



Big Data: How Big is Big? -- Using *de novo* Assembly and Annotation for Goat Rumen Bacterial and Fungal Meta Genomics as Examples

Speaker: Dr. Chet Langin, SIUC, OIT
Date/Time: Thursday, September 27, 4:00 pm
Location: LSIII Auditorium

Abstract: "Big Data" has come to mean anything too large for processing on a workstation, but how big is it? In bioinformatics, a single genome is huge, yet meta-genomics analyzes multiple genomes simultaneously. This presentation relates practical experiences of an NSF-funded project assembling and annotating DNA sequences of multiple unknown microorganisms taken from goats' rumen microbial communities in order to provide a better understanding of how to use various agricultural waste resources to produce important biomolecules. Bacterial and fungal metagenomics and DNA libraries were sequenced using the next generation sequencing technology. The rumen fungal community plays a key role in releasing carbohydrates from cellulosic biomass, and in providing nutrients and energy for these ruminant animals. Results are expected to provide genomic information on fungal communities and to have a great impact on enhancing the research/management capacity in goat farming. In this project, each DNA sequence was about 100 bases long and the sequences were mixed together for all of the unknown microorganisms. One of the goals was to assemble these short sequences into larger ones and then determine what microorganisms were represented by the DNA. A primary challenge was how to handle this "Big Data."

Follow-up Roundtable Discussion: 9:00 – 10:00 am Fri Sep 28
Morris Library Rm752-754

Questions: contact Dr. Jane Geisler-Lee geislerlee@siu.edu



Biography of Dr. Chester (Chet) Langin



Dr. Chet Langin is the OIT Research Coordinator for Southern Illinois University. As such, he supervises graduate students in the support of the SIU BigDawg high performance computing (HPC) cluster for researchers in all academic disciplines at the university, which includes meeting with faculty and other graduate students and assisting them in getting their jobs running on the BigDawg HPC cluster. Chet is an active participant of an NSF-funded national organization of supercomputing professionals called XSEDE (Extreme Science and Engineering Development Environment) and has twice been named a Fellow in the XSEDE Campus Champions program. The SIU BigDawg cluster is one of 38 XSEDE resources in the nation and has assisted numerous SIU faculty and students in their computational research. Chet's Computer Science Ph.D. research involved using Self-Organizing Maps for network intrusion detection, although now he assists research in multiple academic domains.

Financial Support for this Big Data Series Made Available by:

SIUC Office of Vice Chancellor for Research (OVCR)
Sigma Xi Science, Mathematics, Engineering Education (SMEE) Program

