CeMbio - The Caenorhabditis elegans Microbiome Resource

To cite this version:

HAL Id: hal-03046949
https://hal.science/hal-03046949
Submitted on 8 Dec 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
CeMbio - The Caenorhabditis elegans Microbiome Resource

Philipp Dirksen,*1 Adrien Assié,†1 Johannes Zimmermann,† Fan Zhang,† Adina-Malin Tietje,* Sarah Arnaud Marsh,§ Marie-Anne Félix,§ Michael Shapira,** Christoph Kaleta,‡ Hinrich Schulenburg,*1,2 and Buck S. Samuel†1,2

*Department of Evolutionary Ecology and Genetics, Christian-Albrechts University, Kiel, Germany, †Alkek Center for Metagenomics and Microbiome Research, Baylor College of Medicine, Houston TX, ‡Medical Systems Biology, University Medical Center Schleswig-Holstein, Kiel, Germany, §Institute of Biology of the Ecole Normale Supérieure, Paris, France, and **Department of Integrative Biology, University of California, Berkeley CA

ORCID IDs: 0000-0002-0336-8711 (P.D.); 0000-0003-3568-0766 (A.A.); 0000-0002-5041-1954 (J.Z.); 0000-0003-2849-1713 (F.Z.); 0000-0003-4194-043X (S.A.M.); 0000-0002-2635-3141 (M.-A.F.); 0000-0003-1332-2220 (M.S.); 0000-0001-8004-9514 (C.K.); 0000-0002-1413-913X (H.S.); 0000-0002-4347-3997 (B.S.S.)

ABSTRACT The study of microbiomes by sequencing has revealed a plethora of correlations between microbial community composition and various life-history characteristics of the corresponding host species. However, inferring causation from correlation is often hampered by the sheer compositional complexity of microbiomes, even in simple organisms. Synthetic communities offer an effective approach to infer cause-effect relationships in host-microbiome systems. Yet the available communities suffer from several drawbacks, such as artificial (thus non-natural) choice of microbes, microbe-host mismatch (e.g., human microbes in gnotobiotic mice), or hosts lacking genetic tractability. Here we introduce CeMbio, a simplified natural Caenorhabditis elegans microbiota derived from our previous meta-analysis of the natural microbiome of this nematode. The CeMbio resource is amenable to all strengths of the C. elegans model system, strains included are readily culturable, they all colonize the worm gut individually, and comprise a robust community that distinctly affects nematode life-history. Several tools have additionally been developed for the CeMbio strains, including diagnostic PCR primers, completely sequenced genomes, and metabolic network models. With CeMbio, we provide a versatile resource and toolbox for the in-depth dissection of naturally relevant host-microbiome interactions in C. elegans.

KEYWORDS C. elegans Microbiome resource Host-microbe interactions Synthetic communities Metabolic networks

While there is little debate that microbiomes exert broad influence on their hosts (McFall-Ngai 2014; Gilbert et al. 2018), less is known about the mediators of this influence. Often the complexity of the systems renders interrogation impossible. Model hosts address variation by controlling much of the environmental, genetic and dietary drivers of host-microbiome interactions (Fraune and Bosch 2010; Douglas 2019), but often overlook the importance or extent of genetic and functional variation on the part of the microbiome. The greatest advances in understanding have emerged largely from binary tests of one host and one microbe under gnotobiotic conditions (Fischbach 2018). While certainly valuable, these types of experiments likely also oversimplify the system in a manner that limits ability to identify properties that emerge from collaborations and competitions between microbiome members and their natural host. Thus, there is a need to develop well-characterized, tractable systems that faithfully capture the complexity of these interactions and identity of the molecular drivers of microbiome impact.

To this end, C. elegans has emerged as a powerful high-throughput system for studying host-microbiome interactions (Zhang et al. 2017).
This free-living nematode has many inherent strengths including a short life cycle of 3 days and lifespans of 3 weeks, a well-defined and transparent body plan, widely available resources and facile methods for forward and reverse genetics, plus a wealth of understanding of its biology and physiology (Girard et al. 2007; Frézal and Félix 2015). In the wild, *C. elegans* harbors a characteristic gut microbiome community that is recruited from its surrounding environment (Dirksen et al. 2016; Samuel et al. 2016; Berg et al. 2016a). Meta-analyses of these natural microbiomes highlight core membership of over a dozen bacterial families, including Gammaproteobacteria (*Enterobacteriaceae, Pseudomonadaceae*, and *Xanthomonodaceae*) and Bacteroidetes (*Sphingobacteriaceae, Weekellaceae, Flavobacteriaceae*) (Zhang et al. 2017).

Here we establish a publicly available and well-defined model microbiome for use in *C. elegans* (CeMbio). This set is composed of 12 bacteria from 9 different families that represent the core microbiome of *C. elegans* based on analyses and empirical studies of intestinal colonization. These bacterial strains are presented with fully sequenced and annotated genomes, metabolic network reconstructions, and robust protocols for their use in *C. elegans* studies and beyond. All of the bacteria effectively colonize the *C. elegans* gut both alone and as a community, which can impact the growth and development of the host. The CeMbio community has broad application to any aspect of *C. elegans* biology from aging to pathogenesis, development to neurobiology, and any aspect of physiology where a more natural environment is desired. Ultimately, pairing of this well-defined microbiome and highly-traceable host is envisioned to complement other systems (e.g., Fraune and Bosch 2010; Brugiroux et al. 2016; Douglas 2019) in advancing understanding of the mechanisms of microbiome impact on host health and disease.

**METHODS**

**Bacterial collections of natural Caenorhabditis populations**

The CeMbio strains were chosen as described in the next section from a set of previously cultured bacteria from the Félix, Samuel, Schulenburg, and Shapira labs (Montalvo-Katz et al. 2013; Berg et al. 2016b; Dirksen et al. 2016; Samuel et al. 2016; Zimmermann et al. 2020b), plus an additional collection of 139 bacterial strains also isolated from wild *Caenorhabditis* animals in the Félix lab (JUb130-274; Table S1)

For the new Félix lab collection, *Caenorhabditis* animals were collected from rotting fruit and stems from in and around Paris as well as Brittany and Indre (France). Substrate samples were brought back to the laboratory to isolate worms using adapted methods as in Barrière and Félix (2006). Briefly, while working aseptically, samples were plated onto sterile petri plates containing Normal Growth Medium (NGM: Autoclave 3g NaCl, 2.5g Bacto-Peptone, 17g Bacter Agar, 1 L sterile water; after cooled to 55° add 1 ml of 5 mg/ml Cholesterol, 1 ml 1 M CaCl₂, 1 ml 1 M MgSO₄, 25 ml 1 M pH 6 KPO₄ and diacetyl, a chemical attractant (10 ml of 1:30 dilution onto the agar at the opposite of the 90 mm plate)). Nematodes were identified to the genus level immediately by morphology, and to the species level through subsequent crosses and molecular verification as needed (most common around Paris: *C. elegans, C. briggsae, or C. remanei*).

Animals were then surface sterilized following a method similar to that described in Portal-Celhaye and Blaser (2012). Worms were washed off plates using sterile M9, then spun down for 2 min at 3,000 rpm. We removed excess liquid, then transferred worms to 55mm plates with 100 mM Gentamicin in NGM agar. After an hour on these plates, worms were washed off the plate with M9 and spun down for 2 min at 3,000 rpm. Excess liquid was pipetted off and the wash was repeated.

Gut-associated microbes were isolated from the surface sterilized samples above (JUb130-JUb265) or previously frozen nematode strains (JUb266-JUb274) using standard microbiological isolation techniques. We placed three adult worms from each sample into an Eppendorf tube with 500μl sterile water, then deadbeat (Mini-beadbeater, BioSpec Products, Bartlesville, OK, USA) them at maximum speed for two minutes. 100 μl lysed material from each sample was then plated onto each of four types of agar media in 90 mm plates. The liquid was spread and the plates were allowed to dry completely before wrapping them with parafilm. The four types of bacterial culture media included: NGM, Yeast Malt Extract Agar (YMEA), lysogeny broth supplemented with mannitol (LB+M: 5 g NaCl, 10 g Tryptone, 5 g Yeast Extract, 15 g Bacto-agar, 10 g Mannitol, 975 ml sterile water, 50 μl 10N NaOH), and chitin agars (Autoclave: 20 g Agar, 4 g Chitin, 0.75 g K₂HPO₄, 0.5 g MgSO₄ x 7H₂O, 0.35 g KH₂PO₄, 0.01 g FeSO₄ x 7H₂O, 0.001 g MnCl₂ x 4H₂O, 0.001 g ZnSO₄ x 7H₂O and 1 L sterile water). All media were supplemented with antifungals (20 ml/l nystatin and 0.05 g/l cycloheximide) after autoclaving and cooling to 55°. Bacteria were cultured at room-temperature (23°). Colonies were picked 1 to 3 days after bead beating and again 1 to 5 weeks after in an effort to isolate both slow and fast growing bacterial strains. Single colonies were picked again a few days later. Once they were in pure culture, bacterial strains were preserved in 10% glycerol solution and frozen at -80° for long-term storage. Strains were identified by sequencing the 16S rRNA sequence using the following primers: 27f-1492r (Lane 1991), 530f-1391r (Walker and Pace 2007), S-C

**Characterization and maintenance of CeMbio bacteria**

Twelve bacterial isolates were selected as part of the CeMbio resource (Table 1). Maximum-likelihood phylogentic analysis was used to inform candidate selection for the CeMbio resource by comparing a total of 510 sequences from *C. elegans*-related bacterial isolates from the Félix, Samuel, Schulenburg, and Shapira labs to the 12 most common OTUs, which we inferred by repeating our previous meta-analysis (Zhang et al. 2017) with only the natural worm samples (Figure 1A; Table S2). Taxonomic identity of the CeMbio strains was inferred via comparisons of genome-derived 16S rRNA sequences with same-family bacterial type strain sequences from the SILVA ribosomal RNA database project (Quast et al. 2013) as of 2018-11-20 using maximum-likelihood phylogenetic analysis (File S1). In both cases, we constructed multiple sequence alignments of the 16S sequences using the R package DECIPHER (Wright 2016) and the aligner SINA (Pruesse et al. 2012). Phylogenetic tree reconstruction was performed with IQ-TREE (Nguyen et al. 2015) with its implementation of ModelFinder (Kalyaanamoorthy et al. 2017) for maximum-likelihood model selection and a total of 10000 ultra fast bootstrap (Hoang et al. 2018) replicates. The resulting trees were visualized with the R packages ape (Paradis and Schliep 2019), ggplot2 (Wickham 2009), and ggtree (Yu et al. 2017). An isolate was assigned to a particular taxon if it clustered in a clade containing only strains of a single species or genus with at least 75% bootstrap support (10,000 replicates). Using this approach, we could assign all strains to a genus and seven of them to a species (Table 1). Using a
Table 1: Members of the CeMbio v1.0 collection

<table>
<thead>
<tr>
<th>Strain Name</th>
<th>Strain Source (Lab of Origin)</th>
<th>OTU Rank</th>
<th>OTU %</th>
<th>OTU Taxonomy</th>
</tr>
</thead>
<tbody>
<tr>
<td>CEent1</td>
<td>N2 from mesocosm (Shapira)</td>
<td>OTU_01</td>
<td>12.7%</td>
<td>Enterobacteriaceae</td>
</tr>
<tr>
<td>Lelliottia amnigena</td>
<td>JUb66</td>
<td>OTU_01</td>
<td>9.4%</td>
<td>Enterobacteriaceae</td>
</tr>
<tr>
<td>MYb10</td>
<td>wild C. elegans (Félix)</td>
<td>OTU_03</td>
<td>9.4%</td>
<td>Sphingomonadaceae</td>
</tr>
<tr>
<td>JUb134</td>
<td>wild C. elegans (Félix)</td>
<td>OTU_03</td>
<td>9.4%</td>
<td>Sphingomonadaceae</td>
</tr>
<tr>
<td>Stenotrophomonas indicatrix</td>
<td>N2 from mesocosm (Shapira)</td>
<td>OTU_04 / OTU_09</td>
<td>10.6%</td>
<td>Pseudomonadaceae</td>
</tr>
<tr>
<td>MYb11</td>
<td>C. elegans from compost (Schulenburg)</td>
<td>OTU_05</td>
<td>5.4%</td>
<td>Pseudomonadaceae</td>
</tr>
<tr>
<td>BIGb0172</td>
<td>wild C. elegans from rotting apple (Samuel)</td>
<td>OTU_07</td>
<td>3.4%</td>
<td>Comamonadaceae</td>
</tr>
<tr>
<td>BIGb0393</td>
<td>Pantoea nematovictus</td>
<td>OTU_08</td>
<td>3.1%</td>
<td>Enterobacteriaceae</td>
</tr>
<tr>
<td>MYb71</td>
<td>wild C. elegans from compost (Schulenburg)</td>
<td>OTU_10</td>
<td>2.1%</td>
<td>Sphingobacteriaceae</td>
</tr>
</tbody>
</table>

The phylogenetic identity of each isolate was assigned first using 16S rRNA maximum likelihood analysis; all isolates could be assigned to a genus. Then further phylogenomic analysis was performed on the full length 16S rRNA genes to better reflect the functional differences in this family of bacteria. The strains are ordered according to their OTU rank.

CeMbio colonization experiments with C. elegans

We performed three independent experiments to assess the ability and dynamics of the CeMbio strains to colonize the nematode gut. The first experiment characterized colonization by each individual CeMbio strain separately, while the second and third experiments focused on colonization by the CeMbio community. The methods for these experiments are generally similar, yet deviated in particular aspects of the protocols, thereby allowing us to assess robustness of the results. In experiment 1, colonization was assessed by counting colony forming units (CFU) of bacteria isolated from nematodes. In experiments 2 and 3, colonization was characterized through CFU counts for the entire community and separately an analysis of the relative abundance of strains, inferred from 16S-based microbiome sequencing. The experiments were performed with the canonical C. elegans strains N2 (all experiments) and CB4856 (only experiment 2). Nematodes were maintained on nematode growth medium (NGM) seeded with a lawn of Escherichia coli OP50, as previously described (Stiernagle 2006). Below, we describe the methods used for each experiment.

**Experiment 1:** Experiment 1 assessed colonization levels by each CeMbio strain separately and as a community in C. elegans (N2) gut. The experiment generally followed the previously published protocol from the Samuel lab (Zhang et al. 2020) and is available online on protocols.io (DOI: dx.doi.org/10.17504/protocols.io.rtzd6p6). Briefly, each CeMbio strain was grown individually in LB overnight at 25°C. Cultures were harvested by centrifugation, adjusted to a final optical density (OD, 600 nm) of 1 in PBS. Around 50 synchronized C. elegans stage 1 larvae (L1) were raised at 20°C on NGM in 6-well plates, each well inoculated with 60 μl bacteria. Nematodes and bacterial lawns were harvested after 72 h and 120 h with 600 μL of M9-T (M9 + 0.025% Triton X-100) and transferred to a sterile 96-well deep plate. Worms were allowed to settle without centrifugation to separate adults from progeny, then supernatant was removed using an aspiration manifold (V & P scientific, INC.) and worms were resuspended in 1 ml M9-T.
adherent bacteria, worms were washed four more times with 1 ml M9-T. After each washing step, worms were pelleted by centrifugation and aspiration of the supernatant using an aspiration manifold. After the final wash, worms were left in 100 μl M9-T for 10 min in order to enhance digestion or defecation of any transient gut bacteria. 100 μl 10 mM levamisole solution was added to paralyze worms, followed by surface sterilization using 200 μl 4% bleach solution in M9 for 2 min. Thereafter, nematodes were washed twice with PBS to remove excess levamisole and bleach. 300 μl of the worms in PBS were combined with 1 mm sterilized garnet beads, followed by lysis in a Mixer Mill at 25 Hz for 5 min. Lysates were used directly for inference of the number of colony forming units (CFUs). CFU numbers were calculated by adapting a previously published protocol (Hazan et al. 2012). In short, a reference curve for microbiome abundance is generated that relates standardized CFU counts on plates to OD measurements of a corresponding culture in liquid. For both approaches, a dilution series was established for each CeMbio strain and then measured in parallel for the two methods in four replicates. The resulting reference curve was subsequently used to calculate CFU counts from OD measurements for the experimental samples.

**Experiment 2:** Experiment 2 served to assess colonization levels and composition of two *C. elegans* strains, N2 and CB4856, by the CeMbio community. It was based on the same protocols used for experiment 1 and included the following modifications. The CeMbio community inoculum was established by mixing equal volumes of the different bacterial strains, grown and processed as above. Worms were harvested after only 120 h. The obtained lysates were split in two and then either used directly for CFU inference (as above, File S2) or pelleted by centrifugation and frozen at -20°C for later microbiome analysis.

For microbiome analysis, DNA was extracted from frozen lysate pellets. The pellets were resuspended in 200 μl sterile PBS, 0.1 mm sterile zirconia/silica beads were added, and bacterial cells were further lysed in a Mixer Mill for 5 min (25 Hz). 190 μl of the lysis was combined with 10 μl of 20 mg/ml proteinase K in a PCR plate and incubated in a MasterCycler ProS (Eppendorf) for 60 min at 60°C for digestion, followed by 15 min at 95°C to deactivate the proteinase. Barcoded amplicon sequencing was performed according to the Earth Microbiome project (Caporaso et al. 2012; Ul-Hasan et al. 2019) using the V4 region and sequenced by the Center for Metagenomics and Microbiome Research, Houston, Texas, USA.

**Experiment 3:** Experiment 3 used a slightly different approach to similarly study colonization of the *C. elegans* N2 strain by the CeMbio community in the Schuilenburg lab. The CeMbio strains were grown individually in LB medium overnight at 28°C. Cultures were harvested by centrifugation, washed three times with PBS, and adjusted to a final OD600 of 5. The cultures were mixed in equal volumes to produce the CeMbio community inoculum. Synchronized *C. elegans* L1 animals were raised at 20°C on 6 cm plates containing either NGM or peptone-free NGM (PFM) seeded with 250 μl of the CeMbio inoculum. Nematodes and bacterial lawns were harvested after 48 h, 72 h, and 96 h with M9-T. Surface-adherent bacteria were removed using a modification of a previously described method (Dirksen et al. 2016; Papkou et al. 2019) (Figure S3). Briefly, suspended worms were placed onto the top of pipette tips containing a 10 μm filter (SafeSeal-Tips Precision 1000 μl, cat #701081, Biozyme) and repeatedly incubated with washing solutions: 2x 3 min of M9-T with 25 mM tetramizole hydrochloride (to anesthetize worms and prevent subsequent bleach intake); 1x 4 min of M9-T with 2% bleach (equal volumes of 12% sodium hypochlorite and 5 N NaOH, (Stiernagle 2006)); 2x 3 min of M9-T to remove the bleach. After each washing step, the solution was removed by centrifugation of the tip box. The washed worms were pelleted by centrifugation and either frozen at -20°C for microbiome analysis or subjected to immediate CFU extraction.

For the latter, ten L4/adult nematodes were transferred to a 2 ml tube containing 100 μl M9-T and 10–20 1 mm zirconium beads, followed by sample homogenization using a Geno/Grinder 2000 (SPEX SamplePrep, Metuchen, USA) at 1500 strokes/min for 3 min. The homogenate was serially diluted in M9-T and each dilution was plated on LB-agar in triplicates. After 48 h of incubation at 25°C, the plates were imaged and appropriate dilutions counted.

For the microbiome analysis, DNA was isolated from frozen surface-sterilized worm samples or frozen lawn pellets, resuspended in buffer T1 from the NucleoSpin Tissue Kit (Macherey & Nagel), and processed with the additional steps described in the “Support protocol for bacteria” following the manufacturer’s instructions. Barcoded amplicon sequencing of the V3-V4 region of the bacterial 16S rRNA gene was carried out by the Institute for Clinical Molecular Biology, Kiel, Germany, using Illumina MiSeq technology.

**Microbiome data analysis**

Sequencing data were prepared for subsequent statistical analysis by first removing adapter and primer sequences with cutadapt (Martin 2011). OTUs were inferred using the R package dada2 (Callahan et al. 2016) with default parameters except for the following settings: sequence truncation length forward/reverse: 250/200 (longest expected amplicon for V3-V5: 428 nt, for V4-V5: 250nt); taxonomic assignment with silva training set release 132 (Quast et al. 2013); species assignment with a custom reference set of genome-derived 16S sequence variants of the CeMbio strains. The OTU read counts were normalized by the 16S gene copy numbers of the corresponding bacterial strains as predicted by the genome assemblies prior to analysis. The statistical analysis of the OTU data were performed in R using the following packages: DECIPHER (Wright 2016), phyloseq (McMurdie and Holmes 2013), DESeq2 (Love et al. 2014), vegan (Oksanen et al. 2019), ggplot2 (Wickham 2009) and Rcompanion (Mangiafico 2015).

**Developmental timing**

To assess the influence of CeMbio strains on *C. elegans* development, nematodes were raised on single and mixture lawns and the number of adults over time was counted, following a previously published protocol (Samuel et al. 2016). In brief, the 12 CeMbio bacteria were grown in LB at 25°C overnight; *E. coli* OP50 was also assayed for comparison, yet grown at 37°C. Bacteria were concentrated and seeded (50 μl) into 6-well NGM plates. Plates were dried and incubated overnight at 25°C before adding around 100 synchronized N2 or CB4856 L1 worms to the wells containing either a single bacterial strain or the CeMbio mixture. Adult animals were scored on an hourly basis from 46-60 h post L1.
stage. Each combination of *C. elegans* strains and bacteria were performed with two replicates.

**Physiological profiling using EcoPlates**  
The bacterial cells used for this experiment were from the same culture batch as were used for microbiome experiment 3. After the above mentioned step of adjusting the harvested cells to an OD$_{600}$ of 5, the individual bacterial cultures as well as the CeMbio mix culture were diluted to a final OD$_{600}$ of 0.2 and added to EcoPlates (Biolog, Inc.) in aliquots of 100 μl per well in triplicate for each condition. The inoculated plates were incubated at 28°C under constant orbital shaking. Physiological activity was measured 1, 3, and 5 days after

---

**Figure 1** The CeMbio strains. (A) The CeMbio strains (blue) were selected based on a comparison of 510 cultured *C. elegans* microbiome bacteria with the 12 most common OTUs inferred by repeating our previous meta-analysis ([Zhang et al. 2017](#)) with only the natural worm samples. The tree is based on a maximum-likelihood analysis using a TIM3e+R4 model and 10000 bootstraps. Nodes with bootstrap support >75% are denoted with a red dot. Some branches include several highly similar OTUs, as indicated (e.g., +6 more). (B) Fluorescence *in situ* hybridization of *C. elegans* N2 colonized with the CeMbio strains [red, general bacterial probe EUB338; blue, DAPI].
inoculation as described previously (Classen et al. 2003) in a microplate reader (Infinite 200Pro, Tecan). In brief, OD was measured at 590 nm and 750 nm. A bacterium was deemed physiologically active, if OD\textsubscript{590} - OD\textsubscript{750} > 0.1.

**CeMbio genome sequences and metabolic network reconstructions**

Bacterial genomes were sequenced using short (Illumina Nextera XT for CeEnt1, MvB10, MvB11, MvB71, and MSP1m; and all remaining isolates with Illumina MiSeq v3) and long read (PacBio SMRT; all isolates) sequencing. Short read Illumina reads were preprocessed with fastqillumina_filter_0.1 (--keep N -vv) and prinseq-lite 0.20.4 (-min_len 20 -ns_max_n 8 -min_qual_mean 15 -trim_qual_left 12 -trim_qual_right 12) (Schmieder and Edwards 2011), followed by barcode demultiplexing and filtering of the long reads with lima 1.8 (--peek-guess--split-bam-named) (https://github.com/PacificBiosciences/barcoding). The genomes were assembled by combining short and long reads in a hybrid approach, using the following programs: SPAdes v3.13.1 (Bankevich et al. 2012), Canu 1.8 (Koren et al. 2017), MaSuRCA 3.3.4 (Zimin et al. 2013), and the Unicycler pipeline 0.4.8 (Wick et al. 2017). Long read correction was achieved with LoRDEC 0.6 (Salmela and Rivals 2014), proovread 2.14.1 (Hackl et al. 2014), and Canu 1.8 (detailed script with all program calls and parameters is available in the supplement). The quality of genome assemblies was assessed with QUAST 5.0 (Mikheenko et al. 2018). QUAST provides an overview of assembly statistics such as estimated completeness using BUSCO (Seppey et al. 2019), predicted rRNA genes using RNAmmer (Lagesen et al. 2007), tRNA using aragon (Laslett and Canback 2004) and estimated contamination using bloattools (Laetsch and Blaxter 2017). The complete assembly statistics for each bacteria are provided in Table 2. The final genomes were derived after assessing the quality by coverage vs. length plots and by removing low quality contigs with <500bp and <5 coverage (Douglass et al. 2019). Genomes were then annotated using the PROKKA package (Seaman 2014).

The genome assemblies served as input for the reconstruction of metabolic networks, using gapseq 1.0 (Zimmermann et al. 2020a). The reconstruction of metabolic networks consists of systematic annotation of genomic metabolic genes that are then linked to the corresponding enzymatic reactions. The entirety of these reactions and metabolites were used to form the initial metabolic networks for each microbe that were then further improved during the curation process similar to published protocols (Thiele and Palsson 2010). In detail, pathways and transporters were predicted by gapseq find (-b 150), and the draft network was created by gapseq draft (-u 150 -l 50 -a 1). Network gaps were filled with gapseq fill (-b 50). Metabolic networks were thus represented by genome-scale metabolic models and combined with flux balance analysis (Orth et al. 2010), in order to predict growth rates under specified conditions. Gap filling was focused on ensuring bacterial growth in LB medium, which is known to support the growth of all CeMbio organisms in experiments. The metabolic networks were further improved with gapseq 1.0 by integrating experimental data derived from EcоПlаte assays (Biolog, Inc, USA), in which the reduction of a colorimetric tetrazolium dye indicates microbial metabolic activity on selected carbon sources (Bochner 2009), thereby providing empirical information on the metabolic competences of the CeMbio strains. The metabolic network models were subsequently used to predict carbon source utilization by the CeMbio strains, based on flux balance analysis with the recycling of electron carriers (quinones, NADH) as objective function. An organism was predicted to be able to use a certain compound if electron carriers could be recycled under conditions of a minimal medium including this compound as sole energy and carbon source. The inferred metabolic network models are available in the supplement (SBML format; File S3).

**CeMbio phylogenomic reconstructions**

Genome-scale phylogenies were calculated using GToTree V1.4.11 (Lee 2019). Each step of the pipeline was used with the default parameters. In brief, for each CeMbio strain, we downloaded NCBI RefSeq assemblies belonging either to the same genus or the same family depending on the number of published related genomes that were available. For genera with a large amount of genomes available, such as Enterobacteriaceae and Pseudomonas, we downloaded only genomes annotated as complete for representatives. For the less represented genera, we included partial assemblies. Genomes without annotation were scanned for CDS using prodigal (Hyatt et al. 2010), then genes were scanned for single-copy marker genes using HMMER3 (Eddy 2011), genomes with less than 10% of single marker gene redundancies were kept. Then single-copy marker genes were aligned using MUSCLE (Edgar 2004), trimmed with TrimAl (Capella-Gutiérrez et al. 2009) in order to keep sequence overlap and finally phylogenetic tree were calculated using FastTree 2 (Price et al. 2010). An Alphaproteobacteria, Bradyrhizobium diazoefficiens (GCF_000011365.1), was arbitrarily chosen as an outgroup for all trees. Taxonomy was edited on the tree using Taxonkit (Shen and Xiong 2019). For additional details on phylogenomic reconstructions, the phylogenomic tree as well as the code used for the analyses see File S4.

To evaluate the taxonomic affiliation of each CeMbio strain, we compared 16S rRNA phylogeny and phylogenomic reconstructions, then estimated the relatedness of each genome to their close phylogenomic relative using average nucleotide identity (ANI). ANI was calculated using a script available from the Enve-omics package (Rodriguez-R and Konstantinidis 2016). Bacteria with closely related genomes were compared and we used a ANI of 94–96% for the species cutoff, as described in previous studies (Konstantinidis and Tiedje 2005; Richter and Rosselló-Mora 2009). In the case where the 16S rRNA phylogenetic reconstruction provided a closely related named species that were not sequenced and no genomes were closely related to the CeMbio bacteria we relied on the 16S ribosomal phylogeny for strain naming purposes.

Lastly, each CeMbio bacteria was phylogenetically compared to the closest characterized strain in ATCC to determine likely biosafety levels (summarized in File S2 - Table S4.13); all are closely related to BLS1 categorized bacteria with the potential exception of BIGb0170, Sphingobacterium multivorum. However, no pathogenicity was observed for any of the bacteria against the C. elegans nematodes.

**Data availability**

To facilitate broad distribution of CeMbio strains, all isolates are available from the Caenorhabditis Genetics Center (CGC, http://cgc.umn.edu) search for ‘CeMbio’ in strain descriptions). Whole genome sequencing data for the bacteria are available from the European Nucleotide Archive (accession number PRJEB37895). Complete genome sequences for Acinetobacter guillouianiae MvB10, Pseudomonas larida MYb11 and Ochrobactrum vromis MYb71 were deposited previously (PRJNA400855), as well as the genome of Sphingomonas mollissima JUb134 (GCA_004341505.1), Chryseobacterium scophthalmum JUb44 (SAMN13190037) and Lelliottia amnigena JUb66 (GCF_003752235.1). The remaining genomes are
published under the Bioproject PRJNA624308. Microbiome 16S amplicon sequencing data are publicly available under accession numbers PRJEB37101 (experiment 2) and PRJEB37035 (experiment 3). All experimental data are provided in the supplement (see File S2) and have been uploaded to the GSA Figsheare Portal. Supplemental material available at figshare: https://doi.org/10.25387/g3.12580454.

RESULTS AND DISCUSSION

Overview of the CeMbio resource

We here established an ecologically informed model C. elegans microbiome (CeMbio) based on the following key criteria: (i) the chosen community should resemble the broad taxonomic diversity of the natural C. elegans microbiome as closely as possible; (ii) it should be ecologically meaningful and thus originate from natural C. elegans or at least its natural habitat; and (iii) it should include bacteria that are easy to grow and maintain on a standard medium, thus facilitating experiments in different fields of biological research.

As a first step, we repeated our previous microbiome analysis using only the natural C. elegans samples (Zhang et al. 2017), in order to identify the most abundant worm-associated bacterial taxa (Table S2). A total of 12 OTUs was consistently present in natural C. elegans, regardless of origin, isolating laboratory, or other covariates. This set of bacterial OTUs is likely to represent an ecologically relevant part of the natural microbiome (CeMbio) based on the following key criteria: (i) the CeMbio bacteria are able to colonize the C. elegans intestine. 

<table>
<thead>
<tr>
<th>Strain</th>
<th>Genus</th>
<th>Size²</th>
<th>GC²</th>
<th>Contigs²</th>
<th>rRNA³</th>
<th>tRNA³</th>
<th>CMP⁴</th>
<th>CDS⁵</th>
<th>UG⁶</th>
<th>CRISPR⁷</th>
<th>VG²</th>
<th>RG²</th>
</tr>
</thead>
<tbody>
<tr>
<td>CEn1</td>
<td>Enterobacter</td>
<td>4.8</td>
<td>55.3</td>
<td>1</td>
<td>8</td>
<td>89</td>
<td>98</td>
<td>4458</td>
<td>2787</td>
<td>0</td>
<td>38</td>
<td>4</td>
</tr>
<tr>
<td>JUb66</td>
<td>Lelliotti</td>
<td>4.6</td>
<td>52.9</td>
<td>1</td>
<td>7</td>
<td>84</td>
<td>98</td>
<td>4207</td>
<td>2684</td>
<td>3</td>
<td>33</td>
<td>3</td>
</tr>
<tr>
<td>B1G0393</td>
<td>Pantoea</td>
<td>5.2</td>
<td>54.6</td>
<td>2</td>
<td>7</td>
<td>82</td>
<td>98</td>
<td>4667</td>
<td>2540</td>
<td>0</td>
<td>21</td>
<td>1</td>
</tr>
<tr>
<td>Myb10</td>
<td>Acinetobacter</td>
<td>4.6</td>
<td>38.3</td>
<td>1</td>
<td>7</td>
<td>82</td>
<td>98</td>
<td>4244</td>
<td>1611</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>JUb19</td>
<td>Stenotrophomonas</td>
<td>4.6</td>
<td>66.3</td>
<td>2</td>
<td>4</td>
<td>78</td>
<td>97</td>
<td>4079</td>
<td>1625</td>
<td>0</td>
<td>13</td>
<td>5</td>
</tr>
<tr>
<td>Myb11</td>
<td>Pseudomonas</td>
<td>6.1</td>
<td>60.8</td>
<td>1</td>
<td>5</td>
<td>66</td>
<td>100</td>
<td>5456</td>
<td>2248</td>
<td>0</td>
<td>78</td>
<td>1</td>
</tr>
<tr>
<td>MSpm1</td>
<td>Pseudomonas</td>
<td>5.7</td>
<td>62.4</td>
<td>2</td>
<td>4</td>
<td>70</td>
<td>100</td>
<td>5159</td>
<td>2017</td>
<td>1</td>
<td>83</td>
<td>0</td>
</tr>
<tr>
<td>B1G0172</td>
<td>Comamonas</td>
<td>5.2</td>
<td>62.6</td>
<td>1</td>
<td>6</td>
<td>81</td>
<td>96</td>
<td>4595</td>
<td>1719</td>
<td>0</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Myb71</td>
<td>Ochrobactrum</td>
<td>5.4</td>
<td>55.9</td>
<td>3</td>
<td>4</td>
<td>60</td>
<td>97</td>
<td>5191</td>
<td>1814</td>
<td>0</td>
<td>16</td>
<td>2</td>
</tr>
<tr>
<td>JUb134</td>
<td>Sphingomonas</td>
<td>4.1</td>
<td>67.5</td>
<td>4</td>
<td>3</td>
<td>65</td>
<td>97</td>
<td>3816</td>
<td>1292</td>
<td>0</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>B1G0170</td>
<td>Sphingobacterium</td>
<td>6.4</td>
<td>39.9</td>
<td>1</td>
<td>7</td>
<td>85</td>
<td>93</td>
<td>5391</td>
<td>1362</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>JUb44</td>
<td>Chryseobacterium</td>
<td>4.7</td>
<td>33.6</td>
<td>1</td>
<td>7</td>
<td>80</td>
<td>95</td>
<td>4199</td>
<td>1173</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

1 Genome size in Mb.  
2 Percent GC content.  
3 Total no. of contigs of the assembly.  
4 Number of genomic rRNA as predicted by ARAGORN (Laslett and Canback 2004).  
5 Percent genome completeness as predicted by BUSCO (Seppey et al. 2019).  
6 Number of coding sequences as predicted by PROKKA (Seemann 2014).  
7 Number of unique genes per assembly (PROKKA).  
8 Number of genomic CRISPR genes (PROKKA).  
9 Number of virulence and resistance genes as predicted by Abricate (Seemann T, Github https://github.com/tseemann/abricate) using the Virulence Factor Database (Chen et al., 2016).  
10 Number of virulence and resistance genes as predicted by Abricate using the NCBI Bacterial Antimicrobial Resistance Reference Gene Database (Feldgarden et al., 2019), respectively.

Individual CeMbio strains effectively colonize the C. elegans intestine

In experiment 1, we determined whether and to what extent each of the CeMbio strains are able to colonize the C. elegans intestine. Based on previous studies with other non-pathogenic bacteria, we expected colonization of the C. elegans intestine during early adulthood (Dirksen et al. 2016; Vega and Gore 2017). To examine this directly with the CeMbio strains, synchronized L1 animals were exposed to each bacterial strain for 72 h and 120 h. E. coli OP50 colonized the C. elegans intestine after 120 h at a density of 18300 ± 7450 CFUs/worm for N2 nematodes and 7950 ± 3400 CFU/worm for CB4856 nematodes, which is similar to that of other studies (Portal-Celhay et al., 2017). By comparison, all of the wild strains were able to colonize the intestines of C. elegans, in overall agreement with previous studies of these strains (Montalvo-Katz et al. 2013; Berg et al. 2016b 2019; Dirksen et al. 2016; Zimmermann et al. 2020b). However, the extent and persistence of colonization over time varied among the strains. Eleven of the twelve strains exhibited
colonization levels that increased 10-fold over time (from an average of ~500 CFUs/worm at 72 h to ~10000 CFUs/worm at 120 h; Figure 2). Only the Sphingomonas mullosquorum JUb134 exhibited a slow colonization increase over time (threefold), which could be linked to its long doubling time. At the other end of the spectrum, Ochrobactrum vermis MYb71 showed a dramatic 38-fold increase in colonization over time (from an average 2). Only the three strains grouped into three groups: (1) the low colonizers, including all Enterobacteriaceae (Pantoea nematovitrix BIGb0393, Enterobacter hormaechei CEent1, Lelliottia ammigena JUb666) and Acinetobacter guillouiae MYb10; (2) the intermediate colonizers, Sphingomonas mullosquorum JUb134, Comamonas piscis BIGb0172, Sphingobacterium multivorum BIGb0170 and Pseudomonas lurida MYb11; and (3) the high colonizers, Ochrobactrum vermis MYb71, Chryseobacterium scophthalmum JUb44, Pseudomonas berkeleyensis MSPm1, and Stenotrophomonas indicatrix JUb19. We conclude that the selected CeMbio strains are all able to establish themselves in the gut of C. elegans, providing ample opportunity for direct interactions between microbe and worm.

All CeMbio strains colonize the worm gut when they are part of a community

In experiments 2 and 3, we next assessed whether the individual CeMbio strains colonize C. elegans while being part of a community. We performed two independent experiments that were performed in different labs and varied in culture media, sampling time points, and exact processing protocols (see Methods for more details). Overall, the experiments demonstrate that (i) a single C. elegans adult is consistently colonized by at least 1,000 and usually more than 10,000 bacteria (Figures 3C, 4F), (ii) all CeMbio strains can establish themselves in the worm gut as community members (Figures 3B, 4B, 4C), (iii) the exact colonization dynamics depend on bacterial strain, time, and culture medium (Figures 3B, 3E, 4B, 4C, 4E), and (iv) the C. elegans-associated community is clearly distinct in composition and diversity from the corresponding lawn community on the Agar plates (Figures 3B, 3D, 3E, 4B-4E).

The CeMbio strains were generally able to persist as a community both on plates and in nematodes. All 12 strains could be detected in the lawns of both NGM and PFM plates across the two independent experiments, even though some strains were present at very low levels (Figures 3B, 4B, 4C). Using improved protocols for nematode surface sterilization (Supplementary Figure 3; see Methods), we reliably detected all CeMbio strains as colonizers of C. elegans guts when part of a community despite some again appearing only at very low levels (Figures 3B, 4B, 4C).

Using DESeq2 for differential abundance analysis (File S2), we found that the C. elegans microbiome is generally very distinct from the corresponding microbial environment on the plates. This difference is most obvious in the PCoA analyses (Figure 3E, 4E), but also in the relative abundances (Figures 3B, 4B, 4C) and in experiment 3 in the observed diversities (Figure 4D). It is worth noting that the bacterial mixture used to inoculate the plates is best maintained on PFM (Figures 4A-4C, 4E), which does not support bacterial growth and thus appears to enhance experimental control of the source community. Additionally, we could determine that the colonization level significantly contributes to the ordination of the samples in experiment 3 (Effect of different time points and media, generalized additive model, df = 15, F = 2.523, P < 10^-5, R-square = 0.57), but not in experiment 2. In the latter case, there were differences in colonization between host strains, with CB4856 harboring more intestinal bacteria than N2, yet the differences were not significant. Though host strain also had a minor impact on overall microbiome composition, significant differences were observed in a subset of bacterial strains with significantly higher levels of Pseudomonas (MSPm1 and MYb11), Enterobacteria (JUb666 and CEent1), and Comamonas BIGb0172 in CB4856 microbiomes (Figures 3B and 4).

CeMbio bacteria followed distinct types of colonization patterns depending on the experiment, medium, and time points (Figure 4B, 4C, and S5). While the worm gut microbiome grew on NGM did not vary much over time, we observed substantial microbiome changes in nematodes grown on PFM. Bacteria such as Sphingobacterium BIGb0170, Chryseobacterium JUb44, Pseudomonas MSPm1, Pseudomonas MYb11, Enterobacter CEent1, Lelliottia JUb666, Comamonas BIGb0172 and Stenotrophomonas JUb19 all decreased in abundance over time when grown on both Pantoea BIGb0393 and Ochrobactrum MYb71 increased. This community shift may be explained by poor availability of nutrients on PFM and communities starving with time.

Compared to the lawn, in both experiments, two CeMbio strains were generally highly abundant in worms: Stenotrophomonas JUb19, more prevalent in NGM-raised worms (from 18 to 46% of the microbial population), with similar to lawn proportion except at the 96h time point where it was significantly more abundant in NGM worm populations (P-value = 0.02), and the Ochrobactrum MYb71, which was significantly more abundant in worms (from 59 to 97% of the microbial population, with P-values < 0.005 for each comparison to the lawn), especially on PFM. Conversely, some CeMbio bacteria were more abundant in the lawn, such as the Sphingobacterium BIGb0170 (from 28 to 62%, with P-values < 0.05 for each comparison to the lawn), Acinetobacter MYb10 (from 3 to 14%, with P-values < 0.05 except for the 48h time point), and Chryseobacterium JUb44 (from 3 to 39%, with P-values < 0.05 for each comparison to the lawn). The remaining CeMbio members represented a lower but stable fraction of the microbial community present in the worms, some present in lower abundance on NGM, such as Pseudomonas MSPm1, Comamonas BIGb0172 and Sphingomonas JUb134. Pantoea BIGb0393, Pseudomonas Lelliottia JUb666, and Enterobacter CEent1 were also present in lower abundance in the different communities. Pantoea BIGb0393 represented almost 10% of the worm community at 96h on PFM, while the other bacterial abundances decreased over time. A more detailed overview of comparisons is displayed in Figures S4 and S5 and a summary of the statistical tests performed is available in File S2 - Table S4.14.

Taken together, these results suggest that the CeMbio community contains a combination of strong general colonizers (Ochrobactrum MYb71 and Stenotrophomonas JUb19), poor colonizers (Sphingomonas JUb134), and many context-dependent colonizers (e.g., Pantoea BIGb0393 and Sphingobacterium BIGb0170). C. elegans can be colonized by all members of the CeMbio community with different profiles linked to different conditions, thus enabling new research on the effect of the microbiome on nematode biology.

CeMbio strains and community vary in impact on host growth rates

To illustrate the potential for the CeMbio strains to influence C. elegans biology, we characterized its phenotypic impact on nematode development rates. Two C. elegans strains (N2 and CB4856) were raised on single bacteria and the community from L1s and populations were followed over time for maturation to adulthood (Figure 5). In comparison to growth on E. coli OP50, worms developed faster on all single bacteria with three exceptions: on Ochrobactrum MYb71, they developed at a similar pace, while on Chryseobacterium JUb44 and Sphingobacterium BIGb0170, nematodes developed at a significantly slower pace. This is consistent with previous studies that indicate
slower growth of worms on Bacteroidetes strains like JUb44 and BIGb0170 (Samuel et al. 2016). Notably, the CeMbio community as a whole significantly enhanced growth rates for both host strains compared to growth on E. coli OP50. This suggests that the community may contain emergent properties that are produced as a result of interactions between the microbes that promote host development.

Our results also suggest a range of host-bacteria interaction types for the CeMbio members that may have both positive and negative impacts on the host. For example, JUb44 are efficient colonizers individually, yet slow down the developmental rate of its host, possibly indicating a negative effect of the bacteria on C. elegans. Though as part of the community this negative impact is mitigated, and also consistent with previous studies where growth rates improved with when JUb44 lawns included as little as 5% of a Proteobacteria strain (Samuel et al. 2016). Good colonizers, like Stenotrophomonas JUb19, also have a positive effect on host developmental speed, possibly indicating a beneficial association. Other bacteria such as Pantoea BIGb0393 increase developmental speed yet are bad colonizers; thus they may represent a good source of nutrition for the worm despite lacking a more intimate interaction with its host. When colonized by the bacterial community, the apparently negative impact of certain individual bacteria seems to be counterbalanced by the presence of other bacteria.

Using whole genome sequences to reconstruct strain phylogenies

To add to the value of this resource for the community and spur future in-depth analysis of C. elegans-microbiome interactions, we sequenced the genome of each of the 12 isolates. Using a combination of short and long reads and a hybrid approach for genome assembly, complete genome sequences were obtained for all CeMbio members. Details on genome characteristics and assembly quality are provided in Table 2.

Complete genomes allowed us to perform Phylogenomic analysis and assign an accurate species level phylogeny for 10 out of 12 CeMbio species. In brief, in the case of BIGb0172 and JUb134, not enough published genomes were available to reconstruct a reliable phylogenomic tree. We then used the 16S rRNA phylogenies (File S1) to classify those organisms as Comamonas piscis for BIGb0172 and Sphingomonas molluscorum for JUb134. For MYb11, MYb10 and JUb66 strains, phylogenomic reconstruction yielded similar trees as their respective 16S rRNA phylogenetic counterparts and we confirmed their assignment as Pseudomonas lurida, Acinetobacter guillouiae and Lelliottia annigena. The bacteria CEent1, BIGb0170 and JUb19 could be attributed more accurately to a single species, respectively Enterobacter hormaechei, Sphingobacterium multivorum and Stenotrophomonas indicatrix. Our phylogenomic analysis further indicated that BIGb0393, MYb71, JUb44 and MSPm1 were new species of, the Pantoea, Ochrobactrum, Chrysobacterium and Pseudomonas genera, respectively (see File S4; Table 1).

Whole genome sequences reveal diverse metabolic competences of the CeMbio strains

Based on these genomes, we determined the presence or absence of different metabolic pathways and the overall metabolic network for each CeMbio strain. We found that the metabolic potential of the CeMbio bacteria ranges from 186 pathways present in the Chrysobacterium JUb44 to 389 pathways in Enterobacter CEent1 (Figure 6A). Overall, common pathways are present in similar abundance across the genomes while more unique pathways are more unevenly distributed. Both Pseudomonas (MSPm1 and MYb11) and Enterobacteriaceae (Pantoea BIGb0393, Enterobacter CEent1, Lelliottia JUb66) strains have overall more pathways and more unique pathways, while Chrysobacterium JUb44 and Sphingobacterium BIGb0170 have fewer predicted pathways overall and also fewer unique pathways (Figure 6A). A principal component analysis of the metabolic potential of the 12 bacteria shows a clustering related to taxonomy, with distinct groupings for the Enterobacteriaceae, Bacteroidetes and Pseudomonas (Figure 6B).

The greatest differences in the predicted metabolic capabilities among the strains is observed in both biosynthesis and degradation of amino acids, carbohydrates, cofactors and vitamins (Figure S6). The microbial potential of biosynthetic and degradation pathways does not correlate with their ability to colonize the C. elegans gut individually. Chrysobacterium JUb44 reaches twofold higher bacterial density (up to 10000 CFU per worms) at the third day of adulthood in comparison to the Enterobacteriaceae strains. However, predicted metabolic pathways alone cannot explain the colonization potential or the competitive fitness of the different CeMbio bacteria in a community setting. For instance, both Ochrobactrum MYb71 and Stenotrophomonas JUb19 are good colonizers and dominant members of nematode microbiome, despite having fewer metabolic pathways than lower colonizing Enterobacteriaceae or Pseudomonas strains. Future work will be needed to resolve the functional importance of metabolic pathways in microbiome assembly in the nematode gut.

As an illustration of the potential for interactions between the strains, we subsequently used metabolic modeling to predict the range of carbon
sources that each strain can utilize and compared it with experimental results obtained from Biolog microarrays (File S3). We found that 66% of the experimental data on carbon utilization were consistent with the predictions, which was similar to our previous study (Zimmermann et al. 2020b). In addition, the experimental data were subsequently used to further optimize the metabolic models (File S3). These adjustments led to an overlap of 99% between the model predictions of usable carbon sources and the Biolog results. In general, the metabolic network model analysis indicated variation in the metabolic competences of the CeMbio strains. Almost all strains can utilize and likely would compete for specific

Figure 3 Colonization of N2 and CB4856 C. elegans strains by the CeMbio community. (A) Proportion of reads in the initial community assembly used as inoculum for the lawns. (B) Proportion of reads in the C. elegans strains N2 and CB4856 and the corresponding lawn samples. The two Enterobacteriaceae CEent1 and JUb66 share a similar 16S rRNA sequence and the V4 PCR primers used in this 16S amplicon sequencing experiment do not discriminate between the two 16S rRNA sequences over this region. (C) Colony forming units (CFUs) of the CeMbio community isolated from N2 and CB4856 nematodes. (D) Mean observed number of CeMbio members (top) and Inverse Simpson Index (bottom) with standard deviation, indicating richness and diversity of the bacterial communities in N2 and CB4856 worms. (E) Principle coordinate analysis of Bray-Curtis dissimilarities of the microbial communities of nematode and lawn samples with an ellipse representing the 95% confidence interval of the nematode samples. These results are from colonization experiment 2.
carbon sources like pyruvate, xylose, and N-acetyl-D-glucosamine. However, the bacteria varied in their abilities to process other carbon sources such as glycogen, phenylalanine, or benzoate (Figure 6C). Together, this type of metabolic niche-partitioning may explain some of the consistently identified community structure that we have observed when colonizing the *C. elegans* gut vs. what is observed in the lawn.

Metabolic network modeling further indicates that the CeMbio community can provide metabolites important for *C. elegans* growth (File S3). This assessment is in general agreement with previous work. For example, a strain of the soil bacterium *Comamonas* was shown to provide vitamin B12 to the worm, which in turn influences development and fertility through the methionine/S-adenosylmethionine cycle while it also processes propionic acid, thereby removing its toxic

---

**Figure 4** Colonization of *C. elegans* gut by the CeMbio community under different plating conditions. (A) Proportion of reads in the initial community assembly used as inoculum for the lawns. (B) Proportion of reads in NGM worm and lawn samples. (C) Proportion of reads in PFM worm and lawn samples. (D) Mean observed number of CeMbio members (top) and Inverse Simpson Index (bottom) with standard deviation indicating richness and diversity of the communities. Stars indicate significant differences in alpha diversity (P < 0.005). (E) Principle coordinate analysis of Bray-Curtis dissimilarities of the microbial communities in nematode and lawn samples with ellipses representing the 95% confidence intervals of the nematode (dashed) and lawn (solid) samples. (F) Colony forming units (CFUs) of the CeMbio community in single nematodes. These results are from colonization experiment 3.
effects (Watson et al. 2014, 2016). Interestingly, several CeMbio members, including MYb11 and MYb71 that are often enriched in the C. elegans microbiome (Dirksen et al. 2016; Zhang et al. 2017; Zimmermann et al. 2020b; Johnke et al. 2020), also possess the pathways for vitamin B12 production (Zimmermann et al. 2020b) and could thus influence similar C. elegans characteristics as the Comamonas strain. Together, these studies illustrate how CeMbio strain combinations and underlying genetic potential can facilitate interrogation of ecologically relevant influence of microbial metabolites on a wide range of life history characteristics and aspects of C. elegans physiology.

CONCLUSIONS
Here we present a robust and flexible resource for the community that has the potential to bring C. elegans research into a more natural and ecologically relevant microbial setting while retaining its strengths as a model system. We demonstrate that the CeMbio strains, either alone or as a community, can affect a key fitness-affecting trait such as developmental rate. Considering that C. elegans in nature is inhabited by a diverse microbial community (Dirksen et al. 2016; Samuel et al. 2016; Berg et al. 2016a; Johnke et al. 2020), the use of the CeMbio resource in C. elegans research will help us to produce a more realistic understanding of nematode biology. We anticipate that the CeMbio community affects the nematode’s interaction with pathogens, as it contains several strains with immune-protective effects, including the two Pseudomonas strains (MSPm1 and MYb11), the Enterobacter CEn1, and Ochrobactrum MYb71 (Montalvo-Katz et al. 2013; Berg et al. 2016b 2019; Dirksen et al. 2016; Kissoyan et al. 2019). Moreover, previous analyses of the C. elegans transcriptome response to Ochrobactrum MYb71 suggests that the bacteria further affect fertility, energy metabolism, metabolism of specific amino acids, and folate biosynthesis (Yang et al. 2019). We expect that other well-studied C. elegans phenotypes are also influenced by colonization with these bacteria. By coupling this resource to extensive microbial genomic resources and metabolic models and a small set of bacteria, we anticipate that the CeMbio resource will both provide a facile entry point for C. elegans researchers into the more natural world and a nearly limitless arena to explore combinations of these strains together.
ACKNOWLEDGMENTS
This work was supported by: NIH grants DP2DK116645 (to B.S.S), R01OD024780 and R01AG061302 (to M.S.); an JGI/DOE grant CSP-503338 (to B.S.S. and H.S.); the Max-Planck Society (fellowship to H.S.); and the German Science Foundation via the CRC 1182 (project A1 to C.K. and H.S.), the Excellence Cluster Inflammation at Interfaces (EXC 306; C.K., H.S.), and under Germany’s Excellence Strategy EXC 2167–39088401 (C.K. and H.S.). Some strains were provided by the CGC, which is funded by NIH Office of Research Infrastructure Programs (P40 OD010440). We also thank Sabrina Butze, Simon Graspeuntner, Georgios Marinos, Katharina Ratjen, Ateequur Rehman, Daniela Vidal and Wentao Yang for support and advice, plus members of the Samuel and Schulenburg labs for helpful comments.

LITERATURE CITED

Figure 6 Comparison of metabolic pathways between the 12 CeMbio strains. (A) Distribution of unique and shared metabolic pathways across the 12 CeMbio members. Pathways are categorized from the most commonly found (present in 10 to 12 genomes) to unique pathways (present in 1 to 3 genomes). (B) Principal component analysis of the metabolic profiles of the 12 CeMbio members. (C) Summary of carbon source utilization for each CeMbio strain as inferred from the genome-scale metabolic models, additionally trained with Biolog EcoPlate plate data, given as a binary response (circle, ability to utilize the indicated carbon source; no circle, inability to utilize it). The carbon source utilization by the whole CeMbio community is highlighted in yellow.