

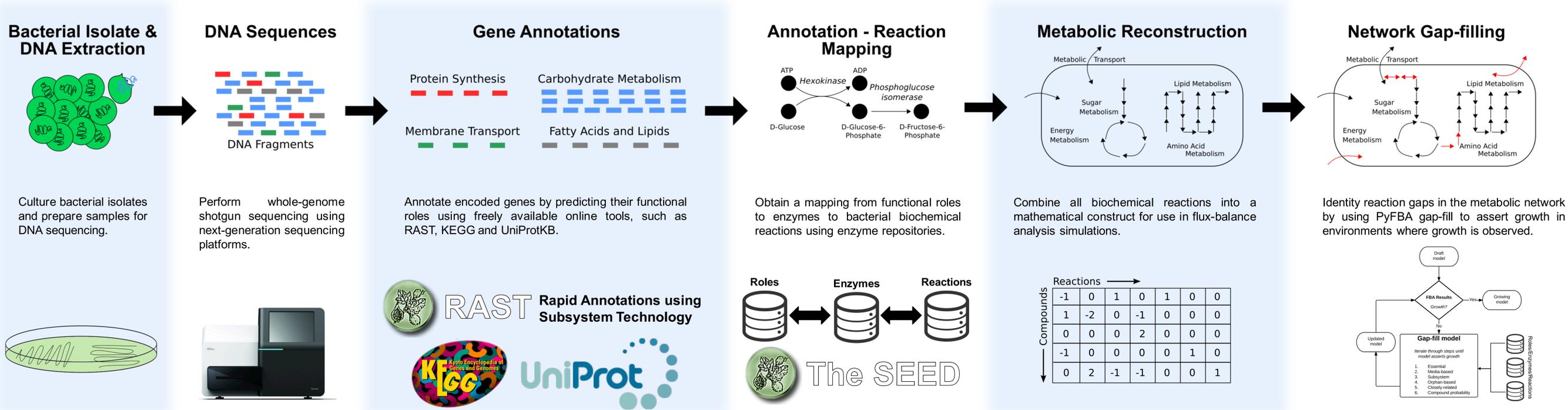
Daniel A Cuevas¹, Taylor G O'Connell², Blaire Robinson², Tucker Lopez³, Kristen Aguinaldo³, Rebecca de Wardt³, Rhaya Alkafaji³, Thiago Bruce³, Elizabeth Dinsdale³, and Robert A Edwards^{1,2,4}

¹Computational Science Research Center, ²Biological and Medical Informatics Research Center, ³Department of Biology, ⁴Department of Computer Science, San Diego State University, CA

Introduction

Genomic data has become exponentially inexpensive to generate and increasingly accessible. Limitations with traditional homology-based bioinformatics algorithms often identify in **gaps in our knowledge**. Genome-scale metabolic models offer rapid analysis and a unique perspective to genomic data by **connecting** an organism's **genome** to its **phenome**. **Can we exploit bacterial metabolic networks to fill these genomic gaps?**

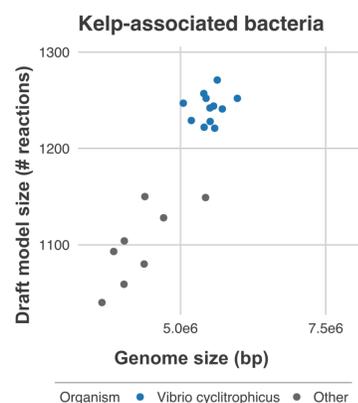
PyFBA Workflow



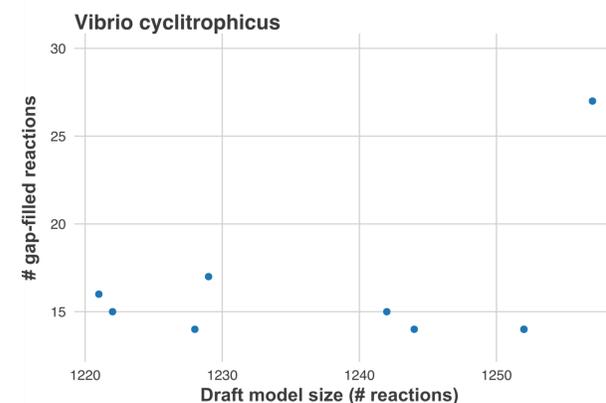
Databases

- Roles:** 4,748 functional roles
- Enzymes:** 4,067 enzyme complexes
- Reactions:** 34,700 reactions
- Bacteria:** 53 diverse bacteria
 - 5,088+ growth profiles
 - 23+ metabolic models

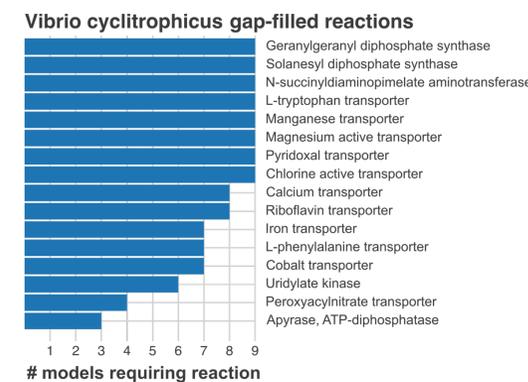
Case Study: Kelp Forest Isolates



Out of our 53 diverse bacteria, 21 are isolated genomes from various Southern California kelp forests and 12 of those genomes share strong similarities to *Vibrio cyclitrophicus*. The size of the genome correlates with the number of reactions a genome's model contains.



Eight *Vibrio cyclitrophicus* models were gap-filled on a rich-nutrient LB media source using the PyFBA gap-fill algorithm. The number of added reactions is consistent.



From the eight gap-filled models, a group of reactions were consistently added to assert growth on LB media. A total of 16 reactions were required by 3 or more models – 11 of those reactions were transporter proteins.

Conclusions

PyFBA provides capabilities to use metabolic functional annotations to build draft FBA models.

Combined with phenotypic growth data, PyFBA can predict the metabolic reactions missing from draft models using biochemical databases and customized gap-filling algorithms.

Generating more metabolic models and growth profiles will help elucidate and establish patterns that describe our missing knowledge.

