Sequence Similarity Resource Adaptation:
Exploring Ebola Virus
By William Tapprich

<table>
<thead>
<tr>
<th>Virus</th>
<th>Accession</th>
<th>GP Accession</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ebola virus 1 outbreak in North Kivu and Ituri Provinces</td>
<td>M5007941.1</td>
<td>AYN74157.1</td>
</tr>
<tr>
<td>Ebola virus 2 outbreak in North Kivu and Ituri Provinces</td>
<td>M5007940.1</td>
<td>AYN74148.1</td>
</tr>
<tr>
<td>Ebola virus 3 outbreak in North Kivu and Ituri Provinces</td>
<td>M5007943.1</td>
<td>AYN74168.1</td>
</tr>
</tbody>
</table>

Module Description:
This week’s featured resource is an adaptation of:


*Note: Kleinschmit et al. (2019) was previously featured as a QUBES Resource of the Week. Learn more about this resource [here](https://doi.org/10.24918/cs.2019.5).*

The first three exercises of Kleinschmit et al. (2019) were assigned without extensive modification as homework to provide students with an introduction to the general concept of similarity, algorithms for quantifying sequence similarity, and web-based tools (e.g., BLAST).

These homework assignments prepared students for a fourth exercise developed specifically for this adaptation. The open-ended Ebola inquiry exercise explores the phylogenetic relationship between Ebola virus strains. It addresses the hypothesis that the ongoing 2018-2019 Ebola outbreak in the Democratic Republic of Congo (DRC) is caused by a known viral strain rather than a new strain. This information is important for vaccine development. Using skills acquired from working the first three exercises, students are asked to address the hypothesis on their own given reference sequences for the viral glycoprotein from all of the known viral strains as well as four sequences from individuals infected in the ongoing DRC outbreak.

Teaching Setting:
This adaptation was developed and implemented in a senior level virology course with ~30 students. Students in this course generally have varied experience with bioinformatics, but all have completed an introductory biology laboratory that includes a simple sequence alignment and ORF finding. There is an accompanying virology laboratory, taught as a CURE, that is not required. Students who opt into the lab portion of the course receive extensive bioinformatics experience including de novo sequence assembly, BLAST, and
The first three exercises were assigned over a three week period. Approximately 30 minutes of in-class time during each of the three weeks was devoted to discussion and completion of the exercise worksheets. Students were allowed to work in groups or on their own. Most students worked in groups. Most students came to class with the worksheets complete and used the time to discuss their answers with colleagues. There were a fair number of students that had misconceptions and needed the in-class time to correct their responses. Several students asked me questions directly. This was helpful for identifying misconceptions and roadblocks. I brought the entire class together to address these issues when they occurred. For students at this level, 30 minutes of class time was sufficient to address concepts, demonstrate tools and correct misconceptions. Following class discussion each week, students turned in their worksheets online for credit. The open ended inquiry exercise 4 was assigned in the fourth week as homework to be turned in online. During class in the fourth week, following the deadline for submitting the homework, students were asked to share their results.

Citation:
Related Materials and Opportunities:

This adaptation was developed by the author during the Bring Bioinformatics to Your Biology Classroom 2019 Faculty Mentoring Network (FMN). In this FMN, participants were asked to adapt and implement one or more of the bioinformatics exercises from Kleinschmit et al. (2019). Participants received support and tips from the lead author of these exercises, Adam Kleinschmit, over the course of the FMN and upon completion, received recognition as Network for Integrating Bioinformatics into Life Science Education (NIBLSE) Education Scholars. Two additional adaptations of the Kleinschmit et al. (2019) resource were produced during the FMN - one with a focus on comparing histone sequences in plants and the other with a focus on comparing human and mouse myostatin proteins.

If you are interested in adapting Kleinschmit et al. (2019) for use in your classroom, please contact Adam Kleinschmit (akleinschmit@dbq.edu) and William Tapprich (wtapprich@unomaha.edu) for a pre-/post-class assessment instrument aimed at helping the authors quantify student impacts. Faculty interested in becoming involved in this assessment instrument can also directly fill out this brief survey to join the project.

QUBES FMNs bring a small group of like-minded educators together to discuss pedagogy; share resources; adapt, implement, and share teaching materials; and create a community around a given topic. If you are interested in participating in an FMN, apply today for one of the Fall 2019 FMNs:

**Biodiversity Literacy in Undergraduate Education (BLUE) FMN**
Do you want to learn more about the data coming from natural history collections? Are you interested in adopting modules that use natural history collections and biodiversity data? Are you looking for data driven exercises and modules that can help deliver ecology and evolutionary biology content? Do you have a data driven exercise based on collections data that you would like to tweak and assess? Apply by July 22 for the BLUE FMN.

**NEON Data Education Fellows FMN**
Interested in adding quantitative reasoning and ecological data to your classroom? Join the NEON Data Education Fellows FMN to implement existing educational materials using data from the National Ecological Observatory Network (NEON). Topics range from plant phenology to earth-atmosphere gas exchange to ecological disturbance. Applications are due July 22.

**Agent/Individual-Based Modeling FMN**
Do you teach a modeling course, and want to include agent-based models (traditionally called 'individual-based models' in ecology), but don’t have a lot of experience with them yet? Would you like an approachable way to introduce math-phobic students to real modeling in your intro or ecology course? Join the Agent/Individual-Based Modeling FMN! Applications are due August 16.

**Teaching with R in Undergraduate Biology FMN**
This FMN is intended for undergraduate biology instructors with prior R programming experience who are interested in learning ways to teach R effectively to students with little to no programming experience. Participants will focus on developing, implementing, and sharing modules for teaching statistical and biological concepts in R with Swirl. Swirl lessons simplify the R learning process by providing a guided, interactive experience through on-screen prompts and exercises which students answer directly in the RStudio console. Apply by August 16.

**Genome Solver Faculty Training in Basic Bioinformatics**
In this FMN, participants will learn about basic bioinformatics tools and resources, all available for free, and most available on the web. We’ll take a look "under the hood" to explain how the tools work, but no coding or scripting is required. This entry-level course can get you started in bioinformatics with a friendly community. Applications are due August 1.
BioQUEST is a transformative, collaborative community empowering educators to drive innovation in STEM education for all students.