

Lawrence Berkeley National Laboratory

Environ Genomics & Systems Bio

Title

Erroneous predictions of auxotrophies by CarveMe

Permalink

<https://escholarship.org/uc/item/3zs8m1w5>

Journal

Nature Ecology & Evolution, 7(2)

ISSN

2397-334X

Author

Price, Morgan

Publication Date

2023-02-01

DOI

10.1038/s41559-022-01936-3

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

Erroneous Predictions of Auxotrophies by CarveMe

Morgan Price, Lawrence Berkeley National Lab

In “Polarization of microbial communities between competitive and cooperative metabolism,” Machado and colleagues claim to identify “cooperative” communities with many amino acid auxotrophies¹. A related paper, with many of the same authors, claims that 26% of bacteria from free-living communities' are amino acid auxotrophs². The putative amino acid auxotrophies were derived from genome sequences via CarveMe's metabolic models³, but were not validated in any way. I checked CarveMe's predictions for prototrophic bacteria and found a high rate of spurious auxotrophies.

I obtained CarveMe's predictions for 127 diverse bacteria that were experimentally demonstrated to grow in a defined medium without amino acids (listed in ref. 4). CarveMe incorrectly predicted that 41 of these bacteria (32%) were auxotrophic for at least one amino acid. Among CarveMe's predictions for prototrophic bacteria, the most common errors were for asparagine (29 cases), methionine (10 cases), and glutamine (3 cases). To identify likely biosynthetic pathways in these bacteria, I used GapMind⁴.

All of the genomes with spurious asparagine auxotrophies encode the tRNA-dependent pathway, which does not involve free asparagine as an intermediate^{5,6}. The failure to allow this pathway might be due to CarveMe's biomass reaction, which includes asparagine instead of asparaginyI-tRNA. (In metabolic modeling, the biomass reaction describes all of the compounds that are required for growth.) In effect, CarveMe assumes that bacteria need asparagine to grow; in fact, bacteria only need asparagine for protein synthesis, and asparaginyI-tRNA suffices.

Of the 10 bacteria that CarveMe incorrectly predicted to be methionine auxotrophs, six appear to use direct sulfuration of aspartate semialdehyde^{7,8}. This pathway may be missing from CarveMe's database.

I could not determine why CarveMe predicted that three of the prototrophs would be glutamine auxotrophs. All three of these genomes contain GlnA-type glutamine synthetases.

Overall, for prototrophic bacteria, the fraction of genomes with spurious auxotrophies (32%) is similar to the proportion of free-living bacteria that CarveMe predicted to be amino acid auxotrophs (26%). This is consistent with my view that most free-living bacteria can make all 20 amino acids⁸. For bacteria from host-associated communities, CarveMe predicts that 46% of bacteria are amino acid auxotrophs², which is not much higher than the false positive rate of 32%, so most of these auxotrophies may be spurious as well. More broadly, automatically-generated metabolic models are not accurate enough for ecological inferences, and their predictions should be treated with caution.

Data sources and availability

I began with a list of 140 diverse bacteria that grow in defined media without any amino acids ⁴. (These were identified from a database of media-isolate combinations ⁹ and from a compilation of nitrogen-fixing microbes ¹⁰.) CarveMe's auxotrophy predictions were available for 127 of those 140 bacteria.

CarveMe's auxotrophy predictions were obtained by using the download.sh script in <https://github.com/cdanielmachado/cooccurrence> (after download, the predictions are in the tab-delimited files in the ancestry subdirectory). CarveMe's models were downloaded from https://github.com/cdanielmachado/embl_gems.

A list of the 127 bacteria, as well as a table comparing the spurious CarveMe "auxotrophies" to the GapMind annotations for those pathways and genomes, is available from figshare (<https://doi.org/10.6084/m9.figshare.20993512.v1>).

The GapMind predictions for the 127 prototrophic bacteria (as well as some other bacteria and archaea) are available from <https://doi.org/10.6084/m9.figshare.9693689.v1> and can be viewed at https://papers.genomics.lbl.gov/cgi-bin/gapView.cgi?set=aa&orgs=orgsDef_150. (GapMind has been updated since the original publication, so the two analyses are not always identical.)

Acknowledgements

This material by ENIGMA- Ecosystems and Networks Integrated with Genes and Molecular Assemblies (<http://enigma.lbl.gov>), a Science Focus Area Program at Lawrence Berkeley National Laboratory is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Biological & Environmental Research under contract number DE-AC02-05CH11231

Bibliography

1. Machado, D. *et al.* Polarization of microbial communities between competitive and cooperative metabolism. *Nat. Ecol. Evol.* **5**, 195–203 (2021).
2. Yu, J. S. L. *et al.* Microbial communities form rich extracellular metabolomes that foster metabolic interactions and promote drug tolerance. *Nat. Microbiol.* **7**, 542–555 (2022).
3. 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).
4. Price, M. N., Deutschbauer, A. M. & Arkin, A. P. Gapmind: automated annotation of amino

- acid biosynthesis. *mSystems* **5**, (2020).
5. Min, B., Pelaschier, J. T., Graham, D. E., Tumbula-Hansen, D. & Söll, D. Transfer RNA-dependent amino acid biosynthesis: an essential route to asparagine formation. *Proc Natl Acad Sci USA* **99**, 2678–2683 (2002).
 6. Nair, N. *et al.* The *Bacillus subtilis* and *Bacillus halodurans* Aspartyl-tRNA Synthetases Retain Recognition of tRNA(Asn). *J. Mol. Biol.* **428**, 618–630 (2016).
 7. Rauch, B. J., Gustafson, A. & Perona, J. J. Novel proteins for homocysteine biosynthesis in anaerobic microorganisms. *Mol. Microbiol.* **94**, 1330–1342 (2014).
 8. Price, M. N. *et al.* Filling gaps in bacterial amino acid biosynthesis pathways with high-throughput genetics. *PLoS Genet.* **14**, e1007147 (2018).
 9. Oberhardt, M. A. *et al.* Harnessing the landscape of microbial culture media to predict new organism-media pairings. *Nat. Commun.* **6**, 8493 (2015).
 10. Dos Santos, P. C., Fang, Z., Mason, S. W., Setubal, J. C. & Dixon, R. Distribution of nitrogen fixation and nitrogenase-like sequences amongst microbial genomes. *BMC Genomics* **13**, 162 (2012).