

1 **The genome-wide transcriptional response to varying RpoS levels in *Escherichia***  
2 ***coli* K-12.**

3

4 Garrett T. Wong<sup>1,2</sup>, Richard P. Bonocora<sup>3</sup>, Alicia N. Schep<sup>1,4</sup>, Suzannah M. Beeler<sup>1,5</sup>,

5 Anna Lee<sup>1,6</sup>, Lauren Shull<sup>1,7</sup>, Lakshmi Batachari<sup>1</sup>, Moira Dillon<sup>1</sup>, Ciaran Evans<sup>8,9</sup>,

6 Johanna Hardin<sup>8</sup>, Joseph T. Wade<sup>3,10,#</sup>, and Daniel M. Stoebel<sup>1,#</sup>

7

8 1 Department of Biology, Harvey Mudd College, Claremont CA

9 2 Current address: Departments of Neuroscience, Genetics and Genome Sciences,

10 Icahn School of Medicine at Mount Sinai.

11 3 Wadsworth Center, New York State Department of Health, Albany NY

12 4 Current address: Department of Genetics, Stanford University

13 5 Current address: Division of Biology and Biological Engineering, California Institute of  
14 Technology

15 6 Current address: a-tek, Washington DC

16 7 Current address: Sackler School of Graduate Biomedical Sciences, Tufts University

17 8 Department of Mathematics, Pomona College, Claremont CA

18 9 Current address: Department of Statistics, Carnegie Mellon University

19 10 Department of Biomedical Sciences, School of Public Health, University at Albany,  
20 SUNY, Albany NY

21

22 # Address correspondence to Joseph T. Wade, [joseph.wade@health.ny.gov](mailto:joseph.wade@health.ny.gov), or Daniel

23 M. Stoebel, [stoebel@g.hmc.edu](mailto:stoebel@g.hmc.edu)

24

25 Running title: Effect of RpoS concentration on *E. coli* transcriptome

26

27        **Abstract**

28        The alternative sigma factor RpoS is a central regulator of a many stress responses in  
29        *Escherichia coli*. The level of functional RpoS differs depending on the stress. The effect  
30        of these differing concentrations of RpoS on global transcriptional responses remains  
31        unclear. We investigated the effect of RpoS concentration on the transcriptome during  
32        stationary phase in rich media. We show that 23% of genes in the *E. coli* genome are  
33        regulated by RpoS level, and we identify many RpoS-transcribed genes and promoters.  
34        We observe three distinct classes of response to RpoS by genes in the regulon: genes  
35        whose expression changes linearly with increasing RpoS level, genes whose  
36        expression changes dramatically with the production of only a little RpoS (“sensitive”  
37        genes), and genes whose expression changes very little with the production of a little  
38        RpoS (“insensitive”). We show that sequences outside the core promoter region  
39        determine whether a RpoS-regulated gene is sensitive or insensitive. Moreover, we  
40        show that sensitive and insensitive genes are enriched for specific functional classes,  
41        and that the (in)sensitivity of a gene to RpoS corresponds to the timing of induction as  
42        cells enter stationary phase. Thus, promoter sensitivity to RpoS is a mechanism to  
43        coordinate specific cellular processes with growth phase, and may also contribute to the  
44        diversity of stress responses directed by RpoS.

45

46        **Importance**

47        The sigma factor RpoS is a global regulator that controls the response to many  
48        stresses in *Escherichia coli*. Different stresses result in different levels of RpoS

49 production, but the consequences of this variation are unknown. We describe how  
50 changing the level of RpoS does not influence all genes equally. The cause of this  
51 variation is likely the action of transcription factors that bind the promoters of the genes.  
52 We show that the sensitivity of a gene to RpoS levels explains the timing of expression  
53 as cells enter stationary phase, and that genes with different RpoS sensitivities are  
54 enriched for specific functional groups. Thus, promoter sensitivity to RpoS is a  
55 mechanism to coordinate specific cellular processes in response to stresses.

56

## 57 **Introduction**

58

59 Genome-wide measurements of RNA levels have revolutionized our understanding  
60 of how cells organize their patterns of transcription. These studies have given us  
61 snapshots of how patterns of gene expression change in response to changes in the  
62 external environment. They have also allowed us to define the regulons controlled by  
63 specific transcription factors. A major weakness of the vast majority of these studies is  
64 that they explore the function of a regulatory protein only by comparing expression of  
65 target genes in a wild-type strain to a gene knock-out, or to a mutant with a single  
66 diminished or increased level of activity. While some genetic regulatory networks are  
67 certainly switch-like (1, 2) and can be fully characterized by only two levels of activity,  
68 many other regulatory proteins vary continuously in their abundance and/or activity.  
69 How regulons respond to a range of regulator levels is a largely unstudied question (3).  
70 Should we expect all genes in a regulon to increase or decrease expression by the  
71 same relative amount following a change in abundance of a regulatory protein? Or

72 should we expect genes to respond in different ways? These questions motivate this  
73 study.

74 A paradigmatic example of a bacterial regulatory protein whose abundance and  
75 activity vary continuously in response to different conditions is the alternative sigma  
76 factor RpoS in *Escherichia coli*. Transcription by RNA polymerase containing RpoS is  
77 responsible for the general stress response (4–6). Under conditions of optimal growth in  
78 the laboratory (such as exponential phase growth in rich media at 37°C), RpoS levels  
79 are nearly undetectable in the model *E. coli* strain K-12. As conditions become poorer  
80 for growth, either because cells begin to starve for various nutrients, or because they  
81 face physical challenges such as low temperature or elevated osmolarity, RpoS levels  
82 rise (4–6). RpoS coordinates the transcription of genes that are critical for the response  
83 to these stresses.

84 RpoS expression is not an all-or-none phenomenon. For example, RpoS levels rise  
85 continuously as cells transition from exponential growth to stationary phase (7).  
86 Moreover, starvation for different nutrients upregulates RpoS to differing levels (8, 9).  
87 This level of control over RpoS levels is accomplished by regulating transcription,  
88 translation, and degradation (4, 7, 10), allowing for careful control over protein levels. In  
89 addition to regulation of protein abundance, RpoS activity can also be directly  
90 modulated by a number of factors, such as CrI (11).

91 Not only do RpoS levels vary across conditions for a single strain, but different  
92 strains of *E. coli* also differ in their patterns of expression of RpoS. For example,  
93 naturally occurring strains can differ in the amount of RpoS they produce during

94 exponential phase (12) or stationary phase (13). All studies that have measured RpoS  
95 levels in naturally occurring strains of *E. coli* have detected variation (13–17), though the  
96 extent and cause of this variation in RpoS between strains is still a matter of some  
97 controversy (16, 17).

98       Microarray studies (18–20) have shown that RpoS controls the expression of at least  
99 500 genes (over 10% of the genome) either directly or indirectly, but the set of RpoS-  
100 regulated genes differs across environmental conditions. For most RpoS-controlled  
101 genes, it is not known whether the gene is directly transcribed by RpoS or regulated  
102 indirectly, as a consequence of RpoS transcribing other genes. Previous studies have  
103 not investigated the impact of changes in RpoS levels on the RpoS regulon, or whether  
104 quantitative differences in RpoS levels between environmental conditions influence the  
105 observed differences in what genes are RpoS-regulated. It is clear that *E. coli* has a  
106 complicated regulatory network to fine-tune RpoS levels to different conditions, but we  
107 do not yet know the consequences of this regulation.

108       In this study, we tested the hypothesis that members of the RpoS regulon vary in  
109 their response to RpoS levels. Using a combination of ChIP-seq and RNA-seq, we  
110 identified RpoS-regulated genes, and we showed that genes vary in their sensitivity to  
111 RpoS levels in a manner dependent on sequences outside of the core promoter region.  
112 Sensitivity of genes to RpoS levels corresponds to the order in which genes are induced  
113 during the transition into stationary phase, and genes with different levels of sensitivity  
114 are enriched for specific functional groups. Thus, the level of sensitivity of genes to  
115 RpoS controls the physiological response to different stress conditions.

## 116 **Materials and Methods**

117

### 118 *Culture conditions*

119 Cells were grown in 5mL of LB broth (1% tryptone, 0.5% yeast extract, 1% NaCl) in  
120 150 x 18 mm tubes shaking at 225 rpm at 37°C in a water bath unless otherwise  
121 specified in the text. When required, antibiotics were used at the following final  
122 concentrations: ampicillin at 100  $\mu\text{g/ml}$  (for plasmids) or 25  $\mu\text{g/ml}$  (for chromosomal  
123 integration); chloramphenicol at 20  $\mu\text{g/ml}$ ; kanamycin at 50  $\mu\text{g/ml}$ .

124

### 125 *Strains and genetic manipulations*

126 Strains and plasmids used in this study are listed in Table S1. The wild-type genetic  
127 background for this all experiments except for ChIP-seq is BW27786, a strain designed  
128 to give a graded transcriptional response to increasing arabinose concentration (21). To  
129 create a strain of this background lacking *rpoS*, the  $\Delta rpoS746::kan$  allele of JW5437  
130 (22) was moved by P1 transduction into BW27786, creating strain DMS2545.

131 The arabinose-inducible RpoS strain was created by PCR amplifying the *kan* gene  
132 and the  $P_{araB}$  promoter of plasmid pAH150 (23) using primers ParaBRpoSRecomb-F  
133 and ParaBRpoSRecomb-R (Table S2). This PCR product was then integrated into the  
134 *nlpD* gene (i.e. 5' of *rpoS*) in a MG1655 background using plasmid pKD46 (24), and P1  
135 transduced into BW27786, creating strain DMS2564. This strain thus lacks both the  
136 native transcriptional and translational control of RpoS.

137 For flow cytometry experiments, plasmid pDMS123 (25), which contains a *otsB-gfp*  
138 transcriptional fusion, was transformed into DMS2564 by the method of Groth et al. (26).

139       ChIP-seq experiments used strain RPB104, an unmarked derivative of MG1655 that  
140 expresses a C-terminally SPA-tagged derivative of RpoS from its native locus. This  
141 strain was constructed by P1 transduction of kan<sup>R</sup>-linked *rpoS*-SPA from a previously  
142 described strain (27). The kan<sup>R</sup> cassette was removed using the pCP20 plasmid that  
143 encodes Flp recombinase (24).

144

#### 145 *Construction and chromosomal integration of lacZ fusions*

146       The *gadB* and *astC* plasmids were built by standard cloning methods. The *gadC*  
147 and *astA* promoter regions with transcription factor binding sites were PCR-amplified  
148 with primers *gadC*promoter +/- and *astA*promoter +/- (Table S2), which included *KpnI*  
149 and *EcoRI* restriction sites for cloning. Cloning of core promoter regions was performed  
150 by annealing oligonucleotides designed to contain the whole RpoS binding region, as  
151 predicted by Fraley *et al.* (28) and Castanie-Cornet and Foster (29). Oligonucleotides  
152 were annealed by heating 1  $\mu$ M of forward and reverse primer for one minute at 100°C  
153 with 5 mM MgCl<sub>2</sub> and 7 mM Tris-HCl and then cooling slowly to room temperature.  
154 Inserts and plasmid (pLFX) were digested with *EcoRI*-HF and *KpnI*-HF (NEB), ligated  
155 with T7 ligase (NEB) and cloned into strain BW23473. Transformants were minipreped  
156 and inserts were verified by Sanger sequencing.

157       The *gadA* and *hdeA* plasmids were built using Gibson assembly (30) with the  
158 NEBuilder HiFi Assembly kit (New England Biolabs). The *gadA* and *hdeA* promoter  
159 regions were PCR amplified with primers *hdeA*HiFi+/- and *gadA*HiFi+/- . The core  
160 promoter was cloned with the single long oligo *hdeA*coreHiFi or *gadA*coreHiFi, as



161 predicted by Arnqvist et al (31) and De Biase et al (32). PCR products or  
162 oligonucleotides were mixed with pLFX digested with *KpnI*-HF and *EcoRI*-HF, and  
163 assembled according to the manufacturers' instructions. Mixtures were cloned into  
164 strain BW23473, transformants were minipreped, and inserts were verified by Sanger  
165 sequencing.

166 *lacZ* fusion plasmids were integrated into strain DMS2564 with helper plasmid  
167 pPFINT (33). Single-copy integrants were confirmed using the PCR assay of Haldimann  
168 and Wanner (23).

169

#### 170 *Quantitative Western Blotting*

171 Quantitative western blotting was used to measure RpoS levels. Cells were  
172 inoculated from frozen cultures into 5 mL of LB, and grown overnight at 37 °C, shaken  
173 at 225 rpm. 5  $\mu$ L of this overnight culture was diluted into 5 mL of LB with the  
174 appropriate concentration of arabinose, and grown for 20 h. 100  $\mu$ L of overnight culture  
175 to be assayed was centrifuged and resuspended in 1X Laemmli sample buffer (Sigma-  
176 Aldrich) and boiled for 5 minutes. Samples were diluted 1:10 in 1X Laemmli, and 10  $\mu$ L  
177 was electrophoresed on a 10% polyacrylamide gel (Bio-Rad) in tris-glycine running  
178 buffer (25 mM Tris base, 250 mM glycine, 0.5% SDS) at 100 V for 90 min at room  
179 temperature. Proteins were transferred to an Immobilon-FL PVDF membrane by  
180 electrophoresis at 100 V for 45 min at 4 °C in transfer buffer (48 mM Tris base, 39 mM  
181 glycine, 20% methanol, and 0.0375% SDS). Membranes were blocked by overnight  
182 incubation in Odyssey blocking buffer (Li-Cor) at 4 °C.

183           The blocked membrane was probed with affinity-purified monoclonal antibodies  
184 to RpoS (clone 1RS1) and RpoD (clone 2G10) (NeoClone) at a final concentration of 0.4  
185  $\mu\text{g} / \text{mL}$  in Odyssey blocking buffer plus 0.2% Tween 20 for 1 hour at room temperature.  
186 The membrane was washed four times for five minutes each with 15 mL of 1X TBST. A  
187 fluorescent secondary antibody (IRDye 800CW goat anti-mouse, Li-Cor) was diluted  
188 1:10,000 in a solution of Odyssey Blocking Buffer plus 0.2% Tween 20 and 0.01% SDS  
189 and incubated in the dark for 1 hour at room temperature. The membrane was washed  
190 as before, dried for 2 hours between sheets of Whatmann 3MM blotting paper, and  
191 imaged on a LiCor Clx fluorescent imager.

192           Band intensity was estimated using Image Studio 2.1 (LiCor). RpoS levels were  
193 divided by RpoD levels to normalize for differences in total protein levels. The ratio of  
194 RpoS to RpoD is biologically meaningful because this ratio, rather than RpoS level  
195 alone, dictates levels of transcription from RpoS-dependent promoters due to sigma  
196 factor competition (34).

197

#### 198           *RNA-seq experiments and analysis*

199           Cells were inoculated from frozen cultures into 5 mL of LB, and grown overnight. 5  
200  $\mu\text{L}$  of this overnight culture was diluted into 5 mL of LB and grown for 20 hours. (For the  
201 26% RpoS condition, a final concentration of  $10^{-4}\%$  arabinose was also added.) RNA  
202 was purified from 200  $\mu\text{L}$  of overnight culture by pelleting and resuspending in 500  $\mu\text{L}$  of  
203 Trizol at 65 °C, followed by purification on a column (Direct-Zol, Zymo Research).  
204 Samples received two 30-minute DNase treatments using TURBO DNA-free (Ambion)

205 following the manufacturer's instructions. RNA samples were then purified on a column  
206 (RNA Clean & Concentrator, Zymo Research). Samples were stored at -80 °C until  
207 used. Three samples were prepped from each culture and pooled to generate sufficient  
208 RNA. Two biological replicates were prepared for each strain or condition of interest.  
209 rRNA depletion, cDNA synthesis, library preparation, and sequencing were performed  
210 by a commercial provider (Otogenetics, Norcross, GA). Paired-end, 100 bp sequences  
211 were generated for 7-15 million reads per sample.

212 Before reads were mapped, the first ten base pairs of each read were trimmed using  
213 FASTX-Toolkit 0.0.13. Reads were mapped to the NCBI K12 reference genome  
214 (NC\_000913.2 Escherichia coli str. K-12 substr. MG1655) using BWA v0.7.5a (35). The  
215 number of read pairs mapped to each gene was counted with HTSeq 0.6.1 (36).  
216 Differential expression analysis was performed with DESeq v2.13 (37). All p-values  
217 were first FDR adjusted using the procedure of Benjamini and Hochberg (38). p-values  
218 were then further Bonferroni-adjusted for the three comparisons between pairs of RpoS  
219 levels. All differential expression p-values reported in this paper reflect both the FDR  
220 and Bonferroni-adjustments.

221 To determine if a gene differed significantly from the null expectation of linearity, we  
222 calculated the probability of the observed read count value at 26% RpoS if the true  
223 expression level was given by the linear prediction between the endpoints. Note that to  
224 calculate both the expected read count and the probability of the observed value under  
225 the null (i.e. the p-value) our model required estimating both the DEseq size factor (for  
226 scaling) and the dispersion (a variance factor) for each of the samples. The negative

227 binomial probability model (routinely used to measure count data, (e.g. (37)) was used,  
228 with the size factors and dispersion estimated from DESeq (37) to calculate the  
229 probability of the observed read count at 26% RpoS.

230 GO term analysis was performed using the topGO package (39) together with the  
231 org.EcK12.eg.db annotation package (40) in R 3.1.0. Enrichment was assessed using  
232 the weight01 algorithm (39) together with the Fisher's-exact test. The GO hierarchy was  
233 pruned to include only nodes with at least five associated genes, as significance tests  
234 can be unstable for GO terms with fewer genes (39).

235 A Venn diagram of the number of significant genes in each condition was prepared  
236 with EulerAPE 3.0.0 (41).

237

### 238 *ChIP-seq analysis*

239 *E. coli* strain RPB104 (MG1655 with C-terminally SPA-tagged *rpoS*) was grown  
240 overnight in M9 minimal medium with 0.4% glycerol at 30 °C and then subcultured 1:100  
241 in the same medium, and grown for 60 hours to saturation (OD<sub>600</sub> of ~3). ChIP-seq  
242 using the M2 monoclonal anti-FLAG antibody was performed as described previously  
243 (42). Regions of enrichment ("peaks") were identified as described previously (43).  
244 Relative enrichment was reported as a "Fold Above Threshold" (FAT) score.

245 We used MEME-ChIP (Version 4.11.2; default parameters) to analyze enriched  
246 regions identified by ChIP-seq (44) Regions within 100 bp were merged. The reported  
247 sequence motif was identified by MEME (Version 4.11.2) (45), run within the MEME-  
248 ChIP environment.

249 We used MEME (45) within the MEME-ChIP environment (Version 4.11.2; default  
250 parameters except that only the given strand was analyzed) to identify enriched  
251 sequence motifs in regions surrounding TSSs associated with RpoS ChIP-seq peaks.  
252 We analyzed sequences from -45 to +5 relative to each TSS.

253 We identified directly RpoS-transcribed genes by requiring that (i) the gene start is  
254 within 300 bp of a ChIP-seq peak, (ii) the gene is positively regulated by RpoS, as  
255 determined by RNA-seq, (iii) no other positively regulated gene starts within 300 bp of  
256 the ChIP-seq peak, (iv) there is no associated sequence motif or TSS (identified in (46))  
257 that would be consistent with transcription in the opposite orientation. For regulated  
258 genes, we determined whether other genes in the same operon were also regulated,  
259 using a published operon list (47) for *E. coli*. Peaks were associated with transcription  
260 start sites from a previous study (46) by identifying transcription start sites within 20 bp  
261 of a peak.

262

### 263 *QPCR*

264 RNA was isolated as for RNA-seq, except that samples underwent three 30-  
265 minute DNase treatments using TURBO DNA-free (Ambion). cDNA was made from the  
266 RNA samples using SuperScript VILO master mix (Invitrogen) and stored at -20 °C for  
267 use in quantitative real-time PCR (qPCR) reactions.

268 qPCR was performed using Power SYBR Green master mix (Invitrogen), 2  $\mu$ L of  
269 cDNA, and primers at 300 nM. Cycling was performed at 95 °C for 10 min; and 40  
270 cycles of 95 °C for 15 sec, 58 °C for 15 sec, and 68 °C for 30 sec. Three control genes

271 (*ftsZ*, *pgm*, and *hemL*) were used in addition genes of interest. These genes were  
272 selected using the approach of Vandesompele et al. (48). Details of this selection  
273 process, including the other seven genes tested, are available in Supplementary Text 1,  
274 Table S3, and Figure S1.

275 For each gene, a standard curve was made using the following amounts of  
276 genomic DNA: 200 ng, 20 ng, 2 ng, 200 pg, 20 pg, and 2 pg. Genomic DNA was  
277 extracted from overnight cultures using a Puregene Kit (Qiagen), following the  
278 manufacturer's instructions. Expression levels for each gene were interpolated from the  
279 standard curve. Expression levels of experimental genes were divided by the geometric  
280 mean of the three control genes.

281 Statistical assessment of sensitivity was performed by a bootstrapping approach.  
282 Bootstrapping, rather than a parametric approach, was appropriate because the data  
283 did not conform to parametric assumptions. In this approach, the unit of resampling was  
284 the RNA isolated from the 0%, 26% and 100% RpoS conditions on an individual day.  
285 For each resampled data set, the median expression at 0% and 100% RpoS was  
286 calculated. The linear fit from the 0% RpoS median and the 100% RpoS median was  
287 then used to predict the level of expression at the intermediate level of RpoS. Repeating  
288 this re-sampling of the data 10,000 times yielded a 95% confidence interval for 26%  
289 RpoS. The observed median for 26% was then compared to the confidence interval to  
290 test for significance.

291

292 *Beta-galactosidase activity*

293           Beta-galactosidase activity was measured using the method of Miller (49). With  
294 lacZ fusion strains, the level of sensitivity was quantified, rather than only categorized as  
295 sensitive, linear, or insensitive. To quantify the level of sensitivity, for each replicate we  
296 calculated the distance between the observed expression at the intermediate RpoS  
297 concentration and the expected level based on a linear pattern, standardized by the  
298 difference in expression between high and low RpoS conditions. Testing if sensitivities  
299 were different from zero used a one-sample t-test, with p-values adjusted for multiple  
300 comparisons using Holm's sequential adjustment method (50). Testing if two  
301 sensitivities were different used a two-sample t-test with the same method of adjustment  
302 for multiple comparisons.

303

#### 304 *Analysis of published RNA-seq data*

305           We analyzed published RNA-seq data for wild-type and  $\Delta rpoS$  *E. coli* over a  
306 time-course of growth into stationary phase (47). Using normalized genome coverage  
307 information extracted from wiggle files, we calculated relative abundance for all genes at  
308 each of the four stationary phase time-points in the growth curves (the last four time-  
309 points for each strain). We arbitrarily selected a threshold coverage value of 500 and  
310 excluded any genes scoring below this threshold at all four time-points in wild-type cells.  
311 This reduced variability associated with low expression levels. We excluded any genes  
312 for which coverage at the final time-point was 0 since this would have prevented  
313 normalization. We also excluded any gene for which the first stationary phase time-point  
314 had the highest expression value of the four time-points for wild-type cells, since RpoS-

315 dependent expression of these genes is likely to be masked by other factors. We then  
316 selected genes whose expression we had found to be induced by RpoS, and separated  
317 these genes into insensitive, linear, and sensitive classes. We calculated expression  
318 levels for each of these genes relative to expression at the final time-point.

319

320

## 321 **Results**

322

323 *The RpoS regulon in late stationary phase.*

324 To understand the role of the RpoS protein in late stationary phase, we used RNA-  
325 seq to compare the transcriptome of wild-type and  $\Delta rpoS$  cells. We observed differential  
326 expression of 1044 genes (23% of genes) between these two conditions ( $p < 0.05$ ). Of  
327 the 1044 genes whose expression is influenced by RpoS, 605 are upregulated, and 439  
328 are downregulated (Table S4).

329 Influencing transcription of 23% of the genome could have many potential  
330 phenotypic effects. To better understand the function of these genes, we examined  
331 which kinds of gene functions, as described by the gene ontology (GO), are more  
332 abundant in the regulon than expected by chance. GO enrichment analysis indicates  
333 that the RpoS regulon includes many genes involved in metabolic processes (Table 1);  
334 17 of 18 significantly enriched GO terms are metabolic terms. This metabolic  
335 reorganization includes the upregulation of genes encoding glycolytic enzymes and  
336 pathways for metabolism of L-arginine to glutamine, or from L-arginine to putrescine and  
337 then into succinate. RpoS also drives the downregulation of genes involved in the TCA



338 cycle. These patterns of metabolic regulation are very similar to those identified in late  
339 stationary phase in *Salmonella enterica* (51). The only significant GO term not explicitly  
340 linked to metabolism (GO:0006970, “response to osmotic stress”) also includes  
341 metabolic genes, such as *otsA* and *otsB*, that are involved in trehalose biosynthesis.

342 Central metabolism is not the only phenotype similarly regulated by RpoS in both *S.*  
343 *enterica* and *E. coli*. Other similarities include transcription of genes involved in  
344 antioxidant activities, iron regulation, and Fe-S cluster assembly, the upregulation of  
345 proteases, and down regulation of porins. As in *S. enterica*, RpoS in *E. coli* influences  
346 the expression of many genes encoding other regulatory proteins, including *csrA*, *arcA*,  
347 *cra*, *fur*, *ihfA*, *hupA* and *hupB*. Not all of the regulation is identical between *E. coli* and *S.*  
348 *enterica*, however. For example, Lévi-Meyrueis *et al.* (51) noted that RpoS appears to  
349 direct switching between many pairs of isozymes. While some pairs show a similar  
350 pattern in *E. coli* (such as *tktA/B* and *acnA/B*), others (such as *fumA/fumC*) do not show  
351 this pattern.

352

### 353 *Genome-wide binding profile of RpoS*

354 While RNA-seq identifies which genes are regulated by RpoS, it cannot distinguish  
355 between direct and indirect effects. To determine sites where RpoS binds (and hence  
356 likely plays a direct role in transcription), we used ChIP-seq to map the association of  
357 RpoS across the *E. coli* chromosome during stationary phase growth in minimal medium.  
358 To facilitate ChIP, RpoS was C-terminally SPA-tagged at its native locus. We reasoned  
359 that RpoS would only be identified with promoter regions since it is likely released from

360 elongating RNAP complexes. We identified 284 peaks of RpoS ChIP-seq signal  
361 covering 260 genomic regions (peaks within 100 bp of each other were merged). 217 of  
362 the RpoS-bound regions are intergenic, and 67 are located within genes. We reasoned  
363 that annotated genes that are transcribed by RpoS would be positioned close to an  
364 RpoS-bound region. Consistent with this, 213 RpoS-bound regions are  $\leq 300$  bp  
365 upstream of an annotated gene start. These 213 regions include 27 that are intragenic.  
366 In 79 cases, we observed an RpoS-bound region  $\leq 300$  bp from the starts of two  
367 divergently transcribed genes.

368 We used MEME to search for enriched sequence motifs within the RpoS-bound  
369 regions. We detected a highly enriched motif in 107 regions; this motif closely  
370 resembles the known -10 hexamer recognized by RpoS (Figure 1A). Moreover,  
371 occurrences of the motif are positionally enriched with respect to the ChIP-seq peak  
372 center (Figure 1B), indicating that the data have high spatial resolution. Note that the  
373 motifs tend to be located just upstream of the peak centers (Figure 1B), as we have  
374 observed previously for the *E. coli* flagellar Sigma factor, FliA (43). This presumably  
375 reflects the fact that the footprint of initiating RNA polymerase associated with RpoS is  
376 not centered on the -10 hexamer.

377 The identification of RpoS binding sites allowed us to better understand the role of  
378 RpoS in both positive and negative regulation. While similar proportions of the RpoS  
379 regulon are positively and negatively regulated by RpoS, it is not clear if this is true at  
380 the level of direct regulation. Only 19 genes within 300 bp of a binding site are  
381 negatively regulated by RpoS. 19 of the 286 such genes is *fewer* than we would expect

382 by chance if binding sites were randomly distributed around the genome, given than 439  
383 of 4513 genes in the genome are negatively regulated. On the other hand, 111 of 286  
384 genes within 300 bp are positively regulated, a highly significant effect (Fisher's exact  
385 test,  $p < 10^{-16}$ ). Thus, the binding profile of RpoS is consistent with direct positive  
386 regulation of many genes, but provides no evidence of direct negative regulation.

387

### 388 *Identification of RpoS-transcribed genes and promoters*

389 We combined the ChIP-seq and RNA-seq data to identify genes that are directly  
390 transcribed by RNA polymerase containing RpoS. Thus, we identified 123 RpoS-  
391 transcribed genes in 99 transcripts (Table S5), and compared them to other published  
392 analyses (18–20, 46, 52–55). In some cases, we identified RpoS-bound regions  
393 upstream of genes that were not detected as being RpoS-regulated by RNA-seq. These  
394 genes may have promoters that bind transcriptionally inactive RNA polymerase (56, 57).  
395 Alternatively, the disparity between RpoS binding and regulation could be explained by  
396 differences in growth conditions between the ChIP-seq and RNA-seq experiments, or by  
397 the possibility that the C-terminal SPA tag affects the function of the C-terminus of the  
398 protein in the response to transcription activators (56, 58).

399 The high spatial resolution of sigma factor ChIP-seq can facilitate the  
400 identification of specific promoters (43, 59) when combined with nucleotide resolution  
401 transcription start site (TSS) maps. Using published TSS data from *E. coli* under  
402 stationary phase conditions similar to those used for the ChIP-seq experiment (46), we  
403 determined all pairwise distances between RpoS ChIP-seq peaks and stationary phase

404 TSSs. We observed a strong enrichment for peak-TSS distances  $\leq 20$  bp (Figure S2).  
405 We infer that these 112 TSSs are RpoS-transcribed. Consistent with this, the putative  
406 RpoS-transcribed TSSs are associated with -10 hexamers that have features expected  
407 of RpoS promoters (Figure 1C). In some cases, RpoS promoters were identified  $\leq 300$   
408 bp upstream of genes that were not RpoS-regulated in the RNA-seq data set. We  
409 presume that these are RpoS-transcribed genes that otherwise escaped detection  
410 because of differences in the growth conditions used for ChIP-seq and RNA-seq.

411

#### 412 *The transcriptome at three different RpoS levels*

413 This first view of the RpoS regulon considers RpoS as either present or absent.  
414 RpoS levels vary continuously across environmental stresses (7), so we sought to better  
415 understand how the level of RpoS in the cell influences transcription of target genes. To  
416 do this, we placed the *rpoS* gene under the control of the arabinose-inducible promoter  
417  $P_{araB}$ . This promoter was integrated just upstream of the native *rpoS* gene, placing  
418 transcription under the control of arabinose concentration and removing the 5' region  
419 that regulates translation of the native mRNA (4).

420 To measure the resulting arabinose-induced expression of RpoS, we employed  
421 quantitative western blotting. RpoS levels increased with increasing arabinose  
422 concentration, from undetectable RpoS levels, to levels similar to those in wild-type cells  
423 (Fig 2a). To confirm that expression was graded and not all-or-none in this system (21),  
424 we used flow cytometry to measure expression in individual cells. We transformed the  
425 arabinose-inducible RpoS strain with the plasmid pDMS123 (25), which contains the

426 RpoS-dependent *otsBA* promoter fused to *gfp*. As expected, *gfp* expression increased  
427 with increasing arabinose concentrations, and at each expression level the population  
428 was unimodal (Fig 2b).

429 We measured the transcriptome in cells with 26% of wild-type levels of RpoS,  
430 achieved with the addition of  $10^{-4}$  % arabinose to cells with our arabinose-inducible *rpoS*  
431 strain. Of the genes that are differentially expressed between 26% RpoS and either 0%  
432 ( $\Delta$ rpoS) or 100% (wild-type), 95% are also differentially expressed between 0% and  
433 100% (Fig 3) ( $p < 0.05$ ).

434 Nearly all genes that are significantly differentially expressed have monotonically  
435 increasing or decreasing patterns of expression between three levels of RpoS. Only  
436 two genes (*ytfR* and *ytfT*) have an expression level at 26% rpoS that is significantly  
437 higher than the expression at both 100% and 0% RpoS. The only two genes with  
438 expression lower in the 26% RpoS condition than either 100% or 0% are *nlpD* and *pcm*.  
439 These genes lie immediately upstream of *rpoS*. *nlpD* was removed from the genome  
440 during the construction the arabinose inducible RpoS strain. The *pcm* gene is still  
441 present, but the level of transcription was lowered by the genetic modification.

442

#### 443 *Classifying quantitative responses to RpoS levels*

444 To explore how genes respond to changing levels of RpoS, we developed a new  
445 metric, *sensitivity*. Our null expectation was that gene expression would increase  
446 linearly with increasing RpoS concentration. We observed many genes in our RNA-seq  
447 data set (such as *osmY*) whose expression at intermediate RpoS levels falls on, or

448 close, to a line drawn between the 0% and 100% RpoS conditions (Fig 4a). We refer to  
449 these genes as *linear* in their response to increasing RpoS levels. Other genes (such as  
450 *astA*) are transcribed more at 26% than would be expected based on their expression  
451 levels at 0% and 100% (Fig 4b). We refer to these genes as *sensitive*, because only a  
452 small amount of RpoS results in relatively high levels of transcription. In contrast, some  
453 genes (like *gadC*) are expressed at intermediate RpoS levels less than expected based  
454 on expression at 0% and 100% RpoS; such genes are referred to as *insensitive* (Fig 4c).

455 We identified 910 linear, 102 sensitive, and 32 insensitive genes. 96% of sensitive  
456 genes and 88% of insensitive genes are positively regulated by RpoS. In contrast, only  
457 53% of linear genes are positively regulated by RpoS, a significant difference (chi-  
458 square test,  $p < 0.001$ ).

459 To determine whether sensitive or insensitive (hereafter *(in)sensitive*) genes are  
460 associated with specific physiological responses to increasing RpoS levels, we again  
461 used GO enrichment. We tested the null hypothesis that the functions of these genes  
462 are a random sample from the entire RpoS regulon (not the whole genome). The GO  
463 terms significantly enriched in the sensitive class are response to osmotic stress,  
464 cellular amino acid catabolic process, and fatty acid oxidation (Table 2). Several genes  
465 encoding regulators are among the sensitive genes, including *arcA*, which encodes a  
466 global regulator of respiratory metabolism, and *rssB*, which encodes the adaptor protein  
467 required for degradation of RpoS by ClpXP.

468 GO enrichment is less useful for understanding the possible function of the  
469 insensitive gene set. Three GO terms are enriched (Table 3), but only a few genes with

470 these annotations are present in the insensitive gene set, and their enrichment probably  
471 reflects the relatively small number of insensitive genes. More strikingly, the insensitive  
472 genes include nearly all of the genes required for acid resistance system 2: the  
473 structural genes *gadA*, *gadB*, and, *gadC*, and the regulator of this system *gadE* (60, 61).  
474 In addition, the genes *yhiM*, *yhiD*, *hdeA*, *hdeB*, *hdeD*, *mdtE*, *mdtF*, all of which have  
475 been described as having roles in acid resistance (60, 61), are insensitive.

476 We used reverse transcription coupled to qPCR to confirm the expression  
477 patterns of two insensitive genes (*gadC* and *gadE*), and three sensitive genes (*prpR*,  
478 *prpD*, and *astA*). All genes were positively regulated by RpoS (Fig 5), and the median  
479 expression at 26% RpoS is consistent with RNA-seq expression patterns for all genes.  
480 We used a bootstrapping approach to assess if expression at 26% was significantly  
481 above or below the linear expectation at 26% RpoS. The expression of *gadC* was  
482 significantly insensitive ( $p < 10^{-4}$ ), and *prpR* expression was significantly sensitive ( $p <$   
483  $10^{-4}$ ). *astA* expression was marginally significantly sensitive ( $p = 0.06$  for sensitivity),  
484 while *gadE* and *prpD* were not significantly different from the linear expectation ( $p =$   
485  $0.10$  and  $p = 0.56$  respectively).

486

#### 487 *Control of (in)sensitivity of expression*

488 What makes one promoter sensitive to RpoS levels and another promoter insensitive  
489 to RpoS levels? We hypothesized three possible mechanisms. First, chromosomal  
490 location could determine the response to RpoS levels, as is known to occur in the  
491 context of total transcription levels (62). Second, it is possible that the DNA sequence of

492 the core promoter drives the response. Finally, it is possible that the binding of  
493 transcription factors upstream of the core promoter influences the response to RpoS  
494 levels. To test these hypotheses, we cloned the promoters (including all upstream  
495 transcription factor binding sites annotated in EcoCyc (63)) of four operons into the *lacZ*  
496 fusion plasmid pLFX (33). The four promoters were the sensitive *astCADBE*, and the  
497 insensitive *gadA*, *gadBC*, and *hdeAB-yhiD*. Plasmid pLFX recombines into the lambda-  
498 attachment site, placing the fusion in a novel genomic context. While we did not detect  
499 binding of RpoS upstream of *astC*, *gadA*, *gadB*, or *hdeA* by ChIP-seq, this is likely due  
500 to the difference in growth conditions, since these genes have been previously shown to  
501 be directly transcribed by RpoS (31, 64, 65).

502 The pattern of transcription of all four fusions was the same as observed for the  
503 respective genes in the RNA-seq data (Fig 6 a-d). *astC* transcription was sensitive to  
504 RpoS levels (one sample t-test,  $p = 0.04$ ), while *gadA*, *gadB*, and *hdeA* transcription  
505 were all insensitive (one sample t-test,  $p = 2 \times 10^{-6}$ ,  $p = 10^{-5}$ ,  $p = 0.04$  respectively).  
506 Since all reporters were placed at the same genomic locus, this result suggests that  
507 genomic location is not the determinant of response to RpoS levels.

508 A second potential mechanism to explain the difference between sensitive and  
509 insensitive genes is interactions between RpoS and the core promoter sequence. For  
510 example, specific nucleotides (or combinations of nucleotides) might tend to confer  
511 (in)sensitive patterns of transcription. The majority of both sensitive and insensitive  
512 genes were not associated with RpoS-bound regions in the ChIP-seq experiment,  
513 suggesting that they are indirectly regulated by RpoS (Table 4). The fact that most



514 sensitive and insensitive genes are not bound by RpoS argues against the hypothesis of  
515 direct RpoS-DNA interactions driving (in)sensitivity. To see if specific sequence motifs  
516 were consistently associated with (in)sensitivity, we used the discriminative motif search  
517 feature of DREME (66) to search for motifs that differed between sensitive and linear, or  
518 insensitive and linear regulatory sequences. There are 28 ChIP-seq peaks associated  
519 with an operon with at least one sensitive gene, and 4 ChIP-seq peaks associated with  
520 an operon with at least one insensitive gene. We found no motifs that distinguished  
521 these sets of sequences, although the small number of sequences would have  
522 restricted the power of such a test.

523 To directly test the hypothesis that the core promoter region of RpoS-transcribed  
524 genes is responsible for determining the sensitivity to RpoS levels, we cloned this short  
525 region from the *astC*, *gadA*, *gadB*, *hdeA* promoters into pLFX and recombined the  
526 plasmids into the chromosome. These core promoters had absolute levels of  
527 transcription much lower than the entire promoter that included upstream transcription  
528 factor binding sites (Fig 6e - h). The core promoter sequence for *astC* alone was  
529 somewhat less sensitive to RpoS levels than the full-length construct (two sample t-test,  
530  $p = 0.08$ ), although not significantly so, probably due to the variability of expression from  
531 the full-length *astC* reporter. The *gadA* and *gadB* core promoters differed significantly  
532 from their full-length promoters (two sample t-test,  $p = 0.01$  and  $p < 10^{-6}$ , respectively).  
533 The *hdeA* core promoter (Fig 6h) is not RpoS-dependent, showing a decline in  
534 expression of approximately 10% in the presence of RpoS. This is consistent with the  
535 previous finding that the ability of RpoD (but not RpoS) to transcribe *hdeA* is repressed

536 by H-NS protein bound upstream of the promoter (67). The core promoter construct  
537 lacks the native H-NS binding sites upstream, and so the selectivity is apparently lost.  
538 Thus, the core promoters do not replicate the RpoS sensitivity of their whole promoter  
539 sequences. In addition, the three RpoS-dependent core promoters (Fig 6 e - g) do not  
540 differ from each other in sensitivity (ANOVA,  $p = 0.32$ ). We conclude that the core  
541 promoter is not responsible for RpoS sensitivity.

542

543 *Sensitivity to RpoS determines the timing of induction during entry into stationary*  
544 *phase*

545 We hypothesized that the degree of sensitivity to RpoS could impact the timing of  
546 expression under conditions when levels of active RpoS increase, such as during entry  
547 into stationary phase. A previous study used RNA-seq to monitor the transcriptome over  
548 a time-course of growth, including four time-points in stationary phase (47). We  
549 analyzed these data to determine whether insensitive, linear, and sensitive genes show  
550 differences in the timing of induction. We selected only those RpoS-induced genes  
551 whose transcription increased upon entry into stationary phase. (In total, there were 250  
552 such linear genes, 90 sensitive, and 19 insensitive.) We then determined the pattern of  
553 expression for each such gene over four time-points beginning at the onset of stationary  
554 phase. Although there is considerable variability in the expression patterns of genes, as  
555 a group the three classes show clear differences in the timing of induction (Figure 7a).  
556 Specifically, sensitive genes are induced most rapidly, followed closely by linear genes.  
557 In both cases, expression peaked early in stationary phase and fell between 30 and 180

558 minutes after entering stationary phase. In contrast, insensitive genes showed relatively  
559 little change in expression until the final time-point, 180 minutes into stationary phase.  
560 To determine the importance of RpoS on the patterns of gene expression, we repeated  
561 the above analysis using data generated from a  $\Delta rpoS$  strain. As expected, the large  
562 difference in timing between the groups of genes was greatly diminished (Figure 7b).  
563 We conclude that the (in)sensitivity of a gene is associated with the timing of expression  
564 during stationary phase in an RpoS-dependent manner.

565

## 566 **Discussion**

567

### 568 *Expanding the known set of RpoS-transcribed genes and promoters*

569 The *E. coli* RpoS regulon has been widely investigated using targeted and genome-  
570 scale approaches. Most genome-scale studies have focused on genes whose  
571 expression is altered in the absence of RpoS (18–20, 53, 54). Hence, these studies  
572 cannot distinguish between genes that are transcribed by RpoS, and those that are  
573 indirectly regulated. ChIP-seq affords a high resolution view of RpoS binding. By  
574 combining ChIP-seq with RNA-seq, we have identified 123 RpoS-transcribed genes with  
575 high confidence, considerably expanding the known RpoS regulon. Previous studies  
576 have suggested a role for RpoS in direct repression of some target genes (46, 56, 68).  
577 While we observed negative regulation of 439 genes by RpoS, there were fewer of  
578 these repressed genes associated with ChIP-seq peaks than expected by chance,  
579 suggesting that direct negative regulation by RpoS is rare.

580 Only two previous studies have used ChIP methods to map RpoS binding genome-  
581 wide in *E. coli*. The first used ChIP-chip to identify 868 RpoS-bound regions (46), many  
582 more identified in our study but with considerably lower resolution (median peak length  
583 of 324 bp for RpoS ChIP-chip). The second used ChIP-seq, but identified relatively few  
584 RpoS-bound regions (52). Of the 63 RpoS-bound regions identified in that study, 41 are  
585 shared with those from our study.

586 The high resolution of ChIP-seq allowed us to identify specific promoter sequences  
587 recognized by RpoS. By combining ChIP-seq data with a TSS map, we identified many  
588 high-confidence RpoS promoters. These promoters are strongly enriched for the  
589 presence of a -10 hexamer, with sequence preferences consistent with several of the  
590 previously described features of RpoS promoters (69). Specifically, we observed a  
591 preference for a C at position -8 (within the -10 hexamer), a C at position -13  
592 (immediately upstream of the -10 hexamer), and a TAA a positions -6 to -4 (immediately  
593 downstream of the -10 hexamer). Previous studies have suggested that RpoS  
594 promoters often contain a -35 hexamer (70), although the spacing relative to the -10  
595 hexamer is considerably more variable than for  $\sigma^{70}$  promoters. However, we did not  
596 detect enrichment of a -35 hexamer-like sequence among RpoS promoters, suggesting  
597 that the requirement for this element is weak.

598

599 *Sensitivity to RpoS affects timing of expression, and groups functionally related*  
600 *genes*

601 As is true for many transcription factors, RpoS levels vary continuously across a  
602 wide range of conditions. Our data show that genes differ in the sensitivity of their  
603 response to RpoS levels. Moreover, whether a gene is (in)sensitive to RpoS levels is  
604 associated with its function, suggesting a physiological rationale for sensitivity. For  
605 example, insensitive genes include many of those involved in the glutamate-dependent  
606 acid resistance 2 system (AR2). These genes are of particular interest because AR2  
607 allows *E. coli* to survive a pH of 2, an important trait for its ability to pass through the  
608 stomach and colonize the gastrointestinal tract (71, 72). To some extent, the shared  
609 sensitivity of functionally related genes can be explained by operon structure, i.e. co-  
610 transcription of multiple functionally related genes from a single promoter. However, the  
611 phenomenon of shared sensitivity for functionally related genes extends beyond  
612 operons. For example, the insensitive genes involved in AR2 are transcribed from at  
613 least five different promoters (63).

614 The fact that functionally related genes often have similar patterns of sensitivity to  
615 RpoS suggests that sensitivity can serve as a mechanism to control the timing of gene  
616 expression, and hence to coordinate specific cellular processes as part of a response to  
617 environmental stresses. Consistent with this idea, we have shown that the sensitivity of  
618 genes to RpoS levels correlates with the timing of their expression. RpoS sensitivity  
619 may drive similar patterns of expression in response to other stresses. Different  
620 environmental stresses are known to upregulate RpoS to varying levels (8), suggesting  
621 that some insensitive genes may only be expressed under certain stresses. In addition  
622 to the effects on gene expression, sensitivity to RpoS may also impact the effects of

623 mutations in *rpoS*, which have been seen to evolve in the lab (73, 74). We expect  
624 mutations attenuating RpoS to have the strongest effect on insensitive genes, and the  
625 weakest effect on sensitive genes.

626

### 627 *Possible mechanisms of RpoS sensitivity*

628 While the connection between RpoS sensitivity and the timing of gene expression is  
629 clear, the molecular basis of sensitivity is less so. Our data indicate that the genomic  
630 location of these operons does not determine the expression pattern. Moreover, several  
631 lines of evidence suggest that direct interactions between RpoS and the core promoter  
632 are also not responsible for determining sensitivity. First, analysis of the ChIP-seq data  
633 for the sensitive and insensitive genes finds no motif that distinguishes between them.  
634 Second, core promoters from both sensitive and insensitive genes do not replicate the  
635 pattern of expression of the full-length promoters. Third, the core promoters of a  
636 sensitive operon (*astC*) and two insensitive operons (*gadA* and *gadB*) have  
637 indistinguishable patterns of sensitivity, suggesting that what was excluded from those  
638 constructs (i.e., binding sites of regulatory proteins) determines the shape of the  
639 relationship.

640 Given our finding that core promoter sequences cannot explain the difference in  
641 (in)sensitivity between promoters, we suggest that sensitivity is largely due to the action  
642 of specific regulatory proteins bound upstream. If this hypothesis is correct, it could also  
643 explain the physiological coherence of these groups. For example, many insensitive  
644 genes are involved in the AR2 phenotype and are also regulated by GadX, GadW, and

645 GadE (75–77). If one or more of these three regulators was directly responsible for the  
646 insensitive pattern of expression, then this could help to explain the physiological  
647 coherence of the insensitive group. The sensitive genes, being a larger group, have no  
648 obvious single regulator, although relatively little is known about regulators that function  
649 in stationary phase.

650 RpoS responds to a wide variety of environmental cues, and regulates genes  
651 responsible for many different kinds of responses. This work has demonstrated that one  
652 facet of that response, the level of RpoS produced, has varying effects across the  
653 entirety of the regulon. The level of RpoS produced in a stress response, together with  
654 the action of other transcription factors, may help to tune the RpoS-dependent stress  
655 response in ways appropriate for individual stresses.

656

### 657 **Acknowledgements**

658 We thank Rachael Kretsch for experimental help, and Robert Drewell, Xuelin Wu,  
659 Jae Hur, Keith Derbyshire and Todd Gray for helpful discussions. This work was  
660 supported by HHMI Undergraduate Science Education award #52007544 to Harvey  
661 Mudd College and by the NIH Director's New Innovator Award Program,  
662 1DP2OD007188 (JTW).

663

### 664 **Literature Cited**

- 665 1. **Novick A, Weiner M.** 1957. ENZYME INDUCTION AS AN ALL-OR-NONE  
666 PHENOMENON. Proc Natl Acad Sci U S A **43**:553–66.

- 667 2. **Ptashne M.** 2004. A Genetic Switch, Third Edition: Phage Lambda Revisited 3rd  
668 edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
- 669 3. **Brinsmade SR, Alexander EL, Livny J, Stettner AI, Segrè D, Rhee KY,**  
670 **Sonenshein AL.** 2014. Hierarchical expression of genes controlled by the *Bacillus*  
671 *subtilis* global regulatory protein CodY. *Proc Natl Acad Sci* **111**:8227–8232.
- 672 4. **Battesti A, Majdalani N, Gottesman S.** 2011. The RpoS-Mediated General Stress  
673 Response in *Escherichia coli*. *Annu Rev Microbiol* 189–213.
- 674 5. **Hengge R.** 2011. Stationary-Phase Gene Regulation in *Escherichia coli*, p. . *In* A.  
675 Böck, R. Curtiss III, J. B. Kaper, P. D. Karp, F. C. Neidhardt, T. Nyström, J. M.  
676 Slauch, C. L. Squires, D. Ussery (eds.), *EcoSal - Escherichia coli and Salmonella:*  
677 *Cellular and Molecular Biology.* Washington, DC, ASM Press.
- 678 6. **Schellhorn HE.** 2014. Elucidating the function of the RpoS regulon. *Future*  
679 *Microbiol* **9**:497–507.
- 680 7. **Lange R, Hengge-Aronis R.** 1994. The cellular concentration of the sigma S  
681 subunit of RNA polymerase in *Escherichia coli* is controlled at the levels of  
682 transcription, translation, and protein stability. *Genes Dev* **8**:1600–1612.
- 683 8. **Mandel MJ, Silhavy TJ.** 2005. Starvation for Different Nutrients in *Escherichia Coli*  
684 Results in Differential Modulation of RpoS Levels and Stability. *J Bacteriol*  
685 **187**:434–442.



- 686 9. **Zafar MA, Carabetta VJ, Mandel MJ, Silhavy TJ.** 2014. Transcriptional occlusion  
687 caused by overlapping promoters. *Proc Natl Acad Sci* **111**:1557–1561.
- 688 10. **Hengge R.** 2009. Proteolysis of sigmaS (RpoS) and the general stress response in  
689 *Escherichia coli*. *Res Microbiol* **160**:667–76.
- 690 11. **Pratt LA, Silhavy TJ.** 1998. Crl stimulates RpoS activity during stationary phase.  
691 *Mol Microbiol* **29**:1225–1236.
- 692 12. **Hryckowian AJ, Battesti A, Lemke JJ, Meyer ZC, Welch RA.** 2014. IraL Is an  
693 RssB Anti-adaptor That Stabilizes RpoS during Logarithmic Phase Growth in  
694 *Escherichia coli* and *Shigella*. *mBio* **5**:e01043–14.
- 695 13. **Chiang SM, Dong T, Edge TA, Schellhorn HE.** 2011. Phenotypic diversity  
696 caused by differential RpoS activity among environmental *Escherichia coli* isolates.  
697 *Appl Environ Microbiol* **77**:7915–23.
- 698 14. **Bhagwat AA, Tan J, Sharma M, Kothary M, Low S, Tall BD, Bhagwat M.** 2006.  
699 Functional heterogeneity of RpoS in stress tolerance of enterohemorrhagic  
700 *Escherichia coli* strains. *Appl Environ Microbiol* **72**:4978–86.
- 701 15. **Ferenci T, Galbiati HF, Betteridge T, Phan K, Spira B.** 2011. The constancy of  
702 global regulation across a species: the concentrations of ppGpp and RpoS are  
703 strain-specific in *Escherichia coli*. *BMC Microbiol* **11**:62.

- 704 16. **Snyder E, Gordon DM, Stoebel DM.** 2012. Escherichia coli Lacking RpoS Are  
705 Rare in Natural Populations of Non-Pathogens. *G3 GenesGenomesGenetics*  
706 **2**:1341–1344.
- 707 17. **Bleibtreu A, Clermont O, Darlu P, Glodt J, Branger C, Picard B, Denamur E.**  
708 2014. The rpoS Gene Is Predominantly Inactivated during Laboratory Storage and  
709 Undergoes Source-Sink Evolution in Escherichia coli Species. *J Bacteriol*  
710 **196**:4276–4284.
- 711 18. **Patten CL, Kirchhof MG, Schertzberg MR, Morton R a, Schellhorn HE.** 2004.  
712 Microarray analysis of RpoS-mediated gene expression in Escherichia coli K-12.  
713 *Mol Genet Genomics MGG* **272**:580–91.
- 714 19. **Weber H, Polen T, Heuveling J, Wendisch VF, Hengge R.** 2005. Genome-wide  
715 analysis of the general stress response network in Escherichia coli: sigmaS-  
716 dependent genes, promoters, and sigma factor selectivity. *J Bacteriol* **187**:1591–  
717 603.
- 718 20. **Dong T, Schellhorn HE.** 2009. Control of RpoS in global gene expression of  
719 Escherichia coli in minimal media. *Mol Genet Genomics* **281**:19–33.
- 720 21. **Khlebnikov A, Datsenko KA, Skaug T, Wanner BL, Keasling JD.** 2001.  
721 Homogeneous expression of the PBAD promoter in Escherichia coli by constitutive  
722 expression of the low-affinity high-capacity AraE transporter. *Microbiology*  
723 **147**:3241–3247.

- 724 22. **Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko KA,**  
725 **Tomita M, Wanner BL, Mori H.** 2006. Construction of Escherichia coli K-12 in-  
726 frame, single-gene knockout mutants: the Keio collection. Mol Syst Biol  
727 **2:2006.0008.**
- 728 23. **Haldimann A, Wanner BL.** 2001. Conditional-replication, integration, excision, and  
729 retrieval plasmid-host systems for gene structure-function studies of bacteria. J  
730 Bacteriol **183:6384–93.**
- 731 24. **Datsenko KA, Wanner BL.** 2000. One-step inactivation of chromosomal genes in  
732 Escherichia coli K-12 using PCR products. Proc Natl Acad Sci U S A **97:6640–**  
733 **6645.**
- 734 25. **Stoebel DM, Hokamp K, Last MS, Dorman CJ.** 2009. Compensatory evolution of  
735 gene regulation in response to stress by Escherichia coli lacking RpoS. PLoS  
736 Genet **5:e1000671.**
- 737 26. **Groth D, Reszka R, Schenk JA.** 1996. Polyethylene glycol-mediated  
738 transformation of Escherichia coli is increased by room temperature incubation.  
739 Anal Biochem **240:302–304.**
- 740 27. **Butland G, Peregrín-Alvarez JM, Li J, Yang W, Yang X, Canadien V,**  
741 **Starostine A, Richards D, Beattie B, Krogan N, Davey M, Parkinson J,**  
742 **Greenblatt J, Emili A.** 2005. Interaction network containing conserved and  
743 essential protein complexes in Escherichia coli. Nature **433:531–537.**

- 744 28. **Fraley CD, Kim JH, McCann MP, Matin A.** 1998. The Escherichia coli Starvation  
745 Gene *cstC* Is Involved in Amino Acid Catabolism. *J Bacteriol* **180**:4287–4290.
- 746 29. **Castanie-Cornet M-P, Foster JW.** 2001. Escherichia coli acid resistance: cAMP  
747 receptor protein and a 20 bp cis-acting sequence control pH and stationary phase  
748 expression of the *gadA* and *gadBC* glutamate decarboxylase genes. *Microbiology*  
749 **147**:709–715.
- 750 30. **Gibson DG, Young L, Chuang R-Y, Venter JC, Hutchison CA, Smith HO.** 2009.  
751 Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat*  
752 *Methods* **6**:343–345.
- 753 31. **Arnqvist A, Olsén A, Normark S.** 1994. Sigma S-dependent growth-phase  
754 induction of the *csgBA* promoter in Escherichia coli can be achieved in vivo by  
755 sigma 70 in the absence of the nucleoid-associated protein H-NS. *Mol Microbiol*  
756 **13**:1021–1032.
- 757 32. **De Biase D, Tramonti a, Bossa F, Visca P.** 1999. The response to stationary-  
758 phase stress conditions in Escherichia coli: role and regulation of the glutamic acid  
759 decarboxylase system. *Mol Microbiol* **32**:1198–211.
- 760 33. **Edwards AN, Patterson-Fortin LM, Vakulskas CA, Mercante JW, Potrykus K,**  
761 **Vinella D, Camacho MI, Fields JA, Thompson SA, Georgellis D, Cashel M,**  
762 **Babitzke P, Romeo T.** 2011. Circuitry Linking the Csr and Stringent Response  
763 Global Regulatory Systems. *Mol Microbiol* **80**:1561–1580.

- 764 34. **Farewell A, Kvint K, Nyström T.** 1998. Negative regulation by RpoS: a case of  
765 sigma factor competition. *Mol Microbiol* **29**:1039–51.
- 766 35. **Li H, Durbin R.** 2009. Fast and accurate short read alignment with Burrows–  
767 Wheeler transform. *Bioinformatics* **25**:1754–1760.
- 768 36. **Anders S, Pyl PT, Huber W.** 2014. HTSeq - A Python framework to work with  
769 high-throughput sequencing data. *bioRxiv*.
- 770 37. **Anders S, Huber W.** 2010. Differential expression analysis for sequence count  
771 data. *Genome Biol* **11**:R106.
- 772 38. **Benjamini Y, Hochberg Y.** 1995. Controlling the False Discovery Rate: A Practical  
773 and Powerful Approach to Multiple Testing. *J R Stat Soc Ser B Methodol* **57**:289–  
774 300.
- 775 39. **Alexa A, Rahnenfuhrer J.** 2010. topGO: Enrichment analysis for Gene Ontology.  
776 R package version 2.22.0.
- 777 40. **Carlson M.** org.EcK12.eg.db: Genome wide annotation for E coli strain K12. R  
778 package version 3.0.0.
- 779 41. **Micallef L, Rodgers P.** 2014. eulerAPE: Drawing Area-Proportional 3-Venn  
780 Diagrams Using Ellipses. *PLoS ONE* **9**:e101717.

- 781 42. **Singh SS, Singh N, Bonocora RP, Fitzgerald DM, Wade JT, Grainger DC.** 2014.  
782 Widespread suppression of intragenic transcription initiation by H-NS. *Genes Dev*  
783 **28**:214–219.
- 784 43. **Fitzgerald DM, Bonocora RP, Wade JT.** 2014. Comprehensive mapping of the  
785 *Escherichia coli* flagellar regulatory network. *PLoS Genet* **10**:e1004649.
- 786 44. **Machanick P, Bailey TL.** 2011. MEME-ChIP: motif analysis of large DNA datasets.  
787 *Bioinformatics* **27**:1696–1697.
- 788 45. **Bailey TL, Elkan C.** 1994. Fitting a mixture model by expectation maximization to  
789 discover motifs in biopolymers. *Proc Int Conf Intell Syst Mol Biol ISMB Int Conf*  
790 *Intell Syst Mol Biol* **2**:28–36.
- 791 46. **Cho B-K, Kim D, Knight EM, Zengler K, Palsson BO.** 2014. Genome-scale  
792 reconstruction of the sigma factor network in *Escherichia coli*: topology and  
793 functional states. *BMC Biol* **12**:4.
- 794 47. **Conway T, Creecy JP, Maddox SM, Grissom JE, Conkle TL, Shadid TM,**  
795 **Teramoto J, San Miguel P, Shimada T, Ishihama A, Mori H, Wanner BL.** 2014.  
796 Unprecedented high-resolution view of bacterial operon architecture revealed by  
797 RNA sequencing. *mBio* **5**:e01442–01414.
- 798 48. **Vandesompele J, Preter KD, Pattyn F, Poppe B, Roy NV, Paepe AD,**  
799 **Speleman F.** 2002. Accurate normalization of real-time quantitative RT-PCR data

- 800 by geometric averaging of multiple internal control genes. *Genome Biol*  
801 **3**:research0034.
- 802 49. **Miller JH**. 1992. A short course in bacterial genetics. Cold Spring Harbor  
803 Laboratory Press.
- 804 50. **Holm S**. 1979. A simple sequentially rejective multiple test procedure. *Scand J Stat*  
805 **6**:65–70.
- 806 51. **Lévi-Meyrueis C, Monteil V, Sismeiro O, Dillies M-A, Monot M, Jagla B,**  
807 **Coppée J-Y, Dupuy B, Norel F**. 2014. Expanding the RpoS/σS-Network by RNA  
808 Sequencing and Identification of σS-Controlled Small RNAs in Salmonella. *PLoS*  
809 *One* **9**:e96918.
- 810 52. **Peano C, Wolf J, Demol J, Rossi E, Petiti L, De Bellis G, Geiselmann J, Egli T,**  
811 **Lacour S, Landini P**. 2015. Characterization of the Escherichia coli σS core  
812 regulon by Chromatin Immunoprecipitation-sequencing (ChIP-seq) analysis. *Sci*  
813 *Rep* **5**.
- 814 53. **Lacour S, Landini P**. 2004. SigmaS-dependent gene expression at the onset of  
815 stationary phase in Escherichia coli: function of sigmaS-dependent genes and  
816 identification of their promoter sequences. *J Bacteriol* **186**:7186–7195.
- 817 54. **Dong T, Kirchhof MG, Schellhorn HE**. 2008. RpoS regulation of gene expression  
818 during exponential growth of Escherichia coli K12. *Mol Genet Genomics* **279**:267–277.  
819

- 820 55. **Salgado H, Peralta-Gil M, Gama-Castro S, Santos-Zavaleta A, Muñiz-Rascado**  
821 **L, García-Sotelo JS, Weiss V, Solano-Lira H, Martínez-Flores I, Medina-Rivera**  
822 **A, Salgado-Osorio G, Alquicira-Hernández S, Alquicira-Hernández K, López-**  
823 **Fuentes A, Porrón-Sotelo L, Huerta AM, Bonavides-Martínez C, Balderas-**  
824 **Martínez YI, Pannier L, Olvera M, Labastida A, Jiménez-Jacinto V, Vega-**  
825 **Alvarado L, Del Moral-Chávez V, Hernández-Alvarez A, Morett E, Collado-**  
826 **Vides J.** 2013. RegulonDB v8.0: omics data sets, evolutionary conservation,  
827 regulatory phrases, cross-validated gold standards and more. *Nucleic Acids Res*  
828 **41:D203–213.**
- 829 56. **Rosenthal AZ, Kim Y, Gralla JD.** 2008. Poising of Escherichia coli RNA  
830 Polymerase and Its Release from the [sigma] 38 C-Terminal Tail for osmY  
831 Transcription. *J Mol Biol* **376:938–949.**
- 832 57. **Reppas NB, Wade JT, Church GM, Struhl K.** 2006. The Transition between  
833 Transcriptional Initiation and Elongation in E. coli Is Highly Variable and Often Rate  
834 Limiting. *Mol Cell* **24:747–757.**
- 835 58. **Huo Y-X, Rosenthal AZ, Gralla JD.** 2008. General stress response signalling:  
836 unwrapping transcription complexes by DNA relaxation via the sigma38 C-terminal  
837 domain. *Mol Microbiol* **70:369–378.**



- 838 59. **Bonocora RP, Smith C, Lapierre P, Wade JT.** 2015. Genome-Scale Mapping of  
839 *Escherichia coli*  $\sigma$  54 Reveals Widespread, Conserved Intragenic Binding. *PLOS*  
840 *Genet* **11**:e1005552.
- 841 60. **Mates AK, Sayed AK, Foster JW.** 2007. Products of the *Escherichia coli* acid  
842 fitness island attenuate metabolite stress at extremely low pH and mediate a cell  
843 density-dependent acid resistance. *J Bacteriol* **189**:2759–2768.
- 844 61. **Kanje U, Houry WA.** 2013. Mechanisms of Acid Resistance in *Escherichia coli*.  
845 *Annu Rev Microbiol* **67**:65–81.
- 846 62. **Bryant JA, Sellars LE, Busby SJW, Lee DJ.** 2014. Chromosome position effects  
847 on gene expression in *Escherichia coli* K-12. *Nucleic Acids Res* **42**:11383–11392.
- 848 63. **Keseler IM, Mackie A, Peralta-Gil M, Santos-Zavaleta A, Gama-Castro S,**  
849 **Bonavides-Martínez C, Fulcher C, Huerta AM, Kothari A, Krummenacker M,**  
850 **Latendresse M, Muñiz-Rascado L, Ong Q, Paley S, Schröder I, Shearer AG,**  
851 **Subhraveti P, Travers M, Weerasinghe D, Weiss V, Collado-Vides J, Gunsalus**  
852 **RP, Paulsen I, Karp PD.** 2013. EcoCyc: fusing model organism databases with  
853 systems biology. *Nucleic Acids Res* **41**:D605–D612.
- 854 64. **Kiupakis AK, Reitzer L.** 2002. ArgR-Independent Induction and ArgR-Dependent  
855 Superinduction of the *astCADBE* Operon in *Escherichia coli*. *J Bacteriol* **184**:2940–  
856 2950.

- 857 65. **Waterman SR, Small PLC.** 2003. Transcriptional expression of *Escherichia coli*  
858 glutamate-dependent acid resistance genes *gadA* and *gadBC* in an *hns rpoS*  
859 mutant. *J Bacteriol* **185**:4644–4647.
- 860 66. **Bailey TL.** 2011. DREME: motif discovery in transcription factor ChIP-seq data.  
861 *Bioinformatics* **27**:1653–1659.
- 862 67. **Shin M, Song M, Rhee JH, Hong Y, Kim Y-J, Seok Y-J, Ha K-S, Jung S-H,**  
863 **Choy HE.** 2005. DNA looping-mediated repression by histone-like protein H-NS:  
864 specific requirement of  $E\sigma 70$  as a cofactor for looping. *Genes Dev* **19**:2388–2398.
- 865 68. **Levi-Meyrueis C, Monteil V, Sismeiro O, Dillies M-A, Kolb A, Monot M, Dupuy**  
866 **B, Duarte SS, Jagla B, Coppee J-Y, Beraud M, Norel F.** 2015. Repressor activity  
867 of the RpoS/  $\sigma$ -dependent RNA polymerase requires DNA binding. *Nucleic Acids*  
868 *Res* **43**:1456–1468.
- 869 69. **Typas A, Becker G, Hengge R.** 2007. The molecular basis of selective promoter  
870 activation by the sigmaS subunit of RNA polymerase. *Mol Microbiol* **63**:1296–306.
- 871 70. **Typas A, Hengge R.** 2006. Role of the spacer between the –35 and –10 regions in  
872  $\sigma$ s promoter selectivity in *Escherichia coli*. *Mol Microbiol* **59**:1037–1051.
- 873 71. **Richard HT, Foster JW.** 2003. Acid resistance in *Escherichia coli*. *Adv Appl*  
874 *Microbiol* **52**:167–186.

- 875 72. **Lund P, Tramonti A, Biase DD**. 2014. Coping with low pH: molecular strategies in  
876 neutralophilic bacteria. *FEMS Microbiol Rev* **38**:1091–1125.
- 877 73. **Zambrano MM, Siegele DA, Almirón M, Tormo A, Kolter R**. 1993. Microbial  
878 competition: *Escherichia coli* mutants that take over stationary phase cultures.  
879 *Science* **259**:1757–60.
- 880 74. **Spira B, de Almeida Toledo R, Maharjan RP, Ferenci T**. 2011. The uncertain  
881 consequences of transferring bacterial strains between laboratories - rpoS  
882 instability as an example. *BMC Microbiol* **11**:248.
- 883 75. **Hommais F, Krin E, Coppée J-Y, Lacroix C, Yeramian E, Danchin A, Bertin P**.  
884 2004. GadE (YhiE): a novel activator involved in the response to acid environment  
885 in *Escherichia coli*. *Microbiol Read Engl* **150**:61–72.
- 886 76. **Tramonti A, De Canio M, Delany I, Scarlato V, De Biase D**. 2006. Mechanisms  
887 of transcription activation exerted by GadX and GadW at the *gadA* and *gadBC*  
888 gene promoters of the glutamate-based acid resistance system in *Escherichia coli*.  
889 *J Bacteriol* **188**:8118–27.
- 890 77. **Seo SW, Kim D, O'Brien EJ, Szubin R, Palsson BO**. 2015. Decoding genome-  
891 wide GadEWX-transcriptional regulatory networks reveals multifaceted cellular  
892 responses to acid stress in *Escherichia coli*. *Nat Commun* **6**:7970.
- 893 78. **Schneider TD, Stephens RM**. 1990. Sequence logos: a new way to display  
894 consensus sequences. *Nucleic Acids Res* **18**:6097–6100.

895 79. **Crooks GE, Hon G, Chandonia J-M, Brenner SE.** 2004. WebLogo: A Sequence  
896 Logo Generator. *Genome Res* **14**:1188–1190.

897

898

899

900 **Table 1** Biological processes enriched in the RpoS regulon.

901

902

GO ID	GO Term	Genes in genome	Observed in regulon	Expected in regulon	p-value
GO:0009063	cellular amino acid catabolic process	69	38	18.5	3.5e-07
GO:0006099	tricarboxylic acid cycle	20	15	5.35	8.3e-06
GO:0006096	glycolytic process	18	14	4.81	8.3e-06
GO:0006094	gluconeogenesis	13	9	3.48	0.0016
GO:0009441	glycolate metabolic process	6	5	1.6	0.0063
GO:0006090	pyruvate metabolic process	21	17	5.61	0.0180
GO:0006970	response to osmotic stress	28	13	7.49	0.0190
GO:0006542	glutamine biosynthetic process	10	6	2.67	0.0270
GO:0009060	aerobic respiration	65	33	17.38	0.0285
GO:0006006	glucose metabolic process	36	20	9.62	0.0317
GO:0016052	carbohydrate catabolic process	276	94	73.79	0.0326
GO:0009101	glycoprotein biosynthetic process	13	7	3.48	0.0342
GO:0019395	fatty acid oxidation	30	13	8.02	0.0354
GO:0000162	tryptophan biosynthetic process	11	6	2.94	0.0463
GO:0009246	enterobacterial common antigen biosynthesis	11	6	2.94	0.0463
GO:0042398	cellular modified amino acid biosynthesis	17	7	4.55	0.0472
GO:0009239	enterobactin biosynthetic process	6	4	1.6	0.0472
GO:0000270	peptidoglycan metabolic process	51	14	13.64	0.0475

903 **Table 2** Biological processes enriched in the sensitive genes.

904

GO ID	GO Term	Genes in genome	Observed in sensitive set	Expected in sensitive set	p-value
GO:0006970	response to osmotic stress	13	4	0.79	0.0055
GO:0009063	cellular amino acid catabolic process	38	6	2.31	0.0210
GO:0019395	fatty acid oxidation	13	3	0.79	0.0385

909

910 **Table 3** Biological processes enriched in the insensitive genes.

GO ID	GO Term	Genes in genome	Observed in insensitive set	Expected in insensitive set	p-value
GO:0006412	translation	21	3	0.6	0.018
GO:0006164	purine nucleotide biosynthetic process	10	2	0.28	0.030
GO:0006805	xenobiotic metabolic process	13	2	0.37	0.049

917

918

919

920

921

922

923 **Table 4** Contingency table of the (in)sensitivity of a gene and if its promoter is or is  
924 not bound by RpoS.

925

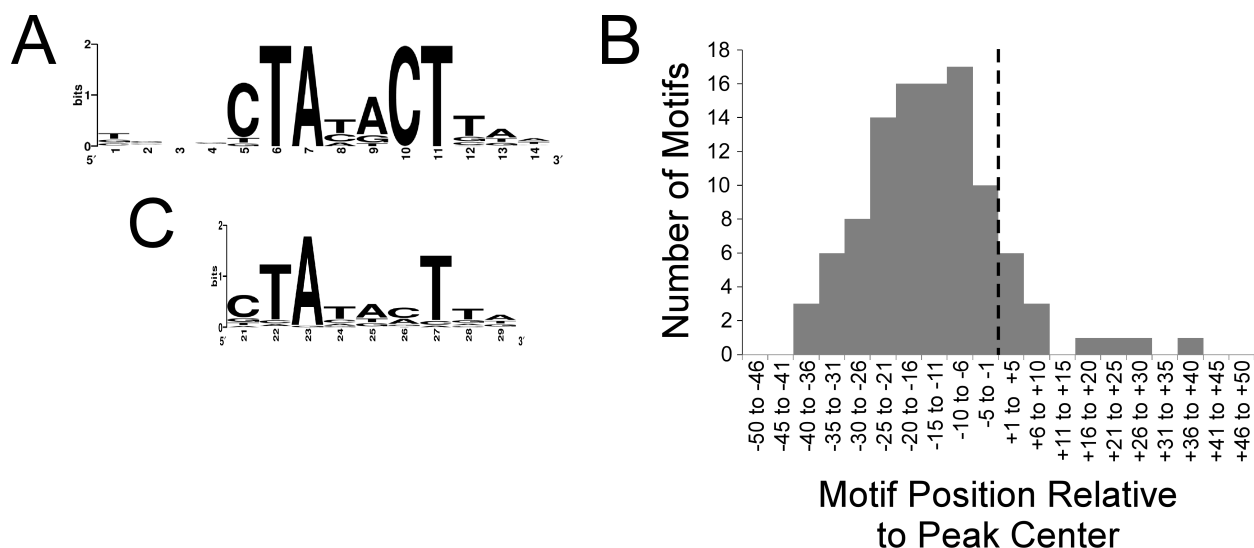
	Sensitive	Linear	Insensitive
Bound by RpoS	31	87	5
Not bound by RpoS	71	823	27

926

927



928  
929

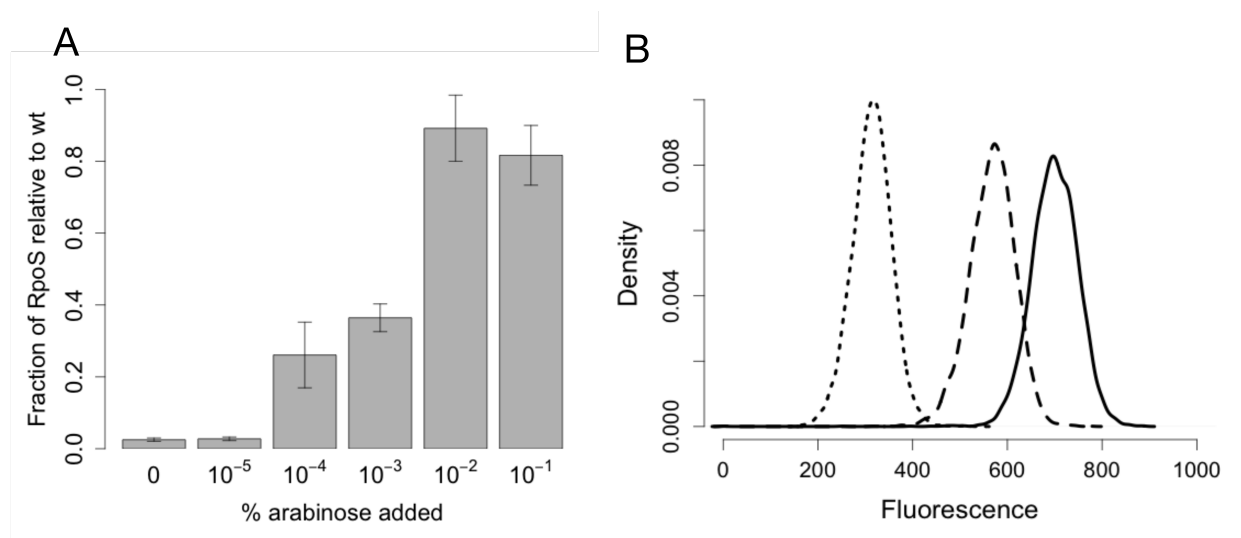


930  
931  
932  
933  
934  
935  
936  
937  
938  
939  
940  
941

**Figure 1. Analysis of RpoS CHIP-seq data.** (A) Sequence logo (78) of the enriched sequence motif associated with RpoS-bound regions (MEME E-value  $2.0e-40$ ). The sequence logo was generated using WebLogo (79) (B) Histogram showing the frequency distribution of distances between identified motifs and CHIP-seq peak centers. The black, vertical, dashed line indicates the peak center position. (C) Sequence logo (78) of the enriched sequence motif associated with TSSs located within 20 bp of RpoS CHIP-seq peak centers (MEME E-value  $2.6e-56$ ). Note that only positions 21-29 of the motif are shown. The sequence logo was generated using WebLogo (79).

942

943



944

945 **Figure 2. Arabinose control of RpoS protein levels in strain DMS2564.** (a)

946 Increasing arabinose results in more RpoS expression in strain DMS2564, as measured

947 by western blotting. n = 4 to 5; error bars represent SEM. (b) Flow cytometric

948 measurement of an *otsB-gfp* fusion under 0%, 10<sup>-4</sup>%, and 10<sup>-1</sup>% arabinose. (Dotted line,

949 dashed line, and solid line, respectively.)

950

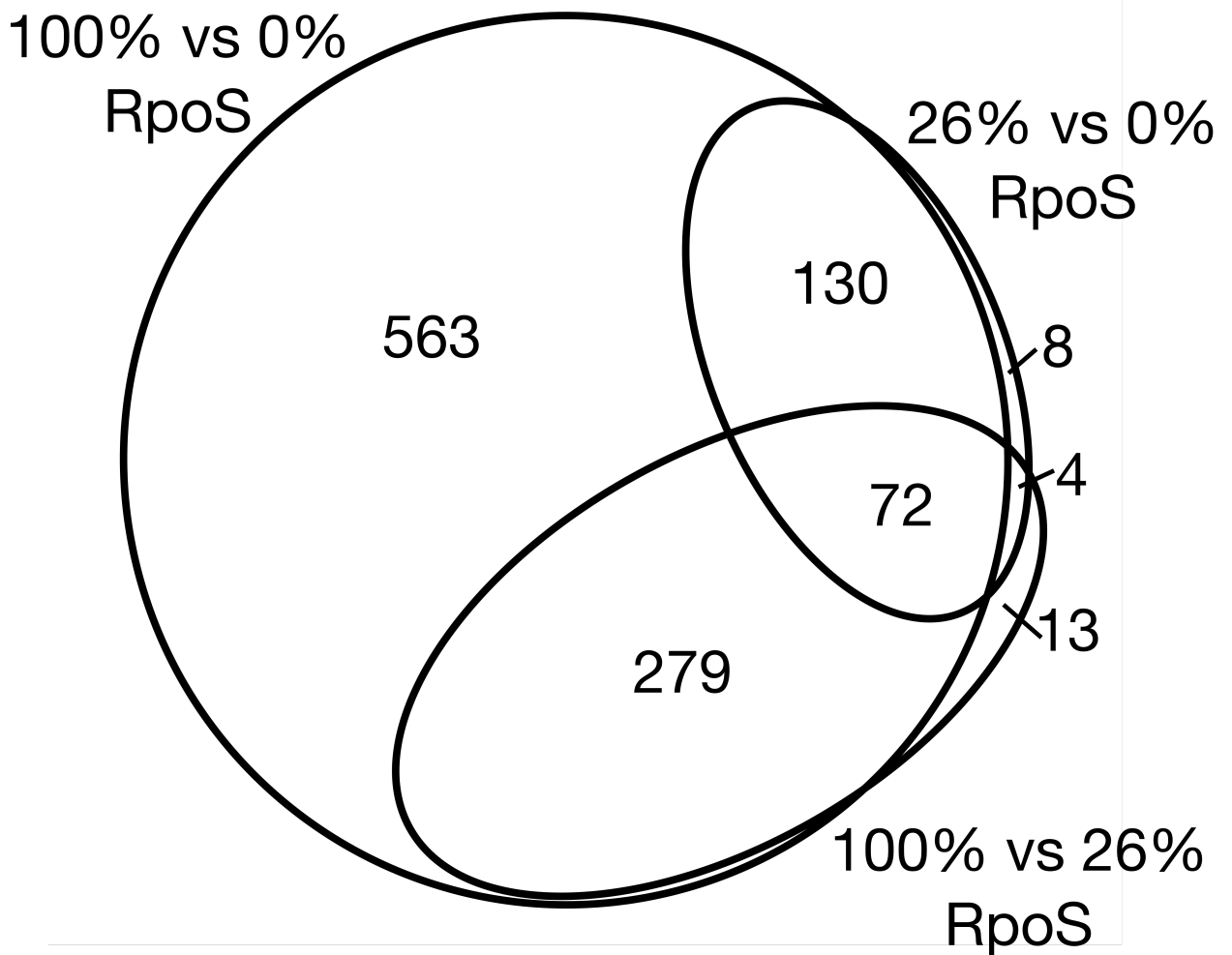
951

952

953

954

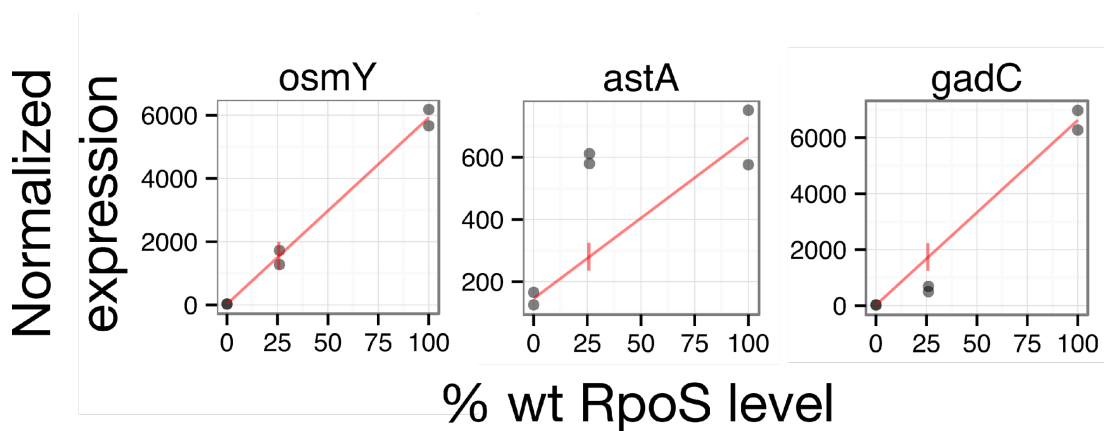
955



956  
957 **Figure 3:** Area proportional Venn diagram showing the number of genes  
958 differentially expressed between each of the three conditions.  
959

960  
961  
962  
963  
964  
965  
966  
967  
968

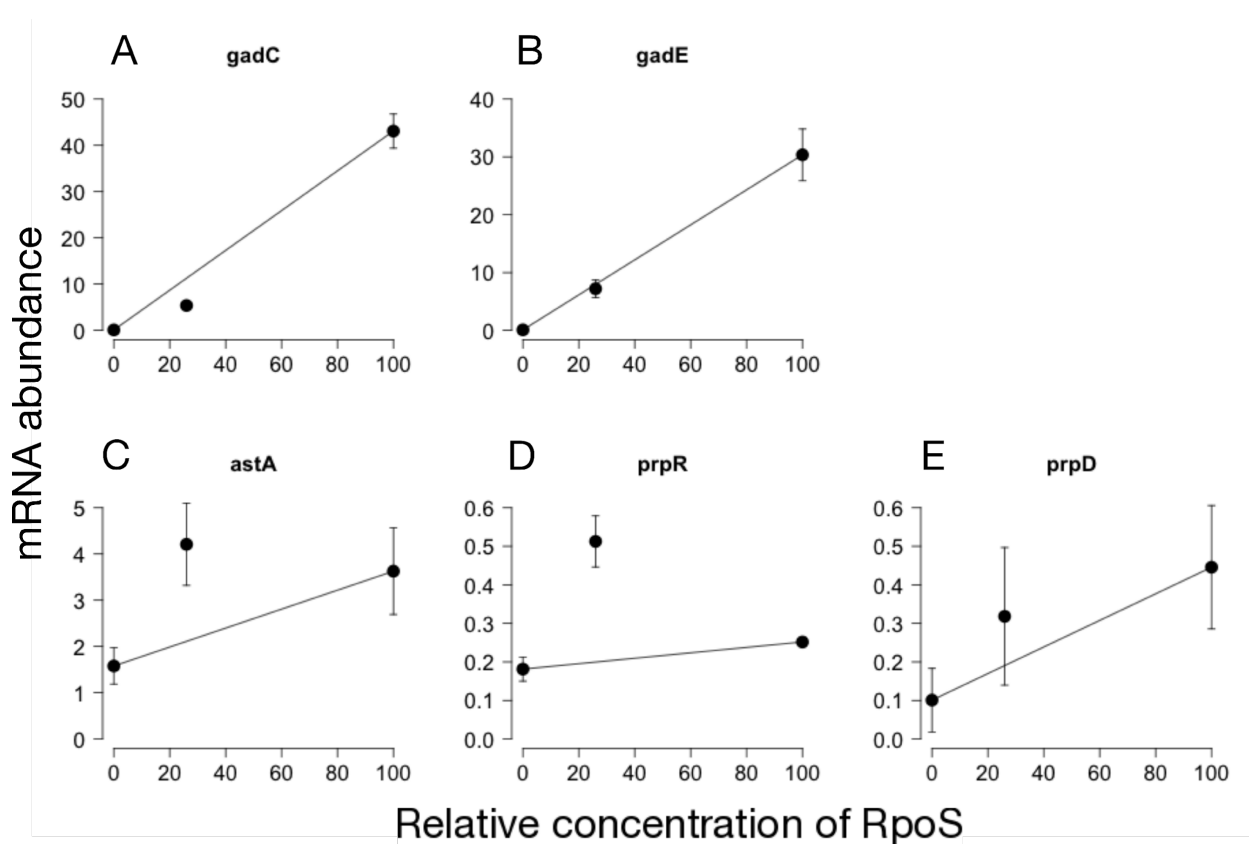
969  
970  
971



972  
973  
974  
975  
976  
977  
978  
979  
980

**Figure 4:** Examples of three classes of patterns of RpoS response. *osmY* expression is linear with RpoS level, *astA* is sensitive to RpoS level, and *gadC* is insensitive to RpoS level. Dots represent normalized expression levels from individual RNA-seq samples. The line is drawn between the mean at 0% and the mean at 100%.

981



982

983

984

985 **Figure 5. Testing of RNA-seq expression patterns using qPCR.** Expression  
986 patterns were measured for (a,b) *gadC* and *gadE*, which were insensitive in the RNA-  
987 seq data, and (c-e) *astA*, *prpR*, and *prpD*, which were sensitive in the RNA-seq data. In  
988 all cases, the level of transcription at 26% RpoS was on the side of the line predicted by  
989 the RNA-seq data, though not always significantly so. *gadC* was significantly insensitive  
990 ( $p < 10^{-4}$ ), *gadE* was not significantly different from linear ( $p = 0.10$ ), *astA* was (barely)  
991 not significantly sensitive ( $p = 0.06$ ), *prpR* was significant sensitive ( $p < 10^{-4}$ ), and *prpD*  
992 was not significant ( $p = 0.56$ ).  $n = 6$ , error bars = SEM.

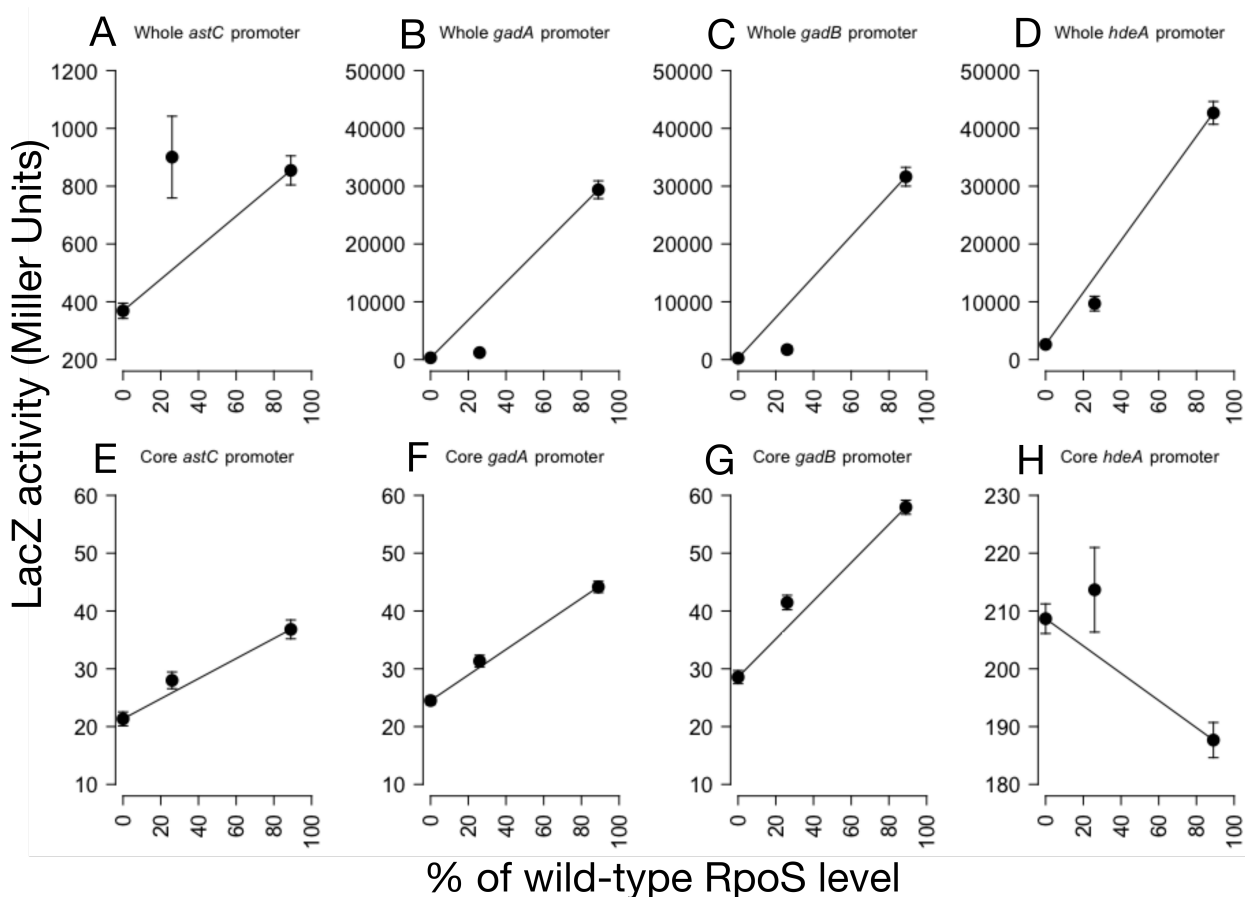
992

993

994

995

996



997

998

999

**Figure 6. Expression patterns of whole promoters and core promoters only.**

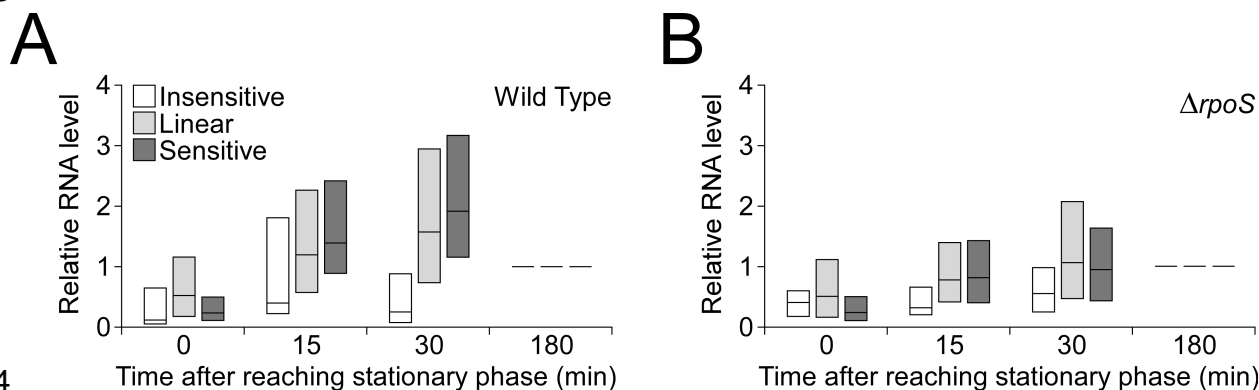
1000 Expression patterns were measured for the (a - d) whole upstream regulatory region  
 1001 and the (e - h) core promoter only of *astC*, *gadA*, *gadB*, and *hdeA*. As expected by  
 1002 RNA-seq data, the whole *astC* fusion was sensitive ( $p = 0.04$ , t-test), and the *gadA*,  
 1003 *gadB*, and *hdeA* fusions were all insensitive ( $p = 2 \times 10^{-6}$ ,  $p = 10^{-5}$ ,  $p = 0.04$ , all by one  
 1004 sample t-test). The p-values are adjusted for multiple comparisons using Holm's  
 1005 sequential adjustment method (50). The core promoters (e - h) have much lower levels  
 1006 of maximal transcription than the full length promoters. All four have much altered  
 1007 patterns of sensitivity, although only *gadA* and *gadB* was significantly different from the  
 1008 whole promoter ( $p = 0.08$  (*astC*),  $p = 0.01$  (*gadA*),  $p = 10^{-6}$  (*gadB*),  $p = 0.21$  (*hdeA*), all  
 1009 by two-sample t-test with Holm's correction.)  $n = 6$  to  $8$ , error bars = SEM.

1010

1011

1012

1013



1014

1015

1016

1017 **Figure 7. Expression profiles of genes when cells enter stationary phase. (A)**

1018 Relative expression levels of RpoS-regulated genes were taken from a published  
1019 study (47) for four time-points following entry into stationary phase. All expression  
1020 values were normalized to those from the final time-point for each gene. The graph  
1021 shows the range of relative expression values for insensitive (white boxes), linear  
1022 (light gray boxes), and sensitive (dark gray boxes) genes for each time-point. Boxes  
1023 represent the 25<sup>th</sup>-75<sup>th</sup> percentile range, and horizontal lines indicate the median  
1024 value. **(B)** As above, but for a  $\Delta rpoS$  strain of *E. coli*.