The transcriptional response to low temperature is weakly conserved across the *Enterobacteriaceae*

Johnson Hoang[1] and Daniel M Stoebel[2]
Department of Biology, Harvey Mudd College, Claremont, California, USA

[1] Present address: Doheny Eye Institute, Pasadena, California, USA
[2] Correspondence to: stoebel@g.hmc.edu

Abstract

Bacteria respond to changes in their external environment like temperature by changing the transcription of their genes, but we know little about how these regulatory patterns evolve. We used RNA-seq to study the transcriptional response of a shift from 37°C to 15°C in wild-type *Escherichia coli, Salmonella enterica, Citrobacter rodentium, Enterobacter cloacae, Klebsiella pneumoniae,* and *Serratia marcescens*, as well as Δ*rpoS* strains of *E. coli* and *S. enterica.* We found that these species change the transcription of between 626 and 1057 genes in response to the temperature shift, but there are only 16 differentially expressed genes in common among the six species. GO enrichment of regulated genes suggests many species-specific phenotypic responses to temperature changes, but terms involved in iron metabolism, central metabolism, and response to osmotic stress are implicated in at least half of the species. The alternative sigma factor RpoS regulates about 200 genes between 37°C and 15°C in both *E. coli* and *S. enterica*, with only 83 genes in common between the two species. Divergence in the RpoS-regulon between the two species is due to both species-specific genes in each genome as well as differences in the regulation of shared genes.

Overall, there is limited conservation of the response to low temperature generally, or the RpoS-regulated part of the response specifically, due both to some genes being species-specific, as well as the species-specific regulation of shared genes. Regulatory responses to a common stress evolve rapidly between closely related species.

Importance

We studied how different species of bacteria change the expression of their genes in response to a change in temperature. We found that the six species in this study change the level of expression of many of their genes in response to a shift from human body temperature (37°C) to a temperature that might be found out of doors (15°C). Surprisingly, there are very few genes that change expression in all six species. This was due to each species possessing many unique genes that no other species has and due to many genes that are present in the genome of each species changing expression in only one species. This study is important to the field because it illustrates that closely related species can share many genes but not use those genes in the same way in response to the same environmental change.

Introduction

Bacteria change the levels of transcription of their genes to respond to changes in their environments. These patterns of transcription, and the role of regulatory proteins and other factors in determining these patterns have been measured thoroughly in a handful of model species, like *Escherichia coli*. The degree to which our understanding of *E. coli* gene regulation can inform our understanding of related species will depend on how similar those species are. We know that the genomes of bacterial species often differ dramatically in content due to horizontal gene transfer (HGT), so it is possible that HGT may drive dramatic differences between related species in what genes are differentially expressed in response to an environmental change. Differences in transcriptome patterns between related species can also be due

51 to shared genes that have different patterns of regulation among species. In general,
52 we have little information about the relative roles of species-specific genes vs. species-
53 specific regulation of shared genes in shaping how regulons differ between relatives.
54 Changes in temperature are one challenge faced by many bacteria (1). Low
55 temperature can affect bacteria by rigidifying the lipid bilayer, by altering the folding of
56 proteins and RNAs, and by altering the assembly and activity of essential
57 macromolecular machinery like ribosomes (1). Low temperature may also serve as a
58 cue that bacteria have moved from one environment (such as a mammalian host) into
59 another environment (such as a surface at ambient temperature) (2).
60 A great deal is known about how *E. coli* responds to low temperature (1, 3). A
61 number of global regulatory factors have been implicated, including cold shock
62 proteins (4, 5), RNA degradation enzymes like RNase R and PNPase (5, 6), nucleoid
63 associated proteins like H-NS (7), global supercoiling levels (8), and the alternative
64 sigma factor RpoS (2, 9). Shifts between optimal temperature (37°C) and a lower
65 temperature have substantial impacts on the transcriptome. In a series of experiments
66 examining shifts between 37°C and 23°C (and vice versa), White-Ziegler and
67 coworkers documented that 423 genes, or about 10% of the *Escherichia coli* K-12
68 genome, are differentially expressed (DE) (2, 10). Further, they identified RpoS and H-
69 NS as two global regulators of these transcriptional patterns (2, 11).
70 How relatives of *E. coli* respond to changes in temperature is much less well
71 documented. The *Enterobacteriaceae* are a diverse group of bacteria (12) that exist in
72 many different thermal environments (13). This group includes a number of genera that
73 can be found in endothermic hosts either causing an infection or in the gastrointestinal
74 tract of healthy hosts (14). Many of these species can also be isolated from
75 ectothermic animals, plants, soil, water, plants (15–19). The presence of these
76 organisms outside of endothermic hosts suggests that the species need to cope with
77 broad changes in temperature.
78 Outside of *E. coli*, transcriptional responses to changing temperatures are much
79 less well understood across the *Enterobacteriaceae*. In general, we have little
80 understanding of how regulatory responses evolve in bacteria and at what rate these

81 regulons evolve. In addition to questions of rate, regulons could evolve because
82 different species lose different genes or gain different genes via horizontal gene
83 transfer (HGT). Regulons could also evolve because shared genes evolve different
84 regulatory controls in different species. We do not understand the relative importance
85 of these two modes of regulatory evolution. Investigating the tempo and mode of
86 regulatory evolution in the *Enterobacteriaceae* will both allow fundamental
87 understanding of bacterial evolution and allow researchers to better understand the
88 degree to which *E. coli* serves as an appropriate model system for related bacteria.
89 Single species studies of the transcriptomic response to low temperature exist
90 for several species of *Enterobacteriaceae*, including *E. coli* (2, 10), *Salmonella enterica*
91 (20), *Klebsiella pneumoniae* (21) and *Enterobacter sp.* (22). Unfortunately, it is difficult
92 to directly compare existing studies of individual species because of the large
93 differences in the culture conditions, the experimental approaches to RNA
94 quantification, and the bioinformatic and statistical analyses in these studies. To
95 directly compare the transcriptional response, the experimental and analytical
96 approaches should be the same. In addition, batch effects in both bacterial growth and
97 RNA-seq steps can be avoided by performing all of the bacterial growth, RNA
98 processing, library construction, and sequencing in parallel with a balanced
99 experimental design.
100 In this study, we studied how five species of bacteria in the *Enterobacteriaceae*
101 plus one outgroup species responded to the shift from 37°C to 15°C, as might occur
102 when bacteria leave a mammalian host and find themselves in a temperate external
103 environment. We used comparative genomic tools to identify homologous genes in
104 each species, and then used RNA-seq to examine the patterns of transcription in each
105 species across this temperature transition. In addition, we studied the role of the
106 alternative sigma factor RpoS in this transition in *E. coli* and *S. enterica.* We performed
107 these studies in parallel for all strains in the study, avoiding possible confounding
108 differences due to experimental or analytical approaches or due to batch effects.
109

Results

Most genes are either shared across all species or are species-specific

This study worked with five species of *Enterobacteriaceae*, (*E. coli, S. enterica, C. rodentium, E. cloacae, K. pneumoniae*) and one outgroup species in the *Yersiniaceae* (*S. marcescens*) (12). Their genomes contained between 3968 and 4709 protein coding sequences (Figure 1A), which were the basis of all of the analyses in this paper. We used xenoGI (23) to understand the evolutionary histories of the six genomes. xenoGI uses both synteny and sequence similarity to define groups of genes called locus families. We defined all genes in a locus family as orthologues.

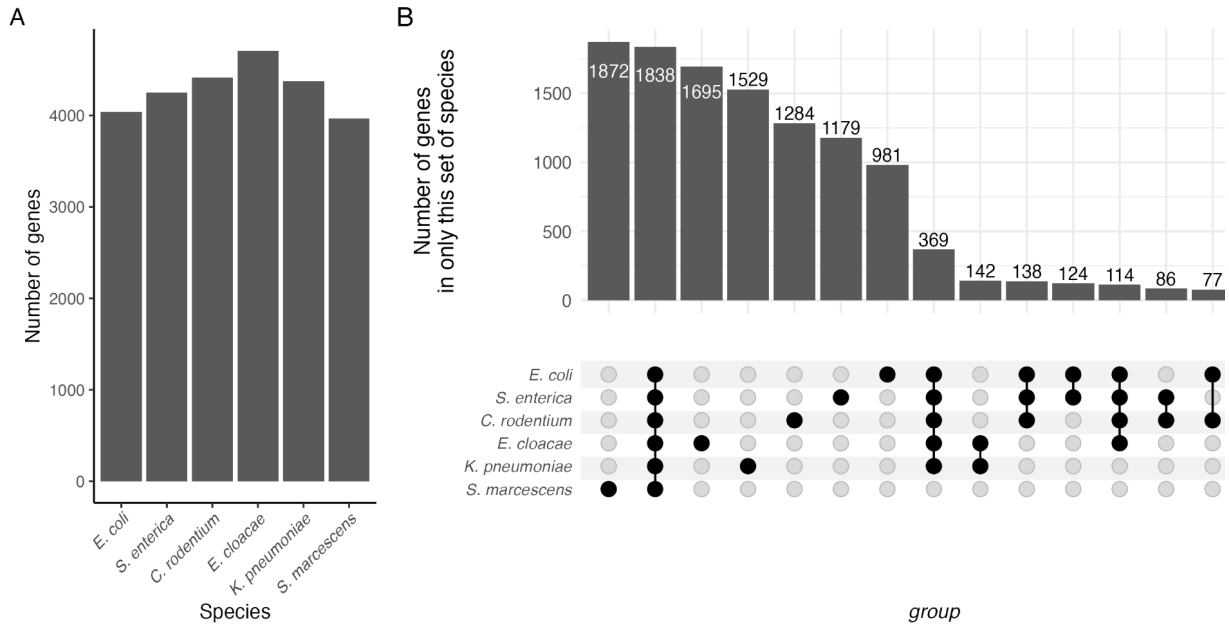


Figure 1: Patterns of shared and unique genome content. (A) The total number of protein coding genes found in the genome of each of the six species. (B) Upset plot of shared and unique genes in these species. Bar heights represent the number of genes found only in this set of species, as indicated by the dots below the bars. For example, the first bar on the left shows the number of genes that are found only in *S. marcescens,* while the second bar represents genes found in all six genomes, the core genome. Only sets with more than 50 genes are shown.

There are 1838 genes found in every genome, a set of genes denoted the core genome. The core genome accounts for 39% to 46% of all genes in a genome. The number of species-specific genes varies from 981 to 1872 (Figure 1B), which accounts for 24% to 47% of all genes in the genome. Most genes in a species are either species-specific or ancestral to the group in this analysis. There are also small subsets of genes that are found in an individual clade, such as the 124 genes found in only *E. coli* and *S. enterica.*

129 ## *All species grow much slower at 15°C than 37°C*

130 We investigated the effect of temperature on growth of the six species in this
131 study in LB medium. We measured the growth rate at 37°C, human body temperature,
132 and the lag phase and growth rate after a shift to 15°C, as might occur when moving
133 from a host into the external environment. All species in this study grow much faster at
134 37°C than at 15°C (Table 1). All species had doubling times under 30 minutes when
135 growing at 37°C, but doubling times were 5 to 14 times as long at 15°C. Cultures all
136 underwent a lag phase of one to two hours before resuming growth. For subsequent
137 studies on gene expression in response to low temperatures, we worked with cultures
138 that had been at 15°C for three hours, as all strains had resumed growth after the
139 temperature shift at this point. After three hours at 15°C, strains should be past the
140 short-term response to cold (involving changes in supercoiling) and into the longer-
141 term response, involving transcriptional changes driven by changing activity of
142 transcription factors (8).
143
144 **Table 1:** Growth parameters at 15°C and 37°C in LB[†]

Strain	Doubling time at 37°C (mins)	Lag time after shift to 15 °C (mins)	Doubling time at 15°C (mins)
E. coli wt	28.5 ± 0.3	63.7 ± 12.2	328.0 ± 15.2
E. coli Δ*rpoS*	25.5 ± 0.2	98.1 ± 3.6	303.5 ± 6.4
S. enterica wt	29.6 ± 1.7	96.2 ± 6.7	175.5 ± 3.8
S. enterica Δ*rpoS*	27.5 ± 1.2	97.8 ± 10.1	188.8 ± 7.4
C. rodentium wt	28.1 ± 0.6	104.0 ± 23.1	342.2 ± 30.9
E. cloacae wt	28.0 ± 3.6	90.3 ± 29.7	394.2 ± 44.6
K. pneumoniae wt	26.8 ± 0.7	119.4 ± 7.7	268.9 ± 19.8
S. marcescens wt	29.9 ± 1.3	123.8 ± 9.8	160.9 ± 7.1

145 [†] Values are the mean ± standard error of the mean. N = 3 for all measurements.

146 *Limited conservation of patterns of differential expression in*
147 *response to temperature shift*

148 To examine how these species alter their transcriptome in response to
149 temperature change, we performed a RNA-seq experiment on samples growing
150 exponentially at 37°C (immediately before a shift to 15°C), and then after three hours
151 growth at 15°C. In the three hours between growth at 37°C and 15°C, patterns of gene
152 expression changed substantially. 626 to 1057 genes were significantly differentially
153 expressed (defined as an FDR-adjusted p-value < 0.01 and an absolute $\log_2$-fold
154 change > 2) between the two temperatures across the different species (Figure 2A).
155 This corresponds to between 14 and 26% of all protein coding genes in the genome
156 being DE. We found little conservation of the set of DE genes in each species (Figure
157 2B,C). For example, each species has between 274 and 505 genes that are uniquely
158 DE in that species alone. In contrast, there are only 16 genes that are DE in every
159 species studied. The transcriptional response to low temperature varies strongly
160 among these species.
161

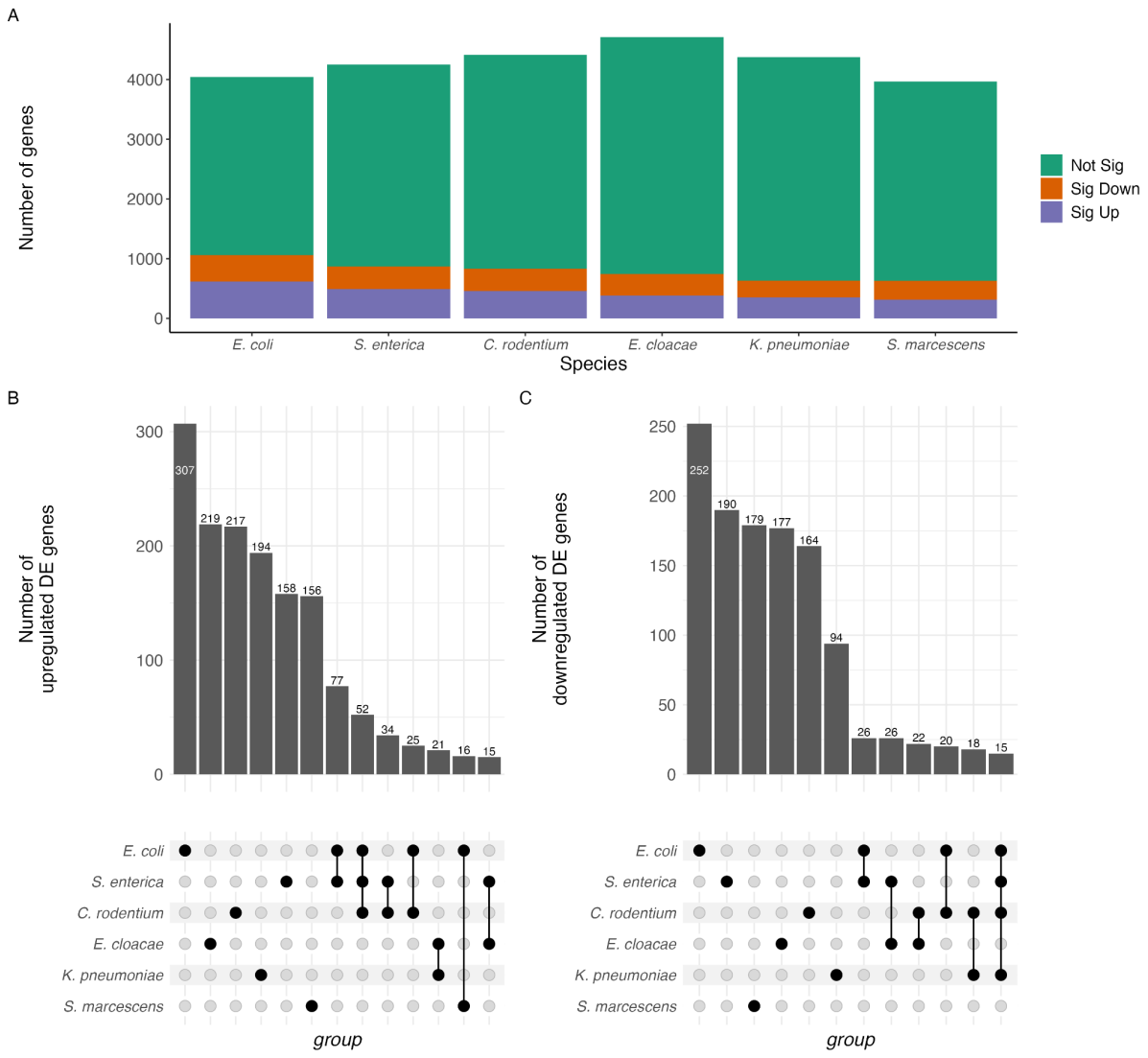


Figure 2: Patterns of differential expression in response to temperature shift of all genes in the genome. (A) The number of genes significantly DE genes (FDR-adjusted $p < 0.01$, absolute $\log_2$-fold change > 2) as a proportion of all genes in the genome for each species. Upregulated genes are expressed more highly at 15°C than 37°C. (B) Upset plot of shared and unique patterns of DE genes in these species. Bar heights represent the number of genes, while the dots below represent the species involved in that group.

171 One factor that could contribute to the limited overlap in DE genes among
172 species is considerable variation in genome content, with each genome containing
173 between 981 to 1872 species specific genes (Figure 1B). Each species-specific gene
174 that is DE is an automatic contribution to the set of unique DE genes of that species.
175 However, these species-specific genes are actually less likely to be DE between
176 temperatures than expected if all genes had equal probability of being DE (Figure 3A;
177 chi-square test, $p < 5 \times 10^{-4}$ for all six species.) Each species' set of unique genes does
178 not disproportionately contribute to the divergence in patterns of transcription.
179 What about the conserved genes? Genes in the core genome are particularly
180 likely to be DE (Figure 3B; chi-square test, $p < 2 \times 10^{-3}$ for all six species). However,
181 while core genes are particularly likely to be DE, which genes are DE in each genome is
182 not well-conserved (Figure 4). For example, of the 561 core genes that are DE in *E. coli*,
183 231 of these genes are not DE in any other species, even though all species in that
184 study have those genes. There is not a total lack of conservation of expression pattern,
185 however. For example, there are 60 core genes that are DE in both *E. coli* and *S.*
186 *enterica* and in no other species. Species-specific expression of core genes plays a
187 major part in giving each species a distinctive transcriptional response to temperature
188 shift.

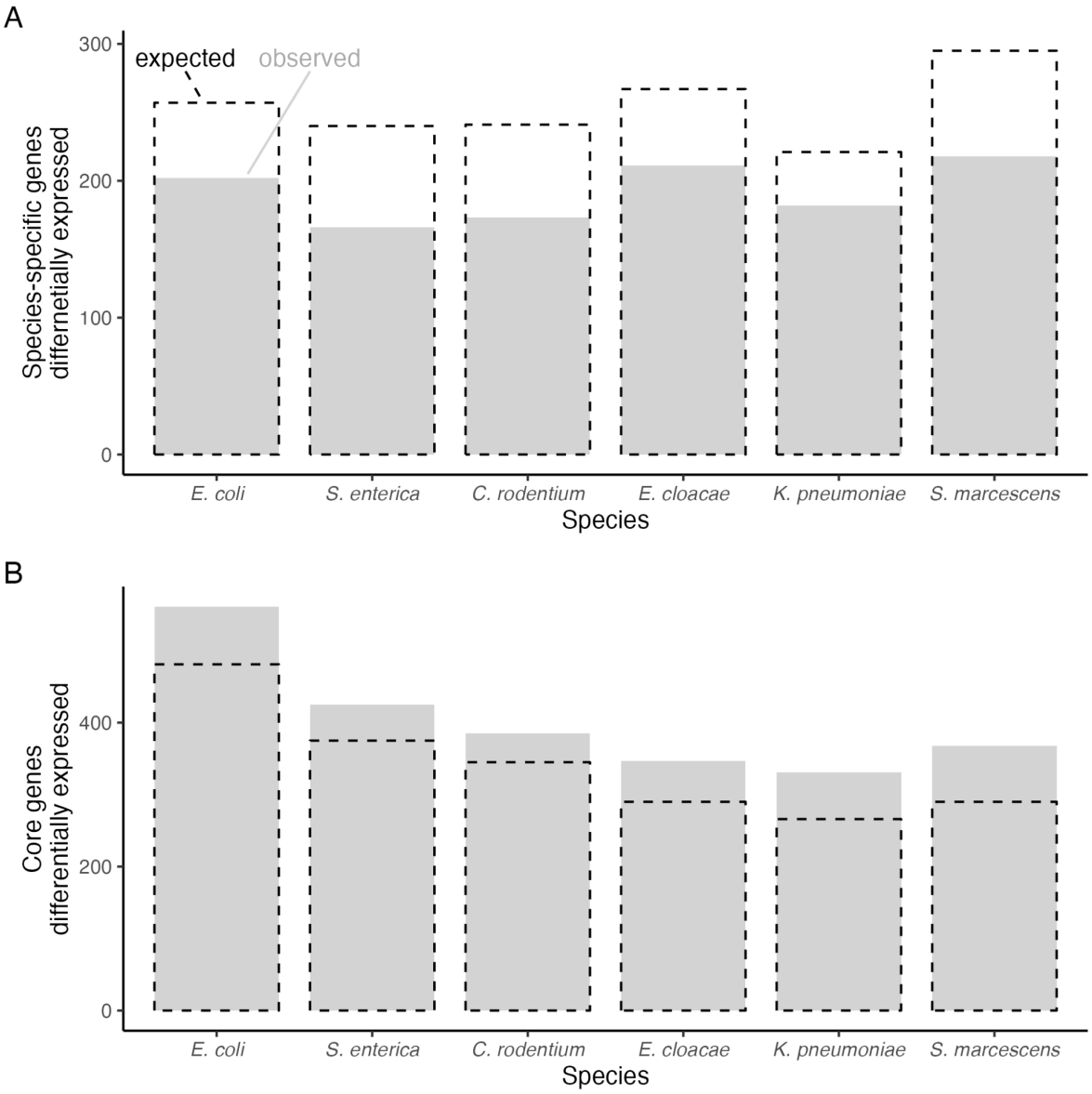


Figure 3: Observed vs. expected number of DE genes. The number of DE genes among the (A) species-specific genes and (B) the core genes. Observed counts are shown as gray bars, while the expected counts are shown with black dashed lines. Expected values were calculated as the number of genes in that category multiplied by the proportion of all genes that were DE. In every case, the observed count is significantly different from the expected count (chi-square test, $p < 5 \times 10^{-4}$ for each comparison in (A), $p < 2 \times 10^{-3}$ for each comparison in (B)).

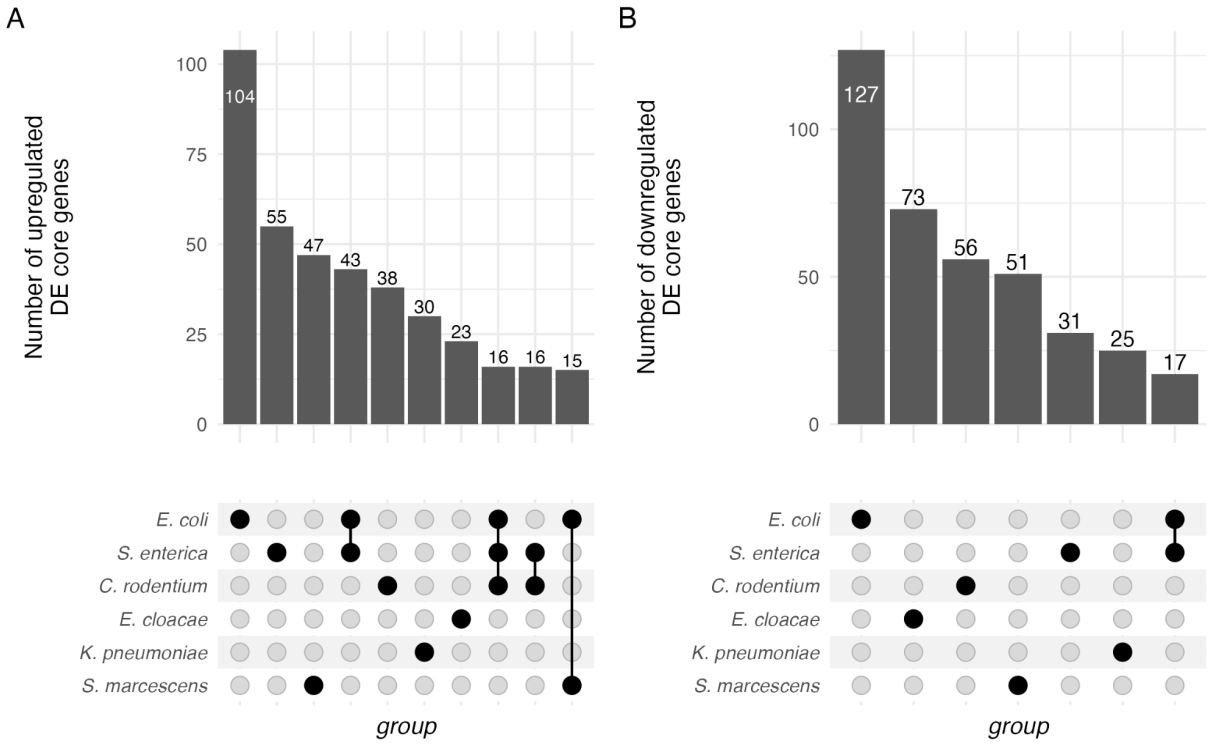


Figure 4: Patterns of differential expression of core genes in response to temperature shift. (A) The number of significantly DE core genes that are (A) upregulated or (B) downregulated in response to temperature shift. Upregulated genes are expressed more highly at 15°C than 37°C. Bar heights represent the number of genes, while the dots below represent the species in that group.

We extended our analysis to the other internal nodes on the phylogeny and examined not just the core genes but all those genes that show some level of sharing between species. At each node on the phylogeny, we considered those genes that have an orthologue in each descendent of that node, but in none of the other species, and are DE in at least one of the species. In every case, we see that if a gene is DE in at least one genome, the most likely outcome is that it is DE in only that one genome (Figure 5) with decreasing frequencies as the number of species increases. Thus, even when genes are shared between multiple species, patterns of differential expression are primarily not conserved.

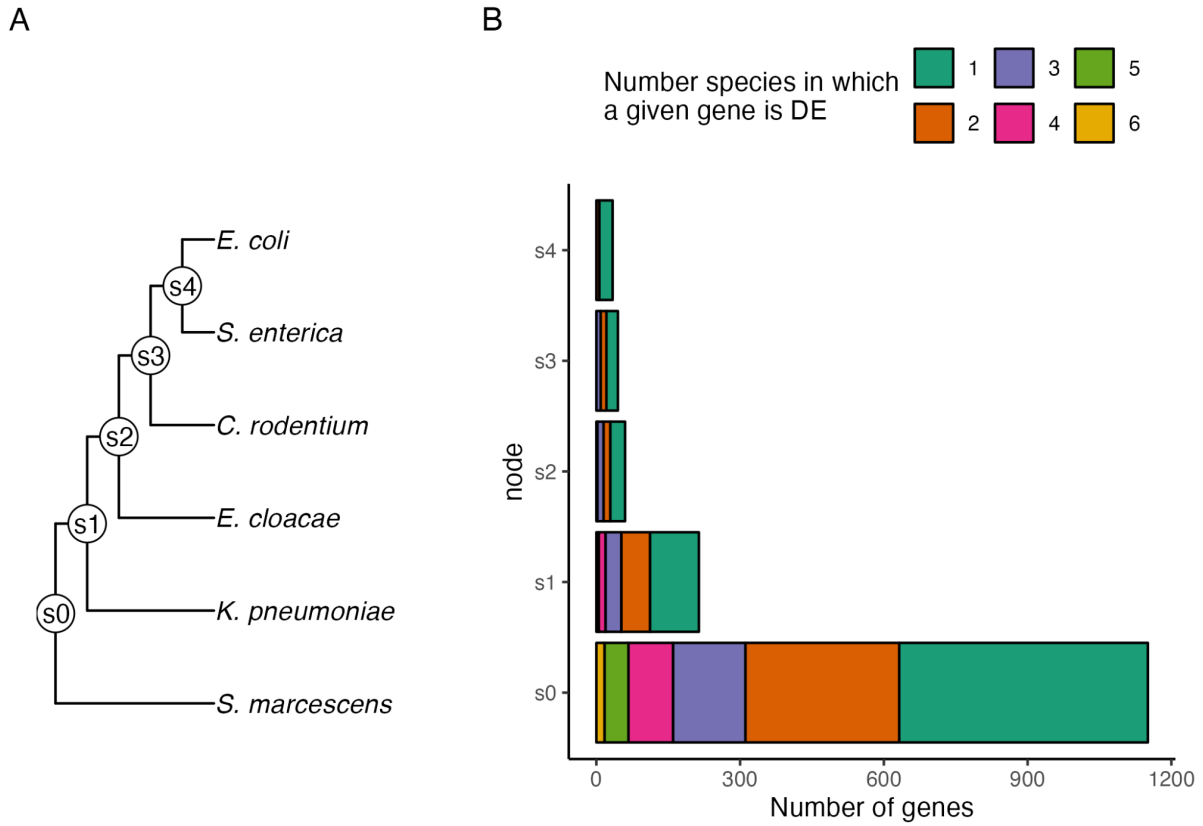


Figure 5: Distribution of differentially expressed genes among members of a clade. (A) Phylogeny of the species in this study, with nodes labeled. (B) For each gene present in all species descended from a given node, total bar length represents the number of genes that are DE in at least one species. Length of colored segments of the bar represent the number of genes for which that many species are DE.

Another way to visualize the overall similarity of the transcriptional response or lack thereof is by principal component analysis (PCA). We calculated the PCA using the 1838 core genes. The first two principal components explain 40% of the variance in normalized gene expression levels. PC2, which accounts for 15% of the variance in gene expression, separates 15°C samples from 37°C samples, as all 37°C samples have larger values of PC2 than the 15°C samples from the same species (Figure 6). (Other principal components not visualized here do not clearly separate samples based on temperature.) Thus, there is some conserved difference in the expression of core genes between 15°C and 37°, but it is a small proportion of all variation in gene expression among the species.

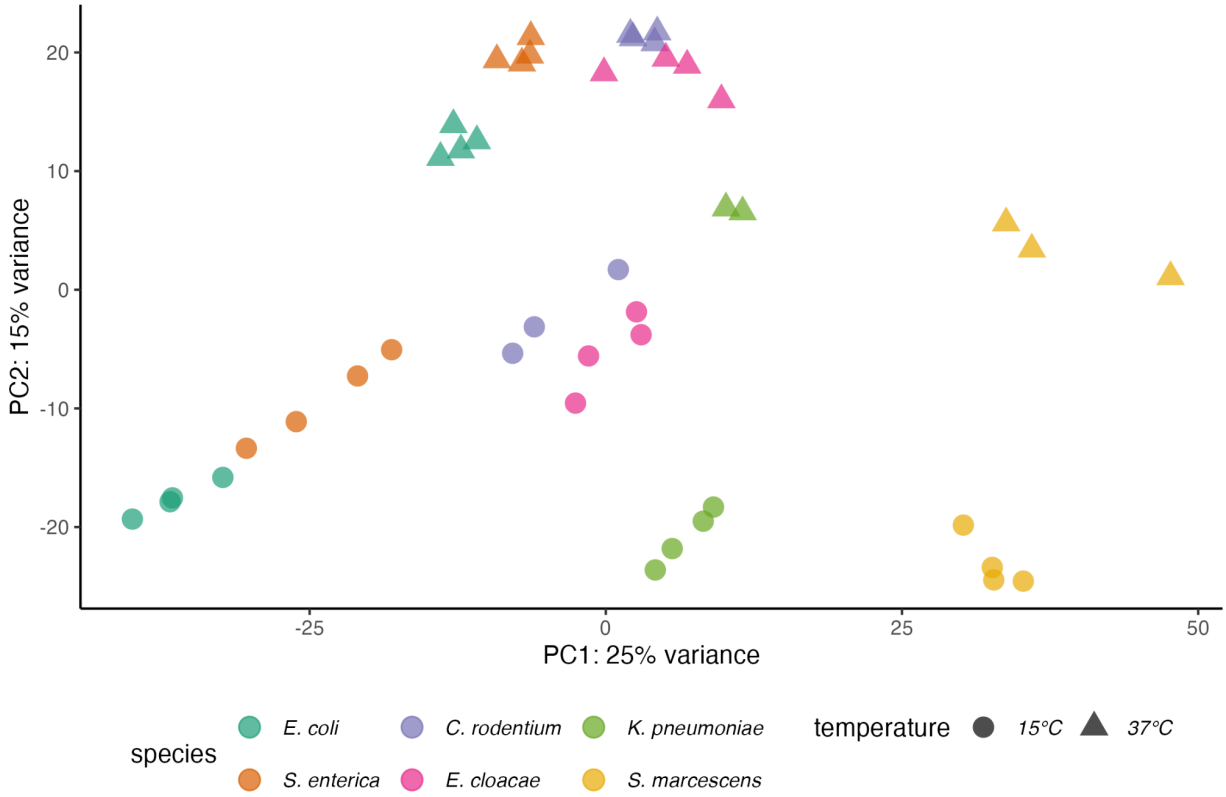


Figure 6 Principal component analysis of expression of the 1838 core genes.

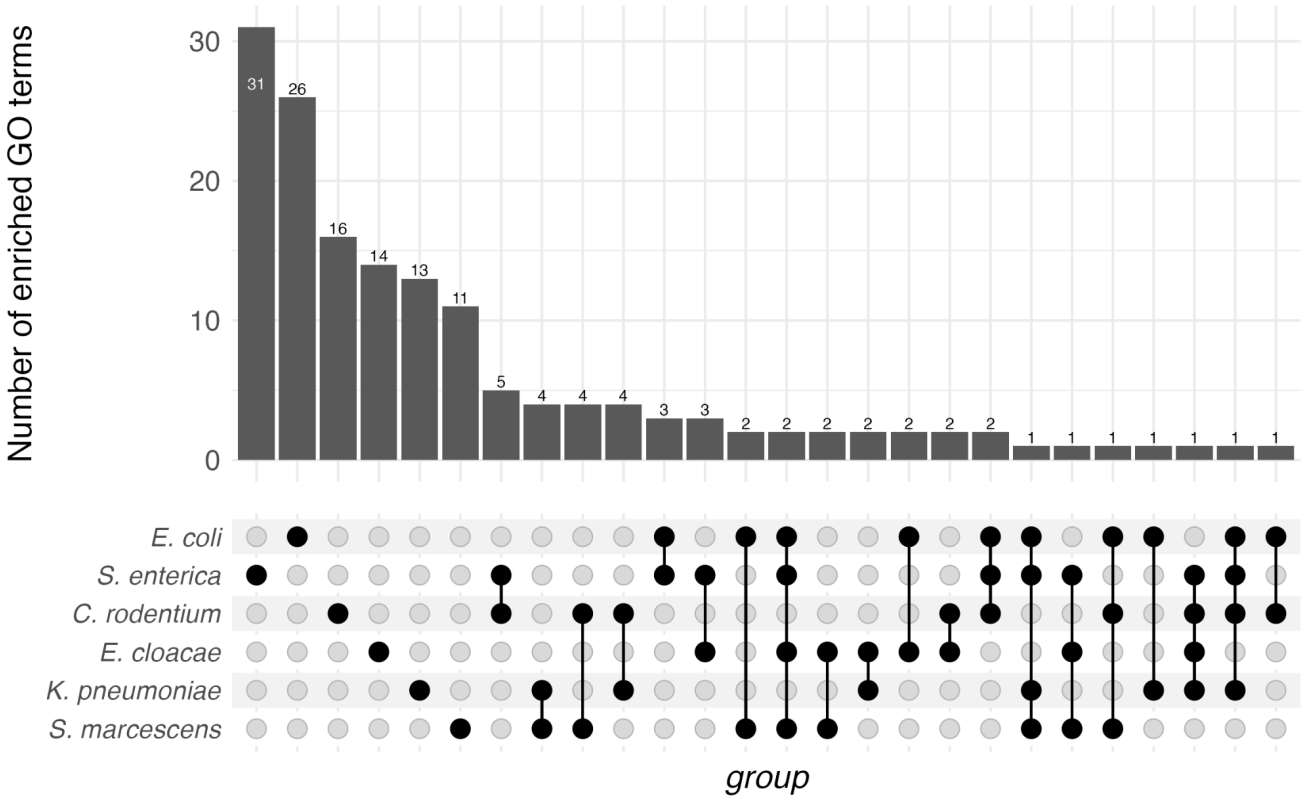
Functions of genes that are DE between warm and cold environments

To better understand the functional consequences of the lack of conservation of DE genes, we used Gene Ontology (GO) enrichment to understand what groups of genes were differentially expressed. We used eggNOG-mapper (24) to assign GO terms to each gene in each genome. (This was done even for genomes like *E. coli* K-12 that had pre-existing GO annotations so that all species had comparable annotations in our analysis.) As a check on the quality of the annotation, we examined the GO terms assigned to each orthologue of the core genome. We found that 1668 of 1838

246 (91%) of core genes had identical GO annotations in all six genomes, suggesting that
247 the annotation worked consistently for most genes. Most of the remaining 118 core
248 genes had identical annotations for 5 of 6 genomes, suggesting that again the
249 annotation worked well but was not perfectly consistent with what we defined as the
250 core genome.

251 We then performed GO enrichment analysis using topGO (25) on the RNA-seq
252 data from each species. Consistent with the lack of conservation of differentially
253 expressed genes, there was overall little conservation of enriched GO terms across the
254 genomes (Figure 7). Of the 151 GO terms that were enriched in at least one genome,
255 only 10 were enriched in at least three genomes (Table S2). 107 of the 151 terms were
256 enriched in only a single genome.

257



258

259 **Figure 7. Shared patterns of GO enrichment.** Bar heights represent the number of GO terms
260 enriched, while the dots below represent the species involved in that comparison.

Iron metabolism genes are affected by temperature in multiple species

One group of GO terms that are enriched in multiple genomes are those associated with iron acquisition and use. Enterobactin is a siderophore that is synthesized by many species of Enterobacteraceae (26). This molecule is synthesized inside the cell and secreted into the environment, where it can bind to ferric iron. The iron-enterobactin complex is then actively transported into the cell where the enterobactin is cleaved, freeing the iron (26). Iron is scarce inside of mammalian hosts and genes that allow for iron scavenging are important virulence factors in many of the species in this study (27). The enterobactin biosynthesis term (GO:0009239) is enriched in *E. coli, S. enterica, C. rodentium,* and *K. pneumoniae*, with the associated genes downregulated at 15°C in these species (Figure 8A). Genes for enterobactin transport (GO:0042930) are enriched in *S. enterica, C. rodentium, E. cloacae* and *K. pneumoniae*, with the associated genes also downregulated (Figure 8B). The *E. coli* results are consistent with previous findings that 37°C increases the expression of iron utilization genes in that organism relative to lower temperatures (10). The overall picture is that *S. enterica, C. rodentium,* and *K. pneumoniae*, like *E. coli,* do less to acquire iron at 15°C than at 37°C. That said, this is not a completely conserved regulatory pattern, as enterobactin biosynthesis levels are largely unchanged in response to temperature in *E. cloacae* and *S. marcescens.*

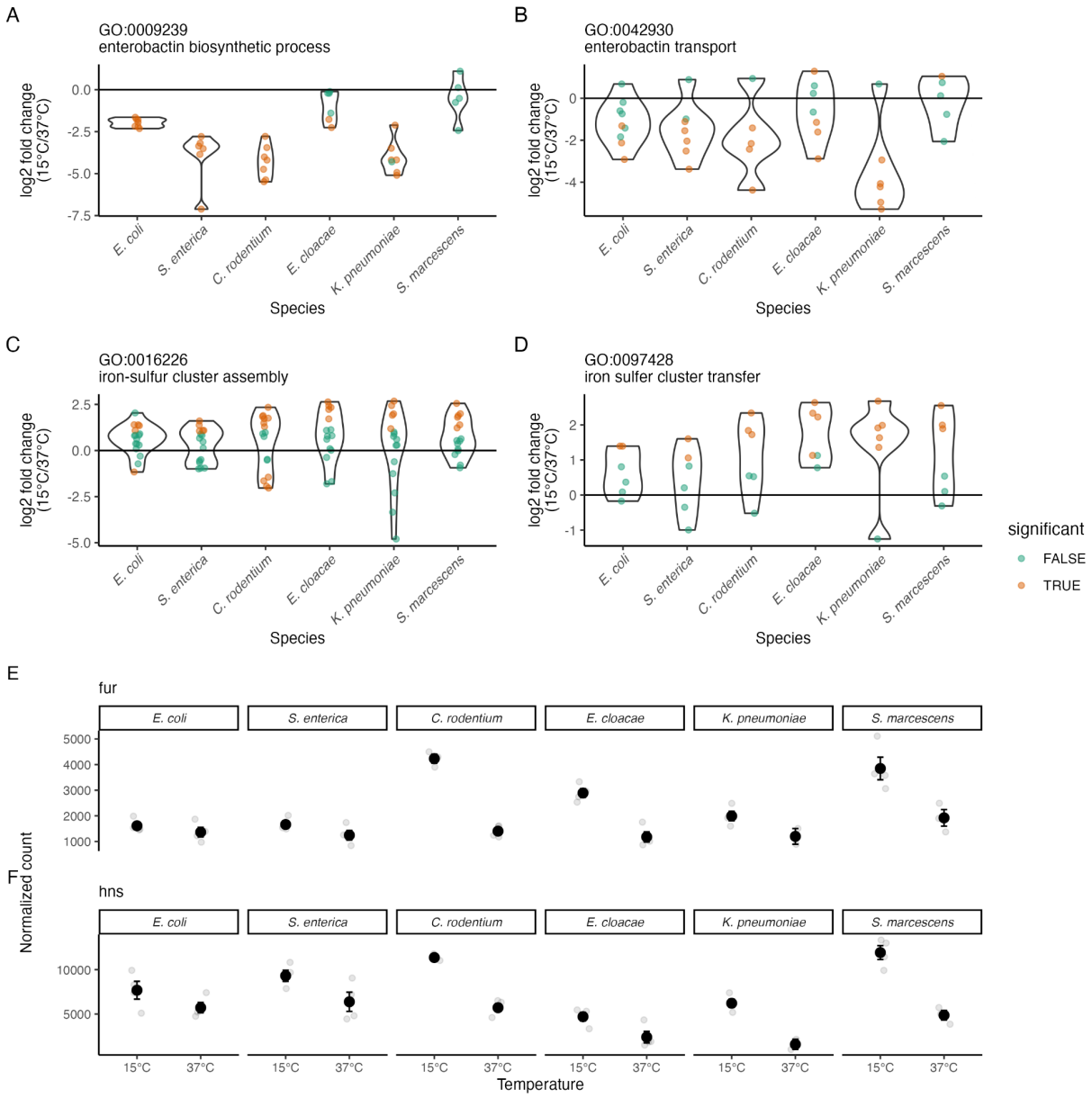


Figure 8: Genes involved in iron metabolism change expression in the cold. Log-2 fold change of gene expression between 15°C and 37°C for genes annotated as involved in several aspects of iron metabolism (A) enterobactin biosynthesis, (B) enterobactin transport, (C) iron-sulfur cluster assembly, and (D) iron-sulfur cluster transfer. Orange dots represent a single DE gene, while green dots are a gene that is not DE. (E,F) The expression of genes coding for two global regulators previously identified as associated with these phenotypes, *fur* and *hns*. Black dots represent the mean, error bars are the SEM, and gray dots are individual measurements.

290 One notable use for iron in enteric bacteria is the production of iron-sulfur
291 clusters. Genes associated with iron-sulfur cluster assembly (GO:0016226) that are
292 significantly DE are all upregulated in 15°C relative to 37°C in most species (Figure 8C).
293 However, there are multiple genes expressed at lower levels in *C. rodentium* at 15°C
294 and one such in *E. coli.* Genes associated with the transfer of mature iron-sulfur
295 clusters into proteins (GO:0097428) that are significantly DE are primarily upregulated
296 (Figure 8D), although the number of DE genes varies among the species. One notable
297 contrast between iron-sulfur cluster synthesis and transfer and enterobactin synthesis
298 and transport is that *E. cloacae* and *S. marcescens* do not have marked responses to
299 temperature in their enterobactin genes but do have similar transcriptional patterns of
300 iron-sulfur cluster metabolism genes as other species.
301 One key regulator of genes associated with iron metabolism in many species is
302 the protein Fur. Fur is primarily a transcriptional repressor (28), but it can serve as a
303 positive regulator in cases where it disrupts repression by another protein, such as its
304 action as an anti-silencer of H-NS at the *ftnA* promoter (29). Fur mRNA levels are
305 significantly higher at 15°C than 37°C only for *C. rodentium* and *E. cloacae* (Figure 8E).
306 *E. coli* and *S. enterica* showed clear repression of enterobactin biosynthesis with non-
307 significant (and very modest) changes in *fur* levels, suggesting that temperature
308 regulation of Fur cannot easily explain the observed patterns in the transcription of
309 enterobactin biosynthesis genes in response to temperature. Similarly, *hns* mRNA
310 levels are only significantly higher at 15°C for *K. pneumoniae* and *S. marcescens*
311 (Figure 8F), suggesting that changes in H-NS abundance do not account for changes
312 in enterobactin biosynthesis genes in response to temperature. While the direction of
313 change in *fur* and *hns* levels are consistent with the changes in enterobactin levels, the
314 modest and non-significant changes in mRNA abundance suggest that other
315 regulatory factors beyond Fur and H-NS regulate this change in iron utilization in some
316 species.

317 ## Other enriched GO terms

318 Genes involved in the TCA cycle and with glycine metabolism were significantly
319 enriched in four genomes each (Table S2). (TCA cycle genes are also annotated as
320 involved in citrate metabolism, GO:0006101). These genes were largely upregulated in
321 response to cold (Figure 9A, B). LB medium contains no fermentable sugars, so the
322 TCA cycle is presumably being fed by metabolizable amino acids (30). This result is
323 consistent with a previous proteomic study that identified several TCA cycle enzymes
324 as upregulated following downshift from 37°C to 20°C (31). Other GO terms enriched
325 in multiple species include the response to osmotic stress (Figure 9C) and single-
326 species biofilm formation (Figure 9D). The overlap between osmotic stress resistance
327 and low temperature response has been previously noted (1). Low temperature has
328 previously been shown to increase biofilm production and the expression of biofilm
329 associated genes in *E. coli* (2), and these results suggest that this response may be
330 conserved in some other members of this group of organisms.
331

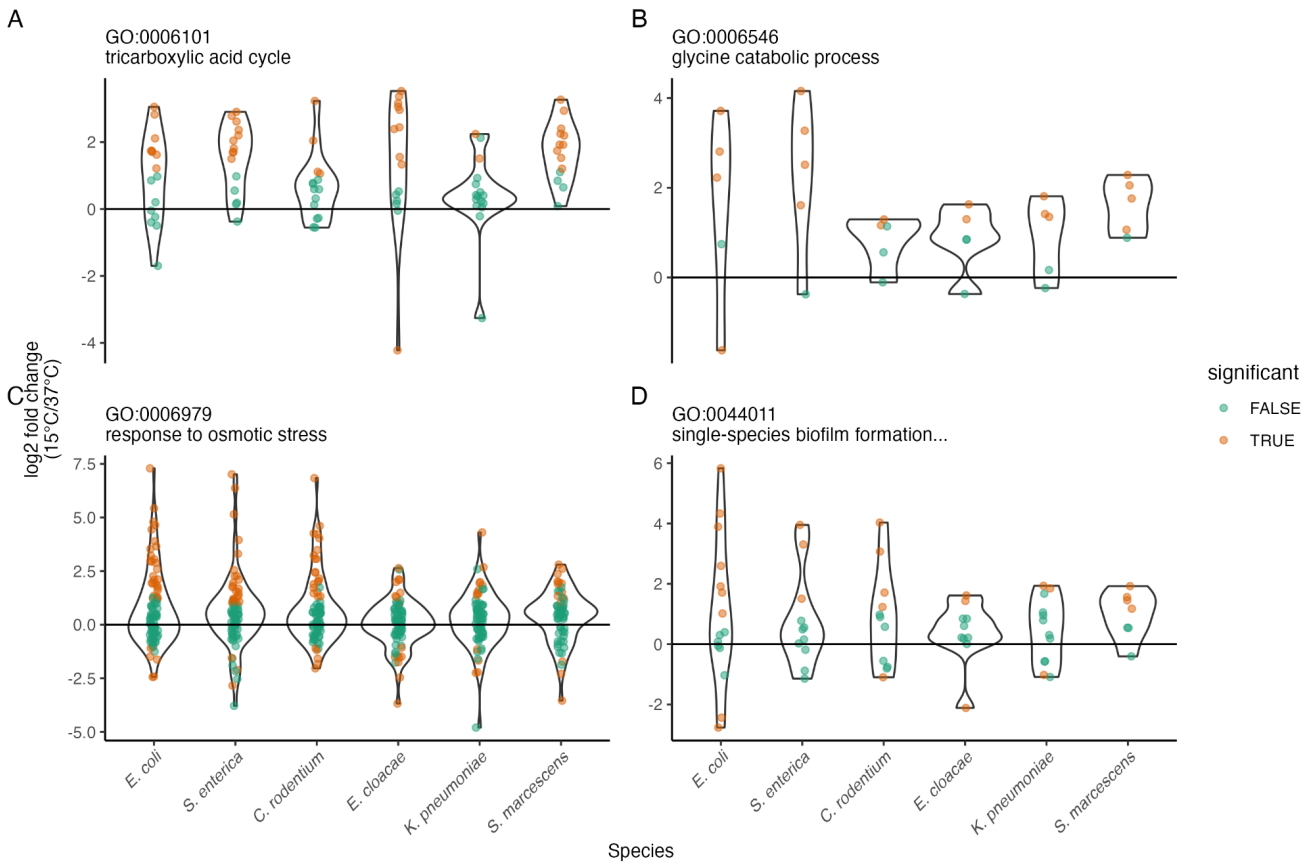


Figure 9: Groups of genes that change in multiple species. Log-2 fold change of gene expression between 15°C and 37°C for genes annotated as involved in (A) the tricarboxylic acid cycle, (B) glycine catabolism, (C) responding to osmotic stress, and (D) single-species biofilm formation on an inanimate surface. Orange dots represent a single DE gene, while green dots are a gene that is not DE.

Species-specific gene expression changes in response to temperature

While there are a few GO terms that are enriched across multiple species and suggest a common phenotypic response to low temperatures, the vast majority of enriched GO terms are species-specific. Here we highlight several species-specific responses. Similar plots of genes associated with a GO term enriched in at least one species are available in the Zenodo archive for interested readers.

Ribosomes and electron coupled proton transport are enriched only in *E. coli*

Several GO terms that were enriched only in *E. coli* are involved with protein production and metabolism. Genes annotated as involved in ribosome assembly (GO:0000027 and GO:0000028, for ribosomal large and small subunit assembly, respectively) are enriched in the DE genes of *E. coli* but not in other species (Figure 10A). (This GO term includes the structural components of the ribosome as well as factors that modulate ribosome assembly.) The mechanistic basis for this difference in ribosomal protein mRNAs is hard to predict because the levels of regulation of ribosomal proteins are so complicated. Ribosomal protein operons are directly regulated by ppGpp + dksA in *E. coli* (32). Ribosomal protein transcripts have very short half-lifes when translationally repressed (32), and both changes in RNA structure and a set of regulatory proteins modulate translation rate in response to lower temperature (1).

Genes coding for proteins that produce the proton motive force via electron transport (GO:0015990) are DE in *E. coli* more than other species (Figure 10B) . These genes, *cyoABCD*, and *nuoLMN* encode for the cytochrome bo3 quinol:oxygen oxidoreductase and NADH:ubiquinone oxidoreductase I respectively. The EcoCyc database (33) lists at least 12 regulatory proteins that may regulate the transcription of one or more of these genes in *E. coli*, so substantial work may be required to determine the basis of this species-specific downregulation.

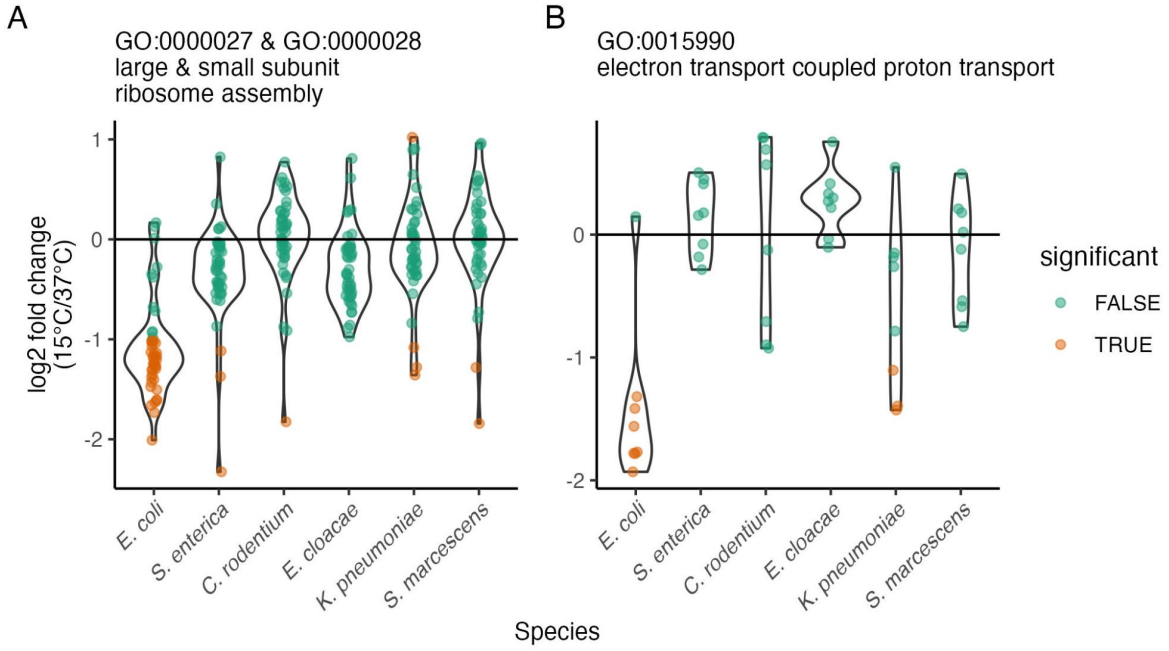


Figure 10: Responses to low temperature unique to *E. coli*. Log-2 fold change of gene expression between 15°C and 37°C for genes annotated as involved in (A) production of the ribosome, and (B) electron transport coupled proton transport. Orange dots represent a single DE gene, while green dots are a gene that is not DE.

Colanic acid biosynthesis is enriched only in *C. rodentium*

Colanic acid is the capsular polysaccharide in the *E. coli* and some relatives (34–36). In our data the genes annotated with the GO term for colanic acid biosynthesis ("GO:0009242") were enriched in *C. rodentium* but not in other species. Because some genes in the *wca* cluster, which are biosynthetic enzymes for colanic acid (35), were not annotated with this GO term, we created a custom gene set that included all genes annotated with GO:0009242 plus all genes from the *wca* cluster that were missing that annotation. (The strains of *K. pneumoniae* and *S. marcescens* we worked with do not have orthologues to the *wca* genes found in the other four species.) 70% genes in the combined set are significantly upregulated in *C. rodentium*, while many fewer (if any) genes are upregulated in each of the other species (Fig 11A).

In *E. coli*, the Rcs phosphorelay system is an important regulator of colanic acid biosynthesis (37). The response regulator RcsB forms a heterodimer with the RcsA

381 protein, and these two proteins are required for maximal capsule expression (38). The
382 gene *rcsA* has an expression pattern that closely matches the pattern of its target
383 genes, being significantly upregulated 10-fold in *C. rodentium,* while being upregulated
384 at most 2-fold in the other five species (Fig 11B). In contrast, RcsB shows a similar
385 pattern of expression across five of the six species, rather than being upregulated only
386 in *C. rodentium* (Fig 11C). The evolution of species-specific regulation of RcsA in *C.*
387 *rodentium* is a promising candidate for the evolutionary change that brought about the
388 species-specific pattern of a large group of genes.
389

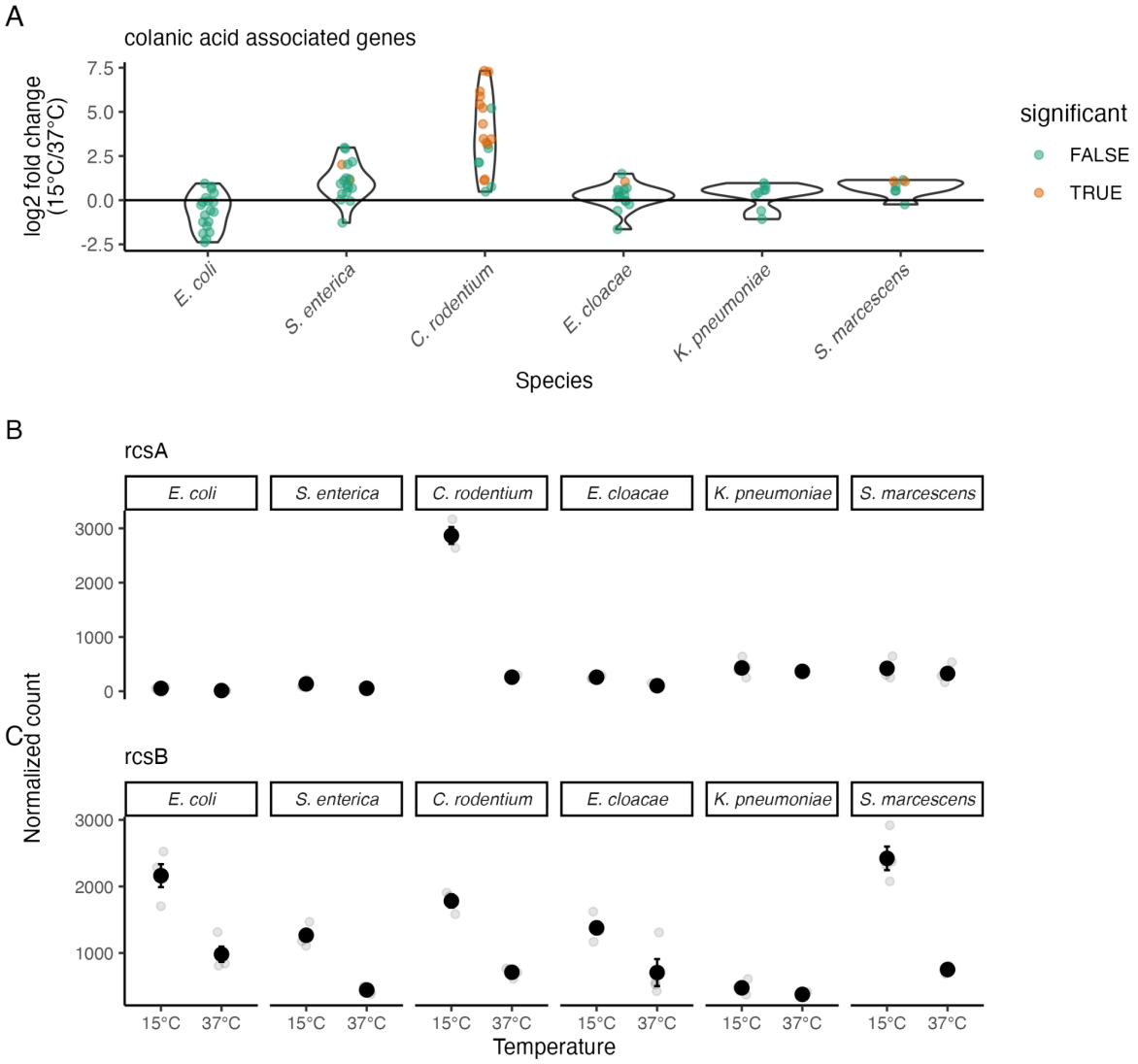


Figure 11: Upregulation of colanic acid related genes only in *C. rodentium.* (A) Log 2-fold change of gene expression between 15°C and 37°C for genes associated with production and use of colanic acid. Orange dots represent a single DE gene, while green dots are a gene that is not DE. (B, C) Expression of two known regulators of this process, *rcsA* and *rcsB*. Black dots represent the mean, error bars are the SEM, and gray dots are individual measurements.

RpoS regulation of the response to temperature shift

The alternative sigma factor RpoS is known to be activated at low temperature in *E. coli* where it regulates the expression of several hundred genes (39). We compared the role of RpoS in *E. coli* and its close relative *S. enterica* during growth at 15°C. We measured RpoS protein levels using western blotting and found that RpoS has a similar pattern of expression in both species, with RpoS accumulating during exponential growth at 15°C to a level like stationary phase at 37°C, but not during exponential growth at 37°C (Figure S1). As expected, the Δ*rpoS* mutants did not produce measurable RpoS.

RNA-seq experiments with Δ*rpoS* strains of *E. coli* and *S. enterica* allowed us to define the RpoS regulon at 15°C*.* During exponential growth at 15°C, 232 genes are DE between wt and Δ*rpoS* in *E. coli.* This represents 22% of the number of genes that are DE between 37°C and 15°C in the wild-type strain. This implies that while RpoS regulates many genes in the cold, it is only a fraction of the genes that are cold-regulated overall. Reassuringly, of the 232 DE genes in *E. coli* in this study, 189 (81%) were previously identified as RpoS-regulated at 15°C by Adams *et al.* (39). In *S. enterica,* 189 genes are DE between wt and Δ*rpoS* strain, 80% of the number of genes that are RpoS-regulated in *E. coli.*

83 genes are DE in both *E. coli* and *S. enterica* (Figure 12). This includes genes previously reported to be RpoS-dependent (40–42) such as *dps, katE, mlrA, osmC, osmY, otsA, otsB, sodC,* and *wrbA.* Consequently, GO terms involved in response to osmotic stress or response to hydrogen peroxide are enriched in both species. While there are 83 genes RpoS-regulated in both species, the majority of the RpoS-regulon in *E. coli* or *S. enterica* is not RpoS-regulated in the other species (Figure 12). For example, of the 232 RpoS-regulated genes in *E. coli,* 149 of those genes (64%) are not RpoS-regulated in *S. enterica.* Those 149 genes are composed of 50 genes that are not present in *S. enterica,* and 99 genes that have an ortholog in *S. enterica*, but that

420 orthlog is not RpoS-regulated in *S. enterica.* Like our earlier analysis with the entire
421 cold-response regulon, divergence of the RpoS-regulon between *E. coli* and *S.*
422 *enterica* is due to a mixture of both different regulation of genes shared between the
423 two species and the presence of species-specific genes.

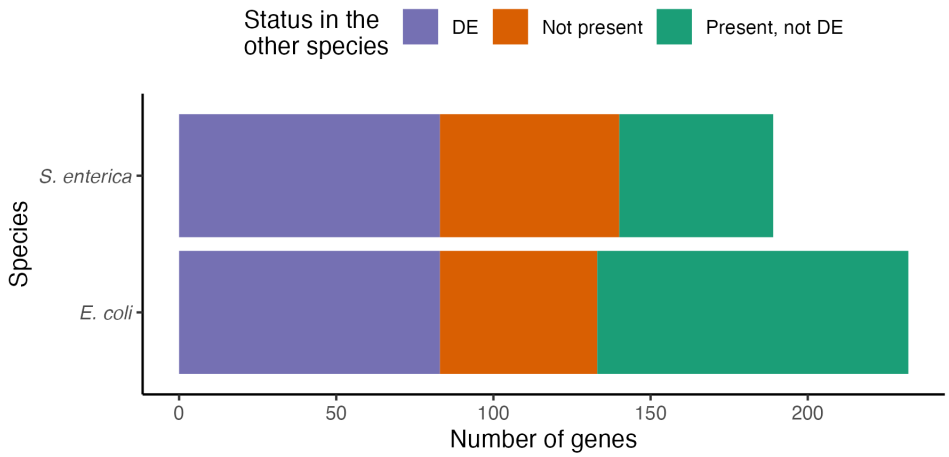


Figure 12: Partial overlap in the RpoS regulon of *E. coli* and *S. enterica*. Bar lengths are the number of RpoS-regulated genes in that species, with colors denoting the status of those genes in the other species. For example, the green part of the *E. coli* bar represents the 99 genes that are RpoS-regulated in *E. coli* that are also present in *S. enterica,* but are not RpoS-regulated in *S. enterica.*

424

425 Genes that are RpoS-regulated only in *E. coli* include *astABCDE,* responsible for
426 arginine catabolism (GO:0006527), *dppABCDF*, involved in di-peptide transport, *gabD*,
427 *astC*, and *ldcC*, involved in lysine metabolism (GO:0006553). These genes have
428 orthologs in *S. enterica* but are not significantly DE in that organism, pointing to a role
429 for RpoS in amino-acid catabolism in *E. coli* that is not present in *S. enterica.* RpoS
430 also has an *E. coli*-specific role in regulation of genes *gadE, gadW, gadX, hdeA, hdeB,*
431 and *slp,* which are part of the acid fitness island not present in related species like *S.*
432 *enterica* (43).

433 Genes that are RpoS-regulated only in *S. enterica* include *glgB, glgC, glgP,* and
434 *glgX*, responsible for the biosynthesis and catabolism of glycogen. (These genes are
435 expressed more highly at 15ºC than 37°C in both *E. coli* and *S. enterica*, but are only
436 RpoS-regulated in *S. enterica*.) There are also genes that are RpoS-dependent and
437 present only in *S. enterica*, like *invA*, *invE*, *prgI, sipD*, and *sprB*. These genes are all
438 located on the Salmonella Pathogenicity Island 1 and are not present in *E. coli* (44).
439 Much like the broader response to low temperature, the RpoS-regulon at 15°C is
440 mostly not conserved between *E. coli* and *Salmonella* due a mix of species-specific
441 genes and different regulation of conserved genes.

442 Discussion

443 All of the species studied respond to a shift in growth temperature from 37°C to
444 15°C by changing the expression of a large proportion of their genes; the number
445 ranges from 626 to 1057 genes, or 14 to 26% of all genes in the genome. For *E. coli*,
446 this is more than twice the number of genes than were previously reported in
447 microarray studies of a shift from 37°C to 23°C (2, 10). The difference can probably be
448 attributed to the combination of the more extreme temperature shift in this study, and
449 the difference in technologies used to measure mRNA abundance.
450 A major finding of this study is the lack of conservation of the transcriptional
451 response to low temperature across the *Enterobacteriaceae*. All the species grow
452 markedly slower at 15°C than at 37°C and all changed expression of hundreds of
453 genes in response to the temperature shift, the identity of these genes is mostly not
454 conserved among species. The lack of conservation of which genes are DE is due to
455 both the fact that each species has unique genome content, and that even when genes
456 are shared, they may be DE in only a subset of species. While species-specific genes
457 do contribute to the diversity of differentially expressed genes in this group, these
458 species-specific genes are actually less likely to be DE than is the core genome.
459 Species respond to a common stress by changing the expression of a different set of
460 their shared genes.

461 This lack of conservation of which genes are differentially expressed extends to
462 a lack of conservation of phenotypes as predicted by enriched GO terms. Most of the
463 enriched GO terms for a species are not shared with other species, which implies that
464 bacteria in this group of species may have different phenotypic responses to the same
465 stress. For example, *E. coli* responds to the lower temperature by down-regulating the
466 mRNAs of ribosomal proteins and proteins involved in electron transport to generate
467 the proton motive force. Surprisingly, this effect is unique to *E. coli*. It is well known
468 that the abundance of ribosomal proteins in *E. coli* is lower at slower growth rates (45),
469 and our data on the abundance of ribosomal protein mRNAs are consistent with it. All
470 of the other species in this study all grow slower at 15°C than at 37°C, and some grow
471 as slowly as *E. coli* at 15°C, but none show the same down-regulation of ribosomal
472 protein mRNA. It is possible that these other species do not achieve down-regulation
473 of ribosomal activity at lower growth rate via the same mechanism, or do not down-
474 regulate ribosomal activity at all.
475 The alternative sigma factor RpoS is one regulatory protein that plays a role in
476 the transcriptional response to temperature shift. Much like the overall cold response,
477 the RpoS-dependent part of the cold response is only partially conserved between *E.*
478 *coli* and *S. enterica.* The divergence between the two species is due to a mix of
479 divergent expression of genes shared between the two species and species-specific
480 genes that are part of the RpoS-regulon. The divergence in RpoS-dependence of
481 genes shared by *E. coli* and *S. enterica* was somewhat surprising given that both
482 species produce similar amounts of RpoS at 15°C. Divergence in RpoS-dependence of
483 shared genes could be due to changes in the RpoS protein itself, in the promoters of
484 these divergently expressed genes, or could be due to changes in the many trans-
485 acting factors that interact with RpoS to determine the transcriptional output of a
486 promoter (42) . Understanding this molecular basis of these changes is an important
487 future area of research.
488 Other authors have also noted that species-specific genes are a substantial part
489 of the divergence of regulons between species pairs. For example, the OmpR protein
490 regulates transcription of a group of genes in response to physical stresses like

491 osmolarity and pH. The OmpR regulon in *E. coli* and *S. enterica* is only weakly
492 conserved, and the authors noted that many species-specific genes were in the
493 regulon (41). The RovA protein of *Yersinia* is involved in transcribing important virulence
494 genes in those organisms. The RovA-regulon of *Yersinia enterocolitica* and *Y. pestis* is
495 also only weakly conserved due at least in part to the species-specific genes regulated
496 by RovA in each species (47). This paper extends this idea to the RpoS-regulon, and
497 shows that while species-specific genes are indeed important for driving divergence of
498 the regulon, differential expression of shared genes is at least as important for this
499 regulon.
500
501 In conclusion, we have studied the transcriptional response to a change in
502 temperature from 37°C to 15°C in five species of the *Enterobacteriaceae* and one
503 outgroup. We have found that there is limited conservation of this response, with each
504 species changing the transcription of hundreds of genes that are unchanged in the
505 other species. This divergence is not due primarily to species-specific genes in each
506 genome, as core genes are more likely than expected to be differentially expressed in
507 response to temperature change. This pattern of evolutionary divergence is also true of
508 RpoS-regulon, where divergence is due more to shared genes being RpoS-regulated in
509 only a single species than it is due to species-specific genes being part of the regulon.
510 These results highlight the importance of the evolution of transcriptional patterns of
511 shared genes in the evolution of bacterial responses to stress.
512

513 Materials and Methods

514 *Strains and strain construction*

515 Six species were used in this study: *E. coli, Salmonella enterica, Citrobacter*
516 *rodentium, Enterobacter cloacae, Klebsiella pneumoniae,* and *Serratia marcescens.* We

517 used strains with wild-type alleles for *rpoS* for each species, along with Δ*rpoS* strains of
518 *E. coli* and *S. enterica* that we had access to (Table S1). To construct the isogenic
519 Δ*rpoS* strain of *S. enterica*, the Δ*rpoS::cm* allele of a *S. enterica* 14028s strain provided
520 by the McClelland Lab (48) was moved by P22 transduction (49) into our lab isolate of
521 *S. enterica* 14028s.
522

523 *Growth conditions*

524 Strains were inoculated from -80°C frozen cultures and grown aerobically
525 overnight in 5 mL of Luria-Bertani (LB) broth (1% tryptone, 0.5% yeast extract, 1%
526 NaCl) at 37˚C with shaking at 225 rpm. The next morning, 250μL of each culture was
527 diluted into 25 mL of LB in a 250 mL borosilicate flask and grown at 37˚C in an
528 incubator shaking at 225 rpm. When cultures reached OD_{600} of 0.25 - 0.35, flasks were
529 transferred to a water bath at 15°C shaking at 225 rpm. (OD_{600} of 0.3 is in the middle of
530 exponential growth in these cultures.) When cultures were grown for RNA and protein
531 isolation, no measurements of OD_{600} were performed at intermediate time points
532 between the shift and sampling of RNA and protein. When cultures were grown to
533 measure growth rate, OD_{600} measurements were taken every 20 to 30 minutes after the
534 shift to 15°C.
535 To minimize confounding factors, we always grew all 8 strains in parallel for
536 each experimental replicate.
537

538 *RNA Isolation*

539 RNA was isolated from the same cultures and at the same time as protein
540 samples. 1.5 mL of bacterial culture was pelleted and resuspended in 500 μL of Trizol
541 pre-heated to 65°C and then stored at -80°C until use. Samples were purified on a
542 column using the Zymo Direct-Zol RNA Miniprep kit. They were then treated twice with

543 DNase I using Turbo DNA-free (Invitrogen) for 30 minutes each at 37°C to ensure that
544 all DNA was degraded. The samples were further purified on a column using the Zymo
545 RNA Clean and Concentrator Kit and again stored at -80°C until sequencing. The
546 concentration of the samples was checked using a NanoDrop spectrophotometer and
547 degradation was checked for by electrophoresing samples on an ethidium bromide
548 agarose gel.
549 N = 4 biological replicate samples were collected for each strain at each
550 temperature.
551

552 *Generation of RNA-Seq data*

553 Illumina cDNA libraries were generated using a modified version of the RNAtag-
554 seq protocol (50). Briefly, 500 ng to 1 µg of total RNA was fragmented by heating to
555 94°C in NEB Alkaline Phosphatase Buffer, depleted of any residual genomic DNA,
556 dephosphorylated, and ligated to DNA adapters carrying 5'-AN_8-3' barcodes of known
557 sequence with a 5' phosphate and a 3' blocking group. Barcoded RNAs were pooled
558 and depleted of rRNA using the RiboZero rRNA depletion kit (Epicentre). Pools of
559 barcoded RNAs were converted to Illumina cDNA libraries in 2 main steps: (i) reverse
560 transcription of the RNA using a primer designed to the constant region of the
561 barcoded adaptor with addition of an adapter to the 3' end of the cDNA by template
562 switching using SMARTScribe (Clontech) as described (51); (ii) PCR amplification using
563 primers whose 5' ends target the constant regions of the 3' or 5' adaptors and whose
564 3' ends contain the full Illumina P5 or P7 sequences. cDNA libraries were sequenced
565 on the Illumina [NextSeq 500] platform at the Broad Institute to generate paired end
566 reads.
567 Each RNA pool consisted of 16 samples, a 15°C and a 37°C sample from each
568 of the 8 strains. N = 4 total pools were constructed and sequenced.
569

570 ## *RNA-seq Data Analysis*

571 Sequencing reads from each sample in a pool were demultiplexed based on
572 their associated barcode sequence using custom scripts. Up to 1 mismatch in the
573 barcode was allowed provided it did not make assignment of the read to a different
574 barcode possible. Barcode sequences were removed from the first read as were
575 terminal G's from the second read that may have been added by SMARTScribe during
576 template switching.

577 Reads were aligned to the reference genomes (Table S1) using BWA v0.7.17
578 (52). Reads were counted with htseq-count v2.0.1 (53). Further analysis was done in R
579 v4.4.0 (54).

580 Before conducting the analysis reported in this paper, we performed initial
581 quality control to ensure that each sample had at least 1,000,000 read-pairs that
582 mapped to protein coding sequences. One sample was below this threshold and
583 removed. In addition, three samples did not cluster with any of the other samples in
584 PCA or a heatmap and were also removed. The samples removed were one from *C.*
585 *rodentium* at 15°C, two from *K. pneumoniae* at 37°C, and one from *S. marcescens* at
586 37°C.

587 Differential expression analysis was performed using the package DESeq2
588 v1.40.2 (55). Genes were considered significant if they had a FDR-adjusted p-value of
589 less than 0.01 and an absolute value of fold-change greater than 2.

590 eggNOG-mapper (24) was used to assign GO terms to each gene in each
591 genome. GO enrichment for each species was performed with the R package topGO
592 (25). The GO tree for each species was pruned to remove terms with fewer than 5
593 genes. The weight01 algorithm was used with Fisher's exact test to calculate p-values
594 on enrichment of GO terms. GO terms were considered significantly enriched if they
595 had a p-value of less than 0.05.
596
597

598 *Protein Isolation & Western Blotting*

599 Total protein was collected immediately before the shift to 15°C and 3 hours
600 after initial cold exposure. Samples with cell density equivalent to 1mL of OD_{600} = 0.3
601 were pelleted at at least 20,000 x g for 30 s, resuspended in 100μL of 2x Laemmli
602 sample buffer (LiCor), heated at 95°C for 5 minutes, and then stored at -20°C until they
603 were used.

604 A quantitative western blot was used to measure the levels of RpoS in the cell at
605 each time point, in a procedure slightly modified from (15). 10μL of each protein
606 sample was electrophoresed on a 4 - 20% gradient polyacrylamide gel (Bio-Rad) in
607 Tris-glycine running buffer (25 mM Tris base, 250 mM glycine, 0.5% SDS) at 200V for
608 35 minutes at room temperature. Electrophoresis was used to transfer proteins to a
609 Immobilon-FL PVDF membrane in transfer buffer (48 mM Tris base, 39 mM glycine,
610 0.0375% SDS, 20% methanol) at 100V for 45 minutes at 4°C. In order to quantify the
611 total amount of protein in the samples, the membranes were stained with 5 mL of
612 REVERT Total Protein Stain solution and imaged using the LiCor Odyssey CLx imager.
613 The membranes were blocked with 10mL of Odyssey Blocking Buffer for 1 hour at
614 room temperature.

615 RpoS was detected using an anti-RpoS monoclonal antibody (clone 1RS1,
616 BioLegend). The blocked membranes were labeled with primary antibody (0.4 μg/mL
617 mouse anti-RpoS, 10mL Odyssey Blocking Buffer, 0.2% Tween 20) at 4°C with
618 shaking at 55 rpm overnight. Once labeled, the membrane was washed four times with
619 10 - 15mL of 1x Tris-buffered saline with Tween 20 for 5 minutes each wash. The
620 membranes were probed with a fluorescent secondary antibody (IRDye 800CW goat
621 anti-mouse, LiCor) that was diluted in 10mL of Odyssey Blocking Buffer with 0.2%
622 Tween 20 and 0.01% SDS. The membranes were incubated with the diluted secondary
623 antibody for 1 hour at room temperature and were washed as described previously,
624 with an additional wash at the end with 1x Tris-buffered saline without Tween 20.
625 Membranes were dried for 2-3 hours between 2 sheets of Whatmann 3MM blotting
626 paper before imaging on a Li-Cor Odyssey CLx imager.

627 RpoS levels were quantified by analyzing the amount of fluorescence of each band
628 with Image Studio v 2.1 software.
629 Data analysis was performed in R v4.4.0 (54). RpoS levels were normalized by
630 dividing them by the total amount of protein at each time point, as quantified by the
631 total protein stain.
632

633 *xenoGI analysis*

634 The program xenoGI v3.1.1 (23) was used to reconstruct the evolutionary history
635 of the genomes used in this study. The phylogeny used by xenoGI was inferred using a
636 concatenation of the amino acid sequences of all of the core genes as identified by
637 xenoGI. The phylogeny was inferred by Astral v5.6.3 (56) using the built-in helper
638 function of xenoGI.
639

640 *Data Availability*

641 The RNA-seq FASTQ files are deposited in the Gene Expression Omnibus with
642 ascension number GSE267531.
643 All other data and scripts used for data analysis are in a Zenodo archive at DOI
644 10.5281/zenodo.11460468
645
646
647

648 # Acknowledgements

649 RNA-Seq libraries were constructed and sequenced at the Broad Institute of
650 MIT and Harvard by the Microbial 'Omics Core and Genomics Platform, respectively.
651 The Microbial 'Omics Core also provided guidance on experimental design and
652 performed the demultiplexing of RNA-Seq data. Thanks to Michael McClelland for

653 sharing the *S. enterica* Δ*rpoS::cm* allele and phage P22, to Eliot Bush for help with
654 xenoGI, and to Danae Schulz and Chris White-Ziegler for feedback and insights.
655 This material is based upon work supported by the National Science Foundation
656 under Grant No. 1716794.
657

658 # References

659
660 1. Zhang Y, Gross CA. 2021. Cold Shock Response in Bacteria. Annual Review of
661 Genetics 55:377–400.
662 2. White-Ziegler CA, Um S, Pérez NM, Berns AL, Malhowski AJ, Young S. 2008. Low
663 temperature (23 degrees C) increases expression of biofilm-, cold-shock- and
664 RpoS-dependent genes in Escherichia coli K-12. Microbiology (Reading) 154:148–
665 166.
666 3. Moon S, Ham S, Jeong J, Ku H, Kim H, Lee C. 2023. Temperature Matters:
667 Bacterial Response to Temperature Change. J Microbiol 61:343–357.
668 4. Xia B, Ke H, Inouye M. 2001. Acquirement of cold sensitivity by quadruple deletion
669 of the cspA family and its suppression by PNPase S1 domain in Escherichia coli.
670 Molecular Microbiology 40:179–188.
671 5. Zhang Y, Burkhardt DH, Rouskin S, Li G-W, Weissman JS, Gross CA. 2018. A
672 Stress Response that Monitors and Regulates mRNA Structure Is Central to Cold
673 Shock Adaptation. Molecular Cell 70:274-286.e7.
674 6. Zangrossi S, Briani F, Ghisotti D, Regonesi ME, Tortora P, Dehò G. 2000.
675 Transcriptional and post-transcriptional control of polynucleotide phosphorylase

676 during cold acclimation in Escherichia coli. Molecular Microbiology 36:1470–1480.
677 7. Dersch P, Kneip S, Bremer E. 1994. The nucleoid-associated DNA-binding protein
678 H-NS is required for the efficient adaptation of Escherichia coli K-12 to a cold
679 environment. Mol Gen Genet 245:255–259.
680 8. Dash S, Palma CSD, Baptista ISC, Almeida BLB, Bahrudeen MNM, Chauhan V,
681 Jagadeesan R, Ribeiro AS. 2022. Alteration of DNA supercoiling serves as a trigger
682 of short-term cold shock repressed genes of E. coli. Nucleic Acids Research
683 50:8512–8528.
684 9. Sledjeski DD, Gupta A, Gottesman S. 1996. The small RNA, DsrA, is essential for
685 the low temperature expression of RpoS during exponential growth in Escherichia
686 coli. EMBO J 15:3993–4000.
687 10. White-Ziegler CA, Malhowski AJ, Young S. 2007. Human Body Temperature
688 (37°C) Increases the Expression of Iron, Carbohydrate, and Amino Acid Utilization
689 Genes in Escherichia coli K-12. Journal of Bacteriology 189:5429–5440.
690 11. White-Ziegler CA, Davis TR. 2009. Genome-wide identification of H-NS-
691 controlled, temperature-regulated genes in Escherichia coli K-12. J Bacteriol
692 191:1106–10.
693 12. Adeolu M, Alnajar S, Naushad S, S. Gupta R. 2016. Genome-based phylogeny
694 and taxonomy of the 'Enterobacteriales': proposal for Enterobacterales ord. nov.
695 divided into the families Enterobacteriaceae, Erwiniaceae fam. nov.,
696 Pectobacteriaceae fam. nov., Yersiniaceae fam. nov., Hafniaceae fam. nov.,
697 Morganellaceae fam. nov., and Budviciaceae fam. nov. International Journal of
698 Systematic and Evolutionary Microbiology 66:5575–5599.

699 13. Brenner DJ, Farmer JJ. 2015. *Enterobacteriaceae*, p. 1–24. *In* Whitman, WB
700 (ed.), Bergey's Manual of Systematics of Archaea and Bacteria, 1st ed. Wiley.
701 14. Guentzel MN. 1996. Escherichia, Klebsiella, Enterobacter, Serratia, Citrobacter,
702 and Proteus, p. . *In* Baron, S (ed.), Medical Microbiology, 4th ed. University of
703 Texas Medical Branch at Galveston, Galveston (TX).
704 15. Popoff MY, Le Minor LE. 2015. Salmonella, p. 1–1. *In* Bergey's Manual of
705 Systematics of Archaea and Bacteria. John Wiley & Sons, Ltd.
706 16. Leimbach A, Hacker J, Dobrindt U. 2013. E. coli as an All-Rounder: The Thin
707 Line Between Commensalism and Pathogenicity, p. 3–32. *In* Dobrindt, U, Hacker,
708 JH, Svanborg, C (eds.), Between Pathogenicity and Commensalism. Springer Berlin
709 Heidelberg.
710 17. Wyres KL, Holt KE. 2018. *Klebsiella pneumoniae* as a key trafficker of drug
711 resistance genes from environmental to clinically important bacteria. Current
712 Opinion in Microbiology 45:131–139.
713 18. Ishii S, Ksoll WB, Hicks RE, Sadowsky MJ. 2006. Presence and Growth of
714 Naturalized Escherichia coli in Temperate Soils from Lake Superior Watersheds.
715 Applied and Environmental Microbiology 72:612–621.
716 19. Rivera SC, Hazen TC, Toranzos GA. 1988. Isolation of fecal coliforms from
717 pristine sites in a tropical rain forest. Appl Environ Microbiol 54:513–517.
718 20. Kröger C, Colgan A, Srikumar S, Händler K, Sivasankaran SK, Hammarlöf DL,
719 Canals R, Grissom JE, Conway T, Hokamp K, Hinton JCD. 2013. An infection-
720 relevant transcriptomic compendium for Salmonella enterica Serovar Typhimurium.
721 Cell Host Microbe 14:683–695.

722 21. Tripathy S, Sen R, Padhi SK, Mohanty S, Maiti NK. 2014. Upregulation of
723 transcripts for metabolism in diverse environments is a shared response associated
724 with survival and adaptation of Klebsiella pneumoniae in response to temperature
725 extremes. Funct Integr Genomics 14:591–601.
726 22. Kumari K, Sharma PK, Singh RP. 2024. Unravelling the transcriptome response
727 of Enterobacter sp. S-33 under varying temperature. Arch Microbiol 206:81.
728 23. Liu N, Gonzalez TA, Fischer J, Hong C, Johnson M, Mawhorter R, Mugnatto F,
729 Soh R, Somji S, Wirth JS, Libeskind-Hadas R, Bush EC. 2023. xenoGI 3: using the
730 DTLOR model to reconstruct the evolution of gene families in clades of microbes.
731 BMC Bioinformatics 24:295.
732 24. Cantalapiedra CP, Hernández-Plaza A, Letunic I, Bork P, Huerta-Cepas J. 2021.
733 eggNOG-mapper v2: Functional Annotation, Orthology Assignments, and Domain
734 Prediction at the Metagenomic Scale. Molecular Biology and Evolution 38:5825–
735 5829.
736 25. Alexa A, Rahnenfuhrer J. 2023. topGO: Enrichment analysis for Gene Ontology.
737 R package version 2.54.0. (2.54.0).
738 26. Mey AR, Gómez-Garzón C, Payne SM. 2021. Iron Transport and Metabolism in
739 Escherichia, Shigella, and Salmonella. EcoSal Plus 9:eESP-0034-2020.
740 27. Holden VI, Bachman MA. 2015. Diverging roles of bacterial siderophores during
741 infection. Metallomics 7:986–995.
742 28. Steingard CH, Helmann JD. 2023. Meddling with Metal Sensors: Fur-Family
743 Proteins as Signaling Hubs. Journal of Bacteriology 205:e00022-23.
744 29. Nandal A, Huggins CCO, Woodhall MR, McHugh J, Rodríguez-Quiñones F,

745 Quail MA, Guest JR, Andrews SC. 2010. Induction of the ferritin gene (ftnA) of
746 Escherichia coli by Fe2+–Fur is mediated by reversal of H-NS silencing and is RyhB
747 independent. Molecular Microbiology 75:637–657.
748 30. Sezonov G, Joseleau-Petit D, D'Ari R. 2007. Escherichia coli Physiology in Luria-
749 Bertani Broth. J Bacteriol 189:8746–8749.
750 31. Kim Y-H, Han KY, Lee K, Lee J. 2005. Proteome response of Escherichia coli
751 fed-batch culture to temperature downshift. Appl Microbiol Biotechnol 68:786–793.
752 32. Lemke JJ, Sanchez-Vazquez P, Burgos HL, Hedberg G, Ross W, Gourse RL.
753 2011. Direct regulation of Escherichia coli ribosomal protein promoters by the
754 transcription factors ppGpp and DksA. Proceedings of the National Academy of
755 Sciences 108:5712–5717.
756 33. Keseler IM, Gama-Castro S, Mackie A, Billington R, Bonavides-Martínez C,
757 Caspi R, Kothari A, Krummenacker M, Midford PE, Muñiz-Rascado L, Ong WK,
758 Paley S, Santos-Zavaleta A, Subhraveti P, Tierrafría VH, Wolfe AJ, Collado-Vides J,
759 Paulsen IT, Karp PD. 2021. The EcoCyc Database in 2021. Front Microbiol
760 12:711077.
761 34. Miajlovic H, Smith SG. 2014. Bacterial self-defence: how Escherichia coli
762 evades serum killing. FEMS Microbiology Letters 354:1–9.
763 35. Stevenson G, Andrianopoulos K, Hobbs M, Reeves PR. 1996. Organization of
764 the Escherichia coli K-12 gene cluster responsible for production of the
765 extracellular polysaccharide colanic acid. J Bacteriol 178:4885–4893.
766 36. Limoli DH, Jones CJ, Wozniak DJ. 2015. Bacterial Extracellular Polysaccharides
767 in Biofilm Formation and Function. Microbiology Spectrum

768 3:10.1128/microbiolspec.mb-0011–2014.
769 37. Wall E, Majdalani N, Gottesman S. 2018. The Complex Rcs Regulatory
770 Cascade. Annu Rev Microbiol 72:111–139.
771 38. Gottesman S, Stout V. 1991. Regulation of capsular polysaccharide synthesis in
772 Escherichia coli K12. Mol Microbiol 5:1599–1606.
773 39. Adams J, Hoang J, Petroni E, Ashby E, Hardin J, Stoebel DM. 2023. The timing
774 of transcription of RpoS-dependent genes varies across multiple stresses in
775 *Escherichia coli* K-12. mSystems 8:e00663-23.
776 40. Wong GT, Bonocora RP, Schep AN, Beeler SM, Fong AJL, Shull LM, Batachari
777 LE, Dillon M, Evans C, Becker CJ, Bush EC, Hardin J, Wade JT, Stoebel DM. 2017.
778 Genome-Wide Transcriptional Response to Varying RpoS Levels in Escherichia coli
779 K-12. J Bacteriol 199:e00755-16.
780 41. Weber H, Polen T, Heuveling J, Wendisch VF, Hengge R. 2005. Genome-wide
781 analysis of the general stress response network in *Escherichia coli*: sigmaS-
782 dependent genes, promoters, and sigma factor selectivity. Journal of bacteriology
783 187:1591–603.
784 42. Bouillet S, Bauer TS, Gottesman S. 2024. RpoS and the bacterial general stress
785 response. Microbiology and Molecular Biology Reviews 0:e00151-22.
786 43. Mates AK, Sayed AK, Foster JW. 2007. Products of the *Escherichia coli* acid
787 fitness island attenuate metabolite stress at extremely low pH and mediate a cell
788 density-dependent acid resistance. J Bacteriol 189:2759–2768.
789 44. Lerminiaux NA, MacKenzie KD, Cameron ADS. 2020. Salmonella Pathogenicity
790 Island 1 (SPI-1): The Evolution and Stabilization of a Core Genomic Type Three

791 Secretion System. Microorganisms 8:576.
792 45. Ingraham JL, Maaløe O, Neidhardt FC. 1983. Growth of the Bacterial Cell.
793 Sinauer Associates, Sunderland, MA.
794 46. Quinn HJ, Cameron ADS, Dorman CJ. 2014. Bacterial Regulon Evolution:
795 Distinct Responses and Roles for the Identical OmpR Proteins of Salmonella
796 Typhimurium and Escherichia coli in the Acid Stress Response. PLOS Genetics
797 10:e1004215.
798 47. Ellison DW, Hinchliffe SJ, Wren BW, Miller VL. 2007. The RovA regulons of
799 Yersinia enterocolitica and Yersinia pestis are distinct: evidence that many RovA-
800 regulated genes were acquired more recently than the core genome. Molecular
801 Microbiology 66:189–205.
802 48. Porwollik S, Santiviago CA, Cheng P, Long F, Desai P, Fredlund J, Srikumar S,
803 Silva CA, Chu W, Chen X, Canals R, Reynolds MM, Bogomolnaya L, Shields C, Cui
804 P, Guo J, Zheng Y, Endicott-Yazdani T, Yang H-J, Maple A, Ragoza Y, Blondel CJ,
805 Valenzuela C, Andrews-Polymenis H, McClelland M. 2014. Defined Single-Gene
806 and Multi-Gene Deletion Mutant Collections in Salmonella enterica sv Typhimurium.
807 PLOS ONE 9:e99820.
808 49. Schmieger H. 1972. Phage P22-mutants with increased or decreased
809 transduction abilities. Mol Gen Genet 119:75–88.
810 50. Shishkin AA, Giannoukos G, Kucukural A, Ciulla D, Busby M, Surka C, Chen J,
811 Bhattacharyya RP, Rudy RF, Patel MM, Novod N, Hung DT, Gnirke A, Garber M,
812 Guttman M, Livny J. 2015. Simultaneous generation of many RNA-seq libraries in a
813 single reaction. Nat Meth 12:323–325.

814 51. Zhu YY, Machleder EM, Chenchik A, Li R, Siebert PD. 2001. Reverse
815 transcriptase template switching: a SMART approach for full-length cDNA library
816 construction. Biotechniques 30:892–897.
817 52. Li H. 2013. Aligning sequence reads, clone sequences and assembly contigs
818 with BWA-MEM. arXiv:13033997 [q-bio].
819 53. Anders S, Pyl PT, Huber W. 2015. HTSeq--a Python framework to work with
820 high-throughput sequencing data. Bioinformatics 31:166–169.
821 54. R Core Team. 2023. R: A language and environment for statistical computing. R
822 Foundation for Statistical Computing, Vienna, Austria.
823 55. Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and
824 dispersion for RNA-seq data with DESeq2. Genome Biology 15:550.
825 56. Zhang C, Rabiee M, Sayyari E, Mirarab S. 2018. ASTRAL-III: polynomial time
826 species tree reconstruction from partially resolved gene trees. BMC Bioinformatics
827 19:153.
828
829
830