



# GENI and GENI-ACT projects provide authentic undergraduate research experiences in genome analysis

L. Scott<sup>1</sup>, A. Ghrist<sup>2</sup>, B. Goodner<sup>3</sup>, K. Houmiel<sup>5</sup>, C. Kerfeld<sup>4</sup>, C. Kvaal<sup>9</sup>, P. Pratt-Szeliga<sup>6</sup>, D. Rhoads<sup>8</sup>, S. Slater<sup>7</sup>, R. Welch<sup>6</sup>, D. Wood<sup>5</sup>

Augustana College<sup>1</sup>, Eastern Iowa Community College District<sup>2</sup>, Hiram College<sup>3</sup>, Berkley Univ.<sup>4</sup>, Seattle Pacific University<sup>5</sup>, Syracuse Univ.<sup>6</sup>, Univ. Wisconsin-Madison<sup>7</sup>, California State University<sup>8</sup>, St. Cloud State Univ.<sup>9</sup>

## Introduction

**The collaboration.** The Microbial Genome Annotation Network (MGAN<sup>1</sup>) is collaborating with the Genomics Education National Initiative (GENI<sup>2</sup>) project and Openmods<sup>3</sup> (Fig. 1) to provide authentic and inexpensive research opportunities to high school and undergraduate students in the field of microbial genome analysis. The available technology includes: 1) an online toolkit that consolidates publically available bioinformatics tools into a single platform, and 2) online collaboration tools and wet-lab resources (e.g., bacterial strains, primers and plasmids) for functional genomics studies.

**Defining gene/genome annotation and functional genomics.** Gene annotation is the process of using computer technology to predict a biological function for a gene or region of a chromosome. There are many different bioinformatics tools currently available, which makes choosing among them a daunting task for a novice annotator. Functional genomics is the process of performing wet-lab experimentation to confirm the function predictions. Numerous approaches are described in the published literature. GENI and GENI-ACT<sup>4</sup> make it easier to implement an authentic research experience in genome analysis by providing training and faculty resources.

### 3 steps to an authentic research experience in genome analysis.

Step 1: Find a microbe whose genome has been sequenced, but for which little published work is available. Start with IMG or NCBI.  
Step 2: Choose an annotation approach. A few of the annotation questions that one might ask using GENI-ACT are these:

- Is the original (automated) gene call accurate?
- Have the genes in a particular region of a chromosome been accurately called?
- Does my organism-of-interest possess the genes for a particular pathway, process or cellular structure?
- What is a predicted function of a gene identified as "hypothetical"?
- Does a particular gene in the chosen bacterial genome show evidence of being acquired by horizontal gene transfer?
- Other questions. A myriad of other questions can be asked that utilize some or all of the bioinformatics tools associated with GENI-ACT.

Step 3: Collect experimental evidence to support the function prediction. GENI focuses on the complementation approach.

**Purpose.** This poster describes the GENI and GENI-ACT systems, faculty resources, and faculty development opportunities that will enable one to implement a microbial genome analysis project.

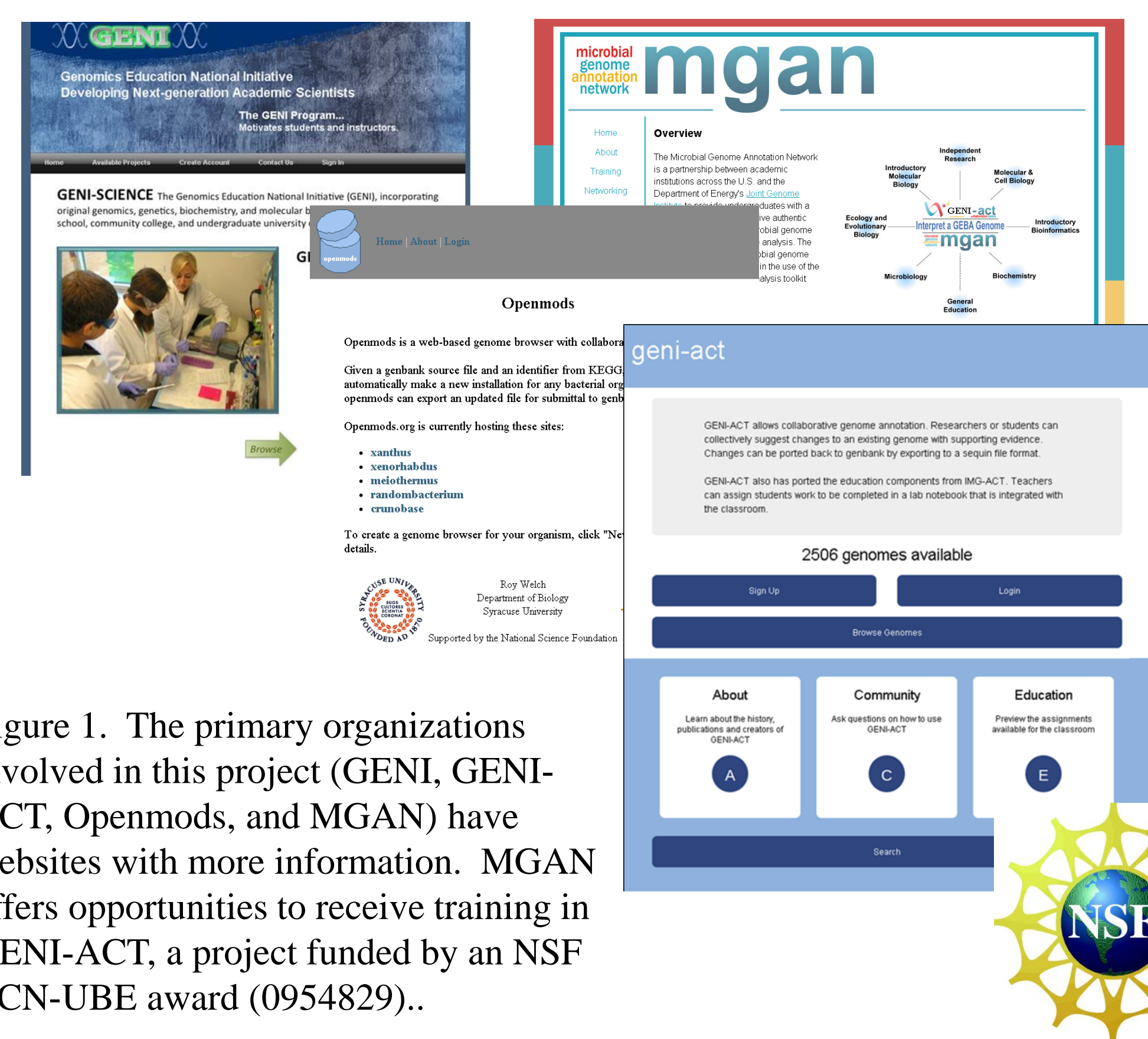


Figure 1. The primary organizations involved in this project (GENI, GENI-ACT, Openmods, and MGAN) have websites with more information. MGAN offers opportunities to receive training in GENI-ACT, a project funded by an NSF RCN-UBE award (0954829).

## Genome annotation using GENI-ACT

Previously known as the DOE Joint Genome Institute's IMG-ACT<sup>5</sup>, the GENI Annotation Collaboration Toolkit (GENI-ACT<sup>4</sup>) builds on the capabilities of the previous toolkit by: 1) expanding to include access to all sequenced genomes in GenBank; 2) allowing for public dissemination of student work; and 3) allowing for multi-instructor collaborations.

**A**

**B**

Gene Locus	Organism	Lab Notebook	Remove Annotation
Smon_0084	Streptobacillus moniliformis DSM 12112 CP001779	<a href="#">Lab Notebook</a>	<a href="#">Remove Annotation</a>
Mrub_0004	Meiothermus ruber DSM 1279 CP001743	<a href="#">Lab Notebook</a>	<a href="#">Remove Annotation</a>
EcolC_1778	Escherichia coli ATCC 8739 CP000946	<a href="#">Lab Notebook</a>	<a href="#">Remove Annotation</a>

Figure 2. GENI-ACT contains course management features (Panel A), and the ability to assign multiple genes from multiple organisms to the same students through an online lab notebook format (Panel B). Panel A is a screenshot of the course management page: students can work individually or a team; and teaching assistants have access to student work. Panel B is a screenshot of the entry into the lab notebook feature: students can be assigned one or multiple genes from different microbes; students can work individually or in teams on the same project.

**Lab Notebook**

Organism: *Streptobacillus moniliformis* DSM 12112 CP001779  
Locus: Smon\_0084

**Instructions**

- Basic Information
- Sequence-based Similarity Data
- Cellular Localization Data
- Alternative Open Reading Frame
- Structure-based Evidence
- Enzymatic Function
- Duplication and Degradation
- Horizontal Gene Transfer
- RNA
- Proposed Annotation

**Publishing Lab Notebook**

This lab notebook is private  
A public lab notebook will be visible to the general public on the gene page for this organism.  
[Make Lab Notebook Public]

The coordinate changes are private  
Publishing coordinate changes will make them visible to the general public on the gene page for this organism.  
[Make Coordinate Changes Public]

The gene products are private  
Publishing gene products will make them visible to the general public on the gene page for this organism.  
[Make Gene Products Public]

Figure 3. Screenshot of the GENI-ACT online lab notebook.

The GENI-ACT lab notebook is organized into modules that use a variety of publically available bioinformatics tools and databases (e.g., NCBI, GenBank, BLAST, CDD, T-COFFEE, TMHMM, SignalP, PSORT-B, Phobius, JGI's Integrated Microbial Genomes (IMG), TIGRFAM, Pfam, PDB, KEGG, MetaCyc, ExPASy Enzyme, IMG Gene Context, GC Heat Map, Phylogeny.fr, Rfam). Collectively, these programs and databases are used to make function predictions of putative open reading frames.

GENI-ACT has the following features:

- instructions on how to use the different bioinformatics programs
- the flexibility to ask a variety of research questions
- a modular design that allows for flexibility in using any combination of modules, as determined by the project question
- faculty development opportunities – instructors and students are learning contemporary bioinformatics tools
- student work is publishable online and available to other GENI-ACT users

## Functional genomics using GENI

As a follow-up to gene annotation studies, the GENI project provides detailed instructions for subsequent functional genomics studies, which is the wet-lab complement to the bioinformatics analysis. Panel A in Fig. 4 outlines the approach of cloning a gene-of-interest using ligase independent cloning and its subsequent complementation using an appropriate *E. coli* Keio null strain.<sup>5</sup>

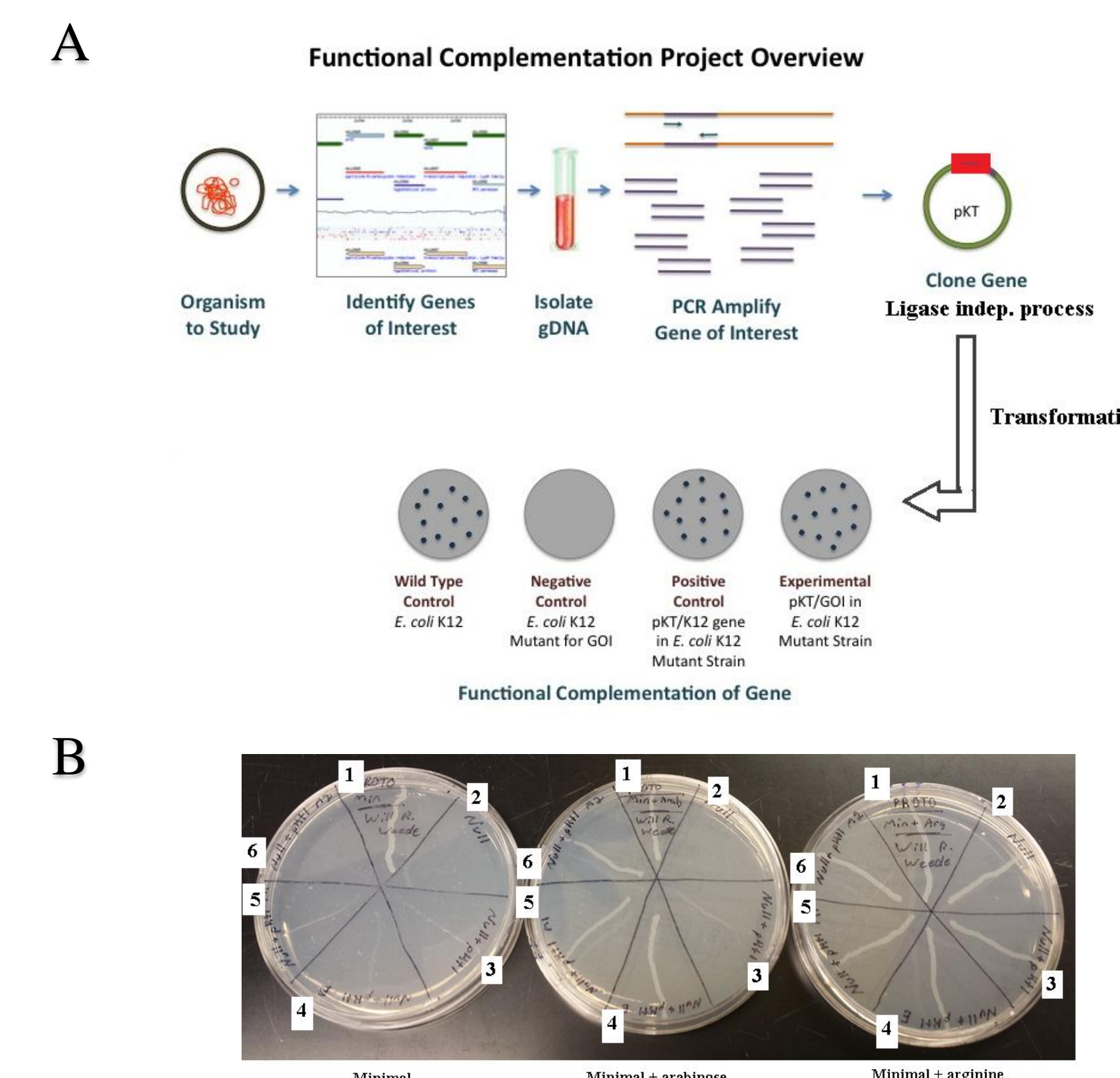
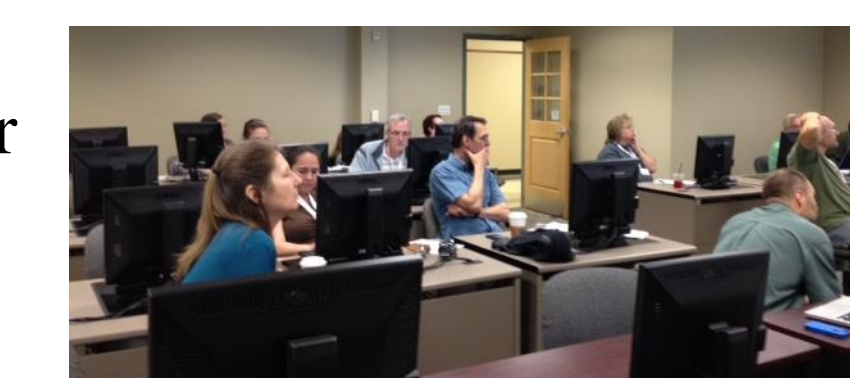


Figure 4. Overview of the ligase independent cloning and complementation approach used in the GENI system<sup>2</sup> (Panel A), and the confirmation of *Meiothermus ruber* gene Mrub\_2778 as acetylglutamate kinase/argB (Panel B) in the arginine biosynthesis pathway. Mrub\_2778 and *E. coli* argB were cloned into the pKt-1 vector and transformed into the *E. coli* argB null strain JW5553-1. Plate Sector 1 = WT *E. coli*, Sector 2 = null *E. coli* strain; Sector 3 = pKt1 in null strain; Sector 4 = *E. coli* argB-pKt1 in null strain; Sectors 5&6 = two Mrub2778-pKt1 clones in *E. coli* null strain.

## Training and contact information

MGAN provides training workshops for the GENI-ACT platform. Contact Lori Scott at loriscott@augustana.edu.



Contact Brad Goodner at GoodnerBW@hiram.edu for GENI-ACT instructor accounts and information about an upcoming functional genomics workshop in the summer of 2015 at Hiram College, OH.

Contact Katie Houmiel at houmik@spu.edu for instructional resources associated with GENI.

## Acknowledgements

Many people have contributed to the development of the tools described in this poster, but only a few could be listed due to space limitations. Please see the various websites for a more comprehensive list of collaborators.

### Literature cited

- <sup>1</sup>Microbial Genome Annotation Network (MGAN) website at <http://mgan-network.org/index.html> accessed 6/13/14
- <sup>2</sup>Genomics Education National Initiative (GENI) website at <http://www.geni-science.org/> (accessed 9/23/14)
- <sup>3</sup>Openmods website at <http://www.openmods.org/> (accessed 6/13/14)
- <sup>4</sup>GENI-Annotation Collaboration Tool (GENI-ACT) website at <http://www.geni-act.org/> (accessed 6/13/14)
- <sup>5</sup>Ditty JL, Kvaal CA, Goodner B, et al. 2010. Incorporating genomics and bioinformatics across the life sciences curriculum. *PLoS Biol.* 8(8):e1000448. <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2919421/> (Accessed 6/13/14)
- <sup>6</sup>Baba, T., T. Ara, M. Hasegawa, et al. 2006. Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol Syst Biol* 2:1-1; *E. coli* null strain provided by the Coli Stock Center (<http://cgsc.biology.yale.edu/Strain.php?ID=99802>)



