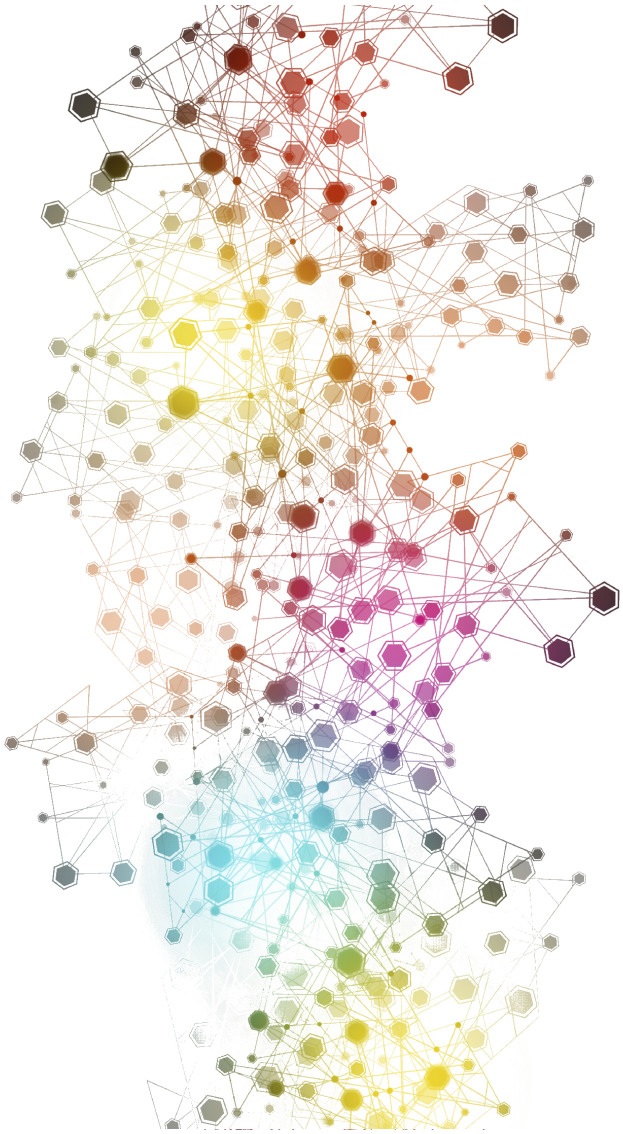


2015 Dry Lab Proposal

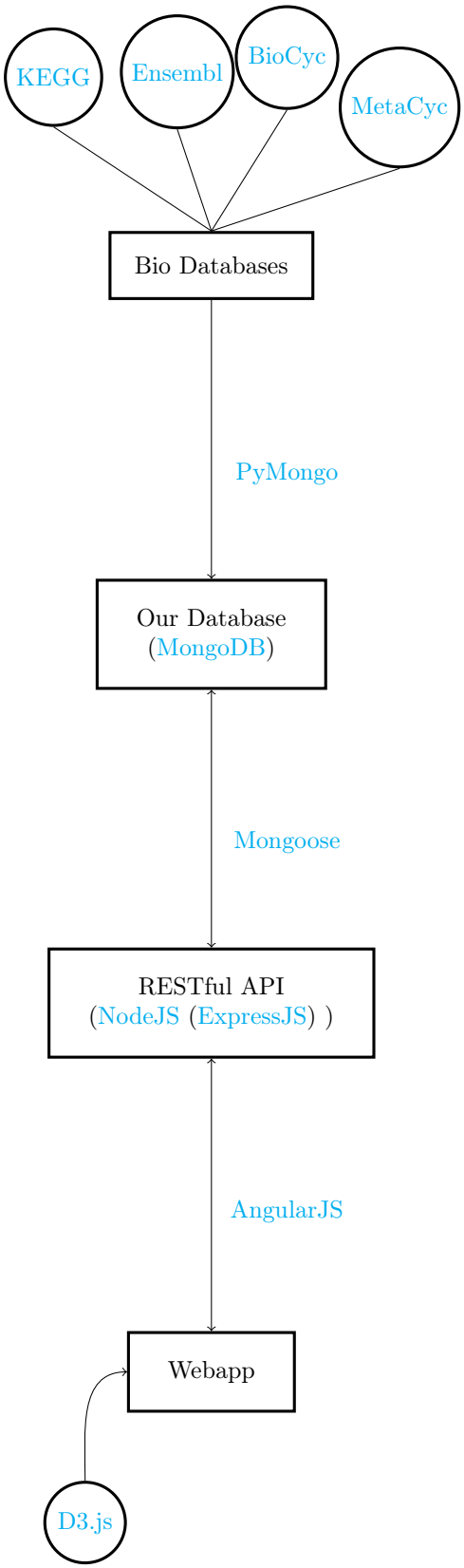
1 Abstract

With the advancement of high throughput sequencing and large-scale proteomic experiments, studying biology computationally and systematically has become extremely useful and even necessary. Rapid advances in network biology has elucidated many pathways and mechanisms important for many applications such as drug development, disease prediction and biological engineering [1]. We propose to create a genetic circuit editor which uses iGEM's Registry of Standard Biological Parts database [2]. This editor will allow the user to intuitively create genetic circuits from any part(s) listed in the registry for any species that exists in our curated database. As a genetic circuit is being built, our tool will also dynamically render a metabolic network representing the metabolome of the species being worked on. Their metabolic network will change dynamically as new parts are added to the circuit through the addition of new nodes, new links, deletion of nodes, deletion of links and changes in flux. Dynamic flux balance analysis will power all metabolic network simulations [3]. Lastly, our tool will allow the user to test their engineered species in a microbiome setting by integrating the species's metabolic network with the metabolic network of a user defined microbial community.

2 Methods

Our tool will be built on a web-interface which uses MongoDB as its database. User interface and graph rendering will be powered by the open source JavaScript library d3.js. We will develop a pipeline for curating data and for creating our database. Our pipeline will start with taking a "point-in-time" dump of the 2010 Registry of Standard Biological Parts database from iGEM's website as an XML file. Using Python, we will parse through the XML file and build our MongoDB database of the registry. We will manually download proteomes from public databases such as KEGG and Ensembl as .fasta files. Using the gene annotations of all species, we will subsequently draft a dataset of all interactions and reactions by using databases such as KEGG, Biocyc and Metacyc. We will then manually validate each reaction and will correct false reactions through manual curation. All gene interaction and reaction data will be stored in our MongoDB database. By constructing stoichiometric matrices from a set of user defined species, our tool will use flux balance analysis to simulate their metabolome.

3 Web Architecture



References

[1] http://www.nature.com/nbt/journal/v28/n12/full/nbt.1711.html

[2] http://parts.igem.org/Registry_API

[3] http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1302231/