

Microbial Database

Blaire Robinson, Dwaine Smith, Daniel Cuevas, Katelyn McNair, Robert Edwards

San Diego State University



SAN DIEGO STATE UNIVERSITY

Abstract

Technological advances in genetic research have enabled scientists to sequence and study microbial genomes. Microbial genomes provide information about specific proteins, how these specific proteins function, and how these proteins are regulated to mechanistically maintain a functioning organism. There is, however, a lack of phenotypic and biochemical information associated with the microbes whose genomes have been sequenced. Phenotypic information includes optimum pH, as well as maximum, minimum, and optimal growth temperatures at which the organisms live.

Creating a database that will collect and combine phenotypic microbial data that can be utilized and accessed in a computational consumable format will help to expand the usability of microbial genomic databases. We have defined the database schema and selected organisms being studied in Edwards' lab were targeted to populate the database with phenotypic data. We are implementing community standard ontologies to facilitate data exchange with other databases. The microbial phenotype database will provide access to phenotypic data that will drive the formulation of correlations between genotype and phenotype that can be tested experimentally.

Problem

The phenotypic data of organisms has not been collected and studied much since sequencing of genomes has become efficient in its cost and easiness to acquire. There are not many databases that collect phenotypic data and the few databases out there are limited. There is also no standard ontology between various databases.

Method

First, we collected as much phenotypic and biochemical information as possible about the 27 microbial strains being studied in Edwards' lab, using various sources such as published scientific papers, existing databases, and Bergey's Manual of Determinative Bacteriology. We then set up a standard of 67 characteristics to define the individual microbial strains.

Genus species	Oxygen requirement	Cellular morphology phenotype	Gelatin hydrolysis, 22°C	Gram stain phenotype	D-Glucose, acid production
Escherichia coli	Facultative anaerobic	Straight rods, 1.1-1.5um diameter 2.0-6.0um long	Negative	Negative	Positive

We are now converting the information in our database into code using a pre-existing but limited ontology called Ontology of Microbial Phenotypes, a web-based collaboration project organized through Texas A&M University. We are collaborating with the developers to expand the ontology to include all 67 defined characteristics to establish a standard for describing microbial characteristics.

Genus species	OMP: 0000039	OMP: 0000071	OMP: 0000090	OMP: 0000191	OMP: 0000167
Escherichia coli	OMP: 0000087	OMP: 0000076 1.1-1.5um diameter 2.0-6.0um long	-1	OMP: 0000189	1

Visualization of Database

We are developing an open-access website that utilizes this microbial database and allows user-submitted expansion using the standardized ontology.

The Microbial Database
Driving the formulation of correlations between genotype and phenotype.

Search Results for "E. coli 406":

Genus species	Escherichia coli
Strain	406
Oxygen Requirement	Facultatively Anaerobic
Cellular Morphology Phenotype	Straight Rods; 1.1-1.5µm in diameter; 2.0-6.0µm long
Gelatin Hydrolysis	Negative
Gram Stain Phenotype	Gram Negative
D-Glucose, Acid Production	Positive

Edwards' Bioinformatics Lab | San Diego State University

Acknowledgements

Special thanks to Debby Siegele and James Hu at Texas A&M University.

