Research-Grade Bioinformatics for High School and Beyond: Identifying Mutations that Produce Adaptations in Bacterial Biofilms

Background. Imagine a mutant bacterium that grows better in a new, stressful environment than its ancestor. How exactly does this mutant differ? What gene or genes change in response to new environmental pressures, and how would you find out? How do these genetic changes produce a fitter mutant phenotype?

The ability to capture and decode vast amounts of DNA sequencing data has revolutionized our ability to identify the genetic basis of adaptations. The EvolvingSTEM evolution-in-action experiment (https://evolvingstem.org/) provides students with the opportunity to watch bacterial populations adapt to form complex biofilm communities over the course of a week-long experiment. During this experiment, they identify mutants with different phenotypes than their ancestor on petri plates. We can isolate the DNA from these mutants and sequence their complete genomes, which should differ from the ancestor at very few nucleotides. Here, you will identify and interpret the differences.

The laboratory experiment in brief. Students culture a probiotic bacterium, *Pseudomonas fluorescens* SBW25, in test tubes that contain growth media and a polystyrene bead. Bacteria colonize the bead and form a biofilm, and the biofilm-covered bead is then transferred to a new test tube with fresh media and a sterile bead. Beads are serially transferred to select for biofilm-adapted mutants. In just a few days, colonies mutate from smooth circles to wrinkly rosettes to fill diverse ecological roles. Their results simulate the process of adaptive radiation often illustrated using Darwin's finches in textbooks; moreover, their findings are medically relevant, as pathogenic bacteria involved in chronic infections often evolve to form biofilms through mutations in the same genetic pathways as the harmless bacteria used in the classroom experiment.

This bioinformatics module takes the evolution experiment to the next level by walking students through the process of identifying the genetic basis of their evolved mutants. Students analyze whole genome sequencing data from bacterial populations evolved during evolution experiments performed by high school students. This allows students to understand that changes to the appearance of bacterial colonies was the result of heritable mutations to the bacteria's DNA. Using modern software, students learn how to quickly analyze the massive amount of data produced during genome sequencing and how to identify the new mutations selected in mutant colonies.

Module Overview. Students are provided with a detailed introduction to the EvolvingSTEM laboratory experiment. This introductory information provides a step-by-step overview of the experiment, along with images from real experimental populations to visualize how the ancestral *P. fluorescens* populations evolved from exclusively smooth colony morphologies to display a range of colony morphologies that included wrinkly spreaders, small colony variants, and smooth colonies.

Students are then provided an overview of the process of whole genome sequencing, including the steps involved in extracting DNA from an interesting mutant to using Illumina whole-genome sequencing to read the DNA sequence of the mutant's entire genome.

The main activity introduces modern bioinformatic analysis techniques. Students are provided with representative images and sequencing read data (fastq files) of 10 wrinkly mutants evolved by students from two high schools: Pittsburgh Science and Technology Academy (Pittsburgh, PA, USA), and Pittsburgh Allderdice High School (Pittsburgh, PA, USA). Students select a mutant and follow a step-by-step guide to use online data analysis tools provided by the PATRIC resource (https://patricbrc.org/). They process this data with Microsoft Excel or Google Sheets to identify a list of causative mutations for their chosen mutant. Students are also directed on how to use the Pseudomonas Genome Database (http://www.pseudomonas.com/) to discover more information about their mutated genes.

Students are provided with a follow-up presentation that describes the genetic pathways involved in producing the mutant colony morphologies and the reason that particular mutations are likely to lead to increased biofilm production in their mutants. This presentation also connects the relevance of these studies to real world medical applications.

Learning Goals.

Goal 1. Students will feel more engaged with science and empowered to practice science.

The EvolvingSTEM bioinformatics module provides students with the opportunity to perform authentic scientific analyses. Students become researchers who need to integrate information from multiple disciplines, use appropriate tools to analyze their data, and communicate their findings. This allows students to build their capacity for argumentation and critical reasoning, as they struggle to understand and explain the phenomena they are studying. We believe investigations such as ours allow students to form a sense of agency to see themselves as scientists and motivate their interest in related careers.

Goal 2. Students will understand why evolution is a unifying principle of biology that is relevant to their everyday lives.

The EvolvingSTEM experiment simulates the process of adaptive radiation often illustrated in textbooks with Darwin's finches and mimics selection for traits associated with adherence that occur in biofilm-associated infections. Adaptations to biofilm growth have been shown to contribute to increased inflammation, prolonged duration of infections, increased health care expense, and worse outcomes.

The bacteria students work with in the classroom, *Pseudomonas fluorescens*, is closely related to the notorious opportunistic pathogen *P. aeruginosa*, which is responsible for >50,000 infections per year, many of which are multidrug-resistant. *P. aeruginosa* is the "P" in ESKAPE, the acronym for the six largest threats to the rising tide of antimicrobial resistance. Importantly, *P. aeruginosa* strains are known to evolve and diversify rapidly, often with worse outcomes for patients. Thus, the evolution of diverse types in biofilms of Pseudomonads may be studied with a safe species (*P. fluorescens*) but can represent the process of pathogenesis in opportunistic infections. Students can therefore connect their classroom experiments to recent findings at the interface of evolutionary biology and medicine. We hypothesize that students will see that basic scientific research impacts their everyday lives with this important example of how laboratory research can produce results that influence the treatment of common medical issues.

Goal 3. Students will demonstrate understanding of microbiology, genetics, and biotechnology through the organizing principle of evolution.

Our NGSS aligned, 3D learning approach connects the activity to the cross-cutting concepts, science and engineering practices, and disciplinary core ideas central to effective student learning. After completing this module, students will be able to demonstrate a deep understanding of fundamental life science concepts.

The following are examples of the Disciplinary Core Ideas supported by this module:

HS-LS1. Cells contain information in DNA. Feedback mechanisms maintain an organism's internal conditions in dynamic environments like biofilms.

HS-LS2. Selection in biofilm ecosystems favors adaptations to inhabit different niches.
HS-LS3. Traits may vary by mutations, which alter DNA and are inherited.
HS-LS4. Evidence of common ancestry/diversity from the founding clone, analyzing relatedness ("tree-thinking"), natural selection on biofilm adaptations, random mutation and genetic drift produce variation among populations.