1 **Wild-type *Caenorhabditis elegans* Isolates Exhibit Distinct Gene Expression Profiles in**

2 **Response to Microbial Infection**

3 Patrick Lansdon, Maci Carlson, and Brian D. Ackley

4 Department of Molecular Biosciences, University of Kansas, Lawrence, KS, 66045

5

6 Running title: Pathogens and *C. elegans* wild isolates

7

8 Keywords: Bacterial Pathogens, Innate Immunity, Transcriptomes, *Caenorhabditis elegans*

9 Corresponding Author:

10 Brian D. Ackley

11 5004 Haworth Hall

12 1200 Sunnyside Ave

13 Lawrence, KS 66045

14 bdackley@ku.edu

15

16 7 Figures, 5 Tables, 15 Supplemental Tables, 1 Supplemental File

17 32 Pages

18 Abstract: 201 Words

19 Introduction: 523 Words

20 Results: 3704 Words

21 Discussion: 1731 Words

Abstract

The soil-dwelling nematode *Caenorhabditis elegans* serves as a model system to study innate immunity against microbial pathogens. *C. elegans* have been collected from around the world, where they, presumably, adapted to regional microbial ecologies. Here we use survival assays and RNA-sequencing to better understand how two isolates from disparate climates respond to pathogenic bacteria. We found that, relative to N2 (originally isolated in Bristol, UK), CB4856 (isolated in Hawaii), was more susceptible to the Gram-positive pathogen, *Staphylococcus epidermidis*, but equally susceptible to *Staphylococcus aureus* as well as two Gram-negative pathogens, *Providencia rettgeri* and *Pseudomonas aeruginosa*. We performed transcriptome analysis of infected worms and found gene-expression profiles were considerably different in an isolate-specific and pathogen-specific manner. We performed GO term analysis to categorize differential gene expression in response to *S. epidermidis*. In N2, genes that encoded detoxification enzymes and extracellular matrix proteins were significantly enriched, while in CB4856, genes that encoded detoxification enzymes, C-type lectins, and lipid metabolism proteins were enriched, suggesting they have different responses to these bacterial pathogens, despite being the same species. Overall, discerning gene expression signatures in an isolate by pathogen manner can help us to understand the different possibilities for the evolution of immune responses within organisms.

39 **Introduction**

40 Antimicrobial resistance is a looming public health crisis as microbial pathogens evade drugs
41 of last resort. When pathogenic microbes bypass the physical barriers of the host (*e.g.,* an
42 epidermis), an organism must rely on cellular and humoral defenses (*e.g.,* the innate immune
43 system) to protect itself from infections that can compromise organismal fitness. While all
44 organisms will inevitably become infected over their lifespan, genetic variation in individuals will
45 influence the outcomes of infections. Thus, our ability to link genetic variation to infection
46 susceptibility may enable more sophisticated strategies for deploying antibiotics or help us
47 enhance host immunity to treat microbial infection.[1,2]

48 Genetic variation in organisms can create tradeoffs in susceptibility to different pathogens.
49 The presence of non-synonymous single nucleotide polymorphisms (SNPs) in the coding
50 sequence of immune effector genes has been shown to alter the propensity of individuals to
51 become infected and develop disease.[3] For instance, a SNP in the *TLR5* gene introduces a
52 premature stop codon, creating a truncated form of the protein and impairing the ability of TLR5
53 to recognize the flagellin protein in flagellated bacteria.[4] Individuals carrying this SNP are more
54 susceptible to infection by *Legionella pneumophila*, the vector for Legionnaires' disease, but not
55 *Salmonella enterica serotype Typhi,* the microbial vector for typhoid fever.[5,6]

56 Animal models have been critical to our understanding of the innate immune system.
57 However, the reliance on canonical laboratory strains often ignores genetic variation in innate
58 immunity and susceptibility to pathogens.[7] The soil-dwelling nematode *Caenorhabditis elegans*
59 is a genetically tractable model organism to examine host-pathogen interactions.[8,9] *C. elegans*

60 possesses a robust microbial defense system with evolutionarily conserved mechanisms in
61 microbial pathogen detection, immune activation and pathogen killing (*e.g.* the induction of
62 reactive oxygen species (ROS), deployment of antimicrobial and antifungal peptides, and
63 expression of bacterial-binding C-type lectins).[10-14] These evolutionarily conserved responses to
64 infection make *C. elegans* an ideal model to study bacterial pathogenesis.[9]

65 *C. elegans* wild isolates collected from geographically discrete regions are apt to vary in their
66 ability to fight infections either because of a history of different selective pressures or simply
67 genetic drift. Many discoveries regarding *C. elegans* host-microbe interactions have come from
68 studying wild isolates, including identifying virus that infect *C. elegans* and the characterization
69 of the *C. elegans* gut microbiome.[15-17] The genomic diversity of wild isolates has also provided a
70 platform for quantitatively examining the genetic variation underlying complex traits, including
71 behavioral differences in foraging as well as pathogen avoidance and susceptibility.[18-23] However,
72 less attention has been paid to the variation in the transcriptomic responses of wild isolates to
73 multiple, discrete microbial pathogens.

74 In this study, we show that the *C. elegans* Hawaiian wild isolate CB4856 is significantly more
75 susceptible to infection by the Gram-positive pathogen *Staphylococcus epidermidis* than the N2
76 lab-conditioned strain. We used high-throughput RNA sequencing (RNA-seq) to identify
77 differentially expressed genes in staged N2 and CB4856 worms exposed to pathogenic bacteria.
78 The differentially expressed genes identified in pathogen-infected animals provide insight into
79 isolate-specific transcriptomic responses of *C. elegans* to pathogenic bacteria. We believe that

80 this knowledge will provide a foundation for better understanding the role of natural genetic

81 variation in the *C. elegans* innate immune response to microbial pathogens.

82 ## Materials and Methods

83 ### *C. elegans* and Bacterial Strains

84 *C. elegans* N2 and CB4856 wild-type strains were maintained at 20°C on nematode growth
85 medium (NGM) agar plates seeded with *Escherichia coli* OP50.[24] *Staphylococcus epidermidis*
86 (EVL2000, this study), *Staphylococcus epidermidis* (1191; ATCC#700562), *Providencia rettgeri*
87 (Dmel1)[25], *Pseudomonas aeruginosa* (PA-14)[26], and *Staphylococcus aureus* (Newman)[27] were
88 used as pathogens in this study. All bacteria strains were grown at 37°C with shaking in low-salt
89 Luria Bertani (LB) medium (10 g Bacto-tryptone, 5 g yeast extract, 5 g NaCl).

90

91 ### Identification of *S. epidermidis* Strain

92 DNA was extracted from 1mL of fresh culture using a QiaQuick miniprep column (Qiagen) and
93 the 16S ribosomal region was amplified using universal primers 27F (5'-
94 AGAGTTTGATCCTGGCTCAG-3') and 1492R (5'-CGGTTACCTTGTTACGACTT-3'). PCR cycling
95 conditions were as follows: 1) 92°C for 2 minutes; 2) 35 cycles at 92°C for 15-sec, 55°C for 15-sec,
96 68°C for 75-seconds; and 3) 68°C for 10 minutes. The amplified PCR product was purified, Sanger
97 sequenced, and the NCBI BLAST tool was used to identify the species.

98

99 ### Survival Assays

100 NGM plates were seeded with 40 µL of bacteria (*E. coli*, *S. epidermidis*, *P. rettgeri*, *P.*
101 *aeruginosa*, *or S. aureus*) and incubated overnight at 37°C. Plates were acclimated to room

102 temperature and 30 late-stage L4 worms were added. Each day, living worms were transferred
103 to fresh plates and the number of dead worms was recorded.
104
105 **Pathogen Avoidance Assays**
106 NGM plates were seeded with a 10 µL dot of pathogenic bacteria (*S. epidermidis*, *P. rettgeri*,
107 *S. aureus*, or *P. aeruginosa*) on one half, a 10 µL dot of non-pathogenic *E. coli* OP50 on the other
108 half and incubated overnight at 37°C. 20 late-stage L4 worms were placed on blank plates for 30
109 minutes to remove any external bacteria while pathogen avoidance plates acclimated to room
110 temperature. Worms were placed at the midline and their location on the plate was recorded at
111 1-hour, 2-hour, 4-hour, 8-hour, and 24-hour time points. To avoid ambiguity in the number of
112 animals associated with the specific bacteria, only animals residing within the spot of bacteria
113 were assigned "preference" for either the control (OP50) or pathogen.
114
115 **Exposure of Nematodes to Pathogen**
116 NGM plates were seeded with 250 µL of bacteria incubated overnight at 37°C. Approximately
117 2,000 L4 stage worms grown on *E. coli* OP50 were transferred to NGM plates seeded with bacteria
118 and incubated at 20°C for 24 hours. After 24 hours, the worms were washed with M9 buffer three
119 times to remove any residual bacteria. The worms were resuspended in 100 µL of M9 buffer and
120 mechanically disrupted in liquid nitrogen using a ceramic mortar and pestle. Frozen tissue was
121 transferred to a microcentrifuge tube and 1 mL TRIZOL reagent was added. The tubes were flash

122 frozen in liquid nitrogen and stored at -80°C. Three biological replicates were collected for each
123 worm and pathogen strain.
124
125 **RNA isolation and Sequencing**
126 Frozen worm tissue was thawed, vortexed for 5 seconds and incubated at room temperature
127 for 5 minutes. After the addition of 470 µL chloroform (mixed by inversion and phase separated
128 for 2 minutes at room temperature), the samples were centrifuged at 15,000 RPM at 4°C for 15
129 minutes. The upper aqueous phase containing RNA (approximately 600 µL) was transferred to a
130 new RNase-free Eppendorf tube. Total RNA extraction was carried out using the Monarch RNA
131 Cleanup Kit (New England Biolabs) according to the manufacturer's instructions. Total RNA was
132 quantified using the Qubit (ThermoFisher Scientific) and RNA quality and integrity was assessed
133 using the Agilent Tapestation 2200 (Agilent Technologies). Sequence libraries were prepared
134 using the TruSeq Stranded mRNA Library Prep Kit (Illumina) and sequenced using single-end 1x75
135 bp sequencing on the Illumina NextSeq 550.
136
137 **RNA-seq Analysis and Quantification of Differentially Expressed Genes**
138 The quality of the 75-bp single-end reads generated using the Illumina NextSeq 550 was
139 assessed using FastQC[28] (v.0.11.5). Illumina sequencing adapters and low-quality bases (Phred
140 score <30) were trimmed from reads using Fastp[29] (v.0.19.8). Reads were first aligned to the *E.*
141 *coli*, *S. epidermidis*, *P. rettgeri*, *S. aureus*, or *P. aeruginosa* genomes to evaluate bacterial
142 contamination, then aligned to the *C. elegans* genome assembly (Wormbase WS273 release)

143 using HISAT2[30] (v.2.1.0) with default settings. Aligned reads were mapped to the *C. elegans*
144 annotated genome (Wormbase WS273 release) and transcript abundances were quantified using
145 featureCounts[31] (v.1.6.0). Data normalization and differential gene expression analysis was
146 performed with DESeq2[32] (v.3.1.0) in R (v.3.6.0) using a false discovery rate (FDR)-adjusted P-
147 value <0.05 and fold change ≥2 as cutoffs. Separate design formulas were used to determine
148 differential gene expression for infection (~Infection), Gram-stain (~Gram-stain), and individual
149 pathogens (~Pathogen). Geneset enrichment analysis was performed with WormCat[33] using an
150 FDR-adjusted P-value <0.05 as the significance cutoff.

151

152 **Statistical Analysis**

153 All statistical analysis was carried out using SigmaPlot 14.0 (Systat Software, Inc.). Log rank
154 (Mantel-Cox) analysis was used for pairwise comparison of survival curves and to calculate
155 median survival (LT_{50}). Chi-squared or Fisher's Exact Test was used to assess statistical
156 significance of the pathogen avoidance data. Spearman's correlation was used to determine the
157 correlation coefficient between $\log_2$ fold changes in gene expression. Statistically significant
158 differences are defined in the figure legend and noted in the figure with asterisks.

159

160 **Data Availability**

161 The raw sequence files used for differential gene expression analysis are available in FASTQ
162 format in the NCBI BioProject repository (BioProject Accession number: PRJNA770277;
163 https://www.ncbi.nlm.nih.gov/sra/PRJNA770277).

164 ## Results

165 **CB4856 and N2 exhibit differences in susceptibility to the pathogen *S. epidermidis***

166 During our studies on susceptibility to microbial pathogens we serendipitously cultured a
167 bacterial contaminant that exhibited differences in a *C. elegans* isolate-specific manner. We
168 identified this bacterium as *Staphylococcus epidermidis* by 16S sequencing (see Materials and
169 Methods) and provided it with the strain designation EVL2000. Survival assays conducted with
170 the EVL2000 *S. epidermidis* strain found CB4856 animals to be significantly more susceptible than
171 N2, with a median survival (LT_{50}) of 11.3 ± 1.6 days, whereas 50% of N2 animals were killed after
172 an average of 18.8 ± 1.5 days (p<0.001) (Figure 1A, Table S1).

173 We then compared N2 and CB4856 susceptibility to other known *C. elegans* pathogens,
174 *Staphylococcus aureus, Providencia rettgeri*, and *Pseudomonas aeruginosa.* In each case, the
175 susceptibility of N2 and CB4856 to these pathogens was not significantly different (Figures 1B-D,
176 Table S1). We also grew animals on non-pathogenic *E. coli* OP50 to evaluate the possibility that
177 the increased susceptibility of *S. epidermidis* was due to a general defect in lifespan. Interestingly,
178 the lifespan of N2 worms was significantly shorter than CB4856, with 50% of animals dying at
179 12.0 ± 0.4 and 16.2 ± 0.6 days for N2 and CB4856 animals, respectively (p<0.001) (Figure 1E, Table
180 S1). While there were differences in lifespan on *E. coli*, our data are consistent with the idea that
181 the increased susceptibility of CB4856 to *S. epidermidis* is not due to a universal decrease in
182 lifespan.

183 Relative to worms fed *E. coli*, infection with the EVL2000 *S. epidermidis* strain actually
184 prolonged the median lifespan of N2 animals (*E. coli* OP50 LT_{50} = 12.0 ± 0.4 days, *S. epidermidis*

185 EVL2000 LT_{50} = 18.8 ± 1.5 days; $p<0.001$). To determine if the differences in susceptibility were
186 unique to the EVL2000 strain, we infected N2 and CB4856 animals with an additional *S.*
187 *epidermidis* strain (1191). CB4856 animals infected with the 1191 strain had a median survival of
188 5.7 ± 1.2 days, whereas 50% of N2 animals were killed after an average of 12.3 ± 0.9 days
189 ($p<0.001$) (Figure 1F, Table S1). Together, our results indicate that CB4856 is more susceptible to
190 *S. epidermidis* infection and that this increase in susceptibility is observed across multiple strains
191 of *S. epidermidis.* Since infection by the *S. epidermidis* EVL2000 strain provided the greatest
192 difference in susceptibility between N2 and CB4856, follow-up experiments used this bacterial
193 strain.
194
195 **Pathogen avoidance of N2 and CB4856 isolates depends on the pathogen and time of**
196 **observation**
197 We performed two-choice pathogen avoidance assays to determine if the difference in
198 susceptibility could be explained in part by differences in animal behavior (Figure 2A). At the 1-
199 hr time point there was a significant difference in *S. epidermidis* avoidance, with 100% of N2 and
200 70% of CB4586 animals on the OP50 lawn (Figure 2B) ($p<0.001$). By 24 hours the difference in
201 pathogen avoidance was no longer significant and nearly 100% of worms from both isolates were
202 found on the OP50 lawn. This suggests that both genotypes perceived *S. epidermidis* as
203 pathogenic and avoided them accordingly.
204 We also examined avoidance behavior in N2 and CB4856 animals exposed to *S. aureus*, *P.*
205 *rettgeri* or *P. aeruginosa* and found differences in the initial avoidance of pathogenic bacteria. At

206 the 1-hr time point avoidance of *S. aureus* was not significantly different between the two
207 isolates, with 80% and 70% of animals observed on the OP50 lawn for N2 and CB4856,
208 respectively (Figure 2C) (p=0.390). Avoidance of *P. rettgeri* at the 1-hr time point was significantly
209 more likely in N2 worms, with 53% of N2 and 19% of CB4856 animals on the OP50 lawn (Figure
210 2D) (p<0.001). The avoidance of *P. aeruginosa* trended in the opposite direction with 46% of N2
211 and 65% of CB4856 animals on the OP50 lawn (Figure 2E) (p=0.051). Neither the N2 nor CB4856
212 genotypes had significant changes in pathogen avoidance between the 1-hr and 24-hr time points
213 for these three pathogens (Figure 2C-E). Natural variation in olfactory preference may partially
214 explain these differences in pathogen avoidance.[34] While the results are interesting and warrant
215 follow-up, they suggest that the avoidance behavior of N2 and CB4856 animals likely has little
216 impact on susceptibility to *S. epidermidis*.
217
218 **Gene expression changes during infection cluster by genotype, rather than pathogen type**
219 To identify gene expression changes due to pathogen we performed RNA sequencing
220 comparing CB4856 and N2 animals fed pathogenic bacteria versus non-pathogenic *E. coli* OP50
221 for 24 hours. *E. coli* OP50 is a standardized food source for laboratory-bred *C. elegans*, and thus
222 represents the nonpathogenic control diet. We analyzed gene expression in N2 and CB4856
223 animals exposed to either *S. epidermidis, S. aureus*, *P. rettgeri* or *P. aeruginosa*.
224 Although both *C. elegans* isolates showed similar susceptibility to three of these pathogens,
225 we hypothesized geographical isolation might have resulted in variations in the immune
226 response, which could be analyzed through transcriptional profiling. Principal component

227 analysis (PCA) of gene expression found that N2 and CB4856 samples exposed to *E. coli*, the
228 Gram-positive pathogen *S. epidermidis* or the pathogenic Gram-negative *P. rettgeri* clustered in
229 an isolate-specific manner (Figure 3A). N2 samples exposed to the virulent Gram-negative *P.*
230 *aeruginosa* were also part of this N2-specific cluster, whereas CB4856 samples exposed to *P.*
231 *aeruginosa* formed their own distinct cluster. Further, both genotypes exposed to the moderately
232 virulent Gram-positive *S. aureus* formed distinct groups that did not cluster with other samples
233 from the same genotype, nor did they cluster with each other (Figure 3A).

234 We also performed PCA of gene expression using samples from a single genotype. For both
235 *C. elegans* isolates *S. epidermidis* and *P. rettgeri* samples again clustered with *E. coli*, whereas. *S.*
236 *aureus* samples formed a separate group along the first principal component axis (PC1). *P.*
237 *aeruginosa* samples also formed individual clusters for both isolates, with separation along the
238 second principal component axis (PC2) in N2 and separation along both the PC1 and PC2 axes in
239 CB4856 (Figures 3B, 3C). Thus, it appears that both nematode genotype and microbial taxonomy
240 contribute to the gene expression differences between N2 and CB4856.

241

242 **Differential Gene Expression is Largely Unique in N2 and CB4856 Animals After Infection with**
243 **Pathogenic Bacteria**

244 We first sought to identify how gene expression changes in these two isolates, as a function
245 of infection. Thus, we identified genes differentially expressed in response to all pathogens (for
246 all analyses, differentially expressed genes (DEGs) were defined as those with an FDR-adjusted P-
247 value <0.05 and fold change ≥2, infected vs. control). We identified 273 and 117 genes that were

248 differentially expressed in N2 and CB4856, respectively, as a function of infection. In either
249 genotype the majority of DEGs were upregulated (N2, 236/273 = 86%; CB4856, 96/117 = 82%)
250 (Table 1, Table 2, Table S2). Interestingly, only six out of the 384 total genes were differentially
251 expressed in both N2 and CB4856 (N2-only, 267/273 = 98%; CB4856-only, 111/117 = 95%) (Figure
252 4A, Table S2).

253 We examined the correlation of gene expression using the $\log_2$ fold changes after infection.
254 Expression levels were weakly, but positively, correlated between N2 and CB4856 ($\rho = 0.097$; $p<$
255 0.001) (Figure 4B). A correlation of baseline gene expression was determined using normalized
256 gene counts from worms exposed to the control bacteria (OP50 *E. coli*). In that context N2 and
257 CB4856 had similar patterns of expression ($\rho = 0.826$), suggesting that gene expression in
258 response to infection diverges from a common baseline. Taken together, these data suggest that
259 while there is some conservation of expression in response to infection, the pattern is relatively
260 weak between these two isolates.

261 We further examined the DEGs identified in both isolates. Of them, only Y65B4BR.1 has any
262 functional annotation relating to infection. Previous data suggests that RNAi knockdown
263 increases susceptibility to infection by *S. aureus.*[35] A single gene, *oac-14*, was identified as
264 differentially expressed in opposing directions, *i.e.*, up in N2 and down in CB4856. RNAi
265 knockdown delays or slows development, suggesting *oac-14* contributes to organismal growth.[36]

266 To take a more holistic approach, we used geneset enrichment analysis to analyze DEGs. The
267 GO terms "Stress response: pathogen", "Transmembrane transport: lipid", and "Metabolism:
268 lipid" were observed to be enriched in N2 (N2-only), whereas genes corresponding to the GO

term "Metabolism: lipid" were overrepresented in CB4856 animals (CB4856-only) (Table 3, Table S3). Taken together, these results indicate that although there was little overlap in the transcriptomic response to microbial pathogen in N2 and CB4856 animals, both genotypes exhibit an enrichment of genes involved in lipid metabolism.

Finally, to examine the change in transcriptional networks, we cross referenced our DEG lists with a table of known *C. elegans* transcription factors (File S1). We found that in N2, the transcription factors *cky-1, nhr-19, nhr-99, nhr-116, nhr-127, peb-1* and *unc-42* were all upregulated in response to infection by pathogens, whereas in CB4856-infected animals, *ham-1* expression was repressed and *klu-2* expression was induced (Table S2). These results are consistent with the divergence of gene expression profiles between N2 and CB4856, suggesting that, upon infection, these isolates rely on different immune pathways.

N2 and CB4856 Transcriptomic Responses Exhibit More Overlap Following Infection with Gram-positive than Gram-negative Pathogens

We then asked how gene expression profiles differed when we grouped pathogens by Gram-stain. We identified 3833 and 2565 genes that were differentially expressed in response to Gram-positive bacteria in N2 and CB4856, respectively. In either genotype the majority of DEGs were upregulated (N2, 2964/3833 = 77%; CB4856, 1962/2565 = 76%) (Table 1, Table 2, Table S4). While most genes were unique to a genotype (N2-only, 2661/3833 = 69%; CB4856-only, 1393/2565 = 54%), 1172 genes were differentially expressed in both N2 and CB4856 worms (Figure 5A, Table S4). Further, the majority of DEGs identified in both genotypes were expressed in the same

direction, while only 1% of genes were differentially expressed in opposing directions, *i.e.*, down in N2 and up in CB4856 (14/1172 = 1%) (Figure 5C, Table S4). These results were reflected in a strong, positive correlation between N2 and CB4856 gene expression levels ($\rho = 0.613$; $p < 0.001$) (Figure 5A).

We again looked for transcription factors that were differentially expressed in both N2 and CB4856 as a function of infection by Gram-positive bacteria. In this list, there were 50 transcription factor genes that were significantly different from baseline (Table S4). As observed for the overall gene expression similarity, for each of the transcription factors, we observed a similar pattern for the direction of expression (*e.g.,* genes upregulated in N2 were upregulated in CB4856, *etc.*).

In response to Gram-negative bacteria, 628 and 1720 genes were differentially expressed in N2 and CB4856 animals, respectively. Interestingly, in both genotypes a minority of DEGs were upregulated (N2, 254/628 = 40%; CB4856, 456/1720 = 27%) (Table 1, Table 2, Table S5). Like the transcriptomic response when all pathogens were grouped together, most genes were unique to a genotype with a total of 124 genes differentially expressed in both N2 and CB4856 (N2-only, 504/628 = 80%; CB4856-only, 1596/1720 = 93%) (Figure 5B, Table S5). Further, the majority of DEGs identified in both genotypes were expressed in opposing directions (82/124 = 66%), leaving only 42 genes that were differentially expressed in both genotypes and in the same direction (Figure 5C, Table S5). Gene expression profiles were again weakly, but positively correlated ($\rho = 0.107$; $p < 0.001$) (Figure 5B). Finally, there were no transcription factors that were coordinately regulated in both isolates in response to Gram-negative bacteria (Table S5).

311 These results highlight the dramatic differences when we compare the transcriptomic
312 response of N2 and CB4856 worms infected with Gram-positive or Gram-negative pathogens.
313 Interestingly, exposure to Gram-positive pathogens resulted in differential gene expression
314 profiles with considerable overlap between the two isolates, suggesting that there are conserved
315 responses to some bacteria. In contrast, the transcriptomic responses of N2 and CB4856 animals
316 to Gram-negative pathogens were relatively divergent, with few differentially expressed genes
317 in common.

318

319 **Overlap in Differentially Expressed Genes between N2 and CB4856 Depends on the Microbial**
320 **Pathogen**

321 Finally, we analyzed gene expression by individual pathogen exposure. This identified 2384
322 (*S. epidermidis*), 8754 (*S. aureus*), 5802 (*P. aeruginosa*), and 1029 (*P. rettgeri*) differentially
323 expressed genes in N2 and 1776 (*S. epidermidis*), 7033 (*S. aureus*), 5171 (*P. aeruginosa*), and 2253
324 (*P. rettgeri*) differentially expressed genes in CB4856 (Table 1, Table 2). In response to *S.*
325 *epidermidis*, upregulated DEGs outnumbered downregulated DEGs at a ratio of 3:1 (N2,
326 1754/2384 = 74%; CB4856, 1296/1776 = 73%), whereas the response to *S. aureus* was split more
327 evenly between upregulated and down regulated genes, although upregulated DEGs still
328 represented a majority (N2, 4901/8754 = 56%; CB4856, 3597/7033 = 51%) (Table 1, Table 2, Table
329 S6, Table S7).

330 Although both belong to the *Staphylococcus* genus, the overlap between N2 and CB4856
331 transcriptomic responses to *S. epidermidis* and *S. aureus* were considerably different. In *S.*

332 *epidermidis*-infected animals the minority of DEGs, 511 in total, were differentially expressed in
333 both N2 and CB4856 (N2-only, 1873/2384 = 79%; CB4856-only, 1265/1776 = 71%) (Figure 6A,
334 Table S6). The correlation of gene expression in response to *S. epidermidis* was ρ = 0.294 ($p<$
335 0.001). In contrast, a majority of differentially expressed genes, 5039 in total, overlapped in N2
336 and CB4856 animals infected with *S. aureus* (N2-only, 3715/8754 = 42%; CB4856-only, 1994/7033
337 = 28%) (Figure 6B, Table S7), with a correlation of ρ = 0.721 ($p<$ 0.001).

338 The majority of DEGs identified in both genotypes following either *S. epidermidis* or *S. aureus*
339 infection were expressed in the same direction (*S. epidermidis*, 410/511 = 80%; *S. aureus*,
340 4928/5039 = 98%) (Figure 6C, Table S6, Table S7). Lastly, we only identified nine transcription
341 factors that were coordinately regulated in *S. epidermidis*-infected animals, whereas 346
342 transcription factors genes were differentially expressed in both genotypes in the same direction
343 after *S. aureus* infection (Table S6, Table S7). Together, these results suggest N2 and CB4856
344 worms exhibit considerable similarity in their transcriptomic response to *S. aureus*, whereas
345 there is much less overlap in their response to *S. epidermidis* infection, which may perhaps
346 explain some of the difference in susceptibility to infection by *S. epidermis*.

347 Exposure to either Gram-negative pathogen resulted in a minority of DEGs being upregulated
348 in both genotypes. In *P. aeruginosa*-infected animals, upregulated genes represented less than
349 one-third of all DEGs in both genotypes (N2, 1804/5802 = 31%; CB4856, 1460/5171 = 28%),
350 whereas, after *P. rettgeri* infection, differential expression analysis found nearly equivalent
351 numbers of upregulated and downregulated genes in N2 and upregulation of less than one-third

352 of all DEGS in CB4856 (N2, 504/1029 =49%; CB4856, 694/2253 =31%) (Table 1, Table 2, Table S8,
353 Table S9).

354 We observed considerable differences in the overlap of gene expression profiles between N2
355 and CB4856 worms infected with *P. aeruginosa* or *P. rettgeri*. In animals infected with *P.*
356 *aeruginosa*, a majority of differentially expressed genes, 3092 in total, overlapped (N2-only,
357 2710/5802 = 47%; CB4856-only, 2079/5171 = 40%) (Figure 7A, Table S8), with a correlation of ρ
358 = 0.539 (p< 0.001). Like infection with *S. epidermidis* or *S. aureus*, the majority of DEGs identified
359 in both genotypes following *P. aeruginosa* infection were expressed in the same direction
360 (2971/3092 = 96%) (Figure 7C, Table S8). In response to *P. rettgeri*, 581 genes were differentially
361 expressed in both genotypes, representing a majority of DEGs in N2 but a minority in CB4856
362 (N2-only, 448/1029 = 44% CB4856-only, 1672/2253 = 74%) (Figure 7B, Table S9). Interestingly,
363 the vast majority of DEGs found in both genotypes following *P. rettgeri* infection were expressed
364 in opposing directions (564/581 = 97%), leaving only 17 genes that were coordinately expressed
365 (Figure 7C, Table S9). Finally, we identified 114 transcription factor genes that were differentially
366 expressed in both genotypes in the same direction after *P. aeruginosa* infection, whereas no
367 transcription factors were coordinately regulated in *P. rettgeri*-infected animals (Table S8, Table
368 S9).

369 Taken together, these results indicate that the N2 and CB4856 transcriptomic responses to *P.*
370 *aeruginosa* have more in common, whereas little overlap in differential gene expression exists
371 when N2 and C4856 are infected with *P. rettgeri*. In fact, of all the gene expression profiles
372 compared, *P. rettgeri* infection was the only one to have a negative correlation (ρ = -0.387; p<

373 0.001). These results are consistent with the responses to infection being driven both by the
374 isolate and the pathogen.
375
376 **Geneset Enrichment Analysis Identifies Unique GO Terms Enriched in N2 and CB4856 Worms**
377 **Exposed to *S. epidermidis***
378 Since N2 and CB4856 exhibited survival differences after *S. epidermidis* infection but not *S.*
379 *aureus* infection, we were interested in identifying biological processes that may be enriched in
380 these animals. We cross-referenced DEGs identified in N2 following *S. epidermidis* and *S. aureus*
381 infection and identified 556 genes that were either exclusively expressed in response to *S.*
382 *epidermidis* or were expressed in response to both pathogens but in opposing directions (*e.g.,*
383 upregulated in *S. epidermidis* and downregulated in *S. aureus*). The same analysis performed for
384 DEGs in CB4856 animals identified 749 genes. Only 73 of these genes were differentially
385 expressed in both N2 and CB4856 (N2-only, 483/556 = 87%; CB4856-only, 676/749 = 90%) (Table
386 S10).
387 We performed geneset enrichment analysis on genes that were exclusive to N2 (N2-only),
388 genes that were exclusive to CB4856 (CB4856-only), and genes that were differentially expressed
389 in both N2 and CB4856 animals (N2 & CB4856). Biological functions pertaining to the extracellular
390 matrix, stress response, and metabolism were enriched in N2 and CB4856 worms following *S.*
391 *epidermidis* infection (Table 4, Table S11). The Gene Ontology (GO) terms “Extracellular material:
392 collagen” and “Extracellular material: cuticlin” were enriched in N2-exclusive DEGs, whereas
393 genes corresponding to the GO terms “Stress response: C-type Lectin” and “Metabolism: lipid”

394 were overrepresented in DEGs exclusive to CB4856. Genes in the GO category "Stress response:
395 detoxification" were overrepresented in both isolates, even if the individual genes were not
396 necessarily the same. Genes that were differentially expressed in both isolates were enriched for
397 a single GO term, "Stress response: pathogen". A separate group of genes that was expressed
398 exclusively in CB4856 animals was also enriched for this GO term. Notably, we did not find
399 significant enrichment of the GO term "Stress response: pathogen" in DEGs exclusive to N2
400 animals. Thus, when only considering genes that are differentially expressed in an *S. epidermidis*-
401 specific manner, our results indicate a common enrichment of biological processes related to
402 stress response and innate immunity in N2 and CB4856 animals. At the same time, we also
403 observe an N2-specific enrichment of extracellular matrix genes and a CB4856-specific
404 enrichment of lipid metabolism genes, suggesting a degree of divergence in the transcriptomic
405 response of these isolates to *S. epidermidis* infection.

406

407 **Comparing DEGs between N2 and CB4856 fed OP50 versus *S. epidermidis***

408 The genomes of N2 and CB4856 animals differ at approximately one polymorphism per kb
409 and gene expression between these two isolates, including the expression of innate immune
410 genes, is known to vary across similar developmental stages, even in the absence of infection.[37-
411 39] Since principal component analysis found that genotype is the greatest predictor of isolate-
412 specific differences in gene expression following *S. epidermidis* infection (Figure 3), we performed
413 an additional gene expression analysis to identify genes that were differentially expressed

414 between N2 and CB4856 animals fed either the non-pathogenic *E. coli* OP50 or the pathogenic *S.*
415 *epidermidis*.

416 Using N2 expression as a baseline, we found 3384 and 1892 genes to be differentially
417 expressed in CB4856 animals fed *E. coli* OP50 and *S. epidermidis*, respectively. Upregulated genes
418 outnumbered those that were downregulated at a ratio of 2:1, regardless of whether the worms
419 were fed OP50 or *S. epidermidis* (OP50 Upregulated, 2121/3384 = 63%; OP50 Downregulated,
420 1263/3384 =37%) (*S. epidermidis* Upregulated, 1245/1892 = 66%; *S. epidermidis* Downregulated,
421 647/1892 =34%) (Table S12, Table S13).

422 Next, we cross-referenced genes that were differentially expressed between N2 and CB4856
423 animals in the OP50 and *S. epidermidis* conditions and classified them into four categories based
424 on changes in gene expression, or lack thereof. Relative to N2, a total of 2280 genes were
425 differentially expressed in CB4856 animals fed OP50 but were no longer differentially expressed
426 when worms were fed *S. epidermidis*. This suggests that the final level of expression for these
427 gene was similar in N2 and CB4856 animals following infection. A total of 1072 genes were
428 differentially expressed between N2 and CB4856 animals fed OP50 and continued to be
429 differentially expressed in the same direction in worms fed *S. epidermidis*, suggesting that they
430 are likely uninvolved in the transcriptional response to the pathogen. A third group of genes, 788
431 in total, were exclusively differentially expressed between N2 and CB4856 animals fed *S.*
432 *epidermidis*. Finally, we identified 32 genes that were differentially expressed between N2 and
433 CB4856 animals fed either OP50 or *S. epidermidis* but in opposing directions, *i.e.*, up in CB4856
434 fed OP50 and down in CB4856 fed *S. epidermidis* (Table S14).

435 The isolate- and pathogen-specific expression pattern of genes from these latter two
436 categories prompted us to perform geneset enrichment analysis to identify biological processes
437 that may be differentially enriched in N2 and CB4856 animals following *S. epidermidis* infection.
438 The 788 genes that were exclusively differentially expressed in CB4856 animals fed *S. epidermidis*
439 were enriched for GO terms pertaining to immunity (Stress response: C-type lectin), cell signaling
440 and transduction (Signaling: phosphatase; Signaling: tyrosine kinase), reproduction (Major sperm
441 protein), posttranslational modification (Protein modification: glycoprotein), and cytoskeletal
442 movement (Cytoskeleton: microtubule). From the 32 genes that were differentially expressed in
443 opposing directions in CB4856 animals fed either OP50 or *S. epidermidis* geneset enrichment
444 analysis identified GO terms corresponding to immunity (Stress response: pathogen) and the
445 extracellular matrix (Extracellular material: collagen) (Table 5) (Table S15). Thus, by considering
446 differences in gene expression between N2 and CB4856 animals fed non-pathogenic *E. coli* OP50,
447 we have identified additional sets of genes that are enriched for GO terms pertaining to, among
448 other categories, stress response, innate immunity, and the extracellular matrix.
449 Finally, to better understand the transcriptional regulators that might be significantly
450 different upon infection by *S. epidermidis,* we looked for transcription factors that were
451 differentially expressed only upon infection. There were 22 protein coding genes (1 pseudogene),
452 that fit these criteria. Interestingly, 13 of the 22 were members of the nuclear hormone receptor
453 (*nhr*) family. There are 266 *nhr* genes in our list of 1077 transcription factors. The incidence of
454 *nhr* genes in our DEG list was significantly enriched (P = 0.007, Fisher's exact test, odds ratio 4.4).
455

456 ## Discussion

457 We conducted survival, behavioral and transcriptional assays using the *C. elegans* strains N2
458 and CB4856 to identify differences in pathogen susceptibility between two wild-type isolates.
459 CB4856 was significantly more susceptible than N2 to infection by the Gram-positive pathogen
460 *S. epidermidis*. CB4856 and N2 were equally susceptible to a related species, *S. aureus,* and two
461 non-*Staphylococcus* pathogens, *P. aeruginosa* and *P. rettgeri.* We confirmed the CB4856
462 susceptibility using a second isolate of *S. epidermidis*.

463 What might underlie the differences in susceptibility? Initially we hypothesized that, perhaps,
464 the CB4856 strain lacked the ability to detect *S. epidermidis* as pathogenic. However, CB4856
465 worms avoided *S. epidermidis,* when given *E. coli* as an alternative, suggesting they were averse
466 to the bacteria. Further, we found robust changes in gene expression in response to the *S.*
467 *epidermis* infection. Unexpectedly, we found that, despite *P. rettgeri* being equally pathogenic to
468 both isolates, CB4856 animals preferred *P. rettgeri* to *E. coli,* while N2 did not. This observation
469 was also highlighted by the fact that changes in gene expression were negatively correlated in
470 these two isolates in response to *P. rettgeri.*

471

472 **Isolates exhibit distinct transcriptomic responses to infection**

473 The pathogens used in this study are diverse both in terms of their genome size and known
474 virulence mechanisms.[40-43] Principal component analysis of the sequencing dataset found that
475 gene expression changes during infection clustered largely by genotype, rather than pathogen,
476 even in instances where susceptibility was equivalent. At a glance, this suggests that the

477 strategies these strains have adopted for dealing with microbial infection have diverged since
478 they were separated from one another.

479 When we looked for genes that were differentially expressed in response to infection, *i.e.,*
480 was identified in response to all pathogens, there were hundreds, compared to the lists for
481 individual pathogen (thousands). Further, we found only six DEGs in common between N2 and
482 CB4856, with those genes largely uncharacterized functionally. Overall, these results suggest that
483 N2 and CB4856 respond to infection by altering gene transcription, but the core infection
484 response represents only a small fraction of the total genes regulated, and that the core
485 transcriptional response in each background has diverged. It also suggests there are still a
486 number of genes involved in immunity that need to be characterized.

487 By categorizing the genes that were differentially expressed in response to all pathogens, we
488 found an overrepresentation of genes involved in lipid metabolism in both N2- and CB4856-
489 infected animals. Lipid metabolism has been shown to be required for proper innate immune
490 system activation and pathogen defense.[44] These genes could play similar roles in both isolates
491 and thus, while they may reflect a divergence in specific gene activation, they could represent a
492 similarity in how N2 and CB4856 animals utilize lipid metabolism enzymes to facilitate the innate
493 immune response.

494

495 **Gram-stain level analysis suggests an absence of conserved master regulatory pathways**

496 In many organisms, the Toll signaling pathway is activated by infection with Gram-positive
497 bacteria and/or fungi, while the IMD pathway is activated in response to Gram-negative bacteria.

498 Thus, we examined the results grouping pathogens by Gram-stain, and then compared between
499 isolates. Our results found that the majority of DEGs were strain-specific, consistent with a
500 divergence of the responses to pathogens, but that, generally, there was a stronger conservation
501 of responses to Gram-positive bacteria, compared to Gram-negative. Our results are consistent
502 with the evidence that Toll and IMD signaling is diminished or absent in *C. elegans* as master
503 regulators of immunity and suggests each isolate deals with Gram-negative and Gram-positive
504 pathogens differently.

505 We examined the list of DEGs for transcription factors that could be contributing to the
506 differential expression profiles in each of the isolates. There were hundreds of transcription
507 factors regulated by infection with Gram-positive pathogens, with 50 overlapping between the
508 N2 and CB4856 isolates. In contrast, for Gram-negative pathogens, N2 had only nine differentially
509 expressed transcription factors, while CB4856 had 97, with none in common.

510 One possible explanation for the large overlap in N2 and CB4856 transcriptomic responses to
511 Gram-positive bacteria is the identity of the microbial pathogens. *S. aureus* and *S. epidermidis*
512 are both members of the *Staphylococcus* genus, share a number of virulence factors and invoke
513 overlapping host immune responses following infection.[41,45,46] Thus, the greater proportion of
514 overlapping DEGs in N2 and CB4856 animals infected with Gram-positive bacteria may reflect a
515 common response to *Staphylococcus* pathogens. Similar results for infection of N2 animals by
516 *Enterococcus faecalis* and *Enterococcus faecium* suggest that, within species of pathogens, there
517 are core transcriptional similarities.[47]

518

519 **Individual pathogens induce different responses in *C. elegans* isolates**

520 Since N2 and CB4856 worms exhibited similar susceptibility to infection by *P. aeruginosa*, *S.*
521 *aureus* or *P. rettgeri*, we posited that RNA sequencing would reveal considerable overlap in gene
522 expression between the two isolates. Indeed, after infecting N2 and CB4856 worms with either
523 *P. aeruginosa* or *S. aureus* we identified thousands of differentially expressed genes, a majority
524 of which were coordinately regulated in both isolates, supporting the idea that infection by either
525 of these pathogens invokes similar kinds of transcriptomic responses in N2 and CB4856 animals.
526 However, animals infected with *P. rettgeri* exhibited very little overlap in gene expression with
527 fewer than 20 genes differentially expressed in both isolates and expressed in the same direction.

528 It is possible that N2 and CB4856 worms have adopted two very different immune responses
529 to *P. rettgeri* infection, with neither providing a perceivable survival advantage. But why do we
530 not observe a similar lack of overlap in gene expression after infection with the Gram-negative
531 pathogen, *P. aeruginosa*, which also showed similar lethality in N2 and CB4856 animals? Analysis
532 of gene expression at regular intervals throughout an infection has demonstrated that the host
533 immune response is a dynamic process which needs to be initiated and sustained to fight an
534 infection.[48,49] At 24 hours post-infection, the timepoint at which RNA was collected for gene
535 expression analysis, animals infected with *P. aeruginosa* were much closer to dying than animals
536 infected with *P. rettgeri* (N2 *P. aeruginosa* LT_{50} = 3.8 ± 0.3 days, CB4856 *P. aeruginosa* LT_{50} = 3.5
537 ± 0.3 days; N2 *P. rettgeri* LT_{50} = 7.8 ± 0.7 days, CB4856 *P. rettgeri* LT_{50} = 7.2 ± 0.6 days). Perhaps
538 animals exposed to *P. aeruginosa* or *P. rettgeri* are at two different stages in combatting their

539 respective infections and the overlap in gene expression between N2 and CB4856 animals grows
540 larger as animals progress from earlier to later stages of infection.
541
542 **Changes in gene expression by infection with *S. epidermidis***
543 We found differences in survival in response to *S. epidermidis.* For each isolate we narrowed
544 down the list of differentially expressed genes to those that were expressed in a *S. epidermidis*-
545 specific manner and performed geneset enrichment analysis on sets of genes that were either
546 exclusive to N2, exclusive to CB4856, or expressed in both N2 and CB4856 animals. We observed
547 an enrichment of genes encoding extracellular matrix (ECM) proteins, such as collagens and
548 cuticlins, exclusively in N2 animals. Cuticular collagens are an integral component of the *C.*
549 *elegans* epidermal barrier and have established roles in the host epidermal immune response to
550 pathogens.[50-52] Further, ECM proteins are likely involved in promoting longevity as changes in
551 collagen gene expression have been used as *in vivo* biomarkers in *C. elegans* aging research.[53]
552 Together, our data reveal isolate-specific changes in ECM gene expression after *S. epidermidis*
553 exposure, suggesting a role for collagen genes in the increased resistance of N2 worms to *S.*
554 *epidermidis* infection. In DEGs exclusive to *S. epidermidis*-infected N2 animals as well as DEGs
555 exclusive to *S. epidermidis*-infected CB4856 animals, our analysis detected an enrichment of
556 genes encoding detoxification enzymes. These enzymes are primarily responsible for
557 metabolizing xenobiotic substances and have demonstrated roles in the *C. elegans* immune
558 response.[54,55] In both N2 and CB4856 animals, the DEGs were either members of the Cytochrome
559 P450 (CYP) family or possessed UDP glucuronosyltransferase (UGT) activity. It is possible that

560 these genes play similar roles in both isolates and thus may reflect a divergence in how N2 and
561 CB4856 animals utilize detoxification enzymes to counteract the microbial infection.

562 Finally, we took advantage of the ability to look for genes with expression differences
563 between N2 and CB4856 on control bacteria. This allowed us to sort differentially expressed
564 genes in the *S. epidermidis* condition into those that started out different, but ended up being
565 similar, and those that changed as a function of infection to end up at different levels of
566 expression. *A priori* we might consider the former group as being less likely to underlie the
567 differences in susceptibility. In those analyses, approximately 90% of the genes were upregulated
568 in response to infection.

569 When we examined the transcription factors that were differentially expressed, we found an
570 overrepresentation of nuclear hormone receptors. What is more, of the 13 *nhr* genes, seven were
571 upregulated, while six were downregulated during infection. Nuclear hormone receptors are
572 used in many contexts to systemically coordinate signaling. In *C. elegans* multiple *nhr* genes have
573 been linked to innate immunity, including *nhr-8, nhr-23, nhr-25, nhr-45, nhr-49, nhr-57, nhr-80,*
574 and *nhr-112.*[56-63] Thus, this is an interesting family of genes to potentially contribute to
575 differences in organismal response and/or susceptibility *in vivo.*

576

577 **Conclusions**

578 In populations, genetic and genomic diversity drive phenotypes. Immune signaling is no
579 exception, and yet, our understanding of *C. elegans* immunity has largely been driven by the
580 analysis of the canonical laboratory strain, N2. We found that the N2 strain and CB4856 wild

581 isolate were significantly different when we measured the susceptibility to *S. epidermidis.* It is
582 possible that the relative abundance of *S. epidermidis* or related pathogens in their respective
583 ecological niches resulted in the shift in immune capacity between the N2 and CB4856 lines.
584 Unfortunately, without sufficient metagenomic sampling of the environment, this would be
585 difficult to address. However, the robust differences observed suggests that isolating the animals
586 and transferring them to the laboratory environment has fixed genetic variation that underlies
587 pathogen-specific susceptibility.

588 Perhaps more surprisingly, we found significantly divergent transcriptomic responses
589 between these isolates, even when pathogen susceptibility was equivalent. These differences
590 could come from protein coding variants in the molecules that detect pathogens or the
591 transcription factors that coordinate immune defenses, non-coding variants in regulatory
592 regions, or, more likely, a combination of the two. By narrowing down the regions of the genome
593 that are important for the observed differences, we will be able to better understand how genetic
594 variation contributes to different immune outcomes.

595

596 **Acknowledgements**

597 We thank Dr. Robert Unckless for his helpful advice in sequencing and identifying the *S.*
598 *epidermidis* strain. We also thank Dr. Erik Lundquist and his laboratory for helpful discussions.
599 PAL was supported by an IRACDA award from the NIH (K12GM63651) and BDA was partially
600 supported by an award from the KU Center for Chemical Biology of Infectious Diseases
601 (P20GM113117). The KU Genome Sequencing Core is partially supported by the KU Center for
602 the Molecular Analysis of Disease Pathways (P20GM103638).

603

604 **Figures**

605

606

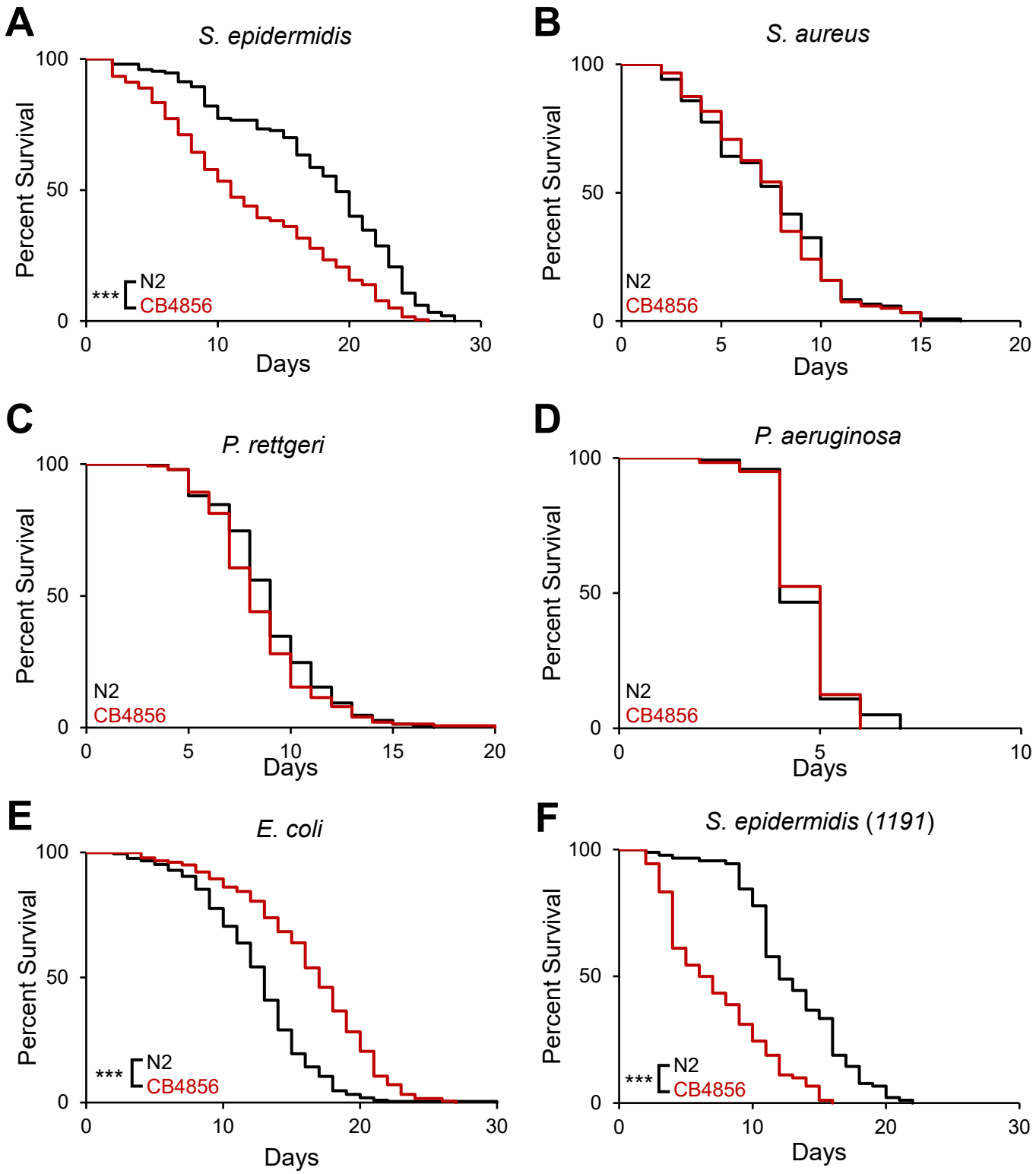


Figure 1: CB4856 is more susceptible to *S. epidermidis,* but not other bacterial pathogens, relative to N2.
(A-E) Survival of the *C. elegans* Hawaiian wild isolate, CB4856, versus N2 on nematode growth medium (NGM) with a lawn of *S. epidermidis* (A), *S. aureus* (B), *P. rettgeri* (C), *P. aeruginosa* (D), or *E. coli* (E). We tested a second *S. epidermidis* strain (1191) to confirm the differences in genotypes (F). In each experiment, 30 worms were placed on the bacterial lawn and transferred every day while scored for survival. Survival curves represent 3-7 independent experiments. For each pathogen, statistical analysis by Mantel-Cox (CB486 vs N2). ***, $p<0.001$.

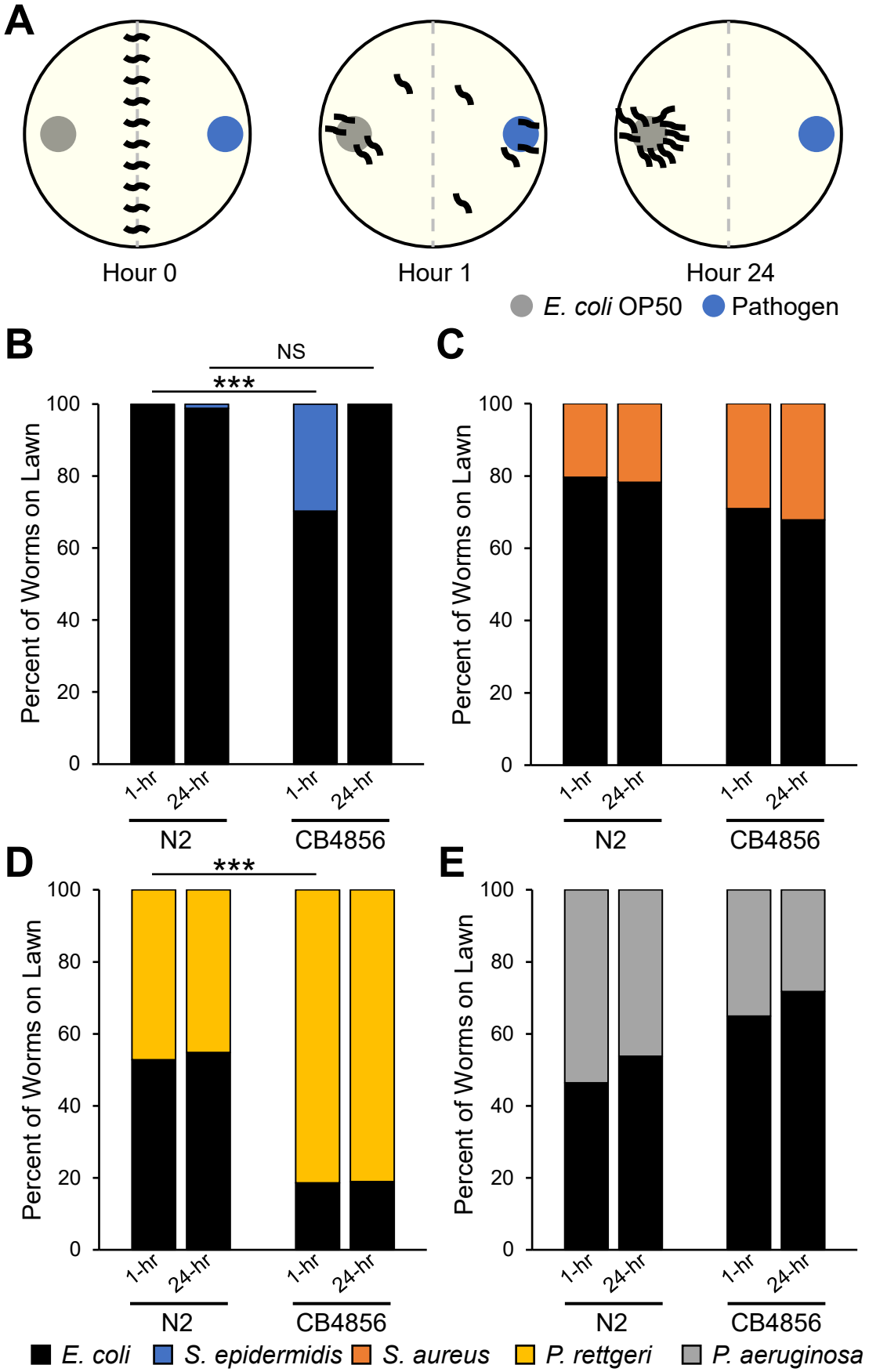


Figure 2: Pathogen avoidance of N2 and CB4856 isolates depends on the pathogen and time of observation.

(A) Two-choice pathogen avoidance assay. L4 worms were washed and then placed at the midline of test plates containing two bacterial lawns, OP50 or pathogen. At 1-hr and 24-hr timepoints the percentage of worms on each lawn was calculated. (B-E) Percentage of N2 and CB4856 isolates on a bacterial lawn when given a choice of *E. coli* OP50 or. *S. epidermidis* (B), *S. aureus* (C), *P. rettgeri* (D) or *P. aeruginosa* (E). In each experiment, 20 worms were placed on NGM plates and the number of worms on the bacterial lawn was recorded at regular intervals. Data represent 5 independent experiments. ***, $p<0.001$; NS, $p>0.05$.

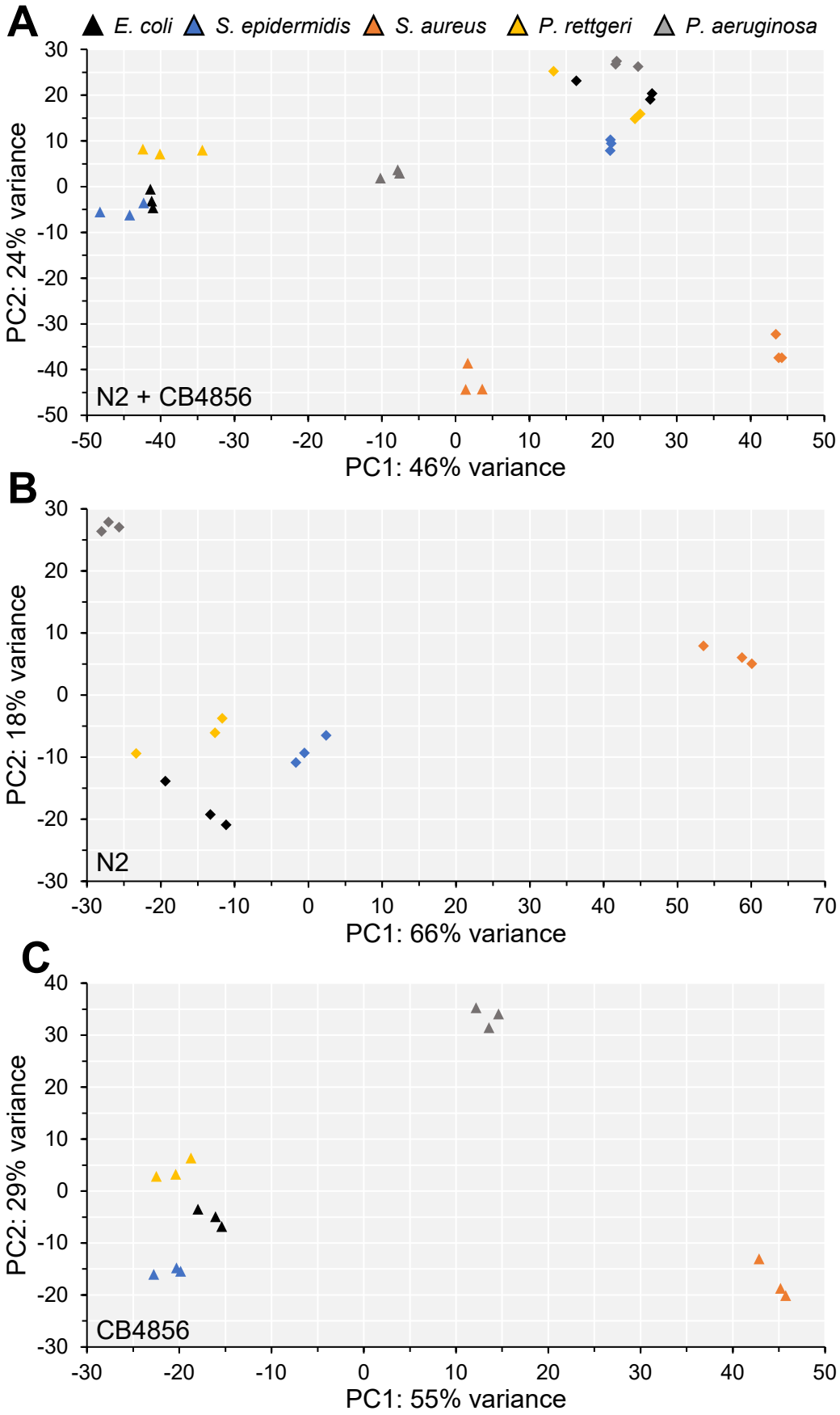


Figure 3: Principal component analysis clusters samples based on genotype, rather than pathogen exposure.
(A-C) Principal component plots of expression data from the Hawaiian wild isolate, CB4856 (triangles) and N2 (diamonds) samples (A), only N2 samples (B), and only CB4856 samples (C).

Table 1: Differentially expressed genes (DEGs) in N2 animals following 24-hour exposure to microbial pathogens.

Condition	[1]Total DEGs	Total Upregulated	Total Downregulated	[2]Total Genes Expressed
All Pathogens	273	236	37	23938
Gram Positive	3833	2964	869	22934
Gram Negative	628	254	374	21646
Staphylococcus epidermidis	2384	1754	630	21124
Staphylococcus aureus	8754	4901	3853	21433
Pseudomonas aeruginosa	5802	1804	3998	20304
Providencia rettgeri	1029	504	525	20295
Escherichia coli	-	-	-	19961

[1]Genes with a false discovery rate (FDR)-adjusted p-value of <0.05 and log2 fold change >1
[2]Total number of genes with >0 counts

Table 2: Differentially expressed genes (DEGs) in CB4856 animals following 24-hour exposure to microbial pathogens.

Condition	[1]Total DEGs	Total Upregulated	Total Downregulated	[2]Total Genes Expressed
All Pathogens	117	96	21	24333
Gram Positive	2565	1962	603	23297
Gram Negative	1720	456	1264	21847
Staphylococcus epidermidis	1776	1296	480	21252
Staphylococcus aureus	7033	3597	3436	21675
Pseudomonas aeruginosa	5171	1460	3711	20963
Providencia rettgeri	2253	694	1559	19731
Escherichia coli	-	-	-	20288

[1]Genes with a false discovery rate (FDR)-adjusted p-value of <0.05 and log2 fold change >1
[2]Total number of genes with >0 counts

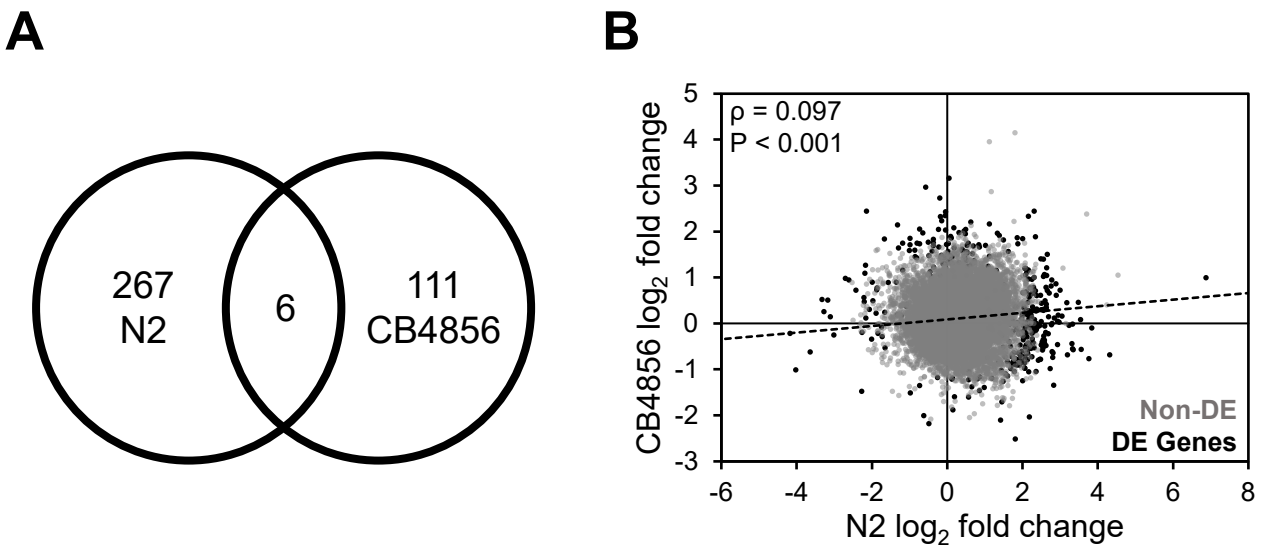
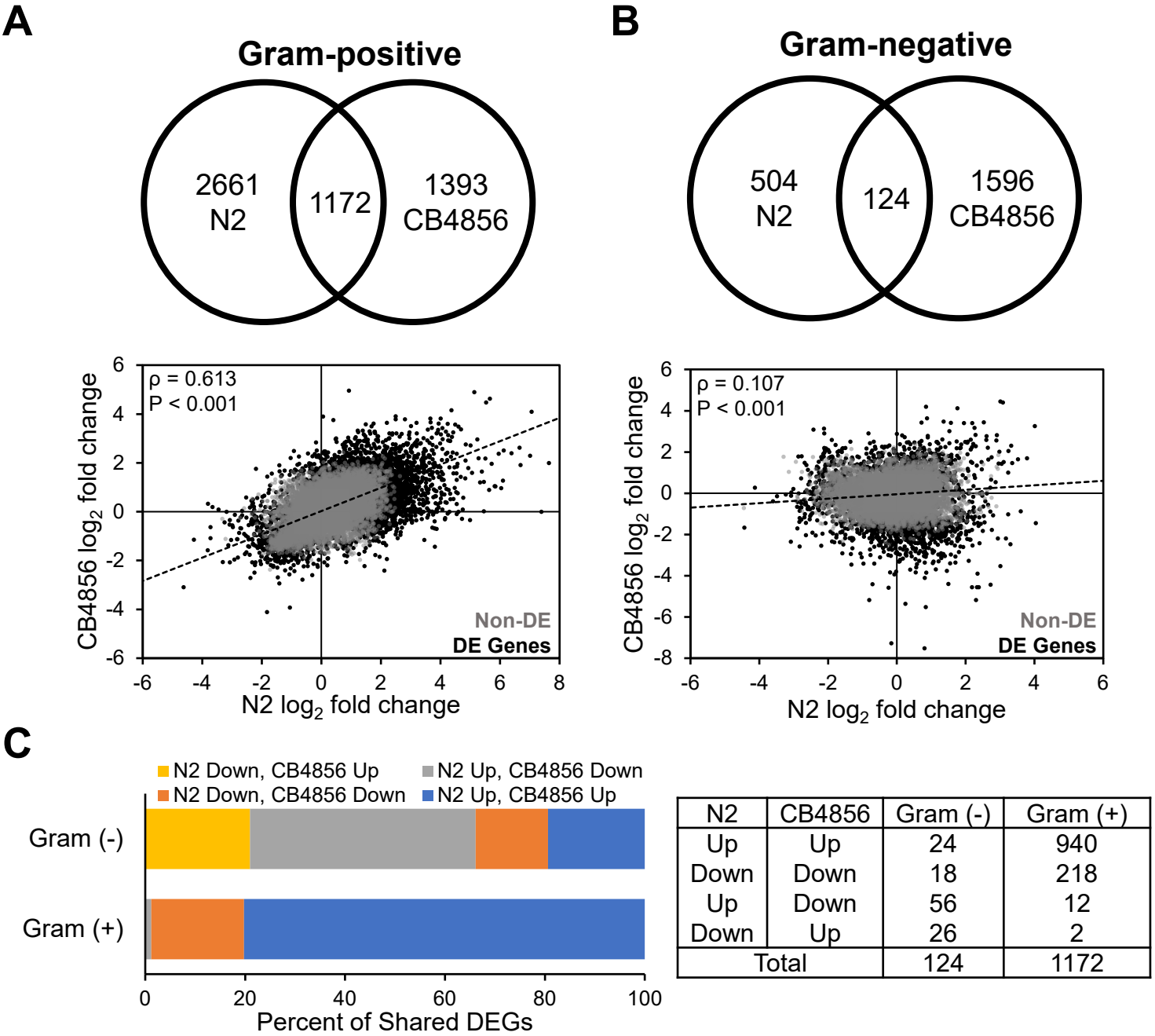


Figure 4: Expression differences in N2 and CB4856 animals after infection with pathogenic bacteria.
(A) Intersection of differentially expressed genes and (B) correlation between $\log_2$ fold changes in N2 and CB4856 animals infected with pathogenic bacteria. Spearman's rho value and P-values are presented in the upper left corner. Gray (Non-DE) and black dots (DE Genes) represent expressed and differentially expressed genes, respectively.

Table 3: Enriched Gene Ontology Terms after Infection with Pathogenic Bacteria

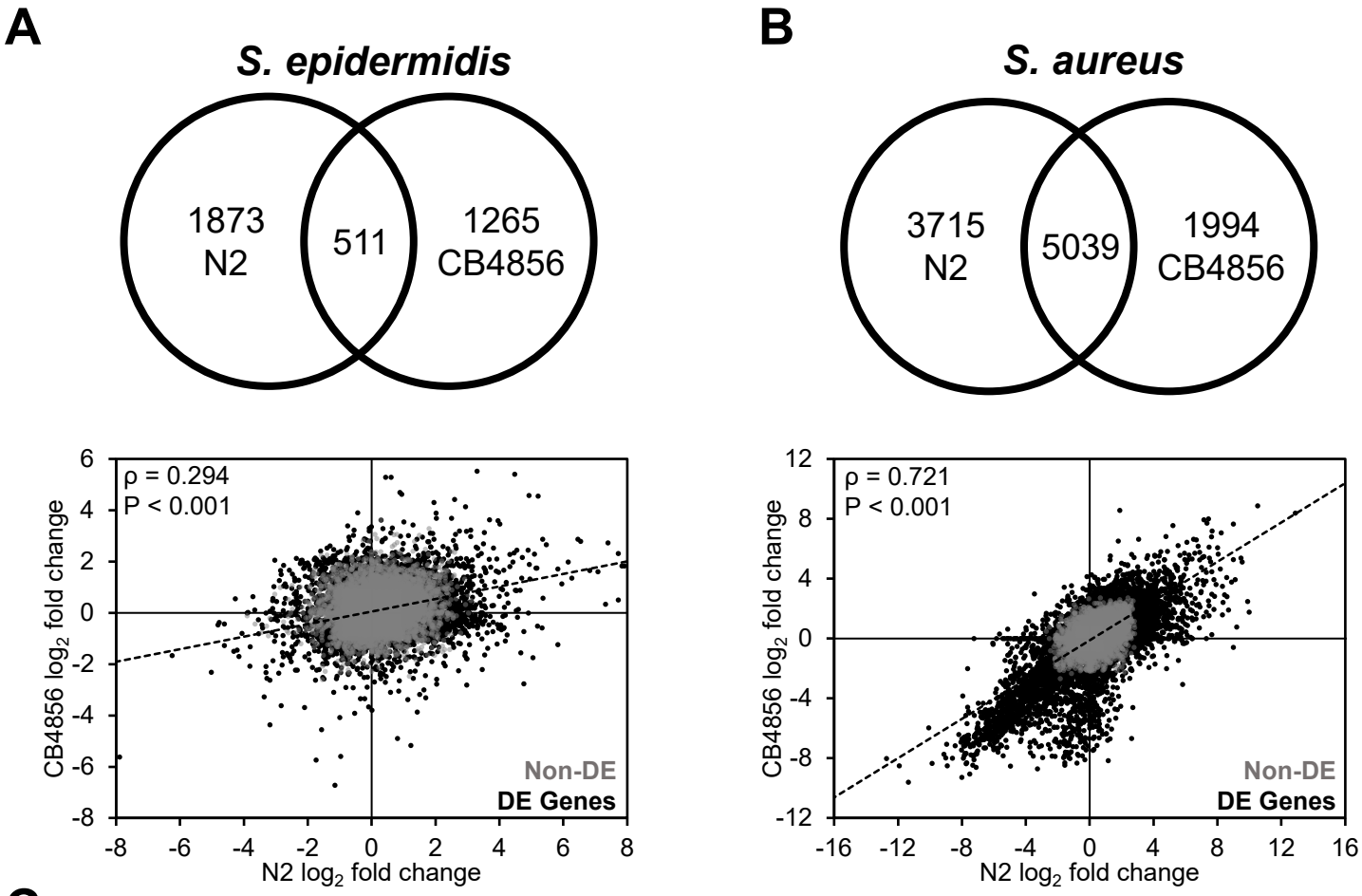
	Gene Ontology Term	Number of Genes	Percent[1]	FDR-adjusted p-value[2]
N2-only	Stress response: pathogen	13	7.3	9.4×10^{-7}
	Transmembrane transport: lipid	4	26.7	1.1×10^{-3}
	Metabolism: lipid	16	3.0	1.7×10^{-3}
CB4856-only	Metabolism: lipid	10	1.9	1.7×10^{-3}

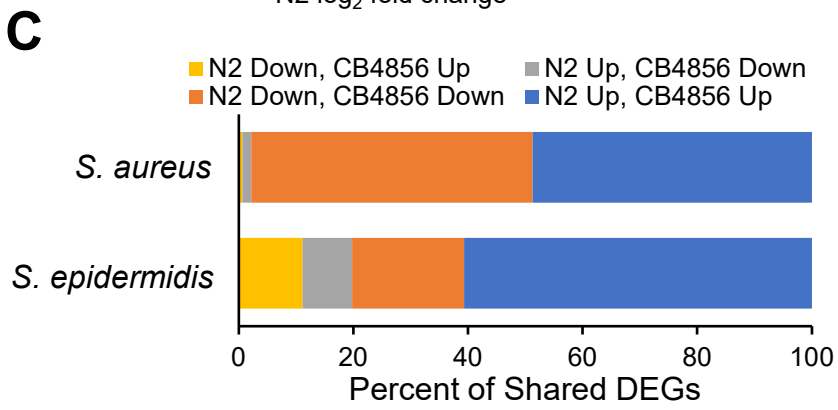
[1]Percentage of genes annotated with the indicated GO term that are differentially expressed
[2]FDR-adjusted *p*-value is a measure of enrichment (Fisher exact test; Bonferroni correction) of the GO term among differentially expressed genes



N2	CB4856	Gram (-)	Gram (+)
Up	Up	24	940
Down	Down	18	218
Up	Down	56	12
Down	Up	26	2
Total		124	1172

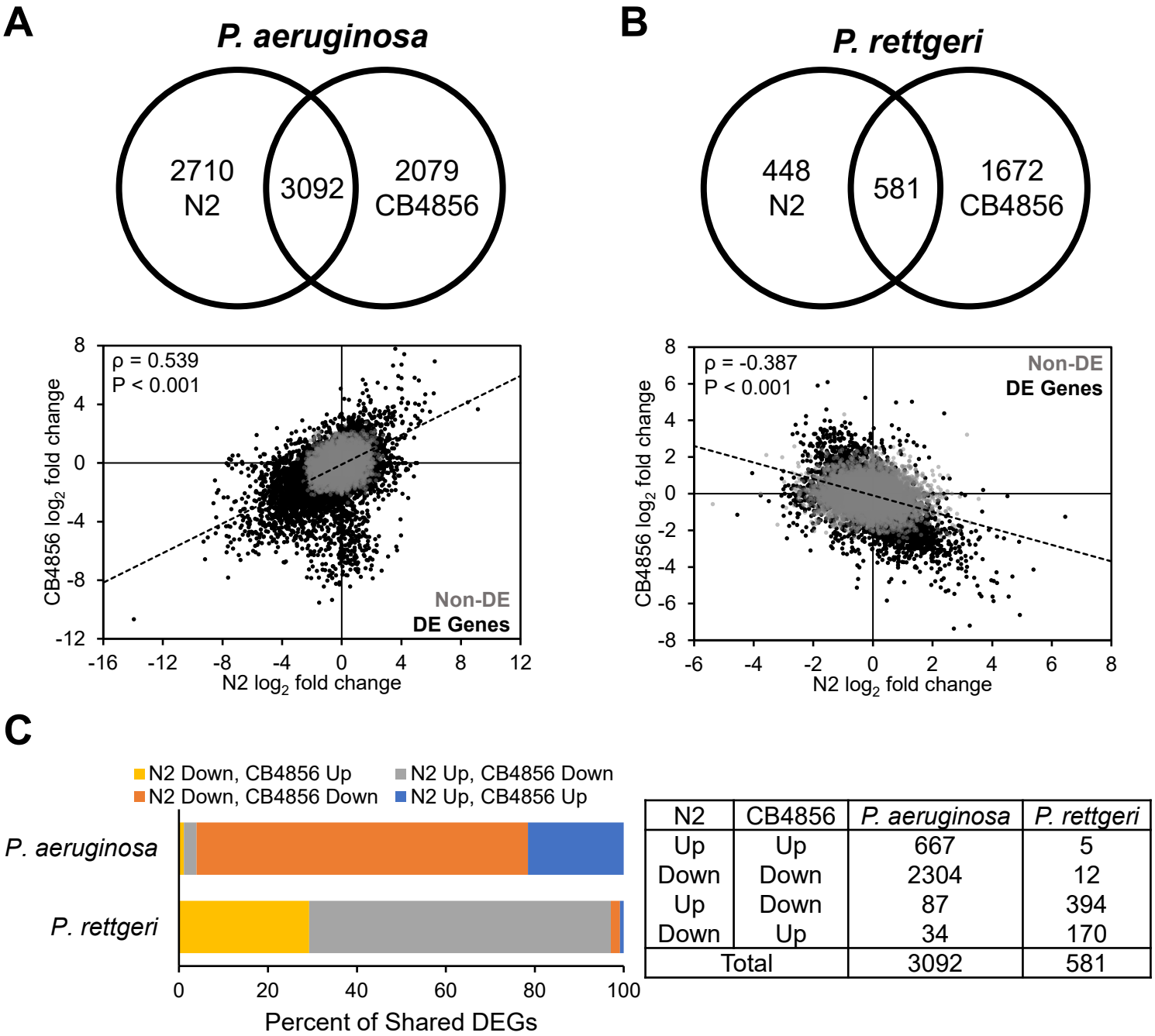
Figure 5: Expression differences in N2 and CB4856 animals infected with Gram-positive or Gram-negative pathogens.
(A) Intersection of differentially expressed genes (top) and correlation between $\log_2$ fold changes (bottom) in N2 and CB4856 animals infected with Gram-positive pathogens (*S. epidermidis, S. aureus*) or (B) Gram-negative pathogens (*P. rettgeri*, *P. aeruginosa*). Spearman's rho value and P-values are presented in the upper left corner. Gray (Non-DE) and black dots (DE Genes) represent expressed and differentially expressed genes, respectively. (C) The distribution of up- and down-regulated genes that are differentially expressed in both N2 and CB4856 animals infected with either Gram-negative or Gram-positive pathogens.





N2	CB4856	*S. aureus*	*S. epidermidis*
Up	Up	2456	310
Down	Down	2472	100
Up	Down	79	44
Down	Up	32	57
Total		5039	511

Figure 6: Expression differences in N2 and CB4856 animals infected with the gram-positive pathogens, *S. epidermidis* or *S. aureus*.
(A) Intersection of differentially expressed genes (top) and correlation between $\log_2$ fold changes (bottom) in N2 and CB4856 animals infected with either *S. epidermidis* or (B) *S. aureus*. Spearman's rho value and P-values are presented in the upper left corner. Gray (Non-DE) and black dots (DE Genes) represent expressed and differentially expressed genes, respectively. (C) The distribution of up- and down-regulated genes that are differentially expressed in both N2 and CB4856 animals infected with either *S. aureus* or *S. epidermidis*.



N2	CB4856	*P. aeruginosa*	*P. rettgeri*
Up	Up	667	5
Down	Down	2304	12
Up	Down	87	394
Down	Up	34	170
Total		3092	581

Figure 7: Expression differences in N2 and CB4856 animals infected with the gram-negative pathogens, *P. rettgeri* or *P. aeruginosa*.
(A) Intersection of differentially expressed genes (top) and correlation between $\log_2$ fold changes (bottom) in N2 and CB4856 animals infected with either *P. rettgeri* or (B) *P. aeruginosa*. Spearman's rho value and P-values are presented in the upper left corner. Gray (Non-DE) and black dots (DE Genes) represent expressed and differentially expressed genes, respectively. (C) The distribution of up- and down-regulated genes that are differentially expressed in both N2 and CB4856 animals infected with either *P. rettgeri* or *P. aeruginosa*.

Table 4: Enriched Gene Ontology Terms in N2 and CB4856 Worms Exposed to *S. epidermidis*

	Gene Ontology Term	Number of Genes	Percent[1]	FDR-adjusted p-value[2]
N2-only	Extracellular material: collagen	55	29.7	4.5×10^{-44}
	Extracellular material: cuticlin	7	20.0	1.5×10^{-4}
	Stress response: detoxification	13	8.6	2.8×10^{-4}
CB485-only	Stress response: pathogen	22	12.4	4.6×10^{-8}
	Stress Response: C-type Lectin	26	10.1	7.8×10^{-8}
	Stress response: detoxification	20	13.2	1.0×10^{-7}
	Metabolism: lipid	30	5.7	5.2×10^{-4}
N2 & CB4856	Stress response: pathogen	7	4.0	1.4×10^{-5}

[1]Percentage of genes annotated with the indicated GO term that are differentially expressed
[2]FDR-adjusted *p*-value is a measure of enrichment (Fisher exact test; Bonferroni Correction) of the GO term among differentially expressed genes

Table 5: Enriched Gene Ontology Terms when Comparing N2 and CB4856 Worms fed OP50 versus *S. epidermidis*

	Gene Ontology Term	Number of Genes	Percent[1]	FDR-adjusted p-value[2]
Genes expressed exclusively in worms fed *S. epidermidis*	Signaling: phosphatase	33	16.8	7.3×10^{-14}
	Cytoskeleton: microtubule	24	18.8	6.3×10^{-11}
	Major sperm protein	12	38.7	3.3×10^{-8}
	Signaling: tyrosine kinase	12	16.0	1.4×10^{-4}
	Protein modification: carbohydrate	15	11.8	2.4×10^{-4}
	Stress response: C-type Lectin	20	7.8	2.1×10^{-3}
Genes expressed in opposing directions in worms fed *E. coli* OP50 or *S. epidermidis*	Extracellular material: collagen	7	3.8	4.5×10^{-8}
	Stress response: pathogen	4	2.3	6.8×10^{-4}

[1]Percentage of genes annotated with the indicated GO term that are differentially expressed
[2]FDR-adjusted *p*-value is a measure of enrichment (Fisher exact test; Bonferroni Correction) of the GO term among differentially expressed genes

607 **Literature Cited**

608 1 Hancock, R. E., Nijnik, A. & Philpott, D. J. Modulating immunity as a therapy for bacterial
609 infections. *Nat Rev Microbiol* **10**, 243-254, doi:10.1038/nrmicro2745 (2012).
610 2 Mahlapuu, M., Hakansson, J., Ringstad, L. & Bjorn, C. Antimicrobial Peptides: An
611 Emerging Category of Therapeutic Agents. *Front Cell Infect Microbiol* **6**, 194,
612 doi:10.3389/fcimb.2016.00194 (2016).
613 3 Schroder, N. W. & Schumann, R. R. Single nucleotide polymorphisms of Toll-like
614 receptors and susceptibility to infectious disease. *Lancet Infect Dis* **5**, 156-164,
615 doi:10.1016/S1473-3099(05)01308-3 (2005).
616 4 Hayashi, F. *et al.* The innate immune response to bacterial flagellin is mediated by Toll-
617 like receptor 5. *Nature* **410**, 1099-1103, doi:10.1038/35074106 (2001).
618 5 Dunstan, S. J. *et al.* Host susceptibility and clinical outcomes in toll-like receptor 5-
619 deficient patients with typhoid fever in Vietnam. *J Infect Dis* **191**, 1068-1071,
620 doi:10.1086/428593 (2005).
621 6 Hawn, T. R. *et al.* A common dominant TLR5 stop codon polymorphism abolishes
622 flagellin signaling and is associated with susceptibility to legionnaires' disease. *J Exp Med*
623 **198**, 1563-1572, doi:10.1084/jem.20031220 (2003).
624 7 Leist, S. R. & Baric, R. S. Giving the Genes a Shuffle: Using Natural Variation to
625 Understand Host Genetic Contributions to Viral Infections. *Trends Genet* **34**, 777-789,
626 doi:10.1016/j.tig.2018.07.005 (2018).
627 8 Balla, K. M. & Troemel, E. R. Caenorhabditis elegans as a model for intracellular
628 pathogen infection. *Cell Microbiol* **15**, 1313-1322, doi:10.1111/cmi.12152 (2013).
629 9 Marsh, E. K. & May, R. C. Caenorhabditis elegans, a model organism for investigating
630 immunity. *Appl Environ Microbiol* **78**, 2075-2081, doi:10.1128/AEM.07486-11 (2012).
631 10 van der Hoeven, R., McCallum, K. C. & Garsin, D. A. Speculations on the activation of
632 ROS generation in C. elegans innate immune signaling. *Worm* **1**, 160-163,
633 doi:10.4161/worm.19767 (2012).
634 11 Chavez, V., Mohri-Shiomi, A., Maadani, A., Vega, L. A. & Garsin, D. A. Oxidative stress
635 enzymes are required for DAF-16-mediated immunity due to generation of reactive
636 oxygen species by Caenorhabditis elegans. *Genetics* **176**, 1567-1577,
637 doi:10.1534/genetics.107.072587 (2007).
638 12 Dierking, K., Yang, W. & Schulenburg, H. Antimicrobial effectors in the nematode
639 Caenorhabditis elegans: an outgroup to the Arthropoda. *Philos Trans R Soc Lond B Biol*
640 *Sci* **371**, doi:10.1098/rstb.2015.0299 (2016).
641 13 Schulenburg, H., Hoeppner, M. P., Weiner, J., 3rd & Bornberg-Bauer, E. Specificity of the
642 innate immune system and diversity of C-type lectin domain (CTLD) proteins in the
643 nematode Caenorhabditis elegans. *Immunobiology* **213**, 237-250,
644 doi:10.1016/j.imbio.2007.12.004 (2008).
645 14 Shivers, R. P. *et al.* Phosphorylation of the conserved transcription factor ATF-7 by PMK-
646 1 p38 MAPK regulates innate immunity in Caenorhabditis elegans. *PLoS Genet* **6**,
647 e1000892, doi:10.1371/journal.pgen.1000892 (2010).

648 15 Felix, M. A. *et al.* Natural and experimental infection of Caenorhabditis nematodes by
649 novel viruses related to nodaviruses. *PLoS Biol* **9**, e1000586,
650 doi:10.1371/journal.pbio.1000586 (2011).
651 16 Berg, M. *et al.* Assembly of the Caenorhabditis elegans gut microbiota from diverse soil
652 microbial environments. *ISME J* **10**, 1998-2009, doi:10.1038/ismej.2015.253 (2016).
653 17 Samuel, B. S., Rowedder, H., Braendle, C., Felix, M. A. & Ruvkun, G. Caenorhabditis
654 elegans responses to bacteria from its natural habitats. *Proc Natl Acad Sci U S A* **113**,
655 E3941-3949, doi:10.1073/pnas.1607183113 (2016).
656 18 de Bono, M. & Bargmann, C. I. Natural variation in a neuropeptide Y receptor homolog
657 modifies social behavior and food response in C. elegans. *Cell* **94**, 679-689,
658 doi:10.1016/s0092-8674(00)81609-8 (1998).
659 19 Chang, H. C., Paek, J. & Kim, D. H. Natural polymorphisms in C. elegans HECW-1 E3 ligase
660 affect pathogen avoidance behaviour. *Nature* **480**, 525-529, doi:10.1038/nature10643
661 (2011).
662 20 Schulenburg, H. & Muller, S. Natural variation in the response of Caenorhabditis elegans
663 towards Bacillus thuringiensis. *Parasitology* **128**, 433-443,
664 doi:10.1017/s003118200300461x (2004).
665 21 Reddy, K. C., Andersen, E. C., Kruglyak, L. & Kim, D. H. A polymorphism in npr-1 is a
666 behavioral determinant of pathogen susceptibility in C. elegans. *Science* **323**, 382-384,
667 doi:10.1126/science.1166527 (2009).
668 22 Balla, K. M., Andersen, E. C., Kruglyak, L. & Troemel, E. R. A wild C. elegans strain has
669 enhanced epithelial immunity to a natural microsporidian parasite. *PLoS Pathog* **11**,
670 e1004583, doi:10.1371/journal.ppat.1004583 (2015).
671 23 Martin, N., Singh, J. & Aballay, A. Natural Genetic Variation in the Caenorhabditis
672 elegans Response to Pseudomonas aeruginosa. *G3 (Bethesda)* **7**, 1137-1147,
673 doi:10.1534/g3.117.039057 (2017).
674 24 Brenner, S. The genetics of Caenorhabditis elegans. *Genetics* **77**, 71-94 (1974).
675 25 Galac, M. R. & Lazzaro, B. P. Comparative genomics of bacteria in the genus Providencia
676 isolated from wild Drosophila melanogaster. *BMC Genomics* **13**, 612, doi:10.1186/1471-
677 2164-13-612 (2012).
678 26 Rahme, L. G. *et al.* Common virulence factors for bacterial pathogenicity in plants and
679 animals. *Science* **268**, 1899-1902, doi:10.1126/science.7604262 (1995).
680 27 Duthie, E. S. & Lorenz, L. L. Staphylococcal coagulase; mode of action and antigenicity. *J*
681 *Gen Microbiol* **6**, 95-107, doi:10.1099/00221287-6-1-2-95 (1952).
682 28 FastQC: a quality control tool for high throughput sequence data. (2010).
683 29 Chen, S., Zhou, Y., Chen, Y. & Gu, J. fastp: an ultra-fast all-in-one FASTQ preprocessor.
684 *Bioinformatics* **34**, i884-i890, doi:10.1093/bioinformatics/bty560 (2018).
685 30 Kim, D., Paggi, J. M., Park, C., Bennett, C. & Salzberg, S. L. Graph-based genome
686 alignment and genotyping with HISAT2 and HISAT-genotype. *Nat Biotechnol* **37**, 907-
687 915, doi:10.1038/s41587-019-0201-4 (2019).

688 31 Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for
689 assigning sequence reads to genomic features. *Bioinformatics* **30**, 923-930,
690 doi:10.1093/bioinformatics/btt656 (2014).
691 32 Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion
692 for RNA-seq data with DESeq2. *Genome Biol* **15**, 550, doi:10.1186/s13059-014-0550-8
693 (2014).
694 33 Holdorf, A. D. *et al.* WormCat: An Online Tool for Annotation and Visualization of
695 Caenorhabditis elegans Genome-Scale Data. *Genetics* **214**, 279-294,
696 doi:10.1534/genetics.119.302919 (2020).
697 34 Glater, E. E., Rockman, M. V. & Bargmann, C. I. Multigenic natural variation underlies
698 Caenorhabditis elegans olfactory preference for the bacterial pathogen Serratia
699 marcescens. *G3 (Bethesda)* **4**, 265-276, doi:10.1534/g3.113.008649 (2014).
700 35 Luhachack, L. G. *et al.* EGL-9 controls C. elegans host defense specificity through prolyl
701 hydroxylation-dependent and -independent HIF-1 pathways. *PLoS Pathog* **8**, e1002798,
702 doi:10.1371/journal.ppat.1002798 (2012).
703 36 Cui, Y., McBride, S. J., Boyd, W. A., Alper, S. & Freedman, J. H. Toxicogenomic analysis of
704 Caenorhabditis elegans reveals novel genes and pathways involved in the resistance to
705 cadmium toxicity. *Genome Biol* **8**, R122, doi:10.1186/gb-2007-8-6-r122 (2007).
706 37 Wicks, S. R., Yeh, R. T., Gish, W. R., Waterston, R. H. & Plasterk, R. H. Rapid gene
707 mapping in Caenorhabditis elegans using a high density polymorphism map. *Nat Genet*
708 **28**, 160-164, doi:10.1038/88878 (2001).
709 38 Swan, K. A. *et al.* High-throughput gene mapping in Caenorhabditis elegans. *Genome Res*
710 **12**, 1100-1105, doi:10.1101/gr.208902 (2002).
711 39 Capra, E. J., Skrovanek, S. M. & Kruglyak, L. Comparative developmental expression
712 profiling of two C. elegans isolates. *PLoS One* **3**, e4055,
713 doi:10.1371/journal.pone.0004055 (2008).
714 40 Olaitan, A. O., Diene, S. M., Assous, M. V. & Rolain, J. M. Genomic Plasticity of
715 Multidrug-Resistant NDM-1 Positive Clinical Isolate of Providencia rettgeri. *Genome Biol*
716 *Evol* **8**, 723-728, doi:10.1093/gbe/evv195 (2016).
717 41 Sabate Bresco, M. *et al.* Pathogenic Mechanisms and Host Interactions in
718 Staphylococcus epidermidis Device-Related Infection. *Front Microbiol* **8**, 1401,
719 doi:10.3389/fmicb.2017.01401 (2017).
720 42 Manukumar, H. M. & Umesha, S. MALDI-TOF-MS based identification and molecular
721 characterization of food associated methicillin-resistant Staphylococcus aureus. *Sci Rep*
722 **7**, 11414, doi:10.1038/s41598-017-11597-z (2017).
723 43 Newman, J. W., Floyd, R. V. & Fothergill, J. L. The contribution of Pseudomonas
724 aeruginosa virulence factors and host factors in the establishment of urinary tract
725 infections. *FEMS Microbiol Lett* **364**, doi:10.1093/femsle/fnx124 (2017).
726 44 Anderson, S. M. & Pukkila-Worley, R. Immunometabolism in Caenorhabditis elegans.
727 *PLoS Pathog* **16**, e1008897, doi:10.1371/journal.ppat.1008897 (2020).
728 45 Gill, S. R. *et al.* Insights on evolution of virulence and resistance from the complete
729 genome analysis of an early methicillin-resistant Staphylococcus aureus strain and a

730 biofilm-producing methicillin-resistant Staphylococcus epidermidis strain. *J Bacteriol*
731 **187**, 2426-2438, doi:10.1128/JB.187.7.2426-2438.2005 (2005).
732 46 Le, K. Y., Park, M. D. & Otto, M. Immune Evasion Mechanisms of Staphylococcus
733 epidermidis Biofilm Infection. *Front Microbiol* **9**, 359, doi:10.3389/fmicb.2018.00359
734 (2018).
735 47 Yuen, G. J. & Ausubel, F. M. Both live and dead Enterococci activate Caenorhabditis
736 elegans host defense via immune and stress pathways. *Virulence* **9**, 683-699,
737 doi:10.1080/21505594.2018.1438025 (2018).
738 48 Guan, Y. *et al.* Whole-genome and time-course dual RNA-Seq analyses reveal chronic
739 pathogenicity-related gene dynamics in the ginseng rusty root rot pathogen Ilyonectria
740 robusta. *Sci Rep* **10**, 1586, doi:10.1038/s41598-020-58342-7 (2020).
741 49 Schlamp, F. *et al.* Dense time-course gene expression profiling of the Drosophila
742 melanogaster innate immune response. *BMC Genomics* **22**, 304, doi:10.1186/s12864-
743 021-07593-3 (2021).
744 50 Mesbahi, H., Pho, K. B., Tench, A. J., Leon Guerrero, V. L. & MacNeil, L. T. Cuticle
745 Collagen Expression Is Regulated in Response to Environmental Stimuli by the GATA
746 Transcription Factor ELT-3 in Caenorhabditis elegans. *Genetics* **215**, 483-495,
747 doi:10.1534/genetics.120.303125 (2020).
748 51 Sandhu, A., Badal, D., Sheokand, R., Tyagi, S. & Singh, V. Specific collagens maintain the
749 cuticle permeability barrier in Caenorhabditis elegans. *Genetics* **217**,
750 doi:10.1093/genetics/iyaa047 (2021).
751 52 Taffoni, C. & Pujol, N. Mechanisms of innate immunity in C. elegans epidermis. *Tissue*
752 *Barriers* **3**, e1078432, doi:10.1080/21688370.2015.1078432 (2015).
753 53 Statzer, C. *et al.* Youthful and age-related matreotypes predict drugs promoting
754 longevity. *Aging Cell*, e13441, doi:10.1111/acel.13441 (2021).
755 54 Pukkila-Worley, R. Surveillance Immunity: An Emerging Paradigm of Innate Defense
756 Activation in Caenorhabditis elegans. *PLoS Pathog* **12**, e1005795,
757 doi:10.1371/journal.ppat.1005795 (2016).
758 55 Simonsen, K. T., Gallego, S. F., Faergeman, N. J. & Kallipolitis, B. H. Strength in numbers:
759 "Omics" studies of C. elegans innate immunity. *Virulence* **3**, 477-484,
760 doi:10.4161/viru.21906 (2012).
761 56 Otarigho, B. & Aballay, A. Cholesterol Regulates Innate Immunity via Nuclear Hormone
762 Receptor NHR-8. *iScience* **23**, 101068, doi:10.1016/j.isci.2020.101068 (2020).
763 57 Sahu, S. N. *et al.* Genomic analysis of immune response against Vibrio cholerae
764 hemolysin in Caenorhabditis elegans. *PLoS One* **7**, e38200,
765 doi:10.1371/journal.pone.0038200 (2012).
766 58 Ward, J. D. *et al.* Defects in the C. elegans acyl-CoA synthase, acs-3, and nuclear
767 hormone receptor, nhr-25, cause sensitivity to distinct, but overlapping stresses. *PLoS*
768 *One* **9**, e92552, doi:10.1371/journal.pone.0092552 (2014).
769 59 Mao, K. *et al.* Mitochondrial Dysfunction in C. elegans Activates Mitochondrial
770 Relocalization and Nuclear Hormone Receptor-Dependent Detoxification Genes. *Cell*
771 *Metab* **29**, 1182-1191 e1184, doi:10.1016/j.cmet.2019.01.022 (2019).

772 60 Sim, S. & Hibberd, M. L. Caenorhabditis elegans susceptibility to gut Enterococcus
773 faecalis infection is associated with fat metabolism and epithelial junction integrity. *BMC*
774 *Microbiol* **16**, 6, doi:10.1186/s12866-016-0624-8 (2016).
775 61 Bellier, A., Chen, C. S., Kao, C. Y., Cinar, H. N. & Aroian, R. V. Hypoxia and the hypoxic
776 response pathway protect against pore-forming toxins in C. elegans. *PLoS Pathog* **5**,
777 e1000689, doi:10.1371/journal.ppat.1000689 (2009).
778 62 Zugasti, O. *et al.* A quantitative genome-wide RNAi screen in C. elegans for antifungal
779 innate immunity genes. *BMC Biol* **14**, 35, doi:10.1186/s12915-016-0256-3 (2016).
780 63 Shapira, M. *et al.* A conserved role for a GATA transcription factor in regulating epithelial
781 innate immune responses. *Proc Natl Acad Sci U S A* **103**, 14086-14091,
782 doi:10.1073/pnas.0603424103 (2006).
783