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

Johnson Hoang[a]* and Daniel M Stoebel[a]#

[a] Department of Biology, Harvey Mudd College, Claremont, California, USA

Running head: Transcriptional response to cold in *Enterobacteriaceae*

*Present address: Doheny Eye Institute, Pasadena, California, USA

#Correspondence to: stoebel@g.hmc.edu

Abstract

Bacteria respond to changes in their external environment, such as temperature, by changing the transcription of their genes. We know little about how these regulatory patterns evolve. We used RNA-seq to study the transcriptional response to 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 were only 16 differentially expressed genes in common among the six species. Species-specific transcriptional patterns of shared genes were a prominent cause of this lack of conservation. GO enrichment of regulated genes suggested many species-specific phenotypic responses to temperature changes, but enriched terms associated with iron metabolism, central metabolism, and biofilm formation were implicated in at least half of the species. The alternative sigma factor RpoS regulated 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. Overall, there was limited conservation of the response to low temperature generally, or the RpoS-regulated part of the response specifically. This study suggests that species-specific patterns of transcription of shared genes, rather than horizontal acquisition of unique genes, are the major reason for the lack of conservation of the transcriptomic response to low temperature.

Importance

We studied how different species of bacteria from the same Family (Enterobacteriacae) change the expression of their genes in response to a decrease in temperature. Using de novo-generated parallel RNA-seq datasets, 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 were very few genes that change expression in all six species. This was due in part to differences in gene

38 content, and in part due to shared genes with distinct expression profiles between the species.
39 This study is important to the field because it illustrates that closely related species can share
40 many genes but not use those genes in the same way in response to the same environmental
41 change.

42 Introduction

43 Bacteria change the levels of transcription of their genes to respond to changes in their
44 environments. These patterns of transcription, and the role of regulatory proteins and other
45 factors in determining these patterns have been measured thoroughly in a handful of model
46 species, like *Escherichia coli*. The degree to which our understanding of *E. coli* gene regulation
47 can inform our understanding of related species depends on how similar those species are. We
48 know that the genomes of bacterial species often differ dramatically in content due to
49 horizontal gene transfer (HGT), so it is possible that HGT may drive dramatic differences
50 between related species in what genes are differentially expressed (DE) in response to an
51 environmental change. Differences in transcriptome patterns between related species can also
52 be due to shared genes that have different patterns of regulation among species. In general,
53 we have little information about the relative roles of species-specific genes vs. species-specific
54 regulation of shared genes in shaping how regulons differ between related species.
55 Changes in temperature are one challenge faced by many bacteria (1). Low
56 temperature can affect bacteria by rigidifying the lipid bilayer, by altering the folding of
57 proteins and RNAs, and by influencing the assembly and activity of essential macromolecular
58 machinery like ribosomes (1). Low temperature may also serve as a cue that bacteria have
59 moved from one environment (such as a mammalian host) into another environment (such as a
60 surface at ambient temperature) (2).
61 A great deal is known about how *E. coli* responds to low temperature (1, 3). Several
62 global regulatory factors have been implicated, including cold shock proteins (4, 5), RNA
63 degradation enzymes like RNase R and PNPase (5, 6), nucleoid associated proteins like H-NS
64 (7), global supercoiling levels (8), and the alternative sigma factor RpoS (2, 9). Shifts between

65 optimal temperature (37°C) and a lower temperature have substantial impacts on the
66 transcriptome. In a series of experiments examining shifts between 37°C and 23°C (and vice
67 versa), White-Ziegler *et al.* documented that 423 genes, or about 10% of the *Escherichia coli* K-
68 12 genome, are DE (2, 10). Further, they identified RpoS and H-NS as two global regulators of
69 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 many
72 different thermal environments (13). This group includes genera that can be found in
73 endothermic hosts either causing an infection or in the gastrointestinal tract of healthy hosts
74 (14). Many of these species can also be isolated from ectothermic animals, plants, soil, water,
75 and plants (15–19). The presence of these organisms outside of endothermic hosts suggests
76 that they naturally experience broad changes in temperature.

77 In general, we have little understanding of how regulatory responses evolve in bacteria
78 and at what rate these regulons evolve. In addition to questions of rate, regulons could evolve
79 because different species lose different genes or gain different genes via horizontal gene
80 transfer (HGT). Regulons could also evolve because shared genes evolve different regulatory
81 controls in different species. We do not understand the relative importance of these two
82 modes of regulatory evolution. Investigating the tempo and mode of regulatory evolution in
83 the *Enterobacteriaceae* will both allow fundamental understanding of bacterial evolution and
84 allow researchers to better understand the degree to which *E. coli* serves as an appropriate
85 model system for related bacteria.

86 Single species studies of the transcriptomic response to low temperature exist for
87 several species of *Enterobacteriaceae*, including *E. coli* (2, 10), *Salmonella enterica* (20),
88 *Klebsiella pneumoniae* (21) and *Enterobacter sp.* (22). Unfortunately, it is difficult to directly
89 compare existing studies of individual species because of the large differences in culture
90 conditions, overall experimental approach (e.g. microarrays versus RNA-seq), experimental
91 details (e.g. methods of library construction for RNA-seq), and bioinformatic and statistical
92 analyses in these studies. To directly compare the transcriptional response, the experimental
93 and analytical approaches should be the same. In addition, batch effects in both bacterial

94 growth and RNA-seq steps can be avoided by performing all the bacterial growth, RNA
95 processing, library construction, and sequencing in parallel with a balanced experimental
96 design.

97 In this study, we studied how five species of bacteria in the *Enterobacteriaceae* plus one
98 closely related species responded to the shift from 37°C to 15°C, as might occur when bacteria
99 leave a mammalian host and find themselves in a temperate external environment. We used
100 comparative genomic tools to identify homologous genes in each species, and then used RNA-
101 seq to examine the patterns of transcription in each species across this temperature transition.
102 In addition, we studied the role of the alternative sigma factor RpoS in this transition in *E. coli*
103 and *S. enterica.* We performed these studies in parallel for all strains, avoiding possible
104 confounding differences due to experimental or analytical approaches or due to batch effects.
105

106 Results

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

108 This study worked with strains of five species of *Enterobacteriaceae*, (*E. coli* K-12, *S.*
109 *enterica* serovar Typhimurium 14028s, *Citrobacter rodentium* DBS100, *Enterobacter cloacae*
110 ATCC13047, *K. pneumoniae* MKP103) and one species in the *Yersiniaceae* (*Serratia marcescens*
111 Db11) (Figure 1A) (12). Their genomes contained between 3968 and 4709 protein coding
112 sequences (Figure 1B), which were the basis of all the analyses in this paper. We used xenoGI
113 (23) to understand the evolutionary histories of the six genomes. xenoGI uses both synteny and
114 sequence similarity to define groups of genes called locus families. We defined all genes in a
115 locus family as orthologues.

116 There were 1838 genes found in every genome, a set of genes denoted the core
117 genome. The core genome accounted for 39% to 46% of all genes in a genome. The number of
118 species-specific genes varied from 981 to 1872 (Figure 1C), which accounted for 24% to 47% of

119 all genes in the genome. Most genes in a species were either species-specific or ancestral to
120 the group in this analysis.

121 *Limited conservation of patterns of differential expression in response to*
122 *temperature shift*

123 We investigated how the six species responded to a change in temperature from 37°C
124 to 15°C. All species underwent a lag phase after the shift to 15°C, but all were growing again
125 after 3 hours at 15°C (Table S1). After three hours at 15°C, strains were expected to be past the
126 short-term response to cold (involving changes in DNA supercoiling) and into the longer-term
127 response, involving transcriptional changes driven by changing activity of transcription factors
128 (8).

129 We then performed an RNA-seq experiment on samples growing exponentially at 37°C
130 (immediately before a shift to 15°C), and then after three hours growth at 15°C. Principal
131 components analysis (PCA) confirmed that all samples from each species at each temperature
132 clustered together (Figure 2A), with differences apparent both among species and between
133 the two temperatures. 626 to 1057 genes (14 and 26% of all protein coding genes in the
134 genome) were significantly differentially expressed between the two temperatures across the
135 species (Figure 2B; File S1). (Differential expression was defined as an FDR-adjusted p-value <
136 0.01 and an absolute $\log_2$-fold change > 2 .) We found little conservation of the set of DE genes
137 in each species (Figure 2C,D). For example, each species had between 274 and 505 genes
138 uniquely DE in that species. In contrast, there were only 16 genes that are DE in every species
139 studied (Table S2). The transcriptional response to low temperature varied strongly among
140 these species.

141 Each genome contained between 981 to 1872 species-specific genes (Figure 1B), which
142 could be a major contributor to the limited conservation of DE genes. However, these species-
143 specific genes were less likely to be DE between temperatures than expected if all genes had
144 equal probability of being DE (Figure 3A; chi-square test, $p < 5 \times 10^{-4}$ for all six species.) In
145 contrast, genes in the core genome were more likely to be DE than expected by chance (Figure
146 3B; chi-square test, $p < 2 \times 10^{-3}$ for all six species). Core genes were particularly likely to be DE in

147 each species, but which core genes were DE in each species was not well-conserved (Figure
148 3C,D). For example, of the 561 core genes that were DE in *E. coli*, 231 were not DE in any other
149 species, even though all species in this study had those genes. There was not a total lack of
150 conservation of expression pattern, however. For example, there were 60 core genes that were
151 DE in both *E. coli* and *S. enterica* and in no other species. Species-specific expression patterns
152 of core genes strongly contributed to the distinctive transcriptional response to temperature
153 shift of each species.

154 We extended our analysis to the other internal nodes on the phylogeny and examined
155 not just the core genes but all those genes that show some level of sharing between species. At
156 each node on the phylogeny, we considered those genes that had an ortholog in each
157 descendent of that node, but in none of the other species, and were DE in at least one of the
158 species. In every case, we found that if a gene was DE in at least one species, the most likely
159 outcome was that it was DE in only that one species (Figure 3E) with decreasing frequencies as
160 the number of species increased. Thus, even when genes were shared between multiple
161 species, patterns of differential expression were primarily not conserved.

162 Another way to visualize the overall similarity of the transcriptional response or lack
163 thereof is PCA. Principal component 2 of the expression levels of the 1838 core genes, which
164 accounted for 15% of the variance in gene expression, separateed 15°C samples from 37°C
165 samples (Figure 2A). (PC1 and other principal components not visualized in Figure 2A did not
166 clearly separate samples based on temperature.) Thus, there was some conserved difference in
167 the expression of core genes between 15°C and 37°, but it was a small proportion of all
168 variation in gene expression among the species.

169 *Functions of genes that are DE between warm and cold environments*

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

175 quality of the annotation, we examined the GO terms assigned to each orthologue of the core
176 genome. We found that 1668 of 1838 (91%) of core genes had identical GO annotations in all
177 six genomes, and most of the remaining 118 core genes had identical annotations for 5 of 6
178 genomes, suggesting that the annotation worked consistently for most core genes.
179 We then performed GO enrichment analysis using topGO (25) on the RNA-seq data
180 from each species. Consistent with the lack of conservation of differentially expressed genes,
181 there was overall little conservation of enriched GO terms across the genomes (Figure 4; File
182 S2). Of the 151 GO terms that were enriched in at least one species, none were enriched in all
183 six species or in five of the six species, and only 10 were enriched in three or four species (Table
184 S2). 107 of the 151 terms were enriched in only a single species.
185

186 ## Iron metabolism genes are affected by temperature in multiple species

187 One group of GO terms that were enriched in multiple genomes were those associated
188 with iron acquisition and use. Enterobactin is a siderophore that is synthesized by many
189 species of Enterobacteraceae (26). This molecule is synthesized inside the cell and secreted
190 into the environment, where it can bind to ferric iron. The iron-enterobactin complex is then
191 actively transported into the cell where the enterobactin is cleaved, freeing the iron (26). Iron is
192 scarce inside of mammalian hosts and genes that allow for iron scavenging are important
193 virulence factors in many of the species in this study (27). The enterobactin biosynthesis term
194 (GO:0009239) was enriched in *E. coli, S. enterica, C. rodentium,* and *K. pneumoniae*, with the
195 associated genes downregulated at 15°C in these species (Figure 5A). Genes for enterobactin
196 transport (GO:0042930) were enriched in *S. enterica*, *C. rodentium, E. cloacae* and *K.*
197 *pneumoniae*, with the associated genes also downregulated (Figure 5B). The *E. coli* results are
198 consistent with previous findings that 37°C increases the expression of iron utilization genes in
199 that organism relative to lower temperatures (10). The overall picture is that *S. enterica, C.*
200 *rodentium,* and *K. pneumoniae*, like *E. coli,* also express genes for iron acquisition less at 15°C
201 than at 37°C. That said, this was not a completely conserved regulatory pattern, as

202 enterobactin biosynthesis levels were largely unchanged in response to temperature in *E.*
203 *cloacae* and *S. marcescens.*
204 One notable use for iron in enteric bacteria is the production of iron-sulfur clusters.
205 Genes associated with iron-sulfur cluster assembly (GO:0016226) that were significantly DE
206 were all upregulated in 15°C relative to 37°C in most species (Figure 5C). However, there were
207 multiple genes expressed at lower levels in *C. rodentium* at 15°C and one such in *E. coli.* Genes
208 associated with the transfer of mature iron-sulfur clusters into proteins (GO:0097428) that
209 were significantly DE were primarily upregulated (Figure 5D), although the number of DE
210 genes varied among the species.
211 One key regulator of genes associated with iron metabolism in many species is the
212 protein Fur. Fur is primarily a transcriptional repressor (28), but it can serve as a positive
213 regulator in cases where it disrupts repression by another protein, such as its action as an anti-
214 silencer of H-NS at the *ftnA* promoter (29). Fur mRNA levels were significantly higher at 15°C
215 than 37°C only for *C. rodentium* and *E. cloacae* (Figure 5E). *E. coli* and *S. enterica* showed clear
216 repression of enterobactin biosynthesis with non-significant (and very modest) changes in *fur*
217 levels. Temperature regulation of *hns* mRNA also varies across species, with levels significantly
218 higher at 15°C only for *K. pneumoniae* and *S. marcescens* (Figure 5F). Fur protein activity is
219 allosterically regulated by iron (28), so there is not a direct relationship between *fur* mRNA
220 levels and Fur activity. Similarly, H-NS activity is influenced by both physico-chemical factors
221 and by other proteins (30–34), so there is no simple relationship between *hns* mRNA levels and
222 H-NS repression. The effect of increasing *fur* and *hns* mRNA levels at low temperature on the
223 transcription of iron uptake genes remains to be determined.

224 Other enriched GO terms

225 Genes involved in the TCA cycle and with glycine metabolism were significantly
226 enriched in four genomes each (Table S2). (TCA cycle genes are also annotated as involved in
227 citrate metabolism, GO:0006101). These genes were largely upregulated in response to cold
228 (Figure S1A, B), consistent with a previous proteomic study (35). Genes involved in single-
229 species biofilm formation were enriched in three species (Table S2, Figure S1C), replicating

230 previous results from *E. coli* (2), and suggesting that this response is conserved in related
231 species. Genes annotated as being involved in the response to cold were only enriched in *E. coli*
232 and *S. enterica* (Table S2, Figure S1D) . Most experimental information used to create GO
233 annotations comes from experiments in *E. coli*, so the lack of enrichment of response to cold
234 genes in more distantly related species is consistent with the lack of conservation in what
235 genes are DE.

236 *Species-specific gene expression changes in response to temperature*

237 While a few GO terms were enriched across multiple species, 70% of enriched GO terms
238 were species-specific. Here we highlight several species-specific responses. Similar plots of
239 genes associated with a GO term enriched in at least one species are available in the Zenodo
240 archive (DOI 10.5281/zenodo.11460468) for interested readers.

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

242 GO terms for ribosome assembly (GO:0000027 and GO:0000028, for ribosomal large
243 and small subunit assembly, respectively) were enriched in the DE genes of *E. coli* but not in
244 other species (Figure S2A). The genes with those GO terms included the structural
245 components of the ribosome as well as factors that modulate ribosome assembly. Those DE
246 ribosomal genes were all downregulated in *E. coli*. Genes coding for proteins that produce the
247 proton motive force via electron transport (GO:0015990) were also enriched in *E. coli* but not
248 the other species (Figure S2B). Those genes, which were all downregulated, were *cyoABCD*,
249 and *nuoLMN*, encoding the cytochrome bo3 quinol:oxygen oxidoreductase and
250 NADH:ubiquinone oxidoreductase I respectively. The mechanism and consequences of this *E.*
251 *coli*-specific regulation remains to be determined.
252

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

254 Colanic acid is the capsular polysaccharide in *E. coli* and some relatives (36–38). Genes
255 annotated with the GO term for colanic acid biosynthesis ("GO:0009242") were enriched in *C.*

256 *rodentium* but not in other species. Some genes in the *wca* cluster, which encodes the
257 biosynthetic enzymes for colanic acid (37), were not annotated with this GO term, so we
258 created a custom gene set that included all genes annotated with GO:0009242 plus all genes
259 from the *wca* cluster that were missing that annotation. (The strains of *K. pneumoniae* and *S.*
260 *marcescens* we worked with do not have orthologues to the *wca* genes found in the other four
261 species.) 70% of genes in the combined set were significantly upregulated in *C. rodentium*,
262 while many fewer (if any) genes were upregulated in each of the other species (Fig 6A).
263 In *E. coli*, the Rcs phosphorelay system is an important regulator of colanic acid
264 biosynthesis (39). The response regulator RcsB forms a heterodimer with the RcsA protein, and
265 these two proteins are required for maximal capsule expression in that organism (40). The
266 gene *rcsA* had an expression pattern that closely matched the pattern of its target genes, being
267 significantly upregulated 10-fold in *C. rodentium,* while being upregulated at most 2-fold in the
268 other five species (Fig 6B). In contrast, RcsB showed a similar pattern of expression across five
269 of the six species, rather than being upregulated only in *C. rodentium* (Fig 6C). The evolution of
270 species-specific regulation of RcsA in *C. rodentium* is a promising candidate for the
271 evolutionary change that brought about the species-specific pattern of the colanic acid
272 biosynthesis genes.
273

274 *RpoS regulation of the response to temperature shift*

275 The alternative sigma factor RpoS is known to be activated at low temperature in *E. coli*
276 where it regulates the expression of several hundred genes (41). We compared the role of RpoS
277 in *E. coli* and its close relative *S. enterica* during growth at 15°C. Western blotting
278 demonstrated that RpoS had a similar pattern of expression in both species. RpoS could
279 accumulate during exponential growth at 15°C to a similar level as in stationary phase at 37°C,
280 but not during exponential growth at 37°C (Figure S3).

281 RNA-seq experiments with Δ*rpoS* strains of *E. coli* and *S. enterica* allowed us to define
282 the RpoS regulon at 15°C. During exponential growth at 15°C, 232 genes were DE

283 between wild-type and Δ*rpoS E. coli.* This represented 22% of the number of genes that were
284 DE between 37°C and 15°C in the wild-type strain. Reassuringly, of the 232 DE genes in *E. coli* in
285 this study, 189 (81%) were previously identified as RpoS-regulated at 15°C by Adams *et al.* (41).
286 In *S. enterica,* 189 genes were DE between the wild-type and Δ*rpoS* strains, 80% of the
287 number of genes that are RpoS-regulated in *E. coli.* RpoS regulates many genes in the cold,
288 although only a fraction of all the genes that are cold-regulated.
289 The role of RpoS in regulating the 15°C response is similar but distinct in *E. coli* and *S.*
290 *enterica.* Principal component analysis of the expression of core genes clearly separated the
291 two species, but also found the same direction of change in PCA space between wild-type and
292 Δ*rpoS* strains of both species (Figure 7A). 83 genes were DE in both *E. coli* and *S. enterica*
293 (Figure 7B). This included genes previously reported to be RpoS-dependent (42–44) such as
294 *dps*, *katE*, *mlrA*, *osmC*, *osmY*, *otsA*, *otsB*, *sodC*, and *wrbA.* Consequently, GO terms involved in
295 response to osmotic stress or response to hydrogen peroxide were enriched in both species.
296 Most of the RpoS-regulon in *E. coli* or *S. enterica* was not composed of the 83 genes that were
297 RpoS-regulated in the other species (Figure 7B). For example, of the 232 RpoS-regulated genes
298 in *E. coli,* 149 of those genes (64%) were not RpoS-regulated in *S. enterica.* Those 149 genes
299 were composed of 50 genes that were not present in *S. enterica,* and 99 genes that had an
300 ortholog in *S. enterica*, but that ortholog was not RpoS-regulated in *S. enterica.* Like the entire
301 cold-response regulon, divergence of the RpoS-regulon between *E. coli* and *S. enterica* was due
302 to a mixture of both different regulation of genes shared between the two species and the
303 presence of species-specific genes.
304 Genes that were RpoS-regulated only in *E. coli* include *astABCDE,* responsible for
305 arginine catabolism (GO:0006527), *dppABCDF*, involved in di-peptide transport, *gabD*, *astC*,
306 and *ldcC*, involved in lysine metabolism (GO:0006553). These genes have orthologs in *S.*
307 *enterica* but were not significantly DE in that organism, pointing to a role for RpoS in amino-
308 acid catabolism in *E. coli* that is not present in *S. enterica.* RpoS also had an *E. coli*-specific role
309 in regulation of genes *gadE*, *gadW*, *gadX*, *hdeA*, *hdeB*, and *slp*, which are part of the acid fitness
310 island not present in related species like *S. enterica* (45).

311 Genes that were RpoS-regulated only in *S. enterica* include *glgB*, *glgC*, *glgP*, and *glgX*,
312 responsible for the biosynthesis and catabolism of glycogen. (These genes are expressed more
313 highly at 15°C than 37°C in both *E. coli* and *S. enterica* but are only RpoS-regulated in *S.*
314 *enterica*.) There were also genes that were RpoS-dependent and present only in *S. enterica*, like
315 *invA*, *invE*, *prgI*, *sipD*, and *sprB*. These genes are all located on the Salmonella Pathogenicity
316 Island 1 and are not present in *E. coli* (46). Much like the broader response to low temperature,
317 the RpoS-regulon at 15°C was mostly not conserved between *E. coli* and *Salmonella* due a mix
318 of species-specific genes and different regulation of conserved genes.

319 Discussion

320 All the species studied in these experiments responded to a shift in growth temperature
321 from 37°C to 15°C by changing the expression of a large proportion of their genes; the number
322 ranged from 626 to 1057 genes, or 14 to 26% of all genes in the genome. For *E. coli*, this was
323 more than twice the number of genes than were previously reported in microarray studies of a
324 shift from 37°C to 23°C (2, 10). The difference can probably be attributed to the combination of
325 the more extreme temperature shift in this study, and the difference in technologies used to
326 measure mRNA abundance.
327 A major finding of this study is the lack of conservation of the transcriptional response
328 to low temperature across the *Enterobacteriaceae*. All the species grew markedly slower at
329 15°C than at 37°C and all changed expression of hundreds of genes in response to the
330 temperature shift. However, the identity of these genes was mostly not conserved among
331 species. The lack of conservation of which genes were DE is due primarily to species-specific
332 transcriptional patterns of shared genes. While species-specific genes did contribute to the
333 diversity of differentially expressed genes in this group, these species-specific genes were less
334 likely to be DE than was the core genome. Horizontal gene transfer is the major way that
335 bacterial genomes become different from other each other, but the transcriptomes in this
336 study mostly became different by species-specific patterns of expression of shared genes.

337 *E. coli* responded to low temperature by changing the expression of more of its genes
338 than any of the other species studied here. It is possible that the K-12 strain of *E. coli* is more
339 lab-adapted than the other strains used in this study (47, 48) and so responds more strongly to
340 low temperature than the other strains. (The history of some of the strains used in this study is
341 not well documented.) It is also possible that K-12 strains are typical of all *E. coli* in the number
342 of genes differentially regulated by temperature shift, or that some of the other strains in this
343 study are just as lab-adapted as *E. coli* K-12. Now that we have documented extensive
344 transcriptional divergence using single strains of each species, future experiments should
345 examine multiple strains sampled from the diversity of a single species to determine the extent
346 of intraspecific variation in transcriptomic response to low temperature.
347 The lack of conservation of which genes were differentially expressed extends to a lack
348 of conservation of phenotypes as predicted by enriched GO terms. Most of the GO terms
349 enriched in a species were not shared with other species, which implies that each species in this
350 study may have different phenotypic responses to the same stress. For example, only *E. coli*
351 responded to the lower temperature by down-regulating the mRNAs of ribosomal proteins. It
352 is well known that the abundance of ribosomal proteins in *E. coli* is lower at slower growth
353 rates (49), and our data on the abundance of ribosomal protein mRNAs are consistent with this
354 finding. The other species in this study all grew more slowly at 15°C than at 37°C, and some
355 grew as slowly as *E. coli* at 15°C, but none showed the same down-regulation of ribosomal
356 protein mRNA. It is possible that these other species do not achieve down-regulation of
357 ribosomal activity at lower growth rate via the same mechanism, or do not down-regulate
358 ribosomal activity at all.
359 The alternative sigma factor RpoS is one regulatory protein that plays a role in the
360 transcriptional response to temperature shift. Much like the overall cold response, the RpoS-
361 dependent part of the cold response was only partially conserved between *E. coli* and *S.*
362 *enterica.* The divergence between the two species was due to a mix of divergent expression of
363 genes shared between the two species and species-specific genes that are part of the RpoS-
364 regulon. The divergence in RpoS-dependence of genes shared by *E. coli* and *S. enterica* was
365 somewhat surprising given that both species produce similar amounts of RpoS at 15°C.

366 Divergence in RpoS-dependence of shared genes could be due to changes in the RpoS protein
367 itself, in the promoters of these divergently expressed genes, or could be due to changes in the
368 many trans-acting factors that interact with RpoS to determine the transcriptional output of a
369 promoter (44). Understanding this molecular basis of these changes is an important future
370 area of research.
371 In other systems, species-specific genes are a substantial part of the divergence of
372 regulons between species pairs. For example, the OmpR protein regulates transcription of a
373 group of genes in response to physical stresses like osmolarity and pH. The OmpR regulon in *E.*
374 *coli* and *S. enterica* is only weakly conserved, and the authors noted that many species-specific
375 genes were in the regulon (50). The RovA protein regulates the transcription of important
376 virulence genes in *Yersinia*. The RovA-regulon of *Yersinia enterocolitica* and *Y. pestis* is also only
377 weakly conserved due at least in part to the species-specific genes regulated by RovA in each
378 species (51). The transcriptional regulator YhaJ regulates different sets of genes in strains of
379 enterohemorrhagic and uropathogenic *E. coli*, some of which are strain-specific (52, 53). This
380 paper extends the observation of regulon divergence to the RpoS-regulon and shows that
381 while species-specific genes do have a role in divergence of the regulon, shared genes are a
382 disproportionate part of the species-specific regulons.
383
384 In conclusion, we studied the transcriptional response to a change in temperature from
385 37°C to 15°C in five species of the *Enterobacteriaceae* and one outgroup. We found that there
386 was limited conservation of this response, with each species changing the transcription of
387 hundreds of genes that are unchanged in the other species. This divergence was not due
388 primarily to species-specific genes in each genome, as core genes were more likely than
389 expected to be differentially expressed in response to temperature change. This pattern of
390 evolutionary divergence was also true of RpoS-regulon, where divergence was due more to
391 shared genes being RpoS-regulated in only a single species than it was due to species-specific
392 genes being part of the regulon. These results highlight the importance of the evolution of
393 transcriptional patterns of shared genes in the evolution of bacterial responses to stress.
394

395 # Materials and Methods

396 ## *Strains and strain construction*

397 Strains of six species were used in this study: *E. coli* K-12, *S. enterica* sv. Typhimurium
398 14028s, *C. rodentium* DBS100, *E. cloacae* ATCC13047, *Klebsiella pneumoniae* MKP103, and
399 *Serratia marcescens* Db11. Strains with wild-type alleles for *rpoS* were used for each species.
400 In addition, strains of *E. coli and S. enterica* with existing Δ*rpoS* alleles were used (Table
401 S3). To construct the isogenic Δ*rpoS* strain of *S. enterica*, the Δ*rpoS::cm* allele of a *S. enterica*
402 14028s strain provided by the McClelland Lab (54) was moved by P22 transduction (55) into our
403 lab isolate of *S. enterica* 14028s.
404

405 ## *Growth conditions*

406 Strains were inoculated from −80°C frozen cultures and grown aerobically overnight in
407 5 mL of Luria-Bertani (LB) broth (1% tryptone, 0.5% yeast extract, 1% NaCl) at 37˚C with
408 shaking at 225 rpm. The next morning, 250µL of each culture was diluted into 25 mL of LB in a
409 250 mL borosilicate flask and grown at 37˚C in an incubator shaking at 225 rpm. When cultures
410 reached OD_{600} of 0.25 - 0.35, flasks were transferred to a water bath at 15°C shaking at 225
411 rpm. (OD_{600} of 0.3 is in the middle of exponential growth in these cultures.) When cultures were
412 grown for RNA and protein isolation, no measurements of OD_{600} were performed at
413 intermediate time points between the shift and sampling of RNA and protein. To minimize
414 confounding factors, we always grew all 8 strains in parallel for each experimental replicate.
415

416 ## *Measurement of growth parameters*

417 When cultures were grown to measure growth rate, the growth conditions mentioned
418 above were used, except that OD_{600} measurements were taken approximately every 15

419 minutes in the 37°C portion of the experiment (from the start of the experiment until an OD_{600}
420 of 0.25 – 0.35 was reached), and every 20 to 30 minutes after the shift to 15°C. (Cultures
421 reached maximum OD_{600} values of less than 1 after three hours at 15°C.) Growth rates at 37°C
422 and 15°C, as well as the length of the lag phase after the shift to 15°C were estimated using the
423 "Growth Rates Made Easy" method of Hall *et al.* (56), as implemented in the R package
424 growthrates (57). To minimize confounding factors, all 8 strains were always grown in parallel
425 for each experimental replicate.
426

427 ## *RNA Isolation*

428 For RNA isolation, 1.5 mL of bacterial culture was pelleted and resuspended in 500 µL of
429 Trizol pre-heated to 65°C and then stored at -80°C until use. Samples were purified on a
430 column using the Zymo Direct-Zol RNA Miniprep kit. They were then treated twice with DNase
431 I using Turbo DNA-free (Invitrogen) for 30 minutes each at 37˚C to ensure that all DNA was
432 degraded. The samples were further purified on a column using the Zymo RNA Clean and
433 Concentrator Kit and again stored at -80˚C until sequencing. The concentration of the samples
434 was checked using a NanoDrop spectrophotometer and degradation was checked for by
435 electrophoresing samples on an ethidium bromide agarose gel. Four biological replicate
436 samples were collected for each strain at each temperature.
437

438 ## *Generation of RNA-Seq data*

439 Illumina cDNA libraries were generated using a modified version of the RNAtag-seq
440 protocol (58), which creates tagged pools of RNA samples for cDNA synthesis and sequencing.
441 Briefly, 500 ng to 1 µg of total RNA was fragmented by heating to 94°C in NEB Alkaline
442 Phosphatase Buffer, depleted of any residual genomic DNA, dephosphorylated, and ligated to
443 DNA adapters carrying 5′-AN_8-3′ barcodes of known sequence with a 5′ phosphate and a 3′
444 blocking group. Barcoded RNAs were pooled and depleted of rRNA using the RiboZero rRNA

445 depletion kit (Epicentre). Pools of barcoded RNAs were converted to Illumina cDNA libraries in
446 2 main steps: (i) reverse transcription of the RNA using a primer designed to the constant
447 region of the barcoded adaptor with addition of an adapter to the 3′ end of the cDNA by
448 template switching using SMARTScribe (Clontech) as described (59); (ii) PCR amplification
449 using primers whose 5′ ends target the constant regions of the 3′ or 5′ adaptors and whose 3′
450 ends contain the full Illumina P5 or P7 sequences. cDNA libraries were sequenced on the
451 Illumina [NextSeq 500] platform at the Broad Institute to generate paired end reads. Each RNA
452 pool consisted of 16 samples, a 15°C and a 37°C sample from each of the 8 strains. Four total
453 pools were constructed and sequenced.
454

455 ## *RNA-seq Data Analysis*

456 Sequencing reads from each sample in a pool were demultiplexed based on their
457 associated barcode sequence using custom scripts. Up to 1 mismatch in the barcode was
458 allowed provided it did not make assignment of the read to a different barcode possible.
459 Barcode sequences were removed from the first read as were terminal G's from the second
460 read that may have been added by SMARTScribe during template switching.
461 Reads were aligned to the reference genomes (Table S1) using BWA v0.7.17 (60). Reads
462 were counted with htseq-count v2.0.1 (61). Further analysis was done in R v4.4.0 (62). Plots
463 were created with ggplot2 3.5.1 (63).
464 Before conducting the analysis reported in this paper, initial quality control was
465 performed to ensure that each sample had at least 1,000,000 read-pairs that mapped to
466 protein coding sequences. One sample below this threshold was removed. In addition, three
467 samples did not cluster with any of the other samples in PCA or a heatmap and were also
468 removed. The samples removed were one from *C. rodentium* at 15°C, two from *K. pneumoniae*
469 at 37°C, and one from *S. marcescens* at 37°C.
470 Differential expression analysis was performed using the package DESeq2 v1.40.2 (64),
471 with the analysis of each species occurring separately. Genes were considered significant if
472 they had an FDR-adjusted p-value of less than 0.01 and an absolute value of fold-change

473 greater than 2. Readers interested in repeating these analyses with different p-values and fold-
474 change can use the data and scripts in the Zenodo archive. Upset plots were created with the
475 package ComplexUpset v 1.3.3 in R (65). For the comparison of expression levels across species
476 in plots of individual genes, all samples from all species were normalized together as one data
477 set.
478 eggNOG-mapper (24) was used to assign GO terms to each gene in each genome. GO
479 enrichment for each species was performed with the R package topGO (25). The GO tree for
480 each species was pruned to remove terms with fewer than 5 genes. The weight01 algorithm
481 was used with Fisher's exact test to calculate p-values on enrichment of GO terms. GO terms
482 were considered significantly enriched if they had a p-value of less than 0.05.
483

484 *Protein Isolation & Western Blotting*

485 Total protein was isolated from the same cultures at the same time as RNA samples.
486 Samples with cell density equivalent to 1mL of OD_{600} = 0.3 were pelleted by centrifugation at
487 least 20,000 x g for 30 s, resuspended in 100µL of 2x Laemmli sample buffer (LiCor), heated at
488 95˚C for 5 minutes, and then stored at -20°C until they were used.
489 A quantitative western blot was used to measure the levels of RpoS in the cell at each
490 time point, in a procedure slightly modified from (15). 10µL of each protein sample was
491 electrophoresed on a 4 - 20% gradient polyacrylamide gel (Bio-Rad) in Tris-glycine running
492 buffer (25 mM Tris base, 250 mM glycine, 0.5% SDS) at 200V for 35 minutes at room
493 temperature. Electrophoresis was used to transfer proteins to a Immobilon-FL PVDF
494 membrane in transfer buffer (48 mM Tris base, 39 mM glycine, 0.0375% SDS, 20% methanol)
495 at 100V for 45 minutes at 4°C. To quantify the total amount of protein in the samples, the
496 membranes were stained with 5 mL of REVERT Total Protein Stain solution and imaged using
497 the LiCor Odyssey CLx imager. The membranes were blocked with 10mL of Odyssey Blocking
498 Buffer for 1 hour at room temperature.
499 RpoS was detected using an anti-RpoS monoclonal antibody (clone 1RS1, BioLegend).
500 The blocked membranes were labeled with primary antibody (0.4 µg/mL mouse anti-RpoS,

501 10mL Odyssey Blocking Buffer, 0.2% Tween 20) at 4°C with shaking at 55 rpm overnight. Once
502 labeled, the membrane was washed four times with 10 - 15mL of 1x Tris-buffered saline with
503 Tween 20 for 5 minutes each wash. The membranes were probed with a fluorescent secondary
504 antibody (IRDye 800CW goat anti-mouse, LiCor) that was diluted in 10mL of Odyssey Blocking
505 Buffer with 0.2% Tween 20 and 0.01% SDS. The membranes were incubated with the diluted
506 secondary antibody for 1 hour at room temperature and were washed as described previously,
507 with an additional wash at the end with 1x Tris-buffered saline without Tween 20. Membranes
508 were dried for 2-3 hours between 2 sheets of Whatmann 3MM blotting paper before imaging
509 on a Li-Cor Odyssey CLx imager. RpoS levels were quantified by analyzing the amount of
510 fluorescence of each band with Image Studio v 2.1 software.
511 Data analysis was performed in R v4.4.0 (62). RpoS levels were normalized by dividing
512 them by the total amount of protein at each time point, as quantified by the total protein stain.
513

514 *xenoGI analysis*

515 The program xenoGI v3.1.1 (23) was used to reconstruct the evolutionary history of the
516 genomes used in this study. The phylogeny was inferred using the built-in helper function of
517 xenoGI. This function built a gene tree for each core gene using FastTree 2.1.11 (66), then
518 reconciled all of the gene trees into a species tree using Astral v5.6.3 (67).
519

520 *Data Availability*

521 The RNA-seq FASTQ files are deposited in the Gene Expression Omnibus with
522 ascension number GSE267531.
523 All other data and scripts used for data analysis are in a Zenodo archive at DOI
524 10.5281/zenodo.11460468
525
526

Acknowledgements

RNA-Seq libraries were constructed and sequenced at the Broad Institute of MIT and Harvard by the Microbial 'Omics Core and Genomics Platform, respectively. The Microbial 'Omics Core also provided guidance on experimental design and performed the demultiplexing of RNA-Seq data. Thanks to Michael McClelland, Colin Manoil, the Salmonella Genetic Stock Center, and the *C. elegans* Stock Center for sharing strains, to Eliot Bush for help with xenoGI, and to Danae Schulz and Chris White-Ziegler and anonymous peer-reviewers for feedback and insights.

This material is based upon work supported by the National Science Foundation under Grant No. 1716794.

References

1. Zhang Y, Gross CA. 2021. Cold Shock Response in Bacteria. Annual Review of Genetics 55:377–400.

2. White-Ziegler CA, Um S, Pérez NM, Berns AL, Malhowski AJ, Young S. 2008. Low temperature (23 degrees C) increases expression of biofilm-, cold-shock- and RpoS-dependent genes in Escherichia coli K-12. Microbiology (Reading) 154:148–166.

3. Moon S, Ham S, Jeong J, Ku H, Kim H, Lee C. 2023. Temperature Matters: Bacterial Response to Temperature Change. J Microbiol 61:343–357.

4. Xia B, Ke H, Inouye M. 2001. Acquirement of cold sensitivity by quadruple deletion of the cspA family and its suppression by PNPase S1 domain in Escherichia coli. Molecular Microbiology 40:179–188.

5. Zhang Y, Burkhardt DH, Rouskin S, Li G-W, Weissman JS, Gross CA. 2018. A Stress Response that Monitors and Regulates mRNA Structure Is Central to Cold Shock Adaptation. Molecular Cell 70:274-286.e7.

6. Zangrossi S, Briani F, Ghisotti D, Regonesi ME, Tortora P, Dehò G. 2000. Transcriptional and post-transcriptional control of polynucleotide phosphorylase during cold acclimation in Escherichia coli. Molecular Microbiology 36:1470–1480.

557 7. Dersch P, Kneip S, Bremer E. 1994. The nucleoid-associated DNA-binding protein H-NS is
558 required for the efficient adaptation of Escherichia coli K-12 to a cold environment. Mol
559 Gen Genet 245:255–259.

560 8. Dash S, Palma CSD, Baptista ISC, Almeida BLB, Bahrudeen MNM, Chauhan V,
561 Jagadeesan R, Ribeiro AS. 2022. Alteration of DNA supercoiling serves as a trigger of
562 short-term cold shock repressed genes of E. coli. Nucleic Acids Research 50:8512–8528.

563 9. Sledjeski DD, Gupta A, Gottesman S. 1996. The small RNA, DsrA, is essential for the low
564 temperature expression of RpoS during exponential growth in Escherichia coli. EMBO J
565 15:3993–4000.

566 10. White-Ziegler CA, Malhowski AJ, Young S. 2007. Human Body Temperature (37°C)
567 Increases the Expression of Iron, Carbohydrate, and Amino Acid Utilization Genes in
568 Escherichia coli K-12. Journal of Bacteriology 189:5429–5440.

569 11. White-Ziegler CA, Davis TR. 2009. Genome-wide identification of H-NS-controlled,
570 temperature-regulated genes in Escherichia coli K-12. J Bacteriol 191:1106–10.

571 12. Adeolu M, Alnajar S, Naushad S, S. Gupta R. 2016. Genome-based phylogeny and
572 taxonomy of the 'Enterobacteriales': proposal for Enterobacterales ord. nov. divided into
573 the families Enterobacteriaceae, Erwiniaceae fam. nov., Pectobacteriaceae fam. nov.,
574 Yersiniaceae fam. nov., Hafniaceae fam. nov., Morganellaceae fam. nov., and
575 Budviciaceae fam. nov. International Journal of Systematic and Evolutionary
576 Microbiology 66:5575–5599.

577 13. Brenner DJ, Farmer JJ. 2015. *Enterobacteriaceae*, p. 1–24. *In* Whitman, WB (ed.), Bergey's
578 Manual of Systematics of Archaea and Bacteria, 1st ed. Wiley.

579 14. Guentzel MN. 1996. Escherichia, Klebsiella, Enterobacter, Serratia, Citrobacter, and
580 Proteus, p. . *In* Baron, S (ed.), Medical Microbiology, 4th ed. University of Texas Medical
581 Branch at Galveston, Galveston (TX).

582 15. Popoff MY, Le Minor LE. 2015. Salmonella, p. 1–1. *In* Bergey's Manual of Systematics of
583 Archaea and Bacteria. John Wiley & Sons, Ltd.

584 16. Leimbach A, Hacker J, Dobrindt U. 2013. E. coli as an All-Rounder: The Thin Line Between
585 Commensalism and Pathogenicity, p. 3–32. *In* Dobrindt, U, Hacker, JH, Svanborg, C (eds.),
586 Between Pathogenicity and Commensalism. Springer Berlin Heidelberg.

587 17. Wyres KL, Holt KE. 2018. *Klebsiella pneumoniae* as a key trafficker of drug resistance
588 genes from environmental to clinically important bacteria. Current Opinion in
589 Microbiology 45:131–139.

590 18. Ishii S, Ksoll WB, Hicks RE, Sadowsky MJ. 2006. Presence and Growth of Naturalized
591 Escherichia coli in Temperate Soils from Lake Superior Watersheds. Applied and
592 Environmental Microbiology 72:612–621.

593 19. Rivera SC, Hazen TC, Toranzos GA. 1988. Isolation of fecal coliforms from pristine sites in
594 a tropical rain forest. Appl Environ Microbiol 54:513–517.

595 20. Kröger C, Colgan A, Srikumar S, Händler K, Sivasankaran SK, Hammarlöf DL, Canals R,
596 Grissom JE, Conway T, Hokamp K, Hinton JCD. 2013. An infection-relevant transcriptomic
597 compendium for Salmonella enterica Serovar Typhimurium. Cell Host Microbe 14:683–
598 695.

599 21. Tripathy S, Sen R, Padhi SK, Mohanty S, Maiti NK. 2014. Upregulation of transcripts for
600 metabolism in diverse environments is a shared response associated with survival and
601 adaptation of Klebsiella pneumoniae in response to temperature extremes. Funct Integr
602 Genomics 14:591–601.

603 22. Kumari K, Sharma PK, Singh RP. 2024. Unravelling the transcriptome response of
604 Enterobacter sp. S-33 under varying temperature. Arch Microbiol 206:81.

605 23. Liu N, Gonzalez TA, Fischer J, Hong C, Johnson M, Mawhorter R, Mugnatto F, Soh R,
606 Somji S, Wirth JS, Libeskind-Hadas R, Bush EC. 2023. xenoGI 3: using the DTLOR model
607 to reconstruct the evolution of gene families in clades of microbes. BMC Bioinformatics
608 24:295.

609 24. Cantalapiedra CP, Hernández-Plaza A, Letunic I, Bork P, Huerta-Cepas J. 2021. eggNOG-
610 mapper v2: Functional Annotation, Orthology Assignments, and Domain Prediction at
611 the Metagenomic Scale. Molecular Biology and Evolution 38:5825–5829.

612 25. Alexa A, Rahnenfuhrer J. 2023. topGO: Enrichment analysis for Gene Ontology. R
613 package version 2.54.0. (2.54.0).

614 26. Mey AR, Gómez-Garzón C, Payne SM. 2021. Iron Transport and Metabolism in
615 Escherichia, Shigella, and Salmonella. EcoSal Plus 9:eESP-0034-2020.

616 27. Holden VI, Bachman MA. 2015. Diverging roles of bacterial siderophores during infection.
617 Metallomics 7:986–995.

618 28. Steingard CH, Helmann JD. 2023. Meddling with Metal Sensors: Fur-Family Proteins as
619 Signaling Hubs. Journal of Bacteriology 205:e00022-23.

620 29. Nandal A, Huggins CCO, Woodhall MR, McHugh J, Rodríguez-Quiñones F, Quail MA,
621 Guest JR, Andrews SC. 2010. Induction of the ferritin gene (ftnA) of Escherichia coli by
622 Fe2+–Fur is mediated by reversal of H-NS silencing and is RyhB independent. Molecular
623 Microbiology 75:637–657.

624 30. Atlung T, Ingmer H. 1997. H-NS: a modulator of environmentally regulated gene
625 expression. Mol Microbiol 24:7–17.

626 31. Amit R, Oppenheim AB, Stavans J. 2003. Increased bending rigidity of single DNA
627 molecules by H-NS, a temperature and osmolarity sensor. Biophys J 84:2467–73.

628 32. Dorman CJ. 2004. H-NS: a universal regulator for a dynamic genome. Nature Reviews
629 Microbiology 2:391–400.

630 33. Stoebel DM, Free A, Dorman CJ. 2008. Anti-silencing: overcoming H-NS-mediated
631 repression of transcription in Gram-negative enteric bacteria. Microbiology (Reading,
632 Engl) 154:2533–2545.

633 34. Choi J, Schmukler M, Groisman EA. 2022. Degradation of gene silencer is essential for
634 expression of foreign genes and bacterial colonization of the mammalian gut.
635 Proceedings of the National Academy of Sciences 119:e2210239119.

636 35. Kim Y-H, Han KY, Lee K, Lee J. 2005. Proteome response of Escherichia coli fed-batch
637 culture to temperature downshift. Appl Microbiol Biotechnol 68:786–793.

638 36. Miajlovic H, Smith SG. 2014. Bacterial self-defence: how Escherichia coli evades serum
639 killing. FEMS Microbiology Letters 354:1–9.

640 37. Stevenson G, Andrianopoulos K, Hobbs M, Reeves PR. 1996. Organization of the
641 Escherichia coli K-12 gene cluster responsible for production of the extracellular
642 polysaccharide colanic acid. J Bacteriol 178:4885–4893.

643 38. Limoli DH, Jones CJ, Wozniak DJ. 2015. Bacterial Extracellular Polysaccharides in Biofilm
644 Formation and Function. Microbiology Spectrum 3:10.1128/microbiolspec.mb-0011–2014.

645 39. Wall E, Majdalani N, Gottesman S. 2018. The Complex Rcs Regulatory Cascade. Annu Rev
646 Microbiol 72:111–139.

647 40. Gottesman S, Stout V. 1991. Regulation of capsular polysaccharide synthesis in
648 Escherichia coli K12. Mol Microbiol 5:1599–1606.

649 41. Adams J, Hoang J, Petroni E, Ashby E, Hardin J, Stoebel DM. 2023. The timing of
650 transcription of RpoS-dependent genes varies across multiple stresses in *Escherichia coli*
651 K-12. mSystems 8:e00663-23.

652 42. Wong GT, Bonocora RP, Schep AN, Beeler SM, Fong AJL, Shull LM, Batachari LE, Dillon
653 M, Evans C, Becker CJ, Bush EC, Hardin J, Wade JT, Stoebel DM. 2017. Genome-Wide
654 Transcriptional Response to Varying RpoS Levels in Escherichia coli K-12. J Bacteriol
655 199:e00755-16.

656 43. Weber H, Polen T, Heuveling J, Wendisch VF, Hengge R. 2005. Genome-wide analysis of
657 the general stress response network in *Escherichia coli*: sigmaS-dependent genes,
658 promoters, and sigma factor selectivity. Journal of bacteriology 187:1591–603.

659 44. Bouillet S, Bauer TS, Gottesman S. 2024. RpoS and the bacterial general stress response.
660 Microbiology and Molecular Biology Reviews 0:e00151-22.

661 45. Mates AK, Sayed AK, Foster JW. 2007. Products of the *Escherichia coli* acid fitness island
662 attenuate metabolite stress at extremely low pH and mediate a cell density-dependent
663 acid resistance. J Bacteriol 189:2759–2768.

664 46. Lerminiaux NA, MacKenzie KD, Cameron ADS. 2020. Salmonella Pathogenicity Island 1
665 (SPI-1): The Evolution and Stabilization of a Core Genomic Type Three Secretion System.
666 Microorganisms 8:576.

667 47. Hobman JL, Penn CW, Pallen MJ. 2007. Laboratory strains of Escherichia coli: model
668 citizens or deceitful delinquents growing old disgracefully? Mol Microbiol 64:881–885.

669 48. Browning DF, Hobman JL, Busby SJW. 2023. Laboratory strains of Escherichia coli K-12:
670 things are seldom what they seem. Microb Genom 9:mgen000922.

671 49. Ingraham JL, Maaløe O, Neidhardt FC. 1983. Growth of the Bacterial Cell. Sinauer
672 Associates, Sunderland, MA.

673 50. Quinn HJ, Cameron ADS, Dorman CJ. 2014. Bacterial Regulon Evolution: Distinct
674 Responses and Roles for the Identical OmpR Proteins of Salmonella Typhimurium and
675 Escherichia coli in the Acid Stress Response. PLOS Genetics 10:e1004215.

676 51. Ellison DW, Hinchliffe SJ, Wren BW, Miller VL. 2007. The RovA regulons of Yersinia
677 enterocolitica and Yersinia pestis are distinct: evidence that many RovA-regulated genes
678 were acquired more recently than the core genome. Molecular Microbiology 66:189–205.

679 52. Connolly JPR, O'Boyle N, Turner NCA, Browning DF, Roe AJ. 2019. Distinct intraspecies
680 virulence mechanisms regulated by a conserved transcription factor. Proceedings of the
681 National Academy of Sciences 116:19695–19704.

682 53. O'Boyle N, Turner NCA, Roe AJ, Connolly JPR. 2020. Plastic Circuits: Regulatory Flexibility
683 in Fine Tuning Pathogen Success. Trends in Microbiology
684 https://doi.org/10.1016/j.tim.2020.01.002.

685 54. Porwollik S, Santiviago CA, Cheng P, Long F, Desai P, Fredlund J, Srikumar S, Silva CA,
686 Chu W, Chen X, Canals R, Reynolds MM, Bogomolnaya L, Shields C, Cui P, Guo J, Zheng
687 Y, Endicott-Yazdani T, Yang H-J, Maple A, Ragoza Y, Blondel CJ, Valenzuela C, Andrews-
688 Polymenis H, McClelland M. 2014. Defined Single-Gene and Multi-Gene Deletion Mutant
689 Collections in Salmonella enterica sv Typhimurium. PLOS ONE 9:e99820.

690 55. Schmieger H. 1972. Phage P22-mutants with increased or decreased transduction
691 abilities. Mol Gen Genet 119:75–88.

692 56. Hall BG, Acar H, Nandipati A, Barlow M. 2014. Growth Rates Made Easy. Molecular
693 Biology and Evolution 31:232–238.

694 57. Petzoldt T. 2022. growthrates: Estimate Growth Rates from Experimental Data.
695 https://CRAN.R-project.org/package=growthrates.

696 58. Shishkin AA, Giannoukos G, Kucukural A, Ciulla D, Busby M, Surka C, Chen J,
697 Bhattacharyya RP, Rudy RF, Patel MM, Novod N, Hung DT, Gnirke A, Garber M, Guttman
698 M, Livny J. 2015. Simultaneous generation of many RNA-seq libraries in a single reaction.
699 Nat Meth 12:323–325.

700 59. Zhu YY, Machleder EM, Chenchik A, Li R, Siebert PD. 2001. Reverse transcriptase
701 template switching: a SMART approach for full-length cDNA library construction.
702 Biotechniques 30:892–897.

703 60. Li H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-
704 MEM. arXiv:13033997 [q-bio].

705 61. Anders S, Pyl PT, Huber W. 2015. HTSeq--a Python framework to work with high-
706 throughput sequencing data. Bioinformatics 31:166–169.

707 62. R Core Team. 2023. R: A language and environment for statistical computing. R
708 Foundation for Statistical Computing, Vienna, Austria.

709 63. Wickham H. 2016. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New
710 York. https://ggplot2.tidyverse.org.

711 64. Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion
712 for RNA-seq data with DESeq2. Genome Biology 15:550.

713 65. Krassowski M. 2022. ComplexUpset. https://zenodo.org/records/7314197. Retrieved 30
714 July 2024.

715 66. Price MN, Dehal PS, Arkin AP. 2010. FastTree 2 – Approximately Maximum-Likelihood
716 Trees for Large Alignments. PLoS ONE 5:e9490.

717 67. Zhang C, Rabiee M, Sayyari E, Mirarab S. 2018. ASTRAL-III: polynomial time species tree
718 reconstruction from partially resolved gene trees. BMC Bioinformatics 19:153.

719

720

Figure legends

Figure 1: Patterns of shared and unique genome content. (A) Phylogeny of the six species in this study. The phylogeny was constructed by inferring the gene tree for each core gene by maximum likelihood with FastTree (66), followed by reconciliation of the gene trees using ASTRAL (67). The inferred phylogeny is consistent with prior systematic work with this group of organisms (12). (B) The total number of protein coding genes found in the genome of each of the six species. (C) 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.

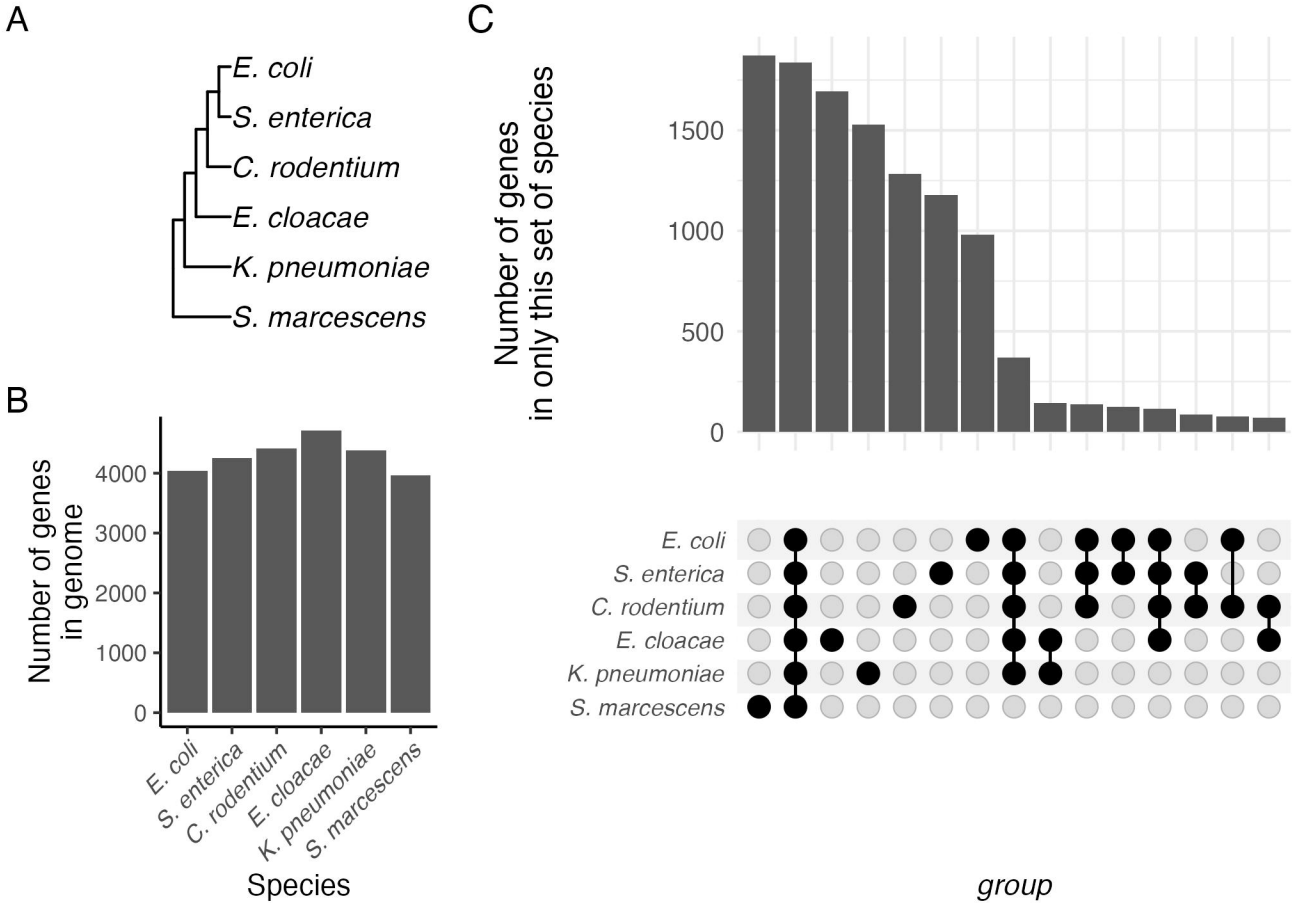
Figure 2: Patterns of differential expression in response to temperature shift of all genes in the genome. (A) Principal component analysis of expression of the 1838 core genes in all six species at 37°C and 15°C. (B) The number of significantly DE genes (FDR-adjusted $p < 0.01$, absolute $\log_2$-fold change > 2). Upregulated genes are expressed more highly at 15°C than 37°C. (C, D) Upset plot of shared and unique patterns of DE genes in these species. Bar heights represent the number of genes that are (C) upregulated or (D) downregulated, while the dots below represent the species in that group. Only sets with 15 or more genes are shown.

Figure 3: Role of core vs species-specific genes in divergence of transcriptomes. (A, B) 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)). (C, D) The number of significantly DE core genes that are (C) upregulated or (D) 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. Only sets with 15 or more genes are shown. (E) Distribution of differentially expressed genes among members of a clade. 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. (Inset) Phylogeny of the species, with nodes labels.

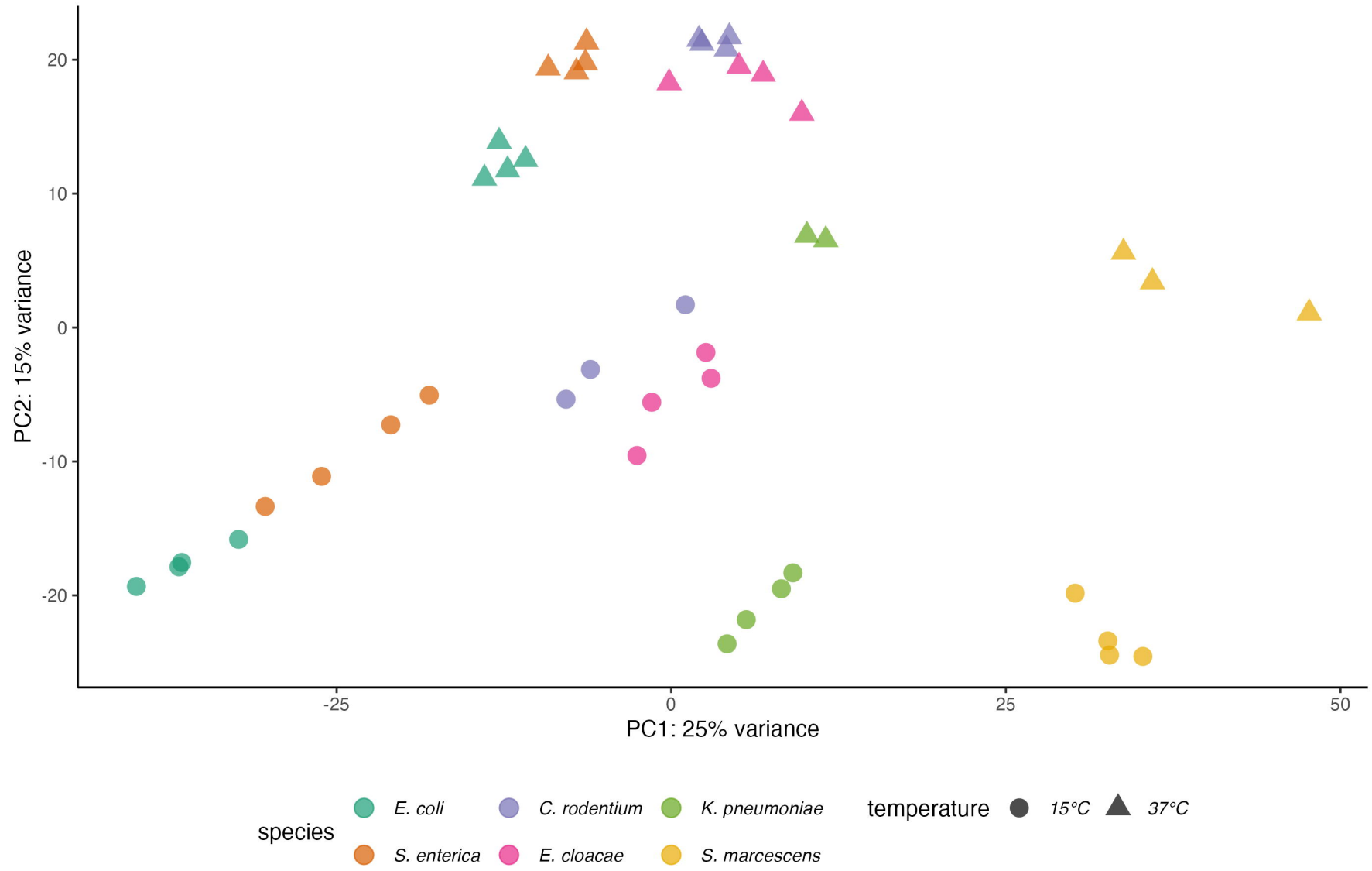
Figure 4. Shared patterns of GO enrichment. Bar heights represent the number of GO terms enriched, while the dots below represent the species involved in that comparison. All observed sets are shown.

Figure 5: Genes involved in iron metabolism change expression in the cold. (A-D) Log_2-fold change of gene expression between 15°C and 37°C for genes annotated as involved in several

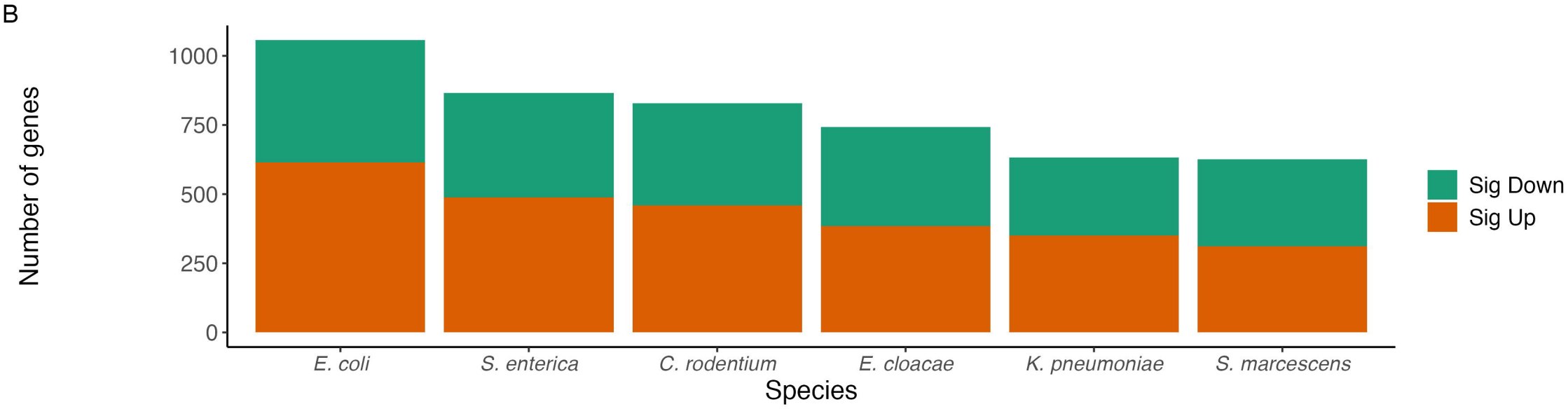
764 aspects of iron metabolism (A) enterobactin biosynthesis, (B) enterobactin transport, (C) iron-
765 sulfur cluster assembly, and (D) iron-sulfur cluster transfer. Green dots represent a single DE
766 gene, while orange dots are a gene that is not DE. Violin plots show the distributions of log_2-
767 fold changes of the individual genes. (E, F) The expression of genes coding for two global
768 regulators, *fur* and *hns,* previously identified as associated with these phenotypes. Black dots
769 represent the mean, error bars are the SEM, and gray dots are individual measurements.
770
771 **Figure 6: Upregulation of colanic acid related genes only in *C. rodentium.*** (A) Log_2-fold
772 change of gene expression between 15°C and 37°C for genes associated with production and
773 use of colanic acid. Green dots represent a single DE gene, while orange dots are a gene that is
774 not DE. Violin plots show the distributions of log_2-fold changes of the individual genes. (B, C)
775 Expression of two known regulators of this process, *rcsA* and *rcsB*. Black dots represent the
776 mean, error bars are the SEM, and gray dots are individual measurements.
777
778 **Figure 7: Partial overlap of the RpoS regulon of *E. coli* and *S. enterica* at 15°C.** (A) Principal
779 component analysis of expression of the 1838 core genes in wild-type and Δ*rpoS E. coli* and *S.*
780 *enterica* at 37°C and 15°C. (B) Size of the RpoS regulon in *E. coli* and *S. enterica.* Bar lengths are
781 the number of RpoS-regulated genes in that species, with colors denoting the status of those
782 genes in the other species. For example, the orange part of the *E. coli* bar represents the 99
783 genes that are RpoS-regulated in *E. coli* that are also present in *S. enterica* but are not RpoS-
784 regulated in *S. enterica.*
785
786
787


A
E. coli
S. enterica
C. rodentium
E. cloacae
K. pneumoniae
S. marcescens
B
Number of genes in genome
4000
3000
2000
1000
0
E. coli
S. enterica
C. rodentium
E. cloacae
K. pneumoniae
S. marcescens
Species
C
Number of genes in only this set of species
1500
1000
500
0
E. coli
S. enterica
C. rodentium
E. cloacae
K. pneumoniae
S. marcescens
group

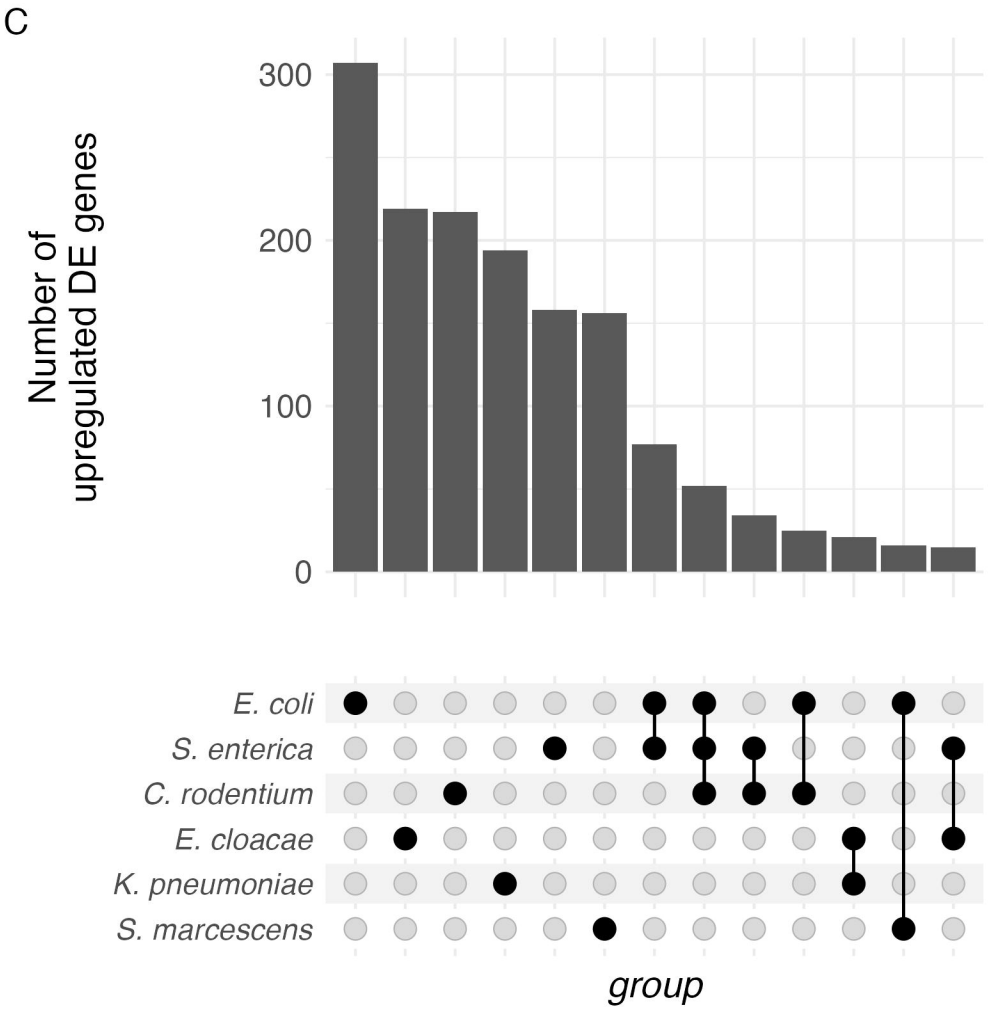
A


PC2: 15% variance
PC1: 25% variance
20
10
0
-10
-20
-25
0
25
50
species
E. coli
S. enterica
C. rodentium
E. cloacae
K. pneumoniae
S. marcescens
temperature
15°C
37°C

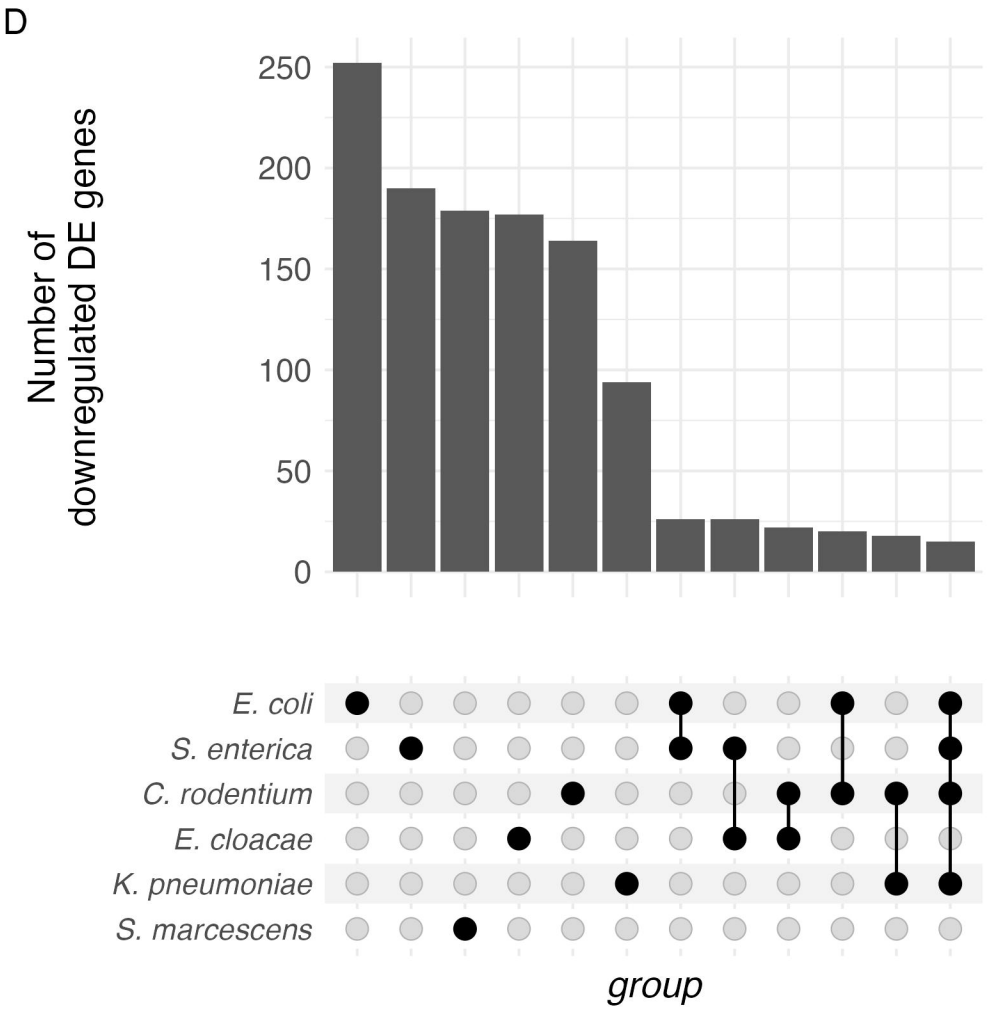
B

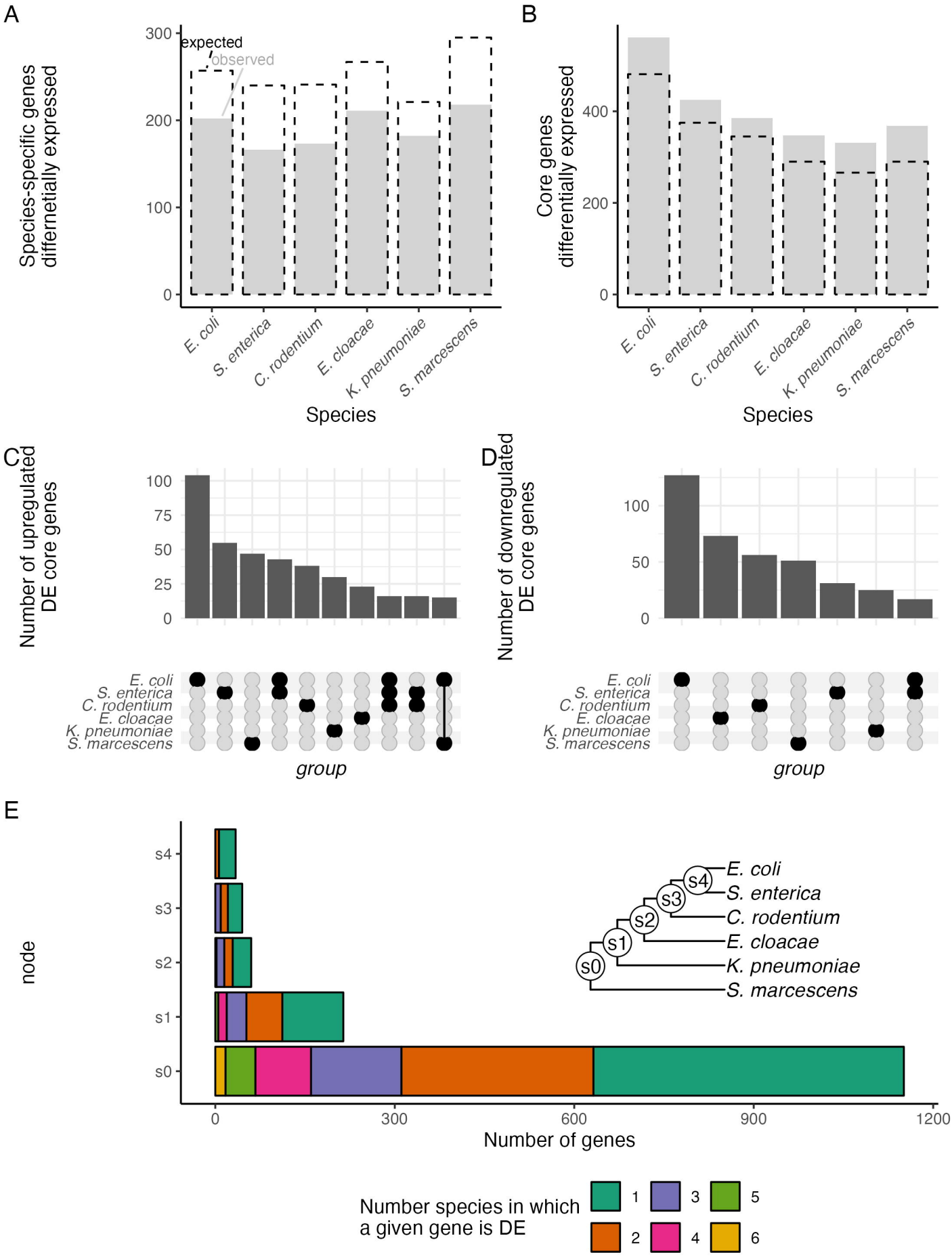

Number of genes
1000
750
500
250
0
E. coli
S. enterica
C. rodentium
E. cloacae
K. pneumoniae
S. marcescens
Species
Sig Down
Sig Up

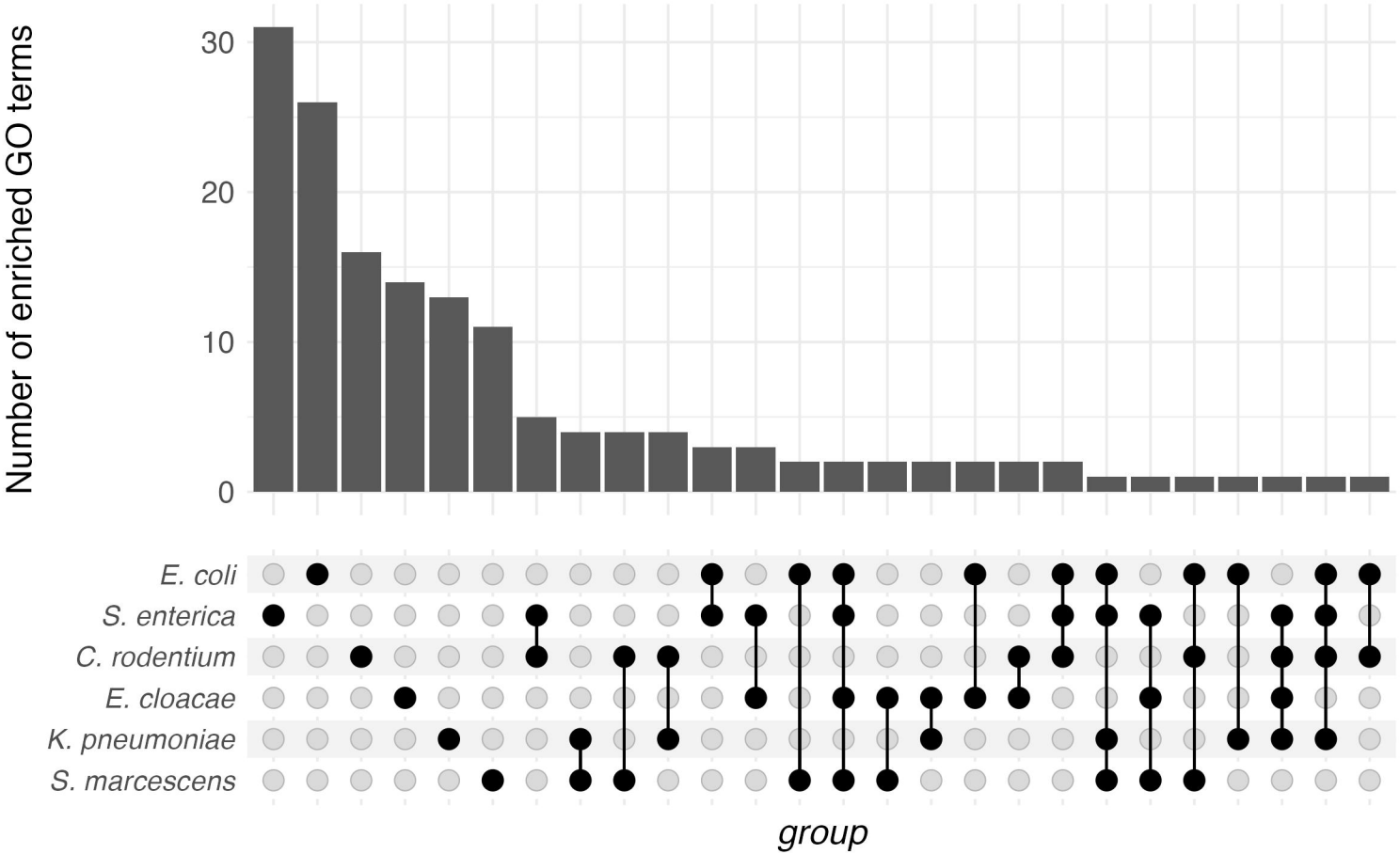
C

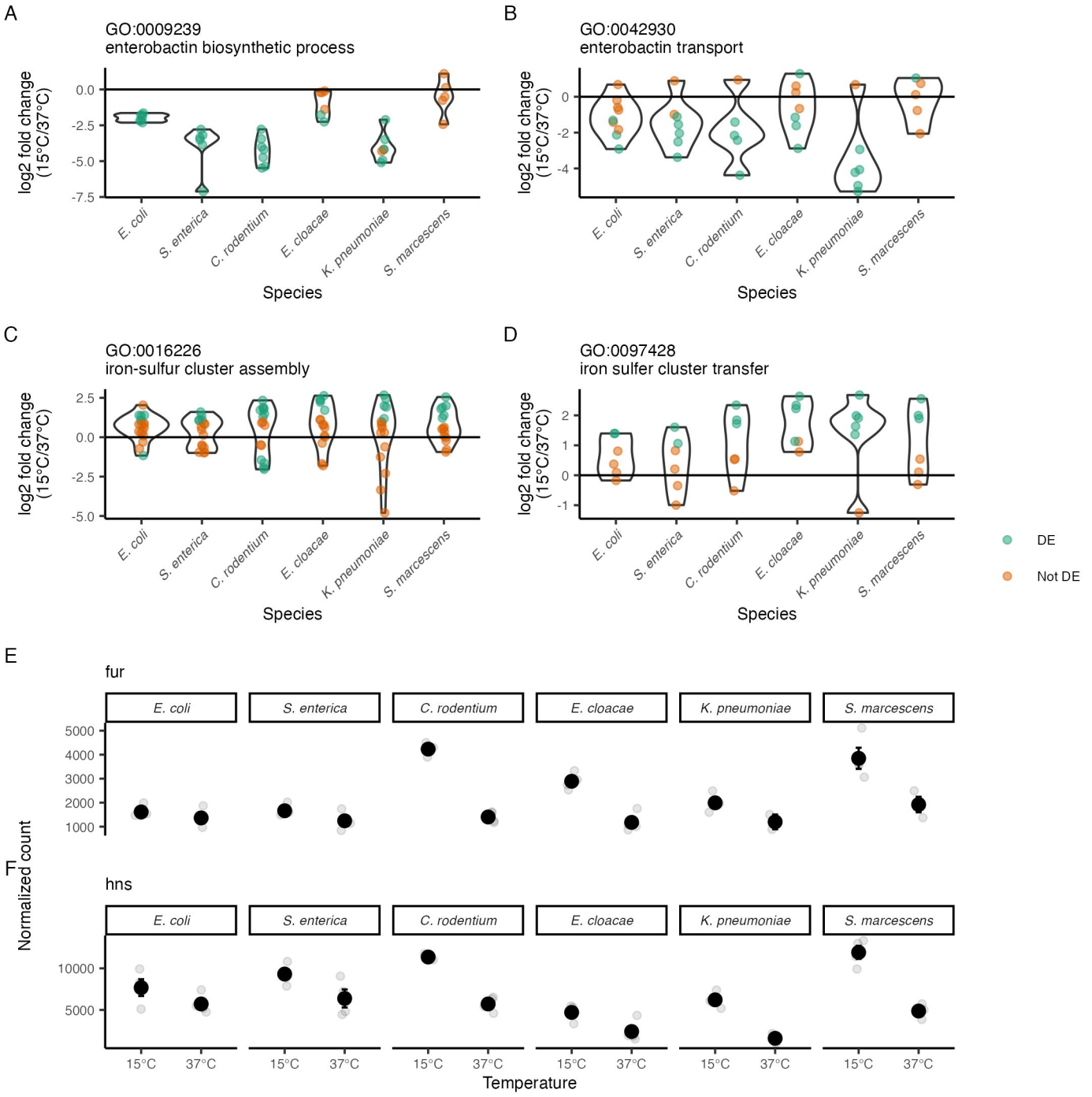

Number of upregulated DE genes
300
200
100
0
E. coli
S. enterica
C. rodentium
E. cloacae
K. pneumoniae
S. marcescens
group

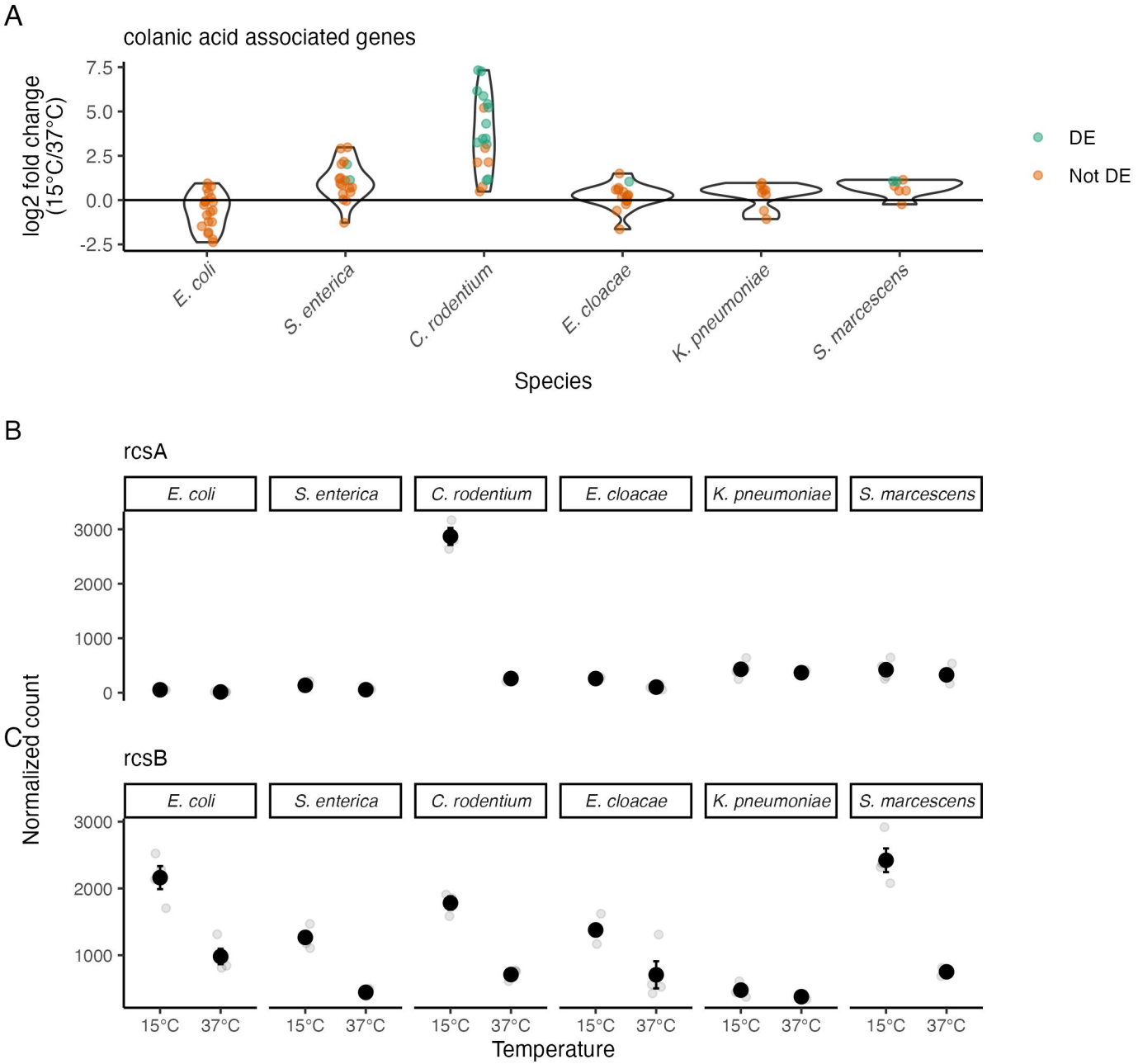
D

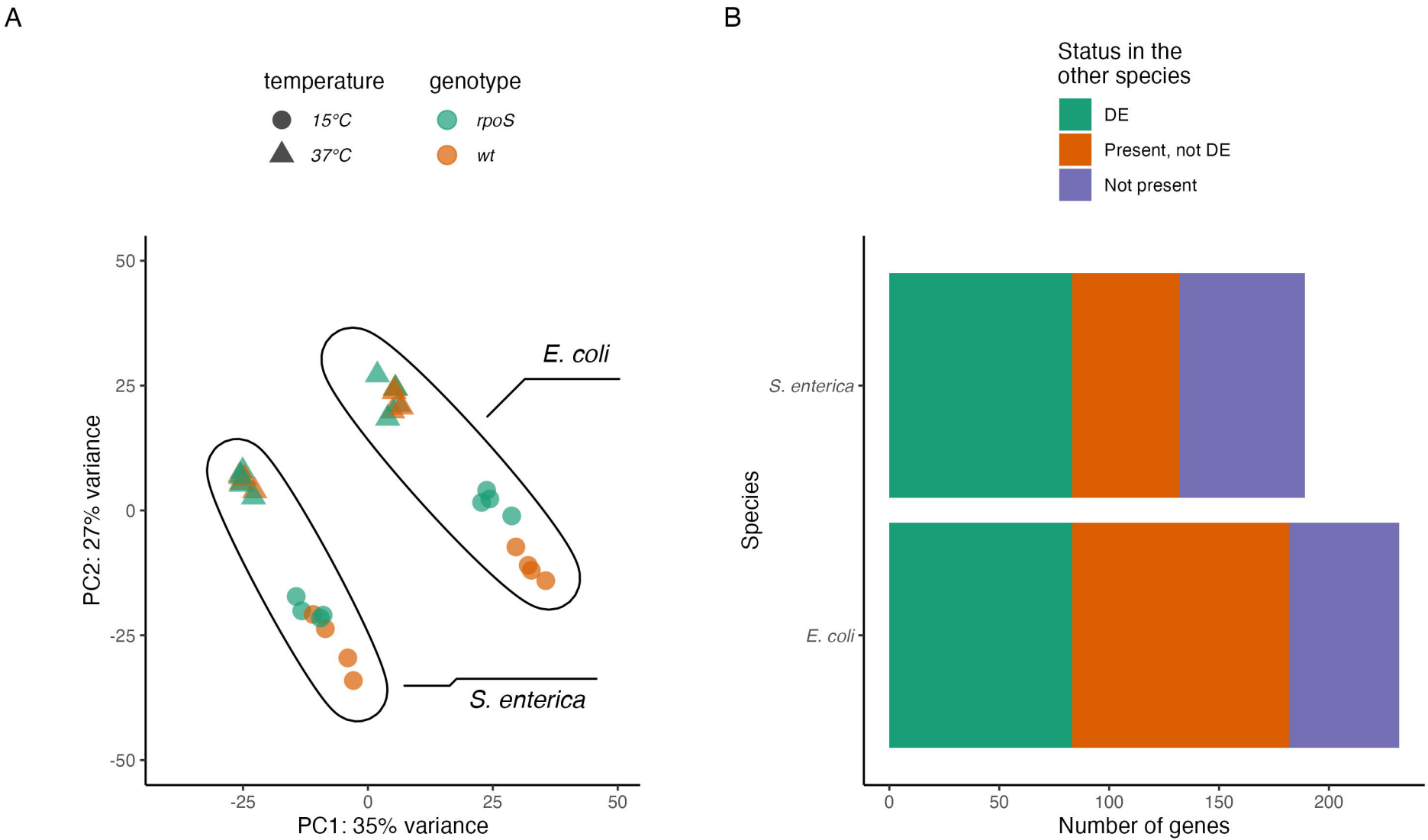

Number of downregulated DE genes
250
200
150
100
50
0
E. coli
S. enterica
C. rodentium
E. cloacae
K. pneumoniae
S. marcescens
group


A
Species-specific genes differnetially expressed
expected
observed
300
200
100
0
E. coli
S. enterica
C. rodentium
E. cloacae
K. pneumoniae
S. marcescens
Species
B
Core genes differentially expressed
400
200
0
E. coli
S. enterica
C. rodentium
E. cloacae
K. pneumoniae
S. marcescens
Species
C
Number of upregulated DE core genes
100
75
50
25
0
E. coli
S. enterica
C. rodentium
E. cloacae
K. pneumoniae
S. marcescens
group
D
Number of downregulated DE core genes
100
50
0
E. coli
S. enterica
C. rodentium
E. cloacae
K. pneumoniae
S. marcescens
group
E
node
s4
s3
s2
s1
s0
0
300
600
900
1200
Number of genes
s4
s3
s2
s1
s0
E. coli
S. enterica
C. rodentium
E. cloacae
K. pneumoniae
S. marcescens
Number species in which a given gene is DE
1
2
3
4
5
6


Number of enriched GO terms
30
20
10
0
E. coli
S. enterica
C. rodentium
E. cloacae
K. pneumoniae
S. marcescens
group


A
GO:0009239
enterobactin biosynthetic process
log2 fold change (15°C/37°C)
0.0
-2.5
-5.0
-7.5
E. coli
S. enterica
C. rodentium
E. cloacae
K. pneumoniae
S. marcescens
Species
B
GO:0042930
enterobactin transport
log2 fold change (15°C/37°C)
0
-2
-4
Species
C
GO:0016226
iron-sulfur cluster assembly
log2 fold change (15°C/37°C)
2.5
0.0
-2.5
-5.0
Species
D
GO:0097428
iron sulfer cluster transfer
log2 fold change (15°C/37°C)
2
1
0
-1
Species
DE
Not DE
E
fur
5000
4000
3000
2000
1000
F
hns
10000
5000
Normalized count
15°C
37°C
Temperature


A
colanic acid associated genes
log2 fold change (15°C/37°C)
7.5
5.0
2.5
0.0
-2.5
E. coli
S. enterica
C. rodentium
E. cloacae
K. pneumoniae
S. marcescens
Species
DE
Not DE
B
rcsA
E. coli
S. enterica
C. rodentium
E. cloacae
K. pneumoniae
S. marcescens
3000
2000
1000
0
Normalized count
C
rcsB
3000
2000
1000
15°C
37°C
Temperature


A
temperature
15°C
37°C
genotype
rpoS
wt
E. coli
S. enterica
PC2: 27% variance
PC1: 35% variance
50
25
0
-25
-50
-25
0
25
50
B
Status in the other species
DE
Present, not DE
Not present
Species
S. enterica
E. coli
0
50
100
150
200
Number of genes