CHEM-643 Intermediary Metabolism
Final Group Research Assignment
Does natural selection operate on the amino acid composition of enzymes involved in amino acid biosynthesis?

Learning Goals
- To integrate concepts from various domains of biochemistry (genetic and metabolic regulation, metabolic pathways, protein structure and function, comparative biochemistry, molecular evolution).
- To develop and execute a strategy to address and resolve a research question.
- To demonstrate an ability to contribute productively with others toward a common goal.

Introduction
How important are each of the 26 letters in our alphabet? Could we dispense with "x" or "z"? What about "e" or "s" that we use a lot more? The French writer, Georges Perec (1936-1982), as a type of discipline and word play, wrote the novel *La Disparition* avoiding the letter "e", the most common letter. How constraining is this? (Two consecutive letters in the alphabet are not used in this problem statement.) What if we reduced usage of some letter to 50% of normal?

In a way, amino acids are to proteins as letters are to language. There are twenty common amino acids that are strung together in proteins. Different orders have different meaning. While it may seem pointless to dispense with a letter in writing, there are situations where selection for fewer of a particular amino might enable certain proteins to function better, for example, the enzymes involved in amino acid biosynthesis.

Typically in autotrophic microorganisms, each amino acid operon is repressed by its respective amino acid. Thus, for example, if lysine is present, the organism does not need the enzymes of lysine biosynthesis and they are not made. However, when lysine concentrations drop, the lysine operon is derepressed and synthesis of the enzymes in the lysine biosynthetic pathway begins. Because lysine is scarce when the organism needs to synthesize it, one might expect that the enzymes for lysine biosynthesis would be synthesized faster to the extent that they do not contain lysine. One can hypothesize that selection would favor mutations that replaced lysine with arginine and disfavor mutations to lysine codons unless lysines, in particular, served specific functions at conserved locations. Similar arguments apply for the enzymes of each amino acid biosynthetic pathway. But is this true?

Assignment
The DNA sequences of the genomes of over 1000 microorganisms are now available and accessible. Some of these organisms live in extreme environments and have metabolic pathways not found in any text. Each group is expected to select an interesting organism for which the genome has been sequenced recently and test the hypothesis in any way you want; however, your report, due Monday, December 3, must present relevant data, provide an in-depth analysis, and come to conclusions. This assignment will display your

ability to define a problem, access relevant information in genome databases, construct logical arguments, and rule out other hypotheses. Your group will present your results to the rest of the class on Monday, December 3 or Wednesday, December 5 or Thursday December 6.

Your group's final report worth 10% of each member's course grade should have the form of a co-authored scientific paper. It should have
- an informative title,
- an abstract summarizing the findings and their importance,
- a short introduction,
- a methods section,
- a results section with tables and graphs, and
- a discussion and conclusion section, followed by cited references.

The grade on the group assignment will apply to all members of a group unless there is need to modify based on peer evaluation.

Because this exercise has probably never been done for the amino acid pathway in the organism you select, your observations and conclusions will be new to science.

Getting started
The Department of Energy's Genome Institute website and their Integrated Microbial Genomes website are good places to start. You may also find this on-line article useful, Ditty et al. *PLoS Biology* **8**(8) e1000448 (2010).