Bacterial Stress Responses Lower mRNA-Protein Level Correlations

Sena Gizem Suer[1,2,3,7], Yi Yang Lim[1,2], Ganeshwari Dhurve[1,4], Rabia Şen[5,7], Jérôme Arnoux[1,2,3,7], Cemal Erdem[5,7], André Mateus[3,4,6,7] and Kemal Avican[1,2,3,7,8*]

[1]Department of Molecular Biology, Umeå University, Umeå, 90187, Sweden.

[2]Integrated Science Lab (IceLab), Umeå University, Umeå, 90187, Sweden.

[3]Umeå Centre for Microbial research (UCMR), Umeå University, Umeå, 90187, Sweden.

[4]Laboratory for Molecular Infection Medicine Sweden (MIMS), Umeå University, Umeå, 90187, Sweden.

[5]Department of Medical Biosciences, Science for Life Laboratory, Umeå University, Umeå, 90187, Sweden.

[6]Department of Chemistry, Umeå University, Umeå, 90187, Sweden.

[7] Science for Life Laboratory, Umeå University, Umeå, 90187, Sweden.

[8]Lead contact

[*]Correspondence: kemal.avican@umu.se (K.A.)

Abstract

Diverse bacterial pathogens have evolved complex regulatory mechanisms to adapt to various environmental stresses during infection. The uncertainty in mRNA-protein levels in response to environmental stressors complicates our understanding of bacterial physiology and their adaptation to stressful environments. To examine this issue, we have integrated transcriptomics and proteomics data on three human bacterial pathogens *Salmonella enterica* Typhimurium, *Yersinia pseudotuberculosis*, and *Staphylococcus aureus* under ten infection-relevant stress conditions. We observed positive correlations between mRNA and protein levels, which were decreased under different stress conditions. Essential genes exhibited higher expression levels with lower variation across the conditions and stronger mRNA-protein correlations compared to non-essential genes, highlighting their critical role in bacterial adaptability and survival. Moreover, we identified a substantial number of genes with stress-induced non-correlating mRNA-protein levels, particularly under conditions triggering strong stress responses. Particularly this level was dramatically lowered for osmotic stress specific genes affected by impaired translational activity under osmotic stress. Our findings highlight the prevalence of non-correlating mRNA-protein levels and the potential role of post-translational modifications in modulating protein levels in response to environmental stressors during infection. This study provides a comprehensive framework for integrating transcriptomics and proteomics data and identifies potential gene products that might significantly impact the ability of diverse bacterial pathogens to adapt to hostile infection environments.

Introduction

Bacterial pathogens have evolved complex mechanisms to adapt to different environmental stresses in the host such as nutrient limitations, temperature or pH shifts, and antimicrobial challenges during infection. They respond to deleterious effects of such stressors via stress response mechanisms, involving complex gene regulatory networks, for survival. Therefore, comprehensive understanding of bacterial stress response mechanisms is critical for development of novel therapeutic approaches to resolve bacterial infections. One useful approach is through investigation of global transcriptomes of diverse bacterial pathogens under a variety of stress conditions that mimic host environments. This provides information on regulation across diverse stress conditions that could be used in attempts to reveal complex gene regulations and identify key stress response genes. We previously performed a large-scale transcriptomic study of 32 human bacterial pathogens, PATHOgenex RNA Atlas[1], aiming to understand bacterial stress responses under 11 *in vivo* mimicking stress conditions. The study revealed a dynamic regulation of gene expression under different stress conditions and suggested a set of genes that has potential to be candidate as new antimicrobial targets. Co-PATHOgenex[2], the continuation of the PATHOgenex study, identifies stress-specific stimulons by generating gene co-expression modules. These stimulons represent sets of genes with unique expression patterns under specific stressors, providing insights into the stress adaptation to different stressors. Yet the main functional units shaping the cellular response are proteins and there is, generally, a weak correlation between mRNA and protein levels[3–11]. This highlights the complex nature of gene regulation and the diverse post-transcriptional factors affecting protein stability, and abundance[12].

Linking protein expression to gene expression can help to identify key proteins in response to specific stimuli or conditions and highlight mechanisms that drive bacterial adaptation, evasion of the immune system, and disease development[13–17]. This is crucial for developing more effective diagnostic tools, treatment strategies, and vaccines for bacterial infections[18,19]. Therefore, integrated transcriptomic and proteomic studies have emerged as powerful tools for understanding the complex molecular mechanisms underlying bacterial pathogenesis[15]. These studies measure mRNA and proteins levels from the same source and combine this information for comprehensive understanding of the interplay between pathogenic bacteria and their host environments, revealing gene products that are essential for bacterial survival, virulence, and resistance[13,20–22]. These efforts have provided valuable insights into the complex mechanisms that bacteria employ to adapt and survive under challenging conditions. Despite the numerous studies integrating transcriptome and proteome data, efforts to identify genes where mRNA and protein levels *do not* correlate are relatively limited[23,24]. Focusing on

such genes could offer valuable insight into the complex post-transcriptional and post-translational regulatory mechanisms that escape detection by analyses addressing correlated gene-expression patterns.

In this study, we employed an integrative transcriptomic and proteomic analysis of *Salmonella enterica* Typhimurium, *Yersinia pseudotuberculosis*, and *Staphylococcus aureus* under ten infection-relevant stress conditions to investigate mRNA-protein level correlations. We revealed that mRNA-protein correlations significantly varied under different stress conditions for the three species. Essential genes consistently showed higher mRNA and protein expression levels and stronger correlations, emphasizing their importance in bacterial survival. Stress conditions such as hypoxia and nutritional downshift exhibited the lowest correlations, suggesting that post-transcriptional and post-translational regulation play a significant role under these conditions. We identified plentiful genes that showed non-correlating mRNA-protein levels in the three species, as well as sets of genes that were species specific, particularly under conditions eliciting strong stress responses. Among all conditions tested, osmotic stress stood out as exhibiting the most pronounced mRNA-protein decoupling. Using the MOBILE[25] framework, we characterized the osmotic stress stimulon and linked the observed mRNA-protein decoupling to a specific impairment of translation initiation.

Materials and Methods

Bacterial strains and growth conditions

Salmonella enterica subsp. enterica serovar Typhimurium SL1344, *Staphylococcus aureus* subsp. aureus MSSA476, and *Yersinia pseudotuberculosi*s serotype O:3 (strain YPIII) species were selected from PATHOgenex database[1] to analyze transcriptomic and proteomic data.

Salmonella enterica SL1344 and *Staphylococcus aureus* MSSA476 strains used in this study did not carry antibiotic resistance markers and were grown in Luria–Bertani (LB) medium (BD Difco™ LB Agar, Lennox) at 37 °C with shaking. The kanamycin-resistant wt *Y. pseudotuberculosis* (YPIII/pIBX) strain was used in this study. Unless otherwise stated, the wild-type strain was grown in LB medium supplemented with 50 µg/mL kanamycin at 26°C with shaking.

Stress exposure and sample collection

Polysome profiling

The wt *Y. pseudotuberculosis* strain was grown overnight in LB, subcultured to an initial optical density (OD_{600}) of 0.05 in 200 mL of LB. When the culture reached an OD_{600} of 0.3, the culture

90 was divided into three 50 mL aliquots, each subjected to one of the following conditions for 10
91 minutes: (1) unexposed control, (2) osmotic stress (0.5 M NaCl), or (3) low-iron stress (250
92 μM 2,2′-Bipyridyl; Sigma-Aldrich). Cells were harvested by filtration, immediately flash-frozen
93 in liquid nitrogen, and stored at -80°C until lysis.

94 Frozen cells were lysed by cryogenic grinding in liquid nitrogen using aluminum oxide. Lysis
95 buffer was added during grinding, and the lysate was allowed to thaw gradually throughout
96 the process. The mixture was transferred into a pre-cooled 1.5 mL reaction tube and
97 immediately placed on ice. Lysates were centrifuged at 18,000 x *g* for 5 minutes at 4°C. The
98 resulting supernatants (~400 μL) were divided into 50 μL aliquots in new, pre-cooled 1.5 mL
99 tubes and stored at −80°C. Total RNA concentration was quantified using the Qubit
100 fluorometer. Each aliquot was used only once for downstream polysome profiling.

101 Linear 15–45% sucrose gradients were prepared using 10% and 50% (w/v) sucrose solutions
102 in 50 mM Tris-HCl (pH 7.5) and supplemented with 100 mM NH_4Cl, 10 mM $MgCl_2$, and 6 mM
103 β-mercaptoethanol. Solutions were sterile filtered and gradients generated using the Gradient
104 Master (BioComp Instruments) following the manufacturer's preprogrammed settings
105 (SW60Ti rotor, short sucrose program, w/v 1st). Prepared gradients were stored at 4°C
106 overnight prior to use.

107 For polysome analysis, 50 μg of cell lysate was carefully layered onto each gradient without
108 disturbing the interface. Samples were centrifuged at 40,000 rpm for 2 hours and 30 minutes
109 at 4°C in an Optima L-100K ultracentrifuge (Beckman Coulter) equipped with an SW60Ti rotor.
110 Immediately following centrifugation, the sucrose gradients were carefully placed on ice to
111 prevent disturbance, and continuous polysome absorbance profiles at 254 nm were recorded
112 using the BioComp gradient fractionation system complemented with a TRIAX flow cell. 14
113 discrete fractions were sequentially collected from the top to the bottom of each gradient using
114 the automated fraction collector, following the manufacturer's instructions and stored at −80°C
115 until further analysis.

116 *Puromycin assay*

117 To assess active translation, the wt strain was cultured as described above for overnight and
118 subculture preparation. Upon reaching exponential phase (OD_{600} of 0.4), the culture was
119 distributed into four separate flasks representing the following treatment conditions: (1)
120 chloramphenicol (50 μg/mL), (2) unexposed control, (3) osmotic stress (0.5 M NaCl), or (4)
121 low-iron stress (250 μM 2,2′-Bipyridyl; Sigma-Aldrich). Cultures grew with their respective
122 conditions for 10 minutes. Subsequently, to abruptly freeze and capture the translational
123 landscape, both 10 μg/mL puromycin and 200 μg/mL chloramphenicol were simultaneously

124 added to the culture, followed by an additional 10-minute incubation period prior to cell lysis.
125 The experiment was performed in four independent biological replicates.

126 To extract proteins, 20 mL of bacterial culture was collected from each treatment group and
127 cells were immediately harvested by centrifugation at 4,000 × *g* for 10 minutes at 4°C. The
128 resulting cell pellets were washed twice with phosphate-buffered saline and resuspended in 1
129 ml of lysis buffer (100 mM Tris-HCl, 10 mM EDTA, 1x cOmplete™ protease inhibitor cocktail
130 (Roche)). The cell suspensions were lysed on ice using sonication for nine 20 second cycles,
131 each cycle comprised 20 second sonication step separated by 30 second intervals with 10%
132 amplitude. Following centrifugation at 13,000 x *g* for 10 minutes at 4°C, the supernatant was
133 collected and stored at -20°C. Protein concentrations were defined using QuBit Protein Assay
134 Kit (Catalog Number: Q33211).

135 *Immunoblotting*

136 Protein concentrations were fixed across samples, and proteins were mixed in 1x sample
137 buffer (2% SDS, 62.5 mM Tris-HCl, 20% glycerol, 0.05% bromophenol blue, 1% β-
138 mercaptoethanol). 17 µl of proteins were separated by sodium dodecyl sulfate (SDS)-PAGE
139 on 4-15% Mini-PROTEAN TGX gels using 80V for 20 minutes, followed by 120V for 1 hour.
140 SDS gel was transferred onto methanol activated Immobilon-P PVDF membrane (Millipore)
141 by semidry electroblotting using Bio-Rad Trans-Blot SD Semi-Dry Transfer Cell (70V for 1
142 hour 10 minutes). Membranes were blocked with 5% milk in 1x Tris-buffered saline with Tween
143 (TBST) and probed with Anti-Puromycin at 1:10,000 dilution (Sigma-Aldrich, MABE343, clone
144 12D10) primary antibody in 1x TBST, overnight at 4°C. Following each antibody incubation,
145 the membranes were washed three times for 10 minutes with 1x TBST. Membranes were
146 incubated with secondary antibody: monoclonal Sheep Anti-Mouse IgG - Horseradish
147 Peroxidase at 1:7,000 dilution (Cytiva, NA931V) in 1x TBST for 1 hour at room temperature.
148 The final blot was developed by chemiluminescence (Immobilon Western Chemiluminescent
149 HRP Substrate, Millipore) for 5 minutes, and the bands were detected with Amersham Imager
150 680. Band intensities were quantified using ImageJ (version 1.54p).

151 *Transcriptomic data retrieval and proteomic data generation*

152 The transcriptome data published for *S. enterica*, *Y. pseudotuberculosis*, and *S. aureus* were
153 obtained from Gene Expression Omnibus (GEO)[26] with accession number GSE152295.

154 Proteomics data was obtained by exposing the three bacterial species to the same stress
155 conditions as mentioned in Avican *et al.*[1], with unexposed control condition and excluding
156 virulence inducing condition. After stress exposures, 3 ml of bacterial cultures were pelleted
157 with 9000 RPM for 2 minutes centrifugation and cell pellets were lysed in 200 µl 2% SDS.

158 Thereafter, 5 µg of each sample were reduced with 20 mM TCEP. Samples were digested
159 with a modified sp3 protocol[27], as previously described[28]. Briefly, samples were added to a
160 bead suspension (10 µg of beads (Sera-Mag Speed Beads, 4515-2105-050250, 6515-2105-
161 050250) in 10 µl 15% formic acid and 30 µl ethanol) and incubated shaking for 15 min at room
162 temperature. Beads were then washed four times with 70% ethanol. Proteins were digested
163 overnight by adding 40 µl of 5 mM chloroacetamide, 1.25 mM TCEP, and 200 ng trypsin in
164 100 mM HEPES pH 8.5. Peptides were eluted from the beads and were dried under vacuum.
165 Peptides were then labelled with TMT11plex (Thermo Fisher Scientific), pooled and desalted
166 with solid-phase extraction using a Waters OASIS HLB µElution Plate (30 µm). Samples were
167 fractionated into 48 fractions on a reversed-phase C18 system running under high pH
168 conditions, with every sixth fraction being pooled together. Samples were analyzed by LC-
169 MS/MS using a data-dependent acquisition strategy on a Thermo Fisher Scientific Vanquish
170 Neo LC coupled with an Thermo Fisher Scientific Orbitrap Exploris 480. Raw files were
171 processed with MSFragger using standard settings for TMT[29] against FASTA databases of
172 each organism obtained from NCBI; for *S. enterica* Typhimurium SL1344 accession numbers
173 NC_016810, NC_017718, NC_017719, NC_017720; for *Y. pseudotuberculosis* accession
174 number NC_010465; for *S. aureus* (MSSA476) accession number NC_002952. The mass
175 spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via
176 the PRIDE partner repository[28] with the dataset identifier PXD055243.

177 Data processing and normalization

178 *Processing of transcriptomic data*

179 Transcriptomic data were imported from PATHOgenex, including statistical values such as
180 LogFC, p-value, and adjusted p-value. Moreover, the differential expression analysis was also
181 already conducted, and thus no further pre-processing was done here.

182 *Processing of proteomic data*

183 Raw protein files containing intensity data from TMT runs were imported into R[30] (version
184 4.4.2). Data were pre-processed to retain only protein locus tag, intensities values, and razor
185 peptide count -defined as peptides shared across multiple proteins. Proteins were then filtered
186 to retain only those quantified by at least two unique peptides and exhibiting non-zero intensity
187 values; entries with zero intensity were considered missing data and excluded. This
188 processing yielded one matrix per species per run, in which each row contained a protein
189 identifier, its total intensity, and its corresponding razor peptide count.

190 To estimate protein abundance, each protein was mapped to its corresponding treatment
191 condition. The relative reporter ion abundance of a given protein for its treatment was divided
192 by the total relative reporter ion abundance as a first estimate of protein abundance. Since
193 longer proteins tend to generate a higher number of razor peptides — and thus more ions —
194 this estimate was further divided by the razor peptide count to obtain the Relative Intensity
195 (RI), ensuring that RI reflects protein abundance independently of protein size. Matrices from
196 the same species were subsequently merged across runs, and proteins absent across all
197 replicates were excluded, yielding three species-specific matrices each containing RI values
198 for every detected protein across the three biological replicates and their associated treatment
199 conditions.

200 To identify the most appropriate normalization strategy, 8 normalization methods were
201 evaluated using the NormalyzerDE[31] package (version 1.23.2): Variance Stabilizing
202 Normalization (VSN), Log2-transformation, Quantile, Global Intensity, Mean, Median, Linear
203 Regression, and Cyclic Loess. Normalized outputs were visualized using principal component
204 analysis (PCA) plots generated with ggplot2[32] (version 4.0.2) to assess replicate clustering
205 within each stress condition. VSN consistently produced the tightest replicate clustering
206 across condition and was therefore applied to the three species-specific matrices to generate
207 VSN-normalized RIs values for downstream analysis.

208 Differential abundance analysis was performed on the normalized proteome data using the
209 limma[33] package (version 3.58.1) with default parameters, applying empirical Bayes
210 moderation to stabilize variance estimates across proteins. Differentially abundant proteins
211 were identified for each species to characterize the proteomic response to every stress
212 condition, and highlight proteins potentially involved in key regulatory mechanisms.

213 Data analysis

214 *Correlation analysis*

215 To assess the correlation between mRNA and protein levels, we used the log2-transformed
216 mean TPM values for mRNAs and the log2-transformed mean RI values for proteins from
217 three replicates of each condition. The Pearson correlation coefficient was calculated for each
218 condition to determine the strength of the correlation. Statistical significance p-value < 0.05
219 was evaluated using Student's t-test.

220 *Identification of gene clusters with discordant mRNA and protein level regulation*

221 To identify genes with discordant mRNA and protein level regulation under 10 stress
222 conditions, the LogFC values of mRNA from control data were combined with the LogFC

223 values of protein, calculated using limma[33] differential expression analysis to create a single
224 data table and visualize using scatter plots separately for each species. Thereafter, the
225 combined transcriptome and proteome data were filtered with *p*-value < 0.05 to select genes
226 whose mRNA or protein level changes upon exposure to stress conditions.

227 Genes were then grouped in 6 different clusters based on the expression patterns as
228 described in Table 1. Cluster 1; mRNA downregulated in at least one stress condition but
229 protein level is not changed, Cluster 2; mRNA downregulated in at least one condition but
230 protein is upregulated, Cluster 3; mRNA level is not changed in any condition but protein is
231 downregulated, Cluster 4; mRNA level is not changed in any condition but protein is
232 upregulated in at least one condition, Cluster 5; mRNA is upregulated in at least one condition
233 but protein is downregulated, and Cluster 6; mRNA is upregulated in at least one condition but
234 protein level is not changed.

235 *Integrative omics analysis with MOBILE*

236 MOBILE pipeline[25] is used for three species (SA, SE, and YP) with 11 samples including both
237 transcriptomic and proteomic datasets. The RI datasets were log2-transformed and
238 normalized. First, they were filtered to only retain the most highly variant features (top 10% for
239 transcripts and 20% for proteins). Next, the MOBILE pipeline (MATLAB version R2023a) is
240 run, using the glmnet package[34] and $Y = \beta.X + \delta$ formalism, where Y matrix is the proteomics
241 dataset and X is the transcriptomics. The objective is to obtain the coefficients matrix β. δ is
242 the y-intercept values, which are almost zero by design. The regression process is repeated
243 10000 times with randomly selected random number generation seeds to yield an ensemble
244 of Lasso matrices. By identifying coefficients that appeared in at least half of the models, a
245 robust Lasso coefficient matrix is selected. This matrix represents the stable associations
246 between analytes across omics layers.

247 To find stress-specific interactions, MOBILE employs a "Leave-One-Group-Out" (LOGO)
248 strategy where each sample exposed to the stress is excluded from the input one-at-a-time.
249 Here, both the osmotic stress and control condition samples were subtracted, and
250 corresponding final matrices were named as OSS-IAN and CTRL-IAN matrices. IANs are the
251 Integrated Association Networks that are coalesced gene-level networks, where nodes
252 represent genes of the matrix analytes and edges stand for the non-zero coefficients.

253 ## Functional annotation

254 *Enrichment analysis*

255 Gene Ontology (GO) enrichment analyses were conducted using the clusterProfiler[35] package
256 v4.10.1 in R. Most significant enriched terms were identified using a cut-off criterion of p-
257 value<0.05. Plots showing the most significant enriched Biological Process (BP) GO terms
258 were generated using the ggplot2 package version 4.0.2.

259 *Over representation analysis (ORA)*

260 The subsequent step involves comparing OSS-IAN and FULL-IAN (all condition matrix), to
261 obtain OSS-dependent interactions. Similarly, comparing CTRL-IAN to FULL-IAN yields
262 CTRL-dependent interactions. Next, pathway enrichment analysis (clusterProfiler[35] version
263 4.14.6) was adopted for the genes list obtained by comparing OSS-dependent list to CTRL-
264 dependent gene list. Gene Ontology gene sets curated for three species (*SA, SE, YP)* were
265 used in gmt format with a minimum gene set size of 15 and a maximum of 500. Significantly
266 enriched gene sets were filtered by q-value < 0.2 and p-value < 0.05.

267 # Results

268 ## Transcriptomic and proteomic analysis of three human bacterial
269 ## pathogens under ten infection relevant stress conditions

270 We analyzed transcriptomic and proteomic datasets of three human bacterial pathogens
271 including Gram-negative *Salmonella enterica* Typhimurium SL1344, *Yersinia*
272 *pseudotuberculosis* YPIII and Gram-positive *Staphylococcus aureus* MSSA476. We used
273 transcriptomes of the three pathogens under 10 infection relevant stress conditions together
274 with unexposed control condition from PATHOgenex RNA Atlas[1] and complemented it with
275 mass-spectrometry proteomic profiling under the same stressors (**Figure 1A**). As mass
276 spectrometry-based proteomics often yields a lower coverage than RNA-seq, we compared
277 detected transcripts and proteins across all conditions in three pathogens, focusing on protein
278 coding sequences (CDSs) in the analysis. While almost all CDSs were detected by RNA-seq
279 for the three pathogens, proteomics could detect between 48% of the CDSs for *S. enterica*
280 Typhimurium and *Y. pseudotuberculosis*, and 32% for *S. aureus* (**Figure 1B**). Detected
281 proteins covered nearly the full spectrum of mRNA expressions, with a bias towards highly
282 expressed mRNAs across the conditions for the three pathogens (**Figure 1C**).

mRNA and protein level correlation under stress

The correlation between mRNA and protein levels has been previously explored in bacteria under individual growth conditions[24,36–38]. Our data allows us to assess whether mRNA levels can serve as an accurate indicator of protein levels in three human bacterial pathogens. To evaluate how different stressors affect the mRNA-protein level correlation, we calculated the intensity relative to razor peptide (RI) **(Supplementary Table 1)** values for protein levels and compared to mRNA levels in Transcript Per Million (TPM)[39] for every stress condition including the unexposed control condition. We have previously shown that the three biological replicates of the RNA-seq samples clustered together[1]. For the proteome dataset generated in this study, we conducted pairwise comparisons of the three replicates of each condition and computed Pearson correlation coefficient values (R-value). The R-value for pairwise comparison within the replicates of the same condition varied from 0.86-0.96 for the three pathogens **(Supplementary Figures 1-3)** showing strong similarities between the replicates. We then measured the correlation between mRNA and protein levels by calculating Pearson correlation coefficient values of the log2 transformed mean TPM and log2 transformed mean RI of each condition for the three pathogens **(Figure 2A, B, and C)**. For these measurements, we limited the analysis to the genes that are detected as mRNA and protein across all tested conditions for each of the pathogens. In the present study, we observed that correlations between mRNA and protein levels under 11 conditions ranged from 0.46-0.65 for *S. enterica* Typhimurium, 0.51-0.71 for *S. aureus* and 0.55-0.72 for *Y. pseudotuberculosis* **(Figure 2A, B, and C)**. These correlation coefficients are comparable to those reported in previous studies across diverse biological systems[24,38,40–42]. The variation in mRNA–protein correlation across stress conditions suggests that distinct molecular mechanisms, activated by each stressor, might modulate translational or post-translational regulation.

Essential genes have higher mRNA-protein level correlations in the three species

The positive correlation between mRNA-protein levels for all detected genes prompted us to investigate subsets of genes with different correlations levels. Essential genes have been shown to have stronger correlations in comparison to non-essentials for *Pseudomonas aeruginosa* during different growth phases[23]. Therefore, we sought to investigate mRNA-protein level correlation profiles of essential genes in the three species. Of the 354, 341, and 445 previously identified essential genes for *S. enterica* Typhimurium[43] *S. aureus*[44] and *Y. pseudotuberculosis*[45] respectively, we used 300 (85%), 278 (82%), and 335 (75%) essential genes detected in both our transcriptome and proteome datasets. We found that mean expression levels of mRNAs and proteins were significantly higher for essential genes in

318 comparison to non-essentials for the three species (**Figure 3A and B).** This was true for the
319 expression levels in every condition **(Supplementary Figures 4 and 5**). Moreover, we
320 observed that essential genes have higher mRNA-protein level correlation than all detected
321 genes for all three species (**Figure 3C**), except for *S. enterica* Typhimurium and *Y.*
322 *pseudotuberculosis* under nutritional downshift (Nd) stress condition. As previously observed
323 in other species[28,46], essential genes showed significantly lower variance on both mRNA and
324 protein expression across the 11 conditions for the three species (**Figure 3D and E**). These
325 findings show that essential genes of *S. enterica* Typhimurium, *S. aureus*, and *Y.*
326 *pseudotuberculosis* have high and stable expression at both mRNA and protein level under a
327 wide range of stress conditions with robust correlations between mRNA and protein levels.
328 This underlines the critical role these genes play in the organisms' adaptability and survival
329 under different stress conditions.

330 Stress responses differentially affect mRNA-protein correlation

331 The variation in correlation under different stress conditions prompted us to investigate
332 possible underlaying reasons. The mRNA-protein levels in Hypoxia (Hyp) and Nutritional
333 downshift (Nd) showed the lowest correlation across three species **(Figure 2A, B, and C)**.
334 Those stress conditions differ from others in the length of the exposure time: 4 hours for Hyp
335 and 30 minutes for Nd, against 10 minutes for others **(Figure 1A)**. To determine whether
336 exposure duration influences mRNA-protein correlation, we examined the relationship
337 between correlation and exposure time across all conditions. Despite a slight negative trend
338 between the mRNA-protein level correlation and the exposure time (R=-0.33; *p-value*=0.078),
339 this relationship was not statistically significant **(Figure 4A).** Additionally, multiple conditions
340 did not follow this trend.

341 To assess whether the degree of stress responses through differentially regulated mRNAs
342 and proteins has an impact on mRNA-protein level correlations, we used differentially
343 expressed genes (DEGs) and differentially expressed proteins (DEPs) for each condition in
344 the three species by using unexposed samples as control. Before identifying DEPs, we
345 evaluated 8 normalization methods from NormalyzerDE[31], on the RI protein abundance data
346 and found that Variance Stabilizing Normalization (VSN) was the most effective in clustering
347 the condition together (**Supplementary Figure 6**). Then we used limma[33] differential
348 expression package[33] for the identification of DEPs with VSN normalized data. For DEGs, we
349 used previously reported differential expression analysis of PATHOgenex RNA Atlas[1]. We
350 employed *p*-value < 0.05 as the threshold to identify DEGs and DEPs for each condition.
351 Notably, the number of DEGs were significantly higher compared to DEPs in all conditions for
352 the three species **(Figure 4B)**. This can be linked to the stronger regulation of mRNA levels

353 than of protein levels in humans[47,48]. Moreover, percentages of DEGs and DEPs were highest
354 under Hyp, Nd, and stationary phase (Sp) conditions for the three pathogens (**Figure 4B**).
355 Next, we tested whether the percentages of DEGs and DEPs have an impact on mRNA-
356 protein level correlations. To do this, we measured the association between the percentages
357 of DEGs and DEPs and the mRNA-protein level correlations with Pearson correlation
358 coefficient. Our results revealed a strong negative correlation between the percentages of
359 DEGs and DEPs and the mRNA-protein level correlation (**Figure 4C**). This suggests that other
360 factors, such as post-transcriptional and post-translational regulation triggered by different
361 stressors, might be influencing protein levels alongside the mRNA expression.

362 Identification of genes whose mRNA-protein levels are discordantly 363 regulated under stress conditions

364 The decreased correlation between mRNA-protein levels with increased number of DEGs and
365 DEPs under specific stress conditions led us to investigate the genes whose mRNA-protein
366 levels are discordantly regulated when the bacteria are exposed to stress conditions. We
367 clustered genes into 6 groups (Table 1: Classification of differentially expressed genes based
368 on mRNA and protein regulation patterns.**Table 1**) corresponding to distinct combinations of
369 differential expression patterns at the mRNA and protein levels (**Figure 5A and B and**
370 **Supplementary Table 2**). In agreement with lowest correlation between mRNA-protein levels
371 observed under Hyp, Nd, and Sp conditions (**Figure 2A-C**), discordantly regulated genes were
372 mostly found under those conditions for the three species (**Supplementary Figure 7**).
373 Noteworthy, Clusters 1 and 6, where mRNAs were regulated but not proteins, had the highest
374 number of discordantly regulated genes. This is in accordance with higher number of
375 differentially regulated mRNAs (DEGs) compared to differentially regulated proteins (DEPs)
376 in all conditions for the three species (**Figure 4B**).

377 *Table 1: Classification of differentially expressed genes based on mRNA and protein*
378 *regulation patterns.*

Cluster	mRNA regulation	Protein regulation	Description
1	▼ Downregulated	• Not changed	mRNA is downregulated in at least one stress condition, while protein level remains unchanged
2	▼ Downregulated	▲ Upregulated	mRNA is downregulated in at least one condition, while protein level is upregulated
3	• Not changed	▼ Downregulated	mRNA level remains unchanged across all conditions, while protein is downregulated
4	• Not changed	▲ Upregulated	mRNA level remains unchanged across all conditions, while protein is upregulated in at least one condition
5	▲ Upregulated	▼ Downregulated	mRNA is upregulated in at least one condition, while protein level is downregulated
6	▲ Upregulated	• Not changed	mRNA is upregulated in at least one condition, while protein level remains unchanged

Discordantly regulated genes are enriched in translation

To investigate the underlying mechanisms leading to lowered mRNA-protein level correlation under stress conditions, we performed Gene Ontology[49] (GO) enrichment analysis for genes exhibiting discordant mRNA and protein levels. Interestingly, we found that biological processes such as translation, RNA processing, RNA modification, proteolysis, and cell division, among other biological processes were commonly enriched for the three species, particularly when mRNA levels were changed but there was no change at the protein level (**Figure 5C and Supplementary Table 3**). This suggests that genes encoding subunits of translation machinery together with other components of translation such as tRNA and rRNA were discordantly regulated under stress conditions for the three pathogens. In detail, proteins in Cluster 1 have stable levels despite mRNA being downregulated, likely due to increased protein stability. In contrast, proteins in Cluster 6 have stable levels despite mRNA being upregulated, suggesting reduced translation or increased proteolysis, which is enriched for *S. enterica* Typhimurium (**Figure 5C**). Moreover, genes encoding proteins involved in central metabolism such as tricarboxylic acid cycle were commonly enriched for the three species in Cluster 5, where mRNAs were upregulated and proteins downregulated. This indicates presence of post-transcriptional regulations such as regulation via small regulatory RNAs involved in carbon metabolism[50,51]or post-translational modifications which have high prevalence in bacterial metabolic enzymes[52].

mRNA-protein level stress-specific stimulon genes under stress

We previously identified stress-specific stimulon genes as sets of genes that are co-regulated in response to distinct stress conditions[2]. To assess whether these stimulon genes exhibit distinct mRNA-protein level correlations compared to other genes, we analyzed mRNA and protein differential expression data across all genes and stress-specific stimulons. No significant differences in mRNA-protein level correlations were observed between stimulon genes and other genes across most conditions except in the case of the osmotic stress-specific stimulon (**Figure 6A**). Notably, genes within the osmotic stimulon showed a markedly lower mRNA-protein level correlation across all three pathogens studied.

To investigate this further, we used MOBILE framework[25] to identify osmotic stress specific associations between mRNA and protein levels. The variation in mRNA–protein correlation across stress conditions suggests that distinct molecular mechanisms, differentially activated by each stressor, modulate translational or post-translational regulation. Through repeated Lasso regression between paired omics layers, MOBILE enables the identification of condition-specific associations without requiring prior knowledge. The transcriptomic and

413 proteomic datasets with all conditions were used as an input for MOBILE, yielding a list of
414 statistically robust associations between proteins and mRNAs. Next, the osmotic stress (Oss)
415 condition data from both datasets were excluded from the input, and MOBILE is re-run to
416 obtain condition-dependent associations list by comparing the all-data run to the latter.
417 Similarly, the control (Ctrl) condition data is held out, and MOBILE re-run yielded another list
418 of associations to identify CTRL-data dependent associations (**Figure 6B**). By comparing the
419 osmotic stress-dependent and control-dependent association lists, we determined by over
420 representation analysis (ORA), that the OSS-specific associations can be used to generate a
421 gene-level network and a weighted genes list based on connectivity of each gene. Using that
422 gene list, we ran a GO enrichment analysis. Interestingly, translation-related processes and
423 components of the translational machinery were enriched only for *Y. pseudotuberculosis*
424 (**Figure 6C**). This suggests a functional link between osmotic stress-specific mRNA-protein
425 decoupling and translational regulation.

426 Based on this finding, we hypothesized that the reduced mRNA-protein correlation observed
427 for osmotic stress-specific stimulon genes might result from impaired translation under stress.
428 To test this, we performed polysome profiling on *Y. pseudotuberculosis* exposed to osmotic
429 stress for 10 minutes, using unstressed bacteria collected at matched time points as controls.
430 The polysome profiles showed no notable differences between osmotic stress and control
431 condition (**Figure 6D**), indicating that global translation remains active under osmotic stress.
432 However, global translational activity may be insufficient to meet the heightened translational
433 demand triggered by the strong transcriptional induction of osmotic stress-specific stimulon
434 genes.

435 To further characterize the translational response, translation efficiency was assessed using
436 a puromycin incorporation assay across osmotic stress, low iron, and unexposed control
437 conditions. Chloramphenicol-treated bacteria served as a translationally inactive control.
438 Since puromycin causes a premature translation termination, its incorporation serves as a
439 direct proxy for active translation. The results showed that translational activity was
440 significantly reduced under osmotic stress compared to both unexposed control and low iron
441 conditions, indicating that osmotic stress specifically impairs translation activity (**Figure 6E**).
442 Therefore, the reduced mRNA-protein correlation for osmotic stress-specific stimulons is not
443 surprising given the overall reduction in translational activity. These stimulon genes exhibit
444 significant changes in mRNA abundance during stress. However, because the translational
445 machinery is impaired, it cannot efficiently process or respond to these altered transcript
446 levels. As a result, the correlation between mRNA and protein decreases. Notably, this
447 decoupling is specific to osmotic stress and was not observed under low iron conditions.

448 # Discussion

449 Bacterial species continuously sense and respond to changes in stressful host conditions and
450 environments. Such stress responses include a series of regulatory mechanisms at both
451 transcriptional and translational levels leading to changes on mRNA and subsequently protein
452 level. However, it is known that transcription and translation could be uncoupled constitutively
453 or under certain conditions, such as stress conditions, at a global or at a gene-specific level[53,54]
454 Therefore, the relationship between mRNA and protein levels can vary depending on the
455 conditions, and this variability has been the subject of extensive research and ongoing
456 debate[12,48,54,55]. Nevertheless, mRNA levels have been shown to be positively correlated with
457 protein levels in many different organisms[8,23,36,40,41,56–60]. Revealing relationship between
458 mRNA and protein level in pathogenic bacteria is therefore critical for understanding bacterial
459 pathogenesis and adaptation mechanism to host conditions during infection. Integrative
460 transcriptome and proteome analyses have previously shown lowered mRNA-protein
461 correlation levels under stress conditions such as stationary phase, heat shock, osmotic
462 stress, and oxidative stress in *P. aeruginosa* and *Mycoplasma pneumoniae*[23,59].

463 We conducted a comprehensive analysis by integrating transcriptomic and proteomic datasets
464 of *S. enterica* Typhimurium, *Y. pseudotuberculosis*, and *S. aureus* under 10 infection-relevant
465 stress conditions, as well as a control condition during exponential growth. The inclusion of
466 diverse pathogens, which are distantly related yet possess conserved genes, enabled us to
467 identify both common and species-specific mRNA-protein relationships across a range of
468 conditions mimicking host environments. Consistent with prior findings in other organisms, our
469 analyses indicate a positive correlation between mRNA and protein levels across the three
470 bacterial pathogens. However, this correlation is significantly reduced under stress conditions,
471 particularly those that trigger a strong stress response.

472 This reduction is not surprising, as stress response proteins might be involved in post-
473 transcriptional or post-translational regulations that can decouple mRNA and protein levels.
474 Such regulations can occur through the stabilization or destabilization of mRNA and proteins
475 as a response to the stress. RNA-binding proteins such as CsrA and ProQ can bind to target
476 mRNAs under specific conditions, either stabilizing or destabilizing them, or preventing their
477 translation[61–66]. Furthermore, the discoordination of transcription and translation elongation
478 rates under stress conditions may contribute to the decrease in mRNA-protein level
479 correlation[67]. For example, in *Escherichia coli* exposed to oxidative stress, there is a significant
480 induction of mRNA expression of the OxyR regulon, which is crucial for survival under
481 oxidative stress, while the translation elongation rate significantly slows down[68,69].

482 In contrast, under nutrient-limited conditions, there has been tight coordination between
483 transcription and translation[70,71]. This suggests that the reduced mRNA-protein level
484 correlation in our nutritional downshift condition might not be solely due to the discoordination
485 of transcription and translation but rather other regulatory mechanisms. Conformational
486 changes in riboswitches upon a change in the environment can also lead to reprogrammed
487 mRNA translation and degradation, as is the case for the lysC riboswitch in the presence and
488 absence of lysine[72].

489 Even though mostly studied in eukaryotes, post-transcriptional nucleotide modifications
490 triggered by stress conditions can also have an impact on mRNA-protein level as it is shown
491 for increased level of m^6A methylation under hypoxic for human cancer cells[73]. On the other
492 hand, m^6A methylation was shown to be lowered under increased temperature for *P.*
493 *aeruginosa*[74]. Similarly post-translational modifications such as methylation, phosphorylation,
494 acetylation, and sumoylation can directly impact protein stability or indirectly through triggered
495 signaling cascade in both eukaryotes and prokaryotes[75,76]. Even though protein stability was
496 not specifically investigated, it was shown that cellular protein phosphorylation was increased
497 during *Streptomyces* differentiation to form spores, which is induced by stress conditions[77].

498 To assess whether the mRNA-protein discrepancies observed for DEGs are a common
499 feature across species, we identified conserved DEGs across the three pathogens and
500 compared their expression patterns. These conserved genes exhibited similar non-correlating
501 expression patterns under the same stress conditions, suggesting that post-transcriptional and
502 post-translational modifications affecting these genes may be conserved across species. This
503 highlights the evolutionary significance of these regulatory mechanisms in bacterial adaptation
504 to stress.

505 Among all conditions investigated for stress stimulons, osmotic stress exhibited the most
506 pronounced mRNA-protein decoupling. This impairment may stem from structural remodeling
507 of the cell envelope. Osmotic stress induces rapid cell depolymerization to manage turgor
508 pressure[78,79], which can physically disrupt membrane integrity and the efficiency of
509 transmembrane transporters. In case of deficient methionine uptake, translation initiation
510 would be affected due to the need of formylmethionine at the start of bacterial protein
511 synthesis[56,57]

512 Our study underscores the importance of integrating transcriptomic and proteomic data to gain
513 a holistic understanding of bacterial responses to infection-relevant stressors. The observed
514 discrepancies between mRNA and protein levels emphasize the need for caution when
515 inferring protein abundance solely from transcriptomic data The precise post-transcriptional
516 and post-translational mechanisms contributing to the observed mRNA-protein discrepancies

517 remain to be elucidated, as the currently available datasets do not allow their direct
518 identification, representing an important avenue for future investigation.

519 In conclusion, our findings provide a comprehensive overview of mRNA-protein relationships
520 under stress conditions in 3 bacterial pathogens, laying the groundwork for further
521 investigations into the regulatory mechanisms mediating bacterial adaptation and survival in
522 hostile host environments. The multi-species, multi-condition nature of this dataset
523 furthermore represents a valuable resource for training predictive models of protein
524 abundance from mRNA levels, particularly under stress conditions where the mRNA-protein
525 relationship deviates substantially from baseline.

526 Funding

527 This work was supported by Swedish Research Council (No. 2021-02466), Kempestiftelserna
528 (JCK22-0017), and the Medical Faculty at Umeå University (FS 2.1.6-281-22) to K. Avican, by
529 Swedish Research Council Excellence Center grant (No. 2022-06543) for the Center for
530 Modeling Adaptive Mechanisms in Living Systems Under Stress to K. Avican

531 CE is funded by the Knut and Alice Wallenberg Foundation under the SciLifeLab and
532 Wallenberg Data Driven Life Science Program (KAW 2020.0239).

533 References

534 1. Avican, K. *et al.* RNA atlas of human bacterial pathogens uncovers stress dynamics linked to
535 infection. *Nat. Commun.* **12**, 3282 (2021).
536 2. Fernandez, L., Rosvall, M., Normark, J., Fallman, M. & Avican, K. Co-PATHOgenex web
537 application for assessing complex stress responses in pathogenic bacteria. *Microbiol Spectr* **12**,
538 e02781-23 (2024).
539 3. Lee, M. J. & Yaffe, M. B. Protein Regulation in Signal Transduction. *Cold Spring Harb Perspect*
540 *Biol* **8**, a005918 (2016).
541 4. Bathke, J., Konzer, A., Remes, B., McIntosh, M. & Klug, G. Comparative analyses of the variation
542 of the transcriptome and proteome of Rhodobacter sphaeroides throughout growth. *BMC*
543 *Genomics* **20**, 358 (2019).
544 5. Zhu, Y. *et al.* Integrative analysis of transcriptome and proteome provides insights into adaptation
545 to cadmium stress in Sedum plumbizincicola. *Ecotoxicol Env. Saf* **230**, 113–149 (2022).
546 6. Liu, S. *et al.* Transcriptome and Proteome of Methicillin-Resistant Staphylococcus aureus Small-
547 Colony Variants Reveal Changed Metabolism and Increased Immune Evasion. *Microbiol Spectr*
548 **11**, e0189822 (2023).
549 7. Grunberger, F. *et al.* Uncovering the temporal dynamics and regulatory networks of thermal stress
550 response in a hyperthermophile using transcriptomics and proteomics. *mBio* **14**, e0217423 (2023).
551 8. Edfors, F. *et al.* Gene-specific correlation of RNA and protein levels in human cells and tissues.
552 *Mol Syst Biol* **12**, 883 (2016).
553 9. Feng, X. & Ma, Q. Transcriptome and proteome profiling revealed molecular mechanism of
554 selenium responses in bread wheat (Triticum aestivum L.). *BMC Plant Biol* **21**, 584 (2021).
555 10. Backman, M. *et al.* Multi-omics insights into functional alterations of the liver in insulin-deficient
556 diabetes mellitus. *Mol Metab* **26**, 30–44 (2019).

557 11. Chen, G. *et al.* Discordant protein and mRNA expression in lung adenocarcinomas. *Mol Cell*
558 *Proteomics* **1**, 304–13 (2002).
559 12. Vogel, C. & Marcotte, E. M. Insights into the regulation of protein abundance from proteomic and
560 transcriptomic analyses. *Nat Rev Genet* **13**, 227–32 (2012).
561 13. Wang, J. *et al.* Combined proteomic and transcriptomic analysis of the antimicrobial mechanism of
562 tannic acid against Staphylococcus aureus. *Front Pharmacol* **14**, 1178177 (2023).
563 14. Kavela, S. *et al.* Use of an Integrated Multi-Omics Approach To Identify Molecular Mechanisms and
564 Critical Factors Involved in the Pathogenesis of Leptospira. *Microbiol Spectr* **11**, e0313522 (2023).
565 15. Huang, L. *et al.* Integration of Transcriptomic and Proteomic Approaches Reveals the Temperature-
566 Dependent Virulence of Pseudomonas plecoglossicida. *Front Cell Infect Microbiol* **8**, 207 (2018).
567 16. Kocharunchitt, C., King, T., Gobius, K., Bowman, J. P. & Ross, T. Integrated transcriptomic and
568 proteomic analysis of the physiological response of Escherichia coli O157:H7 Sakai to steady-state
569 conditions of cold and water activity stress. *Mol Cell Proteomics* **11**, M111.009019 (2012).
570 17. Xiong, Q. *et al.* Integrated transcriptomic and proteomic analysis of the global response of
571 Synechococcus to high light stress. *Mol Cell Proteomics* **14**, 1038–53 (2015).
572 18. Perez-Llarena, F. J. & Bou, G. Proteomics As a Tool for Studying Bacterial Virulence and
573 Antimicrobial Resistance. *Front Microbiol* **7**, 410 (2016).
574 19. Yang, X. *et al.* RNA markers for ultra-rapid molecular antimicrobial susceptibility testing in
575 fluoroquinolone-treated Klebsiella pneumoniae. *J Antimicrob Chemother* **75**, 1747–1755 (2020).
576 20. Ballen, V. *et al.* Antibiotic Resistance and Virulence Profiles of Klebsiella pneumoniae Strains
577 Isolated From Different Clinical Sources. *Front Cell Infect Microbiol* **11**, 738223 (2021).
578 21. Chaffin, D. O., Taylor, D., Skerrett, S. J. & Rubens, C. E. Changes in the Staphylococcus aureus
579 transcriptome during early adaptation to the lung. *PLoS One* **7**, e41329 (2012).
580 22. Liou, G. G., Chao Kaberdina, A., Wang, W. S., Kaberdin, V. R. & Lin-Chao, S. Combined
581 Transcriptomic and Proteomic Profiling of E. coli under Microaerobic versus Aerobic Conditions:
582 The Multifaceted Roles of Noncoding Small RNAs and Oxygen-Dependent Sensing in Global Gene
583 Expression Control. *Int J Mol Sci* **23**, 2570 (2022).
584 23. Zhang, M. *et al.* Impact of Growth Rate on the Protein-mRNA Ratio in Pseudomonas aeruginosa.
585 *mBio* **14**, e0306722 (2023).
586 24. Maier, T. *et al.* Quantification of mRNA and protein and integration with protein turnover in a
587 bacterium. *Mol. Syst. Biol.* **7**, MSB201138 (2011).
588 25. Erdem, C., Gross, S. M., Heiser, L. M. & Birtwistle, M. R. MOBILE pipeline enables identification of
589 context-specific networks and regulatory mechanisms. *Nat. Commun.* **14**, 3991 (2023).
590 26. Barrett, T. *et al.* NCBI GEO: archive for functional genomics data sets—update. *Nucleic Acids Res.*
591 **41**, D991–D995 (2013).
592 27. Hughes, C. S. *et al.* Ultrasensitive proteome analysis using paramagnetic bead technology. *Mol*
593 *Syst Biol* **10**, 757 (2014).
594 28. Mateus, A. *et al.* The functional proteome landscape of Escherichia coli. *Nature* **588**, 473–478
595 (2020).
596 29. Kong, A. T., Leprevost, F. V., Avtonomov, D. M., Mellacheruvu, D. & Nesvizhskii, A. I. MSFragger:
597 ultrafast and comprehensive peptide identification in mass spectrometry-based proteomics. *Nat*
598 *Methods* **14**, 513–520 (2017).
599 30. R Core Team. *R: A Language and Environment for Statistical Computing*. (R Foundation for
600 Statistical Computing, Vienna, Austria, 2024).
601 31. Willforss, J., Chawade, A. & Levander, F. NormalyzerDE: Online Tool for Improved Normalization
602 of Omics Expression Data and High-Sensitivity Differential Expression Analysis. *J Proteome Res*
603 **18**, 732–740 (2019).
604 32. Wickham, H. *Ggplot2: Elegant Graphics for Data Analysis*. (Springer-Verlag New York, 2016).
605 33. Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-sequencing and
606 microarray studies. *Nucleic Acids Res* **43**, e47 (2015).
607 34. Friedman, J. H., Hastie, T. & Tibshirani, R. Regularization Paths for Generalized Linear Models via
608 Coordinate Descent. *J. Stat. Softw.* **33**, 1–22 (2010).
609 35. Wu, T. *et al.* clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innov. Camb*
610 **2**, 100141 (2021).
611 36. Caglar, M. U. *et al.* The E. coli molecular phenotype under different growth conditions. *Sci. Rep.* **7**,
612 45303 (2017).

613 37. Kwon, T., Huse, H. K., Vogel, C., Whiteley, M. & Marcotte, E. M. Protein-to-mRNA Ratios Are
614 Conserved between Pseudomonas aeruginosa Strains. *J. Proteome Res.* **13**, 2370–2380 (2014).
615 38. Jayapal, K. P. *et al.* Uncovering Genes with Divergent mRNA-Protein Dynamics in Streptomyces
616 coelicolor. *PLOS ONE* **3**, e2097 (2008).
617 39. Li, B. & Dewey, C. N. RSEM: accurate transcript quantification from RNA-Seq data with or without
618 a reference genome. *BMC Bioinformatics* **12**, 323 (2011).
619 40. Choi, Y. W., Park, S. A., Lee, H. W., Kim, D. S. & Lee, N. G. Analysis of growth phase-dependent
620 proteome profiles reveals differential regulation of mRNA and protein in Helicobacter pylori.
621 *PROTEOMICS* **8**, 2665–2675 (2008).
622 41. Chen, W.-H. *et al.* Integration of multi-omics data of a genome-reduced bacterium: Prevalence of
623 post-transcriptional regulation and its correlation with protein abundances. *Nucleic Acids Res.* **44**,
624 1192–1202 (2016).
625 42. Lu, P., Vogel, C., Wang, R., Yao, X. & Marcotte, E. M. Absolute protein expression profiling
626 estimates the relative contributions of transcriptional and translational regulation. *Nat. Biotechnol.*
627 **25**, 117–124 (2007).
628 43. Barquist, L. *et al.* A comparison of dense transposon insertion libraries in the Salmonella serovars
629 Typhi and Typhimurium. *Nucleic Acids Res.* **41**, 4549–4564 (2013).
630 44. Chaudhuri, R. R. *et al.* Comprehensive identification of essential Staphylococcus aureus genes
631 using Transposon-Mediated Differential Hybridisation (TMDH). *BMC Genomics* **10**, 291 (2009).
632 45. Willcocks, S. J., Stabler, R. A., Atkins, H. S., Oyston, P. F. & Wren, B. W. High-throughput analysis
633 of Yersinia pseudotuberculosis gene essentiality in optimised in vitro conditions, and implications
634 for the speciation of Yersinia pestis. *BMC Microbiol.* **18**, 46 (2018).
635 46. Sigalova, O. M., Shaeiri, A., Forneris, M., Furlong, E. E. & Zaugg, J. B. Predictive features of gene
636 expression variation reveal mechanistic link with differential expression. *Mol. Syst. Biol.* **16**,
637 MSB209539 (2020).
638 47. Robles, M. S., Cox, J. & Mann, M. In-Vivo Quantitative Proteomics Reveals a Key Contribution of
639 Post-Transcriptional Mechanisms to the Circadian Regulation of Liver Metabolism. *PLOS Genet.*
640 **10**, e1004047 (2014).
641 48. Jovanovic, M. *et al.* Dynamic profiling of the protein life cycle in response to pathogens. *Science*
642 **347**, 1259038 (2015).
643 49. Ashburner, M. *et al.* Gene Ontology: tool for the unification of biology. *Nat. Genet.* **25**, 25–29 (2000).
644 50. Durica-Mitic, S., Göpel, Y. & Görke, B. Carbohydrate Utilization in Bacteria: Making the Most Out
645 of Sugars with the Help of Small Regulatory RNAs. *Microbiol. Spectr.* **6**, 10.1128/microbiolspec.rwr-
646 0013–2017 (2018).
647 51. Raina, M. *et al.* Dual-function AzuCR RNA modulates carbon metabolism. *Proc. Natl. Acad. Sci.*
648 **119**, e2117930119 (2022).
649 52. Pisithkul, T., Patel, N. M. & Amador-Noguez, D. Post-translational modifications as key regulators
650 of bacterial metabolic fluxes. *Curr. Opin. Microbiol.* **24**, 29–37 (2015).
651 53. Gebauer, F. & Hentze, M. W. Molecular mechanisms of translational control. *Nat. Rev. Mol. Cell*
652 *Biol.* **5**, 827–835 (2004).
653 54. Maier, T., Güell, M. & Serrano, L. Correlation of mRNA and protein in complex biological samples.
654 *FEBS Lett.* **583**, 3966–3973 (2009).
655 55. Schwanhäusser, B. *et al.* Global quantification of mammalian gene expression control. *Nature* **473**,
656 337–342 (2011).
657 56. Becker, K. *et al.* Quantifying post-transcriptional regulation in the development of Drosophila
658 melanogaster. *Nat. Commun.* **9**, 4970 (2018).
659 57. Corbin, R. W. *et al.* Toward a Protein Profile of Escherichia coli: Comparison to Its Transcription
660 Profile. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 9232–9237 (2003).
661 58. Gygi, S. P., Rochon, Y., Franza, B. R. & Aebersold, R. Correlation between Protein and mRNA
662 Abundance in Yeast. *Mol. Cell. Biol.* **19**, 1720–1730 (1999).
663 59. Jeacock, L., Faria, J. & Horn, D. Codon usage bias controls mRNA and protein abundance in
664 trypanosomatids. *eLife* **7**, e32496 (2018).
665 60. Ponnala, L., Wang, Y., Sun, Q. & van Wijk, K. J. Correlation of mRNA and protein abundance in
666 the developing maize leaf. *Plant J.* **78**, 424–440 (2014).

667 61. Yakhnin, H. *et al.* CsrA Represses Translation of sdiA, Which Encodes the N-Acylhomoserine-l-
668 Lactone Receptor of Escherichia coli, by Binding Exclusively within the Coding Region of sdiA
669 mRNA. *J. Bacteriol.* **193**, 6162–6170 (2011).
670 62. Pourciau, C., Lai, Y.-J., Gorelik, M., Babitzke, P. & Romeo, T. Diverse Mechanisms and Circuitry
671 for Global Regulation by the RNA-Binding Protein CsrA. *Front. Microbiol.* **11**, (2020).
672 63. Wang, X. *et al.* CsrA post-transcriptionally represses pgaABCD, responsible for synthesis of a
673 biofilm polysaccharide adhesin of Escherichia coli. *Mol. Microbiol.* **56**, 1648–1663 (2005).
674 64. Dubey, A. K. *et al.* CsrA Regulates Translation of the Escherichia coli Carbon Starvation Gene,
675 cstA, by Blocking Ribosome Access to the cstA Transcript. *J. Bacteriol.* **185**, 4450–4460 (2003).
676 65. Holmqvist, E., Li, L., Bischler, T., Barquist, L. & Vogel, J. Global Maps of ProQ Binding *In Vivo*
677 Reveal Target Recognition via RNA Structure and Stability Control at mRNA 3′ Ends. *Mol. Cell* **70**,
678 971-982.e6 (2018).
679 66. Smirnov, A., Wang, C., Drewry, L. L. & Vogel, J. Molecular mechanism of mRNA repression in trans
680 by a ProQ-dependent small RNA. *EMBO J.* **36**, 1029–1045 (2017).
681 67. Irastortza-Olaziregi, M. & Amster-Choder, O. Coupled Transcription-Translation in Prokaryotes: An
682 Old Couple With New Surprises. *Front. Microbiol.* **11**, (2021).
683 68. Zhu, M. & Dai, X. Maintenance of translational elongation rate underlies the survival of Escherichia
684 coli during oxidative stress. *Nucleic Acids Res.* **47**, 7592–7604 (2019).
685 69. Zhu, M. & Dai, X. Bacterial stress defense: the crucial role of ribosome speed. *Cell. Mol. Life Sci.*
686 **77**, 853–858 (2020).
687 70. Vogel, U. & Jensen, K. F. The RNA chain elongation rate in Escherichia coli depends on the growth
688 rate. *J. Bacteriol.* **176**, 2807–2813 (1994).
689 71. Zhu, M., Mori, M., Hwa, T. & Dai, X. Disruption of transcription–translation coordination in
690 Escherichia coli leads to premature transcriptional termination. *Nat. Microbiol.* **4**, 2347–2356
691 (2019).
692 72. Caron, M.-P. *et al.* Dual-acting riboswitch control of translation initiation and mRNA decay. *Proc.*
693 *Natl. Acad. Sci.* **109**, E3444–E3453 (2012).
694 73. Scheithauer, B. W. & Bruner, J. M. Central nervous system tumors. *Clin. Lab. Med.* **7**, 157–179
695 (1987).
696 74. Deng, X. *et al.* Widespread occurrence of N6-methyladenosine in bacterial mRNA. *Nucleic Acids*
697 *Res.* **43**, 6557–6567 (2015).
698 75. Lee, J. M., Hammarén, H. M., Savitski, M. M. & Baek, S. H. Control of protein stability by post-
699 translational modifications. *Nat. Commun.* **14**, 201 (2023).
700 76. Cain, J. A., Solis, N. & Cordwell, S. J. Beyond gene expression: The impact of protein post-
701 translational modifications in bacteria. *J. Proteomics* **97**, 265–286 (2014).
702 77. Manteca, A., Ye, J., Sánchez, J. & Jensen, O. N. Phosphoproteome Analysis of Streptomyces
703 Development Reveals Extensive Protein Phosphorylation Accompanying Bacterial Differentiation.
704 *J. Proteome Res.* **10**, 5481–5492 (2011).
705 78. Bremer, E. & Krämer, R. Responses of Microorganisms to Osmotic Stress. *Annu. Rev. Microbiol.*
706 **73**, 313–334 (2019).
707 79. Mueller, E. A. & Levin, P. A. Bacterial Cell Wall Quality Control during Environmental Stress. *mBio*
708 **11**, 10.1128/mbio.02456-20 (2020).
709

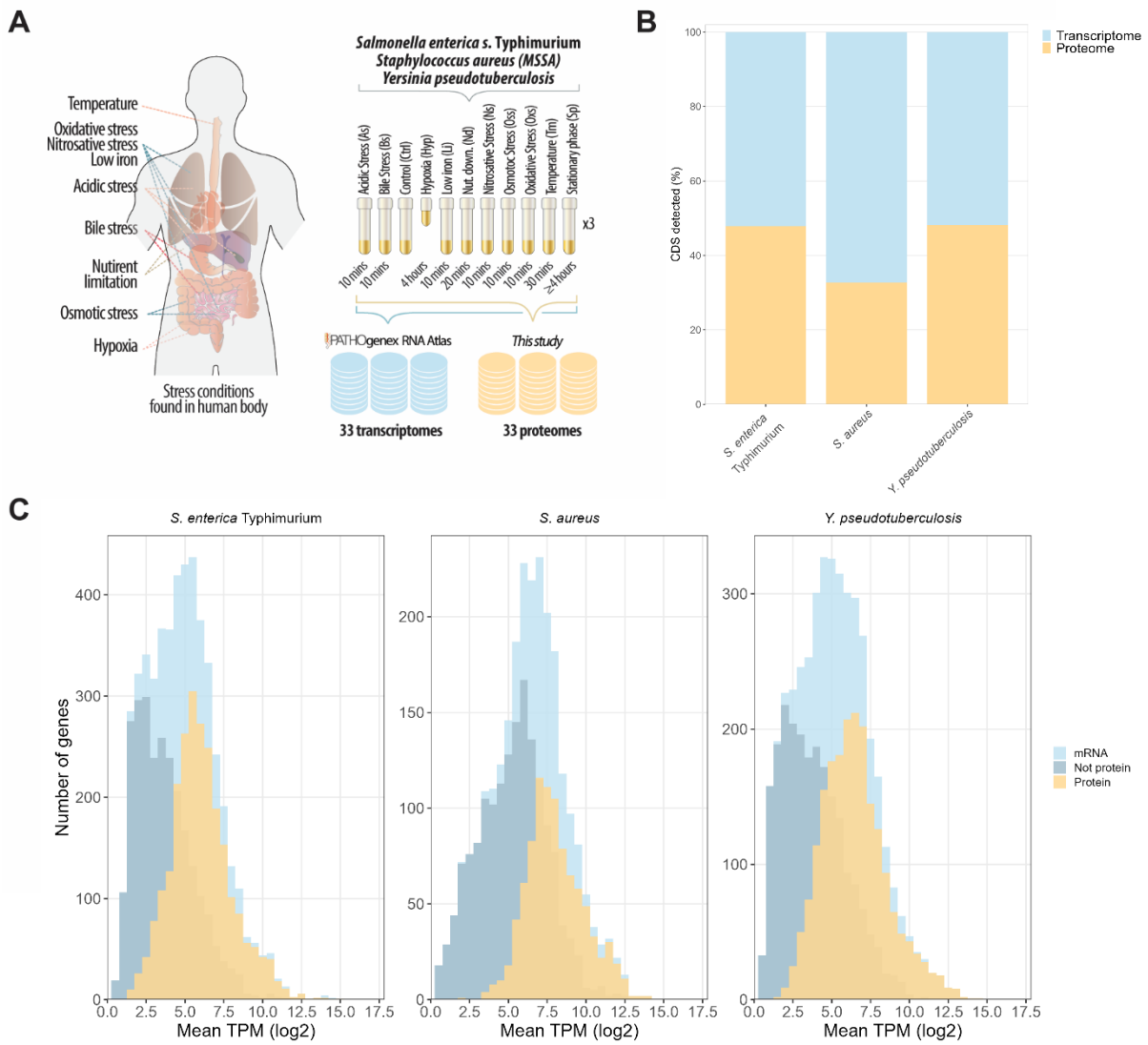


Figure 1. Transcriptomic and proteomic analysis of three human bacterial pathogens under ten infection relevant stress conditions. A) Schematic presentation of data collection and data generation for gene expression on mRNA and protein level under stress conditions. **B)** Percentages of CDSs detected by transcriptomics and proteomics for the three species. **C**) Distribution of transcript abundance across all conditions (light blue) in the three species; yellow indicates the fraction of proteins detected, while dark blue represents the transcripts for which no corresponding protein was detected.

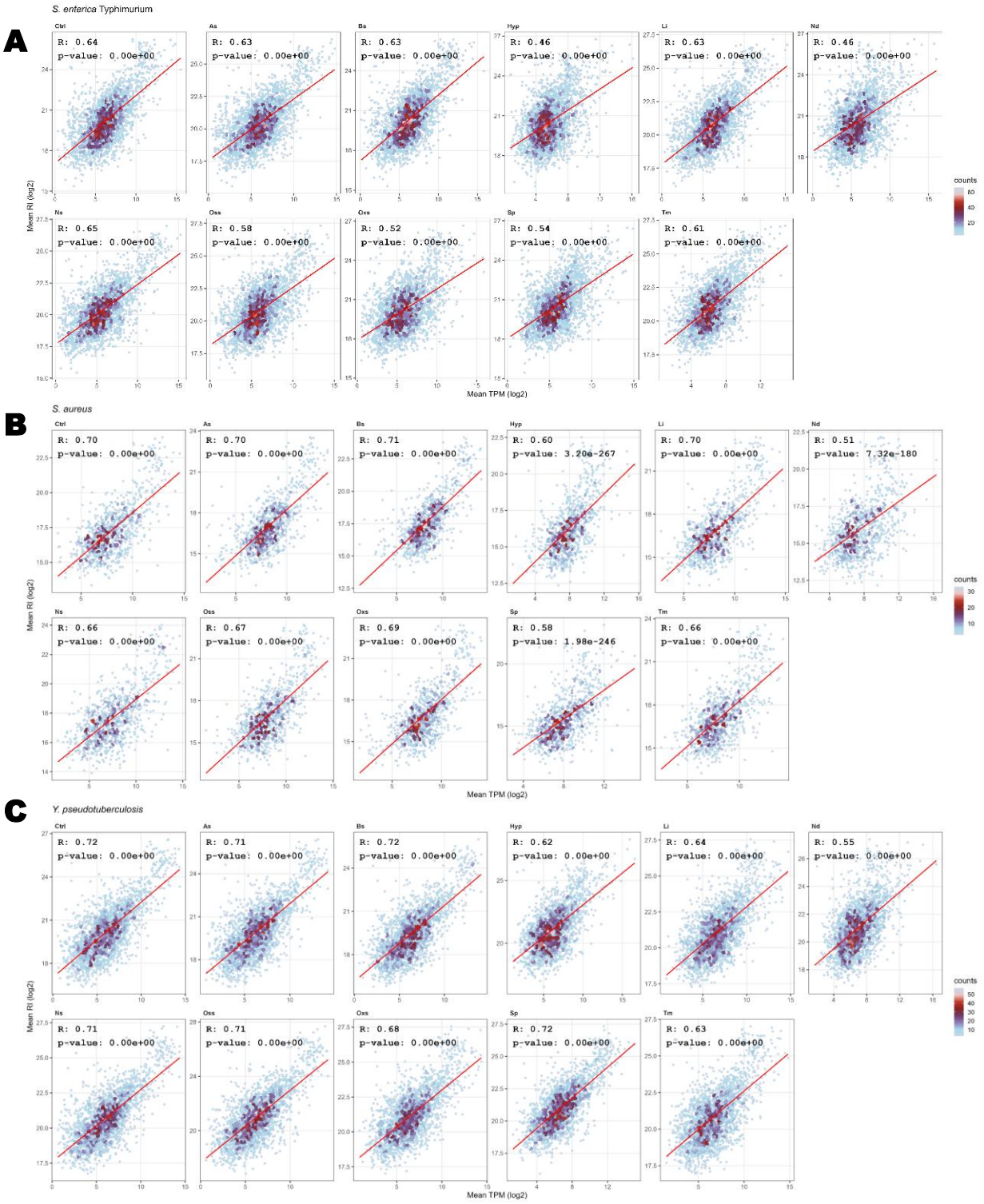
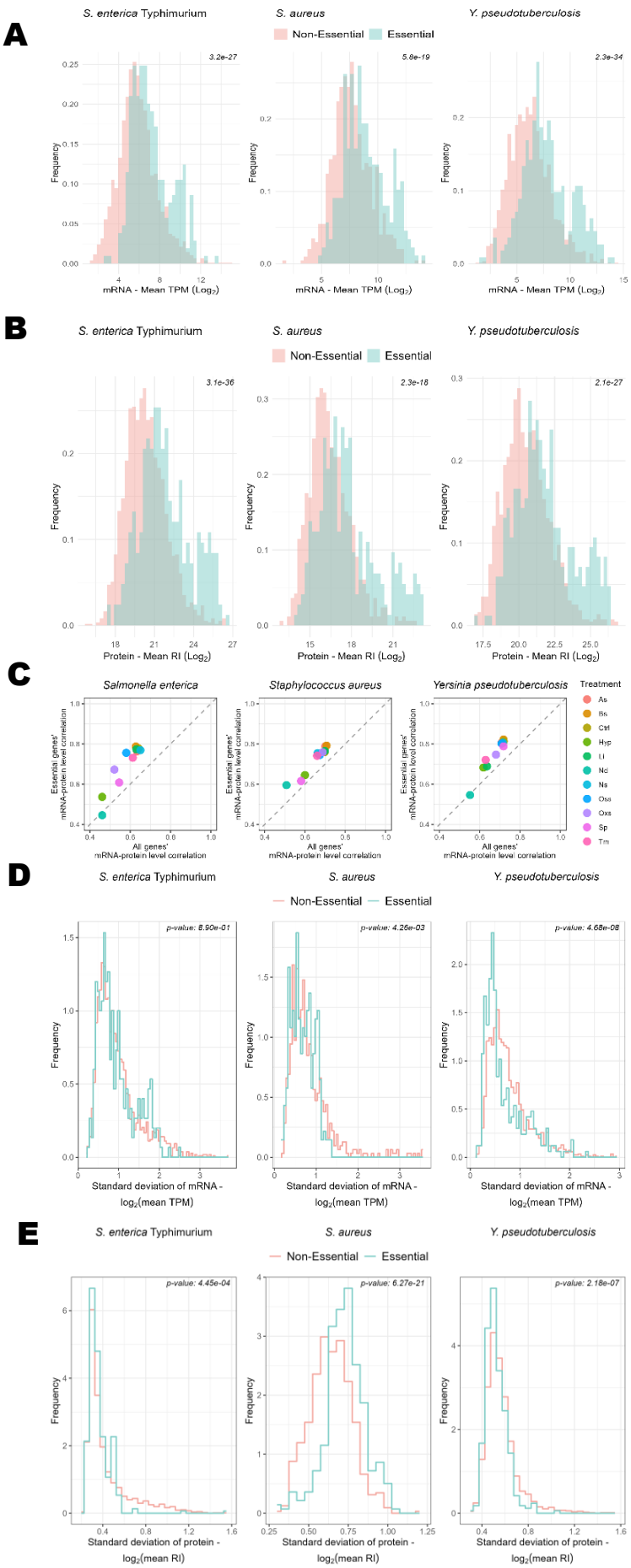


Figure 2. mRNA-protein levels show differing degree of statistically significant positive correlation under different stress conditions and control condition in A) *S. enterica* Typhimurium, **B)** *S. aureus* and **C)** *Y. pseudotuberculosis*. mRNA and protein expression levels were taken as log2 transformed TPM and RI values, respectively. The expression values of genes that are detected in all conditions, as mRNA and protein, were used in this analysis. Pearsson correlation coefficient (R) shows positive correlation under all conditions in the three species. *p* values were calculated using Student's t-test.

727


A
S. enterica Typhimurium
S. aureus
Y. pseudotuberculosis
Non-Essential
Essential
3.2e-27
5.8e-19
2.3e-34
Frequency
mRNA - Mean TPM (Log2)
B
3.1e-36
2.3e-18
2.1e-27
Protein - Mean RI (Log2)
C
Salmonella enterica
Staphylococcus aureus
Yersinia pseudotuberculosis
Treatment
As
Bs
Ctrl
Hyp
Li
Nd
Ns
Oss
Oxs
Sp
Tm
Essential genes' mRNA-protein level correlation
All genes' mRNA-protein level correlation
D
p-value: 8.90e-01
p-value: 4.26e-03
p-value: 4.68e-08
Standard deviation of mRNA - log2(mean TPM)
E
p-value: 4.45e-04
p-value: 6.27e-21
p-value: 2.18e-07
Standard deviation of protein - log2(mean RI)

728 **Figure 3. Essential genes of the three pathogens have higher mRNA and protein**
729 **expressions with higher correlation and lower variance across the 11 conditions in**
730 **comparison to all genes. A)** Mean mRNA expression frequency of essential and non-
731 essential genes in the 11 conditions. **B)** Mean protein expression frequency of essential and
732 non-essential genes in the 11 conditions. **C)** Essential genes and all genes mRNA-protein
733 level correlation under 11 conditions for the three species. **D)** Frequency of standard deviation
734 of mRNA expression under 11 conditions for the three species. **E)** Frequency of standard
735 deviation of protein expression under 11 conditions for the three species. mRNA-protein level
736 correlations were measeured with Pearsson correlation coefficient (R). *p* values were
737 calculated using a Wilcoxon rank sum test.

738

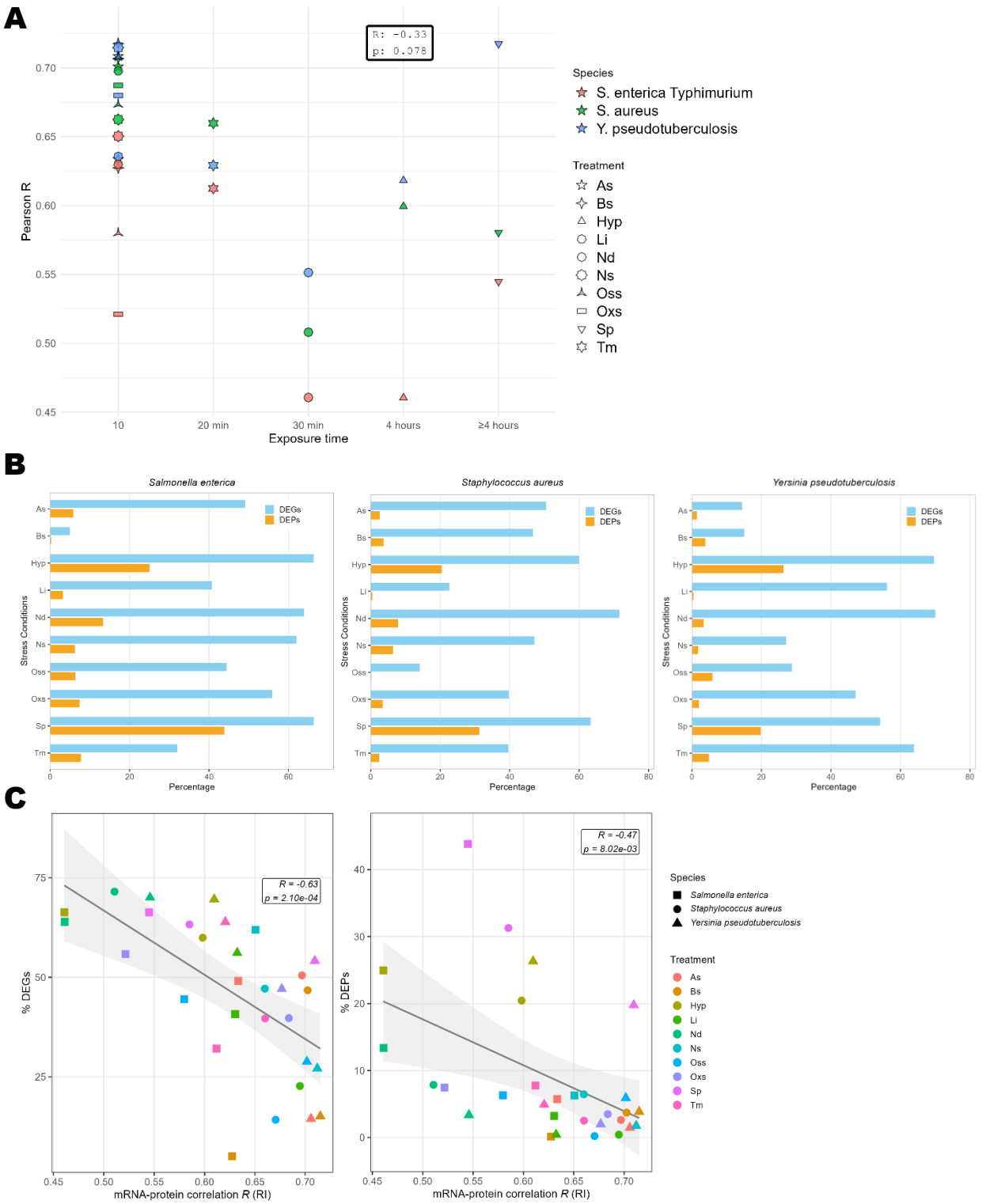


Figure 4. Degree of stress response lowers mRNA-protein level correlations in the three species. A) mRNA-protein level correlations for conditions exposed to stressors for 10 minutes, 20 minutes, 30 minutes, 4 hours, and longer than 4 hours in the three species. **B)** Percentages of differentially expressed genes (DEGs) and proteins (DEPs) under the 10 stress conditions in comparison to unexposed control conditions for the three species. **C)** Scatter plot showing negative correlation between percentages of DEGs (top) and DEPs (bottom) and mRNA-protein level correlation. Correlations were measured with Pearsson correlation coefficient (R). *p*-values were calculated using Student's t-test.

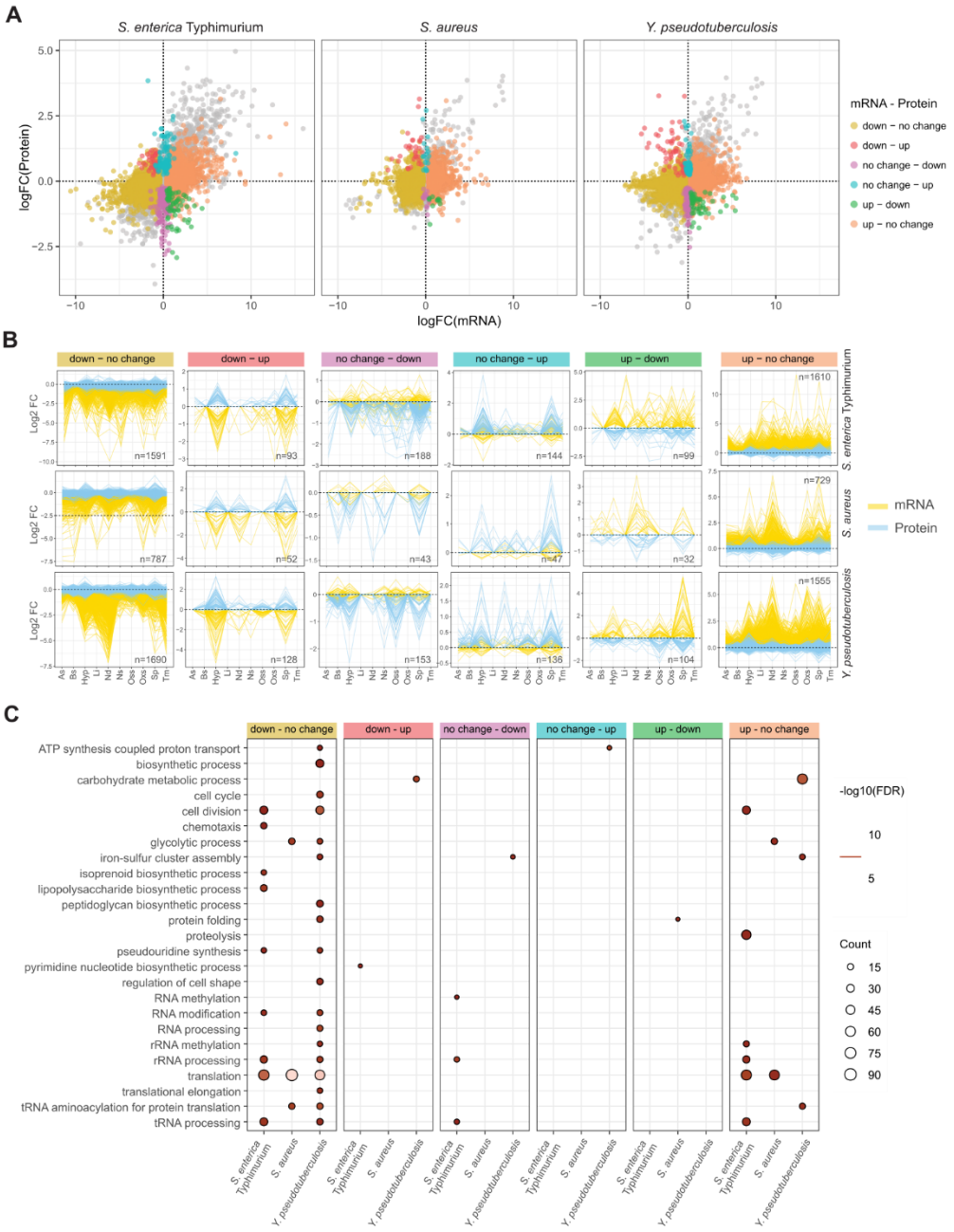


Figure 5. Stress response induces different patterns of discordant mRNA and protein level regulation highlighting key role for translation. A) Scatter plots with fold changes of mRNA and protein levels for each the 10 stress conditions in comparison to control for the three species. The 6 gene clusters whose mRNA and protein level have different discordant regulation patterns are shown in different colors. **B)** mRNA and protein log2 fold change patterns of the 6 discordantly regulated gene clusters under 10 stress conditions for the three species. The number of discordantly regulated genes in each cluster are indicated with 'n' **C)** GO biological process enrichment analysis of 6 discordantly regulated gene clusters for the three species.

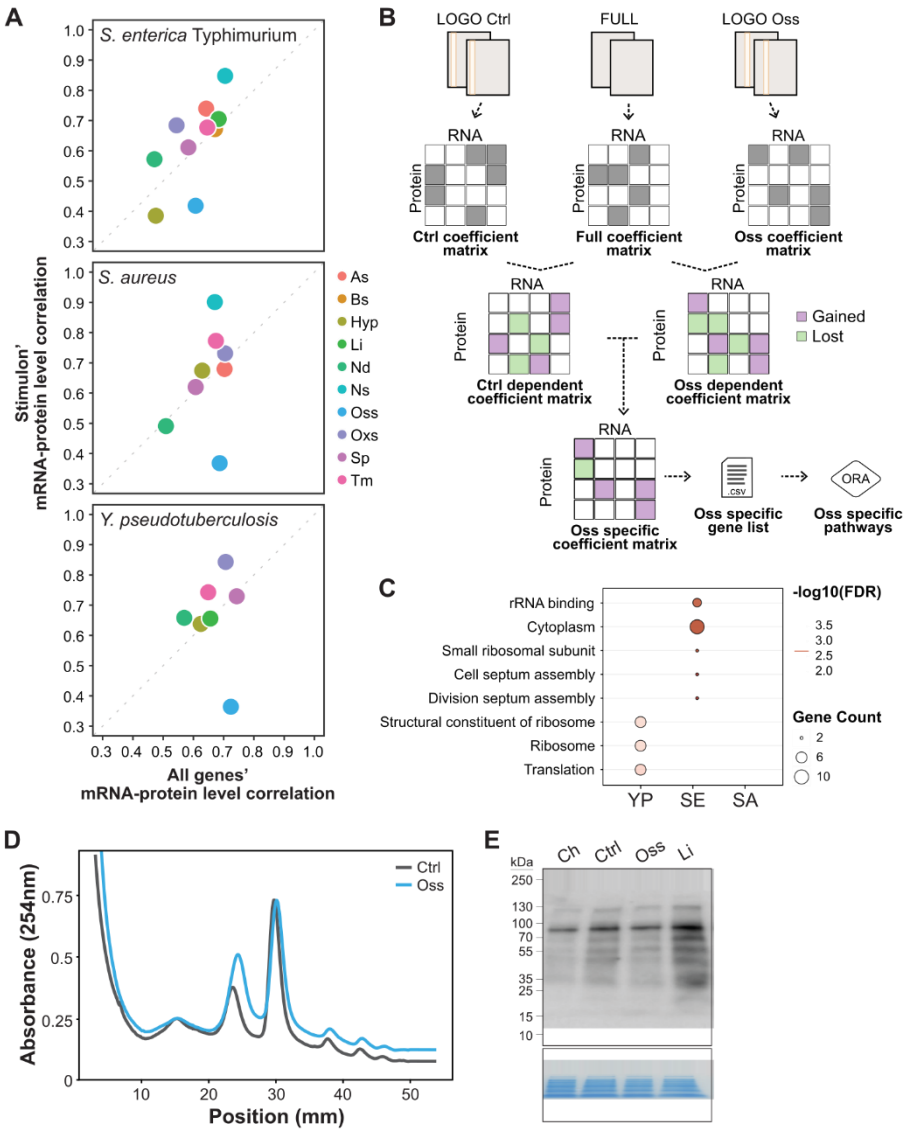


Figure 6. Impaired translation efficiency lowers mRNA-protein level under osmotic stress. A) Stress specific stimulon genes and all genes mRNA-protein level correlation under 10 conditions for the three species. **B)** MOBILE framework to identify osmotic stress specific associations between mRNA and protein levels. **C)** ORA of osmotic stress-specific mRNA–protein associations across three species. **D)** Polysome profiles of wt *Y. pseudotuberculosis* cells under control (Ctrl) and osmotic stress (0.5M NaCl) conditions. **E)** Coomassie blue-stained gel (upper panel) and western blot membrane (lower panel) of total proteins extracted from wt *Y. pseudotuberculosis* cells under negative control (Ch), positive control (Ctrl), osmotic stress (Oss), and low iron (Li). Active translation of nascent proteins was detected using a monoclonal anti-puromycin antibody