DOI:
10.1038/nbt.4213

Document Version
Peer reviewed version

Correspondence

Advancing towards modelling of the human gut microbiome

To the Editor

Recently Magnusdottir et al.[1] described the generation of genome-scale metabolic models (GEMs) for 773 members of the human gut microbiota, referred to as AGORA, and this work represents a valuable contribution towards our ability to perform quantitative analysis of how the many different species of the gut microbiota interact and impact host metabolism. Several studies have shown that the gut microbiota impacts human health, including cardiovascular disease[2] and Type 2 Diabetes[3-5], but so far no studies have shown causal relationships between the gut microbiota composition and human health status. Through mathematical modelling of the gut microbiota, it will be possible to evaluate different hypothesis and hereby gain mechanistic insight into how the gut microbiota composition affects host metabolism. Genome-scale metabolic modelling is particularly well suited for this purpose as it is possible to reconstruct the metabolic networks of gut symbionts based on genomic information and then use constraint-based modelling, often referred to as Flux Balance Analysis (FBA), for simulation of their metabolic functions[6]. This modelling concept has been validated for its ability to correctly simulate the metabolism of colonized bacteria in germ free mice[7] and for being able to predict how the levels of key metabolites in faecal water and plasma are impacted by the gut microbiota composition[8]. Furthermore. Using this concept, computational tools and frameworks have been developed to allocate metabolic resources to individual microbes and infer the ecological interaction, spatial dynamics, as well as the community-level assembly processes [9, 10].

The work of Magnusdottir et al. is therefore important as it represents a resource of GEMs for more than 700 gut microbes, which accounts for most of the dominant species in the human gut microbiota. The work, however, also points to a number of challenges in going forward and using these GEMs as functional models for simulating metabolic interactions between gut symbionts and between the gut microbiota and their host. We performed a quality check of the 773 GEMs in AGORA and found that most of these are incapable of quantitative predictions without further manual curation and support of experimental data. We performed the same analysis on 70 metabolic models from BiGG repository[11] in order to have a fair comparison. It is to be mentioned that the majority of these models are different strains of *E. coli*, however our results were not found to be significantly different while excluding *E. coli* models. First, we evaluated the ability of the models to predict growth. The COBRA toolbox was used to calculate the specific growth rate of the models using the glpk solver. Growth at aerobic, micro-aerobic and anaerobic conditions were simulated by setting the lower bound of oxygen exchange reaction to -1000, -1, and zero $mmol.gr^{-1}_{DW}.h^{-1}$, respectively. In all cases, the upper bound for the oxygen exchange reaction was zero. Other exchange bounds were not altered and taken from AGORA. The specific growth rate at anaerobic condition varies between 0.004 h^{-1} and 255 h^{-1} (**Fig. S1A**). We noticed that the models whose growth rates are in reasonable ranges have very tight constraints on their exchange reactions whereas the models who grow too fast do not have any bounds limiting their metabolite uptake fluxes. This means that the reason for some of the models having reasonable predicted growth rates is due to bounds forced on almost all exchange reactions. Although constraining exchange reactions using available experimental data is conventional in constraint-based modelling, applying constraints to control the growth rates does not justify the quality of a model and boils down to the fact that lower input would lead to lower objective function value which is self-evident. The fact that the value of growth rate is controlled via exchange reactions, rather than internal constraints in the AGORA models, is also reflected in the reduced costs

associated with the exchange and internal reactions. The median of reduced costs associated to the active internal reactions in AGORA is 4.2877e-19, which is several orders of magnitude smaller than the median of reduced costs associated to the active exchange reactions (-1.7227e-04). The same values are -9.1570e-06 for internal and -0.0062 for exchange reactions in the BiGG models. We further tested the effect of constraining the exchange reactions on the specific growth rate predictions. To this end, all exchange reactions that were originally allowed in the AGORA models (with either tight or loose constraints) were set to have a minimum of -10 $mmol.gr^{-1}_{DW}.h^{-1}$ which is the modelling convention for lower bounds of carbon source uptake reactions. Even with this constraint, the majority of the models are growing in biologically infeasible ranges (**Fig. S1B**). The same analysis for the BiGG models, however, shows that their growth rates do not increase substantially when the constraints on the allowed exchange reactions are relaxed. As shown in Figure S1, even when the lower bounds of exchange reactions are increased to -100 $mmol/gr_{DW}/h$ in the BiGG models, the growth rate does not increase as much as for the AGORA models even when the maximum uptake rate is constrained at -10 $mmol/gr_{DW}/h$. Excluding the *E. coli* models did not change these results as shown in Figure S1G. One possible reason for this result could be the reliance of many of the AGORA models on several different carbon sources simultaneously. This might not be a problem by itself, but when the calculated growth rates are divided by the total carbon influx (resulting in calculating the biomass yield), we observed that the growth yields are very low, and much lower than experimentally observed for most bacteria (**Fig. S1C**). This fact indicates that many carbon-containing metabolites are actively taken up from the environment but not used for biomass formation. We investigated the carbon influx performing flux coupling using the F2C2 tool[12] on 100 randomly chosen AGORA models to check the coupling types existing between carbon-containing active exchange reactions and internal active reactions. Some carbon-containing active exchange reactions in AGORA are coupled with as many as 300 internal reactions, pointing to the fact that the carbon sources are used in many different pathways. We performed the same analysis with the BiGG models, and as seen in Figure S2 there is much less flux coupling in these models. We also looked into the essential carbon-containing metabolites that are used by the AGORA models, serving as a carbon source, and their gene associations. The results are shown in Figure S3. This is not the case in most well-curated models. In BiGG models, there are 1 or 2 essential carbon sources and their transport reactions have gene associations. Models that rely on so many carbon sources, should either have genetic support for their transport reactions or internal constraints, and in the AGORA models utilization of the majority of the carbon sources have no gene associations. This might as well be an indicative of high auxotrophy of microbiota, however auxotrophy is massively over-estimated in these models.

Another reason for the great effect of the exchange reactions on the growth rate is that many biomass precursors in the AGORA models are directly taken from the media. We checked the number of such metabolites in all AGORA and BiGG models and the results are shown in Figure S4. This is partly a consequence of using an automatic gap-filling procedure, a mixed integer linear programming (MILP) algorithm where the number of the added reactions to the draft model is minimized. This means that a minimal set of reactions are added to the point that the biomass reaction (or the objective function in general) would have a non-zero flux. In gap filling approach it is possible to penalize the exchange and transport reactions more than internal ones, however, automatic gap filling methods still add exchange/transport reactions more than they should. Furthermore, since all of the models in the AGORA repository have undergone the same procedure of gap-filling, the use of metabolites in these models is highly similar. We calculated the Jaccard similarity score in every pair of these models and the

results are shown in Figure S 4C. From this it is found that the models are very similar in terms of biomass precursors that are consumed directly from the media.

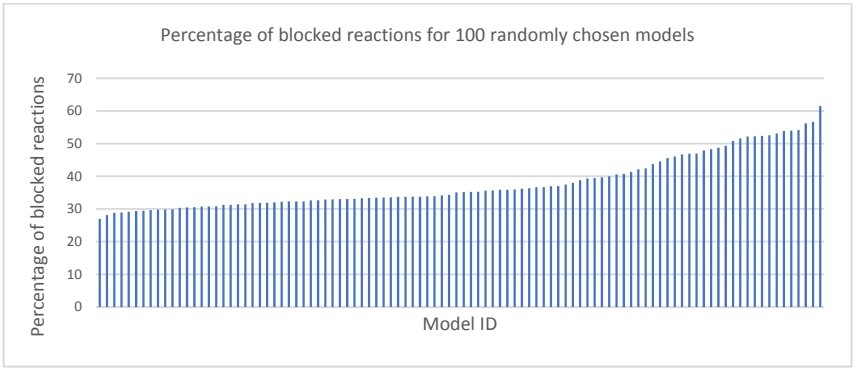
In order to further analyse the AGORA models, we randomly selected 100 models and investigated the blocked reactions, which are the reactions not able to carry any flux, and also the reactions involved in thermodynamically infeasible loops. To identify blocked reactions, flux variability analysis was performed and the reactions with minimum and maximum fluxes below a relevant tolerance (1e-6) were identified. The percentage of blocked reactions for the chosen models are shown in **Fig. 1A**. The average percentage of blocked reactions in the models are 38%, which is a very high percentage for a GEM, but typical for automatically reconstructed models. 65 out of the 100 models showed a percentage higher than 33% (about 1/3) of the reactions being blocked and this shows that these models need manual curation to be more connected. These reactions are most probably associated with dead-end metabolites that are only consumed or produced. We checked the number of dead-end metabolites in AGORA and the results shown in Figure S 5A. Same analysis on the BiGG models showed a median of 16% and an average of 21% for the percentage of blocked reactions (Figure S 5C) and much fewer dead-end metabolites than in the AGORA models (Figure S 5B).

In the AGORA models, reactions participating in thermodynamically infeasible loops are able to carry flux even if all exchange reactions are blocked. We therefore performed loop detection analysis on the same 100 models. To this end, we blocked all exchange reactions (no uptake/excretion was allowed) and performed flux variability analysis. Reactions whose minimum/ maximum fluxes reached their lower/upper bounds were detected. This clearly indicates that these reactions are involved in thermodynamically infeasible loops since they were able to hit the bounds even though no input was provided for the model. The results are depicted in **Fig. 1B**. On average, there are 76 reactions involved in thermodynamically infeasible loops. We also checked their essentiality for growth and found that all of them were essential in all models. These reactions should be curated in order to avoid mathematical artefacts, in particular as all these reactions are essential for growth in AGORA models. We investigated the BiGG models in the same way, and only one model was detected to have 7 reactions involved in thermodynamically infeasible loops. To make sure that our analysis is not favoured towards one particular phyla we evaluated the phyla distribution of the chosen 100 models and compared this with the overall phyla distribution of AGORA. As shown in **Fig. S6**, the phyla distribution of our random sample is similar to the overall phyla distribution and our analysis is therefore not biased.

The gut environment is anaerobic or at best micro-aerobic and isolation of gut symbionts generally requires handling of all samples at strict anaerobic conditions [13, 14]. However, all of the AGORA models have oxygen-related internal reactions, some of which are essential. To evaluate this, we extracted all oxygen-related reactions in the models and investigated their essentiality. In 17 models (**Table S1**), the predicted growth rate would be zero if all oxygen-related reactions were blocked. We could therefore not confirm the statement by the authors that all the models could simulate growth at anaerobic growth conditions. It is clearly necessary to curate these models further so that they can predict growth at strict anaerobic conditions. We also checked the directionality of ATP-related reactions. In the models, the reactions that involve phosphate transfer from ATP to an acceptor molecule should be irreversible. We checked to see whether ATP could be produced as a result of wrong directionality. In order to do so, we fetched all ATP-related reactions and identified the ones that were functioning in the wrong direction. On average, there are 10 reactions in all of the models that do not follow the aforementioned rule and should be further curated. The energy

that is produced in the form of ATP is subsequently used by other internal or transport reactions in the model, including the biomass reaction.

A.



B.

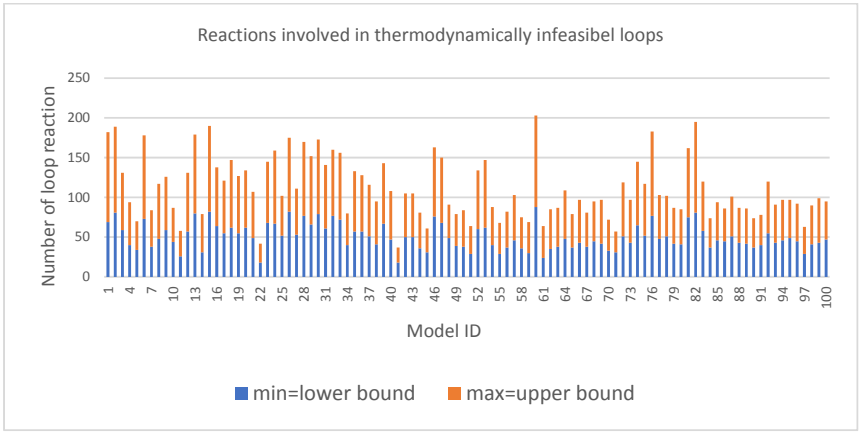


Fig. 1 The blocked reactions and reactions involved in thermodynamically infeasible loops in 100 models. **A.** Percentage of blocked reactions in 100 randomly chosen AGORA models. Models are sorted based on the aforementioned percentage. **B.** The number of reactions involved in thermodynamically infeasible loops. Models are in the same order as part A. The blue bars show the number of the reactions whose minimum hit their lower bound and the orange bars show the reactions whose maximum value hit their upper bound.

Even though the gut microbiota typically consists of more than 100 species it is questionable whether there are really large functional differences. Thus, many different species may in fact have similar metabolic functions. We therefore performed an analysis of the similarity of the AGORA models in terms of functional closeness. The Jaccard similarities were calculated based on the presence or absence of corresponding reactions. The models cluster into 3 main groups (**Fig. 2**). The optimum number of the clusters was calculated with minimum total

intra-cluster variation and average silhouette method (**Fig. S7**). We also calculated 16 indexes for determining the number of clusters. These indexes also support the optimum number of 3 clusters and are summarized in **Table S2.** Group 1 included 401 microbes mainly from Firmicutes, Proteobacteria and Actinobacteria; group 2 included 254 microbes mainly from Firmicutes, Proteobacteria, Actinobacteria and Fusobacteria; while group 3 included 112 microbes from Bacteroidetes (**Fig. S8**). We performed the same analysis for BiGG models and have shown that unlike AGORA, these models cluster in biologically meaningful groups (Fig S9). We have also shown that excluding the *E. coli* models does not change the way similar phylogenetically closer organisms cluster together.

In order to investigate their functional similarity, we used FBA and parsimonious FBA (pFBA) to identify flux-carrying internal reactions. In both cases an active reaction was defined as the one carrying a flux more than 1e-10 (changing this tolerance to higher values, e.g. 1e-4, did not change the results). There was not a significant difference between the results of these two methods as depicted in **Fig. S10**. In pFBA simulations, 83.8% of all possible pairwise comparisons had 50-70% similarity and 15.6% of the cases showed similarities between 70-90%. Based on FBA results in 79.5% of the possible pairwise comparison cases, there was 50-70% similarity and in 19.7% of the cases we observed 70-90% similarities. The active reactions across all 773 models cluster in three main groups. It turns out that this clustering is to a large extent determined by the definition of the biomass reaction in the models. In AGORA, there are 9 different formulations for the biomass equation and these formulations cluster in three groups (**Fig. S11 A**), which is very similar to the grouping according to flux-carrying reactions. This shows that the models which have similar biomass reactions, have similar objective functions and might lead to similar active reaction sets. Phyla distribution in each of the three main groups shows that Actinobacteria and Firmicutes are divided across all groups, while Bacteroidetes and Proteobacteria are separated between group 1 and 3 (**Fig. S11 B** and **S11 C**). This means that there are no single phyla whose members cluster together in one group. Finally, we also investigated the similarity in terms of metabolite use for growth. As observed the metabolites that are used for growth of the AGORA models, show a similar pattern as the reaction similarities (**Fig. S12**). Using the same set of growth-required metabolites is to be expected considering the high degree of similarity of the models.

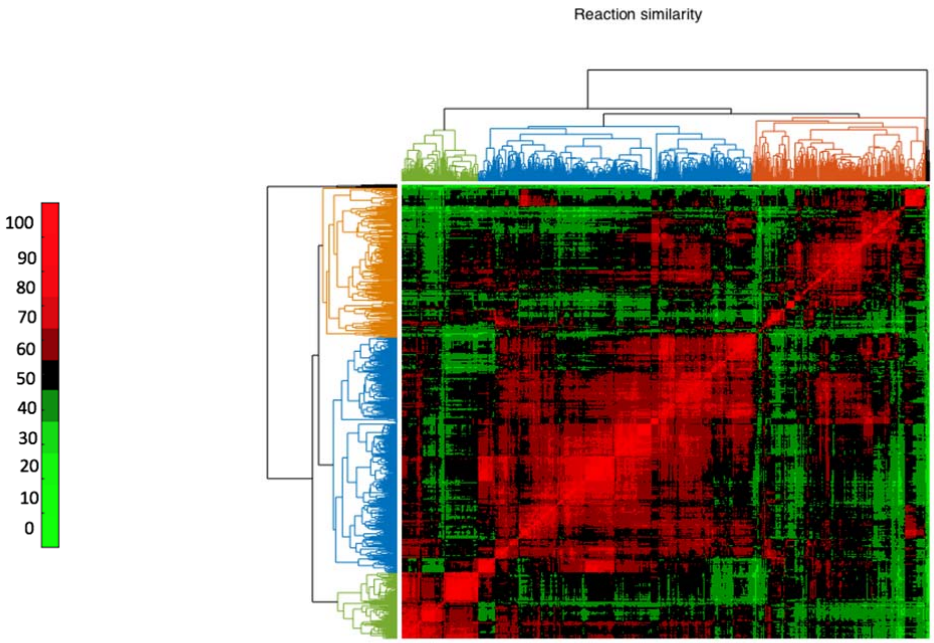


Fig. 2 Clustergram based on Jaccard similarity index for reactions.

In conclusion, our analysis shows that the AGORA models could be divided into two groups based on their *in silico* growth rates. The first group consists of the ones that predict the growth rates in biologically reasonable ranges and the models in the second group predict very high growth rates. The reason behind this behaviour is the bounds used on the exchange reactions. Exchange reactions play an important role in the value of the growth rate, since a large part of biomass metabolites is directly taken from media and consumed by the biomass reaction. While the models in the first group are exposed to very low bounds on their exchange reactions, the models in the second group are not constrained at all. One way to overcome this problem is to divide the growth rate by the total carbon influx, but from our analysis we find that this results in very low biomass yields. Therefore, it is necessary to carefully curate these models before they are used for simulations of growth. It is a major effort to manually curate GEMs, but from our analysis of the model similarity it is also clear that one may not need to actually curate all the 773 GEMs of AGORA as one can select representative species from each of the functional similar groups. It should, however, be mentioned that this similarity analysis may to some extent be driven by the fact that only 9 different biomass reactions are used for the AGORA models. Thus, the similarity in biomass reactions have affected the gap-filling and curation process of the AGORA models, and may therefore have resulted in the clustering of the models that we here present. In order to further advance the field of modelling of the gut microbiota using GEMs, it is therefore necessary to perform more detailed analysis of individual species with the aim of further expanding the set of biomass reactions as well as manual curation of the GEMs with the objective to improve their oxygen and ATP metabolism. Considering the scale of such an effort, this will have to be done at the community level where clear standards are being defined for the quality of individual models.

Parizad Babaei[1], Saeed Shoaie[2], Boyang Ji[1] and Jens Nielsen[1,3]
[1] Department of Biology and Biological Engineering, Chalmers University of Technology, SE41296 Gothenburg, Sweden
[2] Centre for Host–Microbiome Interactions, Dental Institute, King's College London, SE1 9RT London, UK
[3] Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, DK2800 Kgs. Lyngby, Denmark
Author of correspondence: Jens Nielsen, nielsenj@chalmers.se

1. Magnúsdóttir, S., et al., *Generation of genome-scale metabolic reconstructions for 773 members of the human gut microbiota.* Nature Biotechnology, 2016.
2. Karlsson, F.H., et al., *Symptomatic atherosclerosis is associated with an altered gut metagenome.* Nature communications, 2012. **3**: p. 1245.
3. Karlsson, F.H., et al., *Gut metagenome in European women with normal, impaired and diabetic glucose control.* Nature, 2013. **498**(7452): p. 99-103.
4. Qin, J., et al., *A metagenome-wide association study of gut microbiota in type 2 diabetes.* Nature, 2012. **490**(7418): p. 55-60.
5. Forslund, K., et al., *Disentangling type 2 diabetes and metformin treatment signatures in the human gut microbiota.* Nature, 2015. **528**(7581): p. 262-266.
6. Karlsson, F.H., et al., *Prospects for systems biology and modeling of the gut microbiome.* Trends in biotechnology, 2011. **29**(6): p. 251-258.
7. Mardinoglu, A., et al., *The gut microbiota modulates host amino acid and glutathione metabolism in mice.* Molecular systems biology, 2015. **11**(10): p. 834.
8. Shoaie, S., et al., *Quantifying diet-induced metabolic changes of the human gut microbiome.* Cell metabolism, 2015. **22**(2): p. 320-331.
9. Zomorrodi, A.R. and C.D. Maranas, *OptCom: a multi-level optimization framework for the metabolic modeling and analysis of microbial communities.* PLoS Comput Biol, 2012. **8**(2): p. e1002363.
10. Harcombe, W.R., et al., *Metabolic resource allocation in individual microbes determines ecosystem interactions and spatial dynamics.* Cell Rep, 2014. **7**(4): p. 1104-15.
11. King, Z.A., et al., *BiGG Models: A platform for integrating, standardizing and sharing genome-scale models.* Nucleic acids research, 2015. **44**(D1): p. D515-D522.
12. Larhlimi, A., et al., *F2C2: a fast tool for the computation of flux coupling in genome-scale metabolic networks.* BMC bioinformatics, 2012. **13**(1): p. 57.
13. Albenberg, L., et al., *Correlation between intraluminal oxygen gradient and radial partitioning of intestinal microbiota.* Gastroenterology, 2014. **147**(5): p. 1055-1063. e8.
14. Espey, M.G., *Role of oxygen gradients in shaping redox relationships between the human intestine and its microbiota.* Free Radical Biology and Medicine, 2013. **55**: p. 130-140.