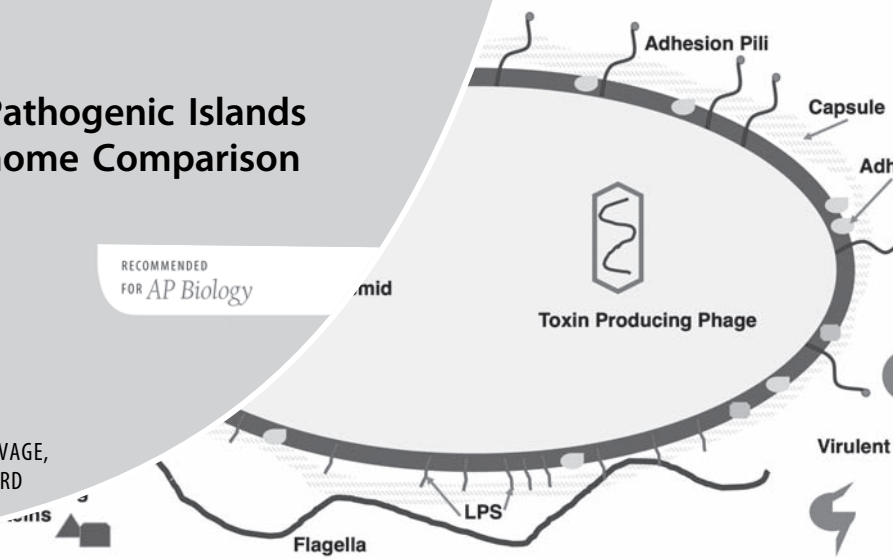


# Identifying Pathogenic Islands through Genome Comparison

RECOMMENDED FOR AP Biology

JO MARIE BACUSMO,  
JULIE BOKOR, KATHY SAVAGE,  
VALÉRIE DE CRÉCY-LAGARD



## ABSTRACT

Bioinformatics, the study of biological data using various computational techniques, is a very important aspect of biology, and its integration would greatly benefit current high school curricula. However, because most bioinformatics tools have not been readily accessible until recently, most high school instructors were not exposed to them during their formative years. We describe a bioinformatics-based module that introduces the application of genome comparison in the identification of “pathogenic islands.” The module also introduces foundational concepts of horizontal gene transfer and the genetic basis of virulence, with a special focus on antibiotic resistance – a theme teachers and students alike can easily connect and relate to. The module takes students on a journey: from conceptualizing the perfect pathogen, to an immersive experience of being a pathogen, and finally the experience of being a research scientist identifying drug-resistant genes and other virulence factors using the bioinformatics tool of genome comparison.

**Key Words:** Antibiotic resistance; pathogenic islands; virulence; gene transfer; comparative genomics; bioinformatics.

## ○ Introduction

One of the constant challenges in health care is the potential inefficacy of existing antimicrobial drugs against new and emerging pathogens (Martínez & Baquero, 2014). Most pathogens evolve at an accelerated rate, thus promoting diversification and rapid distribution of virulence genes (Cabezón et al., 2017; Logan & Weinstein, 2017). These facts place emphasis on the importance of studying the epidemiology of various virulence genes. One of the ways to approach this problem is through rapid identification of the virulence genes in pathogenic strains using bioinformatics tools such as genome comparison. The dynamic field of bioinformatics has made available

*This module presents a gateway to embracing integration of bioinformatics-based tools in the high school curriculum.*

powerful and informative tools for investigating pathogens, without having to perform wet lab experiments (Che et al., 2014). An increasing number of fully sequenced and assembled genomes allows for in-depth examination of these genomes through various bioinformatics tools. Pathogens often contain “signatures” in their genomes where accumulation of virulence and antimicrobial resistance genes are concentrated into specific regions of the genome called “genomic islands.” Genome comparison, a bioinformatics tool, allows for easy identification of these genome signatures and provides information on the various virulence genes contained within the island (Che et al., 2014). This method facilitates identification of genes that contribute to virulence and pathogenicity, thereby leading to a targeted approach in counteracting the pathogen and the disease it causes.

Advancements in research have provided a wealth of information and biological data. The field of bioinformatics has provided us with the necessary tools to effectively mine and make use of the available data, yet bioinformatics has not been readily integrated into most high school curricula. One of the primary reasons for this is the unfamiliarity of most high school teachers with bioinformatics, due to lack of training, exposure, and availability of instructional materials (Madlung, 2018). For the most part, bioinformatics is performed by computer scientists and select biologists who are comfortable with coding and use of these platforms. The module described here, “Identification of Pathogenic Islands Using Comparative Genomics Based Tools,” aims to fill that gap and provide a series of activities in which bioinformatics is seamlessly incorporated into a series of biology lessons in the context of pathogenicity and antibiotic resistance. We utilize a bioinformatics platform that is easy to adapt, integrate, and implement into the existing biology curriculum through student-centered and activity-based teaching and learning methods tailored for high school students.

## ○ Module Overview

The “Pathogenic Islands” module consists of six lessons formatted into 50-minute sessions (Table 1). It was developed by a science teacher/science researcher team to align with Florida’s *Next Generation Sunshine State Standards for Science*. It also embraces three-dimensional learning as espoused by the National Research Council (2012), which calls for students to be actively engaged in the practices of science while exploring disciplinary core ideas and crosscutting concepts. The major learning goals of the module are as follows:

1. List and describe virulence genes that cause the different modes of pathogenicity.
2. Describe the various modes of gene transfer.
3. Demonstrate the concept of gene transfer applied in the process of selection and evolution.
4. Use genome comparison tools such as PATRIC and Island Viewer to identify genomic islands.
5. Apply knowledge of virulence genes and horizontal gene transfer to identify the source of virulence in pathogenic strains.
6. Perform independent research on a set of pathogens and present the location of pathogenic islands on the genome, identify the virulence genes contained in the pathogenic islands and their function, and provide a brief history on the epidemiology of the disease caused by the pathogen.

## ○ Lesson Descriptions

### Lesson 1: Introduction

Lesson 1 introduces the real-life issue of diseases caused by pathogenic microorganisms, such as typhoid fever and bubonic plague, to stimulate student thinking. Students are presented with the various modes by which a pathogen can cause a disease. The main objective of this lesson is to equip students with the foundational knowledge and vocabulary to initiate learning and embark on this six-day journey we have set them on. Lesson 1 begins with a video of *Salmonella* invading the gut (<https://youtu.be/gpLUQ-za4uWw>), followed by a whole-class discussion of the video. The instructor prompts students to suggest features of the pathogen that contribute to its successful invasion of the host. Specific vocabulary is then introduced, and students are encouraged to help classify the virulence factors as either offensive or defensive (Figure 1).

After recording initial ideas and introducing new vocabulary, the instructor replays the video, pausing at each virulence factor presented and discussing it with the class. Once the class is familiar with a few of these virulence factors, students perform independent research on a list of virulence factors. Students are provided with a website as a starting point and invited to use additional web or text resources for answers. This information will be used in the next lesson.

**Table 1. Lesson sequencing guide and summaries. All lessons are based on a 50-minute class period and 24 students per class.**

<b>Day 1</b>	Lesson 1	Students take a pretest over the content presented in these six lessons. They then watch short videos of a pathogenic bacterium invading a host cell to identify the behaviors and biological mechanisms (virulence factors) exhibited by the bacterium that make it successful.
<b>Day 2</b>	Lesson 2	Student groups build a bacterial prototype expressing virulence factors and then compare their prototype to those created by the other groups. Students assess the potential success of each prototype by voting for the most successful and least successful prototype and justifying their choices.
<b>Day 3</b>	Lesson 3, Activity 1	Students watch a video on horizontal gene transfer and answer three questions. The teacher can choose to discuss the answers to these questions after the worksheet has been collected.
	Lesson 3, Activity 2	Students team up and play a teacher-directed game (“Pathogen Survivor”) demonstrating genome diversification via gene transfer, highlighting its impact on bacterial fitness and survival.
<b>Day 4</b>	Lesson 4, Activity 1	The video “The Power of Comparative Genomics” (7:07; <a href="https://www.youtube.com/watch?v=mU9ROpm6d70">https://www.youtube.com/watch?v=mU9ROpm6d70</a> ) introduces comparative genomics as a tool to help scientists focus their research.
	Lesson 4, Activity 2	The video “Comparison of Genomes of Eight Enterococcal <i>E.coli</i> O104:H4 Isolates” (2:07; <a href="https://youtu.be/6VTxmnZQXgU">https://youtu.be/6VTxmnZQXgU</a> ) shows how comparative genomics facilitates identification of genomic islands that contribute to the pathogenicity of disease outbreak strains.
	Lesson 4, Activity 3	Students complete video tutorials on the Pathosystems Resource Integration Center (PATRIC), a web-based comparative genomics tool.
<b>Day 5</b>	Lesson 5, Activity 1	Students work in groups, using PATRIC to research virulent genes and disease outbreaks for an assigned bacterial species.
<b>Day 6</b>	Lesson 5, Activity 2	Students present their research to the class and are graded according to a rubric (Figure 4).

## Lesson 2: Build a Pathogen

In this lesson, students build on their knowledge from Lesson 1 and design a pathogen prototype (Figure 2), indicating the host, mode of infection, and virulence factors. Students then present their work and discuss with the class what makes their pathogen a “superbug.” Here we ask the question “What makes a pathogen a pathogen?” This lesson not only reinforces logical and analytical thinking, but also stimulates creativity and imagination. Students are made to think outside the box to come up with fascinating diseases and modes of infection. Samples of student output for this activity are shown in Figure 3.

## Lesson 3: Pathogen Survivor

Lesson 3 transitions students from designing a pathogen to an immersive experience of “stepping into the shoes” of a pathogen. The driving

Offensive	Defensive
•Flagella or ability to ‘swim’	•Acid resistance
•Attachment	•Antibiotic resistance
•Injection	•Capsule or protective coating
•Toxins	
•Secretion system	

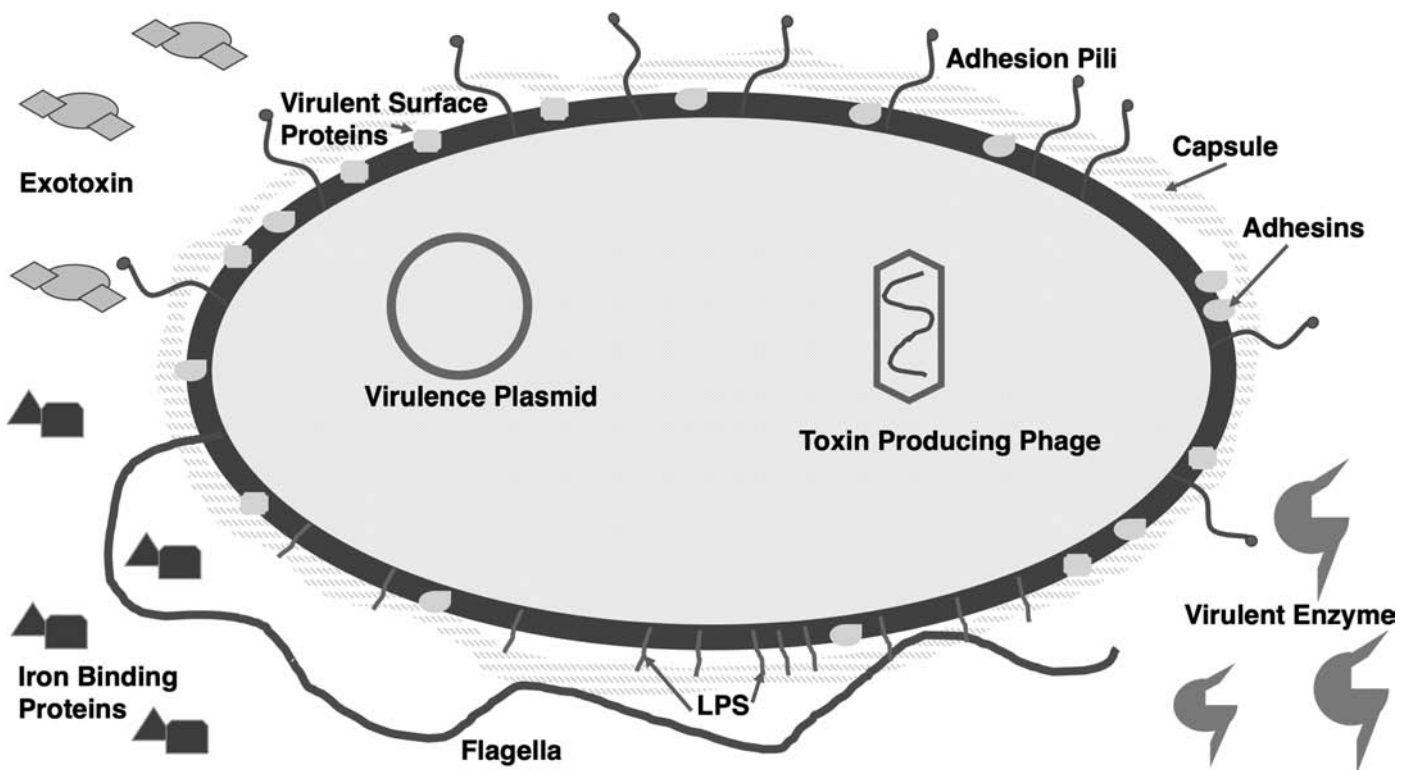
**Figure 1.** Common virulence factors designated as either offensive or defensive features.

question for Lesson 3 is “What factors influence the evolution of the pathogen?” For this lesson, we designed a game (“Pathogen Survivor”) in which each player (either an individual student or a group of students, depending on class size) represents a pathogenic species that must survive a simulated experience of a microorganism’s life cycle in overcoming survival challenges such as avoiding predators, evading a host’s immune system, and surviving external stressors like antibiotics. In Lesson 3, students come to understand that a pathogen does not have a sinister vendetta against human beings but is simply trying to survive in its environment just as all organisms do. The acquired pathogenicity is simply a consequence of the environmental pressures experienced by the pathogen. This activity leads to a group discussion on the use of antibiotics, weighing the pros and cons: exposure to antibiotics increases the probability of developing resistance, while not using antibiotics at all will increase mortality from the disease.

This lesson also introduces the basic modes of gene transfer and the limits of bacterial genome sizes. Students perform the various modes of gene transfer, including conjugation, transformation, and transduction, in order to diversify their genome with better genes that allow them to adapt to environmental stressors while maintaining a reasonable genome size. Lesson 3 enhances learning and retention because students are immediately applying concepts on pathogenesis and virulence and seeing them in action. Further, these concepts are given real-life relevance in the form of their quest for survival as the ultimate pathogen.

## Lesson 4: PATRIC & Island Viewer

Lesson 4 introduces the basic concepts of comparative genomics using videos to illustrate how comparative genomics has been used



**Figure 2.** Example of a pathogen prototype displaying common virulence factors (adapted from <https://preview.tinyurl.com/yc89aw4g>).



**Figure 3.** Examples of student output from the “Build a Pathogen” activity.

**Poster and Presentation Rubric**

**Bacteria Name** \_\_\_\_\_  
**Group Members Names** \_\_\_\_\_

All points earned according to this rubric assume accuracy of content and can be verified by the teacher.

Criteria	Excellent (3)	Good (2)	Fair (1)	Poor (0)
<b>Poster (9 points)</b>	<input type="checkbox"/> Content is well organized <input type="checkbox"/> Excellent balance of text, graphics and color <input type="checkbox"/> No misspellings	<input type="checkbox"/> Content is reasonably organized <input type="checkbox"/> Good balance of text, graphics and color <input type="checkbox"/> No more than one misspelling	<input type="checkbox"/> Minimal organization of content <input type="checkbox"/> Fair balance of text, graphics and color <input type="checkbox"/> No more than two misspellings	<input type="checkbox"/> Lacking in organization (cluttered) <input type="checkbox"/> Poor balance of text, graphics and color <input type="checkbox"/> Many misspelled words
<b>Bacteria (12 points)</b>	<input type="checkbox"/> Morphology is complete <input type="checkbox"/> Photo enhances content <input type="checkbox"/> Genomic information is complete <input type="checkbox"/> Genomic Islands are identified and thoroughly discussed	<input type="checkbox"/> Morphology is missing one item <input type="checkbox"/> Photo relates to content <input type="checkbox"/> Genomic information is missing one item <input type="checkbox"/> Genomic islands are identified and discussed	<input type="checkbox"/> Morphology is missing two items <input type="checkbox"/> Photo only tangentially related to content <input type="checkbox"/> Genomic information is missing two items <input type="checkbox"/> Genomic islands are minimally discussed	<input type="checkbox"/> Morphology is missing more than two items <input type="checkbox"/> Missing photo or graphic <input type="checkbox"/> Genomic information is mostly incomplete <input type="checkbox"/> Genomic Islands are mentioned, but not discussed
<b>Disease (9 points)</b>	Thorough discussion of the following topics: <input type="checkbox"/> Pathology <input type="checkbox"/> Antibiotics <input type="checkbox"/> Epidemiology	Good discussion of the following topics: <input type="checkbox"/> Pathology <input type="checkbox"/> Antibiotics <input type="checkbox"/> Epidemiology	Fair discussion of the following topics: <input type="checkbox"/> Pathology <input type="checkbox"/> Antibiotics <input type="checkbox"/> Epidemiology	Poor or no discussion of the following topics: <input type="checkbox"/> Pathology <input type="checkbox"/> Antibiotics <input type="checkbox"/> Epidemiology
<b>Story (6 points)</b>	<input type="checkbox"/> Outbreak Information is complete	<input type="checkbox"/> Outbreak information is partially complete	<input type="checkbox"/> Outbreak information is mostly incomplete	<input type="checkbox"/> Missing outbreak information
<b>Resources (9 points)</b>	<input type="checkbox"/> Evidence of primary resource that relates directly to lesson content <input type="checkbox"/> Main points on IslandViewer Image are complete and easy to identify <input type="checkbox"/> Typed bibliography in APA format attached	<input type="checkbox"/> Evidence of primary resource that is mostly relatable to the content <input type="checkbox"/> Main points on IslandViewer Image are incomplete <input type="checkbox"/> Typed bibliography attached, but format is inconsistent	<input type="checkbox"/> Primary resource is only tangentially related to lesson content <input type="checkbox"/> Main points on IslandViewer Image are missing <input type="checkbox"/> Bibliography attached, but not typed and inconsistent formatting	<input type="checkbox"/> No primary resource <input type="checkbox"/> Missing IslandViewer Image <input type="checkbox"/> Missing bibliography
<b>Oral Presentation (9 points)</b>	<input type="checkbox"/> Roles are clearly defined <input type="checkbox"/> Each student seems confident and knowledgeable and correctly pronounce all words <input type="checkbox"/> Time limit is met, but not exceeded	<input type="checkbox"/> Roles mostly well defined <input type="checkbox"/> Three students seem confident and knowledgeable <input type="checkbox"/> Time is within 2-min window	<input type="checkbox"/> Roles are not clearly defined <input type="checkbox"/> Two students seem confident and knowledgeable <input type="checkbox"/> Time is within 4-min window	<input type="checkbox"/> Students do not seem to know their role <input type="checkbox"/> An obvious and heavy reliance on one student for knowledge <input type="checkbox"/> Time is within 5-min window
<b>Total Points</b>				

Points Earned: \_\_\_\_\_ /54

Bibliography Grade: \_\_\_\_\_ /  
 IslandViewer Data Table Grade: \_\_\_\_\_ /  
 Total Points Earned: \_\_\_\_\_ / (+ 54)

**Figure 4.** Rubric used to assess student performance in the independent research and presentation.

to identify virulence factors in a pathogenic strain. Next, a video from the *New England Journal of Medicine* compares eight strains of *Escherichia coli* to show the presence of a genomic island causing the pathogenicity of the outbreak strain. A closer examination shows that the virulence factors responsible for the pathogenicity is a set of genes from *Shigella* transferred into *E. coli*. These videos prompt the inquiry of how scientists compare genomes. We then proceed with a brief introduction and demonstration of PATRIC (Pathosystems Resource Integration Center; Wattam et al., 2017). Students learn that the PATRIC database is one of many resources that can be used to perform comparative genomics, allowing them to do research similar to what they saw in the video. Students are introduced to the Genome Comparison tool in PATRIC, using the genomes of antibiotic resistant and nonresistant strains of *Staphylococcus aureus* as an example. Lesson 4 introduces a simple but powerful bioinformatics tool that allows students to find the pathogenic genes they learned about in the previous lessons.

### Lesson 5: Independent Research

In this lesson, students apply everything they have learned from the previous lessons and perform independent research on the history of an assigned pathogen, comparing two bacterial strains to find the genes responsible for its pathogenicity. Students then present their research to the class, which allows for an open discussion on the various modes and causes of pathogenicity. Here students get a better understanding of the nature of disease outbreaks and the contributing factors that may lead to an outbreak. Students are assessed using the rubric shown in Figure 4.

### Implementation

Lessons from this module were implemented in professional development programs for high school teachers. Afterward, a number of teachers expressed interest in integrating select lessons from the module in their classes. This module presents a gateway to embracing integration of bioinformatics-based tools in the high school curriculum. We hope that more projects like this will result in further progress toward seamless adoption of bioinformatics in various curricula.

### Additional Resources

The “Pathogenic Islands” module can be found in its entirety here: <https://bit.ly/2ZqpnTq>.

### Acknowledgments

We thank Houda Darwiche for help and support in the completion of this project, and Geoffrey Hutinet for assisting in the design and mechanics of the “Pathogen Survivor” game. This project was funded by the National Science Foundation (MCB 1412379). We would also like to thank the participants of UF CPET CATALySES and the 2016 MIT CBMM Summer Workshop for their invaluable feedback in improving the module.

### References

- Cabezón, E., de la Cruz, F. & Arechaga, I. (2017). Conjugation inhibitors and their potential use to prevent dissemination of antibiotic resistance genes in bacteria. *Frontiers in Microbiology*, 8, 2329.
- Che, D., Hasan, M.S. & Chen, B. (2014). Identifying pathogenicity islands in bacterial pathogenomics using computational approaches. *Pathogens*, 3, 36–56.
- Logan, L.K. & Weinstein, R.A. (2017). The epidemiology of carbapenem-resistant Enterobacteriaceae: the impact and evolution of a global menace. *Journal of Infectious Diseases*, 215(Supplement 1), S28–S36.
- Madlung, A. (2018). Assessing an effective undergraduate module teaching applied bioinformatics to biology students. *PLoS Computational Biology*, 14, e1005872.
- Martínez, J.L. & Baquero, F. (2014). Emergence and spread of antibiotic resistance: setting a parameter space. *Upsala Journal of Medical Sciences*, 119, 68–77.
- National Research Council (2012). *A Framework for K–12 Science Education*. Washington, DC: National Academies Press.
- Ventola, C.L. (2015). The antibiotic resistance crisis. *Pharmacy and Therapeutics*, 40, 277–283.
- Wattam, A.R., Davis, J.J., Assaf, R., Boisvert, S., Brettin, T., Bun, C., et al. (2017). Improvements to PATRIC, the all-bacterial bioinformatics database and analysis resource center. *Nucleic Acids Research*, 45, D535–D542.

JO MARIE BACUSMO is a Lecturer in the Department of Microbiology and Cell Science, University of Florida, Gainesville, FL 32611. JULIE BOKOR is Associate Director of the Center for Precollegiate Education, University of Florida; e-mail: julie@cpet.ufl.edu. KATHY SAVAGE is a Chemistry and Bioscience Teacher at Oviedo High School, Oviedo, FL 32765. VALÉRIE DE CRÉCY-LAGARD is a Professor in the Department of Microbiology and Cell Science, University of Florida; e-mail: vcrecy@ufl.edu.