

Figure S1. Establishment of gut microbiome analysis pipeline in *C. elegans*, Related to Table 1.

A. Phylogenetic tree is based on the full length of 16S rRNA genes of bacterial strains in this study using maximum likelihood method (Jukes–Cantor correction; see **Table 1** and **Data S1AA**). BIGbiome strains are highlighted in bold and colored by corresponding genus. 16S rRNA gene sequence from *Synechococcus* sp. PCC7902 (AF216946) was used as an outgroup. **B.** BIGbiome strains represent from 49-90% of core microbiome families in *C. elegans* natural habitats. See also **Data S1D**. **C.** Gut microbiome colonization increases from day 1 to day 3 adulthood in *C. elegans* N2 and JU1218 strains. **D.** Gut microbiome composition of N2 and JU1218 is similar on day 1 then diverges on Day 3 adulthood. Relative microbiome abundance was calculated from biological duplicates. See also **Data S1E**.

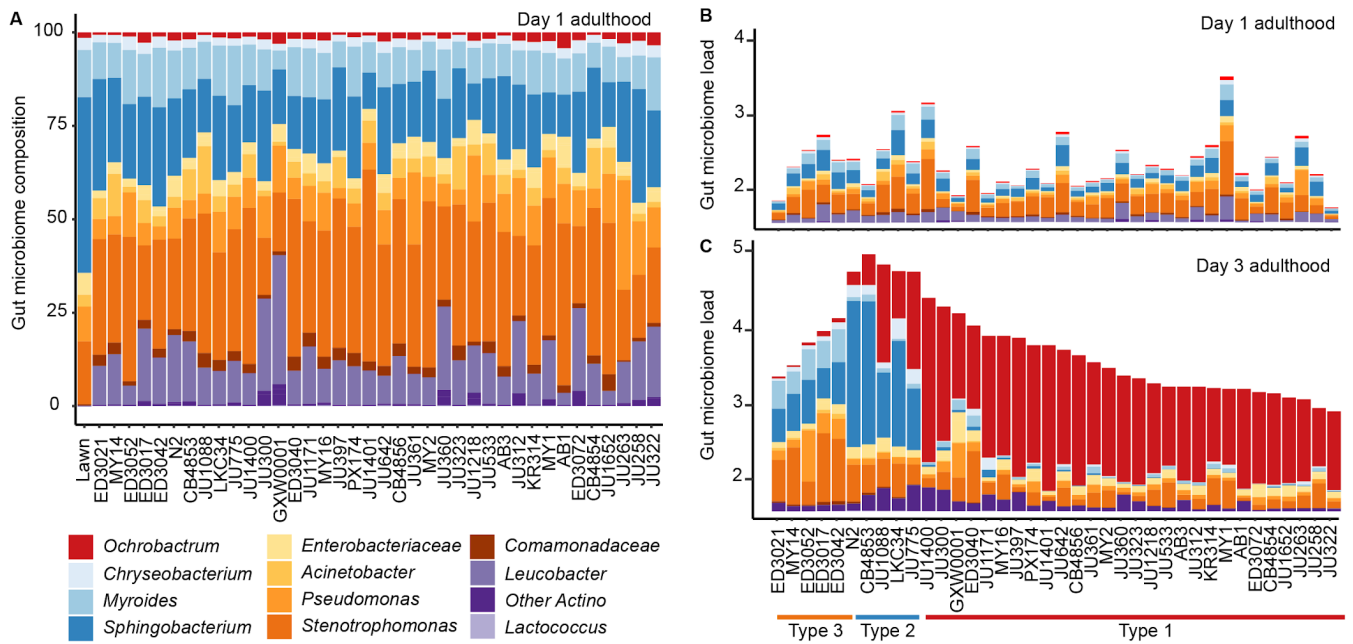


Figure S2. Minimal differences in *C. elegans* gut microbiomes at day 1 of adulthood, Related to Figure 1.

A. Gut microbiome composition of 38 *C. elegans* strains on Day 1 adulthood. Relative microbiome abundance was calculated from biological duplicates for each strain. **B.** Similar gut microbiome colonization levels of 38 *C. elegans* strains on Day 1 adulthood. Bar represented the mean calculated from biological duplicates for each strain. **C.** Gut microbiome colonization of 38 *C. elegans* strains on Day 3 adulthood. Type 2 strains, including lab-adapted N2 strain, carry on average a log higher gut microbiome load compared to Type 1 and 3 strains. Gut microbiome composition and load are calculated from the mean of biological duplicates for each strain. See also **Data S1A** and **Figure 1B-C**.

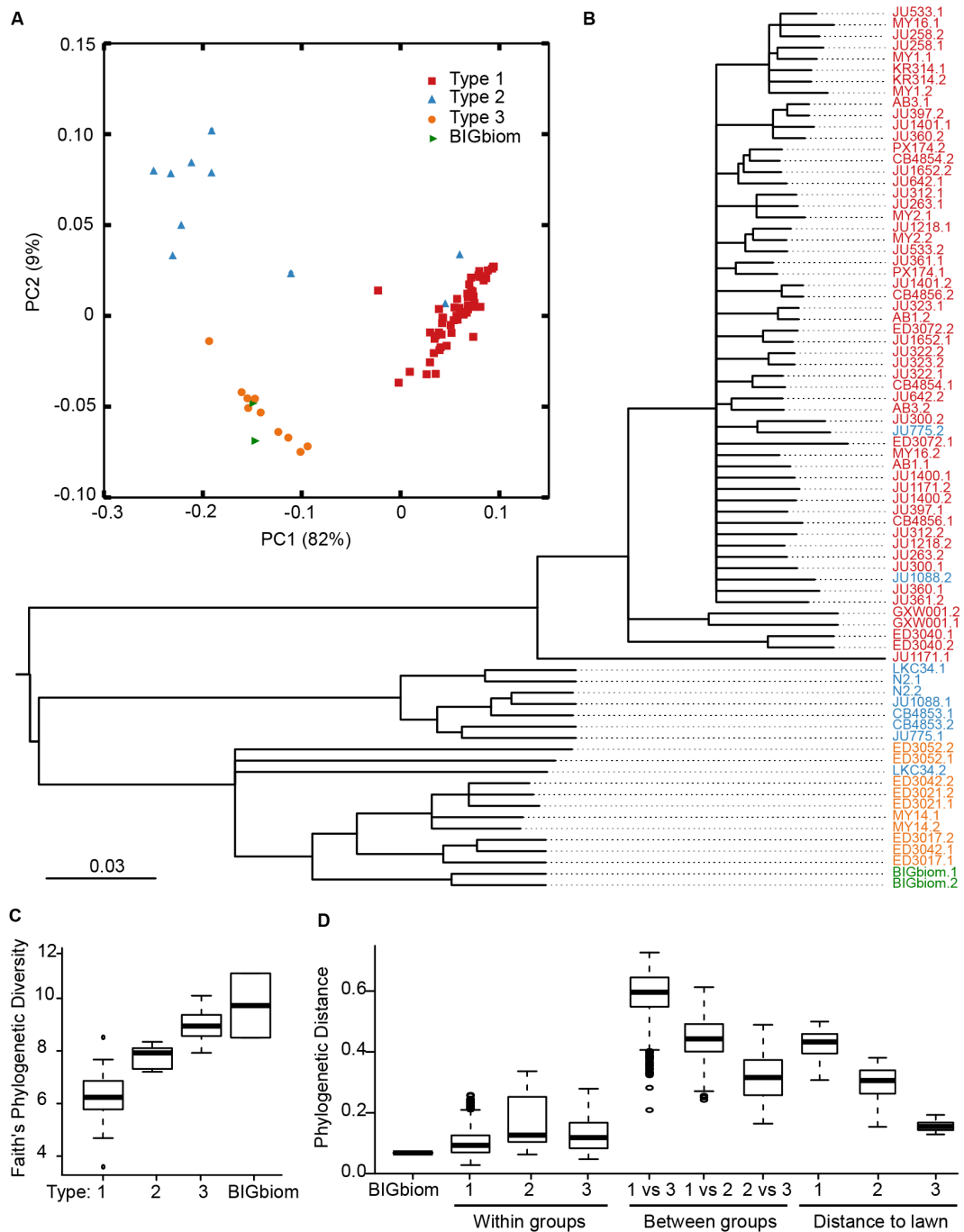


Figure S3. Microbiome diversity on day 3 of adulthood, Related to Figure 1.

A. PCoA of the microbiome composition of 38 *C. elegans* strains in Day 3 adulthood. **B.** Weighted jackknife clustering of microbiome composition of 38 *C. elegans* strains in Day 3 adulthood (colored by microbiome types). Replicates are shown under the same strain name. **C.** Box-whisker plot of Faith's phylogenetic diversity for each of the three microbiome types and BIGbiome lawn. See also **Data S1C**. **D.** Box-whisker plot of weighted UniFRAC distance between the three microbiome types and BIGbiome lawn.

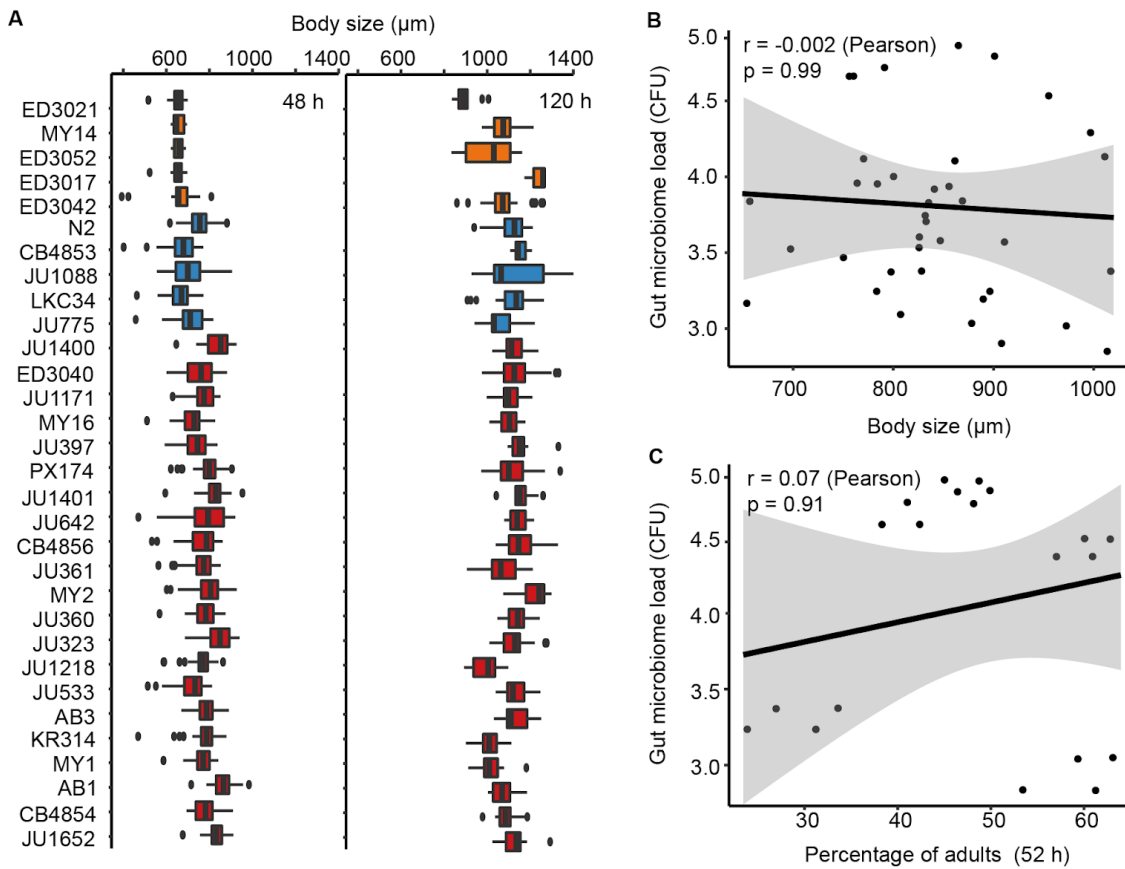


Figure S4. Gut microbiome load does not correlate with host development and body size, Related to Figure 3.

A. Box-whisker plot of body size of 38 *C. elegans* strains on day 1 and day 3 adulthood on BIGbiome. See also **Figure 3C** and **Data S1G**. **B.** No correlation observed between gut microbiome colonization and percentage of adults at 52 h post L1s ($r = -0.002$, $p = 0.99$). **C.** No correlation observed between gut microbiome colonization and worm body size at 48 h post L1s ($r = 0.07$, $p = 0.91$).

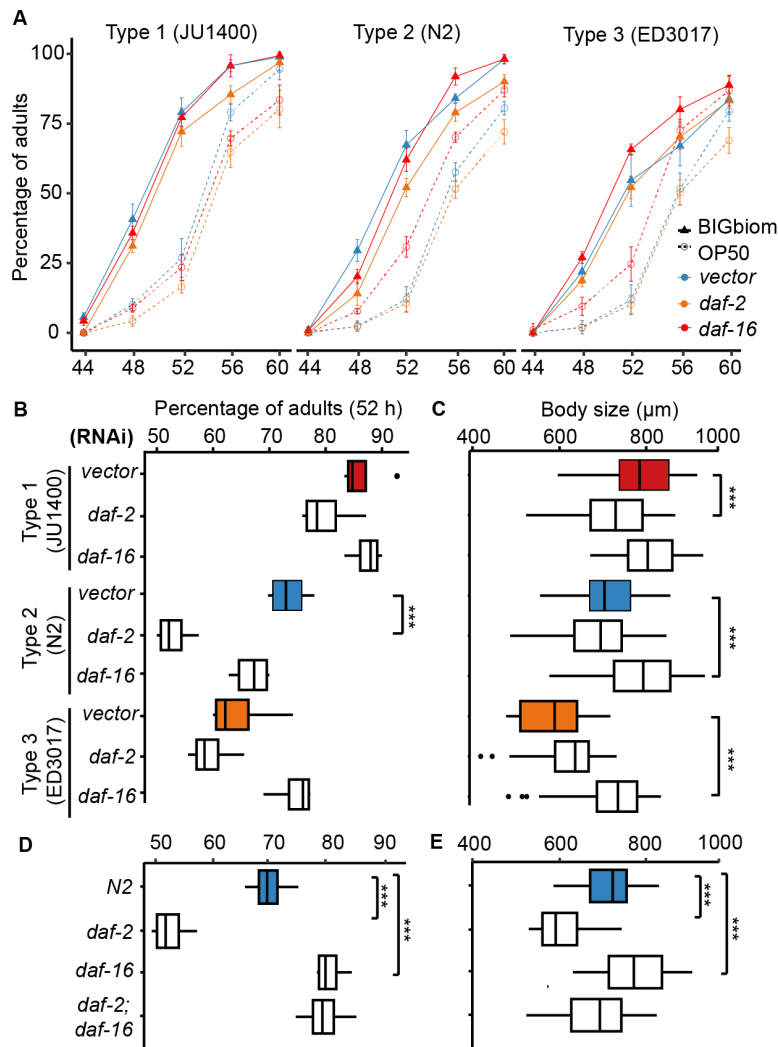


Figure S5. Development and body size of *daf-2* and *daf-16* mutants on BIGbiome, Related to Figure 5.

Additional experiments related to **Figure 5**. **A**. Developmental curves of *daf-2* and *daf-16* RNAi knockdown in representative microbiome type strains grown on the BIGbiome and *E. coli* OP50. Percentage of adults are represented as mean \pm SD with 4 replicates for each condition. **B**. Box-whisker plot of adult percentages of vector (n=4), *daf-2* (n=4), *daf-16* (n=4) knockdown in representative strains of the three microbiome types at 52 h post L1 stage. See also **Data S1Q**. **C**. Box-whisker plot of body size of vector, *daf-2*, *daf-16* knockdown mutants in representative strains of the three microbiome types at 48 h post L1 stage. See also **Data S1O**. **D**. Box-whisker plot of adult percentages of N2, *daf-2*(*e1370*), *daf-16*(*mgDf50*), and *daf-16*(*mgDf50*);*daf-2*(*e1370*) at 52 h post L1 stage (n=4 for each condition). See also **Data S1R**. **E**. Box-whisker plot of body size of N2, *daf-2*(*e1370*), *daf-16*(*mgDf50*), and *daf-16*(*mgDf50*);*daf-2*(*e1370*) at 48 h post L1 stage.

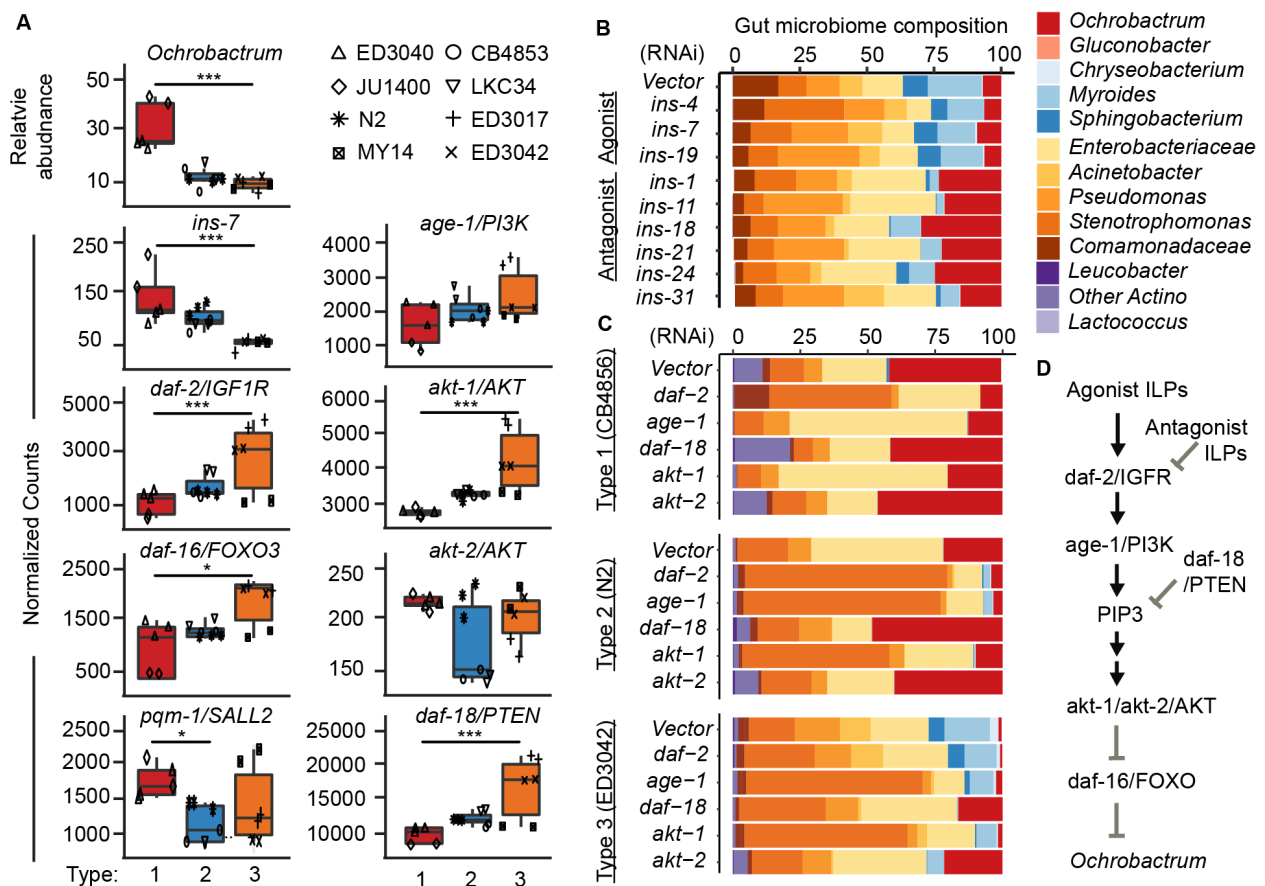


Figure S6. Key insulin signaling pathway genes that mediate recruitment of *Ochrobactrum*, Related to Figure 4.

A. Boxplots showing absolute abundance of *Ochrobactrum*, and expression of *ins-7*, *daf-2*, *daf-16*, *pqm-1*, *age-1*, *akt-1*, *akt-2*, and *daf-18* in each of the three microbiome types. Further analyses of datasets from **Data S1C** related to **Figure 4**. **B.** Gut microbiome sequence showed *Ochrobactrum* colonization increased in the intestinal specific RNAi strain JM45 with *ins-1*, *-11*, *-18*, *-21*, *-24*, *-31*(RNAi). See also **Data S1S**. **C.** *Ochrobactrum* colonization in CB4856 (Type 1) decreased with *daf-2*(RNAi), *age-1*(RNAi), and *akt-1*(RNAi). In contrast, *Ochrobactrum* colonization increased in N2 (Type 2) and ED3042 (Type 3) with *daf-18*(RNAi) and *akt-2*(RNAi) shown by bulk gut microbiome sequence of the corresponding population. See also **Data S1T**. **(B,C)** Relative microbiome abundance is presented here as the mean of biological duplicates for each strain. **D.** Schematic diagram of insulin signaling pathway in *C. elegans*.

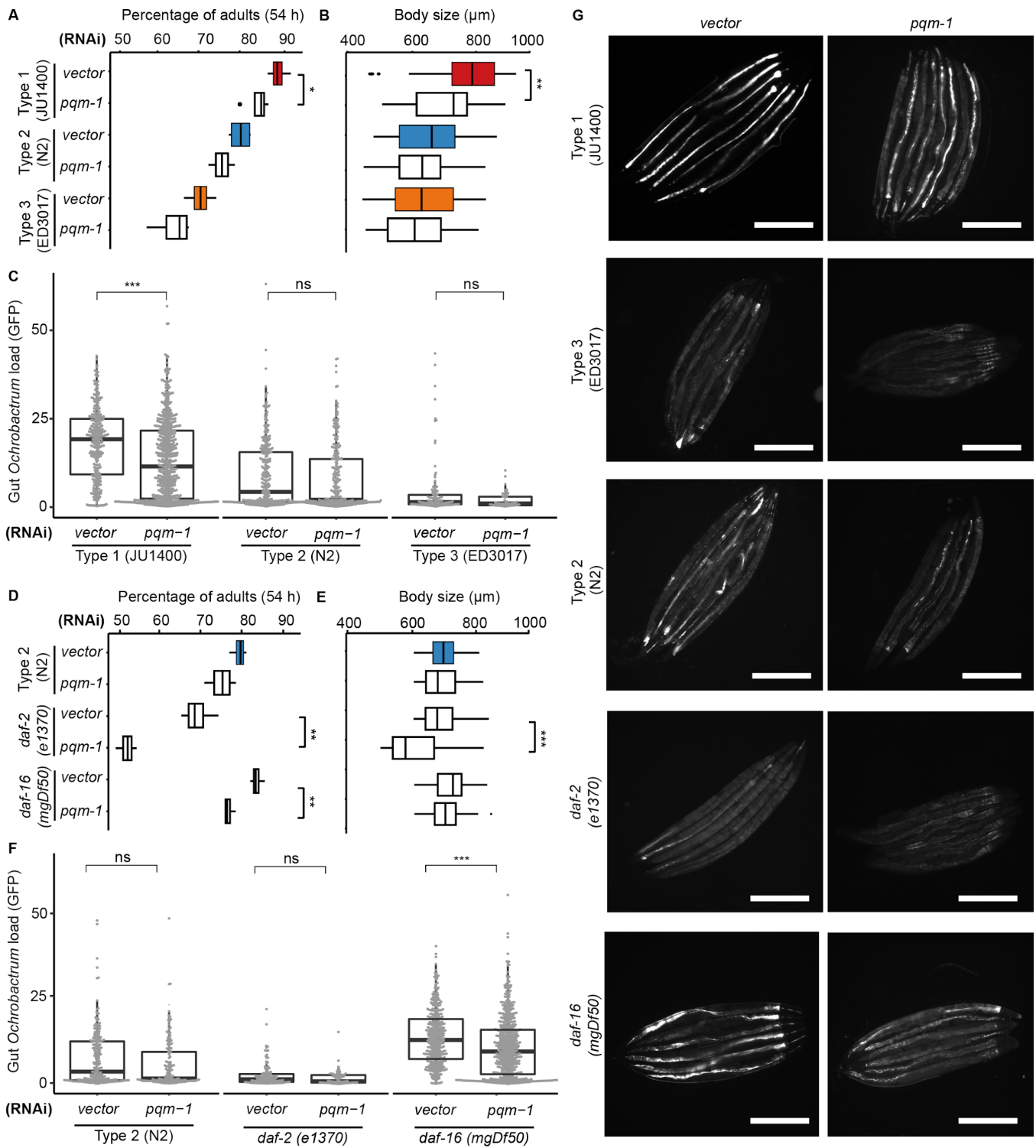


Figure S7. Impact of *pqm-1* on host development, body size, and *Ochromobacter* colonization, Related to Figure 6.

Further analyses of *pqm-1* related to **Figure 6**. **A**. Box-whisker plot of adult percentages of vector and *pqm-1* RNAi knockdown mutants in representative strains of the three microbiome types at 54 h post L1 stage. See also **Data S1V**. **B**. Box-whisker plot of body size of vector and *pqm-1* RNAi knockdown mutants in representative strains of the three microbiome types at 48 h post L1 stage. JU1400(vector): n=66, JU1400(*pqm-1* RNAi): n=85, N2(vector): n=79, N2(*pqm-1* RNAi): n=84, ED3017(vector): n=107, ED3017(*pqm-1*): n=68. See also **Figure 6AB**, **Data S1Y**. **C**. Box-whisker plot of GFP-*Ochromobacter*

colonization of vector and *pqm-1* RNAi knockdown mutants in representative strains of the three microbiome types at 120 h post L1 stage. JU1400(vector): n=306, JU1400(*pqm-1* RNAi): n=657, N2(vector): n=290, N2(*pqm-1* RNAi): n=283, ED3017(vector): n=114, ED3017(*pqm-1* RNAi): n=57. See also **Figure 6C, Data S1W**. **D**. Box-whisker plot of adult percentages of vector and *pqm-1* RNAi knockdown mutants in N2, *daf-2(e1370)*, and *daf-16(mgDf50)* at 54 h post L1 stage. See also **Data S1U**. **E**. Box-whisker plot of body size of vector and *pqm-1* RNAi knockdown mutants in N2, *daf-2(e1370)*, and *daf-16(mgDf50)* at 48 h post L1 stage. N2(vector): n=206, N2(*pqm-1* RNAi): n=267, *daf-2*(vector): n=300, *daf-2*(*pqm-1* RNAi): n=109, *daf-16*(vector): n=136, *daf-16*(*pqm-1* RNAi): n=179. See also **Figure 6DE, Data S1X**. **F**. Box-whisker plot of GFP-*Ochrobactrum* colonization of vector and *pqm-1* RNAi knockdown mutants in N2, *daf-2(e1370)*, and *daf-16(mgDf50)* at 120 h post L1 stage. N2(vector): n=237, N2(*pqm-1* RNAi): n=164, *daf-2*(vector): n=122, *daf-2*(*pqm-1* RNAi): n=75, *daf-16*(vector): n=412, *daf-16*(*pqm-1* RNAi): n=621. See also **Figure 6F, Data S1W**. (**B,E**) n represents the number of individual animals quantified by microscopic images. (**C,F**) n represents the number of individual animals quantified by Biosorter. P-values were generated from student's t-test (***p<0.001, **p<0.01, *p<0.05). **G**. Representative images of GFP-*Ochrobactrum* colonization of vector and *pqm-1* RNAi knockdown mutants in Type 1(JU1400), Type 3(ED3017), Type 2(N2), *daf-2(e1370)*, and *daf-16(mgDf50)* at 120 h post L1 stage (Bar = 500 μ m). Complete set of images from **Figure 6E**.

Strains	City of isolation	Country of isolation	Substrate of isolation
AB1	Adelaide, South Australia	Australia	NA
AB3	Adelaide, South Australia	Australia	NA
CB4853	Altadena, California	U.S.A.	NA
CB4854	Altadena, California	U.S.A.	NA
CB4856	Hawaii	U.S.A.	Pineapple
ED3017	Edinburgh	Scotland	Compost
ED3021	Edinburgh	Scotland	Compost
ED3040	Johannesburg, Gauteng	South Africa	Compost
ED3042	Ceres, Western Cape	South Africa	Compost
ED3052	Ceres, Western Cape	South Africa	Compost
ED3072	Limuru, Kiambu West	Kenya	Compost
GXW0001	Wuhan	China	Kiwi
JU1088	Kakegawa	Japan	Soil
JU1171	Conception	Chile	Compost
JU1218	Santeuil	France	Apple
JU1400	Seville, Andalucia	Spain	Orange
JU1401	Carmona, Andalucia	Spain	Snail
JU1652	Montevideo	Uruguay	Compost
JU258	Ribeiro Frio, Madeira	Portugal	Compost
JU263	Le Blanc, Indre	France	Compost
JU300	Le Blanc, Indre	France	Compost
JU312	Merlet, Lagorce (Ardeche)	France	Snail
JU322	Merlet, Lagorce (Ardeche)	France	Snail
JU323	Merlet, Lagorce (Ardeche)	France	Snail
JU360	Franconville (Val d'Oise)	France	Compost
JU361	Franconville (Val d'Oise)	France	Compost
JU397	Hermanville (Calvados)	France	Compost
JU533	Primel-Trigastrel (Finistere)	France	Compost
JU642	Le Perreux sur Marne	France	Compost
JU775	Lisbon	Portugal	Fruit
KR314	Vancouver, B.C.	Canada	Compost
LKC34	Madagascar	Madagascar	Palm seed
MY1	Lingen, Emsland	Germany	Compost
MY14	Mecklenbeck, Munster	Germany	Compost
MY16	Mecklenbeck, Munster	Germany	Compost
MY2	Roxel, Munster	Germany	Compost
N2	Bristol	U.K.	Compost
PX174	Lincoln City, Oregon	U.S.A.	Compost

Table S1. List of *C. elegans* strains used in this study, Related to Figure 1.

This table contains metadata related to *C. elegans* strains grown on BIGbiome001, including isolation locations and substrates. All strains used in **Figure 1** experiments, and representative strains used throughout the remainder of the study.

Abundance Type	Chromosome	Interval	-LOG10 (P-value)	Microbe	Variance Explained	# of Genes
Absolute	II	II:11915965-12279192	9.14	<i>Gluconobacter</i>	63%	76
Absolute	II	II:2707668-3460063	15.78	<i>Gluconobacter</i>	33%	320
Absolute	IV	IV:1941722-2759510	6.26	<i>Gluconobacter</i>		
Absolute	V	V:16336870-16718573	8.18	<i>Gluconobacter</i>	18%	166
Absolute	III	III:2173687-2578338	0.68	<i>Leucobacter</i>	0%	85
Absolute	III	III:2173687-2578338	0.68	<i>Enterobacteriaceae</i>	1%	85
Absolute	V	V:18365057-18641205	6.02	<i>Chryseobacterium</i>	17%	99
Absolute	X	X:8998713-11420509	9.32	<i>Chryseobacterium</i>	25%	778
Absolute	X	X:14807446-16227725	6.28	<i>Acinetobacter</i>	15%	477
Relative	II	II:11915965-12279192	7.69	<i>Gluconobacter</i>	56%	76
Relative	II	II:12067145-12447805	6.3	<i>Enterobacteriaceae</i>	17%	106
Relative	III	III:2494297-2803454	6.83	<i>Curtobacterium</i>	28%	58
Relative	IV	IV: 2357572-2772696	5.6	<i>Curtobacterium</i>	4%	
Relative	IV	IV: 3697125-4278131	5.6	<i>Curtobacterium</i>	4%	
Relative	V	V:13481602-14777699	6.02	<i>Curtobacterium</i>	17%	505
Relative	V	V:18810891-19550497	6.02	<i>Curtobacterium</i>	17%	221
Relative	V	V:18365057-18641205	5.69	<i>Chryseobacterium</i>	17%	99
Relative	X	X:9334761-11420509	5.69	<i>Chryseobacterium</i>	17%	661

Table S2. Summary of GWAS loci associated with bacterial taxa, Related to Figure 2.

This table is related to **Figure 2B-C** and contains direct output information from the CeNDR GWAS tool including marker, chromosome, position, user-provided trait (microbe and abundance, in this case), $\log_{10}(\text{p-value})$, significance threshold (BF), binary signifier if the trait was above the BF score, *C. elegans* strain, value, allele, var.exp, start and stop position, peak position, peak identification, and interval size (see reference 82 for more information about the CeNDR tool and information outputs).

Microbe	Correlation (Body size)	P-value (Body size)	Correlation (growth rate)	P-value (growth rate)
<i>Curtobacterium</i>	-0.48	0.00	-0.37	0.05
<i>Leucobacter</i>	0.04	0.80	0.49	0.00
<i>Arthrobacter</i>	-0.48	0.00	-0.80	0.00
<i>Rhodococcus</i>	0.06	0.72	0.06	0.70
<i>Chryseobacterium</i>	-0.39	0.02	-0.39	0.02
<i>Myroides</i>	-0.50	0.00	-0.75	0.00
<i>Sphingobacterium</i>	-0.38	0.02	-0.22	0.19
<i>Lactococcus</i>	-0.12	0.46	-0.19	0.24
<i>Ochrobactrum</i>	0.49	0.00	0.74	0.00
<i>Gluconobacter</i>	0.01	0.95	-0.10	0.56
<i>Delftia</i>	-0.39	0.02	-0.73	0.00
<i>Limnohabitans</i>	-0.61	0.00	-0.45	0.00
<i>Ramlibacter</i>	-0.33	0.04	-0.73	0.00
<i>Citrobacter</i>	0.12	0.47	-0.24	0.15
<i>Enterobacter</i>	0.45	0.00	0.24	0.15
<i>Erwinia</i>	0.10	0.54	0.22	0.19
<i>Providencia</i>	-0.30	0.07	-0.73	0.00
<i>Raoultella</i>	0.13	0.42	-0.44	0.01
<i>Yersinia</i>	0.01	0.93	-0.51	0.00
<i>Acinetobacter</i>	-0.33	0.04	-0.81	0.00
<i>Pseudomonas</i>	-0.20	0.22	-0.74	0.00
<i>Stenotrophomonas</i>	-0.41	0.01	-0.87	0.00

Table S3. Pearson correlation of gut microbial abundance with host body size and developmental growth rate, Related to Figure 3.

This table lists the Pearson correlation coefficient between microbial abundance on day 3 adults and two host phenotypes (developmental growth rate/body size) related to **Figure 3D**. p-value was computed at the level of 95% confidence interval. *Ochrobactrum* is the only microbial taxa with positive correlations with both host phenotypes ($p < 0.05$). 9 other microbial taxa show negative correlations with both host phenotypes ($p < 0.05$).