UC San Diego

UC San Diego Previously Published Works

Title

Modeling Meets Metabolomics—The WormJam Consensus Model as Basis for Metabolic Studies in the Model Organism Caenorhabditis elegans

Permalink

https://escholarship.org/uc/item/4fp5d886

Authors

Witting, Michael
Hastings, Janna
Rodriguez, Nicolas
et al.

Publication Date

2018

DOI

10.3389/fmolb.2018.00096

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

frontiers in Molecular Biosciences

ORIGINAL RESEARCH
published: 14 November 2018
doi: 10.3389/fmolb.2018.00096

Modeling Meets Metabolomics—The WormJam Consensus Model as Basis for Metabolic Studies in the Model Organism *Caenorhabditis elegans*

Michael Witting[1,2*], *Janna Hastings*[3], *Nicolas Rodriguez*[3], *Chintan J. Joshi*[4], *Jake P. N. Hattwell*[5], *Paul R. Ebert*[6], *Michel van Weeghel*[7], *Arwen W. Gao*[7], *Michael J. O. Wakelam*[8], *Riekelt H. Houtkooper*[7], *Abraham Mains*[3], *Nicolas Le Novère*[3], *Sean Sadykoff*[4], *Frank Schroeder*[9], *Nathan E. Lewis*[4,10], *Horst-Joachim Schirra*[5], *Christoph Kaleta*[11] and *Olivia Casanueva*[3]

[1] Research Unit Analytical BioGeoChemistry, Helmholtz Zentrum München, Neuherberg, Germany, [2] Chair of Analytical Food Chemistry, Technische Universtität München, Freising, Germany, [3] Epigenetics Department, Babraham Institute, Cambridge, United Kingdom, [4] Department of Pediatrics, University of California, San Diego, La Jolla, CA, United States, [5] Centre for Advanced Imaging, The University of Queensland, Brisbane, QLD, Australia, [6] School of Biological Sciences, The University of Queensland, Brisbane, QLD, Australia, [7] Laboratory Genetic Metabolic Diseases, Amsterdam UMC, University of Amsterdam, Amsterdam Gastroenterology and Metabolism, Amsterdam Cardiovascular Sciences, Amsterdam, Netherlands, [8] Signaling Department, Babraham Institute, Cambridge, United Kingdom, [9] BTI/Cornell University, Ithaca, NY, United States, [10] Novo Nordisk Foundation Center for Biosustainability at University of California, San Diego, La Jolla, CA, United States, [11] Research Group Medical Systems Biology, Institute of Experimental Medicine, Christian-Albrechts-University Kiel, Kiel, Germany

OPEN ACCESS

Edited by:
Reza M. Salek, International Agency For Research On Cancer (IARC), France

Reviewed by:
Etienne Thevenot, Commissariat à l'Energie Atomique et aux Energies Alternatives (CEA), France
Young-Mo Kim, Pacific Northwest National Laboratory (DOE), United States

***Correspondence:**
Michael Witting
michael.witting@helmholtz-muenchen.de

Specialty section:
This article was submitted to Metabolomics, a section of the journal Frontiers in Molecular Biosciences

Received: *31 July 2018*
Accepted: *22 October 2018*
Published: *14 November 2018*

Citation:
Witting M, Hastings J, Rodriguez N, Joshi CJ, Hattwell JPN, Ebert PR, van Weeghel M, Gao AW, Wakelam MJO, Houtkooper RH, Mains A, Le Novère N, Sadykoff S, Schroeder F, Lewis NE, Schirra H-J, Kaleta C and Casanueva O (2018) Modeling Meets Metabolomics—The WormJam Consensus Model as Basis for Metabolic Studies in the Model Organism Caenorhabditis elegans. Front. Mol. Biosci. 5:96. doi: 10.3389/fmolb.2018.00096

Metabolism is one of the attributes of life and supplies energy and building blocks to organisms. Therefore, understanding metabolism is crucial for the understanding of complex biological phenomena. Despite having been in the focus of research for centuries, our picture of metabolism is still incomplete. Metabolomics, the systematic analysis of all small molecules in a biological system, aims to close this gap. In order to facilitate such investigations a blueprint of the metabolic network is required. Recently, several metabolic network reconstructions for the model organism *Caenorhabditis elegans* have been published, each having unique features. We have established the WormJam Community to merge and reconcile these (and other unpublished models) into a single consensus metabolic reconstruction. In a series of workshops and annotation seminars this model was refined with manual correction of incorrect assignments, metabolite structure and identifier curation as well as addition of new pathways. The WormJam consensus metabolic reconstruction represents a rich data source not only for *in silico* network-based approaches like flux balance analysis, but also for metabolomics, as it includes a database of metabolites present in *C. elegans*, which can be used for annotation. Here we present the process of model merging, correction and curation and give a detailed overview of the model. In the future it is intended to expand the model toward different tissues and put special emphasizes on lipid metabolism and secondary metabolism including ascaroside metabolism in accordance to their central role in *C. elegans* physiology.

Keywords: *Caenorhabditis elegans*, metabolic reconstruction, metabolism, metabolomics, metabolic pathways

INTRODUCTION

Metabolism is a key mediator of the biological processes underlying living organisms. Metabolic changes are at the frontline of the cellular response to environmental or physiological changes, and altered metabolism is a hallmark and driver of the pathologies accompanying conditions such as aging and cancer (Finkel, 2015). Nevertheless, our understanding of all the complexities of metabolic processes in different conditions remains incomplete. The model organism *Caenorhabditis elegans* is emerging as a key resource for the study of metabolism in multicellular organisms, as while it shares much of its central metabolic pathways with humans, it is easy to culture in laboratory conditions, can be grown in large populations of isogenic individuals in order to study purely environmental differences, and has a short life span enabling rapid longitudinal data acquisition even across multiple generations (Tissenbaum, 2015; Maglioni and Ventura, 2016; Shen et al., 2018).

Metabolomics evaluates the metabolic state of a given sample by measuring the concentrations of a large number of small molecules simultaneously, allowing metabolic differences between different conditions to be evaluated. Advances in metabolomics involve the development of methods and standards to more accurately detect, distinguish and quantify small molecules from a broad range of pathways in increasingly smaller quantities of sample material. In *C. elegans* the currently detectable metabolome encompasses >1,000 distinct metabolites (not counting lipids), but this is continuously evolving, and the estimated size of the full metabolome under ordinary conditions may be even > 10,000 distinct molecules based on recent metabolomics work uncovering new metabolites (Artyukhin et al., 2018).

Whole-genome metabolic reconstructions are *in silico* representations of all the metabolic reactions in a given organism as a network of metabolites and the reactions in which they are produced or consumed, with associated genes. They are representations of metabolic knowledge in a given organism abstracted to the level of a single cell. These reconstructions allow sophisticated mathematical analysis techniques to make predictions about the dynamic intracellular fluxes under different conditions. In particular, Flux Balance Analysis (FBA) and its derivatives permit the use of whole-genome reconstructions together with experimental molecular phenotypes and biological objective functions in order to obtain optimal fluxes landscapes at steady-states (O'Brien et al., 2015). One can then observe how these landscapes evolve upon mutations and in different environments, or to predict drug targets and biomarkers. For *C. elegans*, several such metabolic reconstructions exist (Büchel et al., 2013; Gebauer et al., 2016; Yilmaz and Walhout, 2016; Ma et al., 2017) (see **Figure 1**). We have been working with the whole community to reconcile and develop a single model, representing the best consensus of known metabolism in *C. elegans*. This community effort has been christened "WormJam" (Hastings et al., 2017).

Metabolic reconstructions and metabolomics characterization in different organisms typically proceed independently; the metabolites that are included in a metabolic reconstruction may be different to the metabolites that can confidently be identified in cutting-edge metabolomics investigations in that organism. This is largely due to the different sources of technical complexity and opportunities in the different types of investigation. The gap that ensues between model and metabolomics is one of the specific areas that we are aiming to address with the WormJam effort, which includes participants both from the metabolomics and metabolic modeling communities.

In this paper, we describe the work we have done to develop and extend the consensus model, including merging pre-existing models and curation of novel pathways for *C. elegans*, and the work that is currently ongoing to bridge between the model and the metabolites that can be detected with metabolomics.

MATERIALS AND METHODS

Reconciling Existing Models

Merging iCEL1273 and ElegCyc

The merging of these two previously published worm reconstructions ElegCyc and iCEL1273 involved using databases and standards for nomenclature. For genes, we first identified a list of unique genes in each model. ElegCyc (Gebauer et al., 2016) and iCEL1273 (Yilmaz and Walhout, 2016) used different gene identifiers, so a table linking the different gene identifiers for *C. elegans* was obtained (http://geneontology.org/page/download-go-annotations). We chose to work with the WormBase gene identifiers (WBGeneXXXXXXXX) for the sake of their simplicity in parsing and formatting, and accompanying ease of access to the online WormBase database (Lee et al., 2018). If the gene was not found in gene mapping obtained from the GO website, Kyoto Encyclopedia of Genes and Genomes (KEGG) (Kanehisa et al., 2017) was queried to obtain gene information. If duplicate gene entries were found because of different naming conventions used in the model, we updated the gene rules and gene-protein-reaction (GPR) association matrix in the model. A gene-protein-reaction (GPR) matrix is a Boolean matrix identifying genes associated to reaction, a.k.a. gene rules. The gene rules are encoded in a text field which describes two properties about the metabolic reactions: (i) the gene products involved in catalyzing a given reaction, and (ii) the relationship between the gene products involved (isoenzymes—OR, multimers—AND).

For metabolites, we identified a unique list of metabolites using 3 different databases: BiGG (King et al., 2016), KEGG (Kanehisa et al., 2017), and MetaNetX (Moretti et al., 2016). For metabolite identifiers that were in the BiGG database format, we extracted the following information: ID from the Chemical Entities of Biological Interest (ChEBI) (Hastings et al., 2013), KEGG Compound ID (Kanehisa et al., 2017), MetaNetX ID (Moretti et al., 2016), charge, and formula. If a metabolite identifier was in KEGG Compound format, we extracted the following information: name, KEGG compound id/drug id/glycan id, and compound formula. These data were used to map the metabolite to its BiGG identifier and retrieve additional information from BiGG. Information for metabolite identifiers in MetaNetX format were extracted as was done with the BiGG database, and we confirmed the information was

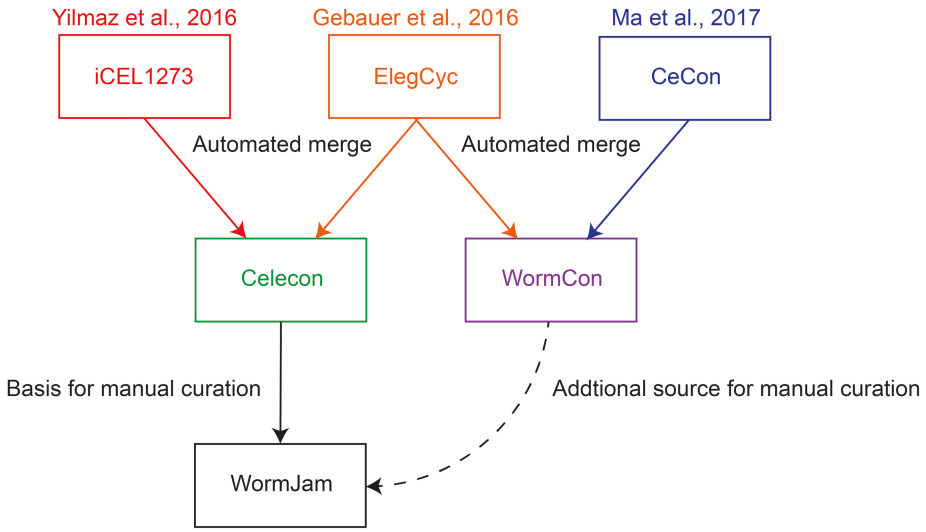

FIGURE 1 | Overview of published *C. elegans* metabolic reconstructions and their relation to consensus models described in this manuscript.

consistent by performing a reverse mapping. If neither KEGG nor BiGG information was found, we performed a manual search. Duplicate metabolite entries within the model were fixed by removing one of the instances and resolving in-model meta information and stoichiometry.

For reactions, we employed a script that does the following steps to identify whether the reactions being compared are the same. All used scripts are available from https://github.com/LewisLabUCSD/celegans_reconciliation. We kept the reactions belonging to the first model ($R_1 = 1893$) and parsed through the reactions in the second model ($r_2 \in [R_1+1, R_C]$ reactions). R_C is the total number of reactions including duplicates in both the models. We performed the following steps on each reaction ($r_1 \in [1, R_1]$) from the first model:

a. Check the *stoichiometry* of r_1^{th} reaction with all reactions (R_1+1 to R_C reactions) (See **Figure 2B**).
   i. If the stoichiometry is the same, check gene-reaction association of the reactions.
      1. If the gene-reaction association is the same, remove the reaction (r_2) coming from the second model.
      2. If the gene-reaction association is different, append the gene rules in the first model using an OR rule and add the relevant gene associations to the gene-reaction association matrix of the first model, then remove the reaction (r_2) coming from the second model
   ii. If the stoichiometry is different, move to r_1+1th reaction.
b. Check the stoichiometry after removing protons to identify if the reactions are the same but have different charge/element balance (See **Figure 2B**).
   i. Perform *elemental balancing* on both reactions to identify the correct version.
   ii. Repeat steps from (a) again.
   iii. *Ensure that in-model meta information is consistent*. The in-model reaction meta information is copied from the source model of the reaction; for e.g., if RXNA (has correct balancing) in the second model and is identified as a duplicate of RXNB (has incorrect balancing) in the first model, all in-model meta information will be set to the reaction values from the second model.

Merging of CeCon and ElegCyc

To begin merging, ElegCyc was obtained in SBML format from the Supplementary Information of Gebauer et al. (2016), and an SBML format of CeCon (Ma et al., 2017) was exported from Pathway Tools. The two models were then translated into the MetaNetX format (Moretti et al., 2016), and imported into MATLAB for merging. The COMMGEN toolbox (van Heck et al., 2016) contains multiple algorithms for the semi-automated merging of genome scale models, which were applied to the two models. After completely merging the two models into a single COMMGEN entity in MATLAB, the biomass reaction of ElegCyc was set as the objective function for flux balance analysis testing of the model's validity, as CeCon, which was previously used for the mapping of —omic data, does not contain a biomass reaction.

Next, the following steps were performed semi-automatically through COMMGEN functions.

1. Merging of duplicate reactions
2. Merging of reactions with similar species
3. Removal of nested reactions
4. Alteration of invalid transport reactions
5. Removal of invalid external reactions
6. Checking of reactions with the same metabolites, but differing stoichiometry
7. Merging of similar transport reactions

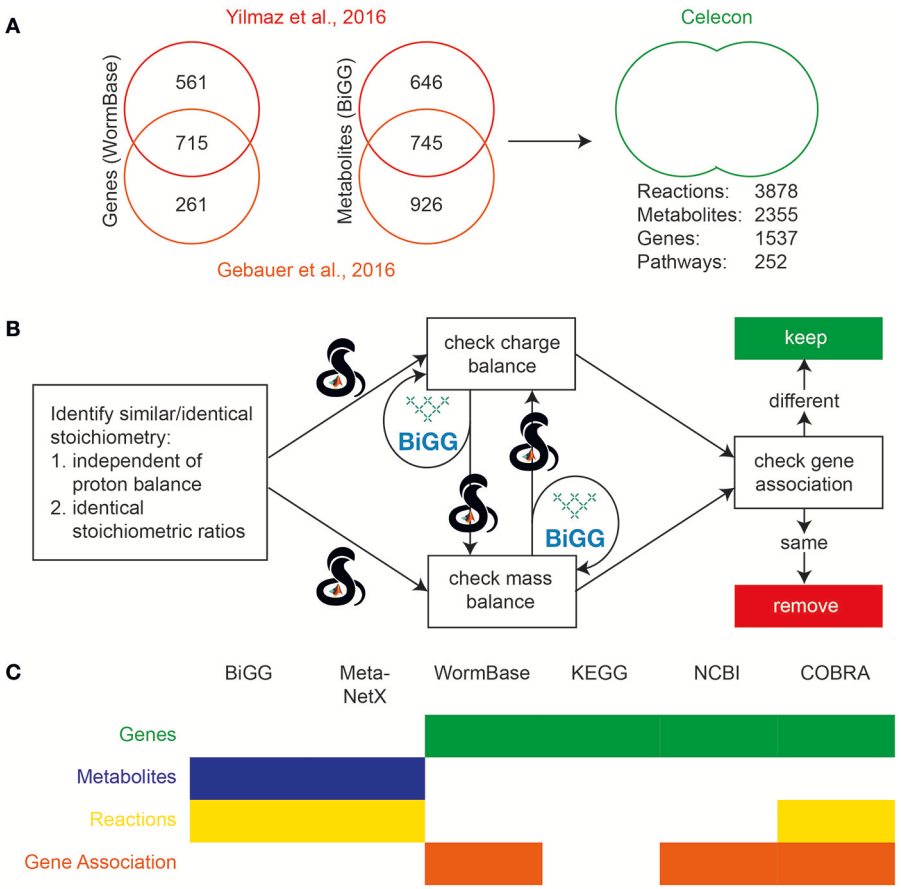

FIGURE 2 | Initial reconciliation and merging of two previously published *C. elegans* reconstructions. **(A)** Genes from the two earlier reconstructions were compared using Wormbase, and metabolites were compared using BiGG. Sizable differences were seen in the scope of the content of the two reconstructions, and so they were curated and merged. **(B)** Duplicate reactions were identified and eliminated using the reaction formulae and information on charge and mass balances using the BiGG and MetaNetX databases. The calculations were performed using the COBRA Toolbox in MATLAB. **(C)** Several databases were used to reconcile various properties of the previous reconstructions. The filled boxes indicate the usage of the database, while color indicates the property reconciled.

8. Merging of duplicate reactions

In each function, COMMGEN suggested potential merge candidates, and a decision was manually made whether to merge the reactions or keep both based on literature and databases such as KEGG (Kanehisa et al., 2017). After every merge, the model was exported to COBRA Toolbox format, and checked for viability by ensuring flux was able to be carried through the merged model's biomass reaction in a "free-growth" simulation. If the merge was inviable, the change would be reverted. At the end of this process, the resulting model, named WormCon, was exported to SBML format, and further merging was halted in favor of manual curation as part of WormJam.

Metabolite Structure Curation

Chemical structures were associated with metabolites included in the consensus reconstruction. For the bulk of the metabolites, structures were found in ChEBI (Hastings et al., 2013). If no structure was available in ChEBI, other databases, e.g., KEGG (Kanehisa et al., 2017), Human Metabolome Database (HMDB) (Wishart et al., 2018), Chemspider (Pence and Williams, 2010), PubChem (Kim et al., 2016), and LipidMaps (Fahy et al., 2009) were queried. Structures not available in any database were drawn in MarvinSketch 18.1.0 (ChemAxon, Budapest, Hungary).

In certain cases, only charged or neutral structures were available in the databases, so the corresponding missing structure was then also drawn in MarvinSketch 18.1.0. All structures missing in ChEBI were submitted via the batch submission tool to assign new ChEBI identifiers (see *Submission of New Structures*).

Structural Similarity Calculation

Structural similarity between metabolites was calculated using the Tanimoto similarity of chemical fingerprints as implemented in JChem for Excel 18.5.0196 (ChemAxon, Budapest, Hungary) in Microsoft Excel 2016. To define cut-offs for structural similarity to identify potential reaction pairs, Tanimoto similarities were calculated from known reaction pairs from the WormJam consensus model.

Secondary Identifier, MS/MS, and NMR Spectra Search

The Chemical Translation Service (http://cts.fiehnlab.ucdavis.edu/) from the Fiehn lab was used to retrieve secondary identifiers from different databases (e.g., KEGG, HMDB, LipidMaps) by the InChIKey if missing. All spectra from MassBank of North America (MoNA, Download 10.June.2018) were downloaded and sorted according to type: LC-MS, GC-MS, CE-MS, *in silico* and others. InChIKey and SPLASH IDs for each spectrum together with the compound name were isolated for comparison against predicted metabolites (Wohlgemuth et al., 2016). NMR spectra were downloaded from HMDB (Download 03.July.2018) and associated with the respective metabolite and its InChIKey.

Sharing of the Reconstruction and Manual Curation

The first public release of the reconstruction, as described in the current paper can be accessed from BioModels (Chelliah et al., 2015) under the accession MODEL1807230002. To facilitate community curation, the WormJam community uses a shared Google Drive. All past reconstructions and versions of the WormJam reconstruction are stored in standard computer-readable and human-readable formats. To exchange and re-use the models, they are provided in SBML (Hucka et al., 2003), both Level 2 (Hucka et al., 2015) with COBRA-specific notes, and Level 3 (Hucka et al., 2018) with the Flux Balance Constraint package (Olivier Brett and Bergmann Frank, 2018). For manual curation, these models are converted into spreadsheets using a customized SBtab format (Lubitz et al., 2016). Interested persons can join the community by subscribing the WormJam Google group (https://groups.google.com/forum/?hl=en#!forum/wormjam).

Curation of Metabolites From Literature

Metabolites detected in different *C. elegans* metabolomics related publications have been manually curated by extracting relevant information from text, tables, figures and Supplementary Information if available. The most unambiguous structure was curated based on details reported in the respective publications. Structural representation in form of SMILES and InChIs were curated together with identifiers from different databases as well as PubMed IDs of the respective publication (**Supplementary Information Table 1**).

RESULTS

Preexisting Reconstructions

Several *C. elegans* metabolic reconstructions exist and are used as basis for the WormJam consensus model. The first reconstruction was published in 2013 in the frame of the Path2Models project and was based on automatic reconstruction from KEGG pathways, MetaCyc (Caspi et al., 2016) and a gap-filling step. In 2016, ElegCyc (Gebauer et al., 2016) and iCEL1273 (Yilmaz and Walhout, 2016) were published side-by-side. ElegCyc used PathwayTools (Karp et al., 2016), while iCEL1273 is based on a new workflow developed by the authors. Finally, in 2017 CeCon, also based on PathwayTools, was published by Ma et al. (2017). **Table 1** summarizes key metrics of each model.

To generate a consensus model, different models have been merged together in several stages. The paragraphs below describe the merging process of two reconciliations, which were the basis for the WormJam consensus model (**Figure 1**).

Initial Automated Model Merging

Merging of iCEL1273 and ElegCyc

The manual curation of the final WormJam model was preceded by finding a consensus and merging two previously published worm models iCEL1273 (Yilmaz and Walhout, 2016) and ElegCyc (Gebauer et al., 2016). The details of these models have been discussed in the previous section. The preliminary merging process was conducted using MATLAB and the COBRA Toolbox. Metabolic models describe a list of reactions which transform chemical compounds (metabolites) using enzymes (gene products). Therefore, we merged the metabolic models by first finding unions of the genes, metabolites, and reactions across all of the models.

Genes

We used the list of genes in the GO website (http://geneontology.org/page/download-go-annotations) and KEGG to map common gene names (nomenclature used in the models) to WormBase gene identifiers. We chose to work with the WormBase gene identifiers (WBGeneXXXXXXXX) for the sake of their simplicity in parsing/formatting, and ease of access to the online WormBase database (Lee et al., 2018). For details on reconciliation of genes, please see section Materials and Methods. We were able to map 980 genes in ElegCyc to 974 unique genes and 1292 genes in iCEL1273 to 1276 unique genes.

Metabolites

We used four databases: BiGG (King et al., 2016), ChEBI (Hastings et al., 2013), KEGG (Kanehisa et al., 2017), and MetaNetX (Moretti et al., 2016). The metabolite nomenclatures in the models was converted to BiGG as it is the most widely used in the field of metabolic modeling. We used these databases to reconcile various metabolite properties such as charge, formula, and in-model meta information. We were able to map and/or process 1626 metabolites in ElegCyc and 1401 metabolites in WormFlux.

Celecon Merge (See Figure 2A)

The above processing resulted in versions of ElegCyc and iCEL1273 with annotations in the same format including their in-model meta information (i.e., reaction notes, reaction subsystems, reaction compartmentalization, metabolite notes, metabolite charge, metabolite formula, metabolite compartmentalization, database specific identifiers, and/or any other supporting/cross-reference information). Identification of metabolites (M) and genes (G) were followed by performing a merge such that $M_C = M_1 + M_2$, $G_C = G_1 + G_2$, and $R_C = R_1 + R_2$. Here, subscript "1" or "2" represents first or second model, respectively. After identification of duplicate sets of metabolites/genes and combining their respective in-model meta

TABLE 1 | Metrics of pre-existing reconstructions.

Name	Compounds	Reactions	References	Notes
BMID000000141468	3207	2272	Büchel et al., 2013	Automatic reconstruction from KEGG PATHWAYS, MetaCyc and gap-filling
ElegCyc axenic	2357	1893	Gebauer et al., 2016	Bacteria-free growth media
ElegCyc *E. coli*	2357	1921	Gebauer et al., 2016	*E. coli* OP50, *E. coli* biomass composition (Orth et al., 2011) serves as the growth media.
iCEL1273	1718	1985	Yilmaz and Walhout, 2016	
CeCon	2166	2085	Ma et al., 2017	

information, we were left with a model with 2355 metabolites (M_C) and 1537 genes (G_C) and 3878 reactions (R_C).

Reactions

Following the initial merge, we identified duplicate reactions using steps described in the section Materials and Methods. We used ElegCyc as the first model and iCEL1273 as the second model. The above process was able to remove 611 reactions. The biomass reactions from both the models were kept for evaluation at a later stage. Therefore, the final merged model had 2355 metabolites (M_{CycCEL}) and 1537 genes (G_{CycCEL}) and 3267 reactions (R'_{CycCEL}), with unique lists of gene, metabolites, and reactions with their respective in-model meta information (See **Figure 2C**).

Merging of CeCon and ElegCyc

WormCon is a consensus model resulting from the merge of CeCon (Ma et al., 2017) and ElegCyc (Gebauer et al., 2016) (**Figure 3A**). Both models were originally designed in the Pathway Tools software (Karp et al., 2016), and later exported to SBML format. An initial comparison was made between the two models using various top-level metrics including number of reactions and metabolites as a benchmark. These metrics are generated by the Pathway Tools software, and a comparison is shown in **Figure 3A**. The versions of ElegCyc and CeCon used for merging contained 1193 and 1923 reactions and 773 and 1394 metabolites, respectively. For merging, both models were converted to the MetaNetX namespace, and then combined in MATLAB 2015a using the COMMGEN tool, a package allowing for semiautomatic merging of genome scale models (van Heck et al., 2016). An initial consensus model was automatically generated, then COMMGEN's functions were used to iteratively refine the model as shown in **Figure 3B**. With the biomass reaction from ElegCyc set as the objective (CeCon does not have a biomass reaction), duplicate reactions were removed from the consensus model. The next two steps were to compare reactions with similar species to determine whether they were the same reaction, just differing by namespace alterations, as well as the removal of nested reactions. Following this, more reactions that could be duplicates were identified. All iterations were performed sequentially, but this resulted in the failing of the model as the merging of reactions vital for the model's function sometimes led to loss of model viability.

To resolve this issue, each stage of merging was performed, then the new version of the model was converted back to SBML, and tested to see if it could produce biomass, and then saved separately if viable. This resulted in a functioning metabolic model, called WormCon can carry flux incorporating reactions from both original models in the process. WormCon contains 2621 genes, 3148 metabolites and 3434 reactions, and has been frozen in its current state for incorporation into the final consensus model.

Curation of the Merged WormJam Model

Initial Manual Curation

The starting point for manual curation was the automatically merged Celecon model, with information from WormCon being an additional source for manual curation (**Figure 1**). Initial manual curation efforts included the further identification and resolution of duplicate reactions that had not been automatically removed (typically because of differences in gene annotation, compartmental localization, or presence/absence of minor cofactors). The charge and ChEBI annotation for metabolites were checked, and duplicate metabolites were also removed (which in some cases led to identification of a set of associated duplicated reactions as well). The metabolic literature was consulted to verify the presence of certain reactions, and in some cases reactions known to be not taking place in *C. elegans* were flagged for removal.

The manual curation effort then further proceeded to enhance and extend the representation of specific aspects of biology not previously included in any of the merged models. This includes the correction of wrong annotations and reactions as well as the addition of new pathways.

Correction of Pathways

Glycogen metabolism

Glycogen metabolism is key to regulate lifespan and healthspan in *C elegans* (Seo et al., 2018). Glycogen is a branched, unbounded, molecule formed by stochastic additions of UDP-glucose molecules forming chains of glucose residues linked by α(1,4) glycosidic bonds from which new chains of glucose residues branch off by α(1,6) glycosidic bonds (Roach et al., 2012). The synthesis is seeded by the protein glycogenin, which also acts as a catalyst for the first glucose extensions. Then glycogen synthase generates the α(1,4) bonds while the glycogen branching enzyme forms α(1,6) bonds. The breakdown of glycogen relies on the split of α(1,4) bonds by glycogen phosphorylase, that produces glucose-1-phosphate and the removal of α(1,6) bonds by the debranching enzyme that produces glucose. Because of the combinatorics born out of the branching and stochastic elongation, it is impossible to model glycogen metabolism accurately and mechanistically.

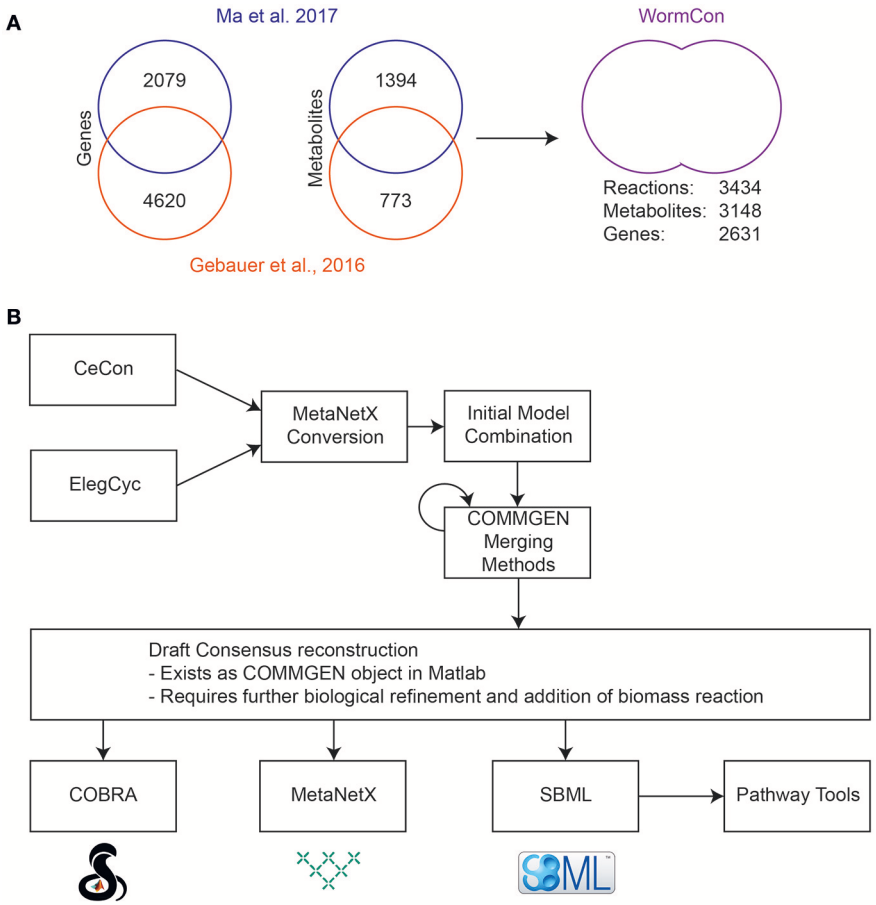

FIGURE 3 | Initial reconciliation and merging of two previously published *C. elegans* reconstructions. **(A)** Metrics obtained from the original Pathway Tools reconstructions of CeCon (Ma et al., 2017) and ElegCyc (Gebauer et al., 2016) and final metrics of the merged model, called WormCon. Data sources for the construction of CeCon and ElegCyc can be found in the original publications. **(B)** The iterative merging process used to generate the draft WormCon consensus model.

More importantly, because of their complexity and the cycles involved, the metabolism of glycogen in existing reconstructions are often incorrect, consuming, and producing different amounts of glucose residues. Instead, we opted for a phenomenological model that consumes the correct amount of UDP glucose, forms the correct ratio of α(1,4) and α(1,6) bonds, then produces the correct ratio of glucose-1-phosphate and glucose. We also removed the reactions relating to maltose metabolism, as they are likely fossil reactions coming from bacterial metabolism reconstructions. The final pathway is quite simple, with 6 forms of glycogen, and 9 reactions. Each synthesis consumes 12 UDP-glucose, and each breakdown produces 11 glucose-1-phosphate and 1 glucose.

Biosynthesis and degradation of fatty acids and BCFA

C. elegans is known to produce monomethyl branched chain fatty acids (BCFA) on its own (Perez and Van Gilst, 2008). The different reconstructions only contained lump reactions for this biosynthetic pathway summing up all individual steps into a single reaction. In contrast, the biosynthesis of straight chain fatty acids was represented in much more detail. We have modified the reactions to the same level of detail, which represents always one round of elongation. Likewise, fatty acid oxidation was adapted to the same level of detail.

Sphingolipid metabolism

Sphingolipids in *C. elegans* were shown to contain branched chain C17iso sphingoid bases (Chitwood et al., 1995; Zhu et al., 2013; Hannich et al., 2017). However, due to homology searches to human genes, the reconstructions contained C18 sphingolipids produced from condensation of palmitic acid with serine. In contrast to this, *C. elegans* uses C15iso fatty acid which is condensed with serine. We have corrected all sphingolipid metabolites and pathways to contain now the C17iso sphingoid base. Additionally, ceramides in *C. elegans* often contain 2-hydroxy fatty acids. We have added pathways to produce all the 2-hydroxy fatty acids shown to be present in the worm by Chitwood et al., and added ceramide biosynthesis reactions for the normal and 2-hydroxy fatty acids (Chitwood et al., 1995; Hannich et al., 2017).

Epigenetics Marks

The expression of *C elegans* genome has been shown to be regulated by a range of epigenetics marks, deposited on histones, DNA, or RNA. Some of these marks were present in the starting reconstructions, albeit not always in the right compartment. We fixed histone acetylation, methylation, demethylation, and trimethylation, by adding missing reactions and fixing wrong compartments as well as associated exchange reactions. Similarly, we added a few epigenetic marks recently discovered. They include DNA nuclear N6-deoxyadenine methylation (Greer et al., 2015), modulated by the methylase DAMT-1 and the demethylase NMAD-1, and the ribosomal RNA Cytidine methylation by NSUN5, which has been shown to modulate lifespan (Schosserer et al., 2015). We also added methylation and demethylation of mitochondrial N6-adenine, by *tfbm-1* based on enzyme homology and subcellular localization. These modifications affect different nucleotides and deoxynucleotide residues. We therefore modified the biomass reactions to disentangle the individual bases.

New Pathways

Different aspects of the worm's metabolism were not covered by previous metabolic models. These pathways and reactions cover *C. elegans* specific metabolic pathways that cannot be inferred from any homology to human genes. We have added two *C. elegans* specific pathways to the model which represent important aspects of the worm's biology.

Maradolipids

Maradolipids, chemically 6,6' diacyltrehaloses, have been identified in dauer larvae of *C. elegans* (Penkov et al., 2010). It is suggested that they are required for desiccation tolerance in the dauer stage. Although the exact biosynthetic pathway and genes are not known yet, a potential reaction sequence was added to the reconstruction. Lysomaradolipids, containing only one fatty acid have been described. Therefore, we suggest a stepwise acylation of the two 6 and 6' position of trehalose using fatty acids CoAs as substrates.

Ascaroside biosynthesis

Ascarosides are important molecules in the biology of *C. elegans* and serve as messenger molecules in worm to worm communication. For example, different ascarosides are responsible for the entry into the dauer stage upon overcrowding, larval dispersion, and male attraction (Izrayelit et al., 2012; Srinivasan et al., 2012; von Reuss et al., 2012; Panda et al., 2017).

They are produced from long chain (LC) and very long chain (VLC) fatty acids, which are either ω- or ω-1-hydroxylated and attached to ascarylose. In several rounds of peroxisomal β-oxidation, involving the genes *acox-1* (WBGene00008564), *maoc-1* (WBGene00017123), *dhs-28* (WBGene00000991), and *daf-22* (WBGene00013284), these long chain ascarosides are broken down in the form of CoA thioesters to shorter versions (von Reuss et al., 2012).

All metabolic reconstructions developed so far did not contain this biosynthetic pathway, as no corresponding human pathway exists. We have added this pathway by manually curating data from different publications. In a similar fashion, structures of ascaroside-CoA thioesters where drawn and their charged version (−4) was used to construct the peroxisomal β-oxidation pathway. Additionally, hydrolysis reactions to free the individual ascarosides have been added together with export reactions for the extracellular export.

Our current knowledge on the biosynthesis of ascarosides is limited. The involvement of *acox-1*, *maoc-1*, *dhs-28,* and *daf-22* in the biosynthesis has been established, which work together for the peroxisomal beta-oxidation of long and very long chain fatty acid versions of ascarosides. Experiments using axenic media have shown that *C. elegans* is still able to produce ascarosides without the presence of bacteria. Therefore, ascarylose biosynthetic genes must be present in *C. elegans*. The exact biosynthetic pathway is not known yet, there are some hints that ascaroside biosynthesis is distinct from the dTDP-rhamnose biosynthesis (Feng et al., 2016). Furthermore, no ascarosyl transferase in *C. elegans* is known. Therefore, a lump reaction generating ascarylose from glucose and adding ω- or ω-1-hydroxyl fatty acids to the sugar moiety was added to the model. Additionally, export reactions from the cytosol to extracellular space were added for the secretion of these molecules. At the current stage only, biosynthetic reactions for basic ascarosides are included. Modified ascarosides have to be added in future versions.

Current Status of the Wormjam Consensus Reconstruction

At present, the WormJam reconstruction encompasses 2833 metabolic compounds distributed across four different compartments—cytosol, mitochondrion, nucleus, and extracellular —, representing 1629 unique chemical species after removing duplicates across compartments. It features 3632 reactions associated with 1524 genes and participated in 70 pathways.

Once the manual curation effort had been completed, the model was rebuilt into computable format from the manually curated spreadsheets in SBtab format for evaluation. Each reaction in the model was tested with respect to its ability to carry flux, and each metabolite whether it could be produced. It is typical for large-scale metabolic models to include some percentage of reactions that cannot carry flux and metabolites that cannot be synthesized. However, it is important that all key pathways are fully functional and biomass can be generated as expected from all cellular precursors. The WormJam model on initial post-curation build contained 1050 reactions that could not carry flux, including the assembly of the primary biomass. However, this was largely due to a few minor errors in key pathways, which were quickly resolved. The current version of the model contains 685 blocked reactions.

Curation of Metabolite Structures

Representation of Metabolite Structures in WormJam

The correct representation of metabolite structures is a key issue, since structures represent the most unique identifier for a given molecule. The problem of correct structural representation in metabolic models are two-fold. Firstly, charged

molecular structures (and their formulas) representing the major microspecies at a given pH (mostly 7.3 in the cytosol) are required for stoichiometrically correct model representation. This approach is also chemically accurate, since acetate is not acetic acid, they are two different molecules and therefore cannot have the same identifier. However, metabolomics scientists who would like to map their results to metabolic pathways in the model are accustomed to working with neutral formulas and molecules, and directly using charged formulas from the metabolic model to calculate m/z ratios of different adducts leads to incorrect results. In the WormJam consensus model we are using ChEBI as primary source for the structures of charged and uncharged molecules for structural identification. ChEBI represents a rich data source with established chemical ontology relationships to link charged and uncharged variants of different molecules. Therefore, it is essential to also supply the correct neutral form of a molecule that can be used in metabolomics for annotation of signals.

We have added chemical structures encoded as InChIs for each metabolite, where possible. Several metabolites represent generic structures, e.g., "Ceramide" containing an R group for different acyl chain lengths. In such cases we report only the formula containing the R group.

In several cases either the neutral or charged version of the molecules were missing in ChEBI. These have been drawn manually in MarvinSketch 18.1.0 and structures were submitted via the batch submission tool to ChEBI. In total we have submitted >500 new metabolite structures either from the model or the curation of metabolites from literature (see below: *Curation of metabolites from C. elegans metabolomics studies*).

Metabolites Detected in *C. elegans* (Metabolomics) Studies

Curation of Metabolites From *C. elegans* Metabolomics Studies

To get an impression which metabolites have been detected so far in *C. elegans* using different metabolomics techniques and how they map onto the metabolic pathways in the WormJam model, we curated metabolites from over 40 publications. This list is not complete and will be extended in future but represents a starting point to investigate how far our consensus model and metabolomics are diverging.

Metabolites were extracted from the text, figures, and supplementary tables if present (see **Supplementary Table 1**). The information found, and the level of curation differs in the different articles, ranging from pure names (e.g., leucine) to reporting of metabolite identifiers (e.g., KEGG numbers). To standardize the manual curation of all the publications, we have taken the closest, most probable matching structure from database (mostly from ChEBI). Reporting of ions was normalized to neutral forms, for example acetate was normalized to acetic acid. We used generic structures (e.g., leucine instead of L-leucine) for metabolites where no stereochemistry was reported. In total 1223 unique metabolites were curated for which a structure could be added. Additionally several lipids with generic lipid identifiers such as PC(40:5) were also collected. The total number was 1622 metabolites and lipids detected in 43 publications.

Different analytical techniques (and combinations thereof) were used for analysis of the worm metabolome. Most studies used NMR (14 publications) or NMR in combination with GC- or LC-MS (3 and 6 publications, respectively), followed by LC-MS (9 publications) and GC-MS (5 publications) alone. Furthermore, for lipid analysis, shotgun lipidomics was employed several times. Several metabolites were detected with different techniques, which might increases the confidence of the different identifications. Thirty seven percent of the metabolite structures were detected at least twice, <5% ten times and more. Glutamic acid (CHEBI:18237) and Glycine (CHEBI:15428) were the most frequently detected metabolites, with each being detected 30 times. However, multiple detection and identification have to be handled with great care, since confidence scores or experiments for identification were not always given. It might also happen that wrong annotations spread across several papers once it has been assumed to be positively identified.

Comparison of "*in silico*" and Measured Metabolites

To get an impression of the coverage of the *C. elegans* metabolome across all included publications, we compared the percentage of metabolites that have been detected in any publication with the metabolite structures present in the WormJam model. Since the structures reported in the curated publications might not reflect the actual stereochemistry, or the used analysis method cannot correctly determine the stereochemical configuration, we compared the first part of the InChI only, which represents the atom connectivity layer. This yields in total 992 unique entities for the WormJam model and 1140 for the literature curated metabolites. These metabolites were mostly derived from the following metabolite classes: amino acids, fatty acids, central carbon metabolism, lipids, and secondary metabolites like bile acids and ascarosides. A detailed list of curated metabolites can be found in **Supplementary Information Table 1**. This approach also merges some different structures into a single entity (e.g., β-D-galactose 6-phosphate, NBSCHQHZLSJFNQ-FPRJBGLDSA-N, CHEBI:41076 and β-D-glucose 6-phosphate, NBSCHQHZLSJFNQ-VFUOTHLCSA-N, CHEBI:17719), but allows a fairer comparison of the predicted and detected metabolites by ignoring stereochemistry. Three hundred eighty-two metabolite entities overlapped between the model and the detected metabolites (**Figure 4A**). Common metabolite entities include organic acids, amino acids, fatty acids, nucleotides, cofactors and several ascarosides.

Interestingly, 758 of the detected entities have no counterpart in the model. We closely examined several of the metabolites. To identify potential connection points to metabolic pathways or metabolites of similar structures in the WormJam model, we used chemical similarity between predicted and detected metabolites not found in the model. Tanimoto similarity was used as measure of chemical similarity (Nobeli et al., 2003). High Tanimoto similarities are found between highly similar

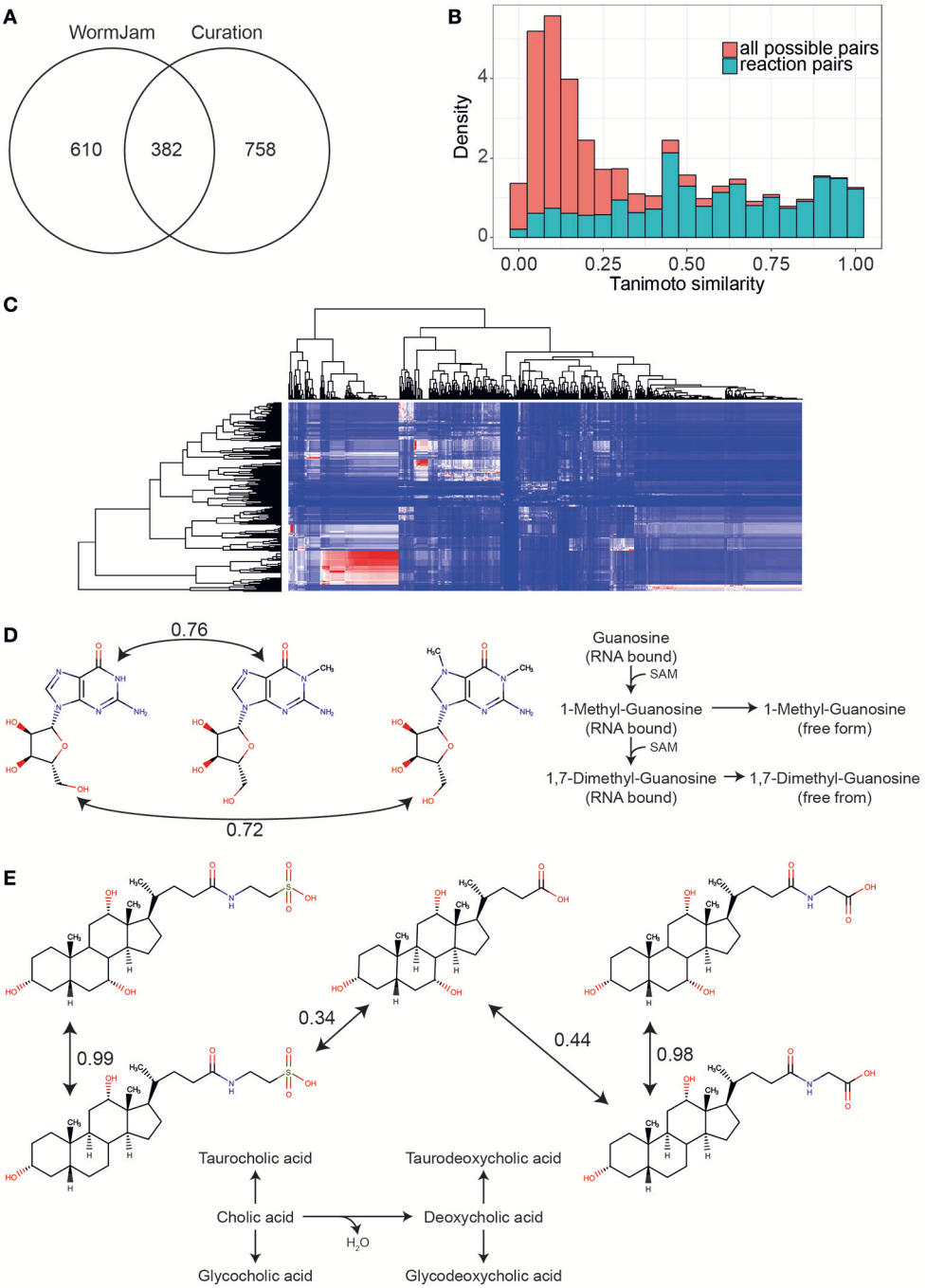

FIGURE 4 | (A) Overlap of unique metabolite entities (first part of InChIKey) between the WormJam model and metabolites curated from literature **(B)** Histogram of Tanimoto similarities between all possible metabolite pairs in the WormJam model (red) and metabolites connected by a biochemical reaction (blue) **(C)** Heatmap of Tanimoto similarities between metabolites detected but not in the model (rows) and metabolites from the WormJam model (columns). Color scale is blue for 0 to red for 1. Whereas, 1 refers to a high structural similarity. The large red cluster contains mostly fatty acids and ascarosides. **(D)** Tanimoto similarity between guanosine and 1-Methylguanosine and 1,7-Dimethylguanosine (both detected but not in the model). These metabolites have high structural similarity indicating a potential relation as substrates and products in biochemical reactions. Both the methylated and dimethylated from are derived from RNA bound modified nucleotides, which are released upon degradation. **(E)** Structural similarity between bile acids from the model (cholic acid, glycocholic, and taurocholic acid) and glycodeoxycholic and taurodeoxycholic acid detected in different studies.

chemical structures. Therefore, neighbors in metabolic pathways should show high similarity values. To identify a good cut-off value for the Tanimoto similarity, we first calculated this value for all substrate-product pairs in all reactions of the WormJam model removing hub metabolites like water, ATP, CoA etc. A background dataset with all combinatorial possible metabolite pairs was calculated to determine which values are most informative. **Figure 4B** shows the histograms for both distributions. Based on this data we decided to use a cut-off >0.4. The pairs contained several generic similarities, e.g., fatty acids are similar to other fatty acids. A large cluster was occupied by fatty acids, sugars and several ascarosides (**Figure 4C**). Several examples are presented in the following paragraphs related to specific aspects of metabolism or the biology of *C. elegans*. This also includes examples where identifications are in contrast to the current knowledge of *C. elegans* metabolism and more experiments for verification of falsification need to be performed.

One example is the detection of creatine and creatinine. Three of the studies have reported creatine and four creatinine using different analytical methods. No evidence for the use of creatine and creatinine in *C. elegans* exists. The worm uses arginine instead of creatine as phosphagen. *C. elegans* harbors a putative ortholog of the human mitochondrial creatine kinase CKMT2 (W10C8.5, WBGene00021128), however this might also act as arginine kinase. This enzyme was expressed at a higher rate, together with muscle related proteins, in *daf-2* worms (Depuydt et al., 2013), but no ortho- or homologs for the enzymes required for the biosynthesis of creatine have been identified. Creatine and creatinine might have been misidentified.

Several small organic acids have been identified in the exometabolome of *C. elegans*. These acids, including e.g., isovaleric acid or 3-hydroxyisovaleric acid are breakdown products of branched chain amino acids and have been measured in the exometabolome of mitochondrial mutants or anoxic conditions. In the WormJam model, they are only present bound to CoA, despite several carnitine derivatives of these organic acids being detected experimentally. Since these organic acids are produced in the mitochondria they also have to be transported, which is achieved by the carnitine transport system. Small hydroxy organic acids are excreted by long-lived mitochondrial mutants (Butler et al., 2010, 2013).

Interestingly, several fatty acids were not present in the model, plus fatty acid ethanolamides. Both play important roles in the biology of *C. elegans*. Fatty acids represent building blocks of lipids and are also linked to production of secondary messengers and signaling molecules. Related to this class of metabolites, several new ascarosides are contained in the list of metabolites curated from literature. These ascarosides contain very long fatty acid side chains, which are currently not covered in the consensus model. Also, the respective fatty acids (>C26) are not present in the model. Currently, only two publications have detected a C30:0 fatty acid (Gao A. et al., 2017; Gao A. W. et al., 2017). There is growing evidence that the biosynthesis of ascarosides is more complex than previously expected (Zhou et al., 2018). So far, only peroxisomal beta-oxidation of simple ascarosides is covered in the model. Several individual lipid species have been reported.

Several modified nucleotides have been detected in one particular study, including 1-Methylguanosine, 1,7-Dimethylguanosine, 1-Methylinosine, 5-Methylcytidine, and N6-Carbamoyl-L-threonyladenosine. These nucleotides are found in RNA and represent degradation products if found in free form. These modified bases have been measured by LC-MS from RNA samples in *C. elegans* (van Delft et al., 2017). These modifications play important roles in fine tuning RNA function, e.g., tRNA contains several modified nucleotides. Additionally, the production of these modified nucleotides consumes distinct other metabolites, e.g., SAM, and therefore represents an important sink for methyl groups. Therefore, they will need to be added in future versions of the model (**Figure 4D**).

Two bile acids, taurodeoxycholic, and glycodeoxycholic acid have been detected, but are not included in the WormJam model. The currently best-studied bile acid-like structures are the dafachronic acids (Mahanti et al., 2014; Aguilaniu et al., 2016). However, recent investigations from two WormJam labs have shown that more steroid derived molecules exist in *C. elegans* than previously expected (Frank Schoeder, personal communication, **Figure 4E**).

Metabolites that are found in the WormJam model but not detected contain mostly CoA derivatives of different molecules or molecules that are not stable enough for detection or only have a short half-life.

Public Availability of MS/MS and NMR Spectra and Reference Standards

Correct identification of metabolites from non-targeted metabolomics plays an important role for biological interpretation and relies on the availability of reference standards and reference spectral libraries. Therefore, we checked the availability of tandem MS spectra in MassBank of North America (http://mona.fiehnlab.ucdavis.edu/). Spectra were sorted into LC-MS, GC-MS, CE-MS, *in silico* and others (no method available in metadata). Metabolites from the WormJam model were used for the search.

We used the InChIKey as search criteria. First, the complete InChIKey was used to identify perfect matches. Two hundred eighty-eight metabolites matched exactly to LC-MS spectra, while 230 were found to have at least one GC-MS related entry in MoNA. Between the two 199 were overlapping. Metabolites for which reference tandem MS spectra are available are mostly from central metabolic pathways and for which reference materials are readily available.

Likewise, we checked for the public availability of NMR spectra. We downloaded all NMR spectra files from HMDB and isolated the HMDB IDs and added the fitting InChIKeys, which were used as search criteria for the WormJam metabolites. In total 81 metabolites had one or more NMR spectra associated in HMDB.

We were next interested which metabolites are publicly available as reference standards for future metabolomics experiments. We queried the catalog of Sigma-Aldrich/Merck, Cayman Chemicals, and other vendors and checked for

availability of substances. Four hundred forty metabolites are available from Sigma-Aldrich/Merck and 131 from additional chemical vendors. For most metabolites for which MS/MS and NMR data is available, a reference standard is also available. Poorly covered areas are different: mostly secondary metabolites, e.g., ascarosides. The Schroeder lab has synthesized a range of reference standards for this substance class, which are also available upon request (Frank Schroeder, personal communication). However, for future metabolite annotation and identification in *C. elegans*, reference spectra will be required in the public domain.

Concentrations of Metabolites in Worms

Although many metabolites have been detected, so far only a few papers report absolute concentrations of known substances. One particular example is presented by Gao A. W. et al. (2017). Results for analysis of fatty acids and amino acids are an exception as they were expressed in nmol/mg of protein. Concentrations from different labs, however might not be directly comparable with each other due to differences in worm cultivation (e.g., different peptone/tryptone). However, a concentration range in which a metabolite might be expected can be derived from this. The curation of this valuable information will become more and more important. Following the example of the human metabolome database, which reports metabolite concentrations in different biofluids for healthy and diseased individuals, concentrations linked with the genotype and phenotype of a worm experiment shall be stored centrally. WormBase represent the ideal candidate for storing this kind of data.

OUTLOOK

Overview of the WormJam Reconstruction

While the current WormJam reconstruction is a good consensus between existing efforts, and the most accurate *C. elegans* metabolic model to date, the task is by no means completed. The accuracy of the metabolic reconstructions is not uniform across biochemistry. For example, the metabolism of carbohydrates and peptides is well-described, but the coverage of nucleic acids and lipids can be improved. The number of different compartments is limited, and in particular some organelles featuring very specific biochemistry, such as lysosomes and peroxisomes, are not explicitly represented. Moreover, *C. elegans* is a multi-cellular organism interacting with a complex environment, and its cellular metabolism cannot be understood in isolation.

Integration With Diet and Omics Data

The typical cultivation of *C. elegans* involves maintenance on three possible Escherichia coli strains OP50, HB101, and HT115. Dietary (bacterial/axenic) and nutritional changes have been known to alter health, lifespan (Fontana and Partridge, 2015) and drug efficacy in nematodes. For example, several recent studies in *C. elegans* have shown that nematodes fed with different bacterial strains respond differently to cancer drugs (García-González et al., 2017; Scott et al., 2017). Another recent study showed that serotonin-increased feeding resulted in increased synthesis of aging-related proteins (Gomez-Amaro et al., 2015). In contrast, in its natural environment, *C. elegans* harbors a rich community of bacterial commensals that are distinct from its environmental microbiota indicating a certain level of host selection of bacteria (Dirksen et al., 2016; Schulenburg and Félix, 2017; Zhang et al., 2017). Similar to the microbiota in higher animals, bacteria can have fitness-modulating effects on the nematode including improved growth compared to cultivation on *E. coli* OP50 (Dirksen et al., 2016; Samuel et al., 2016) and a protection against pathogens (Montalvo-Katz et al., 2013). Based on genomic sequences, we have recently started to reconstruct genome-scale metabolic networks of more than hundred bacterial strains found in association with *C. elegans* (Obeng et al., in preparation). Thus, integration of diet is of particular relevance with respect to the utilization of *C. elegans* on *E. coli* OP50 as a model system for microbiome host-interactions in the context of the influence of the microbiome on the efficacy of pharmaceutic therapies (Cabreiro et al., 2013; Scott et al., 2017). We expect that this resource along with the consensus reconstruction of the *C. elegans* metabolism presented in this work and easy experimental amenability will provide a central cornerstone for the establishment of *C. elegans* as a workhorse in the study of host-microbiome interactions.

Biological variability is imminent when an organism is partitioned into different contexts such as environment, mutation, or cell-type/tissue/organ/life stage. Omics data capture these variations in a gamut of biological molecules (Narayan et al., 2016; Cao et al., 2017; Gao A. W. et al., 2017) and thus can be used to generate context-specific models. One of the previous metabolic reconstructions (Gebauer et al., 2016), integrated transcriptomic data to study aging using two previously published algorithms (Shlomi et al., 2008; Zur et al., 2010; Jensen and Papin, 2011). However, there are an array of different algorithms which include but are not limited to mCADRE (Wang et al., 2012), FASTCORE (Vlassis et al., 2014; Pacheco et al., 2016), INIT (Agren et al., 2012), and CORDA (Schultz and Qutub, 2016). The general methodology of these algorithms involves applying constraints on the metabolic network by binarizing the omics data (Richelle et al., 2018). Though there is no clear verdict on which algorithm is most accurate representation of context-specificity (Shlomi et al., 2008; Jensen and Papin, 2011), the need to include a diverse set of omics data to construct context-specific models is apparent. The WormJam model is being curated by taking into account the list of metabolites that have been detected in *C. elegans*, thus facilitating incorporation of context-specific metabolomics data to provide systemic insights.

Expanding the Breadth of Understanding of the *C. elegans* Metabolome

The refinement of a community network reconstruction of *C. elegans* now provides a context to more comprehensively map out the metabolome. Several further developments will allow us to further grasp this in its full detail. First, the field should strive for better and uniform annotation of metabolites. Efforts such as the WormJam curation will help in this regard, but the actual problem will only be solved once annotation pipelines

that are developed in individual institutes will be shared, compared and become publicly available. Second, improving metabolomics techniques for *C. elegans* will improve detection and annotation of metabolites. One example is the annotation of isobaric and isomeric compounds. As for now, these compounds are often difficult to separate using the most widely used metabolomics platforms. Combining conventional mass spectrometry techniques with additional separation techniques such as ion mobility and NMR spectroscopy will allow the identification of such compounds, and aid in improving the annotation. Third, validation: the rapid expansion of the technical arsenal for metabolomics in *C. elegans* calls for caution when it comes to validating methods between labs. Although routine in clinical environments, the same standards should be adopted in a basic research environment. Dense data sets such as those emerging from a metabolomics screen—often involving >1,000 unique metabolites— make it challenging to establish the limit of detection/quantification and assay variability, but this should nevertheless be the ambition to move the field forward. Finally, in addition to the steady state metabolomics protocols it will be important to develop tools to investigate the dynamics of metabolism. Since metabolism is highly dynamic and adaptable, the analyses should reflect this to prevent biased data based on only a single snapshot. Stable isotope tracing allows for such interpretation. Exposing *C. elegans* to for instance ^{13}C-labeled metabolic substrates at various time points, in combination with analytical MS, standard NMR; or hyperpolarized NMR spectroscopy (Fan and Lane, 2011; Plainchont et al., 2018) will enable the detection of downstream and intermediate products and reconstruction of the metabolic flux distribution. While this strategy is now widely adopted in mammalian cell-based models, only a few papers describe similar efforts in *C. elegans* (Vergano et al., 2014).

A further important point in the future will be the reporting of confidence in metabolite identification. For our curation we did not include any quality metrics of the metabolite annotation/identification supplied by the authors. However, it will be important to make judgements about the quality of the data. A particular example is found in structures that have not been shown in *C. elegans* before (e.g., creatine and creatinine detected in many different publications). Here identifications might be not correct. Different levels of identification have been proposed by the Metabolomics Standard Initiative (MSI), which should be used to report confidence (Sumner et al., 2007; Haug et al., 2013; Rocca-Serra et al., 2015). Also deposition of data, including identified metabolites and their link to the respective organism, e.g., in Metabolights (Haug et al., 2013), will foster the automatic reconstruction of model organism metabolomes (Salek et al., 2017).

Lipids and Lipid Metabolism

Lipids represent a major bottleneck in metabolic models in general, not only in the *C. elegans* model. Lipids cover a large combinatorial space with different fatty acid combinations possible, which is hard to be represented accurately in a stoichiometric model. In addition, *C. elegans* harbors several peculiarities regarding its lipid metabolism (Witting and Schmitt-Kopplin, 2016). First, it can produce, or acquire from its diet, the full range of saturated, MUFAs, PUFAs as well as odd-branched-chain fatty acids on its own and uses them in all lipid classes. Second, it contains several specific lipid classes like the Maradolipids (Mar) (Penkov et al., 2010), long chain ascarosides and phosphoethanolamine glucosylceramides (PEGCs) (Boland et al., 2017). With the curation of the merged model we have undertaken the first steps toward a better representation. However, major parts are still missing, e.g., better biomass compositions for constraint-based modeling including detailed fatty acid profiles for individual lipid classes. Furthermore, even if detailed biosynthesis pathways for each individual lipid could be integrated into the model, there would still be a huge discrepancy between detailed structure in the model and the annotation and identification capabilities of current lipid analysis methods. With standard lipidomics analysis usually neither the position nor the stereochemistry of double bonds can accurately be identified. Therefore, new ways for modeling, but also for the analysis of lipids need to be investigated to make it possible in the future to incorporate lipidomics data.

CONCLUSION

We have shown that there is discrepancy between the model metabolites and metabolites detected by different metabolomics approaches. Specific metabolites from the WormJam model have not been detected so far. Although several are simply not stable enough or have a high turnover rate will be probably never be detected, others might be not accessible with current approaches to low concentrations. On the other side, many detected and reported metabolites could not be found in the model. This is of major interest, since either the identifications are wrong or the model is incomplete. Several examples shown in this paper highlight that we might actually miss several aspects in the metabolism and biology of *C. elegans*. Based on some of our findings the community as whole can now start to explore new metabolic pathways or links to other aspects and their link to metabolism (e.g., epigenetics).

AUTHOR CONTRIBUTIONS

MWi curated the sphingolipid metabolism, metabolite structures, and metabolites detected in the literature. MWi and FS created an initial draft for the ascaroside biosynthesis pathway. MWa and MvW curated the fatty acid metabolic pathways. AG, MvW, and RH helped with metabolite curation from literature. JPNH created WormCon with the advice of H-JS, PE, and CK, and wrote the relevant manuscript sections on WormCon. H-JS assisted in the development of WormCon, curation of metabolic pathways. PE assisted in the development of WormCon, curation of metabolic pathways and editing the manuscript. CJ, SS, and NL performed the merging of ElegCyc and iCEL1273. JPNH, H-JS, PE, and CK performed the merging of CeCon and ElegCyc. JH, NR, AM, OC, MWi, and NLN performed model validation

and have supervised and overseen the manual curation process. All authors wrote and edited the manuscript.

FUNDING

WormJam workshops were funded by the EU GENiE COST network BM1408 and Funding from the Babraham Institute's BBSRC Knowledge Exchange and Commercialisation grant. MWi is funded by the AddOn Fellowship from the Joachim Herz Stiftung. JH is supported by an MRC doctoral training studentship (MR/J003808/1). OC is supported by the ERC (Noisy Aging 638426) and BBSRC (BB/P013406/1). NLN is supported by BBSRC (BB/P013406/1, BB/P013414/1, BB/P013384/1). NR receives fund from the NIH (R01 GM070923 13). CJ and NL were funded by the Keck Foundation, NIH (R35 GM119850) and the Novo Nordisk Foundation (NNF10CC1016517). H-JS has been supported by strategic initiatives funding from the Deputy Vice-Chancellor (Research) of the University of Queensland (DVCR4480). CK acknowledges support by the Excellence Cluster Inflammation at Interfaces (EXC-306). MWa is supported by the BBSRC (BB/P013384/1). RH is supported by an ERC Starting grant (no. 638290).

ACKNOWLEDGMENTS

We would like to acknowledge all members of the WormJam community for their valuable input on different aspects of the *C. elegans* biology and metabolic pathways and in particular.

Andrea Annibal, Marta Artal-Sanz, Bart Braeckman, Ncharam A. Celestina, Artur Lourenço, Anastasios Koutsoumparis, Povila Norvaisas, Christian Pendleton, Sonia Ravanelli, Britta Spanier, Manusnan Suriyalaksh, Safak Yilmaz, Aleksandra Zecic, and Marian Walhout.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2018.00096/full#supplementary-material

Supplementary Information Table 1 | Metabolites curated from literature.

REFERENCES

Agren, R., Bordel, S., Mardinoglu, A., Pornputtapong, N., Nookaew, I., and Nielsen, J. (2012). Reconstruction of genome-scale active metabolic networks for 69 human cell types and 16 cancer types using INIT. *PLoS Comput. Biol.* 8:e1002518. doi: 10.1371/journal.pcbi.1002518

Aguilaniu, H., Fabrizio, P., and Witting, M. (2016). The role dafachronic acid signaling in development and longevity in *Caenorhabditis elegans*: digging deeper using cutting edge analytical chemistry. *Front. Endocrinol.* 7:12. doi: 10.3389/fendo.2016.00012

Artyukhin, A. B., Zhang, Y. K., Akagi, A. E., Panda, O., Sternberg, P. W., and Schroeder, F. C. (2018). Metabolomic "Dark Matter" dependent on peroxisomal β-Oxidation in *Caenorhabditis elegans*. *J. Am. Chem. Soc.* 140, 2841–2852. doi: 10.1021/jacs.7b11811

Boland, S., Schmidt, U., Zagoriy, V., Sampaio, J. L., Fritsche, R. F., Czerwonka, R., et al. (2017). Phosphorylated glycosphingolipids essential for cholesterol mobilization in *Caenorhabditis elegans*. *Nat. Chem. Biol.* 13:647. doi: 10.1038/nchembio.2347

Büchel, F., Rodriguez, N., Swainston, N., Wrzodek, C., Czauderna, T., Keller, R., et al. (2013). Path2Models: large-scale generation of computational models from biochemical pathway maps. *BMC Syst. Biol.* 7:116. doi: 10.1186/1752-0509-7-116

Butler, J. A., Mishur, R. J., Bhaskaran, S., and Rea, S. L. (2013). A metabolic signature for long life in the *Caenorhabditis elegans* Mit mutants. *Aging Cell* 12, 130–138. doi: 10.1111/acel.12029

Butler, J. A., Ventura, N., Johnson, T. E., and Rea, S. L. (2010). Long-lived mitochondrial (Mit) mutants of *Caenorhabditis elegans* utilize a novel metabolism. *FASEB J.* 24, 4977–4988. doi: 10.1096/fj.10-162941

Cabreiro, F., Au, C., Leung, K. Y., Vergara-Irigaray, N., Cochem,é, H. M., Noori, T., et al. (2013). Metformin retards aging in *C. elegans* by altering microbial folate and methionine metabolism. *Cell* 153, 228–239. doi: 10.1016/j.cell.2013.02.035

Cao, J., Packer, J. S., Ramani, V., Cusanovich, D. A., Huynh, C., Daza, R., et al. (2017). Comprehensive single-cell transcriptional profiling of a multicellular organism. *Science* 357, 661. doi: 10.1126/science.aam8940

Caspi, R., Billington, R., Ferrer, L., Foerster, H., Fulcher, C. A., Keseler, I. M., et al. (2016). The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. *Nucleic Acids Res.* 44, D471–D480. doi: 10.1093/nar/gkv1164

Chelliah, V., Juty, N., Ajmera, I., Ali, R., Dumousseau, M., Glont, M., et al. (2015). BioModels: ten-year anniversary. *Nucleic Acids Res.* 43, D542–D548. doi: 10.1093/nar/gku1181

Chitwood, D., Lusby, W. R., Thompson, M. J., Kochansky, J. P., and Howarth, O. W. (1995). The glycosylceramides of the nematode*Caenorhabditis elegans* contain an unusual, branched-chain sphingoid base. *Lipids* 30, 567–573. doi: 10.1007/BF02537032

Depuydt, G., Xie, F., Petyuk, V. A., Shanmugam, N., Smolders, A., Dhondt, I., et al. (2013). Reduced insulin/IGF-1 signaling and dietary restriction inhibit translation but preserve muscle mass in *Caenorhabditis elegans*. *Mol. Cell. Proteomics* 12, 3624–3639. doi: 10.1074/mcp.M113.027383

Dirksen, P., Marsh, S. A., Braker, I., Heitland, N., Wagner, S., Nakad, R., et al. (2016). The native microbiome of the nematode *Caenorhabditis elegans*: gateway to a new host-microbiome model. *BMC Biol.* 14:38. doi: 10.1186/s12915-016-0258-1

Fahy, E., Subramaniam, S., Murphy, R. C., Nishijima, M., Raetz, C. R., Shimizu, T., et al. (2009). Update of the LIPID MAPS comprehensive classification system for lipids. *J. Lipid Res.* 50, S9–S14. doi: 10.1194/jlr.R800095-JLR200

Fan, T. W.-M., and Lane, A. N. (2011). NMR-based stable isotope resolved metabolomics in systems biochemistry. *J. Biomol. NMR* 49, 267–280. doi: 10.1007/s10858-011-9484-6

Feng, L., Shou, Q., and Rebecca Butcher, A. (2016). Identification of a dTDP-rhamnose biosynthetic pathway that oscillates with the molting cycle in *Caenorhabditis elegans*. *Biochem. J.* 473, 1507–1521. doi: 10.1042/BCJ20160142

Finkel, T. (2015). The metabolic regulation of aging. *Nat. Med.* 21:1416. doi: 10.1038/nm.3998

Fontana, L., and Partridge, L. (2015). Promoting health and longevity through diet: from model organisms to humans. *Cell* 161, 106–118. doi: 10.1016/j.cell.2015.02.020

Gao, A., Sterken, M., de Bos, J., Van Creij, J., Kamble, R., Snoek, B., Kammenga, J., et al. (2017). Natural genetic variation in *C. elegans* reveals genomic loci controlling metabolite levels. *bioRxiv*. doi: 10.1101/217729

Gao, A. W., Chatzispyrou, I. A., Kamble, R., Liu, Y. J., Herzog, K., Smith, R. L., et al. (2017). A sensitive mass spectrometry platform identifies metabolic changes of life history traits in *C. elegans*. *Sci. Rep.* 7:2408. doi: 10.1038/s41598-017-02539-w

García-González, A. P., Ritter, A. D., Shrestha, S., Andersen, E. C., Yilmaz, L. S., and Walhout, A. J. M. (2017). Bacterial metabolism affects the *C. elegans* response to cancer chemotherapeutics. *Cell* 169, 431–441.e8. doi: 10.1016/j.cell.2017.03.046

Gebauer, J., Gentsch, C., Mansfeld, J., Schmeißer, K., Waschina, S., Brandes, S., et al. (2016). A genome-scale database and reconstruction of *Caenorhabditis elegans* metabolism. *Cell Syst.* 2, 312–322. doi: 10.1016/j.cels.2016.04.017

Gomez-Amaro, R. L., Valentine, E. R., Carretero, M., LeBoeuf, S. E., Rangaraju, S., Broaddus, C. D., et al. (2015). Measuring food intake

and nutrient absorption in *Caenorhabditis elegans*. *Genetics* 200, 443–454. doi: 10.1534/genetics.115.175851

Greer, E. L., Blanco, M. A., Gu, L., Sendinc, E., Liu, J., Aristizábal-Corrales, D., et al. (2015). DNA Methylation on N^6-Adenine in *C. elegans*. *Cell* 161, 868–878. doi: 10.1016/j.cell.2015.04.005

Hannich, J. T., Mellal, D., Feng, S., Zumbuehl, A., and Riezman, H. (2017). Structure and conserved function of iso-branched sphingoid bases from the nematode *Caenorhabditis elegans*. *Chem. Sci.* 8, 3676–3686. doi: 10.1039/C6SC04831E

Hastings, J., de Matos, P., Dekker, A., Ennis, M., Harsha, B., Kale, N., et al. (2013). The ChEBI reference database and ontology for biologically relevant chemistry: enhancements for 2013. *Nucleic Acids Res.* 41, D456–D463. doi: 10.1093/nar/gks1146

Hastings, J., Mains, A., Artal-Sanz, M., Bergmann, S., Braeckman, B. P., Bundy, J., et al. (2017). WormJam: a consensus C. *elegans* Metabolic Reconstruction and Metabolomics Community and Workshop Series. *Worm* 6:e1373939. doi: 10.1080/21624054.2017.1373939

Haug, K., Salek, R. M., Conesa, P., Hastings, J., de Matos, P., Rijnbeek, M., et al. (2013). MetaboLights—an open-access general-purpose repository for metabolomics studies and associated meta-data. *Nucleic Acids Res.* 41, D781–D786. doi: 10.1093/nar/gks1004

Hucka, M., Bergmann, F. T., Dräger, A., Hoops, S., Keating, S. M., Le Novère, N., et al. (2015). Systems Biology Markup Language (SBML) Level 2 Version 5: structures and facilities for model definitions. *J. Integr. Bioinformatics* 12, 271–271. doi: 10.2390/biecoll-jib-2015-271

Hucka, M., Bergmann, F. T., Dräger, A., Hoops, S., Keating, S. M., Le Novère, N., et al. (2018). The Systems Biology Markup Language (SBML): language specification for level 3 version 2 core. J. *Integr. Bioinform.* 15:20170081. doi: 10.1515/jib-2017-0081

Hucka, M., Finney, A., Sauro, H. M., Bolouri, H., Doyle, J. C., Kitano, H., et al. (2003). The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. *Bioinformatics* 19, 524–531. doi: 10.1093/bioinformatics/btg015

Izrayelit, Y., Srinivasan, J., Campbell, S. L., Jo, Y., von Reuss, S. H., Genoff, M. C., et al. (2012). Targeted metabolomics reveals a male pheromone and sex-specific ascaroside biosynthesis in *Caenorhabditis elegans*. *ACS Chem. Biol.* 7, 1321–1325. doi: 10.1021/cb300169c

Jensen, P. A., and Papin, J. A. (2011). Functional integration of a metabolic network model and expression data without arbitrary thresholding. *Bioinformatics* 27, 541–547. doi: 10.1093/bioinformatics/btq702

Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y., and Morishima, K. (2017). KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res.* 45, D353–D361. doi: 10.1093/nar/gkw1092

Karp, P. D., Latendresse, M., Paley, S. M., Krummenacker, M., Ong, Q. D., Billington, R., et al. (2016). Pathway Tools version 19.0 update: software for pathway/genome informatics and systems biology. *Brief. Bioinformatics* 17, 877–890. doi: 10.1093/bib/bbv079

Kim, S., Thiessen, P. A., Thiessen, E. E., Chen, J., Fu, G., Gindulyte, A., Han, L., et al. (2016). PubChem Substance and Compound databases. *Nucleic Acids Res.* 44, D1202–D1213. doi: 10.1093/nar/gkv951

King, Z. A., Lu, J., Dräger, A., Miller, P., Federowicz, S., Lerman, J. A., et al. (2016). BiGG Models: a platform for integrating, standardizing and sharing genome-scale models. *Nucleic Acids Res.* 44, D515–D522. doi: 10.1093/nar/gkv1049

Lee, R. Y. N., Howe, K. L., Harris, T. W., Arnaboldi, V., Cain, S., Chan, J., et al. (2018). WormBase 2017: molting into a new stage. *Nucleic Acids Res.* 46, D869–D874. doi: 10.1093/nar/gkx998

Lubitz, T., Hahn, J., Bergmann, F. T., Noor, E., Klipp, E., and Liebermeister, W. (2016). SBtab: a flexible table format for data exchange in systems biology. *Bioinformatics* 32, 2559–2561. doi: 10.1093/bioinformatics/btw179

Ma, L., Chan, A. H. C., Hattwell, N. P., Ebert, P. R., and Schirra, H. J. (2017). Systems biology analysis using a genome-scale metabolic model shows that phosphine triggers global metabolic suppression in a resistant strain of *C. elegans*. *bioRxiv [Preprint]*. doi: 10.1101/144386

Maglioni, S., and Ventura, N. (2016). *C. elegans* as a model organism for human mitochondrial associated disorders. *Mitochondrion* 30, 117–125. doi: 10.1016/j.mito.2016.02.003

Mahanti, P., Bose, N., Bethke, A., Judkins, J. C., Wollam, J., Dumas, K. J., et al. (2014). Comparative metabolomics reveals endogenous ligands of DAF-12, a nuclear hormone receptor, regulating *C. elegans* development and lifespan. *Cell Metab.* 19, 73–83. doi: 10.1016/j.cmet.2013.11.024

Montalvo-Katz, S., Huang, H., Appel, M. D., Berg, M., and Shapira, M. (2013). Association with soil bacteria enhances p38-dependent infection resistance in *Caenorhabditis elegans*. *Infect. Immun.* 81, 514–520. doi: 10.1128/IAI.00653-12

Moretti, S., Martin, O., Van Du Tran, T., Bridge, A., Morgat, A., and Pagni, M. (2016). MetaNetX/MNXref – reconciliation of metabolites and biochemical reactions to bring together genome-scale metabolic networks. *Nucleic Acids Res.* 44, D523–D526. doi: 10.1093/nar/gkv1117

Narayan, V., Ly, T., Pourkarimi, E., Murillo, A. B., Gartner, A., Lamond, A. I., et al. (2016). Deep proteome analysis identifies age-related processes in *C. elegans*. *Cell Syst.* 3, 144–159. doi: 10.1016/j.cels.2016.06.011

Nobeli, I., Ponstingl, H., Krissinel, E. B., and Thornton, J. M. (2003). A Structure-based anatomy of the *E. coli* metabolome. *J. Mol. Biol.* 334, 697–719. doi: 10.1016/j.jmb.2003.10.008

O'Brien, E. J., Monk, J. M., and Palsson, B. Ø. (2015). Using genome-scale models to predict biological capabilities. *Cell* 161, 971–987. doi: 10.1016/j.cell.2015.05.019

Olivier Brett, G., and Bergmann Frank, T. (2018). SBML Level 3 Package: Flux Balance Constraints version 2. *J. Integr. Bioinformatics*. 15. doi: 10.1515/jib-2017-0082

Orth, J. D., Conrad, T. M., Na, J., Lerman, J. A., Nam, H., Feist, A. M., et al. (2011). A comprehensive genome-scale reconstruction of *Escherichia coli* metabolism-2011. *Mol. Syst. Biol.* 7:535. doi: 10.1038/msb.2011.65

Pacheco, M. P., Pfau, T., and Sauter, T. (2016). Benchmarking procedures for high-throughput context specific reconstruction algorithms. *Front. Physiol.* 6:410. doi: 10.3389/fphys.2015.00410

Panda, O., Akagi, A. E., Artyukhin, A. B., Judkins, J. C., Le, H. H., Mahanti, P., et al. (2017). Biosynthesis of modular ascarosides in *C. elegans*. *Angew. Chem. Int. Edition* 56, 4729–4733. doi: 10.1002/anie.201700103

Pence, H. E., and Williams, A. (2010). ChemSpider: an online chemical information resource. *J. Chem. Educ.* 87, 1123–1124. doi: 10.1021/ed100697w

Penkov, S., Mende, F., Zagoriy, V., Erkut, C., Martin, R., Pässler, U., et al. (2010). Maradolipids: diacyltrehalose glycolipids specific to dauer larva in *Caenorhabditis elegans*. *Angew. Chem. Int. Edition* 49, 9430–9435. doi: 10.1002/anie.201004466

Perez, C. L., and Van Gilst, M. R. (2008). A ^{13}C isotope labeling strategy reveals the influence of insulin signaling on lipogenesis in *C. elegans*. *Cell Metab.* 8, 266–274. doi: 10.1016/j.cmet.2008.08.007

Plainchont, B., Berruyer, P., Dumez, J. N., Jannin, S., and Giraudeau, P. (2018). Dynamic nuclear polarization opens new perspectives for NMR spectroscopy in analytical chemistry. *Anal. Chem.* 90, 3639–3650. doi: 10.1021/acs.analchem.7b05236

Richelle, A., Joshi, C., and Lewis, N. E. (2018). Assessing key decisions for transcriptomic data integration in biochemical networks. *bioRxiv [Preprint]*. doi: 10.1101/301945

Roach, P. J., Depaoli-Roach, A. A., Hurley, T. D., and Tagliabracci, V. S. (2012). Glycogen and its metabolism: some new developments and old themes. *Biochem. J.* 441, 763–787. doi: 10.1042/BJ20111416

Rocca-Serra, P., Salek, R. M., Arita, M., Correa, E., Dayalan, S., Gonzalez-Beltran, A., et al. (2015). Data standards can boost metabolomics research, and if there is a will, there is a way. *Metabolomics* 12:14. doi: 10.1007/s11306-015-0879-3

Salek, R. M., Conesa, P., Cochrane, K., Haug, K., Williams, M., Kale, N., et al. (2017). Automated assembly of species metabolomes through data submission into a public repository. *Gigascience* 6, 1–4. doi: 10.1093/gigascience/gix062

Samuel, B. S., Rowedder, H., Braendle, C., Félix, M. A., and Ruvkun, G. (2016). *Caenorhabditis elegans* responses to bacteria from its natural habitats. *Proc. Natl. Acad. Sci.* 113, E3941–E3949. doi: 10.1073/pnas.1607183113

Schosserer, M., Minois, N., Angerer, T. B., Amring, M., Dellago, H., Harreither, E., et al. (2015). Methylation of ribosomal RNA by NSUN5 is a conserved mechanism modulating organismal lifespan. *Nat. Commun.* 6:6158. doi: 10.1038/ncomms7158

Schulenburg, H., and Félix, A. M. (2017). The natural biotic environment of *Caenorhabditis elegans*. *Genetics* 206, 55–86. doi: 10.1534/genetics.116.195511

Schultz, A., and Qutub, A. A. (2016). Reconstruction of tissue-specific metabolic networks using CORDA. *PLoS Comput. Biol.* 12:e1004808. doi: 10.1371/journal.pcbi.1004808

Scott, T. A., Quintaneiro, L. M., Norvaisas, P., Lui, P. P., Wilson, M. P., Leung, K. Y., et al. (2017). Host-microbe co-metabolism dictates cancer drug efficacy in *C. elegans*. *Cell* 169, 442–456.e18. doi: 10.1016/j.cell.2017.03.040

Seo, Y., Kingsley, S., Walker, G., Mondoux, M. A., and Tissenbaum, H. A. (2018). Metabolic shift from glycogen to trehalose promotes lifespan and healthspan in *Caenorhabditis elegans*. *Proc. Natl. Acad. Sci. U.S.A.* 115, E2791–E2800. doi: 10.1073/pnas.1714178115

Shen, P., Yue, Y., and Park, Y. (2018). A living model for obesity and aging research: *Caenorhabditis elegans*. *Crit. Rev. Food Sci. Nutr.* 58, 741–754. doi: 10.1080/10408398.2016.1220914

Shlomi, T., Cabili, M. N., Herrgård, M. J., Palsson, B. Ø., and Ruppin, E. (2008). Network-based prediction of human tissue-specific metabolism. *Nat. Biotechnol.* 26, 1003-1010. doi: 10.1038/nbt.1487

Srinivasan, J., von Reuss, S. H., Bose, N., Zaslaver, A., Mahanti, P., Ho, M. C., et al. (2012). A modular library of small molecule signals regulates social behaviors in *Caenorhabditis elegans*. *PLoS Biol.* 10:e1001237. doi: 10.1371/journal.pbio.1001237

Sumner, L. W., Amberg, A., Barrett, D., Beale, M. H., Beger, R., Daykin, C. A., et al. (2007). Proposed minimum reporting standards for chemical analysis. *Metabolomics* 3, 211–221. doi: 10.1007/s11306-007-0082-2

Tissenbaum, H. A. (2015). Using C. *elegans* for aging research. *Invert. Reprod. Dev.* 59, 59–63. doi: 10.1080/07924259.2014.940470

van Delft, P., Akay, A., Huber, S. M., Bueschl, C., Rudolph, K. L. M., Di Domenico, T., et al. (2017). The profile and Dynamics of RNA modifications in animals. *Chembiochem* 18, 979–984. doi: 10.1002/cbic.201700093

van Heck, R. G., Ganter, M., Martins Dos Santos, V. A., and Stelling, J. (2016). Efficient reconstruction of predictive consensus metabolic network models. *PLoS Comput. Biol.* 12:e1005085. doi: 10.1371/journal.pcbi.1005085

Vergano, S. S., Rao, M., McCormack, S., Ostrovsky, J., Clarke, C., Preston, J., et al. (2014). *In vivo* metabolic flux profiling with stable isotopes discriminates sites and quantifies effects of mitochondrial dysfunction in *C. elegans*. *Mol. Genet. Metab.* 111, 331–341. doi: 10.1016/j.ymgme.2013.12.011

Vlassis, N., Pacheco, M. P., and Sauter, T. (2014). Fast reconstruction of compact context-specific metabolic network models. *PLoS Comput. Biol.* 10:e1003424. doi: 10.1371/journal.pcbi.1003424

von Reuss, S. H., Bose, N., Srinivasan, J., Yim, J. J., Judkins, J. C., Sternberg, P. W., et al. (2012). Comparative metabolomics reveals biogenesis of ascarosides, a modular library of small-molecule signals in *C. elegans*. *J. Am. Chem. Soc.* 134, 1817–1824. doi: 10.1021/ja210202y

Wang, Y., Eddy, J. A., and Price, N. D. (2012). Reconstruction of genome-scale metabolic models for 126 human tissues using mCADRE. *BMC Syst. Biol.* 6:153. doi: 10.1186/1752-0509-6-153

Wishart, D. S., Feunang, Y. D., Marcu, A., Guo, A. C., Liang, K., Vázquez-Fresno, R., et al. (2018). HMDB 4.0: the human metabolome database for 2018. *Nucleic Acids Res.* 46, D608–D617. doi: 10.1093/nar/gkx1089

Witting, M., and Schmitt-Kopplin, P. (2016). The *Caenorhabditis elegans* lipidome: a primer for lipid analysis in *Caenorhabditis elegans*. *Arch. Biochem. Biophys.* 589, 27–37. doi: 10.1016/j.abb.2015.06.003

Wohlgemuth, G., Mehta, S. S., Mejia, R. F., Neumann, S., Pedrosa, D., Pluskal, T., et al. (2016). SPLASH, A hashed identifier for mass spectra. *Nat. Biotechnol.* 34, 1099–1101. doi: 10.1038/nbt.3689

Yilmaz, L. S., and Walhout, A. J. (2016). *A Caenorhabditis elegans* genome-scale metabolic network model. *Cell Syst.* 2, 297–311. doi: 10.1016/j.cels.2016.04.012

Zhang, F., Berg, M., Dierking, K., Félix, M. A., Shapira, M., Samuel, B. S., et al. (2017). *Caenorhabditis elegans* as a model for microbiome research. *Front. Microbiol.* 8:485. doi: 10.3389/fmicb.2017.00485

Zhou, Y., Wang, Y., Zhang, X., Bhar, S., Jones Lipinski, R. A., Han, J., et al. (2018). Biosynthetic tailoring of existing ascaroside pheromones alters their biological function in *C. elegans*. *eLife* 7:e33286. doi: 10.7554/eLife.33286

Zhu, H., Shen, H., Sewell, A. K., Kniazeva, M., and Han, M. (2013). A novel sphingolipid-TORC1 pathway critically promotes postembryonic development in *Caenorhabditis elegans*, ed. *J. Ahringer*. 2:e00429. doi: 10.7554/eLife.00429

Zur, H., Ruppin, E., and Shlomi, T. (2010). iMAT: an integrative metabolic analysis tool. *Bioinformatics* 26, 3140–3142. doi: 10.1093/bioinformatics/btq602

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Witting, Hastings, Rodriguez, Joshi, Hattwell, Ebert, van Weeghel, Gao, Wakelam, Houtkooper, Mains, Le Novère, Sadykoff, Schroeder, Lewis, Schirra, Kaleta and Casanueva. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.