

<https://doi.org/10.1038/s41540-024-00384-y>

Comparative analysis of metabolic models of microbial communities reconstructed from automated tools and consensus approaches

Yunli Eric Hsieh^{1,2,3}, Kshitij Tandon³, Heroen Verbruggen³ & Zoran Nikoloski^{1,2}✉

Genome-scale metabolic models (GEMs) of microbial communities offer valuable insights into the functional capabilities of their members and facilitate the exploration of microbial interactions. These models are generated using different automated reconstruction tools, each relying on different biochemical databases that may affect the conclusions drawn from the *in silico* analysis. One way to address this problem is to employ a consensus reconstruction method that combines the outcomes of different reconstruction tools. Here, we conducted a comparative analysis of community models reconstructed from three automated tools, i.e. CarveMe, gapseq, and KBase, alongside a consensus approach, utilizing metagenomics data from two marine bacterial communities. Our analysis revealed that these reconstruction approaches, while based on the same genomes, resulted in GEMs with varying numbers of genes and reactions as well as metabolic functionalities, attributed to the different databases employed. Further, our results indicated that the set of exchanged metabolites was more influenced by the reconstruction approach rather than the specific bacterial community investigated. This observation suggests a potential bias in predicting metabolite interactions using community GEMs. We also showed that consensus models encompassed a larger number of reactions and metabolites while concurrently reducing the presence of dead-end metabolites. Therefore, the usage of consensus models allows making full and unbiased use from aggregating genes from the different reconstructions in assessing the functional potential of microbial communities.

Microbe-microbe interactions play a crucial role in maintaining microbial diversity, influence metabolic phenotypes, and shape community functionality^{1,2}. Therefore, microbial communities and interactions are increasingly studied in agriculture³, synthetic biology⁴, pathology⁵, and ecology⁶. Microbial interactions are in part achieved by the exchange of metabolites, and they are particularly challenging to study in wild communities⁷. As a complementary tool, genome-scale metabolic models (GEMs) provide means to identify and dissect the effect of these interactions.

Constraint-based modeling using GEMs has been used to investigate the activity of different reactions in a metabolic network, including exchange reactions that model interactions between microbes. Numerous studies have employed GEMs to investigate metabolic interactions and functionality within microbial communities, including those found in the human gut⁸,

termite gut⁹, mangrove sediments¹⁰, soil microbial communities¹¹, and plant root¹². Community-scale metabolic models are typically constructed using: (i) the mixed-bag approach, which involves integrating all metabolic pathways and transport reactions into a single model with one cytosolic and one extracellular compartment; (ii) compartmentalization, where multiple GEMs are combined into a single stoichiometric matrix, with each species assigned to a distinct compartment; (iii) costless secretion, wherein models are simulated using a dynamically and iteratively updated medium based on exchange reactions and metabolites within the community^{13,14}. The choice of approach depends on the specific objectives and scenarios. The mixed-bag approach is suitable for analyzing interactions between communities, while the other approaches are more appropriate for understanding interactions between organisms within a community¹⁵.

¹Bioinformatics Department, Institute of Biochemistry and Biology, University of Potsdam, Potsdam, Germany. ²Systems Biology and Mathematical Modeling Group, Max Planck Institute of Molecular Plant Physiology, Potsdam, Germany. ³School of BioSciences, The University of Melbourne, Parkville, VIC, Australia.

✉e-mail: Nikoloski@mpimp-golm.mpg.de

Regardless of the approach used, *in silico* analysis of metabolism of individual organisms in a community requires access to reconstructed GEMs for all species in the community. Several automated approaches are available for GEM reconstruction, including: CarveMe¹⁶, gapseq¹⁷, and KBase¹⁸. Mendoza et al.¹⁹ conducted a systematic evaluation of reconstruction tools, revealing that each tool offers distinct features. For example, CarveMe enables fast model generation due to their ready-to-use metabolic networks, while gapseq incorporates comprehensive biochemical information by employing various data sources during reconstruction. However, selecting different tools can lead to the construction of alternative networks, introducing uncertainty in the predictions resulting from the constraint-based modeling with these GEMs²⁰. This uncertainty could be caused by gene annotation, gene-reaction mapping, biomass composition, and environment specification. The structure of the reconstructed network is significantly influenced by the choice of the database of biochemical reactions, and this variation is potentially caused by mis-annotations²¹ and hypothetical sequences of unknown function²². During reconstruction, the inclusion of specific reactions in the model depends on the genomic evidence and the network context, often omitting certain reactions based on the modeling objectives. Furthermore, the use of different namespaces for metabolites and reactions from various data sources can pose challenges when combining GEMs^{11,23}, leading to further difficulties in building consensus model for predicting metabolic phenotypes of microbial communities.

Consensus models, formed by integrating different reconstructed models of single species from various tools, have the potential to reduce the uncertainty existing in a single model^{24,25} and can be used to estimate interactions in a community¹¹. However, a systematic comparison between consensus models and original models in terms of model structure (i.e. the number of reactions and metabolites), the inclusion of genes in the model, model functionality, and the potential exchange of metabolites at the community scale is currently lacking. This is particularly the case for the scenarios in which metabolic models are reconstructed based on metagenome-assembled genomes (MAGs). Here, we conducted a comprehensive analysis of these features for models reconstructed using three automated tools, namely: CarveMe, gapseq, and KBase, and a recently proposed consensus reconstruction¹¹ using data about MAGs. The rationale behind the selection of these specific reconstruction tools includes: (i) user-friendly interfaces and platforms provided, (ii) the generation of immediately functional models that could implement the subsequent constraint-based modeling (e.g. via flux balance analysis), (iii) the use of distinct databases for model reconstruction, and (iv) the distinction between the top-down (CarveMe) and bottom-up reconstruction approaches (gapseq and KBase). The main difference between these approaches lies in the foundational principle: top-down strategies reconstruct models based upon a well-curated, universal template, carving the reactions with annotated

sequences; in contrast, bottom-up strategies construct draft models through the mapping of reactions based on annotated genomic sequences.

Our specific focus was three-fold: (i) investigating whether the iterative order influences gap-filling solutions, (ii) determining if the consensus community model aids in identifying functional characteristics between two marine bacterial communities, and (iii) evaluating if the consensus community model reveals distinct metabolite interactions within the communities. Our findings shed light on the advantages and limitations of each approach, revealing that consensus models retain the majority of unique reactions and metabolites from the original models, while reducing the presence of dead-end metabolites. Furthermore, consensus models incorporate a greater number of genes, indicating stronger genomic evidence support for the reactions. These characteristics of consensus models demonstrate their enhanced functional capability and capacity for more comprehensive metabolic network models in a community context.

Results

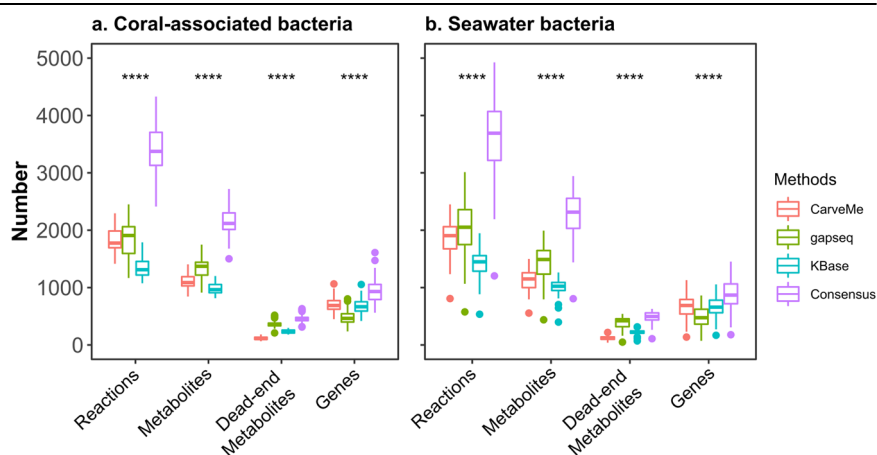
Structural differences in genome-scale metabolic models from two bacterial communities

We utilized a collection of 105 high-quality MAGs derived from coral-associated and seawater bacterial communities described in ref. 26 to construct genome-scale metabolic models. GEM reconstruction used three automated approaches: CarveMe¹⁶, gapseq¹⁷ and KBase¹⁸, to generate draft models, of which the effect of gapseq GEMs on the consensus model generation has not yet been investigated¹¹. Draft models originating from the same MAG were merged to construct draft consensus models by using a recently proposed pipeline¹¹, which was tested with data from species-resolved operational taxonomic units (OTUs) (for which genomes were available) rather than with data about MAGs. Gap-filling of the draft community models was performed using COMMIT¹¹ (see Methods). We note that while metabolic models from species-resolved OTUs provide a gold standard for model comparison, this is not possible to achieve for MAGs which are considered a mixture of organisms.

To compare the structural characteristics of the community models, we examined the number of reactions, metabolites, dead-end metabolites, and genes in the resulting reconstructions (Fig. 1). Genes serve as the fundamental components of GEMs. Inclusion of a gene in the model indicates its association with at least one biochemical reaction, thus affecting the set of metabolites in the models. Our analysis revealed that CarveMe models exhibited the highest number of genes, followed by KBase and gapseq in models of both coral-associated bacterial and seawater bacterial communities. Additionally, gapseq models encompassed more reactions and metabolites compared to CarveMe and KBase models, potentially indicating that many genes in gapseq models are associated with multiple reactions. However, gapseq models also exhibited a larger number of dead-end metabolites, which may affect the functional characteristics of the models.

Fig. 1 | Structural comparison of metabolic models reconstructed using four different approaches.

The metabolic models reconstructed by four approaches, including: CarveMe¹⁶, gapseq¹⁷, KBase¹⁸, and the consensus method used in COMMIT¹¹, were evaluated based on the number of reactions, metabolites, dead-end metabolites, and genes. Statistical analysis was conducted using the Kruskal-Wallis test (**** $p < 0.0001$) to determine significant differences of these characteristics between methods. **a** Metabolic models of 50 coral-associated bacteria. **b** Metabolic models of 55 seawater bacteria, based on MAGs from Robbins et al.²⁶. Each color represents a distinct reconstruction approach, as specified in the legend.



We note that the presence of dead-end metabolites is attributed to gaps in our understanding of the metabolic network and could potentially serve other functions when the organism is modeled jointly with other community members. Therefore, the number of dead-end metabolites does not necessarily imply model inconsistency, but could impact the current functionality of the model.

To assess the similarity of community reconstructions obtained through different approaches, we computed the Jaccard similarity for the sets of reactions, metabolites, dead-end metabolites, and genes in the models derived from the same MAGs (Fig. 2). Our findings revealed that despite being reconstructed from the same MAG, distinct reconstruction approaches yielded markedly different results. The results demonstrated a relatively low similarity between the respective sets resulting from the compared approaches. Specifically, in terms of the overall characteristics, gapseq and KBase models exhibited higher similarity in the composition of reactions and metabolites compared to CarveMe models. On average, the Jaccard similarity for reactions in coral-associated bacteria and seawater bacteria models was 0.23 and 0.24, respectively, while the Jaccard similarity for metabolites was 0.37 for models of both coral-associated and seawater bacterial communities. This observation suggests that the similarity between gapseq and KBase models may be attributed to their shared usage of the ModelSEED database for reconstruction, resulting in a relatively consistent set of reactions and metabolites within the models. However, in terms of gene composition, CarveMe and KBase models exhibited a higher degree of similarity compared to gapseq models. The average Jaccard similarity of the gene sets of coral-associated bacteria and seawater bacteria models was 0.42 and 0.45, respectively. Notably, we found a higher similarity between CarveMe and consensus models, with values of 0.75 and 0.77 for coral-associated bacteria and seawater bacteria models, respectively. This further indicated that the majority of genes included in the consensus models are due to their inclusion in the CarveMe models.

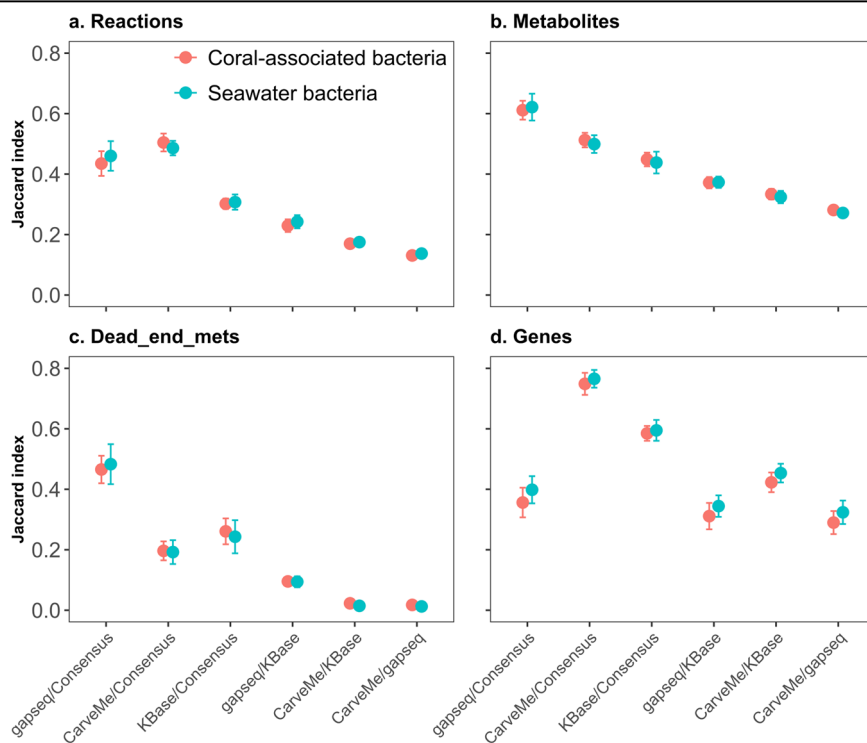
The effects of iterative order on the reconstructed network

During the gap-filling process of the consensus models, we employed an iterative approach based on MAG abundance to specify the ascending/descending order of inclusion of a MAG in the gap-filling step of COMMIT. The

process was initiated with a minimal medium, and after each gap-filling step of single model, permeable metabolites were predicted and used to augment the current medium. These metabolites were then incorporated into subsequent reconstructions by introducing additional uptake reactions in the gap-filling database. To investigate whether the order had an impact on the resulting gap-filling solutions, we conducted an analysis to assess the association between MAG abundance and the obtained solutions. Our findings indicated that the iterative order did not have a significant influence on the number of added reactions in the two communities reconstructed using the four different approaches (Fig. 3a–d, Supplementary Fig. 1a–d, Supplementary Fig. 2a–d, and Supplementary Fig. 3a–d). The results demonstrated that the number of added reactions and abundance of MAGs exhibited only a negligible correlation ($r = 0–0.3$). In addition, although gapseq models exhibited a higher number of reactions compared to CarveMe and KBase models, a considerable number of reactions without genetic support needed to be added to enable simulation of growth with gapseq models (Fig. 4 and Supplementary Fig. 4). This divergence is likely due to distinct reconstruction algorithms employed in draft reconstruction, leading to variations in the number of reactions added during gap-filling. In contrast, the consensus approach demonstrated its ability to significantly reduce the number of required gap-filling solutions, thus minimizing the inclusion of such reactions without genetic support that are necessary for growth simulation.

With regards to the number of imported/exported metabolites (Fig. 3e–l, Supplementary Fig. 1e–l, Supplementary Fig. 2e–l, and Supplementary Fig. 3e–l), the effect of MAG abundance in the order of iterative inclusion varied across different reconstruction approaches, with notable effects observed in the gapseq and KBase models. In CarveMe and consensus approaches, the MAG abundance order did not demonstrate a significant effect on the number of imported/exported metabolites. In contrast, for the KBase models we identified a high negative correlation ($r = -0.7$ to -0.9) between MAG abundance and the number of exported/imported metabolites ($r = -0.76$ and -0.73 for imported/exported metabolites, respectively). In the gapseq models, we found a low negative correlation ($r = -0.3$ to -0.5) between abundance and imported metabolites, while a moderate negative correlation ($r = -0.5$ to -0.7) existed between abundance and the number of exported metabolites. However, when considering the increasing

Fig. 2 | Analysis of similarity of community models derived from different reconstruction approaches. The Jaccard similarity was employed to assess the similarity between each reconstruction, considering: (a) the sets of reactions, (b) metabolites, (c) dead-end metabolites, and (d) genes. Pairwise comparisons were performed among the models reconstructed from the same MAG using different approaches. This comparison was performed on the same models whose characteristics were compared in Fig. 1. The coral-associated bacterial models are represented in red, while the seawater bacterial models are depicted in light blue.



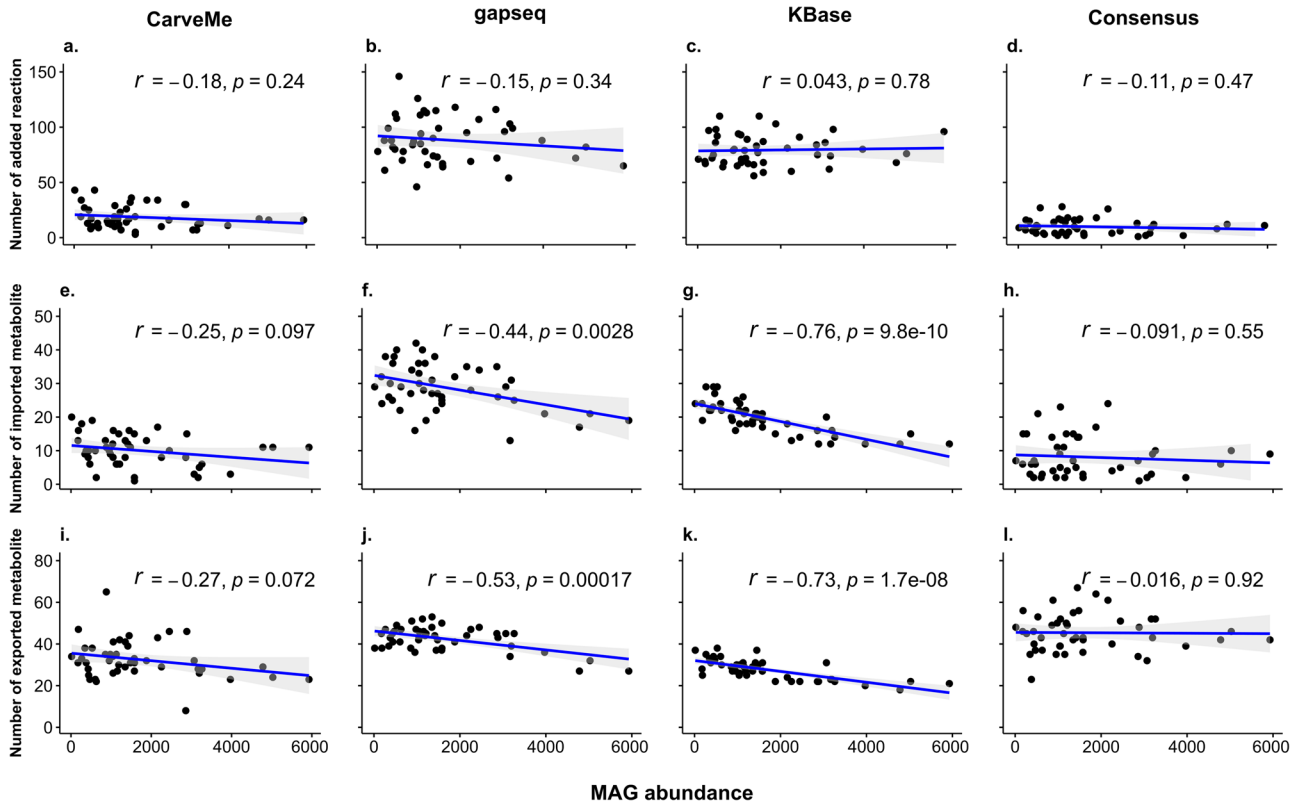
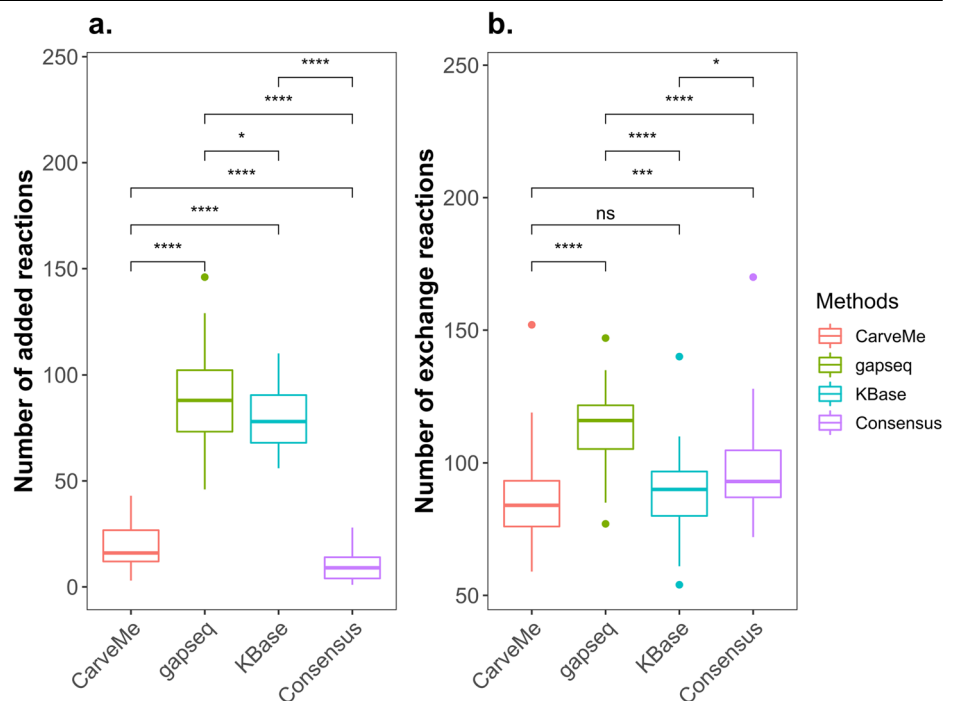


Fig. 3 | Association between MAG abundance and gap-filling results with a descending order in different reconstructions of coral-associated bacterial community model. Pearson correlation coefficient was employed to evaluate the association between MAG abundance and the number of added reactions (a–d),

imported metabolites (e–h), and exported metabolites (i–l), for each of the four reconstruction approaches: CarveMe¹⁶, gapseq¹⁷, KBase¹⁸, and the consensus method used in COMMIT¹¹. The correlation coefficient (*r*) and corresponding *p*-value (*p*) were determined.

Fig. 4 | Comparison of functional models in different reconstructions of the coral-associated bacterial community model. The size of gap-filling solutions and the number of exchange reactions in functional models, that can simulate growth, were compared using the Wilcoxon Rank test (**p* < 0.05; ****p* < 0.001; *****p* < 0.0001; ns *p* > 0.05). Panels a and b represent the size of gap-filling solutions and the number of exchange reactions, respectively.



order of MAG abundance in KBase and gapseq models, the outcomes were reversed, demonstrating a positive correlation between MAG abundance and imported/exported metabolites. Regardless of the iterative order, it was noted that the starting model had a lower number of exchanged metabolites,

while the ending model exhibited a higher number of exchanged metabolites in KBase and gapseq models. These findings suggest effects of reconstruction tools as well as abundance of MAGs on the exchange metabolites in the resulting model.

The quality assessment of functional models

Next, we performed an evaluation of the model quality using the MEMOTE suite of indices (Fig. 5). A higher score within this evaluation indicates better model quality according to the specified indices. The consistency index encompasses assessments of stoichiometric, mass, and charge balance of reactions, as well as metabolite connectivity and unbounded flux within the default medium. Notably, we stress the unbounded flux in the default medium index, as it elucidates the extent to which reactions can carry unlimited flux. This issue often arises due to problems with reaction directionality, missing cofactors, and/or inaccurately defined transport reactions²⁷. A higher score in this index correlates with a reduced number of reactions carrying unlimited flux. Another index we investigated is the reaction annotation index, which evaluates how many reactions in the model are annotated with associated enzyme commission numbers (EC numbers).

We observed that the significant reduction in the total score was primarily attributed to the absence of reaction, metabolite, and gene annotations from databases other than MetaNetX. Regarding the individual scores, we found that KBase obtained the highest average score (62%) for the reaction annotation in the coral-associated bacteria models, while gapseq achieved the highest average score (67%) in the seawater bacteria models. Conversely, CarveMe exhibited the lowest score (54%) in both the coral-associated bacteria and seawater bacteria models. The reaction annotation score of consensus models obtained a medium score in comparison to other approaches. Notably, spontaneous and transport reactions commonly lack an associated EC number, as enzymatic catalysis is not required for these processes. Consequently, the inclusion of such reactions from other models into the consensus models expectedly diluted the proportion of reactions annotated with an EC number. Interestingly, we found considerable variation in each score within the same reconstruction approach, indicating substantial differences in model quality. However, the consensus model demonstrated a noteworthy reduction in the variability of index values across models in comparison to the other approaches.

Functional enrichment in different reconstructions

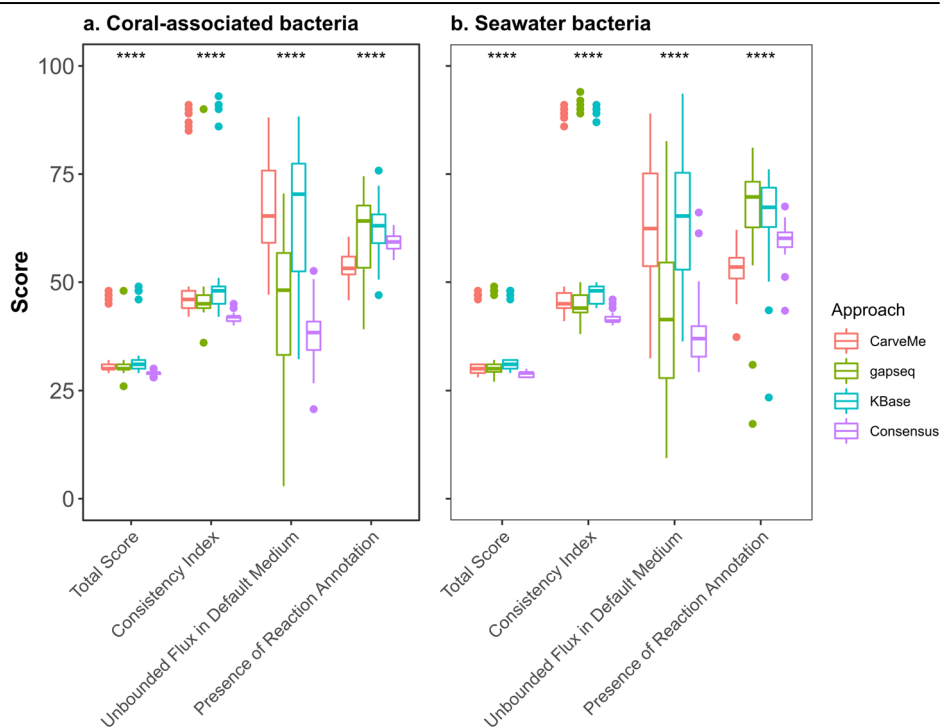
EC numbers provide the means to assess the enzyme functions included in a model in an automated fashion²⁸. For instance, enriched EC numbers can

serve as an indicator of enriched function of metabolic reaction in a metabolic network.

To investigate the enriched functions in the reconstructed models, we performed a comparison of enriched EC numbers for the shared reactions and unblocked shared reactions in the models resulting from the compared approaches (Fig. 6a, Supplementary Fig. 5a). The unblocked shared reactions were identified by performing flux variability analysis (FVA) among all shared reactions between the models. Although gapseq and KBase models exhibit relatively similar sets of reactions, our enrichment analysis revealed distinct enriched functions between these two approaches in terms of shared and unblocked shared reactions. For example, in the shared reactions within gapseq and KBase models, we observed an enrichment of functions related to acyltransferases and carbon-carbon lyases. However, after filtering blocked shared reactions, we found that glycosyltransferases and the enzymes involved in transferring nitrogenous groups and transferring phosphorus containing groups to be enriched. This discrepancy suggests that certain shared reactions in the gapseq and KBase models may not carry flux, thereby contributing to the observed differences.

Conversely, we observed a higher degree of consistency in the enriched functions associated with shared and unblocked shared reactions in CarveMe/gapseq and CarveMe/KBase models. These consistent functions primarily encompassed activities related to carbon-oxygen lyases, glycosyltransferases, and the enzymes involved in transferring nitrogenous groups and transferring phosphorus containing groups. Overall, we found that CarveMe models displayed a greater diversity of enriched functions compared to gapseq and KBase models (Fig. 6b, Supplementary Fig. 5b). Through the integration of reactions from different reconstruction approaches, consensus models presented more comprehensive and less biased metabolic networks that are expected to affect the EC enrichment analyses. Indeed, the results of consensus models displayed more specific enriched functions. Predominantly enriched functions within both bacterial communities were associated with carbon-oxygen lyases and oxidoreductases, specifically those involved in acting on CH-OH and CH-CH group donors. This observation underscores the potential of consensus models to provide a more precise representation of the functional characteristics in bacterial community models. Overall, our results indicated that the seawater bacterial community displayed a higher diversity of enriched

Fig. 5 | Quality assessment using MEMOTE. To assess the quality of models reconstructed from different approaches we used MEMOTE. Statistical analysis was conducted using the Kruskal-Wallis test ($****p < 0.0001$) to determine significant differences of each score between methods. **a** Metabolic models of 50 coral-associated bacteria. **b** Metabolic models of 55 seawater bacteria. Each color represents a distinct reconstruction approach, as specified in the legend.



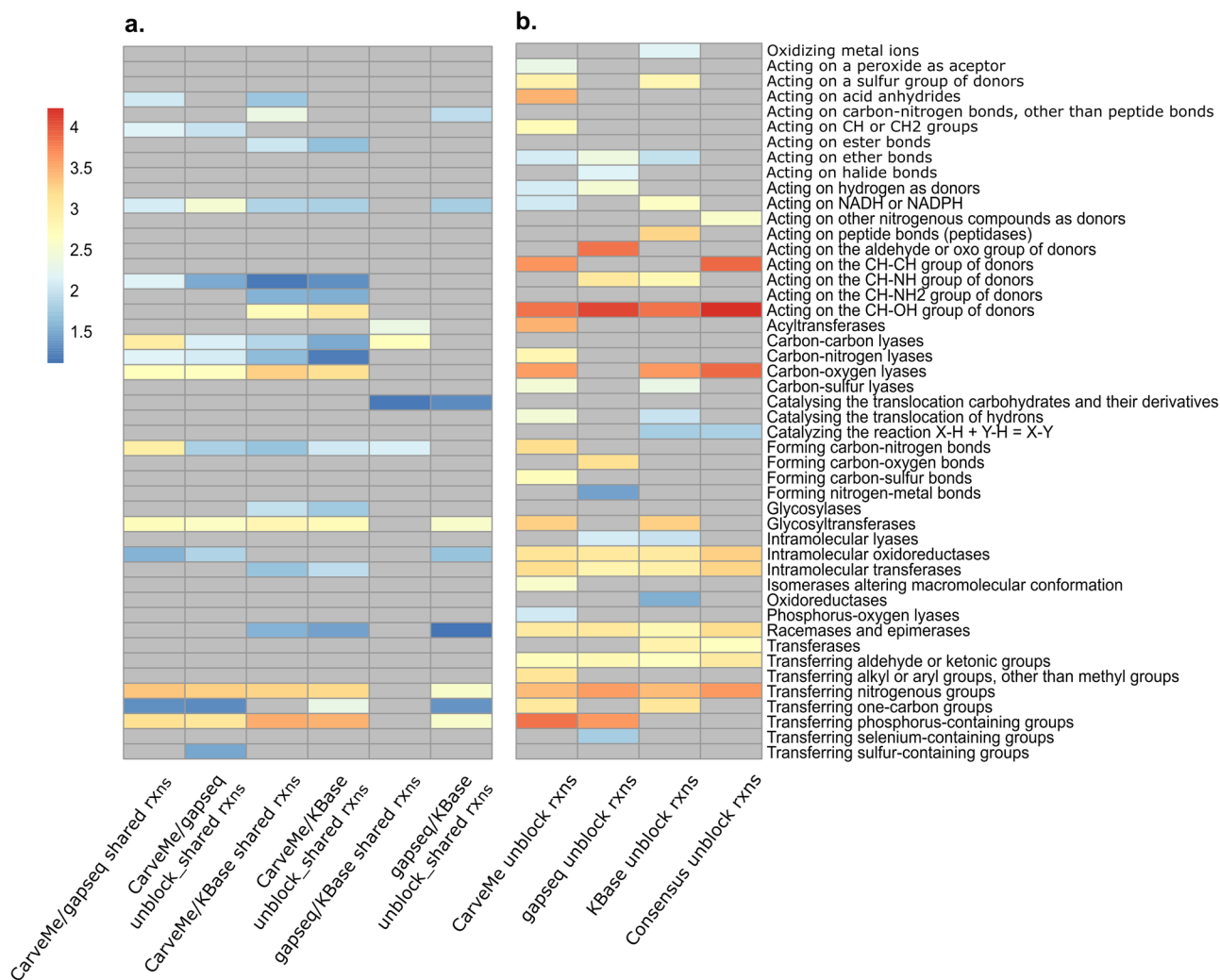


Fig. 6 | Enriched enzyme subclasses in the coral-associated bacterial community model from different reconstructions. The pairwise comparison of enriched enzyme subclasses in (a). shared reactions between each reconstruction and (b). In the community models reconstructed by different approaches, analyzed using the

hypergeometric test. The abundance of enzyme subclasses is represented in a logarithmic scale and depicted using a color scale ranging from blue to red, with higher numbers indicating greater abundance. Grey color indicates the absence of enriched enzyme subclasses.

functions (13 enriched functions) compared to the coral-associated bacteria community (11 enriched functions).

Exchanged metabolites in different reconstructions under community setting

We considered the presence of exchanged metabolites in the community models as a potential indicator of metabolite interactions. Sink reactions and exchange reactions were utilized within the community models to identify exported and imported metabolites, respectively. The intersection of these exported and imported metabolites constituted the set of exchanged metabolites, denoting metabolites that could be both secreted and taken up by members of the bacterial community. Our analysis revealed that consensus community models exhibited higher number of exported metabolites (Table 1). On average, each model secreted 44.8 ± 9.1 and 42.8 ± 6.9 metabolites within the coral-associated bacteria and seawater bacteria community, respectively. However, despite the large number of metabolites available for secretion into the medium within the community, only 64 metabolites were found to be exchanged within the community. The highest number of exchanged metabolites was observed in gapseq models for the coral-associated bacteria community (92 exchanged metabolites) and in CarveMe models for the seawater bacteria community (90 exchanged metabolites).

Regarding the similarity of exchanged metabolites (Supplementary Fig. 6), the gapseq and KBase models exhibited relatively similar sets of exchanged metabolites compared to the CarveMe models in both the coral-associated and seawater bacterial communities (Jaccard index of 0.34 in both communities). This finding suggests that the use of the same database for model reconstruction may contribute to the similarity in exchanged metabolites among these approaches. Furthermore, our results indicate that the types of exchanged metabolites within the community models are highly dependent on the chosen reconstruction approaches and the underlying databases. Interestingly, community models reconstructed using the same approach, even if applied to different communities, displayed more similar sets of exchanged metabolites compared to community models reconstructed using different approaches. However, we note that some models in both communities, reconstructed from the same MAG, did not share the same exported metabolites. Among the models with identical exported metabolites, the maximum and minimum predicted flux of exported metabolites varied between the reconstruction approaches (Supplementary Figs. 7–10). These findings warrant careful consideration of the conclusions drawn from applications of these models to assess the functional relevance of microbial interactions in communities.

Table 1 | The average number of imported, exported metabolites per model, and the number of exchanged metabolites in the community from different reconstruction approaches

	Methods	Imported metabolites	Exported metabolites	Exchanged metabolites
Coral-associated bacteria	CarveMe	10.7 ± 7.1	31.7 ± 9.1	80
	gapseq	28.1 ± 8.1	40.7 ± 7.8	92
	KBase	19.1 ± 7.2	26.4 ± 6.4	49
	Consensus	8.6 ± 8.1	44.8 ± 9.1	64
Seawater bacteria	CarveMe	10.4 ± 7.6	29 ± 6.9	90
	gapseq	24.1 ± 7.8	38.1 ± 5.9	87
	KBase	19.2 ± 7.5	27.2 ± 6.9	50
	Consensus	7.4 ± 8.7	42.8 ± 6.9	64

Discussion

In this study, we employed both top-down and bottom-up approaches for reconstruction of community models on the test case of coral-associated and seawater bacterial communities. The resulting models were subsequently compared with the consensus community models. To minimize the inherent uncertainty associated with each approach, we maintained uniformity by utilizing the same gene annotation tool (RAST) and adopting a universal biomass reaction during the model reconstruction process. However, despite these standardized procedures, we found substantial structural disparities among the resulting community models. We attribute these variations primarily to the gene-reaction mapping in the employed databases, which can significantly impact the model outcomes.

Gene sets are the basis of reconstructing GEMs. The absence of a gene in a model can result from the unavailability of its orthologous gene in the database or a lack of associated reactions within the database. It is generally assumed that models sharing similar gene sets would also exhibit similarity in their sets of reactions. However, CarveMe and KBase models demonstrated contradictory outcomes in terms of the similarity between gene and reaction sets. This finding could be attributed to differences in gene-reaction association information present in the BiGG and ModelSEED databases. This may also be a result of the variation in the number of reactions between CarveMe and KBase models. Additionally, the number of genes does not show positive correlation with the number of reactions or the proportion of reactions supported by genetic evidence. While gapseq models showcased a comparatively smaller number of genes, they encompassed a significant number of reactions, with merely 7.7% and 8.3% of total reactions on average lacking GPR associations in coral-associated and seawater bacterial models, respectively. This divergence might be attributed to the use of a customized database within the gapseq approach, which seemingly provided more comprehensive information regarding gene-reaction associations and resulting in numerous genes being associated with multiple reactions.

In this study, we also applied FVA to identify and filter out blocked reactions within the models, allowing us to investigate the enriched functions in active reactions in the community models. We observed that acyltransferases, which participate in the synthesis of long-chain fatty acids³⁹, was enriched in the shared reactions of gapseq and KBase models. However, this enrichment was not observed in the unblocked shared reactions. We hypothesize that the same reactions may carry different fluxes in the models reconstructed from different approaches, which can influence the enriched functions of models. Interestingly, the consensus approach demonstrated a greater capacity to distinguish the functional characteristics of different community models. This may be attributed to the comprehensive representation of biochemical reactions in consensus models. For instance, enzymes with oxidoreductase activity, acting on X-H and Y-H to form an X-Y bond, with oxygen as acceptor and those transferring aldehyde or ketonic groups were exclusively enriched in the coral-associated bacterial community. Conversely, the enzymes associated with carbon-sulfur lyases,

cleavage of ether bonds, interaction with sulfur group donors, and acyltransferase functions were specific to the seawater bacterial community.

The enriched enzymatic function involving the transfer of aldehyde or ketonic groups suggests a potential linkage to energy metabolism, particularly the reductive pentose phosphate cycle. This finding aligns with the findings of Doering, et al.³⁰, who identified the complete pathway in coral-associated bacteria. Conversely, the presence of acyltransferases, associated with the phosphate acetyltransferase-acetate kinase pathway, in the seawater bacterial community indicates divergent energy-generation pathways between these two bacterial communities. Additionally, our consensus community model of the seawater bacterial community revealed an enriched function of carbon-sulfur lyases, involved in the degradation of dimethylsulfoniopropionate (DMSP). DMSP, a key organic sulfur compound in the marine sulfur cycle, is produced by various marine organisms, including algae, phytoplankton, coral, and marine bacteria^{31–34}. The degradation of DMSP by marine heterotrophic bacteria provides a vital source of organic carbon and reduced sulfur for the bacterial community³⁵. However, further validation of our predictions from the consensus models is needed in the future work.

COMMIT considers permeable metabolites without compromising the growth rate of each model during the construction of the community model. This iterative process allows us to simulate microbial interactions in terms of metabolite exchange and significantly reduces the number of added reactions required for the community networks¹¹. The order of iteration appears to have minimal influence on the number of added reactions across different reconstruction approaches. Nonetheless, in gapseq and KBase models, we observed a negative correlation between the number of imported/exported metabolites and the MAG abundance. Remarkably opposite outcomes were encountered when we applied the iterative process using MAG abundance in increasing order for gapseq and KBase models. These findings underscore the potential impact of the iterative process on the metabolite secretion/take-up capacity of models reconstructed from gapseq and KBase. In contrast, CarveMe and consensus models exhibited no correlation with the iterative order, which might help to mitigate uncertainties related to the metabolite transport capabilities of the models.

Permeable metabolites have been instrumental in studying interspecies interactions in microbial ecosystems and have been suggested as mechanisms for maintaining genetic diversity within communities³⁶. Hence, in this study, we also examined the exchanged metabolites within the community models. We hypothesized that different reconstruction approaches would present distinct interaction outcomes. We observed a considerable variation in the number of exported/imported metabolites per model under different reconstruction approaches, leading to differences in the count of metabolites that could be exchanged within the community. Notably, our findings highlighted that the profile of exchanged metabolites was more influenced by the reconstruction approach used than by the type of bacterial community. This observation suggests a potential bias in predicting metabolite interactions using community GEMs. We note that not all models reconstructed with different approaches using the same MAG in the two communities resulted in the same set of exported metabolites. This stresses the conditional nature of metabolite exchanges depending on the composition of a metabolic community. However, to improve the accuracy of phenotypic predictions from community models, it is crucial to validate the predictions through experimental approaches³⁷, especially given the complexity of the interaction network in marine bacterial communities.

Overall, the consensus approach effectively integrates a majority of the information derived from diverse reconstruction tools into a unified model. For example, the consensus model incorporates all genes present in the models reconstructed from the same MAG, along with a substantial number of reactions and metabolites. This integration effectively mitigates some of the inherent biases in GEM reconstructions, as it concurrently considers reactions from multiple biochemistry databases supported by genetic evidence. Consequently, there is a notable reduction in the required gap-filling solutions, particularly of reactions without genetic support. Indeed, our results demonstrated that, compared to the consensus approach, a

considerable number of reactions without genetic support needed to be added to enable growth simulation when models from the individual approaches are considered. Therefore, the models resulting from the consensus approach are expected to be less biased. Furthermore, when gap-filling is required to render GEMs functional, the consensus approach harnesses the functional capabilities of other community members, which is not the case in other gap-filling solutions. Thus, the essence of the consensus approach underscores the importance of metabolic interactions among community members. However, this does not imply that the consensus approach results in inflation in the number of added exchange reactions; indeed, we found that the number of added exchange reactions, corresponding to the interactions, is not the largest in the consensus approach. In the COMMIT gap-filling approach, the order in which MAGs are considered may affect the added exchange reactions, and thus the predicted interactions in the community. Here, there is no expectation that models reconstructed from MAGs with higher abundance include more exchange reactions; indeed, this is what we found from the consensus approach, in stark contrast to the models from the individual approaches for which bias (towards fewer or more exchange reactions) was observed. Moreover, we found that there are variations in the resulting interactions which were found to strongly depend on the chosen reconstruction approach. In fact, our findings indicated that the reconstruction approach, rather than the community composition, strongly affects the set of exchange reactions in the resulting models. These advantages regarding the structural properties of models from the consensus approach in part contributed to the identification of more specific enriched functions, rendering them more suitable in downstream functional analyses. However, during this incorporation process, the consensus models may also assimilate reactions with unbounded fluxes in the original models. This assimilation, in turn, may result in a reduction in the quality of the model. Despite this potential drawback, we found that the consensus approach results in good quality of models which is shared by the majority of models in the community—a feature which is not typical for the other approaches. By mitigating the variability of model quality, the consensus approach may potentially lead to better prediction of exchange metabolites in the bacterial community.

Methods

Metagenome-assembled genome data

A total of 105 bacterial metagenome-assembled genomes (MAGs) were downloaded from NCBI under Bioproject accession PRJNA545004²⁶ to reconstruct community metabolic models. Altogether 50 MAGs were from the coral tissue of *Porites lutea*, including the bacteria phyla: Acidobacteriota (6 MAGs), Actinobacteriota (2 MAGs), Bacteroidota (3 MAGs), Chloroflexota (14 MAGs), Dadabacteria (1 MAG), Gemmatimonadota (4 MAGs), UBP10 (2 MAGs), Latescibacterota (3 MAGs), Nitrospirota (1 MAG), Poribacteria (7 MAGs), and Proteobacteria (7 MAGs). These phyla broadly represent the taxonomic diversity observed in *P. lutea*²⁶. In addition, 55 MAGs were used to represent the bacteria phyla composition of surrounding seawater around Orpheus Island, Great Barrier Reef, Australia which harbours: Actinobacteriota (3 MAGs), Bacteroidota (10 MAGs), Cyanobacteriota (2 MAGs), Marinisomatota (1 MAG), Patescibacteria (1 MAG), Planctomycetota (4 MAGs), Proteobacteria (30 MAGs), SAR324 (2 MAGs), and Verrucomicrobiota (2 MAGs). Abundance of MAGs was determined using BMap³⁸, which calculated the average read coverage of per contigs per MAG, generating a coverage profile. The abundance of MAGs was then presented as the sum of contig coverages.

Generation of draft and consensus metabolic models

Metabolic reconstruction approaches rely on diverse databases, and the choice between bottom-up and top-down methodologies can lead to variations in the structure of metabolic reconstruction. To provide a comprehensive overview of this discrepancy, we compared three reconstruction approaches: CarveMe¹⁶, gapseq¹⁷, and KBase¹⁸. Among these approaches, CarveMe belongs to top-down reconstruction approach while gapseq and KBase are bottom-up approaches. For the reconstruction, the MAGs were

annotated using Annotate Metagenome Assembly and Re-annotate Metagenome with RASTtk – v1.073 app^{39–41} published on KBase platform.

In the CarveMe reconstruction approach, the manually curated universal bacteria model was used as a template. The annotated sequence was aligned with amino protein sequence in the BiGG database⁴². Subsequently, the reaction scores were derived by associating them with the sequence similarity scores through the gene-protein reaction (GPR) rules. Reactions lacking genetic evidence were assigned negative scores within this framework. During the model carving process, reactions with low scores were eliminated from the universal model to generate the draft models.

In the gapseq tool, draft models were reconstructed using the default settings. The annotated sequences were utilized to predict pathways and subsystem using a customized database, obtained from MetaCyc⁴³, KEGG⁴⁴, and ModelSEED⁴⁵. Additionally, transporters were predicted based on the Transporter Classification Database (TCDB)⁴⁶, which catalogs a wide range of transport proteins and their functional classifications.

In the KBase approach, the metabolic reconstruction process was carried-out using the ModelSEED pipeline⁴⁵. The functional annotation of MAGs obtained from RAST was directly mapped to the corresponding biochemical reactions present in the ModelSEED biochemistry database. The biomass reactions were based on a template biomass reaction and assigned non-universal biomass components, such as cofactors and cell wall components, using the SEED subsystems and RAST functional annotations. Subsequently, the draft models resulting from this process were downloaded for further analysis and refinement.

To build the consensus models, we followed the pipeline provided in COMMIT¹¹. Before merging the models obtained from different reconstructions, we unified the reaction and metabolite IDs in the draft models by mapping them to MNXref IDs using the provided MNXref reference files⁴⁷. The biomass reaction, if present, and exchange reactions were subsequently removed. In the merging process, we used the CarveMe models as the initial component of the consensus models in an iterative fashion (following by gapseq and KBase models). Subsequent reconstructions were compared to this consensus model in a stepwise manner. First, the fields of the models were harmonized to ensure consistency. Next, the gene identifiers were compared, and if necessary, any genes not present in the consensus model were added. Subsequently, the reactions were compared based on various criteria, including reaction IDs, GPR rules, metabolite composition, reversibility, and mass balance. Any duplicate reactions and metabolites were removed from the consensus model to avoid redundancies.

Gap-fill community models obtained by COMMIT

Before the gap-filling, the exchange and biomass reactions were removed from the draft CarveMe, gapseq, and KBase models. Subsequently, a universal biomass reaction, which was adapted from *Escherichia coli* biomass composition⁴⁸ according to the universal biomass components in prokaryotes⁴⁹, was added into the draft CarveMe, gapseq, KBase, and consensus models. To perform the iterative gap-filling, the community models derived from different reconstructions were processed in descending and ascending order given by the species abundance. Initially, a common microbial growth medium (LB media) was provided as the initial media for the gap-filling process. Adjusted M9 media (with glucose and magnesium ion) was then employed for subsequent iterations.

Comparison of community models from different reconstructions

The quality of the models was assessed using MEMOTE²⁷ to evaluate their overall performance. Subsequently, the scores for the corresponding indices were extracted from the output report generated by MEMOTE for subsequent analyses. In addition, several model features, including the number of reactions, metabolites, dead-end metabolites, and genes, were analyzed to compare the structural properties of the models. The similarity between models was determined using the Jaccard similarity coefficient.

To identify enriched functions in the community models, we extracted the enzyme commission numbers (EC numbers) from each reaction

included in the model and used them to the second digit (i.e. enzyme subclass level) in the enrichment analysis. Flux variability analysis (FVA) was applied to distinguish between blocked and unblocked reactions in the models. Subsequently, we conducted a hypergeometric test to identify significantly enriched EC numbers in the shared (unblocked) reactions between two models reconstructed from the same MAG using different approaches. We analyzed the functional characteristics of the community models reconstructed from different approaches by considering all unblocked reactions in the community models.

To identify potential exchange metabolites in the community, we examined the sink and exchange reactions in models, which allowed us to identify the metabolites involved in exchange processes. Sink and exchange reactions were determined using COMMIT during community model reconstruction, defining exported and imported metabolites, respectively. The exchanged metabolites were identified as the intersection of all exported and imported metabolites within the community model. We considered the exchanged metabolites as an indicator of metabolic interaction potential, enabling the evaluation of the metabolic interactions within the community models. However, it is crucial to note that this analysis primarily holds a qualitative nature. To quantify the flux of exchanged metabolites, we further compared the maximum and minimum flux of the same exported metabolites in the model reconstructed from the same MAG using different reconstruction approaches under optimal growth rate conditions.

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The data generated for this manuscript, draft and consensus reconstructions from the individual approaches, all the gap-filled community models are available at <https://doi.org/10.5281/zenodo.10289699>.

Received: 6 January 2024; Accepted: 13 May 2024;

Published online: 23 May 2024

References

1. Konopka, A., Lindemann, S. & Fredrickson, J. Dynamics in microbial communities: unraveling mechanisms to identify principles. *ISME J.* **9**, 1488–1495 (2015).
2. Lawson, C. E. et al. Metabolic network analysis reveals microbial community interactions in anammox granules. *Nat. Commun.* **8**, 15416 (2017).
3. Wang, C.-W. et al. Soil Bacterial Community May Offer Solutions for Ginger Cultivation. *Microbiol. Spectr.* **10**, e01803–e01822 (2022).
4. De Roy, K., Marzorati, M., Van Den Abbeele, P., Van De Wiele, T. & Boon, N. Synthetic microbial ecosystems: an exciting tool to understand and apply microbial communities. *Environ. Microbiol.* **16**, 1472–1481 (2014).
5. Althani, A. A. et al. Human Microbiome and its Association With Health and Diseases. *J. Cell. Physiol.* **231**, 1688–1694 (2016).
6. de Voogd, N. J., Cleary, D. F. R., Polónia, A. R. M. & Gomes, N. C. M. Bacterial community composition and predicted functional ecology of sponges, sediment and seawater from the thousand islands reef complex, West Java, Indonesia. *FEMS Microbiol. Ecol.* **91**, <https://doi.org/10.1093/femsec/fiv019> (2015).
7. Pham, V. H. T. & Kim, J. Cultivation of unculturable soil bacteria. *Trends Biotechnol.* **30**, 475–484 (2012).
8. Magnúsdóttir, S. et al. Generation of genome-scale metabolic reconstructions for 773 members of the human gut microbiota. *Nat. Biotechnol.* **35**, 81–89 (2017).
9. Kundu, P. & Ghosh, A. Genome-scale community modeling for deciphering the inter-microbial metabolic interactions in fungus-farming termite gut microbiome. *Comput. Biol. Med.* **154**, 106600 (2023).
10. Du, H. et al. Microbial active functional modules derived from network analysis and metabolic interactions decipher the complex microbiome assembly in mangrove sediments. *Microbiome* **10**, 224 (2022).
11. Wendering, P. & Nikoloski, Z. COMMIT: Consideration of metabolite leakage and community composition improves microbial community reconstructions. *PLoS Comput. Biol.* **18**, e1009906 (2022).
12. Maitaine, V., Vannier, N., Vandenkoornhuysen, P. & Hacquard, S. Multi-genome metabolic modeling predicts functional interdependencies in the Arabidopsis root microbiome. *Microbiome* **10**, 217 (2022).
13. Henry, C. S. et al. Microbial Community Metabolic Modeling: A Community Data-Driven Network Reconstruction. *J. Cell. Physiol.* **231**, 2339–2345 (2016).
14. Gelbach, P. E. & Finley, S. D. Flux Sampling in Genome-scale Metabolic Modeling of Microbial Communities. *bioRxiv* <https://doi.org/10.1101/2023.04.18.537368> (2023).
15. Ang, K. S., Lakshmanan, M., Lee, N. R. & Lee, D. Y. Metabolic Modeling of Microbial Community Interactions for Health, Environmental and Biotechnological Applications. *Curr. Genomics* **19**, 712–722 (2018).
16. Machado, D., Andrejev, S., Tramontano, M. & Patil, K. R. Fast automated reconstruction of genome-scale metabolic models for microbial species and communities. *Nucleic Acids Res.* **46**, 7542–7553 (2018).
17. Zimmermann, J., Kaleta, C. & Waschina, S. gapseq: informed prediction of bacterial metabolic pathways and reconstruction of accurate metabolic models. *Genome Biology* **22**, <https://doi.org/10.1186/s13059-021-02295-1> (2021).
18. Arkin, A. P. et al. KBase: The United States Department of Energy Systems Biology Knowledgebase. *Nat. Biotechnol.* **36**, 566–569 (2018).
19. Mendoza, S. N., Olivier, B. G., Molenaar, D. & Teusink, B. A systematic assessment of current genome-scale metabolic reconstruction tools. *Genome Biol.* **20**, 158 (2019).
20. Bernstein, D. B., Sulheim, S., Almaas, E. & Segrè, D. Addressing uncertainty in genome-scale metabolic model reconstruction and analysis. *Genome Biol.* **22**, 64 (2021).
21. Schnoes, A. M., Brown, S. D., Dodevski, I. & Babbitt, P. C. Annotation Error in Public Databases: Misannotation of Molecular Function in Enzyme Superfamilies. *PLoS Comput. Biol.* **5**, e1000605 (2009).
22. Lobb, B., Tremblay, B. J.-M., Moreno-Hagelsieb, G. & Doxey, A. C. An assessment of genome annotation coverage across the bacterial tree of life. *Microb. Genomics* **6**, <https://doi.org/10.1099/mgen.0.000341> (2020).
23. Pham, N. et al. Consistency, Inconsistency, and Ambiguity of Metabolite Names in Biochemical Databases Used for Genome-Scale Metabolic Modelling. *Metabolites* **9**, 28 (2019).
24. Chindelevitch, L., Stanley, S., Hung, D., Regev, A. & Berger, B. MetaMerge: scaling up genome-scale metabolic reconstructions, with application to Mycobacterium tuberculosis. *Genome Biol.* **13**, R6 (2012).
25. Aung, H. W., Henry, S. A. & Walker, L. P. Revising the Representation of Fatty Acid, Glycerolipid, and Glycerophospholipid Metabolism in the Consensus Model of Yeast Metabolism. *Ind. Biotechnol.* **9**, 215–228 (2013).
26. Robbins, S. J. et al. A genomic view of the reef-building coral *Porites lutea* and its microbial symbionts. *Nat. Microbiol.* **4**, 2090–2100 (2019).
27. Lieven, C. et al. MEMOTE for standardized genome-scale metabolic model testing. *Nat. Biotechnol.* **38**, 272–276 (2020).
28. Tipton, K. & Boyce, S. History of the enzyme nomenclature system. *Bioinformatics* **16**, 34–40 (2000).
29. Röttig, A. & Steinbüchel, A. Acyltransferases in bacteria. *Microbiol. Mol. Biol. Rev.* **77**, 277–321 (2013).

30. Doering, T. et al. Genomic exploration of coral-associated bacteria: identifying probiotic candidates to increase coral bleaching resilience in *Galaxea fascicularis*. *Microbiome* **11**, 185 (2023).
31. Stefels, J. Physiological aspects of the production and conversion of DMSP in marine algae and higher plants. *J. Sea Res.* **43**, 183–197 (2000).
32. Raina, J.-B. et al. DMSP biosynthesis by an animal and its role in coral thermal stress response. *Nature* **502**, 677–680 (2013).
33. Curson, A. R. J. et al. Dimethylsulfoniopropionate biosynthesis in marine bacteria and identification of the key gene in this process. *Nat. Microbiol.* **2**, 17009 (2017).
34. Zheng, Y. et al. Bacteria are important dimethylsulfoniopropionate producers in marine aphotic and high-pressure environments. *Nat. Commun.* **11**, 4658 (2020).
35. Zhang, X.-H. et al. Biogenic production of DMSP and its degradation to DMS—their roles in the global sulfur cycle. *Sci. China Life Sci.* **62**, 1296–1319 (2019).
36. Morris, J. J. Black Queen evolution: the role of leakiness in structuring microbial communities. *Trends Genet.* **31**, 475–482 (2015).
37. Schäfer, M. et al. Metabolic interaction models recapitulate leaf microbiota ecology. *Science* **381**, eadf5121 (2023).
38. Bushnell, B. BMap: a fast, accurate, splice-aware aligner. (Lawrence Berkeley National Lab.(LBNL), Berkeley, CA (United States), (2014).
39. Aziz, R. K. et al. The RAST Server: Rapid Annotations using Subsystems Technology. *BMC Genomics* **9**, 75 (2008).
40. Overbeek, R. et al. The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic Acids Res.* **42**, D206–D214 (2013).
41. Brettin, T. et al. RASTtk: A modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. *Sci. Rep.* **5**, 8365 (2015).
42. King, Z. A. et al. BiGG Models: A platform for integrating, standardizing and sharing genome-scale models. *Nucleic Acids Res.* **44**, D515–D522 (2015).
43. Caspi, R. et al. The MetaCyc database of metabolic pathways and enzymes - a 2019 update. *Nucleic Acids Res.* **48**, D445–D453 (2019).
44. Kanehisa, M., Sato, Y., Furumichi, M., Morishima, K. & Tanabe, M. New approach for understanding genome variations in KEGG. *Nucleic Acids Res.* **47**, D590–D595 (2018).
45. Henry, C. S. et al. High-throughput generation, optimization and analysis of genome-scale metabolic models. *Nat. Biotechnol.* **28**, 977–982 (2010).
46. Saier, M. H. Jr., Reddy, V. S., Tamang, D. G. & Västermark, Å. The Transporter Classification Database. *Nucleic Acids Res.* **42**, D251–D258 (2013).
47. Moretti, S., Tran, VanDu. T., Mehl, F., Ibberson, M. & Pagni, M. MetaNetX/MNXref: unified namespace for metabolites and biochemical reactions in the context of metabolic models. *Nucleic Acids Res.* **49**, D570–D574 (2020).
48. Orth, J. D. et al. A comprehensive genome-scale reconstruction of *Escherichia coli* metabolism—2011. *Mol. Syst. Biol.* **7**, 535 (2011).
49. Xavier, J. C., Patil, K. R. & Rocha, I. Integration of Biomass Formulations of Genome-Scale Metabolic Models with Experimental Data Reveals Universally Essential Cofactors in Prokaryotes. *Metab. Eng.* **39**, 200–208 (2017).

Acknowledgements

The authors would like to thank the Melbourne-Potsdam PhD Program of the Max Planck Institute of Molecular Plant Physiology, Potsdam, Germany and The University of Melbourne, Parkville, Melbourne for supporting this project.

Author contributions

Y.E.H. and Z.K. designed the research. Y.E.H. reconstructed metabolic models, analyzed the results, and wrote the original draft. Y.E.H., K.T., H.V., and Z.N. reviewed and contributed to the completion of this manuscript.

Funding

Open Access funding enabled and organized by Projekt DEAL.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41540-024-00384-y>.

Correspondence and requests for materials should be addressed to Zoran Nikoloski.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024