| Systems Biology | Research Article

Ribosome profiling reveals the fine-tuned response of *Escherichia coli* to mild and severe acid stress

Kilian Schumacher,[1] Rick Gelhausen,[2] Willow Kion-Crosby,[3,4] Lars Barquist,[3,4] Rolf Backofen,[2,5] Kirsten Jung[1]

AUTHOR AFFILIATIONS See affiliation list on p. 29.

ABSTRACT   The ability to respond to acidic environments is crucial for neutralophilic bacteria. *Escherichia coli* has a well-characterized regulatory network that triggers a multitude of defense mechanisms to counteract excess protons. Nevertheless, systemic studies of the transcriptional and translational reprogramming of *E. coli* to different degrees of acid stress have not yet been performed. Here, we used ribosome profiling and RNA sequencing to compare the response of *E. coli* (pH 7.6) to sudden mild (pH 5.8) and severe near-lethal acid stress (pH 4.4) conditions that mimic passage through the gastrointestinal tract. We uncovered new differentially regulated genes and pathways, key transcriptional regulators, and 18 novel acid-induced candidate small open reading frames. By using machine learning and leveraging large compendia of publicly available *E. coli* expression data, we were able to distinguish between the response to acid stress and general stress. These results expand the acid resistance network and provide new insights into the fine-tuned response of *E. coli* to mild and severe acid stress.

IMPORTANCE   Bacteria react very differently to survive in acidic environments, such as the human gastrointestinal tract. *Escherichia coli* is one of the extremely acid-resistant bacteria and has a variety of acid-defense mechanisms. Here, we provide the first genome-wide overview of the adaptations of *E. coli* K-12 to mild and severe acid stress at both the transcriptional and translational levels. Using ribosome profiling and RNA sequencing, we uncover novel adaptations to different degrees of acidity, including previously hidden stress-induced small proteins and novel key transcription factors for acid defense, and report mRNAs with pH-dependent differential translation efficiency. In addition, we distinguish between acid-specific adaptations and general stress response mechanisms using denoising autoencoders. This workflow represents a powerful approach that takes advantage of next-generation sequencing techniques and machine learning to systematically analyze bacterial stress responses.

T he infective dose of enteropathogens varies significantly among bacterial genera and is dependent on the number and complexity of acid resistance mechanisms (1). *Escherichia coli* is equipped with a high number of defense mechanisms to survive the acidity of the stomach and, correspondingly, can have an infective dose as low as less than 50 cells (2). Enterobacteria that survive the stomach also confront mild acid stress in the colon due to the presence of short-chain fatty acids produced by obligate anaerobes (3). Other neutralophilic bacteria encounter low pH environments in a variety of settings, such as acidic soils, fermented food, or phagosomes within macrophages (1).

The cytoplasmic membrane represents a primary barrier for protons (H^+). Nevertheless, at low pH, H^+ can permeate into the cytoplasm via protonated water chains, ion channels, or damaged membranes (4). Upon acidification, the cytoplasmic pH transiently

decreases but returns to neutral within short time intervals. For example, a reduction in external pH to 5.5 causes a temporary decline in internal pH to ~6.0 (5). Intracellular acidification leads to the protonation of biological molecules, ultimately altering their charge and structure. This, in turn, may lead to protein unfolding, denaturation, and reduced enzymatic activities (6). Furthermore, acidification causes membrane and DNA damage (7, 8).

To maintain a constant internal pH and balance fluctuations in H^+ concentrations, the intrinsic buffering capacity of the cytoplasm is crucial, as protons can be sequestered by side-chains of proteins, inorganic phosphates, polyphosphates, or polyamines (9). Additional protective mechanisms that counteract acid stress and ensure survival in low pH habitats involve proton pumps, membrane remodeling, acid-dependent chemotaxis, chaperones, acid shock proteins, and the induction of enzyme-based H^+-consuming acid resistance (AR) systems (10, 11).

E. coli is equipped with four different AR systems, namely, the glutamate decarboxylase (Gad, AR2), arginine decarboxylase (Adi, AR3), lysine decarboxylase (Cad, AR4), and ornithine decarboxylase (Orn, AR5) systems (12, 13). Each AR system consists of an H^+-consuming amino acid decarboxylase (GadA/GadB, AdiA, CadA, SpeF) and a corresponding antiporter (GadC, AdiC, CadB, PotE), which serves to uptake the amino acid and export the more alkaline reaction product into the surrounding medium. This strategy ensures a simultaneous increase in intracellular and extracellular pH (7, 13, 14). Notably, each AR system is activated at different external pH values and growth phases (7, 15). We have previously found that individual *E. coli* cells exposed to consecutively increasing acid stress activate the Gad, Adi, and Cad systems heterogeneously, resulting in functional diversification and enhanced population fitness (15, 16). The regulatory network of the Gad system is highly complex and involves more than 10 regulatory elements (17), whereas the Adi and Cad systems are each regulated by a single transcriptional regulator, AdiY and CadC, respectively (15). Nevertheless, additional regulatory elements seem to exist that not only link these three ARs in individual cells but also contribute to a fine-tuned response of *E. coli* to different levels of acid stress (15).

Previous systematic studies to determine genome-wide adaptations within the acid stress response of *E. coli* have either been restricted to microarrays (18), two-dimensional gel electrophoresis (19, 20), or have only been performed under mild acidic conditions (pH 5.0–6.0) (21–23). Here, we present the first systemic study of global adaptations to different intensities of acid stress in *E. coli* at both the transcriptional and translational levels. Specifically, we compare gene expression and translation at pH 5.8 and pH 4.4 with pH 7.6, using RNA sequencing (RNA-Seq) and ribosome profiling (Ribo-Seq). Ribo-Seq allows deep-sequencing of ribosome-protected mRNA fragments (RPFs) (24, 25), which are obtained by nuclease digestion of non-ribosome-covered RNA, and has significantly advanced the current understanding of translational regulation (26). The applications of Ribo-Seq are manifold and include monitoring protein synthesis rates across the proteome, the identification of novel small open reading frames (sORFs), as well as determining protein copy numbers per cell during steady-state growth (27, 28). In contrast to mass spectrometry-based approaches, ribosome profiling is independent of protein biochemistry and allows detection of small proteins with less than 50 amino acids in length (29). In recent years, substantial advances in understanding translational events have been made through the utilization of Ribo-Seq, not only in various bacterial species but also in archaea, bacteriophages, and microbiomes (28, 30–35).

Our results reveal hundreds of differentially transcribed and translated genes of *E. coli* K-12 at pH 5.8 and 4.4, as well as examples of pH-dependent changes in translation efficiency. We identified previously undiscovered biological processes in response to acidity, including increased siderophore synthesis, glycerol catabolism, copper efflux, nucleotide biosynthesis, and spermidine/multidrug export, as well as decreased membrane transport and metabolic activities. In addition, we have identified new transcription factors (TFs) as key players during low pH exposure and 18 novel candidate sORFs involved in the response to acid stress. Finally, we differentiated

acid-specific transcriptional adaptations by using machine learning to compare the low pH response to that of other stressors.

RESULTS AND DISCUSSION

Examination of alterations in the translatome and transcriptome of *E. coli* in response to varying degrees of acid stress

To mimic natural stress conditions, such as the passage of *E. coli* through the gastrointestinal tract, we established the following protocol involving a sudden change to low pH, detection of a rapid response, and severe, near-lethal acid stress. Specifically, *E. coli* K-12 MG1655 was cultivated in unbuffered lysogeny broth (LB) medium at pH 7.6 until the exponential growth phase ($OD_{600} = 0.5$). Then, 5 M hydrochloric acid was added directly to expose the cells to a pH of 5.8 or stepwise to a pH of 4.4, corresponding to mild and severe acid stress (Fig. 1A). The final optical densities were comparable (pH 7.6: OD_{600} =
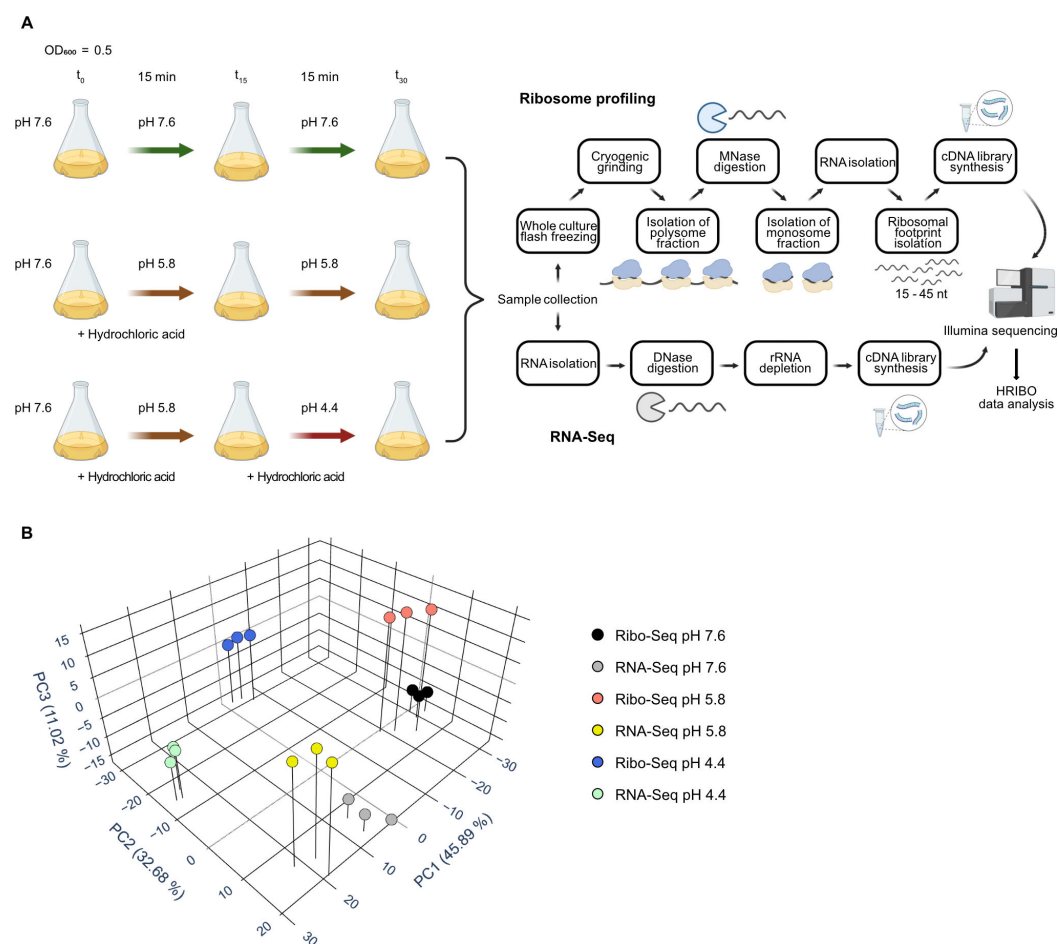

FIG 1 A schematic overview of culture conditions, pH-shift procedures, and the Ribo-Seq and RNA-Seq workflows. (A) Biological triplicates of *E. coli* MG1655 cells were grown in unbuffered LB medium (pH 7.6) to an OD_{600} of 0.5 (t_0). Subsequently, the cultures were either grown for an additional 30 min at pH 7.6, shifted for 30 min to pH 5.8, or shifted first for 15 min to pH 5.8 (t_{15}) and then to pH 4.4 (t_{30}). pH shifts were initiated by the direct addition of 5 M HCl to the cultures. At t_{30}, samples were collected. For RNA-Seq, total RNA was isolated, DNase digested, and ribosomal RNA was depleted prior to cDNA library preparation. For Ribo-Seq, whole cultures were flash-frozen in liquid nitrogen and cryogenically grinded using a freezer mill. Polysome fractions were isolated, and non-ribosome-protected RNA was digested using MNase. Ribosomal footprints were purified from the monosome fraction using a size selection gel and converted to cDNA libraries. Upon Illumina sequencing, the data were analyzed using the HRIBO pipeline (37). (B) Principal component analysis of median-of-ratios normalized and regularized log-transformed (rlog) read count values for *E. coli* Ribo-Seq and RNA-Seq triplicate data at pH 4.4, 5.8, and 7.6. Panel (A) was created with BioRender.com.

~1.1; pH 4.4: OD_{600} = ~0.7) (Table S1), and the pH values hardly changed compared to t_0 (Fig. 1A; Table S2). To investigate whether a 15-min exposure to pH 4.4 was sufficient to induce cellular adaptation to severe acid stress, we examined the temporal dynamics of *adiA* expression by quantitative reverse transcription PCR (RT-qPCR). We detected an increase in *adiA* mRNA levels as early as 15 min after the shift to pH 4.4 and no substantial further increase after 30 or 60 min (Fig. S1). Additionally, we evaluated cell viability at the final experimental time points using propidium iodide (PI) staining (36) and examined colony-forming units (CFUs). The average percentage of dead cells detected by PI staining was less than 1% at pH 7.6, 5% at pH 5.8, and 18% at pH 4.4 (Fig. S2A). A positive control (5-min heat treatment at 80°C) resulted in an average of 97.2% non-viable cells, as determined by PI staining. The CFU count results underline that at least 2×10^8 viable cells were collected irrespective of pH at the moment of sample collection for Ribo- and RNA-Seq (Fig. S2B).

Cells were harvested and lysed as previously described by whole-culture flash freezing and cryogenic grinding in a freezer mill to avoid bias from translation-arresting drugs and filtering (38). The subsequent steps of our ribosome profiling protocol were a combination of methodologies reported by Latif and colleagues (39) and Mohammad and Buskirk (40) (see Materials and Methods for details). Strand-specific Illumina sequencing yielded an average of approximately 30 million cDNA reads per sample for Ribo-Seq and 5–10 million for RNA-Seq. The next-generation sequencing data were analyzed using an extended version of the high-throughput HRIBO data analysis pipeline (37). All samples achieved sufficient coverage with over two million reads, each mapping uniquely to the coding regions. The rRNA contamination was higher in the pH 4.4 Ribo-Seq samples than in other conditions but accounted for less than 15% in all cDNA libraries (Fig. S3). The length distribution of the generated RPFs was broad, ranging from 15 to 45 nucleotides (Fig. S4), consistent with previous observations in other prokaryotic ribosome profiling analyses (38, 39). We did not detect stress-induced increased relative ribosome occupancy in the initiation region of ORFs under acid stress compared to neutral pH. This is in contrast to previous observations in *E. coli* under heat stress (41) and in yeast under oxidative stress (42), which reported increased relative ribosome accumulation at start codons in stressed cells. In fact, ribosome occupancy in the translation initiation regions was slightly reduced at pH 4.4 and 5.8 compared with physiological pH (Fig. S5). This could be explained due to diminished ribosome-RNA complex stability and increased ribosome drop-off under acidic conditions. The biological triplicates for each experimental condition clustered on the first three principal components in a principle component analysis (PCA) plot (Fig. 1B). Notably, the global gene expression profiles were highly distinct at pH 4.4 compared with both pH 5.8 and 7.6.

Coordinated regulation of transcription and translation in response to acid stress

The tool *deltaTE* (43) was used to assess transcriptional and translational changes (i.e., differential expression and differential translation efficiency) in response to mild and severe acid stress. Low-expression transcripts were filtered out, and we focused our analysis on 3,654 genes with mean reads per kilobase per million reads mapped (rpkm) values ≥5 across all investigated conditions. Our findings reveal that 702 transcripts were significantly altered at pH 5.8 compared with physiological pH [absolute mRNA $\log_2$ fold change (FC) ≥1 and false discovery rate (FDR) adjusted $P \leq 0.05$], and 1,030 genes showed significant differences in mRNA levels at pH 4.4 (Fig. 2A). These results suggest that extensive transcriptional reprogramming occurred, which was influenced by the degree of acid stress. As illustrated by the Venn diagram overlaps (Fig. 2), a large number of adaptations occurred regardless of the degree of acid stress. Nonetheless, several hundred genes were differentially expressed exclusively at pH 5.8 or 4.4 (Fig. 2A). This suggests that in addition to universal adaptations at low pH, specific adaptations for mild and severe acid stress occur. We further determined the number of genes with
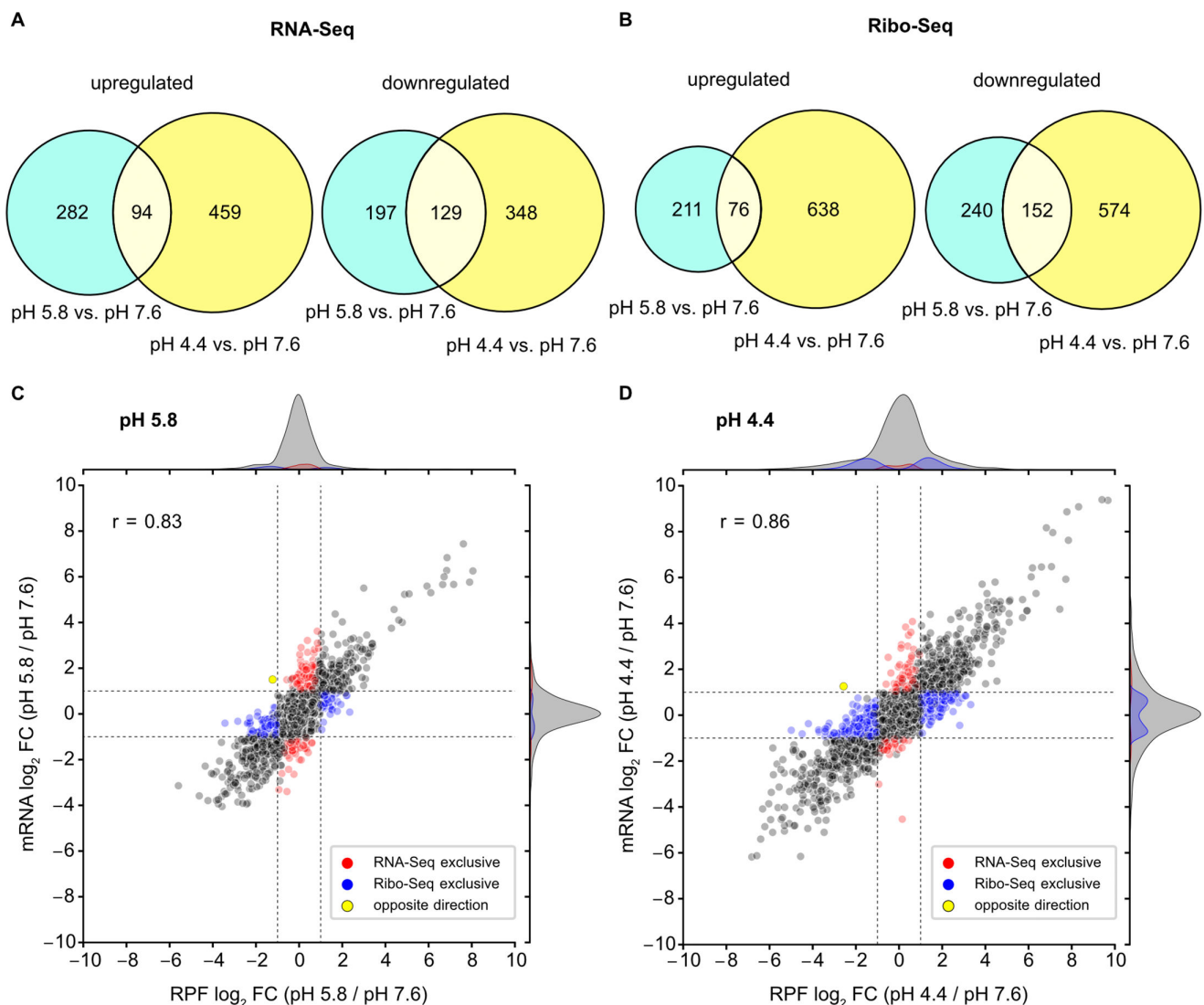
FIG 2 Genome-wide adaptations correlate at the transcriptional and translational levels in *E. coli* under acid stress. Weighted Venn diagrams show the total number and overlap of genes with significant FCs (absolute $\log_2$ FC ≥1 and *P*-adjust ≤0.05) determined by (A) RNA-Seq or (B) Ribo-Seq for cells exposed to pH 5.8 or pH 4.4, compared to the control (pH 7.6). (C) Comparison of global RPF and mRNA $\log_2$ FC values for pH 5.8, or (D) pH 4.4 vs pH 7.6. Dashed lines indicate $\log_2$ fold change values of +1 or −1. Hundreds of genes exhibited differential expression (absolute $\log_2$ FC ≥1 and *P*-adjust ≤0.05) at both the transcriptional and translational levels, whereas others were exclusively detected by either RNA-Seq (red dots) or Ribo-Seq (blue dots) or had significant changes in opposite directions (yellow dots). Values of the Pearson correlation coefficient (*r*) are indicated.

stress-dependent alterations in RPF counts to be 679 at pH 5.8 and 1,440 at pH 4.4 (absolute RPF $\log_2$ FC ≥1 and FDR adjusted $P \leq 0.05$), which was in a similar range compared with the RNA-Seq data (Fig. 2A and B). Accordingly, the global FC values for mRNA and RPF levels showed a high Pearson correlation coefficient (*r*) under both conditions (Fig. 2C and D, gray dots). This indicates that transcriptional regulation of these genes is the predominant response to acid stress. However, a subset of genes exhibited exclusive and significant regulation at either the transcriptional (red dots) or translational (blue dots) level.

Specifically, at pH 5.8, 193 genes were detected to be significantly regulated exclusively by RNA-Seq, while 216 genes were exclusively affected in the Ribo-Seq data (Fig. 2C; Table S3). At pH 4.4, 127 differentially regulated genes were found exclusively by RNA-Seq and 570 genes by Ribo-Seq (Fig. 2D; Table S3). Notably, for *fruA* at pH 5.8 and *yecH* at pH 4.4, opposite changes were observed at the transcriptional and translational

levels (Fig. 2C and D, yellow dots). FruA is the fructose permease of the phosphoenolpyruvate-dependent sugar phosphotransferase system (44), whereas the function of YecH remains unknown.

Next, we investigated translation efficiency (TE) to identify genes that undergo translational regulation in response to acidic conditions. TE provides information regarding ribosome counts per mRNA and is calculated as the ratio of RPFs over transcript counts within a gene's coding sequence normalized to mRNA abundance (43). We identified 22 genes at pH 5.8 and 89 genes at pH 4.4, which displayed significantly altered TEs (absolute $\log_2$ TE fold change ≥1 and P-adjust ≤0.05) (Table S4). The highest increase in TE at pH 4.4 was found for the KpLE2 phage-like element (*topAI*), a hydroxyethylthiazole kinase (*thiM*), and a palmitoleoyl acyltransferase (*lpxP*). In contrast, *yecH* and *yjbE*, both encoding uncharacterized proteins, and *malM* of the maltose regulon showed the most prominent decrease in TE at pH 4.4 (Table S4). At pH 5.8, we noted the largest increase in TE for a ferredoxin-type protein encoded by *napF*, an iron transport protein (*feoA*), and a tagaturonate reductase (*uxaB*). Conversely, the largest decrease was observed for a protein of the fructose-specific phosphotransferase system (*fruA*), a tripartite efflux pump membrane fusion protein (*emrK*), and an HTH-type transcriptional regulator (*ydeO*) (Table S4).

In summary, besides extensive transcriptional reprogramming, dozens of genes exhibit significant FCs either at the transcriptional or translational level in response to acid stress. This underlines that transcription and translation are not always coupled in bacteria. Similar findings were reported by Zhang and colleagues (41), who conducted Ribo-Seq and RNA-Seq analyses for *E. coli* under heat stress (41). Overall, such differential regulation can be explained, for example, by delayed translation relative to transcript synthesis, selective recruitment or release of ribosomes, or regulation during translation initiation, elongation, or ribosome biogenesis (45–49), which could be beneficial under stress conditions.

Functional implications of genes with differential mRNA and RPF levels under mild acid stress

To obtain a more profound understanding of the fine-tuned response of *E. coli* to different degrees of acid stress, we first analyzed all genes with differential mRNA and ribosome coverage levels during mild acid stress (pH 5.8). Under this condition, the top candidates with the highest FC values for mRNA and RPF are as follows: (i) the *cad* operon, encoding the core components of the Cad AR system (see also below); (ii) the *glp* regulon, responsible for glycerol and *sn*-glycerol 3-phosphate uptake and catabolism (50); (iii) the *mdtJI* operon, encoding a heterodimeric multidrug/spermidine exporter (51); and (iv) genes encoding proteins involved in motility and flagella biosynthesis (Table 1). A comprehensive list of normalized read counts, mRNA and RPF FCs, and TEs for all *E. coli* genes is provided in Table S5. We tested a representative selection of differentially expressed genes by RT-qPCR. In all cases, the detected changes in mRNA levels were consistent with the data gathered by RNA-Seq (Fig. S6A). Both *recA* and *secA* were chosen as reference genes for RT-qPCR because their rpkm counts were relatively constant under the conditions tested (Table S5; Fig. S6B).

Next, we performed gene set enrichment analysis (GSEA) using *clusterProfiler* (52) to identify biological processes associated with differentially expressed genes at pH 5.8. Among the most enriched Gene Ontology (GO) terms for biological processes at pH 5.8 was "spermidine transmembrane transport" (Fig. 3), which corresponds to the induction of *mdtJI* (Table 1) and a polyamine ABC transporter encoded by *potABCD* (Table S5). Polyamines are crucial for survival under acid stress, as they reduce membrane permeability by blocking OmpF and OmpC porins (53–55). External spermidine supplementation also improved acid resistance in *Streptococcus pyogenes* (56). On the other hand, overaccumulation of polyamines can be toxic and potentially lethal for *E. coli* (51, 57). Therefore, precise transmembrane transport of polyamines in acidic environments is critical and contributes to survival in acidic conditions.

TABLE 1 Top 20 genes with increased RPF levels at pH 5.8 compared to pH 7.6, sorted in descending order by Ribo-Seq log$_2$ FC values[a]

Gene name	Log$_2$ FC (Ribo)	P-adjust (Ribo)	Log$_2$ FC (RNA)	P-adjust (RNA)	Cellular location	Annotation
glpA	8.05	1×10^{-159}	6.26	2×10^{-38}	C	Anaerobic glycerol-3-phosphate dehydrogenase subunit A
glpB	7.89	2×10^{-141}	5.77	6×10^{-29}	C	Anaerobic glycerol-3-phosphate dehydrogenase subunit B
cadA	7.61	2×10^{-144}	7.44	3×10^{-57}	C	Lysine decarboxylase 1
glpC	7.16	5×10^{-89}	5.67	4×10^{-35}	C, IM	Anaerobic glycerol-3-phosphate dehydrogenase subunit C
cadB	6.86	8×10^{-96}	6.85	5×10^{-35}	IM	Lysine:cadaverine antiporter
glpT	6.84	2×10^{-105}	6.28	2×10^{-39}	IM	sn-Glycerol 3-phosphate:phosphate antiporter
glpD	6.72	4×10^{-126}	6.00	2×10^{-52}	C, IM	Aerobic glycerol 3-phosphate dehydrogenase
mdtJ	6.65	2×10^{-67}	5.67	6×10^{-36}	IM	Multidrug/spermidine efflux pump membrane subunit MdtJ
glpQ	6.11	2×10^{-89}	5.31	2×10^{-29}	P	Glycerophosphoryl diester phosphodiesterase
mdtI	5.93	3×10^{-49}	5.60	9×10^{-34}	IM	Multidrug/spermidine efflux pump membrane subunit MdtI
glpF	5.10	7×10^{-64}	5.26	1×10^{-19}	IM	Glycerol facilitator
yfdX	4.89	2×10^{-24}	5.23	4×10^{-25}	P	Protein YfdX
frc	4.78	1×10^{-36}	4.01	2×10^{-15}	C	Formyl-CoA transferase
glpK	4.62	9×10^{-54}	4.10	1×10^{-25}	C	Glycerol kinase
yfdV	4.40	3×10^{-18}	4.58	1×10^{-12}	IM	Putative transport protein YfdV
oxc	4.28	2×10^{-28}	3.75	1×10^{-14}	C	Xalyl-CoA decarboxylase
flgC	3.41	3×10^{-22}	2.95	3×10^{-9}	P	Flagellar basal-body rod protein FlgC
flgB	3.40	1×10^{-24}	3.09	1×10^{-10}	P	Flagellar basal-body rod protein FlgB
dhaL	3.39	9×10^{-13}	3.08	1×10^{-10}	C	Dihydroxyacetone kinase subunit L
flgD	3.33	7×10^{-25}	2.90	5×10^{-9}	C	Flagellar biosynthesis, initiation of hook assembly

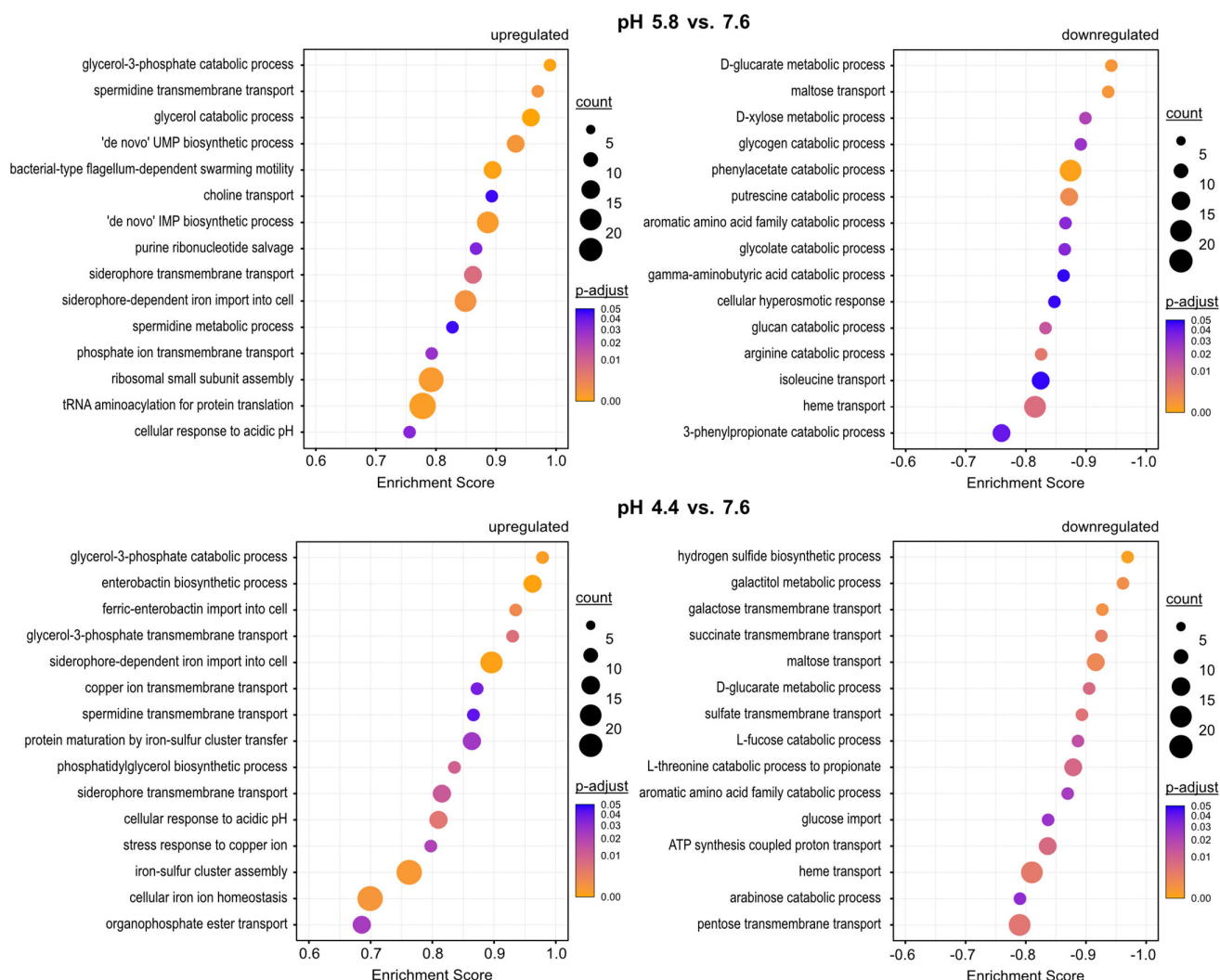[a]C, cytosol; P, periplasm; IM, inner membrane.

FIG 3 Up- and downregulated biological processes in *E. coli* at pH 5.8 and 4.4. GSEA was conducted using the *gseGO* function in the *clusterProfiler* package (52) with the ribosome profiling differential expression data sorted by log$_2$ fold change values as input. GO terms were considered up- or downregulated if *P*-adjust values were ≤0.05. The top 15 non-redundant GO terms were sorted in descending order by the *clusterProfiler* enrichment score and are shown for pH 5.8 vs 7.6 and pH 4.4 vs 7.6. The dot size represents the number of genes associated with each GO term, and the dot color represents adjusted *P*-values corrected for the false discovery rate.

The enrichment of the GO terms "glycerol-3-phosphate catabolic process" and "glycerol catabolic process" at pH 5.8 (Fig. 3) has not yet been associated with acid stress to our knowledge. Notably, of the 14 genes with the largest increase in RPF counts at pH 5.8, 7 belong to the *glp* regulon (Table 1). This regulon is required for the uptake and catabolism of glycerol and *sn*-glycerol 3-phosphate (G3P) (50). In this pathway, G3P is converted to dihydroxyacetone phosphate by membrane-bound dehydrogenases, either aerobically via GlpD or anaerobically by the GlpABC complex (58, 59). Alternatively, dihydroxyacetone phosphate can be produced directly from glycerol by GldA and the protein products of the *dhaKLM* operon (60). The *dhaKLM* operon was also induced at pH 5.8 (Table S5). It remains unclear whether glycerol and G3P catabolism directly contribute to acid tolerance or whether the *glp* regulon is activated as a consequence of other low pH adjustments. Expression of *glp* genes is regulated by the repressor GlpR, which is inactivated upon binding of glycerol or G3P (61). We hypothesize that changes in phospholipid composition under acid stress conditions (62) may release G3P, which in

turn induces the *glp* regulon. Accordingly, the GO term "phosphatidylglycerol biosynthetic process" was enriched under acid stress (Fig. 3).

Another observation is the upregulation of *de novo* biosynthesis pathways for pyrimidine and purine nucleotides at pH 5.8 (Fig. 3). The induction of a large proportion of the PurR-dependent regulon involved in *de novo* nucleotide synthesis (Fig. 3; Table S5) suggests that *E. coli* requires additional nucleotides to cope with the extensive transcriptional reprogramming. Besides, intracellular acidification can lead to DNA damage, such as depurination (63), making enhanced nucleotide biosynthesis a critical compensatory mechanism. Recently, *Oenococcus oeni* was reported to experience a decrease in the abundance of both purines and pyrimidines under acid stress, while nucleotide metabolism and transport increased (64), suggesting a similar phenomenon in this species. Other enriched GO terms under mild acid stress include "choline transport," "siderophore transmembrane transport," "phosphate ion transmembrane transport," "ribosomal small subunit assembly," "tRNA aminoacylation for protein translation," and "bacterial-type flagellum-dependent swarming motility" (Fig. 3). pH-dependent motility has previously been observed in *E. coli*, *Salmonella*, and *Helicobacter* (1). These observations suggest that bacterial cells use an escape strategy to migrate to more favorable pH environments when challenged with acidic conditions.

On the contrary, our findings reveal that many membrane and periplasmic proteins (18 of the 20 genes with the most diminished RPF counts, Table 2) were among the top candidates with decreased mRNA and RPF levels under mild acid stress. This affected, for example, genes encoding ABC transporters (*mal* regulon, *dpp* operon) and symporters (*actP*, *melB*, *gabP*), highlighting the superiority of Ribo-Seq over mass spectrometry-based approaches, namely, its independence of protein biochemistry and higher sensitivity (29). Furthermore, GSEA identified membrane transport and metabolic activities as the most downregulated biological processes in response to mild acid stress. For example, "maltose transport," "isoleucine transport," "heme transport," "putrescine catabolic process," "glycolate catabolic process," and "aromatic amino acid family catabolic process" were among the most downregulated GO terms at pH 5.8 (Fig. 3). Downregulation of H^+-coupled transport processes represents a key mechanism by which *E. coli* restricts proton influx into cells. In addition, the downregulated metabolic processes are in many cases associated with the synthesis and conversion of amino acids and carbon sources. For example, the catabolism of aromatic amino acids and arginine was also reduced at pH 5.8 (Fig. 3). Particularly noteworthy is the downregulation of the arginine catabolic pathway, which involves the protein products of the *astEBDAC* operon. At pH 5.8, hardly any reads were mapped in the *astEBDAC* region, despite detectable expression at pH 7.6 and pH 4.4 (Table S5). Presumably, *E. coli* preserves the intracellular arginine pool at pH 5.8, as this amino acid serves as a substrate for the Adi system during severe acid stress (15, 65).

In summary, the response of *E. coli* to mild acid stress is characterized by the activation of the motility machinery to escape to less acidic habitats, by induction of the *cad* operon, and by genes involved in polyamine transport and glycerol-3-phosphate conversion (Tables 1 and 2; Fig. 3). In addition, *E. coli* restricts the influx of protons and conserves energy by reducing its metabolic activities.

Functional implications of genes with differential mRNA and RPF levels under severe acid stress

Next, we analyzed genes with differential mRNA and ribosome coverage levels in response to severe acid stress (pH 4.4) compared with non-stress (pH 7.6). Genes with the highest number of increased read counts, which were not already upregulated at pH 5.8, were *asr*, encoding an acid shock protein, followed by *bdm*, encoding a biofilm-modulation protein, and *bhsA*, encoding a multiple stress resistance outer membrane protein (Table 3). Originally, Asr was classified as a periplasmic acid shock protein, although its role in acid adaptation remained unclear (66). Recently, Asr was shown to be an intrinsically disordered chaperone that contributes to outer membrane integrity and to

TABLE 2 Top 20 genes with decreased RPF levels at pH 5.8 compared to pH 7.6, sorted in ascending order by Ribo-Seq log$_2$ FC values[a]

Gene name	Log$_2$ FC (Ribo)	P-adjust (Ribo)	Log$_2$ FC (RNA)	P-adjust (RNA)	Cellular location	Annotation
yjcH	-5.59	3×10^{-27}	-3.14	1×10^{-8}	IM	Conserved inner membrane protein YjcH
actP	-4.63	2×10^{-47}	-3.58	1×10^{-11}	IM	Acetate/glycolate:cation symporter
lamB	-4.32	3×10^{-6}	-2.27	6×10^{-6}	OM	Maltose outer membrane channel/phage lambda receptor protein
dppD	-4.23	5×10^{-33}	-3.79	9×10^{-15}	IM	Dipeptide ABC transporter ATP-binding subunit DppD
acs	-4.06	3×10^{-20}	-2.60	3×10^{-6}	C	Acetyl-CoA synthetase (AMP-forming)
malK	-3.96	2×10^{-18}	-2.64	2×10^{-10}	IM	Maltose ABC transporter ATP-binding subunit
dppC	-3.83	1×10^{-28}	-3.78	5×10^{-13}	IM	Dipeptide ABC transporter membrane subunit DppC
malG	-3.82	5×10^{-20}	-2.52	4×10^{-7}	IM	Maltose ABC transporter membrane subunit MalG
dppF	-3.75	4×10^{-34}	-3.93	4×10^{-18}	IM	Dipeptide ABC transporter ATP-binding subunit DppF
dppB	-3.74	1×10^{-31}	-3.74	1×10^{-16}	IM	Dipeptide ABC transporter membrane subunit DppB
malM	-3.69	7×10^{-5}	-2.06	2×10^{-6}	P	Maltose regulon periplasmic protein
melB	-3.60	6×10^{-28}	-1.24	0.03	IM	Melibiose:H$^+$/Na$^+$/Li$^+$symporter
ydcT	-3.59	2×10^{-27}	-3.61	1×10^{-10}	IM	Putative ABC transporter ATP-binding protein YdcT
ydcS	-3.57	1×10^{-31}	-4.05	9×10^{-15}	IM, P	Putative ABC transporter periplasmic–binding protein/polyhy-droxybutyrate synthase
gabP	-3.56	5×10^{-19}	-2.76	9×10^{-8}	IM	4-Aminobutyrate:H$^+$ symporter
ydcV	-3.48	1×10^{-23}	-3.39	1×10^{-8}	IM	Putative ABC transporter membrane subunit YdcV
ydhY	-3.44	5×10^{-16}	-0.40	0.23	IM	Putative 4Fe-4S ferredoxin-type protein
gudP	-3.42	5×10^{-21}	-3.50	4×10^{-7}	IM	Galactarate/glucarate/glycerate transporter GudP
astE	-3.34	2×10^{-15}	-3.47	1×10^{-5}	C	Succinylglutamate desuccinylase
malE	-3.32	1×10^{-12}	-2.52	1×10^{-9}	IM, P	Maltose ABC transporter periplasmic–binding protein

[a]C, cytosol; P, periplasm; IM, inner membrane; OM, outer membrane.

TABLE 3 Top 20 genes with increased RPF levels at pH 4.4 compared to pH 7.6, sorted in descending order by Ribo-Seq log$_2$ FC values[a]

Gene name	Log$_2$ FC (Ribo)	P-adjust (Ribo)	Log$_2$ FC (RNA)	P-adjust (RNA)	Cellular location	Annotation
asr	9.68	1×10^{-214}	9.37	2×10^{-52}	P	Acid shock protein
cadA	9.40	8×10^{-186}	9.39	5×10^{-92}	C	Lysine decarboxylase 1
ydgU	9.22	5×10^{-35}	12.39	3×10^{-16}	IM	Uncharacterized protein YdgU
yhcN	8.32	5×10^{-125}	9.08	3×10^{-94}	P	DUF1471 domain-containing stress-induced protein YhcN
bdm	7.85	4×10^{-71}	7.62	1×10^{-47}	C	Biofilm-dependent modulation protein
cadB	7.78	2×10^{-108}	8.90	2×10^{-51}	IM	Lysine:cadaverine antiporter
yjcB	7.73	4×10^{-45}	5.92	6×10^{-23}	IM	Uncharacterized protein YjcB
mdtJ	7.45	2×10^{-58}	4.62	1×10^{-23}	IM	Multidrug/spermidine efflux pump membrane subunit MdtJ
bhsA	7.12	3×10^{-110}	7.96	2×10^{-54}	OM	DUF1471 domain-containing multiple stress resistance outer membrane protein BhsA
glpD	7.06	2×10^{-122}	6.48	8×10^{-48}	C, IM	Aerobic glycerol 3-phosphate dehydrogenase
azuC	6.83	5×10^{-24}	8.17	1×10^{-26}	IM	Uncharacterized protein AzuC
nrdH	6.60	2×10^{-30}	6.46	1×10^{-29}	C	Glutaredoxin-like protein
yedR	6.35	2×10^{-21}	5.30	8×10^{-20}	IM	Putative inner membrane protein
yhdV	6.18	4×10^{-35}	6.43	3×10^{-18}	IM	Lipoprotein YhdV
nrdI	6.12	3×10^{-42}	6.03	3×10^{-28}	C	Dimanganese-tyrosyl radical cofactor maintenance flavodoxin NrdI
ybiJ	5.94	6×10^{-54}	4.56	2×10^{-21}	P	DUF1471 domain-containing protein YbiJ
fhuF	5.92	9×10^{-38}	4.55	8×10^{-21}	C	Hydroxamate siderophore iron reductase
ycgZ	5.73	3×10^{-50}	5.04	6×10^{-19}	C	Putative two-component system connector protein YcgZ
mdtI	5.72	7×10^{-31}	4.25	7×10^{-18}	IM	Multidrug/spermidine efflux pump membrane subunit MdtI
ycfJ	5.15	7×10^{-48}	5.03	1×10^{-17}	IM	PF05433 family protein YcfJ

[a]C, cytosol; P, periplasm; IM, inner membrane; OM, outer membrane.

act as an aggregase in order to prevent aggregation of proteins with positive charges (67). Our Ribo-seq data clearly illustrate the enormous importance of Asr under severe acid stress in *E. coli*, as it is one of the most abundant proteins in the cell, with approximately 2% of all reads mapping in the *asr* coding region at pH 4.4 (corresponding to an ~1,000-fold upregulation compared to pH 7.6). Strikingly, almost half of the top 20 genes with increased ribosome coverage of transcripts (*ydgU, yhcN, yjcB, yedR, yhdV, ybiJ, ycgZ,* and *ycfJ*) are poorly characterized (Table 3). So far, only YhcN from the above list has been shown to be involved in the response to acid stress (68).

GSEA for biological processes identified the GO terms "enterobactin biosynthetic process," "ferric-enterobactin import into cell," "siderophore-dependent iron import into cell," and "siderophore transmembrane transport" as significantly enriched at pH 4.4 (Fig. 3). Specifically, the complete enterobactin biosynthesis pathway, comprising the *entCEBAH* operon, *entF, entH*, and *ybdZ*, revealed significant enrichment under severe acidic conditions (Table S5). Furthermore, all subunits of the Ton complex (*tonB, exbB, exbD*) and its putative outer membrane receptor encoded by *yncD* exhibited significantly higher RPF and mRNA levels at pH 5.8 and pH 4.4 (Table S5). The Ton complex functions as a proton motive force-dependent molecular motor that facilitates the import of iron-bound siderophores (69, 70). Several other iron uptake systems, including a ferric dicitrate ABC transport system (*fecABCDE*), an iron (III) hydroxamate ABC transport system (*fhuACDB*), a ferric enterobactin ABC transport system (*fepA, fepB, fepCGD*), and a TonB-dependent iron-catecholate outer membrane transporter (*cirA*), were also induced under acidic conditions (Table S5). Moreover, the GO terms "protein maturation by iron-sulfur cluster assembly" and "iron-sulfur cluster assembly" were enriched at pH 4.4

(Fig. 3). Specifically, we detected a fivefold upregulation of all genes of the *isc* and *suf* operons (Table S5), which encode components of the complex machinery responsible for iron-sulfur cluster assembly in *E. coli* (71). In contrast, heme transport was among the most downregulated biological processes at both pH 5.8 and 4.4 (Fig. 3), which could potentially be the cause of iron limitation. Moreover, at low pH, the solubility of iron ions increases, which can destabilize iron-sulfur clusters (72). The iron limitation would be consistent with our data that *E. coli* upregulates the synthesis of iron-chelating sidero-phores and their transporters, as well as the components of the iron-sulfur assembly machinery. Given the better solubility of iron in a low pH environment, the question arises whether *E. coli* synthesizes siderophores to respond to iron limitation, or rather, protects itself against an iron excess. The latter function has been demonstrated for *Pseudomonas aeruginosa*, where siderophores protected cells from the harmful effects of reactive oxygen species. In this case, *P. aeruginosa* no longer secreted siderophores into the extracellular environment but instead stored them intracellularly (73). In conclusion, these results prompt the question of whether the upregulation of the iron uptake machinery counteracts iron limitation or rather provides protection against iron excess under severe acid stress.

We also detected a significant enrichment for the GO terms "cellular response to acidic pH," "stress response to copper ion," and "copper ion transmembrane transport" at pH 4.4 (Fig. 3). These results are in line with previous studies that have suggested an interplay between resistance to copper and acid stress in *Escherichia coli* (74, 75). This overlap between the two stress responses is further emphasized by our findings because at pH 4.4, substantial upregulation of the Cu^+-exporting ATPase CopA and CusA, a component of the copper efflux system, was detected (Table S5). These results are of important physiological relevance, given that copper is an important antibacterial component in the innate immune system (76, 77).

Among the downregulated genes at pH 4.4, the *tnaAB* operon and its leader peptide (*tnaC*) showed the most significant decrease in terms of RPF counts (Table 4). *tnaA* encodes a tryptophanase, which cleaves L-tryptophan into indole, pyruvate, and NH_4^+, whereas *tnaB* encodes a tryptophan:H^+ symporter (78). This finding is particularly intriguing because, in a previous study, persister cell formation in *E. coli* was related to a lower cytoplasmic pH associated with tryptophan metabolism (79). It is important to note that we also detected a substantial upregulation in RPFs for *hipA* (Table S5), which encodes a serine/threonine kinase that plays a role in persistence in *E. coli* (80). Therefore, our data provide further evidence for the link between internal pH and persistence.

The expression of several outer membrane proteins and porins (*ompW*, *ompF*, *nmpC*, *lamB*) was also downregulated at pH 4.4 (Table 4). This observation is consistent with the extensive restructuring of the *E. coli* lipid bilayers to reduce membrane permeability and limit proton entry. Similar to pH 5.8, the majority of the 20 proteins with the most reduced RPF levels compared with physiological pH are membrane proteins (Table 4). Moreover, the GO term "ATP synthesis coupled proton transport" was significantly reduced at pH 4.4 (Fig. 3). This is explained by the reduction in RPF levels of genes encoding subunits of the F_OF_1-ATPase (Table S5). F_OF_1-ATPase uses the electrochemical gradient of protons to synthesize adenosine 5′-triphosphate (ATP) from ADP and inorganic phosphate but can also hydrolyze ATP to pump protons out of the cytoplasm (81, 82). As at pH 5.8, the most downregulated biological processes at pH 4.4 were almost exclusively GO terms related to transport and cellular metabolism (Fig. 3).

In summary, the response of *E. coli* to severe acid stress is dominated by the activation of survival strategies that limit the entry of protons into the cell, prevent protein aggregation, and maintain iron homeostasis. Severe acid stress leads to a reduction in metabolic, transcriptional, and translational activity, thereby preparing *E. coli* for a dormant state. Eventually, these dormant cells may be able to withstand antibiotic attack (i.e., persister cells).

TABLE 4 Top 20 genes with decreased RPF levels at pH 4.4 compared to pH 7.6, sorted in ascending order by Ribo-Seq log$_2$ FC values[a]

Gene name	Log$_2$ FC (Ribo)	P-adjust (Ribo)	Log$_2$ FC (RNA)	P-adjust (RNA)	Cellular location	Annotation
tnaC	-6.82	4×10^{-39}	-6.18	5×10^{-27}	C	tnaAB operon leader peptide
tnaA	-6.53	6×10^{-7}	-6.15	1×10^{-34}	C	Tryptophanase
tnaB	-6.40	6×10^{-47}	-5.40	6×10^{-26}	IM	Tryptophan:H$^+$ symporter TnaB
mglB	-6.31	7×10^{-51}	-4.08	1×10^{-17}	IM, P	D-galactose/methyl-galactoside ABC transporter periplasmic-binding protein
ompW	-6.21	1×10^{-39}	-4.83	2×10^{-13}	OM	Outer membrane protein W
nmpC	-6.05	3×10^{-62}	-3.87	1×10^{-17}	OM	DLP12 prophage; putative outer membrane porin NmpC
treB	-5.94	3×10^{-62}	-4.57	3×10^{-26}	IM	Trehalose-specific PTS enzyme IIBC component
borD	-5.92	1×10^{-25}	-3.36	5×10^{-9}	IM	DLP12 prophage; prophage lipoprotein BorD
garP	-5.86	4×10^{-47}	-5.10	1×10^{-16}	IM	Galactarate/glucarate/glycerate transporter GarP
tdcA	-5.84	9×10^{-32}	-1.27	4×10^{-3}	C	DNA-binding transcriptional activator TdcA
lamB	-5.79	2×10^{-54}	-2.60	4×10^{-8}	OM	Maltose outer membrane channel/phage lambda receptor protein
malK	-5.65	2×10^{-70}	-2.14	7×10^{-5}	IM	Maltose ABC transporter ATP-binding subunit
ompF	-5.59	1×10^{-63}	-4.03	6×10^{-20}	OM	Outer membrane porin F
garD	-5.54	3×10^{-42}	-3.12	2×10^{-6}	C	Galactarate dehydratase
nrfB	-5.47	4×10^{-24}	-2.36	1×10^{-3}	P	Periplasmic nitrite reductase penta-heme c-type cytochrome
xylF	-5.46	3×10^{-22}	-4.54	3×10^{-16}	IM, P	Xylose ABC transporter periplasmic-binding protein
malM	-5.38	3×10^{-65}	-1.93	7×10^{-5}	P	Maltose regulon periplasmic protein
treC	-5.38	1×10^{-44}	-4.47	3×10^{-23}	C	Trehalose-6-phosphate hydrolase
malE	-5.31	8×10^{-35}	-2.53	2×10^{-8}	IM, P	Maltose ABC transporter periplasmic-binding protein
ydeN	-5.31	2×10^{-49}	-5.05	5×10^{-30}	P	Putative sulfatase

[a]C, cytosol; P, periplasm; IM, inner membrane; OM, outer membrane.

Expanding the regulatory network of enzyme-based H$^+$-consuming acid resistance systems

Recently, we have shown that the Adi and Cad AR systems are mutually exclusively activated in individual *E. coli* cells, indicating functional diversification and division of labor under acid stress (15). To gain further insights into the fine-tuned regulation of the three major AR systems, we first studied the mRNA and RPF levels of known enzyme-based H$^+$-consuming AR components. The core components of the Gad system (AR2) (*gadA*, *gadB*, and *gadC*) and several transcriptional components (*gadW*, *gadX*, *gadY*, *phoP*, *phoQ*) showed an increase in mRNA and RPF levels by approximately two- to sixfold at pH 4.4, but not at pH 5.8, whereas the expression of *ydeO* was massively induced at pH 5.8 (particularly at the mRNA level), and RPF levels were decreased at pH 4.4 (Fig. S7). Expression of the core components of the Adi system (AR3), *adiA* and *adiC*, was induced at severe acid stress but not at pH 5.8, consistent with our previous study (15). Upregulation was not detected for regulatory components of the Adi system. A novel finding was that the levels of *adiA* but not *adiC* were significantly higher in the Ribo-Seq data than in the RNA-Seq data (Fig. S7). In fact, *adiA* had the sixth highest increase in TE among all *E. coli* genes at pH 4.4 (Table S4), indicating translational regulation by a thus far unknown mechanism. The only other component of an AR system in *E. coli*, known to be subject to translational regulation, is the major regulator CadC of the Cad (AR4) system. CadC contains a polyproline motif, and its translation therefore depends on the elongation factor P, a process that keeps the copy number of CadC extremely low (83). As expected, expression of the core components of the Cad system (AR4), *cadA* and *cadB*, was tremendously increased at both pH 5.8 and 4.4. Genes of the Orn system (AR5) were not induced in our experimental setup (Fig. S7).

Next, we analyzed the mRNA and RPF levels of all annotated TFs to search for other potential TFs involved in the acid stress response of *E. coli* (Fig. 4A and B). At pH 5.8, YdeO showed by far the strongest induction at the transcriptional and translational levels, but for all other TFs, the expression levels hardly changed (Fig. 4A). At pH 4.4, the expression of numerous TFs was induced, including GadW, YdcI, and the antibiotic resistance-controlling regulator MarR. The strongest upregulation was found for the IclR-type regulator MhpR and the iron-sulfur cluster-containing regulator IscR (Fig. 4B). Notably, while most acid-induced TFs were differentially expressed and displayed constant TE, YdcI exhibited constant mRNA levels but was differentially translated in response to acid stress (Fig. 4B). The contribution of all TFs with high FC values to survival under acid stress (Table S6) was tested in an acid shock assay. Cells of the corresponding knockout mutants (84) and, for comparison, the *rcsB* and *gadE* mutants (each lacking a TF important for acid resistance) were exposed to pH 3 for 1 h. All mutants except *marR* and *ydeO* showed significantly reduced survival compared to the parental strain (Fig. 4C). For *ydeO*, this result was consistent with our finding that transcript abundance and occupancy with ribosomes were upregulated at mild but not severe acid stress (Fig. 4; Fig. S7). Thus, YdeO appears to be only crucial under mild acid stress (Fig. 4A). In contrast, the *mhpR* mutant had a low survival rate comparable to that of *rcsB* and *gadE*, and the survival rates of the *iscR*, *ydcI*, and *gadW* mutants were only slightly higher (Fig. 4C). These results confirm the physiological relevance of these TFs for acid resistance. As controls, we re-introduced the corresponding genes *in trans* using isopropyl-β-D-thiogalactopyranosid (IPTG)-inducible pCA24N plasmids from the ASKA collection (85). Complementation of the *mhpR*, *iscR*, *ydcI*, and *gadW* mutants, as well as *rcsB* and *gadE* controls, resulted in strains with survival rates comparable to the wild-type (WT) strain carrying the pCA24N control vector (Fig. S8).

Subsequently, we tested whether these TFs are involved in the regulation and interconnectivity of the Gad, Adi, and Cad systems. Therefore, we examined the promoter activities of *gadBC*, *adiA*, and *cadBA* in the corresponding knockout mutants (84) using transcriptional reporter plasmids (promoter-*lux* fusions). The cultivation conditions were the same as those used for Ribo-Seq and RNA-Seq (Fig. 1A), and luciferase activity was monitored during growth in microtiter plates. We found that YdcI significantly
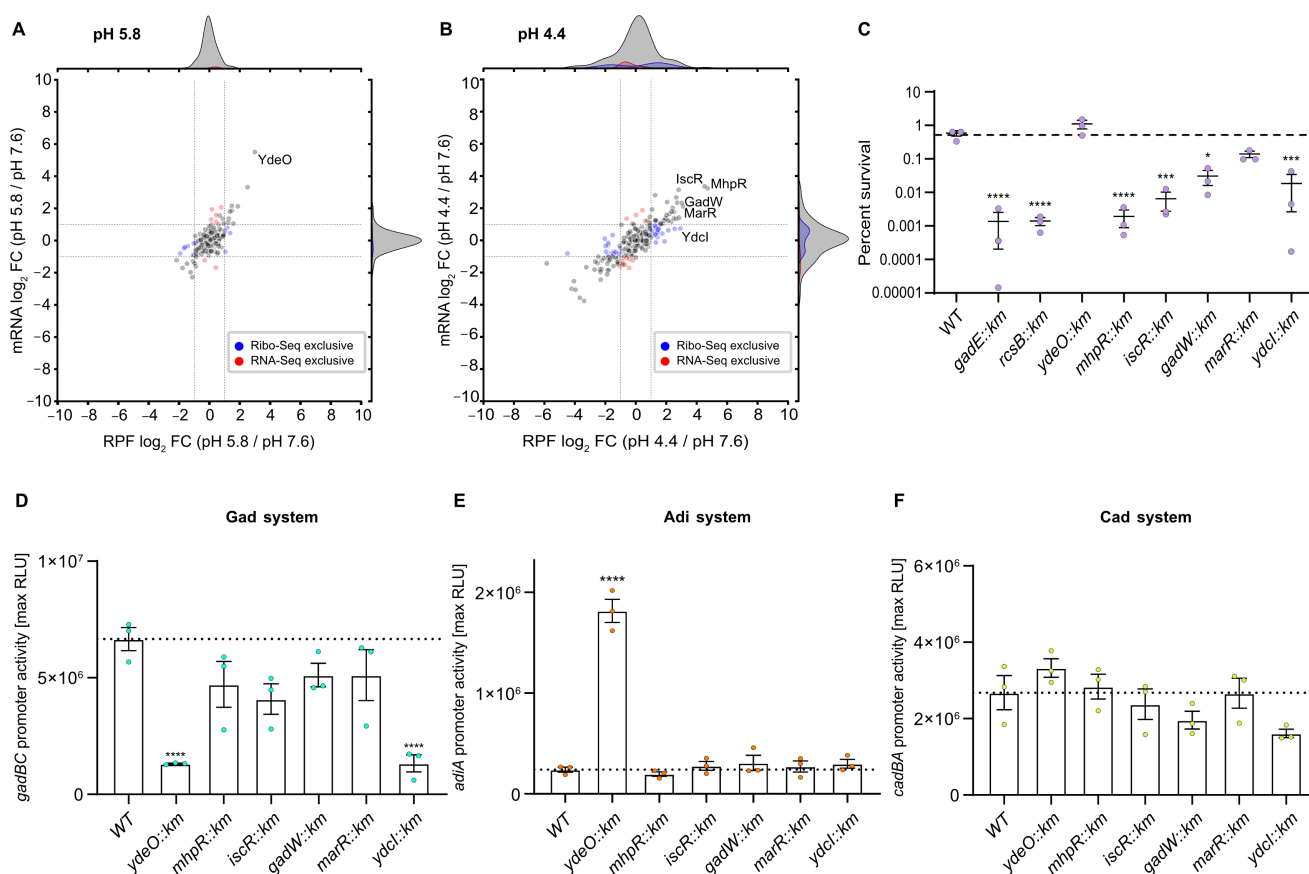
FIG 4 Contribution of TFs to survival and AR induction under acid stress. Comparison of global RPF and mRNA $\log_2$ FC values of transcriptional regulators for (A) pH 5.8 vs pH 7.6 and (B) pH 4.4 vs pH 7.6. Dashed lines indicate $\log_2$ fold changes of +1 or −1. Changes detected exclusively by either RNA-Seq (red dots) or Ribo-Seq (blue dots) are colored. TFs described in panels C-F are highlighted. (C) Acid shock assay to test the survival of *E. coli* BW25113 (WT) and the indicated mutants (84). Cells were grown in LB pH 7.6 to OD_{600} = 0.5. The cultures were split and then either grown at pH 7.6 or stepwise stressed (15 min pH 5.8, 15 min pH 4.4) before being exposed to LB pH 3 for 1 h. Colony-forming units were counted, and the ratio of surviving cells was calculated. The dashed line indicates the average percentage of surviving WT cells. (D through F) Luciferase-based promoter assays. WT and the indicated mutants were transformed either with plasmid pBBR1-P*gadBC:lux* (D), pBBR1-P*adiA:lux* (E), or pBBR1-P*cadBA:lux* (F) and grown in LB medium (pH 7.6) until OD_{600} = 0.5. The medium pH was then adjusted to 5.8 to induce the Cad system or to pH 4.4 to induce the Adi and Gad systems. Luminescence and growth were determined every 10 min in microtiter plates using a CLARIOstar plus plate reader (BMG Labtech). Data are reported as relative light units (RLUs) in counts per second per OD_{600}, with maximal RLU shown. Dashed lines indicate the average maximal RLU values of the WT. C-F, All experiments were performed in biological replicates (*n* = 3), and error bars represent standard deviations of the mean. Analysis of variance, followed by Bonferroni's multiple comparisons test, was used to compare log-transformed max. RLU values between mutant strains and the wild type (BW25113) (*$P \leq 0.05$, ***$P \leq 0.001$, ****$P \leq 0.0001$).

affected the promotor activity of *gadBC* (Fig. 4D). Although the LysR-type regulator YdcI has been shown to affect pH stress regulation in *Salmonella enterica* serovar Typhimurium and *E. coli*, its precise role is still unclear (86–88). Based on the data presented here, we hypothesize that the decreased survival of the *ydcI* mutant under severe acid stress is due to decreased expression of the Gad system. The absence of YdeO resulted in an eightfold stimulation of the *adiA* promoter activity (Fig. 4E). Thus, YdeO not only activates the Gad system (89) but also appears to be a repressor for the Adi system. This implies that the Adi system is regulated not only by the XylS/AraC-type regulator AdiY but also by YdeO. Thus, YdeO is the first example of a transcriptional activator shown to be involved in the regulation of more than one AR system in *E. coli* and might play a role in the heterogeneous activation of the Adi and Gad systems within a population. Although we observed a slight decrease in *cadBA* promoter activity in the *ydcI* mutant, the decrease was not statistically significant. Therefore, none of the tested TFs affected the Cad system (Fig. 4F).

In conclusion, based on the differential expression data and lower survival of mutants during acid shock, we identified two novel TFs, namely, MhpR and IscR, which are crucial under severe acid stress (Fig. 4C), but are not associated with the Gad, Adi, and Cad systems (Fig. 4D through F). This implies that these regulators ensure the survival of *E. coli* in acidic habitats by inducing other defense mechanisms. Of particular interest is MhpR, which had the highest increase in RPFs of all TFs at pH 4.4 (Fig. 4B), and the corresponding mutant had the lowest survival at pH 3 (Fig. 4C). Further studies are needed to determine whether MhpR, which is a specific regulator of the *mhpABCDFE* operon-encoding enzymes for the degradation of phenylpropionate (90, 91), directly or indirectly contributes to acid resistance.

Differential expression of known and novel sORFs under mild and severe acid stress

In recent years, the annotation of many bacterial genomes has been extended by previously unknown small proteins (29), many of which are located in the membrane (92). This progress has been achieved primarily through the development of optimized detection strategies using adapted ribosome profiling and mass spectrometry protocols (28, 33, 93). Recently, additional sORFs were identified in *E. coli* using antibiotic-assisted Ribo-Seq, which captures initiating ribosomes at start codons (94, 95). Advanced detection strategies also revealed novel small proteins in other species, such as the archaeon *Haloferax volcanii*, the nitrogen-fixing plant symbiont *Sinorhizobium meliloti*, *Salmonella* Typhimurium, and *Staphylococcus aureus* (33–35, 96).

Among the previously known sORFs in *E. coli* K-12 and those discovered by Storz and colleagues (94), pH-dependent differential RPF levels were observed in our data sets for 12 and 29 small proteins at pH 5.8 and pH 4.4, respectively (Table S8). These findings validate the expression of these sORFs and highlight their physiological relevance in the acid stress response of *E. coli*. For example, induction of *mdtU*, an upstream ORF of *mdtJI*, was observed under mild acid stress (Fig. 5A) and corresponds to the observed upregulation of the multidrug/spermidine exporter MdtJI (Table 1). A previous study has shown that translation of MdtU is crucial for spermidine-mediated expression of the MdtJ subunit under spermidine supplementation at pH 9 (97). A similar mechanism could operate under acid stress conditions. The strongest induction of sORFs under severe acid stress was detected for *ydgU* (located in the same transcriptional unit as the acid shock protein-encoding gene *asr*) and *azuC* (Fig. 5B). AzuCR acts as a dual-function RNA and encodes a 28-amino acid protein, but it can also base pair as an sRNA (AzuR) with two target mRNAs, including *cadA* (98). AzuCR modulates carbon metabolism through interactions with the aerobic glycerol-3-phosphate dehydrogenase GlpD (98).

In addition to known sORFs, we aimed to uncover further hidden small proteins on the basis that our Ribo-Seq data were acquired under stress conditions to which *E. coli* is exposed in its natural habitat, the gastrointestinal tract. In particular, we searched for novel sORFs that remained undetected in previous Ribo-Seq approaches when *E. coli* was grown at a neutral pH.

Initial predictions for novel sORF candidates were acquired using the neural network-based prediction tool *DeepRibo* (99). All potential candidates were filtered based on coverage (rpkm >30 across all Ribo-Seq samples) and codon count [10–70 amino acids (aa)], with the exception of sORF15 (93 amino acids) (Table S7), which was manually discovered by inspecting the 3′ UTR of *gadW*. To further refine our search, we focused on sORF candidates that were significantly induced at either pH 5.8 or pH 4.4 (RPF $\log_2$ FC >2 and *P*-adjust <0.05) compared to pH 7.6. Predictions that overlapped with annotated genes on the same strand were excluded because Ribo-Seq signals were indistinguishable. This workflow yielded 152 candidates that were visually inspected using the web-based genome browser JBrowse2 (100). Candidates with continuous coverage across the predicted sORF, matching the ORF boundaries, and promising Shine-Dalgarno sequences were considered high-confidence candidates. In total, we identified 18 acid-induced sORF candidates (Table S7) that had not been previously detected. Of note, most of the
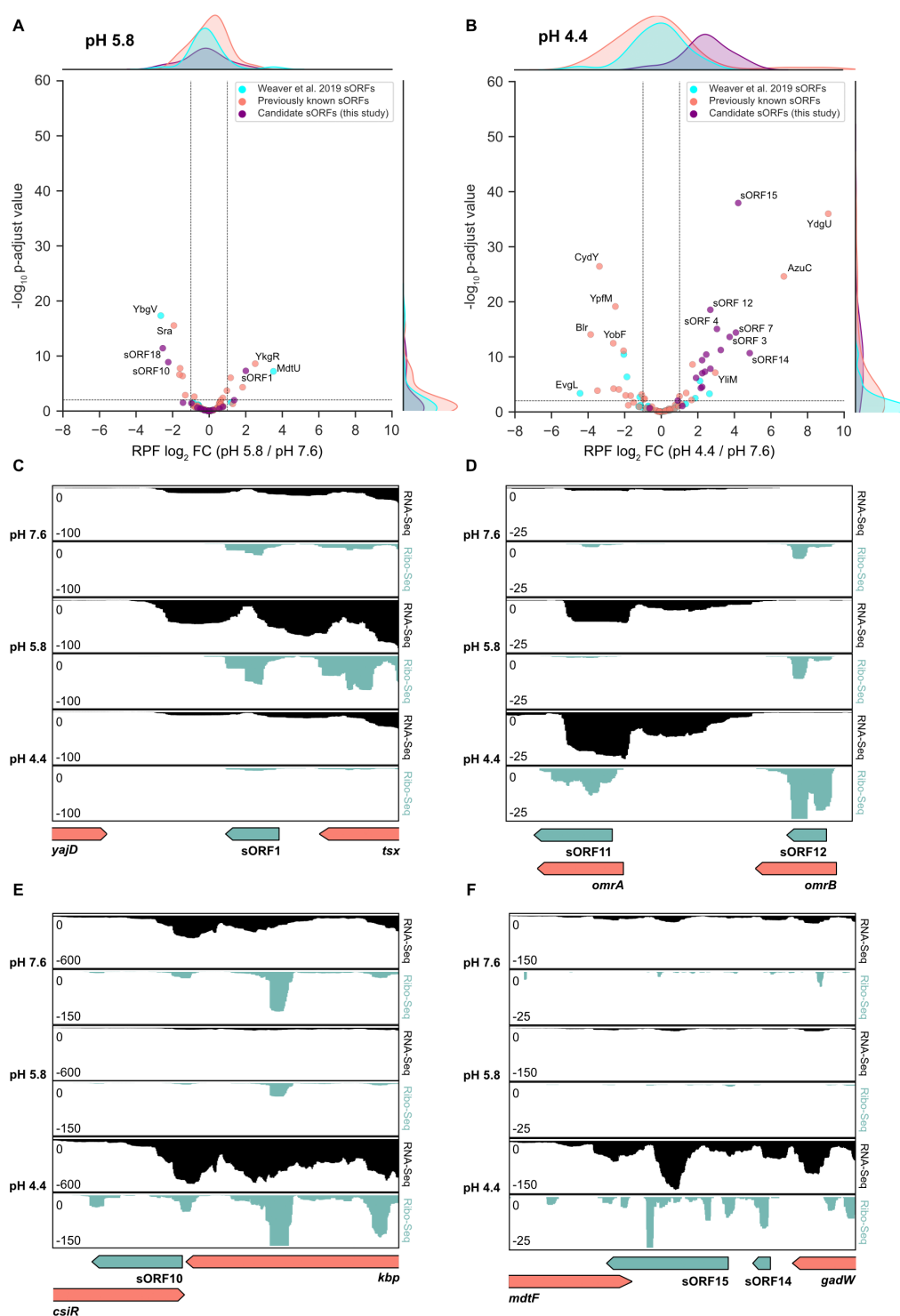
FIG 5 Differentially expressed sORF candidates upon exposure of *E. coli* to acid stress. Volcano plots illustrating differential RPF levels at (A) pH 5.8 vs 7.6 and (B) pH 4.4 vs 7.6 for sORF candidates identified in this study (Table S7), sORFs identified by Weaver *et al.* (94), and previously known sORFs. (C through F) JBrowse2 screenshots of read coverage from Ribo-Seq (green tracks) and RNA-Seq (black tracks) libraries at pH 7.6, pH 5.8, and pH 4.4. The schematic illustrations below indicate the genomic locations of enriched sORF candidates under acidic stress and adjacent annotated genes. Novel sORF candidates sORF1 (13 aa), sORF10 (38 aa), sORF11 (28 aa), sORF12 (11 aa), sORF14 (13 aa), and sORF15 (93 aa) are shown in green. aa, amino acids

candidates are encoded as part of operons or are located in the 3′ UTR of annotated genes. In addition, we detected one independent antisense sORF (sORF2 encoded antisense to *tesA*) and two upstream ORFs (leader peptides): sORF18, located upstream of the translation start site of the periplasmic chaperone encoding *osmY*, and sORF8, located close to the glucokinase-encoding gene *glk* (Table S7).

Of these 18 acid-induced candidate small proteins (Table S7), 17 had higher RPF counts at pH 4.4 than at pH 5.8. This suggests that the contribution of sORFs to acid defense in *E. coli* is more relevant under severe acid stress. Only sORF1 showed a higher expression level in cells exposed to mild acid stress (Fig. 5C). sORF1 is located in the 3′ UTR of *tsx*, which encodes a nucleoside-specific channel-forming protein. This finding is consistent with the observed increased requirement for nucleotides by *E. coli* at pH 5.8 (Fig. 3).

For the first time, we identified two sORF candidates located within genes encoding the redundant small regulatory RNAs OmrA and OmrB (Fig. 5D). *omrA* and *omrB* are highly identical at the 3′ and 5′ ends, differ mainly in their central parts, and regulate the expression of numerous outer membrane proteins (101). Our analysis suggests that both OmrA and OmrB act as dual-function RNAs under severe acid stress and encode small proteins: a 28-amino acid protein OmrA (sORF11) and an 11-amino acid protein OmrB (sORF12) (Fig. 5D). Due to the sequence variation in the central parts, the translation of OmrB ends at an earlier stop codon. Notably, both *omrA* and *omrB* displayed higher RPF levels at pH 4.4, whereas transcription of *omrA* but not *omrB* was induced at pH 4.4 (Fig. 5D). Thus, despite the high sequence similarity, *omrA* and *omrB* do not encode identical small proteins under severe acid stress and are differentially regulated at the transcriptional and translational levels. We also detected an acid-induced sORF candidate (sORF3) in *rybB* (Table S7), another sRNA involved in the regulation of outer membrane proteins (102). To our best knowledge, the presence of OmrA, OmrB, and RybB peptides has not yet been reported.

Three new candidate sORFs potentially involved in the regulation of AR systems were detected. sORF10 is located in the 3′ UTR of a potassium-binding protein encoded by *kbp* and encoded antisense to the transcriptional regulator CsiR (Fig. 5E). The latter might be involved in the regulation of the Adi system (15, 103). Given the significant upregulation of sORF10 at pH 4.4 and its complete complementarity to the 3′ end of the *csiR* mRNA, we hypothesize that sORF10 plays a role in fine-tuning the expression of the Adi system. Strikingly, we also discovered two high-confidence candidates for sORFs located in the relatively long 3′ UTR of GadW, one of the major transcriptional regulators of the Gad system (Fig. 5F). sORF14 and sORF15 exhibit constant coverage across the predicted ORF and contain Shine-Dalgarno sequences (Table S7). These results suggest that the complex Gad system may consist of even more components.

To gain further insight into the subcellular location and features of the newly identified sORF candidates, we used PSORTb (104) and DeepTMHMM (105) for transmembrane topology prediction. Notably, sORF15 is predicted to be located in the inner membrane and has two transmembrane helices, which were predicted with a probability of >90% (Fig. S9A). Additionally, the sORF15 protein structure prediction using Alpha-Fold2 (106) in Google Colab (ColabFold) (107) revealed a potential third helix toward the C-terminal end (Fig. S9B). Using blastp and tblastn (108), we found homologs of sORF15 with >80% identity in *Vibrio*, *Shigella*, *Klebsiella*, *Salmonella*, *Enterococcus*, and *Escherichia* (Fig. S9C) and identified homologs with at least 60% identity for approximately half of the other candidate sORFs (sORF2, 4, 5, 6, 7, 9, 10, 11, 16, and 18). These results strengthen confidence in the correct prediction of these sORFs. However, homologs in other species often only displayed partial matches and were almost exclusively annotated as "hypothetical proteins," as illustrated for sORF15 (Figure S9C). Moreover, we evaluated whether sORF15 is translated in the absence of the upstream gene *gadW*. A pBAD24-sORF15:3xFLAG plasmid, which harbors the native Shine-Dalgarno sequence of sORF15 (Table S11), and a FLAG-tagged version of sORF15 were constructed. sORF15

translation was successfully verified by Western blotting (Fig. S9D), which exemplifies that sORFs detected in this study yield detectable protein products.

In conclusion, we identified 18 high-confidence candidates for novel sORFs that are significantly induced upon exposure of *E. coli* to mild or severe acid stress.

Differentiation of the acid stress and general stress responses using autoencoder-based machine learning

In general, stress response mechanisms can be broadly classified into two categories: global stress responses and adaptations to specific types of stress. Global stress responses can be triggered by various stimuli and provide protection against multiple other unrelated stress factors (109). The global response often involves the activation of alternative sigma factors that affect hundreds of genes. In contrast, adaptations to specific types of stress are tailored to the specific stressor and involve a regulator that senses an environmental cue and modulates the expression of a set of genes, which counteract the stress (109, 110).

Given the large number of differentially regulated genes and pathways in response to acid stress (Fig. 2 and 3), we asked which of these adaptive mechanisms are acid-specific and which are also triggered by other stressors. In order to distinguish acid-specific and general stress responses, we used denoising autoencoders (DAEs), deep learning models designed for meaningful dimensionality reduction (111, 112). DAEs accomplish this by passing data through an encoder that compresses it into activations of a bottleneck layer (Fig. 6A1), with each node in the bottleneck layer interpretable as a coordinated expression program (113). For our analysis, we employed an ensemble of deep DAEs (see Materials and Methods) (113), trained on the *E. coli* K-12 PRECISE 2.0 compendium (114), augmented with additional stress conditions (115), as well as the acid stress conditions of the current study (Fig. 6A1). Using this method and data set, we have conducted a comparative analysis of the transcriptional response of *E. coli* to pH 4.4 and pH 5.8, contrasted against an extensive range of other stress conditions, including heat stress (116), ethanol stress (117), osmotic stress (118), oxidative stress (119, 120), low oxygen (LOX) (115), and exposure to sublethal concentrations of chloramphenicol (CAM) (115) and trimethoprim (TMP) (115).

To identify biological processes associated with a particular stress condition, we passed the associated RNA-seq data set into the encoder of each network and identified bottleneck nodes that were uniquely turned on by that data set (Fig. 6A2). We then manually turned on these nodes to generate gene sets that are associated with that condition, which can be further analyzed through GO term enrichment (Fig. 6A3 and 4). Using this procedure, we identified groups of nodes that uniquely turn on for acid stress conditions and turn off for all other stress conditions, as well as groups that are simultaneously on for both acid and one additional stress condition. We observed that there are many nodes that turn on simultaneously upon both acid and ethanol exposure (Fig. 6B). The overlap between acid stress and ethanol stress responses has been noted previously and can be explained by the fact that ethanol fluidizes the cytosolic membrane and increases the permeability for protons (121). Furthermore, there are indications of an overlap between acid and antibiotic stress (122, 123), reflected in the high number of acid + CAM activating nodes (Fig. 6B).

To pinpoint which cellular adaptations cause acid-specificity for the 48 and 91 specific bottleneck nodes at pH 5.8 and 4.4 (Fig. 6B), respectively, we conducted GSEA for biological processes on each of the gene sets associated with acid-specific node groups (see Materials and Methods). The GO terms that were significantly enriched in the highest number of both pH 5.8- and pH 4.4-specific upregulating node gene sets were "siderophore transmembrane transport," "response to cold," "bacterial-type flagellum assembly," and "chemotaxis" (Fig. 6C), reflecting our previous differential RNA-seq analysis (Fig. 3). The appearance of the GO term "response to cold" might be a result of the lack of cold stress in our compendium of stressors. Additionally, it should be noted that genes associated with this GO term include cold shock proteins, which may have
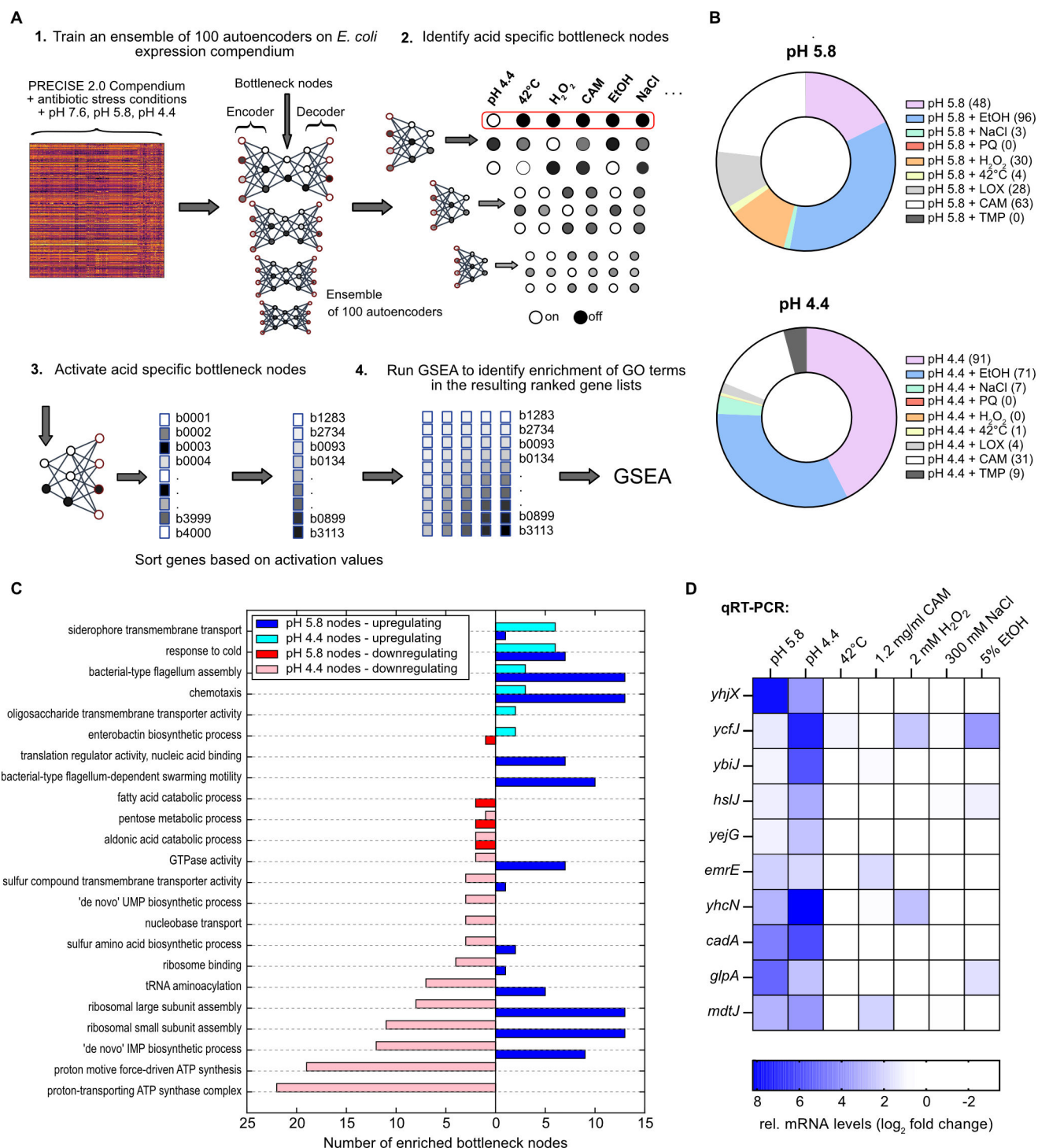
FIG 6 Differentiation between acid stress responses and general stress using autoencoders. (A) Schematic overview of the autoencoder ensemble training and subsequent bottleneck group identification pipeline. (B) Donut charts indicating the proportion of bottleneck nodes that turn on only under the specified conditions. Absolute numbers of specific bottleneck nodes are listed in brackets after each condition. (C) Significantly enriched GO terms associated with pH 4.4- and pH 5.8-specific nodes. Left-facing red bars indicate nodes that downregulate the corresponding GO term, while right-facing blue bars indicate upregulation. (D) Verification of acid stress-specific genes predicted by autoencoders. *E. coli* cells were either grown as indicated in Fig. 1A or exposed to heat (42°C), oxidative (H_2O_2), osmotic (NaCl), chloramphenicol (CAM), or ethanol (EtOH) stress. Total RNA was isolated, and relative mRNA levels were measured by RT-qPCR. Fold change values were determined relative to non-stress conditions and normalized using either *secA* or *recA* as reference genes. Standard deviations were calculated from three replicates ($n = 3$) and accounted for <10% of fold change values in all cases. PQ, paraquat; LOX, low oxygen; TMP, trimethoprim.

broader roles in the survival of stress conditions (124, 125), as well as several prophage genes and ribosome biogenesis factors. We found that mild acid stress turns on nodes, which correspond to gene sets associated with nucleotide and ribosome biosynthesis, including the GO terms "ribosome large subunit assembly," "ribosome small subunit assembly," and *"de novo* IMP biosynthetic process," while severe acid stress turns them off (Fig. 6C). These findings are consistent with our previous observations, namely, that *E. coli* induces nucleotide and ribosome biosynthesis to cope with mild acid stress but enters a metabolically inactive state under severe acid stress. The GO terms significantly affected in the highest number of acid-specific pH 4.4 downregulated bottleneck nodes were "proton motive force-driven ATP synthesis" and "proton-transporting ATP synthase complex" (Fig. 6C). These two GO terms exclusively involve genes encoding subunits of the F_OF_1 ATP synthase and can be considered paradigms for acid-specific adaptations since the F_OF_1 ATP synthase can also pump protons (126).

Considering that we detected a high number of genes induced by severe acid stress with unknown functions (Table 3) and lacking GO associations, we expanded our search for acid-specific adaptations from GO terms to single genes. In order to select acid-specific candidate genes, we investigated all genes associated with acid-specific bottleneck nodes and calculated the $\log_2$ FC between each acid stress and every other above-mentioned stress condition for these genes. Genes with the highest expression values under acidic conditions and $\log_2$ FCs of at least 0.5 for at least 95% of comparisons were then selected. This procedure yielded 10 candidate genes (Fig. 6D). To experimentally validate that these genes are indeed specifically upregulated under acid stress, we exposed *E. coli* to a variety of common stressors and performed qRT-PCR. *E. coli* cells were either grown under acid stress (Fig. 1A) or exposed to heat (42°C), oxidative (H_2O_2), osmotic (NaCl), antibiotic (chloramphenicol), or ethanol (EtOH) stress. For all investigated genes, the strongest upregulation was observed at either pH 4.4 or pH 5.8 relative to non-stress conditions (Fig. 6D), except for *ycfJ*, which was activated at pH 4.4 and under oxidative and ethanol stress. The remaining investigated genes were only upregulated under one other stress condition at most (Fig. 6D). Given that *emrE* and *mdtJ* encode multidrug exporters, the induction upon supplementation with sublethal concentrations of chloramphenicol is not surprising and further underscores the interplay between acid and antibiotic stress. The observed upregulation of *yhcN* under oxidative stress (Fig. 6D) was also reported previously (127). Nevertheless, we uncovered four *bona fide* examples of genes (*ybiJ*, *hslJ*, *yejG*, and *yhjX*) that displayed exclusive pH-dependent expression (Fig. 6D). Induction of *yhjX*, encoding a putative pyruvate transporter, might be related to the deamination of serine, which yields ammonia and pyruvate in uropathogenic *E. coli* (128). The precise molecular functions of YbiJ, YejG, and HslJ in the context of acid stress are currently unclear. These results highlight that our autoencoder pipeline is complementary to differential gene expression analysis, yielding biologically consistent results while also identifying expression patterns that uniquely discriminate acid stress from other stress responses.

Conclusions

Here, we present the first comprehensive study on the global transcriptome- and translatome-wide response of *E. coli* exposed to varying degrees of acid stress. Our investigation goes beyond previous research, which focused on comparing *E. coli* transcriptomes across different pH levels during growth (18, 21). Instead, we report on rapid changes occurring upon sudden pH shifts, which are relevant for bacteria such as *E. coli*, during passage of the gastrointestinal tract (129).

Using both Ribo- and RNA-Seq, we uncovered not only well-known acid defense mechanisms but also numerous previously undiscovered relevant genes and pathways to combat mild and severe acid stress (Fig. 7). The latter include siderophore production, glycerol-3-phosphate conversion, copper export, *de novo* nucleotide biosynthesis, and spermidine/multidrug export (Fig. 3 and 7). A striking number of membrane proteins and H^+-coupled transporters were found to be downregulated under both mild and
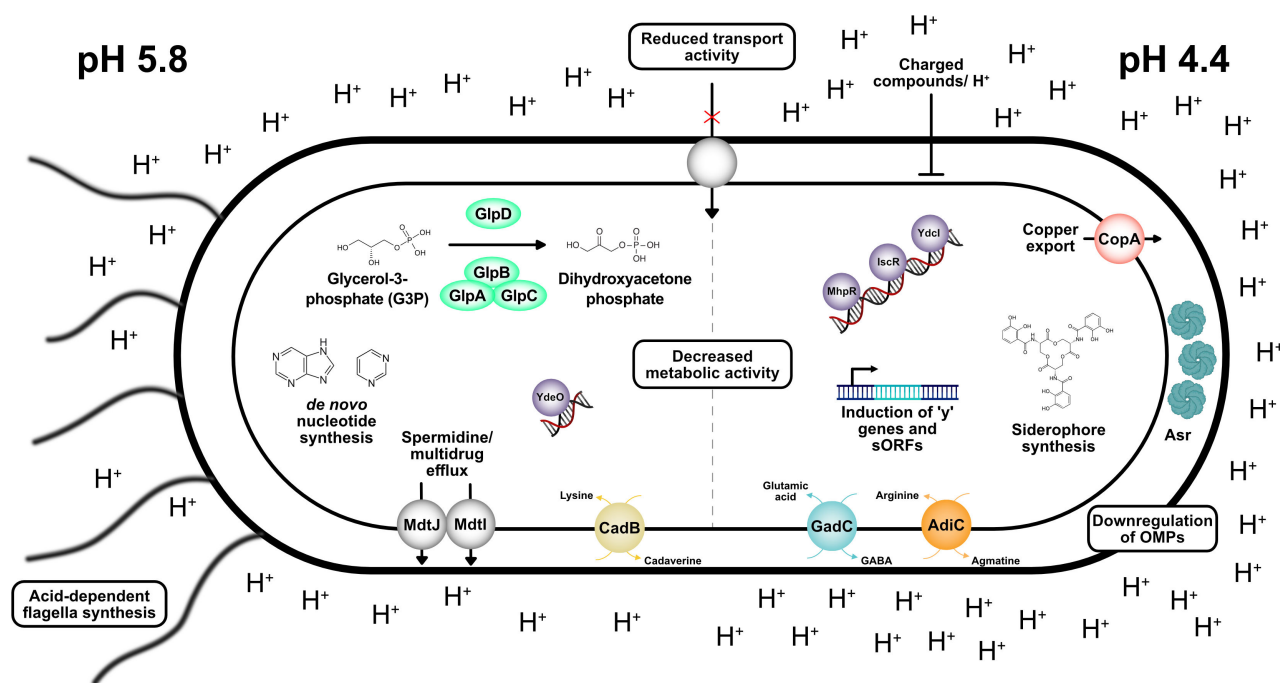
FIG 7 Overview of the fine-tuned response of *E. coli* to mild (pH 5.8) and severe (pH 4.4) acid stress. Acid stress counteracting mechanisms, including those revealed by Ribo-Seq and RNA-seq, are indicated.

severe acid stress (Fig. 7; Tables 2 and 4), underscoring the importance of the cytosolic membrane and its composition as a barrier for protons. Moreover, under severe stress, many outer membrane proteins were downregulated (Fig. 7; Table 4). Notably, a large proportion of genes with yet unknown functions were strongly induced, particularly under severe acid stress (Table 3). Our approach implies that exposing *E. coli* to culture conditions mimicking near-lethal habitats can offer valuable insights into the molecular functions of genes with low expression levels under standard growth conditions.

Our analysis revealed two new TFs, MhpR and IscR, involved in acid stress adaptation. Furthermore, we gained new insights into the role of the TFs YdeO and MarR. YdeO controls not only the transcription of genes in the Gad system but also *adiA* in the Adi system (Fig. 4), suggesting that YdeO connects the regulation of two AR systems in *E. coli*. The observed upregulation of MarR under acid stress, but the low contribution of this TF to acid resistance (Fig. 4B), may provide a link to antibiotic resistance and solvent stress tolerance in *E. coli* (130).

In addition to the pH-dependent differential expression levels of previously identified small proteins, such as YdgU, MdtU, and AzuC, we identified 18 high-confidence, not yet annotated, sORF candidates (Fig. 5). Of particular interest are sORF14 (13 amino acids) and sORF15 (93 amino acids), which are located in a transcriptional unit with *gadW* and *gadX*, suggesting their association with the Gad AR system and a potential involvement in glutamate transport and/or glutamate decarboxylation to gamma-aminobutyrate (GABA). Considering the predicted membrane location of sORF15 and its adjacent gene *mdtF*, an association with either the glutamate/GABA antiporter GadC and/or the multidrug efflux pump MdtF is conceivable.

The autoencoder-based comparison with other common stressors allowed us to distinguish acid stress-specific adaptations from general stress response programs (Fig. 6). Therefore, it was possible to differentiate between direct and indirect effects triggered by protonation and/or cellular damage. Considering the growing volume of next-generation sequence data, denoising autoencoders will be an increasingly important tool for interpreting future studies in the full context of accumulating RNA-seq data sets. Colonizing the intestinal tract is a complex process that includes not only rapid pH

changes but also alterations in oxygen and nutrient availability as well as competition with other bacteria. The ability of pathogenic *E. coli* strains to respond to such rapidly changing environments ensures their fitness advantage. We have shown here, for acid stress, the complexity of the regulatory network for ensuring survival and adaptation. The use of autoencoders, successfully tested here, could allow for the identification of physiological weak points associated with the survival of specific stresses. Targeting such weak points could lead to new classes of antibiotics or antivirulence treatments that take advantage of the unique expression patterns induced by natural stress conditions encountered in the host environment.

MATERIALS AND METHODS

Strains and growth conditions

Bacterial strains and plasmids used in this study are listed in Table S9. *E. coli* strains were cultivated in LB medium (10 g/L tryptone, 5 g/L yeast extract, 10 g/L NaCl) and incubated aerobically in a rotary shaker at 37°C. When appropriate, media were supplemented with 15 µg/mL gentamycin, 100 µg/mL carbenicillin, or 50 µg/mL chloramphenicol.

For ribosome profiling and RNA-Seq experiments, the pH of the medium was adjusted at the indicated time points by the direct addition of 5 M HCl to the growing cultures (Fig. 1A).

For comparison with other stress conditions (heat stress, osmotic stress, oxidative stress, antibiotic stress, and ethanol stress), *E. coli* was initially grown in LB medium to $OD_{600} = 0.5$. Stress conditions were initiated either by moving flasks to a pre-heated 42°C incubator or by the addition of either H_2O_2 (2 mM), NaCl (300 mM), chloramphenicol [1.2 µg/mL (wt/vol)], or ethanol [5% (vol/vol)]. In all cases, samples were collected after 15 min of stress treatment.

Plasmid construction

Molecular methods were performed according to standard protocols or according to the manufacturer's instructions. Kits for the isolation of plasmids and the purification of PCR products were purchased from Süd-Laborbedarf. Enzymes and HiFi DNA Assembly Master Mix were purchased from New England Biolabs. To construct the reporter plasmid pBBR1-MCS5-P*gadBC:lux* (Table S11), 335 nt of the upstream region of *gadBC* were amplified by PCR using primers KSO-0131/KSO-0132 (Table S10) and MG1655 genomic DNA as a template. For the construction of the pBAD24-sORF15:3xFLAG plasmid (Table S11), the sORF15 coding region (excluding the stop codon) and an additional 15 nt upstream of the start codon (harboring the native Shine-Dalgarno sequence) were amplified using primers KSO-183/KSO-184 (Table S10) from MG1655 genomic DNA. After purification, promoter fragments were assembled into PCR-linearized pBBR1-MCS5 or pBAD24-3xFLAG vectors via Gibson assembly (131). The pCA24N-control plasmid was obtained by excising the *ydcI* insert from a pCA24N-*ydcI* vector using restriction enzymes XhoI and SalI (NEB) and subsequent ligation using T4 DNA Ligase (NEB). Correct insertions were verified by colony PCR and sequencing.

Sample collection for Ribo-Seq and RNA-Seq

Three sets of biological triplicates of MG1655 cells were inoculated to a starting OD_{600} of 0.05 from overnight cultures and grown to exponential phase ($OD_{600} = 0.5$) in 200 mL of unbuffered LB medium (pH 7.6). Two sets of cultures were adjusted to pH 5.8 by direct addition of 5 M HCl, while one set was further grown for 30 min at pH 7.6 (Fig. 1A). After 15 min, one set of biological triplicates was further adjusted from pH 5.8 to pH 4.4 by the addition of 5 M HCl, while the other cultures remained at pH 5.8 or pH 7.6, respectively (Fig. 1A). Cells were grown for another 15 min, samples were collected, and Ribo-Seq and RNA-Seq were performed as described below. pH values before and after pH shifts, as

well as final optical densities, were monitored throughout the experiment (TablesS1 and S2).

Ribosome profiling

Whole-culture flash freezing, cell lysis using a freezer mill, pelleting ribosomes over sucrose cushions, and ribosomal footprint isolation using a size selection gel were performed following the published protocol from Mohammad and Buskirk (40). MNase treatment, monosome recovery, RNA isolation, end-labelling by T4 polynucleotide kinase, and cDNA library construction were conducted by adapting the protocol from Latif and colleagues (39).

Briefly, 100 mL of liquid cultures and 10× lysis buffer [200 mM Tris pH 8, 1.5 M $MgCl_2$, 1 M NH_4Cl, 50 mM $CaCl_2$, 4% (vol/vol) Triton X-100, 1% (vol/vol) NP-40] were flash frozen in liquid nitrogen. For each sample, 90 g of frozen cells was mixed with 10 g of frozen 10× lysis buffer and lysed by cryogenic grinding in a SPEX SamplePrep 6875 Freezer/Mill (10 cycles, 10 Hz, 5 min precool, 1 min run, 1 min cool). The pulverized samples were thawed, and the lysate was pre-cleared by centrifugation (9,800 × g, 10 min, 4°C) in a Beckman Coulter Optima XE-90 Ultracentrifuge using a 50.2 Ti Rotor. The supernatant was used to pellet ribosomes over sucrose cushions by centrifugation in a Beckman Coulter Optima XE-90 Ultracentrifuge using a 70.1 Ti Rotor (330,000 × g, 1.5 h, 4°C). After resuspension of pellets in resuspension buffer (20 mM Tris pH 8, 15 mM $MgCl_2$, 100 mM NH_4Cl, 5 mM $CaCl_2$), nuclease digestion was performed using MNase (NEB) (2 h, 25°C). Monosomes were recovered using MicroSpin S-400 HR columns (GE Healthcare), and RNA was isolated using the miRNeasy Mini Kit (QIAGEN) in combination with the RNase-Free DNase Set (QIAGEN). The isolated RNA was loaded on a 15% TBE urea size selection gel, and after staining with SYBR Gold (Invitrogen), ribosomal footprints between 15 and 45 nt were excised from the gel. For elution of RNA, gel pieces were crushed by poking a hole with an 18 G needle in a 0.5-mL tube, placing it in a 1.5-mL tube, and subsequent centrifugation. RNA was recovered by precipitation after adding elution buffer (300 mM NaOAc pH 5.5, 1 mM EDTA pH 8) to the crushed gel, overnight incubation (4°C), and centrifugation in Corning-Costar Spin-X Centrifuge Tube Filters (20,000 × g, 3 min, room temperature). The isolation of RNA fragments corresponding to ribosomal footprints of 15–45 nt size was verified using the RNA 6000 Nano Kit (Agilent) and the Agilent 2100 Bioanalyzer. Then, 5′ phosphorylation of RNA fragments was achieved using T4 polynucleotide kinase (NEB) and ATP (NEB). After RNA recovery using an RNA MinElute Cleanup Kit (QIAGEN), footprint fragments were again evaluated using the Agilent 2100 Bioanalyzer as described above. cDNA libraries were constructed using the NEBNext Small RNA Library Prep Set for Illumina (NEB) with 14 PCR amplification cycles. cDNA library quality was assessed using a High Sensitivity DNA Kit (Agilent). cDNA libraries were purified using the QIAquick PCR Purification Kit (QIAGEN) and sequenced using a HiSeq 1500 machine (Illumina) in single-read mode with a 50-bp read length.

RNA-Seq analysis

After flash-freezing the cultures for Ribo-Seq, 6 mL of the remaining culture volume was mixed with 1.2 mL of Stop Mix solution [95% (vol/vol) ethanol, 5% (vol/vol) phenol] to terminate ongoing transcription and translation. Samples were frozen in liquid nitrogen and stored at −80°C until RNA isolation. Cells were pelleted (3,000 × g, 15 min, 4°C), and total RNA was isolated using the miRNeasy Mini Kit (QIAGEN) in combination with the RNase-Free DNase Set (QIAGEN). The integrity of RNA samples was evaluated using the RNA 6000 Nano Kit (Agilent) and the Agilent 2100 Bioanalyzer. RNA was quantified using the Qubit RNA HS Assay Kit (Invitrogen). Ribosomal RNA depletion was performed using the NEBNext rRNA Depletion Kit for bacteria (NEB), and directional cDNA libraries were prepared using the NEBNext Ultra II Directional RNA Library Prep Kit for Illumina (NEB). cDNA library quality was assessed using a High Sensitivity DNA Kit (Agilent). Finally,

cDNA libraries were sequenced using a NextSeq 1000 machine (Illumina) in single-read mode with a 60-bp read length.

Next-generation sequencing data analysis

E. coli Ribo-seq and RNA-seq raw sequencing libraries were processed and analyzed using the published HRIBO workflow (version 1.6.0) (37), which has previously been used for analysis of bacterial Ribo-seq data (35). HRIBO is a snakemake (132) workflow that downloads all required tools from Bioconda (133) and Singularity (134). All necessary processing steps are automatically determined by the workflow.

Adapter and quality trimming of the reads was performed using cutadapt (version 2.1) (135). The trimmed reads were then mapped against the *E. coli* K-12 MG1655 reference genome (ASM584v2) with segemehl (version 0.3.4) (136). Reads corresponding to ribosomal RNA (rRNA), transfer RNA (tRNA), and multi-mapping reads were removed with SAMtools (version 1.9) (137) using rRNA and tRNA annotations.

Quality control was performed by creating read count statistics for each processing step and RNA class with Subread featureCounts (1.6.3) (138). All processing steps were analyzed with FastQC (version 0.11.8) (139), and results were aggregated with MultiQC (version 1.7) (140). Additionally, a PCA was performed to determine whether the major source of variance in the data stems from the different experimental conditions. The PCA ensures the correctness of the downstream differential expression analysis. To generate the PCA plots, normalized read counts for all samples were generated using DESeq2 (141) normalization. To improve the clustering, a regularized log transform was applied to the normalized read counts. Standard deviations and variance were subsequently calculated using DESeq2, and the first three principle components were plotted using plotly (142).

Read coverage files were generated with HRIBO using different full read mapping approaches (global or centered) and single-nucleotide mapping strategies (5′ or 3′ end). Read coverage files were normalized using the counts per million (mil) normalization. For the mil normalization, read counts were normalized by the total number of mapped reads within the sample and scaled by a per-million factor.

Metagene analysis of ribosome density at start codons was performed as described previously (143). Here, annotated start codons of coding sequences are collected, and the density of the read coverage is determined for every position in a pre-determined window around the collected start codons.

A differential expression and translation analysis was performed using the tool deltaTE (43). The tool combines both Ribo-seq and RNA-seq data by calculating the TE of genes in order to capture changes in translational regulation when comparing different growth conditions.

Gene set enrichment analysis

GSEA was conducted using the tool *clusterProfiler* (52). To this end, the genome-wide annotation database for *E. coli* K-12 MG1655 (144) was used in combination with the results of the differential expression analysis. The analysis was focused on the biological process domain, omitting the cellular component and molecular function domains. Furthermore, the minimum gene set size was set to 5 and the maximum to 30. To avoid redundancy within the results, GO terms were filtered, and only terms at the bottom (lowest branch level) of the GO hierarchy were analyzed.

Detection and differential expression of novel acid-induced sORF candidates

sORF candidates were initially detected using the neural network *DeepRibo* (99). Summary statistics, including TE, rpkm, codon counts, and nucleotide and amino acid sequences, for annotated and potential novel sORFs were computed using HRIBO (version 1.6.0) (37). Moreover, GFF track files were created for detailed manual inspection using the web-based genome browser JBrowse2 (100).

DeepRibo was reported to produce high numbers of false positives (145). Therefore, the high number of initial predictions (>25,000) was reduced by introducing cut-off criteria based on sORF length and Ribo-Seq coverage. Predictions with average rpkm (reads per kilobase of transcript per million mapped reads) values <30 across all Ribo-Seq samples and outside of the codon count range of 10–70 amino acids were excluded from further analysis. To specifically detect sORFs, which are only detectable under acidic conditions and involved in the acid response of *E. coli*, the search was further restricted to predictions with deltaTE Ribo- and RNA $\log_2$ FC values of ≥2, in combination with *P*-adjust values of ≤0.05, at either pH 5.8 or pH 4.4, compared to pH 7.6. Additionally, *DeepRibo* predictions of overlapping annotated genes on the same strand were excluded. The remaining 152 candidates were manually inspected in JBrowse2, and novel sORFs were included in our final candidate list (Table S7) if the coverage was even over the predicted ORF, restricted to the ORF boundaries, and potential Shine-Dalgarno sequences were detectable. The remaining candidates were manually curated, and, in a few cases, alternative sORFs in the genomic vicinity of DeepRibo predictions were selected, which matched better to the Ribo-Seq coverage signal. Differential expression values, comparing pH 7.6, pH 5.8, and pH 4.4, for novel sORF candidates, previously known sORFs, and sORFs detected by Storz and colleagues (94), were calculated using *deltaTE* (43).

Homologs for novel sORF candidates were identified using blastp and tblastn (108) within the "BLAST at NCBI" function in the CLC Main Workbench v. 20.0.4 (QIAGEN). For tblastn, the following parameters were used: an *E*-value (expected value) ≤0.05, a seed length that initiates an alignment (word size) of 6, and the filter for low-complexity regions off. Potential cellular localization was evaluated with PSORTb v3.0.3 (https://www.psort.org/psortb/) (104), and transmembrane topology was assessed using DeepTMHMM v1.0.24 (105). The protein structure of sORF15 was predicted by running AlphaFold2 (106) via Google Colab (ColabFold v1.5.2) (107) with default settings.

RNA isolation and RT-qPCR analysis

RNA was isolated using the Quick-RNA Miniprep Kit (Zymo Research) or miRNeasy Mini Kit (QIAGEN) in combination with the RNase-Free DNase Set (QIAGEN) according to the manufacturer's instructions. Total RNA was DNase digested for 30 min at 37°C using 1 µL TURBO DNase (2 U/µL) (Invitrogen). A 500-ng aliquot of the isolated RNA was converted to cDNA with the iScript Advanced Kit (Bio-Rad) according to the manufacturer's instructions. Then, 1 µL of a 1:10 dilution in nuclease-free water of the cDNA samples was mixed with 5 µL of SsoAdvanced Univ SYBR Green Supermix (Bio-Rad) and 0.8 µL of 5 µM forward and reverse primers (Table S10). The total reaction volume was adjusted to 10 µL with nuclease-free water, dispensed in triplicates in a 96-well PCR plate (Bio-Rad), and subjected to qPCR in a Bio-Rad CFX real-time cycler. Evaluation of the obtained data were performed according to the ΔΔCt method (146), using *recA* or *secA* genes as internal references.

Acid shock assay

Acid resistance was determined based on previously described protocols (16, 147) with the following modifications: *E. coli* BW25113 cells were grown at 37°C in LB pH 7.6 to OD_{600} = 0.5, adjusted to OD_{600} = 1, and then shifted for 15 min each, first to LB pH 5.8 and then to LB pH 4.4. Then, cells were shifted to LB pH = 3 for 1 h at 37°C. As a control, cells were cultivated at pH 7.6 throughout the experiment. After 1 h at pH 7.6, or pH 3, samples were serially diluted in 1× phosphate-buffered saline (PBS) and plated on LB agar plates to count the number of colonies. Percent survival was calculated as the ratio of colony-forming units at pH 3 and pH 7.6. For complementation of mutants from the Keio collection (84), BW25113 strains were transformed with IPTG-inducible pCA24N plasmids from the ASKA collection (Table S11) (85). Acid shock survival rates were determined as described above, with the exception that media were supplemented with 50 µg/mL chloramphenicol and different concentrations of IPTG to mimic native

expression levels. pCA24N-control, *gadE*, and *gadW* vectors were induced with 100 μM IPTG, pCA24N-*rcsB*, and *iscR* with 10 μM IPTG, and pCA24N-*iscR*, *ydcI*, and *mhpR* with 1 μM IPTG.

Promoter activity assay

In vivo promoter activities of *gadBC*, *adiA*, and *cadBA* were determined using luminescence-based reporter plasmids harboring fusions of the respective promoter regions to the *luxCDABE* genes from *Photorhabdus luminescens*. BW25113 wild-type cells or corresponding mutants from the Keio collection (84) were transformed with plasmids pBBR1-P*gadBC:lux*, pBBR1-P*adiA:lux*, or pBBR1-P*cadBA:lux*. All strains were cultivated in LB medium supplemented with gentamycin overnight. The overnight cultures were inoculated to an OD_{600} of 0.05 in fresh LB medium (pH 7.6), aerobically cultivated until exponential phase (OD_{600} = 0.5), and shifted to LB pH 5.8. To assess the promoter activity of *adiA* and *gadBC*, the cultures were shifted again to LB pH 4.4 after 15 min of growth in LB pH 5.8. In the next step, the cells were transferred to a 96-well plate and aerobically cultivated at 37°C in LB medium at different pH values supplemented with gentamycin. Growth and bioluminescence were measured every 10 min in the microtiter plates using a CLARIOstar Plus plate reader (BMG Labtech). Data are reported as relative light units in counts per second of OD_{600}.

Autoencoder-based identification of acid-specific genes and biological processes

The denoising autoencoders in this study were implemented using the Python package Keras (148). Details of the various hyperparameter and network architecture choices were taken from reference (113). In brief, an ensemble of 100 DAEs with two hidden layers of 2,000 and 1,000 nodes between the input layer and the bottleneck was used. The bottleneck layer of each network consists of 50 nodes, as this was found to be within an optimal range for DAEs trained on the PRECISE 2.0 expression compendium (114) by Kion-Crosby and Barquist (113). All layers have sigmoid activation functions; the weights of each layer were randomly initialized based on the Glorot distribution, and the bias vectors were zeros. The weight matrices that make up the decoder of each network are tied together such that they consist of the transpose of the corresponding weight matrices of each encoder.

All networks were trained in each ensemble using the Adam optimization algorithm. Data corruption during training was employed to improve generalizability, such that 10% of the entries of each input data point were randomly set to zero during each training step. Additionally, early stopping was employed during training; the data were randomly portioned into an 80% training set and a 10% validation set, and training was stopped once the validation score began to worsen. A 10% test set was also portioned to determine the optimal training parameters. A local search over all training parameters, including the learning rate and batch size, was performed using the test score as a metric, and training was done with batch shuffling enabled.

For the training of the autoencoder ensemble, the PRECISE 2.0 compendium (114) was used, as well as six additional data points (115) representing *E. coli* K-12 MG1655 strains grown in M9 medium and treated with various antibiotics, in addition to the nine data points from the current study. All data were converted to log transcripts per million, and features were normalized after the train/validation/test split such that the expression of each gene was scaled between 0 and 1.

Following the procedure described by Kion-Crosby and Barquist (113), the encoder of each of the 100 trained networks was used to determine which nodes turn on for the specific stress conditions of interest (e.g., pH 4.4) and simultaneously turn off for alternative stress conditions (e.g., oxidative stress, heat stress, and ethanol). This was done by passing each data point into each encoder while observing the activations of nodes at the bottleneck layer. After identifying which nodes are specific to the stress condition of interest, each corresponding decoder was utilized to generate gene

expression predictions for all *E. coli* genes by activating each of these bottleneck nodes individually and propagating this signal through the decoder. After sorting all genes based on the average decoder output from all identified nodes, the top of this sorted vector was used to define a gene set associated with the condition of interest.

Additionally, GSEA was run on the outputs of each decoder for each node to determine which biological processes were associated with the condition of interest. Since these groups often consist of ~100 nodes and thousands of GO terms are evaluated, a conservative threshold for the adj. *P*-value of 0.0005 was set. The adj. *P*-value was found using the BH method, and each *P*-value was computed based on 1,000,000 permutations. Finally, for the selection of the high-confidence acid-specific gene candidates, first the log_2 FC between each acidic condition and every other stress condition was taken for all genes in the acid-specific gene sets. Genes corresponding to a log_2 FC of at least 0.5 for at least 95% of comparisons were then selected.

Propidium iodide viability staining

Cells were grown either at pH 7.6, 5.8, or 4.4 in the same manner as described for Ribo-Seq and RNA-Seq (Fig. 1A). As a positive control, cells were grown to an OD_{600} of 1 and subsequently heat-shocked for 5 min at 80°C. After cultivation, 1 mL of each culture was centrifuged (15,000 × *g*, room temperature) and washed with PBS. Propidium iodide (Invitrogen) was added with a final concentration of 2 µg/mL, followed by a 5-min incubation at room temperature in the dark to label dead cells. After another wash step using PBS, 2 µL of the culture was spotted on 1% (wt/vol) agarose pads, placed onto microscope slides, and covered with a coverslip. Microscopic images were taken using a Leica DMi8 inverted microscope equipped with a Leica DFC365 FX camera. An excitation wavelength of 546 nm and a 605-nm emission filter with a 75-nm bandwidth were used to detect fluorescence with an exposure of 500 ms, a gain of 5, and 100% intensity in the Leica LAS X 3.7.4 software.

To quantify the relative fluorescent intensities (RF) of single cells, phase contrast and fluorescent images were analyzed using the ImageJ (149) plugin for MicrobeJ (150). Default settings of MicrobeJ were used for cell segmentation (Fit shape, rod-shaped bacteria) apart from the following settings: area, 0.1-max $µm^2$; length, 1.2–5 µm; width, 0.1–1 µm; curvature, 0.–0.15; and angularity, 0.–0.25 for *E. coli* cells. In total, ≥1,000 cells were quantified per strain and condition, and the background of the agarose pad was subtracted from each cell per field of view. Cells with RF values ≥300 after subtraction of the background were considered dead.

Western blot analysis

To verify the translation of sORF15, *E. coli* MG1655 cells harboring a pBAD24-sORF15:3xFLAG plasmid were cultivated in LB medium. Once an OD_{600} of 0.5 was reached, expression of sORF15 was induced for 1 h by supplementation of a final concentration of 0.2% (vol/vol) L-arabinose. Next, culture aliquots normalized to an OD_{600} of 1 were collected by centrifugation (1 min, 13,000 × *g*, 4°C). Supernatants were removed, and pellets were resuspended in 100 µL of protein loading buffer [12% SDS (wt/vol), 30% glycerol (wt/vol), 0.05% Coomassie blue (wt/vol), 150 mM Tris-HCl pH 7]. Prior to loading, protein samples were denatured at 95°C for 5 min and then chilled on ice. Protein fractions were separated by Tricine-SDS-PAGE on a 16% gel containing 6 M urea (151). After separation, proteins were transferred onto a nitrocellulose membrane in a semidry blot chamber. The membrane was blocked for 1 h in 5% (wt/vol) skim milk in 1× TBS. After short washing with Tris-buffered saline with Tween (TBS-T), blots were hybridized with the primary antibody (α-FLAG, Invitrogen) for 1 h at room temperature. After three washing steps in TBS-T for 10 min each, membranes were incubated with an alkaline phosphatase-conjugated secondary antibody (α-rabbit, Rockland Immunochemicals) for 1 h at room temperature. After three washing steps conducted as described above, proteins were visualized using colorimetric detection of alkaline phosphatase activity with 5-bromo-4-chloro-3-indolyl phosphate and nitro blue tetrazolium chloride.

As a ladder, the PageRuler Prestained Protein Ladder (10–180 kDa, Thermo Fisher) was used.

AUTHOR AFFILIATIONS

[1]Faculty of Biology, Microbiology, Ludwig-Maximilians-Universität München, Martinsried, Germany
[2]Bioinformatics Group, Department of Computer Science, Albert-Ludwigs-Universität Freiburg, Freiburg, Germany
[3]Helmholtz Institute for RNA-based Infection Research (HIRI)/Helmholtz Centre for Infection Research (HZI), Würzburg, Germany
[4]University of Würzburg, Faculty of Medicine, Würzburg, Germany
[5]Signalling Research Centres BIOSS and CIBSS, University of Freiburg, Freiburg, Germany

AUTHOR ORCIDs

Kilian Schumacher ⓘ http://orcid.org/0000-0002-4751-301X
Rick Gelhausen ⓘ http://orcid.org/0000-0003-4161-4394
Willow Kion-Crosby ⓘ http://orcid.org/0000-0001-9401-9340
Kirsten Jung ⓘ http://orcid.org/0000-0003-0779-6841

AUTHOR CONTRIBUTIONS

Kilian Schumacher, Conceptualization, Investigation, Methodology, Writing – original draft, Writing – review and editing | Rick Gelhausen, Investigation, Methodology, Software, Writing – review and editing | Willow Kion-Crosby, Investigation, Methodology, Software, Writing – review and editing | Lars Barquist, Funding acquisition, Methodology, Software, Writing – review and editing | Rolf Backofen, Funding acquisition, Methodology, Software, Writing – review and editing | Kirsten Jung, Conceptualization, Funding

acquisition, Methodology, Supervision, Writing – original draft, Writing – review and editing

DATA AVAILABILITY

The raw sequencing data for both the Ribo-seq and RNA-seq were deposited at Gene Expression Omnibus (GEO) under the accession number GSE219022. A JBrowse2 (100) online genome browser instance containing annotation files and read coverage files of the processed data is available via the RIBOBASE (http://www.bioinf.uni-freiburg.de/ribo-base). Additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

ADDITIONAL FILES

The following material is available online.

Supplemental Material

Supplemental material (mSystems01037-23-S0001.docx). Fig. S1 to S9; Tables S1, S2, and S6.
Supplemental resource table (mSystems01037-23-S0002.docx). Table of key resources.
Table S3 (mSystems01037-23-S0003.xlsx). Ribo-Seq and RNA-Seq exclusive genes.
Table S4 (mSystems01037-23-S0004.xlsx). Genes with differential translation efficiency under acid stress.
Table S5 (mSystems01037-23-S0005.xlsx). Complete HRIBO results.
Table S7 (mSystems01037-23-S0006.xlsx). Acid-induced novel candidate sORFs.
Table S8 (mSystems01037-23-S0007.xlsx). Differential expression of sORFs under acid stress.
Table S9 (mSystems01037-23-S0008.xlsx). *Escherichia coli* strains used in this study.
Table S10 (mSystems01037-23-S0009.xlsx). Oligonucleotides used in this study.
Table S11 (mSystems01037-23-S00010.xlsx). Plasmids used in this study.

REFERENCES

1. Schwarz J, Schumacher K, Brameyer S, Jung K. 2022. Bacterial battle against acidity. FEMS Microbiol Rev 46:fuac037. https://doi.org/10.1093/femsre/fuac037
2. Tilden J, Young W, McNamara AM, Custer C, Boesel B, Lambert-Fair MA, Majkowski J, Vugia D, Werner SB, Hollingsworth J, Morris JG. 1996. A new route of transmission for *Escherichia coli*: infection from dry fermented salami. Am J Public Health 86:1142–1145. https://doi.org/10.2105/ajph.86.8_pt_1.1142
3. Hirshfield IN, Terzulli S, O'Byrne C. 2003. Weak organic acids: a panoply of effects on bacteria. Sci Prog 86:245–269. https://doi.org/10.3184/003685003783238626
4. Deamer DW. 1987. Proton permeation of lipid bilayers. J Bioenerg Biomembr 19:457–479. https://doi.org/10.1007/BF00770030
5. Wilks JC, Slonczewski JL. 2007. pH of the cytoplasm and periplasm of *Escherichia coli*: rapid measurement by green fluorescent protein fluorimetry. J Bacteriol 189:5601–5607. https://doi.org/10.1128/JB.00615-07
6. Goto Y, Takahashi N, Fink AL. 1990. Mechanism of acid-induced folding of proteins. Biochemistry 29:3480–3488. https://doi.org/10.1021/bi00466a009
7. Lund P, Tramonti A, De Biase D. 2014. Coping with low pH: molecular strategies in neutralophilic bacteria. FEMS Microbiol Rev 38:1091–1125. https://doi.org/10.1111/1574-6976.12076
8. Jeong KC, Hung KF, Baumler DJ, Byrd JJ, Kaspar CW. 2008. Acid stress damage of DNA is prevented by Dps binding in *Escherichia coli* O157:H7. BMC Microbiol 8:1–13. https://doi.org/10.1186/1471-2180-8-181
9. Slonczewski JL, Fujisawa M, Dopson M, Krulwich TA. 2009. Cytoplasmic pH measurement and homeostasis in bacteria and archaea. Adv Microb Physiol 55:1–79. https://doi.org/%2010.1016/S0065-2911(09)05501-5
10. Arcari T, Feger M-L, Guerreiro DN, Wu J, O'Byrne CP. 2020. Comparative review of the responses of *Listeria monocytogenes* and *Escherichia coli* to low pH stress. Genes 11:1330. https://doi.org/10.3390/genes11111330
11. Foster JW. 2004. *Escherichia coli* acid resistance: tales of an amateur acidophile. Nat Rev Microbiol 2:898–907. https://doi.org/10.1038/nrmicro1021
12. Castanie-Cornet MP, Foster JW. 2001. *Escherichia coli* acid resistance: cAMP receptor protein and a 20 bp *cis*-acting sequence control pH and stationary phase expression of the *gadA* and *gadBC* glutamate decarboxylase genes. Microbiology 147:709–715. https://doi.org/10.1099/00221287-147-3-709
13. Kanjee U, Houry WA. 2013. Mechanisms of acid resistance in *Escherichia coli*. Annu Rev Microbiol 67:65–81. https://doi.org/10.1146/annurev-micro-092412-155708
14. De Biase D, Lund PA. 2015. The *Escherichia coli* acid stress response and its significance for pathogenesis. Adv Appl Microbiol 92:49–88. https://doi.org/10.1016/bs.aambs.2015.03.002
15. Brameyer S, Schumacher K, Kuppermann S, Jung K. 2022. Division of labor and collective functionality in *Escherichia coli* under acid stress. Commun Biol 5:327. https://doi.org/10.1038/s42003-022-03281-4
16. Brameyer S, Hoyer E, Bibinger S, Burdack K, Lassak J, Jung K. 2020. Molecular design of a signaling system influences noise in protein abundance under acid stress in different gammaproteobacteria. J Bacteriol 202:e00121-20. https://doi.org/10.1128/JB.00121-20
17. De Biase D, Pennacchietti E. 2012. Glutamate decarboxylase-dependent acid resistance in orally acquired bacteria: function, distribution and biomedical implications of the *gadBC* operon. Mol Microbiol 86:770–786. https://doi.org/10.1111/mmi.12020

18. Maurer LM, Yohannes E, Bondurant SS, Radmacher M, Slonczewski JL. 2005. pH regulates genes for flagellar motility, catabolism, and oxidative stress in *Escherichia coli* K-12. J Bacteriol 187:304–319. https://doi.org/10.1128/JB.187.1.304-319.2005

19. Blankenhorn D, Phillips J, Slonczewski JL. 1999. Acid- and base-induced proteins during aerobic and anaerobic growth of *Escherichia coli* revealed by two-dimensional gel electrophoresis. J Bacteriol 181:2209–2216. https://doi.org/10.1128/JB.181.7.2209-2216.1999

20. Stancik LM, Stancik DM, Schmidt B, Barnhart DM, Yoncheva YN, Slonczewski JL. 2002. pH-dependent expression of periplasmic proteins and amino acid catabolism in *Escherichia coli*. J Bacteriol 184:4246–4258. https://doi.org/10.1128/JB.184.15.4246-4258.2002

21. Du B, Yang L, Lloyd CJ, Fang X, Palsson BO. 2019. Genome-scale model of metabolism and gene expression provides a multi-scale description of acid stress responses in *Escherichia coli*. PLoS Comput Biol 15:e1007525. https://doi.org/10.1371/journal.pcbi.1007525

22. Seo SW, Kim D, O'Brien EJ, Szubin R, Palsson BO. 2015. Decoding genome-wide GadEWX-transcriptional regulatory networks reveals multifaceted cellular responses to acid stress in *Escherichia coli*. Nat Commun 6:7970. https://doi.org/10.1038/ncomms8970

23. Schmidt A, Kochanowski K, Vedelaar S, Ahrné E, Volkmer B, Callipo L, Knoops K, Bauer M, Aebersold R, Heinemann M. 2016. The quantitative and condition-dependent *Escherichia coli* proteome. Nat Biotechnol 34:104–110. https://doi.org/10.1038/nbt.3418

24. Ingolia NT, Ghaemmaghami S, Newman JRS, Weissman JS. 2009. Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. Science 324:218–223. https://doi.org/10.1126/science.1168978

25. Ingolia NT, Brar GA, Rouskin S, McGeachy AM, Weissman JS. 2012. The Ribosome profiling strategy for monitoring translation *in vivo* by deep sequencing of ribosome-protected mRNA fragments. Nat Protoc 7:1534–1550. https://doi.org/10.1038/nprot.2012.086

26. Ingolia NT, Hussmann JA, Weissman JS. 2019. Ribosome profiling: global views of translation. Cold Spring Harb Perspect Biol 11:1–20. https://doi.org/10.1101/cshperspect.a032698

27. Li GW, Burkhardt D, Gross C, Weissman JS. 2014. Quantifying absolute protein synthesis rates reveals principles underlying allocation of cellular resources. Cell 157:624–635. https://doi.org/10.1016/j.cell.2014.02.033

28. Vazquez-Laslop N, Sharma CM, Mankin A, Buskirk AR. 2022. Identifying small open reading frames in prokaryotes with ribosome profiling. J Bacteriol 204:e0029421. https://doi.org/10.1128/JB.00294-21

29. Gray T, Storz G, Papenfort K. 2022. Small proteins; big questions. J Bacteriol 204:e0034121. https://doi.org/10.1128/JB.00341-21

30. Liu X, Jiang H, Gu Z, Roberts JW. 2013. High-resolution view of bacteriophage lambda gene expression by ribosome profiling. Proc Natl Acad Sci USA 110:11928–11933. https://doi.org/10.1073/pnas.1309739110

31. Gelsinger DR, Dallon E, Reddy R, Mohammad F, Buskirk AR, DiRuggiero J. 2021. Ribosome profiling in archaea reveals leaderless translation, novel translational initiation sites, and ribosome pausing at single codon resolution. Nucleic Acids Res 48:5201–5216. https://doi.org/10.1093/nar/gkaa304

32. Fremin BJ, Sberro H, Bhatt AS. 2020. MetaRibo-seq measures translation in microbiomes. Nat Commun 11:3268. https://doi.org/10.1038/s41467-020-17081-z

33. Hadjeras L, Bartel J, Maier L-K, Maaß S, Vogel V, Svensson SL, Eggenhofer F, Gelhausen R, Müller T, Alkhnbashi OS, Backofen R, Becher D, Sharma CM, Marchfelder A. 2023. Revealing the small proteome of *Haloferax volcanii* by combining ribosome profiling and small-protein optimised mass spectrometry. Microlife 4:uqad001. https://doi.org/10.1093/femsml/uqad001

34. Fuchs S, Kucklick M, Lehmann E, Beckmann A, Wilkens M, Kolte B, Mustafayeva A, Ludwig T, Diwo M, Wissing J, Jänsch L, Ahrens CH, Ignatova Z, Engelmann S. 2021. Towards the characterization of the hidden world of small proteins in *Staphylococcus aureus*, a proteogenomics approach. PLoS Genet 17:e1009585. https://doi.org/10.1371/journal.pgen.1009585

35. Venturini E, Svensson SL, Maaß S, Gelhausen R, Eggenhofer F, Li L, Cain AK, Parkhill J, Becher D, Backofen R, Barquist L, Sharma CM, Westermann AJ, Vogel J. 2020. A global data-driven census of *Salmonella* small

36. Boulos L, Prévost M, Barbeau B, Coallier J, Desjardins R. 1999. LIVE/DEAD baclight: application of a new rapid staining method for direct enumeration of viable and total bacteria in drinking water. J Microbiol Methods 37:77–86. https://doi.org/10.1016/s0167-7012(99)00048-2

37. Gelhausen R, Svensson SL, Froschauer K, Heyl F, Hadjeras L, Sharma CM, Eggenhofer F, Backofen R. 2021. HRIBO: high-throughput analysis of bacterial ribosome profiling data. Bioinformatics 37:2061–2063. https://doi.org/10.1093/bioinformatics/btaa959

38. Mohammad F, Green R, Buskirk AR. 2019. A systematically-revised ribosome profiling method for bacteria reveals pauses at single-codon resolution. Elife 8:1–25. https://doi.org/10.7554/eLife.42591

39. Latif H, Szubin R, Tan J, Brunk E, Lechner A, Zengler K, Palsson BO. 2015. A streamlined ribosome profiling protocol for the characterization of microorganisms. Biotechniques 58:329–332. https://doi.org/10.2144/000114302

40. Mohammad F, Buskirk AR. 2019. Protocol for ribosome profiling in bacteria. Bio Protoc 9:e3468. https://doi.org/10.21769/BioProtoc.3468

41. Zhang Y, Xiao Z, Zou Q, Fang J, Wang Q, Yang X, Gao N. 2017. Ribosome profiling reveals genome-wide cellular translational regulation upon heat stress in *Escherichia coli*. Genomics, Proteomics & Bioinformatics 15:324–330. https://doi.org/10.1016/j.gpb.2017.04.005

42. Gerashchenko MV, Lobanov AV, Gladyshev VN. 2012. Genome-wide ribosome profiling reveals complex translational regulation in response to oxidative stress. Proc Natl Acad Sci U S A 109:17394–17399. https://doi.org/10.1073/pnas.1120799109

43. Chothani S, Adami E, Ouyang JF, Viswanathan S, Hubner N, Cook SA, Schafer S, Rackham OJL. 2019. deltaTE: detection of translationally regulated genes by integrative analysis of Ribo-Seq and RNA-Seq data. Curr Protoc Mol Biol 129:e108. https://doi.org/10.1002/cpmb.108

44. Prior TI, Kornberg HL. 1988. Nucleotide sequence of *fruA*, the gene specifying enzyme II^{fru} of the phosphoenolpyruvate-dependent sugar phosphotransferase system in *Escherichia coli* K12. Microbiology 134:2757–2768. https://doi.org/10.1099/00221287-134-10-2757

45. Dai X, Zhu M, Warren M, Balakrishnan R, Okano H, Williamson JR, Fredrick K, Hwa T, Lee SY, Pedersen S, Record M. 2018. Slowdown of translational elongation in *Escherichia coli* under hyperosmotic stress. mBio 9:1. https://doi.org/10.1128/mBio.02375-17

46. Sin C, Chiarugi D, Valleriani A. 2016. Quantitative assessment of ribosome drop-off in *E. coli*. Nucleic Acids Res 44:2528–2537. https://doi.org/10.1093/nar/gkw137

47. Sharma IM, Woodson SA. 2020. RbfA and IF3 couple ribosome biogenesis and translation initiation to increase stress tolerance. Nucleic Acids Res 48:359–372. https://doi.org/10.1093/nar/gkz1065

48. Vesper O, Amitai S, Belitsky M, Byrgazov K, Kaberdina AC, Engelberg-Kulka H, Moll I. 2011. Selective translation of leaderless mRNAs by specialized ribosomes generated by MazF in *Escherichia coli*. Cell 147:147–157. https://doi.org/10.1016/j.cell.2011.07.047

49. Moll I, Engelberg-Kulka H. 2012. Selective translation during stress in *Escherichia coli*. Trends Biochem Sci 37:493–498. https://doi.org/10.1016/j.tibs.2012.07.007

50. Freedberg WB, Lin ECC. 1973. Three kinds of controls affecting the expression of the *glp* regulon in *Escherichia coli*. J Bacteriol 115:816–823. https://doi.org/10.1128/jb.115.3.816-823.1973

51. Higashi K, Ishigure H, Demizu R, Uemura T, Nishino K, Yamaguchi A, Kashiwagi K, Igarashi K. 2008. Identification of a spermidine excretion protein complex (MdtJI) in *Escherichia coli*. J Bacteriol 190:872–878. https://doi.org/10.1128/JB.01505-07

52. Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, Feng T, Zhou L, Tang W, Zhan L, Fu X, Liu S, Bo X, Yu G. 2021. clusterProfiler 4.0: a universal enrichment tool for interpreting omics data. The Innovation 2:100141. https://doi.org/10.1016/j.xinn.2021.100141

53. Yohannes E, Thurber AE, Wilks JC, Tate DP, Slonczewski JL. 2005. Polyamine stress at high pH in *Escherichia coli* K-12. BMC Microbiol 5:1–8. https://doi.org/10.1186/1471-2180-5-59

54. Iyer R, Delcour AH. 1997. Complex inhibition of OmpF and OmpC bacterial porins by polyamines. J Biol Chem 272:18595–18601. https://doi.org/10.1074/jbc.272.30.18595

55. Samartzidou H, Mehrazin M, Xu Z, Benedik MJ, Delcour AH. 2003. Cadaverine inhibition of porin plays a role in cell survival at acidic pH. J Bacteriol 185:13–19. https://doi.org/10.1128/JB.185.1.13-19.2003

56. Banerji R, Iyer P, Bhagwat A, Saroj SD. 2022. Spermidine promotes lysozyme tolerance and acid stress resistance in Streptococcus pyogenes M3. Microbiology 168. https://doi.org/10.1099/mic.0.001228

57. Fukuchi J, Kashiwagi K, Yamagishi M, Ishihama A, Igarashi K. 1995. Decrease in cell viability due to the accumulation of spermidine in spermidine acetyltransferase-deficient mutant of Escherichia coli. J Biol Chem 270:18831–18835. https://doi.org/10.1074/jbc.270.32.18831

58. Schryvers A, Weiner JH. 1981. The anaerobic sn-glycerol-3-phosphate dehydrogenase of Escherichia coli: purification and characterization. J Biol Chem 256:9959–9965. https://doi.org/10.1016/S0021-9258(19)68723-5

59. Schweizer H, Larson TJ. 1987. Cloning and characterization of the aerobic sn-glycerol-3-phosphate dehydrogenase structural gene glpD of Escherichia coli K-12. J Bacteriol 169:507–513. https://doi.org/10.1128/jb.169.2.507-513.1987

60. Paulsen IT, Reizer J, Jin RZ, Lin ECC, Saier MH Jr. 2000. Functional genomic studies of dihydroxyacetone utilization in Escherichia coli. Microbiology 146 (Pt 10):2343–2344. https://doi.org/10.1099/00221287-146-10-2343

61. Larson TJ, Ye SZ, Weissenborn DL, Hoffmann HJ, Schweizer H. 1987. Purification and characterization of the repressor for the sn-glycerol 3-phosphate regulon of Escherichia coli K12. J Biol Chem 262:15869–15874.

62. Sohlenkamp C. 2017. Membrane homeostasis in bacteria upon pH challenge, p 1–13. In Geiger O (ed), Biogenesis of fatty acids, lipids and membranes. Springer International Publishing, Cham. https://doi.org/10.1007/978-3-319-43676-0

63. Lindahl T, Nyberg B. 1972. Rate of depurination of native deoxyribonucleic acid. Biochemistry 11:3610–3618. https://doi.org/10.1021/bi00769a018

64. Qi Y, Wang H, Chen X, Wei G, Tao S, Fan M. 2021. Altered metabolic strategies: elaborate mechanisms adopted by Oenococcus oeni in response to acid stress. J Agric Food Chem 69:2906–2918. https://doi.org/10.1021/acs.jafc.0c07599

65. Blethen SL, Boeker EA, Snell EE. 1968. Arginine decarboxylase from Escherichia coli. J Biol Chem 243:1671–1677. https://doi.org/10.1016/S0021-9258(18)93498-8

66. Seputiene V, Motiejūnas D, Suziedelis K, Tomenius H, Normark S, Melefors O, Suziedeliene E. 2003. Molecular characterization of the acid-inducible asr gene of Escherichia coli and its role in acid stress response. J Bacteriol 185:2475–2484. https://doi.org/10.1128/JB.185.8.2475-2484.2003

67. Ren C, Zheng Y, Liu C, Mencius J, Wu Z, Quan S. 2022. Molecular characterization of an intrinsically disordered chaperone reveals net-charge regulation in chaperone action. J Mol Biol 434:167405. https://doi.org/10.1016/j.jmb.2021.167405

68. Kannan G, Wilks JC, Fitzgerald DM, Jones BD, Bondurant SS, Slonczewski JL. 2008. Rapid acid treatment of Escherichia coli: transcriptomic response and recovery. BMC Microbiol 8:1–13. https://doi.org/10.1186/1471-2180-8-37

69. Celia H, Noinaj N, Buchanan SK. 2020. Structure and stoichiometry of the Ton molecular motor. Int J Mol Sci 21:1–15. https://doi.org/10.3390/ijms21020375

70. Ratliff AC, Buchanan SK, Celia H. 2022. The Ton motor. Front Microbiol 13:852955. https://doi.org/10.3389/fmicb.2022.852955

71. Vinella D, Brochier-Armanet C, Loiseau L, Talla E, Barras F. 2009. Iron-sulfur (Fe/S) protein biogenesis: phylogenomic and genetic studies of A-type carriers. PLoS Genet 5:e1000497. https://doi.org/10.1371/journal.pgen.1000497

72. Oakley K, Sterling K, Shearer J, Kim E. 2021. Controlled protonation of [2Fe-2S] leading to MitoNEET analogues and concurrent cluster modification. Inorg Chem 60:16074–16078. https://doi.org/10.1021/acs.inorgchem.1c02622

73. Jin Z, Li J, Ni L, Zhang R, Xia A, Jin F. 2018. Conditional privatization of a public siderophore enables Pseudomonas aeruginosa to resist cheater invasion. Nat Commun 9:1383. https://doi.org/10.1038/s41467-018-03791-y

74. Djoko KY, Phan M-D, Peters KM, Walker MJ, Schembri MA, McEwan AG. 2017. Interplay between tolerance mechanisms to copper and acid stress in Escherichia coli. Proc Natl Acad Sci USA 114:6818–6823. https://doi.org/10.1073/pnas.1620232114

75. Kim Y, Lee S, Park K, Yoon H. 2022. Cooperative interaction between acid and copper resistance in Escherichia coli . J Microbiol Biotechnol 32:602–611. https://doi.org/10.4014/jmb.2201.01034

76. Hyre A, Casanova-Hampton K, Subashchandrabose S. 2021. Copper homeostatic mechanisms and their role in the virulence of Escherichia coli and Salmonella enterica. EcoSal Plus 9:eESP00142020. https://doi.org/10.1128/ecosalplus.ESP-0014-2020

77. Djoko KY, Ong CY, Walker MJ, McEwan AG. 2015. The role of copper and zinc toxicity in innate immune defense against bacterial pathogens. J Biol Chem 290:18954–18961. https://doi.org/10.1074/jbc.R115.647099

78. Yanofsky C, Horn V, Gollnick P. 1991. Physiological studies of tryptophan transport and tryptophanase operon induction in Escherichia coli. J Bacteriol 173:6009–6017. https://doi.org/10.1128/jb.173.19.6009-6017.1991

79. Goode O, Smith A, Zarkan A, Cama J, Invergo BM, Belgami D, Caño-Muñiz S, Metz J, O'Neill P, Jeffries A, Norville IH, David J, Summers D, Pagliara S. 2021. Persister Escherichia coli cells have a lower intracellular pH than susceptible cells but maintain their pH in response to antibiotic treatment. mBio 12:e0090921. https://doi.org/10.1128/mBio.00909-21

80. Balaban NQ, Merrin J, Chait R, Kowalik L, Leibler S. 2004. Bacterial persistence as a phenotypic switch. Science 305:1622–1625. https://doi.org/10.1126/science.1099390

81. Mitchell P. 1973. Performance and conservation of osmotic work by proton-coupled solute porter systems. J Bioenerg 4:63–91. https://doi.org/10.1007/BF01516051

82. Richard H, Foster JW. 2004. Escherichia coli glutamate- and arginine-dependent acid resistance systems increase internal pH and reverse transmembrane potential. J Bacteriol 186:6032–6041. https://doi.org/10.1128/JB.186.18.6032-6041.2004

83. Ude S, Lassak J, Starosta AL, Kraxenberger T, Wilson DN, Jung K. 2013. Translation elongation factor EF-P alleviates ribosome stalling at polyproline stretches. Science 339:82–85. https://doi.org/10.1126/science.1228985

84. Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko KA, Tomita M, Wanner BL, Mori H. 2006. Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. Mol Syst Biol 2:2006.0008. https://doi.org/10.1038/msb4100050

85. Kitagawa M, Ara T, Arifuzzaman M, Ioka-Nakamichi T, Inamoto E, Toyonaga H, Mori H. 2005. Complete set of ORF clones of Escherichia coli ASKA library (a complete set of E. coli K-12 ORF archive): unique resources for biological research. DNA Res 12:291–299. https://doi.org/10.1093/dnares/dsi012

86. Romiyo V, Wilson JW. 2020. Phenotypes, transcriptome, and novel biofilm formation associated with the ydcI gene. Antonie van Leeuwenhoek 113:1109–1122. https://doi.org/10.1007/s10482-020-01412-7

87. Gao Y, Yurkovich JT, Seo SW, Kabimoldayev I, Dräger A, Chen K, Sastry AV, Fang X, Mih N, Yang L, Eichner J, Cho B-K, Kim D, Palsson BO. 2018. Systematic discovery of uncharacterized transcription factors in Escherichia coli K-12 MG1655. Nucleic Acids Res 46:10682–10696. https://doi.org/10.1093/nar/gky752

88. Jennings ME, Quick LN, Soni A, Davis RR, Crosby K, Ott CM, Nickerson CA, Wilson JW. 2011. Characterization of the Salmonella enterica serovar Typhimurium ydcI gene, which encodes a conserved DNA binding protein required for full acid stress resistance. J Bacteriol 193:2208–2217. https://doi.org/10.1128/JB.01335-10

89. Ma Z, Masuda N, Foster JW. 2004. Characterization of EvgAS-YdeO-GadE branched regulatory circuit governing glutamate-dependent acid resistance in Escherichia coli. J Bacteriol 186:7378–7389. https://doi.org/10.1128/JB.186.21.7378-7389.2004

90. Torres B, Porras G, Garcia JL, Diaz E. 2003. Regulation of the mhp cluster responsible for 3-(3-hydroxyphenyl)propionic acid degradation in Escherichia coli. J Biol Chem 278:27575–27585. https://doi.org/10.1074/jbc.M303245200

91. Ferrández A, Garciá JL, Díaz E. 1997. Genetic characterization and expression in heterologous hosts of the 3-(3-hydroxyphenyl)propionate

catabolic pathway of *Escherichia coli* K-12. J Bacteriol 179:2573–2581. https://doi.org/10.1128/jb.179.8.2573-2581.1997

92. Yadavalli SS, Yuan J. 2022. Bacterial small membrane proteins: the swiss army knife of regulators at the lipid bilayer. J Bacteriol 204:e0034421. https://doi.org/10.1128/JB.00344-21

93. Petruschke H, Anders J, Stadler PF, Jehmlich N, von Bergen M. 2020. Enrichment and identification of small proteins in a simplified human gut microbiome. J Proteomics 213:103604. https://doi.org/10.1016/j.jprot.2019.103604

94. Weaver J, Mohammad F, Buskirk AR, Storz G. 2019. Identifying small proteins by ribosome profiling with stalled initiation complexes. mBio 10:1–21. https://doi.org/10.1128/mBio.02819-18

95. Meydan S, Marks J, Klepacki D, Sharma V, Baranov PV, Firth AE, Margus T, Kefi A, Vázquez-Laslop N, Mankin AS. 2019. Retapamulin-assisted ribosome profiling reveals the alternative bacterial proteome. Mol Cell 74:481–493. https://doi.org/10.1016/j.molcel.2019.02.017

96. Hadjeras L, Heiniger B, Maaß S, Scheuer R, Gelhausen R, Azarderakhsh S, Barth-Weber S, Backofen R, Becher D, Ahrens CH, Sharma CM, Evgenieva-Hackenberg E. 2023. Unraveling the small proteome of the plant symbiont *Sinorhizobium meliloti* by ribosome profiling and proteogenomics. Microlife 4:uqad012. https://doi.org/10.1093/femsml/uqad012

97. Adams PP, Baniulyte G, Esnault C, Chegireddy K, Singh N, Monge M, Dale RK, Storz G, Wade JT. 2021. Regulatory roles of *Escherichia coli* 5′ UTR and ORF-internal RNAs detected by 3′ end mapping. Elife 10:1–33. https://doi.org/10.7554/eLife.62438

98. Raina M, Aoyama JJ, Bhatt S, Paul BJ, Zhang A, Updegrove TB, Miranda-Ríos J, Storz G. 2022. Dual-function AzuCR RNA modulates carbon metabolism. Proc Natl Acad Sci U S A 119:e2117930119. https://doi.org/10.1073/pnas.2117930119

99. Clauwaert J, Menschaert G, Waegeman W. 2019. DeepRibo: a neural network for precise gene annotation of prokaryotes by combining ribosome profiling signal and binding site patterns. Nucleic Acids Res 47:e36. https://doi.org/10.1093/nar/gkz061

100. Diesh C, Stevens GJ, Xie P, De Jesus Martinez T, Hershberg EA, Leung A, Guo E, Dider S, Zhang J, Bridge C, Hogue G, Duncan A, Morgan M, Flores T, Bimber BN, Haw R, Cain S, Buels RM, Stein LD, Holmes IH. 2022. JBrowse 2: a modular genome browser with views of synteny and structural variation. bioRxiv. https://doi.org/10.1101/2022.07.28.501447

101. Guillier M, Gottesman S. 2008. The 5′ end of two redundant sRNAs is involved in the regulation of multiple targets, including their own regulator. Nucleic Acids Res 36:6781–6794. https://doi.org/10.1093/nar/gkn742

102. Johansen J, Rasmussen AA, Overgaard M, Valentin-Hansen P. 2006. Conserved small non-coding RNAs that belong to the σE regulon: role in down-regulation of outer membrane proteins. J Mol Biol 364:1–8. https://doi.org/10.1016/j.jmb.2006.09.004

103. Aquino P, Honda B, Jaini S, Lyubetskaya A, Hosur K, Chiu JG, Ekladious I, Hu D, Jin L, Sayeg MK, et al. 2017. Coordinated regulation of acid resistance in *Escherichia coli*. BMC Syst Biol 11:1. https://doi.org/10.1186/s12918-016-0376-y

104. Yu NY, Wagner JR, Laird MR, Melli G, Rey S, Lo R, Dao P, Sahinalp SC, Ester M, Foster LJ, Brinkman FSL. 2010. PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. Bioinformatics 26:1608–1615. https://doi.org/10.1093/bioinformatics/btq249

105. Hallgren J, Tsirigos KD, Pedersen MD, Almagro Armenteros JJ, Marcatili P, Nielsen H, Krogh A, Winther O. 2022. DeepTMHMM predicts alpha and beta transmembrane proteins using deep neural networks. Bioinformatics. https://doi.org/10.1101/2022.04.08.487609

106. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Žídek A, Potapenko A, et al. 2021. Highly accurate protein structure prediction with AlphaFold. Nature 596:583–589. https://doi.org/10.1038/s41586-021-03819-2

107. Mirdita M, Schütze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M. 2022. ColabFold: making protein folding accessible to all. Nat Methods 19:679–682. https://doi.org/10.1038/s41592-022-01488-1

108. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389–3402. https://doi.org/10.1093/nar/25.17.3389

109. Gottesman S. 2019. Trouble is coming: signaling pathways that regulate general stress responses in bacteria. J Biol Chem 294:11685–11700. https://doi.org/10.1074/jbc.REV119.005593

110. Battesti A, Majdalani N, Gottesman S. 2011. The RpoS-mediated general stress response in *Escherichia coli*. Annu Rev Microbiol 65:189–213. https://doi.org/10.1146/annurev-micro-090110-102946

111. Vincent P, Larochelle H, Bengio Y, Manzagol P-A. 2008. Extracting and composing robust features with denoising autoencoders. Proceedings of the 25th International Conference on Machine Learning:1096–1103. https://doi.org/10.1145/1390156.1390294

112. Vincent P, Larochelle H, Lajoie I, Bengio Y, Manzagol P-A. 2010. Stacked denoising autoencoders: learning useful representations in a deep network with a local denoising criterion. J Mach Learn Res 11:3371–3408.

113. Kion-Crosby W, Barquist L. 2023. Network depth affects inference of gene sets from bacterial transcriptomes using denoising autoencoders. bioRxiv. https://doi.org/10.1101/2023.05.30.542622

114. Lamoureux CR, Decker KT, Sastry AV, Rychel K, Gao Y, McConn JL, Zielinski DC, Palsson BO. 2021. A multi-scale transcriptional regulatory network knowledge base for *Escherichia coli* . bioRxiv. https://doi.org/10.1101/2021.04.08.439047

115. Bhatia RP, Kirit HA, Predeus AV, Bollback JP. 2022. Transcriptomic profiling of *Escherichia coli* K-12 in response to a compendium of stressors. Sci Rep 12:16865. https://doi.org/10.1038/s41598-022-21625-2

116. Sandberg TE, Pedersen M, LaCroix RA, Ebrahim A, Bonde M, Herrgard MJ, Palsson BO, Sommer M, Feist AM. 2014. Evolution of *Escherichia coli* to 42℃ and subsequent genetic engineering reveals adaptive mechanisms and novel mutations. Mol Biol Evol 31:2647–2662. https://doi.org/10.1093/molbev/msu209

117. Choudhary KS, Kleinmanns JA, Decker K, Sastry AV, Gao Y, Szubin R, Seif Y, Palsson BO. 2020. Elucidation of regulatory modes for five two-component systems in *Escherichia coli* reveals novel relationships. mSystems 5:1–20. https://doi.org/10.1128/mSystems.00980-20

118. Seo SW, Gao Y, Kim D, Szubin R, Yang J, Cho BK, Palsson BO. 2017. Revealing genome-scale transcriptional regulatory landscape of OmpR highlights its expanded regulatory roles under osmotic stress in *Escherichia coli* K-12 MG1655. Sci Rep 7:2181. https://doi.org/10.1038/s41598-017-02110-7

119. Anand A, Chen K, Catoiu E, Sastry AV, Olson CA, Sandberg TE, Seif Y, Xu S, Szubin R, Yang L, Feist AM, Palsson BO. 2020. OxyR is a convergent target for mutations acquired during adaptation to oxidative stress-prone metabolic states. Mol Biol Evol 37:660–667. https://doi.org/10.1093/molbev/msz251

120. Seo SW, Kim D, Szubin R, Palsson BO. 2015. Genome-wide reconstruction of OxyR and SoxRS transcriptional regulatory networks under oxidative stress in *Escherichia coli* K-12 MG1655. Cell Rep 12:1289–1299. https://doi.org/10.1016/j.celrep.2015.07.043

121. Cao H, Wei D, Yang Y, Shang Y, Li G, Zhou Y, Ma Q, Xu Y. 2017. Systems-level understanding of ethanol-induced stresses and adaptation in *E. coli*. Sci Rep 7:1–15. https://doi.org/10.1038/srep44150

122. Reyes-Fernández EZ, Schuldiner S. 2020. Acidification of cytoplasm in *Escherichia coli* provides a strategy to cope with stress and facilitates development of antibiotic resistance. Sci Rep 10:9954. https://doi.org/10.1038/s41598-020-66890-1

123. Mitosch K, Rieckh G, Bollenbach T. 2017. Noisy response to antibiotic stress predicts subsequent single-cell survival in an acidic environment. Cell Systems 4:393–403. https://doi.org/10.1016/j.cels.2017.03.001

124. Yair Y, Michaux C, Biran D, Bernhard J, Vogel J, Barquist L, Ron EZ, Van Laar TA. 2022. Cellular RNA targets of cold shock proteins CspC and CspE and their importance for serum resistance in septicemic *Escherichia coli*. mSystems 7:e00086-22 https://doi.org/10.1128/msystems.00086-22

125. Michaux C, Holmqvist E, Vasicek E, Sharan M, Barquist L, Westermann AJ, Gunn JS, Vogel J. 2017. RNA target profiles direct the discovery of virulence functions for the cold-shock proteins CspC and CspE. Proc Natl Acad Sci USA 114:6824–6829. https://doi.org/10.1073/pnas.1620772114

126. Sun Y, Fukamachi T, Saito H, Kobayashi H. 2012. Respiration and the F_1F_o-ATPase enhance survival under acidic conditions in *Escherichia coli*. PLoS One 7:e52577. https://doi.org/10.1371/journal.pone.0052577

127. Lee J, Hiibel SR, Reardon KF, Wood TK. 2010. Identification of stress-related proteins in *Escherichia coli* using the pollutant *cis*-dichloroethylene. J Appl Microbiol 108:2088–2102. https://doi.org/10.1111/j.1365-2672.2009.04611.x

128. Wiebe MA, Brannon JR, Steiner BD, Bamidele A, Schrimpe-Rutledge AC, Codreanu SG, Sherrod SD, McLean JA, Hadjifrangiskou M. 2022. Serine deamination is a new acid tolerance mechanism observed in uropathogenic *Escherichia coli*. mBio 13:e0296322. https://doi.org/10.1128/mbio.02963-22

129. Fallingborg J. 1999. Intraluminal pH of the human gastrointestinal tract. Dan Med Bull 46:183–196.

130. Martin RG, Rosner JL. 1995. Binding of purified multiple antibiotic-resistance repressor protein (MarR) to *mar* operator sequences. Proc Natl Acad Sci USA 92:5456–5460. https://doi.org/10.1073/pnas.92.12.5456

131. Gibson DG, Young L, Chuang R-Y, Venter JC, Hutchison CA III, Smith HO. 2009. Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat Methods 6:343–345. https://doi.org/10.1038/nmeth.1318

132. Köster J, Rahmann S. 2012. Snakemake-a scalable bioinformatics workflow engine. Bioinformatics 28:2520–2522. https://doi.org/10.1093/bioinformatics/bts480

133. Grüning B, Dale R, Sjödin A, Chapman BA, Rowe J, Tomkins-Tinch CH, Valieris R, Köster J, Bioconda Team. 2018. Bioconda: sustainable and comprehensive software distribution for the life sciences. Nat Methods 15:475–476. https://doi.org/10.1038/s41592-018-0046-7

134. Kurtzer GM, Sochat V, Bauer MW. 2017. Singularity: scientific containers for mobility of compute. PLoS One 12:e0177459. https://doi.org/10.1371/journal.pone.0177459

135. Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet j 17:10. https://doi.org/10.14806/ej.17.1.200

136. Otto C, Stadler PF, Hoffmann S. 2014. Lacking alignments? The next-generation sequencing mapper segemehl revisited. Bioinformatics 30:1837–1843. https://doi.org/10.1093/bioinformatics/btu146

137. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup. 2009. The sequence alignment/map format and SAMtools. Bioinformatics 25:2078–2079. https://doi.org/10.1093/bioinformatics/btp352

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

139. Wingett SW, Andrews S. 2018. FastQ screen: a tool for multi-genome mapping and quality control. F1000Res 7:1338. https://doi.org/10.12688/f1000research.15931.1

140. Ewels P, Magnusson M, Lundin S, Käller M. 2016. MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics 32:3047–3048. https://doi.org/10.1093/bioinformatics/btw354

141. Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-Seq data with DEseq2. Genome Biol 15:1–21. https://doi.org/10.1186/s13059-014-0550-8

142. Inc. P.T. 2015. Collaborative data science. Montreal, QC. Plotly Technologies Inc. Plotly Technologies Inc. https://plot.ly.

143. Becker AH, Oh E, Weissman JS, Kramer G, Bukau B. 2013. Selective ribosome profiling as a tool for studying the interaction of chaperones and targeting factors with nascent polypeptide chains and ribosomes. Nat Protoc 8:2212–2239. https://doi.org/10.1038/nprot.2013.133

144. Carlson M. 2019. Org.Eck12.Eg.dB: genome wide annotation for *E coli* strain K12. https://doi.org/https://doi.org/doi:10.18129/B9.bioc.org.EcK12.eg.db.

145. Gelhausen R, Müller T, Svensson SL, Alkhnbashi OS, Sharma CM, Eggenhofer F, Backofen R. 2022. RiboReport - benchmarking tools for ribosome profiling-based identification of open reading frames in bacteria. Briefings in Bioinformatics 23:1–22. https://doi.org/10.1093/bib/bbab549

146. Schmittgen TD, Livak KJ. 2008. Analyzing real-time PCR data by the comparative CT method. Nat Protoc 3:1101–1108. https://doi.org/10.1038/nprot.2008.73

147. Castanie-Cornet MP, Penfound TA, Smith D, Elliott JF, Foster JW. 1999. Control of acid resistance in *Escherichia coli*. J Bacteriol 181:3525–3535. https://doi.org/10.1128/JB.181.11.3525-3535.1999

148. Cholet F. 2015. Keras. Github. Available from: https://keras.io

149. Schneider CA, Rasband WS, Eliceiri KW. 2012. NIH image to ImageJ: 25 years of image analysis. Nat Methods 9:671–675. https://doi.org/10.1038/nmeth.2089

150. Ducret A, Quardokus EM, Brun YV. 2016. MicrobeJ, a tool for high throughput bacterial cell detection and quantitative analysis. Nat Microbiol 1:1–7. https://doi.org/10.1038/nmicrobiol.2016.77

151. Schägger H. 2006. Tricine-SDS-PAGE. Nat Protoc 1:16–22. https://doi.org/10.1038/nprot.2006.4