



7-21-2016

Mutation Selection on the Metabolic Pathway and the Effects on Protein Co-evolution and the Rate Limiting Steps on the Tree of Life

Katherine S. Porter

Ursinus College, kaporter1@ursinus.edu

Follow this and additional works at: https://digitalcommons.ursinus.edu/math_sum

 Part of the [Bioinformatics Commons](#), [Biology Commons](#), and the [Mathematics Commons](#)

Click here to let us know how access to this document benefits you.

Recommended Citation

Porter, Katherine S., "Mutation Selection on the Metabolic Pathway and the Effects on Protein Co-evolution and the Rate Limiting Steps on the Tree of Life" (2016). *Mathematics Summer Fellows*. 4.

https://digitalcommons.ursinus.edu/math_sum/4

This Paper is brought to you for free and open access by the Student Research at Digital Commons @ Ursinus College. It has been accepted for inclusion in Mathematics Summer Fellows by an authorized administrator of Digital Commons @ Ursinus College. For more information, please contact aprock@ursinus.edu.

Name: Katherine Porter

Title: Mutation selection on the metabolic pathway and the effects on protein co-evolution and the rate limiting steps on the tree of life

Department: Mathematics and Computer Science

Mentor: Dr. Peter Chi and David Liberles (Department of Biology at Temple University)

Abstract

Metabolic pathways are made of a series of reactions by enzymes at different speeds. These pathways include the rate limiting step, which is the slowest step that determines the rate of the overall reaction. To date, one study has examined the pathway of glycolysis, and found no evidence of evolutionary stability of its rate limiting step. In addition, phylogenetic evidence has shown evolution in the pathway over time including gene duplication and positive selection within the pathway. This evidence suggests that there is coevolutionary selection on glycolysis. The evidence from this previous study is simulation-based. The Michaelis-Menten kinetics that describe metabolic pathways can be studied directly from a graphical perspective. This research project shows solution to the set of nonlinear system of differential equations through graphical solutions. Specifically, the graphical solutions will be expressed through design space plots that compensate for the multiple parameters.

Introduction

In the field of molecular biology and biochemistry, scientists want to understand metabolic pathways. These pathways are made of a series of chemical reactions where enzymes are interacting with substrate to synthesize into a product. For example, glycolysis is a metabolic pathway that converts glucose into pyruvate through a series of chemical reactions. The metabolic pathway glycolysis provides necessary energy for cells. In a metabolic pathway, the rate of turnover of molecules through the pathway are measured through the flux reaction. Within the flux reactions, it is best to analyze the slowest step of the reaction, the rate limiting step. The metabolic pathways' flux reaction can be modeled through the Michaelis Menten Kinetics model that describe enzyme kinetics. We want to understand how genes in the pathway evolve and their overall effect.

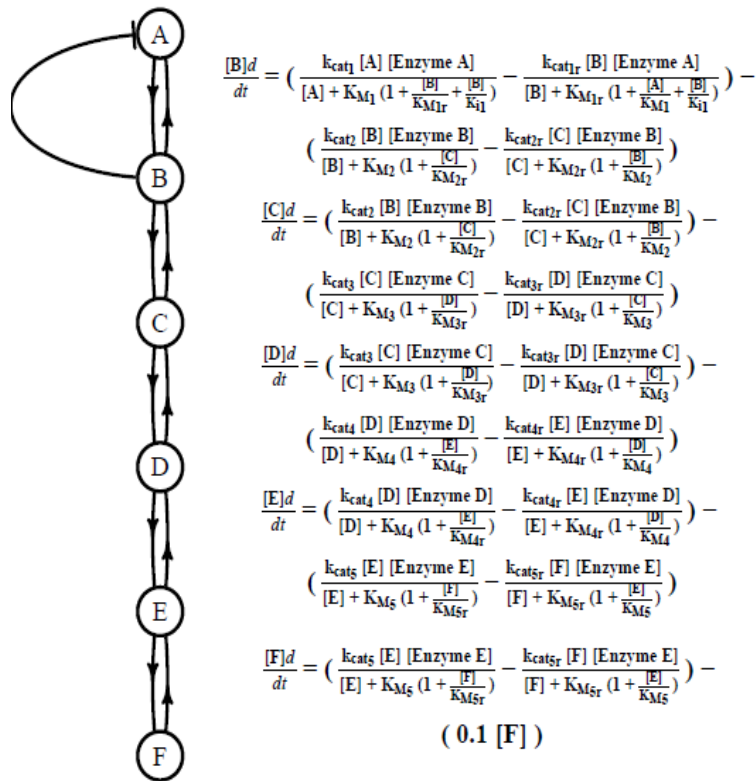
In past research, a software package called COPASI was used to simulate the biochemical pathways evolving over time. The simulations from the COPASI software showed coevolution within the metabolic pathway (Orlenko et.al 2016). In addition, the software produces a set of nonlinear differential equations. A differential equation is an equation containing derivatives of a function or functions. Researchers want to find a solution to these differential equations to further characterize how evolution affects the metabolic pathway. The set of equations contains many different parameters that affect each other. We want to find a solution that accommodate all of the different parameters in the set of differential equations. The set of nonlinear differential equation can be solved analytically, through numerical analysis, or graphically. An analytical solution would allow creates an exact solution and allows for parameters to be adjusted for evolution. In contrast, numerical analysis is an approximation and a graphical approach would be an interpretation. An analytical solution to the set of differential

equations would be the best solution due to the fact that an analytical solution would allow simple analysis of data in an equation that always works, rather than running graphical estimates every time you want to check data in this project. This particular set of differential equations does not follow a model that has been analytically solved. For example, the Lotka Volterra model also known as the predator and prey model can be solved analytically. Unfortunately this model does not follow the format of the set of differential equations we are analyzing. In addition, the metabolic pathway has too many parameters (Saint Olaf Mathematics 2010). A numerical analysis approach such as the Runge Kutta method could be used but this project focuses on a graphical approach. Specifically, this project focuses on a graphical approach using design space plots compensating for multiple parameters. Design space is made of input variables with a combination of interactions and parameters separated through boundaries. The plots express the different parameters through different colors. The change in color shows the change in parameters. In the study, there are distinct boundaries that help show the phenotypic and genotypic boundaries (Savageau et. al 2014). These types of plots has been performed in previous studies. The studies were trying to model biochemical systems with multiple parameters that needed to be expressed using design space plots (Savageau et.al 2010).

We ultimately would like to show the coevolution of proteins and their fitness over time. We will use the software Matlab that is designed as a matrix based platform that helps express computational mathematics. The program can apply to our focus in Bioinformatics. Within the Matlab program we are using the design space toolbox created in previous studies to create the design space plots (Savageau et.al 2010). Since previous studies have shown that biochemical pathways can successfully be described by the design space plots, then we hypothesize that the design space plots can characterize the metabolic pathways and show the fitness over time.

Methods

Characterizing the flux reactions within the metabolic pathway consisted of two major phases, an analytical approach and a graphical approach. The initial problem given was to solve the system of equation created from the previous study into an analytical solution (Figure 1).



(Orlenko et.al 2010)

Figure 1. The system of nonlinear differential equation equations describe the metabolic pathway and was produced using simulations from the COPASI software.

Through evaluation of the given information, we were better able to understand the problem, and as such we were better equipped to find a solution”.

The flux reactions are in the form of the Michaelis Menten equation shown below.

$$V_0 = \frac{V_{\max} [S]}{K_m + [S]}$$

<http://m.wolframalpha.com/input/?i=Michaelis+constant>

Initially Michaelis Menten equation could not be solved through an analytical solution. A mathematician did find a solution called the Lambert W function. When solving the Michaelis Menten equation analytically the Lambert W function is used to get the solution (Figure 2).

$$[S]_t = K_m \cdot W_0 \left\{ \frac{[S]_0}{K_m} \cdot \exp \left(\frac{[S]_0 - V \cdot t}{K_m} \right) \right\}$$

www.wolframalpha.com

Figure 2. The analytical solution to the Michaelis Menten equation using the Lambert W function to solve.

Using WolframAlpha, a computational knowledge engine, the solution was found using the Lambert W function that was derived and helps solve the old Michaelis Menten functions into solutions. Taking the simplest form of one of the differential equations, you substitute constants of 1 for everything except the substrate concentration [S] you find a separable differential equation. Let $[S]=x$ then we have $x(t) = -W(-e^{t+c_1})$. This was only solving one nonlinear differential equation, not a set of differential equation. Few nonlinear set of differential equation are solved analytically. The only model found among texts was the Lotka Volterra model also known as the predator prey model (Figure 3). When using the computational engine from Wolfram Alpha, the following solutions are produced from the computational engine:

| | |
|--|--|
| $\frac{dx}{dt} = x(-c y + \alpha) \quad \left \quad \frac{dy}{dt} = y(x \gamma + \delta)$ | |
| α | prey population growth parameter |
| δ | predator population extinction parameter |
| c | species interaction parameter 1 |
| γ | species interaction parameter 2 |
| x_0 | initial prey population |
| y_0 | initial predator population |

$$\int_{\tau_0}^{x[t]} \frac{1}{\tau} \frac{1}{1+W\left(-\frac{c}{\alpha} \exp\left(\frac{\gamma \tau - c_1}{\alpha}\right) \tau^{-\frac{\delta}{\alpha}}\right)} d\tau = \alpha t + c_2$$

$$y[t] = -\frac{\alpha}{c} W\left(-\frac{c}{\alpha} \exp\left(\frac{\gamma x[t] - c_1}{\alpha}\right) x[t]^{-\frac{\delta}{\alpha}}\right)$$

www.wolframalpha.com

Figure 3. The Lotka Volterra Model that helps model predator and prey model that is defined on the left table. The right shows the function and solution to the predator and prey model.

Differential equation textbooks show that there was one example called the Holling Disk equation that incorporates the predator prey model into the Michaelis Menten function unfortunately this does not apply to the format of the differential equations. Differential Equation books and resource suggest that unless the system is autonomous (Zill,2012) where it does not depend on the variable but depend on time then you should use approximation, numerical analysis, or a graphical approach.

In this project we decided to continue with a graphical approach. Since there are so many parameters to account for in the series of equations, we needed to take a unique approach to show the different parameters for flux reactions. As stated before, research has been done on design space plots and how they can describe biochemical pathways. This study could apply the design space plots to the set of metabolic pathways to create a

graphical solution. Using the software from Matlab and the design space toolbox, the plot should be able to be created. Using the examples provided from the toolbox we analyzed the steps needed to show the plots.

The initial step is to enter the series of differential equation into the program as a function. Then the variables of the initial enzyme substrate must be defined. Then the program needs to be parse the equations, enumerate the cases, and calculate the boundaries. The design space is characterized on a dominant S system, creating the condition of dominance, based on these conditions the boundaries are set. Next the number of positive and negative terms, also known as the signature are calculated. The signature is important to know for analysis. The toolbox suggests, to print the system and boundaries can be printed for each case created. Next it is suggested to test the validity of the boundaries, in some situations, it is suggested that the boundaries can be exclusive to particular cases or are incorrect. Therefore causing issues in the display for the plots. After confirming the boundaries are valid, the program can then go through the process of actually plotting. We want to analyze particular points in the cases to show the design space. Then we want to determine a certain part of space to be evaluated. The toolbox suggested that 500x500 points are evaluated. Once the space and selected points are chosen the design space is evaluated.

```
22 %
23 % The following steps are common in the analysis of design space: parse the
24 % equations, enumerate the cases, and calculate the boundaries.
25
26 - gma = dspace.parse_gma(f);
27 - cas = dspace.enumerate_cases(gma);
28 - bnd = dspace.calculate_boundaries(cas);
29
30 %%
Command Window
Error in dspace.gma.parse (line 78)
K = dspace.util.safe_permute(K,[1,3,2]); % put into standard form

Error in dspace.parse_gma (line 41)
[k,sigma,gma.x,gma.xk] = dspace.gma.parse(f,xd);

Error in demo_example_1 (line 26)
gma = dspace.parse_gma(f);
```

Figure 4. The Matlab gma error message provided in the design space toolbox.

When we used the steps following the toolbox, we ran into errors in the programming language. All of the provided examples in the toolbox were able to run and display the figures. Our system of equations in the simplest forms could not display the figure. The Matlab program found errors in the programming of the toolbox to create the plots. For example, there would always be an error in the gma function (Figure 4). The job of the gma function is to parse the set of equations. Since we could not determine the bug in the language, we believe that a different type of programming should be used. Open source software still can do the same functions as the Matlab software. For example, Python has shown function such as parsing and enumerating matrices.

Future directions

Since there still is no original design space plots for the metabolic pathways, these need to be created. Another program platform should be used and programmed to create the

metabolic pathway. Then parameters can be adjusted to show how the metabolic pathway evolves over time.

Discussion

In an effort to characterize the metabolic pathway we made an effort to characterize and find solutions to system of metabolic pathway. Ultimately we wanted characterize the metabolic pathway further and the effect it has on evolution. We found that since the pathway is a set of nonlinear differential equations and is not in the form of the Lotka Volterra model then an analytical solution cannot be found. Since the analytical solution was not a choice, we tried creating design space plots. These type of plots were one of the few plots that could accommodate all of the different parameters in the metabolic pathway. When trying to use the Matlab software there was a bug that appeared in the toolbox programming that did not accommodate the set of nonlinear differential equations. Since the software does not accommodate the needed parameters, a different program platform such as Python or R should be used to create the design space plots.

Research needs to be continued to create the design space plots. Once the plots are created we can truly see how the metabolic pathway changes over time. The plots ultimately are characterizing the metabolic pathway further. In addition, the plots can help gain understandings for characterizing evolutionary stability and population genetics over time. Ultimately the data will help characterize evolution and help better human quality of life.

Works Cited

- Fasani, R. A., & Savageau, M. A. (2010). Automated construction and analysis of the design space for biochemical systems. *Oxford Journals*.
- Fasani, R. A., & Savageau, M. A. (2014). Evolution of a Genome-Encoded Bias in Amino Acid Biosynthetic Pathways Is a Potential Indicator of Amino Acid Dynamics in the Environment. *Oxford Journals*.
- Mat, S. O. (n.d.). *Lotka-Volterra Two Species Model* . Retrieved from St. Olaf:
<http://www.stolaf.edu/people/mckelvey/envision.dir/lotka-volt.html>
- Orlenko, A. (2015). Flux control in glycolysis varies across the tree of life.
- Orlenko, A., & Liberles, D. (2016). Selection on Metabolic Pathway function in the presence of mutation-selection-drift balance leads to rate-limiting steps that are not evolutionary stable. *Biology Direct*.
- Zill, D. e. (2012). Holling Disk Equation . In *Differential Equations with Boundary-Value Problems*. Cengage Learning.