TITLE: Global transcriptional regulation of innate immunity in *C. elegans*

SHORT TITLE: p38 MAPK – ATF regulated immunity in *C. elegans*

AUTHORS: Marissa Fletcher[1], Erik J. Tillman[1], Vincent L. Butty[2], Stuart S. Levine[2], and Dennis H. Kim[1]*

AUTHOR AFFILIATIONS: [1]Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

[2]BioMicro Center, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

*Corresponding author

E-mail: dhkim@mit.edu

KEYWORDS: *C. elegans*, p38 MAPK signaling, innate immunity, PMK-1, ATF-7

Abstract

The nematode *Caenorhabditis elegans* has emerged as a genetically tractable animal host in which to study evolutionarily conserved mechanisms of innate immune signaling. We previously showed that the PMK-1 p38 mitogen-activated protein kinase (MAPK) pathway regulates innate immunity of *C. elegans* through phosphorylation of the CREB/ATF bZIP transcription factor, ATF-7. Here, we have undertaken a genomic analysis of the transcriptional response of *C. elegans* to infection by *Pseudomonas aeruginosa*, combining genome-wide expression analysis by RNA-seq with ATF-7 chromatin immunoprecipitation followed by sequencing (ChIP-Seq). We observe that PMK-1-ATF-7 activity regulates a majority of all genes induced by pathogen infection, and observe ATF-7 occupancy in regulatory regions of pathogen-induced genes in a PMK-1-dependent manner. Moreover, functional analysis of a subset of these ATF-7-regulated pathogen-induced target genes supports a direct role for this transcriptional response in host defense. The genome-wide regulation through PMK-1– ATF-7 signaling reveals global control over the innate immune response to infection through a single transcriptional regulator in a simple animal host.

Author Summary

Innate immunity is the first line of defense against invading microbes across metazoans. *Caenorhabditis elegans* lacks adaptive immunity and is therefore particularly dependent on mounting an innate immune response against pathogens. A major component of this response is the conserved PMK-1/p38 MAPK signaling cascade, the activation of which results in phosphorylation of the bZIP transcription factor ATF-7. Signaling via PMK-1 and ATF-7 causes broad transcriptional changes including the induction of many genes that are predicted to have antimicrobial activity including C-type lectins and lysozymes. In this study, we show that ATF-7 directly regulates the majority of innate immune response genes upon pathogen infection of *C. elegans*, and demonstrate that many ATF-7 targets function to promote pathogen resistance.

Introduction

Convergent genetic studies of host defense of *Drosophila melanogaster* and mammalian innate immune signaling revealed a commonality in signaling pathways of innate immunity that has helped motivate the study of pathogen resistance mechanisms in genetically tractable host organisms such as *Caenorhabditis elegans* [1]. The simple *C. elegans* host has enabled the genetic dissection of integrative stress physiology orchestrating host defense of *C. elegans*[2-5]. Genetic analysis of resistance of *C. elegans* to infection by pathogenic *Pseudomonas aeruginosa* has defined an essential role for a conserved p38 mitogen-activated protein kinase pathway that acts on a CREB/ATF family bZIP transcription factor, ATF-7, in immune responses [6,7]. A complementary approach to characterizing the host response has been organismal transcriptome-wide characterization of genes induced upon infection by a number of different bacterial pathogens [8-15]. Putative effector genes encoding lysozymes and C-type lectin domain (CTLD)-containing proteins have been identified that have also served as useful markers of immune induction. Here, we report the genome-level characterization of the *C. elegans* response to *P. aeruginosa* that is mediated by ATF-7 activity downstream of PMK-1 activation, combining RNA-seq analysis of pathogen-induced gene expression with ChIP-seq analysis of ATF-7 binding, which suggests global regulation of the immune response of *C. elegans* through a single MAPK- transcription factor pathway.

Results and Discussion

We performed RNA-seq on wild-type (N2), *pmk-1* mutant, or *atf-7* mutant animals exposed to *E. coli* OP50 or *P. aeruginosa* PA14 to identify genes that are

differentially regulated upon infection that also require PMK-1 or ATF-7 for induction (Figure 1A). We found that in wild-type animals, 890 genes were two-fold upregulated (adjusted p-value <0.05), and 803 genes were two-fold downregulated upon *P. aeruginosa* exposure, compared to animals exposed in parallel to *E. coli* (Figure 1B; Table S1). Many of these upregulated genes have been previously implicated in the *C. elegans* immune response and including genes encoding C-type lectin domain (CTLD)-containing genes and lysozymes, corroborating prior microarray-based gene expression studies (Figure 1C) [8-10,12-14]. In contrast, gene ontology analysis of genes that are decreased in expression upon *P. aeruginosa* exposure shows enrichment for genes associated with homeostasis with significant ontology terms consistent with growth, development and reproduction (Figure S1A). Of note, many of the genes upregulated in response to *P. aeruginosa* exposure exhibit relatively low expression when animals are propagated on *E. coli*, whereas genes that are decreased in expression upon *P. aeruginosa* exposure are expressed at a higher basal level during normal growth conditions on *E. coli* (Figure 1D, S1B). In parallel, we analyzed *P. aeruginosa*-mediated gene expression changes in *pmk-1* and *atf-7* mutants to identify the proportion of genes induced by *P. aeruginosa* exposure that required PMK-1 and/or ATF-7 for induction (Figure S2). We observed that 70% of genes significantly induced two-fold or greater by *P. aeruginosa* exposure were no longer fully induced upon loss of *pmk-1,* and that 53% of upregulated genes were no longer fully induced upon loss of *atf-7* (Figure 1E, Table S1). We also found that 41% of genes reduced two-fold or more by *P. aeruginosa* required PMK-1, and 50% required ATF-7 for reduction of expression (Figure S2B, Table S1). These data

implicate a high degree of involvement of PMK-1-ATF-7 signaling in the majority of changes in gene expression induced in response to infection by *P. aeruginosa*.

To evaluate the role of ATF-7 in the direct regulation of genes induced by *P. aeruginosa* infection, we performed chromatin immunoprecipitation followed by sequencing (ChIP-seq) of animals carrying a GFP-tag fused to the C-terminal end of the endogenous *atf-7* locus. Using a GFP polyclonal antibody for immunoprecipitation, we generated ChIP binding profiles for animals in either the wild-type background (*atf-7(qd328[atf-7::2xTY1::GFP])*) or the *pmk-1* mutant background (*pmk-1(km25);atf-7(qd328[atf-7::2xTY1::GFP])*) after a four hour exposure to either *E. coli* OP50 or *P. aeruginosa* PA14, for a total of four conditions, similar to the treatment described in Figure 1A. In all conditions analyzed, ATF-7 exhibited abundant association throughout the genome, with around 9,000 total peaks identified as enriched by MACS2, corresponding to 23% of genes and 25% of transcription start sites (TSSs) (Table S2).

Analysis of the ATF-7 binding profile across all genes associated with enriched TSSs, as well as the subset altered in expression by *P. aeruginosa* in wild-type animals according to our RNA-seq data, revealed that ATF-7 is preferentially located at the promoter regions of genes that are increased in expression by *P. aeruginosa,* and that this enrichment for ATF-7 is lessened by *pmk-1* loss (Figure 2B, Figure S3A). MEME analysis of the most enriched loci identified significant enrichment for the motif GACgTCA, which corresponds to the Jun D bZIP motif expected for ATF-7 (Figure 2A). This motif is present in as many as 80% of the most highly enriched regions of the genome and its abundance is positively correlated with enrichment levels.

113 To identify the most likely immediate downstream targets of ATF-7, we set a
114 peak threshold based on the fraction of peaks containing the bZIP motif after ranking
115 ATF-7 peaks by enrichment (Figure S3B). This resulted in ~1500-4000 highly enriched
116 locations per experiment. Overlap of the of the retained ATF-7 binding profile of *P.*
117 *aeruginoasa*-induced genes compared to the RNA-seq data from *P. aeruginosa* induction
118 was measured using a Gene Set Enrichment Analysis (GSEA), which showed that ChIP
119 peaks were enriched for association with transcripts that are positively changed upon
120 pathogen exposure in both *E. coli* and *P. aeruginosa* ChIP conditions (Figure 2C, Figure
121 S3C). This association remains in the *pmk-1* mutant, although at weaker significance
122 level (Figure 2D, Figure S3D). Moreover, we also evaluated ATF-7 binding at individual
123 genomic loci induced by *P. aeruginosa* infection that were dependent on ATF-7 for full
124 upregulation. Examinations of distinct genetic loci further support the conclusions drawn
125 from the metagene analyses described above (Figure 3). These observations suggest a
126 direct transcriptional regulatory role for ATF-7 in the induction of broad transcriptional
127 changes upon immune challenge involving activation of p38/PMK-1 MAPK signaling in
128 response to *P. aeruginosa* infection.
129 For functional validation of putative ATF-7-regulated immune response target
130 genes, we focused on transcripts that were upregulated at least two-fold by *P. aeruginosa*
131 exposure in an ATF-7-dependent manner and that were also bound by ATF-7 in any of
132 our four ChIP-seq conditions. Included among these putative ATF-7 targets were genes
133 encoding antimicrobial effector molecules, such as CTLD-containing proteins and
134 lysozymes (Table S3). We determined whether RNAi-mediated knockdown of these
135 genes resulted in enhanced susceptibility to killing by *P. aeruginosa* and observed that

RNAi of 13 of 43 genes conferred enhanced sensitivity to killing by *P. aeruginosa*, without affecting survival on non-pathogenic *E. coli* (Table S3, Figure S4).

Our data suggest that ATF-7 is a direct regulator of immune effector genes that is regulated by PMK-1 p38 MAPK. We previously proposed a model in which PMK-1 phosphorylates ATF-7 in response to pathogen infection, switching the activity of ATF-7 from that of a transcriptional repressor to that of an activator, allowing the induction of immune response genes [7]. Our data here are consistent with this model, showing a strong dependence of pathogen-induced gene induction on PMK-1 and ATF-7, and a high degree of occupancy of regulatory regions of pathogen-induced genes by ATF-7 under basal and pathogen-induced conditions, with ATF-7 occupancy of pathogen-induced genes being strongly dependent on PMK-1. Moreover, our data reveal that PMK-1-ATF-7 signaling regulates over half of all pathogen-induced genes at the genome-wide level.

PMK-1 signaling has also been implicated in a number of non-infection contexts in *C. elegans* [2,16,17]. Interestingly, we observed that ATF-7 binds quite strongly to several key regulators of stress response pathways. We found that ATF-7 exhibits binding affinity to regulators of autophagy (*lgg-1*), the Unfolded Protein Response (*xbp-1*), and the oxidative stress response (*skn-1*), as well as several immunity regulators (*hlh-30*, *zip-2*, and interestingly, *atf-7*) (Figure 4). These observations suggest that initiation of other stress responses may be integrated with the immune response. For example, we have previously shown that immune response activation in developing larval is lethal without compensatory XBP-1 activity, establishing an essential role for XBP-1 during activation of innate immunity during infection of *C. elegans* [2]. We speculate that ATF-7 may function to activate anticipatory stress responses that can be activated in concert

159 with innate immunity to promote host survival during microbial infection in a context-
160 dependent manner. Our genomic and genetic findings in the simple, genetically tractable
161 *C. elegans* host reveal a striking degree of global regulation of the organismal response to
162 pathogenic bacteria through a single p38 MAPK-regulated transcriptional regulator. Our
163 data support the idea that host defense, on a genome-wide and organism-wide level, is
164 under the control of a limited number of stress-activated signaling pathways that regulate
165 global regulators of gene transcription.

Materials and Methods

C. elegans Strains

Strains used: N2, ZD386 (*atf-7(qd22 qd130)*), KU25 (*pmk-1(km25)*), ZD1807 (*atf-7(qd328[atf-7::2xTY1::GFP])*), ZD1976 (*atf-7(qd328[atf-7::2xTY1::GFP]);pmk-1(km25)*). *C. elegans* were maintained at 16°C on *E. coli* OP50 as described [18]. The *atf-7(qd328)* allele was generated by the CRISPR-Cas9 system as described [19,20] and verified by Sanger sequencing. GFP expression in ZD1807 (*atf-7(qd328)*) was verified by immunobloting, and pull-down was assessed by IP-IB. The *atf-7(qd238)* allele was confirmed to function as wild-type, as assayed by susceptibility to *P. aeruginosa* strain PA14 in a slow kill assay, and then crossed into the *pmk-1(km25)* mutant background.

Preparation of Animals for Sequencing Experiments

Slow Kill Assay (SKA) plates were prepared as previously described [21]. *P. aeruginosa* strain PA14 or *E. coli* OP50 was grown overnight in Luria Broth (LB), seeded onto SKA media and then grown overnight at 37°C, followed by an additional day at room temperature as previously described [22]. Large populations of animals were synchronized by egg-preparation of gravid adult worms in bleach, followed by L1 arrest overnight in M9 buffer. L1 animals were dropped onto concentrated OP50 lawns seeded onto Nematode Growth Media (NGM) and raised to the L4 larval stage at 20°C (about 40 h). Upon reaching L4, worms were washed off growth plates with M9 and placed on SKA plates prepared as described above, seeded with either PA14 or OP50 and incubated at 25°C for four hours. At this time, worms were harvested by washing for downstream applications.

189

190 **Chromatin Immunoprecipitation Followed by Sequencing**

191 After three washes in M9 buffer, animal pellets were resuspended in an equal
192 volume of PBS + complete ULTRA protease inhibitor tablets (Roche), flash frozen in
193 liquid nitrogen, and stored at -80°C until chromatin immunoprecipitation (ChIP). ChIP
194 was preformed as described [23,24] using Ab290, a ChIP-grade polyclonal GFP antibody
195 (Abcam). Libraries were prepared using the SPRIworks Fragment Library System
196 (Beckman Coulter) and single-end sequenced on an Illumina HiSeq2000 sequencer.
197 Three biological replicates of at least 15,000 animals were prepared and sequenced for
198 each condition, with the exception of only two replicates for *atf-7(qd328)* on PA14, as
199 one of the samples failed to pass quality control.
200 ChIP-seq reads were aligned against the *C. elegans* WBPS9 assembly using bwa
201 v. 0.7.12-r1039 [25] and the resulting bam files were sorted and indexed using samtools
202 v. 1.3 [26]. Sorted bam files were pooled by strain and microbial treatment, and peaks
203 were called using MACS2 (v. 2.1.1.20160309), as follows: callpeak on specific strain
204 bam file ("-t" flag) against the N2_PA14 control sample bam file ("-c" flag) callpeak -c
205 N2_PA14_control.sorted.bam -g ce --keep-dup all --call-summits --extsize 150 -p 1e-3 --
206 nomodel -B. Peak locations were intersected with regions +/-0.5kb around annotated TSS
207 based on the WBPS9/WS258 annotation using bedtools intersect (v2.26.0) [27], and in
208 cases of multiple peaks associated with a given TSS, peaks with maximal enrichment
209 over N2 control were retained. For the purpose of motif identification, peaks were ranked
210 by fold-enrichment over N2 control in descending order and the top 400 peaks were
211 retained, regions +/- 200 bps around the summit were retrieved and sequences were

obtained with bedtools getfasta. MEME-ChIP v. 4.12.0 [28] was used to call motifs using the following parameters: meme-chip -oc . -time 300 -order 1 -db db/JASPAR/JASPAR2018_CORE_nematodes_non-redundant.meme -meme-mod anr -meme-minw 5 -meme-maxw 30 -meme-nmotifs 8 -dreme-e 0.05 -centrimo-local -centrimo-score 5.0 -centrimo-ethresh 10.0 . Presence of the top motifs under each peak called by macs2 was assessed using Mast v.5.0.1 [29] on the same +/- 200bp region around the summit of each peak. The number of peaks with one or more occurrences of the motif was tallied using a 200-peak window, and plotted across all peaks ranked either by log-fold enrichment over N2 or –log-transformed p-values. Inflection points in the motif density function were used to narrow down the number of peaks retained for downstream analyses.

RNA Sequencing

After three washes in M9 buffer, TRIzol™ Reagent (Invitrogen) was added to worm pellets and flash frozen in liquid nitrogen. Following an additional round of freeze-thaw, RNA was isolated using the Direct-zol™ RNA MiniPrep kit (Zymo Research). Libraries were prepared using the Kapa mRNA Hyperprep kit and paired end reads were sequenced on the Illumina NextSeq500 sequencer. Three biological replicates of at least 1,000 animals were prepared and sequenced for each condition, with the exception of only two replicates for *atf-7(qd22 qd130)* on PA14, as one of the samples failed to pass quality control.

Reads were aligned against the *C. elegans* WBPS9 assembly/ WS258 annotation using STAR v. 2.5.3a [30] with the following flags: -runThreadN 16 --runMode

alignReads --outFilterType BySJout --outFilterMultimapNmax 20 --alignSJoverhangMin 8 --alignSJDBoverhangMin 1 --outFilterMismatchNmax 999 --alignIntronMin 10 --alignIntronMax 1000000 --alignMatesGapMax 1000000 --outSAMtype BAM SortedByCoordinate --quantMode TranscriptomeSAM . with --genomeDir pointing to a low-memory footprint, 75nt-junction WBPS9/WS258 STAR suffix array. Gene expression was quantitated using RSEM v. 1.3.0 [31] with the following flags for all libraries: rsem-calculate-expression --calc-pme --alignments -p 8 against an annotation matching the STAR SA reference. Posterior mean estimates (pme) of counts and estimated "transcript-TPMs" were retrieved for genes and isoforms. Subsequently, counts of isoforms sharing a transcription start site (TSS) were summed, and differential-expression analysis was carried out using DESeq2 [32] in the R v3.4.0 statistical environment, building pairwise models of conditions to be compared (microbial exposures within each genotype). Sequencing library size factors were estimated for each library to account for differences in sequencing depth and complexity among libraries, as well as gene-specific count dispersion parameters (reflecting the relationship between the variance in a given gene's counts and that gene's mean expression across samples).

Differences in gene expression between conditions (expressed as log2-transformed fold-changes in expression levels) were estimated under a general linear model (GLM) framework fitted on the read counts. In this model, read counts of each gene in each sample were modeled under a negative binomial distribution, based on the fitted mean of the counts and aforementioned dispersion parameters. Differential expression significance was assessed using a Wald test on the fitted count data (all these

258 steps were performed using the DESeq() function in DESeq2) [32]. P-values were
259 adjusted for multiple-comparison testing using the Benjamini-Hochberg procedure [33].
260
261 **Data availability**
262 Raw data presented in this manuscript have been deposited in NCBI's Gene
263 Expression Omnibus [34] and are accessible through GEO SuperSeries accession number
264 GSE119294, which contains SubSeries GSE119292 (RNA-seq data, including count
265 files) and SubSeries GSE119293 (ChIP-seq data, including wig files and peak calls).
266
267 **Evaluation of ATF-7 binding and modulation of gene expression**
268 Metagene analyses of gene expression and ATF-7 binding enrichment were
269 generated by ngs.plot as described [35], using ChIP .bam files from each condition
270 normalized to N2 control as input. Genes considered two-fold upregulated or
271 downregulated are listed in the "N2_up" and "N2_down" tabs of Table S1, respectively.
272 Correlations between ATF-7 binding and regulation of gene expression were
273 interrogated using the gene set enrichment analysis (GSEA) framework [36]. Briefly, all
274 transcription start sites (TSSs) associated with a protein-coding transcript were ranked
275 based on differential expression results from DESeq2 (log2 fold-changes), which is a
276 measure of the correlation between their expression and the host response to infectious
277 agents. Biases in expression of ATF-7-bound TSSs were assessed using a walk down the
278 list tallying a running-sum statistic, which increases each time a TSS is part of the list and
279 decreases otherwise. The maximum of this metrics (i.e. where the distribution if furthest
280 away from the background) is called the enrichment score (ES). Significance is estimated

using random permutations of the TSSs to generate p-values gauging how often the observed ES can be seen in randomized gene sets, for each direction of the expression biases independently. Multiple-testing correction is addressed using a false-discovery rate calculation on permuted datasets.

Gene Ontology analysis

Genes with adjusted p-values <0.05 were considered for Gene Ontology enrichment analysis using the DAVID online webtool, considering as a background the union of all genes with a non-zero baseMean value across any of the DE comparison, based on unique WormBase IDs.

Killing Assays and Bacterial Strains

PA14 plates were prepared as described as above. N2 animals were grown on NGM, supplemented with 25 ug/mL carbenicillin and 2mM isopropyl b-D-1 thiogalactopyranoside (IPTG), that was seeded with either the *E. coli* HT115 expressing plasmids targeting the gene of interest or the empty L4440 vector backbone for two generations prior to each experiment. Animal populations were synchronized by egg lay. At the L4 larval stage, approximately 30 worms were transferred to prepared SKA plates and incubated at 25°C. Animals were scored for killing twice daily until the majority of animals had died. Within each experiment, three plates were prepared and scored per RNAi treatment. All clones were obtained from the Ahringer [37] or Vidal [38]RNAi libraries and were verified by sequencing. For a list of all RNAi clones used, see Table S4.

Author Contributions

M.F. and E.J.T. performed all experiments. V.B. and S.S.L. performed bioinformatics analysis of RNA-seq and ChIP-seq datasets. M.F. and D.H.K analyzed data, interpreted results, and wrote the paper with input from E.J.T.

Acknowledgements

We thank H.R. Horvitz and the *Caenorhabditis* Genetics Center for providing strains and reagents. This work was funded by NIH grants R01GM084477 (to DHK) and T32GM007287 (to MF and EJT).

Figure 1: Exposure to *P. aeruginosa* prompts gene expression changes.

(A) Schematic of experimental design. Yellow bacterial lawn indicates *E. coli* OP50, green bacterial lawn indicates *P. aeruginosa* PA14. PMK-1 and ATF-7 activation states are indicated below each condition. **(B)** Volcano plot of transcripts corresponding to differentially expressed protein-coding genes by exposure to *P. aeruginosa* PA14 versus *E. coli* OP50 in N2 animals. Orange and blue colored dots denote the top 100 outliers that are increased or decreased, respectively, and are annotated with their gene names. **(C)** Top GO terms and InterPRO classifications of transcripts that are significantly upregulated (adjusted p-value < 0.05) in N2 animals exposed to PA14 versus OP50. **(D)** Average expression (RPM) across the gene body of genes that are two-fold upregulated by exposure to pathogenic PA14. **(E)** Venn diagram of induced genes that are dependent upon *pmk-1* or *atf-7* for complete upregulation.

Figure 2: ATF-7 associates with genes that are differentially expressed upon exposure to pathogenic *P. aeruginosa.*

(A) Motif analysis of ATF-7::GFP ChIP peaks. The top 400 peaks (indicated by red shading) were considered for motif analysis. **(B)** Metagene analysis of ATF-7::GFP binding profile in WT or *pmk-1(km25)* mutant animals exposed to PA14 across genes that are two-fold upregulated (by RNA-seq) upon exposure to PA14. Shading represents standard error among replicates. **(C,D)** Gene Set Enrichment Analysis (GSEA) of transcripts detected by RNA-seq (ranked from most upregulated to most downregulated upon PA14 exposure in N2 animals) for association with ATF-7::GFP peaks in WT (C) or *pmk-1(km25)* mutant (D) animals exposed to PA14.

335 **Figure 3: ATF-7 binding at PA14-induced genes.**

336 Examples of ATF-7::GFP read pileup at individual loci in all four ChIP conditions.

337 Expression in transcripts per kilobase million (TPM) in all RNA-seq conditions are

338 displayed to the right of each locus. For genes with multiple isoforms, the most highly

339 abundant transcript is displayed.

340 **Figure 4: ATF-7 binds to regulatory regions of key regulators of stress physiology.**

341 Examples of ATF-7::GFP read pileups stress response (A-C) and immune (D-F)

342 regulators.

343

Supplemental Figure 1: Expression of genes decreased by PA14 exposure.

(A) Top GO terms and InterPRO classifications of transcripts that are significantly downregulated (adjusted p-value < 0.05) in N2 animals exposed to PA14 versus OP50. **(B)** Average expression (RPM) across the gene body of genes that are two-fold downregulated by exposure to pathogenic PA14. **(C)** Venn diagram of decreased genes that are dependent upon *pmk-1* or *atf-7* for complete downregulation.

Supplemental Figure 2: Differential PA14-induced gene expression in *pmk-1* and *atf-7* mutant animals.

2x2 comparison of genes differentially expressed by exposure to PA14 in N2 animals (x-axis) or upon loss of *pmk-1* (A) or *atf-7* (B) (y-axis). Transcripts highlighted in purple correspond to genes that were significantly different compared to OP50 in both genotypes (adjusted p-value of <0.05). Blue dots indicate genes that are significantly different in the N2 PA14/OP50 comparison only. Red dots represent genes that reach significance in only the mutant condition. Grey dots indicate genes with detected transcripts in at least one condition being compared, but that failed to reach significance cutoffs in either data set.

Supplemental Figure 3: Evaluation of ATF-7::GFP peaks.

(A) Motif analysis of ATF-7::GFP ChIP peaks. The top 400 peaks were considered for motif analysis. **(B)** Metagene analysis of ATF-7::GFP binding in all four ChIP conditions across all genes, genes that are two-fold upregulated, or two-fold downregulated by RNA-seq upon exposure to PA14 in a wild-type (N2) background. Shading represents standard error among replicates. **(C)** Ranked ATF-7::GFP peaks called in animals in all four ChIP conditions. Double red arrow indicates the peaks that were retained for further analysis. **(D,E)** Gene Set Enrichment Analysis (GSEA) of transcripts detected by RNA-seq (ranked from most upregulated to most downregulated upon PA14 exposure in N2 animals) for association with ATF-7::GFP peaks in WT (C) or *pmk-1(km25)* mutant (D) animals exposed to OP50.

Supplemental Figure 4: ATF-7 target genes that effect survival on *P. aeruginosa* PA14.

Representative survival curves of animals treated with RNAi against indicated genes that resulted in a significant (p-value < 0.05 by log-rank test) reduction in survival on PA14 compared to EV controls in 2/2 experiments. Animals were treated with RNAi for two generations prior to exposure to PA14. EV refers to HT115 carrying the Empty Vector control plasmid, L4440.

Table S1: RNA-seq summary

See separate electronic (.xlsx) file.

Table S2: ATF-7::GFP peaks from ChIP-seq.

See separate electronic (.xlsx) file.

Table S3: Genes tested for Esp phenotype by RNAi knockdown.

Protein domains classified using the David 6.8 Functional Annotation Tool. "Yes," indicates a significant (p-value < 0.05 by log-rank test) reduction in survival on PA14 compared to Empty Vector control in 2/2 experiments.

References

1. Schulenburg H, Kurz CL, Ewbank JJ. Evolution of the innate immune system: the worm perspective. Immunol Rev. 2004;198: 36–58.

2. Richardson CE, Kooistra T, Kim DH. An essential role for XBP-1 in host protection against immune activation in C. elegans. Nature. Nature Publishing Group; 2010;463: 1092–1095. doi:10.1038/nature08762

3. Dunbar TL, Yan Z, Balla KM, Smelkinson MG, Troemel ER. C. elegans detects pathogen-induced translational inhibition to activate immune signaling. Cell Host & Microbe. 2012;11: 375–386. doi:10.1016/j.chom.2012.02.008

4. McEwan DL, Kirienko NV, Ausubel FM. Host translational inhibition by Pseudomonas aeruginosa Exotoxin A Triggers an immune response in Caenorhabditis elegans. Cell Host & Microbe. 2012;11: 364–374. doi:10.1016/j.chom.2012.02.007

5. Pellegrino MW, Nargund AM, Kirienko NV, Gillis R, Fiorese CJ, Haynes CM. Mitochondrial UPR-regulated innate immunity provides resistance to pathogen infection. Nature. 2014;516: 414–417. doi:10.1038/nature13818

6. Kim DH, Feinbaum R, Alloing G, Emerson FE, Garsin DA, Inoue H, et al. A conserved p38 MAP kinase pathway in Caenorhabditis elegans innate immunity. Science. American Association for the Advancement of Science; 2002;297: 623–626. doi:10.1126/science.1073759

7. Shivers RP, Pagano DJ, Kooistra T, Richardson CE, Reddy KC, Whitney JK, et al. Phosphorylation of the conserved transcription factor ATF-7 by PMK-1 p38 MAPK regulates innate immunity in Caenorhabditis elegans. Ashrafi K, editor. PLoS Genet. 2010;6: e1000892. doi:10.1371/journal.pgen.1000892

8. Mallo GV, Kurz CL, Couillault C, Pujol N, Granjeaud S, Kohara Y, et al. Inducible Antibacterial Defense System in C. elegans. Current Biology. 2002;12: 1209–1214. doi:10.1016/S0960-9822(02)00928-4

9. Huffman DL, Abrami L, Sasik R, Corbeil J, van der Goot FG, Aroian RV. Mitogen-activated protein kinase pathways defend against bacterial pore-forming toxins. Proceedings of the National Academy of Sciences. 2004;101: 10995–11000. doi:10.1073/pnas.0404073101

10. O'Rourke D, Baban D, Demidova M, Mott R, Hodgkin J. Genomic clusters, putative pathogen recognition molecules, and antimicrobial genes are induced by infection of C. elegans with M. nematophilum. Genome Res. 2006;16: 1005–1016. doi:10.1101/gr.50823006

11. Shapira M, Hamlin BJ, Rong J, Chen K, Ronen M, Tan M-W. A conserved role for a GATA transcription factor in regulating epithelial innate immune responses.

Proceedings of the National Academy of Sciences. 2006;103: 14086–14091. doi:10.1073/pnas.0603424103

12. Troemel ER, Chu SW, Reinke V, Lee SS, Ausubel FM, Kim DH. p38 MAPK Regulates Expression of Immune Response Genes and Contributes to Longevity in C. elegans. PLoS Genet. 2006;2: e183. doi:10.1371/journal.pgen.0020183

13. Wong D, Bazopoulou D, Pujol N, Tavernarakis N, Ewbank JJ. Genome-wide investigation reveals pathogen-specific and shared signatures in the response of Caenorhabditis elegans to infection. Genome Biology. 2007;8: R194. doi:10.1186/gb-2007-8-9-r194

14. Engelmann I, Griffon A, Tichit L, Montañana-Sanchis F, Wang G, Reinke V, et al. A comprehensive analysis of gene expression changes provoked by bacterial and fungal infection in C. elegans. Lehner B, editor. PLoS ONE. 2011;6: e19055. doi:10.1371/journal.pone.0019055

15. Shivers RP, Youngman MJ, Kim DH. Transcriptional responses to pathogens in Caenorhabditis elegans. Curr Opin Microbiol. 2008;11: 251–256. doi:10.1016/j.mib.2008.05.014

16. Inoue H, Hisamoto N, An JH, Oliveira RP, Nishida E, Blackwell TK, et al. The C. elegans p38 MAPK pathway regulates nuclear localization of the transcription factor SKN-1 in oxidative stress response. Genes & Development. 2005;19: 2278–2283. doi:10.1101/gad.1324805

17. Chikka MR, Anbalagan C, Dvorak K, Dombeck K, Prahlad V. The Mitochondria-Regulated Immune Pathway Activated in the C. elegans Intestine Is Neuroprotective. Cell Reports. 2016;16: 2399–2414. doi:10.1016/j.celrep.2016.07.077

18. Brenner S. The genetics of Caenorhabditis elegans. Genetics. Genetics Society of America; 1974;77: 71–94. doi:10.1073/pnas.0910342106

19. Arribere JA, Bell RT, Fu BXH, Artiles KL, Hartman PS, Fire AZ. Efficient marker-free recovery of custom genetic modifications with CRISPR/Cas9 in Caenorhabditis elegans. Genetics. 2014;198: 837–846. doi:10.1534/genetics.114.169730

20. Tillman EJ, Richardson CE, Cattie DJ, Reddy KC, Lehrbach NJ, Droste R, et al. Endoplasmic Reticulum Homeostasis Is Modulated by the Forkhead Transcription Factor FKH-9 During Infection of Caenorhabditis elegans. Genetics. 2018;: genetics.301450.2018. doi:10.1534/genetics.118.301450

21. Killing of Caenorhabditis elegans by Pseudomonas aeruginosa used to model mammalian bacterial pathogenesis. Proceedings of the National Academy of Sciences. 1999;96: 715–720.

22. Meisel JD, Panda O, Mahanti P, Schroeder FC, Kim DH. Chemosensation of Bacterial Secondary Metabolites Modulates Neuroendocrine Signaling and Behavior of C. elegans. Cell. 2014;159: 267–280. doi:10.1016/j.cell.2014.09.011

23. Ercan S, Giresi PG, Whittle CM, Zhang X, Green RD, Lieb JD. X chromosome repression by localization of the C. elegans dosage compensation machinery to sites of transcription initiation. Nat Genet. 2007;39: 403–408. doi:10.1038/ng1983

24. Gerstein MB, Lu ZJ, Van Nostrand EL, Cheng C, Arshinoff BI, Liu T, et al. Integrative analysis of the Caenorhabditis elegans genome by the modENCODE project. Science. American Association for the Advancement of Science; 2010;330: 1775–1787. doi:10.1126/science.1196914

25. Li H. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009;25: 1754–1760. doi:10.1093/bioinformatics/btp324

26. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009;25: 2078–2079. doi:10.1093/bioinformatics/btp352

27. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics. 2010;26: 841–842. doi:10.1093/bioinformatics/btq033

28. Machanick P, Bailey TL. MEME-ChIP: motif analysis of large DNA datasets. Bioinformatics. 2011;27: 1696–1697. doi:10.1093/bioinformatics/btr189

29. Bailey TL, Gribskov M. Combining evidence using p-values: application to sequence homology searches. Bioinformatics. 1998;14: 48–54.

30. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013;29: 15–21. doi:10.1093/bioinformatics/bts635

31. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics. 2011;12: 323. doi:10.1186/1471-2105-12-323

32. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology. BioMed Central Ltd; 2014;15: 550. doi:10.1186/s13059-014-0550-8

33. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society, Series B. 1995;57: 289–300.

34. Edgar R, Domrachev M, Lash AE. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res. Oxford University Press; 2002;30: 207–210.

35. Shen L, Shao N, Liu X, Nestler E. ngs.plot: Quick mining and visualization of next-generation sequencing data by integrating genomic databases. BMC Genomics. 2014;15: 284. doi:10.1186/1471-2164-15-284

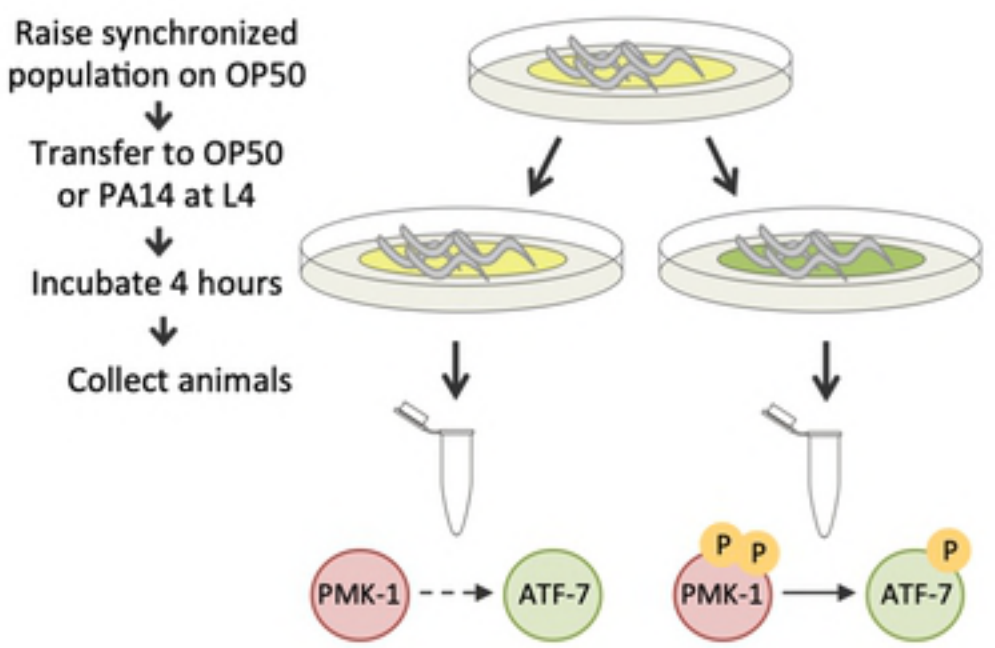
36. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci USA. 2005;102: 15545–15550. doi:10.1073/pnas.0506580102

37. Kamath RS, Fraser AG, Dong Y, Poulin G, Durbin R, Gotta M, et al. Systematic functional analysis of the Caenorhabditis elegans genome using RNAi. Nature. 2003;421: 231–237. doi:10.1038/nature01278

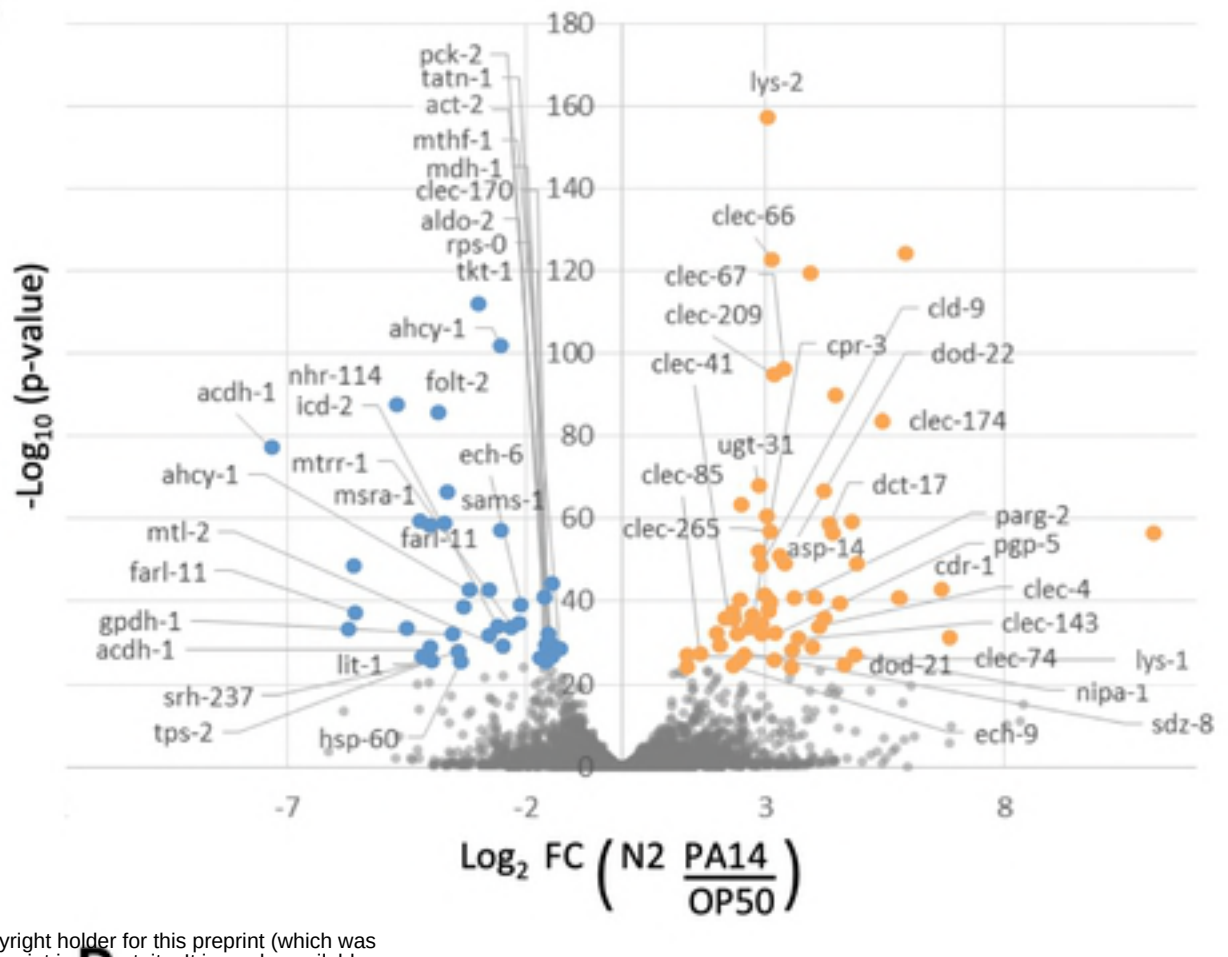
38. Rual J-F, Ceron J, Koreth J, Hao T, Nicot A-S, Hirozane-Kishikawa T, et al. Toward improving Caenorhabditis elegans phenome mapping with an ORFeome-based RNAi library. Genome Res. 2004;14: 2162–2168. doi:10.1101/gr.2505604

Figure 1

A



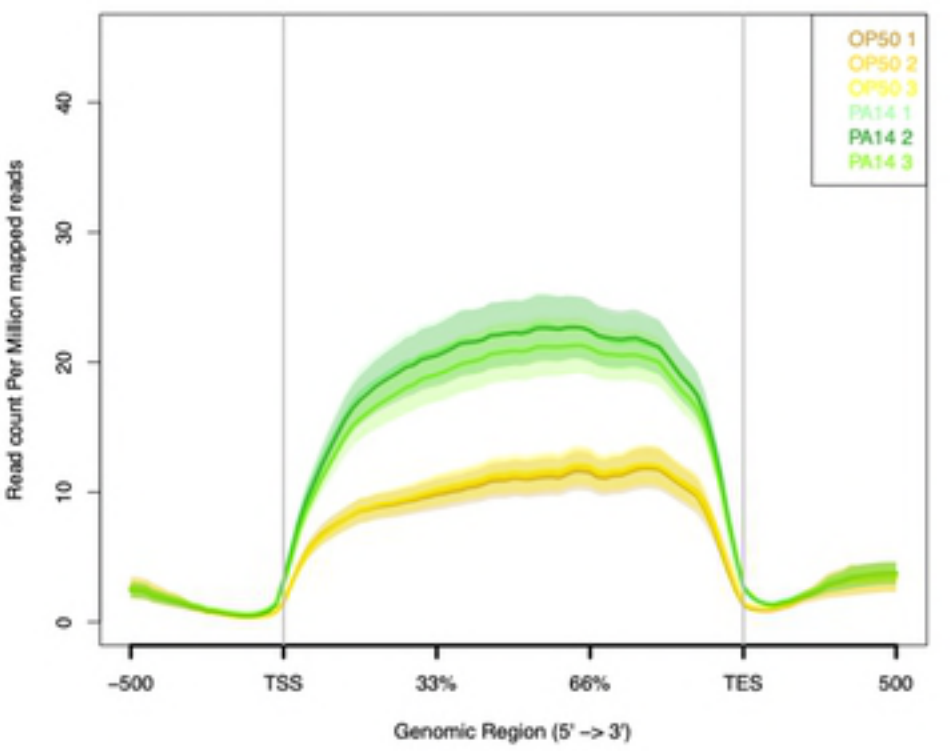
B



C

GO term, InterPRO classification	Fold Enrichment	p-value	FDR
GO:0045087~ innate immune response	7.24	3.57E-61	5.03E-58
IPR003366: CUB-like domain	13.97	1.91E-28	2.74E-25
GO:0050829~ defense response to Gram-negative bacterium	7.55	2.00E-15	2.81E-12
IPR005071: Transmembrane glycoprotein	11.27	4.35E-10	6.24E-07
IPR001304: C-type lectin	3.60	2.05E-07	2.94E-04
IPR002035: von Willebrand factor, type A	6.84	4.03E-07	5.77E-04

D



E

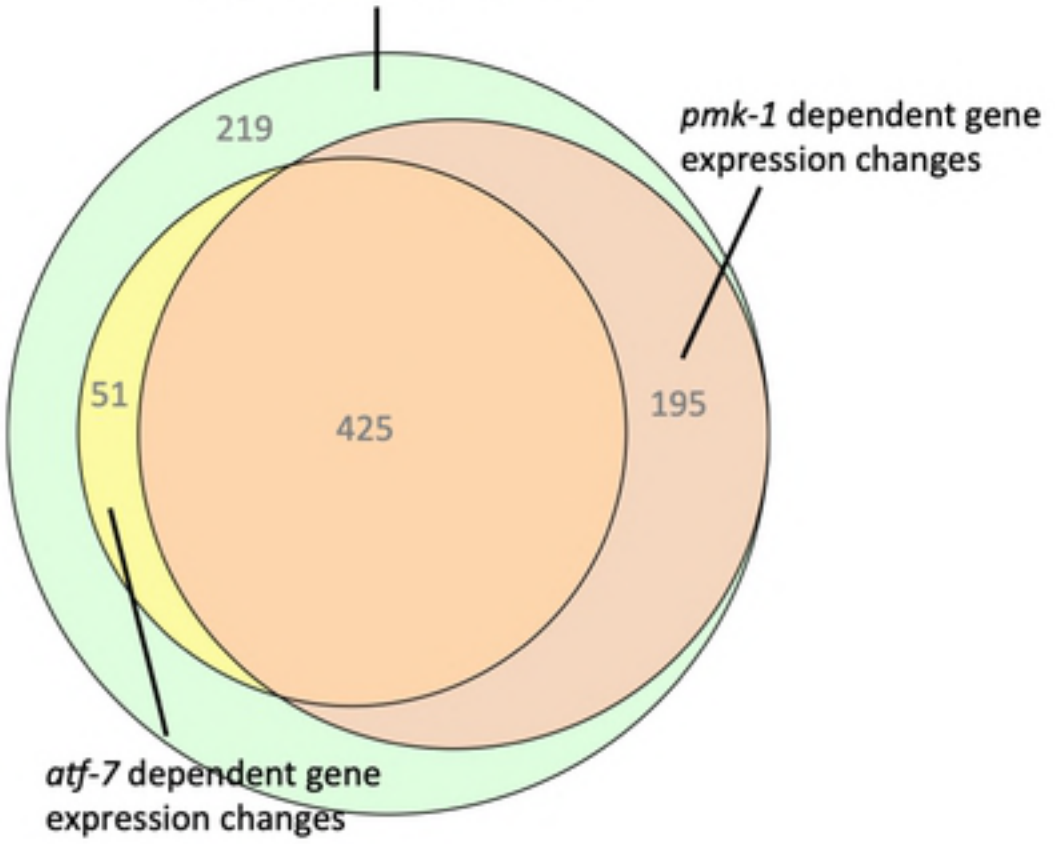
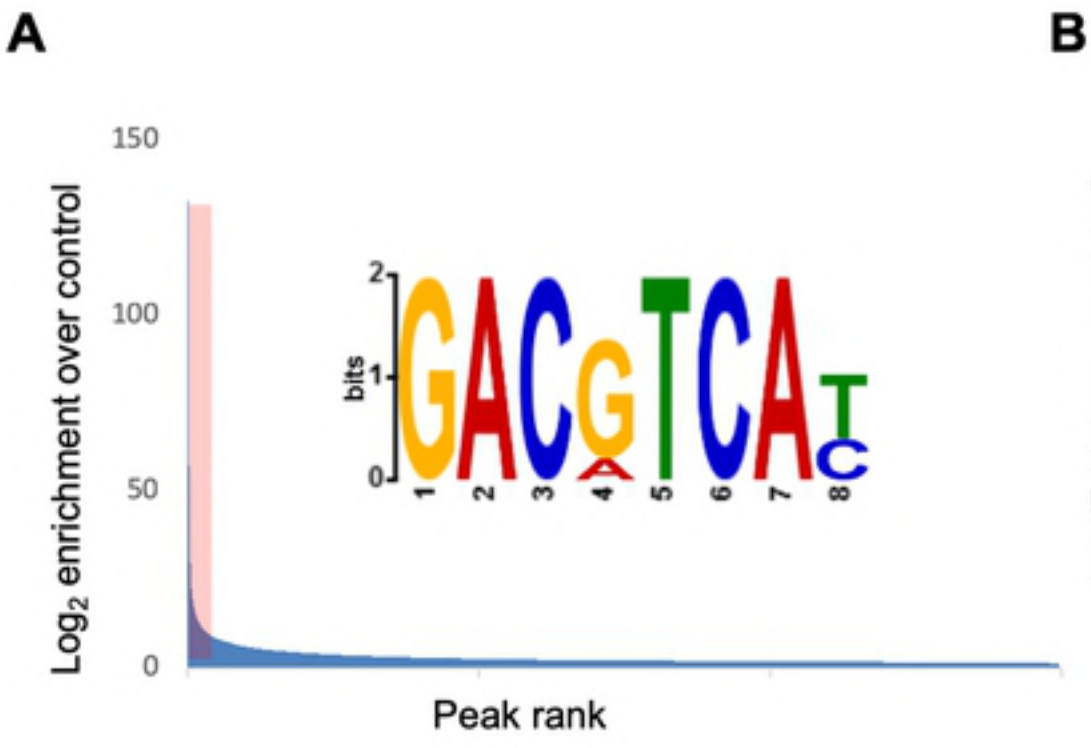
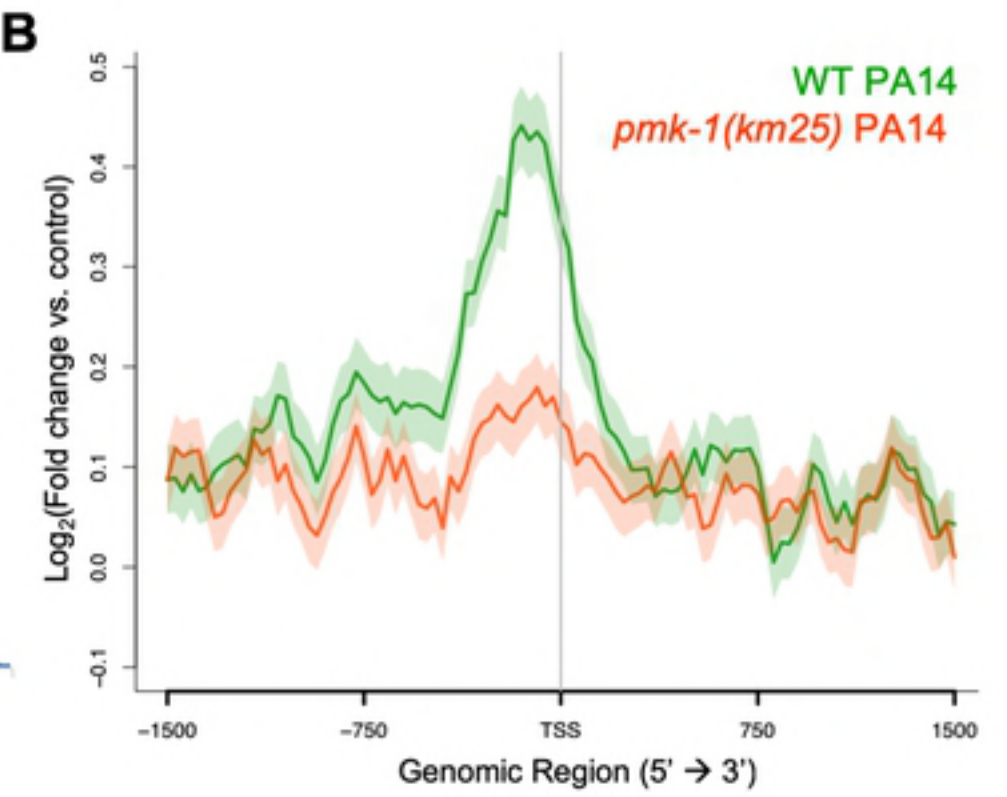
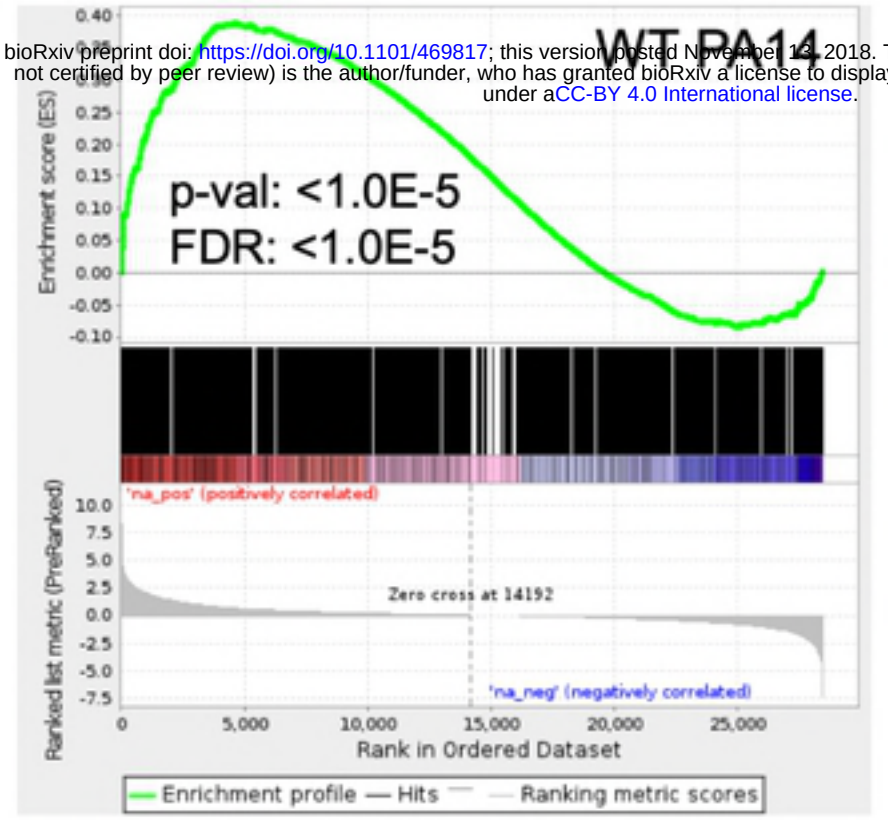


Figure 1

Figure 2







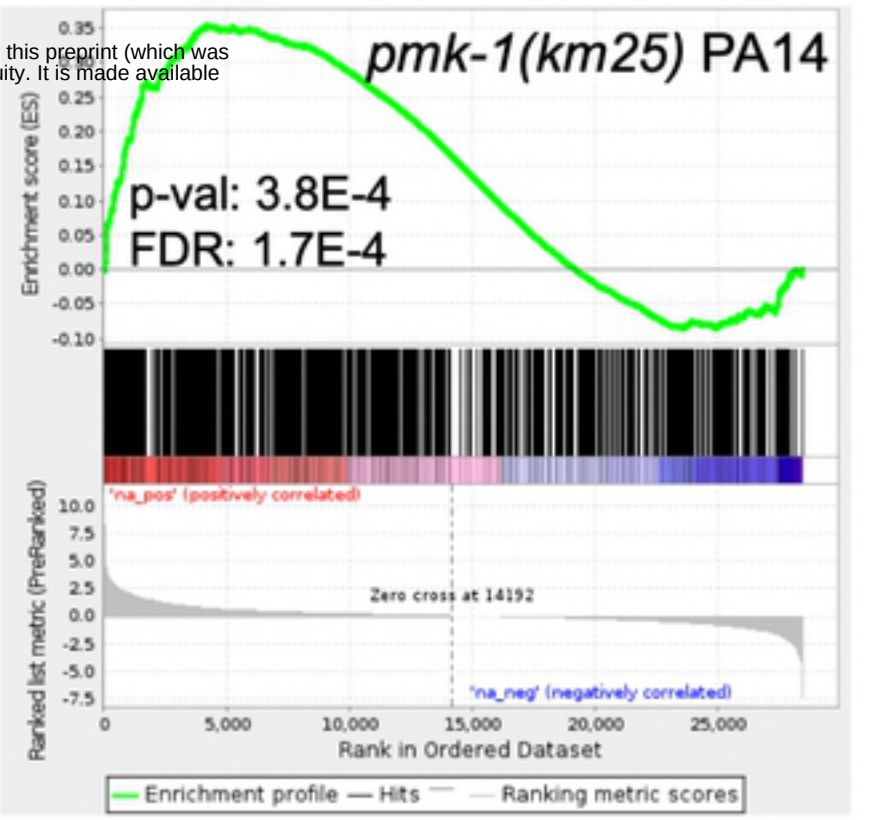
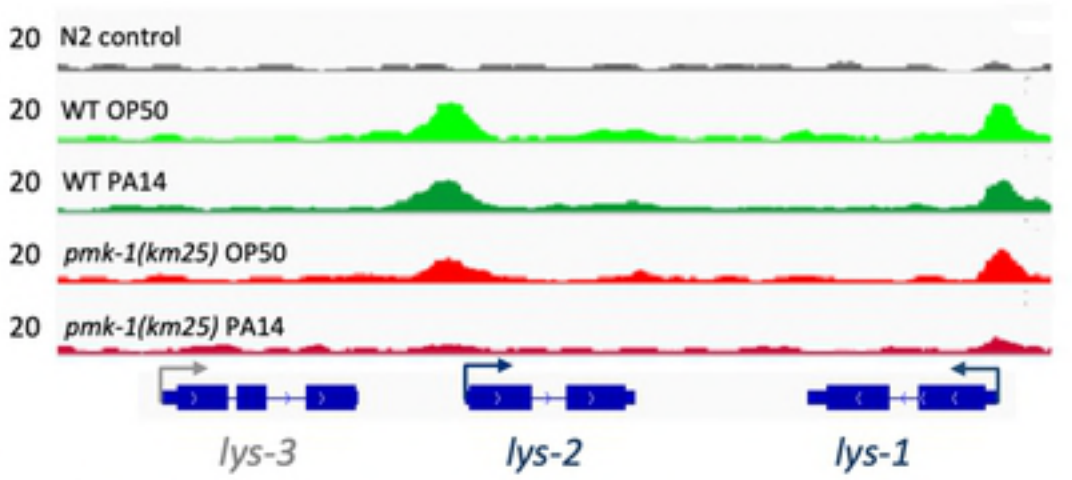


Figure 2

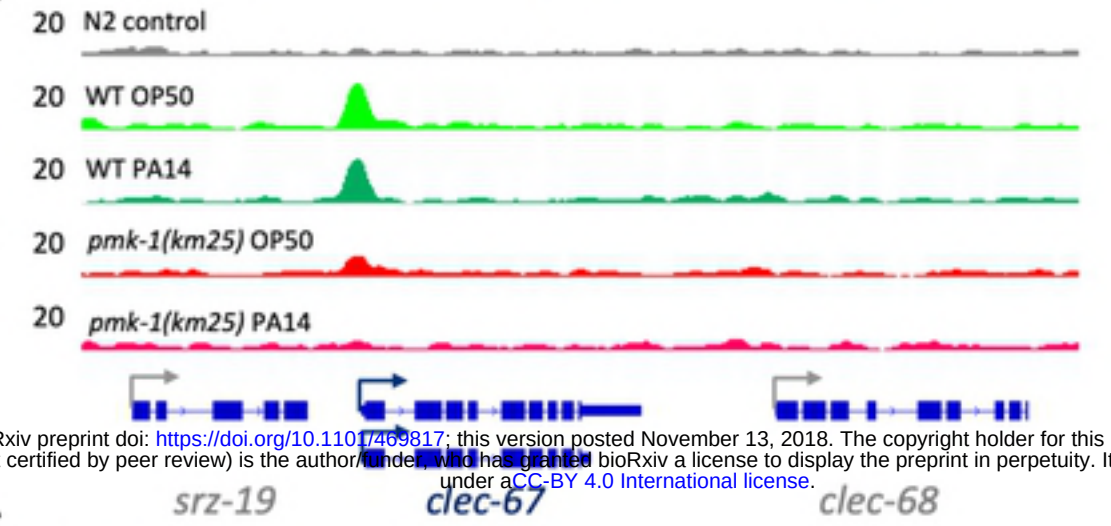
Figure 3

A



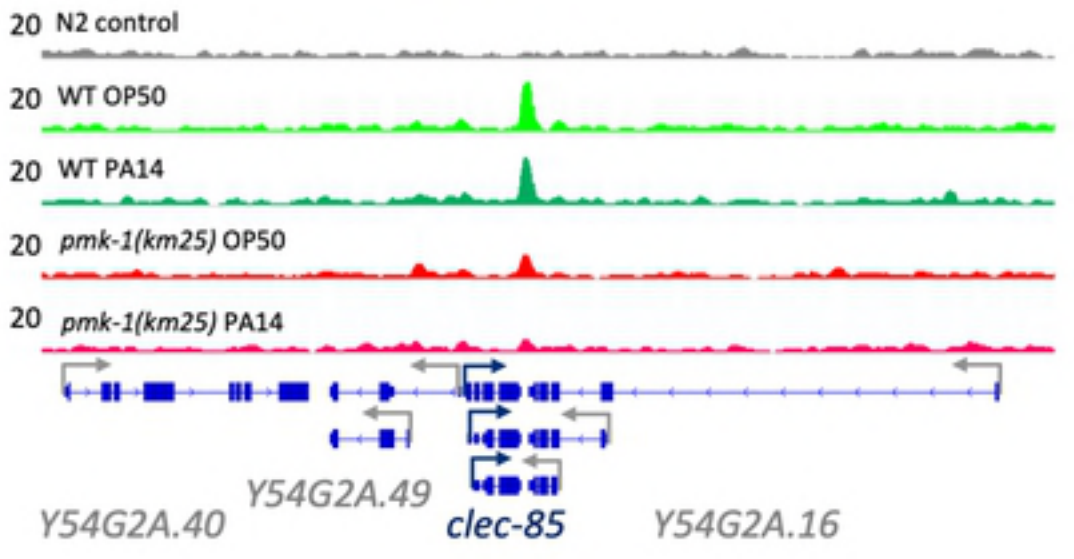
Average TPMs: (OP50 / PA14)	lys-3	lys-2	lys-1
N2	2 / 6	210 / 2032	322 / 1032
pmk-1(km25)	1 / 29	12 / 62	135 / 222
atf-7(qd22 qd130)	3 / 16	40 / 242	254 / 440

B



Average TPMs: (OP50 / PA14)	srz-19	clec-67	clec-68
N2	0 / 0	15 / 189	0 / 3
pmk-1(km25)	0 / 0	0 / 0	0 / 0
atf-7(qd22 qd130)	0 / 0	1 / 1	0 / 0

C



Average TPMs: (OP50 / PA14)	Y54G2A.49	clec-85	Y54G2A.16
N2	18 / 37	49 / 177	1 / 1
pmk-1(km25)	16 / 15	4 / 5	1 / 1
atf-7(qd22 qd130)	13 / 20	24 / 49	1 / 1

Figure 3

Figure 4

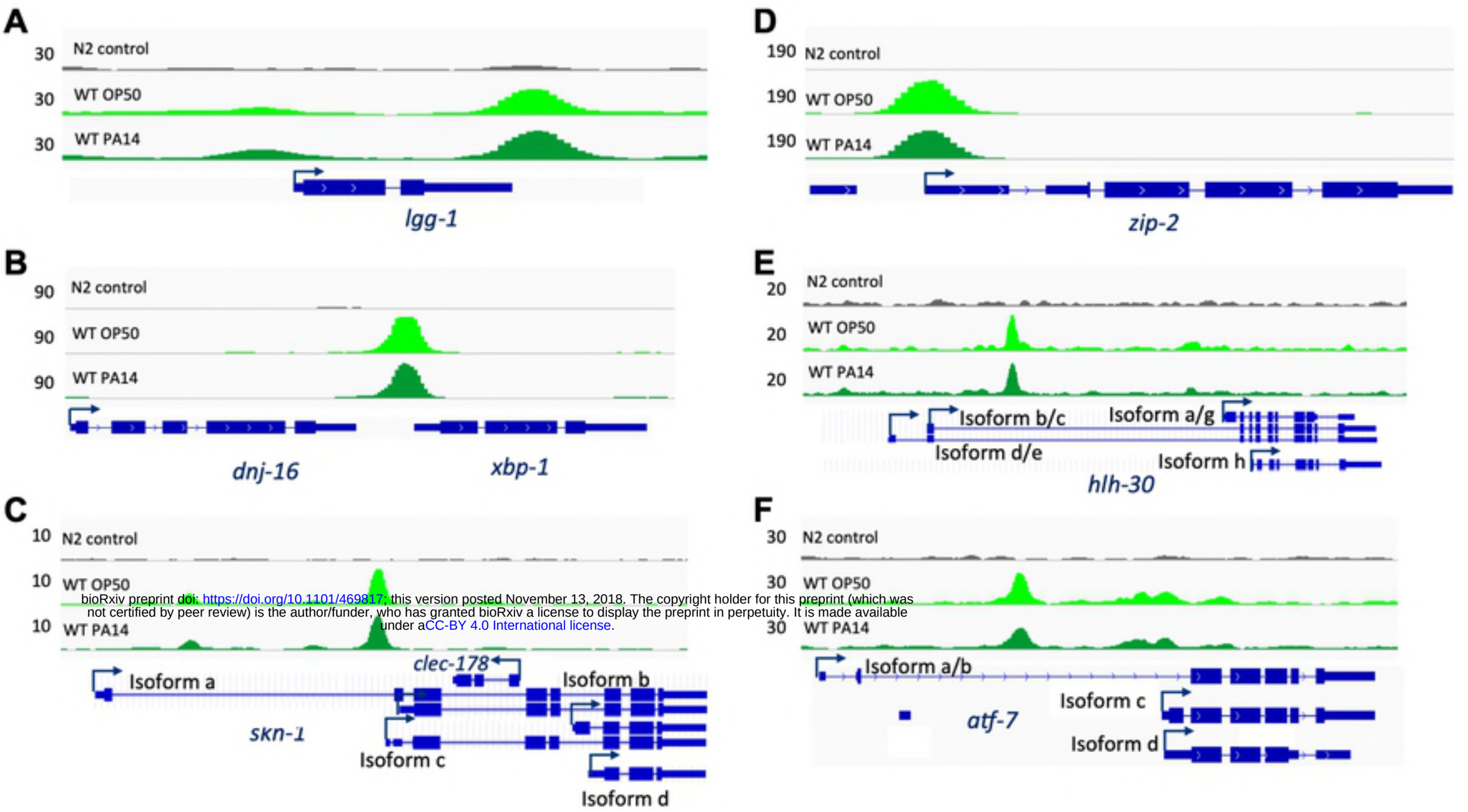


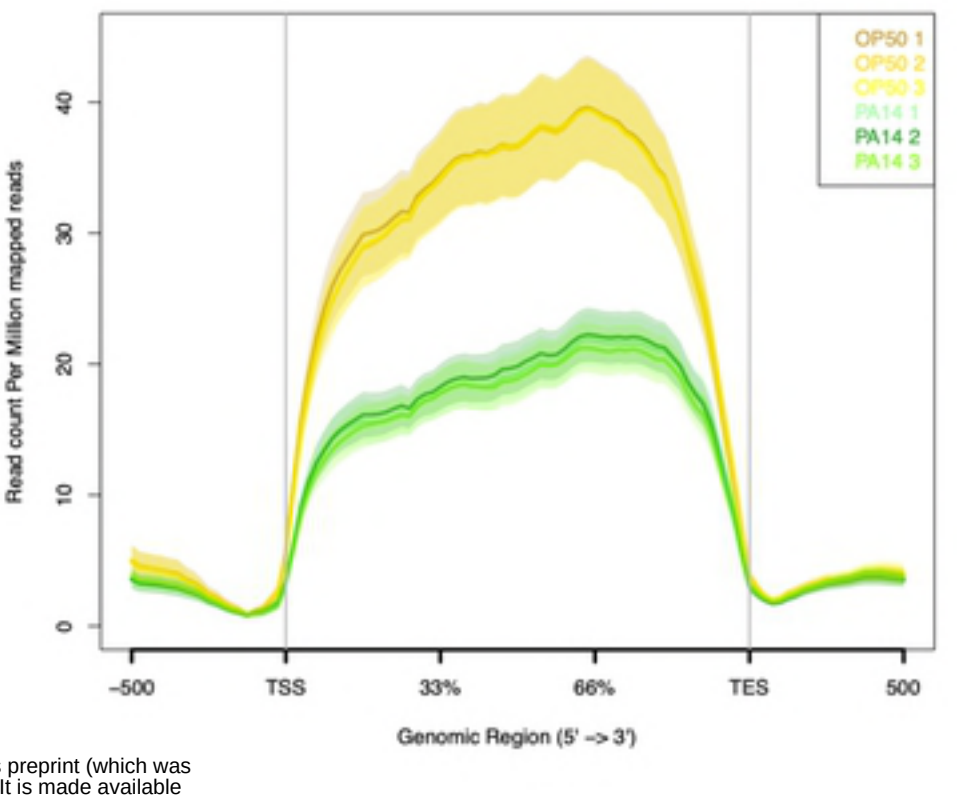
Figure 4

Supplemental Figure 1

A

GO term, InterPRO classification	Fold Enrichment	p-value	FDR
GO:0000003~ reproduction	1.90	4.14E-17	6.13E-14
GO:0006412~ translation	4.34	2.11E-16	3.33E-13
GO:0002119~ nematode larval development	1.89	5.90E-15	8.72E-12
GO:0009792~ embryo development ending in birth or egg hatching	1.59	5.47E-12	8.11E-09
GO:0008340~ determination of adult lifespan	2.04	1.35E-09	2.00E-06
GO:0055114~ oxidation-reduction process	2.51	1.40E-08	2.08E-05
GO:0006915~ apoptotic process	2.49	8.73E-08	1.29E-04
GO:0006096~ glycolytic process	8.92	8.08E-07	1.19E-03

B



C

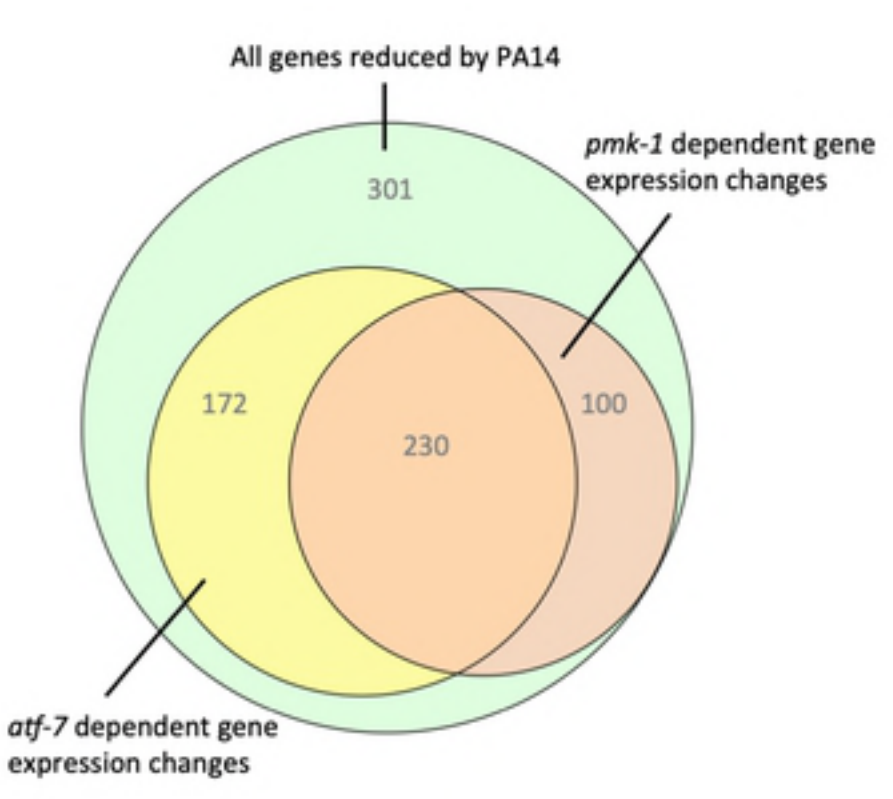


Figure S1

Supplemental Figure 2

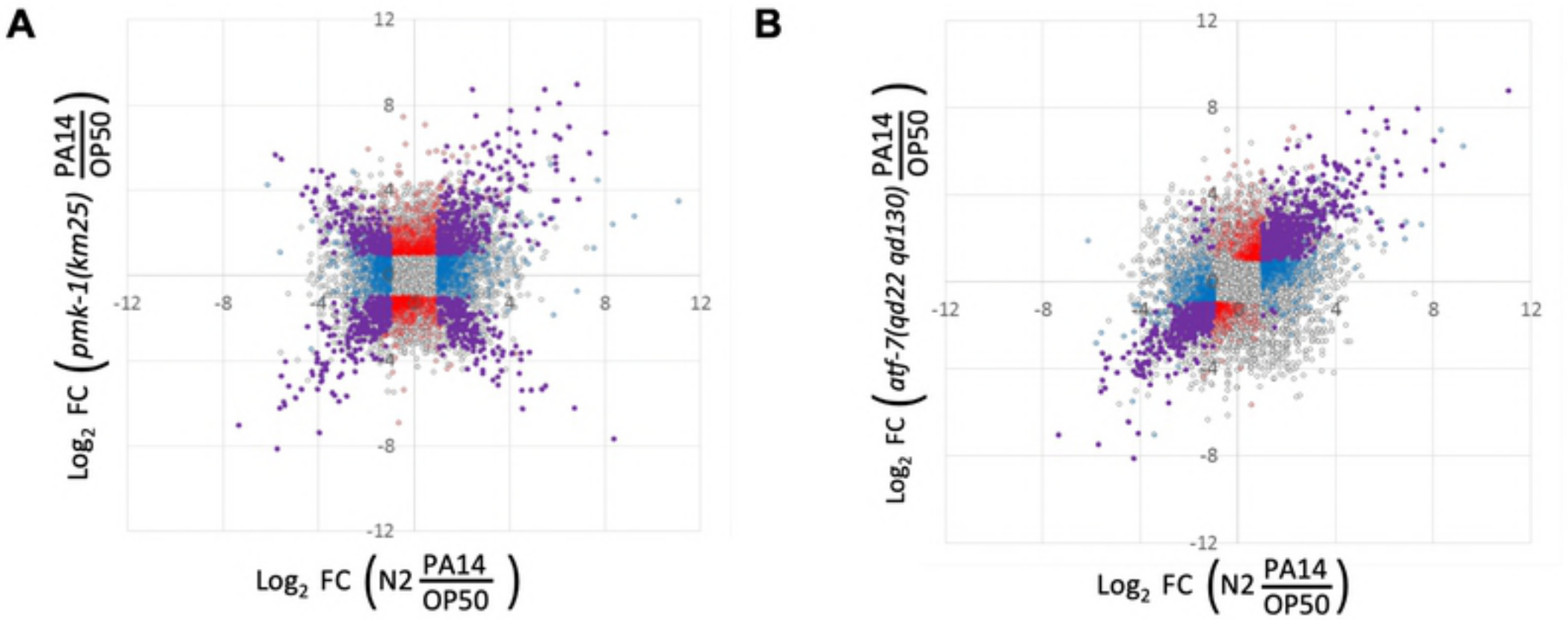


Figure S2

Supplemental Figure 3

A

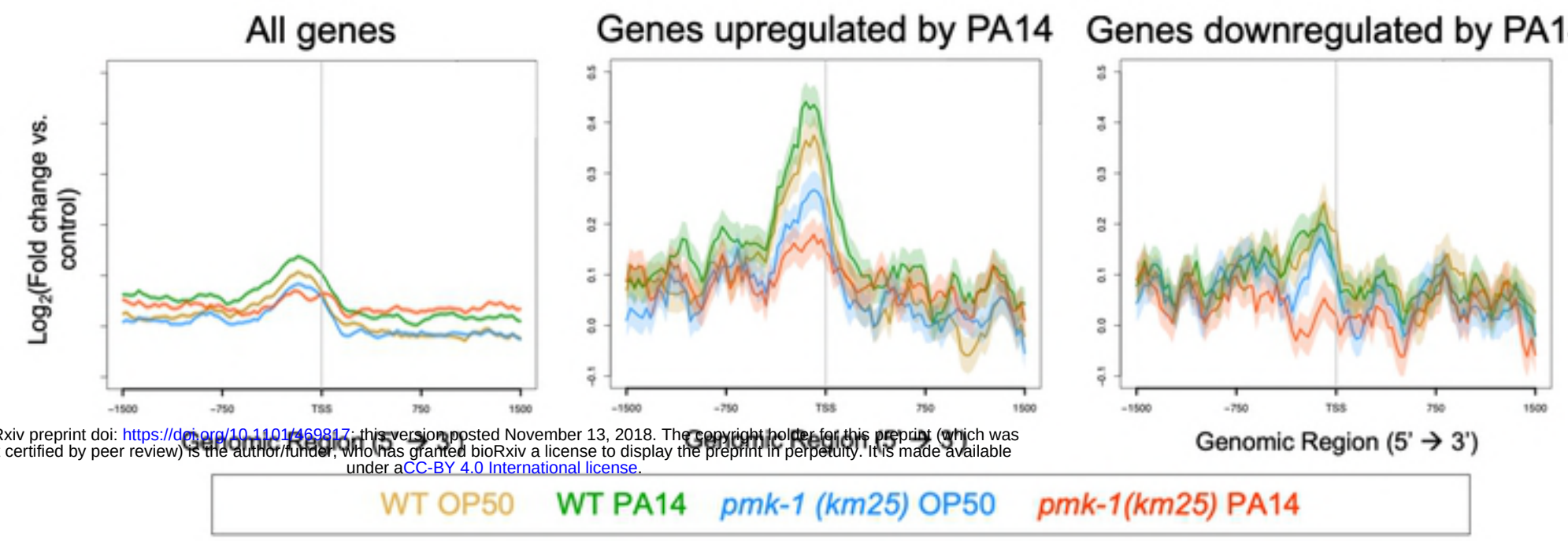
Details

Positives	Negatives	P-value	E-value	Unerased E-value
322 / 399	3 / 399	1.1e-142	2.8e-138	2.8e-138

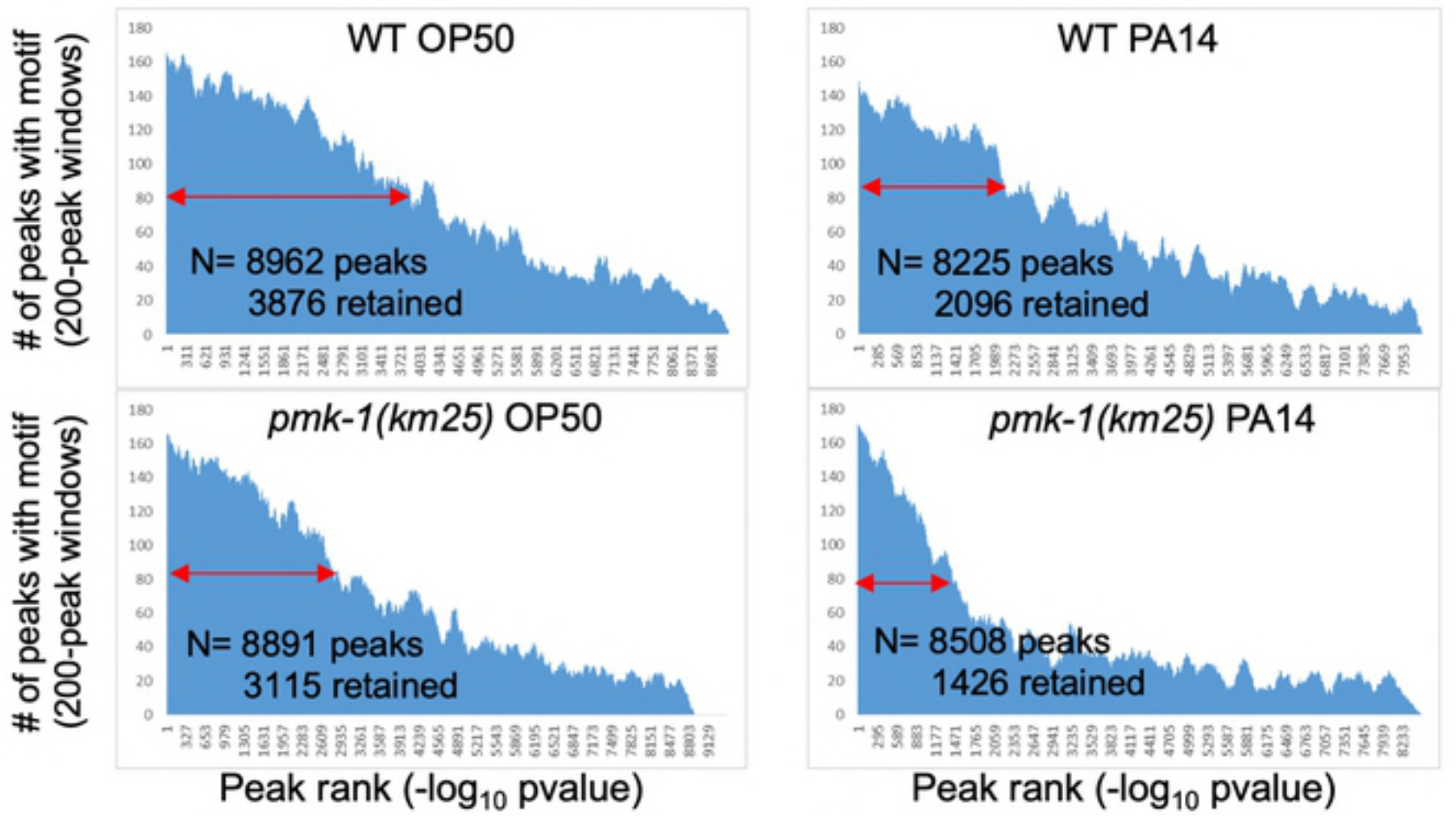
Enriched Matching Words

Word	Positives	Negatives	P-value	E-value
GACGTCAT	233 / 399	0 / 399	3.0e-092	7.7e-088
GACGTCAC	177 / 399	1 / 399	5.5e-063	1.4e-058
GACATCAT	39 / 399	1 / 399	1.5e-011	3.8e-007
GACATCAC	15 / 399	1 / 399	2.3e-004	5.9e+000

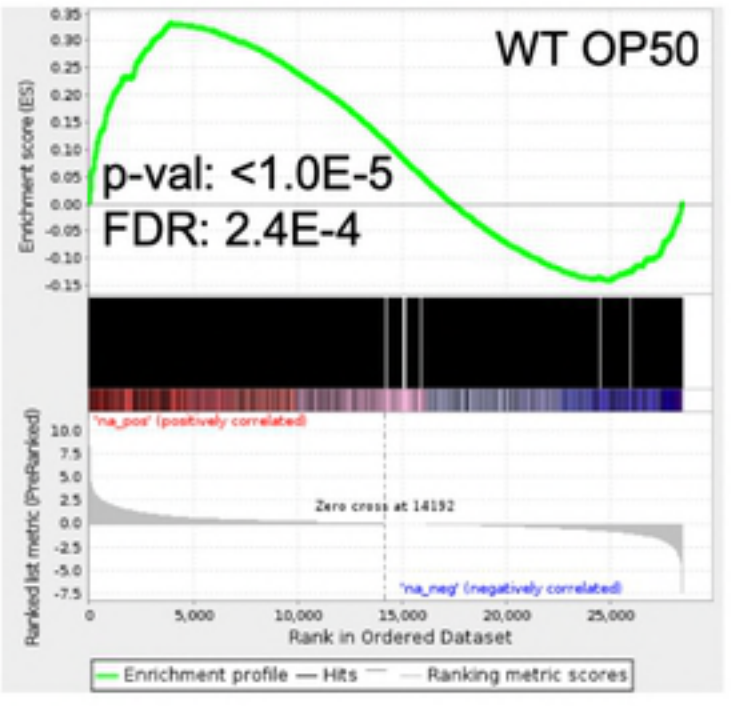
B



C



D



E

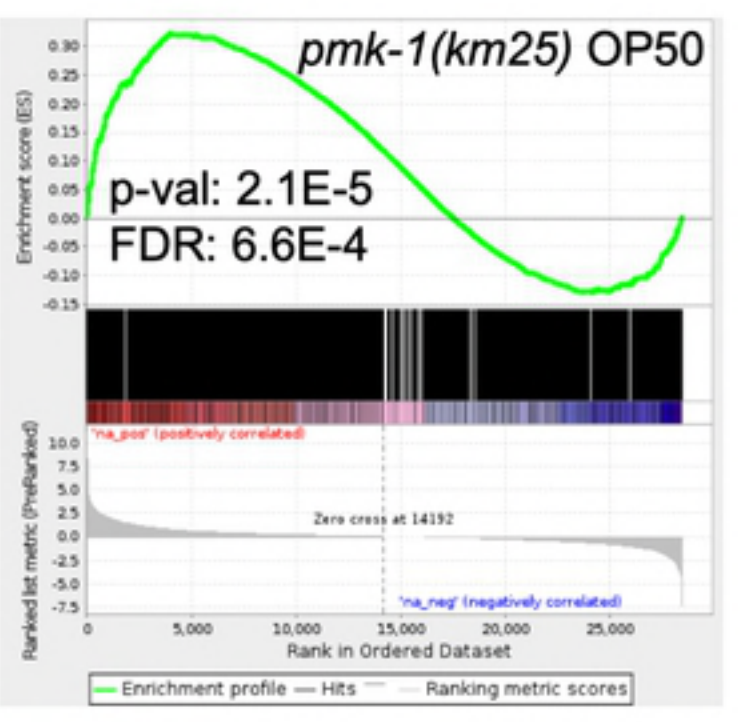
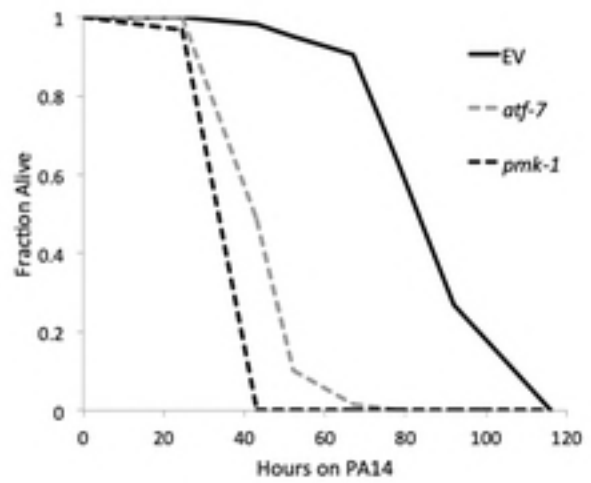
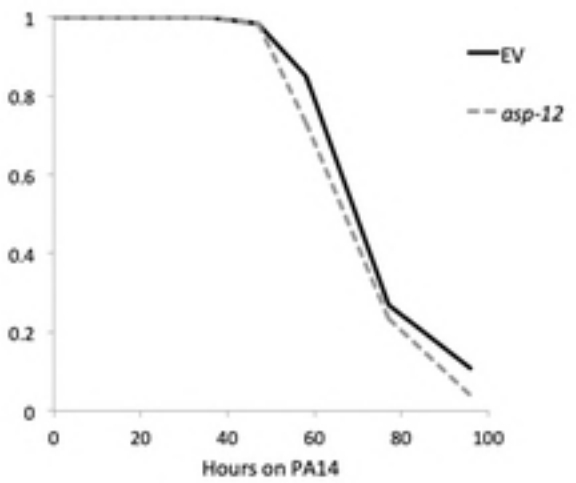
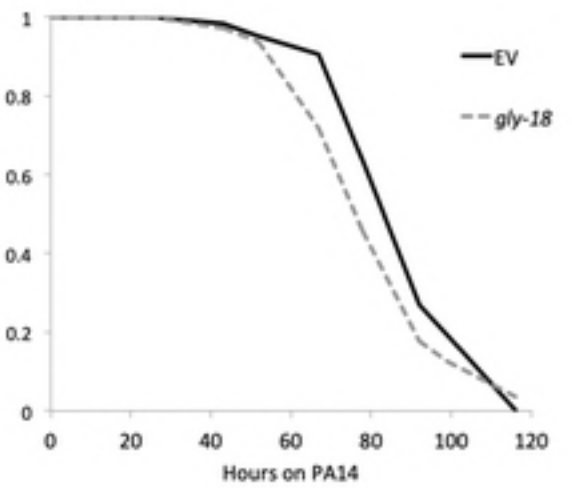


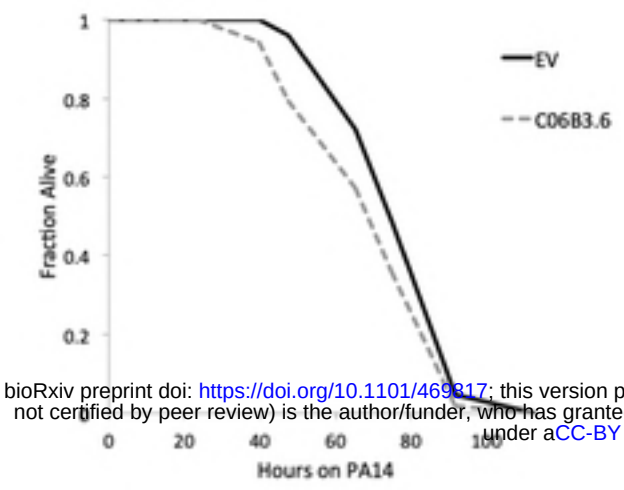
Figure S3

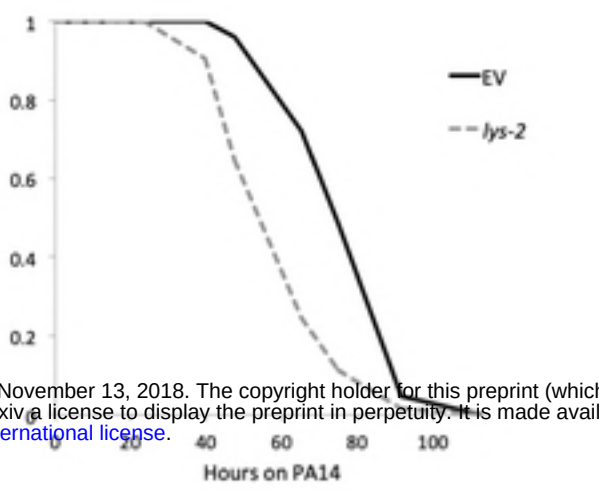
Supplemental Figure 4

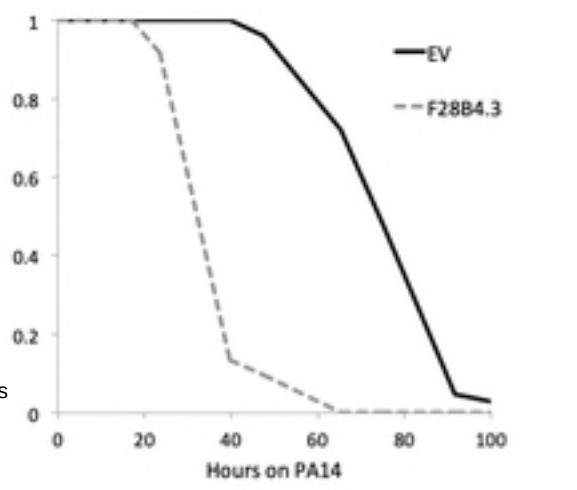


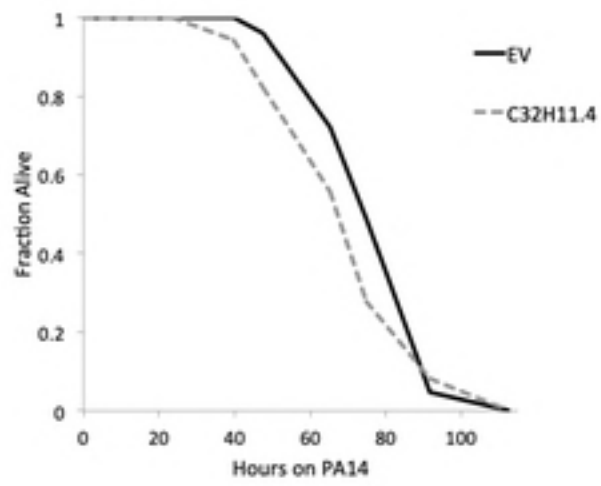


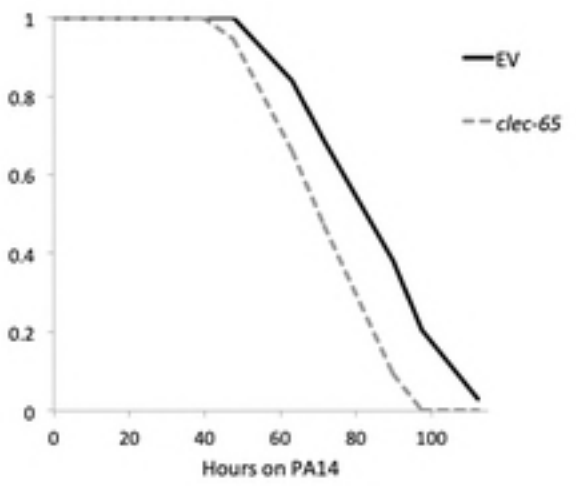


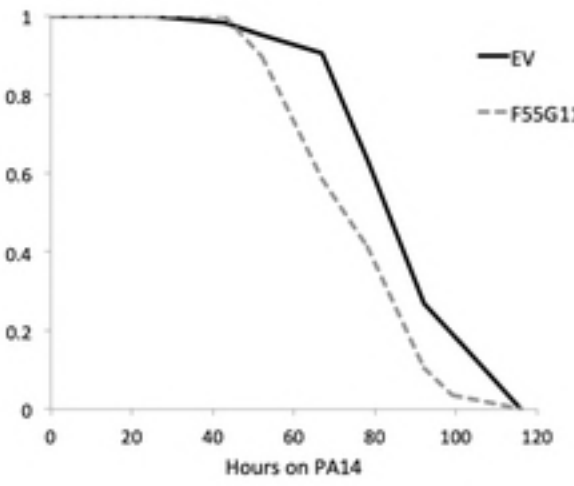


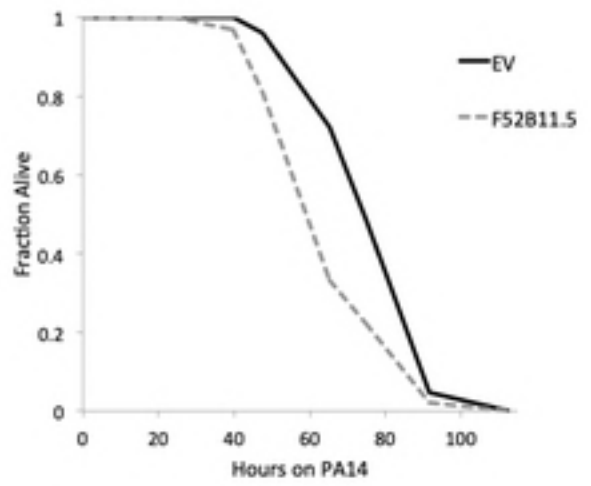


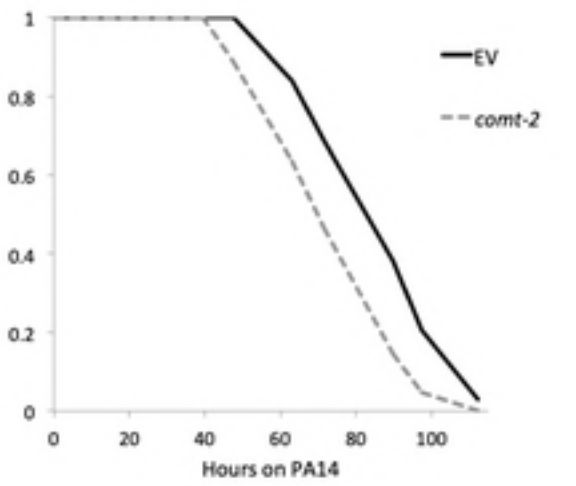


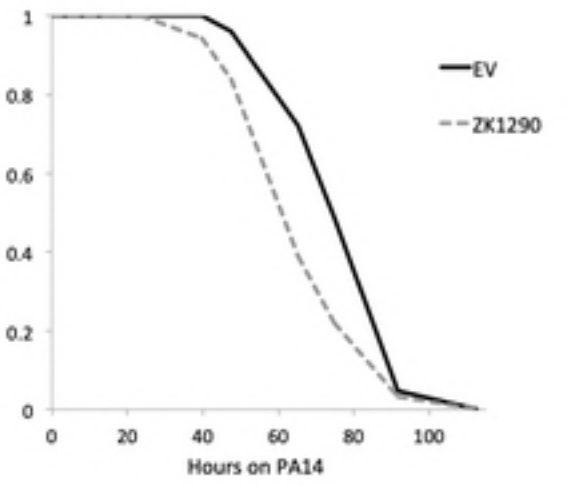


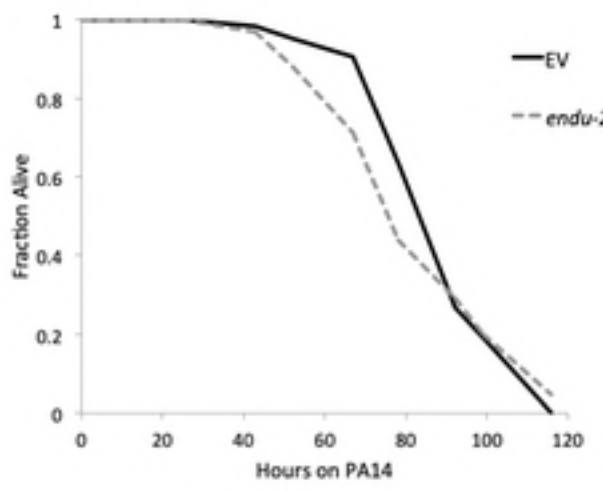












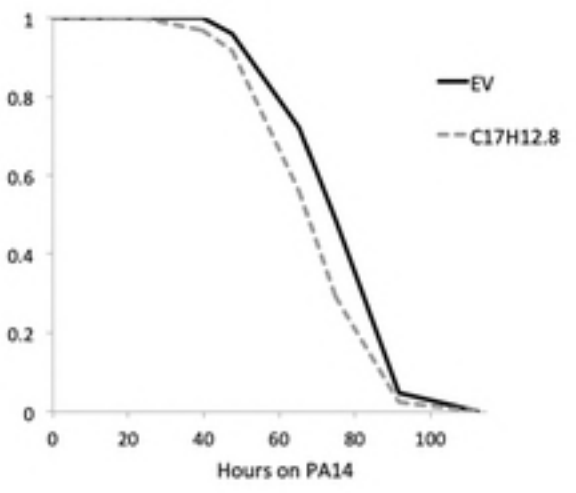


Figure S4