locus of protein quality control (tLPQC), respectively, (Lee et al., 2015; Mercer et al., 2015). To prevent the continuing use of divergent nomenclature, we propose the term transmissible locus of stress tolerance (tLST). This genomic island provides an exceptional example of the mobilization of a number of highly conserved genes to be commonly horizontally transferred among diverse members of beta- and gamma-Proteobacteria from a so far unknown origin (Lee et al., 2015; Mercer et al., 2015). Initially discovered to mediate tolerance toward lethal heat shock in strains of *Cronobacter sakazakii*, *Klebsiella pneumoniae*, *P. aeruginosa*, and *E. coli* (Bojer et al., 2010; Gajdosova et al., 2011; Lee et al., 2015; Mercer et al., 2015), the tLST island was later identified to provide a wide range of tolerance phenotypes towards environmental and anthropogenic stresses including chlorine and other oxidizing chemicals, and high hydrostatic pressure (Li et al., 2020; Wang et al., 2020) and may interfere with the expression of virulence genes (Wang et al., 2020). The initial analysis of gene products, which are often xenologs of chromosomally encoded genes, showed that tLST island gene products are characterized by physiological and biochemical features that are complementary to, expand or replace the function of core gene products and that allow the organism to enhance persistence and transmission (Wang et al., 2020). The archetype likely close to the major pathogen-related ancestral composite variant of this genomic island is the 18–19 kbp tLSTa (Figure 1), however, several other variants with insertions or deletions have been identified, including the 14–15 kbp tLST1 and the 19 kbp tLST2 (Gajdosova et al., 2011; Lee et al., 2015; Boll et al., 2017; Nguyen et al., 2017). This review aims to summarize current knowledge on the tLST island with respect to ecology, evolution, and mechanisms of resistance. We also use the available information to propose hypotheses related to the evolutionary processes that maintain the tLST island in distinct isolates of many species of Proteobacteria, and the role of the tLST island in the resistance of food-borne and nosocomial bacterial pathogens towards antimicrobial interventions used in food processing, (waste) water treatment, and health care settings.

ANTROPHOGENIC SELECTIVE PRESSURES FOR THE tLST

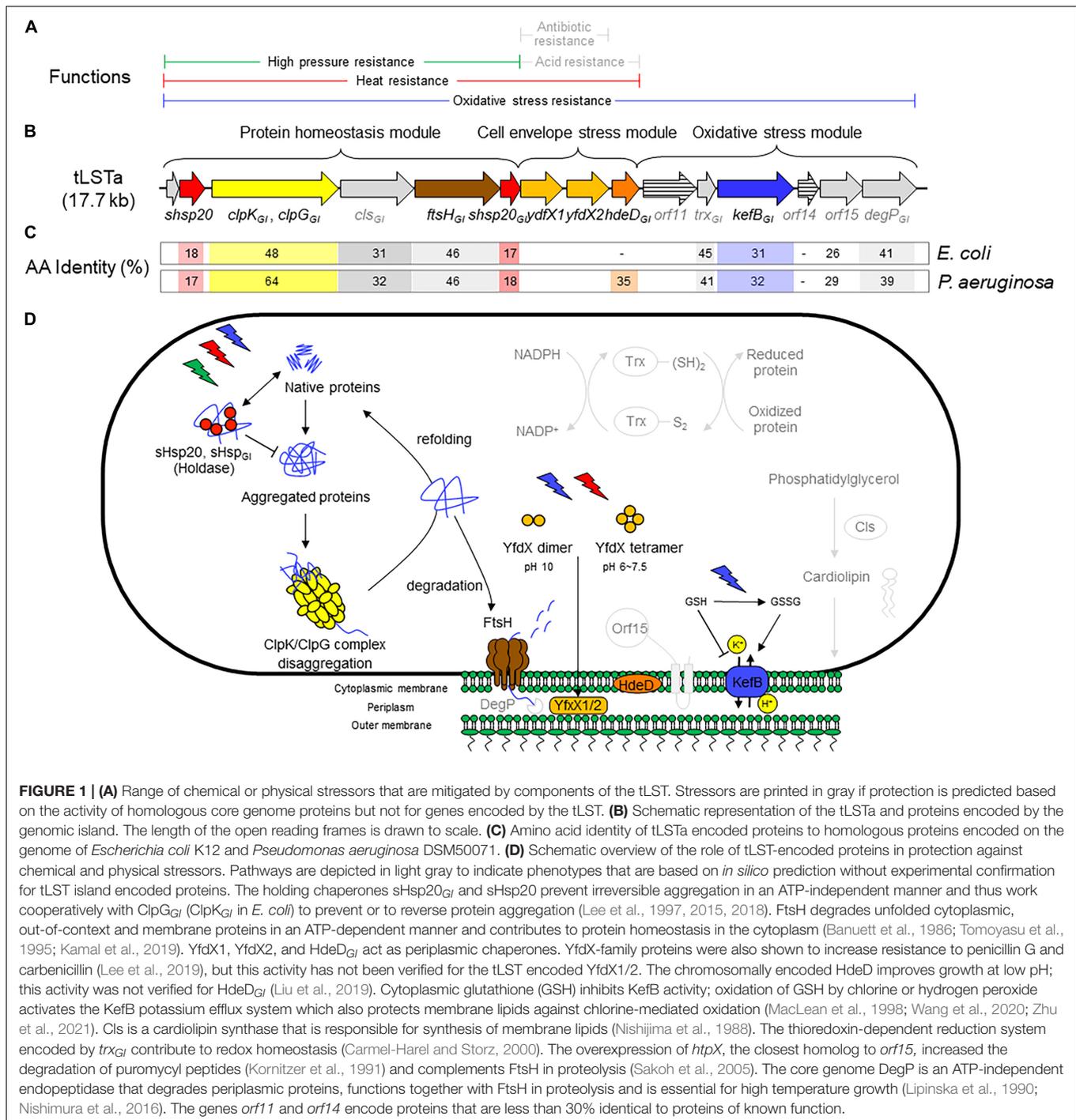
In 1884, Ferdinand Hueppe described the isolation of the *E. coli* strain C from soured cow's milk that later became one of the *E. coli* model organisms for biotechnological purposes and basic scientific studies (Hueppe, 1884; Krol et al., 2019). The now available genome sequence revealed that *E. coli* strain C encodes the tLST, suggesting that tolerance against exposure to stress conditions such as oxidative stress and elevated temperatures is a horizontally transferred feature that predates industrial food production, water sanitation, and antibiotic use

(Figure 1). Previously documented thermotolerant pathogenic bacteria causing outbreaks due to contaminated milk powder span from an outbreak during the second world war caused by *Salmonella enterica* serovar Senftenberg (Goepfert and Biggie, 1968) to recent infections in neonates caused by *C. sakazakii* (Nazarowec-White and Farber, 1997); both pathogens were later identified to carry the tLST island (Gajdosova et al., 2011; Nguyen et al., 2017; Mercer et al., 2017b). Raw milk contains a diverse bacterial microbiota that is beneficially used for the processing of this animal product to cheese and other fermented products, but may also include pathogenic or opportunistic pathogenic organisms (Quigley et al., 2013). Pasteurization or thermization of fluid and cheese milk may provide an evolutionary pressure to select for organisms that carry the tLST island. This concept is supported by the high prevalence of the tLST island in cheese milk after thermization, i.e., heating to 60°C for 30 min (Marti et al., 2016; Boll et al., 2017). Other anthropogenic habitats with a high prevalence of tLST-positive bacteria include chlorinated waste water, where more than 50% of isolates were found tLST positive (Zhi et al., 2016), North American meat processing plants that employ thermal treatments as pathogen interventions on beef carcasses (Zhang et al., 2020; Guragain et al., 2021), and DaQu, a saccharification starter that is used in China for production of cereal beverages and vinegar (Wang et al., 2018). Daqu is produced from spontaneously fermented cereals; with the fermentation microbiota are recruited from plant microbiota which includes plant-associated Enterobacteriaceae (Zheng et al., 2011). During fermentation, the temperature increases to 50–60°C and fermentation conditions select for Gram-positive bacilli that are heat resistant owing to the presence of the *spoVA*^{2mob} operon as well as tLST-positive Enterobacteriaceae (Wang et al., 2018).

The presence of the tLST, however, is not limited to bacteria associated with direct anthropogenic manipulation procedures, indicating that efficient horizontal transfer of the tLST to opportunistic pathogenic microorganisms may occur in alternative ecological niche including the environment. The opportunistic pathogen *P. aeruginosa* is foremost an environmental organism that thrives in soil, water and in association with plants (Lee et al., 2020). In *P. aeruginosa*, at least two ubiquitous clones, groups of closely related strains that can be recovered from the clinical habitat including patients as well as the environment have acquired the tLST (Lee et al., 2015). Likewise, the tLST is found in *Cronobacter* and *Klebsiella* species which are relevant as opportunistic and nosocomial (hospital-acquired) human pathogens, but originate from environmental niches (Schmid et al., 2009; Gajdosova et al., 2011; Mercer et al., 2015). Strong selective pressure for maintenance of the tLST may therefore exist also in environmental habitats that can nevertheless be impacted by human activity such as (waste) water treatment.

THE tLST PROTECTS AGAINST MULTIPLE STRESSES

The tLST island was originally discovered as it mediated thermotolerance to food-derived bacteria and pathogens



(Gajdosova et al., 2011; Lee et al., 2015; Mercer et al., 2015). tLST-mediated thermal tolerance is not incremental but represents a “quantum leap” of superior functionality. For example, the tLST core genes *dna-hsp20-clpG* can provide up to 10-fold higher lethal thermotolerance to genetically unrelated thermosensitive *P. aeruginosa* and *K. pneumoniae* strains (Bojer et al., 2011; Lee et al., 2015). On a similar scale, *E. coli* lacking the entire tLST exhibit a $D_{60^\circ\text{C}}$ -value of less than 1 min while the $D_{60^\circ\text{C}}$ -value

of tLST-positive strains of *E. coli* ranges from 10 min to more than 60 min (Li and Gänzle, 2016); i.e., treatment at 60°C for 10 min reduces cell counts of tLST-negative strains by more than 10 log(cfu/mL), while tLST positive strains resist treatment with a reduction of less than 1 log(cfu/mL). tLST-mediated thermotolerance explains the high prevalence of tLST-positive strains of *E. coli* in the meat and cheese production chains (Dlusskaya et al., 2011; Marti et al., 2016; Boll et al., 2017;

Zhang et al., 2020; Guragain et al., 2021). tLST-positive strains of *E. coli* also are among the most pressure resistant strains of this species (Liu et al., 2015) and cloning of the tLST island confirmed that the genomic island can increase pressure resistance of *E. coli* (Li et al., 2020). In contrast to heat resistance, pressure resistance in *E. coli* can be conferred by alternative genetic alterations that mediate equivalent pressure resistance in tLST-negative strains (Vanlint et al., 2011).

The observation that tLST-positive strains of *E. coli* are highly enriched in chlorinated waste water (Zhi et al., 2016) led to the discovery that the genomic island also mediates resistance to chlorine, hydrogen peroxide, and peroxyacetic acid but not to acrolein or isothiocyanates (Wang et al., 2020). tLST mediated resistance toward other stressors has not been demonstrated experimentally, but is expected from the predicted protein function. For example, homologs of tLST-encoded proteins were shown to protect against acid stress and to increase antibiotic resistance (Figure 1 and below).

Bioinformatic and functional analyses of the tLST mediated stress resistance suggested that the three different parts of the genomic island predominantly function to protect different segments of the bacterial cells. Proteins encoded by the protein homeostasis module of the tLST (Figure 1) have been shown to predominantly prevent or reverse aggregation and oxidation of cytoplasmic and membrane proteins (Lee et al., 2020). Proteins encoded by the cell envelope stress module are periplasmic chaperones and have been shown to prevent oxidation of membrane lipids. Several proteins of the oxidative stress module are predicted to mitigate oxidative stress by various mechanisms including proteolysis and ion antiport, the latter conducted by KefB_{GI}. KefB_{GI} is a H⁺/K⁺ antiporter that maintains an inside acidic membrane potential at alkaline pH in the presence of chlorine (Figure 1; Mercer et al., 2017a; Lee et al., 2018, 2020; Li et al., 2020; Wang et al., 2020; Zhu et al., 2021). The tLST-mediated pressure resistance phenotype of *E. coli* isolates is provided by the protein homeostasis module (Li et al., 2020); the heat resistance phenotype in *E. coli*, *K. pneumoniae*, and *P. aeruginosa* is encoded by the protein homeostasis and cell envelope stress modules and *E. coli* strains carrying only the former are substantially less resistant than those carrying both modules or the full island (Bojer et al., 2010; Mercer et al., 2015, 2017a), while the resistance to chlorine and other inorganic oxidative chemicals requires presence of all three modules (Wang et al., 2020). In *P. aeruginosa*, the three genes *dna-shsp20_{GI}-clp_{GI}* on the protein homeostasis module mediate significant thermotolerance to unrelated thermosensitive strains though (Lee et al., 2015).

EXPRESSION OF GENES ENCODED BY THE tLST ISLAND

A hallmark of the expression of gene products of the tLST island is their production during exponential growth, which is substantially increased upon entry in the stationary phase of growth at environmental or body temperature (Mercer et al., 2017a) or exclusively highly produced in the late growth phase

(Lee et al., 2015). Even more, in *E. coli* and *P. aeruginosa*, tLST-encoded proteins can be among the most abundant proteins (Williams, 1995; Sriramulu et al., 2005; Lee et al., 2015, 2018; Kamal et al., 2019; Li et al., 2020). This expression pattern is distinct from homologous core genome heat shock proteins, which are overexpressed upon exposure to sublethal heat stress. Indeed, a network of proteases has recently been shown to rescue growth arrest of *P. aeruginosa* (Basta et al., 2020). Acquisition of tLST proteases might therefore aid recovery of widely distributed *P. aeruginosa* clones from this environmentally relevant physiological status (Bergkessel, 2020). tLST-mediated heat resistance is enhanced by the presence of 4% NaCl but this effect is mediated by accumulation of compatible solutes rather than over-expression of tLST-encoded proteins as addition of up to 4% NaCl did not increase expression from the tLST promoter that is located 63 bps upstream of the *orf1* (alternatively named *dna*), a Mer-like transcriptional regulator (Figure 1; Pleitner et al., 2012; Mercer et al., 2017a). Expression of genes encoded by the tLST in *E. coli* at alkaline, but not at neutral pH was reported to be repressed by the Cpx two-component regulatory system which mitigates cell envelope stress during growth at alkaline pH (Zhu et al., 2021). The *in silico* prediction of multiple promoters that respond to diverse environmental stimuli suggests that constitutive expression of tLST island encoded proteins is mediated by multiple factors that have not been fully elucidated (Nguyen, 2019). As a trans mediated cross-reactive physiological trait, the tLST mediated protection against oxidative stress and interferes with the induction of prophages carrying genes coding for the Shiga toxin in the late phage protein region production (Marti et al., 2016; Wang et al., 2020).

FUNCTION OF INDIVIDUAL GENE PRODUCTS ENCODED ON THE tLST ISLAND

A major hallmark of the tLSTa is that the majority of the gene products are xenologs, distantly related homologs of evolutionary highly conserved core genes from a phylogenetically distant bacterial species (Lee et al., 2015, 2016; Mercer et al., 2015). Gene duplication is known as a concept to widen the physiological and metabolic capabilities of organisms including the human pathogen *Mycobacterium tuberculosis* (Tekaia and Gordon, 1999). With respect to their physiological impact, the preservation of gene products such as the proteases FtsH, HtpX, and DegP, small heat shock protein holding chaperones and the redox protein thioredoxin Trx in all organisms including humans emphasizes a central role of these proteins in basic physiological functions such as protein homeostasis and redox balance that are core survival mechanisms of cellular organisms. An overview of the function of proteins that are encoded by the tLSTa is shown in Figure 1. Several gene products including the small heat shock protein sHsp20_{1GI}, the disaggregase Clp_{GI}/Clp_{KGI} and the protease FtsH_{GI} (FtsH2) have been genetically and biochemically characterized (Lee et al., 2015, 2018; Kamal et al., 2019). Information on the distinct function of KefB_{GI} is derived from studies with *kefB* deletion mutants of tLST-expressing *E. coli*

(Zhu et al., 2021). Information on the function of other tLSTa encoded genes is also derived from the expression of plasmid-encoded (fragments of the) tLST in *E. coli* in combination with *in silico* prediction of protein functions that is based on the function of core-genome homologs. As would be expected for xenologous gene products, the amino acid identity of several tLSTa encoded proteins to core genome proteins in *E. coli* or *P. aeruginosa* is low, and even below 30%, i.e., *orf14* shares less than 30% homology to PsiE and *orf15* shares less than 30% homology to the M48 type protease HtpX. A respective functionality and catalytic activity is, though, predicted based on the conservation of the domain and the respective amino acid signatures (Wanner, 1986). For other proteins including the membrane bound protease FtsH, the periplasmic chaperone protease DegP and thioredoxin Trx, the homology to core genome proteins is below 45%, and the catalytic activity can be reliably predicted based on the conservation of signature amino acids. Details on the regulation of the catalytic activity and the respective substrate specificity need to be unraveled by future experimentation.

Proteins Encoded by the Protein Homeostasis Module

The presence of the protein homeostasis module is required for pressure, heat, and chlorine resistance (Mercer et al., 2015; Lee et al., 2016; Li et al., 2020; Wang et al., 2020). The protein homeostasis module includes the two small heat shock proteins sHsp20_{1GI} and sHsp20_{2GI}, a cardiolipin synthase, the protease FtsH_{GI}, and the disaggregase ClpG_{GI}, which is termed ClpK_{GI} in *E. coli* and *K. pneumoniae* (Figure 1). Disaggregases transform aggregated proteins by ATPase driven force into linear and refoldable peptide chains (Lee et al., 2018; Mogk et al., 2018). The ability of the disaggregase ClpG_{GI} to process aggregates that are formed at higher temperature or initial protein concentration is superior to the core genome ClpB-DnaK co-disaggregation system (Mogk et al., 1999; Motohashi et al., 1999; Zolkiewski, 1999; Lee et al., 2018). Furthermore, ClpG_{GI} directly binds its substrates, protein aggregates, through an extension of the N-terminal domain and displays a high intrinsic ATPase activity (Lee et al., 2018), in contrast to ClpB where the co-chaperone DnaK delivers aggregates to ClpB with subsequent activation of its ATPase activity. *P. aeruginosa* but not *E. coli* has the core genome equivalent ClpG to ClpG_{GI} with identical domain structure in addition to the more distantly related disaggregase ClpB. Preliminary one-dimensional protein profiles of respective *P. aeruginosa* SG17M mutants indicated a distinct substrate pattern for each of the three genome encoded disaggregases, ClpB, ClpG, and ClpG_{GI} (Lee et al., 2018). Although the phylogenetic origin of ClpG_{GI} can be diverse, the ClpG_{GI} family consists of highly conserved proteins encoded exclusively by tLST-like gene clusters of pathogens and environmental bacteria of diverse evolutionary origin (Figures 1, 2A,B). The recent information expansion of genome sequences has, however, unraveled several additional ClpG_{GI} subgroups that are characterized by distinct N- and C-terminal domains present in various genetic context (Figure 2).

sHsp20 proteins are holding chaperones that prevent irreversible aggregation of proteins in their initial state of unfolding and therefore functionally cooperate with disaggregases (Sun and MacRae, 2005). The core of sHsp20 proteins consists of seven anti-parallel aligned beta-strands. The divergent structurally disordered N- and C-terminal extensions and the intramolecular unstructured loop between the fourth and fifth β -strand of sHsp20 proteins (with the occurrence of up to two sHsp20 proteins on the tLST island) might broaden the substrate range and point to distinct mechanisms of substrate recognition and stabilization equally as subunit homo- and hetero-oligomerization, which is considered the inactive status of sHsp20 proteins (Figure 3; Sun and MacRae, 2005). Of the two sHsp20 proteins encoded by the tLST, only sHsp20_{1GI} has been characterized to provide holding chaperone activity to the thermolabile model substrate citrate synthase (Lee et al., 2015).

The membrane-bound protease FtsH is required for optimal growth and is involved in the turnover of proteins equally as it degrades truncated, out-of-context, and disordered proteins (Banuett et al., 1986; Tomoyasu et al., 1995). Substrate pull-down assays with FtsH proteins that possess a trapping prone ATPase active AAA+ domain, but lack proteolytic activity revealed that the core genome protease FtsH of the aquatic isolate *P. aeruginosa* SG17M distinctively bound and processed substrates compared to FtsH_{GI}. Selective degradation of PhzC, a key enzyme for the biosynthesis of the redox-active secondary metabolite phenazine and of the heat shock sigma factor RpoH, a major known target of FtsH in *E. coli*, was verified by *in vivo* degradation assays (Kamal et al., 2019). A contribution of the tLSTa-encoded FtsH_{GI} to stress tolerance in *E. coli* has not been demonstrated and isogenic strains of *E. coli* carrying tLST2, which encodes a functional version of *ftsH_{GI}* and tLST1, were the gene coding for FtsH_{GI} is truncated, have comparable heat resistance (Boll et al., 2017). In *P. aeruginosa* clone C, *ftsH_{GI}* backs up core genome *ftsH* with regards to a number of phenotypes including heat and antibiotic resistance and biofilm formation (Kamal et al., 2019). The core genomes of *E. coli* and *P. aeruginosa* encode three and four, respectively, cardiolipin synthases with overlapping function (Lee et al., 2020). Cls mediated changes of the composition of membrane fatty acids improved stationary phase survival of *E. coli* (Nishijima et al., 1988; Hiraoka et al., 1993).

Proteins Encoded by the Cell Envelope Stress Module

The cell envelope stress module encodes the two predicted periplasmic proteins YfdX1 and YfxX2 and the integral membrane protein HdeD_{GI}. The module is required for heat and chlorine resistance, but not for pressure resistance in *E. coli* (Mercer et al., 2015; Li et al., 2020; Wang et al., 2020). Each of the three proteins of the cell envelope stress module is required for heat resistance (Mercer et al., 2017a; Li et al., 2020). The periplasmic stress chaperones YfdX1 and YfdX2 are predicted to act as periplasmic chaperones to control protein quality. YfdX in *S. Typhi* was also shown to increase resistance to penicillin G and carbenicillin (Lee et al., 2019; Liu et al., 2019) but this activity has not been verified for the tLST encoded YfdX1/2. The chromosomally encoded integral membrane protein HdeD

improves growth at low pH in *E. coli* (Mates et al., 2011); however, currently available data do not provide evidence for a contribution of the tLST or the tLST-encoded HdeD_{GI} to enhance growth or survival at low pH.

Proteins Encoded by the Oxidative Stress Module

The function of the oxidative stress module remains poorly characterized when compared to the protein homeostasis and cell envelope stress modules (Figure 1). The oxidative stress module is required for resistance to chlorine but not for heat or pressure resistance. As mentioned earlier, the genes

orf11, *orf14*, and *orf15* encode proteins that are less than 30% identical to proteins of known function; the expression of the tLHR-encoded *orf15* in *E. coli* too low to be detected by proteome analysis (Li et al., 2020). DegP encoded on the core genome of *E. coli* is a periplasmic chaperone and endopeptidase that is essential for growth at high temperature, aids in protein homeostasis, and activates expression from the alternative sigma factor σ^E (Lipinska et al., 1990; Mecas et al., 1993; Sklar et al., 2007). The *orf16* tLST-encoded DegP is, however, only 41% identical to the core genome DegP; in addition, the tLST island does not improve growth of *E. coli* at high temperature (Ruan et al., 2011) or at alkaline pH (Zhu et al., 2021).

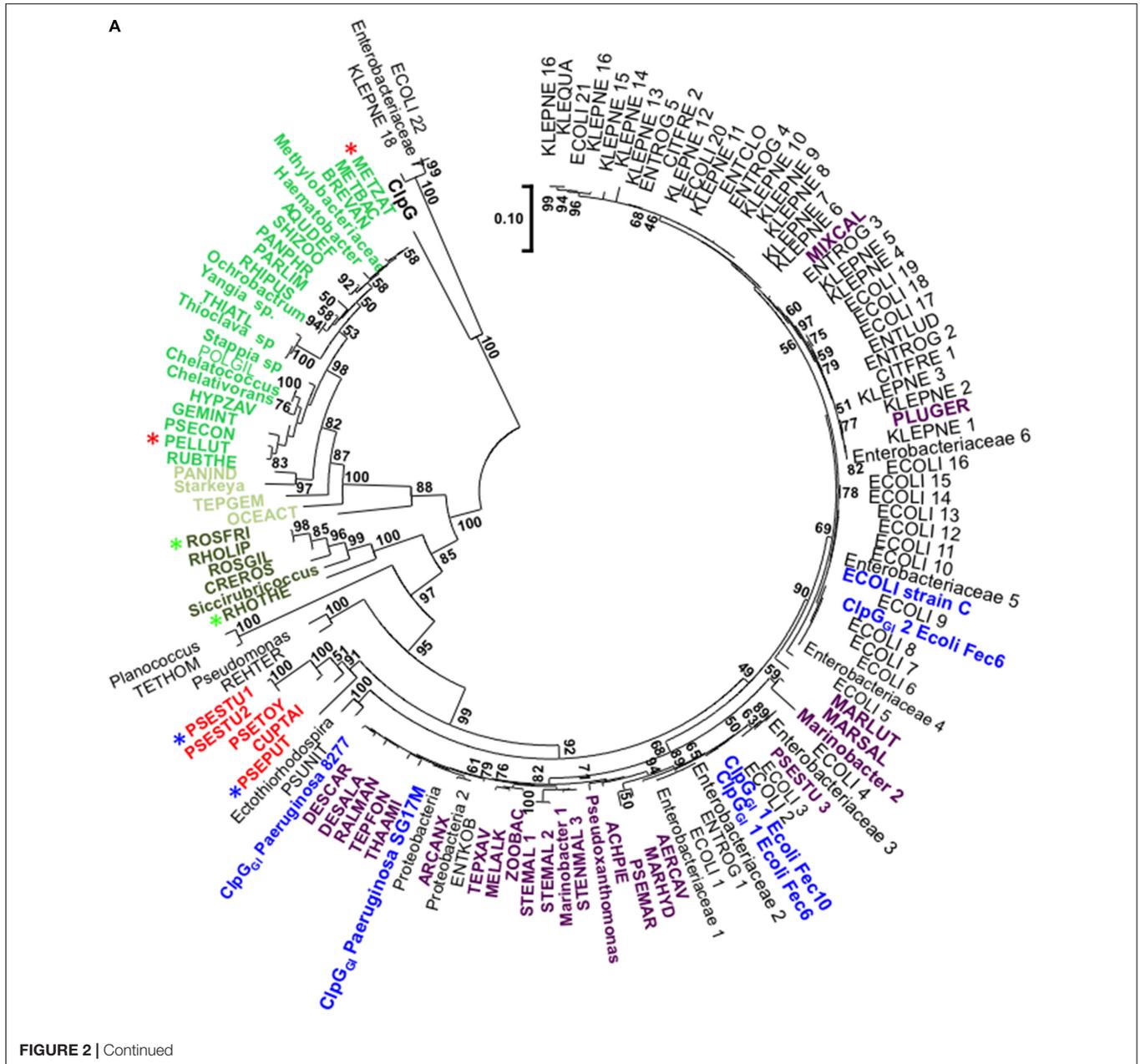
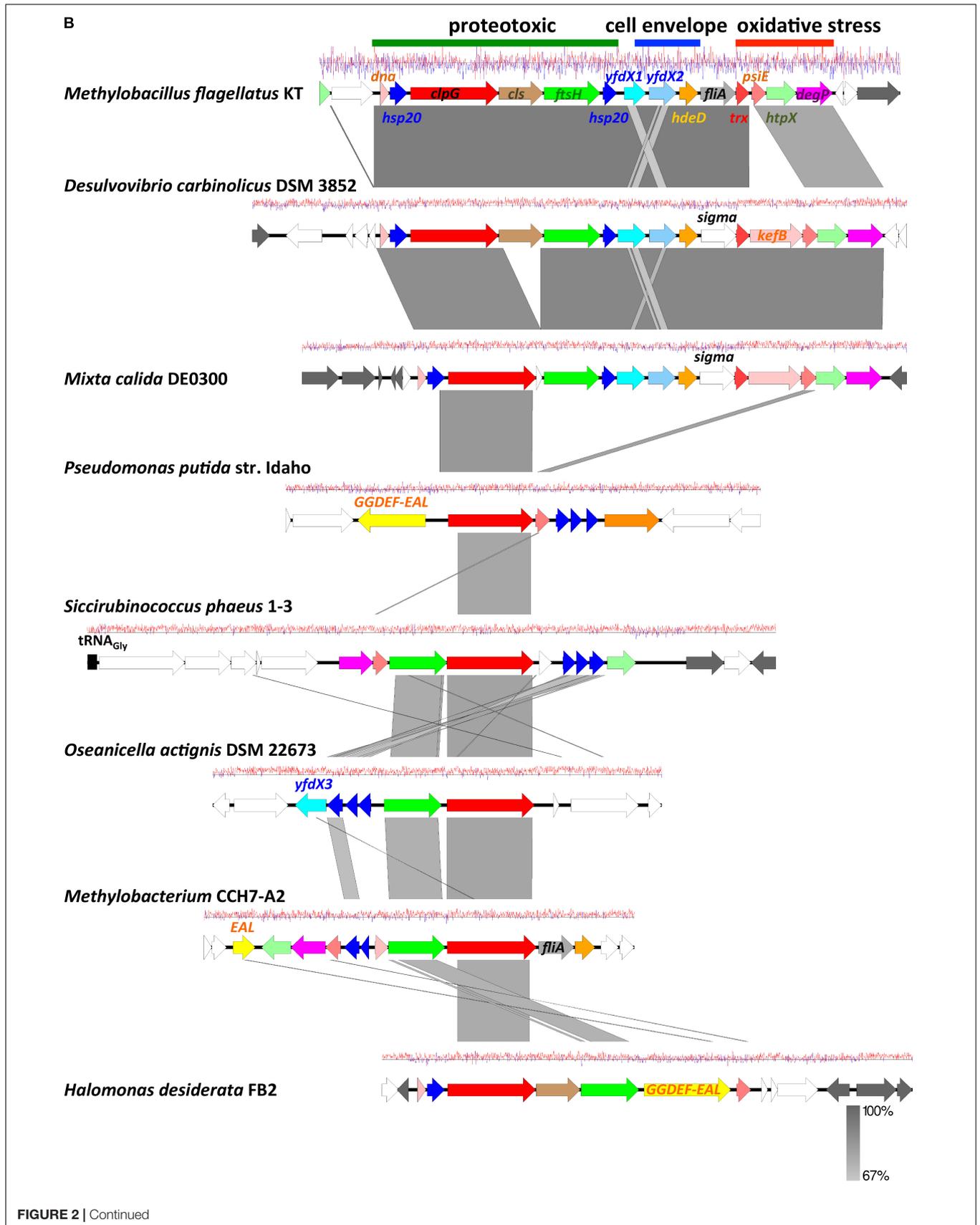


FIGURE 2 | Continued



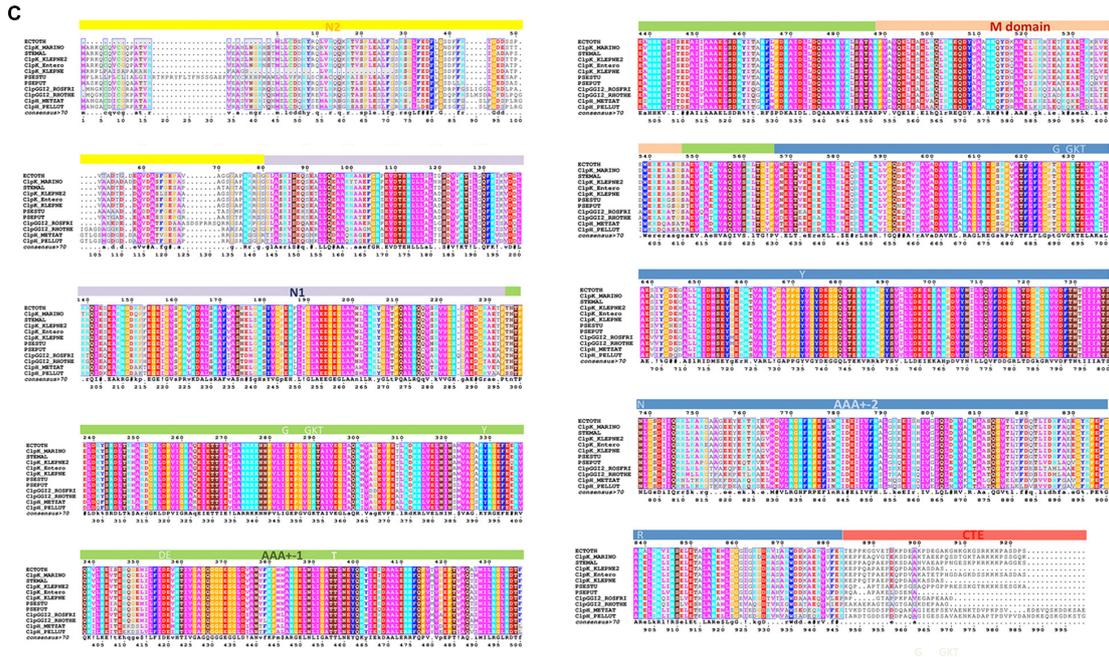


FIGURE 2 | Phylogenetic tree of ClpG_{GI} proteins. **(A)** ClpG_{GI} proteins most closely related to ClpG_{GI} from the urinary isolate *P. aeruginosa* 8277 as query were collected after Blast search with standard parameters at the NCBI homepage (Altschul et al., 1990). Core genome ClpG from *P. aeruginosa* SG17M was used as outgroup. Using the available Blast acquired data, there is a clear distinction in similarity values between ClpG and the ClpG_{GI} group family members. However, among the proteins most closely related to ClpG_{GI}, at least two potentially novel subgroups of ClpG proteins, ClpG2_{GI} (in dark green) and ClpH (in green and light green) can be discriminated. ClpG_{GI} proteins previously investigated or mentioned in the text in blue, ClpG_{GI} proteins from environmental species in violet and an additional potentially distinct novel ClpG_{GI} subgroup (ClpG3_{GI}) in red. Abbreviations used in the tree are defined in **Supplementary Table 1**. The protein sequences were aligned with ClustalX2.1 (Aiyar, 2000) using standard parameters and in MEGA7 (Kumar et al., 2016), the evolutionary relationship using the standard Maximum Likelihood protocol was determined with 100 bootstraps and the phylogenetic tree subsequently displayed. The branch lengths correspond to the number of substitutions per site. Stars indicate the protein sequences used for the alignment in **Figure 2C**. **(B)** Gene synteny around representative *clpG_{GI}* genes from different *clpG_{GI}* subfamilies from archetypical environmental isolates. As a reference, the tLST of the α -proteobacterium *Methylobacillus flagellatus* KT is depicted. The δ -proteobacterium *Desulfovibrio carbinolicus* DSM 3852 encodes a *clpG_{GI}* gene member [DESCAR in **(A)**] of the core *clpG_{GI}* family closely related to the *P. aeruginosa* clone C *clpG_{GI}* gene. On the other hand, *Mixta calida* DE0300 contains a *clpG_{GI}* member [MIXCAL in **(A)**] closely related to most *K. pneumoniae* and *E. coli* *clpG_{GI}* genes. *Pseudomonas putida* str. Idaho (PSEPUT), *Siccirubricoccus phaeus* 1–3 (Siccirubricoccus), *Oseanicella actignis* DSM 22673 (OSEACT) and *Methylobacterium* CCH7-A2 (METBAC) encode representatives of the subfamilies ClpG3_{GI}, a yet undefined subfamily, ClpG2_{GI} and ClpH families, respectively. *ClpG_{GI}* of *Halomonas desiderata* FB2 is 97.9% identical to *clpG_{GI}* of *P. aeruginosa* SG17M. Gene rearrangements and insertion of novel genes occur within tLST. The pathogen-related archetypical composite tLST as depicted in **Figure 1** is only present in a subgroup of isolates. While *D. carbinolicus* and γ -proteobacteria *M. calida* possess a close to archetypical tLST, members of the α -proteobacteria such as *O. actignis* DSM 22673, *Methylobacterium* CCH7-A2 and *S. phaeus* 1–3, but also γ -proteobacteria *P. putida* str. Idaho and *H. desiderata* FB2 contain mostly the protein homeostasis part of the tLST gene cluster with an occasional expansion of genes for small heat shock proteins. The gene product of *yfdX3* from *O. actignis* DSM22673 is distinct from *yfdX1* and *yfdX2* and most closely related to a *Paracoccus* representative. As the environmental source, *S. phaeus* 1–3 has been isolated from oil soil, *D. desiderata* from water and *O. actignis* DSM 22673 from the water of a hot spring. The isolation of *Methylobacterium* CCH7-A2 from a hospital shower hose biofilm and *H. desiderata* FB2 and *M. flagellatus* KT from a sewage treatment plant suggests opportunities for horizontal transmission to other microorganisms. *P. putida* str. Idaho is a unique organic solvent tolerant strain. The gene arrangement is centered around the most conserved *clpG_{GI}* gene. Other colored genes possess core genome or novel functionality, gray colored genes are related to transposition or phage function. The graph reports the G + C content at a window of ± 10 nucleotides (red, above average; blue, below average). The figure has been drawn with Easyfig 2.2 (Sullivan et al., 2011). **(C)** As ClpG_{GI}, the ClpG2_{GI}, ClpG3_{GI}, and ClpH proteins are characterized by an AAA +1 ATPase domain-M-domain-AAA +2 ATPase domain structure with conserved sequence motifs Walker A and B nucleotide binding motifs, the conserved pore loop residue tyrosine (Y), sensor 1 motifs with conserved threonine (T) and asparagine (N) and sensor 2 motif arginine (R), but possess distinct N (N2 and N1)- and C (CTE) -terminal domains. The alignment was created with ClustalX 2.1 (Aiyar, 2000) and visualized with ESPript 3 (Gross et al., 2007).

Kefb_{GI} is predicted to function as potassium-proton antiporter; this ion exchange acidifies the cytoplasm at alkaline conditions (Elmore et al., 1990). The core genome KefB of *E. coli*, which is 31% identical to Kefb_{GI}, is inhibited by cytoplasmic glutathione and activated by glutathione adducts (Roosild et al., 2010). In accordance with a predicted K⁺/H⁺ antiport activity, the *orf13* gene product Kefb_{GI} protected *E. coli* against alkaline pH in presence of chlorine, which depletes cytoplasmic glutathione, but not at solely alkaline pH (Zhu et al., 2021).

OCCURRENCE OF tLST OR ITS COMPONENTS IN PROTEOBACTERIA AND OTHER GRAM-NEGATIVE BACTERIA

The tLST with the core genes *dna-hsp20-clpG* encoded in genomes of beta- and gamma-Proteobacteria is generally flanked by mobile elements such as transposases or phage derived

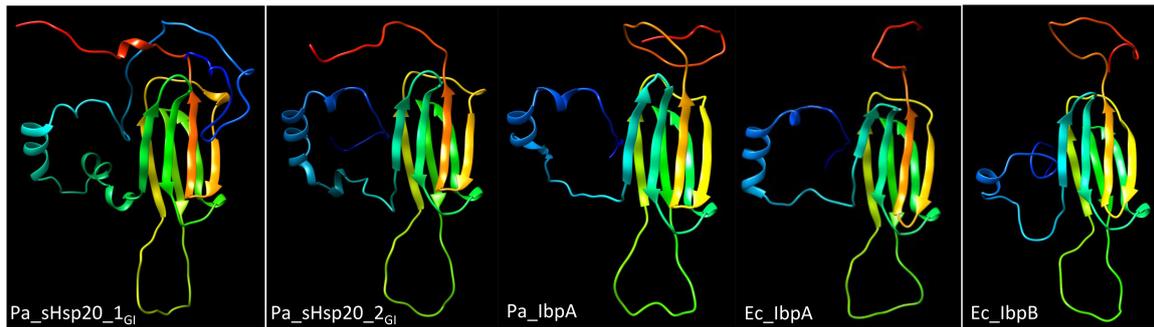
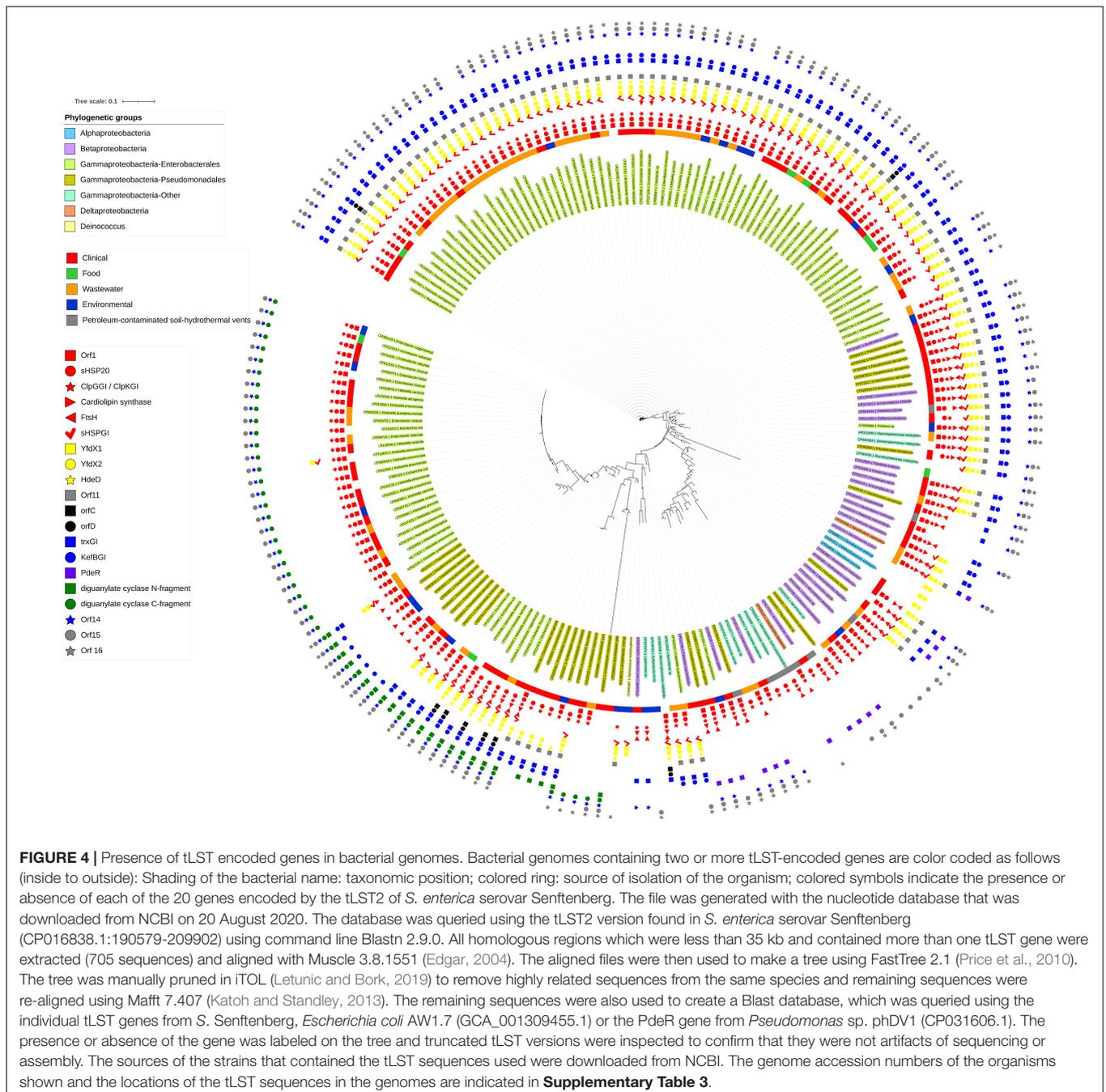


FIGURE 3 | Structural models of sHSP proteins show distinct structures for tLST island and core genome sHsp20 proteins. From left to right, *P. aeruginosa* sHsp20_{1GI}, *P. aeruginosa* sHsp20_{2GI}, *P. aeruginosa* core genome lbpA, *E. coli* core genome lbpA and *E. coli* core genome lbpB structural model of the monomer. All sHsp20 proteins possess a core of seven antiparallel β -strands that are characteristic for sHsp20 proteins including the paralogous α crystallin reference protein with distinct N-, intramolecular and C-terminal loops. The models were created with Phyre2 (Kelley et al., 2015) based on the structure of the *Triticum aestivum* HSP16.9 protein (1GME chain A; van Montfort et al., 2001) as the best-fit model.

genes (Mercer et al., 2015; Lee et al., 2016). Different insertion elements such as IS5, IS3, and Tn7 elements can flank the tLST (**Figure 2B**), although the data on complete genomes are too scarce to allow a systematic analysis. The genomic island can be either plasmid or chromosomally encoded; in addition, bacterial genomes can harbor more than one genetically distinct tLST island on the chromosome and on plasmids (Boll et al., 2017; Nguyen et al., 2017; Kamal et al., 2021). The >80% conservation of the tLST on the DNA level throughout even distantly related species suggests a singular and relatively recent source of island mobilization. The G + C content of the tLST is also relatively constant at approximately 61%, whereas the core genomes of its hosts such as *P. aeruginosa*, *E. coli*, *Aeromonas*, *Stenotrophomonas*, and *Acinetobacter* spp. can range from 39 to 70%. This further supports the hypothesis of promiscuous horizontal transfer of the genomic island (**Supplementary Table 2**). However, current data is insufficient to reliably indicate the origin of the genomic island.

The initial description of the archetypical composite tLST in beta- and gamma-Proteobacteria were based on the ~19 kbp tLSTa (previously: TLPQC) in the water isolate *P. aeruginosa* SG17M, which lacks the cell envelope stress module and adjacent up- and downstream genes (**Figure 1**; Lee et al., 2015) and the ~15 kbp tLST1 (previously: LHR1) in food-derived *E. coli*, which lacks the cardiolipin synthase and FtsH_{GI} (Mercer et al., 2015). Since 2015, additional tLST variants were described which include insertions in the oxidative stress module and were termed tLST2 (previously: LHR2; Boll et al., 2017; Nguyen et al., 2017). Previous database searches suggested that tLST variants are present mainly in gamma- and beta-Proteobacteria (Mercer et al., 2015). Among the *Enterobacteriales*, the tLST has been found virtually exclusive to the *Enterobacteriaceae*. The *Enterobacteriaceae* include environmental or insect-associated bacteria, organisms that are associated with plants but also persist in the intestine of vertebrates and are also of relevance as nosocomial pathogens, e.g., *Klebsiella*, *Enterobacter*, and *Cronobacter* species, and

vertebrate- or human-adapted pathogens such as *Shigella* species and *S. enterica* (Adeolu et al., 2016). Remarkably, the tLST has not been detected in insect associated *Enterobacteriaceae* e.g., *Trabulsilla* species (Mercer et al., 2015); and is also absent in the human pathogenic *Shigella* species, *S. enterica*, Shiga-toxin producing *E. coli* and the pandemic ST131 *E. coli* clone. The tLST occurs, however, in non-pathogenic strains of *E. coli* as well as *Klebsiella* and *Cronobacter* species (Mercer et al., 2015; Wang et al., 2020; Kamal et al., 2021). Beyond beta- and gamma-Proteobacteria an archetypical composite island is present, for example, in the α -proteobacterium *Methylobacillus flagellatus* KT isolated from a sewage plant (**Figure 2B**). As a hallmark, a sigma 24 like transcription factor is encoded by the island, but an integrated KefB H⁺/K⁺ transporter is missing (Gajdosova et al., 2011). We further interrogated the distribution of tLST variants by nucleotide BLAST analysis against the NCBI database in August 2020, using each of the genes encoded in the tLST1 or tLST2 as BLAST queries (**Figure 4**). Genomes that are deposited in the NCBI database predominantly originate from organisms that relate to human activity, particularly human pathogens, and thus do not allow reliable quantification of the distribution of the tLST. To address sampling bias, the NCBI database was used as a qualitative resource and only one sequence variant of the tLST for each bacterial species was randomly chosen (**Figure 4**). Sequences that include tLST-encoded proteins were predominantly recovered from organisms of the beta- and gamma-Proteobacteria, but alpha- and delta-Proteobacteria and *Deinococcus* of the class *Deinococci* were also represented (**Figures 2, 4**). Strains encoding for the tLST were isolated not only from anthropogenic sources, i.e., clinical sources, food or food-processing facilities, waste water, and soils contaminated by metal mining or oil extraction, but also from environmental sources including fresh water and hydrothermal vents where the tLST or major components thereof can also reside on a plasmid (**Figures 2, 4** and unpublished data citation). Strong selective pressure for transmission and maintenance of the tLST therefore also exists in habitats that are seemingly not impacted by human activity (**Figures 2B, 4**). Irrespective of sampling



bias for genomes deposited in the NCBI database, the high frequency of tLST-positive nosocomial pathogens in conjunction with the virtual absence in human-adapted pathogens such as *Shigella*, *Salmonella* Typhi, and toxin-producing intestinal organisms such as STEC provides evidence to prior suggestions that the tLST does not contribute to the ecological fitness of these pathogens, but can increase virulence or persistence of commensal *E. coli* and organisms in hospitals that are predominantly of environmental origin, but also opportunistic or nosocomial pathogens (Lee et al., 2020; Wang et al., 2020; Kamal et al., 2021).

CYCLIC di-GMP SECOND MESSENGER SIGNALING CONTRIBUTES TO tLST ISLAND VARIABILITY

The tLST1 that encodes for *orf1* (*dna*), *shsp20*_{1GI}, *clpK*_{GI}, *shsp20*_{2GI}, *yfdX1*, *yfdX2*, *hdeD*_{GI}, *orf11*, *trxB*_{GI}, *kefB*_{GI}, *orf14*, *orf15*, and *degP*_{GI} has been detected only in *Enterobacteriaceae* (Figure 4 position 11:00 to 2:00). Some of these tLST1 variants include *ftsH*_{GI} and individual sequences exclude genes encoded by the oxidative stress module and/or the cell envelope stress

module (**Figure 4**). The tLSTa (**Figure 1**) is represented mainly by sequences retrieved from beta-Proteobacteria including *Burkholderia* and *Achromobacter* species (**Figure 4** at position 2:30). Substantial variability in gene content and amino acid sequence of gene products of the archetypical tLST is found. Several sequences in beta- and gamma-Proteobacteria that are closely related to the tLSTa lack genes of the cell envelope stress module (**Figure 4** at position 3:30). The tLST2 (previously: LHR2) which includes the cardiolipin synthase, *ftsH_{GI}*, *orfC*, *orfD*, and a diguanylate cyclase has also been exclusively represented by *Enterobacteriaceae* including *E. coli*, *Klebsiella* and *Citrobacter* species and the only strain of *Salmonella* that even harbors two tLST islands, *S. enterica* serovar Senftenberg ATCC 43845 (**Figure 4** at position 7:00, Nguyen et al., 2017). Finally, multiple *Enterobacteriaceae* encode a tLST variant termed truncated tLST or tLSTt which encodes only the MerR-like regulator Orf1 (Dna), a putative Dna binding protein; sHSP20_{1GI} and ClpG_{GI} of the protein homeostasis module, the C-fragment of the diguanylate cyclase, *orf14*, and *orf15* of the oxidative stress module (**Figure 4** at position 8:00 to 10:00). Further more, although tLST gene products are highly conserved, their nevertheless present substantial sequence variability can lead to initial diversification of functional properties and thus provide a study subject for protein evolution in vivo (Kamal et al., 2021).

Genes whose products function as turnover enzymes for the bacterial second messenger cyclic di-GMP can be inserted or replace core tLST genes. For example, tLST variants of *E. coli* carry a disrupted *orf14*, but encode *orfE*, a predicted diguanylate cyclase that has a putative function in starvation survival (Boll et al., 2017), while in another tLST variant from *P. aeruginosa* clone C the gene for the potassium-proton antiporter KefB_{GI} is replaced by a putative diguanylate cyclase-phosphodiesterase protein (Lee et al., 2020). Cyclic di-GMP is a ubiquitous second messenger in bacteria that directs the fundamental life style transition between sessility (biofilm formation) and motility as well as between chronic and acute infections (Römling et al., 2013). Recently, the TdcA gene product inserted downstream of the *kefB_{GI}* antiporter gene in *P. aeruginosa* was characterized as a thermo-responsive diguanylate cyclase (Almblad et al., 2021). The response to temperature changes is mediated by a Per-Arnt-Sim (PAS) sensory domain, which led to high biofilm formation at body temperature. Increased biofilm formation in response to temperature might thereby aid chronic colonization by *P. aeruginosa*.

In the section above, we emphasized that different modules of the tLST differ in their protection against the various chemical or physical stressors, and protect different segments of bacterial cells (**Figure 1**). The phylogenetic analysis shown in **Figure 4** confirms that this tLST mediated protection against multiple stressors improves the ecological fitness of many organisms, and provides selective pressure for maintenance of those tLST encoded core genes that are present in tLSTa, tLST1, and tLST2 (**Figures 1, 4**). The high level of expression of tLST encoded proteins (Lee et al., 2015; Kamal et al., 2019; Li et al., 2020) likely also imposes substantial fitness cost, which is apparently offset by the increased resistance to heat, oxidative stress, and additional but yet unknown environmental insults, but can explain the

presence of the tLST in only distinct genetic backgrounds of the different species. The gene content of the tLST is altered, however, by deletions or insertions in many bacterial species (**Figure 4**), indicating that the genomic island is “customized” to match the selective pressures that bacterial species encounter in their respective ecological niches.

RELEVANCE OF THE tLST FOR NOSOCOMIAL PATHOGENS AND ANTIBIOTIC RESISTANCE

Is the tLST a relevant contributor to persistence of nosocomial pathogens in hospitals and aids their antibiotic resistance? In the section above, we have emphasized that organisms that adapted to a human intestinal pathogenic lifestyle, as is the case for Shiga-toxin producing *E. coli*, *Shigella* species, and *S. enterica*, rarely encode the tLST. Heat resistant and tLST-encoding commensal and meat derived strains of *E. coli* are predominantly of phylotype A, while urinary tract pathogens bearing virulence factors consistently exclude the tLST island consistent with the antivirulence function of the YfdX proteins (Wang et al., 2020; Kamal et al., 2021). The tLST island is frequently present in organisms that are found in environmental or plant-associated niches but can also be nosocomial or opportunistic pathogens (**Figure 4**; Struve and Krogfelt, 2004; Lee et al., 2017, 2020; Wang et al., 2020). While this pattern is consistent, the mechanisms for tLST maintenance are unknown. However, biofilms of *K. pneumoniae* survive significantly better upon ClpGgi/ClpK production (Bojer et al., 2011) and an endoscope associated outbreak was caused by *P. aeruginosa* clone C strains which usually bear tLST (Fernández-Cuenca et al., 2020). The presence of the tLST in nosocomial pathogens may therefore reduce dispersal limitation, i.e., improve the ability of organisms to survive in water or on surfaces after sanitation, or it may relate to increased virulence upon infection by improved biofilm formation, which can be triggered by thermosensitive cyclic di-GMP modules, or through resistance to antimicrobial compounds generated by immune cells of the host. While a direct contribution of the tLST to antibiotic resistance has not been demonstrated experimentally, the predicted function of several tLST encoded proteins suggests a contribution to the resistance to therapeutic antibiotics (**Figure 4**).

CONCLUSION AND PERSPECTIVES

The tLST is a composite horizontally transferred genetic element that provides exceptional resistance against a number of clinically and environmentally relevant chemical and physical stressors including heat and oxidative stress. Although those stresses can be occasionally present in a number of ecological niches, selection for tLST positive strains of organisms such as *E. coli*, *P. aeruginosa*, *K. pneumoniae*, and *C. sakazakii* is promoted by human activities including thermal interventions in food processing, hospital-based disinfection procedures and common water and waste water sanitation procedures indicating the

need for widening the substrate spectrum for proteases and disaggregases by horizontal transfer of xenologs. The sequence diversity of tLST variants in different bacterial backgrounds implies either adaptation to the host conditions and/or multiple unknown stressors that select for maintenance of the tLST in natural habitats remain unknown, as is the original host for the archetypical tLST from which this composite genomic island was disseminated. Although core activities are well known, the distinct physiological function of multiple tLST encoded proteins remains to be elucidated. In particular, it is unclear why tLST elements can encode up to three distinct small heat shock proteins (Figure 2B), why multiple variants of the tLST integrate cyclic di-GMP turnover proteins and how these proteins function in stress resistance.

Irrespective of our ability to answer these questions, the tLST appears to contribute to the success of multi-drug-resistant bacteria as nosocomial pathogens and should be considered in efforts to reduce their spread and persistence. Conversely, the demonstrated ability of tLST-encoded proteins to prevent protein aggregation, or to restore aggregated proteins to their native state may be an asset in biotechnological applications that aim to achieve high-yield production of heterologously expressed proteins (Guzzo, 2012). Those characteristics might be extended to clinical applications as ClpG also prevents toxicity of protein substrates involved in neurodegenerative diseases (March et al., 2020). Moreover, while core genome heat shock proteins are often essential genes, tLST encoded heat shock proteins are accessory

genes and thus provide an excellent tool to study the role of protein homeostasis in bacterial stress resistance and ecology.

AUTHOR CONTRIBUTIONS

SK, DS, MG, and UR: conceptualization. SK, MG, and UR: writing. SK, ZW, DS, and UR: data analysis and preparation of figures. SK, DS, ZW, MG, and UR: review and editing of the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

SK has been partially supported by a scholarship from Future University in Kairo, Egypt. UR acknowledges funding from the Swedish Research Council for Natural Sciences and Engineering (2017-04465) and the Karolinska Institutet. MG acknowledges support from the Canada Research Chairs Program. DS and ZW are supported by project funding from the Alberta Agriculture Funding Consortium.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2021.696522/full#supplementary-material>

REFERENCES

- Adeolu, M., Alnajjar, S., Naushad, S., and Gupta, R. S. (2016). Genome-based phylogeny and taxonomy of the 'Enterobacteriales': proposal for *Enterobacteriales* ord. nov. divided into the families *Enterobacteriaceae*, *Erwiniaceae* fam. nov., *Pectobacteriaceae* fam. nov., *Yersiniaceae* fam. nov., *Hafniaceae* fam. nov., *Morganellaceae* fam. nov., and *Budviciaceae* fam. nov. *Int. J. Syst. Evol. Microbiol.* 66, 5575–5599. doi: 10.1099/ijsem.0.001485
- Aiyar, A. (2000). The use of CLUSTAL W and CLUSTAL X for multiple sequence alignment. *Methods Mol. Biol.* 132, 221–241.
- Almblad, H., Randall, T. E., Liu, F., Leblanc, K., Groves, R. A., Kittichotirat, W., et al. (2021). Bacterial cyclic diguanylate signaling networks sense temperature. *Nat. Commun.* 12:1986. doi: 10.1038/s41467-021-22176-2
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410.
- Banuett, F., Hoyt, M. A., McFarlane, L., Echols, H., and Herskowitz, I. (1986). *hflB*, a new *Escherichia coli* locus regulating lysogeny and the level of bacteriophage lambda cII protein. *J. Mol. Biol.* 187, 213–224. doi: 10.1016/0022-2836(86)90229-9
- Basta, D. W., Angeles-Albores, D., Spero, M. A., Ciemniecki, J. A., and Newman, D. K. (2020). Heat-shock proteases promote survival of *Pseudomonas aeruginosa* during growth arrest. *Proc. Natl. Acad. Sci. U.S.A.* 117, 4358–4367.
- Berendsen, E. M., Boekhorst, J., Kuipers, O. P., and Wells-Bennik, M. H. J. (2016). A mobile genetic element profoundly increases heat resistance of bacterial spores. *ISME J.* 10, 2633–2642. doi: 10.1038/ismej.2016.59
- Bergkessel, M. (2020). Regulation of protein biosynthetic activity during growth arrest. *Curr. Opin. Microbiol.* 57, 62–69.
- Bojer, M. S., Krogfelt, K. A., and Struve, C. (2011). The newly discovered ClpK protein strongly promotes survival of *Klebsiella pneumoniae* biofilm subjected to heat shock. *J. Med. Microbiol.* 60, 1559–1561. doi: 10.1099/jmm.0.032698-0
- Bojer, M. S., Struve, C., Ingmer, H., Hansen, D. S., and Krogfelt, K. A. (2010). Heat resistance mediated by a new plasmid encoded Clp ATPase, ClpK, as a possible novel mechanism for nosocomial persistence of *Klebsiella pneumoniae*. *PLoS One* 5:e15467. doi: 10.1371/journal.pone.0015467
- Boll, E. J., Marti, R., Hasman, H., Overballe-Petersen, S., Stegger, M., Ng, K., et al. (2017). Turn up the heat - food and clinical *Escherichia coli* isolates feature two transferrable loci of heat resistance. *Front. Microbiol.* 8:579. doi: 10.3389/fmicb.2017.00579
- Carmel-Harel, O., and Storz, G. (2000). Roles of the glutathione- and thioredoxin-dependent reduction systems in the *Escherichia coli* and *Saccharomyces cerevisiae* thioredoxin-dependent reduction systems responses to oxidative stress. *Annu. Rev. Microbiol.* 54, 439–461.
- Dlusskaya, E. A., McMullen, L. M., and Gänzle, M. G. (2011). Characterization of an extremely heat-resistant *Escherichia coli* obtained from a beef processing facility. *J. Appl. Microbiol.* 110, 840–849. doi: 10.1111/j.1365-2672.2011.04943.x
- Edgar, R. C. (2004). MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinform.* 5:113. doi: 10.1186/1471-2105-5-113
- Elliott, S. J., Wainwright, L. A., McDaniel, T. K., Jarvis, K. G., Deng, Y., Lai, L.-C., et al. (2002). The complete sequence of the locus of enterocyte effacement (LEE) from enteropathogenic *Escherichia coli* E2348/69. *Mol. Microbiol.* 28, 1–4. doi: 10.1046/j.1365-2958.1998.00783.x
- Elmore, M. J., Lamb, A. J., Ritchie, G. Y., Douglas, R. M., Munro, A., Gajewska, A., et al. (1990). Activation potassium efflux from *Escherichia coli* by glutathione metabolites. *Mol. Microbiol.* 4, 405–412. doi: 10.1111/j.1365-2958.1990.tb00607.x
- Fernández-Cuenca, F., López-Cerero, L., Cabot, G., Oliver, A., López-Méndez, J., Recacha, E., et al. (2020). Nosocomial outbreak linked to a flexible gastrointestinal endoscope contaminated with an amikacin-resistant ST17 clone of *Pseudomonas aeruginosa*. *Eur. J. Clin. Microbiol. Infect. Dis.* 39, 1837–1844. doi: 10.1007/s10096-020-03915-7

- Gajdosova, J., Benedikovicova, K., Kamodyova, N., Tothova, L., Kaclikova, E., Stuchlik, S., et al. (2011). Analysis of the DNA region mediating increased thermotolerance at 58°C in *Cronobacter* sp. and other enterobacterial strains. *Antonie Leeuwenhoek* 100, 279–289.
- Goepfert, J. M., and Biggie, R. A. (1968). Heat resistance of *Salmonella typhimurium* and *Salmonella senftenberg* 775W in milk chocolate. *Appl. Microbiol.* 16, 1939–1940. doi: 10.1128/aem.16.12.1939-1940.1968
- Gross, L., Bourgouin, L., Hale, A. J., and Mühlhaus, H.-B. (2007). Interface modeling in incompressible media using level sets in Escript. *Phys. Earth Planet. Inter.* 163, 23–34.
- Guragain, M., Brichta-Harhay, D. M., Bono, J. L., and Bosilevac, J. M. (2021). Locus of Heat Resistance (LHR) in meat-borne *Escherichia coli*: screening and genetic characterization. *Appl. Environ. Microbiol.* 87:e02343-20. doi: 10.1128/aem.02343-20
- Guzzo, J. (2012). Biotechnical applications of small heat shock proteins from bacteria. *Int. J. Biochem. Cell Biol.* 44, 1698–1705. doi: 10.1016/j.biocel.2012.06.007
- Herold, S., Karch, H., and Schmidt, H. (2004). Shiga toxin-encoding bacteriophages - genomes in motion. *Int. J. Med. Microbiol.* 294, 115–121. doi: 10.1016/j.ijmm.2004.06.023
- Hiraoka, S., Matsuzaki, H., and Shibuya, I. (1993). Active increase in cardiolipin synthesis in the stationary growth phase and its physiological significance in *Escherichia coli*. *FEBS Lett.* 336, 221–224. doi: 10.1016/0014-5793(93)80807-7
- Hueppe, F. (1884). Ueber die Zersetzungen der Milch und die biologischen Grundlagen der Gährungsphysiologie (Fortsetzung aus No. 48.). *Dtsch. Medizinische Wochenschrift* 10, 796–797. doi: 10.1055/s-0028-1143388
- Kamal, S. M., Cimdins-Ahne, A., Lee, C., Li, F., Martín-Rodríguez, A. J., Seferbekova, Z., et al. (2021). A recently isolated human commensal *Escherichia coli* ST10 clone member mediates enhanced thermotolerance and tetrathionate respiration on a P1 phage-derived IncY plasmid. *Mol. Microbiol.* 115, 255–271. doi: 10.1111/mmi.14614
- Kamal, S. M., Rytbke, M. L., Nimtz, M., Sperlein, S., Giske, C., Trèek, J., et al. (2019). Two FtsH proteases contribute to fitness and adaptation of *Pseudomonas aeruginosa* clone C strains. *Front. Microbiol.* 10:1372. doi: 10.3389/fmicb.2019.01372
- Katoh, K., and Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–780. doi: 10.1093/molbev/mst010
- Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., and Sternberg, M. J. E. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* 10, 845–858. doi: 10.1038/nprot.2015.053
- Kornitzer, D., Teff, D., Altuvia, S., and Oppenheim, A. B. (1991). Isolation, characterization, and sequence of an *Escherichia coli* heat shock gene, *htpX*. *J. Bacteriol.* 173, 2944–2953.
- Krol, J. E., Hall, D. C. Jr., Balashov, S., Pastor, S., Sibert, J., McCaffrey, J., et al. (2019). Genome rearrangements induce biofilm formation in *Escherichia coli* C - an old model organism with a new application in biofilm research. *BMC Genom.* 20:767. doi: 10.1186/s12864-019-6165-4
- Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* 33, 1870–1874. doi: 10.1093/molbev/msw054
- Lan, R., Lumb, B., Ryan, D., and Reeves, P. R. (2001). Molecular evolution of large virulence plasmid in *Shigella* clones and enteroinvasive *Escherichia coli*. *Infect. Immun.* 69, 6303–6309. doi: 10.1128/IAI.69.10.6303-6309.2001
- Lee, C., Franke, K. B., Kamal, S. M., Kim, H., Lünsdorf, H., Jäger, J., et al. (2018). Stand-alone ClpG disaggregase confers superior heat tolerance to bacteria. *Proc. Natl. Acad. Sci. U.S.A.* 115, E273–E282. doi: 10.1073/pnas.1712051115
- Lee, C., Klockgether, J., Fischer, S., Trcek, J., Tümmeler, B., and Römling, U. (2020). Why? - Successful *Pseudomonas aeruginosa* clones with a focus on clone C. *FEMS Microbiol. Rev.* 44, 740–762. doi: 10.1093/femsre/fuaa029
- Lee, C., Wigren, E., Lünsdorf, H., and Römling, U. (2016). Protein homeostasis - more than resisting a hot bath. *Curr. Opin. Microbiol.* 30, 147–154. doi: 10.1016/j.mib.2016.02.006
- Lee, C., Wigren, E., Trèek, J., Peters, V., Kim, J., Hasni, M. S., et al. (2015). A novel protein quality control mechanism contributes to heat shock resistance of worldwide-distributed *Pseudomonas aeruginosa* clone C strains. *Environ. Microbiol.* 17, 4511–4526. doi: 10.1111/1462-2920.12915
- Lee, C. R., Lee, J. H., Park, M., Park, K. S., Bae, I. K., Kim, Y. B., et al. (2017). Biology of *Acinetobacter baumannii*: pathogenesis, antibiotic resistance mechanisms, and prospective treatment options. *Front. Cell. Infect. Microbiol.* 7:55. doi: 10.3389/fcimb.2017.00055
- Lee, G. J., Roseman, A. M., Saibil, H. R., and Vierling, E. (1997). A small heat shock protein stably binds heat-denatured model substrates and can maintain a substrate in a folding-competent state. *EMBO J.* 16, 659–671.
- Lee, H. S., Lee, S., Kim, J. S., Lee, H. R., Shin, H. C., Lee, M. S., et al. (2019). Structural and physiological exploration of *Salmonella* Typhi YfdX uncovers its dual function in bacterial antibiotic stress and virulence. *Front. Microbiol.* 9:3329. doi: 10.3389/fmicb.2018.03329
- Letunic, I., and Bork, P. (2019). Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res.* 47, W256–W259. doi: 10.1093/nar/gkz239
- Li, H., and Gänzle, M. (2016). Some like it hot: heat resistance of *Escherichia coli* in food. *Front. Microbiol.* 7:01763. doi: 10.3389/fmicb.2016.01763
- Li, H., Mercer, R., Behr, J., Heinzlmeier, S., McMullen, L. M., Vogel, R. F., et al. (2020). Heat and pressure resistance in *Escherichia coli* relates to protein folding and aggregation. *Front. Microbiol.* 11:111. doi: 10.3389/fmicb.2020.00111
- Lin, X. B., Lohans, C. T., Duar, R., Zheng, J., Vederas, J. C., Walter, J., et al. (2015). Genetic determinants of reutericyclin biosynthesis in *Lactobacillus reuteri*. *Appl. Environ. Microbiol.* 81, 2032–2041. doi: 10.1128/AEM.03691-14
- Lipinska, B., Zylicz, M., and Georgopoulos, C. (1990). The HtrA (DegP) protein, essential for *Escherichia coli* survival at high temperatures, is an endopeptidase. *J. Bacteriol.* 172, 1791–1797. doi: 10.1128/jb.172.4.1791-1797.1990
- Liu, C. J., Lin, C. T., Chiang, J. D., Lin, C. Y., Tay, Y. X., and Fan, L. C. (2019). RcsB regulation of the YfdX-mediated acid stress response in *Klebsiella pneumoniae* CG43S3. *PLoS One* 14:e0212909. doi: 10.1371/journal.pone.0212909
- Liu, Y., Gill, A., McMullen, L., and Gänzle, M. (2015). Variation in heat and pressure resistance of verotoxigenic and nontoxicogenic *Escherichia coli*. *J. Food Prot.* 78, 111–120. doi: 10.4315/0362-028X.JFP-14-267
- MacLean, M. J., Ness, L. S., Ferguson, G. P., and Booth, I. R. (1998). The role of glyoxalase I in the detoxification of methylglyoxal and in the activation of the KefB K+ efflux system in *Escherichia coli*. *Mol. Microbiol.* 27, 563–571.
- March, Z. M., Sweeney, K., Kim, H., Yan, X., Castellano, L. M., Jackrel, M. E., et al. (2020). Therapeutic genetic variation revealed in diverse Hsp104 homologs. *eLife* 9:e57457.
- Marti, R., Muniesa, M., Schmid, M., Ahrens, C. H., Naskova, J., and Hummerjohann, J. (2016). Heat-resistant *Escherichia coli* as potential persistent reservoir of extended-spectrum β -lactamases and Shiga toxin-encoding phages in dairy. *J. Dairy Sci.* 99, 8622–8632.
- Mates, A. K., Sayed, A. K., and Foster, J. W. (2011). Products of the *Escherichia coli* acid fitness island attenuate metabolite stress at extremely low pH and mediate a cell density-dependent acid resistance. *J. Bacteriol.* 189, 2759–2768. doi: 10.1128/JB.01490-06
- Mecas, J., Rouviere, P. E., Erickson, J. W., Donohue, T. J., and Gross, C. A. (1993). The activity of $\sigma(E)$, an *Escherichia coli* heat-inducible σ -factor, is modulated by expression of outer membrane proteins. *Genes Dev.* 7, 2618–2628. doi: 10.1101/gad.7.12b.2618
- Mercer, R. G., Nguyen, O., Ou, Q., McMullen, L., and Gänzle, M. G. (2017a). Functional analysis of genes encoded by the locus of heat resistance in *Escherichia coli*. *Appl. Environ. Microbiol.* 83:e01400-17. doi: 10.1128/AEM.01400-17
- Mercer, R. G., Walker, B. D., Yang, X., McMullen, L. M., and Gänzle, M. G. (2017b). The locus of heat resistance (LHR) mediates heat resistance in *Salmonella enterica*, *Escherichia coli* and *Enterobacter cloacae*. *Food Microbiol.* 64, 96–103. doi: 10.1016/j.fm.2016.12.018
- Mercer, R. G., Zheng, J., Garcia-Hernandez, R., Ruan, L., Gänzle, M. G., and McMullen, L. M. (2015). Genetic determinants of heat resistance in *Escherichia coli*. *Front. Microbiol.* 6:932. doi: 10.3389/fmicb.2015.00932
- Mogk, A., Bukau, B., and Kampinga, H. H. (2018). Cellular handling of protein aggregates by disaggregation machines. *Mol. Cell* 69, 214–226. doi: 10.1016/j.molcel.2018.01.004
- Mogk, A., Tomoyasu, T., Goloubinoff, P., Rudiger, S., Roder, D., Langen, H., et al. (1999). Identification of thermolabile *Escherichia coli* proteins: prevention

- and reversion of aggregation by DnaK and ClpB. *EMBO J.* 18, 6934–6949. doi: 10.1093/emboj/18.24.6934
- Motohashi, K., Watanabe, Y., Yohda, M., and Yoshida, M. (1999). Heat-inactivated proteins are rescued by the DnaK-J-GrpE set and ClpB chaperones. *Proc. Natl. Acad. Sci. U.S.A.* 96, 7184–7189. doi: 10.1073/pnas.96.13.7184
- Nazarowec-White, M., and Farber, J. M. (1997). *Enterobacter sakazakii*: a review. *Int. J. Food Microbiol.* 34, 103–113. doi: 10.1016/S0168-1605(96)01172-5
- Nguyen, O. T. H. (2019). *Genetic Regulation of the Locus of Heat Resistance in Escherichia coli*. Edmonton: University of Alberta.
- Nguyen, S. V., Harhay, G. P., Bono, J. L., Smith, T. P. L., and Harhay, D. M. (2017). Genome sequence of the thermotolerant foodborne pathogen *Salmonella enterica* serovar Senftenberg ATCC 43845 and phylogenetic analysis of loci encoding increased protein quality control mechanisms. *mSystems* 2:e0190-16. doi: 10.1128/msystems.00190-16
- Nishijima, S., Asami, Y., Uetake, N., Yamagoe, S., Ohta, A., and Shibuya, I. (1988). Disruption of the *Escherichia coli* *cls* gene responsible for cardiolipin synthesis. *J. Bacteriol.* 170, 775–780.
- Nishimura, K., Kato, Y., and Sakamoto, W. (2016). Chloroplast proteases: updates on proteolysis within and across suborganellar compartments. *Plant Physiol.* 171, 2280–2293.
- Pleitner, A., Zhai, Y., Winter, R., Ruan, L., McMullen, L. M. L. M., and Gänzle, M. G. M. G. (2012). Compatible solutes contribute to heat resistance and ribosome stability in *Escherichia coli* AW1.7. *Biochim. Biophys. Acta Proteins Proteom.* 1824, 1351–1357. doi: 10.1016/j.bbapap.2012.07.007
- Price, M. N., Dehal, P. S., and Arkin, A. P. (2010). FastTree 2 - approximately maximum-likelihood trees for large alignments. *PLoS One* 5:e9490. doi: 10.1371/journal.pone.0009490
- Quigley, L., O'Sullivan, O., Stanton, C., Beresford, T. P., Ross, R. P., Fitzgerald, G. F., et al. (2013). The complex microbiota of raw milk. *FEMS Microbiol. Rev.* 37, 664–698. doi: 10.1111/1574-6976.12030
- Römling, U., Galperin, M. Y., and Gomelsky, M. (2013). Cyclic di-GMP: the first 25 years of a universal bacterial second messenger. *Microbiol. Mol. Biol. Rev.* 77, 1–52. doi: 10.1128/mmbbr.00043-12
- Roosild, T. P., Castronovo, S., Healy, J., Miller, S., Pliotas, C., Rasmussen, T., et al. (2010). Mechanism of ligand-gated potassium efflux in bacterial pathogens. *Proc. Natl. Acad. Sci. U.S.A.* 107, 19784–19789. doi: 10.1073/pnas.1012716107
- Ruan, L., Pleitner, A., Gänzle, M. G., and McMullen, L. M. (2011). Solute transport proteins and the outer membrane protein NmpC contribute to heat resistance of *Escherichia coli* AW1.7. *Appl. Environ. Microbiol.* 77, 2961–2967. doi: 10.1128/AEM.01930-10
- Sakoh, M., Ito, K., and Akiyama, Y. (2005). Proteolytic activity of HtpX, a membrane-bound and stress-controlled protease from *Escherichia coli*. *J. Biol. Chem.* 280, 33305–33310. doi: 10.1074/jbc.M506180200
- Schmid, M., Iversen, C., Gontia, I., Stephan, R., Hofmann, A., Hartmann, A., et al. (2009). Evidence for a plant-associated natural habitat for *Cronobacter* spp. *Res. Microbiol.* 160, 608–614. doi: 10.1016/j.resmic.2009.08.013
- Shintani, M. (2017). The behavior of mobile genetic elements (MGEs) in different environments. *Biosci. Biotechnol. Biochem.* 81, 854–862. doi: 10.1080/09168451.2016.1270743
- Sklar, J. G., Wu, T., Kahne, D., and Silhavy, T. J. (2007). Defining the roles of the periplasmic chaperones SurA, Skp, and DegP in *Escherichia coli*. *Genes Dev.* 21, 2473–2484. doi: 10.1101/gad.1581007
- Sriramulu, D. D., Nimtz, M., and Römling, U. (2005). Proteome analysis reveals adaptation of *Pseudomonas aeruginosa* to the cystic fibrosis lung environment. *Proteomics* 5, 3712–3721. doi: 10.1002/pmic.200401227
- Struve, C., and Krogfelt, K. A. (2004). Pathogenic potential of environmental *Klebsiella pneumoniae* isolates. *Environ. Microbiol.* 6, 584–590.
- Sullivan, M. J., Petty, N. K., and Beatson, S. A. (2011). Easyfig: a genome comparison visualizer. *Bioinformatics* 27, 1009–1010.
- Sun, Y., and MacRae, T. H. (2005). Small heat shock proteins: molecular structure and chaperone function. *Cell. Mol. Life Sci. C* 62, 2460–2476.
- Tekaia, F., and Gordon, S. V. (1999). Analysis of the proteome of *Mycobacterium tuberculosis* in silico. *Res. Institut. Pasteur. Tuberc. Lung Dis.* 79, 329–342.
- Tomoyasu, T., Gamer, J., Bukau, B., Kanemori, M., Mori, H., Rutman, A. J., et al. (1995). *Escherichia coli* FtsH is a membrane-bound, ATP-dependent protease which degrades the heat-shock transcription factor sigma 32. *EMBO J.* 14, 2551–2560.
- van Montfort, R. L. M., Basha, E., Friedrich, K. L., Slingsby, C., and Vierling, E. (2001). Crystal structure and assembly of a eukaryotic small heat shock protein. *Nat. Struct. Biol.* 8, 1025–1030.
- Vanlint, D., Mitchell, R., Bailey, E., Meersman, F., McMillan, P. F., Michiels, C. W., et al. (2011). Rapid acquisition of gigapascal-high-pressure resistance by *Escherichia coli*. *mBio* 2:e0130-10. doi: 10.1128/mBio.0130-10
- Wang, Z., Fang, Y., Zhi, S., Simpson, D. J., Gill, A., McMullen, L. M., et al. (2020). The locus of heat resistance confers resistance to chlorine and other oxidizing chemicals in *Escherichia coli*. *Appl. Environ. Microbiol.* 86:e02123-19. doi: 10.1128/AEM.02123-19
- Wang, Z., Li, P., Luo, L., Simpson, D. J., and Gänzle, M. G. (2018). *Daqu* fermentation selects for heat-resistant *Enterobacteriaceae* and bacilli. *Appl. Environ. Microbiol.* 84:e1483. doi: 10.1128/AEM.01483-18
- Wanner, B. L. (1986). Novel regulatory mutants of the phosphate regulon in *Escherichia coli* K-12. *J. Mol. Biol.* 191, 39–58. doi: 10.1016/0022-2836(86)90421-3
- Williams, C. H. Jr. (1995). Mechanism and structure of thioredoxin reductase from *Escherichia coli*. *FASEB J.* 9, 1267–1276.
- Zhang, P., Tran, F., Stanford, K., and Yang, X. (2020). Are antimicrobial interventions associated with heat-resistant *Escherichia coli* on meat? *Appl. Environ. Microbiol.* 86:e0512-20. doi: 10.1128/AEM.00512-20
- Zheng, X.-W., Tabrizi, M. R., Nout, M. J. R., and Han, B.-Z. (2011). *Daqu* - a traditional Chinese liquor fermentation starter. *J. Inst. Brew.* 117, 82–90. doi: 10.1002/j.2050-0416.2011.tb00447.x
- Zhi, S., Banting, G., Li, Q., Edge, T. A., Topp, E., Sokurenko, M., et al. (2016). Evidence of naturalized stress-tolerant strains of *Escherichia coli* in municipal wastewater treatment plants. *Appl. Environ. Microbiol.* 82, 5505–5518. doi: 10.1128/AEM.00143-16
- Zhu, T., Wang, Z., McMullen, L. M., Raivio, T., Simpson, D. J., and Gänzle, M. G. (2021). Contribution of the Locus of Heat Resistance to growth and survival of *Escherichia coli* at alkaline pH and at alkaline pH in the presence of chlorine. *Microorganisms* 9:701. doi: 10.3390/microorganisms9040701
- Zolkiewski, M. (1999). ClpB cooperates with DnaK, DnaJ, and GrpE in suppressing protein aggregation. A novel multi-chaperone system from *Escherichia coli*. *J. Biol. Chem.* 274, 28083–28086. doi: 10.1074/jbc.274.40.28083

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Kamal, Simpson, Wang, Gänzle and Römling. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.