

Title: Prokaryotic Comparative Genomics and Metabolic Engineering

Choong-Min Kang* and James J. Youngblom*

Affiliations:

Department of Biological Sciences, California State University, Stanislaus, Turlock, CA 95382.

* Corresponding authors: ckang1@csustan.edu, jyoungblom@csustan.edu

Type of Manuscript: *CourseSource* Lesson Manuscript

Funding & Conflict of Interest Statement: This work was supported by a grant from the Next-Generation BioGreen 21 Program (No.PJ00900705), Rural Development Administration, Republic of Korea. We do not have any conflict of interest.

Abstract Page

Current students majoring in biology often lack the bioinformatics skills required for a variety of employment opportunities in the 21st century. To address this challenge, we have developed a hands-on prokaryotic comparative genomics module. In our module, students compared genomic sequences of wild-type and a lysine (Lys)-overproducing strains of *Corynebacterium glutamicum*. Comparison of the two genomes provided information on genetic mutations that contributed to the Lys overproduction. Through this bioinformatic exercise, students learned how to assemble, annotate, and analyze prokaryotic next generation sequencing data. In the analysis portion of the module, students connected the genetic changes to metabolic outcomes that lead to Lys overproduction. Based on this information, new strains were created by specific genetic alterations designed to enhance production of Lys. The effect of the alterations was assayed. To further enhance the metabolite production, additional alterations and assays will be conducted. This module is applicable to anyone who is interested in creating any prokaryotic strains with specific metabolic changes. The requirement of the module includes two prokaryotic genome sequences to analyze and students' access to bioinformatic tools.

Scientific Teaching Context Page

Learning Goal(s): *Students will know how to analyze whole microbial genome sequencing data and apply the knowledge to bacterial metabolic engineering*

Learning Objective(s):

- Demonstrate the ability to convert raw DNA sequencing data into assemblies and annotated bacterial genome sequences
- Document the genetic changes between the two strains by comparative genomic analysis
- Predict how the genetic changes will affect the metabolism of the strain
- Clone a gene predicted to contribute to metabolic changes

1. Introduction:

Studies have shown that undergraduate classroom research experience is as effective as summer undergraduate research experience for student retention, motivation, and inspiring post-graduate interest. However, not many students have an opportunity to participate in summer research. We propose in this module exercises requiring the formation of hypotheses, experimental designs, and hands-on experiments using prokaryotic comparative genomics and bioinformatic techniques to address real scientific questions. In addition to providing students with a real research experience, this module will also introduce the bioinformatics skills required for a variety of employment opportunities in the 21st century. Bioinformatics skills especially as related to whole genome assembly and analysis are often desired skills for students graduating in biology whether heading to graduate school or the work force. We believe these subjects are currently insufficient in many undergraduate biology programs and that the need for these skills is increasing everyday.

This module investigates bacterial production of lysine. Lysine is an important additive to animal feed because it is an essential amino acid for many animals.^{1,2} Lysine is industrially produced by microbial fermentation, mainly by *Corynebacterium glutamicum*.³ To improve the efficiency of Lys production in this microbe, a lysine-overproducing strain of *Corynebacterium glutamicum* was generated by random mutagenesis. In our module, students compared genomic sequences of wild-type strain and the Lys-overproducing mutant strain of *Corynebacterium glutamicum*. As a first step, students assembled contigs of DNA sequencing data from the two strains by using the NextGene program at the Juniata GCAT-Seek server. Students then annotated the primary assembly using the RAST service at the website of National Microbial Pathogen Data Resource. After the annotation, a certain number of genes was assigned to each student to analyze and identify mutations that may have contributed to higher production of Lys in the mutant strain. Students identified these mutated genes through literature search, brainstorming sessions, and group discussions. Student achievement was evaluated by the quality of their reports describing the evidence for the involvement of candidate genes in Lys over-production. To determine whether these genes contributed to the higher production of Lys, each gene was cloned in a *Corynebacterium* expression vector through PCR, restriction digestion, and ligation. Screening of recombinant clones was performed via plasmid prep, restriction digestion analysis, and sequencing. Successful clones were introduced into wild-type *Corynebacterium glutamicum* and the effect of the clone on Lys biosynthesis was assayed by HPLC analysis.

This teaching module is intended for upper division biology major students. The approximate lab times required for this module will be 3 hours a week for 10 weeks. Prerequisite course is general genetics class. This module is applicable to anyone who is interested in creating any prokaryotic strains with specific metabolic changes. The requirement of the module includes two prokaryotic genome sequences to analyze and students' access to bioinformatic tools.

2. Scientific Teaching Themes:

- *Active Learning*: Research is active learning. In this module, students will actively seek answers to questions for which the answer is unknown and each student will be seeking the answer to a different question. Students need to form a hypothesis and research plan that will benefit from group feedback. Regular brainstorming sessions will foment shared ideas.
- *Assessment*: Students will be evaluated by the following criteria:
 1. Analysis of weekly input to student lab notebook
 2. Mid-term report on the genes analyzed by each student
 3. Summarization of appropriate scientific literature
 4. Final oral presentation on student lab achievements
 5. Comprehensive written final exam
- *Inclusive Teaching*: Since every student is assigned a certain number of genes to analyze, students are guaranteed equal participatory roles in this module. In addition, this module will introduce many female students who are routinely underrepresented to the areas of computational biology.

3. Lesson Plan:

Week	Activity
1	- Distribution of whole genome sequences of the two bacterial strains - Introduction to the contig assembly software (NextGene) - Submission of the data to NextGene
2	- Down NextGene assembly - Upload the assembly to RAST to begin automatic annotation - Assign genes to students - Identify mutations predicted to affect metabolite over-production
3	- Continue to identify mutations - Literature search, brainstorming, and group discussion for the identification of appropriate mutations
4	- Student report clone candidate genes - Design primers for PCR candidate genes
5	- PCR analysis of candidate genes, restriction enzyme digestion of PCR products and expression vector
6	- Ligation and transformation to <i>E. coli</i>
7	- Plasmid prep and restriction digestion to screen the clones
8	- Analysis of the restriction digestion and send out for sequencing
9	- Analyze the sequence and electroporation into expression host
10	- Assay for metabolite production

4. Teaching Discussion: Unavailable at this time. This module is scheduled to be taught during the fall semester, 2014.

Acknowledgments

This work was supported by Central Valley Math and Science Alliances (CVMSA), California State University, Stanislaus. This work was also supported by a grant to C.M. Kang from the Next-Generation BioGreen 21 Program (No.PJ00900705), Rural Development Administration, Republic of Korea.

References

1. Smriga, Kameishi, Uneyama, and Torii. 2002. "Dietary L-Lysine Deficiency Increases Stress-Induced Anxiety and Fecal Excretion in Rats". *The Journal of Nutrition* **132** (12): 3744–6
2. Chen C, Sander JE, Dale NM. 2003. "The effect of dietary lysine deficiency on the immune response to Newcastle disease vaccination in chickens". *Avian Dis.* **47**(4): 1346–51.
3. Hermann T. 2003. Industrial production of amino acids by *Corynebacterium* bacteria. *J. Biotechnol.* 104:155-172