Interspecies systems biology links bacterial metabolic pathways to nematode gene expression, chemotaxis behavior, and survival

Marina Athanasouli[1,2], Tobias Loschko[1], Christian Rödelsperger[1,*]

[1]Department for Integrative Evolutionary Biology, Max Planck Institute for Biology, Max-Planck-Ring 9, 72076 Tübingen, Germany

[2]Present address: Department of Internal Medicine I and M3 Research Institute, University Clinic Tübingen, Otfried-Müller-Strasse 37, Tübingen, 72076, Germany

[*]Corresponding author: Email: christian.roedelsperger@tuebingen.mpg.de

Running title: Host microbe interactome of *Pristionchus pacificus*

Abstract

All animals live in tight association with complex microbial communities, yet studying the effects of individual bacteria remains challenging. Bacterial feeding nematodes are powerful systems to study host microbe interactions as worms can be grown on monoxenic cultures. Here, we present three different types of resources that may assist future research of cross-species interactions in the nematode *Pristionchus pacificus*, but also in other organisms. First, by sequencing the genomes of 84 *Pristionchus*-associated bacteria, we establish a genomic basis to study host microbe interactions and we demonstrate its utility to identify candidate pathways in the bacteria affecting chemotaxis behavior and survival in the nematodes. Second, we generated nematode transcriptomes of *P. pacificus* nematodes on 38 bacterial diets and characterized 60 coexpression modules with differential responses to environmental microbiota. Third, we link the microbial genome and host transcriptome data by predicting a global map of more than 2,800 metabolic interactions. These interactions represent statistical associations between variation in bacterial metabolic potential and differential transcriptomic responses of coexpression modules in the nematode. Analysis of the interactome identifies several intestinal modules as the primary response layer to diverse microbiota and reveals a number of broadly conserved metabolic interactions. In summary, our study establishes a multiomic framework for future mechanistic studies in *P. pacificus* and may also be conceptually transferred and reimplemented in other organisms in order to investigate the evolution of the host microbe interactome.

Keywords: Microbiome, interactome, *Caenorhabditis elegans*, *Pristionchus pacificus*, comparative genomics

Introduction

The human gut microbiome exemplifies the intricate associations between organisms and bacteria, highlighting the complexity of cross-kingdom interactions and their impact on host development and disease. These interactions are influenced by the heterogeneity of the gut microbiome, as well as variations in host diet and genetic background, making them challenging to study. In recent years, the nematodes *Caenorhabditis elegans* and *Pristionchus pacificus* have emerged as a powerful model for investigating host–microbiota interactions. These models allow for detailed examination of individual interactions through the use of monoxenic bacterial cultures, where worms are grown on single bacterial strains (Zhang et al. 2017; Akduman et al. 2020). Both nematodes engage with bacteria in multiple ways: they are bacterial feeders, typically consuming the *Escherichia coli* OP50 strain in laboratory settings, and the bacterial diet significantly influences its development, as exemplified by the vitamin B_{12} production by *Comamonas* that impacts *C. elegans* development (Watson et al. 2014). Moreover, both nematodes harbor complex microbiomes in their natural habitat, with certain bacteria enhancing the host's fitness under stress (Dirksen et al. 2016; Slowinski et al. 2020; Lo et al. 2022; Meyer et al. 2017). While many of the previous studies focused on understanding individual interactions at a mechanistic level (Watson et al. 2014; Iatsenko et al. 2014; Akduman et al. 2020), we lack a broad overview of the global landscape of interactions between nematodes and bacteria. Recently, we have started to characterize transcriptomic profiles of *P. pacificus* nematodes in response to diverse

microbiota and used them to test the hypothesis that the adaptation to novel microbiota might be facilitated by the evolution of novel genes (Athanasouli et al. 2023). This revealed evidence that novel genes might be preferentially integrated into specific environmentally responsive networks. However, we have very little understanding about what environmental signals these networks are sensing. In this study, we characterize bacterial metabolic variation as well as variation in nematode gene networks that respond to these bacteria. Applying a systems biology approach, we compile the first cross-species interactome of *P. pacificus* by statistically linking bacterial genomic and nematode transcriptomic data.

Results

Sequencing of 84 bacteria establishes the genomic basis to study host microbe interactions

In order to establish a phylogenomic framework that could be used to study differential effects of bacteria on *P. pacificus* nematodes, we sequenced 84 strains from a large bacterial collection isolated previously from *Pristionchus*-associated environments (Akduman et al. 2018). This collection mostly comprised Proteobacteria as well as some members of the phyla Firmicutes, Actinobacteria, and Bacteroidetes which are among the most abundant bacteria found in *Pristionchus* nematodes and their beetle hosts (Meyer et al. 2017). 84 bacterial strains were selected for whole genome sequencing based on criteria such as fast growth, ease of DNA extractions, and low propensity to contamination. (Supplemental Table S1). The resulting genome assemblies were highly complete as indicated by a median single-copy BUSCO completeness of 99% with an

interquartile range (IQR) of 98-100%. The largest parts of the bacterial genomes could be assembled into contigs spanning more than 100 kb. Specifically, the N50 value, which indicates the minimum contig length among the largest contigs that account for at least half of the genome, was 324 kb (IQR=107-498 kb) and median genome assembly size was 4.9 Mb (IQR=4.4-5.6 Mb). Gene annotation yielded a median of 4,562 gene models per assembly (IQR=4,184-5,225) with a median single-copy BUSCO completeness level of 99% (IQR=98-100%). A complete overview of these different quality measures can be found in Supplemental Table S1. This set of bacterial genomes builds the basis for the current and future studies investigating host microbe interactions in the *Pristionchus* system.

***Pristionchus*-associated microbiota harbor previously uncharacterized bacterial strains**

The predicted proteomes of all 84 genomes were taken to reconstruct a bacterial phylogeny (see Methods) (Fig. 1). The different lineages were mostly consistent with previous genus assignments based on 16S sequencing (Akduman et al. 2018). To more accurately classify the strains, we performed homology searches against the non-redundant version of NCBI and determined the nomenclature of our bacteria by assigning each strain to the bacterial genus with the majority of best hits (Supplemental Table S2). During this process, we observed a low median percentage of identity (<90%) for the strains *Sphingobacterium* L2, *Pseudomonas* L74 and *Erwinia* V71 to their respective best hits in NCBI. In addition, we were unable to definitively determine the genus for the closely related strains V69 and V91, despite a median identity with available *Phytobacter* genomes of 90.6% and

96% respectively. We hypothesized that these bacterial strains could either represent novel isolates or have no associated whole genome sequencing data in the NCBI database. Therefore, we searched the literature for recent, comprehensive phylogenies of these genera (Brady et al. 2022; Girard et al. 2021; Kakumanu et al. 2021; Smits et al. 2022) and recreated species phylogenies including our strains and their best hits in the NCBI database. The reconstructed phylogenies for the three bacterial strains with low median identity (L2, L74, V71) revealed that our strains are outgroups to the clusters containing their best hits in their respective phylogenetic trees but were still more closely related to them compared to any other genus (Supplemental Fig. S1). The reconstructed phylogeny including the two strains V69 and V91 exhibited the same patterns and indicated that both bacteria should be classified as *Phytobacter* strains (Supplemental Fig. S1). These findings support that *Pristionchus*-associated microbiota constitute a previously undersampled environment that harbors novel bacterial strains and species.

More than 1000 metabolic pathways vary across the bacterial collection

In order to characterize the variation of metabolic potential among members of the *Pristionchus*-associated microbiota, we reconstructed metabolic networks for the 84 bacterial genomes with the gapseq approach using existing pathway information from the MetaCyc database (Caspi et al. 2012; Zimmermann et al. 2021). MetaCyc was preferred to KEGG due to the higher detail it offers with regards to the available pathways. The gapseq pipeline considers a pathway to be present if at least 80% of the reaction or at least two thirds of the key reactions can be found (Zimmermann et al. 2021). This can lead to false negative calls in cases of incompletely annotated

147 pathways. Similarly, relying on genomic data can lead to false positive calls in cases
148 where a pathway is present, but not active under the tested condition. However, both
149 effects should be mitigated by the high number of bacterial genomes in our
150 collection. We then investigated the absence/presence patterns of metabolic
151 pathways among the 84 bacteria and defined metabolic pathways groups (MPGs)
152 based on identical patterns (Fig. 2). This reduced the number of total metabolic
153 pathways from 2,902 (Supplemental Table S3) to 715 MPGs. The two most
154 abundant MPGs denote 1,832 and 45 individual pathways that are either absent or
155 present in all strains (Fig. 2), respectively. While most pathways that were not
156 detected in our data correspond to pathways that are known only from eukaryotes,
157 examples of the core metabolic pathways that were annotated in all strains relate to
158 the TCA cycle, biosynthesis of several amino acids as well as fatty acid biosynthesis
159 and elongation pathways, pointing towards necessary pathways for the survival of
160 bacterial cells (Supplemental Table S3). The third most abundant pathway group
161 MPG 3 denotes 13 pathways that were not found in *Enterococcus casseliflavus*
162 strain TSA3A which is one of only three Gram positive strains in our collection (Fig.
163 1). Its genome assembly has good contiguity (N50=234kb) and is highly complete
164 (BUSCO completeness = 98%), which suggests that this pattern is unlikely due to a
165 lower quality of this particular genome. Other MPGs show strong phylogenetic
166 signatures such as MPG 7 and 13 that are restricted to the genera *Achromobacter*
167 and *Pseudomonas*, respectively. We would therefore conclude that most of the
168 variation in metabolic potential reflects genome evolution rather than technical
169 differences. To explore the amount of metabolic variation across the bacterial
170 phylogeny, we counted the number of pathway differences for pairwise comparison
171 within and across genera. This showed significantly fewer differences in the

metabolic potential for strains of the same genus (Supplemental Fig. S2). However, even strains of the same genus can differ in dozens of metabolic pathways, which can potentially explain the previously observed variability in the transcriptomic profiles and other phenotypes of *P. pacificus* nematodes in response to different bacteria from the same genus (Akduman et al. 2018; Athanasouli et al. 2023). Overall, this analysis revealed 1,025 metabolic pathways that exhibit absence presence variation in our bacterial collection.

Interspecies association studies identify candidate pathways affecting nematode survival and chemotaxis behavior

During the initial culture-based characterization of *Pristionchus*-associated microbiota, Akduman et. al. performed assays for chemotaxis behavior and survival with over a hundred bacterial strains (Akduman et al. 2018), many of which are part of our collection (Fig. 3A). In a pilot experiment, we employed our phylogenomic data to screen for candidate pathways, and their associated metabolic products, that could possibly impact nematode survival and chemotaxis behavior. To this end, we tested for associations between the presence or absence of an MPG with the behavioral and survival phenotypic data mentioned above. With this strategy, we identified 24 MPGs exhibiting significant association with chemotaxis ($P < 0.05$, Wilcoxon-test, Supplemental Fig. S3; Supplemental Table S4) and 40 MPGs that are associated with survival (Fig. 3B; Supplemental Fig. S3; Supplemental Table S5). Among these candidate pathways, chitin degradation represents one of the pathways that has previously been shown to contribute to pathogenicity of certain bacteria (Chen et al. 2015; Zhang et al. 2014). However, the most dramatic difference in survival is associated with the presence or absence of the

paerucumarin and rhabduscin biosynthesis pathways. Both pathways are partially overlapping and rhabduscin as well as coumarin derivatives have previously been reported to exhibit nematicidal activity (Abebew et al. 2022; Guo et al. 2018; Molnar et al. 2017; Ibrahim et al. 2024). However, to our knowledge, bacterial production of coumarin derivatives has never been linked to pathogenicity against nematodes (Kurz et al. 2003). To complement this analysis with an additional screen for other virulence factors, we combined the predicted proteomes of the 84 bacteria with the virulence factor database (VFDB) (Liu et al. 2022) and performed a similar association study. This yielded 15 orthologous clusters with significant association ($P < 0.05$, Wilcoxon-test), among which members of the paerucumarin biosynthesis gene cluster showed the strongest and most consistent effect (Supplemental Fig. S4). Closer examination of the presence of these genes in our data set revealed that these orthologs are exclusively present in some strains of the genus *Serratia.* Since the paerucumarin gene cluster has been initially characterized in *Pseudomonas* (Clarke-Pearson and Brady 2008), it is not known if the homologs of those genes in *Serratia* will produce exactly the same compound. Thus, hypothesizing that paerucumarin, rhabduscin, or similar metabolites could contribute to the pathogenicity of *Serratia* strains, we tested three coumarin-derivatives: 4-hydroxycoumarin, 7-methoxycoumarin, and Xanthotoxin for their effect against *P. pacificus.* All three compounds are commercially available and have previously been shown to possess some nematicidal activity in other species (Abebew et al. 2022; Guo et al. 2018; Molnar et al. 2017). Among those, 7-methoxycoumarin and Xanthotoxin caused substantial lethality at 1mM concentrations ($P < 0.05$, t-test, Fig. 3C). This supports that some coumarin derivatives have nematicidal activity against *P. pacificus* and that such compounds could potentially contribute to the increased

pathogenicity of *Serratia* strains. However, further experimental validation is needed to elucidate the underlying pathways and confirm causality.

Transcriptomic profiles of *P. pacificus* worms on 38 bacteria establish a catalog of molecular phenotypes

Having demonstrated the utility of the bacterial genomes to uncover candidate pathways impacting nematode survival and behavior, we complement the existing genomic data with phenotypic data that reflects the regulatory and metabolic response networks in the worm. To this end, we generated nematode transcriptome profiles for a subset of 38 dietary bacteria that belong to the three classes, Alpha-, Beta- and Gammaproteobacteria, and represent 16 genera. These 38 bacteria were selected as they supported complete development from egg to adult for at least two generations, which was not the case for all the strains. For each bacterial strain, we generated two biological replicates by collecting F1 worms 72 hours after egg-laying (see *Methods*). While most animals should have reached adulthood at this time under standard *E. coli* OP50 diet (Sun et al. 2021), some bacterial diets either accelerate or slow down development leading to a shift of the transcriptomic profiles towards either older or younger developmental stages (Athanasouli et al. 2023; Akduman et al. 2020). We decided to sequence the transcriptomes of mixed-stage populations as we had previously observed that even when adult worms are manually picked from plates, a strong developmental signature is still visible in the resulting RNA-seq data. This is because worms will still develop at different speeds and chronological age might not necessarily correspond to developmental age (Athanasouli et al. 2023). As a consequence, differences in stage composition likely explain the larger extent of transcriptomic variation in our study, i.e. quite a number

of samples showed correlation coefficients r < 0.9 (Supplemental Fig. S5), which was not the case in our previous study (Athanasouli et al. 2023). Generally, we observed that biological replicates of the same bacterial diet exhibit significantly less variation than transcriptomic profiles from different bacterial diets (Supplemental Fig. S5). With regard to the taxonomic grouping of the bacteria, we observe that bacterial strains of the same genus show more similar transcriptomic responses in the worms than members of different genera (Supplemental Fig. S5). In summary, our transcriptomic data represents molecular phenotypes of *P. pacificus* nematodes that exhibit considerable variation in response to diverse microbiota.

Specific network modules respond to signals from diverse bacteria

While the above-mentioned transcriptomic analysis captures overall transcriptomic similarities, specific effects of individual bacteria might be overshadowed by global patterns. To investigate the effects of individual bacteria on smaller gene sets that may represent specific regulatory or metabolic network modules, we constructed a gene coexpression network from the newly generated RNA-seq data following our previously established protocol (see *Methods*). Overall, our coexpression network captures 24,375 (84%) of *P. pacificus* genes, of which 13,972 are found in the 60 largest modules with more than 20 genes (Supplemental Table S6). To test for the consistency between the current and the previously generated coexpression network (Athanasouli et al. 2023), we compared the gene set overlap between modules of both studies (Supplemental Fig. S6). Although the current study spans a wider developmental range than our previous study, the majority of modules from the current network match only a single module in the previous network. Specifically, the

two largest modules (modules 1 and 2) of both networks appear to correspond to each other. This supports that the construction of coexpression networks is largely robust with regard to the underlying transcriptomic data set. Visualization of expression values in these modules across the different microbiota demonstrated that individual modules are induced by specific bacteria (Fig. 4). For example, the second largest module 2 shows high expression on most *Pseudomonas* strains whereas module 7 exhibits highest expression only on the *Pseudomonas protegens* strain LRB88. Note that some of these patterns may represent indirect effects of altered development on diverse bacterial diets. This is shown by distinct expression signatures of specific coexpression modules when overlaying them with the developmental transcriptomic data of *P. pacificus* (Supplemental Fig. S7) (Sun et al. 2021). For example, module 1 peaks late in development close to adulthood and is slightly preceded by a peak of module 2. However, no matter if the expression of those genes is a direct or indirect consequence, it can be interpreted as a specific response of *P. pacificus* nematodes to a distinct microbiota (Athanasouli et al. 2023). Thus, different bacterial diets induce specific transcriptomic responses of diverse modules.

Most coexpression modules can be functionally labeled

The biological interpretation of coexpression modules requires their functional annotation. Given that roughly one third of *P. pacificus* genes have no detectable homologs in *C. elegans* (Athanasouli et al. 2023) and that detailed experimental studies have revealed distinct functions of highly conserved genes (Theska and Sommer 2024; Moreno et al. 2016), the ability to functionally annotate gene sets based on conservation alone is limited. Therefore, we integrated multiple data sets to

functionally annotate the top 60 modules of the network by performing enrichment analyses. This included overrepresentation of protein domains (Mistry et al. 2021), regionally enriched genes as identified from spatial transcriptomics (Rödelsperger et al. 2021), KEGG pathways (Kanehisa and Goto 2000), and previous gene expression studies focusing on sex-biased genes (Rödelsperger et al. 2018), intestinal transcriptome (Lightfoot et al. 2016), and developmental oscillations (Sun et al. 2021) (Supplemental Table S7). This allowed the labeling of 50 modules with biological terms based on the strongest enrichment patterns (Fig. 5; Supplemental Table S8). For example, the largest module 1 (OOGEN_1) is overrepresented in hermaphrodite-biased and gonad-enriched genes, which indicates that it largely represents oogenesis (Fig 5; Supplemental Table S8). This is further supported by the expression of germline markers like *spo-11*, *dmc-1*, *hcp-4* (*cenp-c*), *syp-4*, and *hop-1* (Rillo-Bohn et al. 2021). However, module 1 also contains neuropeptides and head-enriched genes suggesting that also a portion of neuronal genes are captured in this module. The second largest module 2 (SPERM_2) seems to be more homogenous, as all overrepresented terms point towards spermatogenesis (Fig. 6). The assignments of module 1 and 2 are also in line with the module comparison with our previous study (Supplemental Fig. S6) and the developmental signature of both modules (Supplemental Fig. S7) as sperm are generated during late larval development whereas oocytes are produced during adulthood (Rudel et al. 2005). While many modules could only be associated with broad terms such as developmental oscillations or hermaphrodites, some individual modules can be assigned to very specific pathways or anatomical regions. For example, module 36 (GLAND_36) is strongly associated with gland cell specific expression (Sieriebriennikov et al. 2020). We used the transcriptomic responses to different

bacteria to arbitrarily label a handful of modules that did not show any enrichment in the above-mentioned data set (Supplemental Table S8). Thus, module 13 (LRB7_13) is strongly upregulated on *Serratia quinivorans* strain LRB7 and module 27 (LRB111_27) exhibits highest expression on *Proteus terrae* strain LRB111 (Fig. 4). Altogether, the characterization of the 60 largest coexpression modules functionally annotates a large portion of the *P. pacificus* gene set which includes many lineage-specific genes for which no homologs are known in *C. elegans* (Lo et al. 2024; Athanasouli et al. 2023). This data will be helpful to link specific transcriptomic responses to biological processes in the worm.

Linking bacterial pathways to coexpression modules predicts a global map of metabolic interactions

A main goal of this study was to test if we can trace the transcriptomic response of individual coexpression modules in *P. pacificus* to candidate metabolic pathways in the bacteria. Hereby, we define such links as statistical associations between variations in bacterial metabolic potential with differential responses of coexpression modules in the worms. To this end, we combined the bacterial MPGs and the nematode coexpression modules into a bipartite network and calculated the interaction of each MPG with a given coexpression module based on the significance of the overlap between the genes within a coexpression module and the genes that exhibit differential expression in response to the absence or presence of an MPG. This predicted a cross-species interactome of 2,852 metabolic interactions (6.7% of all possible interactions) that involved 620 (86%) MPGs and 42 coexpression modules (Supplemental Table S9; Fig. 6A). Given that such a high number of MPGs are associated with a transcriptomic response of at least one coexpression module

suggests that nematodes are highly sensitive to changes in dietary composition and that altered concentrations of most metabolites will have molecular consequences. The bacterial biosynthesis pathway that had the highest number of interactions produced phosphatidylcholine (PWY-6826) which was recently proposed as a key player regulating lipid homeostasis in *C. elegans* (Laranjeira et al. 2024). Within the interactome, we identified other metabolites and associated pathways that had already been implicated in regulating various traits in *C. elegans* and other nematodes. Thus, arginine and ornithine have been recently shown to affect reproductive fitness in *C. elegans* and parasitic nematodes (Venzon et al. 2022) and are also predicted to affect module OOGEN_1 (Supplemental Table S9). Similarly, intermediate metabolites of the pyrimidine pathways have been shown to affect reproduction and lifespan (Wan et al. 2019). Thus, although the complete cross-species interactome just represents predicted interactions that need to be validated using classical genetic approaches, it captured specific metabolites that have been previously shown to affect different biological processes in *C. elegans* and other nematodes. This suggests that the interactome displays a certain degree of conservation making it a potentially valuable resource also outside the *P. pacificus* research community.

Host microbe interactions may not be the dominant force driving the formation of novel genes in the *P. pacificus* genome

While the detailed functional investigation of individual interactions is beyond the scope of our study, we further characterize global properties of the interactome. Using the number of predicted interactions as a proxy for the environmental

responsiveness of a given coexpression module, we can identify the main modules that are responsible in sensing and processing bacterial metabolites. Notably, all of the top five modules and 11 out of the top 15 modules that interact with the highest number of bacterial pathways are associated with the intestine, detoxification or specific metabolic pathways (Fig. 6B), which is consistent with the hypothesis that variation in metabolic potential of dietary bacteria should primarily affect the nematode gut. On the contrary, several hermaphrodite-enriched modules showed very low numbers of interactions pointing towards a certain robustness of specific reproduction-associated processes. In summary, the inferred cross-species interactome identifies intestinal and detoxification related modules as components of the primary response layer to diverse microbiota. This further supports the general structure of our interactome as the module labels were inferred independently from the cross-species interactions. Furthermore, overlaying the number of interactions with the module size (Fig. 6C) and the distribution of phylostrata (gene age classes) (Athanasouli et al. 2023) across modules (Fig. 6D) shows that the most responsive modules do not contain the largest numbers of genes in total and are also not enriched in diplogastrid-specific orphan genes (Athanasouli et al. 2023). Notably, the largest module (SPERM_2) which exhibits an enrichment of diplogastrid-specific orphan genes only has an intermediate number of interactions. These findings suggest that host microbe interactions are unlikely the dominant force driving the emergence of novel genes in *Pristionchus* nematodes and point towards a major contribution of sexual conflict reflecting similar findings in flies (VanKuren et al. 2024).

Discussion

All animals live in tight associations with microbial communities and bacteria can have a significant impact on the biology of their host. However, quantifying the effects of individual bacteria remains a major challenge in many study systems. In this study, we took advantage of bacterial feeding nematodes, where individual interactions can be studied in isolation (Zhang et al. 2017). This distinguishes our work from most metagenomic studies that investigate the aggregate effects of highly complex microbial communities on their hosts (Salazar-Jaramillo et al. 2024; Shalev et al. 2022; Lo et al. 2024). In addition, in contrast to most other studies focussing on the effect of single bacteria on specific nematode traits (Iatsenko et al. 2014; Akduman et al. 2020; O'Donnell et al. 2020; Lo et al. 2022; Watson et al. 2014), we undertake a systems level approach to gain broad insights into the metabolic interactions between nematodes and their bacterial diets. We generated expression profiles of *P. pacificus* nematodes on 38 different bacterial diets. This represents to our knowledge the largest nematode transcriptomic study in response to different microbiota and this data will be of substantial value for future studies in *P. pacificus* that explore the environmental responsiveness of specific genes or gene sets. By integrating these transcriptomic and additional phenotypic data with 84 newly sequenced bacterial genomes, our study illuminates both sides of this cross-kingdom interaction and allows us to associate phenotypic and transcriptomic variation in the nematode to the underlying metabolic variation in the bacteria. This is conceptually similar to the study of Zimmermann et al. (Zimmermann et al. 2020), who associated metabolic potential of members of the *C. elegans* microbiome with specific traits, but extends this idea by considering the response of specific coexpression networks in the worm as molecular phenotypes. This established a first global map of more than

2,800 metabolic interactions between *P. pacificus* and its bacterial diets. Although the structure of this network is likely influenced by sampling of bacteria and correlated phylogenetic patterns across MPGs, it can be taken as a starting point for future studies investigating the effect of individual bacteria or metabolic pathways on specific nematode traits. Two main features of our interactome reinforce the assertion that it reflects true biological signals. First, given that the functional annotation of coexpression modules was completely independent from the association with bacterial metabolic pathways, it appears striking that most of the most responsive modules are associated with the intestine and detoxification processes. This is consistent with the assumption of the intestine as the primary response layer and thus supports the overall structure of the interactome. Second, the fact that many of the prominently highlighted metabolic pathways or metabolites like chitin degradation (Chen et al. 2015), coumarins (Guo et al. 2018), phosphatidylcholine (Laranjeira et al. 2024), arginine and ornithine (Venzon et al. 2022) have been shown to affect various biological processes in *C. elegans* or other nematodes points towards some degree of conservation of the map of metabolic interactions. The conserved nature of the metabolic interaction map suggests that our findings in *P. pacificus* may extrapolate to diverse host-microbe systems, informing mechanistic studies across taxa. Finally, to our knowledge our work represents the first application of such a systems level approach to predict a cross-species interactome in nematodes. Thus, this approach may be transferred to virtually any bacterial feeding nematode in order to establish species-specific maps of metabolic interactions and to explore the evolution of cross-species interactomes.

Materials and Methods

Bacterial culture conditions

The bacterial strains used in this correspond to the strains isolated previously (Akduman et al. 2018). All strains were initially restreaked from a glycerol stock on Lysogeny Broth (LB) plates and grown overnight in LB. All bacterial isolates were grown at 30°C according to previous protocols (Akduman et al. 2018) while *Escherichia coli* OP50 was cultured at 37°C.

Nematode growth conditions

The wild type strain of *Pristionchus pacificus* (PS312) was maintained at 20°C on nematode growth medium (NGM) seeded with 300 uL *E. coli* OP50 before use. Every five days, three adults were transferred to fresh NGM plates with a worm pick. Both NGM and *E. coli* OP50 were sourced from in-house stock.

Whole genome sequencing of bacterial strains

All bacterial strains to be sequenced were grown overnight in LB in 15 mL falcon tubes and DNA samples were obtained using the Epicentre MasterPure Complete DNA and RNA Purification Kit (Illumina, San Diego, USA). DNA libraries were prepared using the Illumina DNA Prep kit according to the manufacturer's guidelines. The libraries were quantified using both a Qubit 2.0 Fluorometer (Thermo Fisher Scientific, Waltham, USA) and a Bioanalyzer (Agilent Technologies, Santa Clara, California) and normalized to 2.5 nM. Samples were sequenced as 150bp single-end

reads on multiplexed lanes of an Illumina HiSeq 3000 in our in-house sequencing facility.

Genome assembly and metabolic potential identification of bacterial strains

The bacterial genomes were assembled from the raw reads using SPAdes (Prjibelski et al. 2020) (version 3.15.1, parameters: --careful -o -1 -2 -t 30 –cov-cutoff 'auto') and annotated with PROKKA (Seemann 2014) (version 1.14.6, parameters: --addgenes) and eggNOG-Mapper (Cantalapiedra et al. 2021) (version 2.1.0-1, eggnog DB version: 5.0.2). Genome assembly and annotation were evaluated using BUSCO (Manni et al. 2021) and QUAST (Gurevich et al. 2013). The predicted bacterial protein sequences were used as input to detect orthogroups and reconstruct the phylogeny with OrthoFinder (Emms and Kelly 2015). The bacterial genomes were used to predict the presence or absence of metabolic pathways with gapseq (Zimmermann et al. 2021) (version 1.2, parameters: find -p all, MetaCyc (Caspi et al. 2012) as the default database) as well as the presence of 67 bacterial phenotypes with the implementation of Traitar (Weimann et al. 2016) in Python (version 3.0.1, parameters: predict) . The bacterial metabolic pathways were grouped (MPGs) based on their presence/absence patterns across the bacterial strains to reduce redundancy. This decreased the number of bacterial metabolic pathways to be assessed from 2902 to 712. The nomenclature of the bacterial strains was determined by applying DIAMOND (Buchfink et al. 2021) (version 2.1.4.158, parameters: makedb and blastp) to compare the predicted bacterial protein sequences against the non-redundant version of the NCBI database (August 2023 version).

Association of MPGs with chemotaxis and survival

In the initial study, the chemotaxis index ranges from -1 (repulsion) to 1 (attraction) and the survival index from 0 to 100, the latter reflecting the percentage of worms that were alive on the 5^{th} day post transfer (Akduman et al. 2018) For our analysis, we binarized the assay data by assigning the values below 40 percent survival and 0.2 chemotaxis as zero and the rest as one. The threshold for the binarization was determined based on the distribution of values. To test the association between each MPG and the nematode phenotypes, we divided the assay data to two vectors, representing the set of values when the nematode was maintained on bacterial strains with and without the MPG and performed a Wilcoxon-test, followed by Bonferroni-correction ($P < 0.05$). To associate bacterial virulence factors with survival, we downloaded the 4,236 core proteins from the Virulence Factor DataBase (Liu et al. 2022) (2024-07-31) and performed pairwise all-against-all BLASTP searches for all bacterial protein sets combined with the VFDB data (e-value < 10^{-5}). The widely used clustering algorithm MCL was used to cluster the data into 1605 orthogroups. These orthogroups were then tested for association with survival using a Wilcoxon-test with FDR correction ($P < 0.1$).

Supplementation experiments

To check the nematicidal effect of coumarin derivatives *in vivo*, we supplemented 4-hydroxycoumarin (Sigma-Aldrich, CAS: 1076-38-6, purity: >97.5%), 7-methoxycoumarin (Sigma-Aldrich, CAS: 531-59-9, purity: >98%) and Xanthotoxin (TCI Deutschland, CAS: 298-81-7, purity: >98%) to the nematode. The coumarins were added directly to liquid NGM to a final concentration of 1 mM and the solution

was subsequently poured to 6 cm plates (11 mL per plate). The plates were seeded with 300 µL of *E. coli* OP50 the next day and left to grow a bacterial lawn for two days. After that time, 20 young adult hermaphrodites were transferred on each plate and their survival was tracked daily for five days. The surviving nematodes were transferred on new plates containing the coumarins on the third day after the initial transfer to avoid food depletion and misidentification from their offspring. Mortality was determined by lack of response to prodding with the pick. For the control and each coumarin, we performed five replicates.

Dietary experiments

Eggs from *P. pacificus* adults were obtained and spotted on NGM plates seeded with 75 µL bacterial overnight cultures to produce the parental generation (P0) on each diet. 25 gravid hermaphrodites from the P0 were transferred to a new plate seeded with the same bacterial strain and were left to lay eggs for five hours, after which point they were removed. The F1 worms were collected 72 hours after the initial transfer. Collection of the worms required washing the plate with autoclaved $M9^-$ buffer, centrifugation at 500g for a minute and discarding the supernatant in order to remove as much of the collected bacterial lawn as possible. The worm pellets were flash frozen in liquid N_2 and were used for RNA extraction or stored at -80^oC for later processing.

RNA sequencing

For total RNA extraction, the frozen worm pellets were treated with TRIzole followed by purification with the Zymo RNA Clean & Concentrator 25 Kit according to the manufacturer's instructions. The extracted RNA was quantified and quality assessed with a NanoDrop ND1000 spectrometer (PeqLab, Erlangen, Germany) and a Qubit 2.0 Fluorometer. The samples were shipped to Novogene for library preparation and messenger RNA (mRNA) sequencing. Libraries were sequenced as 150-bp paired-end reads on an Illumina NovoSeq 6000 platform.

RNA-seq and coexpression network analysis

For the coexpression analysis, we only used newly generated RNA-seq data as the samples from our previous study were obtained using different protocols and were only sequenced as single replicates. The raw reads were aligned to the reference *Pristionchus pacificus* genome (version El Paco) with STAR (Dobin et al. 2013) (version 2.7.1a) and quantified with featureCounts (Liao et al. 2014) from the Subread R package (version 2.0.1) based on the latest annotations (Athanasouli et al. 2020). After filtering the count matrix to remove genes with less than 10 reads total according to the reference manual, we reduced the number of *P. pacificus* genes to be assessed from 28,896 to 24,335. The read counts across different conditions and replicates were normalized with the DESeq2 (Love et al. 2014) counts function (option: normalized = TRUE) and were input in the Markov CLustering software (MCL) (van Dongen and Abreu-Goodger 2012)(version 22-282, r=0.7, I=2) to create a gene coexpression network. The network modules were tested for the enrichment of protein domain annotations (Pfam) (Mistry et al. 2021), metabolic pathways (KEGG) (Kanehisa and Goto 2000) and gene expression sets

559 from previous studies (Athanasouli et al. 2023) with Fisher's exact-test followed by
560 Bonferroni correction ($P < 0.05$) in R (version 4.4.0) (R Core Team 2024).
561 Phylostrata were defined as described previously (Athanasouli et al. 2023): in short,
562 BLASTP searches (e-value < 0.001) were carried out against predicted proteomes of
563 related nematodes (Prabh et al. 2018; Howe et al. 2017) and phylostrata were
564 defined based on the presence of homologs in the most distantly related species.
565 The network modules were assessed further for their interactions with the bacterial
566 metabolic potential. Specifically, for each MPG, the normalized counts of the genes
567 in each module were separated into two vectors based on whether the presence or
568 absence of the MPG in the bacterial strain in order to represent the expression of a
569 specific gene in a module in the presence or absence of a MPG. The presence and
570 absence expression vectors for each gene were compared using Wilcoxon-test as
571 well as to calculate the fold change in expression. To determine the interaction of a
572 module with a MPG, we retained the genes where the fold change was lower than
573 0.5 or higher than 2 and performed multiple hypothesis testing for the module genes
574 that fulfilled the condition (Bonferroni < 0.1). The list of genes identified as interacting
575 with the MPGs were tested for module enrichment using Fisher's exact test
576 (Bonferroni-adjusted p-value<0.05) to determine the interaction between the MPGs
577 and the coexpression modules. The bipartite network was implemented using Python
578 (version 3.10.12).

579 **Data access:** All sequencing data generated in this study have been submitted to
580 the European Nucleotide Archive under the accession number PRJEB80633.

581 **Acknowledgments**

We would like to thank Nicholas Youngblut, Boris Macek, Thorsten Langner, and all members of the Sommer lab for helpful discussions. This work was funded by the Max Planck Society.

Author contributions: Conceptualization: C.R. Investigation: M.A., and T.L. Data Curation: M.A. Formal Analysis: M.A. Visualization: M.A. and C.R. Writing – Original Draft: M.A. Writing – Review & Editing: C.R. Supervision: C.R.

Competing interests: The authors declare no competing interests.

Figure Legends

Fig. 1. Phylogenomics of 84 bacterial genomes. The tree depicts the bacterial phylogeny acquired using RAxML after whole genome sequencing, genome assembly and orthogroup detection. Phenotype prediction with traitar3 revealed that the overwhelming majority of bacterial strains sequenced are aerobic with the exception of TSA3A and LRB41 who are facultative anaerobic as well as the complete lack of anaerobes. Our collection contains only three Gram positive strains while the rest are gram negative. Genome sizes may vary even within a genus as is the case with *Acinetobacter*, where we observe genomes ranging between 3 and 5 Mb.

Fig. 2. Variation in bacterial metabolic potential. The heatmap shows the predicted presence or absence of the most abundant MPGs across our genome collection (grouped according to the bacterial phylogeny). This revealed the metabolic potential of the individual strains. The majority of MPGs detected were

absent from all strains with a core of 45 bacterial pathways predicted always present. The rest of the MPGs exhibit distinct patterns even among the same genus.

Fig. 3. Association of nematode survival with the bacterial pathways. (A) The tree shows the bacterial phylogeny together with previously determined survival data (Akduman et al. 2018). **(B)** An association study between the MPGs and mean survival data narrowed down the MPGs with a significant effect on the trait from hundreds to 41, a subsample of which is shown here. **(C)** Supplementation tests with three types of coumarins showed that 7-methoxycoumarin and xanthotoxin affect nematode survival ($P<0.05$, *t*-test).

Fig. 4. Expression of modules across the different bacterial diets. The *Z*-score normalized expression of the top 60 coexpression modules across the different diets was visualized as a heatmap. The expression of the biological replicates was averaged in order to produce a single expression profile for each of the 38 diets. Modules with more than 50 genes were randomly downsampled to aid the visualization of smaller modules and the comparisons between them. Expression patterns were unique for each module and depended on the bacterial diet, as was the case with modules 2, 7 and 15 on LRB88.

Fig. 5. Enrichment analysis of the gene coexpression modules using existing transcriptomic datasets and functional annotation. The networks summarize the results of the overrepresentation analysis for specific coexpression modules (Supplemental Table S7). Based on the strongest enrichment patterns, we connected 50 of the processed 60 modules with biological processes and protein

domains, allowing the functional labeling of the majority of the modules (Supplemental Table S8).

Fig. 6. Interactions between bacterial pathways and coexpression modules.

(A) The heatmap shows a representative interaction between pathway PWY-5651 (L-tryptophan degradation to 2-amino-3carboxymuconate semialdehyde) and module GUT_5. Twenty randomly chosen genes from this module are shown together with their expression across different bacterial diets. These genes exhibit much higher expression on bacteria that have the PWY-5651 pathway. **(B)** The bars shows the numbers of interactions per coexpression module. All of the top five and 11 out of the top 15 modules are associated with the intestine, detoxification or specific metabolic pathways. The number of genes per module **(C)** and the phylostratigraphic distribution across modules **(D)** show that the five modules exhibiting the highest number of interactions are not the largest modules and are also not enriched in diplogastrid-specific orphan genes (phylostratum ≤ 9).

References

Abebew D, Sayedain FS, Bode E, Bode HB. 2022. Uncovering Nematicidal Natural Products from *Xenorhabdus* Bacteria. *J Agric Food Chem* **70**: 498–506.

Akduman N, Lightfoot JW, Röseler W, Witte H, Lo W-S, Rödelsperger C, Sommer RJ. 2020. Bacterial vitamin B_{12} production enhances nematode predatory behavior. *ISME J* **14**: 1494–1507.

Akduman N, Rödelsperger C, Sommer RJ. 2018. Culture-based analysis of Pristionchus-associated microbiota from beetles and figs for studying nematode-bacterial interactions. *PLoS One* **13**: e0198018.

655 Athanasouli M, Witte H, Weiler C, Loschko T, Eberhardt G, Sommer RJ,
656 Rödelsperger C. 2020. Comparative genomics and community curation further
657 improve gene annotations in the nematode *Pristionchus pacificus*. *BMC*
658 *Genomics* **21**, 708.

659 Athanasouli M, Akduman N, Röseler W, Theam P, Rödelsperger C. 2023.
660 Thousands of Pristionchus pacificus orphan genes were integrated into
661 developmental networks that respond to diverse environmental microbiota.
662 *PLoS Genet* **19**: e1010832.

663 Brady C, Kaur S, Crampton B, Maddock D, Arnold D, Denman S. 2022. Transfer of
664 *Erwinia toletana* and *Erwinia iniecta* to a novel genus *Winslowiella* gen. nov. as
665 *Winslowiella toletana* comb. nov. and *Winslowiella iniecta* comb. nov. and
666 description of *Winslowiella arboricola* sp. nov., isolated from bleeding cankers
667 on broadleaf hosts. *Front Microbiol* **13**: 1063107.

668 Buchfink B, Reuter K, Drost H-G. 2021. Sensitive protein alignments at tree-of-life
669 scale using DIAMOND. *Nat Methods* **18**: 366–368.

670 Cantalapiedra CP, Hernández-Plaza A, Letunic I, Bork P, Huerta-Cepas J. 2021.
671 eggNOG-mapper v2: Functional Annotation, Orthology Assignments, and
672 Domain Prediction at the Metagenomic Scale. *Mol Biol Evol* **38**: 5825–5829.

673 Caspi R, Altman T, Dreher K, Fulcher CA, Subhraveti P, Keseler IM, Kothari A,
674 Krummenacker M, Latendresse M, Mueller LA, et al. 2012. The MetaCyc
675 database of metabolic pathways and enzymes and the BioCyc collection of
676 pathway/genome databases. *Nucleic Acids Res* **40**: D742–53.

677 Chen L, Jiang H, Cheng Q, Chen J, Wu G, Kumar A, Sun M, Liu Z. 2015. Enhanced
678 nematicidal potential of the chitinase pachi from Pseudomonas aeruginosa in
679 association with Cry21Aa. *Sci Rep* **5**: 14395.

680 Clarke-Pearson MF, Brady SF. 2008. Paerucumarin, a new metabolite produced by
681 the pvc gene cluster from Pseudomonas aeruginosa. *J Bacteriol* **190**: 6927–
682 6930.

683 Dirksen P, Marsh SA, Braker I, Heitland N, Wagner S, Nakad R, Mader S, Petersen
684 C, Kowallik V, Rosenstiel P, et al. 2016. The native microbiome of the nematode
685 Caenorhabditis elegans: gateway to a new host-microbiome model. *BMC Biol*
686 **14**: 38.

687 Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson
688 M, Gingeras TR. 2013. STAR: ultrafast universal RNA-seq aligner.
689 *Bioinformatics* **29**: 15–21.

690 Emms DM, Kelly S. 2015. OrthoFinder: solving fundamental biases in whole genome
691 comparisons dramatically improves orthogroup inference accuracy. *Genome*
692 *Biol* **16**: 157.

693 Girard L, Lood C, Höfte M, Vandamme P, Rokni-Zadeh H, van Noort V, Lavigne R,
694 De Mot R. 2021. The Ever-Expanding *Pseudomonas* Genus: Description of 43

New Species and Partition of the *Pseudomonas putida* Group. *Microorganisms* **9**. http://dx.doi.org/10.3390/microorganisms9081766.

Guo Q-Q, Du G-C, Li Y-X, Liang C-Y, Wang C, Zhang Y-N, Li R-G. 2018. Nematotoxic coumarins from *Angelica pubescens* Maxim. *f. biserrata* Shan *et* Yuan roots and their physiological effects on *Bursaphelenchus xylophilus*. J Nematol. *J Nematol* **50**: 559–568.

Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUAST: quality assessment tool for genome assemblies. *Bioinformatics* **29**: 1072–1075.

Howe KL, Bolt BJ, Shafie M, Kersey P, Berriman M. 2017. WormBase ParaSite - a comprehensive resource for helminth genomics. *Mol Biochem Parasitol* **215**: 2–10.

Iatsenko I, Yim JJ, Schroeder FC, Sommer RJ. 2014. B. subtilis GS67 protects C. elegans from Gram-positive pathogens via fengycin-mediated microbial antagonism. *Curr Biol* **24**: 2720–2727.

Ibrahim H, Nchiozem-Ngnitedem V-A, Dandurand L-M, Popova I. 2024. Naturally-occurring nematicides of plant origin: two decades of novel chemistries. *Pest Manag Sci*. http://dx.doi.org/10.1002/ps.8504.

Kakumanu ML, Marayati BF, Wada-Katsumata A, Wasserberg G, Schal C, Apperson CS, Ponnusamy L. 2021. *Sphingobacterium phlebotomi* sp. nov., a new member of family *Sphingobacteriaceae* isolated from sand fly rearing substrate. *Int J Syst Evol Microbiol* **71**. http://dx.doi.org/10.1099/ijsem.0.004809.

Kanehisa M, Goto S. 2000. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* **28**: 27–30.

Kurz CL, Chauvet S, Andrès E, Aurouze M, Vallet I, Michel GPF, Uh M, Celli J, Filloux A, De Bentzmann S, et al. 2003. Virulence factors of the human opportunistic pathogen Serratia marcescens identified by in vivo screening. *EMBO J* **22**: 1451–1460.

Laranjeira AC, Berger S, Kohlbrenner T, Greter NR, Hajnal A. 2024. Nutritional vitamin B12 regulates RAS/MAPK-mediated cell fate decisions through one-carbon metabolism. *Nat Commun* **15**: 8178.

Liao Y, Smyth GK, Shi W. 2014. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**: 923–930.

Lightfoot JW, Chauhan VM, Aylott JW, Rödelsperger C. 2016. Comparative transcriptomics of the nematode gut identifies global shifts in feeding mode and pathogen susceptibility. *BMC Res Notes* **9**: 142.

Liu B, Zheng D, Zhou S, Chen L, Yang J. 2022. VFDB 2022: a general classification scheme for bacterial virulence factors. *Nucleic Acids Res* **50**: D912–D917.

Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and

734 dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**: 550.

735 Lo W-S, Han Z, Witte H, Röseler W, Sommer RJ. 2022. Synergistic interaction of gut
736 microbiota enhances the growth of nematode through neuroendocrine signaling.
737 *Curr Biol* **32**: 2037–2050.e4.

738 Lo W-S, Sommer RJ, Han Z. 2024. Microbiota succession influences nematode
739 physiology in a beetle microcosm ecosystem. *Nat Commun* **15**: 5137.

740 Manni M, Berkeley MR, Seppey M, Simão FA, Zdobnov EM. 2021. BUSCO Update:
741 Novel and Streamlined Workflows along with Broader and Deeper Phylogenetic
742 Coverage for Scoring of Eukaryotic, Prokaryotic, and Viral Genomes. *Mol Biol*
743 *Evol* **38**: 4647–4654.

744 Meyer JM, Baskaran P, Quast C, Susoy V, Rödelsperger C, Glöckner FO, Sommer
745 RJ. 2017. Succession and dynamics of Pristionchus nematodes and their
746 microbiome during decomposition of Oryctes borbonicus on La Réunion Island.
747 *Environ Microbiol* **19**: 1476–1489.

748 Mistry J, Chuguransky S, Williams L, Qureshi M, Salazar GA, Sonnhammer ELL,
749 Tosatto SCE, Paladin L, Raj S, Richardson LJ, et al. 2021. Pfam: The protein
750 families database in 2021. *Nucleic Acids Res* **49**: D412–D419.

751 Molnar M, Komar M, Brahmbhatt H, Babić J, Jokić S, Rastija V. 2017. Deep Eutectic
752 Solvents as Convenient Media for Synthesis of Novel Coumarinyl Schiff Bases
753 and Their QSAR Studies. *Molecules* **22**.

754 Moreno E, McGaughran A, Rödelsperger C, Zimmer M, Sommer RJ. 2016. Oxygen-
755 induced social behaviours in Pristionchus pacificus have a distinct evolutionary
756 history and genetic regulation from Caenorhabditis elegans. *Proc Biol Sci* **283**:
757 20152263.

758 O'Donnell MP, Fox BW, Chao P-H, Schroeder FC, Sengupta P. 2020. A
759 neurotransmitter produced by gut bacteria modulates host sensory behaviour.
760 *Nature* **583**: 415–420.

761 Prabh N, Roeseler W, Witte H, Eberhardt G, Sommer RJ, Rödelsperger C. 2018.
762 Deep taxon sampling reveals the evolutionary dynamics of novel gene families
763 in Pristionchus nematodes. *Genome Res* **28**: 1664–1674.

764 Prjibelski A, Antipov D, Meleshko D, Lapidus A, Korobeynikov A. 2020. Using
765 SPAdes De Novo Assembler. *Curr Protoc Bioinformatics* **70**: e102.

766 R Core Team. 2024. R: a language and environment for statistical computing. R
767 Foundation for Statistical Computing, Vienna https://www.R-project.org/.

768 Rillo-Bohn R, Adilardi R, Mitros T, Avşaroğlu B, Stevens L, Köhler S, Bayes J, Wang
769 C, Lin S, Baskevitch KA, et al. 2021. Analysis of meiosis in Pristionchus
770 pacificus reveals plasticity in homolog pairing and synapsis in the nematode
771 lineage. *Elife* **10**. http://dx.doi.org/10.7554/eLife.70990.

772 Rödelsperger C, Ebbing A, Sharma DR, Okumura M, Sommer RJ, Korswagen HC.

2021. Spatial Transcriptomics of Nematodes Identifies Sperm Cells as a Source of Genomic Novelty and Rapid Evolution. *Mol Biol Evol* **38**: 229–243.

Rödelsperger C, Röseler W, Prabh N, Yoshida K, Weiler C, Herrmann M, Sommer RJ. 2018. Phylotranscriptomics of Pristionchus Nematodes Reveals Parallel Gene Loss in Six Hermaphroditic Lineages. *Curr Biol* **28**: 3123–3127.e5.

Rudel D, Riebesell M, Sommer RJ. 2005. Gonadogenesis in Pristionchus pacificus and organ evolution: development, adult morphology and cell-cell interactions in the hermaphrodite gonad. *Dev Biol* **277**: 200–221.

Salazar-Jaramillo L, de la Cuesta-Zuluaga J, Chica LA, Cadavid M, Ley RE, Reyes A, Escobar JS. 2024. Gut microbiome diversity within is negatively associated with human obesity. *mSystems* **9**: e0062724.

Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* **30**: 2068–2069.

Shalev O, Karasov TL, Lundberg DS, Ashkenazy H, Pramoj Na Ayutthaya P, Weigel D. 2022. Commensal Pseudomonas strains facilitate protective response against pathogens in the host plant. *Nat Ecol Evol* **6**: 383–396.

Sieriebriennikov B, Sun S, Lightfoot JW, Witte H, Moreno E, Rödelsperger C, Sommer RJ. 2020. Conserved nuclear hormone receptors controlling a novel plastic trait target fast-evolving genes expressed in a single cell. *PLoS Genet* **16**: e1008687.

Slowinski S, Ramirez I, Narayan V, Somayaji M, Para M, Pi S, Jadeja N, Karimzadegan S, Pees B, Shapira M. 2020. Interactions with a Complex Microbiota Mediate a Trade-Off between the Host Development Rate and Heat Stress Resistance. *Microorganisms* **8**.

Smits THM, Arend LNVS, Cardew S, Tång-Hallbäck E, Mira MT, Moore ERB, Sampaio JLM, Rezzonico F, Pillonetto M. 2022. Resolving taxonomic confusion: establishing the genus Phytobacter on the list of clinically relevant Enterobacteriaceae. *Eur J Clin Microbiol Infect Dis* **41**: 547–558.

Sun S, Rödelsperger C, Sommer RJ. 2021. Single worm transcriptomics identifies a developmental core network of oscillating genes with deep conservation across nematodes. *Genome Res* **31**: 1590–1601.

Theska T, Sommer RJ. 2024. Feeding-structure morphogenesis in “rhabditid” and diplogastrid nematodes is not controlled by a conserved genetic module. *Evol Dev* **26**: e12471.

van Dongen S, Abreu-Goodger C. 2012. Using MCL to extract clusters from networks. *Methods Mol Biol* **804**: 281–295.

VanKuren NW, Chen J, Long M. 2024. Sexual conflict drive in the rapid evolution of new gametogenesis genes. *Semin Cell Dev Biol* **159-160**: 27–37.

Venzon M, Das R, Luciano DJ, Burnett J, Park HS, Devlin JC, Kool ET, Belasco JG,

812 Hubbard EJA, Cadwell K. 2022. Microbial byproducts determine reproductive
813 fitness of free-living and parasitic nematodes. *Cell Host Microbe* **30**: 786–
814 797.e8.

815 Wan Q-L, Meng X, Fu X, Chen B, Yang J, Yang H, Zhou Q. 2019. Intermediate
816 metabolites of the pyrimidine metabolism pathway extend the lifespan of C.
817 elegans through regulating reproductive signals. *Aging (Albany NY)* **11**: 3993–
818 4010.

819 Watson E, MacNeil LT, Ritter AD, Yilmaz LS, Rosebrock AP, Caudy AA, Walhout
820 AJM. 2014. Interspecies systems biology uncovers metabolites affecting C.
821 elegans gene expression and life history traits. *Cell* **156**: 759–770.

822 Weimann A, Mooren K, Frank J, Pope PB, Bremges A, McHardy AC. 2016. From
823 Genomes to Phenotypes: Traitar, the Microbial Trait Analyzer. *mSystems* **1**.
824 http://dx.doi.org/10.1128/mSystems.00101-16.

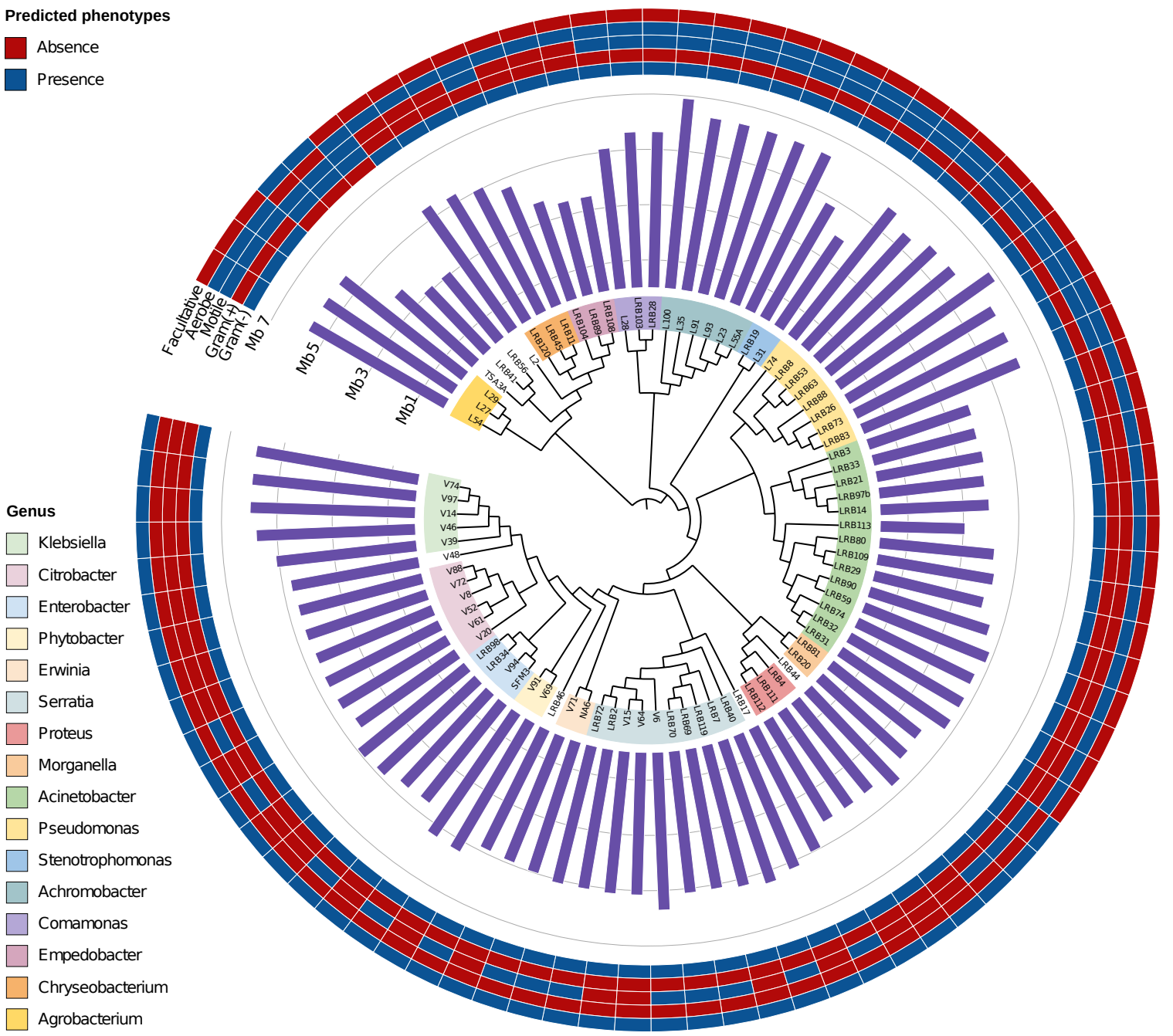
825 Zhang J, Holdorf AD, Walhout AJ. 2017. C. elegans and its bacterial diet as a model
826 for systems-level understanding of host-microbiota interactions. *Curr Opin*
827 *Biotechnol* **46**: 74–80.

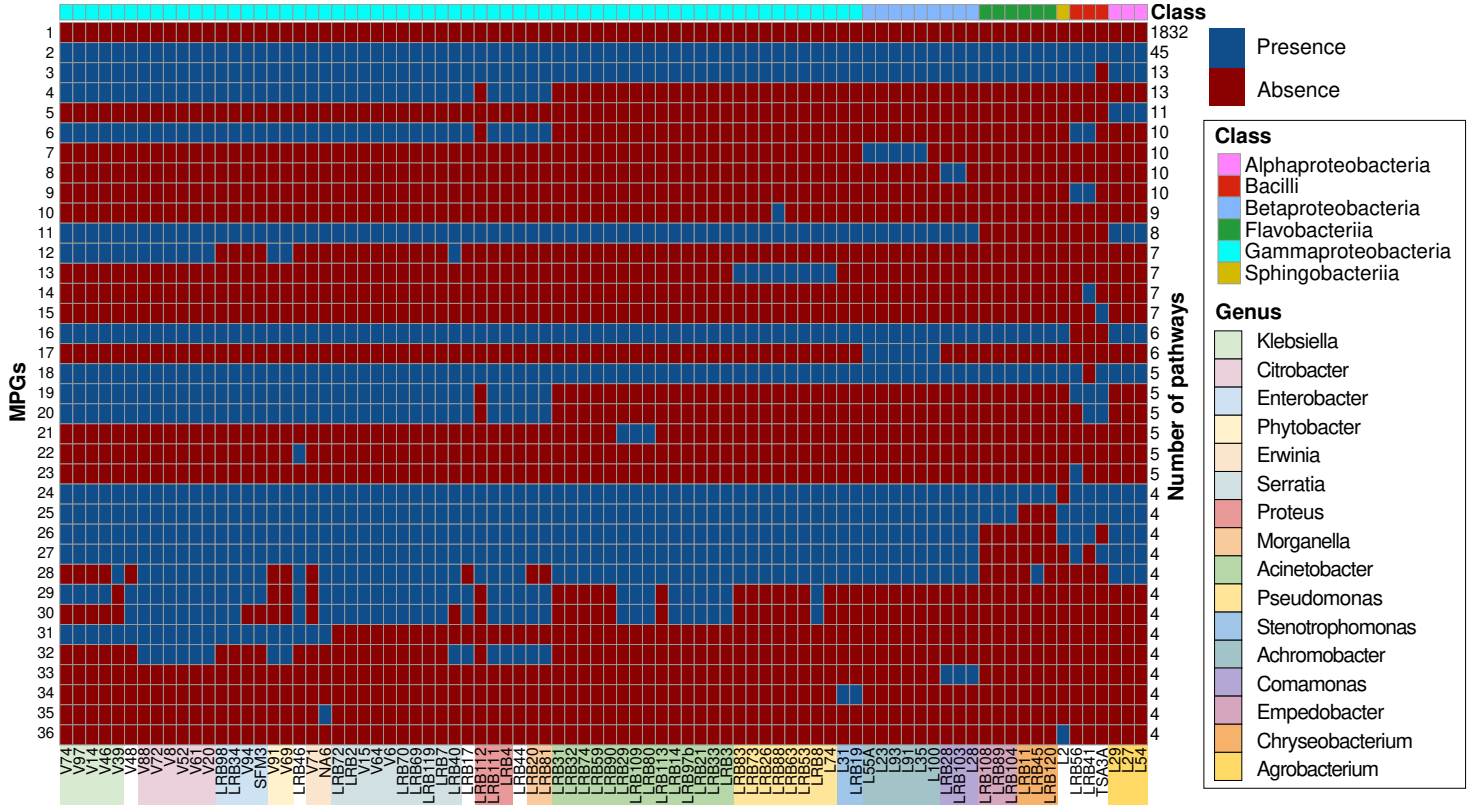
828 Zhang L, Yu J, Xie Y, Lin H, Huang Z, Xu L, Gelbic I, Guan X. 2014. Biological
829 activity of Bacillus thuringiensis (Bacillales: Bacillaceae) chitinase against
830 Caenorhabditis elegans (Rhabditida: Rhabditidae). *J Econ Entomol* **107**: 551–
831 558.

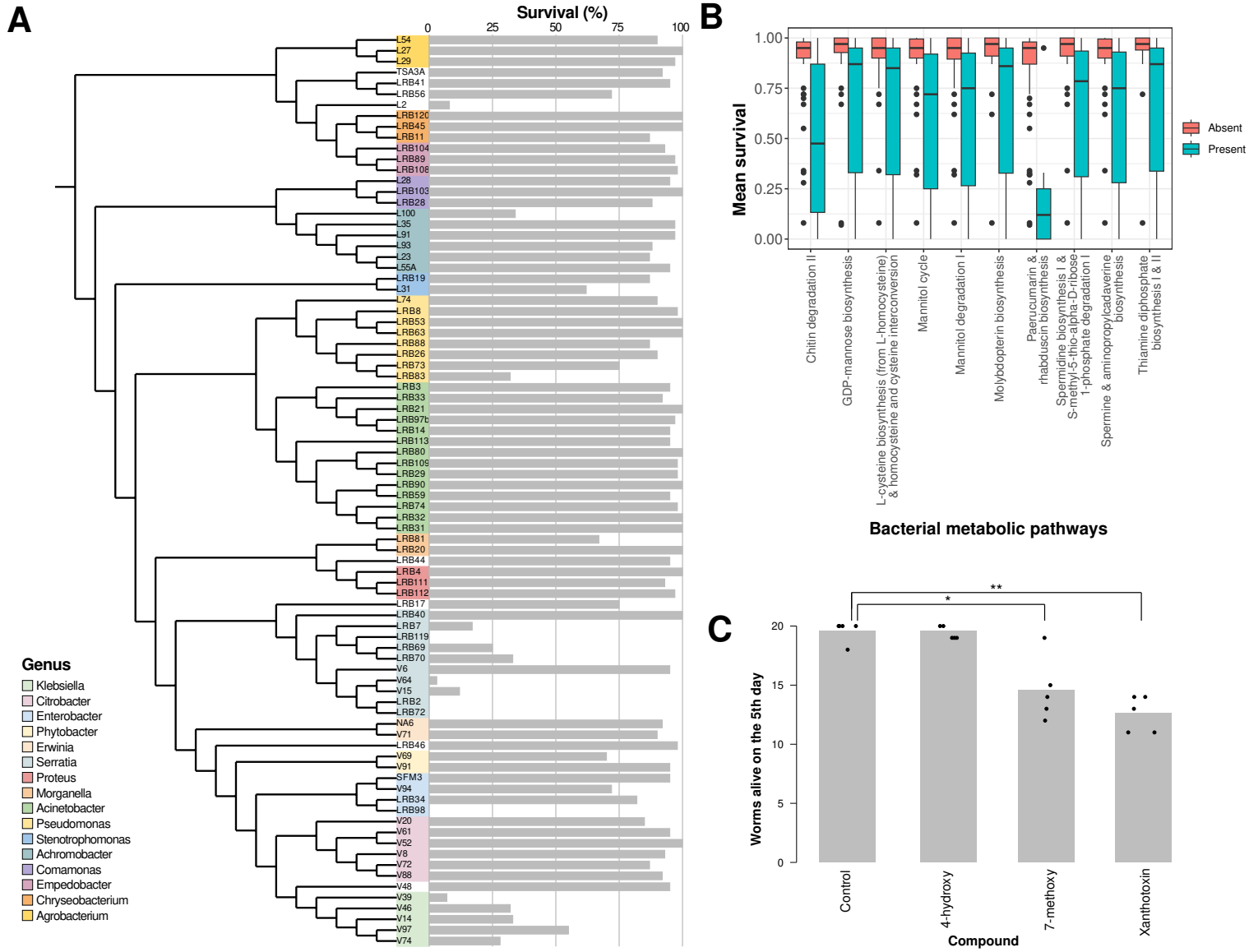
832 Zimmermann J, Kaleta C, Waschina S. 2021. gapseq: informed prediction of
833 bacterial metabolic pathways and reconstruction of accurate metabolic models.
834 *Genome Biol* **22**: 81.

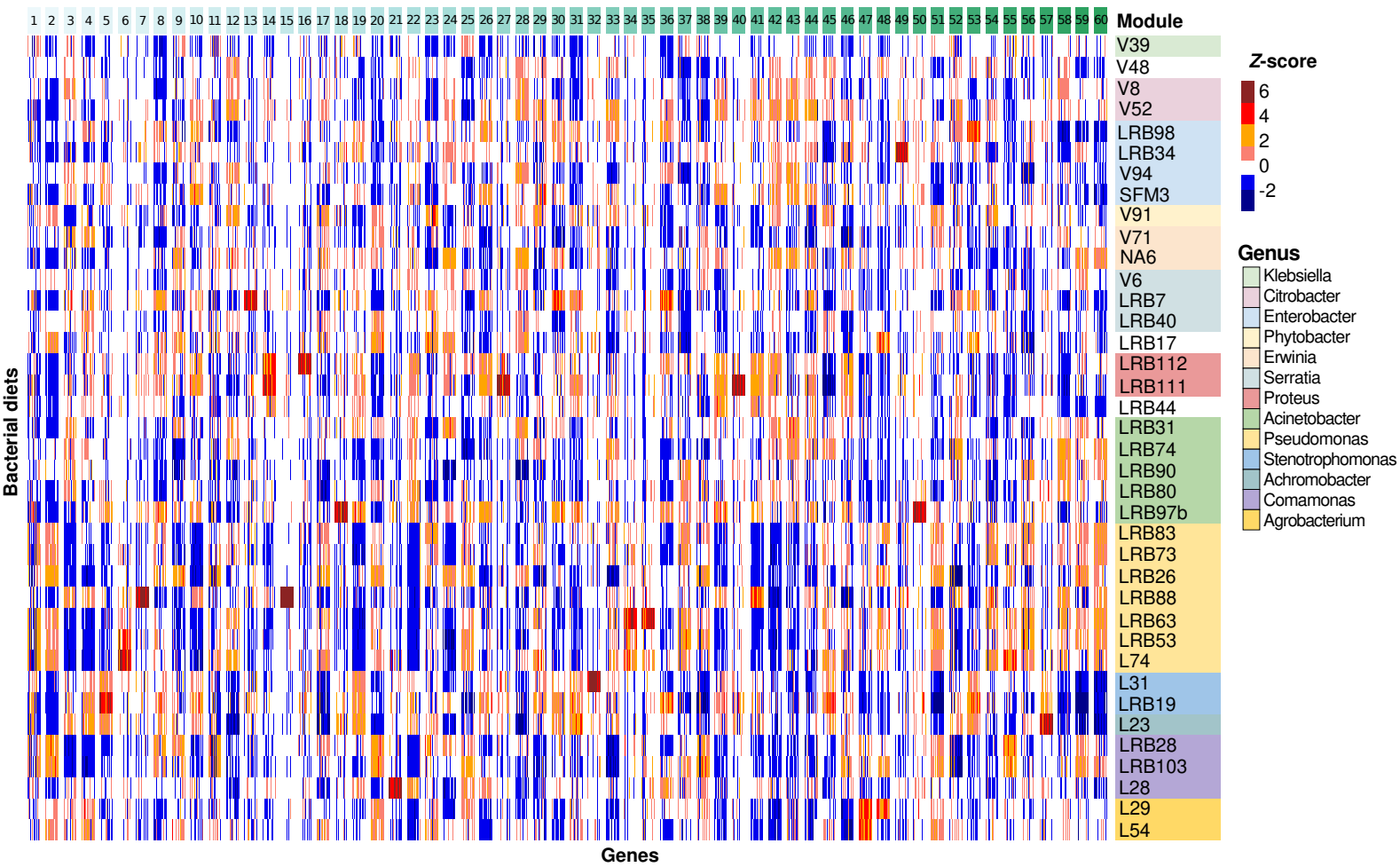
835 Zimmermann J, Obeng N, Yang W, Pees B, Petersen C, Waschina S, Kissoyan KA,
836 Aidley J, Hoeppner MP, Bunk B, et al. 2020. The functional repertoire contained
837 within the native microbiota of the model nematode Caenorhabditis elegans.
838 *ISME J* **14**: 26–38.

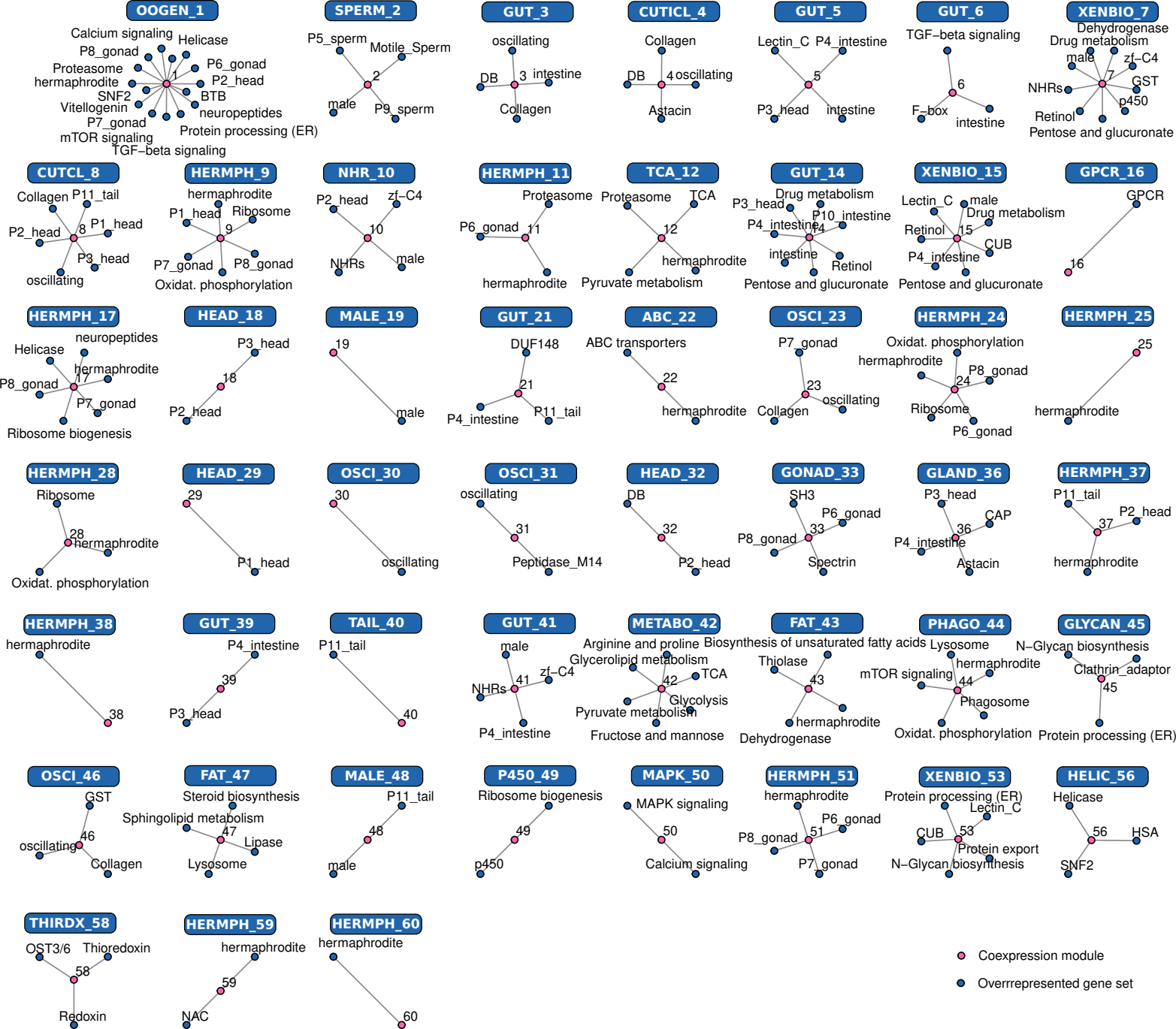
839

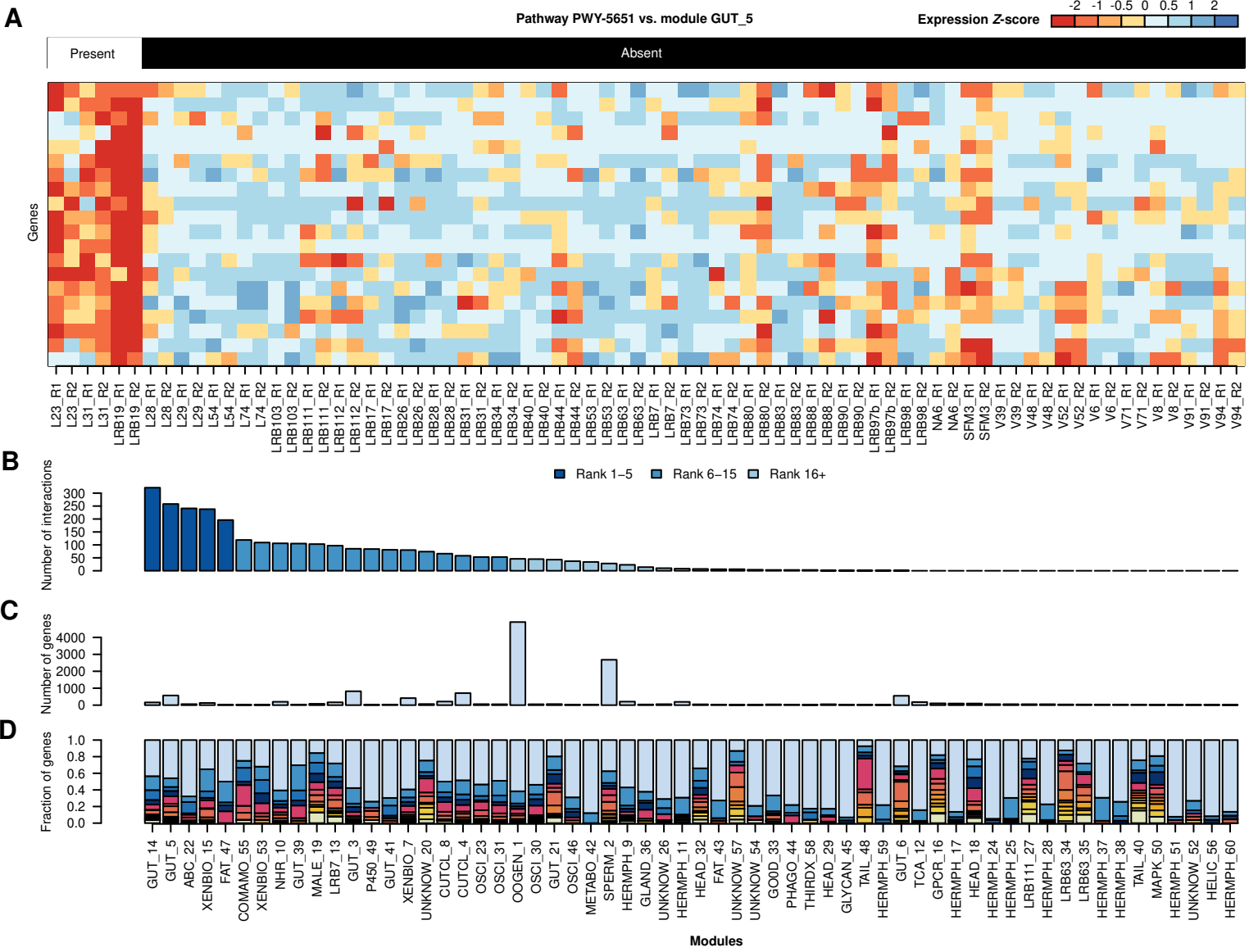

Predicted phenotypes
Absence
Presence
Facultative Aerobe
Motile
Gram(+)
Gram(-)
Mb 7
Mb5
Mb3
Mb1
Genus
Klebsiella
Citrobacter
Enterobacter
Phytobacter
Erwinia
Serratia
Proteus
Morganella
Acinetobacter
Pseudomonas
Stenotrophomonas
Achromobacter
Comamonas
Empedobacter
Chryseobacterium
Agrobacterium


Class
MPGs
Number of pathways
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
1832
45
13
13
11
10
10
10
10
9
8
7
7
7
6
6
5
5
5
5
5
4
4
4
4
4
4
4
4
4
4
4
4
4
V74
V97
V14
V46
V39
V48
V88
V72
V8
V52
V20
RB98
LRB34
V94
SFM3
V91
V69
LRB46
V71
NA6
LRB72
LRB2
V15
V64
V6
LRB70
LRB69
LRB119
LRB7
LRB40
LRB17
LRB112
LRB111
LRB4
LRB44
LRB42
LRB20
LRB31
LRB32
LRB74
LRB59
LRB90
LRB29
LRB109
LRB80
LRB113
LRB14
LRB97b
LRB21
LRB83
LRB3
LRB83
LRB73
LRB26
LRB88
LRB63
LRB53
LRB8
L74
L31
LRB19
L55A
L23
L91
L35
L100
LRB28
LRB103
L28
LRB108
LRB89
LRB104
LRB11
LRB45
LRB120
LRB56
LRB41
TSA3A
L29
L27
L54
Presence
Absence
Class
Alphaproteobacteria
Bacilli
Betaproteobacteria
Flavobacteriia
Gammaproteobacteria
Sphingobacteriia
Genus
Klebsiella
Citrobacter
Enterobacter
Phytobacter
Erwinia
Serratia
Proteus
Morganella
Acinetobacter
Pseudomonas
Stenotrophomonas
Achromobacter
Comamonas
Empedobacter
Chryseobacterium
Agrobacterium


A
Survival (%)
0
25
50
75
100
L54
L27
L29
TSA3A
LRB41
LRB56
L2
LRB120
LRB45
LRB11
LRB104
LRB89
LRB108
L28
LRB103
LRB28
L100
L35
L91
L93
L23
L55A
LRB19
L31
L74
LRB8
LRB53
LRB63
LRB88
LRB26
LRB73
LRB83
LRB3
LRB33
LRB21
LRB97b
LRB14
LRB113
LRB80
LRB109
LRB29
LRB90
LRB59
LRB74
LRB32
LRB31
LRB81
LRB20
LRB44
LRB4
LRB111
LRB112
LRB17
LRB40
LRB7
LRB119
LRB69
LRB70
V6
V64
V15
LRB2
LRB72
NA6
V71
LRB46
V69
V91
SFM3
V94
LRB34
LRB98
V20
V61
V52
V8
V72
V88
V48
V39
V46
V14
V97
V74
Genus
Klebsiella
Citrobacter
Enterobacter
Phytobacter
Erwinia
Serratia
Proteus
Morganella
Acinetobacter
Pseudomonas
Stenotrophomonas
Achromobacter
Comamonas
Empedobacter
Chryseobacterium
Agrobacterium
B
Mean survival
1.00
0.75
0.50
0.25
0.00
Absent
Present
Chitin degradation II
GDP-mannose biosynthesis
L-cysteine biosynthesis (from L-homocysteine) & homocysteine and cysteine interconversion
Mannitol cycle
Mannitol degradation I
Molybdopterin biosynthesis
Paerucumarin & rhabduscin biosynthesis
Spermidine biosynthesis I & S-methyl-5-thio-alpha-D-ribose 1-phosphate degradation I
Spermine & aminopropylcadaverine biosynthesis
Thiamine diphosphate biosynthesis I & II
Bacterial metabolic pathways
C
**
*
Worms alive on the 5th day
20
15
10
5
0
Control
4-hydroxy
7-methoxy
Xanthotoxin
Compound


1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60
Module
V39
V48
V8
V52
LRB98
LRB34
V94
SFM3
V91
V71
NA6
V6
LRB7
LRB40
LRB17
LRB112
LRB111
LRB44
LRB31
LRB74
LRB90
LRB80
LRB97b
LRB83
LRB73
LRB26
LRB88
LRB63
LRB53
L74
L31
LRB19
L23
LRB28
LRB103
L28
L29
L54
Z-score
6
4
2
0
-2
Genus
Klebsiella
Citrobacter
Enterobacter
Phytobacter
Erwinia
Serratia
Proteus
Acinetobacter
Pseudomonas
Stenotrophomonas
Achromobacter
Comamonas
Agrobacterium
Bacterial diets
Genes


OOGEN_1
Calcium signaling
Helicase
P8_gonad
Proteasome
P6_gonad
hermaphrodite
1
P2_head
SNF2
BTB
Vitellogenin
neuropeptides
P7_gonad
Protein processing (ER)
mTOR signaling
TGF−beta signaling
SPERM_2
P5_sperm
Motile_Sperm
2
male
P9_sperm
GUT_3
oscillating
DB
3
intestine
Collagen
CUTICL_4
Collagen
DB
4
oscillating
Astacin
GUT_5
Lectin_C
P4_intestine
5
P3_head
intestine
GUT_6
TGF−beta signaling
6
F−box
intestine
XENBIO_7
Dehydrogenase
Drug metabolism
male
zf−C4
7
GST
NHRs
p450
Retinol
Pentose and glucuronate
CUTCL_8
Collagen
P11_tail
P2_head
8
P1_head
P3_head
oscillating
HERMPH_9
hermaphrodite
P1_head
Ribosome
9
P7_gonad
P8_gonad
Oxidat. phosphorylation
NHR_10
P2_head
zf−C4
10
NHRs
male
HERMPH_11
Proteasome
P6_gonad
11
hermaphrodite
TCA_12
Proteasome
TCA
12
hermaphrodite
Pyruvate metabolism
GUT_14
Drug metabolism
P3_head
P10_intestine
P4_intestine
14
intestine
Retinol
Pentose and glucuronate
XENBIO_15
Lectin_C
male
Drug metabolism
Retinol
15
CUB
P4_intestine
Pentose and glucuronate
GPCR_16
GPCR
16
HERMPH_17
neuropeptides
Helicase
hermaphrodite
P8_gonad
17
P7_gonad
Ribosome biogenesis
HEAD_18
P3_head
18
P2_head
MALE_19
19
male
GUT_21
DUF148
21
P4_intestine
P11_tail
ABC_22
ABC transporters
22
hermaphrodite
OSCI_23
P7_gonad
23
oscillating
Collagen
HERMPH_24
Oxidat. phosphorylation
hermaphrodite
P8_gonad
24
Ribosome
P6_gonad
HERMPH_25
25
hermaphrodite
HERMPH_28
Ribosome
28
hermaphrodite
Oxidat. phosphorylation
HEAD_29
29
P1_head
OSCI_30
30
oscillating
OSCI_31
oscillating
31
Peptidase_M14
HEAD_32
DB
32
P2_head
GONAD_33
SH3
P6_gonad
P8_gonad
33
Spectrin
GLAND_36
P3_head
CAP
36
P4_intestine
Astacin
HERMPH_37
P11_tail
P2_head
37
hermaphrodite
HERMPH_38
hermaphrodite
38
GUT_39
P4_intestine
39
P3_head
TAIL_40
P11_tail
40
GUT_41
male
zf−C4
41
NHRs
P4_intestine
METABO_42
Arginine and proline
Glycerolipid metabolism
TCA
42
Glycolysis
Pyruvate metabolism
Fructose and mannose
FAT_43
Biosynthesis of unsaturated fatty acids
Thiolase
43
hermaphrodite
Dehydrogenase
PHAGO_44
Lysosome
hermaphrodite
mTOR signaling
44
Phagosome
Oxidat. phosphorylation
GLYCAN_45
N−Glycan biosynthesis
Clathrin_adaptor
45
Protein processing (ER)
OSCI_46
GST
oscillating
46
Collagen
FAT_47
Steroid biosynthesis
Sphingolipid metabolism
47
Lipase
Lysosome
MALE_48
P11_tail
48
male
P450_49
Ribosome biogenesis
49
p450
MAPK_50
MAPK signaling
50
Calcium signaling
HERMPH_51
hermaphrodite
P6_gonad
P8_gonad
51
P7_gonad
XENBIO_53
Protein processing (ER)
Lectin_C
CUB
53
Protein export
N−Glycan biosynthesis
HELIC_56
Helicase
56
HSA
SNF2
THIRDX_58
OST3/6
Thioredoxin
58
Redoxin
HERMPH_59
hermaphrodite
59
NAC
HERMPH_60
hermaphrodite
60
Coexpression module
Overrrepresented gene set


A
Pathway PWY-5651 vs. module GUT_5
Expression Z-score
-2 -1 -0.5 0 0.5 1 2
Present
Absent
Genes
L23_R1 L23_R2 L31_R1 L31_R2 LRB19_R1 LRB19_R2 L28_R1 L28_R2 L29_R1 L29_R2 L54_R1 L54_R2 L74_R1 L74_R2 LRB103_R1 LRB103_R2 LRB111_R1 LRB111_R2 LRB112_R1 LRB112_R2 LRB17_R1 LRB17_R2 LRB26_R1 LRB26_R2 LRB28_R1 LRB28_R2 LRB31_R1 LRB31_R2 LRB34_R1 LRB34_R2 LRB40_R1 LRB40_R2 LRB44_R1 LRB44_R2 LRB53_R1 LRB53_R2 LRB63_R1 LRB63_R2 LRB7_R1 LRB7_R2 LRB73_R1 LRB73_R2 LRB74_R1 LRB74_R2 LRB80_R1 LRB80_R2 LRB83_R1 LRB83_R2 LRB88_R1 LRB88_R2 LRB90_R1 LRB90_R2 LRB97b_R1 LRB97b_R2 LRB98_R1 LRB98_R2 NA6_R1 NA6_R2 SFM3_R1 SFM3_R2 V39_R1 V39_R2 V48_R1 V48_R2 V52_R1 V52_R2 V6_R1 V6_R2 V71_R1 V71_R2 V8_R1 V8_R2 V91_R1 V91_R2 V94_R1 V94_R2
B
Rank 1–5
Rank 6–15
Rank 16+
Number of interactions
300 250 200 150 100 50 0
C
Number of genes
4000 3000 2000 1000 0
D
Fraction of genes
1.0 0.8 0.6 0.4 0.2 0.0
GUT_14 GUT_5 ABC_22 XENBIO_15 FAT_47 COMAMO_55 XENBIO_53 NHR_10 GUT_39 MALE_19 LRB7_13 GUT_3 P450_49 GUT_41 XENBIO_7 UNKNOW_20 CUTCL_8 CUTCL_4 OSCI_23 OSCI_31 OOGEN_1 OSCI_30 GUT_21 OSCI_46 METABO_42 SPERM_2 HERMPH_9 GLAND_36 UNKNOW_26 HERMPH_11 HEAD_32 FAT_43 UNKNOW_57 UNKNOW_54 GOOD_33 PHAGO_44 THIRDX_58 HEAD_29 GLYCAN_45 TAIL_48 HERMPH_59 GUT_6 TCA_12 GPCR_16 HERMPH_17 HEAD_18 HERMPH_24 HERMPH_25 LRB111_27 HERMPH_28 LRB63_34 LRB63_35 HERMPH_37 HERMPH_38 TAIL_40 MAPK_50 HERMPH_51 UNKNOW_52 HELIC_56 HERMPH_60
Modules

Interspecies systems biology links bacterial metabolic pathways to nematode gene expression, chemotaxis behavior, and survival

Marina Athanasouli, Tobias Loschko and Christian Roedelsperger

Genome Res. published online August 5, 2025
Access the most recent version at doi:10.1101/gr.280848.125