

Using Evolutionary Data in Developing Phylogenetic Trees: A Scaffolded Approach with Authentic Data

● KD DAVENPORT, KIRSTIN JANE MILKS,
REBECCA VAN TASSELL



ABSTRACT

Analyzing evolutionary relationships requires that students have a thorough understanding of evidence and of how scientists use evidence to develop these relationships. In this lesson sequence, students work in groups to process many different lines of evidence of evolutionary relationships between ungulates, then construct a scientific argument for a particular set of relationships as modeled in a cladogram. Visual and verbal scaffolds are used throughout the lessons to address common misconceptions and points of difficulty for students.

Key Words: Tree thinking; scientific argumentation; phylogeny; evolution; science practices; data analysis.

○ Introduction

Evolution continues to be the light that illuminates the study of biology from a “pile of sundry facts” to a richly “satisfying and inspiring science” (Dobzhansky, 1973). Biology teachers often try to address students’ misconceptions about evolution by using fictional or simplified data sets that allow students to quickly reach the desired conclusions. However, imaginary organisms and pruned data sets provide limited opportunities to meaningfully grapple with information that is often incomplete, confusing, or even misleading – common challenges faced by real evolutionary biologists trying to generate cladograms. In this activity, we provide students an opportunity to explore authentic data on evolution, challenging them to grapple with complexity and draw conclusions in ways that parallel the work done by scientists.

Analyzing evolutionary relationships, including those depicted in phylogenetic trees, is a critical skill for high school biology. To do this analysis meaningfully, students need to structure their knowledge of evolutionary processes and relationships into scientific explanations. The practice of scientific argument – a type of explanation that uses scientific

evidence to create and justify a conclusion – improves learning and engagement in science (Duschl et al., 2007) and supports students’ understanding of authentic scientific practice (National Research Council, 2012).

It may seem daunting to give your students purposely confusing data in a lesson on evolution – isn’t teaching evolution hard enough? Will sifting through ambiguous data give credence to the idea that evolution is “just a theory”? To the contrary, we have found in designing, teaching, and refining this lesson that students appreciate the challenge of drawing the best conclusion they can from imperfect data and, through doing so, come to understand how evolutionary biologists study a process (like mammalian evolution) that cannot be directly observed. In fact, we find that allowing students to wrestle with real, rather than simulated, data will build their confidence that evolutionary biology gives us useful tools for understanding life on Earth.

This activity is the result of 4 years of collaboration as part of our teaching fellowship with the Knowles Science Teaching Foundation. Modeling our work on the Japanese practice of lesson study (Stigler & Hiebert, 2009), we collaboratively planned and refined instruction based on national standards, research into pedagogy, design protocols, and video and work samples over 4 years. We have each successfully implemented this lesson at levels ranging from 9th-grade introductory classes to advanced 12th-grade AP and IB biology classes. The instructions that follow give suggestions for modifying the lesson for courses from middle-grades life science to early college.

Analyzing evolutionary relationships, including those depicted in phylogenetic trees, is a critical skill for high school biology.

○ Objective

Students will be able to describe major categories of evidence for evolutionary relationships and determine whether evidence supports or refutes a hypothesized relationship in order to compose and evaluate a scientific explanation about evolutionary relationships among living things.

How can we represent evolutionary relationships by arranging these animals on this phylogenetic tree?

Students analyze 7 data sets in small groups by ...

- summarizing data in writing **(A)**
- representing how data support/contradict possible trees by line-drawing **(B)**

... and then use their analyses to draw and justify a final tree

Day 1: Data analysis using line drawing.
Data sets: skulls, fur/skin, locomotion

Day 2: Applying technique to “modern” data
Data sets: range, foot/digestive morphology, DNA inserts, protein sequence

Day 3: Constructing “best” tree and writing scientific explanations

Data 1:

Description of data (WHAT I SEE):

Humpback has no visible teeth.

Camel and elk's teeth fit together and eye is far back above rounded jaw.

Hippo and camel have tusks.

(A)

Tree 1

Tree 2

————— These data support the tree.

————— These data **STRONGLY** support the tree.

—X— These data **CONTRADICT** (go against) the tree.

- - - - These data neither support nor contradict the tree.

Figure 1. Overview of lesson series and sample student work. This lesson was designed for three or more traditional class periods but can be adapted for different schedules. Sample work (summary and line-drawing) has been reproduced in gray to demonstrate the scaffolding that students use as they analyze data sets.

○ Details of Activity

In this lesson sequence (Figure 1), students begin by evaluating two different (incorrect) phylogenetic hypotheses using multiple sources of data. Students then generate their own hypothesis to explain the

data and compose a scientific argument to defend it. The lesson engages students in an iterative experience in which they receive additional data that provide evidence that may be contradictory or supportive of previously identified relationships. Students must

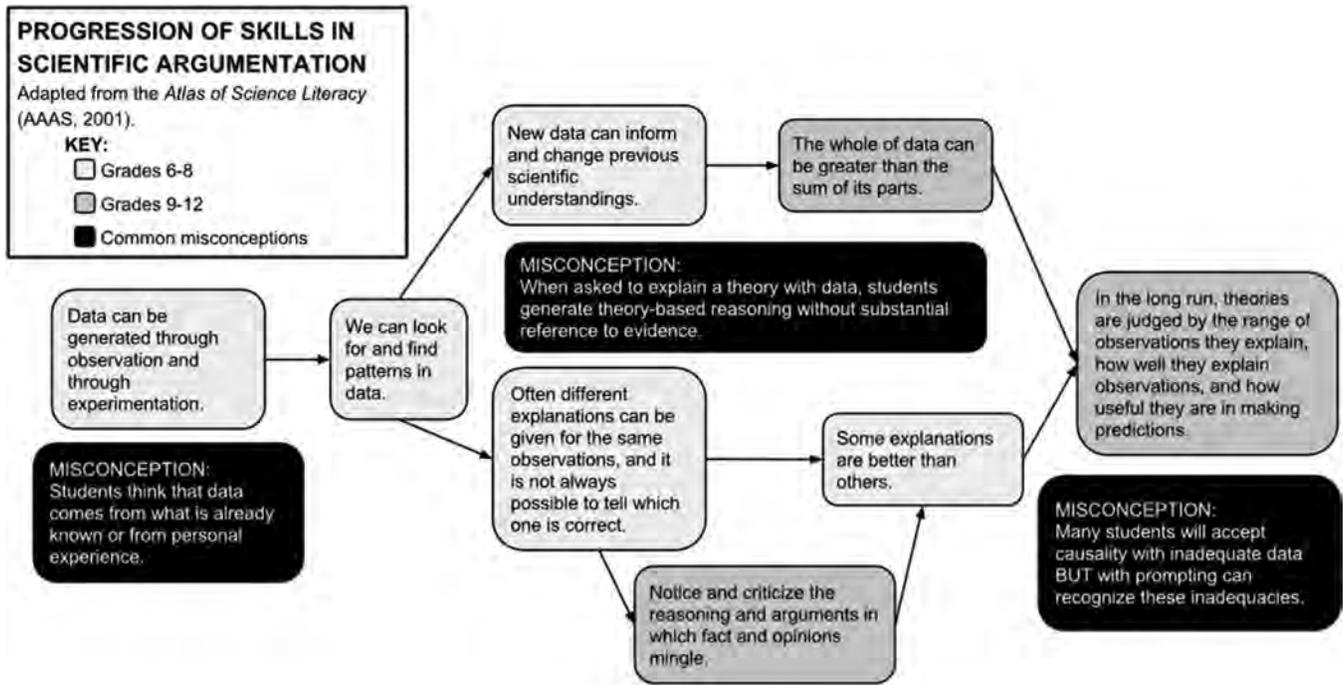


Figure 2. Progression of skills in scientific argumentation. This progression, along with included common misconceptions, inspired the scaffolding developed in this lesson series.

rationalize these conflicting data sets to develop a phylogenetic tree that represents the most likely evolutionary relationships.

Prerequisite Knowledge

Before engaging in this lesson sequence, students should have a working knowledge of how phylogenetic trees are constructed to represent relationships among organisms and their common ancestors. We use the activity featured in “Investigating Tree Thinking & Ancestry with Cladograms” (in the previous issue of *ABT*) to lead students to uncover the idea of common ancestry and develop skills in reading phylogenetic trees. A more thorough review of the body of research related to teaching and learning about evolutionary relationships is included in that article.

Collaborative skills are also required for successful completion of this lesson series. Compared to how we had taught this content in the past, we found in teaching this lesson series in our classroom that students were enthusiastic about discussion of and collaborative reasoning about the data set and its implications. One AP student wrote that she wasn’t “sure [she] would have spent half the time thinking so carefully” about her final tree without her group’s discussion as they struggled together to articulate their reasoning. If you are introducing your students to this type of work, consider setting norms as a class or suggesting specific behaviors (explain your thinking, ask questions, etc.) that will foster student success. Acknowledge the discomfort that students may feel working with data that are sometimes unclear or contradictory, and encourage students to explore that uncertainty. To avoid anxiety about “correct” answers, your assessment of student work should focus on how students analyze data, draw conclusions, and revise their ideas, not on how quickly they deduce the “correct” phylogeny.

Finally, it might be useful for your students, particularly those new to biology, to have recent experiences in molecular genetics and

evolution and, therefore, a stronger scientific basis on which to prioritize data. More experienced students might only need a well-timed teacher question to prompt their thinking about these topics. See “Assessments” below for sample questions.

Day 1: Representing Data Analysis Using Line Drawing

Distribute the worksheet and describe the task to your students: Each student group will receive a series of data that will provide information on the evolutionary relationship of eight mammals. For each data set, students first observe and summarize trends or relationships. (This step is included because middle and high school students often generate reasoning without substantial reference to evidence; see Figure 2 for this and other design considerations.) Students must then decide whether the data support, strongly support, contradict, or neither support nor contradict two hypothetical trees. The purpose of this activity is for students to think about how to determine whether a particular datum supports a particular hypothesis, not to identify the correct solution; in fact, both hypotheses shown are incorrect. However, because the data are fragmentary, it is possible that one piece of data supports one *part* of an overall incorrect cladogram. Through this activity, students should come to understand that, while pieces of a complex model may be supportable, they must consider the totality of the evidence in their evaluation.

To visually represent these relationships and help students use data specifically and holistically, students draw specific lines (as illustrated in Figure 1) in a structure first introduced to us by Ravit Golan Duncan and colleagues (Chinn & Buckland, 2012). Consider modeling lines for an unrelated data set and competing claims with the whole class; you might also assign a homework assignment that features this scaffold in analyzing other data sets and arguments.

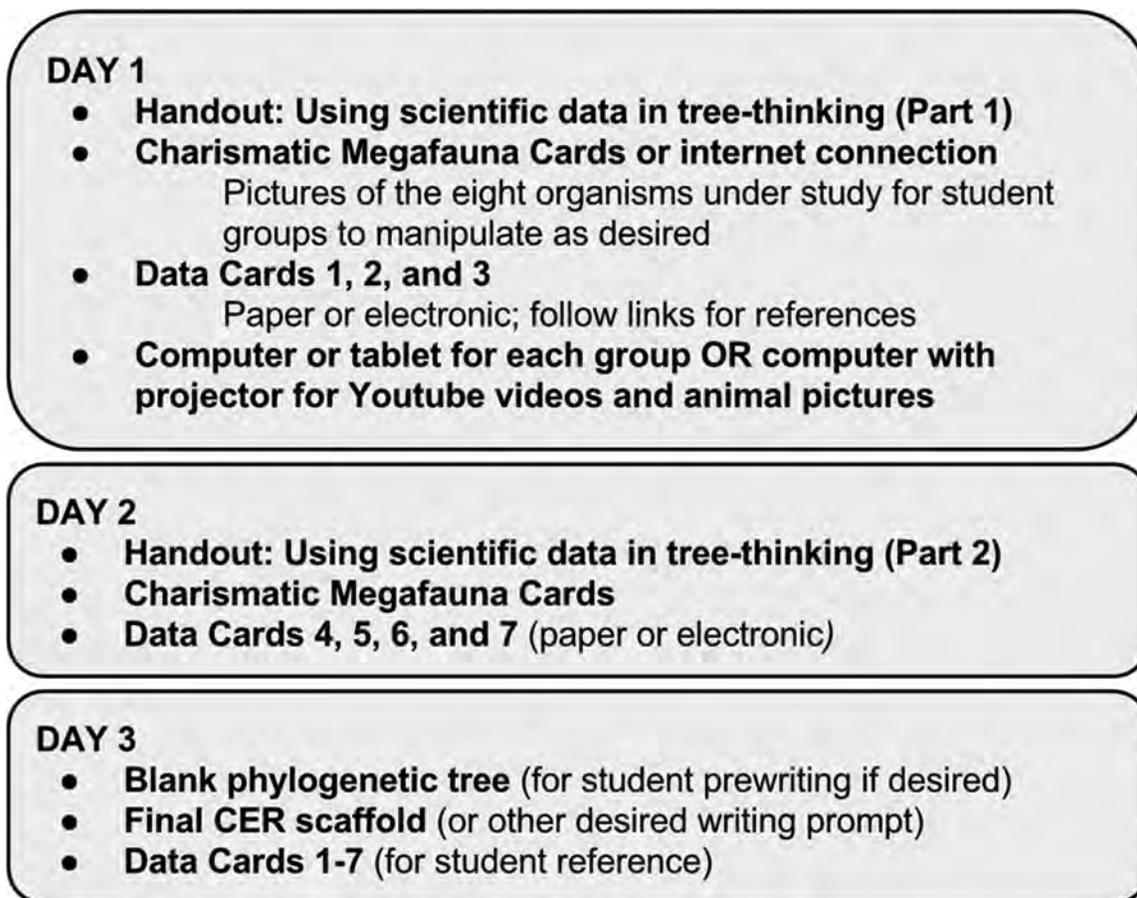


Figure 3. Materials required for this lesson series.

The first day's data (skull shape, fur/skin color, and gait; see Figure 3) are easily accessible to students. This lowers the cognitive challenge of sense-making so that students can familiarize themselves with the line-drawing task before working with the more challenging molecular data on Day 2. We have found that bright students in the class will quickly challenge the fur-color data as irrelevant (as, of course, they are); if students do not bring this up on their own, be sure to provoke the question yourself. This may prompt conversations about how some explanations in science are better than others – a conversation that grows richer during Day 2.

Before engaging with more data, student groups discuss and provide written answers to questions that provide scaffolding for students' reasoning:

1. According to the lines you drew, which tree is more well supported by the pieces of evidence?
2. Explain your choice.
3. Was there any evidence that made you confident that one of the trees was scientifically valid (*correct*)? If yes, please explain which data card and why. If no, please explain why not.

These questions scaffold arguments by asking students for specific data that support or do not support various trees and providing explanations of why.

As students work, they will need frequent reminders of good group behaviors and guidance in writing detailed summaries of the data. After students have completed Day 1's reflection questions, facilitate a whole-class discussion to model how data can be used to support or contradict the trees, with a focus on citing specific data and highlighting reasoning.

Day 2: Applying the Technique to “Modern” Data

To begin Day 2, students generate their own cladogram based on Day 1's work. The tree structure is provided, so groups need only arrange organisms on branch tips. In Day 2, students move through new data, considering their own tree (which can change after each data card) and two new hypothetical trees (neither of which represents the most parsimonious final answer). Students continue the now familiar summarization and line-drawing protocol to determine which tree features best represent relationships shown in the data they select as evidence. As students progress through Day 2, you continue to facilitate as groups attempt to reconcile and revise existing claims using new data.

Day 2's data cards are designed to challenge students to think about what data are most important. The data are intentionally contradictory and force students to make decisions about which support the most reasonable and likely explanations of evolutionary relationships. Data Card 5 shows foot and digestive-system morphologies that, if interpreted without considering convergent evolution, suggest

Data #1: Skulls

 <p>Camel skull http://www.dundee.ac.uk/museum/lib/img/zoology/skulls/DUNUC1929.jpg</p>	 <p>Elk skull http://www.valleyanatomical.com/catalog/images/R-S378-American-Elk-skull-(.jpg</p>	 <p>Hippo skull http://upload.wikimedia.org/wikipedia/commons/7/71/Hippo_skull_dark.jpg</p>	 <p>Humpback whale skull http://www.rescast.com/Images/specimens/Humpback_skull.jpg</p>
--	--	--	---

Data #2: Color -- Check out pictures online to see these animals' colors.

Data #3: Videos of movement of selected animals

Check out the videos listed below (playlist at

<http://www.youtube.com/playlist?list=PL1wclsxSNFg4JOGpHSpVR9CRkMGuyVcah>) to see these animals in motion!

- Chacoan Peccary: <http://www.youtube.com/watch?v=FDapL0WjBIM>
- Humpback Whale: http://www.youtube.com/watch?v=1_2O-87BxJ8
- Hippopotamus: <http://www.youtube.com/watch?v=fDJ3LNDdygY>
- Domesticated Pig: <http://www.youtube.com/watch?v=-LeOoohGrM>
- Mouse Deer: <http://www.youtube.com/watch?v=WTdNsC3Z30A>
- Dromedary Camel: <http://www.youtube.com/watch?v=Svt5BpwxAfk>
- Elk: <http://www.youtube.com/watch?v=Knev7wPWC2Q>
- Bottlenose Dolphin: <http://www.youtube.com/watch?v=mXtDghLPuZ0>

Data #4: Range

 <p><i>Cervus canadensis</i> (Elk) http://upload.wikimedia.org/wikipedia/commons/thumb/e/e8/Range_map_cervus_canadensis.svg/800px-Range_map_cervus_canadensis.svg.png</p>	 <p><i>Catagonus wagneri</i> (Peccaries) http://upload.wikimedia.org/wikipedia/commons/d/d0/Catagonus_wagneri_distribution_map.PNG</p>	 <p><i>Camelus dromedaries</i> (Camel) http://upload.wikimedia.org/wikipedia/commons/thumb/4/47/Dromedary_Range.png/800px-Dromedary_Range.png</p>	 <p><i>Moschiola indica</i> (Chevrotain) http://www.iucnredlist.org/apps/redlist/images/range/thumbs/136585.png</p>
---	---	--	---

Figure 4. Data cards for Days 1 and 2.

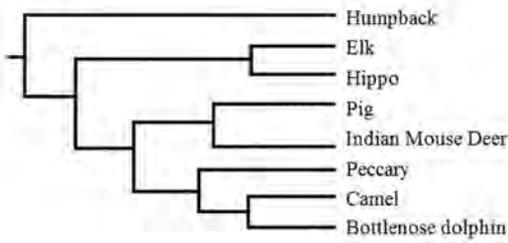
that the camel and pecoran groups are most closely related. In Data Card 6, however, students analyze DNA insert data derived from the work of Nikaido et al. (1999; see Figure 4) that show which groups of organisms share nine different noncoding DNA sequences. (Since no protein or RNA is encoded, the presence of these sequences in distinct groups can be explained on the basis of homology through common descent and not by analogy resulting from selection for any phenotype.) It is clear that the camel is an outgroup because it shares

none of the DNA inserts with any other group. Students must decide: Is it more likely that the camel somehow lost six separate DNA insert sequences that it should have in common with the pecorans, or that the camel and pecoran feet are analogous and not actually indicative of relatedness?

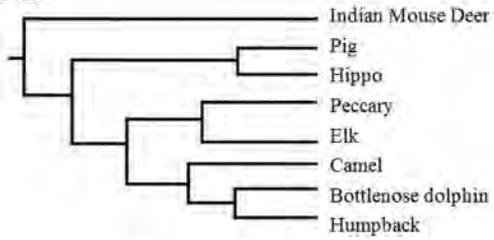
We have found that, at this point, students still prioritize data that are more familiar or easier to understand. For example, despite the DNA data clearly showing that hippos share all the DNA inserts

Trees for student use on Day 1:

Tree 1

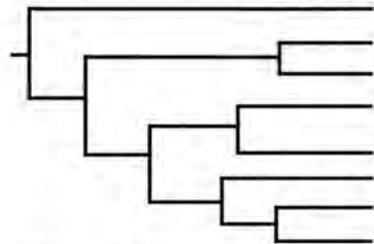


Tree 2

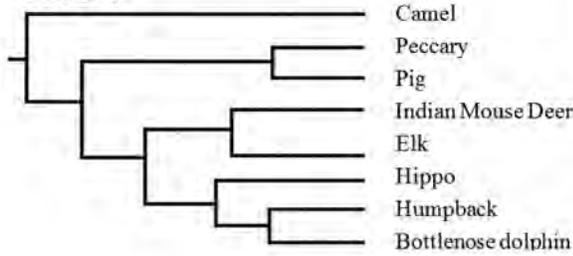


Trees for student use on Day 2 and sample worksheet layout:

Tree 3: Your current tree



Tree 4



Tree 5

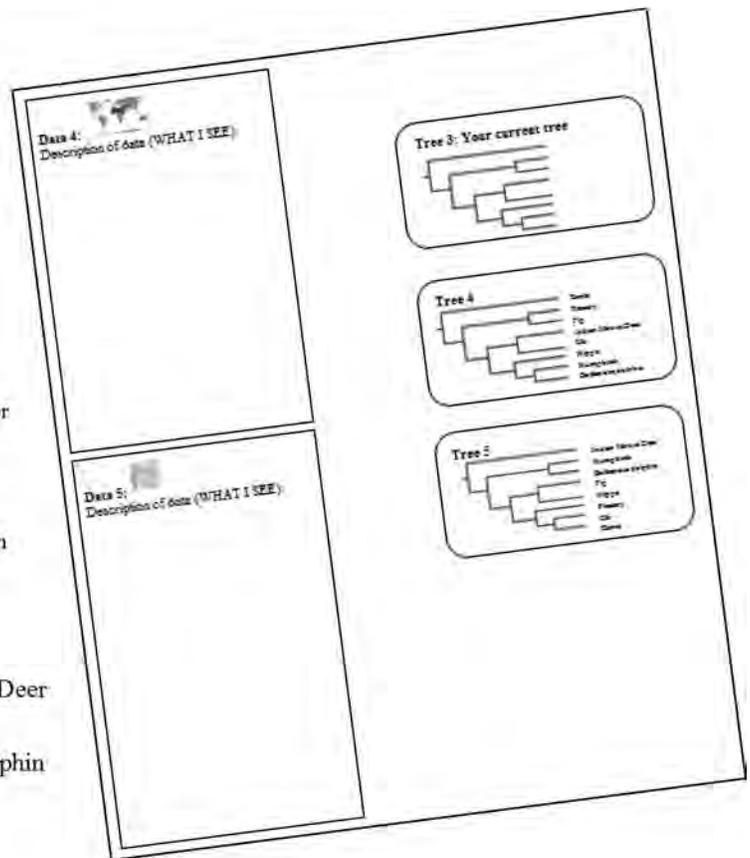
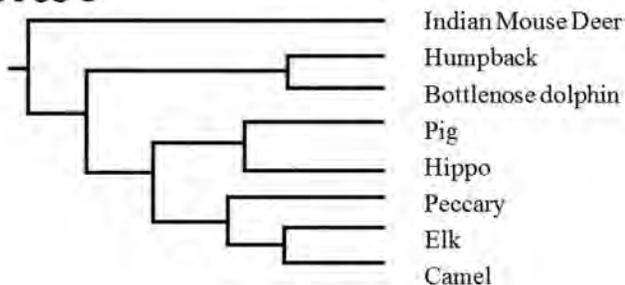


Figure 5. Trees used on Days 1 and 2.

evidence, draw lines, and construct responses on the handout means that teachers can look over groups' shoulders for a window into student thinking throughout the lessons. If desired, you can collect handouts at the end of a class to quickly review them. This will allow

you to determine what scaffolds or hints students might need at the beginning of your next class period.

Our preferred way of analyzing student thinking, however, is careful questioning as students are working in small groups. Probing

questions whose answers require students to verbalize their thinking process can provide useful information for group members and instructors. Teachers might ask:

- *Why did you decide to draw this type of line instead of that type of line?*
- *Is either tree better supported by the data? Tell me why.*

If your students are just starting out in biology, you might ask them to describe the process of determining how evidence supports a claim:

- *Describe what this data card tells you about the animals featured.*
- *Are these data important for determining evolutionary relationships? Why/why not?*
- *Does fur color really tell us about how closely animals are related?*

Conversely, more advanced students are ready to clearly articulate reasoning using appropriate scientific vocabulary, particularly for ambiguous evidence or the later data cards that indicate the tree positions of the closely related pairs:

- *What evolutionary phenomena or processes might explain why the camel's relatedness to other organisms looks so different in these two evidence cards?*
- *You've identified pairs or sets of animals that are very closely related. How can the DNA data be used as evidence to place those sets on the tree? Why can you make these claims?*

Follow-up questions or hints can allow scaffolding tailored specifically for individual groups' needs and can allow for extension if a group has worked particularly quickly.

Summative Assessment

As described above, we ask students to construct their final tree, then write justifications for their choices using the CER framework. We provide the data cards for student reference during the writing process. The following assignment, given after students have assigned the positions on a blank tree as letters A through H, shows one example of how we've prompted students to develop claims about particular nodes on the tree, rather than approaching the entire tree all at once.

PROMPT 1: You are claiming that "B and C are more closely related to each other than to any other animal on the tree."

1. Restate your claim, using the sentence structure above, replacing the names of the animals for B and C.
2. Describe the evidence that supports this claim – that is, tell me what the data *are*.
3. Explain your reasoning about why evidence supports the claim – in other words, convince me of your idea about what the data *mean*. Why does that evidence lead you to the conclusion that B and C are more closely related to each other than to any other animal?

We prompt students to construct similar justifications for placement of the other pairs/triads and the outlier; depending on your context, you may choose to have students construct an argument that focuses only on a subset of the organisms or provide further organization to help support the argumentation process. We encourage more advanced students to fully explain their reasoning about the evidence they chose as the basis for their tree and why they prioritized that

evidence. Advanced students also might consider a 2011 AP Biology exam prompt (College Board, 2011) that addresses relationships between these organisms.

Using a technique described extensively in the Five Practices approach (Cartier, 2013), we often preview student writing and, during the next class period, run a discussion highlighting student work samples. After this discussion, you may ask students to revise work and resubmit. For example, a student in one of our courses provided the following argument for positions of animals on the tree: "The hippo and the dolphin share many identical amino acids in hemoglobin C. The whale and dolphin live in the same environment, they swim in the same way and have similar color." After this argument was shared with the class, students identified it as partially correct. Students then revised this sample argument to either remove the color and locomotion data or use the evidence as a counterclaim, then shared these revisions with the class. After modeling this revision sequence with the whole class, using student work as examples, all students were given the opportunity to revise their arguments for credit. Through this discussion-and-revision format, students can see a variety of explanations, varying in their thoroughness and eloquence, and get indirect feedback on their work as they prepare their final justifications.

○ Conclusions

Considering complex evidence about real organisms allows students authentic experience with practicing the type of thinking – and scientific communication – that actual biologists must use to solve evolutionary puzzles. We developed and refined the sequencing and scaffolds described in this article to give a wide variety of students access to this important work, recommended by national biology standards documents from middle school to college, because we believe strongly that developing the practices of science is vital for learners of any level. By intertwining science content and skill development in our students' classroom experiences, we educators can help train the next generations of scientists and scientifically literate citizens to think critically in a variety of contexts.

Acknowledgments

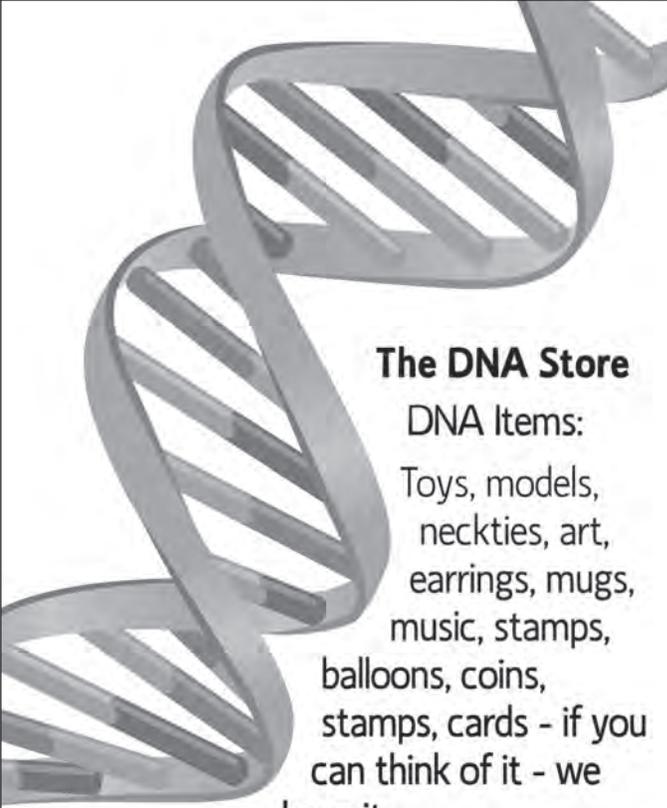
We thank the Knowles Science Teaching Foundation (KSTF) for introducing us to lesson study and supporting this work. We also thank the anonymous reviewers of the manuscript for extremely helpful suggestions. We are also grateful to early readers of the manuscript, particularly Michele Cheyne at KSTF, and our students for their thoughtful and helpful feedback.

References

- AAAS. (2001). *Atlas of Science Literacy, vols. 1 and 2*. Washington, DC: AAAS Project 2061 and NSTA Press.
- AAAS. (2011). *Vision and Change in Undergraduate Biology Education: A Call to Action*. Washington, DC: AAAS.
- Berland, L.K. & McNeill, K.L. (2010). A learning progression for scientific argumentation: understanding student work and designing supportive instructional contexts. *Science Education*, 94, 765–793.

- Cartier, J. (2013). *5 Practices for Orchestrating Task-Based Discussions in Science*. Arlington, VA: NSTA Press.
- Chinn, C.A. & Buckland, L.A. (2012) Model-based instruction: fostering change in evolutionary conceptions and in epistemic practices. In K.S. Rosengren, S.K. Brem, E.M. Evans, and G.M. Sinatra (Eds.), *Evolution Challenges: Integrating Research and Practice in Teaching and Learning about Evolution* (pp. 211–232). New York, NY: Oxford University Press.
- College Board. (2011). AP Biology 2011 scoring guidelines, form B. Available online at http://apcentral.collegeboard.com/apc/public/repository/ap11_biology_form_b_scoring_guidelines.pdf.
- College Board. (2013). AP Biology course and exam description, revised edition. Available online at http://media.collegeboard.com/digitalServices/pdf/ap/IN120084785_BiologyCED_Effective_Fall_2012_Revised_lkd.pdf.
- Dobzhansky, T. (1973). Nothing in biology makes sense except in the light of evolution. *American Biology Teacher*, 35, 125–129.
- Duschl, R.A., Schweingruber, H.A. & Shouse, A.W., Eds. (2007). *Taking Science to School: Learning and Teaching Science in Grades K–8*. Washington, DC: National Academies Press.
- Keeley, P.D. (2005). *Science Curriculum Topic Study: Bridging the Gap between Standards and Practice*. Thousand Oaks, CA: Corwin.
- International Baccalaureate Organization. (2007). Diploma programme biology: guide. Cardiff, Wales: IBO.
- Irwin, D.M., Kocher, T.D. & Wilson, A.C. (1991). Evolution of the cytochrome b gene of mammals. *Journal of Molecular Evolution*, 32, 128–144.
- McNeill, K.L. (2009). Teachers' use of curriculum to support students in writing scientific arguments to explain phenomena. *Science Education*, 93, 233–268.
- McNeill, K.L., Lizotte, D.J., Krajcik, J. & Marx, R.W. (2006). Supporting students' construction of scientific explanations by fading scaffolds in instructional materials. *Journal of the Learning Sciences*, 15, 153–191.
- National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010). *Common Core State Standards*. Washington, DC: Authors.
- National Research Council. (2012). *A Framework for K–12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*. Washington, DC: National Academies Press.
- Nikaido, M., Rooney, A.P. & Okada, N. (1999). Phylogenetic relationships among cetartiodactyls based on insertions of short and long interspersed elements: hippopotamuses are the closest extant relatives of whales. *Proceedings of the National Academy of Sciences USA*, 96, 10261–10266.
- Sampson, V. & Grooms, J. (2009). Promoting and supporting scientific argumentation in the classroom: the evaluate-alternatives instructional model. *Science Scope*, 33, 66–73.
- Stigler, J.W. & Hiebert, J. (2009). *The Teaching Gap: Best Ideas from the World's Teachers for Improving Education in the Classroom*. New York, NY: Free Press.

KD DAVENPORT is a teacher at Central High School, 1700 W. Olney Ave., Philadelphia, PA 19141; e-mail: davenportbiology@gmail.com. KIRSTIN JANE MILKS is a teacher at Bloomington High School South, 1965 S. Walnut St., Bloomington, IN 47401; e-mail: kmilks@mccsc.edu. REBECCA VAN TASSELL taught most recently at Horace Greeley High School, Chappaqua, NY 10514; e-mail: rebeccavantassell@gmail.com.



The DNA Store
 DNA Items:
 Toys, models,
 neckties, art,
 earrings, mugs,
 music, stamps,
 balloons, coins,
 stamps, cards - if you
 can think of it - we
 have it.

Visit www.TheDNAStore.com