You've Caught Them All...Now What?

Using Pokémon to Teach Bioinformatics

Janelle Nunez-Castilla[1], Emily J. Warschefsky[1], Laura Timm[1*]

Affiliations:
[1]Florida International University, 3000 NE 151st St, MSB330, North Miami, FL 33181
*Florida International University, 3000 NE 151st St, MSB330, North Miami, FL 33181
ltimm004@fiu.edu

Type of Manuscript: *CourseSource* Lesson Manuscript

Funding and Conflict of Interest: None of the authors have a financial, personal, or professional conflict of interest related to this work.

Title and Description of Primary Image: You've Caught Them All...Now What? Using Pokémon to Teach Bioinformatics. The image included is the tree of life (Fig 1). This image represents the union between biological data and computational resources, in short, bioinformatics.

Abstract

The accumulation of genomic sequence data has revolutionized the field of phylogenetics from one that relies on morphological data to one that relies on sequence data, and consequently computational tools, to reconstruct evolutionary histories. In order to address increasingly complex evolutionary questions, bioinformatic tools are essential. In fact, as biology progresses more towards becoming a big-data science, the use of bioinformatics is becoming ubiquitous across disciplines. Despite the prevalence of bioinformatics in science today, the importance of computational tools as they pertain to biology is not often discussed at the high school or introductory biology level. In this lesson, a series of activities are presented to help highlight this importance. Using Pokémon, students first reconstruct a phylogeny using morphological characteristics, illustrating the drawbacks of inferring evolutionary relationships with this method. When provided with fictitious DNA sequence data belonging to a Pokémon, students use web-based tools to assemble a dataset of closely related sequences, generate a sequence alignment, and reconstruct a phylogenetic tree. Prior to computationally aligning the sequences in their dataset, students perform an exercise to manually align a small set of sequences. After phylogeny reconstruction, students investigate the species in their tree and place their Pokémon in the tree of life. This lesson plan aims to introduce students to bioinformatics through authentic, inquiry-driven activities and provides formal and informal assessments at several points throughout to ensure completion of these aims.

Article Context

Course

- o Biochemistry
- o Cell Biology
- o Developmental Biology
- X Genetics
- o Microbiology
- X Molecular Biology
- X Introductory Biology
- X Bioinformatics
- X Evolution
- o Ecology
- o Anatomy-Physiology
- o Neurobiology
- o Plant Biology
- X Science Process Skills

Course Level

- X Introductory
- o Upper Level
- o Graduate
- X High School
- o Other

Class Type

- o Lecture
- X Lab
- o Seminar
- o Discussion Section
- o On-line
- o Other

Audience

- X Life Sciences Major
- X Non-Life Science Major
- X Non-Traditional Student
- X 2-year College
- X 4-year College
- X University
- X Other

Class Size

- X 1 – 50
- o 51 – 100
- o 101+

Assessment Type

- X Assessment of individual student performance
- o Assessment of student groups/teams
- o Assignment
- o Exam/quiz, in class
- o Exam/quiz, take home
- o Homework
- o Answer clicker-type question(s)
- o Answer essay question(s)
- o Answer fill in the blank question(s)
- o Answer multiple choice question(s)
- o Answer short answer questions(s)
- o Answer true/false question(s)
- o Create a concept map
- o Create a diagram, drawing, figure, etc.
- o Create a website
- o Create graph, table etc. to present data
- o Design an experiment or research study
- o Design/present a poster
- o Give an oral presentation
- o Informal in-class report
- o Interpret data
- o Order items (e.g. strip sequence)
- X Participate in discussion
- o Peer evaluation
- X Post-test
- X Pre-test
- o Produce a video or video response
- o Respond to metacognition/reflection prompt
- X Self-evaluation
- o Solve problem(s)
- o Written assignment: One minute paper
- o Written assignment: Brochure
- o Written assignment: Essay
- o Written assignment: Figure and or figure legend
- o Written assignment: Lab report
- o Written assignment: Literature review
- o Other

Lesson Length

- o Portion of one class period
- X One class period
- o Multiple class periods
- o One term (semester or quarter)
- o One year
- o Other

Key Scientific Process Skills

- o Reading research papers
- o Reviewing prior research
- o Asking a question
- X Formulating hypotheses (how confident are you in the validity the morphological tree?)
- o Designing/conducting experiments
- X Predicting outcomes (when shown tree of life with blanks, where do you think your Pokémon fits in?)
- o Gathering data/making observations
- X Analyzing data (MAFFT alignment and tree-building)
- X Interpreting results/data (What is your Pokémon related to? Where does it fit in the larger tree?)
- X Displaying/modeling results/data (Place your Pokémon in the tree)
- X Communicating results (insert Pokémon into the tree of life)

Pedagogical Approaches

- o Think-Pair-Share
- o Brainstorming
- o Case Study
- o Clicker Question
- X Collaborative Work (work in a group)
- o One Minute Paper
- o Reflective Writing
- o Concept Maps
- X Strip Sequence
- X Computer Model (build a phylo tree using a computer)
- o Physical Model
- X Interactive Lecture (discussion-based lecture combined with group/individual activities)
- X Pre/Post Questions (survey before and after)
- o Other

Bloom's Cognitive Level (based on learning objectives & assessments)

- o Foundational: factual knowledge & comprehension
- X Application & Analysis (morphological data & molecular data to assess relationships)
- X Synthesis/Evaluation/Creation (morphological data & molecular data to determine relationships and generate a novel phylogeny)
- o Principles of how people learn
- X Motivates student to learn material
- X Focuses student on the material to be learned (requires small group/individual work on computers)
- o Develops supportive community of learners
- o Leverages differences among learners
- X Reveals prior knowledge (basic knowledge of how organisms are related)
- X Requires student to do the bulk of the work (computer work is largely done independently)

Vision and Change Core Concepts

- X Evolution (phylogenies!)
- o Structure and Function
- X Information flow, exchange and storage (FASTA files, alignment, stored in BLAST database)
- o Pathways and transformations of energy and matter
- o Systems

Vision and Change Core Competencies

X Ability to apply the process of science (prior information [morphology], hypotheses [confidence in tree], data generation and analysis [fasta alignments], results & interpretation)
o Ability to use quantitative reasoning
X Ability to use modeling and simulation (phylogenetic trees simulate multiple evo relationships)
X Ability to tap into the interdisciplinary nature of science (biology + computers = bioinformatics)
o Ability to communicate and collaborate with other disciplines
o Ability to understand the relationship between science and society

Key Words: List 3 to 10 key words that are relevant for the Lesson (e.g. mitosis; meiosis; reproduction; egg; etc.)

1. Bioinformatics
2. Evolution
3. DNA Sequence
4. Phylogeny
5. Morphology

Scientific Teaching Context

Learning Goals

- Recognize the interdisciplinary nature of science by exploring how computational and math-based tools can be used to answer biological questions.
- Appreciate the immense amounts of data generated and analyzed through evolutionary and genetic studies to synthesize answers to on-going questions.
- Increase feelings of self-efficacy in relation to computers/computational resources.

Learning Objectives

- Describe how bioinformatics tools can be used to answer evolutionary questions
- Hypothesize evolutionary relationships based on morphology
- Use DNA sequences from an organism to identify its closest relatives
- Align and analyze DNA sequences from different organisms to determine how they are related to one another
- Compare and contrast different ways that you can infer the evolutionary history of an organism
- Follow precisely a complex multistep procedure when carrying out experiments, taking measurements, or performing technical tasks; analyze the specific results based on explanations in the text.
- Explain how scientists can answer other questions using these same techniques

Introduction

A lack of formal bioinformatics training

As biology increasingly relies on big-data sets, the use of bioinformatics is becoming ubiquitous across disciplines. In the case of phylogenetics, the accumulation of genomic sequence data has revolutionized the field, turning it from one that relies on morphological data to one that relies on genetic and genomic sequence data, and consequently computational tools, to reconstruct evolutionary histories. The enormous amounts of genomic data available make it possible to answer more complex evolutionary questions, but also necessitate bioinformatic tools. Despite the prevalence of bioinformatics in science today, the importance of computational tools as they pertain to biology is not often discussed at the high school or introductory biology level (Machluf and Yarden, 2013). In 2008, a letter to the prestigious journal *Nature* directly addressed the lack of formal bioinformatics curricula, concluding that the failure to include bioinformatics training in higher education is contributing to a growing divide between the skills taught in science classrooms and those necessary to succeed in the workforce (Donovan, 2008).

Bringing bioinformatics into the classroom

Recent work has emphasized the importance of introducing the field of bioinformatics into high school and undergraduate classrooms, and integrating this important field into the curriculum (Donovan, 2008; Machluf and Yarden, 2013; Machluf et al., 2017). Scientific journals, such PLoS Computational Biology, have begun to recognize this effort by including a subsection of their Education section aimed at teaching bioinformatics and computational biology in secondary school classrooms (Lewitter and Bourne 2011). Historically, bioinformatics has been largely self-taught (Machluf and Yarden, 2013), however current educational research calls for a more focused, academic approach (Donovan, 2008).

A study by Machluf and Yarden (2013) found that bioinformatics was most effectively taught through computer-based scientific inquiry. By exposing students to current research tools and allowing them to use these tools to answer current scientific questions, students were able to parallel research trends and develop a working understanding of computational biology. Current work further supports the application of this inquiry-based process, citing this “big question” based approach as a critical means of linking bioinformatics to the broader science curricula (Machluf et al., 2017).

Using Pokémon to teach bioinformatics

Here, we build on previous efforts to produce lesson plans appropriate for the high school and undergraduate level (e.g. Gallagher et al. 2011, Machluf and Yarden 2013), by providing a lesson plan that focuses on the role of bioinformatics in evolutionary biology and phylogenetic analysis. This lesson incorporates many of the guidelines set forth by Form and Lewitter (2011) in their article “Ten Simple Rules for Teaching Bioinformatics at the High School Level.” Form and Lewitter (2011) contend that students learn best when they are actively working towards a goal. We have structured this lesson so that students are given fictitious molecular data from Pokémon which they must use to achieve their goal of determining the relationship of the Pokémon to each other and their context in the tree of life. Our lesson has students use morphological and molecular information from Pokémon to build phylogenies, first starting with morphological characteristics before progressing to DNA sequence data. The use of Pokémon in this activity aligns with the recommendation for students to explore examples that are familiar to them. Additionally, progressing from morphological to molecular data is consistent with the suggestion to use activities that build on one another. It is also proposed that students should work out computational processes (such as sequence alignment) using pen and paper before using the computational tool itself. As such, students in this lesson manually align a small set of sequences prior to using web-based tools for sequence alignment. Finally, after phylogeny reconstruction, students investigate the species in their tree and place their Pokémon in the tree of life, which provides students with a product (phylogeny) that can be shared with the class.

The lesson plan outlined below fits into the pre-existing curriculum requirements for both high school and undergraduate science courses. In particular, it addresses the following standards:

Common Core State Standards

- Follow precisely a complex multistep procedure when carrying out experiments, taking measurements, or performing technical tasks; analyze the specific results based on explanations in the text. (CCSS.ELA-LITERACY.RST.11-12.3)
- Determine the meaning of symbols, key terms, and other domain-specific words and phrases as they are used in a specific scientific or technical context (CCSS.ELA-LITERACY.RST.11-12.4)

Next Generation Sunshine State Standards (Florida)

- Describe how and why organisms are hierarchically classified and based on evolutionary relationships (SC.912.L.15.4)
- Represent and understand natural phenomena using modeling and simulation (SC.912.CS-CS.1.5)
- Analyze data and identify real-world patterns through modeling and simulation (SC.912.CS-CS.1.1)
- Explain how data analysis is used to enhance the understanding of complex natural and human systems (SC.912.CS-CS.1.3)
- Define a problem based on a specific body of knowledge and use tools to gather, analyze, and interpret data (SC.912.N.1)
- Compare techniques for analyzing massive data collections (SC.912.CS-CS.1.4)
- Analyze and manipulate data collected by a variety of data collection techniques to support a hypothesis (SC.912.CS-CP.1.3)
- Perform advanced searches to locate information and/or design a data-collection approach to gather original data (e.g., qualitative interviews, surveys, prototypes, and simulations) (SC.912.CS-CP.1.2)

Vision & Change AAAS

- *Core Concepts for Biological Literacy* 1. Evolution: The diversity of life evolved over time by processes of mutation, selection, and genetic change.
- *Core Competencies and Disciplinary Practice* 2. Ability to use quantitative reasoning. Biology relies on applications of quantitative analysis and mathematical reasoning.
- *Core Competencies and Disciplinary Practice* 3. Ability to use modeling and simulation: Biology focuses on the study of complex systems.
- *Core Competencies and Disciplinary Practice* 4. Ability to tap into the interdisciplinary nature of science: Biology is an interdisciplinary science.

In summary, the lesson presented addresses the growing need to provide formal bioinformatics training to students, while also explicitly meeting multiple Common Core State Standards (2), Next-Generation Sunshine State Standards (8), and American Association for the Advancement of Science (AAAS) Standards (4).

Intended Audience

This lesson is suitable for a range of students, from junior/senior high school students to undergraduates (major, non-major, and non-traditional). The only prerequisite is an introductory understanding of DNA as the mode of inheritance and minimum exposure to the Tree of Life concept.

Required Learning Time

This lesson requires 45-60 min for completion (see Recommended Timeline, Table 1)

Pre-requisite Student Knowledge

The only prerequisite is an introductory understanding of DNA as the mode of inheritance and minimum exposure to the Tree of Life concept.

Pre-requisite Teacher Knowledge

Some experience with computers will be helpful. A basic understanding of genetics and evolution is required (the authors suggest a brief guide, provided in English and Spanish by UC Berkeley: http://evolution.berkeley.edu/evolibrary/article/0_0_0/evo_01). Teachers should also be fairly comfortable with troubleshooting potential computer-based issues that students may have. All other requisite knowledge is contained in the lesson. It is strongly suggested that the instructor review the instructions given in the Teacher Solutions sheets and perform all the computer-based activities prior to performing this class activity. Detailed step-by-step instructions are provided in the Teacher's packet and should be followed during the lesson.

Scientific Teaching Themes

The lesson plan presented here relates to the scientific teaching themes of active learning, assessment, and inclusive teaching. The specific learning objectives and larger learning goals are approached and explored through five hands-on activities, including two computational activities. These activities closely parallel the process by which scientists recapitulate evolutionary relationships. Successful learning will be assessed formally, through the completion of Activity sheets and the generation of a Pokémon phylogeny, and informally through leading questions and discussion at multiple levels (one-on-one with instructor, small group, and class-level). The activities presented here aim to make the learning objectives accessible to a variety of learning styles by including a variety of approaches. This increases the likelihood of successful completion of the learning goals.

Active Learning

Students will actively engage in learning the concepts through five activities that mirror real phylogenetic studies. All five activities can be completed individually or in small groups.

1. *Morphology Matters*: Includes individual problem-solving, small group discussion of Pokémon morphology, and large group tree-building. This is a variation of think-pair-share.
2. *Bioinformatics Is a BLAST*: Includes small group/individual computer work with the goal of identifying the closest relative(s) of a given Pokémon. The results of the analysis enable students to identify the logical basis of their hypotheses.
3. *Aligning Sequences - Hands on or Hands Off?*: Requires the use of sequential thinking to align DNA sequences by hand. This activity promotes analogical thinking to connect/compare/contrast morphological data with molecular data.
4. *Computing Relationships*: Students, working individually or in small groups, use computer programs to align and analyze DNA sequences to determine how their Pokémon is related to other real-life organisms. This promotes computer literacy and self-efficacy through the use of the web-based MAFFT software.
5. *Tree of Life*: In this activity, students place their Pokémon in the Tree of Life, helping them visually connect this new information with their prior knowledge of biological diversity. This final activity requires the students to draw conclusions about the evolution of their Pokémon and relate this back to the hypotheses formed in the second activity.

Assessment

The lesson will begin with a pre-test focused on the students' understanding of bioinformatics and evolution. The lesson concludes with a post-test of the same questions.

1. *Morphology Matters*: Understanding of the concepts presented in this activity will be formally assessed through the successful completion of a morphological character matrix. Informal assessment will include group/class discussion.
   Specific Learning Objectives:
   - Hypothesize evolutionary relationships based on morphology
   - Compare and contrast different ways that you can infer the evolutionary history of an organism
2. *Bioinformatics Is a BLAST*: Successful internalization of this activity's learning objectives will be formally assessed through correct identification of the Pokémon's closest relative in the BLAST database. Informally, learning will be measured through one-on-one discussions with students during the activity.
   Specific Learning Objectives:
   - Describe how bioinformatics tools can be used to answer evolutionary questions
   - Use DNA sequences from an organism to identify its closest relatives
   - Compare and contrast different ways that you can infer the evolutionary history of an organism
   - Increase feelings of self-efficacy in relation to using computers and computational resources for problem-solving
   - Follow precisely a complex multistep procedure when carrying out experiments, taking measurements, or performing technical tasks; analyze the specific results based on explanations in the text.
3. *Aligning Sequences - Hands on or Hands Off?*: Understanding of the concepts presented in this activity will be formally assessed through the successful completion of a DNA sequence alignment. Informal assessment will include one-on-one discussions with students during the activity.
   Specific Learning Objectives:
   - Align and analyze DNA sequences from different organisms to determine how they are related to one another
4. *Computing Relationships*: Learning will be formally measured by the completeness and correctness of the resulting phylogeny generated through the MAFFT website. Informally, learning will be gauged through prompting questions (What does this tree tell you? What other questions could you answer using this approach?)
   Specific Learning Objectives:
   - Describe how bioinformatics tools can be used to answer evolutionary questions
   - Align and analyze DNA sequences from different organisms to determine how they are related to one another
   - Compare and contrast different ways that you can infer the evolutionary history of an organism
   - Increase feelings of self-efficacy in relation to using computers and computational resources for problem-solving
   - Follow precisely a complex multistep procedure when carrying out experiments, taking measurements, or performing technical tasks; analyze the specific results based on explanations in the text.

5. *Tree of Life*: Formal assessment of learning will consist of correct placement of all Pokémon in the Tree of Life. Informal assessments will be completed through class discussion of hypotheses, hypothesis testing, and comparing/contrasting data types (morphological and molecular), and expanding these questions to other real-life scenarios.
   Specific Learning Objectives:
   - Describe how bioinformatics tools can be used to answer evolutionary questions
   - Compare and contrast different ways that you can infer the evolutionary history of an organism
   - Explain how scientists can answer other questions using these same techniques

Inclusive Teaching

This lesson accommodates different learning types by incorporating elements of reading and writing as well as aural, visual, and kinesthetic activities. By showcasing the role of computing in biological science, we help students understand the need for scientists with diverse skillsets that they may not traditionally associate with biology.

Lesson Plan

Table 1. Recommended Timeline

Activity	Description	Estimated Time	Notes
Pre-class Preparation			
Preparation for class	1. Make one copy of Activity sheets for each student 2. Make one copy of Screen Shots for each group 3. Print out copy of Teacher Solutions 4. Load Pokémon_sequence file on to computers that will be used by students 5. Go through exercises step-by-step to familiarize yourself with each step 6. Print Pre- and Post-Assessment	1.5 hours	· Activity sheets for each part (I-V; Supporting Files S2-6) should be printed so that they can be passed out independently of one another · Screen Shots with step-by-step instructions are found in Supporting File S7 · Teacher Solutions are found in Supporting File S8 · Make sure the Pokémon_sequence (Supporting File S9) file is saved in an easily accessible location (e.g. the Desktop)
3-4 Days Before the Workshop: Pre-Assessment			
Pre-Assessment	1. Pass out pre-assessment to each student and have them fill it out	15 min	· Pre-assessment is in Supporting File S10
Part I: Morphology Matters			
Morphological table and hypothesized relationships between species	1. Introduce the topic, pass out overview of today's lesson 2. Pass out Activity sheet for Part I: Morphology Matters, explain premise of activity 3. Have each student fill in morphological table	7-10 min	· Overview of today's activity is in Supporting File S1 · Activity sheet for Part I is in Supporting File S2 · Teacher solutions are found in Supporting File S8

	4. Together, draw the tree on the board according to the Teacher Solutions 5. *Transition:* Discuss ambiguity and conflicts in the ways that the tree can be drawn		
Part II: Bioinformatics is a BLAST			
Students use the BLAST database to find DNA sequences of relatives of their organism	1. Form groups of 2-4 students 2. Pass out instructions for Part II: Bioinformatics is a BLAST to each student, explain premise of activity 3. Pass out one Screen Shots Activity sheet to each group of students 4. Check in with groups & students frequently to help troubleshoot 5. *Transition*: How do we compare DNA sequences?	10min - 12min	· Instructions for Part II is in Supporting File S3 · Screen_Shots instructions is in Supporting File S7
Part III: Aligning Sequences – hands on or hands off?			
Sequence alignment activity	1. Pass out Part III Activity sheet, explain premise of activity 2. Have students solve alignment on their own 3. *Transition:* Discuss – would this be possible to do by hand with longer sequences?	7-10 min	· Activity sheet for Part III: Aligning Sequences is available in Supplementary File S4 · Possible solutions are found in the Teacher Solutions guide (Supplementary File S8)
Part IV: Computing Relationships			
Computing relationships & building an evolutionary tree	1. Pass out Part IV Activity sheet, explain premise of activity 2. Have students follow the steps in groups 3. Check in with groups & students frequently to help troubleshoot	10-12 min	· Activity sheet for Part IV: Computing Relationships is available in Supplementary File S5

	4. As students finish, pass out the Activity sheet for Part V		
Part V: Tree of Life			
Placing Pokémon on the Tree of Life and comparing our results to our morphological hypothesis	1. As students begin to fill in the Activity sheet for Part V, have them come to the board to show where their Pokémon falls on the larger Tree of Life 2. Compare the result to the previous hypothesis based on morphology 3. Discuss that we can use this same process for real organisms	10-12 min	· Activity sheet for Part V: Tree of Life is available in Supplementary File S6
3-4 Days After the Workshop: Post-Assessment			
Post-Assessment	1. Pass out post-assessment to each student and have them fill it out	15 min	· Post-assessment is in Supporting File S10

Pre-class Preparation

Required equipment

Note that this lesson requires computers with internet access, preferably enough for one per 1-4 students.

Familiarization with lesson

Before beginning class, go through the exercises for Parts I-V to ensure you understand the steps.

Preparation of materials

Based on your class size, determine how many groups you will have for the group activities (we recommend 1-4 students per group). Print the appropriate resource materials: Overview of today's activity (S1), one per student; Parts I-V (S2-6), one per student; Screen Shots (S7), one per group; Teacher Solutions (S8), one per teacher. To save time, it is helpful to load the Pokémon_sequence file (S9) on to the computers that will be used by students. If this is not possible, have the sequence file available on a USB drive for students to copy it to their computer's desktop at the beginning of Part II. You will also need one red and one green post-it note for each group.

Pre-assessment

Have students complete the pre-assessment activity sheet prior to the beginning of the class period, preferably at least one day in advance.

Part I: Morphology Matters

Key Point

We can hypothesize evolutionary relationships based on morphological traits.

Introduction to activity

Pass out the overview of today's lesson (S1), which includes a picture of the tree of life, and prompt students to think about how we have determined the evolutionary history of organisms.

Morphology Matters Activity

Pass out the Activity sheet for Part I: Morphology Matters (S2) to each student, giving a brief explanation (written instructions are included on the Activity sheet). Have each student fill in their morphological table. As students finish, ask them to pair up with 1-2 students near them to discuss their thoughts and answer the questions on the Activity sheet. After a few minutes, bring the class together as a whole to talk about their answers and draw a tree based on the morphology (see Teacher Solutions, S7) - emphasize that this is our current hypothesis.

Potential Stumbling Points

There will almost certainly be questions about why the tree was drawn this particular way, and whether certain morphological traits should be more important than others (i.e. why did we start with leaves/flowers and not mouth?) - use these questions to address the problems with using morphology and as a way to introduce DNA sequences as an alternative.

Part II: Bioinformatics is a BLAST

Key Point

The BLAST database is an enormous bioinformatic resource compiled by scientists from around the world that can be used to answer lots of different types of questions.

Introduction to activity

Give a very brief explanation about the activity, explaining that we now have DNA sequences from Pokémon, and we want to figure out what other species each one is related to. Introduce the BLAST database as an enormous resource of biological information that we will use to answer this question.

Bioinformatics is a BLAST activity

Form groups of 2-4 students. Pass out one red and one green post-it note to each group. Explain to students that, while they are working without trouble, they should have the green "flag" on their computer, and if they have problems or are stuck, they should put up the red "flag". Pass out one copy of the Part II Activity sheet (S3) to each student, and one copy of the Screen Shots (S7) to each group. Have groups begin the activity. Regardless of whether there are red flags up, be sure to check in with groups frequently to make sure they are progressing and are not frustrated.

Potential Stumbling Points

Despite having instructions in front of them, students will likely keep asking you "What do I do now?" - remind them that they have the instructions and screen shots to follow, and red flags should be used when they are lost or have encountered a problem.

Part III: Aligning Sequences - hands on or hands off?

Key Point

Make the connection to prior knowledge: differences in the sequences including "gaps" arise from mutations in DNA replication process.

Introduction to activity

Explain that, now that we have the DNA characters for their Pokémon's relatives, we need to figure out how to compare them, just as we did with the morphological characters. This activity will give them an understanding of how that happens.

Aligning sequences activity

Pass out one copy of the Part III Activity sheet (S4) to each student. Have students work on the Activity sheet individually. As they finish, ask students to compare their answers to the other students in their group and answer the other questions on the Activity sheet. As a whole class, talk about the answers to the Activity sheet questions: Where do the gaps come from? Why is alignment important? Could you do this by hand for a very long DNA sequence?

Potential Stumbling Points

Students may be concerned about finding the correct answer. Be sure to explain to them that there may not be one "correct" answer and that is okay - it is something real scientists still struggle with. The table in the Teacher Solutions (S8) shows two possible solutions that students may come up with.

Part IV: Computing Relationships

Key Point
Scientists often use computational tools (computer programs) to help them analyze large datasets, in this case, to help align and compare DNA sequences.

Introduction to activity
Explain that scientists have developed computational tools, like MAFFT, to help them with tasks like DNA alignment.

Computing relationships activity
Have students return to their groups. Remind groups to use their red flags to indicate when they are having trouble. Pass out Part IV Activity sheet (S5) to each student, remind them of the screen shots Activity sheet, and have groups begin the activity. Regardless of whether there are red flags up, be sure to check in with groups frequently to make sure they are progressing and are not frustrated.

Potential Stumbling Points
Students may take different amounts of time to complete computing portions. As students finish, they can begin learning more about the organisms that are closely related to their Pokémon. You can also pass out the Part V Activity sheet (S6) at this point, particularly if time is running short.

Part V: Tree of Life

Key Point
Bioinformatic and computational tools allow scientists to analyze very large datasets and refine their hypotheses.

Introduction to activity
Explain to students that we can now determine the relationships between Pokémon based on where they fit in the larger tree of life.

Tree of Life activity
Draw the simplified tree of life provided in the Teacher Solutions (S8) on the board. Make sure to leave the names that are in parentheses blank; these spaces will later be filled in by the students. As students complete Part IV, pass out the Part V Activity sheet (S6). Guide students as they place their Pokémon on the larger tree of life on the board. Discuss with students how the relationships shown here compare to the hypothesized relationships based on morphology. Emphasize that we can use the same process to answer questions about real organisms.

Potential Stumbling Points

Because reading phylogenetic trees is not always intuitive, students may have difficulty comparing the relationships depicted by the trees. Guide students in the comparison and explain that the order of names is not as important as the branching pattern itself. Overall, emphasize the value of bioinformatic and computational resources in helping scientists answer questions about evolutionary history and relatedness.

Post-Assessment

Have students complete the post-assessment activity sheet preferably at least one day after the workshop.

Supporting Materials

S1. Overview of Today's Lesson. Overview and vocabulary for the day's lesson. (S1-6_StudentActivitySheets)

S2. Part I Activity sheet. Exercise to develop hypothesis of evolutionary relationships based on morphology. (S1-6_StudentActivitySheets)

S3. Part II Activity sheet. Instructions for navigating BLAST database. (S1-6_StudentActivitySheets)

S4. Part III Activity sheet. Exercise to have students gain hands-on experience with DNA alignment. (S1-6_StudentActivitySheets)

S5. Part IV Activity sheet. Instructions for MAFFT DNA sequence and tree-building exercise. (S1-6_StudentActivitySheets)

S6. Part V Activity sheet. Wrap-up exercise to identify where students' organisms fall on the broader tree of life. (S1-6_StudentActivitySheets)

S7. Screen Shots. Step-by-step screen shots for Parts II and IV. (S7_ScreenShots)

S8. Teacher Solutions. Examples of solutions to each exercise, key points to emphasize for each component. (S8_TeacherSolutions)

S9. Pokémon Sequence File. Fictitious DNA sequences for Pokémon. (S9_PokémonSequence)

S10. Pre & Post Assessment. Short pre- and post-assessment questions to see if this activity was effective at changing students' perspectives on bioinformatics, computing, and science. (S10_PrePostAssess)

Acknowledgments

The authors would like to acknowledge and thank Ms. Penny Lane, Ms. L. Castilla, Mr. J. Nunez, Mr. J. Faust Lopez, Dr. A. Schulze, Dr. H. Bracken-Grissom, Mr. J. Warschefsky, Ms. B. Warschefsky, Ms. R. Timm, Mr. J. Ahrens, Mr. C. Gonzalez, Dr. J. Siltberg-Liberles, and an extensive list of other teachers and influential individuals. This is your fault.

References

Donovan S. (2008). Big data: teaching must evolve to keep up with advances. *Nature*, *455*(7212): 461-461.

Form D & Lewitter F. (2011). Ten simple rules for teaching bioinformatics at the high school level. *PLoS Computational Biology*, *2011(*7(10)): e1002243. doi:10.1371/journal.pcbi.1002243.

Gallagher SR, Coon W, Donley K, Scott A & Goldberg DS. (2011). A first attempt to bring computational biology into advanced high school biology classrooms. *PLoS Computational Biology, 2011*(7(10)): e1002244. doi:10.1371/journal.pcbi.1002244.

Lewitter F & Bourne PE. (2011). Teaching bioinformatics at the secondary school level. *PLoS Computational Biology, 2011*(7(10)): e1002242. doi:10.1371/journal.pcbi.1002242.

Machluf Y, Gelbart H, Ben-Dor S & Yarden A. (2017). Making authentic science accessible—the benefits and challenges of integrating bioinformatics into a high-school science curriculum. *Briefings in Bioinformatics, 18.1*(2017): 145–159.

Machluf Y & Yarden A. (2013). Integrating bioinformatics into senior high school: design principles and implications. *Briefings in Bioinformatics, 14*(5): 648-660.

Table and Figure Legends

Tables

Table 1. Recommended Timeline

Figures

Figure 1. Tree of Life

Supporting Materials

S1. Overview of Today's Lesson. Overview and vocabulary for the day's lesson. (S1-6_StudentActivitySheets)
S2. Part I Activity sheet. Exercise to develop hypothesis of evolutionary relationships based on morphology. (S1-6_StudentActivitySheets)
S3. Part II Activity sheet. Instructions for navigating BLAST database. (S1-6_StudentActivitySheets)
S4. Part III Activity sheet. Exercise to have students gain hands-on experience with DNA alignment. (S1-6_StudentActivitySheets)
S5. Part IV Activity sheet. Instructions for MAFFT DNA sequence and tree-building exercise. (S1-6_StudentActivitySheets)
S6. Part V Activity sheet. Wrap-up exercise to identify where students' organisms fall on the broader tree of life. (S1-6_StudentActivitySheets)
S7. Screen Shots. Step-by-step screen shots for Parts II and IV. (S7_ScreenShots)
S8. Teacher Solutions. Examples of solutions to each exercise, key points to emphasize for each component. (S8_TeacherSolutions)
S9. Pokémon Sequence File. Fictitious DNA sequences for Pokémon. (S9_PokémonSequence)
S10. Pre & Post Assessment. Short pre- and post-assessment questions to see if this activity was effective at changing students' perspectives on bioinformatics, computing, and science. (S10_PrePostAssess)