Supporting Information

Chun *et al*. 10.1073/pnas.0802369105

SI Text

Squid Colonization Experiments. The breeding stock of adult *Euprymna scolopes* was maintained and bred under laboratory conditions as previously described (1). Because the juveniles used in this study were the progeny of a number of field-caught animals, and were hatched over the course of two weeks, we sought to minimize the effects of natural genetic variation. To achieve this aim, the animals that hatched on a given day were equally divided among the six treatments. As a result of this strategy, the pooled RNA from a given treatment was derived from six to eight separate clutches, laid by three to four different field-caught females.

Four biological replicates (i.e., separate light-organ RNA extractions of independently collected light organs, each containing 90 organs for each of the six colonization treatments) were performed on the same day with the same reagents. Thus, a total of 540 organs were sampled for each of the four biological replicates.

The newly hatched juvenile animals were colonized by exposure to either wild-type, *luxA* or *luxI* strains of *Vibrio fischeri*. The *luxA* mutant has a deletion in a gene encoding bacterial luciferase, and is completely defective in bioluminescence, while the *luxI* mutant is unable to synthesize 3-oxo-hexanoyl-L-homoserine lactone (AI), an autoinducer of luciferase. The *luxI* mutant makes only a reduced level of luminescence (2). Because AI spontaneously inactivates in the alkaline conditions of seawater (3), we determined the rate of loss under our experimental conditions (Fig. S2). The resulting data indicated that by supplementing the seawater with 6 μM AI, the concentration of this luminescence inducer remained above 100 nM, the approximate level found in light organs colonized by wild-type cells (4), throughout the incubation (i.e., from 6 to18 h post inoculation). The efficacy of this AI-addition protocol was additionally confirmed by the restoration of the luminescence of animals colonized by the *luxI* mutant to a level close to that characteristic of colonization by wild-type *V. fischeri* (Fig. S3).

Microarray Hybridizations. The spotted microarray contained 13,962 cDNAs (5) applied two times on each glass slide, for a total of 27,924 sample spots (GEO accession: Platform GLP3825; I. V. Koroleva, B. J. Brown, E. Snir, H. Almabrazi, T. L. Casavant, M. B. Soares, and M. McFall-Ngai; Squid EST 30K cDNA array; public on June 2, 2006). The microarray slides had five positive-control spikes prepared using sequences isolated from *Xenopus leavis*, *Anopheles gambiae, Schistosoma mansoni*, and *Apis mellifera* (6–8). The chosen sequences did not cross-hybridize with any of the cDNAs from the *E. scolopes* light organ EST database (data not shown). Negative controls were spots of (*i*) buffer alone, (*ii*) polyA oligonucleotides, or (*iii*) no template.

Pools of light organs, obtained from 90 juvenile squid subjected to the same treatment, were homogenized (Polytron 1200C, Brinkmann Instruments Inc.), and total RNA was extracted using the MasterPure Purification Kit (Epicentre Biotechnologies), followed by the RNeasy Mini Kit with on-column DNase digestion (Qiagen Inc.). The resulting RNA sample was concentrated with a Microcon 30 (Millipore). The concentration and quality of the samples were determined spectrophotometrically, and their purity was estimated by agarose gel electrophoresis. Each biological replicate of 90 light organs contained on average 20 μg of total RNA, and only samples with a 260/280 nm optical ratio of between 1.9 and 2.0 were used. Visual inspection revealed two distinct ribosomal bands and no evidence of degradation products (data not shown). Total RNA samples were processed and indirectly labeled with 3DNA 350HS Expression Array Detection Kit (Genisphere). Slides were hybridized overnight, and washed using a Lucidia SlidePro hybridization station (GE Healthcare). The arrays were then scanned using GenePix 4000B (Molecular Devices). The raw data were stored in the database, and used for subsequent analysis.

On each slide we used a runoff reference, which is a mixture of vector-primed transcription products derived from an equal mixture of the 13,962 *E. scolopes* cDNA clones, labeled with the same protocol as the experimental samples (see below). This approach allowed us to compare each experimental condition to a single standardized reference, and to use a robust normalization method, the two-way semilinear model (TW-SLM) (9), that has been specifically optimized to identify genes in our experimental design. Each spotted glass-slide microarray was hybridized with two samples: (*i*) the experimental cDNA; and, (*ii*) the run-off reference, using a procedure described previously (10) except that sheared *E. scolopes* genomic DNA was used to block repetitive elements. This genomic DNA was isolated from a freshly dissected adult squid using the Blood and Cell Culture DNA Maxi Kit (Qiagen Inc.). As a control for technical variation inherent in the fluorescent-probe chemistry, the Cy3 and Cy5 labeling of the experimental sample and the run-off reference were alternated in at least two slides per replicate (i.e., three replicates per condition) in each of the six conditions. After normalization of the arrays' hybridization signals, Volcano plot analyses and ANOVA (see below) were performed on data from the six treatment conditions. These analyses identified 781 RNA transcripts/genes that were significantly differentially regulated between the conditions (Table S4).

Microarray Significance Analysis. To determine the confidence level of the signals from the hybridized glass-slide microarrays, we first accounted for the dye-swap technical replication, followed by a further independent normalization using two different techniques: per spot per chip (PSPC) and two-way semilinear model (TW-SLM) (9). All statistics described below were performed using GeneSpring GX software (Agilent Technologies). To eliminate systematic error due to either inconsistency of replicates or intensity levels below a reliable range, both raw and control data were filtered with 50% confidence according to the cross-gene error model for normalization conditions, both with and without the consideration of spike controls (11). The resulting 7,503 transcripts were analyzed using two statistical methods: (*i*) Volcano plot analysis with a fold-change threshold of 2.0 and a P value $\leq$0.05; and (*ii*) analysis of variance (ANOVA) with a P value $\leq$0.05, using Benjamini and Hochberg false discovery rate (FDR) multiple-testing correction (12) followed by Tukey's pair-wise comparison (Table S1).

To assess whether the hybridization signals that had been detected were specific to cDNA spots, we compared the mean spot intensity (meanS) with the mean of the median local background intensity (meanB) for each slide. Each spot was given a reliability score ranging from 0 to 4 (0 = meanS $\leq$ meanB; one point is given for each standard deviation (SD) greater than the background, and 4 = meanS $>$ meanB + 3SD). Scores for all replicates were averaged. Spots with scores of 3 and 4 are considered highly reliable. Of the 13,962 sequences tested on the microarray, 12,616 sequences (90.4%) had an average score of 4 and 1130 sequences (8.1%) had an average score of

3 (Table S4). Approximately, 98% (765 of 781) transcripts found to be significantly differentially regulated in at least one pairwise comparison of the microarray analyses had an average reliability score of 3 or 4.

Quantitative AI Assay. To determine the amount of biologically active AI predicted to be remaining in HOSW (pH 8.2 at 23°C) throughout the 18-h incubation of animals to be used in the microarray experiment, we used a quantitative bioassay with *Escherichia coli* (pHV200I$^-$) as described (3). Levels of AI were monitored over a 12 h period in either the presence or absence of uncolonized juvenile squid. Synthetic AI (Sigma-Aldrich) was resuspended as a solution and stored in acidified ethyl acetate (EtAc). Immediately before use, the AI solution was dispensed to a glass vial, dried under nitrogen gas and dissolved in 2 ml of HOSW to a final concentration of 6 μM. The pH and temperature of the water was measured at the start and end of each experiment. At the onset and at various times throughout the experiment, 6 μl of the HOSW containing AI were removed from the sample and added to 100 μl of EtAc. These samples were stored in airtight glass vials at -20°C before bioassay (3).

Bioluminescence Colonization Assay. The minimum concentration of exogenously added AI necessary to complement the luminescence of the *V. fischeri luxI* mutant to wild-type levels was determined as follows. Sets of 12 newly hatched squids were placed in HOSW inoculated with either the wild type or the *luxI* mutant strain of *V. fischeri*. After 6 h, the animals were transferred to fresh HOSW containing different concentrations of AI, and the level of squid bioluminescence that subsequently developed was monitored for 28 h (Fig. S3) using an automated photometer (13). Data presented are representative of two independent trials, and indicated that AI added to a concentration of 200 nM would induce and maintain the luminescence of the *luxI* mutant at nearly wild-type levels for at least 12 additional hours.

Quantitative Real-Time PCR. To verify the differential expression of a subset of genes identified in the microarray analyses, QRT-PCR was performed on selected transcripts (Table S3). To ensure that differential gene regulation identified in the microarray consistently occurs (i.e., irrespective of the cohort of animals or the reagent lot), the RNA for QRT-PCR analysis was derived from light organs that had been isolated from a fourth cohort of animals at the same time as the three sets of samples collected for the microarrays. Actin-specific primers were used as a control (Table S3) because previous work (14), as well as this study (data not shown), indicated that the levels of actin transcript do not change during the first few days of light-organ development. Standard curves were created using a 10-fold cDNA dilution series with each primer set. The Pfaffl method (15) was used to calculate the fold-change in transcript abundance between each condition. The efficiencies of all QRT-PCR reactions were between 90% and 105%, although the range between any two reactions used to determine fold-change was $<$10%. All reactions used to determine fold-change were constructed from the same set of cDNA dilutions.

QRT-PCR was performed using iQSYBR Green Supermix in an iCycler Thermal Cycler (Bio-Rad Laboratories). Gene-specific primers were designed to create a product between 83 and 148 bp (Table S3), and amplification was performed under the following conditions: 95°C for 5 min, followed by 50 cycles of 95°C for 15 sec, a specific annealing temperature (Table S3) for 15 sec, and 72°C for 15 sec. Each reaction was performed in triplicate and contained 0.2 μM primers and 3.0 mM $MgCl_2$. To determine whether a single amplicon resulted from this PCR, the presence of only one optimal dissociation temperature was assayed by incrementally increasing the temperature every 10 sec from 60 to 89.5°C.

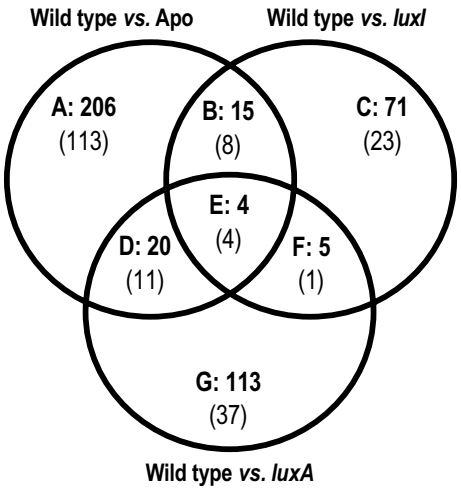
Antibody Production and Immunocytochemistry with Antibodies to EsLBP. For the production of an antibody to EsLBP, we analyzed the derived amino acid sequence of the open-reading frame of the transcript to identify a peptide region of high antigenicity, surface probably, and hydrophilicity. The resulting candidate, a 20 amino acid peptide (DNKTDCNGEQDGRHECENSQ), was conjugated to bovine gamma globulin and injected intramuscularly into chickens for production of polyclonal hen-egg antibodies. In addition, before injection with the antigen, eggs were collected for the preimmune controls. Antibodies were concentrated by polyethylene-glycol precipitation from both the eggs of preimmunized and immunized hens. The EsLBP antibody was characterized by western-blot analysis, which showed that the antibody recognizes a single peptide at the molecular mass predicted for the derived amino acid sequence corresponding to the gene encoding the protein.

Immunocytochemistry (ICC) was performed as described (14). Briefly, juvenile squid were anesthetized in 2% ethanol in seawater, and fixed overnight at 4°C in 4% paraformaldehyde in marine PBS (mPBS) consisting of 50 mM sodium phosphate buffer with 0.45 M NaCl, pH 7.4. Animals were then rinsed 4 times for 30 min in mPBS. The samples were then permeabolized for 2 days at 4°C in 1 ml of 1% Triton X-100 in mPBS with mixing. They were then blocked overnight at 4°C in a solution of 1% Triton X-100, 2% goat serum, and 0.5% BSA in mPBS. The samples were incubated with 1:1000 dilution of the anti-EsLBP polyclonal antibody in blocking solution for two weeks at 4°C. Samples were then rinsed 4 times for 1 h in 1% Triton X-100 in mPBS, and incubated overnight in blocking solution at 4°C. Fluorescein isothiocyanate (FITC)-conjugated to goat anti-chicken secondary antibody (Jackson Immunoresearch) was added at a 1:250 dilution to fresh blocking solution containing a1:2500 dilution of a rhodamine phalloidin counterstain, which labels filamentous actin, and the samples were incubated in the dark overnight at 4°C. Samples were rinsed 4 times at 30 min in 1% Triton X-100 in mPBS, followed by two final rinses in mPBS. The tissues were then counterstained with the nucleic acid TOTO-3 (Invitrogen, Inc.) following the manufacturer's instructions. Samples were then mounted on glass slides in Vectashield (Vector Laboratories), a mounting medium that retards photobleaching. Preimmune chicken antibodies (at a dilution of 1:1,000) were used as a control for non-specific binding of chicken antibodies to host tissues. To determine whether the EsLBP antibody adheres non-specifically to bacteria, we treated culture-grown cells of *V. fischeri* to the above protocol with both preimmune and immune hen-egg antibodies. The cells were then counterstained with 500 nM propidium iodide. All confocal experiments were performed on a Zeiss LSM 510 system.

Statistical Treatment of Comparisons with Mouse and Zebrafish Microarray Data. To determine whether there was a significance to the overlap between symbiosis-induced genes in the vertebrate and the squid expression studies, we calculated the probability that the 16-gene overlap in these studies could have occurred by chance. Specifically, the question was: Given that 462 of the 7503 expressed squid genes vary significantly between the symbiotic and the aposymbiotic conditions, what is the likelihood of observing 16 symbiosis-expressed squid genes among the 45 total genes shared between the squid library and the subset of vertebrate genes that are differentially expressed by interaction with the normal microbiota? To determine this likelihood, we applied both the Fisher's Exact test and the χ^2 test; because one specific gene set was analyzed, a correction for multiple testing was not needed. The finding that 16 of 45 genes are shared in the two comparisons showed high levels of significance by both analyses (Fisher's Exact test, $P = 3.3 \times 10^{-7}$; χ^2, $P = 1.2 \times 10^{-10}$).

1. Doino J, McFall-Ngai M (1995) Transient exposure to competent bacteria initiates symbiosis-specific squid light organ morphogenesis. *Biol Bull* 189:347–355.
2. Visick K, Foster J, Doino J, McFall-Ngai M, Ruby E (2000) *Vibrio fischeri* lux genes play an important role in colonization and development of the host light organ. *J Bacteriol* 182:4578–4586.
3. Schaefer AL (2000) Detection, purification, and structural elucidation of the acylhomoserine lactone inducer of *Vibrio fischeri* and other related molecules. *Methods Enzymol* 305:288–301.
4. Boettcher KJ, Ruby EG (1995) Detection and quantification of *Vibrio fischeri* autoinducer from symbiotic squid light organs. *J Bacteriol* 177:1053–1058.
5. Chun, C, *et al.* (2006) An annotated cDNA library of juvenile *Euprymna scolopes* with and without colonization by the symbiont *Vibrio fischeri. BMC Genomics* 7:154.
5. Blackshear PJ, *et al.* (2001) The NIEHS *Xenopus* maternal EST project: interim analysis of the first 13,879 ESTs from unfertilized eggs. *Gene* 267:71–87.
6. Franco GR, *et al*. (1995) Identification of new *Schistosoma mansoni* genes by the EST strategy using a directional cDNA library. *Gene* 152:141–147.
7. Nunes FM, *et al.* (2004) The use of Open Reading frame ESTs (ORESTES) for analysis of the honey bee transcriptome. *BMC Genomics* 5:84.
8. Wang D, Huang J, Xie H, Manzella L, Soares MB (2005) A robust two-way semi-linear model for normalization of cDNA microarray data. *BMC Bioinformatics* 6:14.
9. Whitfield CW, *et al.* (2002) Annotated expressed sequence tags and cDNA microarrays for studies of brain and behavior in the honey bee. *Genome Res* 12:555–566.
10. Peng X, *et al*. (2003) Statistical implications of pooling RNA samples for microarray experiments. *BMC Bioinformatics* 4:26.
11. Benjamini Y, Drai D, Elmer G, Kafkafi N, Golani I (2001) Controlling the false discovery rate in behavior genetics research. *Behav Brain Res* 125:279–284.
12. Millikan DS, Ruby EG (2004) Vibrio fischeri flagellin A is essential for normal motility and for symbiotic competence during initial squid light organ colonization. *J Bacteriol* 186:4315–4325.
13. Kimbell J, McFall-Ngai M (2004) Symbiont-induced changes in host actin during the onset of a beneficial animal-bacterial association. *Appl Environ Microbiol* 70:1434–1441.
14. Pfaffl M (2001) A new mathematical model for relative quantification in real-time RTPCR. *Nucleic Acids Res* 29:e45.
15. Handley S, Dube P, Miller V (2006) Histamine signaling through the H(2) receptor in the Peyer's patch is important for controlling *Yersinia enterocolitica* infection. *Proc Natl Acad Sci USA* 103:9268–9273.

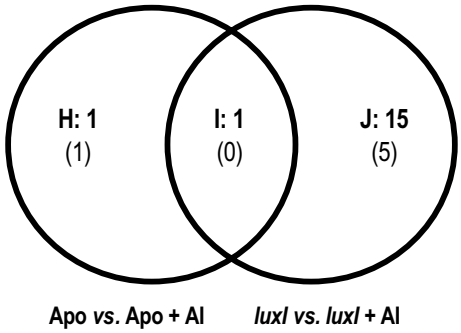
A



Number of light-organ transcripts regulated in response to:

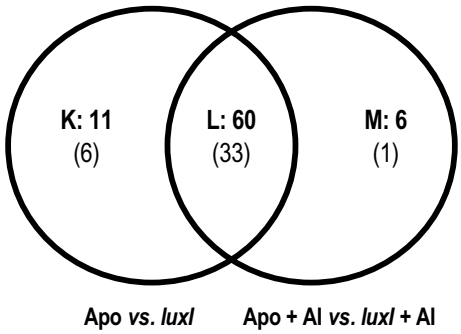
A – the presence of symbionts, regardless of light or AI
B – AI, regardless of symbionts
C – AI, in the presence of symbionts
D – light, regardless of symbionts
E – AI or light, regardless of symbionts
F – AI or light, in the presence of symbionts
G – light, in the presence of symbionts

B



H -- added AI, in the absence of *luxI* symbionts
I -- added AI, regardless of light or *luxI* symbionts
J -- added AI, in the presence of *luxI* symbionts

C



K -- the presence of *luxI* symbionts, in the absence of AI
L -- the presence of *luxI* symbionts, regardless of AI
M -- the presence of *luxI* symbionts, in the presence of added AI

Fig. S1. Venn diagrams summarizing transcriptional responses of the light organ to (*A*) colonization by symbionts that either do or do not produce light or AI, and (*B* and *C*) colonization by the *luxI* mutant in the presence or absence of added AI. The responses defined by each quadrant (expression classes A-M) are listed. The numbers of total transcripts, as well as those with known annotations (in parentheses), that are differentially regulated ≥2 fold in all of the relevant comparisons are given.

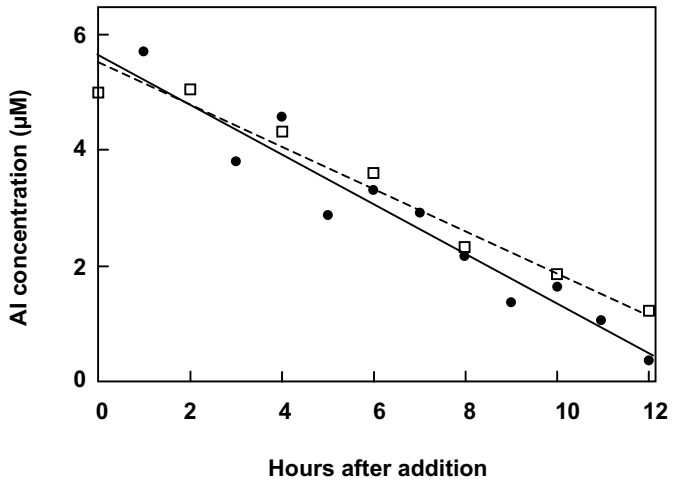


Fig. S2. Rate of spontaneous degradation of AI in seawater. AI concentration, as determined by bioassay (see *SI Text*) decreased linearly when synthetic AI (6 μM) was added to HOSW. Data are means of triplicate samples from three independent trials. Similar results were found when aposymbiotic squid were either absent (closed circles) or present (open squares) in the HOSW. Rates of degradation, calculated for these two conditions by linear regression, are indicated by the solid and dotted lines, respectively.

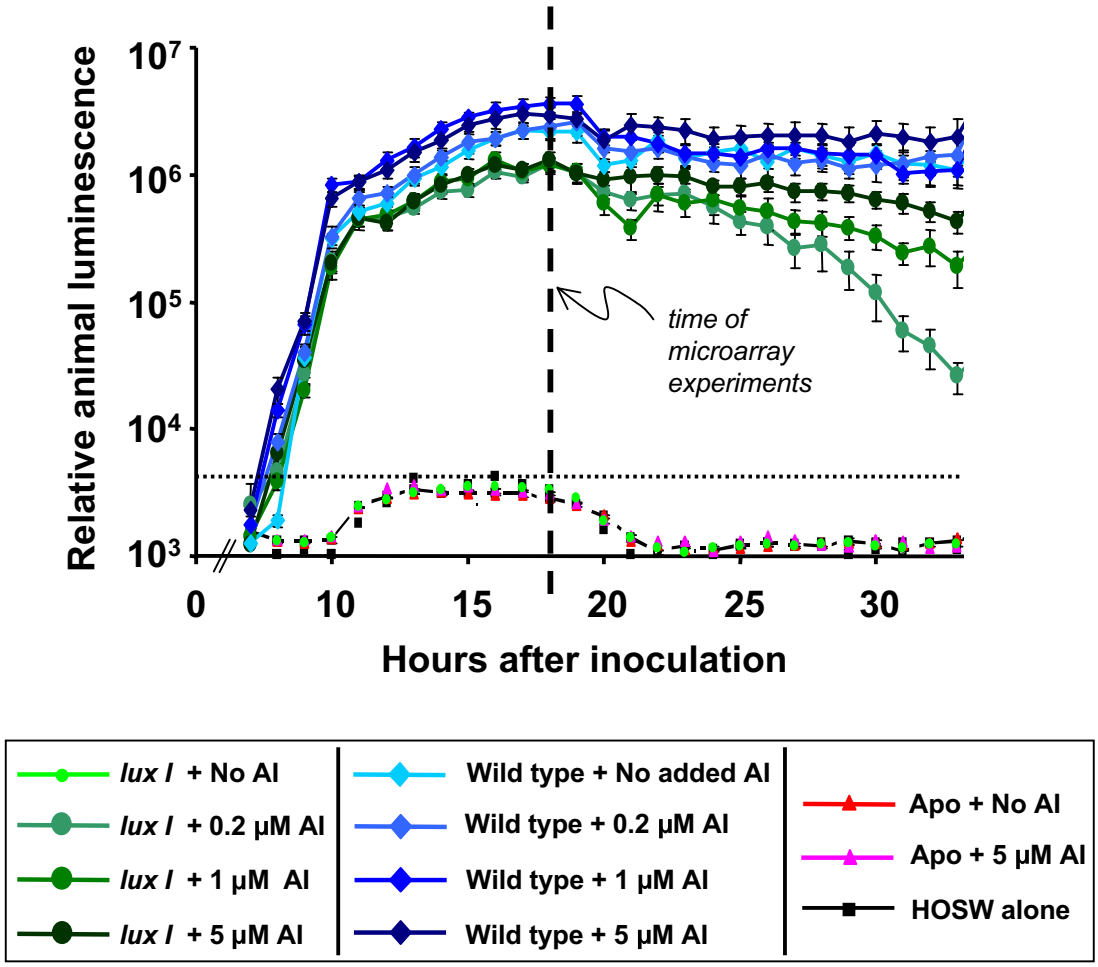


Fig. S3. Complementation of the symbiotic luminescence of a *V. fischeri luxI* mutant by the addition of synthetic AI. Sets of 12 newly hatched juvenile squid were placed in HOSW inoculated with either *V. fischeri* wild type (diamonds) or a *luxI* mutant strain (circles). After 6 h, the animals were transferred to uninoculated HOSW to which different concentrations of synthetic AI were added, and the development of symbiotic luminescence was followed. The luminescence level was determined every hour and averaged over the sets. Data shown represent the results of one of two independent trials, both of which showed the same trend. The time (18 h postinoculation) at which animals were sampled for the microarray experiments (see *SI Text*) is indicated by the vertical dashed line. Luminescence of aposymbiotic (uninoculated) animals, and animals inoculated with the *luxI* mutant in the absence of added AI, remained below the maximum level of background light (dotted line). The small increase in background luminescence observed between 10 and 22 h was due to an instrumental light leak during the daily 12-h period that the room lights were on (the animal's ''day''). More sensitive measurements indicated that the *luxI* mutant (without AI) produced bioluminescence at ≈0.1% of wild-type levels (ref. 2 and data not shown).

Table S1. Summary of ranges for light-organ transcript differentially regulated under different conditions

Comparisons*	Total number of transcripts†	Range of fold-changes of differentially regulated transcripts‡			Expression class§
		Known	Unknown/Hypo	No hits	
Category conditions					
Symbiont vs. no symbiont	132	8.6 – (−3.7)	9.4 – (−3.4)	2.8 – (−5.0)	N
Light vs. no light	27	4.7 – (−1.8)	4.9 – (−2.3)	2.3 – (−2.5)	P
AI vs. no AI	10	3.4 – 1.5	2.6 – 1.5	n.a.¶	O
Individual conditions					
Wild type vs. apo	462	130 – (−20)	29 – (−13)	100 – (−10)	A, B, D, E
Wild type vs. *luxA*	211	69 – (−48)	64 – (−27)	100 – (−10)	D, E, F, G
Wild type vs. *luxI*	131	21 – (−25)	54 – (−180)	8.9 – (−37)	B, C, E, F
Wild type vs. *luxI* + AI	24	11 – (−7.1)	12 – (−50)	4.2 – (−2.3)	Q
Apo vs. apo + AI	13	2.4	27	n.a.	H
Apo vs. *luxI*	154	3.5 – (−33)	4.3 – (−25)	3.6 – (−20)	K, L
Apo + AI vs. *luxI* + AI	148	4.5 – (−10)	3.3 – (−4.8)	20 – (−3.1)	L, M
luxI vs. *luxI* + AI	32	3.4 – (−11)	26 – (−8.3)	24 – (−9.1)	I, J

*Differences between conditions sharing the same category of exposure condition. ''Symbiont'' includes wild type, *luxI*, *luxI* + AI, and *luxA*; ''no symbiont'' includes apo and apo + AI; ''light'' includes wild type and *luxI* + AI; ''no light'' includes apo, apo + AI, and *luxA*; ''AI'' includes wild type, *luxA, luxI* + AI, and apo + AI; and ''no AI'' includes apo and *luxI.*

†Number of transcripts differentially regulated between conditions (see *Methods*).

‡Known, related to a described gene; unknown/hypo, undescribed or hypothetical protein; no hits, no significant homology to the nonredundant database of Genbank as determined by BLASTX analysis.

§Letter designation indicates the expression class in Fig. S1 and/or Table S4.

¶n.a., not applicable.

Table S2. Transcripts regulated in grouped conditions (the presence of symbionts, luminescence and/or AI)

Identifier	Annotation of transcript*	Fold change† S/NS	L/NL	A/NA‡
Require symbionts producing luminescence and AI§ (*n* = 8)				
SQabh-m-16	Lipopolysaccharide binding protein (LBP/BPI)	8.6	4.4	3.4
SQaaf-h-03	Peptidoglycan recognition protein 1 (PGRP1)	7.3	4.7	2.2
SQaab-n-03	Galaxin-1 (invertebrate protein; unknown function)	4.8	3.3	2.6
SQabd-a-21	Galaxin-2 (invertebrate protein; unknown function)	4.1	2.9	2.2
SQaay-n-03	Tetraspanin 3 (TSPAN3) regulator of membrane protein trafficking	2.2	1.9	1.5
SQaaq-p-05	Unknown	9.4	4.9	2.6
SQabd-h-03	Unknown	4.1	3.5	1.9
SQaaw-d-09	Hypothetical	2.8	2.4	1.5
Require symbionts producing luminescence (no AI)				
SQaab-c-08	Rhomboid domain-containing 1 (RHBDD1) membrane protease	3.4	2.7	
SQaao-n-19	Glutamate receptor, ionotropic, N-methyl D-aspartate-assoc. 1 (GRINA)	2.2	2.4	
SQaae-d-01	M2 (small) subunit of ribonucleotide reductase (RRM2) (DNA synthesis)	2.0	2.5	
SQaal-j-05	Guanylate cyclase 1, soluble, beta 2 (GUCY1B2), retina-associated	−2.4	−1.8	
SQaap-o-17	Unknown	2.7	2.1	
SQabc-k-15	No significant hits	2.5	2.0	
SQaaa-g-18	Hypothetical	−2.4	−2.1	
SQabf-h-04	Unknown	−2.7	−2.3	
SQaaf-i-15	No significant hits	−3.2	−2.5	
Require symbionts producing AI (no luminescence) *(n = 2)*				
SQabb-g-13	CCAAT enhancer binding protein (CEBPB)	2.4		1.7
SQaac-i-20	DC12 (unknown function)	2.3		1.7
Require symbionts (no luminescence, no AI) *(n = 51)*				
SQabe-n-18	ETS-family transcription factor (ELF3)	3.3		
SQaao-p-19	Fas apoptotic inhibitory molecule 2 (FAIM2)	3.2		
SQaaj-n-15	Calsequestrin 1 (CASQ1) calcium-binding protein	2.9		
SQaak-m-10	Transposase	2.6		
SQaac-f-23	Vertebrate adenosine A3 receptor (ADORA3)	2.4		
SQaba-a-08	Cullin-3 (CUL-3) (proteasome/ubiquitin pathway of protein degradation)	2.3		
SQabi-l-15	Integrin, beta 1 (ITGB1) extracellular-matrix protein	2.3		
SQaaa-n-06	Solute carrier family 1 (GLAST), glutamate transporter	2.3		
SQaar-j-17	Super cysteine-rich protein (SCRP)	2.2		
SQabd-d-05	Low-density lipoprotein receptor-rel. prot. 2 (Megalin, gp330) (endocytosis)	2.2		
SQaaq-f-07	Peptidoglycan recognition protein 2 (PGRP2)	2.0		
SQaaw-l-02	Annexin A7 (Synexin) calcium-binding protein	2.0		
SQaai-b-17	Receptor (TNFRSF)-interacting serine-threonine kinase 1	−2.0		
SQabf-n-17	Dynein heavy chain 7 (DHC7), cilia associated	−2.0		
SQaas-p-14	Testis specific gene A2 (TSGA2) (assoc. with motility, and egg interaction)	−2.0		
SQaab-b-03	TBC (Tre-2/Bub2/Cdc16) domain (GTPase activation)	−2.0		
SQaag-l-03	AKAP-associated sperm protein (ASP) (phosphorylation; sperm motility)	−2.0		
SQaaz-m-12	Doublecortin and CaM kinase-like 3 (DCAMKL3), microtubule-associated	−2.0		
SQabk-g-07	Golgi autoantigen, golgin subfamily a, 4 (membrane trafficking)	−2.1		
SQabc-h-02	Uroporphyrin-III C-methyltransferase (regulation of porphyrin synthesis)	−2.1		
SQaau-o-20	Stromal protein associated with thymus and lymph nodes isoform 2	−2.1		
SQaab-j-01	Zonadhesin isoform 2 (ZAN), sperm protein associated with binding to egg	−2.1		
SQabe-l-09	Mitogen-act. prot. kinase kinase kinase 9 (MAP3K9) (c-Jun/JNK apoptosis)	−2.2		
SQabk-o-09	Sperm-tail protein SHIPPO1 isoform 2 (ODF3)	−2.2		
SQaat-f-20	Reticulocyte binding protein 2b	−2.3		
SQabb-h-05	Lamin (LAM) nuclear-associated intermediate filament	−2.3		
SQaae-h-11	Excision repair complementing rodent repair deficiency, comp. (ERCC8)	−2.4		
SQaaw-l-14	Annexin A7 (Synexin) calcium binding protein	−2.6		
SQabc-e-07	Ciliary rootlet coiled-coil, rootletin protein	−2.6		
SQaae-j-15	Axonemal dynein heavy chain 7 (DNAHC7) sperm/ciliary microtubule prot.	−2.7		
SQaac-h-19	Dual-specificity tyrosine-(Y)-phosphorylation reg. kinase4 (DYRK4) (testes)	−3.7		
SQaaw-i-02	No significant hits	2.8		
SQaaq-c-05	No significant hits	2.8		
SQabj-d-01	Unknown	2.7		
SQaaq-d-04	No significant hits	2.6		
SQaam-h-22	Hypothetical	2.2		
SQaap-m-24	No significant hits	2.0		
SQaaf-h-16	Unknown	−2.0		
SQaaa-f-22	No significant hits	−2.0		
SQaaw-d-08	No significant hits	−2.1		

Identifier	Annotation of transcript*	Fold change†		
		S/NS	L/NL	A/NA‡
SQaar-l-19	Unknown	−2.1		
SQaax-c-19	Hypothetical	−2.1		
SQaaj-e-07	Unknown	−2.3		
SQaar-n-21	No significant hits	−2.3		
SQaac-h-09	Hypothetical	−2.3		
SQaat-b-08	Hypothetical	−2.3		
SQaao-l-06	No significant hits	−2.5		
SQabc-l-03	Hypothetical	−2.6		
SQaad-j-22	No significant hits	−2.8		
SQaab-e-07	Hypothetical	−3.4		
SQaag-n-09	No significant hits	−5.0		
Require luminescence only *(n = 2)*				
SQabf-o-02	No significant hits	2.3		
SQabh-m-05	Unknown	2.1		
Total number of transcripts in each category:		70	19	10

*Classified based on whether their responses required one, two, or all three of the conditions; no transcripts were found that responded either solely to the presence of AI, or to light and AI.

†Those transcripts up- or down-regulated ≥2-fold in all four normalizations, calculated as the first condition relative to the second; S/NS, symbionts vs. no symbionts; L/NL, light vs. no light; or A/NA, AI vs. no AI.

‡Transcripts in the A/NA category were not always significantly differentially regulated in all four normalizations.

§Category of grouped conditions; *n* = total number of transcripts.

Table S3. QRT-PCR confirmation of microarray transcript regulation

Gene	Fold change		Primer sequences	Annealing temperature/product size*
	Microarray	QRT-PCR		
WT vs. Apo†				
Transcription factor IIB (TFIIB)	−1.9	−1.8	Forward: 5′AAT GCC GAT GCG TCT TGA TGA TGG Reverse: 5′ AAT TGC TGC CAT AAG CTC TGC GTG	58°C/83 bp
NADPH oxidase (gp91phox) (CYBB)	−1.8	−2.0	Forward: 5′GCC AAC ACC TGA CCA ACT TCC AAT Reverse: 5′TTC CCG TGC CGA TAA ATA CGT CCA	58°C /104 bp
Matrix metalloproteinase 17 Preproprotein (MMP17)	3.2	2.7	Forward: 5′GCC AGA TTG GTT GGC TTT CCT CTG Reverse: 5′GAC GCA GCC ATT TCG TCC GAT AAC	61°C /113 bp
Low density lipoprotein receptor-related protein 2 (Megalin, gp330)	4.3	4.0	Forward: 5′TTC AAT GCG CGC ACT AAT TGG AGG Reverse: 5′ACT TAG CCG CCA CTA TGA AGC TGA	58°C /122 bp
M2 (small) subunit of ribonucleotide reductase (RRM2)	2.8	2.6	Forward: 5′AGA ATT GTC GCC TTT GCT GCT GTG Reverse: 5′CCG GCA TCA CAG AGC GTT TCT TTA	58°C / 91 bp
LBP	11.5	10.5	Forward: 5′CTG ACT GCA ATG GAG AGC AAG ACG Reverse: 5′ CAC TGA CTG CCT TAC ACT GGC AAC	62°C / 84 bp
Glutathione peroxidase	−1.9	−1.7	Forward: 5′CCA GAT GAA TGA GCT GGT CG Reverse: 5′CCA GGA CGG ACA TAG CAA AG	61°C /132 bp
WT vs. *luxA*†				
Hemocyanin	68.9	8.7	Forward: 5′CAG TAG TCG GTC TGT TCC AAG GCT Reverse: 5′TTA GTC CAG AGA CGA TGA CCG CAC	62°C /122 bp
MMP19	−2.3	2.3	Forward: 5′ TCC ACC GAC TAC AAC CAC GAA CAA Reverse: 5′ CCT TTG CAT CTG TGA AGG CTG CTT	62°C / 92 bp
Colony stimulating factor	15.4	5.0	Forward: 5′TCG CCC GTG GAA ATT ACG ATC CTG Reverse: 5′GAT GGC GCG TGT TTG TTC AGC TTC	61°C /103 bp
Control transcript	not	not	Forward: 5′GAG CGT AAA TAC TCT GTC	
Actin	regulated	regulated	Reverse: 5′GAG AAT TTG TAG AGT AGC G	56°C /148 bp

*Annealing temperature at which efficiencies were between 90 and 105%.
†Conditions being compared.

Table S6. Symbiosis-regulated genes shared among different hosts

Regulated in zebrafish and mouse, but not in *E. scolopes*

1. angiogenin 4
2. C-reactive protein, pentraxin-related
3. deleted in malignant brain tumors 1; crp-ductin
4. deoxythymidylate kinase (thymidylate kinase)
5. farnesyl diphosphate synthetase
6. fat specific gene 27
7. RNA-binding protein FUS
8. glial cell line derived neurotrophic factor family receptor alpha 1
9. interferon-related developmental regulator 1
10. immunoresponsive gene 1
11. peroxisome proliferative activated receptor, alpha
12. scinderin
13. serum/glucocorticoid regulated kinase
14. transferrin receptor

Regulated in zebrafish and mouse; present in *E. scolopes,* but not regulated

1. angiotensin converting enzyme
2. apolipoprotein B
3. arginase 2
4. ATP-binding cassette, sub-family C (CFTR/MRP), member 2
5. B-cell leukemia/lymphoma 6
6. complement component 3
7. cytochrome P450, family 7, subfamily A, polypeptide 1
8. exostoses (multiple) 1
9. fasting-induced adipose factor; angiopoietin-like 4
10. FK506 binding protein 5
11. flap structure specific endonuclease 1
12. four and a half LIM domains 1
13. fructose-1,6-bisphosphatase 1
14. growth arrest and DNA-damage-inducible, beta
15. H2A histone family, member X
16. heterogeneous nuclear ribonucleoproteins methyltransferase-like 2 (S. cerevisiae)
17. interferon-induced protein with tetratricopeptide repeats 1
18. lactate dehydrogenase B
19. lectin, galactoside-binding, soluble, 9 (galectin 9)
20. nuclear receptor subfamily 1, group D, member 2; Rev-ErbA-beta
21. peptidylprolyl isomerase C-associated protein
22. phenylalanine hydroxylase
23. phosphatidylinositol 3-kinase, regulatory subunit, polypeptide 1 (p85 alpha)
24. serum amyloid A1
25. suppressor of cytokine signaling 3
26. tropomyosin 3, gamma
27. tryptophanyl-tRNA synthetase
28. tyrosyl-tRNA synthetase
29. VAMP (vesicle-associated membrane protein)-associated protein A, 33kDa

Regulated in zebrafish and mouse; present in *E. scolopes,* and symbiont-regulated

1. proteasome (prosome, macropain) subunit, alpha type 5
2. ADP-ribosylation factor 1
3. tumor necrosis factor, alpha-induced protein 2
4. E74-like factor 3 (ETS domain transcription factor, ELF3, epithelial-specific)
5. ribonucleotide reductase M2
6. solute carrier family 31 (copper transporters), member 1
7. arginine-rich, mutated in early stage tumors
8. phosphogluconate dehydrogenase
9. hydroxysteroid (17-beta) dehydrogenase 2
10. cathepsin L
11. N-sulfotransferase
12. solute carrier family 34 (sodium phosphate), member 2
13. calcium binding protein 5 (centrin)
14. glutathione peroxidase 2 (gastrointestinal)
15. cysteine rich protein 2
16. vacuolar protein sorting 35

Vertebrate–species annotation shown.

Table S7. Squid, zebrafish, and mouse transcripts differentially regulated in response to symbiont colonization

Transcript annotation	Predicted function	Fold change*		
		Squid	Zebrafish	Mouse
Proteasome (prosome, macropain) α 1 (PMSA1)†	NF-kappaB pathway	130	2.5	2.8
ADP-ribosylating factor 6 interacting protein (ARL6AIP)†‡	Cell cycle arrest and apoptosis	97	2.6	2.0
Tumor necrosis factor, alpha-induced protein (TNFAIP8)†‡	NF-kappaB pathway	12	2.3	5.6
ETS-family transcription factor (ELF3)‡§	Epithelial differentiation, apoptosis, iNOS	4.6	−5.3	2.6
Ribonucleotide reductase M2 (RRM2)§	Cellular proliferation, induces NF-kappaB activity	2.8	3.1	2.3
Solute carrier family (SLC1)†‡	Enhanced cellular uptake of long-chain fatty acids	2.0	−2.5	−4.0
Arginine-rich mutated in early stage tumors (ARMET)§	Unknown	1.7	3.8	2.5
Phosphogluconate dehydrogenase§	Pentose phosphate shunt/reactive oxygen species	1.7	2.3	2.6
Hydroxysteroid (17-beta) dehydrogenase 12b†‡	Metabolism	1.6	−2.3	−2.5
Cathepsin L (CTSL)‡§	Acid-dependent lysosomal cysteine protease	1.5	2.3	2.8
Sulfotransferase‡§	Development, immune response	−1.7	2.0	−2.5
Centrin†‡	Cell division	−1.8	2.8	3.0
Glutathione peroxidase (GPX)†	Oxidative burst pathway	−1.9	2.7	1.9
Ataxin (super cysteine-rich protein 2) (SCRP)†‡	Stress granule and P-body assembly, apoptosis	−1.9	3.2	3.2
Solute carrier family (SLC 25)†‡	Shuttle metabolites	−2.0	2.6	2.6
Vacuolar protein sorting 18 (VPS18)†‡	Ubiquitin (E3) ligase	−4.3	2.0	2.5

*Fold change of differentially regulated transcripts calculated as colonized over uncolonized conditions.
†Transcripts annotated as members of a similar gene family in squid, zebrafish and mouse; squid annotation is shown.
‡Transcripts, either homologs or belonging to a similar gene family, differentially regulated during infections with *Yersinia enterocolitica* [Handley S, Dube P, Miller V (2006) Histamine signaling through the H(2) receptor in the Peyer's patch is important for controlling *Yersinia enterocolitica* infection. *Proc Natl Acad Sci USA* 103:9268–9273.].
§Transcripts annotated as homologs in squid, zebrafish and mouse.

Other Supporting Information Files

Table 4(XLS)
Table 5(XLS)