

Constructing a genome scale metabolic model of bacteria: What is the metabolic capability of bacteria?

PATRIC Workshop

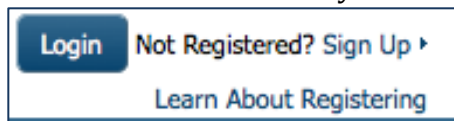
Argonne National Laboratory - June 16, 2015

A genome-scale metabolic network of chemical reactions occurring within a living organism is primarily reconstructed from protein functional annotations originally derived from literature and subsequently propagated from genome to genome by sequence similarity. When a genome is functionally annotated, its metabolic genes are mapped onto biochemical reactions. This information is integrated with data on reaction stoichiometry, subcellular localization, biomass composition, estimation of energy requirements (directionality of reactions), and other constraints into a detailed stoichiometric model of metabolism. This model can be subsequently used for detailed analysis of the metabolic potential of the organism using constraint-based modeling approaches such as Flux Balance Analysis (FBA). So what you can do with metabolic models? Using metabolic models we are able address important biological questions such as: 1.) Can the organism grow on a certain media formulation? 2.) What is the minimal media condition for particular strain? 3) Can an organism be optimized to produce an important biofuel or a medically important drug molecule? What biochemical pathways being utilized when grown under certain environment?

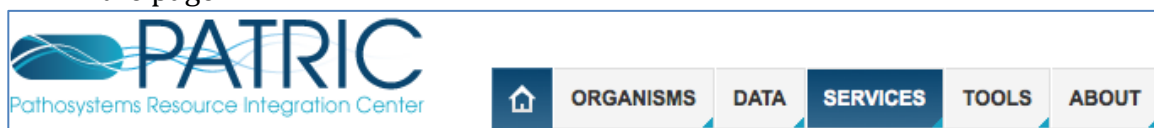
The model construction methodology in PATRIC is primarily based on the Model SEED framework[1] and the RAST genome annotation pipeline[2]. Biochemistry data used in the model construction pipeline have been gathered from multiple published metabolic models and from public databases such as KEGG, MetaCyc, EcoCyc, Plant BioCyc, Plant Metabolic Networks, and Gramene. Currently, the biochemistry database in PATRIC includes more than 13,000 biochemical reactions and over 16,000 metabolic compounds.

Constructing a genome scale metabolic model in PATRIC

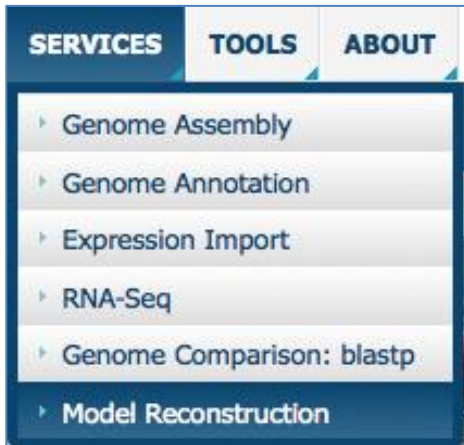
1. Login to the PATRIC website so that you can use your workspace in the downstream analysis.



2. On the PATRIC homepage (patricbrc.org), open the Services tab at the top of the page.



- Using the dropdown menu that appears click on Model Reconstruction. (Highlighted in dark blue below).



- This will open up the model reconstruction app where user will select
 - Genome – select a genome from a drop down list of publically available genomes where a model will be constructed
 - Media - select a predefined media formulation from a dropdown list of media formulations that this model be *gapfilled and simulated
 - Output Name – A user given output name for the newly constructed metabolic model.

A screenshot of a web form titled 'Reconstruct Metabolic Model'. The form has a subtitle 'Reconstruct a Metabolic Model'. It contains three main sections: 'Select a Genome', 'Optional Parameters', and two buttons at the bottom. The 'Select a Genome' section has a dropdown menu with 'e.g. Mycobacterium tuberculosis H37Rv' selected. A callout box points to this dropdown with the text 'Specify genome name.'. The 'Optional Parameters' section has a 'MEDIA' dropdown menu with 'GMM' selected and an 'OUTPUT NAME' text input field with 'Output Name' entered. At the bottom, there are two buttons: 'Reset' and 'Reconstruct'.

*Gapfilling - Draft metabolic models usually have missing reactions (gaps) due to incomplete or incorrect functional genome annotations. As a result, these models are unable to generate biomass on media where the organism typically is capable of growing. Gapfilling algorithms can be used to overcome this problem. These algorithms tentatively fill metabolic gaps by identifying the minimal number of biochemical reactions to add to the draft metabolic model to enable it to produce biomass in a specified media. Gapfilling is an optimization procedure that can produce multiple solutions

5. Upon selecting the fields click the 'Reconstruct' button to construct your metabolic model. You will notice the 'Model Reconstruction Job has been started' message in green at the bottom of the app suggesting you have successfully submitted the model reconstruction app.

Reconstruct Metabolic Model
Reconstruct a Metabolic Model

Select a Genome ⓘ

GENOME
▼ Vibrio cholerae O1 biovar El Tor

Optional Parameters ⓘ

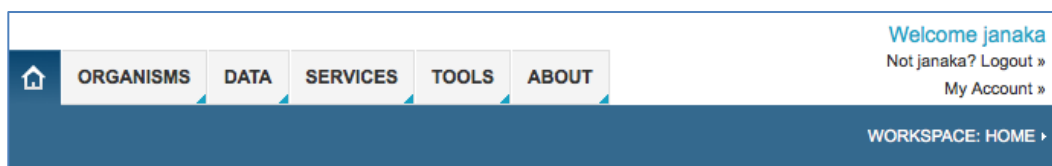
MEDIA
GMM

OUTPUT NAME
Output Name

Model Reconstruction Job has been queued.

Reset Reconstruct

6. Now you have constructed a new metabolic model from your selected genome, you can explore the details of the metabolic model such as reactions and compounds. Your newly constructed model will be stored in 'models' folder in your workspace. To navigate to your model folder click the workspace:home link at the top right corner of the page.



7. Upon selecting WORKSPACE:HOME link you will see multiple folders appear in your screen including 'models' folder where your models are stored (Highlighted with dark blue).

Name	Size	Owner	Created
Parent Folder			
Experiment Groups		janaka	6/4/15, 5:26 PM
Experiments		janaka	6/4/15, 5:26 PM
Feature Groups		janaka	6/4/15, 5:26 PM
Genome Groups		janaka	6/4/15, 5:26 PM
models		janaka	6/11/15, 4:00 PM

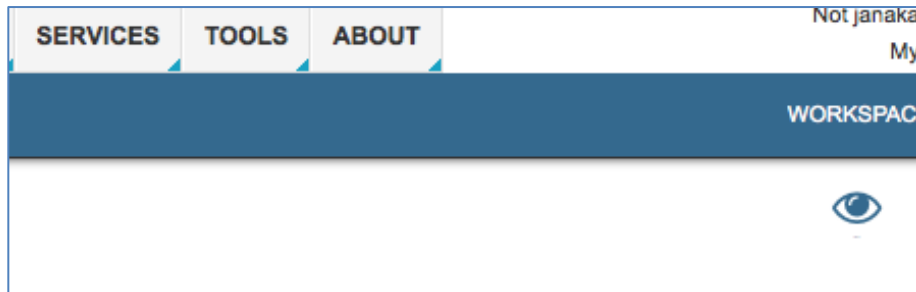
8. Select the models folder, you can find all of your metabolic model files listed by the user assigned names

Name
Parent Folder
Mico_SUBG005_model
Vibrio_cholerae_model
Vibrio_colerae_model
mycobacterium_tuberculosis_model

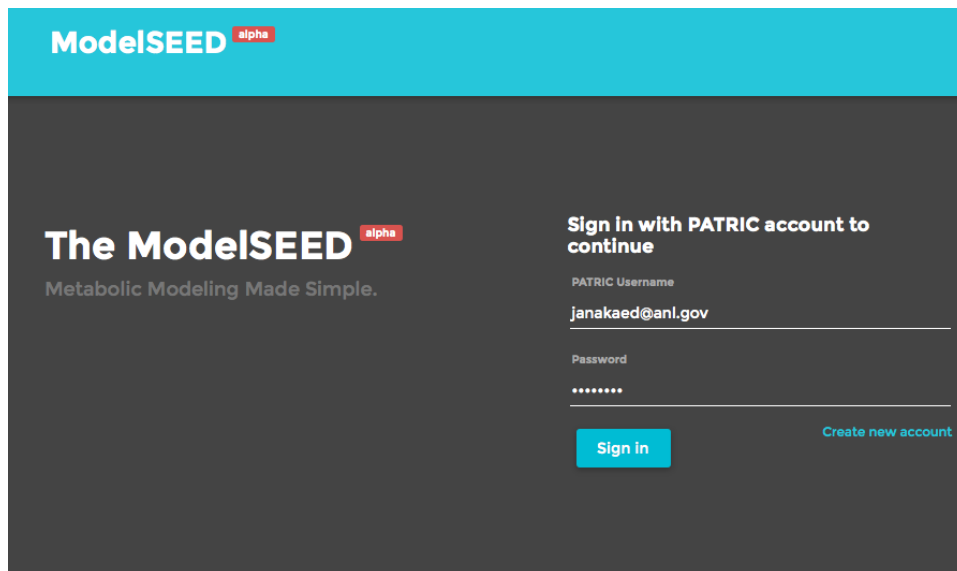
9. Select the desired model by clicking on the model file. A table will be displayed consisting of model information (such as organism name, number of genes, reactions, compounds and biomass equations in the model), and the downloadable links to a.) SBML version of the model b.) Reaction table and the c.) Compound table.

janaka / home / models / Vibrio_cholerae_model	
Metabolic Model - Vibrio_cholerae_model	
Organism	Vibrio cholerae O1 biovlar El Tor
File Name	Vibrio_cholerae_model
Reactions	1127
Compounds	1114
Genes	876
Biomasses	1
Source	PATRIC
Downloads	
File	Size
Vibrio_cholerae_model.sbml	1.6 MB
Vibrio_cholerae_model.cpdtbl	69.6 kB
Vibrio_cholerae_model.rxn.tbl	349.7 kB

10. In addition, users can explore the metabolic model in detail by a user-friendly searchable table by clicking the ‘eye’ sign that appears on the top right corner of the page.



11. You will be navigated to the modelSEED page. Login to modelSEED site if you already have an account or create a new account. Then sign in with your PATRIC credentials.



12. This will take you to a table consist of multiple tabs representing reactions, compounds, genes, compartments of the model and the weighted components of the biomass. Each column in the table is sortable by clicking the heading of the column. Any information of the model is searchable by using a search bar that appears on the top right corner of the table

My Models / Vibrio_cholerae_model

Model

Reactions		Compounds	Genes	Compartments	Biomass	Gapfilling	ETC
Show	10	entries	Reaction	Name	Equation	Genes	Search Reactions
			rxn00001[c0]	Pyrophosphate phosphohydrolase	$H_2O[c0] + PPi[c0] \Rightarrow (2) H^+[c0] + (2) Phosphate[c0]$	fig 686.9.peg.1098, fig 686.9.peg.3007	
			rxn00002[c0]	Urea-1-carboxylate amidohydrolase	$(3) H^+[c0] + Allophanate[c0] + H_2O[c0] \Rightarrow (2) NH_3[c0] + (2) CO_2[c0]$	fig 686.9.peg.1658, fig 686.9.peg.1659	
			rxn00006[c0]	hydrogen-peroxide-hydrogen-peroxide oxidoreductase	$(2) H_2O_2[c0] \Rightarrow (2) H_2O[c0] + O_2[c0]$	fig 686.9.peg.2400, fig 686.9.peg.2424	
			rxn00011[c0]	pyruvate:thiamin diphosphate acetaldehydetransferase	$2\text{-Hydroxyethyl-ThPP}[c0] + CO_2[c0] \rightleftharpoons H^+[c0] + Pyruvate[c0] + TPP[c0]$	fig 686.9.peg.2396, fig 686.9.peg.2896, fig 686.9.peg.3703, fig 686.9.peg.2897, fig 686.9.peg.932	
			rxn00016[c0]	Hexosaminidase	$H_2O[c0] + Chitobiose[c0] \Rightarrow (2) N\text{-Acetyl-D-glucosamine}[c0]$	fig 686.9.peg.743	
			rxn00029[c0]	5-Aminolevulinatase(hydro-lyase(adding 5-aminolevulinatase and	$(2) 5\text{-Aminolevulinatase}[c0] \Rightarrow H^+[c0] + (2) H_2O[c0] + Porphobilinogen[c0]$	fig 686.9.peg.2705	

References

1. Henry, C.S., et al., *High-throughput generation, optimization and analysis of genome-scale metabolic models*. Nat Biotechnol, 2010. **28**(9): p. 977-82.
2. Overbeek, R., et al., *The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST)*. Nucleic Acids Res, 2014. **42**(Database issue): p. D206-14.